Background

• *Idea:* A method used to calculate values of quantities that depend on probability distributions $\rho(\nu)$ for random variables ν, such as mean values $\langle\phi\rangle = \sum_\nu \phi(\nu)\rho(\nu)$ of observables ϕ; in statistical mechanics, $\rho(\nu)$ is usually the canonical $e^{-\beta E_\nu}/Z$. If the calculation is too difficult to carry out analytically, either because ρ is not known sufficiently explicitly or because the sum over all ν's cannot be done, it may be convenient to do it using a numerical simulation. The general idea is that one generates a large number n of ν's at random, distributed according to $\rho(\nu)$, and then evaluates the averages in that sample of states, $\langle\phi\rangle = n^{-1}\sum_{i=1}^n \phi(\nu_i)$.

• *Possible situations:* When $\rho(\nu)$ is a function of one real variable x in some interval $[a,b]$ or it has an otherwise simple form, for example it depends on several variables but it factorizes into 1-variable functions or it is a uniform distribution on a discrete space, one can directly generate ν's with the right distribution.[1] In the cases of interest here, however, one has to use a different method: One starts by generating states with a different, easier distribution to work with, in practice the uniform one, and then uses some criterion to decide whether to accept or reject each configuration.

• *Straight Monte Carlo:* Generate configurations ν uniformly at random and independently of each other, then apply an acceptance/rejection criterion. Used when it is not too computationally intensive to generate a whole new configuration ν from scratch every time, and one knows how to compute $\rho(\nu)$ (this involves knowing Z in statistical mechanics), but the summation needed to obtain $\langle\phi\rangle$ cannot be done.

Metropolis Algorithm

• *Idea:* Choose an initial configuration ν_0, which in principle can be anything, then start generating small random steps that take the system to new configurations ν_I, in practice setting up a fictitious dynamical system. After generating each candidate step, apply an acceptance/rejection criterion adapted to the desired probability distribution to decide whether to keep it or not, then repeat the process. If the step is generated, and acceptance or rejection decided based solely on the configuration at the current time step and the candidate new configuration, then what we get is an example of a Markov process.

• *Markov process:* A Markov process is a stochastic process in which each transition probability $P_{\nu\nu'}$ only depends on the two states involved. The set of such probabilities must satisfy, by consistency, the master equation

$$\dot{P}_\nu = \sum_{\nu'} (-w_{\nu\nu'}\, P_\nu + w_{\nu'\nu}\, P_{\nu'})\ .$$

A system described by a stochastic process is in equilibrium if $\dot{P}_\nu = 0$; in particular, this is the case if detailed balance, $w_{\nu\nu'}\, P_\nu = w_{\nu'\nu}\, P_{\nu'}$, is satisfied $\forall\nu'$, which is a condition both on the system and the state.

• *Procedure:* Before you start, decide how each step will be generated, making sure that those steps can lead anywhere in configuration space;
– Generate the first ν_0 (in any way, random or not);
– Set $\nu = \nu_0$; – Calculate $\rho(\nu)$ // E_ν in the canonical ensemble case;
– Choose a possible step $\nu \mapsto \nu'$;
– Calculate $\rho(\nu')$ // $E_{\nu'}$ in the canonical ensemble case;
– If $\rho(\nu')/\rho(\nu) > 1$, accept the change, set $\nu = \nu'$ and repeat // $\Delta E := E_{\nu'} - E_\nu < 0$;
– If $\rho(\nu')/\rho(\nu) < 1$, accept it with probability $p = \rho(\nu')/\rho(\nu)$ // $\Delta E > 0$ and probability $p = e^{-\beta\Delta E}$.

• *Why this works:* Based on the procedure above,

$$w_{\nu\nu'} = \begin{cases} 1 & \text{if } \rho_{\nu'} \geq \rho_\nu \\ \rho_{\nu'}/\rho_\nu & \text{if } \rho_{\nu'} < \rho_\nu \end{cases}\ , \qquad w_{\nu'\nu} = \begin{cases} \rho_\nu/\rho_{\nu'} & \text{if } \rho_{\nu'} \geq \rho_\nu \\ 1 & \text{if } \rho_{\nu'} < \rho_\nu \end{cases}\ .$$

In either case, $w_{\nu\nu'}/w_{\nu'\nu} = \rho_{\nu'}/\rho_\nu$, which means that detailed balance is satisfied with $P_\nu = \rho_\nu$ for all ν.

Relevant Sections: [Phys 731]; Chandler, Ch 6; Halley, pp 158-159; [Reif]; [Schwabl].

---

[1] In the $\rho(x)$ case, generate a value y uniformly at random in [0,1], then calculate the value of x as $F^{-1}(y)$, where $F(x) := \int_a^x dx\, \rho(x)$; since y has unit probability density $\mu_y(y) = 1$, the resulting probability density for x is $\mu_x(x) = |dy/dx|\, \mu(y(x)) = \rho(x)$, as desired.