Lecture Notes on Stochastic Processes

PD Dr. Benjamin M. Friedrich

November 14, 2018
# Contents

## 1 Diffusion & Random Walk
1.1 Random Walker ................................................. 1
1.2 Continuum Limit: Diffusion Equation ......................... 2
1.3 Random Force Model .......................................... 4

## 2 Probability Theory
2.1 Mathematical Foundations ...................................... 9
2.2 Probability in Physics ....................................... 10
2.3 Important Probability Distributions ......................... 11
2.4 Normal Approximation ...................................... 13
2.5 Stochastic Processes ....................................... 17

## 3 Langevin Equation and Fokker-Planck Equation
3.1 Langevin equation ........................................... 19
3.2 Fokker-Planck-Equation .................................... 20
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

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

for averaging over a certain probability distribution.

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

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

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

$$\sigma^2 = \langle \vec{x}^2 \rangle - \langle \vec{x} \rangle^2 = \langle \vec{x}^2 \rangle = (\Delta \vec{x})^2$$

$$= \left\langle \left( \sum_{n=1}^{N} \vec{x}_n \right)^2 \right\rangle = \sum_{n,m=1}^{N} \langle \vec{x}_n \cdot \vec{x}_m \rangle = Nl^2$$

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 $\Delta \vec{y}$:

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

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

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

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

(1.5)

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

$$P_N(\vec{x}) \approx \int d^d \Delta y P(\Delta \vec{y}) \left[ P_{N-1}(\vec{x}) - \Delta y_i \partial_i P_{N-1}(\vec{x}) + \frac{1}{2} \Delta y_i \Delta y_j \partial_i \partial_j P_{N-1}(\vec{x}) \right]$$

$$= P_{N-1}(\vec{x}) + \frac{\langle (\Delta \vec{y})^2 \rangle}{2d} \nabla^2 P_{N-1}(\vec{x})$$

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

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

(1.6)

and now we can take the limit

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

(1.7)

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

**Diffusion Equation**

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

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

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

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

**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

\[
p(\vec{k}, t) = \mathcal{F}(p(\vec{x}, t)) = \exp(-D\vec{k}^2 t) \tag{1.11}
\]

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 Newton’s law

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

(1.13)

For long time scales $\tau_m \gg \frac{m}{\gamma}$, 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 $\langle f(t)f(t + \tau) \rangle \rightarrow 0$ for $\tau \rightarrow \tau_m$

for averaging over a certain probability distribution.

An important property of the random force is stationarity

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

(1.14)

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

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

$$x(t) = x(0) + \frac{1}{\gamma} \int_{0}^{t} dt_1 f(t_1)$$

(1.15)
1.3 Random Force Model

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

$$\langle \Delta x(t) \rangle = \frac{1}{\gamma} \int_0^t dt_1 \langle f(t_1) \rangle = 0$$ (1.16)

by symmetry and for the mean square deviation

$$\langle \Delta x^2 \rangle = \frac{1}{\gamma^2} \left( \left( \int_0^t dt_1 f(t_1) \right) \left( \int_0^t dt_2 f(t_2) \right) \right)$$

$$= \frac{1}{\gamma^2} \int_0^t dt_1 \int_0^t dt_2 \langle f(t_1)f(t_2) \rangle$$

$$= \frac{1}{\gamma^2} \int_0^t dt_1 \int_{-t_1}^{t+t_1} d\tau \langle f(t_1)f(t_1+\tau) \rangle$$

$$= \frac{1}{\gamma^2} \int_0^t dt_1 \int_{-\infty}^{\infty} d\tau \langle f(t_1)f(t_1+\tau) \rangle + \mathcal{O}(D\tau_m)$$

$$= \frac{1}{\gamma^2} \int_0^t dt_1 \gamma^2 2D = 2Dt$$

Calculating $D = D(T)$

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

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

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

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

The solution to this scenario is given by

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

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

$$x(t) = \int_0^\infty d\tau f(t-\tau) \chi(\tau)$$ (1.20)
with $\langle x(t) \rangle = 0$ by symmetry and

$$
\langle \Delta x^2 \rangle = \frac{1}{\gamma^2} \left( \int_0^\infty d\tau_1 f(t-\tau_1)\chi(\tau_1) \right) \left( \int_0^\infty d\tau_2 f(t-\tau_2)\chi(\tau_2) \right)
$$

$$
= \int_0^\infty d\tau_1 \int_0^\infty d\tau_2 \langle f(t-\tau_1) f(t-\tau_2) \rangle \frac{\chi(\tau_1)\chi(\tau_2)}{\gamma^2} \exp(-\frac{\tau_1 + \tau_2}{\sigma})
$$

$$
= \int_0^\infty d\tau_1 \int_{-\infty}^{\tau_1} d\tau \langle f(t-\tau_1) f(t-\tau_1 - \tau) \rangle \frac{1}{\gamma^2} \exp\left(-\frac{2\tau_1 + \tau}{\sigma}\right)
$$

$$
= \int_0^\infty d\tau_1 \int_{-\infty}^{\tau_1} d\tau \langle f(t-\tau_1) f(t-\tau_1 - \tau) \rangle \frac{1}{\gamma^2} \exp\left(-\frac{2\tau_1 + \tau}{\sigma}\right) + \mathcal{O}(D\tau_m)
$$

$$
\approx \frac{1}{\gamma^2} \int_0^\infty d\tau_1 \exp\left(-\frac{2\tau_1}{\sigma}\right) \int_{-\infty}^{\infty} d\tau \langle f(t-\tau_1) f(t-\tau_1 - \tau) \rangle \exp\left(-\frac{\tau}{\sigma}\right)
$$

$$
= \frac{1}{\gamma^2} \frac{\sigma}{2} 2D\gamma^2 = \frac{\gamma}{k} D
$$

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

$$
\langle k_x^2 \rangle = \frac{k_B T}{2} \quad \tag{1.21}
$$

As we have

$$
\langle k_x^2 \rangle = \frac{k_B T}{2} D = \frac{k_B T}{2} \quad \tag{1.22}
$$

we obtain the Stokes-Einstein-relation

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

**Repetition: Equipartition Theorem**

In thermal equilibrium, the systems energy, given by a Hamiltonian $H$, is distributed

Made with LaTeX by Benjamin Wolba
on its degrees of freedom $x_n$ via

$$\left< x_m \frac{\partial H}{\partial x_n} \right> = \delta_{mn} k_B T$$  \hspace{1cm} (1.24)

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

2.1 Mathematical Foundations

Concept: Probability

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

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

The probability density is analogously a function $p(x) : \mathbb{R} \to \mathbb{R}^+$ and it relates to the probability itself by

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

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

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

Important properties 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

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

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

(2.5)

2.2 Probability in Physics

Usually, probability is regarded as relative frequency of an event \( A \) occurring \( N_A \) times for the total number of measurements being \( N \)

\[
P(A) = \lim_{N \to \infty} \frac{N_A}{N}
\]

(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? \( \rightarrow \) 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: \( P(R \mid \text{Dresden}) = \frac{116}{365} = 18\% \). 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: \( P(R \mid \text{Dresden, October}) = \frac{8}{31} = 25.8\% \).

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 \( P(R \mid \text{current local weather}) = x \) and \( P(R \mid \text{rain today}) = 44\% \), \( P(R \mid \text{dry today}) = 17\% \).

Last but not least we can sample macrostate that is consistent with measurement data and calculate the probabilities for rain from deterministic models (Navier-Stokes-equations / mathematical forecasting) \( P(R \mid \text{current global weather}) \).
2.3 Important Probability Distributions

Normal Distribution

The normal distribution is given by

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

with the moments and cumulants being

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

Bernoulli Distribution

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

\[ P(k, n) = \binom{n}{k} p^k (1-p)^{n-k} \]  \hspace{1cm} (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. Therefore we introduce a time \( t_j = \frac{j}{n}T = j \Delta t \) with \( \Delta t = \frac{T}{n} \) and \( \lambda = np \) being the total number of expected events. The event rate is given by \( r = \frac{\lambda}{T} = \frac{p}{\Delta t} \) with \( [r] = s^{-1} \).

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

\[ p(k, \lambda) = \exp(-\lambda) \frac{\lambda^k}{k!} \]  \hspace{1cm} (2.9)

with \( \mu = \langle k \rangle = \lambda, \sigma^2 = \langle k^2 \rangle - \langle k \rangle^2 = \lambda \). An approximation is \( p(k, \lambda) = N(\lambda, \lambda) \) for very large \( \lambda \).
Remark: Why is the concept of time being used here?

The Poisson distribution is an example of a stochastic Poisson process

\[ f(t) = \sum_{-\infty}^{\infty} \delta(t - t_j) \quad (2.10) \]

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

Power-law distribution

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

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

It has some unpleasant properties such as \( \sigma = \infty \) for \( \alpha < 3 \).
2.4 Normal Approximation

Proof I Using Stirling’s Approximation

We would like to approximate the Bernoulli distribution

\[ p(k, n) = \binom{n}{k} p^k (1-p)^{n-k} \]  \hspace{1cm} (2.12)

by means of a normal distribution. Therefore we introduce a small deviation \( \varepsilon \) such that \( q = 1-p, k = np + n\varepsilon \) and \( p(k, n) \approx 0 \) for \( \varepsilon \gg \frac{1}{\sqrt{N}} \).

Trick number one in order to continue is to use Stirling’s approximation

\[ n! \approx \sqrt{2\pi n} \left( \frac{n}{e} \right)^n \]  \hspace{1cm} (2.13)

which leads us to

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

To do the second trick and apply \( x^k = \exp(k \ln(x)) \) 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 means
\[
\left( \frac{np}{k} \right)^k \left( \frac{nq}{n-k} \right)^{n-k} \approx \exp \left( k \left[ -\frac{\varepsilon}{p} + \frac{1}{2} \left( \frac{\varepsilon}{q} \right)^2 \right] + (n-k) \left[ \frac{\varepsilon}{p} - \frac{1}{2} \left( \frac{\varepsilon}{q} \right)^2 \right] \right)
\]
\[
= 0 \cdot \varepsilon - \frac{1}{2} n \varepsilon^2 - \frac{1}{2} n \varepsilon^2 + O(\varepsilon^3)
\]
\[
= -\frac{1}{2} \frac{\varepsilon^2 (p+q)}{pq} = -\frac{1}{2} \frac{(k-np)^2}{n pq}
\]

Thus

\[
p(k, n) = \frac{1}{\sqrt{2\pi \sigma^2}} \exp \left( -\frac{(k-np)^2}{2\sigma^2} \right)
\]

(2.14)

with \( \sigma^2 = npq \)

**Proof II Using the Central-Limit-Theorem**

**Central-Limit-Theorem**

Consider sequence of \( x_1, \ldots, x_n \) independent, identically distributed, random variables with mean \( \mu \) and variance \( \sigma^2 \). We define the empirical mean by

\[
x = \frac{1}{n} (x_1 + \cdots + x_n) = \text{empirical mean}
\]

(2.15)

We normalize it to a random variable with expectation value zero

\[
z = \frac{x - \mu}{\sigma / \sqrt{n}}
\]

(2.16)

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 \( z \in \mathbb{R} \)

As a second, more elegant proof we consider \( n \) independent random variables \( x_j \) with \( j = 1, \ldots, n \) and

\[
x_j = \begin{cases} 
1 & \text{with probability } p \\
0 & \text{with probability } q = 1 - p
\end{cases}
\]

(2.17)
2.4 Normal Approximation

As we define the empirical mean via

\[ \bar{x} = \frac{1}{n}(x_1 + \cdots + x_n) = \text{empirical mean} \]  

with \( k = n\bar{x} \) we get \( p(k, n) = p(\bar{x}) \sim N(np, npq) \) by the Central-Limit-Theorem. The idea of the proof is to compute the cumulants of \( x_j \)

\[ k_1 = \mu, \; k_2 = \sigma^2 \ldots \]  

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

\[ k_1 = 0, \; k_2 = 1, \; k_3 \sim \frac{1}{\sqrt{n}}, \; k_4 \sim \frac{1}{n} \ldots \]  

and

\[ \lim_{n \to \infty} \ln \langle zt \rangle = \lim_{n \to \infty} \sum_{l=0}^{\infty} k_lt^l = 1 - \frac{1}{2}t^2 \]  

then we have \( z \to N(0, 1) \). To show this behavior of the cumulants we take a look at the functions

\[ C_x(t) = \langle xt \rangle \quad \text{with} \quad C_{\alpha x}(t) = C_x(\alpha t), \; \alpha \in \mathbb{R} \]  

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

\[ C_{\bar{x}}(t) = C_{x_1} \left( \frac{t}{n} \right) C_{x_2} \left( \frac{t}{n} \right) \cdots C_{x_n} \left( \frac{t}{n} \right) = C \left( \left( \frac{t}{n} \right)^n \right) \]

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

giving us

\[ \ln(C_z(t)) = n \ln \left( C_x \left( \frac{t}{\sigma/\sqrt{n}} \right) \right) - \frac{\mu t}{\sigma/\sqrt{n}} \]  

Made with \texttt{B\TeX} by Benjamin Wolba
and so we get

\[ k_{z,j} = n \left( \frac{1}{\sigma \sqrt{n}} \right)^j k_{x,j} \]  

(2.24)
2.5 Stochastic Processes

Stochastic Process

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

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 \( t_j \)

\[
f(t) = \sum_{j=-\infty}^{\infty} \delta(t - t_j)
\]

(2.25)

The property \( x = \int_{\theta}^{T} f(t) \) counts events and yields a Poisson distribution for \( \lambda = rt \).
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

\[
P(t \geq \theta + dt) = P(t \geq \theta) - r \, dt \, P(t \geq \theta)
\]

so that

\[
\frac{d}{dt} P(t \geq \theta) = -r \, dt \, P(t \geq \theta) \quad \Rightarrow \quad P(t \geq \theta) \sim \exp(-rt)
\]
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) $\int_{t_1}^{t_2} dt \xi(t) \sim N(0, 2D[t_2 - t_1])$

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

Remark: Gaussian White Noise and Mathematics

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

$$W(t) = \int_{0}^{t} dt' \xi(t') \quad (2.26)$$

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

\[ \dot{x} = f(x) + \sqrt{2D} \xi(t) \]  

(3.1)

- \( \xi(t) \) represents Gaussian white noise
- it describes diffusion in an effective potential \( U(x) = - \int_0^x dx' f(x') \)

**Generalization**

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

(3.2)

with \( i = 1, \ldots, n \) and \( \xi_j(t) \) being independent Gaussian white noise functions \( \langle \xi_j(t) \xi_i(t') \rangle = \delta_{ji} \delta(t - t') \)

**Example 1: Double-well Potential**

**Example 2: Escape over a Barrier**

**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

$$\hat{x}_{n+1} = \hat{x}_n + f(\hat{x}_n) \, dt + \sqrt{2D \, dt} \, N_n$$

(3.3)

with $D = D_0$, $t_i = i \, dt$, $x_i = x(t_i)$, $N_n \sim N(0,1)$ and $|\hat{x}_n - x_n| \sim \mathcal{O}(dt^{3/2})$

3.2 Fokker-Planck-Equation

Derivation of Fokker-Planck-Equation

Repetition: Ordinary Diffusion

For the example of ordinary diffusion

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

(3.4)

the probability density is given by

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

(3.5)

fullfilling the Diffusion equation

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

(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

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

(3.7)

Therefore we discretize time and take a look how a sub-ensemble of $p(x, t)$ at $x_n$ will evolve
### 3.2 Fokker-Planck-Equation

during a time step from \( p(x, t_n) \) to \( p(x, t_{n+1}) \). For this we are using the Markov-Property:

\[
p(x, t_{n+1}|x_0, t_0) = \int dx_n p(x, t_{n+1}, x_n, t_n|x_0, t_0) \\
= \int dx_n p(x, t_{n+1}|x_n, t_n) p(x_n, t_n|x_0, t_0) \\
= \int dx_n N(x_n + f(x_n), 2Ddt)p(x_n, t_n|x_0, t_0)
\]

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 few 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 have units of inverse length squared

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

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

\[
p(x, t_n) = \int dx_n I(x_n, y)|_{y=x-n} \quad \text{with} \quad I(x_n, y) = p(x_n)n(x_n, y), \\
\text{and} \quad n(x_n, y) = N(f(x_n)dt, 2Ddt)
\]

The integrand \( I(x, y) \) will contribute only for small \( y = O(dt) \), which means \( x_n \approx x \), so we can Taylor expand \( I(x_n, y) \) in \( x_n \) around \( x \):

\[
I(x_n, y) = I(x, y) + \frac{\partial I(x_n, y)}{\partial x_n}|_{x_n=x}(x_n-x) + \frac{\partial^2 I(x_n, y)}{\partial x_n^2}|_{x_n=x}\frac{(x_n-x)^2}{2} \tag{3.8}
\]
Inserting this into the convolution integral leads to

\[ p(x, t_{n+1}) = \int \mathrm{d}y \ I(x, y)|_{x_n=x-y} \]

\[ = \int \mathrm{d}y \left( p(x)n(x, y) - \frac{\partial}{\partial x} \left( p(x) n(x, y) y \right) + \frac{\partial^2}{\partial x^2} \left( p(x) n(x, y) \frac{y^2}{2} \right) \right) \]

\[ = p(x) \int \mathrm{d}y \ n(x, y) - \frac{\partial}{\partial x} \left( p(x) \int \mathrm{d}y \ n(x, y) y \right) + \frac{\partial^2}{\partial x^2} \left( p(x) \int \mathrm{d}y \ n(x, y) \frac{y^2}{2} \right) \]

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

\[ \int \mathrm{d}y \ n(x, y) = 1 \]

\[ \int \mathrm{d}y \ n(x, y) y = f(x) \ dt \]

\[ \int \mathrm{d}y \ n(x, y) \frac{y^2}{2} = D \ dt + \frac{1}{2} [f(x)]^2 = D \ dt + \mathcal{O}(dt^2) \]

which gives us

\[ p(x, t_{n+1}) = p(x, t_n) - \frac{\partial}{\partial x} [p(x, t_n) f(x) \ dt] + \frac{\partial^2}{\partial x^2} [p(x, t_n) D] \ dt \]  \hspace{1cm} (3.9) \]

and thus

\[ \frac{p(x, t_{n+1}) - p(x, t_n)}{dt} = - \frac{\partial}{\partial x} [p(x, t_n) f(x)] + \frac{\partial^2}{\partial x^2} [p(x, t_n) D] \]  \hspace{1cm} (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, which reads

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

\hspace{1cm} (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!).

Made with \texttt{\LaTeX} by Benjamin Wolba
3.2 Fokker-Planck-Equation

Application to the Diffusion Potential

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

$$\dot{x} = -\frac{1}{\gamma} \frac{\partial U}{\partial x} + \xi$$  \hspace{1cm} (3.12)

and look for the steady state $\frac{\partial U}{\partial t} = 0$. Hence, the Fokker-Planck equation reads

$$0 = \nabla\left[\frac{1}{\gamma} \nabla U p\right] + D \nabla^2 p = \nabla\left[\frac{1}{\gamma} \nabla U p + D \nabla p\right]$$

$$\Rightarrow c = \frac{1}{\gamma} \nabla U p + D \nabla p$$

thus, if $c = 0$, we get

$$\frac{\partial}{\partial x} \ln p = \frac{\nabla p}{p} = -\frac{1}{\gamma} \frac{\nabla U}{p}$$  \hspace{1cm} (3.13)

and

$$p \sim \exp\left(-\frac{U}{\gamma D}\right) = \exp\left(-\frac{U}{k_B T}\right)$$  \hspace{1cm} (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

$$\dot{p} = -\nabla J \quad \text{with} \quad J = \frac{1}{\gamma} \nabla U p + D \nabla p$$  \hspace{1cm} (3.15)

of the current $J$. At equilibrium, the current must vanish and thus we have

$$\lim_{t \to \infty} J = c = 0$$  \hspace{1cm} (3.16)
3 Langevin Equation and Fokker-Planck Equation

Eigenvalue Spectrum of $\hat{L}$

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

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

through

$$p(x,t) = \sum a_n \phi_n(x) \exp(\lambda_n t) \quad (3.18)$$

If $\lambda_0 = 0$, then this corresponds to a steady state $\phi_0$ and the slowest decaying mode determines hopping rates.

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

$$\left< \hat{L} g, h \right> = \int dx (\hat{L} g) h = \int dx g \hat{L}^* h = \left< g, \hat{L}^* h \right> \quad \forall g(x), h(x) \quad (3.19)$$

so by partial integration we see that

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

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

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

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

Backward Fokker-Planck Equation

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

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

**Boundary Conditions Matter**

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

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

\[
J(x_1) = J(x_2) = 0 \quad (3.23)
\]

and

\[
\int_{x_1}^{x_2} dx \, p(x, t) = 1 \quad (3.24)
\]

so the steady-state distribution \( p^*(x) = \phi_0(x) \) exists. This is similar for a confinement potential \( \lim_{x \to x_1, x_2} U(x) \to \infty \)

2) Absorbing Boundary Conditions

We have \( p(x_2, t) = 0 \) (Dirichlet Boundary Conditions) and therefore

\[
0 > \frac{d}{dt} \int dx \, p(x, t) = \int dx \, \frac{dp(x, t)}{dt} = \int_{-\infty}^{x_2} \frac{\partial J}{\partial x} = -J(x_2) \quad (3.25)
\]

So 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.