heat bath until the system reaches the temperature of the heat bath. If we imagine a large number of copies of a system at fixed volume V and number of particles N in equilibrium at temperature T, then the probability P_s that the system is in microstate s with energy E_s is given by (15.4)

We can use the Boltzmann distribution (15.4) to obtain the ensemble average of the physical quantities of interest. For example, the mean energy is given by

$$\langle E \rangle = \sum_s E_s P_s = \frac{1}{Z} \sum_s E_s e^{-\beta E_s}. \tag{15.11}$$

Note that the energy fluctuates in the canonical ensemble.

How can we simulate a system of N particles confined in a volume V at a fixed temperature T? Because we can generate only a finite number m of the total number of M microstates, an estimate for the mean value of a physical quantity A would be given by

$$\langle A \rangle \approx A_m = \frac{\sum_{s=1}^m A_s e^{-\beta E_s}}{\sum_{s=1}^m e^{-\beta E_s}}, \tag{15.12}$$

where A_s is the value of the physical quantity A in microstate s. A crude Monte Carlo procedure is to generate a microstate s at random, calculate E_s, A_s, and $e^{-\beta E_s}$, and evaluate the corresponding contribution of the microstate to the sums in (15.12). However, a microstate generated in this way would be very improbable and, hence, contribute little to the sums. Instead, we use an *importance sampling* method and generate microstates according to the probability distribution function π_s, which we will choose in the following.

To introduce importance sampling, we rewrite (15.12) by multiplying and dividing by π_s:

$$A_m = \frac{\sum_{s=1}^m (A_s/\pi_s) e^{-\beta E_s} \pi_s}{\sum_{s=1}^m (1/\pi_s) e^{-\beta E_s} \pi_s} \quad \text{(no importance sampling)}. \tag{15.13}$$

If we generate the microstates (configurations) with probability π_s, then (15.13) becomes

$$A_m = \frac{\sum_{s=1}^m (A_s/\pi_s) e^{-\beta E_s}}{\sum_{s=1}^m (1/\pi_s) e^{-\beta E_s}} \quad \text{(importance sampling)}. \tag{15.14}$$

That is, if we average over a biased sample generated according to π_s, we need to weight each microstate by $1/\pi_s$ to eliminate the bias. Although any form of π_s could be used, the form of (15.14) suggests that a reasonable choice of π_s is the Boltzmann probability itself, that is,

$$\pi_s = \frac{e^{-\beta E_s}}{\sum_{s=1}^m e^{-\beta E_s}}. \tag{15.15}$$

This choice of π_s implies that the estimate A_m of the mean value of A can be written as

$$A_m = \frac{1}{m} \sum_{s=1}^m A_s, \tag{15.16}$$

where each state is sampled according to the Boltzmann distribution. The choice (15.15) for π_s is due to Metropolis et al.

Although we discussed the Metropolis sampling method in Section 11.7 in the context of the numerical evaluation of integrals, it is not necessary to read Section 11.7 to understand the Metropolis algorithm in the present context. The Metropolis algorithm can be summarized in the context of the simulation of a system of spins as follows. The extension to other types of systems is straightforward.

1. Establish an initial microstate. (The energy of the initial microstate is not important.)
2. Choose a spin at random and make a trial flip.
3. Compute $\Delta E \equiv E_{\text{trial}} - E_{\text{old}}$, the change in the energy of the system due to the trial flip.
4. If ΔE is less than or equal to zero, accept the new microstate and go to step 8.
5. If ΔE is positive, compute the quantity $w = e^{-\beta \Delta E}$.
6. Generate a uniform random number r in the unit interval [0, 1].
7. If $r \leq w$, accept the new microstate; otherwise, retain the previous microstate.
8. Determine the value of the desired physical quantities.
9. Repeat steps (2) through (8) to obtain a sufficient number of microstates.
10. Periodically compute averages over the microstates.

Steps (2) to (7) lead to a transition probability that the system moves from microstate $\{s_i\}$ to $\{s_j\}$ proportional to

$$W(i \to j) = \min(1, e^{-\beta \Delta E}) \quad \text{(Metropolis algorithm)}, \tag{15.17}$$

where $\Delta E = E_j - E_i$. Because it is necessary to evaluate only the ratio $P_j/P_i = e^{-\beta \Delta E}$, it is not necessary to normalize the probability. Note that because the microstates are generated with a probability proportional to the desired probability, all averages become arithmetic averages as in (15.16). However, because the constant of proportionally is not known, it is not possible to estimate the partition function Z in this way.

Although we chose π_s to be the Boltzmann distribution, other choices of π_s are possible and are useful in some contexts. In addition, the choice (15.17) of the transition probability is not the only one that leads to the Boltzmann distribution. It can be shown that if W satisfies the *detailed balance* condition

$$W(i \to j) e^{-\beta E_i} = W(j \to i) e^{-\beta E_j} \quad \text{(detailed balance)}, \tag{15.18}$$

then the corresponding Monte Carlo algorithm generates a sequence of states distributed according to the Boltzmann distribution. The proof that the Metropolis algorithm generates states with a probability proportional to the Boltzmann probability distribution after a sufficient number of steps does not add much to our physical understanding of the algorithm. Instead, in Problems 15.8 and 15.9 we apply the algorithm to the ideal classical gas and to a classical magnet in a magnetic field, respectively, and verify that the Metropolis algorithm yields the Boltzmann distribution after a sufficient number of trial changes have been made.

Note that we have implicitly assumed in our discussion of the demon and Metropolis algorithms that the system is ergodic. That is, we have assumed that the important microstates of the system are being sampled with the desired probability. The existence of ergodicity depends on the way the trial moves are made and on the nature of the energy barriers between microstates. For example, consider a one-dimensional lattice of Ising spins with all