Statistical Mechanics and Thermodynamics

Let us go back to our original problem.

Example: A gas. Consider a gas (fluid) of N molecules in a box of volume V. The configurations of the gas are specified by the set of positions $\mathbf{r_1}, \mathbf{r_2}, \ldots$ (each vector defines a position inside the box) and of momenta $\mathbf{p_1}, \mathbf{p_2}, \ldots$ The configuration space is the phase space available to the system. The integral of momenta can be done analytically. We end up with configurational integrals.

For instance, we wish to compute the average potential energy at constant volume defined by

$$E = \frac{1}{Z} \int [dr] U e^{-\beta U},$$

where the integral is over the whole space, $\beta = 1/(k_B T)$ and Z is the configurational partition function.

We rewrite the integral using the probability (density) $\pi(r) = 1/V^N$ (normalized uniform probability in the box), as

$$E = \frac{\int [dr] \pi(r) U e^{-\beta U}}{\int [dr] \pi(r) e^{-\beta U}} = \frac{\langle U e^{-\beta U} \rangle_\pi}{\langle e^{-\beta U} \rangle_\pi}$$

In principle, one could imagine of using the following MC algorithm to compute E:

1. First, we generate a configuration randomly in the box (i.e., with a probability density $\pi(r)$). If the box is cubic, of size L, let $\mathbf{r}_1 = (LU_1, LU_2, LU_3)$, where U_i are random numbers uniformly distributed in [0,1] . Repeat the same for all N particles. We obtain a configuration $(\mathbf{r}_1, \ldots \mathbf{r}_N)$.
2. Compute $n_i = U(r_i) e^{-\beta U(r_i)}$ and $d_i = e^{-\beta U(r_i)}$, which are the arguments of the two integrals.
3. Repeat steps 1 and 2 n_{iter} times (i goes from 1 to n_{iter}). An estimate of E is $\sum_i n_i / \sum_i d_i$. Indeed,

$$\text{estimate of } \langle U e^{-\beta U} \rangle_\pi = \frac{1}{n_{\text{iter}}} \sum_i n_i$$

$$\text{estimate of } \langle e^{-\beta U} \rangle_\pi = \frac{1}{n_{\text{iter}}} \sum_i d_i$$

This algorithm is correct, but, for any reasonable n_{iter}, the error on the estimate is enormous.

The method does NOT work for the very same reason why thermodynamics works.

Consider a system of N particles at fixed temperature T.
The energy distribution is peaked around the mean energy E.
The relative width of the distribution is of order $1/\sqrt{N}$.

It implies that the configurations that essentially contribute to the integral are those for which

$$E(1-w/\sqrt{N}) < U < E(1+w/\sqrt{N}),$$

where w is the width.

But, our algorithm generates configurations with any value of U: most of them will be outside the relevant interval (which shrinks as N increases) and thus they will not be relevant for the computation of E. The result will depend on the very few configurations that belong to the relevant interval and thus the error will be large.

An example. Suppose we wish to compute

$$I = \int_0^1 e^{-10000x}\, dx$$

The function is essentially zero outside a small interval around 0, of width of the order of $w/10000$, with $w \approx 5 - 10$.

We estimate I as $\langle e^{-10000x} \rangle_\pi$ using the uniform distribution in [0,1].

I performed 1000 iterations, obtained the following numbers (8 different simulations): $4.23 \cdot 10^{-6}, 3.30 \cdot 10^{-5}, 2.98 \cdot 10^{-22}, 9.13 \cdot 10^{-8}, 1.57 \cdot 10^{-8}, 5.80 \cdot 10^{-7}, 1.90 \cdot 10^{-5}, 4.71 \cdot 10^{-4}$.
The correct result is $10^{-4} = 0.0001 = 10^{-4}$.
These results are very poor estimates of I.
Indeed, one can easily compute the error for N interations: $(\text{Var } f/N)^{1/2} = 0.007/\sqrt{N}$.
For $N = 1000$, the error is $2.2 \cdot 10^{-4}$: at this level of precision we could simply set $I = 0$, without any simulation.

To get a reasonable result, one should have $N_{\text{iter}} \gtrsim 10000$. This inequality is quite intuitive. If we generate the points randomly in [0,1], typically only w points every 10000 fall in the relevant region around 0 of width $w/10000$.
If we use $N_{\text{iter}} = 100000$, there are typically $10w$ points in the relevant region: the result is essentially determined by those $10w$ points. If we set $w = 5 - 10$, this means that the result depends on 50-100 points, which is not much. The other $N_{\text{iter}} - 10w$ points are simply irrelevant: we have wasted our time generating them.

WAY OUT: use **importance sampling**. Configurations are not generated randomly, but according to the canonical measure.
Use

$$\pi(\mathbf{r}) = \frac{1}{Z} e^{-\beta U(\mathbf{r})}$$

Unfortunately, it is very difficult to do it directly, except in a very few systems: dilute polymers (growth algorithms, dimerization), Gaussian models, percolation...

We need a new method: the **dynamic Monte Carlo method**, which mathematically is nothing but a **Markov process**.