22.106 Neutron Interactions and Applications (Spring 2006)
Lecture 5 (2/23/06)

Particle Simulation Methods I: Monte Carlo in Statistical Physics and Radiation Transport

---

References cited appear at the end of the lecture notes.

"The year was 1945. Two earth-shaking events took place: the successful test at Alamogordo and the building of the first electronic computer. Their combined impact was to modify qualitatively the nature of global interactions between Russia and the West. No less perturbative were the changes wrought in all of academic research and in applied science. On a less grand scale these events brought about a renascence of a mathematical technique known to the old guard as statistical sampling; in its new surroundings and owing to its nature, there was no denying its new name of the Monte Carlo method." -- from "The Beginning of the Monte Carlo Method", N. Metropolis, in *Los Alamos Science*, Special Issue 1987, p. 125.

Monte Carlo Sampling

Monte Carlo (MC) is a very general computational technique that can be used to do numerical integration or carry out sampling of distributions. In all cases one makes use of random numbers. Thus one way to define Monte Carlo is any calculation that involves random numbers. Historical accounts place the naming of this method in March 1947, when N. Metropolis suggested it in a jestful spirit in reference to an uncle of his colleague Stanislaw Ulam who would borrow money from relatives because he "just had to go to Monte Carlo". For our discussion here we will use Monte Carlo to sample the atomic configurations of the system at a given temperature.

There are three areas of statistical computation where MC is used. First is the numerical evaluation of multi-dimensional integrals. Second is the simulation of random walk processes (Markov chains) in statistical mechanics and condensed matter physics. The third area is that of particle and radiation transport. The essence of MC, as used in

statistical physics, is the Metropolis method of sampling which we will describe below. Tracking particles and radiation (neutrons, photons, charged particles) in transport problems is another significant area, which we will mention rather briefly at the end of the lecture. There is considerable literature on this subject [1-4]

Sampling Distributions

Let x be a random variable, that is, a variable whose value is prescribed by a certain probability distribution function. For example, the velocity of a particle is a random variable the distribution of which is given by the Maxwell-Boltzmann distribution. We denote the probability distribution of x by p(x),

p(x)dx = probability that the random variable will have a value lying between x and x + Δx

with normalization

$$\int_0^\infty p(x)dx = 1 \tag{5.1}$$

Let the corresponding cumulative distribution P(x) be

$$P(x) = \int_0^x P(t)dt \tag{5.2}$$

The relationship between these two quantities is shown in Fig. 1. Since the probability

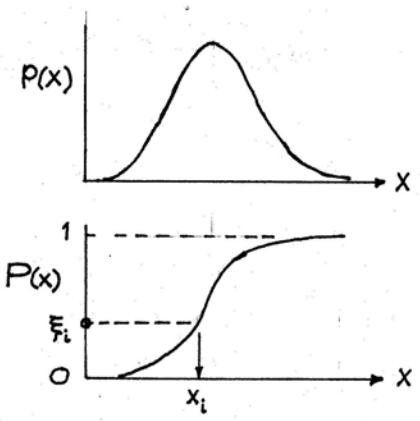

Fig. 1. A probability distribution p(x) and its cumulative distribution P(x). Using a random number ξ_i one samples the random variable x_i.

distribution is normalized to unity, the cumulative distribution is bounded in the range (0,1). Notice also that P(x) rises most sharply when x is in the region where p(x) has its peak. The is the behavior that makes it simple to understand the sampling, given in Eq.(5.3) below.

A special case is where p(x) is a uniform distribution in the range (0,1), then the corresponding random variable is called the **random number** ξ. In other words, random numbers are uniformly distributed in the interval (0,1); the sketch for p(x) is then just a constant over this interval and zero everywhere else, and P(x) is a straight line at 45^o angle.

When we say we want to sample a given distribution p(x), what we mean is that we will choose N_c values of a random variable x such that the resulting values, when plotted as a histogram, will give an outline resembling the shape of the distribution p(x). How closely they agree will depend on how many values one samples and how efficient is the sampling. To carry out the sampling, we take N_c random numbers, ξ_i, i=1,2,…,N_c, and set

$$P(x_i) = \xi_i \tag{5.3}$$

to obtain the N_c values of x_i, one after another. That (5.3) does give the desired sampled values can be seen from Fig. 1 and noting that there is a one-to-one correspondence between ξ_i and x_i. Intuitively, one expects that the region where P(x) is changing the most should be the region where x_i is most likely to occur, or in other words the region where p(x) has the largest value will be favored in the sampling.

Importance Sampling

In statistical physics one is often interested in finding the average of a property $A(\{\underline{r}_N\})$ in a system that is in thermodynamic equilibrium,

$$< A >\equiv \frac{\int d^{3N} r A(\{\underline{r}_N\}) e^{-\beta U(\{\underline{r}_N\})}}{\int d^{3N} r e^{-\beta U(\{\underline{r}_N\})}} \tag{5.4}$$

The calculation involves averaging the dynamical variable of interest, A, which depends on the positions of all the particles in the system, over an appropriate thermodynamic ensemble. Usually the ensemble chosen is the canonical ensemble which is represented by the Boltzmann factor $\exp(-U/k_B T)$, where U is the potential energy of the system, k_B the Boltzmann's constant, and T is the temperature. Integration is over the positions of all the particles (N particles, 3N coordinates). The denominator in (4) appears because of normalization; it is an important quantity in itself in thermodynamics, being known as the partition function.

We imagine there are two ways to perform the indicated integral. One approach is to sample N_c configurations **randomly** and then obtain <A> by carrying out Eq.(8.4) as a sum over a set of particle positions sampled according to the canonical distribution

$$< A >= \sum_{j=1}^{N_c} A(\{\underline{r}_N\}_j) e^{-\beta U(\{\underline{r}_N\}_j)} / \sum_{j=1}^{N_c} e^{-\beta U(\{\underline{r}_N\}_j)} \tag{5.5}$$

In practice this procedure is inefficient because it is quite easy to get a high-energy configuration ($U >> k_B T$) in which case the exponential makes the contribution negligible. The net result is then only a few configurations determine the value of <A> which is clearly undesirable.

To get around this difficulty, one has the second approach where the sampled configurations are not obtained randomly, but from the distribution $\exp(-\beta U)$. Then <A> is determined by weighing the contributions from each configuration equally,

$$< A >= \frac{1}{N_c} \sum_{j=1}^{N_c} A(\{\underline{r}_N\}'_j) \tag{5.6}$$

where $\{\underline{r}_N\}_j^{'}$ are configurations sampled from the distribution $e^{-\beta U(\{\underline{r}_N\}_j^{'})}$. How does one do this? One way is to adopt a procedure developed by N. Metropolis and colleagues in 1953 [5]. This procedure is an example of the concept of importance sampling in Monte Carlo methods.

Metropolis Sampling [5]

This is quite a famous procedure; it is best explained by considering a particle making a displacement in 2D. Let the initial position of the particle be (x,y) and the system potential energy U which depends on the particle position. Imagine now displacing the particle from its initial position to a trial position $(x+\alpha\xi_1^{'}, y+\alpha\xi_2^{'})$, where α is a constant to be adjusted, and $\xi_i^{'} = 2\xi_i - 1$, i = 1 or 2. Notice that $\xi_i^{'}$ is a random number in the interval (-1,1). With this move the system goes from configuration $\{\underline{r}_N\}_j^{'} \rightarrow \{\underline{r}_N\}_{j+1}^{'}$. The Metropolis procedure now consists of 4 steps.

1. *Move system in the way just described.*

2. *Calculate* $\Delta U = U(final) - U(initial) = U_{j+1} - U_j$. Note ΔU is the **energy gain** from the move.

3. *If* $\Delta U < 0$*, accept the move.* This means leaving the particle in its new position.

4. *If* $\Delta U > 0$*, still accept the move provided* $e^{-\beta\Delta U} > \xi$*, where* ξ *is a third random number in the present sequence (1 - 4).*

The **novel feature** of the method is step 4. It is simply a way to make the system **go uphill** from time to time. If not for step 4, step 3 would always let the system go downhill, which would mean that if the particle (system) were ever trapped in some local energy minimum, it has no way of getting out.

Proof of Metropolis Sampling [5]

By this we mean that one can show that the Metropolis procedure allows one to sample the distribution $\exp(-\beta U)$. Consider 2 states (configurations) of the system, r and s, and let $U_r > U_s$. According to the Metropolis procedure, the probability of an ($r \rightarrow s$) transition is $\nu_r P_{rs}$, where ν_r is the probability that the system is in state r, and P_{rs} is the transition probability that given the system is in state r it will go to state s. Similarly, the probability of $s \rightarrow r$ transition is $\nu_s P_{sr} e^{-\beta(U_r - U_s)}$. At equilibrium the two transitions must be equal (otherwise the probability of population in one state versus the other will be piling and the system will not be in equilibrium). Thus,

$$\nu_r P_{rs} = \nu_s P_{sr} e^{-\beta(U_r - U_s)} \tag{5.7}$$

Now $P_{rs} = P_{sr}$ by virtue of microscopic reversibility, then (5.7) gives

$$\frac{\nu_r}{\nu_s} = \frac{e^{-\beta U_r}}{e^{-\beta U_s}}, \quad \text{or} \qquad \nu_r \propto e^{-\beta U_r} \tag{5.8}$$

This completes the proof of the Metropolis sampling method. Stated again, the Metropolis method is an efficient way to sample the Boltzmann factor which has the same form as the canonical distribution in thermodynamics. It is worthwhile to note that this method can be used in any optimization problem where one is interested in finding the global minimum of a multidimensional energy space. The method is better than the standard energy minimization methods such as conjugate gradient because it allows the system to go uphill every now and then in the search for the global minimum. This is the basis of an algorithm in optimization called 'simulated annealing' [6].

Since simulated annealing has become a very powerful technique, we quote here the summary of ref. 6 --

> "There is a deep and useful connection between statistical mechanics (the behavior of systems with many degrees of freedom in thermal equilibrium at a finite temperature) and multivariate or combinatorial optimization (finding the minimum of a given

function depending on many parameters). A detailed analogy with annealing in solids provides a framework for optimization of the properties of very large and complex systems. This connection to statistical mechanics exposes new information and provides unfamiliar perspective on traditional optimization problems and methods."

Kinetic Interpretation of MC [3]

It may appear that MC is able to give only equilibrium properties averaged over a thermodynamic ensemble. This interpretation is unnecessarily restrictive as MC can be used to study time-dependent phenomena . Let P(x,t) be the probability that the system configuration is x at time t. Then P(x,t) satisfies the equation

$$\frac{dP(x,t)}{dt} = -\sum_{x'} W(x \to x')P(x,t) + \sum_{x'} W(x' \to x)P(x',t) \qquad (5.9)$$

where $W(x \to x')$ is the transition probability per unit time of going from x to x' (W is analogous to P_{rs} in the Metropolis method above). Eq.(5.9) is called the Master equation. For the system to be able to reach equilibrium, the transition probability must satisfy the condition (cf. Eq.(5.7)),

$$P_{eq}(x)W(x \to x') = P_{eq}(x')W(x' \to x) \qquad (5.10)$$

which is a relation known as the *principle of detailed balance*. At equilibrium, P(x,t) = P_{eq}(x) and dP(x,t)/dt = 0. Since

$$P_{eq}(x) = \frac{1}{Z} e^{-\beta U(x)} \qquad (5.11)$$

where Z is the partition function, $Z = \sum e^{-\beta U(x)}$, (5.10) gives

$$\begin{aligned} W(x \to x') &= e^{-\beta[U(x')-U(x)]} && \text{U(x') – U(x) > 0} \\ &= 1 && \text{U(x') – U(x) < 0} \end{aligned} \qquad (5.12)$$

which corresponds to the Metropolis procedure. Thus we see that in adopting the Metropolis sampling one is in effect solving the master equation at equilibrium.

Simulation of particle and radiation transport

MC is quite extensively used to track the individual particles as each moves through the medium of interest, streaming and colliding with the atomic constituents of the medium. To give a simple illustration, we consider the trajectory of a neutron as it enters a medium, as depicted in Fig.2. Suppose the first interaction of this neutron is a scattering collision at point 1. After the scattering the neutron moves to point 2 where it is absorbed, causing a fission reaction which emits two neutrons and a photon. One of the neutrons streams to point 3 where it suffers a capture reaction with the emission of a photon, which in turn leaves the medium at point 6. The other neutron and the photon from the fission event both escape from the medium, to points 4 and 7 respectively, without undergoing any further collisions. By sampling a trajectory we mean that process in which one determines the position of point 1 where the scattering occurs, the outgoing neutron direction and its energy, the position of point 2 where fission occurs, the outgoing directions and energies of the two fission neutrons and the photon, etc. After tracking many such trajectories one can estimate the probability of a neutron penetrating the medium and the amount of energy deposited in the medium as a result of the reactions induced along the path of each trajectory. This is the kind of information that one needs in shielding calculations, where one wants to know how much material is needed to prevent the radiation (particles) from getting across the medium (a biological shield), or in dosimetry calculations where one wants to know how much energy is deposited in the medium (human tissue) by the radiation.

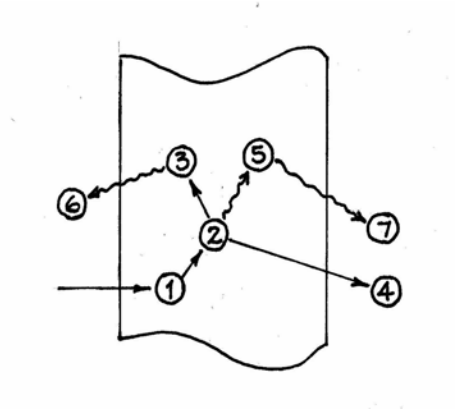

Fig. 2. Schematic of a typical particle trajectory simulated by Monte Carlo. By repeating the simulation many times one obtains sufficient statistics to estimate the probability of radiation penetration in the case of shielding calculations, or the probability of energy deposition in the case of dosimetry problems, etc.

References :

[1] H. Gould and J. Tobochnik, *An Introduction to Computer Simulation Methods* (Addison-Wesley, Reading, 1988), Part 2, chaps 10 - 12, 14, 15.

[2] D. W. Hermann, *Computer Simulation Methods* (Springer-Verlag, Berlin, 1990), 2nd ed., chap 4.

[3] K. Binder and D. W. Hermann, *Monte Carlo Simulation in Statistical Physics, An Introduction* (Springer-Verlag, Berlin, 1988).

[4] E. E. Lewis and W. F. Miller, *Computational Methods of Neutron Transport* (American Nuclear Society, La Grange Park, IL, 1993), chap 7.

[5] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, "Equation of State Calculations by Fast Computing Machines", *Journal of Chemical Physics* **21**, 1087 (1953).

[6] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi, "Optimization by Simulated Annealing", Science **220**, 671 (1983).