Lecture Notes on Stochastic Processes

PD Dr. Benjamin M. Friedrich

April 17, 2019

Contents

1 Diffusion & Random Walk

1.1 Random Walker

The random walk can be used to model a variety of different phenomena just like

- the motion of a particle during diffusion
- the spread of mosquito infestation in a forest
- propagation of sound waves in a heterogeneous material
- money flow

Model: Random Walker

A random walker can be considered as particle moving in steps of length l, while choosing each time a random, uncorrelated direction. Uncorrelated means that

$$\langle \vec{x}_n \cdot \vec{x}_m \rangle = l^2 \delta_{nm} \tag{1.1}$$

for averaging over a certain probability distribution.

Thus, the displacement of a random walker after n steps is given by

$$\vec{x} = \sum_{n=1}^{N} \vec{x}_n \quad \text{with} \quad \langle \vec{x} \rangle = \vec{0} \tag{1.2}$$

The mean square displacement $(\Delta\vec{x})^2$ equals the variance σ^2

$$\begin{aligned} \sigma^2 &= \left\langle \vec{x}^2 \right\rangle - \langle \vec{x} \rangle^2 = \left\langle \vec{x}^2 \right\rangle = (\Delta\vec{x})^2 \\ &= \left\langle \left(\sum_{n=1}^{N} \vec{x}_n \right)^2 \right\rangle = \sum_{n,m=1}^{N} \langle \vec{x}_n \cdot \vec{x}_m \rangle = N l^2 \end{aligned}$$

As we have $(\Delta\vec{x})^2 \sim N$ and $\Delta t \sim N$, the relation $\frac{(\Delta\vec{x})^2}{\Delta t}$ is a constant in the continuum limit, which is quite unusual that a square term in the numerator appears.

1.2 Continuum Limit: Diffusion Equation

We now consider the step sizes $\Delta\vec{y}$ of a random walker becoming infinitesimally small, with $p(\Delta\vec{y})$ being the probability for step $\Delta\vec{y}$:

$$\langle \Delta\vec{y}_i \rangle = \int \mathrm{d}^d\Delta y \, [\Delta y_i \, p(\Delta\vec{y})] = 0 \tag{1.3}$$

$$\langle \Delta\vec{y}_i \, \Delta\vec{y}_j \rangle = \int \mathrm{d}^d\Delta y \, [\Delta y_i \, \Delta y_j \, p(\Delta\vec{y})] = \left\langle (\Delta\vec{y})^2 \right\rangle \frac{\delta_{ij}}{d} \tag{1.4}$$

for $i, j = 1, 2, \ldots, d$ vector components.

We can express the probability for a displacement of $\vec{x}$ after N steps $p_N(\vec{x})$ through the elementary relation

$$P_N(\vec{x}) = \int \mathrm{d}^d\Delta y \, P_{N-1}(\vec{x} - \Delta\vec{y}) P(\Delta\vec{y}) \tag{1.5}$$

Now we do a Taylor expansion of $P_N(\vec{x})$

$$\begin{aligned} P_N(\vec{x}) &\approx \int \mathrm{d}^d\Delta y \, P(\Delta\vec{y}) \left[P_{N-1}(\vec{x}) - \Delta y_i \partial_i P_{N-1}(\vec{x}) + \frac{1}{2} \Delta y_i \Delta y_j \partial_i \partial_j P_{N-1}(\vec{x}) \right] \\ &= P_{N-1}(\vec{x}) + \frac{\langle (\Delta\vec{y})^2 \rangle}{2d} \vec{\nabla}^2 P_{N-1}(\vec{x}) \end{aligned}$$

We can define a continuum probability density $p(\vec{x}, t)$ after the time $t = N\Delta t$

$$p(\vec{x}, t) = p(\vec{x}, N\Delta t) := P_N(\vec{x}) \tag{1.6}$$

and now we can take the limit

$$\frac{\partial p}{\partial t} = \lim_{\Delta t \to 0} \frac{P_N(\vec{x}) - P_{N-1}(\vec{x})}{\Delta t} = D\vec{\nabla}^2 p \quad \text{with} \quad D = \frac{\langle (\Delta\vec{y})^2 \rangle}{2d\Delta t} \tag{1.7}$$

This continuum limit exists, if D can be treated as a constant, i.e. if $\frac{\left\langle (\Delta\vec{y})^2 \right\rangle}{\Delta t}$ is finite for $\Delta t \to 0$. The resulting equation is known as the diffusion equation.

Diffusion Equation

The diffusion equation for a probability density $p(\vec{x}, t)$ reads

$$\frac{\partial p(\vec{x}, t)}{\partial t} = D\vec{\nabla}^2 p(\vec{x}, t) \tag{1.8}$$

which we can also rewrite with the the definition of a current $\vec{J} = -D\vec{\nabla}p$

$$\frac{\partial p}{\partial t} = -\vec{\nabla} \cdot \vec{J} \tag{1.9}$$

Solving the diffusion equation

The diffusion equation can be solved e.g. by doing a Fourier transformation of both sides

$$\frac{\partial p}{\partial t} = D\vec{k}^2 p \tag{1.10}$$

leading to the k-space solution

$$p(\vec{k}, t) = \mathcal{F}(p(\vec{x}, t)) = \exp\left(-Dk^2 t\right) \tag{1.11}$$

A Fourier transform backwards gives the fundamental solution (Green's function)

$$p(\vec{x}, t) = \frac{1}{(4\pi kt)^{d/2}} \exp\left(-\frac{x^2}{4kt}\right) \tag{1.12}$$

with the mean square spread $\sigma^2 \sim kt$

1.3 Random Force Model

Another way to approach diffusion is by considering a colloidal particle suspended in a fluid experiencing random forces $f(t)$ due to the interaction with the fluid molecules.

Model: Random Forces

The motion of a particle under random forces $f(t)$ in one dimension can be described by Newtons law

$$m\ddot{x} + \gamma\dot{x} = f(t) \tag{1.13}$$

For long time scales $\tau_m >> \frac{m}{\gamma}$, inertia is negligible and we just have $\gamma\dot{x} = f(t)$. The random force is characterized through

- $\langle f(t) \rangle = 0$ (by symmetry)
- a vanishing correlation $\langle f(t)f(t+\tau) \rangle \to 0$ for $\tau \to \tau_m$

for averaging over a certain probability distribution.

An important property of the random force is stationarity

$$\frac{1}{\gamma^2} \int_{-\infty}^{\infty} \mathrm{d}\tau \, \langle f(t)f(t+\tau) \rangle = 2D \tag{1.14}$$

with $[D] = \mathrm{m}^2\,\mathrm{s}^{-1}$

Concept: Diffusion

Diffusion is a net movement of particles from a region of high to a region of low concentration due to random motion of the single particles.

Formal Solution

A formal solution to the equation of motion without taking inertia into account reads

$$x(t) = x(0) + \frac{1}{\gamma} \int_0^t \mathrm{d}t_1 \, f(t_1) \tag{1.15}$$

For $\Delta x = x(t) - x(0)$ we get for the mean deviation

$$\langle \Delta x(t) \rangle = \frac{1}{\gamma} \int_0^t \mathrm{d}t_1 \, \langle f(t_1) \rangle = 0 \tag{1.16}$$

by symmetry and for the mean square deviation

$$\begin{aligned}
\langle \Delta x^2 \rangle &= \frac{1}{\gamma^2} \left\langle \left(\int_0^t \mathrm{d}t_1 \, f(t_1) \right) \left(\int_0^t \mathrm{d}t_2 \, f(t_2) \right) \right\rangle \\
&= \frac{1}{\gamma^2} \int_0^t \mathrm{d}t_1 \int_0^t \mathrm{d}t_2 \, \langle f(t_1) f(t_2) \rangle \\
&= \frac{1}{\gamma^2} \int_0^t \mathrm{d}t_1 \int_{-t_1}^{t+t_1} \mathrm{d}\tau \, \langle f(t_1) f(t_1 + \tau) \rangle \\
&= \frac{1}{\gamma^2} \int_0^t \mathrm{d}t_1 \int_{-\infty}^{\infty} \mathrm{d}\tau \, \langle f(t_1) f(t_1 + \tau) \rangle + \mathcal{O}(D\tau_m) \\
&= \frac{1}{\gamma^2} \int_0^t \mathrm{d}t_1 \, \gamma^2 2D = 2Dt
\end{aligned}$$

Calculating $D = D(T)$

In order to calculate $D(T)$ we do a trick and add an elastic spring to the model

$$kx + \gamma \dot{x} = f(t) \tag{1.17}$$

So at first we might ask what happens in reaction to a pulse response?

$$kx + \gamma \dot{x} = \rho_0 \delta(t) \quad \text{with} \quad x(t) = 0 \,|\, t < 0 \tag{1.18}$$

The solution to this scenario is given by

$$x(t) = \rho_0 \chi(t), \quad \chi(t) = \frac{1}{\gamma} \exp\left(-\frac{t}{\sigma} \right) \Theta(t), \quad \sigma = \frac{\gamma}{k} \tag{1.19}$$

We get back to the full problem, where the formal solutions reads

$$x(t) = \int_0^\infty \mathrm{d}\tau \, f(t - \tau) \chi(\tau) \tag{1.20}$$

with $\langle x(t)\rangle = 0$ by symmetry and

$$
\begin{aligned}
\left\langle \Delta x^2 \right\rangle &= \frac{1}{\gamma^2}\left\langle \left(\int_0^\infty \mathrm{d}\tau_1\, f(t-\tau_1)\chi(\tau_1)\right)\left(\int_0^\infty \mathrm{d}\tau_2\, f(t-\tau_2)\chi(\tau_2)\right)\right\rangle \\
&= \int_0^\infty \mathrm{d}\tau_1 \int_0^\infty \mathrm{d}\tau_2 \left\langle f(t-\tau_1)f(t-\tau_2)\right\rangle \underbrace{\chi(\tau_1)\chi(\tau_2)}_{=\frac{1}{\gamma^2}\exp\left(-\frac{\tau_1+\tau_2}{\sigma}\right)} \\
&= \int_0^\infty \mathrm{d}\tau_1 \int_{-\tau_1}^\infty \mathrm{d}\tau \left\langle f(t-\tau_1)f(t-\tau_1-\tau)\right\rangle \frac{1}{\gamma^2}\exp\left(-\frac{2\tau_1+\tau}{\sigma}\right) \\
&= \int_0^\infty \mathrm{d}\tau_1 \int_{-\infty}^\infty \mathrm{d}\tau \left\langle f(t-\tau_1)f(t-\tau_1-\tau)\right\rangle \frac{1}{\gamma^2}\exp\left(-\frac{2\tau_1+\tau}{\sigma}\right) + \mathcal{O}(D\tau_m) \\
&= \frac{1}{\gamma^2}\int_0^\infty \mathrm{d}\tau_1 \exp\left(-\frac{2\tau_1}{\sigma}\right) \int_{-\infty}^\infty \mathrm{d}\tau \left\langle f(t-\tau_1)f(t-\tau_1-\tau)\right\rangle \underbrace{\exp\left(-\frac{\tau}{\sigma}\right)}_{\approx 1,\, \tau_m << \sigma} \\
&= \frac{1}{\gamma^2}\frac{\sigma}{2}2D\gamma^2 = \frac{\gamma}{k}D
\end{aligned}
$$

At this point we would like to make use of the equipartition theorem

$$
\left\langle \frac{k}{2}x^2 \right\rangle = \frac{k_B T}{2} \tag{1.21}
$$

As we have

$$
\left\langle \frac{k}{2}x^2 \right\rangle = \frac{k}{2}\frac{\gamma}{k}D = \frac{k_B T}{2} \tag{1.22}
$$

we obtain the Stokes-Einstein-relation

$$
D = \frac{k_B T}{\gamma} \tag{1.23}
$$

Repetition: Equipartition Theorem

In thermal equilibrium, the systems energy, given by a Hamiltonian H, is distributed on its degrees of freedom x_n via

$$\left\langle x_m \frac{\partial H}{\partial x_n} \right\rangle = \delta_{mn} k_B T \tag{1.24}$$

This holds for a microcanonical and canonical ensemble and relates temperature to the systems average energies.

2 Sort Review of Probability Theory

2.1 Mathematical Foundations

Concept: Probability

Let X be a set of states, then we have the following axioms of probability:

- $P(A)$ is a function with $0 \leq P(A) \leq 1$ defined for some subset $A \subseteq X$
- Additivity: $P(A) + P(B) = P(A \cup B) - P(A \cap B)$

If $X = \mathbb{R}$, we can consider a probability density as a function $p(x) : \mathbb{R} \to \mathbb{R}^+$, which relates to the probability by

$$P(A) = \int_A \mathrm{d}x \, p(x) \tag{2.1}$$

for $A \subseteq \mathbb{R}$. If $[x] = \mathrm{m}$, then $[p] = \mathrm{m}^{-1}$. Note that $p(x)$ is also called probability density function $\mathrm{PDF}(x) = p(x)$. The cumulative density function is given by

$$\mathrm{CDF}(x) = \int_{-\infty}^{x} \mathrm{d}x' \, p(x') \tag{2.2}$$

Important characteristics of probability distributions are its moments and its cumulants.

Moments

The moments of a probability distributions $p(x)$ are given by

$$\mu_n = \langle x^n \rangle = \int_{-\infty}^{\infty} x^n p(x) \tag{2.3}$$

with the characteristic function

$$\langle \exp(tx) \rangle = \sum_{n=0}^{\infty} \mu_n \frac{t^n}{n!} \tag{2.4}$$

Cumulants

The cumulants are given by the mean value $k_1 = \mu_1$, the variance $k_2 = \mu_2 - \mu_1^2$ and higher order cumulants such as $k_3 = \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3$. More generally we have

$$\ln \langle \exp(tx) \rangle = \sum_{n=0}^{\infty} k_n \frac{t^n}{n!} \tag{2.5}$$

2.2 Probability in Physics

Usually, probability is regarded as relative frequency of an event A occurring N_A times while the total number of measurements is N

$$P(A) = \lim_{N \to \infty} \frac{N_A}{N} \tag{2.6}$$

Practically, a probability is determined by

- the experiment being repeated very often with the same initial macrostate
- replacing the physical system by an idealized model for stochastic simulations

(Talk by Jan Nagel: Gott würfelt nicht. Oder doch? –> Uncertainty in initial conditions leads to a dice producing a stochastic behavior.)

Example Weather Forecast

For an event R = "rain tomorrow" we know that it is raining 116 out of 365 days in Dresden: $P(R \,|\, \mathrm{Dresden}) = \frac{116}{365} = 18\%$. Our forecast is getting more accurate if we consider also seasonal changes and thus specific the month being October with 8 days of rain out of 31 in total: $P(R \,|\, \mathrm{Dresden, October}) = \frac{8}{31} = 25.8\%$

Another approach is based on persistence of conditions, i.e. to make a rain prediction for tomorrow based on the weather today. E.g. according to Caskey 1963 we have $P(R \,|\, \mathrm{current\,local\,weather}) = x$ and $P(R \,|\, \mathrm{rain\,today}) = 44\%$, $P(R \,|\, \mathrm{dry\,today}) = 17\%$

Last but not least we can sample macrostate that is consistent with measurement data an calculate the probabilities for rain from deterministic models (Navier-Stokes-equations / mathematical forecasting) $P(R \,|\, \mathrm{current\,global\,weather})$.

2.3 Important Probability Distributions

Normal Distribution

The normal distribution $N(\mu, \sigma^2)$ with mean μ and variance σ^2 is given by

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] = N(\mu, \sigma^2) \tag{2.7}$$

The moments and cumulants read

$$\mu_1 = \mu,\ \mu_2 = \mu^2 + \sigma^2,\ \mu_3 = \mu^3 + 3\mu\sigma^2$$
$$k_1 = \mu,\ k_2 = \sigma^2,\ k_j = 0 \quad \text{for} \quad j \geq 3$$

Bernoulli Distribution

For a Bernoulli trial you have two outcomes with respective probabilities p and $1-p$. If you now perform n independent trials you will get k times the first outcome with probability

$$P(k, n) = \binom{n}{k} p^k (1-p)^{n-k} \tag{2.8}$$

with $\langle k \rangle = np$ and $\langle k^2 \rangle - \langle k \rangle^2 = np(1-p)$. In many practical cases one can do a normal approximation by $P(k, n) = N(np, np(1-p))$.

Poisson Distribution

We consider the continuous time limit of the Binomial distribution. So we introduce discrete central times $t_j = \frac{j}{n}T = j\,\mathrm{d}t$ with $\mathrm{d}t = \frac{T}{n}$ and $\lambda = np$ being the total number of expected events. And we assume that at each time step one Bernoulli trial is conducted. The event rate is given by $r = \frac{\lambda}{T} = \frac{p}{\mathrm{d}t}$ with $[r] = \mathrm{s}^{-1}$ in units of an inverse time.

Now take the limit $n \to \infty$ with $\lambda = \text{const}$ and $p = \frac{\lambda}{n}$, so we get the Poisson distribution

$$p(k, \lambda) = \exp(-\lambda) \frac{\lambda^k}{k!} \tag{2.9}$$

with $\mu = \langle k \rangle = \lambda$, $\sigma^2 = \langle k^2 \rangle - \langle k \rangle^2 = \lambda$. An approximation is $p(k, \lambda) = N(\lambda, \lambda)$, valid for very large λ.

Remark: Why is the concept of time being used here?

The Poisson distribution is an example of a stochastic Poisson process

$$f(t) = \sum_{-\infty}^{\infty} \delta(t - t_j) \tag{2.10}$$

and so $k = \int_0^T f(t)$.

Power-law distribution

E.g. the jump distribution of animals pursuing food foraging (Lévy walk) or to describe the distribution of Facebook contacts ($\alpha = 2.2$) are described by a power-law distribution of the form

$$p(x) \sim x^{-\alpha} \quad \text{for} \quad x \gg 1 \tag{2.11}$$

It has some unpleasant mathematical properties such as a divergent variance $\sigma^2 \to \infty$ for $\alpha < 3$.

2.4 Normal Approximation

In this section we are going to show, that the Bernoulli distribution

$$p(k,n) = \binom{n}{k} p^k (1-p)^{n-k} \tag{2.12}$$

can be approximated by using a normal distribution

$$p(k,n) \approx \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(k-np)^2}{2\sigma^2}\right) \tag{2.13}$$

with $\sigma^2 = npq$.

Proof I Using Stirling's Approximation

For this proof we introduce a small deviation ε such that $q = 1-p$, $k = np + n\varepsilon$ and $p(k,n) \approx 0$ for $\varepsilon \gg \frac{1}{\sqrt{N}}$. The first trick is then to use Stirling's approximation

$$n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \tag{2.14}$$

which leads us to

$$\begin{aligned} p(k,n) &= \frac{\sqrt{2\pi n}}{\sqrt{2\pi k}\sqrt{2\pi(n-k)}} \frac{n^n}{k^k(n-k)^{n-k}} p^k q^{n-k} \\ &= \left[\frac{1}{\sqrt{2\pi pqn}} + \mathcal{O}(\varepsilon)\right] \left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{(n-k)} \end{aligned}$$

The second trick is to apply $x^k = \exp(k\ln(x))$, so we need to evaluate the following two expressions

$$\ln\left(\frac{np}{k}\right) = \ln\left(\frac{p}{p-\varepsilon}\right) = -\ln\left(1+\frac{\varepsilon}{p}\right) \approx -\frac{\varepsilon}{p} + \frac{1}{2}\left(\frac{\varepsilon}{q}\right)^2$$

$$\ln\left(\frac{nq}{n-k}\right) = \ldots \approx \frac{\varepsilon}{p} - \frac{1}{2}\left(\frac{\varepsilon}{q}\right)^2$$

which implies

$$\left(\frac{np}{k}\right)^k\left(\frac{nq}{n-k}\right)^{n-k} \approx \exp\left(k\left[-\frac{\varepsilon}{p}+\frac{1}{2}\left(\frac{\varepsilon}{q}\right)^2\right]+(n-k)\left[\frac{\varepsilon}{p}-\frac{1}{2}\left(\frac{\varepsilon}{q}\right)^2\right]\right)$$
$$= 0\cdot\varepsilon - \frac{1}{2}n\frac{\varepsilon^2}{p} - \frac{1}{2}n\frac{\varepsilon^2}{q} + \mathcal{O}(\varepsilon^3)$$
$$= -\frac{1}{2}\frac{\varepsilon^2(p+q)}{pq} = -\frac{1}{2}\frac{(k-np)^2}{npq}$$

Thus,

$$p(k,n) \approx \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(k-np)^2}{2\sigma^2}\right) \tag{2.15}$$

with $\sigma^2 = npq$.

Proof II Using the Central-Limit-Theorem

A second proof uses the central-limit-theorem.

Central-Limit-Theorem

Consider a sequence of $x_1, ...x_n$ of independent, identically distributed, random variables with mean μ and variance σ^2. We define the empirical mean by

$$\bar{x} = \frac{1}{n}(x_1 + \cdots + x_n) = \text{empirical mean} \tag{2.16}$$

We normalize it to a random variable with expectation value zero

$$z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \tag{2.17}$$

Then the probability distribution $p(z) \to N(0,1)$ for large n ("convergence in distribution") or equivalently $CDF(z) \to Erf(z)$ for almost all $z \in \mathbb{R}$

Now for our second, more elegant proof we consider n independent random variables x_j

with $j = 1, ..., n$, and

$$x_j = \begin{cases} 1 \,|\, \text{with probability } p \\ 0 \,|\, \text{with probability } q = 1 - p \end{cases} \tag{2.18}$$

The empirical mean

$$\bar{x} = \frac{1}{n}(x_1 + \cdots + x_n) = \text{empirical mean} \tag{2.19}$$

is directly related to the number k of positive outcomes $k = n\bar{x}$. we get $p(k, n) = p(\bar{x}) \sim N(np, npq)$ by the Central-Limit-Theorem.

The idea of the proof is now to compute the cumulants of x_j

$$k_1 = \mu,\, k_2 = \sigma^2 \ldots \tag{2.20}$$

and then to show that the cumulants of z_j are given by

$$k_1 = 0,\, k_2 = 1,\, k_3 \sim \frac{1}{\sqrt{n}},\, k_4 \sim \frac{1}{n} \ldots \tag{2.21}$$

and

$$\lim_{n\to\infty} \ln \langle zt \rangle = \lim_{n\to\infty} \sum_{l=0}^{\infty} k_l t^l = 1 - \frac{1}{2}t^2 \tag{2.22}$$

Then we have $p(z) \to N(0, 1)$. To show this behaviour of the cumulants we take a look at the functions

$$C_x(t) = \exp(\langle xt \rangle) \quad \text{with} \quad C_{\alpha x}(t) = C_x(\alpha t),\, \alpha \in \mathbb{R} \tag{2.23}$$

which satisfy $\langle (\alpha x)^j \rangle = \alpha^j \langle x^j \rangle$ and $k_{\alpha x,j} = \alpha^j k_{x,j}$ for all cumulants, i.e. $\forall j \in \mathbb{N}$. Use this for $\bar{x}, z$:

$$C_{\bar{x}}(t) = C_{x_1}\left(\frac{t}{n}\right) C_{x_2}\left(\frac{t}{n}\right) \dots C_{x_n}\left(\frac{t}{n}\right) = \left(C\left(\frac{t}{n}\right)\right)^n$$

$$C_z(t) = C_{\bar{x}}\left(\frac{t}{\sigma/\sqrt{n}}\right) \exp\left(-\frac{\mu t}{\sigma/\sqrt{n}}\right) = \left(C_x\left(\frac{t}{\sigma/\sqrt{n}}\right)\right)^n \exp\left(-\frac{\mu t}{\sigma/\sqrt{n}}\right)$$

giving us

$$\ln(C_z(t)) = n \ln\left(C_x\left(\frac{t}{\sigma/\sqrt{n}}\right)\right) - \frac{\mu t}{\sigma/\sqrt{n}} \tag{2.24}$$

and so we get

$$k_{z,j} = n\left(\frac{1}{\sigma\sqrt{n}}\right)^j k_{x,j} \tag{2.25}$$

2.5 Stochastic Processes

Stochastic Process

A stochastic process is a random function $f(t) : \mathbb{R} \to \mathbb{R}$, i.e. a family of random variables parameterised by $t \in \mathbb{R}$.

Terminology

- conditional probability density: $p(f(t_2) = f_2 | f(x_1) = f_1)$
- Markov property: For $t_3 > t_2 > t_1$ it holds that $p(f(t_3) = f_3 | f(x_2) = f_2, f(x_1) = f_1) = p(f(t_3) = f_3 | f(x_2) = f_2) \; \forall t_j, f_j$, example: diffusion, counter-example: random draw from an urn without replacements
- Martingales: Markov processes with the property $\langle f(t_2) | f(t_1) = t_1 \rangle = f_1$, example: diffusion, counter-example: diffusion with drift

Example: Poisson Process

Poisson Process

For a Poisson process events occur independently with rate r at random times t_j

$$f(t) = \sum_{j=-\infty}^{\infty} \delta(t - t_j) \tag{2.26}$$

with the property $\langle f(t) \rangle = r$. The quantity $x = \int_0^T f(t)$ counts events and yields a Poisson distribution with parameter $\lambda = rt$. The waiting times $t = t_{j+1} - t_j$ are exponentially distributed, i.e. $p(t) = r \exp(-rt)$.

The last property can be proven by taking a look at the CDF, for which we have

$$P(t \geq \theta + \mathrm{d}t) = P(t \geq \theta) - r \, \mathrm{d}t \, P(t \geq \theta)$$

so that

$$\frac{\mathrm{d}}{\mathrm{d}t} P(t \geq \theta) = -r \, \mathrm{d}t \, P(t \geq \theta) \quad \Rightarrow \quad P(t \geq \theta) \sim \exp(-rt)$$

Example: Gaussian White Noise

Poisson Process

Gaussian White Noise is described by a function $\xi(t) : \mathbb{R} \to \mathbb{R}$ with the following properties

i) $\langle \xi(t) \rangle = 0$

ii) $\langle \xi(t)\xi(t') \rangle = 2D\delta(t - t')$

iii) $\int_{t_1}^{t_2} \mathrm{d}t\, \xi(t) \sim N(0, 2D[t_2 - t_1])$

Gaussian white noise can be considered as the idealization of thermal random forces, corresponding to the limit of vanishing correlation time, $\tau_c \to 0$.

Remark: Gaussian White Noise and Mathematics

Strictly speaking, ξ itself cannot be defined mathematically. Instead mathematicians define a so-called Wiener process

$$W(t) = \int_0^t \mathrm{d}t'\, \xi(t') \tag{2.27}$$

so that $W(t)$ exists and is continuous with probability 1.

3 Langevin Equation and Fokker-Planck Equation

3.1 Langevin equation

Langevin theory describes non-equilibrium systems by postulating a stochastic process, thus adding a noise term to fundamental equations. In its original form, Langevin theory was used to describe Brownian motion, e.g. of a particle suspended in a fluid.

Definition of the Langevin equation

The Langevin equation is a stochastic differential equation for the particle velocity

$$\dot{x} = \underbrace{f(x)}_{\text{drift}} + \underbrace{\sqrt{2D}\,\xi(t)}_{\text{random noise}} \tag{3.1}$$

- $\xi(t)$ represents Gaussian white noise
- $f(x)$ describes diffusion in an effective potential $U(x) = -\int_0^x \mathrm{d}x'\, f(x')$

Generalisation to m variables

$$\dot{x}_i = f_i(\vec{x}) + \sum_{j=1}^{m} g_{ij}(\vec{x})\xi_j(t) \tag{3.2}$$

with $i = 1, \ldots, n$ and $\xi_j(t)$ being independent Gaussian white noise functions $\langle \xi_j(t)\xi_l(t')\rangle = \delta_{jl}\delta(t-t')$, $j, l = 1, \ldots, m$.

<u>Example 1: Double-well Potential</u>

<u>Example 2: Escape over a Barrier</u>

Numerics for the Langevin Equation

Euler Scheme

The Langevin equation $\dot{x} = f(x) + \sqrt{2D}\,\xi(t)$ leads, using the Euler scheme, to the following update-rule

$$\hat{x}_{n+1} = \hat{x}_n + f(\hat{x}_n)\,\mathrm{d}t + \sqrt{2D\,\mathrm{d}t}N_n \tag{3.3}$$

with $D = D_0$, $t_i = i\,\mathrm{d}t$, $x_i = x(t_i)$ and the independent Gaussian variables $N_n \sim N(0,1)$. The numerical error scales as $|\hat{x}_n - x_n| \sim \mathcal{O}(\mathrm{d}t^{3/2})$.

3.2 Fokker-Planck-Equation

Derivation of Fokker-Planck-Equation

Repetition: Ordinary Diffusion

For the example of ordinary diffusion in one space dimension

$$\dot{x} = \xi(t),\, x(0) = 0 \quad \text{with} \quad \langle x(t)\rangle = 0,\, \left\langle x^2(t)\right\rangle = 2Dt \tag{3.4}$$

the probability density is given by

$$p(x,t) = \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2Dt}}\exp\left(-\frac{x^2}{4Dt}\right) \tag{3.5}$$

fulfilling the Diffusion equation

$$\frac{\partial p(x,t)}{\partial t} = D\frac{\partial^2 p(x,t)}{\partial x^2} \tag{3.6}$$

Considering the general case $\dot{x} = f(x) + \sqrt{2D}\,\xi(t)$, we would like to find an operator $\hat{L}$ such that

$$\frac{\partial p(x,t)}{\partial t} = \hat{L}p(x,t) \tag{3.7}$$

Therefore, we discretize time and take a look how a sub-ensemble of $p(x,t)$ at x_n will evolve during a time step from $p(x,t_n)$ to $p(x,t_{n+1})$. For this we are using the Markov-Property:

$$\begin{aligned}
p(x,t_{n+1}|x_0,t_0) &= \int \mathrm{d}x_n\, p(x,t_{n+1},x_n,t_n|x_0,t_0) \\
&= \int \mathrm{d}x_n\, p(x,t_{n+1}|x_n,t_n)\, p(x_n,t_n|x_0,t_0) \\
&= \int \mathrm{d}x_n\, N(x_n+f(x_n),2D\,\mathrm{d}t)p(x_n,t_n|x_0,t_0)
\end{aligned}$$

This is already an implicit solution in terms of a convolution of the probability density with a family of normal distributions, but it is of little practical use.

A Remark about Units

Unlike probabilities, probability densities for positions have units of inverse length! Therefore we are integrating over a two-point probability density which has units of inverse length squared

$$\begin{aligned}
[p(x,t_{n+1}|x_0,t_0)] &= \mathrm{m}^{-1} \\
[p(x,t_{n+1},x_n,t_n|x_0,t_0)] &= \mathrm{m}^{-2}
\end{aligned}$$

So, let us define the following abbreviations in order to evaluate this convolution further

$$p(x,t_n) = \int \mathrm{d}x_n I(x_n,y)|_{y=x-x_n} \quad \text{with} \quad I(x_n,y) = p(x_n)n(x_n,y),$$
$$\text{and} \quad n(x_n,y) = N(f(x_n)\mathrm{d}t, 2D\mathrm{d}t)$$

The integrand $I(x,y)$ will contribute only for small $y=\mathcal{O}(\mathrm{d}t)$, which means $x_n \approx x$, so we can Taylor expand $I(x_n,y)$ in x_n around x:

$$I(x_n,y) = I(x,y) + \frac{\partial I(x_n,y)}{\partial x_n}\Big|_{x_n=x}(x_n-x) + \frac{\partial^2 I(x_n,y)}{\partial x_n^2}\Big|_{x_n=x}\frac{(x_n-x)^2}{2} \tag{3.8}$$

Inserting this into the convolution integral leads to

$$\begin{aligned}
p(x, t_{n+1}) &= \int \mathrm{d}y \, I(x,y)|_{x_n = x-y} \\
&= \int \mathrm{d}y \left(p(x) n(x,y) - \frac{\partial}{\partial x}(p(x)\, n(x,y)\, y) + \frac{\partial^2}{\partial x^2}(p(x)\, n(x,y)\, \frac{y^2}{2}) \right) \\
&= p(x) \int \mathrm{d}y \, n(x,y) - \frac{\partial}{\partial x}\left(p(x) \int \mathrm{d}y \, n(x,y)\, y \right) + \frac{\partial^2}{\partial x^2}\left(p(x) \int \mathrm{d}y \, n(x,y)\, \frac{y^2}{2} \right)
\end{aligned}$$

The integrals that are occurring in this step are know as Kramers-Moyal coefficients:

$$\begin{aligned}
&\int \mathrm{d}y \, n(x,y) = 1 \\
&\int \mathrm{d}y \, n(x,y)\, y = f(x)\, \mathrm{d}t \\
&\int \mathrm{d}y \, n(x,y)\, \frac{y^2}{2} = D\, \mathrm{d}t + \frac{1}{2}[f(x)]^2\, \mathrm{d}t^2 = D\, \mathrm{d}t + \mathcal{O}(\mathrm{d}t^2)
\end{aligned}$$

which give us

$$p(x, t_{n+1}) = p(x, t_n) - \frac{\partial}{\partial x}[p(x, t_n) f(x)\, \mathrm{d}t] + \frac{\partial^2}{\partial x^2}[p(x, t_n) D]\, \mathrm{d}t \tag{3.9}$$

and thus

$$\frac{p(x, t_{n+1}) - p(x, t_n)}{\mathrm{d}t} = -\frac{\partial}{\partial x}[p(x, t_n) f(x)] + \frac{\partial^2}{\partial x^2}[p(x, t_n) D] \tag{3.10}$$

Taking the time step to zero, we have finally derived the Fokker-Planck equation.

Fokker-Planck equation

The Fokker-Planck equation is a partial differential equation for a probability density $p(x)$, which reads for an ensemble governed by the Langevin equation (3.1) as follows

$$\frac{\partial}{\partial t} p(x,t) = -\frac{\partial}{\partial x}[p(x,t)\, f] + D \frac{\partial^2}{\partial x^2} p(x,t) \tag{3.11}$$

The structure of the Fokker-Planck equation is similar to the Schrödinger equation, i.e. solution methods from QM can be borrowed (take a look at the Risken book!).

Application to Diffusion in a Potential

We consider diffusion in a potential $U(x)$ (now we care about physical units!)

$$\dot{x} = -\frac{1}{\gamma}\frac{\partial U}{\partial x} + \xi \tag{3.12}$$

and ask for the steady state $\frac{\partial}{\partial t}p(x,t) = 0$. Hence, the Fokker-Planck equation reads

$$\begin{aligned} 0 &= \vec{\nabla}[(\frac{1}{\gamma}\vec{\nabla}U)p] + D\vec{\nabla}^2 p = \vec{\nabla}[\frac{1}{\gamma}\vec{\nabla}Up + D\vec{\nabla}p] \\ \Rightarrow c &= \frac{1}{\gamma}\vec{\nabla}Up + D\vec{\nabla}p \end{aligned}$$

thus, if $c = 0$, we get

$$\frac{\partial}{\partial x}\ln p = \frac{\vec{\nabla}p}{p} = -\frac{1}{\gamma}\frac{\vec{\nabla}U}{p} \tag{3.13}$$

and

$$p \sim \exp\left(-\frac{U}{\gamma D}\right) = \exp\left(-\frac{U}{k_B T}\right) \tag{3.14}$$

with $D = \frac{k_B T}{\gamma}$, i.e. we recover the Boltzmann distribution. If c would not be zero, the solution could not be normalized. Another explanation, why $c = 0$, is based on the Fokker-Planck-equation being interpreted as conservation equation

$$\dot{p} = -\vec{\nabla}\vec{J} \quad \text{with} \quad \vec{J} = \frac{1}{\gamma}\vec{\nabla}Up + D\vec{\nabla}p \tag{3.15}$$

of the current $\vec{J}$. At equilibrium, the current must vanish and thus we have

$$\lim_{t\to\infty}\vec{J} = c = 0 \tag{3.16}$$

Eigenvalue Spectrum of $\hat{L}$

The probability density can be expressed in terms of eigenfunctions of the operator $\hat{L}$

$$\hat{L}\phi_n(x) = \lambda_n \phi_n(x) \tag{3.17}$$

through

$$p(x,t) = \sum a_n \phi_n(x) \exp(\lambda_n t) \tag{3.18}$$

If $\lambda_0 = 0$, then the corresponding eigenfunction is a steady state ϕ_0. The slowest decaying mode determines hopping rates.

But why are the λ_n real? We have $\hat{L} \neq \hat{L}^*$, which means $\hat{L}$ is not Hermitian.

$$\left\langle \hat{L}g, h \right\rangle = \int \mathrm{d}x\, (\hat{L}g)h = \int \mathrm{d}x\, g\hat{L}^* h = \left\langle g, \hat{L}^* h \right\rangle \quad \forall g(x), h(x) \tag{3.19}$$

By partial integration we see that the adjoint operator has the form

$$\hat{L}^* h = f\frac{\partial h}{\partial x} + D\frac{\partial^2 h}{\partial x^2} \tag{3.20}$$

If $f(x) = -\frac{\partial U(x)}{\partial x}$, we can define a Hermitian operator by

$$A = T^{-1}LT \quad \text{with} \quad T = \exp\left(+\frac{\beta U}{2}\right),\, \beta = \frac{1}{D} \tag{3.21}$$

The operator is self-adjoint $\hat{A} = \hat{A}^*$ and thus all eigenvalues are real. $\hat{A}$ and $\hat{L}$ have the same eigenvalues.

Backward Fokker-Planck Equation

$p = p(x_1, t|x_0, 0) = p(x_1, 0|x_0, -t)$, which gives the backward Fokker-Planck equation

$$\dot{p} = \hat{L}_{x_1} p = \hat{L}^*_{x_0} p = \left[+f(x)\frac{\partial}{\partial x_0} + D\frac{\partial^2}{\partial x_0^2}\right] p(x_1, 0\,|\, x_0, -t) \tag{3.22}$$

Boundary Conditions Matter

1) Reflecting Boundary Conditions (No-Flux / Robin B.C.)

The probability current $\dot{p} = -J$ vanishes

$$J(x_1) = J(x_2) = 0 \tag{3.23}$$

and

$$\int_{x_1}^{x_2} \mathrm{d}x \, p(x,t) = 1 \tag{3.24}$$

In this case the steady-state distribution $p^*(x) = \phi_0(x)$ exists. Similar results hold for a confinement potential $\lim_{x \to x_1, x_2} U(x) \to \infty$.

2) Absorbing Boundary Conditions

For absorbing boundary conditions we impose $p(x_2, t) = 0$ (Dirichlet boundary conditions). Therefore

$$0 > \frac{\mathrm{d}}{\mathrm{d}t} \int \mathrm{d}x \, p(x,t) = \int \mathrm{d}x \, \frac{\mathrm{d}p(x,t)}{\mathrm{d}t} = \int_{-\infty}^{x_2} -\frac{\partial J}{\partial x} = -J(x_2) \tag{3.25}$$

No steady-state solution exists (non-trivial / normalizable to one) and all eigenvalues are strictly negative.

Boundary Conditions and Functional Analysis

Changing the boundary conditions changes also the eigenvalues and the adjoint operator (boundary terms might pop up) and thus you will get each time a different operator in terms of functional analysis.

4 Dynkin Equation

4.1 Mean First Passage Times and Dynkin Equation

We consider diffusion in some potential landscape $\gamma\dot{x} = -\frac{\partial U}{\partial x} + \xi(t)$ with initial conditions $p(x,0) = \delta(x - x_1)$ and boundary conditions $p(x_2, t) = 0$.

Mean First Passage Time (MFPT)

$$\tau(x_2|x_1) = \int_0^\infty \mathrm{d}t\, tJ(x_2, t|x_1, 0) \tag{4.1}$$

Our aim is to derive an equation for τ. If Δt is small and kept constant (and we ask which positions can we reach within Δt) we have

$$\tau(x_2|x_1) = \Delta t + \int_{-\infty}^{x_2} \mathrm{d}x'\, \tau(x_2|x')p(x', \Delta t|x_1, 0)$$

and we take the derivative with respect to Δt

$$0 = 1 + \int_{-\infty}^{x_2} \mathrm{d}x'\, \tau(x_2|x')\hat{L}_{x'}p = 1 + \int_{-\infty}^{x_2} \mathrm{d}x'\, \hat{L}^*_{x'}\tau(x_2|x')p$$

so if $\Delta t \to 0$ then $p(x', \Delta t|x_1, 0) \to \delta(x - x_1)$ and we get the Dynkin equation

$$-1 = \hat{L}^*_{x_1}\tau(x_2|x_1) \tag{4.2}$$

Application to Diffusion

Let consider once again the example of diffusion

$$\gamma\dot{x} = -\frac{\partial U}{\partial x} + \xi(t) \quad \langle\xi(t)\xi(t')\rangle = 2D\delta(t - t') \tag{4.3}$$

with the initial condition $p(x,0) = \delta(x - x_1)$ and boundary conditions $p(x_2, t) = 0$. Let $v = \frac{\partial}{\partial x_1}\tau(x_2, x_1)$ so the Dynkin equation reads

$$-1 = Dv' - \frac{U'}{\gamma}v \tag{4.4}$$

which we multiply by $\frac{1}{D}\exp(-\beta U)$

$$-\frac{1}{D}\exp(-\beta U) = v'\exp(-\beta U) - \beta v \exp(-\beta U) = \frac{\mathrm{d}}{\mathrm{d}x_1}[v\exp(-\beta U)] \tag{4.5}$$

to get

$$v = -\frac{1}{D}\exp(-\beta U)\left[\int_{-\infty}^{x_1}\mathrm{d}x'\exp(-\beta U) + c\right] \tag{4.6}$$

If we assume $\lim_{x\to-\infty} U(x) = +\infty$, then $|v| < \infty$ and $c = 0$. With one more integration we get

$$\tau(x_2, x_1) = \frac{1}{D}\int_{x_1}^{x_2}\exp(\beta U(x'))\left[\int_{-\infty}^{x'}\mathrm{d}x''\exp(-\beta U(x''))\right] \tag{4.7}$$

Note that the integration constant if the second integration must be zero due to $\tau(x_2, x_2) = 0$

4.2 Kramers Escape Rate Theory

We consider the escape of particles over an energy barrier ΔE and assume $\beta\Delta E \gg 1$ to calculate $\tau(x_2|x_1)$. $\int \mathrm{d}x''$ is sizeable only nearby x_a, $\int \mathrm{d}x'$ is sizeable only nearby x_b. We do a standard trick: quadratic expansion around x_a and x_b

$$U(x'') = U(x_a) + \frac{1}{2}U''(x_a)(x'' - x_a)^2 + \dots \tag{4.8}$$

with $U''(x_a) = k_a = \gamma/\tau_a$, which introduces a time-scale. Similarly

$$U(x') = U(x_b) + \frac{1}{2}U''(x_b)(x' - x_b)^2 + \dots \tag{4.9}$$

with $U''(x_b) = -k_b = -\gamma/\tau_b$. So lets evaluate our integrals

$$\int_{-\infty}^{x'} \mathrm{d}x'' \exp\left(-\frac{1}{2}\beta U''(x_a)(x''-x_a)^2\right) \approx \int_{-\infty}^{\infty} \mathrm{d}x'' \exp\left(-\frac{1}{2}\beta U''(x_a)(x'-x_a)^2\right)$$
$$= \sqrt{2\pi\sigma^2}$$

with $\sigma^2 = \frac{\tau_a}{\beta\gamma}$ and

$$\int_{-\infty}^{x'} \mathrm{d}x'' \exp\left(+\frac{1}{2}\beta U''(x_b)(x'-x_b)^2\right) \approx \int_{-\infty}^{\infty} \mathrm{d}x'' \exp\left(+\frac{1}{2}\beta U''(x_b)(x'-x_b)^2\right)$$
$$= \sqrt{2\pi\frac{\tau_b}{\beta\gamma}}$$

so

$$\tau(x_2, x_1) = \frac{1}{D}\frac{2\pi\sqrt{\tau_a\tau_b}}{\beta\gamma}\exp(\beta\Delta E) = 2\pi\sqrt{\tau_a\tau_b}\exp(\beta\Delta E) \tag{4.10}$$

Kramers escape rate

$$r = \frac{1}{\tau(x_2, x_1)} \sim \underbrace{\exp(-\beta\Delta E)}_{\text{Arrheniusfactor}} \tag{4.11}$$

4.3 Diffusion to Capture

As an example, we consider a diffusing particle released between two absorbing plates. The question is: What is the probability of getting absorbed at either of the two plates?

$$P(x, t=0) = \delta(x - x_0)$$
$$P(x_1, t) = P(x_2, t) = 0$$

The probability of becoming absorbed at $x = x_1$ when starting at x_0 reads $\pi_1(x_0)$. We have $\pi_1(x_1) = 1$ and $\pi_1(x_2) = 0$.

We will now consider a time step Δt as we did for the derivation of the Dynkin equation in order to find an explicit expression for $\pi_1(x_0)$:

$$\pi_1(x_0) = \int_{x_1}^{x_2} \mathrm{d}x\, \pi_1(x) P(x, \Delta t | x_0, 0)$$

Now we take the partial derivative with respect to Δt

$$0 = \int_{x_1}^{x_2} \mathrm{d}x\, \pi_1(x) \underbrace{\frac{\partial}{\partial \Delta t} P(x, \Delta t | x_0, 0)}_{\hat{L} P}$$

and perform partial integration

$$0 = \int_{x_1}^{x_2} \mathrm{d}x\, \hat{L}^* \pi_1(x) \underbrace{P(x, \Delta t | x_0, 0)}_{\to \delta(x - x_0) \text{ for } \Delta t \to 0}$$

so we obtain

$$0 = \hat{L}^* \pi_1(x)$$

Thus, $\pi_1(x_0)$ must be a linear function and taking the boundary conditions into account we conclude $\pi_1(x_0) = \frac{x_2 - x_0}{x_2 - x_1}$.

Another way to compute this is the method of images. So

$$P(x, t) = N(x_0, 2Dt) - N(2x_1 - x_0, 2Dt) - N(2x_2 - x_0, 2Dt) \tag{4.12}$$

and π_1 could be calculated directly. (Stream of anti-particles is released and annihilates particles at the boundary).

4.4 Polya's theorem

We consider diffusion in $\mathbb{R}^d$ to a d-dimensional absorbing ball. We ask for the probability $p(R_0)$ for a particle initially released at distance R_0 to hit the target ball. For $d = 1$ and $d = 2$ we have $p(R_0) = 1$, but for $d = 3$ $p(R_0) = \frac{R_1}{R_0}$

Note that for $d = 3$, the characteristic arrival time must scale with $\sqrt{R_0^2/D}$ with a power-law tail $\sim t^{-3/2} \exp(-(R_0 - R_1)^2/4Dt)$ and the mean first passage time diverges.

5 Synchronisation

5.1 Active Oscillators

An active oscillator is a non-conservative oscillator, e.g. energy is lost through damping. An example is the

Van-der-Pol oscillator

$$m\ddot{x} - \gamma(\frac{1}{4}\Lambda - x^2)\dot{x} + kx = 0 \tag{5.1}$$

The Van-der-Pol oscillator undergoes a so-called Hopf bifurcation as a function of parameter Λ: For $\Lambda < 0$ $x = 0$ is stable and for $\Lambda > 0$ limit-cycle oscillations occur. Equation (5.1) can be transformed into Hopf normal form.

Other examples are the

Hopf oscillator

An active oscillator with non-linear damping of the form

$$\dot{z} = i\omega_0 z + \mu(\Lambda - |z|^2)z \quad \text{with} \quad z \in \mathbb{C} \tag{5.2}$$

It is used e.g. to describe the limit cycle of electric circuits involving a vacuum tube or in biology to model the activation potential of neurons.

and a phase oscillator with $\dot{\varphi} = \omega_0$.

Hopf normal form of Van-der-Pol oscillator

We set $y = \dot{x}$, $\omega = \sqrt{k/m}$. The idea is to introduce $z \approx x - \frac{i}{\omega} y$, so we do the ansatz (in order to avoid quartic terms / the method is called Center Manifold technique)

$$\begin{aligned} z &= \sum_k^\infty \sum_l^\infty d_{k,l} x^l y^{k-l} \\ &= x - \frac{i}{\omega} y + d_{10} y + d_{33} x^3 + d_3 2 x^2 y + d_{31} x y^2 + d_{30} y^3 + \dots \end{aligned}$$

The back transformation is given by

$$\begin{aligned} x &= \frac{z + \bar{z}}{2} + e_1 z^3 + e_2 z^2 \bar{z} + e_3 z \bar{z}^2 + e_4 \bar{z}^3 \dots \\ y &= i\omega \frac{z - \bar{z}}{2} + f_1 z^3 + f_2 z^2 \bar{z} \dots \end{aligned}$$

so that

$$\dot{z} = h(z, \bar{z}) = Fz + Gz^2\bar{z} + \text{h.o.t.} \tag{5.3}$$

and for appropriate $d_{k,l}$ we have i) no quadratic terms, ii) no term in $\bar{z}$ and iii) no terms proportional to $z^3, z\bar{z}^2, \bar{z}^3$. We find that

$$\begin{aligned} F &= i\omega_0 + \frac{\gamma}{8m}\Lambda + \mathcal{O}(\Lambda^2) \\ G &= \frac{\gamma}{8m} + \mathcal{O}(\Lambda) \end{aligned}$$

and we get

$$\dot{z} = i(\omega_c - \omega_1 |z|^2) z + \mu(\Lambda - |z|^2) \tag{5.4}$$

with $\omega_c = \omega_0$, $\omega_1 = \mathcal{O}(\Lambda)$ and $\mu = \frac{\gamma}{8m}$

5.2 Hopf-oscillator with noise

We now add a noise term to the Hopf-oscillator

$$\dot{z} = i\omega_0 z + \mu(\Lambda - |z|^2)z + (i\xi_\varphi + \xi_A)z \tag{5.5}$$

with

$$\langle \xi_\varphi(t)\xi_\varphi(t')\rangle = 2D_\varphi \delta(t-t') \quad \langle \xi_A(t)\xi_A(t')\rangle = 2D_A \delta(t-t') \quad \langle \xi_\varphi(t)\xi_A(t')\rangle = 0$$

and map z on a phase φ and amplitude A via $z = Ae^{i\varphi}$ so that

$$\left(\frac{\dot{A}}{A} + i\dot{\varphi}\right) z = \dot{z} = \dots \tag{5.6}$$

and

$$\frac{\dot{A}}{A} + i\dot{\varphi} = i\omega_0 + \mu(A_0^2 - A^2) + i\xi_\varphi + \xi_A \tag{5.7}$$

Assuming $\Lambda > 0$ and $\Lambda = A_0^2$, we get a noisy phase oscillator

$$\dot{\varphi} = \omega_0 + \xi_\varphi \tag{5.8}$$

and an Ornstein-Uhlenbeck process for the amplitude

$$\begin{gathered} A = A_0 + a \\ \dot{a} = \mu(A_0 + a)(-2aA_0 + a^2) + \xi_A = -2\mu A_0 a + \xi_A + \mathcal{O}(a^2) \end{gathered} \tag{5.9}$$

with the properties

$$\langle a(t)\rangle = 0 \quad \langle a(t)a(t')\rangle = D_A \tau \exp\left\{-\frac{|t-t'|}{\tau}\right\} \quad \tau = \frac{1}{2\mu A_0}$$

Remark

If we consider an ensemble average, the amplitude fluctuations will decay with τ:

$$\bar{a}(t) = \langle a(t) \rangle \quad \text{with} \quad \frac{\mathrm{d}}{\mathrm{d}d}\bar{a} = -\frac{\bar{a}}{\tau}$$

Manifestation of Phase Noise

We can characterize noisy oscillations in terms of a phase correlation function

$$C(t) = \langle \exp(i\varphi(t_0)) \exp(-i\varphi(t_0 + t)) \rangle \tag{5.10}$$

with $|C(t) = \exp(-D_\varphi t)|$, so

$$\frac{z(t_0)}{A_0} \frac{\bar{z}}{A_0} \approx \exp(\varphi(t_0) - \varphi(t_0 + t)) \to \exp(i\omega_0 t) \quad \text{if} \quad D_\varphi = 0$$

And we can have a look at the power spectral density

$$S_y(\omega) = |\tilde{y}(\omega)|^2 \tag{5.11}$$

with $y = \exp(i\varphi)$ and its Fourier transform $\tilde{y}(\omega)$.

Two coupled oscillators

Two phase oscillators are coupled by the coupling c leading to the ODE system

$$\begin{aligned} \dot{\varphi}_L &= \omega_L + c(\varphi_L - \varphi_R) \\ \dot{\varphi}_R &= \omega_R + c(\varphi_R - \varphi_L) \end{aligned} \tag{5.12}$$

We introduce a phase difference of $\delta = \varphi_L - \varphi_R$

$$\dot{\delta} = \Delta\omega + c(\delta) - c(-\delta) \tag{5.13}$$

with frequency mismatch $\Delta\omega = \omega_L - \omega_R$. A Fourier expansion of the coupling function yields

$$c(\delta) = c(\delta + 2\pi) = \sum_n C'_n \cos(n\delta) + C''_n \sin(n\delta) \tag{5.14}$$

Only the odd coupling terms contribute to synchronization, often $c(\delta)$ is dominated by the first Fourier mode and we end up at the Adler equation ($\lambda = -2c_1''$)

$$\dot{\delta} = \Delta\omega - \lambda \sin(\delta) \tag{5.15}$$

If $|\Delta\omega| < |\lambda|$, we have fixed points for $\delta^* = \sin^{-1}\left(\frac{\Delta\omega}{\lambda}\right)$. The stability of the fixed points is determined by $\gamma\dot{\delta} = -\frac{\partial U}{\partial \delta}$ and the effective potential $U = -\gamma\Delta\omega\delta - \gamma\lambda\cos(\delta)$.

Images missing!

Synchronization in the Presence of Noise

If we consider two coupled phase oscillators with noise

$$\dot{\varphi}_1 = \omega_1 - \frac{\lambda}{2}\sin(\varphi_1 - \varphi_2) + \xi_1(t) \tag{5.16}$$

$$\dot{\varphi}_2 = \omega_2 - \frac{\lambda}{2}\sin(\varphi_2 - \varphi_1) + \xi_2(t) \tag{5.17}$$

we obtain the noises Adler equation for the phase difference $\delta = \varphi_1 - \varphi_2$

$$\dot{\delta} = \Delta\omega - \lambda \sin(\delta) + \xi \tag{5.18}$$

Here, and Gaussian white noise $\xi(t)$ represents Gaussian white noise with $\langle \xi(t)\xi(t')\rangle = 2(D_L + D_R)\delta(t - t')$.

Remark: How to add two noise terms

$$\xi(t) = \xi_L(t) - \xi_R(t)$$

with $\langle \xi(t)\rangle = 0$ and

$$\begin{aligned}
\langle \xi(t)\xi(t')\rangle &= \langle \xi_L(t)\xi_L(t')\rangle + \langle \xi_R(t)\xi_R(t')\rangle + \langle \xi_L(t)\xi_R(t')\rangle \\
&= 2D_L\delta(t - t') + 2D_R\delta(t - t') + 0 \\
&= 2(D_L + D_R)\delta(t - t')
\end{aligned}$$

We can reinterpret the noisy Adler equation as the overdamped dynamics of a diffusing particle in a potential $U(x)$

$$\gamma\dot{\delta} = -\frac{\partial U}{\partial \delta} + \xi \tag{5.19}$$

and $U/\gamma = -\Delta\omega\delta - \lambda\cos(\delta)$. So what is the effect of noise? The steady state probability density reads

$$p^*(\delta) \sim \exp\left(-\frac{U(\delta)}{k_B T_{\text{eff}}}\right) = \frac{1}{2\pi I_0(ND)} \exp\left(-\frac{\lambda}{D}\cos(\delta)\right) \tag{5.20}$$

with $D = k_B T_{\text{eff}}\gamma$ and $\Delta\omega = 0$. So the first effect of noise is, that steady states are smeared out. The second effect are phase slips that occur

$$\begin{aligned} \delta \approx 0 \longrightarrow \delta \approx 2\pi \qquad &\text{with rate } G_+ \\ \delta \approx 0 \longrightarrow \delta \approx -2\pi \qquad &\text{with rate } G_- \end{aligned}$$

We can compute $G_\pm$ using Kramers escape rate theory

$$\frac{\gamma}{\tau_a} = U''|_{\delta=\delta_a} \Rightarrow \tau_a = \frac{1}{\sqrt{\lambda^2 - \Delta\omega^2}}$$

$$\frac{\gamma}{\tau_b} = U''|_{\delta=\delta_b} \Rightarrow \tau_b = \tau_a$$

and so

$$G_+ = 2\pi\tau_a \exp\left(\frac{-\Delta E}{D/\gamma}\right) \tag{5.21}$$

The calculation for G_- can be done analogously

$$\frac{G_+}{G_-} = \exp(+2\pi\Delta\omega/D) \tag{5.22}$$

For $\Delta\omega = 0$ we get

$$G_+ = G_- = \frac{\lambda}{2\pi}\exp\left(-\frac{2\lambda}{D}\right) \tag{5.23}$$

The theory can be also extended to more than two oscillators.

6 Itō versus Stratonovich Calculus

If we are given an ODE, e.g. $\dot{x} = f(x)$, what does this mean? To answer this question, we are going to take a constructive approach and interpret the ODE as a rule to construct the solution. So we estimate the values $x_i = x(i\,\mathrm{d}t)$ and then take the limit $\mathrm{d}t \to 0$.

6.1 Numerical Motivation

Deterministic ODE

For a deterministic ODE, we have various options to chose a scheme in order to solve them numerically. One could use either an explicit scheme like the Euler scheme

$$x_i = x_{i-1} + f(x_{i-1})\,\mathrm{d}t \tag{6.1}$$

and implicit scheme

$$x_i = x_{i-1} + f(x_i)\,\mathrm{d}t \tag{6.2}$$

or a mixed scheme

$$x_i = x_{i-1} + \frac{1}{2}[f(x_{i-1}) + f(x_i)] \tag{6.3}$$

and all schemes will converge to the same limit.

Stochastic Differential Equations

Also for stochastic differential equations, such as $\dot{x} = f(x) + \sqrt{2D(x)}\xi$, we may chose either an explicit scheme (Itō)

$$x_i = x_{i-1} + f(x_{i-1})\,\mathrm{d}t + \sqrt{2D(x_{i-1})}N_i\sqrt{\mathrm{d}t} \tag{6.4}$$

with $N_i \in N(0,1)$. Alternatively, we could consider a mixed scheme (Stratonovich)

$$x_i = x_{i-1} + \frac{1}{2}[f(x_{i-1}) + f(x_i)]\,\mathrm{d}t + \frac{1}{2}[\sqrt{2D(x_{i-1})} + \sqrt{2D(x_i)}]N_i\sqrt{\mathrm{d}t} \tag{6.5}$$

It is important to note that this time both schemes are different. (A purely implicit scheme for SDE is not discussed, because such schemes are rarely used in practice.) We can see this by doing the expansion

$$x_i = x_{i-1} + f(x_{i-1})\,\mathrm{d}t + \mathcal{O}(\mathrm{d}t^{3/2}) + g(x_{i-1})N_i\sqrt{\mathrm{d}t} + g'(x_{i-1})g(x_{i-1})N_i^2\,\mathrm{d}t \tag{6.6}$$

with $\langle N_i^2 \rangle = 1$, so that the last term can not be neglected!

6.2 Different Interpretations

Having a look at the chain rule, one can see that the Itō and Stratonovich interpretation are indeed two different sorts of calculus. In Stratonovich interpretation the ordinary chain rule holds

$$\text{(S)} \quad y = y(x), \quad \dot{y} = \frac{\partial y}{\partial x}\dot{x} \tag{6.7}$$

By contrast, in Itō interpretation we have $\dot{x}_k = f_k + g_{kl}\xi_l$ with $\langle \xi_k(t)\xi_l(t')\rangle = \delta_{kl}\delta(t-t')$ and the Itō chain rule applies

$$\text{(I)} \quad y = y(x), \quad \dot{y} = \frac{\partial y}{\partial x_j}\dot{x}_j + \frac{1}{2}\frac{\partial^2 y}{\partial x_k \partial x_l} g_{km}g_{ml} \tag{6.8}$$

Switching between Itō and Stratonovich

We consider the same stochastic dynamic $x(t)$ represented by a Langevin equation in either Itō or Stratonovich calculus.

In Itō and Stratonovich calculus, respectively, we have

$$\begin{aligned} \text{(S)} \quad \dot{x}_k &= h_k^S + g_{kl}\xi_l \\ \text{(I)} \quad \dot{x}_k &= h_k^I + g_{kl}\xi_l \end{aligned}$$

with $h_k^I = h_k^S + \frac{1}{2}\frac{\partial g_{kl}}{\partial x_m} g_{ml}$. The Fokker Planck Equation reads for these cases

$$\dot{P} = \frac{\partial}{\partial x_k}\left[-\left(h_k^{I/S} + \alpha \frac{\partial g_{kl}}{\partial x_m} g_{ml}\right) P + \frac{1}{2}\frac{\partial}{\partial x_m}\left(g_{kl} g_{ml} P\right)\right] \tag{6.9}$$

with $\alpha = 0$ for Itō and $\alpha = 1/2$ for Stratonovich calculus.

> **Wong-Zakai Theorem**
>
> If $\dot{x} = f(x) + g(x)\xi$ is a SDE with coloured noise of finite correlation τ, then taking the limit $\tau \to 0$ yields a Stratonovich SDE with Gaussian white noise.

Example of Colored Noise (Ornstein-Uhlenberg process)

$\tau\dot{\xi} = -\xi + \eta$ and $\langle \eta(t)\eta(t')\rangle = \delta(t-t') \Rightarrow \langle \eta(t)\eta(t')\rangle \sim \exp\left(-\frac{-|t-t'|}{\tau}\right)$

Toy example I: Geometric Brownian Motion

We consider the example of (I) $\quad \dot{x} = x\xi$ (*), which corresponds to (S) $\quad \dot{x} = x\xi - Dx$. Now we ask about the time evolution of the first moment $m(t) = \langle x(t)\rangle$? In Itō calculus, we have

$$\frac{\mathrm{d}}{\mathrm{d}t} m(t) = \langle \dot{x}\rangle \overset{(I)}{=} \langle x\xi\rangle = \langle x\rangle \underbrace{\langle \xi\rangle}_{=0} = 0 \tag{6.10}$$

so $m(t) = m_0$. Note that $y = \ln(x) \Rightarrow \dot{y} = \xi - D \Rightarrow y(t) \sim N(-Dt, Dt)$

Toy example II

Next, let's do something forbidden and literally read the Itō SDE (*) as Stratonovich SDE (S) $\quad \dot{x} = x\xi$. This Stratonovich SDE would actually correspond to the Itō SDE (I) $\quad \dot{x} = x\xi + Dx$. Now

$$\frac{\mathrm{d}}{\mathrm{d}t} m(t) = \langle \dot{x}\rangle \overset{(I)}{=} \langle x\xi + Dx\rangle = 0 + Dm \tag{6.11}$$

hence, in Itō calculus this time we get $m = m_0 \exp(Dt)$.

Example: Rotational diffusion in 2D

We have for the stochastic dynamics of the azimuthal angle $\dot{\varphi} = \xi$. For the dynamics of the material frame vectors

$$\vec{e}_1 = \begin{pmatrix} \cos(\varphi) \\ \\ \sin(\varphi) \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} -\sin(\varphi) \\ \\ \cos(\varphi) \end{pmatrix}$$

we have

$$\text{(S)} \quad \dot{e}_1 = \xi \vec{e}_2, \quad \dot{e}_2 = \xi \vec{e}_1 \tag{6.12}$$

In order to rewrite the SDE in Itō interpretation we introduce

$$\vec{e}_1 = \begin{pmatrix} x_1 \\ \\ x_2 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} x_3 \\ \\ x_4 \end{pmatrix}$$

with

$$\dot{\vec{x}} = \vec{g}\xi, \quad \vec{g} = (x_3 x_4 - x_1 - x_2)^T$$

and with $\sum_m \frac{\partial g_k}{\partial x_m} g_m = 2D(-\vec{x})$, we find

$$\text{(I)} \quad \dot{e}_1 = \xi \vec{e}_2 - D\vec{e}_1, \quad \dot{e}_2 = \xi \vec{e}_1 - D\vec{e}_2 \tag{6.13}$$

Extended Example: persistent random walk (2D)

Consider $\vec{r} = v_0 \vec{e}_1$ where $(\vec{e}_1, \vec{e}_2)$ is subject to rotational diffusion. Our proposition is that

$$C(t) = \langle \vec{e}_1(t) \cdot \vec{e}_1(t) \rangle = \exp(-Dt) \tag{6.14}$$

with persistence time $t_p = \frac{1}{D}$ and persistence length $l_p = v_0 t_p$.

Proof: $\frac{\mathrm{d}}{\mathrm{d}t} C(t) = \left\langle \vec{e}_1(t) \cdot \dot{\vec{e}}_1(t) \right\rangle = \vec{e}_1(t) \cdot [\xi \vec{e}_2(t) - D\vec{e}_1(t)] = 0 - D$

6.3 Rotational Diffusion in 3D

As another example we consider rotational diffusion in 3D with the rotational diffusion coefficient (instance of the Fluctuation-Dissipation-Theorem!)

$$D_{\text{rot}} = \frac{k_B T}{8\pi\eta r^3} \tag{6.15}$$

and the parameterization

$$\begin{aligned}
\vec{h}_3 &= (\cos(\psi), \sin(\psi)\cos(\vartheta), \sin(\psi)\sin(\vartheta))^T \\
\vec{g}_1 &= -\frac{\partial \vec{h}_3}{\partial \psi} \quad \vec{g}_2 = -\vec{h}_3 \times \vec{g}_1 \\
\vec{h}_1 &= \cos(\varphi)\vec{g}_1 + \sin(\varphi)\vec{g}_2 \quad \vec{h}_2 = \vec{h}_3 \times \vec{h}_1
\end{aligned}$$

The equations of motion are given by the Frenet-Serret equations for Stratonovich calculus

$$(S) \quad \begin{aligned}
\dot{\vec{h}}_3 &= \xi_2 \vec{h}_1 - \xi_1 \vec{h}_2 \\
\dot{\vec{h}}_1 &= \xi_3 \vec{h}_2 - \xi_2 \vec{h}_3 \\
\dot{\vec{h}}_2 &= \xi_1 \vec{h}_3 - \xi_3 \vec{h}_1
\end{aligned}$$

with $\langle \xi_i(t)\xi_j(t')\rangle = \delta_{i,j}\delta(t-t')2D_{\text{rot}}$ and

$$\text{(S)} \quad \dot{\psi} = \sin(\varphi)\xi_1 + \cos(\varphi)\xi_2 \tag{6.16}$$

which is equivalent to

$$\text{(I)} \quad \dot{\psi} = \underbrace{\sin(\varphi)\xi_1 + \cos(\varphi)\xi_2}_{=\xi(t)} + D_{\text{rot}}\cot(\psi) \tag{6.17}$$

in Ito calculus. We can replace the multiplicative noise term by $\xi(t)$ because

$$\begin{aligned}
\langle \xi(t)\xi(t')\rangle &= \langle [\sin(\varphi(t))\xi_1 + \cos(\varphi(t))\xi_2][\sin(\varphi(t'))\xi_1 + \cos(\varphi(t'))\xi_2]\rangle \\
&= \sin(\varphi(t))\sin(\varphi(t'))\langle \xi_1(t)\xi_1(t')\rangle + \cos(\varphi(t))\cos(\varphi(t'))\langle \xi_2(t)\xi_2(t')\rangle \\
&= [\sin^2(\varphi) + \cos^2(\varphi)]2D_{\text{rot}}\delta(t-t') = 2D_{\text{rot}}\delta(t-t')
\end{aligned}$$

We know that the steady-state distribution must be isotropic, so let's check this. The question is, what is $P^*(\psi)$ for isotropic distribution of $\vec{h}_3$? We have the height $h = 1 - \cos(\psi)$, so $A = 2\pi r h$, $\mathrm{d}A = 2\pi\,\mathrm{d}h$. Thus $P^*(h) = \frac{1}{2}$. Furthermore, $P^*(h)\,\mathrm{d}h = P^*(\psi)\,\mathrm{d}\psi$ with $\mathrm{d}h = \sin(\psi)\,\mathrm{d}\psi$ and so $P^*(h) = \frac{1}{2}\sin(\psi)$

The equation of motion can be also rewritten introducing a potential U

$$(\mathrm{I}) \quad \dot{\psi} = D_{\mathrm{rot}}\cot(\psi) + \xi = -\frac{1}{\gamma}\frac{\partial}{\partial\psi}U + \xi \tag{6.18}$$

with $U = -D_{\mathrm{rot}}\gamma\ln(\sin(\psi)) = k_B T\ln(\sin(\psi))$ and $\gamma = 8\pi\eta r^3$. Thus,

$$P^*(\psi) \sim \exp\left(-\frac{U}{k_B T}\right) \sim \exp(\ln(\sin(\psi))) \sim \sin(\psi) \tag{6.19}$$

An Interpretation of $U(\psi)$ is obtained by taking a look at the entropy $S = k_B\ln(\sin(\psi))$ and the free energy $F = -TS = -D_{\mathrm{rot}}\gamma\ln(\sin(\psi)) = U$. Here, knowing $\vec{h}_3$ corresponds to the microstate and knowing h to the macro state.

6.4 How to derive a correct Langevin equation?

1) can be considered as a limit case of coloured noise $\tau_c \to 0$, then employ Wong-Zakai-theorem

2) small number fluctuations (e.g. for chemical reactions, so suppose you have N particles which can transit from 1 to 2 with rate r_2 and from 2 to 1 with rate r_1, so one can derive a continuum limit of a master equation)

3) only thermal fluctuations $T = \text{const}$ and then use $P^* \sim \exp(-\beta U)$

Master Equation for a Two-State System

Let $P(n)$ be the probability, that n entities are in state 2. Then, the Dynamic equation / Master equation for $P(n)$ is given by

$$\begin{aligned}\dot{P}(n,t) &= r_1(n+1)P(n+1,t) - r_1 nP(n,t) + r_2(N-(n-1))P(n-1,t) - r_2(N-n)P(n,t) \\ &= r_1(E^+ - 1)nP + r_2(E^- - 1)(N-n)P\end{aligned}$$

with the shift operators $E^{\pm}$

$$(E^{+}f)(n) = f(n+1) \quad \text{and} \quad (E^{-}f)(n) = f(n-1)$$

In order to go to a continuum limit we let $x = \frac{n}{N}$ and treat x as a continuous variable. Next, we do a Taylor expansion of our fancy step operators

$$(E^{\pm}f)(x) = f(x \pm \frac{1}{N}) = f(x) \pm f'(x)\frac{1}{N} + \frac{1}{2}f''(x)\frac{1}{N^2} + \dots$$

which we feed back so that we get

$$\begin{aligned}\dot{P}(x,t) &= r_1\frac{\partial}{\partial x}(xP) + \frac{r_1}{2N}\frac{\partial^2}{\partial x^2}(xP) - r_2\frac{\partial}{\partial x}[(1-x)P] + \frac{r_2}{2N}\frac{\partial^2}{\partial x^2}[(1-x)P] \\ &= (r_1+r_2)\frac{\partial}{\partial x}[(x-x^*)P] + \frac{1}{2N}\frac{\partial^2}{\partial x^2}[(r_1+(r_1-r_2)x)P]\end{aligned}$$

with $x^* = \frac{r_2}{r_1+r_2}$. In the steady state we have $r_0 = r_1 = r_2$ and the master equation $\dot{P} = -\vec{\nabla}J$ with $J = 0$ at equilibrium, thus $P^*(x) \sim \exp\left(-\frac{(x-\frac{1}{2})^2}{2\sigma^2}\right)$ and $\sigma^2 = \frac{1}{4N}$

Langevin equation

$$(\mathrm{I}) \quad \dot{x} = (r_1+r_2)(x^*-x) + \underbrace{\sqrt{\frac{r_1x + r_2(1-x)}{2N}}}_{=g(x)}\xi \tag{6.20}$$

$$(\mathrm{S}) \quad \dot{x} = (r_1+r_2)(x^*-x) + \underbrace{\sqrt{\frac{r_1x + r_2(1-x)}{2N}}}_{=g(x)}\xi - \frac{1}{2}\frac{r_1-r_2}{4N} \tag{6.21}$$

In the limit $N \gg 1$ we have $x \approx x^*$. Thus, $g(x) \approx g(x^*)$ and $P^*(x) = N(x^*, \sigma^2)$, $\sigma^2 = \frac{1}{N}\frac{r_1r_2}{(r_1+r_2)^2}$.

6.5 Numerical Integration of nonlinear SDE

To numerically integrate an Itō SDE (I) $\dot{x} = f(x) + g(x)\xi(t)$, $\langle \xi(t)\xi(t')\rangle = \delta(t-t')$, we can use the Euler-Maruyama scheme

$$\begin{aligned} x_{t+\Delta t} &= x_t + f(x_t)\Delta t + g(x_t)N_t, \qquad && N_t \sim N(0,\Delta t) \\ x_{t+\Delta t} &= x_t + f(x_t)\Delta t + g(x_t)N_t'\sqrt{\mathrm{d}t}, \qquad && N_t' \sim N(0,1) \end{aligned}$$

For the integration of an Stratonovich SDE (S) $\dot{x} = f(x) + g(x)\xi(t)$ we can use the Euler-Heun scheme

$$x_{t+\Delta t} = x_t + f(x_t)\Delta t + \tfrac{1}{2}[g(x_t) + g(\bar{x}_t)]N_t, \qquad N_t \sim N(0,1)$$

where $\bar{x}_t = x_t + g(x_t)N_t$

7 Fluctuation-Dissipation-Theorem

7.1 Historical Examples

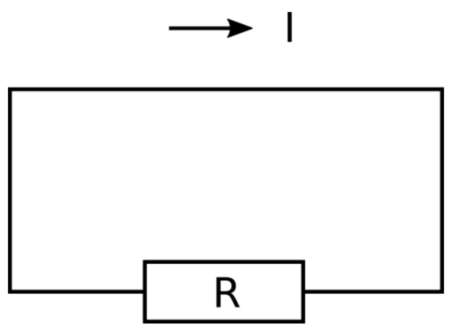
Example 1: Diffusion (Einstein 1905)

The relation found be Einstein for ordinary diffusion in 1905

$$D = \frac{k_B T}{\gamma} \tag{7.1}$$

is an instance of the fluctuation dissipation theorem. The diffusion coefficient D characterises the mean square displacement $\langle x^2(t) \rangle = 2Dt$ (fluctuations) and the right-hand side is related to the dissipated energy via the hydrodynamic mobility $\frac{1}{\gamma} = \frac{1}{6\pi\eta a}$ so that the velocity is given by $v = \frac{1}{\gamma} F$.

Example 2: Electrothermal noise (Johnson, Nyquist 1927)

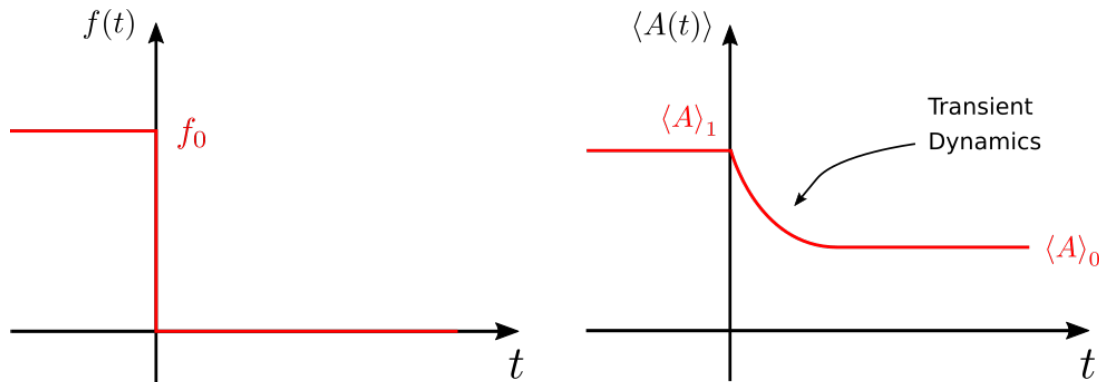


It was found that even a shorted circuit consisting of just one resistor does show a finite current, which is zero on average $\langle I \rangle = 0$, but has a non-zero fluctuation spectrum

$$S_I^{(\omega)} = 2\frac{k_B T}{R\pi} \tag{7.2}$$

Here, we assume the classical limit $\hbar\omega \ll k_B T$. The inverse resistance plays the role of a linear response coefficient $I = \frac{1}{R} U$.

7.2 FDT for classical systems



Lets consider a system described by the Hamiltonian $H_1 = H_0 - fA$ for times $t < 0$, with the probability density $p_1(x) \sim \exp(-\beta H_1)$. At $t = 0$ we switch off the external field f coupled to the observable A. Thus, $p(x, t) \to \exp(-\beta H_0)$ for $t \to \infty$. The ensemble-average of A is given by $\langle A \rangle = \int \mathrm{d}x\, A(x) p(x, t)$ and we integrate over microstates $x = (p_1, \ldots, p_N, q_1, \ldots, q_N)$.

Fluctuation Dissipation Theorem

The FDT relates the fluctuation spectrum on the left side to the dissipative response to an external field on the right side of

$$S_A(\omega) = \frac{2k_B T}{\omega} \operatorname{Im}(\tilde{\chi}_A(\omega)) \tag{7.3}$$

In order to show that the Fluctuation Dissipation Theorem holds we need to key concepts:

- Boltzman distribution $p_0 \sim \exp(-\beta H_0)$ with $\beta = \frac{1}{k_B T}$
- time propagator $P(x_1, t_1 | x_0, t_0)$

Fluctuation Spectrum

The auto-correlation function is given by

$$C_A(\tau) = \langle A(t) A(t + \tau) \rangle - \langle A \rangle^2 \tag{7.4}$$

which is independent of t at thermal equilibrium. The correlation function is an even function $C_A(\tau) = C_A(-\tau)$. It is related to the time propagator by

$$C_A(\tau) = \int \mathrm{d}x_0\, \mathrm{d}x_1\, A(x_0)A(x_1)p_0(x_0)P(x_1, t+\tau|x_0, t) - \langle A \rangle^2 \tag{7.5}$$

The power spectral density is then the Fourier transform

$$S_A(\omega) = \tilde{C}_A(\omega) = \int \mathrm{d}\tau\, C_A(\tau)e^{i\omega\tau} \tag{7.6}$$

where we use the non-unitary Fourier transform with angular frequency.

Wiener-Kinchin Theorem

The Fourier transform exists and has the usual properties.

Formally, we have $\left\langle \tilde{A}(\omega)\tilde{A}^*(\omega') \right\rangle = S_A(\omega)\delta(\omega - \omega')$. Note, however that $\tilde{A}(\omega)$ is not in a strict mathematical sense defined.

Linear Response Function

Let a system possess the Hamiltonian $H(x, t) = H_0(x) - A(x)f(t)$. Then, the linear response is expressed by

$$\langle A(t) \rangle = \langle A \rangle_0 + \int_{-\infty}^{\infty} \mathrm{d}\tau\, \chi_A(\tau)f(t-\tau) + \mathcal{O}(f^2) \tag{7.7}$$

which defines the linear response function $\chi_A(\tau)$. Causality implies that $\chi_A(\tau) = 0$ for all $\tau < 0$.

The Fourier transform reads

$$\tilde{\chi}_A(\omega) = \int_{-\infty}^{\infty} \mathrm{d}\tau\, \chi_A(\tau)e^{i\omega\tau} \tag{7.8}$$

Example: Oscillating Field

$$f(t) = f_0 \cos(\omega t) = \mathrm{Re}\, f_0 e^{i\omega t} \tag{7.9}$$

then

$$\langle A(t) \rangle = \langle A \rangle_0 + [\mathrm{Re}\, \tilde{\chi}_A(\omega)] f_0 \cos(\omega t) - [\mathrm{Im}\, \tilde{\chi}_A(\omega)] f_0 \sin(\omega t) \tag{7.10}$$

so $f(t)$ oscillates with the frequency of driving with amplitude $f_0 |\tilde{\chi_A(\omega)}|$ and with phase lag $\arg(\tilde{\chi_A(\omega)})$. The power performed by the external field is given by $R = -f(t) \frac{\mathrm{d}}{\mathrm{d}t} A(x(t))$ with the time-average $\langle R \rangle = \frac{1}{2} \omega f_0^2 \,\mathrm{Im}\, \tilde{\chi_A(\omega)}$. Thus, the imaginary part $\mathrm{Im}\, \tilde{\chi}_A(\omega)$ characterises the dissipative response of the system.

Derivation of the fluctuation-dissipation-theorem

Let $f(t) = f_0 \Theta(-t)$. We first compute the partition function $Z_1 = \int \mathrm{d}x \exp\{-\beta H_1\}$ with

$$f_1(x) = \frac{1}{Z_1} \exp\{-\beta H_1\} \approx p_0(x)[1 + \beta f_0 (A(x) - \langle A \rangle_0)] \tag{7.11}$$

For $t \geq 0$ we have

$$\begin{aligned}
\langle A(t) \rangle &= \int \mathrm{d}x\, A(x) p(x,t) = \int \mathrm{d}x\, A(x) \int \mathrm{d}x_0\, P(x,t|x_0,t_0) p_1(x_0) \\
&= \int \mathrm{d}x\, A(x) \int \mathrm{d}x_0\, P(x,t|x_0,t_0) p_0(x_0) [1 + \beta f_0 A(x) - \beta f_0 \langle A \rangle_0] \\
&= \langle A \rangle_0 + \beta f_0 \langle A(t) A(0) \rangle - \beta f_0 \langle A \rangle_0^2 \\
&= \langle A \rangle_0 + \beta f_0 C_A(t)
\end{aligned}$$

We also know that

$$\langle A(t) \rangle = \langle A \rangle_0 + \int_{-\infty}^{\infty} \mathrm{d}\tau\, \chi_A(\tau) f(t - \tau) \tag{7.12}$$

The derivative with respect to time reads

$$\chi_A(t) = \begin{cases} \beta \frac{\mathrm{d}}{\mathrm{d}t} C_A(t) & for \quad t \geq 0 \\ 0 & for \quad t < 0 \end{cases}$$

Remark: Even and Odd Functions

Every function $F(t)$ can be separated into an even and an odd part

$$F(t) = \begin{cases} F'(t) = \frac{1}{2}[F(t) + F(-t)] \,(\text{even}) & \Rightarrow \quad \tilde{F}'(\omega) = \operatorname{Re} \tilde{F}(\omega) \\ F''(t) = \frac{1}{2}[F(t) - F(-t)] \,(\text{odd}) & \Rightarrow \quad \tilde{F}'(\omega) = i \operatorname{Im} \tilde{F}(\omega) \end{cases}$$

Caution: The prime $'$ indicates the even part, not a derivative!

so $C_A(t)$ is even, thus $\frac{\mathrm{d}}{\mathrm{d}t} C_A(t)$ is odd and as we take only the odd parts $\chi_A''(t) = \frac{1}{2}\beta \frac{\mathrm{d}}{\mathrm{d}t} C_A(t)$ and so $i \operatorname{Im} \tilde{\chi}_A(\omega) = \frac{1}{2}\beta(+i\omega)\delta_A(\omega)$

In classical mechanics we have $\frac{1}{\beta} = k_B T$, in quantum mechanics we have $\hbar\omega = \coth \frac{\beta\hbar\omega}{2}$

Example: Optical Trap

An optical trap can be described by

$$kx + \gamma \dot{x} = \gamma \xi(t) \quad \text{with} \quad \langle \xi(t) \rangle = 0 \tag{7.13}$$

The fluctuation dissipation theorem is telling us that

$$S_x(\omega) = \frac{2k_B T}{\omega} \operatorname{Im} \tilde{\chi}_x(\omega) = \frac{2k_B T/\gamma}{(k/\gamma)^2 + \omega^2} \tag{7.14}$$

Thus, one can estimate k by measuring $S_x(\omega)$. As a generalised example, we consider

$$\sum_{k=0}^{n} a_k x^{(k)}(t) = \xi(t) \tag{7.15}$$

for which we do a a Fourier transformation in order to get

$$\underbrace{\sum_{k=0}^{n} a_k (i\omega)^k}_{=\tilde{\chi}_A^{-1}(\omega)} \tilde{\chi}(\omega) = \tilde{\xi}(\omega) \tag{7.16}$$

so

$$\begin{aligned} \tilde{\chi}(t) &= \tilde{\chi}_A(t)\tilde{\xi}(\omega) \\ \chi(t) &= \int_0^\infty \mathrm{d}\tau \, \tilde{\chi_A}(\tau)\xi(t-\tau) \end{aligned}$$

We have

$$\begin{aligned} S_x(\omega)\delta(\omega-\omega') &= \langle \tilde{\chi}(\omega)\tilde{\chi}^*(\omega') \rangle \\ &= \tilde{\chi}_A(\omega)\tilde{\chi}_A^*(\omega') \left\langle \tilde{\xi}(\omega)\tilde{\xi}(\omega') \right\rangle \\ &= |\tilde{\chi}_A(\omega)|^2 2D\delta(\omega-\omega') \end{aligned}$$

so

$$S_x(\omega) = |\tilde{\chi}_A(\omega)|^2 2D = \frac{2k_B T}{\omega} \operatorname{Im} \tilde{\chi}_A(\omega) \tag{7.17}$$

$$2D = \frac{2k_B T}{\omega} \frac{\operatorname{Im} \tilde{\chi}_A(\omega)}{|\tilde{\chi}_A(\omega)|^2} \tag{7.18}$$

For the special case $a_1 = \gamma$, but $(a_{2k+1} = 0)$ for $k > 0$ we have

$$\tilde{\chi}_A = \frac{1}{R(\omega) - i\omega\gamma} \quad \Rightarrow \quad \operatorname{Im} \tilde{\chi}_A = \frac{\omega\gamma}{R^2(\omega) + \omega^2\gamma^2} \tag{7.19}$$

Thus, $D = k_B T \gamma$

8 A Link to Statistical Physics

The fluctuation-dissipation theorem is a hallmark of equilibrium systems. Living systems can violate the fluctuation-dissipation theorem.

8.1 Detailed Balance

A system that can reach equilibrium andhas a zero net current at equilibrium obeys to so-called detailed balance. It means, that it is not possible to distinguish whether a dynamics is played forwards or backwards in time. It defines reversible Markov chains.

We can formulate the condition of detailed balance both for continuous and discrete state space descriptions.

Condition for Detailed Balance

We say the dynamics obeys "detailed balance" if

- there exists an equilibrium distribution P^* or P_j^* in the discrete case
- the joint probability is symmetric $P^*(x', \tau|x, 0) = P^*(x, \tau|x', 0)$ and $P^*(i, \tau|j, 0) = P^*(j, \tau|i, 0)$ / $L_{ji}P_j^* = L_{i,j}P_i^*$, respectively

Its behaviour is governed for continuous state space by the Fokker-Planck equation

$$\frac{\mathrm{d}}{\mathrm{d}t}p(x,t) = \hat{L}p(x,t) \tag{8.1}$$

and for discrete state space by the Master equation

$$\frac{\mathrm{d}}{\mathrm{d}t}P_j(t) = P_i(t)L_{ij} \tag{8.2}$$

This means that there is zero net current at equilibrium $i \rightleftharpoons j$. So for example if the transition matrix is symmetric a system obeys detailed balance.

Example: Boltzmann distribution

A simple example is the Boltzmann distribution for a canonical ensemble, where you have states $0, 1, 2, \ldots$ with energies $E_0, E_1, E_2, \ldots$ so

$$P_i^* = \frac{1}{Z}\exp(-\beta E_i) \tag{8.3}$$

and

$$\frac{L_{ji}}{L_{ij}} = \exp(-\beta(E_i - E_j)) \tag{8.4}$$

A counter example would be a circular current $1 \to 2 \to 3 \to 1$ with rate r giving

$$\hat{L} = \begin{pmatrix} -r & r & 0 \\ 0 & -r & r \\ r & 0 & -r \end{pmatrix} \tag{8.5}$$

with eigenvalue $\lambda_1 = 0$ with corresponding eigenvector $\vec{e}_1 = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)^T$ and $\lambda_2 = \lambda_3^* = \left(-\frac{3}{2} + i\frac{\sqrt{3}}{2}\right) r$ resulting in a net current at equilibrium, thus breaking detailed balance.

Proof of Detailed Balance for Hamiltonian Systems

We consider a system characterised by some Hamiltonian H obeying the Hamilton equations

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \quad \dot{q}_i = \frac{\partial H}{\partial p_i} \tag{8.6}$$

with the macroscopic observable $y = Y(q, p)$. The detailed balance holds if

(i) H is even in p_i

(ii) Y is even in p_i

Then the time propagator T fulfills the condition

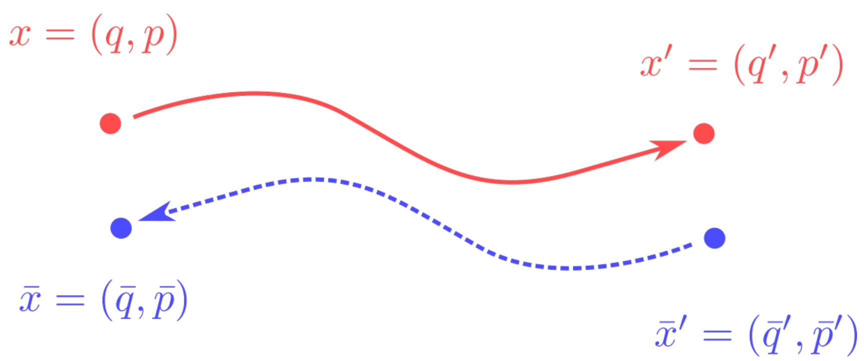
$$P(y', \tau | y, 0) = T_\tau(y'|y)P^*(y) = T_\tau(y|y')P^*(y') = P(y, \tau | y', 0) \tag{8.7}$$

Nota Bene

We always have

$$T_\tau(y'|y)P^*(y) = T_{-\tau}(y|y')P^*(y) \tag{8.8}$$

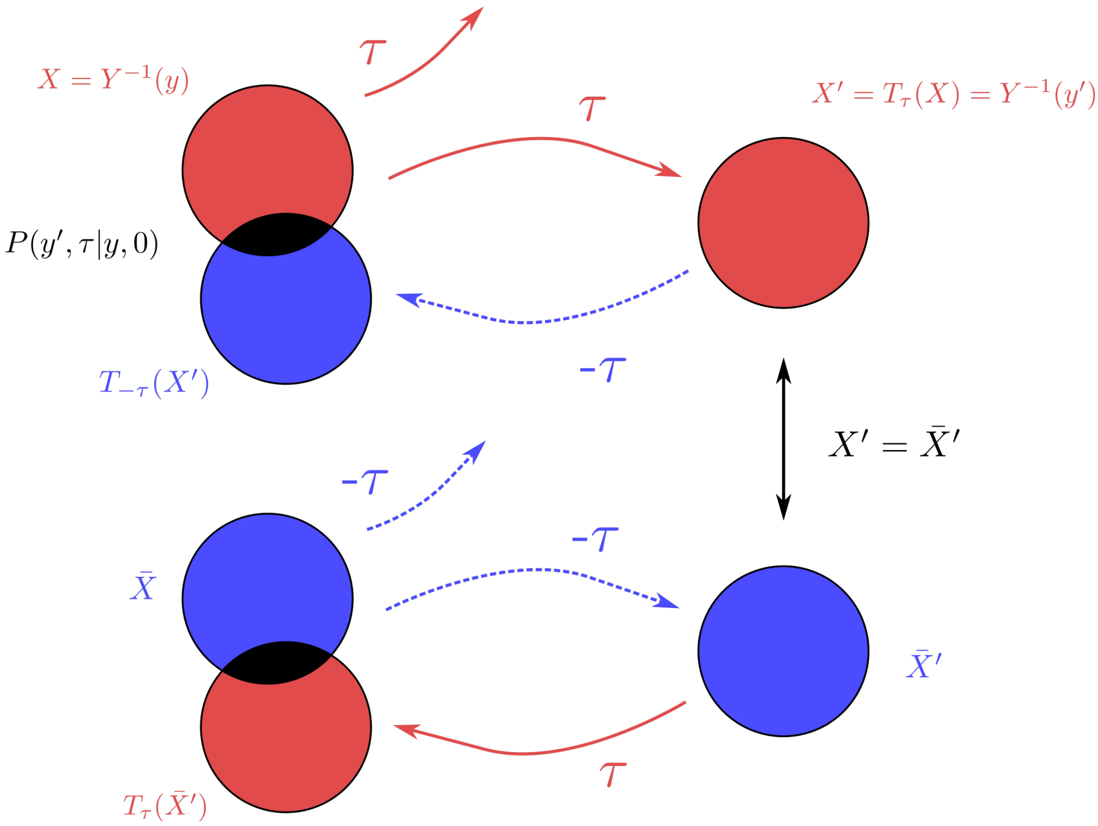
as we can play backwards the dynamics in time.



For the proof of this we make use of time reversal notation

$$\bar{t} = -t \quad \bar{q}_i = q_i \quad \bar{p}_i = -p_i \tag{8.9}$$

We start by looking at a trajectory in (q, p)-phase space and for every point $x' = (q', p')$ we apply time reversal $\bar{x}' = (\bar{q}', \bar{p}')$. By (i) we conclude that $H(x) = H(\bar{x})$ and thus $P^*(x) = P^*(\bar{x})$ (even in p_i means in our case symmetric in time!). Also, we have $X = Y^{-1}(y)$ and by (ii) $X = \bar{X}$ as Y is even.



We can express the probability to observe y' at time τ after observing y at time 0 by the integral over the phase space region $X \cap T_{-\tau}(X')$ of the equilibrium probabilities $P^*(x)$

$$T_\tau(y'|y)P^*(y) = P(y', \tau | y, 0) = \int_{X \cap T_{-\tau}(X')} \mathrm{d}x \, P^*(x)$$

As we have $P^*(x) = P^*(\bar{x})$ we can also change the area of integration in phase space to $\overline{X \cap T_{-\tau}(X')}$

$$\int_{X \cap T_{-\tau}(X')} \mathrm{d}x \, P^*(x) = \int_{\overline{X \cap T_{-\tau}(X')}} \mathrm{d}x \, P^*(x)$$

From the diagram above we see that the states in the lower red circle $T_\tau(\bar{X}')$ are mapped to the states in the upper blue circle $T_{-\tau}(X')$ under time reversal

$$\overline{T_{-\tau}(X')} = T_\tau(\bar{X}') \tag{8.10}$$

and so we have

$$\overline{X \cap T_{-\tau}(X')} = \bar{X} \cap \overline{T_{-\tau}(X')} = \bar{X} \cap T_\tau(\bar{X}') = X \cap T_\tau(X') \tag{8.11}$$

We conclude

$$\int_{\overline{X \cap T_{-\tau}(X')}} \mathrm{d}x\, P^*(x) = \int_{X \cap T_\tau(X')} \mathrm{d}x\, P^*(x) = P(y, \tau | y', 0) = T_\tau(y|y')P^*(y') \tag{8.12}$$

So at equilibrium we cannot distinguish whether a dynamics is played forward or backward in time.

8.2 Increase of Relative Entropy

We consider a Master equation for a Markov chain

$$P_j^{(n+1)} = \sum_i P_i^n T_{ij} \tag{8.13}$$

with the probability P_i^n to be in state i at time $t = t_n$ and a matrix of transition probabilities (T_{ij}) fulfilling $\sum T_{ij} = 1$. For a stationary distribution with $P_j^* > 0$ so that $P_j^* = \sum_i P_i^* T_{ij}$ for all j, we define the relative entropy (Kullberg-Leibler divergence) as

$$D_n = KL(P^n || P^{-k}) = \sum P_i^n \ln\left(\frac{P_i^n}{P_i^*}\right) \tag{8.14}$$

Theorem

$D_{n+1} \leq D_n$

This theorem is a direct consequence of the convexity of D_n. For $P_i^* = \frac{1}{N}$, one finds that $D_n = -\sum P_i \ln(P_i) - \ln(N)$, i.e. the entropy increases with time.

8.3 Equilibrium vs Non-Equilibrium

Signs of Equilibrium

- FDT
- detailed balance $L_{ji}P_j^* = L_{ij}P_i^*$
- equipartition theorem

All of these constrain the Langevin equation to approach thermal equilibrium. At equilibrium we have the Boltzmann distribution as a maximum entropy distribution.

Let us consider a macrostate y with microstates $x = Y^{-1}(y)$. For every probability distribution you can think of we can define the entropy

$$S = \int p(x|y)\ln(p(x|y)) \tag{8.15}$$

that is the relative information of x with respect to y (dimensionless / natural units / units of bits). So $\frac{S}{\ln(2)}$ is the average number of yes / no questions to infer x if only y is known.

In information theory you define a piece of information as 1 nat $= \ln(2)^{-1}$ bits. Rescaling by $\ln(2)$ is equivalent to defining the entropy instead using the binary logarithm.

Properties of Non-Equilibrium Systems

- non-generic steady states possible (e.g. circular currents)
- small number of non-equilibrium fluctuation theorems is available

Derivation of the Boltzmann distribution

Lets consider a system with $\langle E\rangle = U$ for the canonical ensemble, which is contact with some heath bath of temperature T. As a trick we map the it to a microcanonical ensemble of N independent systems with N_i system in an energy state E_i.

Now we have two constraints:

i) $\sum_i N_i = N$

ii) $\sum_i N_i E_i = NU$

Then we introduce the weight function W from statistical physics and count how many compatible microstates there are for a given set $\{N_i\}$. Doing simple combinatorics we get

$$W = \frac{N!}{N_1!N_2!\dots} \tag{8.16}$$

The macrostate with maximum W is most probable, but we need to take into account the constraints. So the way to go is to introduce Lagrange multipliers. We know that

$$0 = \mathrm{d}\ln(W) = \sum_i \frac{\partial \ln(W)}{\partial N_i}\,\mathrm{d}N_i + \alpha \sum \mathrm{d}N_i - \beta \sum E_i\,\mathrm{d}N_i \tag{8.17}$$

so

$$\frac{\partial \ln(W)}{\partial N_i} + \alpha - \beta E_i = 0 \tag{8.18}$$

with

$$\frac{\partial \ln(W)}{\partial N_i} \overset{\text{Stirling}}{=} \frac{\partial N_j \ln(N_j)}{\partial N_i} - \sum_j \frac{\partial N \ln(N)}{\partial N_i} = \dots = -\ln\left(\frac{N_i}{N}\right) \tag{8.19}$$

so $p_i = \frac{N_i}{N} = \exp(\alpha - \beta E_i)$. You can play this game also for the grand canonical ensemble with a third condition for the mean particle number.

8.4 Thermal Fluctuations: Space-dependent Diffusion

Our diffusion coefficient now depends on the position $D = D(x)$ and we write down a Langevin equation

$$\gamma \dot{x} = -\frac{\partial U}{\partial x} + \gamma \xi(t) \tag{8.20}$$

and we have Gaussian white noise with a position-dependent noise strength $\langle \xi(t)\xi(t')\rangle = 2D(x)\delta(t-t')$. So far we can't tell whether to use Stratonovich or Ito calculus.

Starting with the Einstein relation $D = \frac{k_B T}{\gamma}$ we can distinguish two cases:

i) $\gamma = \gamma(x)$ and $T = T_0$ (equilibrium system, just passive obstacles by γ!)

ii) $\gamma = \gamma_0$ and $T = T(x)$ (non-equilibrium system!)

α-calculus

$(\alpha)\,\dot{x} = f(x) + \alpha g(x)$ with $\alpha = 0$ for Ito and $\alpha = \frac{1}{2}$ for Stratonovich. This determines what the Fokker-Planck equation looks like

$$\dot{P} = \frac{\partial}{\partial x}\left[-(f + \frac{\alpha}{2}\frac{\partial^2 g}{\partial 2^2})P + \frac{1}{2}\frac{\partial}{\partial x}(g^2 P)\right] \tag{8.21}$$

Case i): Equilibrium System

For case i) it can be shown that $\alpha = 1$ is correct (isothermal interpretation) (see Lau & Lubensky paper PRE).

$$\dot{P} = \frac{\partial}{\partial x}\left[-(f + \frac{\partial^2 g}{\partial 2^2})P + \frac{1}{2}\frac{\partial}{\partial x}(g^2 P)\right] \tag{8.22}$$

with $f = -\frac{1}{\gamma}\frac{\partial U}{\partial x}$, $g = \sqrt{2D}$ so

$$\dot{P} = -\vec{\nabla}(fP) + \vec{\nabla}(D(x)\vec{\nabla}P) \tag{8.23}$$

so this is the correct generalization of Fick's law for equilibrium systems with $\gamma = \gamma(x)$. So at the steady state we must have $\dot{P}^* = 0$ and $fP) + D(x)\vec{\nabla}P = \text{const}$ and we get the Boltzmann distribution $P^* \sim \exp(-\beta U)$.

Nota Bene

In Ito interpretation with $\alpha = 0$ we would have obtained the Fokker-Planck equation

$$\dot{P} = -\vec{\nabla}(fP) + \vec{\nabla}^2(DP) \tag{8.24}$$

which gives rise to different physics and different steady state distributions. So it is really important where to write D!

Case ii): Position-dependent Temperature

These systems behave non-generically and are characterized by thermophoreses (Sorret effect) and the Fokker-Planck equation

$$\dot{P} = -\vec{\nabla}(f - D_T\vec{\nabla}T)P + \vec{\nabla}(D(x)\vec{\nabla}P) \tag{8.25}$$

with the Sorret coefficient $S_T = \frac{D_T}{D}$, which depends on molecular interaction potentials. The Sorret effect is an example of an non-equilibrium phenomenon.

8.5 Entropy Production

A hallmark of non-equilibrium systems is entropy production, which we know from the second law of thermodynamics $\langle \Delta S \rangle \geq 0$.

As an example lets consider to systems of temperature T_1 and T_2 that are brought into contact resulting in an entropy change

$$\frac{\Delta S}{k_B} = \Delta Q(\frac{1}{k_B T_2} - \frac{1}{k_B T_1}) \tag{8.26}$$

Or consider a second example of two ideal gases with N_A and N_B particles that are mixing: $N = N_A + N_B$, $x = \frac{N_A}{N}$ and

$$\frac{\Delta S}{k_B} = -N[x\ln(x) + (1-x)\ln(1-x)] \tag{8.27}$$

as the mixing entropy.

Fluctuation Theorem of Non-Equilibrium Systems

The change of entropy will be a stochastic variable with

$$\frac{P\left(\frac{\Delta S}{k_B} = I\right)}{P\left(\frac{\Delta S}{k_B} = -I\right)} = \exp(I) \tag{8.28}$$

for a colloidal particle driven by optical tweezers a decrease in entropy might be observed, for larger system it gets really unlikely due to the exponential factor $\exp(I)$.

Crooke Fluctuation Theorem

Lets consider a macroscopic observable y so we can ask how likely it is to observe the time-reversed macroscopic dynamics / trajectory

$$\frac{P(y)}{P(\bar{y})} = \exp(\Delta S[y(t)]/k_B) \tag{8.29}$$

Jarzynski Relation

For an isolated system there is always an adiabatic process connecting y_0 and y_1 costing the work $W = \Delta F$ as change in the free energy. Any non-adiabatic process will take more work $\langle \Delta F \rangle \leq W$ and we have

$$\exp(-\Delta F) = \langle \exp(-\beta W) \rangle \tag{8.30}$$