Stochastic simulations

Peter Olsson

0   Introduction

These notes will shortly introduce five different methods that can be used to simulate a gas of interacting particles. This introductory chapter has got number zero to make it possible to use numbers 1 through 5 for the chapters for these five different methods.

0.1   Adding interactions to the ideal gas

The simplest description of a gas is the Ideal gas law,

$$pV = Nk_BT,$$

which relates pressure times volume to the number of particles, Boltzmann's constant, and the temperature (measured in Kelvin) in a gas where the interactions between the molecules may be neglected. This is often a good approximation, e.g. at room temperature and a pressure of 1 atm. This equation is sometimes written $pV = nRT$, where $n = N/N_A$ is the number of moles of the substance (N_A is Avogadros constant) and $R = N_A k_B$ is the gas constant.

The properties of an ideal (non-interacting) gas may be calculated analytically but that is no longer possible when *interactions* between particles are included. (The effects of the interactions are however captured in an approxmative way in the van der Waals gas.)

We now approximate the interaction energy by a sum of pair interactions:

$$\begin{aligned} U(\{\mathbf{r}\}) &= \frac{1}{2}\sum_i \sum_j u(|\mathbf{r}_i - \mathbf{r}_j|) \\ &= \sum_i \sum_{j>i} u(|\mathbf{r}_i - \mathbf{r}_j|). \end{aligned}$$

0.2   The Lennard-Jones interaction

A common way to model the particle interaction is the Lennard-Jones interaction which is a pair interaction that is attractive, $u_{\mathrm{LJ}}(r) \sim -r^{-6}$, at large distances and repulsive at short distances. A common choice is to let the short distance repulsion be $\sim r^{-12}$. Taken together this gives the Lennard-Jones potential,

$$u_{\mathrm{LJ}}(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right], \tag{1}$$

where σ and ϵ are the length and energy scales of the interaction.

The minimum of the potential is at

$$r_{\min} = 2^{1/6}\sigma, \tag{2}$$

and the interaction energy is there given by

$$u_{\mathrm{LJ}}(r_{\min}) = -\epsilon. \tag{3}$$

0.3   Length and time scales

In the following we will use simulation units and will not be very much concerned with realistic values of energies, distances, and velocities. The point with this section is therefore only to give a feeling for the characteristic length, time, and energy scales of a real gas.

The parameters of the LJ interaction for Neon atoms are

$$\epsilon = 4.18 \times 10^{-22}\mathrm{J} \quad \text{and} \quad \sigma = 3.12 \times 10^{-10}\mathrm{m}.$$

The typical velocity may be obtained from (see Sec. 1.1)

$$m\langle \mathbf{v}^2 \rangle = 3k_B T,$$

where "3" is the dimensionality. With $k_B \approx 1.38 \times 10^{-23}$ J/K, $m \approx 28 \times 1.67 \times 10^{-26}$kg for N_2 molecules (typical of air), and $T = 300$K (which is $27\,^\circ C$), this gives the typical velocity

$$\sqrt{\langle \mathbf{v}^2 \rangle} \approx 10^3 \mathrm{m/s}.$$

With a potential energy that changes on scales of 10^{-10}m this gives a time scale of 10^{-13}s $= 0.1$ps, which is therefore the order of the time step that should be used when integrating the equations of motion for a gas at room temperature.

0.4   Simulation units

To use simulations units we let

$$\frac{r}{\sigma/\mathbf{m}} \to r \quad \text{and} \quad \frac{u_{\mathrm{LJ}}(r)}{\epsilon/\mathrm{J}} \to u_{\mathrm{LJ}}(r),$$

where $\sigma/\mathbf{m}$ (meter) and ϵ/J (Joule) are pure numbers. This means that r and u_{LJ} still have dimensions of length and energy, respectively, but that $r = \sigma$ in ordinary units is now $r = 1$. We also often set $m = 1$. For the temperature we let $k_B T/(\epsilon/\mathrm{J}) \to T$, which means that temperature now has the dimension of energy.

0.5 The Boltzmann distribution

Physical properties are often discussed for systems at constant temperature and the set of possible states are called the canonical ensemble. A constant temperature means that the energy may vary. This could at first seem to complicate things but it is rather the opposite. Including states irrespective of energy often simplifies both analytical calculations and numerical computations.

A key quantity in the canonical ensemble is the Boltzmann factor, $e^{-E_\nu/k_B T}$, and it is important since the probability of a state is proportional to the Boltzmann factor,

$$P_\nu \propto e^{-E_\nu/k_B T}.$$

The next section is a derivation of the Boltzmann factor.

0.5.1 Derivation of the Boltzmann factor

To discuss the properties of the canonical ensemble one considers the system, S, to be in thermal contact with a reservoir, R, such that both energies fluctuate, but that the total energy $E = E_R + E_S$ is a constant. The reservoir is taken to be much bigger than the system, such that the relative fluctuations in the energy of the reservoir are negligable. The temperature of the reservoir may therefore be considered constant.

The *micro*canonical ensemble is the collection of states with a certain energy. The basic assumption is then that each accessible state is equally probable. An important quantity in the microcanonical ensemble is the number of states with a certain energy, $\Omega(E)$. The entropy is related to this through

$$S(E) = k_B \ln \Omega(E),$$

where k_B is Boltzmann's constant. The temperature is in turn related to the entropy through

$$\frac{1}{T} = \frac{\partial S}{\partial E}.$$

3

Making use of the above in an expansion of $\ln \Omega$ for the reservoir gives

$$\ln \Omega(E - E_\nu) \approx \ln \Omega(E) - E_v \frac{\partial \ln \Omega}{\partial E} = \ln \Omega(E) - \frac{E_\nu}{k_B T}.$$

We now consider the probability for a certain state ν with energy E_ν of the system. This is related to the number of states of the reservoir with energy $E - E_\nu$:

$$P_\nu \propto \Omega(E - E_\nu) = e^{\ln \Omega(E - E_\nu)} \propto e^{-E_\nu/k_B T}, \tag{4}$$

and we recover the "Boltzmann factor" introduced above.

0.6 Expression for the pressure – the virial theorem

In a simulation with molecular dynamics it is possible to determine the pressure from the change of momentum the particles get when they bounce off the walls. In a Monte Carlo simulation this isn't possible and one instead has to determine the pressure from a kind of correlation function that is derived by means of the virial theorem. Note that this equation is only valid in ensembles with a fixed volume, which follows since the starting point for the derivation is the differential for the Helmholtz free energy.

The starting point is thus the differential

$$dF = -SdT - pdV + \mu dN,$$

which gives

$$p = -\left(\frac{\partial F}{\partial V}\right)_{N,T}.$$

Since $\beta F = -\ln Z$ we get

$$\beta p = \left(\frac{\partial \ln Z}{\partial V}\right)_{N,\beta},$$

and since the Hamiltonian separates into two parts we may write

$$Z = Z_{\text{kin}} Z_{\text{conf}},$$

where the kinetic part is independent of volume and is equal to $Z_{\text{kin}} = 1/v_Q^N$ and

$$Z_{\text{conf}} = \frac{1}{N!} \int d\mathbf{r}_1 \dots d\mathbf{r}_N e^{-\beta U(r^N)},$$

4

where the integration for each $\mathbf{r}_i$ should be performed over the volume V. Since only the configurational part depends on V, the pressure becomes

$$\beta p = \frac{\partial}{\partial V} \ln \left[\int d\mathbf{r}_1 \dots d\mathbf{r}_N e^{-\beta U(r^N)} \right].$$

This differentiation is tricky since the limits of integration depend on V. To get around this, change coordinates to

$$\mathbf{x} = \mathbf{r}/V^{1/d} \quad \Rightarrow \quad dr^N = V^N dx^N.$$

We then get

$$\begin{aligned}
\beta p &= \frac{\partial}{\partial V} \ln \left[V^N \int dx^N e^{-\beta \sum_{i<j} u(V^{1/d} x_{ij})} \right] \\
&= \frac{d}{dV} \ln V^N + \frac{\partial}{\partial V} \ln \left[\int dx^N e^{-\beta U} \right] \\
&= \frac{N}{V} + \left[\int dx^N \left(-\beta \frac{\partial U}{\partial V} \right) e^{-\beta U} \right] \Big/ \left[\int dx^N e^{-\beta U} \right].
\end{aligned}$$

(The integrations are here understood to be over a unit volume.) We also have

$$\frac{\partial U}{\partial V} = \sum_{i<j} \frac{\partial}{\partial V} u(V^{1/d} x_{ij}) = \sum_{i<j} \frac{\partial V^{1/d}}{\partial V} x_{ij} u'(V^{1/d} x_{ij}) = \frac{1}{Vd} \sum_{i<j} r_{ij} u'(r_{ij})$$

which follows since

$$\frac{\partial V^{1/d} x_{ij}}{\partial V} = \frac{1}{d} V^{1/d-1} x_{ij} = \frac{1}{Vd} V^{1/d} x_{ij} = \frac{1}{Vd} r_{ij}$$

Putting things together we get the virial theorem

$$\beta p = \frac{N}{V} - \frac{\beta}{Vd} \left\langle \sum_i \sum_{j>i} r_{ij} u'(r_{ij}) \right\rangle,$$

which in terms of the interparticle force becomes

$$pV = Nk_B T + \frac{1}{d} \left\langle \sum_i \sum_{j>i} r_{ij} F_{ij} \right\rangle. \tag{5}$$

Note that this simplifies to the ideal gas law in the absence of interactions, $F_{ij} = 0$.

0.7   Different kinds of simulations

This section summarizes five different methods that may be used for simulating a gas. There are differences both in the way to *describe* the system and in the kind of *dynamics* to use in the simulation. The numbers "1" through "5" refer to the section numbers in these notes.

i) Describe the gas in terms of $\{\mathbf{r}\}$ and $\{\mathbf{v}\}$.

    1) Molecular dynamics: ODE – deterministic.

    2) Langevin dynanics: ODE with a stochastic term.

ii) Describe the gas with $\{\mathbf{r}\}$, only.

    3) Brownian dynamics: $\Delta\mathbf{r}_i = c\mathbf{F}_i+$ noise.

    4) Monte Carlo: No real dynamics. Iterate the following for each particle i:

        * suggest a change $\Delta\mathbf{r}_i$ by random,
        * calculate the change in energy, ΔE_i,
        * accept this change with a probability that depends on $\Delta E_i/T$.

iii) Use a discrete space with positions $\mathbf{r}_\ell = (0,0,0), (0,0,1),\ldots(0,0,L-1),\ldots(L-1,L-1,L-1)$.

    5) The lattice gas: Each position can be either occupied or empty, $n_\ell = 0,\,1$. The number of particles fluctuates in the simulation.

Going upwards in this list one finds a more realistic simulation; going downwards the efficiency is higher.

1 Considerations for molecular dynamics

We have stressed in the earlier lectures that a given system may usually be simulated in a number of different ways. First, there is a possiblity of choosing different models (as we have already seen in the predator-prey example) and second it is often possible to simulate a given model with different techniques. In these lectures the focus will be on stochastic simulations — stochastic differential equations and Monte Carlo simulations. Molecular dynamics is only considered briefly as a background to the other methods.

1.1 Elementary physics

The behavior of a classical gas is fully specified by the Hamiltonian, which is the sum of the potential and kinetic energies,

$$\mathcal{H}(\{\mathbf{r}\}, \{\mathbf{v}\}) = U(\{\mathbf{r}\}) + K(\{\mathbf{v}\}),$$

which are, in turn, given by

$$U = \frac{1}{2} \sum_{i,j} u(\mathbf{r}_i - \mathbf{r}_j),$$

and

$$K = \sum_i \frac{m_i \mathbf{v}_i^2}{2}.$$

The dynamics is given by Newton's second law:

$$m_i \dot{\mathbf{v}}_i = \mathbf{F}_i \equiv -\nabla_i U(\{\mathbf{r}\}),$$

where ∇_i denotes differentiation with respect to $\mathbf{r}_i$.

The collisions in a gas means that the molecules exchange momentum ($m\mathbf{v}$) which after a short relaxation time, through the central limit theorem, leads to a normal distribution of the velocities. For each velocity component the probability distribution is given by[1]

$$P(v) \propto e^{-mv^2/2T},$$

and accordingly

$$m\langle v^2 \rangle = T. \tag{6}$$

The total energy is a constant of the motion but U and K fluctuate (in opposite directions) during the simulation.

---

[1]From now on we are using $T = k_B T^{\text{real}}$.

1.2 Cutoff in the interaction potential

If we were to do an "exact" simulation of a Lennard-Jones gas we would use the potential out to arbitrary distances. This is however very inefficient and it is therefore customary to introduce a cutoff, r_c, such that the interaction is zero for $r > r_c$. Such a cutoff introduces a step in the potential which gives an infinite force right at r_c and this is problematic. One therefore shifts the potential for $r < r_c$:

$$u_s(r) = \begin{cases} u_{\text{LJ}}(r) - u_{\text{LJ}}(r_c), & r < r_c, \\ 0, & r > r_c. \end{cases}$$

This is the expression we will use in the computer lab. In cases when one also needs a well-behaved second derivative, the above expression may be modified further to

$$u_s(r) = \begin{cases} u_{\text{LJ}}(r) - u_{\text{LJ}}(r_c) + (r_c - r)\left.\frac{du_{\text{LJ}}}{dr}\right|_{r_c}, & r < r_c, \\ 0, & r > r_c. \end{cases}$$

1.3 Periodic boundary conditions

The simulations always take place in a simulation cell of finite extension. The questions then arises what do to with the walls. To just let the particles disappear when they reach the boundary would be disastrous and to let the particles bounce off the wall isn't any ideal method either. The best method is to use periodic boundary conditions. To visualize how that works one can think of several identical simulation cells placed one after the other with the same particles. This is however only a mental picture, in the simulations each particle only exists once, but associated with each particle is an image in each cell.

There are then a few consequences of this arrangement:

- When a particle leaves the system through a wall it simultanously enters the system through the opposite wall. Each time a particle coordinate changes one should therefore check if the new position is outside the cell, $x < 0$ or $x > L$. If that is the case, the particle should be translated, $x + L \to x$ or $x - L \to x$, to the interval $[0, L)$.

- We only consider the interaction of a certain particle i with a single image of each other particle j; it is the nearest image that should

be used. (This applies to short range interactions. For long range interactions there are different techniques.) To find the nearest image one applies the following method for each of the directions x, y, and z: calculate $x_{ij} = x_i - x_j$, if $|x_{ij}| > L/2$ either add or subtract L such that $|x_{ij}| < L/2$. The distance vector is then

$$\mathbf{r}_{ij} = (x_{ij}, y_{ij}, z_{ij}).$$

- The nearest image convention also has implications for the smallest possible size of the simulation cell: The method may only be used consistently if each given particle has at most one image within the range of the interaction, which means that the linear size should always be $L \geq 2r_c$.

1.4   Integration method and time step

The basic equations we want to integrate numerically are

$$\dot{\mathbf{r}} = \mathbf{v},$$
$$m\dot{\mathbf{v}} = \mathbf{F}.$$

A good integration method should have time reversal symmetry. One possibility is velocity Verlet and another one is the leap-frog method where positions and forces are defined at times $t = n\Delta_t$ whereas velocities are defined at the intermediate times, $t = (n + 1/2)\Delta_t$. With the notation $\mathbf{r}_n \equiv \mathbf{r}(n\Delta_t)$, and similarly for force and velocity, we have

$$\mathbf{v}_{n+1/2} = \mathbf{v}_{n-1/2} + (1/m)\mathbf{F}_n\Delta_t,$$
$$\mathbf{r}_{n+1} = \mathbf{r}_n + \mathbf{v}_{n+1/2}\Delta_t.$$

A given simulation method is only numerically stable for a certain range of Δ_t. Beside this one should consider criteria based on the physical system in consideration. In the present case this means that we can only expect a good precision if the steps are such that the change in interaction potential in a single step is small. If the interaction potential changes on the scale ℓ, this implies that the time step Δ_t should be chosen such that

$$\sqrt{\langle v^2 \rangle}\, \Delta_t \ll \ell, \quad \text{or} \quad \Delta_t \ll \ell\sqrt{\frac{m}{T}}.$$

One should test the simulations with different time steps and choose a time step in the range where the measured results only depend very weakly on Δ_t.

2 Stochastic differential equation — Langevin dynamics

Langevin dynamics was first invented to simulate bigger molecules in a solvent with a large number of smaller molecules. The idea was to include the effect of these smaller molecules in an average way instead of keeping track of all the collisions. This can mean a huge speed up because (1) a large fraction of the molecules may be eliminated from the calculations, and (2) the heavier molecules typically move more slowly, which means that one can do with a bigger time step in the simulations.

However, in most cases one uses Langevin dynamics for other systems than mixtures of bigger and smaller molecules and the method is then mostly a way to implement a thermostat.

For Langevin dynamics one adds two terms to the simulations beside the ordinary ones from molecular dynamics: a random force and a friction term. The first term adds energy into the system whereas the second term has the effect to dissipate energy. The equation of motion then becomes

$$m\dot{\mathbf{v}} = \mathbf{F} - \alpha\mathbf{v} + \boldsymbol{\zeta}. \tag{7}$$

where $\boldsymbol{\zeta}$ is a vector noise.

We will shortly examine how the magnitude of $\boldsymbol{\zeta}$ should be chosen to do a simulation at a desired temperature, but before doing that we need to consider how the noise should be treated.

2.1 Properties of the noise

Each component of the noise is an independent random process with $\langle \zeta(t) \rangle = 0$ and the correlation between ζ at times t and t' given by the δ function,

$$\langle \zeta(t)\zeta(t') \rangle = A\delta(t - t').$$

It becomes clear that this noise process has very unusual properties as it jumps infinitely quickly between minus and plus infinity. To remedy the uneasy feeling one gets when contemplating this process, we instead consider a time interval Δ and the average noise during that time interval:

$$\zeta_\Delta(t) = \frac{1}{\Delta} \int_t^{t+\Delta} dt' \zeta(t').$$

It its then possible to calculate the expectation value, $\langle \zeta_\Delta^2 \rangle$, which shows that the amplitude of the noise depends on the time interval Δ in an unusual way:

$$
\begin{aligned}
\langle \zeta_\Delta^2 \rangle &= \frac{1}{\Delta^2} \int_t^{t+\Delta} dt' \int_t^{t+\Delta} dt'' \langle \zeta(t') \zeta(t'') \rangle \\
&= \frac{1}{\Delta^2} \int_t^{t+\Delta} dt' \int_t^{t+\Delta} dt'' A\delta(t' - t'') = \frac{A}{\Delta}.
\end{aligned}
$$

The amplitude of the noise is thus *bigger* for a smaller time step.

2.2   Magnitude of the noise

We are then ready to consider how the amplitude of the noise should be chosen. The key is to note that Eq. (7) contains both a damping term $-\alpha\mathbf{v}$ that dissipates energy and the noise term $\boldsymbol{\zeta}$ that adds energy. In equilibrium they should cancel each other out. The requirement is therefore that the noise should be chosen such that the expectation value of the kinetic energy doesn't change with time.

In one dimension and in the absence of external forces Eq. (7) gives

$$
mv(t + \Delta_t) = mv(t) - \alpha\Delta_t v(t) + \Delta_t \zeta.
$$

In equilibrium $\langle mv^2 \rangle$ at times $t + \Delta_t$ and t should be the same:

$$
m^2 \langle v^2(t + \Delta_t) \rangle = (m - \alpha\Delta_t)^2 \langle v^2(t) \rangle + \Delta_t^2 \langle \zeta^2 \rangle. \tag{8}
$$

To first order in Δ_t this leads to the requirement

$$
0 = -2m\alpha\Delta_t \langle v^2(t) \rangle + \Delta_t^2 \langle \zeta^2 \rangle.
$$

(Note that the second term, in spite of appearance, is to first order in Δ_t since $\langle \zeta^2 \rangle \sim 1/\Delta_t$.) This tells how to choose the magnitude of the noise,

$$
\langle \zeta^2 \rangle = \frac{2m\alpha \langle v^2(t) \rangle}{\Delta_t},
$$

and together with $m\langle v^2 \rangle = T$ this finally becomes

$$
\langle \zeta^2 \rangle = \frac{2\alpha T}{\Delta_t}. \tag{9}
$$

2.3 Effects of the damping α

When using Langevin dynamics as a way to implement a thermostat it is interesing to examine how a change in α affects the system. It turns out that the static, i.e. *time-independent*, quantities remain the same independent of α whereas the *time-dependent* quantites are affected. This means that if we consider snapshots only, and quantities that may be determined from snapshots, there is no way to tell what α was used in the simulation. If we, on the other hand, make a movie of the motion of a single particle we would see a clear difference.

One can understand this if one considers a dilute gas where the particles collide only seldom. In molecular dynamics (or in Langevin dynamics with very small α) the particles would then move in straight lines between abrupt collisions. On the other hand, if α is sizable the noise would make the particles deviate from the straight line behavior more or less all the time.

2.3.1 Time-dependent correlations

A good example of a time-dependent correlation function is the velocity autocorrelation function, which essentially measures how long a particle moves in the same direction. The definition is

$$g_v(t) = \langle \mathbf{v}_i(t' + t) \cdot \mathbf{v}_i(t') \rangle,$$

where the expectation value is meant as an average over all particles i, and many initial times, t'. The velocity autocorrelation function is important for the behavior of the system which e.g. can be seen by the relation to the diffusion constant.

At long times the particles diffuse around with

$$\langle (\mathbf{r}(t' + t) - \mathbf{r}(t'))^2 \rangle = Dt, \tag{10}$$

and it turns out that this constant—the diffusion constant, D—is related to the velocity autocorrelation function through

$$D = \int_{-\infty}^{\infty} g_v(t) dt. \tag{11}$$

2.3.2 Static quantities – the pair correlation function

A useful quantity to probe that static properies is the potential energy, which should be independent of α for sufficiently small time steps. More detailed

information about the behavior of the gas may be found in the pair correlation function, $g(r)$. The definition of $g(r)$ is

> The probability to find a particle at $\mathbf{r}$, granted that there is a particle at the origin.

To calculate $g(r)$ one collects a histogram over particle separations, $h(r)$, with resolution Δr. In two dimensions the relation between $h(r)$ and the pair correlation function is

$$h(r) = 2\pi r \; g(r)\Delta r,$$

which is easily inverted to get $g(r)$. At large distances $g(r)$ becomes equal to the density, whereas the behavior at short distances shows a peak at $r \approx r_{\min}$ (c.f. Eq. (2)), and possibly some more peaks at larger distances.

2.3.3 The time step and the damping

We finally consider the relation between α and the acceptable values of Δ_t.

It turns out that we have to decrease the time step when increasing α to avoid introducing errors in the results. To show this we keep terms proportional to Δ_t^2 in Eq. (8). This gives

$$m^2 \langle v^2(t + \Delta_t) \rangle = (m^2 - 2m\alpha\Delta_t + \alpha^2\Delta_t^2)\langle v^2(t) \rangle + \Delta_t^2 \langle \zeta^2 \rangle.$$

Assuming a time-independent kinetic energy, $\langle v^2(t + \Delta_t) \rangle = \langle v^2(t) \rangle$, we find

$$\langle v^2 \rangle (2m\alpha\Delta_t - \alpha^2\Delta_t^2) = 2\alpha T \Delta_t,$$

which gives

$$\langle v^2 \rangle = \frac{T}{m - \alpha\Delta_t/2}.$$

This estimate shows that the error depends on $\alpha\Delta_t$, which means that one needs smaller time step when the damping is larger.

3   Brownian dynamics

In the Langevin equation of motion, Eq. (7), the friction term introduces a time scale, m/α. To see this, consider a single particle in one dimension that initially moves with velocity $v(0)$. The dynamics $m\dot{v} = -\alpha v$ gives $v(t) = v(0)e^{-t/(m/\alpha)} +$ noise. Beside the noise—which has zero average—there is thus an exponential decay with the time scale m/α. (Note that this is for a single particle; in the presence of other particles the collisions will also affect the behavior.)

Consider now the limit where this time scale is very small: If $m/\alpha \ll \Delta_t$ the velocity decays quickly to zero and the dynamics is if there were no inertia—so called overdamped dynamics. One way to approach this limit would be to take $m \to 0$ and Eq. (7) then becomes

$$0 = \mathbf{F} - \alpha \mathbf{v} + \boldsymbol{\zeta}, \quad \Rightarrow \quad \mathbf{v} = \frac{1}{\alpha}\mathbf{F} + \frac{\boldsymbol{\zeta}}{\alpha}.$$

In a dynamics where the velocity is quickly damped out there is no need to remember the velocity and one can turn to a description with only the position coordinates. This then becomes

$$\dot{\mathbf{r}} = \frac{1}{\alpha}\mathbf{F} + \boldsymbol{\eta},$$

where the components of the noise variable are now characterized by

$$\langle \eta^2 \rangle = \frac{2T}{\alpha \Delta_t}.$$

3.1   The Fokker-Planck equation

From the above sketchy motivation the Brownian dynamics is only expected to be valid in rather special cases. We now want to demonstrate that Brownian dynamics actually is of more general validity and is always expected to give the correct equilibrium distribution, Eq. (4), in the limit of small Δ_t.

To simplify the notation we consider a particle in a one dimensional potential $U(x)$ at temperature T. We then require that the distribution should be

$$P(x) \propto e^{-U(x)/T}.$$

3.1.1 Derivation of the Fokker-Planck equation

Consider starting from some probability distribution $P(x, 0)$ at time $t = 0$. The change to this distribution with time comes from two terms:

$$P(x, \Delta_t) = P(x, 0) \quad - \quad \int d\tilde{x} \; D(x - \tilde{x}, \Delta_t | x, 0) \; P(x, 0)$$

$$+ \quad \int d\tilde{x} \; D(x, \Delta_t | x - \tilde{x}, 0) \; P(x - \tilde{x}, 0). \qquad (12)$$

The first integral contains $D(x - \tilde{x}, \Delta_t | x, 0)$ which describes the dynamics and is the probability that the particle which was at x at $t = 0$ is at $x - \tilde{x}$ at time Δ_t. The second integral is the probability that the particle originally at $x - \tilde{x}$ will be found at x at time Δ_t.

The first integral is simplified with the use of

$$\int d\tilde{x} \; D(x - \tilde{x}, \Delta_t | x, 0) = 1, \qquad (13)$$

which is just a statement that the particle originally at x has to be somewhere a time Δ_t later. With $\partial P(x, 0) / \partial t \approx [P(x, \Delta_t) - P(x, 0)] / \Delta_t$ one finds

$$\frac{\partial P(x, 0)}{\partial t} \Delta_t = -P(x, 0) + \int d\tilde{x} \; D(x, \Delta_t | x - \tilde{x}, 0) \; P(x - \tilde{x}, 0).$$

With

$$f(x) = D(x + \tilde{x}, \Delta_t | x, 0) \; P(x, 0),$$

this is

$$\frac{\partial P(x, 0)}{\partial t} \Delta_t = -P(x, 0) + \int d\tilde{x} \; f(x - \tilde{x}).$$

With the expansion

$$f(x - \tilde{x}) = \sum_{n=0}^{\infty} \frac{(-\tilde{x})^n}{n!} \frac{\partial^n}{\partial x^n} f(x),$$

the equation becomes

$$\frac{\partial P(x, 0)}{\partial t} \quad = \quad -\frac{P(x, 0)}{\Delta_t} + \sum_n \frac{(-1)^n}{n!} \frac{\partial^n}{\partial x^n} \left[P(x, 0) \frac{1}{\Delta_t} \int d\tilde{x} \; \tilde{x}^n D(x + \tilde{x}, \Delta_t | x, 0) \right]$$

$$\approx \quad -\frac{\partial}{\partial x} \left[P(x, 0) M_1 \right] + \frac{1}{2} \frac{\partial^2}{\partial x^2} \left[P(x, 0) M_2 \right], \qquad (14)$$

15

where
$$M_n = \frac{1}{\Delta_t} \int d\tilde{x}\ \tilde{x}^n D(x + \tilde{x}, \Delta_t | x, 0), \qquad (15)$$
and the $n = 0$ term in the sum cancels the $P(x, 0)$ term. Eq. (14) is the Fokker-Planck equation.

3.1.2 Application of the Fokker-Planck equation

We now want to use Eq. (14) to find the stationary probability distribution for a particle in a one dimensional potential $U(x)$ at temperature T, governed by Brownian dynamics. To that end we need to evaluate the integrals over $D(x + \tilde{x}, \Delta_t | x, 0)$ which are related to the dynamics,
$$x(\Delta_t) = x(0) + \frac{\Delta_t}{\alpha} F + \eta \Delta_t.$$
The dynamics may be written as a delta function,
$$D(x + \tilde{x}, \Delta_t | x, 0) = \delta \left(x + \frac{\Delta_t}{\alpha} F + \eta \Delta_t - (x + \tilde{x}) \right) = \delta \left(\frac{\Delta_t}{\alpha} F + \eta \Delta_t - \tilde{x} \right).$$
The relevant quantities are M_1 and M_2 *averaged over the random noise*:
$$M_n = \frac{1}{\Delta_t} \left\langle \int d\tilde{x}\ \tilde{x}^n\ \delta \left(\frac{\Delta_t}{\alpha} F + \eta \Delta_t - \tilde{x} \right) \right\rangle = \frac{1}{\Delta_t} \left\langle \left(\frac{\Delta_t}{\alpha} F + \eta \Delta_t \right)^n \right\rangle.$$
Using $F = -\partial U/\partial x$, $\langle \eta \rangle = 0$, and $\langle \eta^2 \rangle = 2T/\alpha \Delta_t$ we get, to lowest order in Δ_t,
$$M_1 = -\frac{1}{\alpha} \frac{\partial U}{\partial x} \quad \text{and} \quad M_2 = \frac{2T}{\alpha}.$$

3.1.3 Stationary solution

We now want to demonstrate that $P \propto e^{-U/T}$ is the stationary solution to the F-P equation, i.e. that it gives $\partial P/\partial t = 0$. We then plug in
$$P \propto e^{-U/T}, \quad \frac{\partial P}{\partial x} = -\frac{1}{T} \frac{\partial U}{\partial x} P, \quad \frac{\partial^2 P}{\partial x^2} = \left[\frac{1}{T^2} \left(\frac{\partial U}{\partial x} \right)^2 - \frac{1}{T} \frac{\partial^2 U}{\partial x^2} \right] P,$$
at the right hand side of Eq. (14). It is seen from inspection that the terms cancel one another out
$$-\left(\frac{1}{T} \frac{\partial U}{\partial x} P \right) \left(\frac{1}{\alpha} \frac{\partial U}{\partial x} \right) + P \left(\frac{1}{\alpha} \frac{\partial^2 U}{\partial x^2} \right) + \frac{T}{\alpha} \left[\frac{1}{T^2} \left(\frac{\partial U}{\partial x} \right)^2 - \frac{1}{T} \frac{\partial^2 U}{\partial x^2} \right] P = 0,$$

and the conclusion is that $P \propto e^{-U/T}$ is the stationary solution to the Fokker-Planck equation with Brownian dynamics.

3.1.4 Stochastic differential equations and Monte Carlo

The next section is about Markov chain Monte Carlo which is a method to generate configurations from arbitrary distributions (e.g. $e^{-U/T}$), but we already state the important result that a simulation of a stochastic differential equation is equivalent to a Monte Carlo simulation as far as the static quantities are concerned.

4 Markov chain Monte Carlo

The theoretical treatment of Monte Carlo simulations starts from the concept of a Markov chain but since this is very general and rather abstract we will instead start by introducing Monte Carlo by considering how the method is used for simulating a gas. We will then turn to the formulation in terms of Markov chains with the example of a gas as a useful point of reference.

4.1 Monte Carlo for an interacting gas

4.1.1 One sweep

In the standard implementation one loops over the particles, $i = 1, \ldots N$:

1. Suggest a change of the position of particle i:

$$\mathbf{r}'_i = \mathbf{r}_i + \boldsymbol{\delta}_i,$$

   where the $\delta_i^{(x)}$, $\delta_i^{(y)}$, and $\delta_i^{(z)}$, usually are from a rectangular distribution centered around zero, $\in [-b, b)$.

2. We now let ν to denote the initial configuration and μ the configuration where particle i is at $\mathbf{r}'_i$. Calculate the energy difference $\Delta U = U_\mu - U_\nu$ and the acceptance probability, from Eq. (18),

$$\alpha_{\nu \to \mu} = \min \left(e^{-\Delta U/T}, 1 \right).$$

3. Accept this new position with probability $\alpha_{\nu \to \mu}$, i.e. generate a random number $\xi \in [0, 1)$ and accept the change if $\xi < \alpha_{\nu \to \mu}$.

4.1.2 Remarks

To get a reasonably efficient simulation the range of the suggested changes $\boldsymbol{\delta}_i$ should be chosen such that the acceptance ratio is not too far from 50%, say between 30 and 70%.

   Note also that there is no need to calculate the full potential energy (from $\sum_{i<j} u(\mathbf{r}_{ij})$, which is a computation of order N^2) to compute the change in energy due to the change in position of particle i. One may instead use

$$\Delta U \equiv U_\mu - U_\nu = \sum_j [u(\mathbf{r}'_i - \mathbf{r}_j) - u(\mathbf{r}_i - \mathbf{r}_j)],$$

which is a computation of order N and therefore much faster.

4.2 Markov chains

To explain why the above works we now turn to a mathematical formulation in terms of Markov chains. This is an abstract formulation which is meant to work for all kinds of variables but it could be helpful to think of the concrete example we have discussed above. Also note that the matrices that are discussed below are never used as such in a computer. They are simply much too big for that to be possible.

Suppose now that there is a finite number of possible configurations or multidimensional variables $x^{(\nu)}$. A Markov chain is a random chain of these variables, x_1, x_2,... produced by means of a transition matrix $p_{\nu\mu}$ with the following properties:

1. $p_{\nu\mu} \geq 0$,

2. $p_{\nu\nu} \neq 1$,

3. $\sum_{\mu} p_{\nu\mu} = 1$ for all ν.

Note that (1) is necessary since probabilities have to be non-negative, (2) means that the chain may never come to a halt, and (3) means that the total probability to go to some state must always be equal to unity.

If the system is left to perform a random walk according to these transition probabilities and we can be guaranteed that the chain is ergodic—that we are not stuck in a limited region of the phase space— this will lead to a certain probability distibution $\pi_\nu \equiv \pi(x^{(\nu)})$. This is a profound result which makes Markov chains very useful. We have here for simplicity assumed a discrete set of possible configurations but Markov chains may also be defined for the more general case of a continuous configuration space.

4.3 Construct the transition matrix

The task is now to choose the transition matrix for the Markov chain such that a certain probability distribution $\pi(x)$ is obtained. A sufficient (but not necessary) condition is to require *detailed balance*, i.e.

$$\pi_\nu p_{\nu\mu} = \pi_\mu p_{\mu\nu}. \tag{16}$$

That this gives the required probability distribution may be seen by considering the probability to be in state μ after a given step, which is

$$\sum_\nu \pi_\nu p_{\nu\mu} = \pi_\mu \sum_\nu p_{\mu\nu} = \pi_\mu.$$

We have here made use of the detailed balance condition followed by condition (3) above.

The transition probability may be thought of to consist of two parts. We write

$$p_{\nu\mu} = q_{\nu\mu}\alpha_{\nu\mu},$$

where

$$
\begin{aligned}
q_{\nu\mu} &= \text{probability for the transition to be suggested} \\
\alpha_{\nu\mu} &= \text{probability for the transition to be accepted}
\end{aligned}
$$

To connect back to our configurations in Sec. 4.1 we note that in that case $q_{\nu\mu}$ is non-zero only for a very small subset of all possible configurations μ, i.e. the configurations with $\mathbf{r}_i^{(\mu)} = \mathbf{r}_i^{(\nu)}$ for $N - 1$ particles and $\mathbf{r}_i^{(\mu)} = \mathbf{r}_i^{(\nu)} + \boldsymbol{\delta}$ for the remaining particle. Here $\boldsymbol{\delta} = (\delta_i^{(x)}, \delta_i^{(y)}, \delta_i^{(z)})$ where $\delta_i^{(x)}$, $\delta_i^{(y)}$, and $\delta_i^{(z)}$ are from a rectangular distribution centered around zero.

Detailed balance may e.g. be fulfilled with the following choice for $\alpha_{\nu\mu}$:

$$\alpha_{\nu\mu} = \min\left(1, \frac{\pi_\mu q_{\mu\nu}}{\pi_\nu q_{\nu\mu}}\right) \tag{17}$$

The Metropolis algorithm

One often has a symmetric targeting probability $q_{\nu\mu} = q_{\mu\nu}$,

This is the case with the algorithm in Sec. 4.1. If the step $\nu \to \mu$ is possible the reverse step $\mu \to \nu$ is always possible, and equally likely to be suggested, as the step $\nu \to \mu$. This follows since the $\delta_i^{(x)}$, $\delta_i^{(y)}$, and $\delta_i^{(z)}$, are from a rectangular distribution centered around zero.

and the acceptance probability then simplifies to

$$\alpha_{\nu\mu} = \min\left(1, \frac{\pi_\mu}{\pi_\nu}\right), \quad \nu \neq \mu. \tag{18}$$

This is the choice in the *Metropolis* algorithm.

We will now explicitly demonstrate that the choice of $\alpha_{\nu\mu}$ according to Eq. (18) actually does fulfill detailed balance, i.e. that it fulfills $\pi_\nu p_{\nu\mu} = \pi_\mu p_{\mu\nu}$. To make this simple we show it for the case $q = q_{\nu\mu} = q_{\mu\nu}$; it is straightforward to show the general validity of the result.

$$\pi_\nu p_{\nu\mu} = \pi_\nu q \; \alpha_{\nu\mu} = \pi_\nu q \; \min(1, \pi_\mu/\pi_\nu) = q \; \min(\pi_\nu, \pi_\mu).$$

$$\pi_\mu p_{\mu\nu} = \pi_\mu q \; \alpha_{\mu\nu} = \pi_\mu q \; \min(1, \pi_\nu/\pi_\mu) = q \; \min(\pi_\mu, \pi_\nu).$$

The equality of these expression demonstrates that detailed balance actually is fulfilled.

4.3.1 Requirements for a valid Monte Carlo method

Many different methods have been proposed for doing Monte Carlo in all kinds of models. To prove that a method is correct one needs to show that two conditions are fulfilled:

1. The simulation is ergodic, i.e. that it is possible to reach any state through a finite number of steps.

2. That the Monte Carlo steps fulfill detailed balance, Eq. (17).

4.4 Monte Carlo and expectation values

We now consider the numerical calculation of expectation values. The definition of the expectation value of a quantity A is

$$\langle A \rangle = \sum_\nu A_\nu P_\nu,$$

where P_ν is the probability for the system to be in state ν. For an interacting gas this kind of formula cannot be used directly since there is a continuum of states[2] We are then bound to generating a limited number of states, and this may mainly be done in two different ways:

---

[2]In systems where the states constitute a discrete set, they are typically far too many for this sum to be performed on a computer.

- Generate n configurations, $\{\mathbf{r}^{(\nu)}\}$, for $\nu = 1 \ldots n$, randomly, with the $\mathbf{r}_i^{(\nu)}$ independent of one another. The expectation value is then

$$\langle A \rangle = \frac{\sum_{\nu=1}^{n} A_\nu e^{-E_\nu/T}}{\sum_{\nu=1}^{n} e^{-E_\nu/T}}, \tag{19}$$

However, attempting to do this, one will find that most configurations have a rather large energy and therefore a very small Boltzmann factor, $e^{-E_\nu/T}$. This means that most of the configurations contribute a negligable amount to the sums. The reason for this is that there is a high probability for at least one pair of these randomly generated particles to be very close to one another, which is enough to give the configuration a very high energy. This means that the configurations typical of low temperatures will almost never show up which has disastrous consequences for the possibility to calculate the properties at most temperatures of interest. Except for very small systems this is therefore an entirely impracticable method.

- Use Markov Chain Monte Carlo to generate the configurations with a probability $\pi_\nu \propto e^{-E_\nu/T}$. The expectation values should then be calculated as a direct average,

$$\langle A \rangle = \frac{1}{n} \sum_{\nu=1}^{n} A_\nu. \tag{20}$$

This is a useful method that allows for a simple calculation of many quantities that would otherwise be very difficult—if not altogether impossible—to calculate. The only drawback is that the configurations, and thereby the A_ν, are not independent of one another. One therefore needs to perform simulations for times that are very much longer than the memory of the system to get a large number of independent configurations.

4.4.1 Program structure

The program usually consists of a few different parts:

1. **Initialization** – Assign initial values to the positions in some way. Common alternatives are just random values or to read in the positions of an earlier run from a file.

22

2. **Equilibration** – We have to allow for some time before the Markov chain starts to produce configurations according to the correct probability distribution. This is achieved by calling the `mc_update` routine a (often fairly large) number of times. (This is often called "thermalization".)

3. **Production** – The actual run when data are collected. Results could also be written to a data file for later analysis. The below is a good structure for the simulation. In this example `val`, `vblock`, `v1sum`, and `v2sum` are structured variables with fields as epot and `pressure`.

```
for iblock =... nblock {
  vblock->epot = 0.0;
  for isamp = ... nsamp {
    mc_update(pos);
    measure( , pos, val);
    vblock->epot += val->epot; // the same also for other quantities
  }
  write_data(vsum);
  write_conf(pos);
  v1sum->epot += vblock->epot;
  v2sum->epot += vblock->epot * vblock->epot;
}
```

4. **Print out results** – The averages are obtained from `v1sum`. With `v2sum` it also becomes possible to determine the standard error.

5 The lattice gas and the Ising model

As already mentioned it is common with models that are defined on a lattice instead for on the continuum. An important reason for discretizing the system is that the computations then may be much simpler and thereby also considerably faster. It is immediately clear that a lattice model of a gas cannot reproduce all results from the continuum, but it turns out that certain features are not sensitive to the details of the model and if the goal is to examine such properties, the lattice model is superior to the continuum gas.

5.1 Coexistence region and the critical point

The gas changes to a liquid at temperatures and densities where the attractive interaction becomes important. Roughly speaking it prefers the liquid state for temperatures $T < \epsilon$. (Recall that the energy at the minimum is $-\epsilon$). One way to examine this transition is to start from the gas phase and then decrease the volume at a fixed temperature. At a certain specific volume (volume per particle or volume per mol) there appears a small volume with a higher density – a liquid. We say that we have coexistence between liquid and gas, the region in parameter space where this is possible is called the coexistence region. The specific volume of the gas, v_{gas}, is then higher than the specific volume of the liquid, v_{liq}. As the total volume decreases further, the liquid fraction increases further and the gas phase eventually disappears. This process takes place at a given constant pressure.

If the same process is repeated at different temperatures (and pressures) one finds that the difference in specific volume,

$$\Delta v = v_{\mathrm{gas}} - v_{\mathrm{liq}},$$

decreases when the temperature increases and actually vanishes at a given temperature, the "critical temperature", T_c. This is one of the coordinates that describe the "critical point", the other are v_c and p_c. When approaching the critical temperature the volume difference vanishes with a power law,

$$\Delta v \sim (T_c - T)^{\beta},$$

where $\beta \approx 0.3$ is a "critical exponent". In the region close to the critical point there are large regions with a lower density coexisting with other regions with a higher density. This is a behavior that is both difficult to analyze with

analytical methods and difficult to study with high precision in simulations. It turns out that one may capture this behavior much more efficiently by using a different model for simulation of the gas – the lattice gas.

5.2   The lattice gas

In the lattice gas we give up the continuum and instead let the particles sit on a lattice. This is a good idea if we want to capture the behavior at large length scales, though it of course doesn't reproduce the properties at small scales.

The variables in the lattice gas are $n_i = 0, 1$, corresponding to the presence or absence of a particle at lattice point i. The repulsion is given by the fact that two particles cannot be on the same lattice point and the attraction is given by assigning the energy $-\epsilon$ for each neighboring pair of particles. The energy for a certain configuration is then

$$E = -\epsilon \sum_{\langle ij \rangle} n_i n_j,$$

where $\sum_{\langle ij \rangle}$ denotes a sum over all nearest neighbors. Unlike the models we have discussed so far, the number of particles is not a constant in this model. The large collection of states which is possible when both energy and number of particles vary is called the grand canonical ensemble and one also introduces a chemical potential μ that controls the particle number, much as the temperature controls the fluctuating energy. The Boltzmann factor is then complemented by another term with the effect to keep $\langle n \rangle$ down.

5.3   The Ising model

The Ising model is usually thought of as a model of a magnet. This is in many ways similar to the lattice gas, the variables (the "spins") are now $s_i = \pm 1$, corresponding to spin up and spin down. The energy for the Ising model is[3]

$$E = -\epsilon \sum_{\langle ij \rangle} s_i s_j.$$

---

[3]It is also common to consider the effect of an applied magnetic field, this is included by adding the term $-h \sum_i s_i$ to this expression for the energy.

The Ising model has been called the Drosophila[4] of statistical mechanics because it has been so thoroughly studied and manipulated in all conceivable ways. One can put this model on lattices in different dimensions, $D = 1$, 2, 3,..., and it turns out that the behavior changes dramatically with the dimensionality. The behavior in one dimension is easy to calculate analytically. The properties of the phase transition in 2D is terribly more difficult to calculate. This problem was solved by the norwegian physicist Lars Onsager in 1944—an impressing tour de force, which by some is considered one of the greatest scientific achievements of the 20th century. The behavior in 3D is, in turn, immensely more complicated and continues to attract researchers from both physics and mathematics.

Literature

- *An introduction to Computer Simulation Methods*, Gould & Tobochnik, Addison-Wesley, 1996.

- *Stochastic simulation in physics*, P. Kevin MacKeown, Springer, 1997.

- *A guide to Monte Carlo simulations in statistical physics*, Landau & Binder, Cambridge University Press, 2000.

- *Monte Carlo Methods in Statistical Physics*, Newman & Barkema, Oxford University Press, 1999.

---

[4]A fruit fly that has been thoroughly studied in genetics by innumerable experiments with different kinds of mutations.