Nonequilibrium Statistical Physics

Lecturer: Prof Ramin Golestanian

Rudolf Peierls Centre for Theoretical Physics, University of Oxford

(Dated: January 22, 2020)

Hilary Term (16 lectures)

Contents

I. CHARACTERIZING DYNAMICS WITH RANDOMNESS

Statistical Physics is concerned with the developments of tools that can make a bridge between the *microscopic* world, which is controlled by fluctuations, and the *macroscopic* world that is deterministic. While the ensemble theory has been extremely successful in achieving this for equilibrium systems, many interesting physical phenomena involve nonequilibrium processes, such as relaxing to equilibrium, fluctuating around a nonequilibrium steady state, or undergoing a non-stationary nonequilibrium process. As technological advances allow us to probe the microscopic world with higher resolution every day, it is natural to ask how we can process the information we obtain using such experiments and how we can understand them using Statistical Physics. This question is what we aim to address in this course. We will develop dynamical formulations that will allow us to describe the statistical evolution of microscopic systems, and will use them to study a variety of interesting phenomena.

A. Brownian motion

Since the invention of microscope, many people observed random motion of microscopic objects, and this motion was always attributed to being a manifestation of life. In 1827, the botanist Robert Brown published the results of his observations (in the Edinburgh Philosophical Journal) in an article entitled "A brief account of microscopical observations on the particles contained in the pollen of plants ...", where he reported irregular jittery motion of small (clay) particles in pollen grains. Although several people worked on this phenomenon afterwards, a comprehensive physical explanation for it had to wait for nearly 80 years. A major breakthrough came with the work of Einstein in 1905, when he formulated the problem of a number of particles that are driven slightly out of equilibrium and showed that they will evolve to restore equilibrium. Einstein showed that the motion of the particles can be described by the diffusion equation

$$\frac{\partial}{\partial t}\rho(\boldsymbol{r},t) = D\boldsymbol{\nabla}^2\rho(\boldsymbol{r},t) \quad , \tag{I.1}$$

where $\rho(\boldsymbol{r},t)$ is the concentration profile of the particles and the diffusion coefficient D is given as

$$D = \frac{RT}{6\pi\eta a N_{\rm av}} \quad , \tag{I.2}$$

where R is the ideal gas constant, $N_{\rm av}$ is the Avogadro number, η is the viscosity of the liquid medium (water), and a is the radius of the particles. Here, $\zeta = 6\pi\eta a$ is the friction coefficient of a spherical particle moving in the fluid (Stokes drag*), and controls the *dissipation* in the system during the course of the random motion. The RT in the numerator signifies the source of energy for the stochastic motion: molecular thermal *fluctuations*. Equation (I.2) is a simple example of a *fluctuation-dissipation relation*, which makes a correspondence between *equilibrium* fluctuations in the system, as generally characterized by *correlation functions*, and the dissipation in the system, as characterized by *response functions*. Equation (I.2) was the basis of the first precise measurement of the Avogadro number by Jean Perrin in 1908.

While the probabilistic description of such a stochastic process is a natural choice, in 1908 Paul Langevin proposed an alternative formalism in terms of the stochastic dynamics of the particles. The virtue of this formalism—as proclaimed by Langevin in the original paper—is that "it is infinitely simpler than Einstein's approach". The calculation is as follows. Consider a small spherical particle of radius a and

* In general for an arbitrarily shaped object of size (i.e. the largest characteristic length) L, drag is of the form $\zeta \sim \eta L$

mass m that undergoes Brownian motion in a fluid of viscosity η, which is kept at temperature T. In the absence of other forces, we can write Newton's equation of motion for the position of particle $\boldsymbol{r}(t)$ as

$$m\frac{\mathrm{d}^2\boldsymbol{r}}{\mathrm{d}t^2} = -\zeta\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} + \boldsymbol{f}(t) \quad , \tag{I.3}$$

where the hydrodynamic drag force involves the friction coefficient ζ, and $\boldsymbol{f}(t)$ is the random force that the particle experiences due to the thermal buffeting by the fluid molecules. This force is a representation of the momentum transfer that the random collisions of the fluid molecules impart on the surface of the particle (see the discussion below about central limit theorem). The force is a random quantity, and due to symmetry, we expect that its average will be zero: $\langle \boldsymbol{f}(t) \rangle = 0$.

We can use Eq. (I.3) to calculate the mean-squared displacement (MSD) of the particle after a time t without specifying any other details of the random force if we assume that the solvent is in equilibrium.

Taking the scalar product of Eq. (I.3) with $\boldsymbol{r}$, we can rewrite it as

$$\frac{1}{2}m\frac{\mathrm{d}^2}{\mathrm{d}t^2}\boldsymbol{r}^2 + \frac{1}{2}\zeta\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{r}^2 - m\dot{\boldsymbol{r}}^2 = \boldsymbol{r}\cdot\boldsymbol{f}(t) \quad . \tag{I.4}$$

Taking the average of Eq. (I.4) and using the fact that there is no correlation between the random force and the position of the particle at any given time, i.e. $\langle \boldsymbol{r}(t)\cdot\boldsymbol{f}(t)\rangle = 0$, we find

$$\frac{1}{2}m\frac{\mathrm{d}^2}{\mathrm{d}t^2}\left\langle \boldsymbol{r}^2\right\rangle + \frac{1}{2}\zeta\frac{\mathrm{d}}{\mathrm{d}t}\left\langle \boldsymbol{r}^2\right\rangle - m\left\langle \dot{\boldsymbol{r}}^2\right\rangle = 0 \quad . \tag{I.5}$$

Implementing the equipartition theorem*,

$$\frac{1}{2}m\left\langle \dot{\boldsymbol{r}}^2\right\rangle = \frac{3}{2}\,k_{\mathrm{B}}T \quad , \tag{I.6}$$

we find

$$m\frac{\mathrm{d}^2}{\mathrm{d}t^2}\left\langle \boldsymbol{r}^2\right\rangle + \zeta\frac{\mathrm{d}}{\mathrm{d}t}\left\langle \boldsymbol{r}^2\right\rangle = 6k_{\mathrm{B}}T \quad . \tag{I.7}$$

Using the initial conditions $\boldsymbol{r}(0) = 0$, we can solve Eq. (I.7) to get the following expression for the MSD

$$\left\langle \boldsymbol{r}(t)^2\right\rangle = \frac{6k_{\mathrm{B}}T}{\zeta}\left[t + \tau\left(e^{-t/\tau} - 1\right)\right] \quad , \quad \tau = \frac{m}{\zeta} \quad . \tag{I.8}$$

We can understand Eq. (I.8) by examining its limiting behaviour

$$\left\langle \boldsymbol{r}(t)^2\right\rangle = \begin{cases} \dfrac{3k_{\mathrm{B}}T}{m}\,t^2 = \left\langle \boldsymbol{v}^2\right\rangle t^2 & , \quad t \ll \tau \\ \dfrac{6k_{\mathrm{B}}T}{\zeta}\,t \equiv 6Dt & , \quad t \gg \tau \quad , \end{cases} \tag{I.9}$$

where we have set

$$D = \frac{k_{\mathrm{B}}T}{\zeta} \quad , \tag{I.10}$$

---

* This is equivalent to assuming that the solvent molecules which are colliding with the particle are in equilibrium at temperature T.

which is identical to Eq. (I.2) and is called the Einstein relation.

The characteristic timescale $\tau = m/\zeta$ gives the cross-over from a ballistic regime, where $\langle \boldsymbol{r}(t)^2 \rangle = \langle v^2 \rangle t^2$ and the dynamics is controlled by inertia [the first term on the lhs of Eq. (I.7)], to a diffusive regime where $\langle \boldsymbol{r}(t)^2 \rangle = 6Dt$ and the dynamics is controlled by viscous drag [the second term on the lhs of Eq. (I.7)]. It is interesting to note that for micron sized particles (with density comparable to water), we find $\tau \simeq 10^{-7}$s. The characteristic time scales as $\tau \sim a^2$ with size, so it is considerably smaller for molecules of sub-nanometre size, namely $\tau \simeq 10^{-14}$s. Therefore, for many experiments the viscous regime (achieved by completely ignoring the inertial term in the Langevin equation) gives a very good account of the dynamics, as long as we are not interested in what is happening at timescales less than τ.

It should also be clear that, since the number of degrees of freedom for the equipartition theorem is equal to the dimension, the MSD as a function of time for motion in d dimensions should be written

$$\langle \boldsymbol{r}(t)^2 \rangle = 2\, d\, D\, t \quad . \tag{I.11}$$

B. Random walk

We can understand the $\sqrt{t}$ behaviour for diffusive motion using a simple model for random walk. Imagine a walker that takes steps at a regular pace in totally random directions. If we denote the displacement vector at step i by $\boldsymbol{b}_i$, we can write the total displacement vector after N steps as

$$\boldsymbol{R}_N = \sum_{i=1}^{N} \boldsymbol{b}_i \quad , \tag{I.12}$$

and the corresponding MSD as

$$\langle \boldsymbol{R}_N^2 \rangle = \sum_{i,j=1}^{N} \overbrace{\langle \boldsymbol{b}_i \cdot \boldsymbol{b}_j \rangle}^{b^2 \delta_{ij}} = b^2 N \quad , \tag{I.13}$$

where we have used the assumption that individual steps are independent from one another. Using this result, we can calculate the density of the path that is generated by the random walk (i.e. left behind) in a d-dimensional space as follows

$$\rho = \frac{N}{R_N^d} = \frac{N}{(b\sqrt{N})^d} = \frac{1}{b^d}\, N^{1-d/2} \quad , \tag{I.14}$$

which tells us that for $d > 2$ the density goes to zero as N goes to infinity, whereas it remains a constant or diverges for $d \leq 2$. This shows that the random walk is *recurrent* for $d \leq 2$, i.e. it goes over the same points over and over again, whereas for $d > 2$ this is not the case. Given that humans walk on (two dimensional) surfaces, the recurrence property is not terribly good news for "genuine" random walkers!

C. Gaussian distribution and white noise

This section is a quick reminder of the basic properties of Gaussian random variables that we will use throughout.

If a random variable is taken from a Gaussian distribution with average λ and variance σ^2, then the

probability of having a value y is given by

$$p(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\frac{(y-\lambda)^2}{\sigma^2}\right] \quad . \tag{I.15}$$

It is easy to show that

$$\int_{-\infty}^{\infty} \mathrm{d}x\, p(x) \;=\; 1 \quad . \tag{I.16}$$

$$\langle x \rangle = \int_{-\infty}^{\infty} \mathrm{d}x\, x\, p(x) \;=\; \lambda \quad . \tag{I.17}$$

$$\left\langle x^2 \right\rangle = \int_{-\infty}^{\infty} \mathrm{d}x\, x^2\, p(x) \;=\; \lambda^2 + \sigma^2 \quad . \tag{I.18}$$

$$\left\langle [x - \langle x \rangle]^2 \right\rangle = \left\langle x^2 \right\rangle - \langle x \rangle^2 \;=\; \sigma^2 \quad . \tag{I.19}$$

The Gaussian distribution can also be used to represent a Dirac delta function:

$$\delta(x-\lambda) = \lim_{\sigma \to 0} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\frac{(x-\lambda)^2}{\sigma^2}\right] \quad , \tag{I.20}$$

which has the required properties, i.e. $\int_{-\infty}^{\infty} \mathrm{d}x\, \delta(x-\lambda) = 1$ and $\delta(x-\lambda) = 0$ everywhere except at $x = \lambda$. We can check that as σ becomes smaller, the peak gets higher and narrower while the total area underneath remains constant.

We can obtain another representation of the Dirac delta function using Gaussian integration:

$$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-\lambda)^2}{\sigma^2}} = \int_{-\infty}^{\infty} \frac{dk}{2\pi}\, e^{-\frac{1}{2}\sigma^2 k^2 + ik(x-\lambda)} \quad . \tag{I.21}$$

We can now take the limit $\sigma \to 0$ on both sides of Eq. (I.21) to obtain the very useful result

$$\delta(x-\lambda) = \int_{-\infty}^{\infty} \frac{dk}{2\pi}\, e^{ik(x-\lambda)} \quad . \tag{I.22}$$

It is now straightforward to write an expression for the joint probability distribution of many Gaussian distributed random variables which are *independent* of each other by simply making the product of their individual Gaussian distributions.

Consider N independent Gaussian variables $x_1, x_2, \ldots, x_N \equiv \{x_i\}$. The probability of having a particular set of $\{x_i\}$ is

$$p\left[\{x_i\}\right] = \prod_{i=1}^{N} p_i(x_i) = \prod_{i=1}^{N} \left(\frac{1}{\sqrt{2\pi\sigma_i^2}}\right) \exp\left[-\frac{1}{2}\sum_{i=1}^{N}\frac{(x_i-\lambda_i)^2}{\sigma_i^2}\right] \quad . \tag{I.23}$$

From Eq. (I.23), it follows

$$\langle x_i \rangle \;=\; \lambda_i \quad . \tag{I.24}$$

$$\langle x_i x_j \rangle \;=\; \lambda_i \lambda_j + \sigma_i^2 \delta_{ij} \quad . \tag{I.25}$$

$$\int \prod_{i=1}^{N} \mathrm{d}x_i\, p\left[\{x_i\}\right] \;=\; 1 \quad . \tag{I.26}$$

This can be generalized to the case where the distribution is not factorized, i.e.

$$p\left[\{x_i\}\right] = \frac{1}{Z}\,\exp\left[-\frac{1}{2}\sum_{i,j=1}^{N} M_{ij}x_i x_j + \sum_{i=1}^{N} h_i x_i\right] \quad , \tag{I.27}$$

where the normalization reads

$$\begin{aligned} Z &= \int \prod_{i=1}^{N} \mathrm{d}x_i\,\exp\left[-\frac{1}{2}\sum_{i,j=1}^{N} M_{ij}x_i x_j + \sum_{i=1}^{N} h_i x_i\right] \quad , \\ &= (2\pi)^{N/2}\,(\det M)^{-1/2}\,\exp\left[\frac{1}{2}\sum_{i,j=1}^{N} M^{-1}_{ij} h_i h_j\right] \quad . \end{aligned} \tag{I.28}$$

We will have

$$\langle x_i\rangle = \sum_{k=1}^{N} M^{-1}_{ik} h_k \quad , \tag{I.29}$$

$$\langle x_i x_j\rangle = M^{-1}_{ij} + \sum_{k,l=1}^{N} M^{-1}_{ik} h_k M^{-1}_{jl} h_l \quad . \tag{I.30}$$

Now we can take the continuum limit and produce a probability distribution *functional*, which is the probability of finding a particular random *function*. This enables us to quantify the statistical properties of the random force $\boldsymbol{f}(t)$ in the Langevin equation, which drives the dynamics of Brownian motion. For a random function $\chi(t)$ (e.g. one component of $\boldsymbol{f}(t)$), which is Gaussian distributed with zero mean and variance σ^2, we can write the probability distribution functional as

$$p\left[\chi(t)\right] = \frac{1}{\mathsf{Z}}\exp\left[-\frac{1}{2\sigma^2}\int_0^{\mathsf{T}} \mathrm{d}t\,\chi(t)^2\right] \quad , \tag{I.31}$$

where

$$\mathsf{Z} = \int \mathcal{D}\chi(t)\exp\left[-\frac{1}{2\sigma^2}\int_0^{\mathsf{T}} \mathrm{d}t\,\chi(t)^2\right] \quad , \tag{I.32}$$

and T is a time interval set by the experimental observation time.

The continuum limit [transforming Eq. (I.23) to Eqs. (I.31) and (I.32)] is obtained by

$$x_i \;\to\; \chi(t) \tag{I.33}$$

$$\sum_{i=1}^{N} \;\to\; \int_0^{\mathsf{T}} \mathrm{d}t \tag{I.34}$$

$$\delta_{ij} \;\to\; \delta(t-t') \tag{I.35}$$

$$\int \prod_i \mathrm{d}x_i \;\to\; \int \mathcal{D}\chi(t) \equiv \int \prod_t d\chi(t) \quad . \tag{I.36}$$

If we define $t = n\Delta t$ such that $t = 0 \Leftrightarrow n = 0$ and $t = \mathsf{T} \Leftrightarrow n = N$ (i.e. $N = \mathsf{T}/\Delta t$), the continuum limit involves taking the limits $N\to\infty$ and $\Delta t\to 0$ such that $\mathsf{T} = N\Delta t$ remains finite.

We can use the probability distribution functional to calculate averages using the definition

$$\langle \cdots \rangle = \int \mathcal{D}\chi(t) \ \cdots \ p\left[\chi(t)\right] = \frac{1}{\mathsf{Z}} \int \mathcal{D}\chi(t) \ \cdots \ \exp\left[-\frac{1}{2\sigma^2} \int_0^{\mathsf{T}} \mathrm{d}t \, \chi(t)^2 \right] \quad , \tag{I.37}$$

which yields

$$\langle \chi(t) \rangle \ = \ 0 \quad , \tag{I.38}$$

$$\langle \chi(t) \, \chi(t') \rangle \ = \ \sigma^2 \delta(t - t') \quad . \tag{I.39}$$

A random function $\chi(t)$ which is distributed according to Eq. (I.31) is called *Gaussian white noise*. This occurs because the correlations are only instantaneous and the variance is independent of time; thus in Fourier space all frequencies are equally weighted (in analogy to white light that has all the colours of the spectrum). If the noise auto-correlation function is not local in time (while maintaining the time-translation invariance)

$$\langle \chi(t) \, \chi(t') \rangle = C(t - t') \quad , \tag{I.40}$$

then some of the frequencies in Fourier space will be more highly weighted than others, leading to *coloured* noise.

D. Cumulant expansion and central limit theorem

For an arbitrary probability distribution $p(x)$, a useful quantity called the characteristic function (or the generating function) can be defined via Fourier transformation

$$\left\langle e^{ikx} \right\rangle = \int \mathrm{d}x \, e^{ikx} \, p(x) = \tilde{p}(k) \quad , \tag{I.41}$$

which can be readily used to calculate all moments of the distribution as follows

$$\langle x^n \rangle = \frac{1}{i^n} \frac{\mathrm{d}^n}{\mathrm{d}k^n} \, \tilde{p}(k) \bigg|_{k=0} \quad . \tag{I.42}$$

The generating function defines the *cumulants* of the distribution through the following expansion

$$\ln \left\langle e^{ikx} \right\rangle = \sum_{n=1}^{\infty} \frac{(ik)^n}{n!} \, \langle x^n \rangle_c \quad , \tag{I.43}$$

where $\langle x^n \rangle_c$ is the n-th cumulant. For example, we have $\langle x \rangle_c = \langle x \rangle$ for the 1st cumulant, $\left\langle x^2 \right\rangle_c = \left\langle x^2 \right\rangle - \langle x \rangle^2$ for the 2nd cumulant, $\left\langle x^3 \right\rangle_c = \left\langle x^3 \right\rangle - 3 \left\langle x^2 \right\rangle \langle x \rangle + 2 \langle x \rangle^3$ for the 3rd cumulant, and so on. For a Gaussian distribution, all cumulants for $n \geq 3$ are identically zero.

It may appear that the choice of a Gaussian noise to model the random forces due to the collisions is arbitrary. However, this choice can indeed be justified using the *central limit theorem*. The theorem states that the sum of a large number of independent random variables $x_i, \ i \in 1, \ldots, N$ has a distribution function that asymptotically approaches a Gaussian for large N. To prove it, we need to start from a cumulant expansion of the quantity of interest, $\frac{1}{\sqrt{N}} \sum_{i=1}^{N} x_i$, and evaluate the scaling of each term with

respect to the large number N:

$$\begin{aligned}\left\langle \exp\left[ik\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}x_i\right)\right]\right\rangle &= \left\langle \exp\left[i\,\frac{k}{\sqrt{N}}\,x\right]\right\rangle^N \quad , \\ &= \exp\left[N\sum_m \frac{i^m}{m!}\left(\frac{k}{\sqrt{N}}\right)^m \langle x^m\rangle_c\right] = \exp\left[\sum_m \frac{(ik)^m}{m!}\,N^{1-m/2}\,\langle x^m\rangle_c\right] \quad , \\ &= \exp\left[ik\,N^{1/2}\,\langle x\rangle_c + \frac{(ik)^2}{2}\,\langle x^2\rangle_c + \frac{(ik)^3}{6}\,N^{-1/2}\,\langle x^3\rangle_c + \cdots\right] \quad .\end{aligned} \tag{I.44}$$

This is true for a large class of probability distributions $p(x_i)$ of the individual variables. The only restriction is that the cumulants need to have finite values. If this is the case, we can conclude that

$$\lim_{N\to\infty} P\left(y = \frac{1}{\sqrt{N}}\left[\sum_{i=1}^{N}x_i - N\,\langle x\rangle\right]\right) = \frac{1}{\sqrt{2\pi\,\langle x^2\rangle_c}}\exp\left[-\frac{1}{2}\frac{y^2}{\langle x^2\rangle_c}\right] \quad , \tag{I.45}$$

for 'microscopic' variables with arbitrary probability distributions.

This means that even though we do not in general know the probability distribution of the 'random' force exerted on a Brownian particle by an individual collision, if we coarse-grain to sufficiently long timescales then we can safely assume that the net force exerted on the particle is well approximated by Gaussian-distributed random variables, due to the large number of collisions during the coarse-grained time-step.

E. Diffusion equation and its general properties

So far we have focused on the stochastic trajectories of individual particles. A complementary approach is to study the probability $\mathcal{P}(\boldsymbol{r},t)$ of finding a particle at a position $\boldsymbol{r}$ at time t, which satisfies the diffusion equation

$$\frac{\partial}{\partial t}\mathcal{P}(\boldsymbol{r},t) = D\boldsymbol{\nabla}^2\mathcal{P}(\boldsymbol{r},t) \quad . \tag{I.46}$$

We can demonstrate the equivalence to the stochastic approach using the following calculation

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\langle \boldsymbol{r}(t)^2\right\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\int \mathrm{d}^d\boldsymbol{r}\;\boldsymbol{r}^2\,\mathcal{P}(\boldsymbol{r},t) = \int \mathrm{d}^d\boldsymbol{r}\;\boldsymbol{r}^2\,\frac{\partial}{\partial t}\mathcal{P}(\boldsymbol{r},t) = \int \mathrm{d}^d\boldsymbol{r}\;\boldsymbol{r}^2\,D\boldsymbol{\nabla}^2\mathcal{P}(\boldsymbol{r},t) = 2dD \quad , \tag{I.47}$$

which gives us Eq. (I.11).

We can solve for the full probability distribution by using Fourier transformation

$$\mathcal{P}(\boldsymbol{r},t) = \int \frac{d^d\boldsymbol{k}}{(2\pi)^d}\,e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\,\tilde{\mathcal{P}}(\boldsymbol{k},t) \quad . \tag{I.48}$$

The transformation converts Eq. (I.46) to $\partial_t\tilde{\mathcal{P}}(\boldsymbol{k},t) = -D\boldsymbol{k}^2\tilde{\mathcal{P}}(\boldsymbol{k},t)$, which can be integrated as $\tilde{\mathcal{P}}(\boldsymbol{k},t) = \tilde{\mathcal{P}}(\boldsymbol{k},0)\,e^{-D\boldsymbol{k}^2t}$. Implementing the initial condition that the Brownian particle starts from the origin at time $t=0$, i.e. $\mathcal{P}(\boldsymbol{r},0) = \delta^d(\boldsymbol{r})$, we obtain

$$\tilde{\mathcal{P}}(\boldsymbol{k},t) = e^{-D\boldsymbol{k}^2t} \quad . \tag{I.49}$$

Finally, Fourier transformation gives

$$\mathcal{P}(\boldsymbol{r},t) = \frac{1}{(4\pi D t)^{d/2}} \exp\left(-\frac{\boldsymbol{r}^2}{4Dt}\right) \quad . \tag{I.50}$$

We can also calculate the mean displacement, the MSD, and other moments, using the characteristic function

$$\left\langle e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \right\rangle = \int d^d\boldsymbol{r}\, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\, \mathcal{P}(\boldsymbol{r},t) = \tilde{\mathcal{P}}(\boldsymbol{k},t) \quad . \tag{I.51}$$

The moments are now calculated straightforwardly

$$\langle 1 \rangle = \tilde{\mathcal{P}}(\boldsymbol{k},t)\Big|_{\boldsymbol{k}=0} = 1 \quad , \tag{I.52}$$

$$\langle r_j \rangle = i\,\frac{\partial}{\partial k_j}\tilde{\mathcal{P}}(\boldsymbol{k},t)\Big|_{\boldsymbol{k}=0} = 0 \quad , \tag{I.53}$$

$$\langle r_i r_j \rangle = -\frac{\partial^2}{\partial k_i \partial k_j}\tilde{\mathcal{P}}(\boldsymbol{k},t)\Big|_{\boldsymbol{k}=0} = 2D\delta_{ij}t \quad , \tag{I.54}$$

which are compatible with what was obtained earlier.

1. Smoluchowski capture rate

How can we estimate reaction rates between non-interacting molecules? Smoluchowski argued that this limit will be set by the only means of transport they have to gain access to each other, i.e. diffusion. In order to have a concrete result, he considered the problem of a bath of diffusing particles that can be captured by a spherical object of radius R. Since there are many particles in the system (say N), we should use the concentration field $C(\boldsymbol{r},t) = N\mathcal{P}(\boldsymbol{r},t)$, which also satisfies the diffusion equation

$$\partial_t C = D\boldsymbol{\nabla}^2 C \quad , \tag{I.55}$$

instead of the probability distribution. The diffusive flux of particles is given as $\boldsymbol{J} = -D\boldsymbol{\nabla}C$. We seek a stationary distribution ($\partial_t C = 0$) and set the boundary conditions that the concentration is C_∞ far from the object and 0 on its surface, since we assume that any particle that reaches the surface will be captured by the spherical particle sink. This yields a solution $C(r) = C_\infty\,(1 - R/r)$, which gives a total flux of $J_r = -D\partial_r C(r=R) = -DC_\infty/R$, and the resulting capture rate

$$I = kC_\infty \quad ; \qquad k = 4\pi DR \quad . \tag{I.56}$$

For two species a and b with diffusion coefficients D_a and D_b and radii R_a and R_b, which are maintained at concentrations C_a and C_b, the same calculation can be repeated using the relative and centre of mass coordinates, to obtain

$$k_{ab} = 4\pi(D_a + D_b)(R_a + R_b) \quad , \tag{I.57}$$

with $R_a + R_b$ indicating the distance between the two molecules when they make contact. The rate of the reaction in this diffusion-limited regime will be $k_{ab}C_aC_b$, i.e. proportional to the concentrations of both species. This result can be generalized to the case where the particles interact with each other with a potential $V(\boldsymbol{r})$ (see Sec. IV.A).

F. Markov Processes

A stochastic process is called a Markov process if its time evolution at any instant of time is independent of its previous history. For a stochastic process where a quantity x evolves with time, we can define the probability distribution $\mathcal{P}(x,t)$, which gives the probability of finding the value x at time t, and the conditional probability $\mathcal{P}(x,t|x_0,t_0)$, which gives the probability of finding the value x at time t given that we had the value x_0 at time t_0. These quantities can be used to define higher order structures, such as the two-time joint probability $\mathcal{P}(x,t;x_0,t_0) = \mathcal{P}(x,t|x_0,t_0)\mathcal{P}(x_0,t_0)$*.

The Markov property of the process relates the conditional probabilities at different observation times through the so-called *Chapman-Kolmogorov equation*, which for the conditional probability defined above is written as

$$\mathcal{P}(x,t|x_0,t_0) = \int \mathrm{d}x_1\, \mathcal{P}(x,t|x_1,t_1)\mathcal{P}(x_1,t_1|x_0,t_0) \quad , \tag{I.58}$$

where $t > t_1 > t_0$. It is also possible to formulate the Markov property using the joint probability:

$$\mathcal{P}(x,t) = \int \mathrm{d}x_0\, \mathcal{P}(x,t;x_0,t_0) = \int \mathrm{d}x_0\, \mathcal{P}(x,t|x_0,t_0)\mathcal{P}(x_0,t_0) \quad , \tag{I.59}$$

where both equalities come from the definition of the joint probability. The Markov property is reflected in the fact that $\mathcal{P}(x,t)$ is independent of the history at any arbitrary prior time t_0. Thinking of the evolution of the stochastic quantity, the Chapman-Kolmogorov equation is understood as providing a discrete time evolution framework. When the configuration space of the Markov process itself is discrete, i.e. the system involves discrete states and stochastic jumps between them, the system is called a Markov chain. An example of a Markov chain is a random walker that takes discrete steps on a lattice.

---

* A more elaborate notation is sometimes used to distinguish these quantities: $\mathcal{P}_1(x,t)$ for the single-time probability, $\mathcal{P}_2(x,t;x_0,t_0)$ for the two-time joint probability, and $\mathcal{P}_{1|1}(x,t|x_0,t_0)$ for the conditional probability, plus straightforward generalizations. Our simple notation should be sufficient for the purpose of the material we cover, with the help of the context.

II. NONEQUILIBRIUM DYNAMICS: PROBABILISTIC VERSUS STOCHASTIC DESCRIPTION

A. Solving the Langevin equation

Using the statistical properties of Gaussian random variables, we can solve the Langevin equation, Eq. (I.3), and directly calculate average quantities. On timescales $t \gg \tau = m/\zeta$, we can ignore the inertial term (see Section I.A), and write Eq. (I.3) as

$$\zeta \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \boldsymbol{f}(t) \quad \Rightarrow \quad \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \boldsymbol{u}(t) \quad , \tag{II.1}$$

where $\boldsymbol{u}(t)$ is a random velocity related to the random force by $\boldsymbol{u}(t) = \dfrac{1}{\zeta}\,\boldsymbol{f}(t)$.

We model $\boldsymbol{u}$ as Gaussian white noise with a distribution functional

$$p[\boldsymbol{u}(t)] = \frac{1}{\mathsf{Z}} \exp\left(-\frac{1}{4D}\int \mathrm{d}t\, \boldsymbol{u}(t)^2\right) \quad , \tag{II.2}$$

The constant D is the diffusion coefficient and $\mathsf{Z} = \int \mathcal{D}\boldsymbol{u}(t)\, e^{-\frac{1}{4D}\int \mathrm{d}t\, \boldsymbol{u}(t)^2}$ is the normalization constant. We have chosen the variance of the distribution so as to recover the diffusive regime results we obtained earlier using the equipartition theorem, as shown below. Averages of fluctuating quantities (functionals of the noise, $\Omega[\boldsymbol{u}(t)]$) are obtained by averaging over the random velocity $\boldsymbol{u}(t)$ as

$$\langle \Omega[\boldsymbol{u}(t)] \rangle = \frac{1}{\mathsf{Z}} \int \mathcal{D}\boldsymbol{u}(t)\; \Omega[\boldsymbol{u}(t)]\; e^{-\frac{1}{4D}\int \mathrm{d}t\, \boldsymbol{u}(t)^2} \quad , \tag{II.3}$$

which therefore yields

$$\langle u_i(t) \rangle = 0 \quad , \tag{II.4}$$

$$\langle u_i(t) u_j(t') \rangle = 2D\delta_{ij}\delta(t-t') \quad . \tag{II.5}$$

Integrating Eq. (II.1) we obtain

$$\boldsymbol{r}(t) - \boldsymbol{r}(0) = \int_0^t \mathrm{d}t_1 \boldsymbol{u}(t_1) \quad , \tag{II.6}$$

from which the MSD can be calculated

$$\left\langle [\boldsymbol{r}(t) - \boldsymbol{r}(0)]^2 \right\rangle = \delta_{ij} \int_0^t \mathrm{d}t_1 \int_0^t \mathrm{d}t_2\, \langle u_i(t_1) u_j(t_2) \rangle = 6Dt \tag{II.7}$$

Explicit averaging over noise can be used to calculate the probability of finding the Brownian particle at position $\boldsymbol{x}$ at time t given the initial position of $\boldsymbol{x}_0$ at time $t = 0$, which we denote as $\mathcal{P}(\boldsymbol{x}, t|\boldsymbol{x}_0, 0)$. By the definition, this probability samples the relevant stochastic trajectories only, and hence can be written as

$$\mathcal{P}(\boldsymbol{x}, t|\boldsymbol{x}_0, 0) = \left\langle \delta^3\left(\boldsymbol{x} - \boldsymbol{r}(t)\right)\right\rangle\big|_{\boldsymbol{r}(0)=\boldsymbol{x}_0} \quad . \tag{II.8}$$

Using integral representation of the delta function together with Eq. (II.6), we have

$$\mathcal{P}(\boldsymbol{x}, t|\boldsymbol{x}_0, 0) = \left\langle \int \frac{\mathrm{d}^3\boldsymbol{k}}{(2\pi)^3}\, e^{i\boldsymbol{k}\cdot[\boldsymbol{x}-\boldsymbol{r}(t)]} \right\rangle\Bigg|_{\boldsymbol{r}(0)=\boldsymbol{x}_0} = \int \frac{\mathrm{d}^3\boldsymbol{k}}{(2\pi)^3}\, e^{i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{x}_0)} \left\langle \exp\left(-i\int_0^t \mathrm{d}t_1 \boldsymbol{k}\cdot\boldsymbol{u}(t_1)\right)\right\rangle \quad . \tag{II.9}$$

We can calculate the trajectory average by performing the Gaussian functional integral

$$\left\langle \exp\left(-i\int_0^t \mathrm{d}t_1 \boldsymbol{k}\cdot\boldsymbol{u}(t_1)\right)\right\rangle = \frac{1}{\mathsf{Z}}\int \mathcal{D}\boldsymbol{u}(t_1)\, \exp\left(-\frac{1}{4D}\int \mathrm{d}t_1\, \boldsymbol{u}(t_1)^2 - i\int_0^t \mathrm{d}t_1 \boldsymbol{k}\cdot\boldsymbol{u}(t_1)\right) \quad ,$$
$$= e^{-Dk^2 t} \quad . \tag{II.10}$$

Putting Eq. (II.10) in (II.9) and performing the Gaussian integration over $\boldsymbol{k}$ yields

$$\mathcal{P}(\boldsymbol{x},t|\boldsymbol{x}_0,0) = \frac{1}{(4\pi D t)^{3/2}} \exp\left\{-\frac{(\boldsymbol{x}-\boldsymbol{x}_0)^2}{4Dt}\right\} \quad , \tag{II.11}$$

which is the same as Eq. (I.50).

We can now generalize the basic framework that has been developed for Brownian motion of a particle in the absence of external forces, to study microscopic particles subject to additional forces as well as the random thermal kicks from their environment. The force balance (Newton's) equations must be modified to take into account these extra forces.

For a particle undergoing Brownian motion under the influence of an external potential $U(\boldsymbol{r})$, we write the Langevin equation as

$$\zeta\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = -\boldsymbol{\nabla}U + \boldsymbol{f}(t) \quad \Rightarrow \quad \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = -\frac{1}{\zeta}\boldsymbol{\nabla}U + \boldsymbol{u}(t) \quad , \tag{II.12}$$

where $\boldsymbol{u}(t)$ is a Gaussian white noise with the distribution functional given in Eq. (II.2).

We can rewrite the Langevin Eq. (II.12) as

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \boldsymbol{v}(\boldsymbol{r}(t)) + \boldsymbol{u}(t) \quad , \tag{II.13}$$

with a deterministic component of the velocity $\boldsymbol{v}(\boldsymbol{r}) = -\frac{1}{\zeta}\boldsymbol{\nabla}U(\boldsymbol{r})$ as well as a stochastic component $\boldsymbol{u}(t)$. The stochastic velocity has zero mean which means that on average the particle follows the deterministic trajectory while fluctuating about it. Although we have assumed that the deterministic velocity is derived from a conservative force, many of our future results will be valid for nonconservative forces as well.

1. Ornstein-Uhlenbeck process: the harmonic trap

Consider a particle under the influence of a parabolic potential $U(\boldsymbol{r}) = \frac{1}{2}k\boldsymbol{r}^2$, or a harmonic spring*. The Langevin equation for this particle is

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = -\frac{k}{\zeta}\,\boldsymbol{r} + \boldsymbol{u}(t) \quad , \tag{II.14}$$

---

* A colloidal particle in an optical trap is a perfect example of an Ornstein-Uhlenbeck process. Optical tweezers are highly focussed laser beams which can be used to trap microscopic particles in three dimensions. The trap works due to the fact that colloids (e.g. polystyrene beads) with a higher dielectric constant than the solvent they are suspended in are attracted to the centre of the beam (where the light intensity is highest). A good model of such a trap is a parabolic potential $U(\boldsymbol{r}) = \frac{1}{2}k\boldsymbol{r}^2$, where k measures the strength of the trap.

where $\boldsymbol{u}(t)$ is a Gaussian white noise. The equation can be solved by rearranging as

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[e^{kt/\zeta}\,\boldsymbol{r}(t)\right] = \boldsymbol{u}(t)\,e^{kt/\zeta} \quad , \tag{II.15}$$

which can be easily integrated to obtain

$$\boldsymbol{r}(t) = \boldsymbol{r}(0)\,e^{-kt/\zeta} + \int_0^t \mathrm{d}t_1\,\boldsymbol{u}(t_1)\,e^{-k(t-t_1)/\zeta} \quad . \tag{II.16}$$

Averaging over the noise, we find a mean trajectory described by

$$\langle \boldsymbol{r}(t)\rangle = \boldsymbol{r}(0)\,e^{-kt/\zeta} \quad , \tag{II.17}$$

indicating a decay to equilibrium from the initial position over a characteristic time scale $\tau_s = \zeta/k$ (the relaxation time).

The fluctuations around the mean trajectory are controlled by the noise term and can be calculated as

$$\begin{aligned}\left\langle [\boldsymbol{r}(t) - \langle \boldsymbol{r}(t)\rangle]^2 \right\rangle &= \int_0^t \mathrm{d}t_1 \int_0^t \mathrm{d}t_2\; e^{-k(2t-t_1-t_2)/\zeta}\;\langle u_i(t_1)u_i(t_2)\rangle = 6D\int_0^t \mathrm{d}t_1\; e^{-2k(t-t_1)/\zeta} \quad , \\ &= \frac{3D\zeta}{k}\left[1 - e^{-2\,t/\tau_s}\right] \quad . \end{aligned} \tag{II.18}$$

We can understand the results better if we look at the limiting behaviour

$$\left\langle [\boldsymbol{r}(t) - \langle \boldsymbol{r}(t)\rangle]^2 \right\rangle = \begin{cases} 6D\,t \quad , & t \ll \tau_s \\ \dfrac{3D\zeta}{k} = \dfrac{3k_{\mathrm{B}}T}{k} \quad , & t \gg \tau_s \quad , \end{cases} \tag{II.19}$$

where we have used the Einstein relation Eq. (I.10) to simplify the long time limit. We see that the long time limit only depends on equilibrium quantities and can thus be characterized by state variables only.

2. Nonequilbrium Ornstein-Uhlenbeck process

While the exact calculation of Sec. II.A.1 informs us of how the system achieves equilibration at long times, the same framework can be used to calculate the properties of intrinsically nonequilibrium processes, such as the case when we have a time dependent stiffness $k(t)$.

The method of integrating factors can be used to solve the Langevin equation

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = -\frac{k(t)}{\zeta}\,\boldsymbol{r} + \boldsymbol{u}(t) \quad , \tag{II.20}$$

and obtain

$$\boldsymbol{r}(t) = \boldsymbol{r}(0)\,e^{-\int_0^t \mathrm{d}t_1 k(t_1)/\zeta} + \int_0^t \mathrm{d}t_1\,\boldsymbol{u}(t_1)\,e^{-\int_{t_1}^t \mathrm{d}t k(t)/\zeta} \quad . \tag{II.21}$$

This leads to the following expression for the fluctuations

$$\left\langle [\boldsymbol{r}(t) - \langle \boldsymbol{r}(t)\rangle]^2 \right\rangle = 6D\int_0^t \mathrm{d}t_1\; e^{-2\int_{t_1}^t \mathrm{d}t k(t)/\zeta} \quad . \tag{II.22}$$

Therefore, for a specific protocol of the time-dependence of $k(t)$ (say oscillatory), we will obtain the resulting non-stationary behaviour of the mean-squared fluctuations at long times.

B. Probabilistic description of stochastic dynamics: from trajectory to probability

The information content of the stochastic trajectories can be represented via the probability distribution

$$\mathcal{P}(\boldsymbol{x},t) = \left\langle \delta^3\left(\boldsymbol{x}-\boldsymbol{r}(t)\right)\right\rangle \quad , \tag{II.23}$$

and we can formulate the stochastic dynamics such that we study the time evolution of the probability. This is equivalent to going from the path integral formulation of quantum mechanics to the Schrödinger description. The probability (distribution) is normalized by construction, $\int \mathrm{d}^3\boldsymbol{x}\,\mathcal{P}(\boldsymbol{x},t) = 1$, testifying that exactly one particle exists somewhere is space at any instant of time. This will not necessarily be valid in *open systems* where particles can enter or exit the observation region. The distribution is defined such that it can be used to calculate the noise-average [see Eq. (II.3)] of any quantity as

$$\langle \mathcal{O}(\boldsymbol{r}(t))\rangle = \int \mathrm{d}^3\boldsymbol{x}\,\mathcal{P}(\boldsymbol{x},t)\,\mathcal{O}(\boldsymbol{x},t) \quad . \tag{II.24}$$

To construct a probabilistic formulation, we start with the general Langevin equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{r}(t) = \boldsymbol{v}\left(\boldsymbol{r}(t)\right) + \boldsymbol{u}(t) \quad , \tag{II.25}$$

and integrate it over a finite, but small, time interval Δt to obtain

$$\Delta\boldsymbol{r}(t) \equiv \boldsymbol{r}(t+\Delta t) - \boldsymbol{r}(t) = \int_t^{t+\Delta t} \mathrm{d}t_1 \boldsymbol{v}(\boldsymbol{r}(t_1)) + \int_t^{t+\Delta t} \mathrm{d}t_1 \boldsymbol{u}(t_1) \quad , \tag{II.26}$$

which highlights the complicated form of the deterministic velocity when evaluated at the stochastic trajectory.

Next, we perform the following Taylor expansion

$$\delta^3\left(\boldsymbol{x}-\boldsymbol{r}(t+\Delta t)\right) = \delta^3\left(\boldsymbol{x}-\boldsymbol{r}(t)\right) - \Delta r_i(t)\,\partial_i\delta^3\left(\boldsymbol{x}-\boldsymbol{r}(t)\right) + \frac{1}{2}\,\Delta r_i(t)\Delta r_j(t)\,\partial_i\partial_j\delta^3\left(\boldsymbol{x}-\boldsymbol{r}(t)\right) \quad , \tag{II.27}$$

and neglect the $O(\Delta\boldsymbol{r}^3)$ corrections, where $\partial_i \equiv \frac{\partial}{\partial x_i}$. Averaging over the noise distribution as defined in Eq. (II.3), and using the definition of Eq. (II.23), we obtain

$$\begin{aligned}\mathcal{P}(\boldsymbol{x},t+\Delta t) - \mathcal{P}(\boldsymbol{x},t) &= \left\langle \delta^3\left(\boldsymbol{x}-\boldsymbol{r}(t+\Delta t)\right)\right\rangle - \left\langle \delta^3\left(\boldsymbol{x}-\boldsymbol{r}(t)\right)\right\rangle \\ &= -\partial_i\left\langle \Delta r_i(t)\,\delta^3\left(\boldsymbol{x}-\boldsymbol{r}(t)\right)\right\rangle + \frac{1}{2}\,\partial_i\partial_j\left\langle \Delta r_i(t)\Delta r_j(t)\,\delta^3\left(\boldsymbol{x}-\boldsymbol{r}(t)\right)\right\rangle \quad .\end{aligned} \tag{II.28}$$

We now expand the left hand side of Eq. (II.28) in Δt and keep the leading order, which gives $\mathcal{P}(\boldsymbol{x},t+\Delta t) - \mathcal{P}(\boldsymbol{x},t) \simeq \Delta t\,\partial_t\mathcal{P}(\boldsymbol{x},t)$. We then use the expressions

$$\left\langle \Delta r_i(t)\,\delta^3\left(\boldsymbol{x}-\boldsymbol{r}(t)\right)\right\rangle = \left\langle \Delta r_i(t)\right\rangle \left\langle \delta^3\left(\boldsymbol{x}-\boldsymbol{r}(t)\right)\right\rangle = \left\langle \Delta r_i(t)\right\rangle \mathcal{P}(\boldsymbol{x},t) \quad , \tag{II.29}$$

$$\left\langle \Delta r_i(t)\right\rangle = \int_t^{t+\Delta t} \mathrm{d}t_1 \underbrace{\left\langle v_i(\boldsymbol{r}(t_1))\right\rangle}_{\simeq v_i(\boldsymbol{r}(t)) = v_i(\boldsymbol{x})} + \int_t^{t+\Delta t} \mathrm{d}t_1 \underbrace{\left\langle u_i(t_1)\right\rangle}_{=0} \simeq \Delta t\, v_i(\boldsymbol{x}) \quad , \tag{II.30}$$

and

$$\left\langle \Delta r_i(t)\Delta r_j(t)\,\delta^3\left(\boldsymbol{x}-\boldsymbol{r}(t)\right)\right\rangle \;=\; \langle \Delta r_i(t)\Delta r_j(t)\rangle\,\left\langle \delta^3\left(\boldsymbol{x}-\boldsymbol{r}(t)\right)\right\rangle = \langle \Delta r_i(t)\Delta r_j(t)\rangle\;\mathcal{P}(\boldsymbol{x},t) \quad ,\text{(II.31)}$$

$$\langle \Delta r_i(t)\Delta r_j(t)\rangle \;\simeq\; \int_t^{t+\Delta t} \mathrm{d}t_1 \int_t^{t+\Delta t} \mathrm{d}t_2\; \underbrace{\langle u_i(t_1)u_j(t_2)\rangle}_{2D\delta_{ij}\,\delta(t_1-t_2)} = \Delta t\,(2D)\delta_{ij} \quad , \tag{II.32}$$

that are correct to the leading order in Δt, to obtain the following equation

$$\partial_t \mathcal{P}(\boldsymbol{x},t) = -\boldsymbol{\nabla}\cdot\left[\boldsymbol{v}(\boldsymbol{x})\,\mathcal{P}(\boldsymbol{x},t)\right] + D\,\boldsymbol{\nabla}^2\mathcal{P}(\boldsymbol{x},t) \quad , \tag{II.33}$$

when we take the limit $\Delta t \to 0$. The governing equation for the probability distribution, which is called the *Fokker-Planck equation* or the *Smoluchowski equation* depending on the context, is a conservation law:

$$\partial_t \mathcal{P}(\boldsymbol{x},t) + \boldsymbol{\nabla}\cdot\boldsymbol{J}(\boldsymbol{x},t) = 0 \quad . \tag{II.34}$$

where the Probability Flux is defined as

$$\boldsymbol{J}(\boldsymbol{x},t) = \overbrace{\boldsymbol{v}(\boldsymbol{x})\,\mathcal{P}(\boldsymbol{x},t)}^{\text{deterministic}}\overbrace{-D\,\boldsymbol{\nabla}\mathcal{P}(\boldsymbol{x},t)}^{\text{stochastic}} \quad . \tag{II.35}$$

1. Detailed balance and equilibrium

It is instructive to consider the case of a particle in a potential, $U(\boldsymbol{x})$, where

$$\boldsymbol{v}(\boldsymbol{x}) = -\mu\boldsymbol{\nabla}U(\boldsymbol{x}), \tag{II.36}$$

and the corresponding Fokker-Planck equation

$$\partial_t \mathcal{P}(\boldsymbol{x},t) + \boldsymbol{\nabla}\cdot\left[\left(-\mu\boldsymbol{\nabla}U(\boldsymbol{x})\right)\,\mathcal{P}(\boldsymbol{x},t) - D\,\boldsymbol{\nabla}\mathcal{P}(\boldsymbol{x},t)\right] = 0 \quad . \tag{II.37}$$

Note that we have used the mobility μ instead of the friction coefficient ζ; the two are related as $\mu = 1/\zeta$.

We can examine the equilibrium properties of the system, which can be obtained by setting $\partial_t\mathcal{P}(\boldsymbol{x},t) = 0$ and $\boldsymbol{J} = 0$. This is equivalent to the condition of detailed balance we will consider for Master equation in Sec. III.A. We find

$$-\mu\boldsymbol{\nabla}U(\boldsymbol{x})\,\mathcal{P}_{\mathrm{eq}}(\boldsymbol{x}) - D\,\boldsymbol{\nabla}\mathcal{P}_{\mathrm{eq}}(\boldsymbol{x}) = 0 \tag{II.38}$$

which gives an equilibrium distribution of states

$$\mathcal{P}_{\mathrm{eq}}(\boldsymbol{x}) = \frac{1}{\mathcal{Z}}\exp\left[-\frac{U(\boldsymbol{x})}{k_{\mathrm{B}}T}\right] \quad , \quad \mathcal{Z} = \int \mathrm{d}^3\boldsymbol{x}\, e^{-U(\boldsymbol{x})/k_{\mathrm{B}}T} \quad . \tag{II.39}$$

In the above, we have used the Einstein relation

$$D = \mu k_{\mathrm{B}}T = \frac{k_{\mathrm{B}}T}{\zeta} \quad , \tag{II.40}$$

which has led us to the Boltzmann weight corresponding to the *canonical ensemble*. So the Fokker-Planck equation has equilibrium statistical mechanics as a steady-state (the steady-state with vanishing current, $\boldsymbol{J} = 0$). Note that a stationary state does not necessarily correspond to equilibrium, since

we can have non-zero currents while maintaining the stationarity condition $\partial_t \mathcal{P}(\boldsymbol{x}, t) = 0$; we can also describe the behaviour of systems that are driven out of equilibrium.

2. Many-particle systems

We can easily generalize the above analysis to a system of N *non-interacting* particles. Given N particles at positions $\boldsymbol{r}^\alpha(t)$, $\alpha = 1, \ldots, N$, each undergoing Brownian motion in a potential $U(\boldsymbol{r})$, the Langevin equation for each particle reads

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{r}^\alpha(t) = \boldsymbol{v}(\boldsymbol{r}^\alpha(t)) + \boldsymbol{u}^\alpha(t) \quad , \tag{II.41}$$

where $\boldsymbol{v}(\boldsymbol{r}) = -\mu \nabla U(\boldsymbol{r})$. The random velocities are characterized by

$$\langle \boldsymbol{u}^\alpha(t) \rangle = 0 \quad ; \quad \left\langle u_i^\alpha(t) u_j^\beta(t') \right\rangle = 2D\, \delta^{\alpha\beta} \delta_{ij} \delta(t - t') \quad . \tag{II.42}$$

We can define a stochastic number density $\sum_{\alpha=1}^N \delta^3(\boldsymbol{x} - \boldsymbol{r}^\alpha(t))$ and a related concentration (number of particles per unit volume)

$$\mathcal{C}(\boldsymbol{x}, t) = \left\langle \sum_{\alpha=1}^N \delta^3(\boldsymbol{x} - \boldsymbol{r}^\alpha(t)) \right\rangle \quad , \quad \int \mathrm{d}^3\boldsymbol{x}\; \mathcal{C}(\boldsymbol{x}, t) = N \quad . \tag{II.43}$$

Following the analysis from Eq. (II.26) to Eq. (II.33), we obtain the following equation for the concentration

$$\partial_t \mathcal{C}(\boldsymbol{x}, t) + \boldsymbol{\nabla} \cdot [\boldsymbol{v}(\boldsymbol{x}) \mathcal{C}(\boldsymbol{x}, t) - D \boldsymbol{\nabla} \mathcal{C}(\boldsymbol{x}, t)] = 0 \quad , \tag{II.44}$$

which is our Smoluchowski equation. We find that the equilibrium solution is given by

$$\mathcal{C}_{\mathrm{eq}}(\boldsymbol{x}) = \frac{N}{\mathcal{Z}} \exp\left[-\frac{U(\boldsymbol{x})}{k_{\mathrm{B}} T} \right] \quad ; \quad \mathcal{Z} = \int \mathrm{d}^3\boldsymbol{x}\, e^{-U(\boldsymbol{x})/k_{\mathrm{B}}T} \quad , \tag{II.45}$$

as consistent with the canonical ensemble.

For *interacting* particles, the situation will be more involved as the equation cannot be written solely in terms of the single-particle density. Let us consider a system of N particles that experience the external potential $U(\boldsymbol{r})$, and interact with each other via the pair-wise potential $\mathcal{W}(\boldsymbol{r}, \boldsymbol{r}')$. The Langevin equation for the αth particle now includes the pairwise interaction (with all the other particles):

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{r}^\alpha(t) = \boldsymbol{v}^{(1)}(\boldsymbol{r}^\alpha(t)) + \sum_{\beta \neq \alpha} \boldsymbol{v}^{(2)}(\boldsymbol{r}^\alpha(t), \boldsymbol{r}^\beta(t)) + \boldsymbol{u}^\alpha(t) \quad , \tag{II.46}$$

where we have introduced the notation $\boldsymbol{v}^{(1)}(\boldsymbol{r}) = -\mu \boldsymbol{\nabla} U(\boldsymbol{r})$ and $\boldsymbol{v}^{(2)}(\boldsymbol{r}, \boldsymbol{r}') = -\mu \boldsymbol{\nabla} \mathcal{W}(\boldsymbol{r}, \boldsymbol{r}')$ to make a distinction between the one-particle and two-particle contributions to the drift velocity. Following the same procedure as above, we obtain the Smoluchowski equation

$$\partial_t \mathcal{C}(\boldsymbol{x}, t) + \boldsymbol{\nabla} \cdot \left[\mathcal{C}(\boldsymbol{x}, t) \boldsymbol{v}^{(1)}(\boldsymbol{x}) + \int \mathrm{d}^3\boldsymbol{x}'\; \mathcal{C}^{(2)}(\boldsymbol{x}, \boldsymbol{x}', t) \boldsymbol{v}^{(2)}(\boldsymbol{x}, \boldsymbol{x}') - D \boldsymbol{\nabla} \mathcal{C}(\boldsymbol{x}, t) \right] = 0 \quad , \tag{II.47}$$

where

$$\mathcal{C}^{(2)}(\boldsymbol{x},\boldsymbol{x}',t)=\left\langle\sum_{\alpha,\beta=1}^{N}\delta^3(\boldsymbol{x}-\boldsymbol{r}^{\alpha}(t))\,\delta^3(\boldsymbol{x}'-\boldsymbol{r}^{\beta}(t))\right\rangle \quad , \tag{II.48}$$

is the two-particle distribution function. This equation is not *closed*, as it involves a new quantity. When we seek the corresponding equation for $\mathcal{C}^{(2)}(\boldsymbol{x},\boldsymbol{x}',t)$, we will discover that it depends on the three-particle distribution function, and this *hierarchy* of coupled equations will continue until we reach the N-particle distribution. This is an example of the so-called Bogoliubov-Born-Green-Kirkwood-Yvon (BBGKY) hierarchy in nonequilibrium statistical physics.

A standard treatment involves a truncation of the hierarchy, which is formally equivalent to a mean-field approximation scheme. To see this, consider the following approximation

$$\mathcal{C}^{(2)}(\boldsymbol{x},\boldsymbol{x}',t)\simeq\left\langle\sum_{\alpha=1}^{N}\delta^3(\boldsymbol{x}-\boldsymbol{r}^{\alpha}(t))\right\rangle\left\langle\sum_{\beta=1}^{N}\delta^3(\boldsymbol{x}'-\boldsymbol{r}^{\beta}(t))\right\rangle=\mathcal{C}(\boldsymbol{x},t)\mathcal{C}(\boldsymbol{x}',t) \quad , \tag{II.49}$$

which essentially ignores the correlation between the two particles. This so-called *closure* relation gives the (unknown) pair-distribution function in terms of the concentration, which essentially closes the Smoluchowski equation as follows

$$\partial_t\mathcal{C}(\boldsymbol{x},t)+\boldsymbol{\nabla}\cdot\left[\mathcal{C}(\boldsymbol{x},t)\boldsymbol{v}^{(1)}(\boldsymbol{x})+\mathcal{C}(\boldsymbol{x},t)\int \mathrm{d}^3\boldsymbol{x}'\;\mathcal{C}(\boldsymbol{x}',t)\boldsymbol{v}^{(2)}(\boldsymbol{x},\boldsymbol{x}')-D\boldsymbol{\nabla}\mathcal{C}(\boldsymbol{x},t)\right]=0 \quad . \tag{II.50}$$

The above equation is *nonlinear*, and in general much more difficult to solve than the equation for non-interacting particles.

3. Backward Fokker-Planck equation

The procedure developed above for the transition from trajectory to probability can be used to derive a Fokker-Planck equation for the backward time evolution of the initial condition. The resulting backward Fokker-Planck equation will be a very useful tool, e.g. for calculating the mean first passage time.

The aim is to derive a governing time evolution for the probability of finding a particle at position $\boldsymbol{x}$ at time t given the initial position of $\boldsymbol{x}_0$ at time t_0, denoted and defined as $\mathcal{P}(\boldsymbol{x},t|\boldsymbol{x}_0,t_0)=\left\langle\delta^3\left(\boldsymbol{x}-\boldsymbol{r}(t)\right)\right\rangle\big|_{\boldsymbol{r}(t_0)=\boldsymbol{x}_0}$. The trajectory can be formally represented through an integral of Eq. (II.25) over the entire time domain

$$\boldsymbol{r}(t)=\boldsymbol{x}_0+\int_{t_0}^{t}\mathrm{d}t_1\left[\boldsymbol{v}(\boldsymbol{r}(t_1))+\boldsymbol{u}(t_1)\right] \quad , \tag{II.51}$$

which can be used to calculate the infinitesimal change in initial position when the initial time is changed from t_0 to $t_0-\Delta t_0$. We obtain

$$\begin{aligned}\mathcal{P}(\boldsymbol{x},t|\boldsymbol{x}_0,t_0-\Delta t_0)-\mathcal{P}(\boldsymbol{x},t|\boldsymbol{x}_0,t_0) &= \left\langle\Delta r_i(t_0)\,\partial_{0i}\delta^3\left(\boldsymbol{x}-\boldsymbol{r}(t)\right)\right\rangle+\frac{1}{2}\left\langle\Delta r_i(t_0)\Delta r_j(t_0)\,\partial_{0i}\partial_{0j}\delta^3\left(\boldsymbol{x}-\boldsymbol{r}(t)\right)\right\rangle\\ &= \left\langle\Delta r_i(t_0)\right\rangle\,\partial_{0i}\mathcal{P}(\boldsymbol{x},t|\boldsymbol{x}_0,t_0)+\frac{1}{2}\left\langle\Delta r_i(t_0)\Delta r_j(t_0)\right\rangle\,\partial_{0i}\partial_{0j}\mathcal{P}(\boldsymbol{x},t|\boldsymbol{x}_0,t_0) \quad ,\end{aligned}$$

where

$$\Delta\boldsymbol{r}(t_0)\equiv\boldsymbol{r}(t_0)-\boldsymbol{r}(t_0-\Delta t_0)=\int_{t_0-\Delta t_0}^{t_0}\mathrm{d}t_1\left[\boldsymbol{v}(\boldsymbol{r}(t_1))+\boldsymbol{u}(t_1)\right] \quad . \tag{II.53}$$

Inserting the averages from Eqs. (II.30) and II.32, we find

$$-\partial_{t_0}\mathcal{P}(\boldsymbol{x},t|\boldsymbol{x}_0,t_0) = \boldsymbol{v}(\boldsymbol{x}_0)\cdot\boldsymbol{\nabla}_0\mathcal{P}(\boldsymbol{x},t|\boldsymbol{x}_0,t_0) + D\,\boldsymbol{\nabla}_0^2\mathcal{P}(\boldsymbol{x},t|\boldsymbol{x}_0,t_0) \quad , \tag{II.54}$$

Note the differences between the backward equation above and Eq. (II.33).

C. Stochastic dynamics of orientation

We have so far considered translational stochastic motion where the main dynamical variables were coordinates describing the position of a particle in space. We now consider the dynamics of orientation, which can be represented by a unit vector $\hat{\boldsymbol{n}}$. The constraint $\hat{\boldsymbol{n}}^2 = 1$ defines a spherical surface (in any dimension d), and thus the Brownian motion of orientation can be regarded as a random walk that takes place on the surface of a sphere. Assigning a diffusion coefficient D_r (henceforth referred to as the rotational diffusion coefficient) to each remaining angular degree of freedom, the infinitesimal MSD for displacements in orientation can then be written as

$$\left\langle \Delta\hat{\boldsymbol{n}}^2 \right\rangle = \left\langle [\hat{\boldsymbol{n}}(t+\Delta t) - \hat{\boldsymbol{n}}(t)]^2 \right\rangle = 2(d-1)D_r\Delta t \quad , \tag{II.55}$$

for free particles.

Now, consider a particle that undergoes orientational Brownian motion subject to the potential $U(\hat{\boldsymbol{n}})$. Due to the orientational potential, the particle will experience a torque $\boldsymbol{T}$, which can be calculated as follows. In 3D, a small (vectorial) angular displacement $\delta\boldsymbol{\theta}$ will lead to the rotation of the orientation from $\hat{\boldsymbol{n}}$ to $\hat{\boldsymbol{n}} + \delta\boldsymbol{\theta}\times\hat{\boldsymbol{n}}$ and an accompanying change in the potential energy

$$U(\hat{\boldsymbol{n}} + \delta\boldsymbol{\theta}\times\hat{\boldsymbol{n}}) - U(\hat{\boldsymbol{n}}) = (\delta\boldsymbol{\theta}\times\hat{\boldsymbol{n}})\cdot\frac{\partial}{\partial\hat{\boldsymbol{n}}}U(\hat{\boldsymbol{n}}) = \delta\boldsymbol{\theta}\cdot\left(\hat{\boldsymbol{n}}\times\frac{\partial}{\partial\hat{\boldsymbol{n}}}U(\hat{\boldsymbol{n}})\right) \quad , \tag{II.56}$$

which is to be provided by the work done on the system $-\boldsymbol{T}\cdot\delta\boldsymbol{\theta}$. Therefore, the torque is given as

$$\boldsymbol{T} = -\boldsymbol{\mathcal{R}}U(\hat{\boldsymbol{n}}) \quad , \tag{II.57}$$

where

$$\boldsymbol{\mathcal{R}} \equiv \hat{\boldsymbol{n}}\times\frac{\partial}{\partial\hat{\boldsymbol{n}}} \quad , \tag{II.58}$$

is an orientational gradient operator. The operator $\boldsymbol{\mathcal{R}}$ inherits properties of differential operators; for example $\int \mathrm{d}\hat{\boldsymbol{n}}\ [\boldsymbol{\mathcal{R}}A(\hat{\boldsymbol{n}})]\,B(\hat{\boldsymbol{n}}) = -\int \mathrm{d}\hat{\boldsymbol{n}}\ A(\hat{\boldsymbol{n}})\,[\boldsymbol{\mathcal{R}}B(\hat{\boldsymbol{n}})]$. Moreover, using the definition one can show that $\mathcal{R}_i\hat{n}_j = -\epsilon_{ijk}\,\hat{n}_k$ holds, and consequently, we also have $\boldsymbol{\mathcal{R}}^2\hat{\boldsymbol{n}} = -2\hat{\boldsymbol{n}}$.

We can think of the differential operator $\boldsymbol{\mathcal{R}}$ as the generator of rotation, or the angular momentum in a quantum mechanical analogy, as the definition [Eq. (II.58)] suggests. Knowing that rotation operations do not commute, we can use the operator to construct the canonical form for any Taylor expansion as

$$f(\hat{\boldsymbol{n}} + \delta\boldsymbol{\theta}\times\hat{\boldsymbol{n}}) = e^{\delta\theta_i\mathcal{R}_i}f(\hat{\boldsymbol{n}}) = f(\hat{\boldsymbol{n}}) + \delta\theta_i\,\mathcal{R}_i f(\hat{\boldsymbol{n}}) + \frac{1}{2}\,\delta\theta_i\delta\theta_j\,\mathcal{R}_i\mathcal{R}_j f(\hat{\boldsymbol{n}}) + \cdots \quad , \tag{II.59}$$

corresponding to finite rotations.

Assuming dissipative dynamics, the torque will lead to an angular velocity $\boldsymbol{\omega} = \boldsymbol{T}/\zeta_r$, where ζ_r is the rotational friction coefficient, which depends on the geometric characteristics of the particle. For example, for spherical particle of radius a moving in a fluid of viscosity η, we have $\zeta_r = 8\pi\eta a^3$. In analogy to the translational case, we can define a mobility $\mu_r = 1/\zeta_r$. The equation for the angular velocity now reads $\boldsymbol{\omega} = -\mu_r\boldsymbol{\mathcal{R}}U(\hat{\boldsymbol{n}})$.

The Langevin equation for the orientation can be constructed as

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\boldsymbol{n}}(t) = \boldsymbol{\omega} \times \hat{\boldsymbol{n}} + \boldsymbol{\eta} \times \hat{\boldsymbol{n}} \quad , \tag{II.60}$$

where the random angular velocity $\boldsymbol{\eta}$ represents a Gaussian white noise with the variance of $2D_r$ for each component. Note that the form of Eq. (II.60) preserves $\hat{\boldsymbol{n}}^2 = 1$ by construction, and ensures that only two independent noise components drive the stochastic angular displacements.

We can formulate a probabilistic description for the trajectory-averaged probability distribution

$$\mathcal{P}(\hat{\boldsymbol{n}}, t) = \langle \delta\left(\hat{\boldsymbol{n}} - \hat{\boldsymbol{n}}(t)\right) \rangle \quad , \tag{II.61}$$

which is normalized by construction: $\int \mathrm{d}\hat{\boldsymbol{n}}\, \mathcal{P}(\hat{\boldsymbol{n}}, t) = 1$. To construct the Fokker-Planck equation, we need to consider a finite rotation over a small time interval Δt. We can integrate Eq. (II.60) over the time interval Δt to obtain

$$\Delta\hat{\boldsymbol{n}}(t) \equiv \hat{\boldsymbol{n}}(t + \Delta t) - \hat{\boldsymbol{n}}(t) = \int_t^{t+\Delta t} \mathrm{d}t_1\, \boldsymbol{\omega}(\hat{\boldsymbol{n}}(t_1)) \times \hat{\boldsymbol{n}}(t_1) + \int_t^{t+\Delta t} \mathrm{d}t_1\, \boldsymbol{\eta}(t_1) \times \hat{\boldsymbol{n}}(t_1) \quad , \tag{II.62}$$

which contains the director at different times. The corresponding expression in terms of the rotation angle is

$$\Delta\boldsymbol{\theta}(t) \simeq \int_t^{t+\Delta t} \mathrm{d}t_1\, \boldsymbol{\omega}(\hat{\boldsymbol{n}}(t_1)) + \int_t^{t+\Delta t} \mathrm{d}t_1\, \boldsymbol{\eta}(t_1) \quad . \tag{II.63}$$

The next step is to perform a Taylor expansion, according to Eq. (II.59):

$$\delta\left(\hat{\boldsymbol{n}} - \hat{\boldsymbol{n}}(t + \Delta t)\right) = \delta\left(\hat{\boldsymbol{n}} - \hat{\boldsymbol{n}}(t)\right) - \Delta\theta_i(t)\, \mathcal{R}_i \delta\left(\hat{\boldsymbol{n}} - \hat{\boldsymbol{n}}(t)\right) + \frac{1}{2}\, \Delta\theta_i(t)\Delta\theta_j(t)\, \mathcal{R}_i\mathcal{R}_j \delta\left(\hat{\boldsymbol{n}} - \hat{\boldsymbol{n}}(t)\right) \quad , \tag{II.64}$$

and neglect the $O(\Delta\boldsymbol{\theta}^3)$ corrections.

Averaging over the noise distribution, we obtain

$$\mathcal{P}(\hat{\boldsymbol{n}}, t + \Delta t) - \mathcal{P}(\hat{\boldsymbol{n}}, t) = -\mathcal{R}_i \left\langle \Delta\theta_i(t)\, \delta\left(\hat{\boldsymbol{n}} - \hat{\boldsymbol{n}}(t)\right) \right\rangle + \frac{1}{2}\, \mathcal{R}_i\mathcal{R}_j \left\langle \Delta\theta_i(t)\Delta\theta_j(t)\, \delta\left(\hat{\boldsymbol{n}} - \hat{\boldsymbol{n}}(t)\right) \right\rangle \quad . \tag{II.65}$$

We now expand the left hand side of Eq. (II.65) in Δt and keep the leading order, which gives $\mathcal{P}(\hat{\boldsymbol{n}}, t + \Delta t) - \mathcal{P}(\hat{\boldsymbol{n}}, t) \simeq \Delta t\, \partial_t \mathcal{P}(\hat{\boldsymbol{n}}, t)$, and use Eq. (II.62) to obtain the following expressions

$$\left\langle \Delta\theta_i(t)\, \delta\left(\hat{\boldsymbol{n}} - \hat{\boldsymbol{n}}(t)\right) \right\rangle \;=\; \left\langle \Delta\theta_i(t) \right\rangle \left\langle \delta\left(\hat{\boldsymbol{n}} - \hat{\boldsymbol{n}}(t)\right) \right\rangle = \left\langle \Delta\theta_i(t) \right\rangle \mathcal{P}(\hat{\boldsymbol{n}}, t) \quad , \tag{II.66}$$

$$\left\langle \Delta\theta_i(t) \right\rangle \;\simeq\; \omega_i(\hat{\boldsymbol{n}})\, \Delta t \quad , \tag{II.67}$$

and

$$\left\langle \Delta\theta_i(t)\Delta\theta_j(t)\, \delta\left(\hat{\boldsymbol{n}} - \hat{\boldsymbol{n}}(t)\right) \right\rangle \;=\; \left\langle \Delta\theta_i(t)\Delta\theta_j(t) \right\rangle \left\langle \delta\left(\hat{\boldsymbol{n}} - \hat{\boldsymbol{n}}(t)\right) \right\rangle = \left\langle \Delta\theta_i(t)\Delta\theta_j(t) \right\rangle \mathcal{P}(\hat{\boldsymbol{n}}, t) \quad , \tag{II.68}$$

$$\left\langle \Delta\theta_i(t)\Delta\theta_j(t) \right\rangle \;\simeq\; 2D_r\, \delta_{ij}\, \Delta t \quad , \tag{II.69}$$

that are correct to the leading order in Δt. Putting them altogether, we obtain the following equation

$$\partial_t \mathcal{P}(\hat{\boldsymbol{n}}, t) + \boldsymbol{\mathcal{R}} \cdot \left[\boldsymbol{\omega}(\hat{\boldsymbol{n}})\, \mathcal{P}(\hat{\boldsymbol{n}}, t) - D_r\, \boldsymbol{\mathcal{R}} \mathcal{P}(\hat{\boldsymbol{n}}, t)\right] = 0 \quad , \tag{II.70}$$

when we take the limit $\Delta t \to 0$. This is the orientational Fokker-Planck equation, which is a statement

of the conservation of probability in the orientation space.

1. Equilibrium distribution

Putting $\boldsymbol{\omega}(\hat{\boldsymbol{n}}) = -\mu_r \boldsymbol{\mathcal{R}} U(\hat{\boldsymbol{n}})$ in the Fokker-Planck equation, it reads

$$\partial_t \mathcal{P}(\hat{\boldsymbol{n}}, t) + \boldsymbol{\mathcal{R}} \cdot \left[\left(-\mu_r \boldsymbol{\mathcal{R}} U(\hat{\boldsymbol{n}}) \right) \mathcal{P}(\hat{\boldsymbol{n}}, t) - D_r \, \boldsymbol{\mathcal{R}} \mathcal{P}(\hat{\boldsymbol{n}}, t) \right] = 0 \quad . \tag{II.71}$$

Setting the probability flux to zero

$$-\mu_r \boldsymbol{\mathcal{R}} U(\hat{\boldsymbol{n}}) \, \mathcal{P}(\hat{\boldsymbol{n}}, t) - D_r \, \boldsymbol{\mathcal{R}} \mathcal{P}(\hat{\boldsymbol{n}}, t) = 0 \quad , \tag{II.72}$$

gives the equilibrium distribution

$$\mathcal{P}_{\mathrm{eq}}(\hat{\boldsymbol{n}}) = \frac{1}{\mathcal{Z}} \exp\left[-\frac{\mu_r U(\hat{\boldsymbol{n}})}{D_r} \right] \quad , \quad \mathcal{Z} = \int \mathrm{d}\hat{\boldsymbol{n}} \, e^{-\mu_r U(\hat{\boldsymbol{n}})/D_r} \quad , \tag{II.73}$$

which leads to the orientational Einstein relation

$$D_r = \mu_r k_{\mathrm{B}} T = \frac{k_{\mathrm{B}} T}{\zeta_r} \quad . \tag{II.74}$$

2. Free-particle dynamics

Let us consider the Fokker-Planck equation in the absence of an external potential (or angular velocity):

$$\partial_t \mathcal{P}(\hat{\boldsymbol{n}}, t | \hat{\boldsymbol{n}}_0, 0) = D_r \, \boldsymbol{\mathcal{R}}^2 \mathcal{P}(\hat{\boldsymbol{n}}, t | \hat{\boldsymbol{n}}_0, 0) \quad , \tag{II.75}$$

where we have made explicit reference to the initial condition, i.e. $\mathcal{P}(\hat{\boldsymbol{n}}, 0 | \hat{\boldsymbol{n}}_0, 0) = \delta(\hat{\boldsymbol{n}} - \hat{\boldsymbol{n}}_0)$. We aim to calculate the orientation auto-correlation function $\langle \hat{\boldsymbol{n}}(t) \cdot \hat{\boldsymbol{n}}(0) \rangle$, which can be defined as follows due to the Markov property of the Brownian motion

$$\langle \hat{\boldsymbol{n}}(t) \cdot \hat{\boldsymbol{n}}(0) \rangle = \int \mathrm{d}\hat{\boldsymbol{n}} \mathrm{d}\hat{\boldsymbol{n}}' \; \hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{n}}' \, \mathcal{P}(\hat{\boldsymbol{n}}, t | \hat{\boldsymbol{n}}', 0) \, \mathcal{P}_{\mathrm{eq}}(\hat{\boldsymbol{n}}') \quad . \tag{II.76}$$

We have used $\mathcal{P}_{\mathrm{eq}}(\hat{\boldsymbol{n}}') = 1/(4\pi)$ to ensure that the auto-correlation function applies to any randomly selected initial state of the system and does not require special preparations.

Applying the time derivative and invoking Eq. (II.75), we find

$$\begin{aligned}
\partial_t \langle \hat{\boldsymbol{n}}(t) \cdot \hat{\boldsymbol{n}}(0) \rangle &= \int \mathrm{d}\hat{\boldsymbol{n}} \mathrm{d}\hat{\boldsymbol{n}}' \; \hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{n}}' \, \partial_t \mathcal{P}(\hat{\boldsymbol{n}}, t | \hat{\boldsymbol{n}}', 0) \, \mathcal{P}_{\mathrm{eq}}(\hat{\boldsymbol{n}}') \quad , \\
&= D_r \int \mathrm{d}\hat{\boldsymbol{n}} \mathrm{d}\hat{\boldsymbol{n}}' \; \hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{n}}' \, \boldsymbol{\mathcal{R}}^2 \mathcal{P}(\hat{\boldsymbol{n}}, t | \hat{\boldsymbol{n}}', 0) \, \mathcal{P}_{\mathrm{eq}}(\hat{\boldsymbol{n}}') \quad , \\
&= -2 D_r \int \mathrm{d}\hat{\boldsymbol{n}} \mathrm{d}\hat{\boldsymbol{n}}' \; \hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{n}}' \, \mathcal{P}(\hat{\boldsymbol{n}}, t | \hat{\boldsymbol{n}}', 0) \, \mathcal{P}_{\mathrm{eq}}(\hat{\boldsymbol{n}}') \quad , \\
&= -2 D_r \langle \hat{\boldsymbol{n}}(t) \cdot \hat{\boldsymbol{n}}(0) \rangle \quad ,
\end{aligned} \tag{II.77}$$

where we have also used the properties of $\boldsymbol{\mathcal{R}}$ discussed above. Solving this subject to the appropriate initial condition yields

$$\langle \hat{\boldsymbol{n}}(t) \cdot \hat{\boldsymbol{n}}(0) \rangle = e^{-2 D_r t} \quad . \tag{II.78}$$

We thus find that the particle loses its orientational *memory* with an exponential decay that is controlled by the time scale $1/(2D_r)$. This result can be rewritten in the form

$$\left\langle [\hat{\boldsymbol{n}}(t) - \hat{\boldsymbol{n}}(0)]^2 \right\rangle = 2\left[1 - e^{-2D_r t}\right] \quad , \tag{II.79}$$

which gives the small-time asymptotic form of $\left\langle [\hat{\boldsymbol{n}}(t) - \hat{\boldsymbol{n}}(0)]^2 \right\rangle \simeq 4D_r t$ in agreement with Eq. (II.55). Interestingly, Eq. (II.79) shows a saturation in the long time limit, which is a manifestation of the compactness of the space that is available to the Brownian motion.