Kinetic Monte Carlo
Day 1

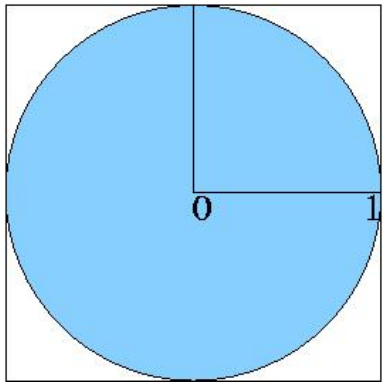
Wolfgang Paul
Institut für Physik
Johannes-Gutenberg Universität
55099 Mainz

Outline of Day 1

- Basics of the MC method
- Markov processes and the master equation
- Random walks and diffusion processes
- The Focker-Planck equation and the Itô SDE
- Diffusion in the double-well potential
- From the double-well to the egg-tray

Basics of the MC method

MC started and is mostly seen as a numerical way to calculate integrals



If we draw pairs of random numbers (x_i, y_i) where the x_i and the y_i are uniformly distributed in $[0, 1]$ the probability for the point to lie within the circle is equal to $\pi/4$.

When we turn to integration over a volume in a high-dimensional space, it becomes advantageous to use integration grid points that are randomly (uniformly) distributed in the volume.
For a regular grid and a finite number of total grid points N an increasing number of grid points lies on the surface of the volume.

$$\frac{\text{surface}}{\text{volume}} \propto \frac{d}{N^{1/d}}$$

Statistical Physics

Let us consider the canonical partition function as an example of such high-dimensional integrals

$$Z = \int_{\mathcal{C}} \mathrm{dx}\, \mathrm{e}^{-\beta U(x)}$$

for a Hamiltonian $H = \sum_i \frac{p_i^2}{2m_i} + U(x)$ and with $\beta = \frac{1}{k_B T}$.

problem: only regions with small $U(x)$ contribute significantly $\Rightarrow$ a uniform sampling "wastes" a lot of points to irrelevant regions of configuration space.

solution: importance sampling

Importance Sampling

Rewrite the partition function

$Z = \int_{\mathcal{C}} \mathrm{d}\mu(x)$ with the Gibbs measure $\mathrm{d}\mu(x) = \mathrm{e}^{-\beta U(x)}\mathrm{dx}$

The expectation value for an observable is

$$\langle O \rangle = \frac{1}{Z} \int_{\mathcal{C}} O(x) \mathrm{d}\mu(x) = \lim_{n \to \infty} \sum_{i=1}^{n} O(x_i)\mu(x_i) / \sum_{i=1}^{n} \mu(x_i)$$

When we generate points in configuration space according to their equilibrium distribution $\rho_{\mathrm{eq}}(x) = \frac{1}{Z}\mathrm{e}^{-\beta U(x)} \approx \mu(x_i) / \sum_{i=1}^{n} \mu(x_i)$ we get

$$\langle O \rangle \approx \frac{1}{n} \sum_{i=1}^{n} O(x_i)$$

Importance Sampling

How do we do this?

Idea: Simulate a Markov process which has $\rho_{\rm eq}(x)$ as unique stationary distribution.

Consider the conditional probabilities for a stochastic process

$$P_n(x_n,t_n;x_{n-1},t_{n-1};\ldots;x_1,t_1) = P_{1|n-1}(x_n,t_n|x_{n-1},t_{n-1};\ldots;x_1,t_1) \\ P_{n-1}(x_{n-1},t_{n-1};\ldots;x_1,t_1)$$

For a Markov process one has

$$P_n(x_n,t_n;x_{n-1},t_{n-1};\ldots;x_1,t_1) = P_{1|1}(x_n,t_n|x_{n-1},t_{n-1}) \\ P_{n-1}(x_{n-1},t_{n-1};\ldots;x_1,t_1)$$

Markov Chains

In discrete time we call this a Markov chain. The conditional probabilities of Markov processes obey the Chapman-Kolmogorov equation

$$P_{1|1}(x_3,t_3|x_1,t_1) = \int \mathrm{dx}_2 P_{1|1}(x_3,t_3|x_2,t_2)P_{1|1}(x_2,t_2|x_1,t_1)$$

For a stationary Markov process, we can write for its two defining functions

$$\begin{aligned} p_1(x,t) &= p_{\mathrm{eq}}(x)\ , \\ p_{1|1}(x_2,t_2|x_1,t_1) &= p_t(x_2|x_1)\ , t = t_2 - t_1\ , \end{aligned}$$

where we want p_t to denote a transition probability within the time interval t from state x_1 to state x_2.

Markov Chains

Using the Chapman–Kolmogorov equation for p_t, we get

$$p_{t+t'}(x_3|x_1) = \int dx_2\, p_{t'}(x_3|x_2) p_t(x_2|x_1\ ) \ .$$

When we consider a discrete probability space for x_i it is easily seen that this is a matrix multiplication, the p_t being matrices transforming one discrete state into another.

We now want to derive the differential form of the Chapman–Kolmogorov equation for stationary Markov processes. For this we consider the case of small time intervals t' and write the transition probability in the following way:

$$p_{t'}(x_3|x_2) = (1 - w_{\mathrm{tot}}(x_2)t')\, \delta(x_3 - x_2) + t' w(x_3|x_2) + \mathsf{o}(t') \ .$$

Markov Chains

This equation defines $w(x_3|x_2)$ as the transition rate (transition probability per unit time) from x_2 to x_3. $(1 - w_{\rm tot}(x_2)t')$ is the probability to remain in state x_2 up to time t', that is

$$w_{\rm tot}(x_2) = \int dx_3\, w(x_3|x_2) \ .$$

Inserting this into the Chapman–Kolmogorov equation results in

$$p_{t+t'}(x_3|x_1) = (1 - w_{\rm tot}(x_3)t')\, p_t(x_3|x_1) + t' \int dx_2\, w(x_3|x_2) p_t(x_2|x_1) \ ,$$

and so

$$\frac{p_{t+t'}(x_3|x_1) - p_t(x_3|x_1)}{t'}$$
$$= \int dx_2\, w(x_3|x_2) p_t(x_2|x_1) - \int dx_2\, w(x_2|x_3) p_t(x_3|x_1) \ ,$$

Markov Chains

in which we have used the definition of w_{tot}. In the limit $t' \to 0$ we arrive at the master equation, which is the differential version of the Chapman–Kolmogorov equation.

$$\frac{\partial}{\partial t} p_t(x_3|x_1) = \int dx_2\, w(x_3|x_2) p_t(x_2|x_1) - \int dx_2\, w(x_2|x_3) p_t(x_3|x_1)\ .$$

It is an integro-differential equation for the transition probabilities of a stationary Markov process.

When we do not assume stationarity and choose a $p_1(x_1, t) \neq p_{\mathrm{eq}}(x)$ but keep the assumption of time-homogeneity of the transition probabilities, i.e., they only depend on time differences, we can multiply this equation by $p_1(x_1, t)$ and integrate over x_1 to get a master equation for the probability density itself:

Markov Chains

$$\frac{\partial}{\partial t} p_1(x_3, t)$$
$$= \int dx_2 \, w(x_3|x_2) p_1(x_2, t) - \int dx_2 \, w(x_2|x_3) p_1(x_3, t) \; .$$

Let us change the notation: $p_1 \to p$, $x_3 \to x$ and $x_2 \to x'$

$$\frac{\partial}{\partial t} p(x, t) = \int dx' \, w(x|x') p(x', t) - \int dx' \, w(x'|x) p(x, t) \; .$$

On a discrete state space we would write this equation in the following way:

$$\frac{\partial}{\partial t} p(x, t) = \sum_{x'} \left(w(x|x') p(x', t) - w(x'|x) p(x, t) \right) \; .$$

Markov Chains

Going back to discrete time as in a simulation

$$t = n\Delta t, \quad \frac{\partial p(x,t)}{\partial t} \to \frac{p(x,n+1) - p(x,n)}{\Delta t}, \quad \tilde{w} = w\Delta t$$

we get

$$p(x,n+1) - p(x,n) = \sum_{x'} \left(\tilde{w}(x|x')p(x',n) - \tilde{w}(x'|x)p(x,n) \right) .$$

In the stationary state we have

$$p(x,n+1) = p(x,n) = p_{eq}(x)$$

so that

$$\sum_{x'} \left(\tilde{w}(x|x')p_{eq}(x') - \tilde{w}(x'|x)p_{eq}(x) \right) = 0 .$$

Detailed Balance

One way to fulfill this equation is to require detailed balance, i.e., the net probability flux between every pair of states in equilibrium is zero.

$$\frac{\tilde{w}(x|x')}{\tilde{w}(x'|x)} = \frac{p_{eq}(x)}{p_{eq}(x')}$$

For thermodynamic averages in the canonical ensemble we have $p_{eq}(x) = \frac{1}{Z}\exp\{-\beta H(x)\}$ and

$$\frac{\tilde{w}(x|x')}{\tilde{w}(x'|x)} = \exp\{-\beta(H(x) - H(x'))\}$$

When we use transition probabilities in our Monte Carlo simulation that fulfill detailed balance with the desired equilibrium distribution we are sure to have

$$\lim_{n\to\infty} p(x,n) = p_{eq}(x) \ .$$

Detailed Balance

The choice of transition rates is therefore not unique.
Common choices for these rates are

- the Metropolis rate

$$w(x'|x) = w_0(x'|x) min(1, \exp\{-\beta(H(x') - H(x))\}$$

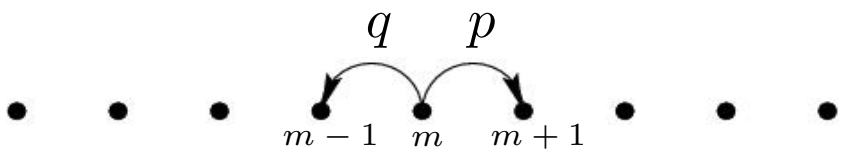
- the Glauber rate

$$w(x'|x) = w_0(x'|x)\frac{1}{2}(1 - \tanh\left(\exp\{-\beta(H(x') - H(x))\}\right)$$

In both of these rates we assumed $w_0(x'|x) = w_0(x|x')$ to be the probability to choose a certain pair of states which are connected through the selected set of moves, i.e., the flip of a single spin in an Ising model simulation.

Random Walk

Let us assume that a walker can sit at regularly spaced positions along a line that are a distance Δx apart; so we can label the positions by the set of whole numbers. Furthermore we require the walker to be at position 0 at time 0. After fixed time intervals Δt the walker either jumps to the right with probability p or to the left with probability $q = 1 - p$; so we can work with discrete time points, labeled by the natural numbers including zero.



What is the probability $p(m, N)$ that the walker will be at position m after N steps?

We can set up a rate equation

$$p(m, N+1) = p\, p(m-1, N) + q\, p(m+1, N)$$

Random Walk

Let us introduce the drift velocity, v, and the diffusion coefficient, D, into this equation:

$$v = (2p-1)\frac{\Delta x}{\Delta t}, \quad D = 2pq\frac{(\Delta x)^2}{\Delta t}$$

We can write

$$\begin{aligned} q &= (D - vq\Delta x)\frac{\Delta t}{(\Delta x)^2} \\ p &= (D + vp\Delta x)\frac{\Delta t}{(\Delta x)^2} \ . \end{aligned}$$

Inserting this into the rate equation and subtracting $p(m,N)$ we get

Random Walk

$$\frac{p(m,N+1)-p(m,N)}{\Delta t} = -vp\,\frac{p(m,N)-p(m-1,N)}{\Delta x}$$

$$-vq\,\frac{p(m+1,N)-p(m,N)}{\Delta x}$$

$$+D\frac{p(m+1,N)-2p(m,N)+p(m-1,N)}{(\Delta x)^2}$$

$$+\left(\frac{2D}{(\Delta x)^2}-\frac{1}{\Delta t}+\frac{vp}{\Delta x}-\frac{vq}{\Delta x}\right)p(m,N)\;.$$

The last term vanishes identically. When we now perform the continuum limit of this equation keeping v and D constant, we arrive at the Fickian diffusion equation.

Random Walk

$$\frac{\partial}{\partial t}p(x,t) = -v\frac{\partial}{\partial x}p(x,t) + D\frac{\partial^2}{\partial x^2}p(x,t) \ .$$

To derive this equation we had to perform a special continuum limit in time and space in which both $\Delta x/\Delta t$ and $(\Delta x)^2/\Delta t$ stay finite. This has to be understood in the sense that the average displacement for many walkers scales as Δt and the average squared displacement scales as Δt as well.

But what type of equation of motion for the walker in continuous time gives rise to such behavior?

For $v = 0$ the solution to above equation is

$$p(x,t) = \frac{1}{\sqrt{2\pi 2Dt}} \exp\left[-\frac{1}{2}\frac{x^2}{2Dt}\right] \ ,$$

Brownian Motion

This stochastic process is called Browian motion or Wiener process. The Wiener process (like every diffusion process) has continuous sample paths

$$\begin{aligned}
\mathrm{Prob}\big[\,|\mathrm{x(t+\Delta t)-x(t)}|>\mathrm{k}\big] &= \int_k^\infty dx\,\frac{2}{\sqrt{2\pi\Delta t}}\exp\left[-\frac{x^2}{2\Delta t}\right], \\
\lim_{\Delta t\to 0}\mathrm{Prob}\big[\,|\mathrm{x(t+\Delta t)-x(t)}|>\mathrm{k}\big] &= 0 \quad \forall\, k\,,
\end{aligned}$$

since the argument of the integral is a representation of the δ distribution as $\Delta t \to 0$.

The sample paths are, with probability one, nowhere differentiable:

$$\begin{aligned}
\mathrm{Prob}\left[\left|\frac{\mathrm{x(t+\Delta t)-x(t)}}{\mathrm{\Delta t}}\right|>\mathrm{k}\right] &= \int_{k\Delta t}^\infty dx\frac{2}{\sqrt{2\pi\Delta t}}\exp\left[-\frac{x^2}{2\Delta t}\right], \\
\lim_{\Delta t\to 0}\mathrm{Prob}\left[\left|\frac{\mathrm{x(t+\Delta t)-x(t)}}{\mathrm{\Delta t}}\right|>\mathrm{k}\right] &= 1 \quad \forall\, k\,.
\end{aligned}$$

Brownian Motion

The equation for Brownian motion is a Itô stochastic differential equation

$$dx(t) = \sqrt{2D}dW(t) \; .$$

The Wiener increments are Gaussian random variables with

$$\langle dW(t) \rangle = 0, \quad \langle dW(t)dW(t') \rangle = 0 \, \text{for} \, t \neq t', \quad \langle dW^2(t) \rangle = dt$$

From this we easily see

$$\langle dx(t) \rangle = 0, \quad \langle dx^2(t) \rangle = 2Ddt$$

which was our definition for the diffusion coefficient of the random walk for $p = q = 1/2$.

For our starting case including a drift term, the equation would be

$$dx(t) = vdt + \sqrt{2D}dW(t) \; .$$

When is MC kinetics reasonable?

For the random walk we have thus seen an example where we can perform a Monte Carlo simulation of a master equation (our original rate equation) and obtain a diffusive transport mechanism for large time ($\Delta t \to 0$).

More general we can say that MC kinetics is reasonable for relaxational processes and transport processes on "mesoscopic" to macroscopic time scales.

"Mesoscopic" meaning a time scale where vibrational motion has been damped out. This in turn means that the degrees of freedom we are interested in have to be coupled to a heat bath and subject to some friction mechanism.

Diffusion processes

Diffusive transport is equivalently described by

- a Fokker-Planck equation for the probability distribution

$$\frac{\partial}{\partial t}p(x,t) = -\frac{\partial}{\partial x}\left[a_1(x)p(x,t)\right] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[a_2(x)p(x,t)\right]$$

- a Langevin equation (Itô SDE) for the sample paths

$$dx(t) = a_1(x,t)dt + \sqrt{a_2(x,t)}dW(t)$$

The numerical solution of the Langevin equation involves generating random numbers to simulate the Wiener process. One can show, that one can use any distribution reproducing the first $2n$ moments of the Gaussian distribution of $dW(t)$ for an algorithm of order n in dt.

Itô Formula

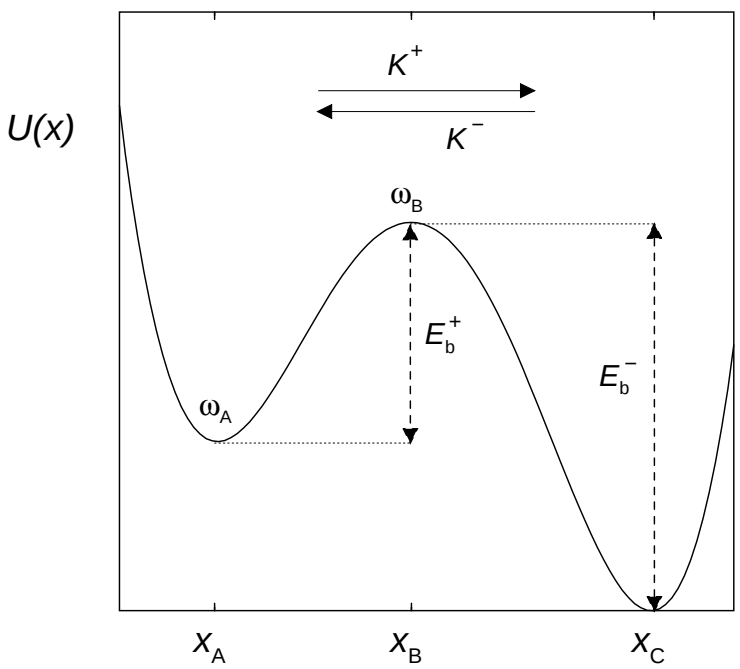
Suppose that $x(t)$ fulfills the Langevin equation of the previous page and that f is a functional of x and t. Then one has

$$
\begin{aligned}
& df[x(t),t] \\
&= \left[\frac{\partial}{\partial t} f[x(t),t] + a[x(t),t]\frac{\partial}{\partial x} f[x(t),t] + \frac{1}{2} b^2[x(t),t]\frac{\partial^2}{\partial x^2} f[x(t),t]\right] dt \\
&\quad + b[x(t),t]\frac{\partial}{\partial x} f[x(t),t]\, dW(t)\ .
\end{aligned}
$$

To derive this one has to use the special property of the diffusion processes that "$[dW(t)]^2 = dt$".

This transformation property will be of importance in the discussion of the application of Monte Carlo methods to finance problems.

Diffusion in the double well



External potential $U(x)$ as a function of some general position coordinate x. The particle starts in the metastable minimum around x_{A} and crosses the barrier to the equilibrium state at x_{C} by a thermally activated stochastic process

What is the average time the system needs to go from x_{A} to x_{B}?
This will be a thermally activated process.

Diffusion in the double well

We treat this problem in the overdamped case and only consider the position of the particle and not its velocity. The Langevin equation then reads

$$dx = \frac{F(x)}{\gamma M} dt + \sqrt{\frac{2k_{\mathrm{B}}T}{\gamma M}} dW(t) \; .$$

The corresponding Fokker–Planck equation is also called Smoluchowski equation because Smoluchowski treated free Brownian motion with this approach

$$\frac{\partial}{\partial t} p(x,t) = -\frac{\partial}{\partial x} \left[\frac{F(x)}{\gamma M} p(x,t) \right] + \frac{1}{2} \frac{\partial^2}{\partial x^2} \left[\frac{2k_{\mathrm{B}}T}{\gamma M} p(x,t) \right] \; .$$

What is the mean first passage time for going from x_{A} to x_{C}.

Diffusion in the double well

For the general form of the Fokker-Planck equation

$$\frac{\partial}{\partial t}p(x,t) = -\frac{\partial}{\partial x}\left[A(x)p(x,t)\right] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[B(x)p(x,t)\right] \ .$$

the result for the mean first passage time is

$$\overline{\tau}(x) = 2\int_x^{x_{\mathrm{C}}} dx'\phi^{-1}(x')\int_{-\infty}^{x'} dx''\frac{\phi(x'')}{B(x'')} \ .$$

with the abbreviation

$$\phi(x) = \exp\left[\int_{-\infty}^{x} dx'\frac{2A(x')}{B(x')}\right]$$

Diffusion in the double well

Inserting

$$A(x) = \frac{-U'(x)}{M\gamma} \quad \text{and} \quad B(x) = \frac{2k_{\mathrm{B}}T}{M\gamma} \ ,$$

into this solution, we get

$$\langle \tau_{\mathrm{xC}} \rangle = \frac{M\gamma}{k_{\mathrm{B}}T} \int_{x}^{x_{\mathrm{C}}} dx' \exp\left[\frac{U(x')}{k_{\mathrm{B}}T}\right] \int_{-\infty}^{x'} dx'' \exp\left[-\frac{U(x'')}{k_{\mathrm{B}}T}\right] \ .$$

With a quadratic expansion of the potential around its stationary points we arrive at Kramers result for the transition rate

$$\frac{1}{K^{+}} = \langle \tau_{\mathrm{AC}} \rangle \approx \frac{2\pi\gamma}{\omega_{\mathrm{B}}\omega_{\mathrm{A}}} \exp\left[\frac{E_{\mathrm{b}}^{+}}{k_{\mathrm{B}}T}\right] \ .$$

From the double well to the egg-tray

An egg-tray is the naive picture for a surface with a simple square lattice of adsorbtion sites and barriers between these sites which an adsorbed atom has to surmount to perform surface diffusion. This problem will be treated in much more detail on the third day of the tutorial.

This type of modeling can also be used to treat the diffusion of solutes through a glassy polymer matrix. The glassy matrix contains preferred adsorption sites with barriers between them. These sites and the paths connecting them define a random 3d network and the solute diffusion can be modeled as a random walk on this network.

From the double well to the egg-tray

For simplicity, we will study a 1d version of this problem, i.e., we are back to our random walk!

Suppose we identify a time interval

$$\Delta t = \frac{2\pi\gamma}{\omega_{\mathrm{B}}\omega_{\mathrm{A}}}$$

with each random walk step. Following Kramers' result we can write the probability for a jump to the right within 1 MCS:

$$p = \exp\left[-\frac{E_{\mathrm{b}}^{+}}{k_{\mathrm{B}}T}\right] \quad \text{and to the left} \quad q = \exp\left[-\frac{E_{\mathrm{b}}^{-}}{k_{\mathrm{B}}T}\right] \ .$$

With a probability $1 - p - q$ the RW remains at its position.

From the double well to the egg-tray

We can now look at the effect of different assumptions for the barrier heights

- constant barrier heights,
- alternating large and small barriers,
- random barrier heights followin a Gaussian distribution,
- random barrier heights following a Cauchy distribution.

This is the task for this afternoon.

Literature

1. W. Paul, J. Baschnagel, *Stochastic Processes: From Physics to Finance*, Springer, Berlin (2000).
2. N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, North Holland, Amsterdam (1987).
3. C. W. Gardiner, *Handbook of Stochastic Methods*, Springer, Berlin (1997).
4. K. Binder ed., *Monte Carlo Methods in Statistical Physics*, Springer, Berlin (1986).
5. K. Binder, D. W. Heermann, *Monte Carlo Simulation in Statistical Physics: An Introduction*, Springer, Heidelberg (1997).
6. D. Landau, K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics*, Cambridge University Press, Cambridge (2000).
7. D. Frenkel, B. Smit, *Understanding Molecular Simulation - From Algorithms to Applications*, Academic Press, San Diego (1996).