Models of Theoretical Physics

Prof. *Amos Maritan*, *Marco Baiesi*
University of Padova

Written by: *Francesco Manzali*

Academic Year 2019/20

Compiled on October 13, 2020

Contents

Introduction

Francesco Manzali, 20/02/2019

Part I

Maritan's Lectures

Chapter 1

The Diffusion problem

(Lesson 1 of 3/10/2019) Compiled: October 13, 2020

1.1 Introduction

In classical mechanics, if we know all forces $\boldsymbol{F}$ that act on a certain particle, along with its initial condition (e.g. position $\boldsymbol{x}(t=0)$ and velocity $\boldsymbol{v}(t=0)$), we can compute its trajectory $\boldsymbol{x}(t)\,\forall t$ by integrating the equations of motion.
This is indeed true even for ensembles of particles - but it becomes very impractical for macroscopic objects. For example, a drop of water contains something in the order of 10^{23} molecules, and so to completely describe its motion it is needed to integrate six times that many equations (3 for position, 3 for velocity for each single particle). Even if we had the computational capacity to do so, it would not be possible to know the necessary initial conditions with the required precision.

*The **goal** of statistical mechanics*

On the other hand, it is not very interesting to solve this kind of problem, because one could not possibly understand the intricacy of this motion, and so the task doesn't give much insight in the relevant physics. In fact, often we are most interested in the *macroscopic* properties of the object. That is the aim of *statistical mechanics*.

1.2 Diffusion: a macroscopic approach

The diffusion problem

In this course, we will examine one of the most important problems in statistical mechanics: the **diffusion problem**. Take a drop of ink immersed in water and it will mix over time, apply heat to the edge of a bar and it will propagate to the entire thing. Spray a bit of perfume and it will spread over the entire room, place a sugar cube in a cub of tea and it will dissolve without the need of stirring it. The diffusion mechanism is key to many aspects of everyday life, and it is yet one of the most striking *direct* consequences of the invisible microscopical motion of molecules. Thus, studying diffusion can provide a link between these two very different worlds.
The first advances in the analysis of the diffusion motion were made in the 19th century, and were all based on a macroscopic approach. For example Fick's Law, that roughly motivates diffusion as the *motion of fluids from regions of high concentrations to regions of low concentration*, dates back to 1856.

The link with the microscopical world, however, was made only in 1905, in a groundbreaking article on Brownian motion published by Einstein - which also served as a striking proof of the atomic nature of matter.

Brownian motion

Brownian motion is the *erratic motion* exhibited by granules of fine powder when suspended in still water. It was already known that this was due to physical reasons, as repeated experiments ruled out every possible explanation requiring living organism.
Einstein proposed a solution based on molecules, and *statistics*. If we assume that water is composed of particles, the single grains of powder behave like large objects hit by smaller particles. The number of hits on each side is almost the same, so the total force which acts on the large object is **almost** 0. However, if the grains are sufficiently small, the slight unbalance in the number of collisions can produce a significant acceleration, leading to a kind of *random motion*.

Motion due to statistical fluctuations

For example, let's consider a spherical grain submerged in the liquid. Let's call U the upper hemisphere, and L the lower one. Denote with $\bar{N}_c$ the average number of collisions per second per surface unit. Then the number of hits on U is *almost* the same to that of L, up to a certain (binomial) error:

$$\bar{N}_c \cdot U = \bar{N}_c \cdot L \pm \sqrt{\bar{N}_c}$$

Thus, the relative error is given by:

$$\frac{\sqrt{\bar{N}_c}}{\bar{N}_c} = \frac{1}{\sqrt{\bar{N}_c}}$$

Note that if the grains are small, $\bar{N}_c$ will be small too, and so the relative error will be high.

1.2.1 The diffusion equation

Let's try to give a quantitative description of this kind of motion. We start by specifying the *initial conditions* as a *starting distribution*, i.e. a function $\rho\colon \mathbb{R}^3 \times \mathbb{R} \to R$ such that $\rho(\boldsymbol{r}, t)$ is the probability to find a particle in position $\boldsymbol{r}$ at the instant t.

Starting distribution

1. For a discrete, point particle we have $\rho(\boldsymbol{r}, 0) = \delta^3(\boldsymbol{r} - \boldsymbol{r}_0)$, i.e. the particle is at the starting position with certainty.

2. For some quantity of matter (for example a droplet of ink), we have some uniform initial density, such as:

$$\rho(\boldsymbol{r}, 0) = \rho_0(\boldsymbol{r}) = \begin{cases} \bar{\rho}_0 & |\boldsymbol{r}\,| < R \\ 0 & \text{otherwise} \end{cases}$$

Note that $\rho(\boldsymbol{r}, t)$ is a probability density, and not a usual density of matter. The difference is merely of normalization. If N is the total number of particles in ink, then $N\rho(\boldsymbol{r}, t)$ is the density of ink particles at *the specific* position $\boldsymbol{r}$ and time t, which will be denoted with $\rho_n(\boldsymbol{r}, t)$:

$$1 = \int_V \mathrm{d}^3 r\, \rho(\boldsymbol{r}, t)$$
$$N = \int_V \mathrm{d}^3 r\, \underbrace{N\rho(\boldsymbol{r}, t)}_{\text{density at } \boldsymbol{r}}$$

The meaning of a point-wise density can be understood as a limit:

Point-wise density

$$N\rho(\boldsymbol{r}, t) = \text{density at } \boldsymbol{r}, \text{ time } t = \lim_{\Delta V \downarrow 0} \frac{\Delta N}{\Delta V}$$

Consider a patch of liquid of volume ΔV, that contains a number ΔN of ink particles. By letting it shrink "enough", $\Delta N/\Delta V$ reaches a constant value - that is the density in a macroscopically small patch of liquid. Of course, ΔV cannot reach 0, because in that case $\Delta N = 0$. So, the limit is to be interpreted in a macroscopical sense (ΔV is macroscopically vanishing, $\Delta V \downarrow 0$) and not in a mathematical sense ($\Delta V \to 0$).

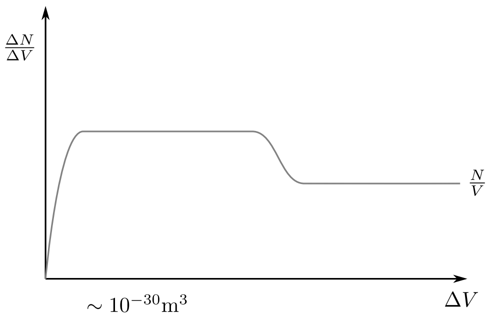


Figure (1.1) – Density (ratio $\Delta N/\Delta V$) as function of patch size ΔV for a region centered around the ink distribution ρ_0 ($|\boldsymbol{r}| < R$ at $t = 0$). If ΔV is sufficiently large, the patch comprises also some space without ink, and so the density is lower.

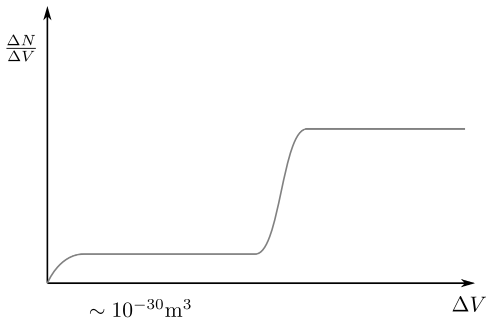


Figure (1.2) – Density for a patch centered on a point $|\boldsymbol{r}| > R$. Here the density is higher for high ΔV, as in these cases the patch comprises also the ink's initial distribution (ρ_0).

We now want to compute $\rho(\boldsymbol{r}, t)$ for $t > 0$, given $\rho(\boldsymbol{r}, 0)$.
We start by considering the *continuity equation*. The idea is that particles do not move by "jumping" between far positions, but travel in a *continuous way*.

Continuity equation

Consider a box of volume V, that contains a fixed number N of particles, with (matter) density:

$$N\rho(\boldsymbol{r}, t) \equiv \rho_n(\boldsymbol{r}, t)$$

Let A be a patch of V, with boundary ∂A. The number of particles inside A at time t is given by the integral of the density:

$$\int_A \mathrm{d}^3 r\, \rho_n(\boldsymbol{r}, t) = N_A(t) \tag{1.1}$$

And at a later time $t + \Delta t$:

$$\int_A \mathrm{d}^3 r \, \rho_n(\boldsymbol{r}, t + \Delta t) = N_A(t + \Delta t) \tag{1.2}$$

Let's introduce a new quantity, the *current* $\boldsymbol{j}(\boldsymbol{r}, t)$ at position $\boldsymbol{r}$ and time t. Consider a small area $\mathrm{d}S$ centered on a point $\boldsymbol{r}$, with $\hat{n}(\boldsymbol{r}) \perp \mathrm{d}S$. The number of particles flowing through $\mathrm{d}S$ during an interval Δt is defined as:

$$\Delta t \boldsymbol{j}(\boldsymbol{r}, t) \cdot \hat{n}(\boldsymbol{r}) \, \mathrm{d}S$$

and this can be used to compute $\boldsymbol{j}$.
For example, for a uniform flow of particles with density ρ_n and velocity $\boldsymbol{v}$, the *current* is $\boldsymbol{j} = \rho_n \boldsymbol{v}$.

Returning to the problem, we note that the *change* of N_A over time is explained by the **flux** of particles through the closed boundary ∂A, i.e. the surface integral of the current $\boldsymbol{j}$:

$$N_A(t + \Delta t) - N_A(t) = -\int_{\partial A} \mathrm{d}S \, \hat{n} \cdot \boldsymbol{j}(\boldsymbol{r}, t) \Delta t \tag{1.3}$$

Here we define, by convention, the sign of $\boldsymbol{j}(\boldsymbol{r}, t)$ to be positive if the current is *outward*, that is from A to $V \setminus A$. So, a positive current means that particles are *leaving* A, and this explains the $-$ in (1.3).

Substituting (1.1) and (1.2) in (1.3) we arrive at:

$$\int_A \mathrm{d}^3 r \, \frac{1}{\Delta t} \left[\rho_n(\boldsymbol{r}, t + \Delta t) - \rho_n(\boldsymbol{r}, t)\right] = -\int_{\partial A} \mathrm{d}S \, (\boldsymbol{r}) \boldsymbol{j}(\boldsymbol{r}, t) \cdot \hat{n}(\boldsymbol{r}) \cancel{\Delta t}$$

Taking the limit $\Delta t \to 0$:

$$\int_A \mathrm{d}^3 r \, \frac{\partial}{\partial t} \rho_n(\boldsymbol{r}, t) = -\int_{\partial A} \mathrm{d}S \, \hat{n} \cdot \boldsymbol{j}(\boldsymbol{r}, t) \underset{(a)}{=} -\int_A \mathrm{d}^3 r \, \boldsymbol{\nabla} \cdot \boldsymbol{j}(\boldsymbol{r}, t)$$

where in (a) we applied the Gauss divergence theorem.
Rearranging:

$$\int_A d^3 r [\dot{\rho}_n(\boldsymbol{r}, t) + \boldsymbol{\nabla} \cdot \boldsymbol{j}(\boldsymbol{r}, t)] = 0$$

Integral form

This is the **continuity equation** in integral form. Note that it holds for any choice of volume $A \subseteq \mathbb{R}^3$. So, knowing that $\boldsymbol{j}$ and $\dot{\rho}$ are continuous functions, by the fundamental theorem of calculus we know that the same relation must hold everywhere *for the integrand*, meaning that:

$$\dot{\rho}_n(\boldsymbol{r}, t) + \boldsymbol{\nabla} \cdot \boldsymbol{j}(\boldsymbol{r}, t) = 0 \qquad \forall \boldsymbol{r}, \forall t \tag{1.4}$$

Differential form

That is the **continuity equation** in differential form .

Now we need a formula to compute the current $\boldsymbol{j}(\boldsymbol{r}, t)$ produced by the diffusion motion. If there are no other fields (EM, gravity, etc.), but we still observe a non-zero $\boldsymbol{j}$, where could it possibly be from?

The only other relevant physical vector in this situation, i.e. not depending on an arbitrary choice of reference frame, is the "spatial" rate of change of density, i.e. its gradient $\boldsymbol{\nabla}\rho_n$. In fact, it is observed that particles tend to move *opposite* to that gradient - from regions where there are more particles to regions where there are less. This can be summarized by **Fick's Law**:

Fick's Law

$$\boldsymbol{j}(\boldsymbol{r}, t) = -D\boldsymbol{\nabla}\rho_n(\boldsymbol{r}, t) \tag{1.5}$$

Of course, there could be some other terms in this expression:

$$\boldsymbol{j}(\boldsymbol{r}, t) = -D\boldsymbol{\nabla}\rho_n(\boldsymbol{r}, t) + C\boldsymbol{\nabla}(\boldsymbol{\nabla}\rho_n) + \dots$$

However, by dimensional analysis, $\partial_x^k \rho_n \sim \rho_n / L^k$, where L is the macroscopic dimension of the container. So, the higher order terms can be considered negligible.
Substituting (1.5) in (1.4) we arrive finally at the **diffusion equation**:

Diffusion Equation

$$\dot{\rho}_n(\boldsymbol{r}, t) = \boldsymbol{\nabla}(D \cdot \boldsymbol{\nabla}\rho_n(\boldsymbol{r}, t)) \tag{1.6}$$

Knowing the initial density $\rho_n(\boldsymbol{r}, 0)$ and some macroscopical details for the fluids (all contained in the *diffusion parameter* D), we can now compute the density after a small interval Δt. For example, we can start by expanding $\rho_n(\boldsymbol{r}, \Delta t)$ around $\Delta t = 0$:

$$\rho_n(\boldsymbol{r}, \Delta t) = \rho_n(\boldsymbol{r}, 0) + \Delta t \dot{\rho}_n(\boldsymbol{r}, 0) + O(\Delta t^2)$$

Ignoring the higher order terms, we can use (1.6) and compute $\rho_n(\boldsymbol{r}, \Delta t)$. This is the gist of the Euler algorithm for numerically approximating differential equations.

This may be more or less doable depending on the form of D, that can depend on both $\boldsymbol{r}$ and t. The $\boldsymbol{r}$-dependence is characteristic of problems that are not translational invariant (e.g. a crystal). In fact, if D **does not** depend on $\boldsymbol{r}$, the diffusion equation becomes:

Translational invariance

$$\dot{\rho}_n(\boldsymbol{r}, t) = D\nabla^2\rho_n(\boldsymbol{r}, t) \tag{1.7}$$

Because the only spatial derivatives are of second order, then if $\rho(\boldsymbol{r}, t)$ is a solution, also $\rho(\boldsymbol{r} + \boldsymbol{R}, t)$ is a solution, for any choice of $\boldsymbol{R}$.

Note that (1.7) is quite similar to the Schrödinger equation for a free particle:

Quantum correspondence

$$-i\,\partial_t\psi = +\frac{\hbar}{2m}\nabla^2\,\psi$$

The yellow term is analogous to D, and the only difference is given by the green term. This can be resolved by a substitution $\tau = it$ (passing to "imaginary time").

Example 1 (Particle diffusing in $\boldsymbol{d=1}$)**:**

Consider the simplest case of a single particle moving in one dimension, with D constant. Let $\rho(x,0)=\delta(x)$, that is consider the particle as being perfectly localized in $x=0$ at the start.
The diffusion equation in $d=1$ is:

$$\dot{\rho}(x,t)=D\rho''(x,t) \tag{1.8}$$

The macroscopic quantities of interest are the expected position and velocity, defined as:

$$\langle x\rangle_t=\int_{-\infty}^{+\infty}\rho(x,t)x\,\mathrm{d}x \qquad \frac{d\langle x\rangle_t}{dt}=\int_{-\infty}^{+\infty}\dot{\rho}(x,t)x\,\mathrm{d}x$$

From the normalization condition:

$$\int_{-\infty}^{+\infty}\rho(x,t)dx=1$$

we note that $\rho(\pm\infty,t)=0$, and also $\rho'(x,t)\to 0$ for $|x|\to\infty$ (otherwise, the density would diverge).
These limits allow us to compute the velocity by repeated integration by parts:

$$\begin{aligned}\frac{\mathrm{d}\langle x\rangle_t}{\mathrm{d}t}&=\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{+\infty}\rho(x,t)x\,\mathrm{d}x=\int_{-\infty}^{+\infty}\dot{\rho}(x,t)x\,\mathrm{d}x\underset{(1.8)}{=}D\int_{-\infty}^{+\infty}\rho''(x,t)x\,\mathrm{d}x=\\&=\underbrace{x\rho'(x,t)\Big|_{x=-\infty}^{x=+\infty}}_{=0}\underbrace{-\left(\frac{\mathrm{d}x}{\mathrm{d}x}\right)\rho(x,t)\Big|_{x=-\infty}^{x=+\infty}}_{=0}+D\int_{-\infty}^{+\infty}\rho(x,t)\underbrace{\left(\frac{\mathrm{d}^2x}{\mathrm{d}x^2}\right)}_{=0}\mathrm{d}x=0\end{aligned}$$

Note that a similar calculation can be done in the more general case of computing the expected value of the time derivative of any function $f(x)$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle f(x)\rangle=D\int_{-\infty}^{\infty}\rho(x,t)\left(\frac{\mathrm{d}^2f(x)}{\mathrm{d}x^2}\right)\mathrm{d}x \tag{1.9}$$

We found that the mean velocity is 0, meaning that the mean position must be constant:

$$\langle x\rangle_t=\langle x\rangle_{t=0}=\int_{-\infty}^{+\infty}\mathrm{d}x\,x\rho(x,0)=0$$

However, if we consider $f(x)=x^2$, thanks to (1.9) we arrive at:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle x^2\rangle_t=\int_{-\infty}^{+\infty}\dot{\rho}(x,t)x^2\,\mathrm{d}x=D\int_{-\infty}^{+\infty}2\rho(x,t)\,\mathrm{d}x$$

As $\rho(x,0)=\delta(x)$, we have:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle x^2\rangle_0=D\int_{-\infty}^{+\infty}2\delta(x)\,\mathrm{d}x=2D$$

And integrating with respect to t:

$$\langle x^2 \rangle_t = 2Dt + \langle x^2 \rangle_0 = 2Dt$$

This allows us to compute the **variance** of x:

$$\text{Var}(x)_t = \langle x^2 \rangle_t - \langle x \rangle_t^2 = 2Dt$$

So the *width* of the distribution of x, which is $\sqrt{\text{Var}(x)}$, expands $\propto \sqrt{Dt}$. The dependence on $\sqrt{t}$ is a defining characteristic of the *diffusion motion*.

(Lesson 2 of 07/10/19) Compiled: October 13, 2020

1.3 Microscopical approach

Let's tackle the diffusion problem with a different approach, studying the motion of *single particles* rather than changes of densities in an ensemble. The correspondence with the results obtained in the previous section will be key to understand the link between the microscopic and the macroscopic - that is the main goal of statistical mechanics.

Consider a particle moving in $d = 1$. To simplify the problem, we allow only discrete steps, both in time and position:

Discretization

$$x_i \equiv i \cdot l \qquad t_n \equiv n \cdot \varepsilon \tag{1.10}$$

In other words, the particle may occupy only points in this defined lattice - and nothing in between. We also look at the system evolution after *discrete* time steps, each of length ϵ.

We already discussed how the diffusion process is intrinsically stochastic, meaning that the motion of grains is given by collisions at the microscopical level, which are essentially random.
So, suppose that the particle lies in a certain known position at $t = 0$. After an instant, the particle may have moved to the right (with probability P_+) or to the left (P_-), or have remained in the same position as before (P_0). As these cases cover all the possibilities, it holds:

$$P_+ + P_- + P_0 = 1$$

Denote with $w_i(t_n)$ the probability that the particle lies at position x_i at time t_n. The probability for the next timestep is then given by the **Master Equation**:

Master Equation

$$w_i(t_{n+1}) = P_0 w_i(t_n) + P_+ w_{i-1}(t_n) + P_- w_{i+1}(t_n) \tag{1.11}$$

In fact, if the particle were at position i at time t_n, then it will remain in the same position with probability P_0. Otherwise, it could have been one position left and moved to the right (P_+), or one position right and moved to the left (P_-).
Here we supposed that ε is sufficiently small, so that the particle will only take

one step at a time.

Stochastic systems for which the state at a certain time depends only on the state one instant before are called **Markov's Processes**.

Markov's Processes

Note that, as the particle cannot "escape from the system", its probability to be in *any* position is conserved at any given time:

Probability conservation

$$\sum_{i=-\infty}^{\infty} w_i(t_{n+1}) = \sum_{i=-\infty}^{\infty} w_i(t_n) = \cdots = \sum_{i=-\infty}^{\infty} w_i(0)$$

Suppose that the particle "always moves", that is $P_0 = 0$, and also that it does so without any preferred direction ($P_+ = P_- = 0.5$). Then, the final position i at time t_n is given by the number of steps to the right n_+ minus the number of steps to the left $n_- \in \mathbb{N}$:

$$\mathbb{Z} \ni i = n_+ - n_-$$

This process can be simulated by flipping a coin at each timestep: if it lands on heads the particle will move to the right, otherwise to the left. So, denoting the total number of steps as $n = n_+ + n_-$, then the probability for the particle to be in position x_i is given by a **binomial distribution**:

$$w_i(t_n) = \binom{n}{n_+} \frac{1}{2^{n_+}} \frac{1}{2^{n_-}} = \frac{1}{2^n} \binom{n}{n_+} = \binom{n}{n_-} \frac{1}{2^n} \tag{1.12}$$

This can be generalized to the case where $P_+ \neq P_-$:

$$w_i(t_n) = \binom{n}{n_-} P_+^{n_+} P_-^{n_-} \tag{1.13}$$

Note that (1.12) satisfies the Master Equation (1.11), that is:

$$w_i(t_{n+1}) = \frac{1}{2}(w_{i+1}(t_n) + w_{i-1}(t_n))$$

We start by noting that if $i = n_+ - n_-$ and $n = n_+ + n_-$, then:

$$n_+ = \frac{n+i}{2} \qquad n_- = \frac{n-i}{2} \tag{1.14}$$

And so:

$$\frac{1}{2}(w_{i+1}(t_n) + w_{i-1}(t_n)) = \frac{1}{2^{n+1}} \left[\binom{n}{\frac{n+i+1}{2}} + \binom{n}{\frac{n+i-1}{2}} \right] \tag{1.15}$$

Recall now the recurrence relation for the binomial coefficient:

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$$

which leads to the desired result:

$$(1.15) = \frac{1}{2^{n+1}} \binom{n+1}{\frac{n+i+1}{2}} = w_i(t_{n+1})$$

1.3.1 Probability Generating Functions

Let's introduce a useful mathematical tool to deal with the binomial coefficient. Let X be a discrete random variables taking values in the non-negative integers ($\mathbb{N}$). The **probability generating function**[1] of X is defined as:

$$G(z) \equiv \mathbb{E}[z^X] = \sum_{x=0}^{+\infty} p(x) z^x \tag{1.16}$$

where p is the *probability mass function* of X, i.e. $p(x)$ is the probability that $X = x$ ($p(x) = \mathbb{P}(X = x)$).

$G(z)$ is useful because we can retrieve $p(k)$ for any $k \in \mathbb{N}$ by simply differentiating k times $G(z)$ with respect to z and setting $z = 0$. In fact, by expanding the sum in the definition (1.16) and then differentiating:

Probability generation

$$G(z) = p(0) + p(1)z + p(2)z^2 + \dots \Rightarrow \frac{\mathrm{d}^k}{\mathrm{d}z^k} G(z)\Big|_{z=0} = p(k)$$

Note that $G(1) = 1$ because of the normalization:

$$G(1) = \sum_{x=0}^{+\infty} p(x) 1^x = \sum_{x=0}^{+\infty} p(x) = 1$$

This suggests a way to use $G(z)$ to compute also the **moments** of X. In fact, if we evaluate the first derivative for $z = 1$ we get:

$$G'(1) = p(1) + 2p(2)z + 3p(3)z^2 + \dots\Big|_{z=1} = \sum_{x=1}^{+\infty} p(x) x z^{x-1}\Big|_{z=1} = \tag{1.17}$$

$$= \sum_{x=1}^{+\infty} x p(x) = (0 \cdot p(0)) + 1 \cdot p(1) + 2 \cdot p(2) + \dots = \mathbb{E}[X] \tag{1.18}$$

However, the second derivative of G evaluated at $z = 1$ does not give the second moment:

$$G''(1) = 2 \cdot 1p(2) + 3 \cdot 2p(3)z + 4 \cdot 3p(4)z^2 + \dots\Big|_{z=1} = \sum_{x=2}^{+\infty} x(x-1) z^{x-2} p(x)\Big|_{z=1} =$$
$$= \mathbb{E}(X(X-1))$$

More generally:

Factorial moment generation

$$G^{(k)}(1) = \mathbb{E}(X(X-1)\dots(X-k+1)) = \mathbb{E}\left(\frac{X!}{(X-k)!}\right)$$

[1] ∧Not to be confused with the *moment generating function* of a **real**-valued random variable X (i.e., *not* discrete), which is defined as $\mathbb{E}(e^{tX})$, with $t \in \mathbb{R}$

which is called the k-th factorial moment of X.
But how can we get the "usual" moments from G? One possibility is to "compensate" the difference between a factorial moment and a usual one by adding other terms. For example, note that:

$$G''(1) = \mathbb{E}(X(X-1)) = \mathbb{E}(X^2) - \mathbb{E}(X)$$

and so:

$$G''(1) + G'(1) = (\mathbb{E}(X^2) - \mathbb{E}(X)) + \mathbb{E}(X) = \mathbb{E}(X^2)$$

A more clever way is to consider the operator $\theta(z)$ defined as:

$$\theta(z) \equiv z\frac{\partial}{\partial z}$$

on G. In fact:

$$\theta(z)G(z) = z\frac{\partial}{\partial z}\sum_{x=0}^{+\infty} p(x)z^x = z\sum_{x=1}^{+\infty} xp(x)z^{x-1} = \sum_{x=1}^{+\infty} xp(x)z^x$$

And setting $z = 1$ leads back to the $\mathbb{E}[X]$. If we apply $\theta(z)$ again, however, something interesting happens:

$$\theta(z)^2 G(z) = \left(z\frac{\partial}{\partial z}\right)\left(z\frac{\partial}{\partial z}\right)G(z) = z\frac{\partial}{\partial z}\sum_{x=1}^{+\infty} xp(x)z^x = \sum_{x=1}^{+\infty} x^2 p(x)z^x \quad (1.19)$$

Now setting $z = 1$ leads to $\mathbb{E}[X^2]$. In general: *Moment generation*

$$\theta(z)^k G(z)\Big|_{z=1} = \mathbb{E}[X^k]$$

Note how the exponent of z never changes, as it is lowered by 1 by the ∂_z, and then rised back by the z factor. So, every new application of the $\theta(z)$ operator merely brings down another x factor, rising the x exponent inside the sum - which is exactly what we want to compute moments.

1.3.2 Moments of the diffusion distribution

Let's focus on our specific (discrete) case, with the particle moving on a discretized line. At any given time t_n we can compute the mass probability function $W_{t_n}: \mathbb{Z} \to \mathbb{R}_+$, with $W_{t_n}(x_i) \equiv w_i(t_n)$. In other words, this is the function that maps every position x_i to the probability of containing a particle at time t_n (we focus on the spatial distribution at a *fixed time* rather than the temporal distribution at a *fixed position*).
We are interested in knowing the *shape* of W_{t_n}, that is its moments:

$$\langle x^q \rangle_{t_n} = \sum_{i=-\infty}^{+\infty} W_{t_n}(x_i)x_i^q \underset{(1.10)}{=} \sum_{i=-\infty}^{+\infty} w_i(t_n)(l \cdot i)^q \qquad q \in \mathbb{N}$$

The first moment ($q = 1$) gives the average position: *First moment of x as function of $\langle n_+ \rangle$*

$$
\begin{aligned}
\langle x\rangle_{t_n} &= l\cdot\sum_{i=-\infty}^{+\infty} w_i(t_n)i = \\
&= l\sum_{i=-\infty}^{+\infty} w_i(t_n)(2n_+ - n) = \\
&\underset{(a)}{=} l\left(2\langle n_+\rangle_{t_n} - n\right) \qquad (1.20)
\end{aligned}
$$

where in (a) we used the normalization condition ($\forall n \in \mathbb{N}, \sum_i w_i(t_n) = 1$). Thus, we found that the average position $\langle x\rangle_t$ of the particle at time t_n is related to the value of n_+.

So, let n_+ be the random variable of interest. Recall that n_+ is sampled from a binomial distribution (1.13), and that $n_+ = (n+i)/2$ (1.14) and so $i = 2n_+ - n$.
Then, the *probability generating function* of n_+ is given by:

$$
\begin{aligned}
\widetilde{W}(z,n) &\underset{(1.16)}{\equiv} \sum_{n_+=0}^{n} z^{n_+} w_i(t_n)\big|_{i=2n_+-n} \underset{(1.13)}{=} \sum_{n_+=0}^{n} z^{n_+}\binom{n}{n_+}P_+^{n_+}P_-^{n-n_+} = \\
&\underset{(a)}{=} (P_+z+P_-)^n
\end{aligned}
$$

where in (a) we used the binomial theorem.

We can now use the property (1.18) of the probability generating function to compute $\langle n_+\rangle$:

First moment of n_+

$$
\langle n_+\rangle = \frac{\partial}{\partial z}\widetilde{W}(z,n)\Big|_{z=1} = n(P_+z+P_-)^{n-1}P_+\Big|_{z=1} = n\underbrace{(P_++P_-)}_{=1}P_+ = nP_+ \qquad (1.21)
$$

For computing the second moment, we apply the $\theta(z)$ operator, as seen in (1.19):

Second moment of n_+

$$
\begin{aligned}
\langle n_+^2\rangle &= \left(z\frac{\partial}{\partial z}\right)^2\widetilde{W}(z,n)\Big|_{z=1} = z\frac{\partial}{\partial z}zn(P_+z+P_-)^{n-1}P_+\Big|_{z=1} = \\
&= z(n(P_+z+P_-)^{n-1}P_+ + znP_+^2(n-1)(P_+z+P_-)^{n-2})\Big|_{z=1} = \\
&= nP_+(1+(n-1)P_+) \qquad (1.22)
\end{aligned}
$$

We can now compute $\mathrm{Var}(n_+)$ recalling that:

$$
\mathrm{Var}[X] = \mathbb{E}[X^2] - (E[X])^2
$$

Thus:

$$
\mathrm{Var}[n_+] = \langle n_+^2\rangle - \langle n_+\rangle^2 = nP_+(1-P_+) \qquad (1.23)
$$

We now go back to $\langle x\rangle_{t_n}$, recalling the relation (1.20):

Moments of x

$$
\langle x\rangle_{t_n} = l(2\langle n_+\rangle_{t_n} - n) \underset{(1.21)}{=} nl(2P_+-1) \underset{(a)}{=} nl(P_+-P_-)
$$

where in (a) we used $P_+ + P_- = 1$.
For the variance, recall that:

$$\mathrm{Var}[aX + b] = a^2 \,\mathrm{Var}[X]$$

and so, starting again from (1.20):

$$\mathrm{Var}[x]_{t_n} = 4l^2 \,\mathrm{Var}[n_+] \underset{(1.23)}{=} 4nl^2 P_+(1 - P_+) = 4nl^2 P_+ P_-$$

Note that the variance is always proportional to time (n), even if $P_+ \neq P_-$. However, if we go back and compute the $\langle x^2 \rangle_{t_n}$, we will note that it is not linear in time:

$$\begin{aligned}\langle x^2 \rangle_{t_n} &= \mathrm{Var}[x]_{t_n} + \langle x \rangle_{t_n}^2 = 4nl^2 P_+(1 - P_+) + (nl(P_+ - P_-))^2 = \\ &= nl^2(4P_+P_- + n(P_+ - P_-)^2)\end{aligned}$$

Let's evaluate the previous quantities for the simple **symmetrical** case, where $P_+ = P_- = 1/2$:

$$\begin{aligned}&\langle n_+ \rangle = \frac{n}{2} \qquad &&\langle n_+^2 \rangle = \frac{n}{4}(n+1) \qquad &&\mathrm{Var}[n_+] = \frac{3}{4}n \\ &\langle x \rangle_{t_n} = 0 &&\langle x^2 \rangle_{t_n} = nl^2 &&\mathrm{Var}(x)_{t_n} = nl^2\end{aligned}$$

As expected, the average number of steps to the right is half the total steps (as $P_+ = 1/2$), and the average position is 0.

Alternative derivation for the x moments. These last results can be also obtained in a simpler way.
The idea is to represent the final state of the random walk at time t_n as the sum of n steps:

$$x(t_n) = u_1 + u_2 + \cdots + u_n$$

Each step can be on the right or on the left according to some probability distribution. In other words, u_i is a random variable. If we suppose steps of *unit length* symmetrically distributed (i.e. $P_+ = P_- = 1/2$) we get:

$$u_n = \begin{cases} +1 & p = 1/2 \\ -1 & p = 1/2 \end{cases}$$

We can now compute the average position (first moment):

$$\begin{aligned}\langle x(t_n) \rangle &= n \langle u \rangle = 0 \\ \langle u \rangle &= \frac{1}{2}(+1) + \frac{1}{2}(-1) = 0\end{aligned}$$

And the second moment:

$$\langle x^2(t_n) \rangle = \langle (u_1 + \cdots + u_n)^2 \rangle = n \cdot \underbrace{\langle u^2 \rangle}_{1} + \sum_{i \neq j} \underbrace{\langle u_i \cdot u_j \rangle}_{\langle u_i \rangle \langle u_j \rangle = 0} = n$$

$$\langle u^2\rangle = \frac{1}{2}(+1)^2 + \frac{1}{2}(-1)^2 = 1$$

Note that $\langle u_i \cdot u_j\rangle = \langle u_i\rangle\langle u_j\rangle$ because u_i and u_j are statistically independent. So we showed that $\langle x^2\rangle_{t_n}$ is linear in the number of time steps (n). Here, the l^2 factor from (1.24) is missing because the spatial step size is set to 1.

1.3.3 Continuum Limit

Recall that $t_n = n\varepsilon$. Inserting this relation in the results we got for the x moments in the previous section (for the symmetrical case $P_+ = P_-$) we get:

$$\langle x\rangle_{t_n} = 0; \qquad \langle x^2\rangle_{t_n} = l^2 n = \frac{l^2}{\varepsilon} t_n \tag{1.24}$$

The analysis of the diffusion equation in $d = 1$ showed that, for a particle starting at $x(t = 0) = 0$:

$$\langle x^2\rangle_t = 2Dt \tag{1.25}$$

The correct *continuum limit* should reproduce the result of (1.25) from (1.24). Notice that if we simply let $\varepsilon \to 0$ and $l \to 0$, $\langle x^2\rangle$ becomes undefined. Therefore, we need to fix the ratio $l^2\varepsilon^{-1}$ during the limit. If we define this ratio as:

Correct continuum limit

$$\frac{l^2}{\varepsilon} \equiv 2D$$

then the limit of (1.24) leads to (1.25) as desired:

$$\langle x^2\rangle_{t_n} = \frac{l^2}{\varepsilon} t_n \xrightarrow[l,\varepsilon\to 0]{l^2/\varepsilon = 2D} 2Dt = \langle x^2\rangle_t$$

Note that $[D] = \mathrm{m}^2\,\mathrm{s}^{-1}$, and so the previous expression is dimensionally correct.

We now know the basic *shape* that the distribution must have in the continuum limit - but we still don't know its explicit form. So, let's start by considering the spatial distribution at a fixed time t_n:

$$W_{t_n}(x_i) \equiv w_i(t_n) = \frac{n!}{\left(\frac{n+i}{2}\right)!\left(\frac{n-i}{2}\right)!}\frac{1}{2^n} \tag{1.26}$$

For $n = 0$ (starting time), all the particles are at x_0. Then, after each timestep the probability distribution "expands", meaning that more and more positions have a non-zero probability of being explored.

In particular, recall that:

i and n have the same parity

$$i = 2n_+ - n$$

Note how i and n must have the same parity, as $2n_+$ is always even. So the particle will always be at an even position (x_i with i even) after an even number n of time steps, and at an odd x_i after an odd time t_n.

To proceed, we note that in the continuum limit, as the timestep ϵ vanishes, every *finite time* t will be reached after a *really big* number of steps. So, we want to examine the **asymptotic behaviour** of (1.26) as $n \to \infty$. We start by computing its logarithm:

$$\ln w_i(t_n) = -n \ln 2 + \ln n! - \ln \left(\frac{n+i}{2}\right)! - \ln \left(\frac{n-i}{2}\right)!$$

In this way, we can use the Stirling approximation:

$$\begin{aligned} \ln k! &= \ln k + \ln(k-1) + \dots + \ln 2 + \ln 1 = \\ &\approx k \ln k - k + \frac{1}{2} \ln(2\pi k) \end{aligned}$$

Thus arriving at a complicated expression:

$$\begin{aligned} \ln w_i(t_n) \underset{n \gg 1}{\approx} & -n \ln 2 + n \ln n - \cancel{n} + \frac{1}{2} \ln(2\pi n) \\ & - \frac{n+i}{2} \ln \left(\frac{n+i}{2}\right) + \cancel{\frac{n+i}{2}} - \frac{1}{2} \ln \left(2\pi \frac{n+i}{2}\right) \\ & - \frac{n-i}{2} \ln \left(\frac{n-i}{2}\right) + \cancel{\frac{n-i}{2}} - \frac{1}{2} \ln \left(2\pi \frac{n-i}{2}\right) \end{aligned}$$

Let's gradually simplify it. We start by collecting all the n:

$$\begin{aligned} & n \left(-\ln 2 + \ln n - \frac{1}{2} \ln \left(\frac{n+i}{2}\right) - \frac{1}{2} \ln \left(\frac{n-i}{2}\right)\right) && = \\ &= n \ln \left(\frac{n}{2\sqrt{\frac{n+i}{2}}\sqrt{\frac{n-i}{2}}}\right) = n \ln \left(\frac{n}{\sqrt{n^2 - i^2}}\right) = n \ln \left(\frac{1}{\sqrt{1 - i^2/n^2}}\right) && = \\ &\underset{(a)}{=} n \ln \left(1 + \frac{1}{2} \frac{i^2}{n^2} + O\left(\frac{i^4}{n^4}\right)\right) \underset{(b)}{=} \frac{n}{2} \frac{i^2}{n^2} + O\left(\frac{i^4}{n^4}\right) \approx \frac{i^2}{2n} && (1.27) \end{aligned}$$

where in (a) and in (b) we used respectively the following Taylor expansions (as $n \to \infty$ and so $1/n \to 0$):

$$(1 \pm x)^n = 1 \pm nx + O(x^2) \qquad \ln(1+x) = x + O(x^2) \tag{1.28}$$

Then we collect the $i/2$:

$$\begin{aligned} & -\frac{i}{2} \left[\ln \left(\frac{n+i}{2}\right) - \ln \left(\frac{n-i}{2}\right)\right] = -\frac{i}{2} \ln \left(\frac{n+i}{n-i}\right) && = \\ &= -\frac{i}{2} \ln \left(\frac{1 + i/n}{1 - i/n}\right) \underset{(a)}{=} -\frac{i}{2} \ln \left(\left(1 + \frac{i}{n}\right)\left(1 + \frac{i}{n} + O\left(\frac{i^2}{n^2}\right)\right)\right) && = \\ &= -\frac{i}{2} \ln \left(1 + \frac{2i}{n} + O\left(\frac{i^2}{n^2}\right)\right) \underset{(b)}{=} -\frac{i}{2} \left(\frac{2i}{n} + O\left(\frac{i^2}{n^2}\right)\right) \approx -\frac{i^2}{n} && (1.29) \end{aligned}$$

And finally we consider the remaining terms:

$$\frac{1}{2} \left[\ln(2\pi n) - \frac{1}{2} \ln \left(2\pi \frac{n+i}{2}\right) - \ln \left(2\pi \frac{n-i}{2}\right)\right] \qquad =$$

$$= \frac{1}{2}\ln\left(\frac{2\pi n}{2\pi\frac{n+i}{2}2\pi\frac{n-i}{2}}\right) = \frac{1}{2}\ln\left(\frac{2n}{\pi(n^2-i^2)}\right) = \frac{1}{2}\ln\left(\frac{2}{\pi n}\frac{1}{1-\frac{i^2}{n^2}}\right) =$$

$$= \frac{1}{2}\ln\left(\frac{2}{\pi n} + O\left(\frac{i^2}{n^2}\right)\right) \approx \frac{1}{2}\ln\left(\frac{2}{\pi n}\right) \tag{1.30}$$

Putting it all back together:

$$\ln w_i(t_n) \underset{n\gg 1}{\approx} (1.27) + (1.29) + (1.30) = \frac{1}{2}\ln\left(\frac{2}{\pi n}\right) - \frac{i^2}{2n}$$

And by exponentiating we get:

$$w_i(t_n) \underset{n\gg 1}{\approx} \sqrt{\frac{2}{\pi n}}\exp\left(-\frac{i^2}{2n}\right) \tag{1.31}$$

We now want to obtain a *continuous* pdf from the mass probability $w_i(t_n)$. Note that if we regard $w_i(t_n)$ as a function of position at a fixed time $W_{t_n}(x_i)$, and extend the domain to all $\mathbb{R}$ we get a really "bumpy" function, as it is non-zero only on $x_i = l \cdot i$ with $i \in \mathbb{Z}$. However, if we *integrate* over every *small patch* of $W_{t_n}(x_i)$, we can "smooth" all the "bumpyness", and get a nice pdf - especially in the continuum limit.
Let's formalize that more carefully. Starting from $W_{t_n}(x_i)$, we can compute the probability to find a particle in an interval $I \subseteq \mathbb{R}$ by simply summing the mass probabilities $w_j(t_n)$ for all the $x_j \in I$.
The idea is now to define a *continuous* pdf $W(x,t)$ as follows:

$$W(x,t_n)\Delta x = \mathbb{P}\left(x \in \left[x - \frac{\Delta x}{2}; x + \frac{\Delta x}{2}\right], t_n\right) \qquad l \ll \Delta x \ll 1$$

That is, $W(x,t_n)\Delta x$ is the probability that a particle lies "near" a certain position $x \in \mathbb{R}$ at an instant t_n (i.e. within an interval I centered on x with width Δx sufficiently small, but large with respect to the discretization). We will then "cure" the discreteness of time by considering the asymptotic behaviour for $t \to \infty$.
By expanding the previous expression we get:

$$\mathbb{P}\left(x \in \left[x - \frac{\Delta x}{2}, x + \frac{\Delta x}{2}\right]\right) \underset{(a)}{\approx} \mathbb{P}\left(i \in \left[i_0 - \frac{\Delta x}{2l}; i_0 + \frac{\Delta x}{2l}\right]\right) = \sum_{j=i_0-\Delta x/(2l)}^{i_0+\Delta x/(2l)} w_j(t_n) \tag{1.32}$$

where in (a) i_0 is such that x_{i_0} is closest to x_i, that is $i_0 = \lfloor x/l \rceil$ (recall that $x_i = il \Rightarrow i = x_i/l$).
Note that this specific choice of I contains $\Delta x/l$ points (supposing $\Delta x \gg l$). However, depending on the parity of n (fixed by the choice of the instant t_n), only half of the positions x_j can be explored, as n and j must have the same parity. This means that half of the $w_j(t_n)$ with x_j inside the interval, $w_j(t_n) = 0$. For the other half, we suppose that $w_j(t_n)$ does not vary much

inside the small interval, and so we approximate their value with the center point $j = i_0$, i.e. $w_j(t_n) \approx w_{i_0}(t_n)$. So, averaging over these two halves:

$$(1.32) \approx \frac{\Delta x}{l}\left(\frac{1}{2}\cdot 0 + \frac{1}{2}w_{i_0}(t_n)\right) = \frac{\Delta x}{2l}w_{i_0}(t_n) \tag{1.33}$$

We have now an expression for $W(x, t_n)$, which is continuous with respect to x:

$$W(x, t_n)\cancel{\Delta x} = \frac{\cancel{\Delta x}}{2l}w_i(t_n) \tag{1.34}$$

If we now take the limit $n \to \infty$ we can substitute (1.31) in (1.34), leading to: *Continuous distribution*

$$W(x, t) = \frac{1}{2l}\sqrt{\frac{2}{\pi n}}\exp\left(-\frac{i^2}{2n}\right)$$

Substituting $x = il$ and $t = n\varepsilon$ we get:

$$W(x, t) = \sqrt{\frac{2\varepsilon}{4l^2\pi t}}\exp\left(-\frac{x^2}{2\frac{l^2}{\varepsilon}t}\right) \tag{1.35}$$

As $l^2\varepsilon^{-1} = 2D$:

$$W(x, t) = \frac{1}{\sqrt{4\pi D t}}\exp\left(-\frac{x^2}{4Dt}\right) \tag{1.36}$$

We can now compute the first two moments of x:

$$\begin{aligned}
\langle x\rangle_t &= \int_{\mathbb{R}} W(x, t)x \,\mathrm{d}x = 0 \\
\langle x^2\rangle_t &= \int_{\mathbb{R}} W(x, t)x^2 \,\mathrm{d}x = 2Dt
\end{aligned}$$

This last integral can be done in many ways. For example, recall the gaussian integral:

$$I = \sqrt{\frac{\pi}{\mu}} = \int_{-\infty}^{\infty} e^{-\mu y^2}\,\mathrm{d}y$$

Differentiating (according to Leibniz integral rule) with respect to μ:

$$\frac{\partial I}{\partial \mu} = -\int_{\mathbb{R}} e^{-\mu y^2} y^2 \,\mathrm{d}y = -\frac{1}{2}\sqrt{\frac{\pi}{\mu^3}} \tag{1.37}$$

and so:

$$\begin{aligned}
\langle x^2\rangle_t &= \frac{1}{\sqrt{4\pi D t}}\int_{-\infty}^{+\infty} x^2 \exp\left(-\frac{x^2}{4Dt}\right)\mathrm{d}x = \\
&= \frac{1}{\sqrt{4\pi D t}}\int_{-\infty}^{+\infty} e^{-\mu x^2}x^2\,\mathrm{d}x\,\Big|_{\mu=(4Dt)^{-1}} \underset{(1.37)}{=} \frac{1}{\sqrt{4\pi D t}}\frac{1}{2}\sqrt{\pi(4Dt)^3} = 2Dt
\end{aligned}$$

(Lesson 4 of 14/10/19) Compiled: October 13, 2020

1.4 The Link between Macroscopic and Microscopic

We will now show that the continuum limit of the Master Equation (1.11) produces the diffusion equation (1.6), in the case of constant D, thus establishing a link between the interpretation in terms of *densities* and that in term of *paths of random motion*.
Then, we will show that (1.36) is the solution of that equation for a starting distribution of δx (particle initially at 0), and derive the general solution for *any* initial condition.

So, we start by recalling that, for a fine discretization, $w_i(t_n)$ is approximately equal to the probability of being around a generic (x,t) (i.e. $W(x,t)\Delta x$), up to a normalization constant:

$$W(x_0, t_n)\Delta x = \mathbb{P}(x \in [x_0 - \Delta x/2, x_0 + \Delta x/2]) \approx \frac{\Delta x}{2l} w_{i_0}(t_n) \qquad i_0 = \lfloor \Delta x/l \rceil$$

And so, with a slight abuse of notation:

$$w_i(t_n) \approx 2l W(x,t) \qquad i = \lfloor x/l \rceil, n = \lfloor t/\epsilon \rceil$$

Substituting in the Master Equation (1.11) leads to:

$$\cancel{2l} W(x, t+\epsilon) = \cancel{2l} \frac{1}{2} (W(x-l, t) + W(x+l, t)) \tag{1.38}$$

which means that an analogous Master Equation holds even for $W(x,t)$, which is a continuous pdf, and thus can be differentiated.
The idea is now to use Taylor expansions to express everything in terms of the derivatives of W evaluated at the *same* point (x,t). So, we compute $W(x,t+\epsilon)$ in terms of $W(x,t)$ (and derivatives) by expanding around $\epsilon = 0$:

$$W(x, t+\epsilon) = W(x,t) + \epsilon \frac{\partial}{\partial \tau} W(\chi, \tau)\Big|_{(x,t)} + \frac{\epsilon^2}{2} \frac{\partial^2}{\partial \tau^2} W(\chi, \tau)\Big|_{(x,t)} + O(\epsilon^3) \tag{1.39}$$

The same is done for $W(x \pm l, t)$ by expanding around $l = 0$:

$$W(x \pm l, t) = W(x,t) \pm l \frac{\partial}{\partial \chi} W(\chi, \tau)\Big|_{(x,t)} + \frac{l^2}{2} \frac{\partial^2}{\partial \chi^2} W(\chi, \tau)\Big|_{(x,t)} + O(l^3) \tag{1.40}$$

We then introduce the following notation for the space and time derivatives:

$$\dot{W}(x,t) = \frac{\partial}{\partial \tau} W(\chi, \tau)\Big|_{(x,t)} \qquad W'(x,t) = \frac{\partial}{\partial \chi} W(\chi, \tau)\Big|_{(x,t)}$$

so that a space derivative is denoted with a' (a'' for the second derivative), and a time derivative with $\dot{a}$ ($\ddot{a}$ for the second derivative).
We can now substitute everything back in (1.38). We start with the right side:

$$W(x+l, t) + W(x-l, t) = 2W(x,t) + l^2 W''(x,t) + O(l^4)$$

where the $O(l^4)$ is given by the cancellation of the odd powers (including l^3). Equating to the left side of (1.38) leads to:

$$\cancel{W(x,t)} + \epsilon\dot{W}(x,t) + \frac{\epsilon^2}{2}\ddot{W}(x,t) = \cancel{W(x,t)} + \frac{l^2}{2}W''(x,t) + O(l^4)$$

Dividing by ϵ:

$$\begin{aligned}\dot{W}(x,t) + \frac{\epsilon}{2}\ddot{W}(x,t) &= \underbrace{\frac{l^2}{2\epsilon}}_{D} W''(x,t) + O\left(\frac{l^4}{\epsilon}\right) \\ &= DW''(x,t) + O(4\epsilon D^2)\end{aligned}$$

If we now take the continuum limit, then $\epsilon, l \to 0$ with the ratio $D = l^2/(2\epsilon)$ fixed, both $\ddot{W}(x,t)$ and the error term vanish, leading to the **diffusion equation**:

$$\dot{W}(x,t) = DW''(x,t) \tag{1.41}$$

Which is indeed the same[2] of (1.6) with D constant.

1.5 Solution of the Diffusion Equation

We want now to solve (1.41), and show that the solution will be the same we previously derived in (1.36).
So, we start from:

$$\partial_t W(x,t) = D\partial_x^2 W(x,t)$$

This is a second order partial differential equation. To be able to solve it, we must first define its **boundary conditions**. In this case, we suppose that the particle is unconstrained, and so the spatial domain coincides with $\mathbb{R}$.
As $W(x,t)$ is a pdf, the following conditions must hold:

$$W(x,t) \geq 0 \quad \forall (x,t) \qquad\qquad \int_{\mathbb{R}} W(x,t) = 1$$

From the normalization, it follows that $W(x,t)$ - and its spatial derivative $W'(x,t)$ - must vanish as $|x| \to \infty$, so that the integral does not diverge:

$$\lim_{|x|\to\infty} W(x,t) = 0 \qquad\qquad \lim_{|x|\to\infty} W'(x,t) = 0 \tag{1.42}$$

However, it is not obvious that $W(x,t) \geq 0$ will always hold, assuming we choose an initial condition $W(x,t_0) \geq 0$. This will be obvious *a posteriori* - and in fact can justified by the peculiar properties of this differential equation.

To solve (1.41), as the spatial domain is all $\mathbb{R}$, one standard technique is that of the Fourier integral transform, which allows us to "remove" derivatives by replacing $\partial_x\psi \to ik\psi$, $\partial_{xx} \to -k^2\psi$. Thus, if we can "remove" the second-order space derivative, we will be left with a much more simpler first order differential equation in the time variable.

[2] ∧Almost: here we deal with the probability distribution, while in (1.6) we have a physical density ρ_n. Effectively the two differ only by a normalization factor, as previously noted

Translational invariance. This approach is suggested by the translational invariance of solutions of (1.41). In fact, if $W(x,t)$ is a solution, then also $\tilde{W}(x,t) = W(x-a,t)$ is a solution.
The generator of the translations is the *momentum*, and its eigenfunctions are the *plane waves*, i.e. the Fourier basis. So, by expressing a function in this base, we will harness the equation's symmetry - simplifying the problem.
In other words, the Fourier basis diagonalizes the Laplacian operator which appears in (1.41):

$$\partial_x^2 \varphi_k(x) = \lambda_k \varphi_k(x) \qquad \lambda_k \equiv -k^2 \qquad \varphi_k(x) = A_k e^{\pm ikx}, k \in \mathbb{R}$$

In a general case, the Fourier integral trick can be tried for every variable, starting from the one with the higher order derivative, and then the case which leads to the most simplification can be pursued.

We start by rewriting $W(x,t)$ as a (infinite) linear combination of vectors of the Fourier basis:

$$W(x,t) = \int_{\mathbb{R}} \frac{\mathrm{d}k}{2\pi} e^{ikx} c_k(t) \tag{1.43}$$

where the 2π factor is just a normalization convention.
Let $\varphi_k(x) = e^{ikx}$. Then, as the Fourier basis is orthonormal, the following holds (recalling the Fourier transform of the δ function):

$$\langle \varphi_k, \varphi_{k'} \rangle = \int_{\mathbb{R}} \mathrm{d}x\, \varphi_k^*(x) \varphi_{k'}(x) = \int_{\mathbb{R}} \mathrm{d}x\, e^{i(k'-k)x} = 2\pi\delta(k-k') \tag{1.44}$$
$$\langle \tilde{\varphi}_x, \tilde{\varphi}_{x'} \rangle = \int_{\mathbb{R}} \mathrm{d}k\, \varphi_k^*(x) \varphi_k(x') = 2\pi\delta(x-x')$$

We then apply a Fourier transform to both members of (1.43), by multiplying by $e^{-ik'x}$ and integrating over x:

$$\int_{\mathbb{R}} W(x,t) e^{-ik'x}\, \mathrm{d}x = \int_{\mathbb{R}} \frac{\mathrm{d}k}{2\pi} \int_{\mathbb{R}} \mathrm{d}x\, e^{i(k-k')x} c_k(t)$$

If we now apply the ON relation (1.44) we can solve the integral in the right side:

$$\int_{\mathbb{R}} W(x,t) e^{-ik'x}\, \mathrm{d}x = \int_{\mathbb{R}} \mathrm{d}k\, \delta(k-k') c_k(t) = c_{k'}(t)$$

And substituting $k' \to k$ we arrive at an expression for $c_k(t)$:

$$c_k(t) = \int_{\mathbb{R}} \mathrm{d}x\, e^{-ikx} W(x,t) \tag{1.45}$$

Starting from (1.41) we can write a corresponding differential equation for the coefficients $c_k(t)$ in the Fourier basis, and then solve it.

Braket notation. Let the solution be $|W(t)\rangle$, so that $\langle x|W(t)\rangle = W(x,t)$. Then in (1.43) we just did a change of basis (by using Dirac completeness):

$$|W(t)\rangle = \mathbb{I}\,|W(t)\rangle = \int_k |k\rangle \underbrace{\langle k|W(t)\rangle}_{c_k(t)}$$

where $|k\rangle$ are elements of the Fourier basis ($\langle x|k\rangle = e^{ikx}$) and so:

$$c_k(t) = \langle k|W(t)\rangle = \int_x \langle k|x\rangle \langle x|W(t)\rangle = \int_{\mathbb{R}} \mathrm{d}x\, e^{-ikx} W(x,t)$$

So the initial differential equation (1.41) is expressed in the position basis, while the following equation involving $c_k(t)$ is expressed in the Fourier basis.

So, we start by differentiating (1.45) with respect to t:

$$\begin{aligned}\dot{c}_k(t) &= \int_{\mathbb{R}} \mathrm{d}x\, e^{-ikx} \dot{W}(x,t) \underset{(a)}{=} D \int_{-\infty}^{\infty} e^{-ikx} W''(x,t)\, \mathrm{d}x = \\ &\underset{(b)}{=} \cancel{DW'(x,t)e^{-ikx}}\Big|_{-\infty}^{\infty} - D\int_{-\infty}^{\infty} \partial_x(e^{-ikx}) W'(x,t)\,\mathrm{d}x = \\ &\underset{(c)}{=} -D\,\underbrace{(\partial_x e^{-ikx})}_{-ike^{-ikx}}\cancel{W(x,t)}\Big|_{-\infty}^{\infty} + D\int_{\mathbb{R}} \underbrace{\partial_x^2(e^{-ikx})}_{-k^2e^{-ikx}} W(x,t)\,\mathrm{d}x = \\ &= -Dk^2 \underbrace{\int_{-\infty}^{+\infty} \mathrm{d}x\, e^{-ikx} W(x,t)}_{c_k(t)} = -Dk^2 c_k(t)\end{aligned}$$

where in (a) we substituted (1.41), and in (b) and (c) we performed two integrations by parts. Note that the $W(x,t)$ and $W'(x,t)$ terms vanish because of the boundary conditions (1.42).
Summarizing:

$$\dot{c}_k(t) = \int_{\mathbb{R}} \mathrm{d}x\, e^{-ikx} \dot{W}(x,t) = -Dk^2 c_k(t)$$

This is a first-order ordinary differential equation, which can be solved by separation of variables:

$$\frac{\mathrm{d}}{\mathrm{d}t} c_k(t) = -Dk^2 c_k(t) \Rightarrow \int \frac{\mathrm{d}c_k(t)}{c_k(t)} = \int -Dk^2\,\mathrm{d}t \Rightarrow \ln c_k(t) = -Dk^2 t + C$$

And rearranging:

$$c_k(t) = Ae^{-Dk^2 t} \tag{1.46}$$

To find the integration constant A we impose the initial conditions, i.e. that $c_k(t)$ be equal to a known $c_k(t_0)$ at time t_0:

$$c_k(t_0) \overset{!}{=} Ae^{-Dk^2 t_0} \Rightarrow A = c_k(t_0) e^{Dk^2 t_0} \tag{1.47}$$

And substituting (1.47) back in (1.46) we arrive at the general integral:

$$c_k(t) = c_k(t_0) e^{-Dk^2(t-t_0)} \tag{1.48}$$

We can now go back to $W(x,t)$ by plugging (1.48) into (1.43):

$$W(x,t) = \int_{\mathbb{R}} \frac{\mathrm{d}k}{2\pi} e^{ikx} c_k(t) = \int_{\mathbb{R}} \frac{\mathrm{d}k}{2\pi} e^{ikx - Dk^2(t-t_0)} c_k(t_0) =$$

$$\underset{(1.45)}{=} \int_{\mathbb{R}} \frac{\mathrm{d}k}{2\pi} e^{ikx - Dk^2(t-t_0)} \int_{\mathbb{R}} \mathrm{d}y\, e^{-iky} W(y, t_0) =$$
$$= \int_{\mathbb{R}} \mathrm{d}y\, W(y, t_0) \int_{\mathbb{R}} \frac{\mathrm{d}k}{2\pi} \exp\left(-Dk^2(t-t_0) + ik(x-y)\right) \tag{1.49}$$

Recall that:

$$\int_{-\infty}^{+\infty} \frac{\mathrm{d}k}{2\pi} e^{-iak^2 - ibk} = \frac{1}{\sqrt{4\pi a i}} \exp\left(\frac{ib^2}{4a}\right)$$

and so with $ia = D(t - t_0)$ and $b = -(x - y)$ we arrive at:

$$\int_{\mathbb{R}} \frac{\mathrm{d}k}{2\pi} \exp\left(-Dk^2(t-t_0) + ik(x-y)\right) = \frac{1}{\sqrt{4\pi D(t-t_0)}} \exp\left(-\frac{(x-y)^2}{4D(t-t_0)}\right)$$

Substituting back in (1.49):

$$W(x,t) = \frac{1}{\sqrt{4\pi D(t-t_0)}} \int_{\mathbb{R}} \mathrm{d}y\, W(y, t_0) \exp\left(-\frac{(x-y)^2}{4D(t-t_0)}\right) \tag{1.50}$$

Note that with $t_0 = 0$ and $W(y, t_0) = \delta(y)$ we retrieve the solution (1.36) that we already found.

1.5.1 Propagators

Suppose we know with certainty that the particle is in $y = x_0$ at time $t = t_0$, that is:

$$W(y, t_0) = \delta(y - x_0)$$

Then, substituting in (1.50) leads to:

$$W(x, t|x_0, t_0) \equiv \frac{1}{\sqrt{4\pi D(t-t_0)}} \exp\left(-\frac{(x-x_0)^2}{4D(t-t_0)}\right) \tag{1.51}$$

where with $W(x, t|x_0, t_0)$ we denote the probability that the particle will be *around* position x at time t, **given** it was certainly in x_0 at time t_0. $W(x, t|x_0, t_0)$ is also called **propagator**, as it "propagates" the particle from (x_0, t_0) to (x, t) as a sort of *continuous transition probability*. This is much more evident if we rewrite (1.50) as follows (with $y \to x_0$ for simplicity):

$$W(x,t) = \int_{\mathbb{R}} \mathrm{d}x_0\, W(x, t|x_0, t_0) W(x_0, t_0) \tag{1.52}$$

Let's explorer some properties of (1.52).

1. **ESCK property**. Let's propagate a particle from a starting point (x_0, t_0) to two different end points (x_1, t_1) and (x_2, t_2): *ESCK property*

$$W(x_1, t_1) = \int_{\mathbb{R}} \mathrm{d}x_0\, W(x_1, t_1|x_0, t_0) W(x_0, t_0) \tag{1.53}$$

$$W(x_2, t_2) = \int_{\mathbb{R}} \mathrm{d}x_0 \, W(x_2, t_2|x_0, t_0) W(x_0, t_0) \tag{1.54}$$

We can also propagate to (x_2, t_2) starting from (x_1, t_1):

$$W(x_2, t_2) = \int_{\mathbb{R}} \mathrm{d}x_1 \, W(x_2, t_2|x_1, t_1) W(x_1, t_1) \tag{1.55}$$

Now, if we substitute (1.53) in (1.55) we get:

$$W(x_2, t_2) = \iint_{\mathbb{R}^2} \mathrm{d}x_1 \, \mathrm{d}x_0 \, W(x_2, t_2|x_1, t_1) W(x_1, t_1|x_0, t_0) W(x_0, t_0)$$

By comparing this expression with (1.54), we find that:

$$W(x_2, t_2|x_0, t_0) = \int_{\mathbb{R}} \mathrm{d}x_1 \, W(x_2, t_2|x_1, t_1) W(x_1, t_1|x_0, t_0)$$

That is, the propagator between two *points* A and B can be obtained by multiplying the propagators between $A \to C$ and $C \to B$ and summing over *all possible choices* of C. This property is the Einstein-Smoluchowski-Kolmogorov-Chapman relation (**ESCK**).

2. **Correlator**. Consider two instants $t_1 \neq t_2$, and suppose we want to compute $\langle x(t_2) x(t_1) \rangle$, supposing that the particle started in $x = 0$ at $t = 0$. Applying the definition of an expected value:

Two-point correlator

$$\langle x(t_2) x(t_1) \rangle = \iint_{\mathbb{R}^2} \mathrm{d}x_1 \, \mathrm{d}x_2 \, \mathbb{P}(x_2, t_2; x_1, t_1|0, 0) x_2 x_1$$

where $\mathbb{P}(x_2, t_2; x_1, t_1|0, 0)$ is the *joint pdf* of a particle being around x_1 at t_1 and around x_2 at t_2, given the *initial position* in $x = 0$ at $t = 0$. Recall from probability theory that:

$$\begin{aligned}
\mathbb{P}(x_2, t_2; x_1, t_1; 0, 0) &= \mathbb{P}(x_2, t_2; x_1, t_1|0, 0) \mathbb{P}(0, 0) \\
\Rightarrow \ \mathbb{P}(x_2, t_2; x_1, t_1|0, 0) &= \frac{\mathbb{P}(x_2, t_2; x_1, t_1; 0, 0)}{\mathbb{P}(0, 0)} = \\
&= \frac{W(x_2, t_2|x_1, t_1) W(x_1, t_1|0, 0) \cancel{W(0, 0)}}{\cancel{W(0, 0)}} = \\
&= W(x_2, t_2|x_1, t_1) W(x_1, t_1|0, 0)
\end{aligned}$$

Recalling the result in (1.51) we can now compute:

$$\langle x(t_2) x(t_1) \rangle = \iint_{\mathbb{R}^2} \mathrm{d}x_1 \, \mathrm{d}x_2 \, x_1 x_2 \frac{\exp\left(-\frac{(x_2 - x_1)^2}{4D(t_2 - t_1)}\right)}{\sqrt{4\pi D(t_2 - t_1)}} \frac{\exp\left(-\frac{x_1^2}{4Dt_1}\right)}{\sqrt{4\pi D t_1}}$$

By changing variables ($x_1 = y_1$, $x_2 - x_1 = y_2$) we arrive at:

$$\begin{aligned}
&= \frac{1}{\sqrt{4\pi D(t_2 - t_1)}} \frac{1}{\sqrt{4\pi D t_1}} \iint_{\mathbb{R}^2} \mathrm{d}y_1 \, \mathrm{d}y_2 \, y_1 (y_1 + y_2) \cdot \\
&\quad \cdot \exp\left(-\frac{y_2^2}{4D(t_2 - t_1)} - \frac{y_1^2}{4Dt_1}\right) =
\end{aligned}$$

$$\underset{(a)}{=} \frac{1}{\sqrt{4\pi D(t_2-t_1)}} \frac{1}{\sqrt{4\pi D t_1}} \int_{\mathbb{R}} \mathrm{d}y_1 \, y_1^2 \exp\left(-\frac{y_1^2}{4Dt_1}\right) \cdot$$
$$\cdot \int_{\mathbb{R}} \mathrm{d}y_2 \exp\left(-\frac{y_2^2}{4D(t_2-t_1)}\right) =$$
$$\underset{(b)}{=} \frac{1}{\cancel{\sqrt{4\pi D(t_2-t_1)}}} \frac{1}{\cancel{\sqrt{4\pi D t_1}}} \left(2Dt_1 \cancel{\sqrt{4\pi D t_1}}\right) \left(\cancel{\sqrt{4\pi D(t_2-t_1)}}\right) = 2Dt_1$$

In (a) we note that by expanding $y_1(y_1+y_2)$, the term with $y_1 y_2$ is an odd function integrated over a symmetric domain, that results in 0. So, only the term with y_1^2 remains, allowing the integral's factorization. Then, in (b), we compute the Gaussian integrals, *supposing* $t_1 < t_2$ (so that $t_2 - t_1 > 0$) and recalling:

$$\int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2}ax^2\right) \mathrm{d}x = \sqrt{\frac{2\pi}{a}}$$
$$\int_{-\infty}^{+\infty} x^2 \exp\left(-\frac{1}{2}ax^2\right) = -2\frac{\mathrm{d}}{\mathrm{d}a} \int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2}ax^2\right) \mathrm{d}x = \sqrt{\frac{2\pi}{a}} \frac{1}{a}$$

The case when $t_1 > t_2$ leads to a similar result, with $t_1 \leftrightarrow t_2$. Thus, in general:

$$\langle x(t_1) x(t_2) \rangle = 2D \min(t_1, t_2)$$

By using the propagator we can compute the probability of passing through a set of points x_i at instants t_i:

Probability of a discrete path

$$\mathbb{P}(x_i, t_i; i = 0, \dots, n) = \mathbb{P}(x_n, t_n; x_{n-1}, t_{n-1}; \dots; x_1, t_1; x_0, t_0) =$$
$$= \prod_{i=1}^{n} W(x_i, t_i | x_{i-1}, t_{i-1}) W(x_0, t_0)$$

This is the joint probability for a *discrete trajectory*, meaning that we care only about what happens at certain discrete times.
This formula is useful to compute the average value of a generic function f of the trajectory points:

$$\langle f(x(t_n), x(t_{n-1}), \dots, x(t_0)) \rangle = \int_{\mathbb{R}^{n+1}} \left(\prod_{i=0}^{n} \mathrm{d}x_i \, W(x_i, t_i | x_{i-1}, t_{i-1} \right) f(x_n, x_{n-1}, \dots, x_0)$$

The need to extend this formula to an *infinite* number of intermediate points - that is for a *path in the continuum* will lead to the notion of **path integral**, that will be explored in detail in the next chapter.

Chapter 2

The Wiener Path Integral

2.1 Average over paths

(Lesson 4 of 24/10/19) Compiled: October 13, 2020

Consider an *unconstrained* Brownian particle, moving on the real line, starting in x_0 at t_0. By solving the diffusion equation we found that the probability of finding the particle in $[x, x + \mathrm{d}x]$ at time $t > t_0$ is given by the **propagator**:

$$\begin{aligned}\mathbb{P}\{x(t) \in [x, x+\mathrm{d}x] | x(t_0) = x_0\} = W(x,t|x_0,t_0)\,\mathrm{d}x = \\ = \frac{1}{\sqrt{4\pi D(t-t_0)}} \exp\left(-\frac{(x-x_0)^2}{4D(t-t_0)}\right)\mathrm{d}x\end{aligned} \tag{2.1}$$

By integrating (2.1) we can then find the probability of finding the particle inside an interval $[A, B]$ at time t:

$$\mathbb{P}\{x(t) \in [A,B] | x(t_0) = x_0\} = \int_A^B \mathrm{d}x\, W(x,t|x_0,t_0) \qquad t > t_0$$

We are now interested in computing the expected value $\langle f \rangle$ of **functionals** f of the trajectory, i.e. of quantities depending on several (or all) points of the trajectory $x(\tau)$ of a Brownian particle.

- The simplest example is the **correlation function**, which is defined as the product of the particle's position at two different times $t_1 < t_2$:

$$f(\{x(t_1), x(t_2)\}) = x(t_1)x(t_2) \quad t_1 < t_2$$

- A more general (and difficult) case is given by a function of the *entire* trajectory, such as:

$$f(\{x(\tau) \colon 0 < \tau \leq t\}) = g\left(\int_0^t x(\tau)a(\tau)\,\mathrm{d}\tau\right) \qquad a,\, g \colon \mathbb{R} \to \mathbb{R}$$

In other words, we want to compute the average of a function f over an ensemble of *random paths*. Every point of the path that is needed to compute f is a *dimension* of the integral for the average. So, if we need the *entire path*, we will need infinite points, leading to an integral *over infinite dimensions* - the **path integral**. We will now formalize it one step at a time.

2.1.1 Functions of a discrete number of points

Let's start from the simplest case, and consider the **correlation function**:

$$f(\{x(t_1), x(t_2)\}) = x(t_1)x(t_2) \qquad t_1 < t_2$$

To compute $\langle f \rangle$ we will need the *joint probability distribution* $g(x_1, x_2)$ that gives the probability of $x(t_1)$ being "close to" x_1 **and** $x(t_2)$ "close to" x_2 *for the same trajectory.* Let us denote the three events of interest:

$$\begin{aligned}
&A\colon \text{Particle starts in } x_0 \text{ at } t_0\\
&B\colon \text{Particle is close to } x_1 \text{ at } t_1 \ (x(t_1) \in [x_1, x_1 + \mathrm{d}x_1])\\
&C\colon \text{Particle is close to } x_2 \text{ at } t_2 \ (x(t_2) \in [x_2, x_2 + \mathrm{d}x_2])
\end{aligned}$$

We are interested in the joint probability $\mathbb{P}(C, B|A)$ (the order is defined by $t_2 > t_1 > t_0$). From probability theory:

$$\mathbb{P}(C, B|A) = \mathbb{P}(C|B, A)\mathbb{P}(B|A)$$

We already know how to compute probabilities like $\mathbb{P}(B|A)$, but not like $\mathbb{P}(C|B, A)$. Fortunately, that is not needed.
Recall, in fact, that Brownian motion is a *Markovian process*, meaning that the *future* depends only on the *present state*, i.e. the particle *has no memory.* So, **subsequent displacements are independent**: the probability of the particle going from x_1 to x_2 is the same whether it has started at x_0 or at any other point $\tilde{x}_0$. In other words, if we take the *present state* as the particle being in x_1 at t_1, the future (position at $t_2 > t_1$) depends only on that, and not on the past (position at t_0). So:

$$\mathbb{P}(C|B, A) = \mathbb{P}(C|B)$$

leading to:

$$\mathbb{P}(C, B|A) = \mathbb{P}(C|B)\mathbb{P}(B|A)$$

Inserting the *propagators* (2.1):

$$\mathrm{d}\mathbb{P}_{t_1,t_2}(x_1, x_2|x_0, t_0) \equiv W(x_2, t_2|x_1, t_1)W(x_1, t_1|x_0, t_0)\,\mathrm{d}x_1\,\mathrm{d}x_2$$

This is the joint probability we need to compute $\langle f \rangle$. Of course, nothing stops us from considering N "jumps" instead of only 2:

$$\begin{aligned}
\mathrm{d}\mathbb{P}_{t_1,\dots,t_n}(x_1, \dots, x_n|x_0, t_0) &\equiv W(x_n, t_n|x_{n-1}, t_{n-1}) \cdots W(x_1, t_1|x_0, t_0)\,\mathrm{d}x_1\,\mathrm{d}x_2 \dots \mathrm{d}x_n =\\
&= \exp\left(-\sum_{i=1}^n \frac{(x_i - x_{i-1})^2}{4D\Delta t_i}\right) \prod_{i=1}^n \frac{\mathrm{d}x_i}{\sqrt{4\pi D\Delta t_i}} \qquad (2.2)
\end{aligned}$$

Then, the average of a generic function $f(x(t_1), \dots, x(t_n))$ of the positions of the particle at times $t_1 < t_2 < \dots < t_n$ is defined as:

$$\langle f(x(t_1), \dots, x(t_n)) \rangle_W = \int_{\mathbb{R}^n} f(x_1, \dots, x_n)\,\mathrm{d}\mathbb{P}_{t_1,\dots,t_n}(x_1, \dots, x_n|x_0, t_0)$$

2.1.2 Functionals of the whole trajectory

The quantity in (2.2) can be interpreted as the *infinitesimal volume element* spanned by all the *trajectories* passing through a set of *tiny gates*, as represented in figure 2.1.

The underlying idea is that *probabilities* satisfy the axioms of *measures*, that is *functions that assign a measure*, i.e. a generalization of "size", *to all sets included in a specific collection.*

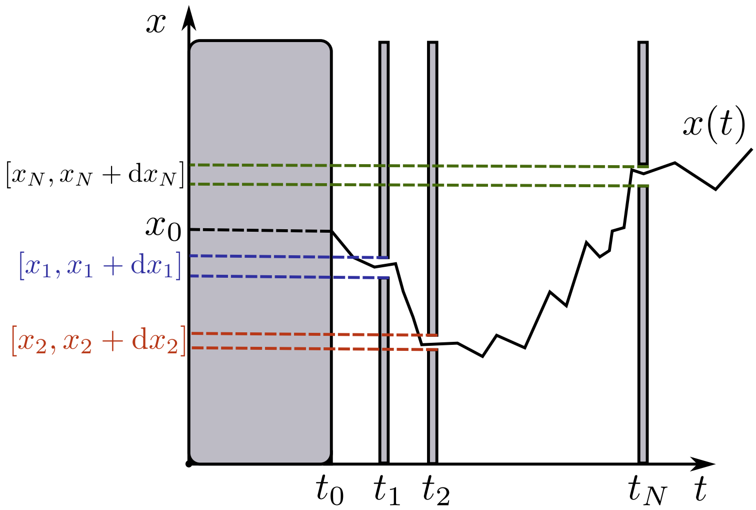


Figure (2.1) – All trajectories that pass through the set of *gates* $[x_i, x_i + dx_i]$ at times t_i (such as the $x(t)$ here represented) contribute to the *volume* $d\mathbb{P}_{t_1,\ldots,t_n}(x_1, \ldots, x_N)$

.

We now try to formalize this idea in order to extend the results of the previous section to the case of functions depending on a *infinite* number of trajectory points.

Path integral formalization

1. **Space definition**. Let $T \subset \mathbb{R}$ (**index set**), denote with $\mathbb{R}^T$ the set of all functions (**stochastic processes**) $k\colon T \to \mathbb{R}$. The idea is that an element of $\mathbb{R}^T$ is a collection of *random variables* indexed by T.

   In our case T is a collection of time instants (e.g. $T = [0, +\infty)$) and a generic element of $\mathbb{R}^T$ is made of *all the traversed points of a trajectory at times* T:

   $$\{x(t)\colon t \in T\} \in \mathbb{R}^T$$

2. **Probability measure on finite points**. The expression in (2.2), as observed, allows us to *measure* the volumes spanned by trajectories traversing a set of gates. Let's formalize this idea. Consider a *finite* set of times $T = \{t_i\}_{i=1,\ldots,n}$ with $n \in \mathbb{N}, t_i \in \mathbb{R}$ and $t_1 < t_2 < \ldots t_n$, each associated to a *gate* $H_i = [a_i, b_i]$, with $a_i, b_i \in \mathbb{R}$ and $a_i < b_i$. All the trajectories $\mathbb{R}^T$ traversing each H_i at a time $t_i \in T$ span a **cylindrical set** A of the form:

   $$A = \{x(t)\colon x(t_1) \in H_1, \ldots, x(t_n) \in H_n\} \subset \mathbb{R}^T$$

Using (2.2) and integrating over the *gates* we can define the measure of A-like sets as:

$$P_W(A) \equiv \int_{\mathbb{R}^n} \mathrm{d}\mathbb{P}_{t_1,\dots,t_n}(x_1,\dots,x_n|x_0,t_0)\mathbb{I}_{H_1}(x_1)\dots\mathbb{I}_{H_n}(x_n)$$

where $\mathbb{I}_{H_i}(x_i)$ are **characteristic functions** of the gates:

$$\mathbb{I}_{H_i}(x) = \begin{cases} 1 & x \in H_i \\ 0 & \text{otherwise} \end{cases}$$

In our case H_i are just intervals, and so:

$$P_W(A) \equiv \mathbb{P}_{t_1,\dots,t_n}(A) \underset{(2.2)}{=} \int_{H_1} \mathrm{d}x_1 \int_{H_2} \mathrm{d}x_2 \cdots \int_{H_n} \mathrm{d}x_n \left(\prod_{i=1}^{n} \frac{1}{\sqrt{4\pi D \Delta t_i}} \right) \cdot \qquad (2.3)$$

$$\cdot \exp\left(-\frac{(x_i - x_{i-1})^2}{4D\Delta t_i} \right)$$

with $\Delta t_i = t_i - t_{i-1}$.

3. **Generalization on infinite points**. Note that (2.3) holds for any n. So, using A-like sets, we can construct a σ-algebra[1] $\mathcal{F}$ of $\mathbb{R}^T$. Then, by applying **Kolmogorov extension theorem** we can extend the measure P_W we just found to the entire $\mathcal{F}$.

4. **Probability space**. We now have a set of *all possible outcomes* $\mathbb{R}^T$ (in our case, all the possible trajectories that can be produced by a Brownian motion). We also have the collection of all *events* $\mathcal{F}$, that is subsets of $\mathbb{R}^T$ for which is meaningful to assign a *probability measure* $P_W : \mathcal{F} \to [0,1]$. The triad $(\mathbb{R}^T, \mathcal{F}, P_W)$ forms a **probability space**, that gives a rigorous meaning to the concept of "computing the probability of a trajectory".

The **measure** so obtained is called **Wiener measure**, and denoted as the following:

Wiener measure

$$P_W(A) \equiv \int_A \mathrm{d}_W\, x(\tau)$$

Then we can compute expected values. For example, if $f(\{x(\tau) : \tau \in T\})$ is a function depending on the points traversed at times in a set T, then:

$$\langle f \rangle_W \equiv \int_{\mathbb{R}^T} f(x(\tau))\, \mathrm{d}_W\, x(\tau) \qquad T = [0, \infty)$$

Note that the Weiner measure *exists* and it's well defined (Kolmogorov's theorem), but we know it explicitly only in specific *finite cases*. So, to compute the expected value of functionals $F(\{x(t)\})$ over *continuous trajectories* we first *discretize* the trajectory, and then take a *continuum limit*.

Main technique to compute path integrals

[1] ∧A σ-algebra on a set X is a collection Σ of subsets of X that includes X itself, is closed under complement, and is closed under countable unions

1. Suppose we have a functional $F(\{x(\tau)\colon 0 < \tau < t\})$, and we want to compute $\langle F \rangle$.

2. We *discretize* the problem by *arbitrarily* subdividing the time interval $[0,t]$ in n parts $0 = t_0 < t_1 < t_2 < \cdots < t_n = t$. Then we consider an *approximated* functional $F_N(\{x(t_0),\dots,x(t_n)\})$ (for example approximating the path $x(\tau)$ with a piecewise linear function, depending only on $x(t_0),\dots,x(t_n)$), so that:

$$F = \text{“}\lim_{N\to\infty}\text{''} F_N$$

where $N \to \infty$ means that $\max \Delta t_i \to 0$, with $\Delta t_i = t_i - t_{i-1}$. This limit needs to be properly defined (by using the Weiner measure to define a norm in a space of *integrable* functionals, etc.), but we will not do that here.

3. Then the **Weiner path integral** is defined as:

$$\begin{aligned}\langle F \rangle_W &= \int_{\mathbb{R}^T} \mathrm{d}_W\, x(\tau) F(\{x(\tau)\colon 0<\tau<t\}) \equiv \text{“}\lim_{N\to\infty}\text{''}\langle F_N\rangle_W = \\ &= \text{“}\lim_{n\to\infty}\text{''} \int_{\mathbb{R}^n} \mathrm{d}\mathbb{P}_{t_1,\dots,t_n}(x_1,\dots,x_n|x_0,t_0) F_N(x(t_0),\dots,x(t_n))\end{aligned}$$

Geometrically, we are evaluating F for every possible Brownian path $x(\tau)$, and then averaging all these results, each weighted by the *probability* of the corresponding path.

Example 2 (Correlation function and ESCK property)**:**

As expected, the more general definition of the Weiner measure - involving the continuum limit $N \to \infty$ - reduces to (2.2) when evaluated for a function depending only on a finite set of particle's positions.
For example, consider the expected value of the *correlation function* (assume the particle starting in 0 at time 0 for simplicity):

$$\begin{aligned}\langle x(t_1')x(t_2')\rangle &= \int_{\mathbb{R}^T} \mathrm{d}_W x\, x_1(t_1')x_2(t_2') = \qquad T=[0,t],\ t_1'<t_2'<t \\ &= \text{“}\lim_{N\to\infty}\text{''}\int_{\mathbb{R}^N} \mathrm{d}\mathbb{P}_{t_1,\dots,t_N}(x_1,\dots,x_N|0,0)x(t_k)x(t_n) =\end{aligned}$$

where we chose the discretization so that $t_k = t_1'$ and $t_n = t_2'$. Then, by expanding the measure and applying the ESCK property we get (omitting the limit):

$$\begin{aligned}&= \int_{\mathbb{R}^N} \mathrm{d}x_1 \dots \mathrm{d}x_N\, W(x_N,t_N|x_{N-1},t_{N-1})\dots W(x_1,t_1|0,0)x_k x_n = \\ &\underset{(a)}{=} \int_{\mathbb{R}^2} \mathrm{d}x_k\,\mathrm{d}x_n\, W(x_n,t_n|x_k,t_k)W(x_k,t_k|0,0)x_k x_n\end{aligned}$$

where in (a) we used the ESCK property to compute all the integrals on $\mathrm{d}x_i$ with $i \neq n,k$, which evaluate all to 1.

We note that the same result can be obtained by direct application of (2.2):

$$\langle x(t_1')x(t_2')\rangle = \int \mathrm{d}x_1'\,\mathrm{d}x_2'\, W(x_2',t_2'|x_1',t_1')W(x_1',t_1'|0,0)x_1'x_2'$$

2.1.3 Change of random variables

In practice, to compute path integrals it will be useful to perform change of random variables. The idea is that we know the pdf for an increment Δx_i, and so we can compute - when needed - the pdf of functions of Δx_i.

So, consider a random variable $X \sim q(x)$, with $q(x)$ being a generic distribution (e.g. $q(x) = \mu e^{-\mu x}$). Now consider a function $y(x)$, e.g. $y(x) = x^2$. Y is then a new random variable, with a certain distribution $p(y)$. We now want to compute $p(y)$ starting from $q(x)$ and $y(x)$.
Suppose that $y(x)$ is invertible. Then, if we extract a value from X, it will be inside $[x, x + \mathrm{d}x]$ with a probability $q(x)\,\mathrm{d}x$. Knowing X, we can use the relation $y(x)$ to uniquely determine Y, that will be in $[y, y + \mathrm{d}y]$ with the same probability. So, the following holds:

$$q(x)\,\mathrm{d}x = p(y)\,\mathrm{d}y \tag{2.4}$$

We can compute $\mathrm{d}y$ by *nudging* $y(x)$, and expanding in Taylor series:

$$y(x + \mathrm{d}x) \equiv y + \mathrm{d}y + O(\mathrm{d}y^2) = y(x) + \underbrace{\mathrm{d}x\, y'(x)}_{\mathrm{d}y} + O(\mathrm{d}x^2)$$

and so $\mathrm{d}y = \mathrm{d}x\, y'(x)$. Substituting in (2.4) we get:

$$q(x)\,\mathrm{d}x = p(y)\,\mathrm{d}y = p(y(x))y'(x)\,\mathrm{d}x \Rightarrow p(y) = q(x(y))\frac{\mathrm{d}x}{\mathrm{d}y} \tag{2.5}$$

Consider now a more general change of variables $y = y(x)$ (not necessarily invertible), with $x \sim q(x)$. We start from the expected value of a function f in terms of $q(x)$:

$$\begin{aligned}
\langle f(y)\rangle &= \int_{\mathbb{R}} \mathrm{d}x\, f(y(x))q(x) = \\
&= \int_{\mathbb{R}} \mathrm{d}x\, f(y(x))q(x) \underbrace{\int_{\mathbb{R}} \mathrm{d}z\, \delta(z - y(x))}_{=1} = \\
&\underset{(a)}{=} \int_{\mathbb{R}} \mathrm{d}z\, f(z) \underbrace{\int_{\mathbb{R}} \mathrm{d}x\, q(x)\delta(z - y(x))}_{\langle\delta(z-y(x))\rangle_{q(x)}}
\end{aligned} \tag{2.6}$$

where in (a) we used the fact that $\delta(z - y(x)) = 1$ only when $z = y(x)$, and it's 0 otherwise, and so:

$$f(y(x)) = \int_{\mathbb{R}} \mathrm{d}z\, f(z)\delta(z - y(x))$$

Of course we can rewrite $\langle f\rangle$ directly in terms of $p(y)$:

$$\langle f(y)\rangle = \int_{\mathbb{R}} \mathrm{d}y\, f(y)p(y) \tag{2.7}$$

Comparing (2.6) with (2.7) and renaming $y \to z$ leads to:

$$p(z) = \int_{\mathbb{R}} \mathrm{d}x\, q(x)\delta(z - y(x)) = \langle\delta(z - y(x))\rangle_{q(x)} \tag{2.8}$$

which, in general, is not the same as the previously obtained result:

$$p(z) \neq q(x(z)) \frac{\mathrm{d}x(z)}{\mathrm{d}z}$$

To retrieve this special case we must assume $y(x)$ to be invertible, with inverse $x(y)$. This means that $\operatorname{sgn} y'(x) = A$, with $A \in \mathbb{R} \setminus \{0\}$ constant.
We want now to compute $\delta(z - y(x))$ in this case. Recall that $\delta \circ g$, if g is a continuously differentiable function with $g(x_0) = 0$ and $g'(x) \neq 0 \, \forall x$ is:

$$\delta(g(x)) = \frac{\delta(x - x_0)}{|g'(x_0)|}$$

So, if we let $g(x) = z - y(x)$, the only zero is at $x = x(z)$, as then $y(x(z)) = z$.
So:

$$\delta(z - y(x)) = \frac{\delta(x - x(z))}{|y'(x(z))|}$$

Substituting back in (2.8):

$$\begin{aligned} p(z) &= \langle \frac{\delta(x - x(z))}{|y'(x(z))|} \rangle_{q(x)} = \int_{\mathbb{R}} \mathrm{d}x \, q(x) \frac{\delta(x - x(z))}{|y'(x(z))|} = q(x(z)) |y'(x(z))|^{-1} = \\ &= q(x(z)) \frac{\mathrm{d}x(y)}{\mathrm{d}y} \Big|_{y = x(z)} \end{aligned} \tag{2.9}$$

which is the same rule found in (2.5).

2.2 Examples of path integrals

We now see some examples of explicit calculation of Wiener path integrals, that will be useful for the upcoming applications.

2.2.1 Transition probabilities

Thanks to the Wiener measure we have a way to assign probabilities to *paths* $x(\tau)$. We can recover from this the *transition probabilities* we started from, by considering the functional that *evaluates* a path at an instant t: $x(\tau) \mapsto x(t) \equiv x_t$. Then, by applying (2.8) we can compute the distribution followed by x_t:

Path with a constrained end-point

$$p(x_t) = W(x_t, t | 0, 0) = \langle \delta(x_t - x(\tau)) \rangle_W = \int_{\mathbb{R}^T} \mathrm{d}_W x(\tau) \, \delta(x_t - x(\tau)) \tag{2.10}$$

(The starting condition $x(0) = 0$ is *contained* in the definition of the measure $\mathrm{d}_W \, x(\tau)$).

So we can now write:

Path integral for a transition probability

$$\begin{aligned} W(x, t | 0, 0) &= \int_{\mathbb{R}^T} \mathrm{d}_W x \, \delta(x(t) - x) = \\ &= \text{“} \lim_{N \to \infty} \text{”} \int_{\mathbb{R}^{N+1}} \prod_{i=1}^{N+1} \frac{\mathrm{d}x_i}{\sqrt{4\pi D \Delta t_i}} \exp\left(- \sum_{i=1}^{N+1} \frac{(x_i - x_{i-1})^2}{4 D \Delta t_i} \right) \delta(x_{N+1} - x) \end{aligned}$$

where $t_n = t$, $x(t_n) = x_{N+1}$.

We already computed this result. In fact, recall that:

$$W(x_t, t|0,0) = \frac{1}{\sqrt{4\pi D t}} \exp\left(-\frac{x_t^2}{4Dt}\right)$$

If we set $x_t = 0$ (for simplicity), we get:

$$W(0, t|0,0) = \frac{1}{\sqrt{4\pi D t}} \tag{2.11}$$

As an exercise to get some familiarity with Wiener integrals, we will now *re-derive* this result, by evaluating the Weiner path integral in (2.10), with $x_t = 0$:

$$W(0,t|0,0) = \langle \delta(0 - x(\tau))\rangle_W = \int_{\mathbb{R}^T} \mathrm{d}_W x(\tau)\, \delta(x(\tau)) \equiv I_1 \tag{2.12}$$

First, it is convenient to establish some **additional notation**.

Notation for path ensembles

Let $T = [0, \infty)$. We denote with $\mathcal{C}\{0,0;t'\}$ the subset of trajectories in $\mathbb{R}^T$ starting from $x = 0$ at $t = 0$, and lasting a time span t'. Then, $\mathcal{C}\{0,0;x',t'\}$ is the subset of $\mathcal{C}\{0,0;t'\}$ when even the end-point is fixed to be $x(t') = x'$. The following **normalization property** holds:

$$\langle 1\rangle_W = \int_{\mathcal{C}\{0,0;t\}} 1 \cdot \mathrm{d}_W x(\tau) = 1$$

We can then rewrite (2.12) as:

$$I_1 = \int_{\mathcal{C}\{0,0;0,t\}} \mathrm{d}_W x(\tau)$$

Geometrically, $W(0,t|0,0) = I_1$ is the probability that a Brownian particle starting at the origin *returns* in it after a finite amount of time t.

Discretization

The standard way to compute a Wiener integral is to *discretize* it, and then take a *continuum* limit. So, consider for simplicity a **uniform** time discretization $\{t_i\}_{i=1,\dots,N+1}$, with instants ϵ-apart from each other, so that:

$$t_i - t_{i-1} \equiv \epsilon = \frac{t}{N+1} \qquad \forall i = 1, \dots, N+1$$

Note that the end-points are $x_0 = x_{N+1} = 0$.

1. Discretized path integral

We can rewrite (2.12) as the *continuum limit* of its discretization:

$$I_1 \equiv \lim_{\substack{\epsilon\to 0\\ N\to\infty}} I_1^{(N)} \tag{2.13}$$

$$I_1^{(N)} \equiv \frac{1}{(\sqrt{4\pi D\epsilon})^{N+1}} \int_{-\infty}^{+\infty} \mathrm{d}x_1 \int_{-\infty}^{+\infty} \mathrm{d}x_2 \cdots \int_{-\infty}^{+\infty} \mathrm{d}x_N \exp\left(-\frac{1}{4D\epsilon}\sum_{i=0}^{N}(x_{i+1} - x_i)^2\right) \tag{2.14}$$

where we already computed the integral over $\mathrm{d}x_{N+1}$ involving the δ, by just setting $x_{N+1} \equiv 0$.

Let's focus on the summation in the exponential:

2. Matrix form

$$\sum_{i=0}^{N}(x_{i+1}-x_i)^2 = x_1^2 + \cancel{x_0}^2 - 2x_0x_1 + x_2^2 + x_1^2 - 2x_1x_2 + \dots + \cancel{x_{N+1}}^2 + x_N^2 - \cancel{2x_N x_{N+1}} =$$
$$= 2(x_1^2 + \dots + x_N^2) - 2(x_1x_2 + x_2x_3 + \dots + x_{N-1}x_N) =$$
$$= 2\left(\sum_{i=1}^{N} x_i^2\right) - 2\left(\sum_{i=1}^{N-1} x_i x_{i+1}\right)$$

This is a *quadratic form*, i.e. a polynomial with all terms of order 2. So, it can be written in *matrix form*:

$$= \sum_{k,l=1}^{N} x_k A_{kl} x_l = \boldsymbol{x}^T A_N \boldsymbol{x}$$

for an appropriate choice of entries A_{kl} of the $N \times N$ matrix A_N:

$$A_{kk} = 2;\ A_{kl} = -(\delta k, l+1 + \delta_{k+1,l}) \Rightarrow A_N = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 2 & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{pmatrix}$$

Substituting back in (2.14):

$$I_1^{(N)} = \frac{1}{(\sqrt{4\pi D\epsilon})^{N+1}} \int_{-\infty}^{+\infty} \mathrm{d}x_1 \cdots \int_{-\infty}^{+\infty} \mathrm{d}x_N \exp\left(-\frac{\boldsymbol{x}^T A_N \boldsymbol{x}}{4D\epsilon}\right)$$

Recall the multivariate Gaussian integral:

3. Multivariate Gaussian

$$\int_{-\infty}^{+\infty} \mathrm{d}x_1 \cdots \mathrm{d}x_N \exp\left(-\sum_{ij}^{N} B_{ij} x_i x_j\right) = \frac{(\sqrt{\pi})^N}{\sqrt{\det B}}$$

with $B = A_N/(4D\epsilon)$, leading to:

$$I_1^{(N)} \underset{(a)}{=} \frac{1}{(\sqrt{4\pi D\epsilon})^{N+1}} \sqrt{\frac{\pi^N}{\det(A_N)\left[\frac{1}{4D\epsilon}\right]^N}} = \frac{1}{(\sqrt{4\pi D\epsilon})^{N+1}} \frac{\sqrt{4\pi D\epsilon}^N}{\sqrt{\det A_N}} =$$
$$= \frac{1}{\sqrt{4\pi D\epsilon}} \frac{1}{\sqrt{\det A_N}} \tag{2.15}$$

where in (a) we used the property of the determinant $\det(cA) = c^n \det(A)\, \forall c \in \mathbb{R}$.

Now, all that's left is to compute the determinant of A_N. Fortunately, as A_N is a *tri-diagonal* matrix, there is a recurrence relation in terms of the leading principal minors of A_N, which turns out to be multiples of the determinants of A_{N-1} and A_{N-2}.

4. Determinant of a tri-diagonal matrix

Explicitly, consider A_N:

$$\det A_N \equiv \begin{vmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 2 & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{vmatrix}_{N\times N}$$

and start computing the determinant following the last column. The only non-zero contributions are:

$$\det A_N = \underbrace{(-1)^{(N-1)+N}(-1)}_{+1} \begin{vmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 2 & -1 \\ 0 & \cdots & 0 & 0 & -1 \end{vmatrix}_{(N-1)\times(N-1)} + (-1)^{2N}(2)\det A_{N-1} =$$

$$= (-1)^{2(N-1)}(-1)\det A_{N-2} + 2\det A_{N-1} = 2\det A_{N-1} - \det A_{N-2} \tag{2.16}$$

where the terms in blue are just the alternating signs from the determinant expansion, and the other colours identify the matrix entries that are being used.

Then, it is just a matter of computing the first two terms of the succession ($|A_N| = \det A_N$ for brevity):

$$|A_1| = 2 \qquad |A_2| = \begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} = 4 - 1 = 3$$

And now we can use (2.16) to iteratively compute all $|A_N|$, e.g. $|A_3| = 2 \cdot 3 - 2 = 4$. To find $|A_N|$ for a *generic* N, we need to make an hypothesis, and then verify that it is compatible with (2.16). In this case, note that $|A_N| = N+1$ $(*)$ for all the examples we explicitly computed. Then, by induction:

$$|A_{N+1}| \underset{(2.16)}{=} 2 \cdot |A_N| - |A_{N-1}| \underset{(*)}{=} 2 \cdot (N+1) - (N-1+1) = 2N + 2 - N = (N+1)+1$$

which is indeed compatible with $(*)$. So, substituting back in (2.15) we get:

$$I_1^{(N)} = \frac{1}{\sqrt{4\pi D\epsilon}} \frac{1}{\sqrt{N+1}} \underset{(a)}{=} \frac{1}{\sqrt{4\pi Dt}}$$

where in (a) we used $\epsilon = t/(N+1) \Rightarrow N+1 = t/\epsilon$ from the discretization. Note that this result is *constant* with respect to ϵ or N (recall that t is fixed beforehand) and so taking the *continuum* limit leads immediately to I_1 (2.13):

$$I_1 \equiv \lim_{\substack{\epsilon\to 0\\ N\to\infty}} \frac{1}{\sqrt{4\pi Dt}} = \frac{1}{\sqrt{4\pi Dt}}$$

which is coherent with the result we previously computed (2.11).

2.2.2 Integral functional

(Lesson 6 of 28/10/19) Compiled: October 13, 2020

Consider a Brownian trajectory $x(\tau)$ (from now on, we will assume that all trajectories start in $x = 0$ at $t = 0$), and a functional that *weights* every traversed point $x(\tau)$ with a function $a\colon \mathbb{R} \to \mathbb{R}$, and then applies another function $F\colon \mathbb{R} \to \mathbb{R}$ to the total integral:

$$F[x(\tau)] = F\left(\int_0^t a(\tau)x(\tau)\,\mathrm{d}\tau\right)$$

For simplicity, we set $D = 1/4$, so that:

$$\mathrm{d}\mathbb{P}_{t_1,\ldots,t_n}(x_1,\ldots,x_n|0,0) = \exp\left(-\sum_{i=1}^n \frac{(x_i - x_{i-1})^2}{\Delta t_i}\right)\prod_{i=1}^n \frac{\mathrm{d}x_i}{\sqrt{\pi\Delta t_i}}$$

This is equivalent to a time rescaling $t \to \tau = 4Dt$.
We want now to compute $\langle F\rangle$:

$$I_3 \equiv \langle F[x(\tau)]\rangle_w = \int_{\mathcal{C}\{0,0;t\}} \mathrm{d}_W x(\tau)\, F[x(\tau)]$$

Note: the next computations will follow the book. Prof. Maritan's method for evaluating I_3 is quicker, but more advanced, and will be presented at the end.

Then we start by discretizing, by choosing a time grid $0 = t_0 < t_1 < \cdots < t_N = t$:

1. Discretized path integral

$$I_3 = \lim_{N\to\infty} I_3^{(N)}$$

$$I_3^{(N)} = \int_{-\infty}^{+\infty} \frac{\mathrm{d}x_1}{\sqrt{\pi\Delta t_1}}\cdots\int_{-\infty}^{+\infty}\frac{\mathrm{d}x_N}{\sqrt{\pi\Delta t_N}} F\left(\sum_{i=1}^N a_i x_i \Delta t_i\right)\exp\left(-\sum_{i=1}^N \frac{(x_i - x_{i-1})^2}{\Delta t_i}\right)$$

$$a_i \equiv a(t_i) \qquad x_i \equiv x(t_i)$$

This integral can be evaluated by transforming it to a *gaussian integral* that we already know. So we start by **changing variables**:

2. Change of variables

$$x_i - x_{i-1} = y_i \qquad i = 1,\ldots,N \tag{2.17}$$

Note that:

$$\sum_{j=1}^i y_j = \cancel{x_1} - \underbrace{x_0}_{=0} + x_2 - \cancel{x_1} + \cdots + x_i - \cancel{x_{i-1}} = x_i \qquad 1 \le i \le N$$

So, when we compute the transformation of the volume element:

$$\det\left|\frac{\partial\{x_i\}}{\partial\{y_j\}}\right| = \det\begin{vmatrix} 1 & 0 & 0 & 0\\ 1 & 1 & 0 & 0\\ \vdots & \vdots & \ddots & 0\\ 1 & 1 & \cdots & 1\end{vmatrix}_{N\times N} = 1$$

as the determinant of a lower triangular matrix is equal to the product of the diagonal entries.
All that's left is to transform the argument of F. Let's start by writing the first terms of the sum and apply the change of variables:

$$\begin{aligned}\sum_{i=1}^{N} a_i x_i \Delta t_i &= a_1 x_1 \Delta t_1 + a_2 x_2 \Delta t_2 + \cdots = \\ &= a_1 (y_1) \Delta t_1 + a_2 (y_1 + y_2) \Delta t_2 + \cdots = \\ &= y_1 \left(\sum_{j=1}^{N} a_j \Delta t_j \right) + y_2 \left(\sum_{j=2}^{N} a_j \Delta t_j \right) + \cdots + y_N a_N \Delta t_N = \\ &= \sum_{i=1}^{N} y_i \underbrace{\left(\sum_{j=i}^{N} a_j \Delta t_j \right)}_{A_i} \equiv \sum_{i=1}^{N} A_i y_i \end{aligned} \tag{2.18}$$

Substituting everything back:

$$I_3^{(N)} = \int_{-\infty}^{+\infty} \frac{\mathrm{d}y_1}{\sqrt{\pi \Delta t_1}} \cdots \int_{-\infty}^{+\infty} \frac{\mathrm{d}y_N}{\sqrt{\pi \Delta t_N}} F\left(\sum_{i=1}^{N} A_i y_i \right) \exp\left(-\sum_{i=1}^{N} \frac{y_i^2}{\Delta t_i} \right) \qquad A_i = \sum_{j=i}^{N} a_j \Delta t_j$$

We can simplify this integral a bit more by rescaling the y_i:

$$z_i = A_i y_i \qquad \mathrm{d}y_i = \frac{\mathrm{d}z_i}{A_i}$$

As each y_i is transformed independently, the jacobian is diagonal.

$$I_3^{(N)} = \int_{-\infty}^{+\infty} \frac{\mathrm{d}z_1}{\sqrt{\pi A_1^2 \Delta t_1}} \cdots \int_{-\infty}^{+\infty} \frac{\mathrm{d}z_N}{\sqrt{\pi A_N^2 \Delta t_N}} F(z_1 + \cdots + z_N) \exp\left(-\sum_{i=1}^{N} \frac{z_i^2}{A_i^2 \Delta t_i} \right)$$

This is the expected value of a function of the *sum* of N normally distributed random variables $\{z_i\}$. The idea is now to *isolate* one of them from the argument of F, integrate over it, and reiterate. This is done by changing variables yet again:

3. Second change of variables

$$\begin{cases} \eta = z_1 + z_2 \\ \xi = z_2 \end{cases} \Rightarrow \begin{cases} z_1 = \eta - \xi \\ z_2 = \xi \end{cases} \Rightarrow \det \left| \frac{\partial\{z_1, z_2\}}{\partial\{\eta, \xi\}} \right| = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1$$

leading to:

$$\begin{aligned} I_3^{(N)} = &\int_{-\infty}^{+\infty} \frac{\mathrm{d}\eta}{\sqrt{\pi A_1^2 \Delta t_1}} \int_{-\infty}^{+\infty} \frac{\mathrm{d}\xi}{\sqrt{\pi A_2^2 \Delta t_2}} \int_{-\infty}^{+\infty} \frac{\mathrm{d}z_3}{\sqrt{\pi A_3^2 \Delta t_3}} \cdots \int_{-\infty}^{+\infty} \frac{\mathrm{d}z_N}{\sqrt{\pi A_N^2 \Delta t_N}} \cdot \\ &\cdot F(\eta + z_3 + \cdots + z_N) \exp\left(-\frac{(\eta - \xi)^2}{A_1^2 \Delta t_1} - \frac{\xi^2}{A_2^2 \Delta t_2} - \sum_{i=3}^{N} \frac{z_i^2}{A_i^2 \Delta t_i} \right) \end{aligned}$$

Note how ξ does not enter in the F argument, and so we can integrate over it:

$$I_\xi = \int_{-\infty}^{+\infty} \mathrm{d}\xi \, \frac{1}{\sqrt{\pi A_1^2 \Delta t_1} \sqrt{\pi A_2^2 \Delta t_1}} \exp\left(-\frac{(\eta - \xi)^2}{A_1^2 \Delta t_1} - \frac{\xi^2}{A_2^2 \Delta t_2} \right) =$$

$$= \int_{-\infty}^{+\infty} \mathrm{d}\xi\,(\cdots)\exp\left(-\frac{\xi^2(A_1^2\Delta t_1 + A_2^2\Delta t_2) - \xi(2\eta A_2^2\Delta t_2) - (-\eta^2 A_2^2 \Delta t_2)}{A_1^2 A_2^2 \Delta t_1 \Delta t_2}\right)$$

Recall the gaussian integral formula:

4. Gaussian integral

$$\int_{-\infty}^{+\infty} \exp\left(-a x^2 + b x + c\right)\mathrm{d}x = \sqrt{\frac{\pi}{a}}\exp\left(\frac{b^2}{4a} + c\right) \tag{2.19}$$

which evaluates to:

$$I_\xi = \frac{1}{\sqrt{\pi(A_1^2\Delta t_1 + A_2^2\Delta t_2)}}\exp\left(-\frac{\eta^2}{A_1^2\Delta t_1 + A_2^2\Delta t_2}\right)$$

and substituting back in $I_3^{(N)}$:

$$I_3^{(N)} = \int_{-\infty}^{+\infty}\frac{\mathrm{d}\eta}{\sqrt{\pi A_1^2\Delta t_1 + \pi A_2^2 \Delta t_2}}\int_{-\infty}^{+\infty}\frac{\mathrm{d}z_3}{\sqrt{\pi A_3^2\Delta t_3}}\cdots\int_{-\infty}^{+\infty}\frac{\mathrm{d}z_N}{\sqrt{\pi A_N^2 \Delta t_N}}\cdot$$
$$\cdot F(\eta + z_3 + \cdots + z_N)\exp\left(-\frac{\eta^2}{A_1^2\Delta t_1 + A_2^2\Delta t_2} - \sum_{i=3}^N \frac{z_i^2}{A_1^2\Delta t_i}\right)$$

We can now reiterate this procedure until only one integration is left:

$$I_3^{(N)} = \int_{-\infty}^{+\infty}\frac{\mathrm{d}z}{\sqrt{\pi\sum_{i=1}^N A_i^2\Delta t_i}}F(z)\exp\left(-\frac{z^2}{\sum_{i=1}^N A_i^2 \Delta t_i}\right)$$

We are now finally ready to take the continuum limit $\Delta t_i \to 0$, $N \to \infty$. Note that:

5. Continuum limit

$$\lim_{\Delta t_i \to 0} A_i = \int_\tau^t a(s)\,\mathrm{d}s = A(\tau) \tag{2.20}$$

as the discrete sum goes from $t_i = \tau$ to $t_N = t$. Then:

$$R \equiv \lim_{\Delta t \to 0}\sum_{i=1}^N A_i^2 \Delta t_i = \int_0^t \mathrm{d}\tau\left(\int_\tau^t \mathrm{d}s\, a(s)\right)^2$$

and so:

$$I_3 = \lim_{N\to\infty} I_3^{(N)} = \int_{-\infty}^{+\infty}\mathrm{d}z\,\frac{F(z)}{\sqrt{\pi R}}\exp\left(-\frac{z^2}{R}\right)$$

And to recover D we can just substitute $R \to 4DR$.

Alternative method

We consider now a different (quicker) technique to compute I_3. We start again from:

$$I_3 \equiv \langle F[x(\tau)]\rangle_w = \int_{C\{0,0;t\}}\mathrm{d}_W x(\tau)\, F\left(\int_0^t a(\tau)x(\tau)\,\mathrm{d}\tau\right)$$

It is convenient to apply the change of variables we did in (2.18). We can do *before* discretizing, by defining $A(\tau)$ as in (2.20):

1. Auxiliary function $A(\tau)$

$$A(\tau) \equiv \int_\tau^t a(s)\,\mathrm{d}s \tag{2.21}$$

Note that $\dot{A}(\tau) = -a(\tau)$, and so the argument of F becomes:

$$\int_0^t a(\tau)x(\tau)\,\mathrm{d}\tau = -\int_0^t \partial_\tau A(\tau)x(\tau)\,\mathrm{d}\tau$$

Integrating by parts, noting that $A(t) = 0$ and $x(0) = 0$ leads to:

$$= -\cancel{[x(\tau)A(\tau)]_{\tau=0}^{\tau=t}} + \int_0^t A(\tau)\dot{x}(\tau)\,\mathrm{d}\tau$$

And now we discretize the path over the instants $0 = t_0 < t_1 < \cdots < t_N$, so that:

2. Path integral discretization

$$\begin{aligned}\int_0^t A(\tau)\dot{x}(\tau)\,\mathrm{d}\tau &= \lim_{\Delta t_i \to 0} \sum_{i=1}^N A(t_i)\frac{x(t_i) - x(t_{i-1})}{\Delta t_i}\Delta t_i = \\ &= \lim_{N\to\infty} \sum_{i=1}^N A_i(x_i - x_{i-1}) = \lim_{N\to\infty}\sum_{i=1}^N A_i \Delta x_i \qquad \begin{matrix} x_i \equiv x(t_i) \\ A_i \equiv A(t_i)\end{matrix}\end{aligned}$$

Substituting back (here $D = 1/4$ for simplicity):

$$\begin{aligned} I_3 &= \lim_{N\to\infty} I_3^{(N)} \\ I_3^{(N)} &= \int_{\mathbb{R}^N} \left(\prod_{i=1}^N \frac{\mathrm{d}x_i}{\sqrt{\pi \Delta t_i}}\right) \exp\left(-\sum_{i=1}^N \frac{(\Delta x_i)^2}{\Delta t_i}\right) F\left(\sum_{i=1}^N A_i \Delta x_i\right)\end{aligned}$$

The idea is now to *apply* a change of random variable, rewriting the average $\langle F[x(\tau)]\rangle_w$ (according to the distribution of *paths*) as the average $\langle F(z)\rangle_{p(z)}$, where $p(z)$ is the distribution followed by the argument of F:

$$\sum_{i=1}^N A_i \Delta x_i$$

So, we begin by inserting the appropriate δ:

3. Change of random variables

$$I_3^{(N)} = \int_{\mathbb{R}^N} \left(\prod_{i=1}^N \frac{\mathrm{d}x_i}{\sqrt{\pi \Delta t_i}}\right) \exp\left(-\sum_{i=1}^N \frac{(\Delta x_i)^2}{\Delta t_i}\right) F\left(\sum_{i=1}^N A_i \Delta x_i\right) \underbrace{\int_{\mathbb{R}} \mathrm{d}z\,\delta\left(z - \sum_{i=1}^N A_i\Delta x_i\right)}_{=1}$$

Exchanging the integrals leads to:

$$\begin{aligned}&\langle F\left(\sum_{i=1}^N A_i \Delta x_i\right)\rangle_w = \langle F(z)\rangle_{p(z)} = \\ &= \int_{\mathbb{R}} \mathrm{d}z\,F(z) \underbrace{\int_{\mathbb{R}^N} \left(\prod_{i=1}^N \frac{\mathrm{d}x_i}{\sqrt{\pi\Delta t_i}}\right)\delta\left(z - \sum_{i=1}^N A_i \Delta x_i\right)\exp\left(-\sum_{i=1}^N \frac{(\Delta x_i)^2}{\Delta t_i}\right)}_{p(z)}\end{aligned}$$

We can evaluate $I_3^{(N)}$ by transforming it to a *gaussian integral*. First, we remove the δ with a Fourier transform:

4. Fourier transform

$$2\pi\delta(x) = \int_{\mathbb{R}} e^{i\alpha x}\, \mathrm{d}\alpha$$

which, in this case, leads to:

$$\delta\left(z - \sum_{i=1}^{N} A_i \Delta x_i\right) = \int_{\mathbb{R}} \frac{\mathrm{d}\alpha}{2\pi} \exp\left(i\alpha\left(z - \sum_{i=1}^{N} A_i \Delta x_i\right)\right)$$

Substituting back:

$$I_3^{(N)} = \int_{\mathbb{R}} \frac{\mathrm{d}\alpha}{2\pi} \int_{\mathbb{R}} \mathrm{d}z\, F(z) e^{i\alpha z} \int_{\mathbb{R}^N} \left(\prod_{i=1}^{N} \frac{\mathrm{d}x_i}{\sqrt{\pi \Delta t_i}}\right) \exp\left(-\sum_{i=1}^{N} \frac{\Delta x_i^2}{\Delta t_i} - i\alpha \sum_{i=1}^{N} A_i \Delta x_i\right)$$

We see that the last term is similar to a multivariate gaussian with a imaginary term, that we know how to integrate. We just need to remove the *differences* in the exponential with a change of variables (as in (2.17)):

5. Change of variables

$$\begin{aligned}
y_1 &= \Delta x_1 = x_1 - \overbrace{x_0}^{=0} = x_1 \\
y_2 &= \Delta x_2 = x_2 - x_1 \\
&\vdots \\
y_N &= \Delta x_N = x_N - x_{N-1}
\end{aligned}$$

The volume element will be transformed by the determinant of the Jacobian:

$$J = \det \frac{\partial(x_1 \dots x_N)}{\partial(y_1 \dots y_N)} = \left[\det \frac{\partial(y_1 \dots y_N)}{\partial(x_1 \dots x_N)}\right]^{-1} = \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & 0 & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{vmatrix}^{-1} = 1$$

where we used the fact that $\det A^{-1} = (\det A)^{-1}$, and that the determinant of a lower triangular matrix is just the product of the diagonal entries.
The integral then becomes:

$$\begin{aligned}
I_3^{(N)} &= \int_{\mathbb{R}} \frac{\mathrm{d}\alpha}{2\pi} \int_{\mathbb{R}} \mathrm{d}z\, F(z) e^{i\alpha z} \int_{\mathbb{R}^N} \left(\prod_{i=1}^{N} \frac{\mathrm{d}y_i}{\sqrt{\pi \Delta t_i}}\right) \exp\left(-\sum_{i=1}^{N} \frac{y_i^2}{\Delta t_i} - i\alpha \sum_{i=1}^{N} A_i y_i\right) = \\
&= \int_{\mathbb{R}} \frac{\mathrm{d}\alpha}{2\pi} \int_{\mathbb{R}} \mathrm{d}z\, F(z) e^{i\alpha z} \left[\prod_{i=1}^{N} \int_{\mathbb{R}} \frac{\mathrm{d}y_i}{\sqrt{\pi \Delta t_i}} \exp\left(-\frac{y_i^2}{\Delta t_i} - i\alpha A_i y_i\right)\right]
\end{aligned}$$

The terms in the product are all independent gaussian integrals. Recall that:

6. Gaussian integral

$$\int_{\mathbb{R}} \mathrm{d}k\, e^{-iak^2 - ibk} = \sqrt{\frac{\pi}{ia}} \exp\left(\frac{ib^2}{4a}\right) \tag{2.22}$$

So, with $ia = 1/\Delta t_i$ and $b = \alpha A_i$ we get:

$$\int_{\mathbb{R}} \frac{\mathrm{d}y_i}{\sqrt{\pi \Delta t_i}} \exp\left(-\frac{y_i^2}{\Delta t_i} - i\alpha A_i y_i\right) = \exp\left(-\frac{\alpha^2 A_i^2 \Delta t_i}{4}\right)$$

and substituting back in the integral leads to:

$$\begin{aligned} I_3^{(N)} &= \int_{\mathbb{R}} \frac{\mathrm{d}\alpha}{2\pi} \int_{\mathbb{R}} \mathrm{d}z\, F(z) e^{i\alpha z} \left[\prod_{i=1}^{N} \exp\left(-\frac{\alpha^2 A_i^2 \Delta t_i}{4}\right)\right] = \\ &= \int_{\mathbb{R}} \frac{\mathrm{d}\alpha}{2\pi} \int_{\mathbb{R}} \mathrm{d}z\, F(z) e^{i\alpha z} \exp\left(-\frac{\alpha^2}{4} \sum_{i=1}^{N} A_i^2 \Delta t_i\right) \end{aligned}$$

Applying the continuum limit ($N \to \infty$, $\Delta t_i \to 0$), the exponential argument becomes the limit of a Riemann sum, i.e. a integral:

7. Continuum limit

$$\sum_{i=1}^{N} A(t_i)^2 \Delta t_i \xrightarrow[N\to\infty]{} \int_0^t A^2(\tau)\,\mathrm{d}\tau \underset{(2.21)}{=} \int_0^t \mathrm{d}\tau \left(\int_\tau^t \mathrm{d}s\, a(s)\right)^2 \equiv R(t)$$

Substituting back:

$$I_3 \equiv \langle F\left(\int_0^t a(\tau)x(\tau)\right)\rangle = \lim_{N\to\infty} I_3^{(N)} = \int_{\mathbb{R}} \mathrm{d}z\, F(z) \int_R \frac{\mathrm{d}\alpha}{2\pi} \exp\left(-\frac{\alpha^2}{4} R(t) + i\alpha z\right)$$

All that's left is to evaluate the last gaussian integral thanks to (2.22) with $ia = R(t)/4$ and $b = -z$, leading to:

$$I_3 = \int_{\mathbb{R}} \mathrm{d}z\, F(z) \frac{1}{2\pi} \sqrt{\frac{4\pi}{R(t)}} \exp\left(-\frac{z^2}{R(t)}\right) = \frac{1}{\sqrt{\pi R(t)}} \int_{\mathbb{R}} \mathrm{d}z\, F(z) \exp\left(-\frac{z^2}{R(t)}\right)$$

So, we showed that:

$$\langle F\left(\int_0^t a(\tau)x(\tau)\,\mathrm{d}\tau\right)\rangle_w = \sqrt{\frac{1}{\pi R(t)}} \int_{\mathbb{R}} \mathrm{d}z\, F(z) \exp\left(-\frac{z^2}{R(t)}\right); \qquad R(t) \equiv \int_0^t \mathrm{d}\tau \left(\int_\tau^t a(s)\,\mathrm{d}s\right)^2 \tag{2.23}$$

Example 3 (Generating function):

Let $F(z) = e^{hz}$. Inserting in (2.23) results in:

$$\langle \exp\left(h\int_0^t a(\tau)x(\tau)\,\mathrm{d}\tau\right)\rangle_w = \frac{1}{\sqrt{\pi R}}\int_{\mathbb{R}} \mathrm{d}z \exp\left(-\frac{z^2}{R} + hz\right) \underset{(a)}{=} \exp\left(\frac{h^2 R}{4}\right) \equiv G(h) \tag{2.24}$$

where in (a) we used formula (2.19) with $a = 1/R$ and $b = h$.
Note that $G(h)$ is the **moment generating function** (see (5.1) at pag. 118) of the integral:

$$I = \int_0^t a(\tau)x(\tau)\,\mathrm{d}\tau$$

We can then retrieve the n-th moment of I by computing the n-th derivative of $G(h)$:

$$\frac{\mathrm{d}^n}{\mathrm{d}h^n} G(h)\Big|_{h=0} = \langle I^n \rangle_w$$

We can see this by differentiating the left side of (2.24):

$$G'(h) = \langle \int_0^t a(\tau)x(\tau)\,\mathrm{d}\tau \exp\left(h\int_0^\tau ax\,\mathrm{d}\tau\right)\rangle_w$$

and then setting $h = 0$:

$$G'(0) = \langle \int_0^t a(\tau)x(\tau)\,\mathrm{d}\tau\rangle_w = \langle I \rangle_w$$

Then, differentiating the right side of (2.24) we have immediately the result:

$$\langle I \rangle_w = G'(h)\Big|_{h=0} = \frac{h}{2} R \exp\left(\frac{h^2 R}{4}\right)\Big|_{h=0} = 0$$

If we differentiate again we get the second moment:

$$G''(h) = \frac{R}{2}\exp\left(\frac{h^2 R}{4}\right) + \frac{h^2}{4} R^2 \exp\left(\frac{h^2 R}{4}\right) \Rightarrow G''(0) = \langle I^2 \rangle_w = \frac{R}{2}$$

Consider now a generic odd moment:

$$\langle \left(\int_0^t a(\tau)x(\tau)\,\mathrm{d}\tau\right)^{2k+1} \rangle_w = 0 \qquad \forall k \in \mathbb{N}$$

In fact, if we expand $G(h)$, we get:

$$G(h) = \sum_{n=0}^{\infty} \left(\frac{R}{4}\right)^n \frac{1}{n!} h^{2n}$$

Since all the powers are even, if we differentiate an odd number of times and set $h = 0$ we are "selecting" an odd power - which just is not there - and so the result will be 0.
On the other hand, an even moment leads to:

$$\langle \left(\int_0^t a(\tau)x(\tau)\,\mathrm{d}\tau\right)^{2k} \rangle_w = \left(\frac{R}{2}\right)^k \frac{(2k)!}{2^k k!}$$

(computations omitted).

2.2.3 Potential-like functional

We consider now the following functional:

$$F[x(\tau)] = \exp\left(-\int_0^t \mathrm{d}\tau\, P(\tau)x^2(\tau)\right)$$

As before, we wish to compute $\langle F\rangle_w$. We start by discretizing the path over a **uniform**[2] grid $0 = t_0 < t_1 < \cdots < t_N = t$ so that $\Delta t_i = t_i - t_{i-1} \equiv \epsilon = t/N$.

1. Discretized path integral

$$I_4 \equiv \int_{\mathcal{C}\{0,0;t\}} \mathrm{d}_W x(\tau) \exp\left(-\int_0^t \mathrm{d}\tau\, P(\tau)x^2(\tau)\right) = \lim_{N\to\infty} I_4^{(N)}$$

$$I_4^{(N)} = \int_{-\infty}^{+\infty} \frac{\mathrm{d}x_1}{\sqrt{\pi\epsilon}} \cdots \int_{-\infty}^{+\infty} \frac{\mathrm{d}x_N}{\sqrt{\pi\epsilon}} \exp\left(-\sum_{i=1}^N P_i x_i^2 \epsilon - \sum_{i=1}^N \frac{(x_i - x_{i-1})^2}{\epsilon}\right) \qquad \begin{aligned} x_i &\equiv x(ti) \\ P_i &\equiv P(t_i)\end{aligned} \tag{2.25}$$

The exponential argument is a **quadratic form**:

2. Rewrite the exponential in matrix form

$$\begin{aligned}
&-\epsilon(P_1 x_1^2 + \cdots + P_N x_N^2) - \frac{1}{\epsilon}[\cancel{x_0^2} + x_1^2 - \cancel{2x_0}x_1 + x_1^2 + x_2^2 - 2x_1x_2 + \cdots + \\
&\quad + \ldots x_{N-1}^2 + x_N^2 - 2x_{N-1}x_N] = \\
&= -\epsilon\sum_{i=1}^N P_i x_i^2 - \frac{1}{\epsilon}\left[2\sum_{i=1}^{N-1} x_i^2 + x_N^2 - 2\sum_{i=1}^N x_{i-1}x_i\right] = \\
&= -\left[x_1^2\left(\epsilon P_1 + \frac{2}{\epsilon}\right) + \cdots + x_{N-1}^2\left(\epsilon P_{N-1} + \frac{2}{\epsilon}\right) + x_N^2\left(\epsilon P_N + \frac{1}{\epsilon}\right) - \frac{2}{\epsilon}\sum_{i=1}^N x_i x_{i-1}\right] = \\
&= -\sum_{i,j=1}^N A_{ij} x_i x_j
\end{aligned}$$

where A_{ij} are matrix elements of a matrix A_N:

$$A_{ij} = \delta_{ij} a_i - \frac{1}{\epsilon}(\delta_{i,j-1} + \delta_{i-1,j}) \qquad a_i = P_i\epsilon + \frac{1}{\epsilon}(2 - \delta_{iN})$$

$$A_N = \begin{pmatrix} a_1 & -\epsilon^{-1} & 0 & \ldots & 0 \\ -\epsilon^{-1} & a_2 & -\epsilon^{-1} & 0 & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & -\epsilon^{-1} & a_{N-1} & -\epsilon^{-1} \\ 0 & 0 & 0 & -\epsilon^{-1} & a_N \end{pmatrix}$$

Note how we "split in half" the green term, making A_N a symmetric matrix.

We can now rewrite $I_4^{(N)}$ as:

$$I_4^{(N)} = \int_{\mathbb{R}^N} \left(\prod_{i=1}^N \frac{\mathrm{d}x_i}{\sqrt{\pi\epsilon}}\right) e^{-\boldsymbol{x}^T A_N \boldsymbol{x}} \qquad \boldsymbol{x}^T = (x_1, \ldots, x_N)$$

This is the integral of a multivariate gaussian, and evaluates to:

$$I_4^{(N)} = \frac{1}{\epsilon^{N/2}(\det A_N)^{1/2}} = \frac{1}{(\det(\epsilon A_N))^{1/2}}$$

as for a $N\times N$ matrix we have $\det(\epsilon A_N) = \epsilon^N \det A_N$. This has the advantage

(Scaling the matrix to remove denominators)

[2] ∧The same result can be proved without this assumption, but with a much more heavy notation.

of removing all denominators in A_N.

To compute this determinant we use a method suggested by Gelfand and Yaglom (1960). We start by denoting with $D_k^{(N)}$ the determinant of the matrix obtained by removing the first $k-1$ rows and columns from ϵA_N:

*3. **Gelfand-Yoglom method***

$$D_k^{(N)} \equiv \begin{vmatrix} \epsilon a_k & -1 & 0 & \dots & 0 \\ -1 & \epsilon a_{k+1} & -1 & 0 & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & a_{N-1} & -1 \\ 0 & \dots & 0 & -1 & \epsilon a_N \end{vmatrix}$$

So that $D_1^{(N)} = \det \epsilon A_N$ is the determinant we want to compute (because here we remove $1-1=0$ rows).

Expanding $D_k^{(N)}$ from the first row we get:

$$D_k^{(N)} = \epsilon a_k D_{k+1} - (-1) \begin{vmatrix} -1 & -1 & 0 & \dots & 0 \\ 0 & \epsilon a_{k+2} & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & a_{N-1} & -1 \\ 0 & 0 & 0 & -1 & \epsilon a_N \end{vmatrix} =$$

$$\underset{(a)}{=} \epsilon a_k D_{k+1}^{(N)} + (-1) D_{k+2}^{(N)} = \epsilon \left(\epsilon P_k + \frac{2}{\epsilon} \right) D_{k+1}^{(N)} - D_{k+2}^{(N)} =$$

$$= (\epsilon^2 P_k + 2) D_{k+1}^{(N)} - D_{k+2}^{(N)}$$

where in (a) we expanded the last determinant following the first column. Rearranging:

$$\frac{D_k^{(N)} - 2D_{k+1}^{(N)} + D_{k+2}^{(N)}}{\epsilon^2} = P_k D_{k+1}^{(N)} \tag{2.26}$$

We introduce now the variable $\tau = (k-1)t/N$, representing the *fraction* of removed rows/columns in each determinant, rescaled to the final time t. Performing a continuum limit $N \to \infty$ we can then map $D_k^{(N)} \xrightarrow[N\to\infty]{} D(s)$. Then, the relation (2.26) becomes a *differential equation*:

A determinant as a differential equation

$$\frac{\mathrm{d}^2 D(\tau)}{\mathrm{d}\tau^2} = P(\tau) D(\tau) \tag{2.27}$$

In fact, note that the first term of (2.26) is a second derivative in the *finite difference* approximation. This can be shown by Taylor expanding a generic function $f(x)$ to get the points immediately before and after:

$$f(x + \Delta x) = f(x) + f'(x)\Delta x + \frac{1}{2} f''(x)(\Delta x)^2 + O((\Delta x)^3)$$

$$f(x-\Delta x) = f(x) - f'(x)\Delta x + \frac{1}{2}f''(x)(\Delta x)^2 + O((\Delta x)^3)$$

Summing side by side, and denoting $f(x) \equiv f_i$, $f(x-\Delta x) \equiv f_{i-1}$ and $f(x+\Delta x) = f_{i+1}$:

$$f_{i+1} + f_{i-1} = 2f_i + f_i''(\Delta x)^2 + O((\Delta x)^3)$$

Rearranging, shifting $i \to i+1$ and ignoring the higher order terms leads to:

$$\frac{f_{i+2} - 2f_{i+1} + f_i}{(\Delta x)^2} = f''_{i+1}$$

Analogously, this can be seen by computing the second derivative as *the derivative of the first derivative* in terms of incremental ratios:

$$\begin{aligned} f''(x) &= \lim_{\Delta x \to 0} \frac{1}{\Delta x}\left(\frac{f(x+\Delta x) - f(x)}{\Delta x} - \frac{f(x) - f(x-\Delta x)}{\Delta x}\right) = \\ &= \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - 2f(x) + f(x-\Delta x)}{(\Delta x)^2} \end{aligned}$$

Returning to the problem, we note that the determinant of the full matrix, in the continuum limit, is given by:

$$\det(\epsilon A_N) = D_1^{(N)} \xrightarrow[N\to\infty]{} D(0)$$

(as $\tau = (k-1)t/N\big|_{k=1} \equiv 0$). So, we just need to solve (2.27) and evaluate it at $\tau = 0$.
To do this, we first need *two* boundary conditions, as (2.27) is a second order differential equation.
Noting that $D_N^{(N)}$ is just the last diagonal entry, we have:

$$D_N^{(N)} = \epsilon a_N = \epsilon^2 p_N + 1 \xrightarrow[\substack{N\to\infty \\ \epsilon\to 0}]{} 1$$

As $\tau = (k-1)t/N\big|_{k=N} = t$ for $N \to \infty$, this means that:

$$D(t) = 1$$

For the second boundary condition, we search a relation for the first derivative at $\tau = t$:

$$\left.\frac{\mathrm{d}D(\tau)}{\mathrm{d}\tau}\right|_{\tau=t} = \lim_{N\to\infty} \frac{D_N^{(N)} - D_{N-1}^{(N)}}{\epsilon}$$

$D_{N-1}^{(N)}$ can be computed directly:

$$D_{N-1}^{(N)} = \begin{vmatrix} P_{N-1}\epsilon^2 + 2 & -1 \\ -1 & P_N\epsilon^2 + 1 \end{vmatrix} = P_{N-1}P_N\epsilon^4 + \epsilon^2(P_{N-1} + 2P_N) + 1$$

leading to:

$$\left.\frac{\mathrm{d}D(\tau)}{\mathrm{d}\tau}\right|_{\tau=t} = \lim_{\epsilon\to 0} \frac{\epsilon^2 P_N + 1 - P_{N-1}P_N\epsilon^4 - \epsilon^2(P_{N-1} + 2P_N) - 1}{\epsilon} = 0$$

Summarizing, we found that:

$$I_4 \equiv \langle \exp\left(-\int_0^t \mathrm{d}\tau\, P(\tau)x^2(\tau)\right)\rangle_w = \frac{1}{\sqrt{D(0)}}$$

where $D(\tau)$ is the solution of the differential equation:

$$\frac{\mathrm{d}^2 D(\tau)}{\mathrm{d}\tau^2} = P(\tau)D(\tau)$$

with the following boundary conditions:

$$\begin{cases} D(t) = 1 \\ \dot{D}(t) = \left.\frac{\mathrm{d}D(\tau)}{\mathrm{d}\tau}\right|_{\tau=t} = 0 \end{cases}$$

Example 4 ($\boldsymbol{P(\tau) = k^2}$, free end-point):

Let's compute I_4 with the choice of $P(\tau) = k^2$. The differential equation becomes:

$$\frac{\mathrm{d}^2 D(\tau)}{\mathrm{d}\tau^2} = k^2 D(\tau)$$

which is that of a *harmonic repulsor*. The solution is a linear combination of exponentials:

$$D(\tau) = Ae^{k\tau} + Be^{-k\tau} \tag{2.28}$$

Differentiating:

$$\dot{D}(\tau) = k(Ae^{k\tau} - Be^{-k\tau})$$

We can now impose the boundary conditions:

$$\begin{cases} D(t) \overset{!}{=} 1 = Ae^{kt} + Be^{-kt} & (a) \\ \dot{D}(t) \overset{!}{=} 0 = \cancel{k}(Ae^{kt} - Be^{-kt}) & (b) \end{cases}$$

leading to:

$$\begin{aligned} (a) + (b)\colon 2Ae^{kt} &= 1 \Rightarrow A = \frac{1}{2}e^{-kt} \\ (a) - (b)\colon 2Be^{-kt} &= 1 \Rightarrow B = \frac{1}{2}e^{kt} \end{aligned}$$

So the solution is:

$$D(\tau) = \frac{1}{2}\left[e^{k(t-\tau)} + e^{-k(t-\tau)}\right] = \cosh(k(t-\tau)) \tag{2.29}$$

from which:

$$I_4 = \overset{(N)}{\lim_{N\to\infty}} = \frac{1}{\sqrt{D(0)}} = \frac{1}{\sqrt{\cosh(kt)}}$$

Fixed end-point

We consider now a small variation of I_4, where we integrate instead on paths with a fixed *end-point* $x(t) \equiv x_t$:

$$\hat{I}_4 = \langle \exp\left(-\int_0^t P(\tau)x^2(\tau)\,\mathrm{d}\tau\right)\delta(x - x(t))\rangle_w = \int_{\mathcal{C}\{0,0;x_t,t\}} \exp\left(-\int_0^t P(\tau)x^2(\tau)\,\mathrm{d}\tau\right)$$

First, we rewrite the δ in terms of a Fourier transform:

$$I_4' = \int_{-\infty}^{+\infty} \frac{\mathrm{d}\alpha}{2\pi} e^{i\alpha x} \langle \exp\left(-\int_0^t P(\tau)x^2(\tau)\,\mathrm{d}\tau\right) e^{-i\alpha x(t)} \rangle_w$$

1. Fourier transform to remove the end-point δ

Then we discretize the path as before, with $0 = t_0 < t_1 < \cdots < t_N = t$ uniformly distributed ($\Delta t_i = t_i - t_{i-1} \equiv \epsilon = t/N$):

2. Discretized path integral

$$\hat{I}_4 = \lim_{N\to\infty} \hat{I}_4^{(N)}$$

$$\hat{I}_4^{(N)} = \int_{\mathbb{R}} \frac{\mathrm{d}\alpha}{2\pi} e^{i\alpha x} \int_{\mathbb{R}^N} \left(\prod_{i=1}^N \frac{\mathrm{d}x_i}{\sqrt{\pi\epsilon}}\right) \exp\left(-\sum_{i=1}^N P_i x_i^2 \epsilon - \sum_{i=1}^N \frac{(x_i - x_{i-1})^2}{\epsilon} - i\alpha x_N\right)$$

where the red terms are the only differences from (2.25). We can rewrite the quadratic form with the matrix A_N as before:

$$\hat{I}_4^{(N)} = \int_{\mathbb{R}} \frac{\mathrm{d}\alpha}{2\pi} e^{i\alpha x} \int_{\mathbb{R}^N} \left(\prod_{i=1}^N \frac{\mathrm{d}x_i}{\sqrt{\pi\epsilon}}\right) \exp\left(-\boldsymbol{x}^T A_N \boldsymbol{x} - i\alpha x_N\right)$$

3. Exponential argument in matrix form

Also, we can express $i\alpha x_N$ as a scalar product between $\boldsymbol{x} = (x_1, \dots, x_N)^T$ and a certain vector $\boldsymbol{h} \in \mathbb{R}^N$ with components h_l given by:

$$i\alpha x_N = \boldsymbol{h}^T \boldsymbol{x} \qquad h_l = \delta_{lN}(-i\alpha)$$

So that we can now use the gaussian integral:

$$\int_{\mathbb{R}^N} \mathrm{d}^N\boldsymbol{x} \exp\left(-\frac{1}{2}\boldsymbol{x}^T A\boldsymbol{x} + \boldsymbol{b}\cdot\boldsymbol{x}\right) = \exp\left(\frac{1}{2}\boldsymbol{b}\cdot A^{-1}\boldsymbol{b}\right)(2\pi)^{N/2}(\det A)^{-1/2}$$

4. First N gaussian integrals

with $A = 2A_N$ and $\boldsymbol{b} = \boldsymbol{h}$:

$$\begin{aligned} I' &\equiv \frac{1}{\sqrt{(\pi\epsilon)^N}} \int_{\mathbb{R}^N} \mathrm{d}^N\boldsymbol{x} \exp\left(-\boldsymbol{x}^T A_N \boldsymbol{x} + \boldsymbol{h}^T\boldsymbol{x}\right) = \\ &= \frac{1}{\sqrt{(\pi\epsilon)^N}} \exp\left(\frac{1}{4}\boldsymbol{h}^T A_N^{-1}\boldsymbol{h}\right)(\cancel{2}\pi)^{N/2}(2^{\cancel{N}}\det A_N)^{-1/2} = \\ &= \frac{1}{\sqrt{(\cancel{\pi}\epsilon)^N}} \sqrt{\frac{\cancel{\pi^N}}{\det A_N}} \exp\left(\frac{1}{4}(-i\alpha)^2 (A_N^{-1})_{NN}\right) = \underbrace{\sqrt{\frac{1}{\epsilon^N \det A_N}}}_{I_0} \exp\left(-\frac{1}{4}\alpha^2 (A_N^{-1})_{NN}\right) \end{aligned}$$

where $(A_N^{-1})_{NN}$ is the last diagonal element of the inverse matrix of A_N. Substituting back:

$$\hat{I}_4^{(N)} = I_0 \int_{\mathbb{R}} \frac{\mathrm{d}\alpha}{2\pi} \exp\left(i\alpha x - \frac{1}{4}\alpha^2 (A_N^{-1})_{NN}\right)$$

which is again a gaussian integral, and following formula (2.22) with $a = (A_N^{-1})_{NN}/4$ and $b = -x$ leads to:

5. Last gaussian integral

$$\hat{I}_4^{(N)} = \frac{I_0}{\cancel{2}\pi}\sqrt{\frac{\cancel{4}\pi}{(A_N^{-1})_{NN}}}\exp\left(-\frac{x^2}{(A_N^{-1})_{NN}}\right) = \frac{I_0}{\sqrt{\pi(A_N^{-1})_{NN}}}\exp\left(-\frac{x^2}{(A_N^{-1})_{NN}}\right) \tag{2.30}$$

All that's left is to compute $(A_N^{-1})_{NN}$ and take the continuum limit. Recall from linear algebra that:

$$A_{ij}^{-1} = \frac{1}{\det A}C_{ji}$$

where C_{ij} are the *cofactors* of A, i.e. the determinants of the matrices obtained from A by removing the i-th row and j-th column. In our case:

$$(A_N^{-1})_{NN} = \frac{C_{NN}}{\det A_N}$$

6. Gelfand-Yoglom method (bottom-top variant)

Before, we obtained $\det A_N$ by means of $D_k^{(N)}$, i.e. the determinants of the matrices obtained by removing the first $k-1$ rows and columns, so that $D_1^{(N)} = \epsilon^N \det A_N$. This leads to:

$$(A_N^{-1})_{NN} = \frac{\epsilon^N}{D_1^{(N)}}C_{NN}$$

For C_{NN} we have to compute the determinant of the $(N-1)\times(N-1)$ matrix $A_*^{(N-1)}$, obtained by removing the last row and column from A_N. Note that $A_*^{(N-1)} \neq A^{(N-1)}$, as they differ for the *last diagonal element* which is:

$$(A_*^{(N-1)})_{N-1,N-1} = P_{N-1}\epsilon + \frac{2}{\epsilon} \neq (A_{N-1,N-1}^{(N-1)}) = P_{N-1}\epsilon + \frac{1}{\epsilon} \tag{2.31}$$

We proceed in a similar manner, defining $\hat{D}_k^{(N-1)}$ to be the determinant of the matrix obtained by removing the first $k-1$ rows and columns from $\epsilon A_*^{(N-1)}$ (again, we multiply by ϵ to remove denominators):

$$\hat{D}_k^{(N-1)} = \begin{vmatrix} \epsilon a_k & -1 & 0 & \dots & 0 \\ -1 & \epsilon a_{k-1} & -1 & 0 & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & -1 & \epsilon a_{N-2} & -1 \\ 0 & \dots & \dots & -1 & \epsilon a_{N-1} \end{vmatrix}$$

So $\hat{D}_1^{(N-1)} = \epsilon^{N-1}\det A_*^{(N-1)} = \epsilon^{N-1}C_{NN}$ leading to:

$$(A_N^{-1})_{NN} = \frac{\epsilon^N}{D_1^{(N)}}\frac{1}{\epsilon^{N-1}}\hat{D}_1^{(N-1)} = \epsilon\frac{\hat{D}_1^{(N-1)}}{D_1^{(N)}}$$

For simplicity, it is convenient to define $\tilde{D}_1^{(N-1)} \equiv \epsilon \hat{D}_1^{(N-1)}$, so that:

$$(A_N^{-1})_{NN} = \frac{\tilde{D}_1^{(N-1)}}{D_1^{(N)}} \tag{2.32}$$

Repeating the steps for the continuum limit, we get the same differential equation for $\tilde{D}(\tau)$:

Determinant as a differential equation

$$\partial_\tau^2 \tilde{D}(\tau) = P(\tau)\tilde{D}(\tau)$$

However, due to (2.31), the boundary conditions are now different:

$$\tilde{D}_{N-1}^{(N-1)} = \epsilon(\epsilon^2 P_{N-1} + 2) = P_{N-1}\epsilon^3 + 2\epsilon \xrightarrow[\epsilon\to 0]{} 0 = \tilde{D}(t)$$

$$\tilde{D}_{N-2}^{(N-1)} = \epsilon \begin{vmatrix} P_{N-2}\epsilon^2 + 2 & -1 \\ -1 & P_{N-1}\epsilon^2 + 2 \end{vmatrix} = \epsilon(P_{N-1}P_{N-2}\epsilon^4 + 2(p_{N-1} + P_{N-2})\epsilon^2 + 3)$$

$$\frac{\tilde{D}_{N-1}^{(N-1)} - \tilde{D}_{N-2}^{(N-1)}}{\epsilon} = -1 + O(\epsilon^2) \xrightarrow[\epsilon\to 0]{} -1 = \left.\frac{\mathrm{d}\tilde{D}(\tau)}{\mathrm{d}\tau}\right|_{\tau=t}$$

Then, substituting (2.32) in (2.30) we get:

$$\hat{I}_4^{(N)} = \frac{I_0}{\sqrt{\pi (A_N^{-1})_{NN}}} \exp\left(-x^2 \frac{D_1^{(N)}}{\tilde{D}_1^{(N-1)}}\right) \qquad I_0 = \frac{1}{\sqrt{\epsilon^N \det A_N}} = \frac{1}{\sqrt{D_1^{(N)}}}$$

$$I_4 = \lim_{N\to\infty} \hat{I}_4^{(N)} = \frac{1}{\sqrt{\pi \tilde{D}(0)}} \exp\left(-x^2 \frac{D(0)}{\tilde{D}(0)}\right) \tag{2.33}$$

Where $D(\tau)$ and $\tilde{D}(\tau)$ are solutions of the following differential equations with the following boundary conditions:

$$\tilde{D}''(\tau) = P(\tau)\tilde{D}(\tau) \qquad \begin{cases} \tilde{D}(t) = 0 \\ \left.\frac{\mathrm{d}\tilde{D}(\tau)}{\mathrm{d}\tau}\right|_{\tau=t} = -1 \end{cases}$$

$$D''(\tau) = P(\tau)D(\tau) \qquad \begin{cases} D(t) = 1 \\ \left.\frac{\mathrm{d}D(\tau)}{\mathrm{d}\tau}\right|_{\tau=t} = 0 \end{cases}$$

Example 5 ($\boldsymbol{P(\tau) = k^2}$ with fixed end-point):

Let $P(\tau) = k^2$, with $k \in \mathbb{R}$ constant. We already solved the equation for $D(\tau)$ with the right boundary conditions in (2.29):

$$D(\tau) = \cosh(k(t-\tau))$$

For $\tilde{D}(\tau)$ we start from the general integral (2.28) and impose the appropriate boundary conditions:

$$\begin{cases} \tilde{D}(t) = \tilde{A}e^{kt} + \tilde{B}e^{-kt} = 0 & (a) \\ \left.\frac{\mathrm{d}\tilde{D}(\tau)}{\mathrm{d}\tau}\right|_{\tau=t} = k(\tilde{A}e^{kt} - \tilde{B}e^{-kt}) = -1 & (b) \end{cases}$$

so that:

$$k(a)+(b)\colon 2\tilde{A}ke^{kt} = -1 \Rightarrow \tilde{A} = -\frac{1}{2k}e^{-kt}$$
$$k(a)-(b)\colon 2Bke^{-kt} = 1 \Rightarrow \tilde{B} = \frac{1}{2k}e^{kt}$$

leading to the solution:

$$\tilde{D}(\tau) = \frac{1}{2k}(e^{k(t-\tau)} - e^{-k(t-\tau)}) = \frac{1}{k}\sinh(k(t-\tau))$$

Finally, using the result we found in (2.33):

$$\begin{aligned}
&\langle \exp\left(-k^2\int_0^t x^2(\tau)\,\mathrm{d}\tau\,\delta(x-x(t))\right)\rangle_w = \\
&= \int_{\mathcal{C}\{0,0;x_t,t\}} \exp\left(-k^2\int_0^t x^2(\tau)\,\mathrm{d}\tau\right)\mathrm{d}_W x(\tau) = \\
&= \sqrt{\frac{k}{\pi\sinh(kt)}}\exp\Big(-kx_t^2\coth(kt)\Big)
\end{aligned}$$

2.3 Properties of Brownian Paths

The Wiener measure allows us to compute the probabilities of *paths* produced by the diffusion process, and also highlight some of their defining characteristics. We now show that all Brownian paths with non-zero Wiener measure (i.e. paths that "can happen") are **everywhere continuous**, but **nowhere differentiable**.

(Lesson 6 of 04/11/19) Compiled: October 13, 2020

2.3.1 Continuity

Consider a particle starting in $x=0$ at $t=0$, and traversing N points $\{x_i\}_{i=1,\dots,N}$ such that all increments $\Delta x_i = x_i - x_{i-1}$ are *independent* and described by a *gaussian pdf*. The density function for such a trajectory $\{x_i\}$ is the usual product of transition probabilities:

$$\mathrm{d}\mathbb{P}_{t_1,\dots,t_N}(x_1,\dots,x_N) = \left(\prod_{i=1}^N \frac{\mathrm{d}x_i}{\sqrt{4\pi\Delta t_i D}}\right)\exp\left(-\sum_{i=1}^N\frac{(\Delta x_i)^2}{4D\Delta t_i}\right) \quad \begin{smallmatrix}\Delta t_i = t_i - t_{i-1}\\ \Delta x_i = x_i - x_{i-1}\end{smallmatrix} \tag{2.34}$$

We now show that, taking the continuum limit $\max_i \Delta t_i \to 0$ leads to paths $\{x(\tau)\}$ that are *almost surely continuous*. In other words, for any interval $T\subseteq\mathbb{R}$, the subset $N\subset\mathbb{R}^T$ of functions that are discontinuous has 0 Wiener measure.
Mathematically, we want to show that, as $\Delta t_i\to 0$, the probability that Δx_i is close to 0 approaches certainty:

$$\lim_{\Delta t_i\to 0}\mathbb{P}(|\Delta x_i|<\epsilon) = 1 \quad \forall\epsilon>0$$

This is just the probability that, during time Δt_i, the particle makes a jump of size lower than ϵ:

$$\begin{aligned}\mathbb{P}(|\Delta x_i| < \epsilon) &= \mathbb{P}(x_{i-1} - \epsilon < x_i < x_{i-1} + \epsilon | x(t_{i-1}) = x_i) = \\ &= \int_{x_{i-1}-\epsilon}^{x_{i-1}+\epsilon} \frac{\mathrm{d}x_i}{\sqrt{4\pi D \Delta t_i}} \exp\left(-\frac{(x_i - x_{i-1})^2}{4D\Delta t_i}\right) = \\ &\underset{(a)}{=} \int_{-\epsilon}^{+\epsilon} \frac{\mathrm{d}\Delta x_i}{\sqrt{4\pi D\Delta t_i}} \exp\left(-\frac{(\Delta x_i)^2}{4D\Delta t_i}\right)\end{aligned}$$

where in (a) we translated the variable of integration $\Delta x_i = x_i - x_{i-1}$. With another change of variables:

$$\frac{(\Delta x_i)^2}{\Delta t_i} = z^2 \Rightarrow z = \frac{\Delta x_i}{\sqrt{\Delta t_i}} \Rightarrow \mathrm{d}\Delta x_i = \mathrm{d}z \sqrt{\Delta t_i}$$

we get:

$$\mathbb{P}(|\Delta x_i| < \epsilon) = \int_{|z| < \epsilon/\sqrt{\Delta t_i}} \frac{\mathrm{d}z \cancel{\sqrt{\Delta t_i}}}{\sqrt{4\pi D \cancel{\Delta t_i}}} \exp\left(-\frac{z^2}{4D}\right)$$

And taking the continuum limit leads to:

$$\lim_{\Delta t_i \to 0} \mathbb{P}(|\Delta x_i| < \epsilon) = \int_{-\infty}^{+\infty} \frac{\mathrm{d}z}{\sqrt{4\pi D}} \exp\left(-\frac{z^2}{4D}\right) = 1$$

2.3.2 Differentiability

With a very similar calculation (here omitted) we can also show that:

$$\lim_{\Delta t_i \downarrow 0} \left(\left| \frac{\Delta x_i}{\Delta t_i} \right| > k \right) = 1 \quad \forall k > 0$$

meaning that Brownian paths are *almost surely everywhere non-differentiable.*

Nonetheless, it is sometimes useful to consider "formal derivatives" of a Brownian path, that acquire a definite meaning only when considering a *finite discretization.* For example, we can start from (2.34) and rewrite it as:

$$d\mathbb{P}_{t_1,\dots,t_N}(x_1,\dots,x_N) = \left(\prod_{i=1}^N \frac{\mathrm{d}x_i}{\sqrt{4\pi\Delta t_i}}\right) \exp\left(-\frac{1}{4D}\sum_{i=1}^N {\color{red}\Delta t_i} \left(\frac{\Delta x_i}{{\color{red}\Delta t_i}}\right)^2\right)$$

Then, in the continuum limit $\Delta t_i \to 0$, the sum in the exponential argument becomes a Riemann integral:

$$\sum_{i=1}^N \Delta t_i \left(\frac{\Delta x_i}{\Delta t_i}\right)^2 \xrightarrow[\Delta t \to 0]{} \int_0^t \mathrm{d}\tau \underbrace{\left(\frac{\mathrm{d}x_i}{\mathrm{d}\tau}\right)^2}_{\dot{x}^2(\tau)} \qquad t = t_N$$

where $t = t_N$. Substituting it back leads to:

$$\mathrm{d}x_w(\tau) = \prod_{\tau=0^+}^{t} \frac{\mathrm{d}x(\tau)}{\sqrt{4\pi D\,\mathrm{d}\tau}} \exp\left(-\frac{1}{4D}\int_0^t \dot{x}^2(\tau)\,\mathrm{d}\tau\right)$$

This expression has no rigorous meaning in this form ($\dot{x}(\tau)$ *does not* exists!) but can be *formally manipulated* into other expressions that *have* a definite meaning, thus proving useful for the discussion.

Chapter 3

Diffusion with Forces

We want now to generalize the framework we previously obtained to the case of a diffusing particle subject to *external forces*, e.g. a drop of ink diffusing through a water medium in the presence of gravity.
To do this, we first return to the beginning, deduce a Master Equation for a more general *evolution*, and then choose the right probability distribution reproducing the behaviour in presence of forces.

3.1 Fokker-Planck equation

So, let's start by considering a particle moving on a *uniform* one-dimensional lattice ($x_i = i \cdot l$, $t_n = n \cdot \epsilon$), and satisfying the Markovian property, meaning that the probability $W_i(t_{n+1})$ of being at the position labelled by i at the *next time-step* t_{n+1} depends only on the current state t_n, that is on the current probabilities $W_j(t_n)\,\forall j$ and on the current transition probabilities $W_{ij}(t_n)$ from j to i:

$$W_i(t_{n+1}) = \sum_{j=-\infty}^{+\infty} W_{ij}(t_n) W_j(t_n) \tag{3.1}$$

Previously, we assumed that:

$$W_{ij}(t_n) = \delta_{j,i-1} P_+ + \delta_{j,i+1} P_-$$

Which means that the particle only jumps from adjacent positions, one step at a time, and cannot remain at the same place. This Master Equation leads, in $d = 3$ and in the continuum limit, to the usual Diffusion Equation:

$$\frac{\partial}{\partial t} W(\boldsymbol{x}, t|\boldsymbol{x_0}, t_0) = \nabla^2 W(\boldsymbol{x}, t|\boldsymbol{x_0}, t_0)$$

We now consider a more general case, where we drop the discretization of the space domain, allowing *jumps* of *any size* in $\mathbb{R}$. Then (3.1) becomes:

Generalized Master Equation (Jumps of any size)

$$W(x, t_{n+1})\,\mathrm{d}x = \int_{-\infty}^{+\infty} \mathrm{d}z\, W(x, t_{n+1}|x - z, t_n) W(x - z, t_n) \tag{3.2}$$

The integrand is the probability of the particle being in $[x-z, x-z+\mathrm{d}x]$ at time t_n and making a jump of size z to reach $[x, x+\mathrm{d}x]$ at time t_{n+1}. By summing over *all possible jump sizes* we compute the total probability of the particle being *near* the arrival position.
If we require *jumps* to be **independent** of each other[1], as it is physically evident by the problem's symmetry, then the *jump* probabilities $W(x, t_{n+1}|x-z, t_n)$ depend only on the *jump size* z.
Assuming a *isolate system*, as the particle cannot *escape*, **probability is conserved**:

Conservation of probability

$$\begin{aligned}
\int_{\mathbb{R}} \mathrm{d}x\, W(x, t_{n+1}) \overset{!}{=} \int_{\mathbb{R}} \mathrm{d}y\, W(y, t_n)& \\
&\underset{(3.2)}{=} \int_{\mathbb{R}} \mathrm{d}z \int_{\mathbb{R}} \mathrm{d}x W(x, t_{n+1}|x-z, t_n)\, W(x-z, t_n) = \\
&\underset{(a)}{=} \int_{\mathbb{R}} \mathrm{d}z \int_{\mathbb{R}} \mathrm{d}y\, W(y+z, t_{n+1}|y, t_n) W(y, t_n) = \\
&\underset{(b)}{=} \left(\int_{\mathbb{R}} \mathrm{d}z\, W(\bar{y}+z, t_{n+1}|\bar{y}, t_n)\right) \left(\int_{\mathbb{R}} \mathrm{d}y\, W(y, t_n)\right) \qquad \forall \bar{y} \in \mathbb{R}
\end{aligned}$$

where in (a) we changed variables $x \mapsto y = x - z$, with $\mathrm{d}y = \mathrm{d}x$, and in (b) we used the *independent increments* property ($\bar{y}$ is a arbitrary constant). Comparing the first and last lines leads to:

$$\int_{\mathbb{R}} \mathrm{d}z\, W(y+z, t_{n+1}|y, t_n) = 1$$

Intuitively, if the particle *cannot disappear*, it *must make a jump*.
Here on, for notation simplicity, we denote:

$$W(y+z, t_{n+1}|y, t_n) \equiv W(+z|y, t_n)$$

Starting from (3.2) and taking the continuum limit in time we can write a *more general* diffusion equation. We start by constructing the difference quotient:

Generalized diffusion equation

$$\begin{aligned}
W(x, t_{n+1}) - W(x, t_n) &= \int_{\mathbb{R}} \mathrm{d}z\, W(+z|x-z, t_n) W(x-z, t_n) - W(x, t_n) = \\
&= \int_{\mathbb{R}} \mathrm{d}z\, W(+z|x-z, t_n) W(x-z, t_n) - \underbrace{\int \mathrm{d}z\, W(+z|x, t_n)}_{=1} W(x, t_n) = \\
&= \int_{\mathbb{R}} \mathrm{d}z \left[\underbrace{W(+z|x-z, t_n) W(x-z, t_n)}_{F_z(x-z)} - \underbrace{W(+z|x, t_n) W(x, t_n)}_{F_z(x)}\right] = \\
&= \int_{\mathbb{R}} \mathrm{d}z\, [F_z(x-z) - F_z(x)] = \\
&\underset{(a)}{=} \int_{\mathbb{R}} \mathrm{d}z \left[\cancel{F_z(x)} - z\frac{\partial}{\partial x} F_z(x) + \frac{z^2}{2}\frac{\partial^2}{\partial x^2}[F_z(x)] + \cdots - \cancel{F_z(x)}\right] = \\
&= -\int_{\mathbb{R}} \mathrm{d}z\, z\frac{\partial}{\partial x}[F_z(x)] + \frac{1}{2}\int_{\mathbb{R}} \mathrm{d}z\, z^2 \frac{\partial^2}{\partial x^2}[F_z(x)] + \cdots =
\end{aligned}$$

[1] ∧This is a stronger requirement than the Markovian property. In fact, *independent increments* imply a *Markov process*, but the converse is not true. See `http://statweb.stanford.edu/~adembo/math-136/Markov_note.pdf`

$$\underset{(b)}{=} -\frac{\partial}{\partial x}\left[\underbrace{\left(\int_{\mathbb{R}} \mathrm{d}z\, z\, W(+z|x,t_n)\right)}_{\mu_1(x,t_n)} W(x,t_n)\right] +$$
$$+\frac{1}{2}\frac{\partial^2}{\partial x^2}\left[\underbrace{\left(\int_{\mathbb{R}} \mathrm{d}z\, z\, W(+z|x,t_n)\right)}_{\mu_2(x,t_n)} W(x,t_n)\right] + \dots$$

where $F_z(x)$ is the probability of a *jump* of size z from the position x. In (a) we expanded F_z about x, and in (b) we exchanged the order of integrals and derivatives. Then we define the k-th moment of the *jump* pdf as follows:

$$\mu_k(x,t) = \int_{\mathbb{R}} \mathrm{d}z\, z^k\, W(+z|x,t)$$

This allows us to rewrite the above difference in a more compact form:

$$W(x,t_{n+1}) - W(x,t_n) = \sum_{k=1}^{+\infty} \frac{(-1)^k}{k!}\frac{\partial^k}{\partial x^k}\left(\mu_k(x,t_n)W(x,t_n)\right)$$

Physically, as probability is conserved, by the continuity equation, the *change* in probability density equals the divergence of a **flux**, which is just the x derivative in this one-dimensional case. So, if we *extract* a derivative, we can write the flux explicitly:

Flux

$$\begin{aligned} &= \frac{\partial}{\partial x}\left(\sum_{k=1}^{+\infty} \frac{(-1)^k}{k!}\frac{\partial^{k-1}}{\partial x^{k-1}}(\mu_k(x,t_n)W(x,t_n))\right) \\ &\equiv -\frac{\partial}{\partial x} J(x,t_n) \end{aligned}$$

where $J(x,t_n)$ is the *outward flux* at x, meaning that if $J > 0$, then $W(x,t_{n+1}) < W(x,t_n)$ (the particle *escapes* from x to another place), and otherwise if $J < 0$ we have $W(x,t_{n+1}) > W(x,t_n)$ (the particle is *sucked in* x).
If we integrate both sides over x and apply the probability conservation we get the boundary conditions for the flux:

Boundary conditions for the flux

$$\begin{aligned} \int_{\mathbb{R}} (W(x,t_{n+1}) - W(x,t_n))\,\mathrm{d}x &= \int_{\mathbb{R}} \mathrm{d}x \left(-\frac{\partial}{\partial x} J(x,t_n)\right) \\ 1 - 1 = -J(x,t_n)\Big|_{-\infty}^{+\infty} &= J(-\infty,t_n) - J(+\infty,t_n) \end{aligned}$$

This means that, in a *isolate system*, the *flux* at $\pm\infty$ must be the same.
Finally, normalizing by the time interval we get the complete difference quotient, which will become a time derivative in the continuum limit.

$$\frac{W(x,t_{n+1}) - W(x,t_n)}{t_{n+1}-t_n} = \frac{\partial}{\partial x}\left\{\sum_{k=1}^{\infty} \frac{(-1)^k}{k!}\frac{\partial^{k-1}}{\partial x^{k-1}}\frac{\mu_k(x,t_n)W(x,t_n)}{t_{n+1}-t_n}\right\} \tag{3.3}$$

Letting $t_{n+1} - t_n = \epsilon$, in the limit $\epsilon \to 0$ the left side will be $\dot{W}(x,t)$.

All that's left is to find an explicit definition for the *jump pdf* $W(+z|x,t)$. Previously, we assumed a **gaussian** pdf for the displacements:

$$z \sim \frac{1}{\sqrt{4\pi D\epsilon}} \exp\left(-\frac{(\Delta x)^2}{4D\epsilon}\right)$$

With this choice, the first two moments become:

$$\mu_1 = 0 \qquad \mu_2 = 2D\epsilon$$

And the variance:

$$\mathrm{Var}(z) = \mu_2 - \mu_1^2 = 2D\epsilon \propto \epsilon$$

However, for a particle subject to a force we would expect to have a *preferred jump direction*, leading to a *constant velocity motion* in the direction of the force. So we require a different μ_1:

$$\langle z \rangle = \mu_1 = \int_{\mathbb{R}} z W(+z|x,t) \propto \epsilon f(x)$$

We still want to fix the variance to be proportional to ϵ, as it is expected in a diffusion process.

An appropriate choice for such a distribution is given by:

Jump distribution

$$W(+z|x,t) = \frac{1}{\sqrt{\epsilon \hat{D}(x,t)}} F\left(\frac{z - \epsilon f(x,t)}{\sqrt{\epsilon \hat{D}(x,t)}}\right) \tag{3.4}$$

with $F, \hat{D}\colon \mathbb{R} \to \mathbb{R}$ functions, satisfying certain conditions, and with a physical meaning that we will now see.

1. Correct normalization

First of all, we check the normalization:

$$1 \overset{!}{=} \int_{\mathbb{R}} \mathrm{d}z\, W(+z|x,t) = \frac{1}{\sqrt{\epsilon \hat{D}(x,t)}} \int_{\mathbb{R}} \mathrm{d}z\, F\left(\frac{z - \epsilon f(x,t)}{\sqrt{\epsilon \hat{D}(x,t)}}\right) \underset{(a)}{=} \int_{\mathbb{R}} \mathrm{d}y\, F(y)$$

where in (a) we changed variables:

$$y = \frac{z - \epsilon f(x,t)}{\sqrt{\epsilon \hat{D}(x,t)}} \qquad \mathrm{d}z = \sqrt{\epsilon \hat{D}(x,t)}\, \mathrm{d}y \tag{3.5}$$

2. First moment $\propto$ force

Then we compute the first moment:

$$\begin{aligned}
\langle z \rangle = \mu_1(x,t) &= \int_{\mathbb{R}} \mathrm{d}z\, z\, F\left(\frac{z - \epsilon f(x,t)}{\sqrt{\epsilon \hat{D}(x,t)}}\right) \frac{1}{\sqrt{\epsilon \hat{D}(x,t)}} = \\
&\underset{(3.5)}{=} \int_{\mathbb{R}} \mathrm{d}y \left(\epsilon f(x,t) + y\sqrt{\epsilon \hat{D}(x,t)}\right) F(y) = \\
&= \epsilon f(x,t) \underbrace{\int_{\mathbb{R}} F(y)\, \mathrm{d}y}_{=1} + \sqrt{\epsilon \hat{D}(x,t)} \int_{\mathbb{R}} y F(y) \overset{!}{=} \epsilon f(x,t)
\end{aligned}$$

So, in order to have the right normalization and the desired $\langle z \rangle$ we need:

$$\begin{cases} \int_{\mathbb{R}} \mathrm{d}y\, F(y) = 1 \\ \int_{\mathbb{R}} \mathrm{d}y\, y F(y) = 0 \end{cases}$$

Both conditions are satisfied, for example, by all even normalized functions.

For the second moment: *3. Variance $\propto$ time*

$$
\begin{aligned}
\mu_2(x,t) &= \frac{1}{\sqrt{\epsilon\hat{D}(x,t)}}\int_{\mathbb{R}} \mathrm{d}z\, z^2 F\left(\frac{z-\epsilon f(x,t)}{\sqrt{\epsilon\hat{D}(x,t)}}\right) = \\
&\underset{(3.5)}{=} \int_{\mathbb{R}} \mathrm{d}y\, (\epsilon f(x,t) + y\sqrt{\epsilon\hat{D}(x,t)})^2 F(y) = \\
&= \int_{\mathbb{R}} \mathrm{d}y\, F(y)[\epsilon^2 f^2 + y^2\hat{D}\epsilon + \cancel{2\epsilon\sqrt{\epsilon\hat{D}}fy}] = \\
&= \epsilon^2 f^2 + \hat{D}\epsilon\int_{\mathbb{R}} \mathrm{d}y\, y^2 F(y) = \epsilon^2 f^2 + \hat{D}\epsilon\langle y^2\rangle_{F(y)}
\end{aligned}
$$

And so the variance becomes:

$$
\mathrm{Var}(z) = \mu_2 - \mu_1^2 = \epsilon\hat{D}\langle y^2\rangle_{F(y)} \propto \epsilon
$$

which is proportional to ϵ as desired. For notational simplicity, we introduce a new function $D\colon \mathbb{R}\to\mathbb{R}$ such that:

$$
\mathrm{Var}(z) = \epsilon\hat{D}\langle y^2\rangle_{F(y)} \equiv 2D(x,t)\epsilon \Rightarrow \mu_2(x,t) = \epsilon^2 f^2 + 2D(x,t)
$$

We note that higher order moments are all of order $O(\epsilon^{3/2})$. For example, the third moment is: *4. Vanishing higher moments*

$$
\begin{aligned}
\mu_3(x,t) &= \frac{1}{\sqrt{\epsilon\hat{D}(x,t)}}\,\mathrm{d}z\int_{\mathbb{R}} z^3 F\left(\frac{z-\epsilon f(x,t)}{\sqrt{\epsilon\hat{D}(x,t)}}\right) = \\
&\underset{(3.5)}{=} \int_{\mathbb{R}} \mathrm{d}y\, (\epsilon f(x,t) + y\sqrt{\epsilon\hat{D}(x,t)})^3 F(y) = \\
&= \int_{\mathbb{R}} \mathrm{d}y\, \left(\epsilon^3 f^3 + y^3(\epsilon\hat{D})^{3/2} + \cancel{3\epsilon^2 f^2 y\sqrt{\epsilon\hat{D}}} + 3\epsilon^2 f\hat{D}y^2\right) F(y) = \\
&= \epsilon^3 f^3 + (\epsilon\hat{D})^{3/2} + 3\epsilon^2 f\hat{D}\langle y^2\rangle_{F(y)} = O(\epsilon^{3/2})
\end{aligned}
$$

Substituting back (3.4) in (3.3) we arrive to:

$$
\begin{aligned}
\frac{W(x,t_{n+1}) - W(x,t_n)}{\epsilon} &= -\frac{\partial}{\partial x}\left[W(x,t_n)\underbrace{\frac{\mu_1(x,t_n)}{\epsilon}}_{f(x,t)}\right] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[\underbrace{\frac{\mu_2(x,t_n)}{\epsilon}}_{\epsilon f^2 + 2D(x,t)} W(x,t_n)\right] + \\
&\quad + \underbrace{\frac{1}{3!}\frac{\partial^3}{\partial x^3}\left[W(x,t_n)\frac{\mu_3(x,t_n)}{\epsilon}\right] + \dots}_{O(\epsilon^{1/2})}
\end{aligned}
$$

Taking the limit $\epsilon \to 0$, we are left with:

Fokker-Planck equation

$$
\begin{aligned}
\frac{\partial W(x,t)}{\partial t} &= -\frac{\partial}{\partial x}[f(x,t)W(x,t)] + \frac{1}{\cancel{2}}\frac{\partial^2}{\partial x^2}[\cancel{2}D(x,t)W(x,t)] = \\
&= -\frac{\partial}{\partial x}\left[f(x,t)W(x,t) - \frac{\partial}{\partial x}\big(D(x,t)W(x,t)\big)\right]
\end{aligned}
$$

This is the **Fokker-Planck equation**, describing the diffusion process in the presence of a *force* $f(x,t)$, and a diffusion parameter $D(x,t)$.
Note that, in absence of forces $f(x,t) \equiv 0$ and with a constant diffusion $D(x,t) \equiv D$ we retrieve the usual *diffusion equation*:

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

3.2 Langevin equation

The Fokker-Planck equation involves *probability distributions*, meaning that it describes the behaviour of *ensembles of trajectories* at once. However, we can find an equivalent description by focusing on a *single path*.
We start with a Wiener process, that is a stochastic process with *independent* and *gaussian* increments and *continuous* paths. Considering a time discretization $\{t_i\}$, the evolution of a single trajectory is described by:

$$x(t_{i+1}) = x(t_i) + \Delta x(t_i) \tag{3.6}$$

where each increment $\Delta x(t_i)$ is sampled from a gaussian pdf:

$$\Delta x_i(t_i) \sim \frac{1}{\sqrt{4\pi D \Delta t_i}} \exp\left(-\frac{(\Delta x)^2}{4D\Delta t_i}\right)$$

To simplify notation, we change variables, so that:

$$\frac{\Delta B^2}{2} = \frac{\Delta x^2}{4D} \Rightarrow \Delta B = \frac{\Delta x}{\sqrt{2D}}$$

If $x \sim p(x)$, and $y = y(x) \sim g(y)$, then by the rule for a change of random variables we have:

$$g(y) = p(x(y)) \frac{\mathrm{d}x(y)}{\mathrm{d}y}$$

In this case:

$$\Delta B \sim \frac{1}{\sqrt{4\pi D \Delta t_i}} \exp\left(-\frac{(\Delta B)^2}{2\Delta t_i}\right) \underbrace{\frac{\mathrm{d}\Delta x}{\mathrm{d}\Delta B}}_{\sqrt{2D}} = \frac{1}{\sqrt{2\pi \Delta t_i}} \exp\left(-\frac{(\Delta B)^2}{2\Delta t_i}\right)$$

"Standard" Brownian path

Note that now $\langle \Delta B^2(t_i) \rangle = \Delta t_i$, leaving out the D - so, in a sense, it is the "standard" Brownian path, and any specific Brownian motion can be obtained by rescaling it.
Substituting in (3.6) and rearranging we get:

$$x(t_{i+1}) - x(t_i) = \sqrt{2D} \Delta B(t_i) \tag{3.7}$$

We want now to form a time derivative in the left side, in order to arrive a (stochastic) differential equation for paths. To do this, we first extract a Δt_i factor from $\Delta B(t_i)$ by performing another change of variables:

$$\Delta B(t_i) \equiv \Delta t_i \xi(t_i) \tag{3.8}$$

so that $\Delta x_i = \sqrt{2D}\Delta t_i \xi_i$, and all the *randomness* is now contained in the random variable ξ, which is distributed according to:

White noise

$$\xi(t_i) \sim \frac{1}{\sqrt{2\pi\Delta t_i}} \exp\left(-\frac{\Delta t_i^2 \xi_i^2}{2\Delta t_i}\right) \underbrace{\frac{\mathrm{d}\Delta B_i}{\mathrm{d}\xi(t_i)}}_{\Delta t_i} = \sqrt{\frac{\Delta t_i}{2\pi}} \exp\left(-\frac{\Delta t_i}{2}\xi_i^2\right) \qquad \xi_i \equiv \xi(t_i)$$

Substituting back in (3.7) and dividing by Δt_i leads to:

$$\frac{x(t_{i+1}) - x(t_i)}{\Delta t_i} = \sqrt{2D}\xi(t_i)$$

And by taking the continuum limit $\Delta t_i \to 0$ we get the **Langevin equation** for a Brownian particle:

Langevin equation

$$\dot{x}(t) = \sqrt{2D}\xi(t) \tag{3.9}$$

We can see $\xi(t)$ as a *highly irregular, quickly varying function*, which, in a certain sense, expresses the result of Brownian collisions at a certain instant. In particular, the following holds:

$$\langle \xi(t) \rangle = 0 \qquad \langle \xi(t)\xi(t') \rangle = \delta(t - t')$$

meaning that the values of $\xi(t)$ at different instants are completely *independent*.

Note that, as we saw previously, Brownian paths are not differentiable - and so $\dot{x}(t)$ does not exist, and this is just a *formal* equation, with a definite meaning only in a given *discretization*. Also, note that $\xi(t)$ is a random variable, and so this is an example of a **stochastic differential equation**. It is not clear how to find a solution to such an equation, or even how to *define* what a solution should be - and this will be the main topic of the next section.

Stochastic Differential Equations

We can rewrite (3.9) in a more *rigorous* form by "multiplying by dt", i.e. performing the change of variables (3.8), which - in the continuum limit - is $\mathrm{d}B = \xi\,\mathrm{d}t$, leading to:

$$\mathrm{d}x(t) = \sqrt{2D}\,\mathrm{d}B \qquad \mathrm{d}B \sim \frac{1}{\sqrt{2\pi\,\mathrm{d}t}} \exp\left(-\frac{\mathrm{d}B^2}{2\,\mathrm{d}t}\right)$$

Before moving on, we want to generalize this equation to the presence of *external forces*. As we saw previously, this just results in adding a *constant velocity motion* to the particle, leading to the full **Langevin equation**:

$$\dot{x}(t) = f(x,t) + \sqrt{2D(x,t)}\xi(t)$$

$$\mathrm{d}x(t) = f(x,t)\,\mathrm{d}t + \sqrt{2D(x,t)}\,\mathrm{d}B \qquad \mathrm{d}B \sim \frac{1}{\sqrt{2\pi\,\mathrm{d}t}} \exp\left(-\frac{\mathrm{d}B^2}{2\,\mathrm{d}t}\right) \tag{3.10}$$

The *physical* meaning of $f(x,t)$ and $D(x,t)$ can be more clearly seen by comparing (3.10) to the equation of motion of the Brownian particle.

Derivation from physical arguments

Consider a particle of mass m immersed in a fluid, with a radius a that is much larger than the surrounding molecules (typically $\sim 10^{-9}$ to 10^{-7} m). The forces acting on it will be that of *viscous friction* $-\gamma\dot{\boldsymbol{r}}$, eventual *external forces* $\boldsymbol{F}_{\text{ext}}$ (e.g. gravity), and a rapidly varying and *random* term $\boldsymbol{F}_{\text{noise}}$, encompassing the effect of the large number of collisions ($\sim 10^{12}$/s) with the smaller fluid particles:

$$m\ddot{\boldsymbol{r}}(t) = -\gamma\dot{\boldsymbol{r}} + \boldsymbol{F}_{\text{ext}} + \boldsymbol{F}_{\text{noise}}(t)$$

Dividing both sides by γ:

$$\frac{m}{\gamma}\ddot{\boldsymbol{r}}(t) = -\dot{\boldsymbol{r}} + \frac{\boldsymbol{F}_{\text{ext}}(\boldsymbol{r},t)}{\gamma} + \frac{\boldsymbol{F}_{\text{noise}}(t)}{\gamma} \tag{3.11}$$

Assuming a spherical particle, γ is given by Stokes law to be $6\pi a\eta$, where η is the viscosity of the surrounding fluid.
Note that, if we ignore the external force and the random term, the equation becomes:

$$\frac{\mathrm{d}\dot{\boldsymbol{r}}(t)}{\mathrm{d}t} = -\frac{\gamma}{m}v(t)$$

which has solution:

$$\dot{\boldsymbol{r}}(t) = \exp\left(-\frac{t}{\tau_B}\right)\dot{\boldsymbol{r}}(0) \qquad \tau_B = \frac{m}{\gamma}$$

τ_B is in the scale of 10^{-3} s, and represents the timescale of reaching equilibrium, i.e. 0 velocity. So, for Brownian motion to happen, $\boldsymbol{F}_{\text{noise}}$ is necessary. Also, if we are interested in the motion on the scale of seconds, we can neglect the acceleration term. This is the **overdamped limit** (in analogy to a damped oscillator with high loss of energy due to attrition, so that it quickly reaches equilibrium without ever "overshooting"). Given that assumption, (3.11) becomes: *Overdamped limit*

$$\dot{\boldsymbol{r}} = \frac{\boldsymbol{F}_{\text{ext}}}{\gamma} + \frac{\boldsymbol{F}_{\text{noise}}}{\gamma}$$

Which, for a particle moving in one dimension, reduces to:

$$x(t) = \underbrace{\frac{F_{\text{ext}}}{\gamma}}_{f(x,t)} + \underbrace{\frac{F_{\text{noise}}}{\gamma}}_{\sqrt{2D(x,t)}\xi(t)}$$

Comparing with (3.10) gives the physical meaning of $f(x,t)$ and $D(x,t)$.

3.3 Summary

(Lesson 7 of 07/11/19) Compiled: October 13, 2020

Summary of the previous lectures. We considered a more general stochastic process, a *Markov Process*, when the future only depends on the present. We wrote a Master Equation, and taking the continuum limit we get a second order partial differential equation, with two coefficients depending on the first two moments of the transition rate: f and D. We would want them to represent the *force* and *diffusion rate*, but we can't find their physical meaning. So we consider the Langenvin equation, reaching the desired physical meaning. There, the increment depends on a *deterministic term* f and a *noise term*:

$$\mathrm{d}x(t) = f(x(t), t)\,\mathrm{d}t + \sqrt{2D(x(t), t)}\,\mathrm{d}B(t) \qquad f = \frac{F_{\text{ext}}}{\gamma}$$

If we discretize this equation, passing to finite differences, we get:

$$\Delta x(t) = f(x(t), t)\Delta t + \sqrt{2D(x(t), t)}\Delta B(t) \qquad \Delta B(t) \sim \frac{1}{\sqrt{2\pi\Delta t}} \exp\left(-\frac{\Delta B^2}{2\Delta t}\right)$$

This is needed because $\mathrm{d}x(t) / \mathrm{d}t$ is ill-defined (as we saw in the previous lecture). Note that $\Delta x(t) = x(t + \Delta t) - x(t)$.
We want to show that this kind equation leads to the same Fokker-Planck equation that we saw previously, and that was derived from the Master Equation. Then we would like to examine how much the stochastic amplitude (coefficient of $\mathrm{d}B(t)$) is related to *temperature*. In fact, we know already that f depends on F_{ext}, with $\boldsymbol{F}_{\text{ext}} = -\boldsymbol{\nabla} V$. We would like that, at constant temperature, the pdf of the stationary state will tend to the *Maxwell-Boltzmann distribution*:

$$\mathbb{P}(x, t) \xrightarrow[t\to\infty]{} \frac{1}{z} \exp\left(-\frac{V(x)}{k_B T}\right)$$

3.4 Stochastic integrals

We arrived at the Langevin equation:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x, t) + \sqrt{2D(x, t)}\xi(t) \tag{3.12}$$

where $\xi(t)$ is a "rapidly varying, highly irregular function", i.e. such that for $t \neq t'$, $\xi(t)$ and $\xi(t')$ are statistically independent. As $\langle \xi(t) \rangle = 0$, this means that:

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

Equation (3.12) does not make much sense, as $\dot{x}(t)$ does not exist anywhere. Even changing variables to $\mathrm{d}B$ (i.e. "multiplying" both sides by $\mathrm{d}t$) and integrating, we are left with the following equation:

$$x(t) = x(0) + \int_0^t f(x(\tau), \tau)\,\mathrm{d}\tau + \int_0^t \sqrt{2D(x(\tau), \tau)}\,\mathrm{d}B(\tau)$$

It is not clear how the last integral is defined, as it involves a *stochastic term* $\mathrm{d}B$.

So, before tackling the full problem, we take a step back and study the theory behind **stochastic calculus**. Let's introduce a *generic* integral of that kind:

$$S_t = \int_0^t G(\tau)\,\mathrm{d}B(\tau)$$

Intuitively, we could see this as an *infinite sum*, where each term $G(\tau)$ is weighted by the outcome of a random variable $B(\tau)$.

So, to compute it, an idea is to first introduce a *time discretization* $\{t_j\}_{j=0,\dots,n}$, with $t_n = t$, leading to:

$$S_n = \sum_{i=0}^{n} G(\tau_i)[B(t_i) - B(t_{i-1})] \qquad t_{i-1} \leq \tau_i \leq t_i \tag{3.13}$$

and then take the continuum limit for $n \to \infty$. This, however, proves to be more difficult than expected, for the following reasons:

- First of all, the increments $B(t_i) - B(t_{i-1})$ are chosen *at random*. This means that S_n is a **random variable**. In fact, we could see S_t as the sum of points from $G(\tau)$, each *weighted* with a *randomly chosen weight*. So it is necessary to define what it means to take the limit of a sequence of random variables S_n. As we will see, there is no unique definition.

- It is not clear how to choose the *sampling instants* τ_i for $G(\tau)$ in the discretization (3.13). We could hope that in the limit of $n \to \infty$, any choice would lead to the same final result. This would be indeed true if $B(\tau)$ were a differentiable function - except it is only *continuous* and *nowhere differentiable*. So we need to pay attention to the *specific* (and arbitrary) rule to be used in computing the discretization.

3.4.1 Limits of sequences of random variables

Some basic definitions. Recall that a probability space is defined by a triple $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is a set of outcomes (*sample space*), $\mathcal{F}$ is a σ-algebra on Ω, containing all possible *events*, that is *sets of outcomes*, and $\mathbb{P}\colon \mathcal{F} \to [0, 1]$ is the *probability measure*. Then, a **random variable** is a *measurable function* $X\colon \Omega \to S$, with S denoting a *state space*.
For example, let Ω be the set of all possible results of rolling two dice, i.e. the set of ordered pairs (x_1, x_2) with $x_1, x_2 \in \{1, 2, 3, 4, 5, 6\}$. Then $\mathcal{F}$ is the set of *all possible subsets* of Ω (including both Ω and $\varnothing$) and $\mathbb{P}\colon \mathcal{F} \ni f \mapsto \mathbb{P}(f)$ is given by:

$$\mathbb{P}(f) = \frac{|f|}{36}$$

where $|f|$ is the cardinality of the set f.
A random variable can be, for example, the *sum* of the two dice:

$$X(\omega) = x_1 + x_2 \quad \forall \omega = (x_1, x_2) \in \Omega$$

Then, we can compute the probability of X assuming a certain value by measuring with $\mathbb{P}$ the preimage set of X:

$$\mathbb{P}(X = 2) = \mathbb{P}(\omega \in \Omega | X(\omega) = 2) = \mathbb{P}(\{1, 1\}) = \frac{1}{36}$$

For discrete one-dimensional variables such as these all of this formalism does not lead to much gain, as there is an immediate and natural choice for $(\Omega, \mathcal{F}, \mathbb{P})$, which is usually denoted by the saying “random”. However, in more complex cases it becomes imperative to precisely define Ω, $\mathcal{F}$ and $\mathbb{P}$, so to avoid ambiguous results (see Bertrand's paradox).

Consider a sequence $\{X_n\}_{n\in\mathbb{N}}$ of random variables in a certain probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Suppose that X is another random variable, and we would like to give meaning to the concept of X_n “tending to” X:

$$X_n \xrightarrow[n\to\infty]{} X$$

There are several possibilities, here stated from the weakest to the strongest:

1. **Convergence in distribution**. In this case, we simply require that the distribution of S_n approaches that of S as $n \to \infty$. Let F_n and F be the cumulative distributions of S_n and S, respectively. Then:

   $$X_n \xrightarrow[n\to\infty]{D} X \Leftrightarrow \lim_{n\to\infty} F_n(x) = F(x) \qquad \forall x \in \mathbb{R} | F \text{ is continuous at } x$$

   (The cumulative distribution, or cdf, is defined as $F_X(x) = \mathbb{P}(X \leq x)$).

   Note that, as we are merely comparing functions, there is no need for X_n or X to be defined on the *same probability space*. Also, here the focus is on *integral properties* of the random variables, so there is no guarantee that sampling X_n and X will lead to *close* results, even for a large n. For example, consider X_n to be a sequence of standard gaussians, which obviously converges to a standard gaussian (X) in the distribution sense. If we sample a number from X_{100} and one from X, they could be arbitrarily far away from each other with a non-zero probability, that remains the same for all n. If we want to *exclude* that possibility we need a *stronger requirement*, which leads to the next definition.

2. **Convergence in probability** (*Stochastic limit*). If the probability of values of X_n being *far* from values of X vanishes as $n \to \infty$, then X_n converges *in probability* to X:

   $$X_n \xrightarrow[n\to\infty]{P} X \Leftrightarrow \lim_{n\to\infty} \mathbb{P}(|X_n - X| > \epsilon) = 0$$

   Expanding the definition, this means that:

   $$\forall \epsilon > 0, \forall \delta > 0, \exists N(\epsilon, \delta) \text{ s.t. } \forall n \geq N, \mathbb{P}(|X_n - X| > \epsilon) < \delta$$

   In other words, the probability of “a significant discordance” between values sampled from X_n and X vanishes as $n \to \infty$. Intuitively, X_n and X are *strongly related*, i.e. they not only distribute similarly, but also come from *similar processes*. For example, let X be the *true length* of a stick chosen *at random* from a population of sticks, and X_n be a measurement of that length made with an instrument that is more and more precise as $n \to \infty$. Then, for large n, it is clear that X_n will have a value that is really close to that of X. In this case, we say that X_n converges *in probability* to X, as $n \to \infty$.

3. **Almost sure convergence**. An even stronger limit requires that:

$$X_n \xrightarrow[n\to\infty]{\text{a.s.}} X \Leftrightarrow \mathbb{P}\left(\liminf_{n\to\infty}\{\omega \in \Omega\colon |X_n(\omega) - X(\omega)| < \epsilon\}\right) = 1 \qquad \forall \epsilon > 0$$

Here, the $\liminf$ of a sequence of sets A_n is defined as:

$$\liminf_{n\to\infty} A_n = \bigcup_{N=1}^{\infty} \bigcap_{n\geq N} A_n$$

A member of $\liminf A_n$ is a member of *all* sets A_n, *except* a *finite* number of them (i.e. it's *definitively* a member of the A_n, as it is $\in A_n$ for all $n \geq \bar{n}$). So the term inside the parentheses is the set of all outcomes $\omega \in \Omega$ for which $X_n(\omega)$ is *definitively* close to $X(\omega)$, i.e. it *covers* all events resulting in a sequence of X_n that *converges* to X.

If we take X_n and X to be *real-valued* random variables, then the definition is simpler:

$$X_n \xrightarrow[n\to\infty]{\text{a.s.}} X \Leftrightarrow \mathbb{P}\left(\omega \in \Omega\colon \lim_{n\to\infty} X_n(\omega) = X(\omega)\right) = 1$$

Or, in other words:

$$\lim_{n\to\infty} X_n(\omega) = X(\omega) \qquad \forall \omega \in \Omega \setminus A$$

where $A \subset \Omega$ has 0 measure.

Almost sure convergence vs probability convergence. The difference between the two definitions is subtle, and can be somewhat seen from the following example, taken from `http://bit.ly/2u2E9Rk` and `http://bit.ly/2Zy66vO`.

Consider a sequence $\{X_n\}$ of *independent* random variables with only two possible values, 0 and 1, such that:

$$\mathbb{P}(X_n = 1) = \frac{1}{n} \qquad \mathbb{P}(X_n = 0) = 1 - \frac{1}{n}$$

For $\epsilon > 0$:

$$\mathbb{P}(|X_n| \geq \epsilon) = \begin{cases} \frac{1}{n} & 0 < \epsilon \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

As $n \to \infty$, $\mathbb{P}(|X_n| \geq \epsilon) \to 0$, and so $X_n \xrightarrow[n\to\infty]{P} 0$.

However, X_n **does not** converge almost surely to 0. Consider a *realization* of the sequence X_n, i.e. the measured outcomes of all X_n during "one run" of the experiment. This will be a binary sequence, like $000101001\ldots$. Now, consider an *ensemble* of such sequences. What is the average number of ones in them? We can estimate it by summing the probability to have a 1 in the first place, in the second, and so on:

$$\sum_{n=1}^{\infty} \frac{1}{n} = +\infty$$

This in fact implies, by the second Borel Cantelli theorem[a], that the probability of getting $X_n = 1$ *infinitely often* (i.o.) is 1, and so X_n *cannot* converge *almost surely* to 0.

[a]∧See a proof at `http://bit.ly/2tcfZU4` The main idea is that, given a set of *independent* events ($X_n = 1$), the sum of their probabilities diverges, then *surely* an infinite number of them *do indeed occur*. Formally: if $\sum_{n=1}^{+\infty} \mathbb{P}(X_n = 1) = \infty$, then $\mathbb{P}(\limsup_{n\to\infty}\{X_n = 1\}) = \mathbb{P}(\cap_{N=1}^{\infty} \cup_{n\geq N} \{X_n = 1\}) = \mathbb{P}(\{X_n = 1\} \text{ i.o.}) = 1$

It can be proven that *almost sure convergence* implies *convergence in probability*, which implies *convergence in distribution*. However, for our purposes we are interested in *another kind of convergence*:

- ***L^q* convergence**:

$$X_n \xrightarrow[n\to\infty]{L^q} X \Leftrightarrow \lim_{n\to\infty} \langle |X_n - X|^q \rangle = 0 \qquad q \in \mathbb{N}$$

  Note that this implies *convergence in probability*. In fact:

$$\mathbb{P}(|X - X_n| > \epsilon) = \langle \mathbb{I}_{|X-X_n|>\epsilon} \rangle \leq \langle \underbrace{\mathbb{I}_{|X-X_n|>\epsilon}}_{0\leq\odot\leq 1} \underbrace{\left| \frac{X - X_n}{\epsilon} \right|^q}_{\geq 1} \rangle \tag{3.14}$$

  where $\mathbb{I}$ is a *characteristic function*, i.e. the random variable that is 1 when $|X - X_n| > \epsilon$ and 0 otherwise - so that the second term is always ≥ 1 when it is not killed by the first one. Then, by substituting $\mathbb{I}$ with its *maximum* 1 we get a *greater* term:

$$(3.14) \leq \langle |X - X_n|^q \rangle \frac{1}{\epsilon^q} \xrightarrow[n\to\infty]{} 0 \qquad \forall \epsilon > 0$$

  where we used the linearity of the average to extract the constant ϵ^q, and then the L^q convergence (assumed by hypothesis).

  Also, L^q convergence implies the convergence (in the usual sense) of the q-th moment:

$$X_n \xrightarrow[n\to\infty]{L^q} X \Rightarrow \lim_{n\to\infty} \langle |X_n|^q \rangle = \langle |X|^q \rangle \tag{3.15}$$

  If we choose $q = 2$, we obtain **mean square convergence**:

$$X_n \xrightarrow[n\to\infty]{\text{m.s.}} X \Leftrightarrow \lim_{n\to\infty} \langle |X_n - X|^2 \rangle = 0$$

  In this case it is easy to prove (3.15) by using the Cauchy-Schwarz inequality:

$$(\mathbb{E}(XY))^2 \leq \mathbb{E}(X^2)\mathbb{E}(Y^2)$$

  If we let $X = X_n - X$ and $Y = 1$, and assume that X_n converges to X in mean square, we obtain:

$$0 \leq (\mathbb{E}(X_n - X))^2 \leq \mathbb{E}((X_n - X)^2)\mathbb{E}(1) \xrightarrow[n\to\infty]{} 0$$

  And so:

$$\mathbb{E}(X_n - X) = \mathbb{E}(X_n) - \mathbb{E}(X) \xrightarrow[n\to\infty]{} 0 \Rightarrow \lim_{n\to\infty} \mathbb{E}(X_n) = \mathbb{E}(X) \qquad \square$$

Hölder inequality. Cauchy inequality is, in this case, a special case of the more general Hölder inequality. Consider a measure space (S, Σ, μ) (where S is the space, Σ a σ-algebra and μ a measure), and two measurable functions $f, g\colon S \to \mathbb{R}$:

$$\|fg\|_1 \leq \|f\|_p \|g\|_p \qquad \|\cdot\|_p = \left(\int_S |\cdot|^p \,\mathrm{d}\mu\right)^{1/p}$$

To compute a stochastic integral, we will proceed like the following:

- Discretize the integral as a finite (Riemann) sum, obtaining a sequence of *finer* and *finer* random variables $\{S_n\}_{n\in\mathbb{N}}$
- Use a *mean square* limit to compute the limit S of the sequence $\{S_n\}$

3.4.2 Prescriptions

All that's left is to choose a *rule* for the mid-points in the terms of the discretized sum. As we will see in the following example, there are several different possibilities, each leading to *different results*.

Example 6 (A simple stochastic integral)**:**

Suppose $G(\tau) = B(\tau)$, and consider the following integral:

$$S = \int_0^t B(\tau)\,\mathrm{d}B(\tau)$$

If $B(\tau)$ where differentiable, then we could simply change variables and solve:

$$S = \int_0^t B(\tau)\frac{\mathrm{d}B(\tau)}{\mathrm{d}\tau}\,\mathrm{d}\tau = \frac{1}{2}B^2(\tau)\Big|_0^t = \frac{B^2(t) - B^2(0)}{2} \quad \text{if } \exists \frac{\mathrm{d}B}{\mathrm{d}\tau}$$

However, here $B(\tau)$ is a *rapidly varying irregular function*, which is nowhere differentiable.
So, following our plan, we first discretize:

$$S_n = \sum_{i=1}^n B(\tau_i)[B(t_i) - B(t_{i-1})] \qquad t_0 \equiv 0;\ t_n \equiv t;\ t_{i-1} \leq \tau_i \leq t_i \tag{3.16}$$

We now need a rule for choosing the τ_i. The simplest possibility is to fix them in the "same relative position" in every interval $[t_{i-1}, t_i]$, that is:

$$\tau_i = \lambda t_i + (1-\lambda)t_{i-1} \qquad \lambda \in [0,1] \tag{3.17}$$

Depending on the value of λ, the limit S will be different. We can quickly check this *before* computing S, by focusing on the *expected values*. In fact, we know that if $S_n \xrightarrow[n\to\infty]{\text{m.s.}} S$, then $\langle S_n\rangle \xrightarrow[n\to\infty]{} S$ in the usual sense. So, we compute the average of S_n:

$$\langle S_n\rangle = \sum_{i=1}^n \langle B(\tau_i)(B(t_i) - B(t_{i-1}))\rangle = \sum_{i=1}^n (\langle B(\tau_i)B(t_i)\rangle - \langle B(\tau_i)B(t_{i-1})\rangle)$$

We already computed the correlator function for the Brownian noise $B(t)$:

$$\langle B(t)B(t')\rangle = \min(t,t') \tag{3.18}$$

And so, as $t_{i-1} \leq \tau_i \leq t_i$, we get:

$$\langle S_n\rangle = \sum_{i=1}^n (\tau_i - t_{i-1})$$

Substituting the choice for τ (3.17):

$$\langle S_n\rangle = \lambda \sum_{i=1}^n (t_i - t_{i-1}) = \lambda t_n = \lambda t$$

Which does not depend on n, making the limit trivial:

$$\langle S\rangle = \lim_{n\to\infty} \langle S_n\rangle = \lambda t$$

This dependence on the **prescription** of τ_i is an important difference from ordinary calculus, meaning that many common results cannot be directly translated to stochastic calculus.
In practice, there are many possibilities for λ. The two most common are:

$$\lambda = \begin{cases} 0 & \text{Ito's prescription} \\ \frac{1}{2} & \text{Stratonovich's prescription (also called middle-point prescription)} \end{cases}$$

Leading to, as we will see:

$$S_n \xrightarrow[n\to\infty]{\text{m.s.}} S = \begin{cases} \frac{B^2(t)-B^2(0)}{2} - \frac{t}{2} & \lambda = 0 \\ \frac{B^2(t)-B^2(0)}{2} & \lambda = 1/2 \end{cases}$$

The Stratonovich prescription gives exactly the same result as ordinary calculus. However, note that it involves a dependence *on the future*, i.e. the next step of a path depends on the point that is a *half-step* later. This has no a real physical meaning (in a certain sense, it "violates causality"). That's why many physicists prefer the Ito's prescription.
Let's explicitly compute both results.

Ito's prescription. We want to prove the following result:

$$\sum_{i=1}^n B(t_{i-1})(B(t_i) - B(t_{i-1})) \xrightarrow[n\to\infty]{\text{m.s.}} \frac{B^2(t) - B^2(0)}{2} - \frac{t}{2} \tag{3.19}$$

Denoting:

$$B(t_i) = B_i; \qquad \Delta B_i = B_i - B_{i-1}$$

we can rewrite (3.16) as:

$$S_n = \sum_{i=1}^n B_{i-1}\Delta B_i$$

First of all, we *split* that product in a sum of terms, with the double-product trick:

$$ab = \frac{1}{2}[(a+b)^2 - a^2 - b^2]$$

So that:

$$\begin{aligned} S_n &= \sum_{i=1}^n B_{i-1}\Delta B_i = \frac{1}{2}\sum_{i=1}^n \Big[\underbrace{(B_{i-1}+\Delta B_i)^2}_{B_i^2} - B_{i-1}^2 - (\Delta B_i)^2\Big] = \\ &= \frac{1}{2}\sum_{i=1}^n \Big[\, B_i^2 - B_{i-1}^2 - (\Delta B_i)^2\Big] = \frac{1}{2}(\, B_n^2 - B_0^2\,) - \frac{1}{2}\sum_{i=1}^n (\Delta B_i)^2 = \\ &= \frac{1}{2}(\, B^2(t) - B^2(0)\,) - \frac{1}{2}\sum_{i=1}^n (\Delta B_i)^2 \end{aligned}$$

Now (3.19) becomes:

$$\frac{B^2(t)-B^2(0)}{2} - \frac{1}{2}\sum_{i=1}^n (\Delta B_i)^2 \xrightarrow[n\to\infty]{\text{m.s.}} \frac{B^2(t)-B(0)}{2} - \frac{t}{2} \qquad t_n = t;\ t_0 = 0$$

Applying the definition of *mean square limit*, this is equivalent to showing that:

$$\left\langle \left| \frac{B^2(t) - \cancel{B^2(0)}}{2} - \frac{1}{2}\sum_{i=1}^n (\Delta B_i)^2 - \left[\frac{B^2(t)-\cancel{B^2(0)}}{2} - \frac{t}{2}\right]\right|^2 \right\rangle \xrightarrow[n\to\infty]{?} 0 \quad (3.20)$$

Expanding:

$$\begin{aligned} \frac{1}{4}\left\langle\left[-\sum_{i=1}^n (\Delta B_i)^2 + t\right]^2\right\rangle &= \frac{1}{4}\left\langle\left[t - \sum_{i=1}^n (\Delta B_i)^2\right]^2\right\rangle \underset{(a)}{=} \frac{1}{4}\left\langle\left[\sum_{i=1}^n (\Delta t_i - \Delta B_i^2)\right]^2\right\rangle = \\ &\underset{(b)}{=} \frac{1}{4}\sum_{i,j=1}^n \langle [\Delta t_i - (\Delta B_i)^2][\Delta t_j - (\Delta B_j)^2]\rangle \qquad (3.21) \end{aligned}$$

where in (a) we used $t = \sum_{i=1}^n \Delta t_i$, and in (b) $(\sum_i a_i)^2 = \sum_{ij} a_i a_j$. We can rewrite the sum highlighting the case where $i = j$:

$$(3.21) = \frac{1}{4}\left[\sum_{i=1}^n \langle [\Delta t_i - (\Delta B_i)^2]^2\rangle + \sum_{i\neq j}^n \langle [\Delta t_i - (\Delta B_i)^2][\Delta t_j - (\Delta B_j)^2]\rangle\right] \qquad (3.22)$$

Noting that the ΔB_i come from *independent gaussians*, we have that the expected values integrals factorize:

$$\langle A\rangle = \int \mathrm{d}\Delta B_i \dots \mathrm{d}\Delta B_n\, A \prod_{i=1}^n \frac{1}{\sqrt{2\pi\Delta t_i}} \exp\left(-\frac{(\Delta B_i)^2}{2\Delta t_i}\right)$$

In other words, this means that the average of the product is just the product of the averages:

$$\langle (\Delta t_i - (\Delta B_i)^2)(\Delta t_j - (\Delta B_j)^2)\rangle = \langle (\Delta t_i - (\Delta B_i)^2)\rangle \langle (\Delta t_j - (\Delta B_j)^2)\rangle =$$

$$= [\Delta t_i - \langle(\Delta B_i)^2\rangle][\Delta t_j - \langle(\Delta B_j)^2\rangle]$$

We already computed the second moment of that gaussian:

$$\langle(\Delta B_i)^2\rangle = \int \frac{d\Delta B_i}{\sqrt{2\pi\Delta t_i}} \Delta B_i^2 \exp\left(-\frac{\Delta B_i^2}{2\Delta t_i}\right) = \Delta t_i$$

and so:

$$\langle(\Delta t_i - (\Delta B_i)^2)\rangle = 0$$

So we are left only with the first term of (3.22):

$$(3.22) = \frac{1}{4}\sum_{i=1}^n \langle[\Delta t_i - (\Delta B_i)^2]^2\rangle = \frac{1}{4}\sum_{i=1}^n \left[\Delta t_i^2 - 2\Delta t_i \underbrace{\langle(\Delta B_i)^2\rangle}_{\Delta t_i} + \langle\Delta B_i^4\rangle\right] \tag{3.23}$$

Recall that, for a random variable x sampled from a gaussian $\mathcal{N}(0,\sigma)$:

$$\langle x^{2n}\rangle = \sigma^{2n}\frac{(2n)!}{2^n n!} = \begin{cases} \sigma^2 & n = 1 \\ \sigma^4 \frac{4!}{4\cdot 2!} = 3\sigma^4 & n = 2 \end{cases}$$

In our case, this means that $\langle(\Delta B_i)^4\rangle = \Delta t_i^2$, leading to:

$$(3.23) = \frac{1}{2}\sum_{i=1}^n \Delta t_i^2$$

When taking the limit of the mesh ($n \to \infty$), the number of summed terms become infinite, but also the size of each of them vanishes:

$$\max_i \Delta t_i \xrightarrow[n\to\infty]{} 0$$

To resolve that limit we need to use the fact that the end-point is fixed ($t_n \equiv t$) and so:

$$\frac{1}{2}\sum_{i=1}^n \Delta t_i^2 \leq \frac{1}{2}\left(\sum_{i=1}^n \Delta t_i\right)^2 = \frac{1}{2}\left(\sum_{i=1}^n \Delta t_i\right)\underbrace{\left(\sum_{j=1}^n \Delta t_j\right)}_{t} \leq \frac{t}{2}\left(\max_i \Delta t_i\right) \xrightarrow[n\to\infty]{} 0$$

This proves (3.20), and so the desired result (3.19).

Stratonovich's prescription. In this case, we want to show that:

$$S_n = \sum_{i=1}^n B\left(\frac{t_i + t_{i-1}}{2}\right)[B(t_i) - B(t_{i-1})] \xrightarrow[n\to\infty]{\text{m.s.}} \frac{B^2(t) - B^2(0)}{2}$$

Note that now we need a set of *middle points* in the mesh, which leads to some complications.

One trick is to simply *double* the "resolution" of the discretization, and choose the *middle points* to be the *odd* indices. We then define:

$$S'_{2n} = \sum_{i=1}^{2n} B_{2i-1}(B_{2i} - B_{2(i-1)})$$

with $t_{2i-1} \equiv (t_{2i} + t_{2(i-1)})/2$, while the t_{2i} may be distributed arbitrarily. The full computation is very long and tedious, and not much enlightening, and is therefore omitted.

A shorter way to compute that, but not as rigorous, is by stating that:

$$S_n = \sum_{i=1}^{n} \frac{B(t_i) + B(t_{i-1})}{2} (B(t_i) - B(t_{i-1}))$$

However it is not obvious that is possible to *approximate* a midpoint of B with an average, as $B(t_i)$ are all random variables. In fact, it is possible to show that the two expressions *have the same distribution*, but they *are not the same random variable*! In any way, if we do this, the thesis immediately follows:

$$= \frac{1}{2} \sum_{i=1}^{n} (B^2(t_i) - B^2(t_{i-1}))$$

3.4.3 Ito's calculus

In our calculations, we will be usually concerned with the following kinds of stochastic integrals $G(t)$:

1. $\int_0^t F(B(\tau))\, \mathrm{d}B(\tau)$
2. $\int_0^t g(\tau)\, \mathrm{d}B(\tau)$
3. $\int_0^t g(\tau)\, \mathrm{d}\tau$ (usual integrals)

These $G(t)$ are called **non-anticipating functions**, because they are independent of $B(t') - B(t)$ for $t' > t$, meaning that they do not dependent on what happens in the Brownian motion at times later than t (i.e. they do not depend on the future). So, by using Ito's prescription (I.p.) in the discretization and *mean square* (m.s.) for the continuum limit we get:

$$\int_0^t F(B(\tau))\, \mathrm{d}B(\tau) \underset{\text{m.s.}}{\overset{\text{I.p.}}{=}} \sum_{i=1}^{n} F(B_{i-1}) \Delta B_i$$

Note how $F(B_{i-1})$ and ΔB_i are independent of each other, simplifying the calculations. (Note that the Stratonovich prescription here causes *troubles* during evaluation, as it introduces some interdependence between different terms).

3.5 Stochastic Differential Calculus

(Lesson 8 of 11/11/19) Compiled: October 13, 2020

3.5.1 Ito's rules of integration

We now consider a more general stochastic integral, and show that, using Ito's prescription:

$$\int_0^t H(B(\tau),\tau)(\mathrm{d}B(\tau))^k \underset{m.s.}{\overset{\text{I.p.}}{=}} \sum_{i=1}^n H(B_{i-1},\tau_{i-1})(\Delta B_i)^k =$$

$$= \begin{cases} \int_0^t H(B,\tau)\,\mathrm{d}B(\tau) & k=1 \\ \int_0^t H(B(\tau),\tau)\,\mathrm{d}\tau & k=2 \\ 0 & k>2 \end{cases}$$

This leads to the following "rules" for *Ito integrals*:

$$(\mathrm{d}B)^n = \begin{cases} \mathrm{d}B & n=1 \\ \mathrm{d}t & n=2 \\ 0 & n>2 \end{cases} \tag{3.24}$$

We already showed an example for $k=1$, and we now proceed with the other two cases.

Example 7 (Integral in $\mathbf{d}\boldsymbol{B}^{\mathbf{2}}$)**:**

Consider a *non-anticipating* function $G(\tau)$, and the following stochastic integral:

$$I = \int_0^t G(\tau)(\mathrm{d}B\,(\tau))^2$$

With *non-anticipating* we mean that $G(\tau)$ does not depend on $B(s) - B(\tau)\,\forall s>\tau$, i.e. it does not depend on the *future*. Discretizing:

$$I = \underset{n\to\infty}{\overset{\text{m.s.}}{\lim}} I_n = \underset{n\to\infty}{\overset{\text{m.s.}}{\lim}} \sum_{i=1}^n G(t_{i-1})\Delta B_i^2$$

For simplicity, denote:

$$G_i \equiv G_i \qquad \Delta B_i \equiv B_i - B_{i-1} \qquad \Delta t_i = t_i - t_{i-1}$$

We want to prove that:

$$\int_0^t G(\tau)(\mathrm{d}B(\tau))^2 \overset{?}{=} \int_0^t G(\tau)\,\mathrm{d}\tau = \lim_{n\to\infty}\sum_{i=1}^n G_{i-1}\Delta t_i$$

Applying the definition of a *mean square* limit, this is equivalent to:

$$\left\langle \left(\sum_{i=1}^n G_{i-1}\Delta B_i^2 - \sum_{i=1}^n G_{i-1}\Delta t_i \right)^2 \right\rangle \xrightarrow[n\to\infty]{?} 0$$

Expanding the square as a product of two sums over i and j, and then highlighting the case with $i = j$:

$$\langle \left[\sum_{i=1}^{n} G_{i-1}[(\Delta B_i)^2 - \Delta t_i] \right]^2 \rangle = \sum_{i,j=1}^{n} \langle G_{i-1}[(\Delta B_i)^2 - \Delta t_i] G_{j-1}[(\Delta B_j)^2 - \Delta t_j] \rangle =$$
$$= \sum_{i=1}^{n} \langle G_{i-1}^2 [(\Delta B_i)^2 - \Delta t_i]^2 \rangle + 2 \sum_{i<j}^{n} \langle G_{i-1}[(\Delta B_i)^2 - \Delta t_i] G_{j-1} \, [(\Delta B_j)^2 - \Delta t_j] \rangle \tag{3.25}$$

As $i < j$, note that the yellow term *does not depend* on $\Delta B_j = B_j - B_{j-1} = B(t_j) - B(t_{j-1})$. In fact, as G is *non-anticipating*, G_{j-1} depends only on the previous steps. Thus, the yellow and blue terms are *independent* of each other, and so we can factorize the average:

$$(3.25) = \sum_{i=1}^{n} \langle G_{i-1}^2 [(\Delta B_i)^2 - \Delta t_i]^2 \rangle + 2 \sum_{i<j}^{n} \langle G_{i-1}[(\Delta B_i)^2 - \Delta t_i] G_{j-1} \rangle \, \langle (\Delta B_j)^2 - \Delta t_j \rangle$$

Recall that:

$$\langle (\Delta B_j)^2 - \Delta t_j \rangle = \langle (\Delta B_j)^2 \rangle - \Delta t_j = 0$$

and so only the first term of (3.25) remains. Again, noting that G_{i-1} does not depend on ΔB_i, as it is *non-anticipating*, can factorize the average:

$$(3.25) = \langle \sum_{i=1}^{n} G_{i-1}^2 [(\Delta B_i)^2 - \Delta t_i]^2 \rangle = \sum_{i=1}^{n} \underbrace{\langle G_{i-1}^2 \rangle}_{G_{i-1}^2} \langle [(\Delta B_i)^2 - \Delta t_i]^2 \rangle \tag{3.26}$$

Expanding the stochastic term:

$$\langle [(\Delta B_i)^2 - \Delta t_i]^2 \rangle = \langle (\Delta B_i)^4 - 2\Delta t_i (\Delta B_i)^2 \rangle + \Delta t_i^2 =$$
$$= \underbrace{\langle (\Delta B_i)^4 \rangle}_{3(\Delta t_i)^2} - 2\Delta t_i \underbrace{\langle (\Delta B_i)^2 \rangle}_{\Delta t_i} + \Delta t_i^2 = 2\Delta t_i^2$$

And substituting back into the sum and taking the limit completes the proof:

$$(3.26) = 2 \sum_{i=1}^{n} G_{i-1}^2 \Delta t_i^2 \leq 2 \left(\max_{i \leq j \leq n} \Delta t_j \right) \sum_{i=1}^{n} G_{i-1}^2 \Delta t_i \xrightarrow[n \to \infty]{} 2 \cdot 0 \cdot \int_0^t G^2(\tau) \, d\tau = 0$$

This proves that $(dB)^2 = dt$.

Example 8 (The case with $\boldsymbol{n > 2}$)**:**

We want now to show that:

$$\int_0^t G(\tau)(\mathrm{d}B(\tau))^n = \underset{n\to\infty}{\overset{\text{m.s.}}{\lim}} \sum_{i=1}^n G_{i-1}(\Delta B_i)^n = 0$$

By definition, we want to show that:

$$\left\langle \left(\sum_{i=1}^n G_{i-1}(\Delta B_i)^n \right)^2 \right\rangle \xrightarrow[n\to\infty]{} 0$$

Expanding the square, and factorizing the averages (as G is *non-anticipating*) leads to:

$$\begin{aligned}\left\langle \left(\sum_{i=1}^n G_{i-1}(\Delta B_i)^n \right)^2 \right\rangle &= \sum_{i=1}^n \langle G_{i-1}^2 (\Delta B_i)^{2n} \rangle + 2\sum_{i<j}^n \langle G_{i-1}G_{j-1}(\Delta B_i)^n (\Delta B_j)^n \rangle = \\ &= \sum_{i=1}^n G_{i-1}^2 \langle (\Delta B_i)^{2n} \rangle + 2\sum_{i<j}^n \langle G_{i-1}G_{j-1}(\Delta B_i)^n \rangle \langle (\Delta B_j)^n \rangle\end{aligned} \tag{3.27}$$

Now, recall that the p-th central moment of $X \sim \mathcal{N}(\mu,\sigma)$ can be computed with Isserlis theorem, resulting in:

$$\mathbb{E}[(X-\mu)^p] = \begin{cases} 0 & p \text{ is odd} \\ \sigma^p (p-1)!! & p \text{ is even} \end{cases}$$

where $p!! = p\cdot(p-2)\cdot \dots \cdot 1$ is a *double factorial*, that can be rewritten in terms of factorials as follows:

$$p!! = \begin{cases} 2^k k! & p = 2k \text{ even} \\ \dfrac{(2k)!}{2^k k!} & p = 2k-1 \text{ odd} \end{cases} \tag{3.28}$$

So, if $\boldsymbol{n}$ **is odd**, the blue term in (3.27) vanishes. Let's suppose, for simplicity, that G is bounded, i.e. $|G(\tau)| < K \ \forall \tau \in \mathbb{R}$. Then:

$$\begin{aligned}(3.27) &= \sum_{i=1}^n G_{i-1}^2 (\Delta t_i)^n (2n-1)!! = \sum_{i=1}^n G_{i-1}^2 (\Delta t_i)^n \frac{(2n)!}{2^n n!} \leq \frac{K^2 (2n)!}{2^n n!} \sum_{i=1}^n (\Delta t_i)^n \\ &\leq \frac{K^2(2n)!}{2^n n!} \left(\max_{i\leq j\leq n} (\Delta t)^{n-1} \right) \underbrace{\sum_{i=1}^n \Delta t_i}_{t} \xrightarrow[n\to\infty]{} 0\end{aligned}$$

On the other hand, if $\boldsymbol{n}$ **is even**, the blue term in (3.27) is not null. However, the same argument for n odd can be applied to the first term, which vanishes in the limit. So we only need to study the blue term:

$$(3.27) = 2\sum_{i<j}^n \langle \underbrace{G_{i-1}G_{j-1}}_{\leq K^2} (\Delta B_i)^n \rangle \langle (\Delta B_j)^n \rangle \tag{3.29}$$

Here, as n is even:

$$\langle (\Delta B_i)^n \rangle = (\Delta t_i)^{n/2}(n-1)!! = (\Delta t_i)^{n/2}\left(2\frac{n}{2}-1\right)!! \underset{(3.28)}{=} (\Delta t_i)^{n/2}\frac{n!}{2^{n/2}(n/2)!}$$

And so:

$$\begin{aligned}(3.29) &\leq 2K^2\left(\frac{n!}{2^{n/2}(n/2)!}\right)^2 \sum_{i<j}^n \Delta t_i^{n/2}\Delta t_j^{n/2} \\ &\leq 2K^2\left(\frac{n!}{2^{n/2}(n/2)!}\right)^2 \left(\max_{i\leq l\leq n}\Delta t_l\right)^{2(n/2-1)} \underbrace{\sum_{i<j}^n \Delta t_i \Delta t_j}_{\leq t^2} \xrightarrow[n\to\infty]{} 0 \qquad \square\end{aligned}$$

Example 9 (Other cases)**:**

Ito's rules allow us to consider even more general integrals. For example:

$$\int_0^t G(\tau)\,\mathrm{d}B(\tau)\,\mathrm{d}\tau = 0$$

In fact, as $(\mathrm{d}B)^2 = \mathrm{d}\tau$, $\mathrm{d}B\,\mathrm{d}\tau = 0$ because $(\mathrm{d}B)^n = 0\ \forall n > 2$.

Example 10 (Integration of polynomials)**:**

By using Ito's rules we can find a formula for integrating *powers* of the Brownian motion:

$$\int_0^t (B(\tau))^n\,\mathrm{d}B(\tau)$$

We first differentiate a polynomial, and then recover the rule for integration by performing the inverse operation.
Recall that, in general, a differential is *the increment* of a function after a small *nudge* of its argument:

$$\mathrm{d}f(t) = f(t+\mathrm{d}t) - f(t)$$

The same holds in the stochastic case. In particular:

$$\begin{aligned}\mathrm{d}(B(t))^n &= [B(t+\mathrm{d}t)]^n - (B(t))^n = [B(t)+\mathrm{d}B(t)]^n - (B(t))^n = \\ &\underset{(a)}{=} \sum_{k=0}^n \binom{n}{k}(\mathrm{d}B(t))^k (B(t))^{n-k} - (B(t))^n = \\ &= \cancel{(B(t))^n} + \sum_{k=1}^n \binom{n}{k}(\mathrm{d}B(t))^k (B(t))^{n-k} - \cancel{(B(t))^n} = \\ &\underset{(b)}{=} \underbrace{n(\mathrm{d}B(t))(B(t))^{n-1}}_{k=1} + \underbrace{\frac{n(n-1)}{2}\overbrace{(\mathrm{d}B(t))^2}^{\mathrm{d}t}(B(t))^{n-2}}_{k=2} + \underbrace{0}_{k>2}\end{aligned}$$

where in (a) we used Newton's binomial formula, and in (b) the previously found Ito's rules for integration (3.24). Letting $m = n - 1$ and isolating $\mathrm{d}B(t)$ leads to:

$$(m+1)(B(t))^m \,\mathrm{d}B(t) = (\mathrm{d}B(t))^{m+1} - \frac{m(m+1)}{2}(B(t))^{m-1}\,\mathrm{d}t$$

Finally, dividing by $m+1$ and integrating leads to the desired formula:

$$\begin{aligned}
\int_0^\tau (B(t))^m \,\mathrm{d}B(t) &= \frac{1}{m+1}\int_0^\tau \mathrm{d}(B(t))^{m+1} - \frac{m}{2}\int_0^\tau (B(t))^{m-1}\,\mathrm{d}t = \\
&= \frac{1}{m+1}(B(t))^{m+1}\Big|_0^\tau - \frac{m}{2}\int_0^\tau (B(t))^{m-1}\,\mathrm{d}t = \\
&= \frac{(B(\tau))^{m+1} - (B(0))^{m+1}}{m+1} - \frac{m}{2}\int_0^\tau (B(t))^{m-1}\,\mathrm{d}t
\end{aligned}$$

And in the case $m = 1$ we retrieve the previously obtained result:

$$\int_0^\tau B(t)\,\mathrm{d}B(t) = \frac{B^2(\tau) - B^2(0)}{2} - \frac{\tau}{2}$$

Example 11 (General differentiation rule)**:**

Because $(\mathrm{d}B)^2 = \mathrm{d}t$, when computing differentials from a Taylor expansion up to $O(\mathrm{d}t^2)$ one must compute even the terms of order $\mathrm{d}B^2$. For example, consider a generic function $f(B(t), t)$:

$$\begin{aligned}
\mathrm{d}f(B(t), t) = &\frac{\partial f}{\partial t}\,\mathrm{d}t + \frac{\partial f}{\partial B}\,\mathrm{d}B(t) + \underbrace{\frac{1}{2}\frac{\partial^2 f}{\partial t^2}^2}_{O([\mathrm{d}t]^2)} + \frac{1}{2}\frac{\partial^2 f}{\partial B^2}\underbrace{[\mathrm{d}B(t)]^2}_{\mathrm{d}t} + \\
&+ \frac{\partial^2 f}{\partial B(t)\partial t}\underbrace{\mathrm{d}t\,\mathrm{d}B(t)}_{0} + O([\mathrm{d}t]^2) = \\
= &\frac{\partial f}{\partial t}\,\mathrm{d}t + \frac{\partial f}{\partial B}\,\mathrm{d}B(t) + \frac{1}{2}\frac{\partial^2 f}{\partial B^2}\,\mathrm{d}t + O([\mathrm{d}t]^2)
\end{aligned}$$

3.6 Derivation of the Fokker-Planck equation

Starting from the Master Equation and taking the continuum limit we arrived at the Fokker-Planck equation:

$$\dot{W}(x, t) = -\frac{\partial}{\partial x}\left[f(x,t)W(x,t) - \frac{\partial}{\partial x}W(x,t)D(x,t)\right] \tag{3.30}$$

At the same time, if we consider the dynamics of a *single path*, adding a *stochastic term* to the second law of motion, we arrive at the Langevin equation (in the overdamped limit):

$$\mathrm{d}x(t) = f(x(t), t)\,\mathrm{d}t + \sqrt{2D(x(t), t)}\,\mathrm{d}B(t) \tag{3.31}$$

We want now to show that these two formulations are equivalent, by deriving (3.30) from (3.31). The main idea is to introduce a *test function* $h(x(t))$, and compute its expected value at the instant t over *all possible points* that can be reached by the trajectory $x(t)$, thus obtaining a value that will depend on the *global* probability distribution $W(x,t)$. Then, we can use Langevin equation to describe the dynamics of each *single path*. In this way, we will obtain a relation between a quantity involving $W(x,t)$ and the parameters $f(x,t)$ and $D(x,t)$ appearing in (3.31), which will hopefully be (3.30).
So, let's start by computing the average of $h(x(t))$ at a fixed time:

$$\langle h(x(t))\rangle = \int_{\mathbb{R}} \mathrm{d}x\, W(x,t)h(x)$$

As we seek to construct a *time derivative*, we start by differentiating:

$$\mathrm{d}\langle h(x(t))\rangle = \left(\frac{\partial}{\partial t}\int_{\mathbb{R}} \mathrm{d}x\, W(x,t)h(x)\right)\mathrm{d}t = \mathrm{d}t \int_{\mathbb{R}} \mathrm{d}x\, \dot{W}(x,t)h(x) \tag{3.32}$$

And then dividing by $\mathrm{d}t$ leads to:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle h(x(t))\rangle = \int_{\mathbb{R}} \mathrm{d}x\, \dot{W}(x,t)h(x) \tag{3.33}$$

However, we could also start by differentiating $h(x(t))$:

$$\begin{aligned}
\mathrm{d}h(x(t)) &= h(x(t)+\mathrm{d}x\,(t)) - h(x(t)) = &(3.34)\\
&\underset{(a)}{=} h'(x(t))\,\mathrm{d}x\,(t) + \frac{1}{2}h''(x(t))[\mathrm{d}x\,(t)]^2 + O([\mathrm{d}x\,(t)]^2) &(3.35)
\end{aligned}$$

where in (a) we used a Taylor expansion for the first term. From (3.31), and applying Ito's rules, we can obtain explicit expressions for the $[\mathrm{d}x(t)]^n$:

$$\begin{aligned}
[\mathrm{d}x(t)]^2 &= f^2[\mathrm{d}t]^2 + 2D\,\overbrace{[\mathrm{d}B(t)]^2}^{\mathrm{d}t} + f\sqrt{2D}\,\overbrace{\mathrm{d}B(t)\,\mathrm{d}t}^{0}\\
[\mathrm{d}x(t)]^3 &= O([\mathrm{d}t]^2)
\end{aligned}$$

And substituting in (3.35) leads to:

$$\begin{aligned}
\mathrm{d}h(x(t)) &= h'[f\,\mathrm{d}t + \sqrt{2D}\,\mathrm{d}B] + \frac{1}{2}h''2D\,\mathrm{d}t + O([\mathrm{d}t]^2) =\\
&= \mathrm{d}t\,[h'f + h''D] + h'\sqrt{2D}\,\mathrm{d}B
\end{aligned}$$

Taking the expected value:

$$\begin{aligned}
\mathrm{d}\langle h(x(t))\rangle &= \langle \mathrm{d}t\,[h'f + h''D]\rangle + \langle h'\sqrt{2D}\,\mathrm{d}B\rangle =\\
&\underset{(a)}{=} \langle \mathrm{d}t\,[h'f + h''D]\rangle + \langle\sqrt{2D}h'\rangle\,\underbrace{\langle \mathrm{d}B\rangle}_{0} =\\
&= \langle \mathrm{d}t\,[h'f + h''D]\rangle
\end{aligned}$$

where in (a) we used the fact that $D(x(t),t)$ is *non-anticipating*, allowing to factor the average.

Dividing by $\mathrm{d}t$ and expanding the average leads to:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\langle h(x(t))\rangle &= \int_{\mathbb{R}} \mathrm{d}x\, W(x,t)[h'(x)f(x,t) + h''(x)D(x,t)] = \\
&= \int_{\mathbb{R}} \mathrm{d}x\, W(x,t)f(x,t)h'(x) + \int_{\mathbb{R}} \mathrm{d}x\, W(x,t)D(x,t)h''(x) = \\
&\underset{(a)}{=} \cancel{Whf\Big|_{-\infty}^{+\infty}} - \int_{\mathbb{R}} \mathrm{d}x\, h\frac{\partial}{\partial x}(Wf) + \\
&+ \cancel{WDh'\Big|_{-\infty}^{+\infty}} - \cancel{h\frac{\partial}{\partial x}(DW)\Big|_{-\infty}^{+\infty}} + \int_{\mathbb{R}} \mathrm{d}x\, h\frac{\partial^2}{\partial x^2}(WD) = \\
&= \int_{\mathbb{R}} \mathrm{d}x\, h(x)\left[\frac{\partial^2}{\partial x^2}(W(x,t)D(x,t)) - \frac{\partial}{\partial x}(W(x,t)f(x,t))\right] \quad (3.36)
\end{aligned}$$

where in (a) we integrated by parts the first integral once, and the second one twice.
Finally, equating (3.33) and (3.36) leads to:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle h(x(t))\rangle = \int_{\mathbb{R}} \mathrm{d}x\, \frac{\partial}{\partial t}W(x,t)h(x) = \int_{\mathbb{R}} \mathrm{d}x\, h(x)\left[\frac{\partial^2}{\partial x^2}(W(x,t)D(x,t)) - \frac{\partial}{\partial x}(W(x,t)f(x,t))\right]$$

As this relation holds for *any* test function $h(x)$, it means that the *integrands* are equal. So, by collecting a derivative, we retrieve the the Fokker-Planck equation (3.30):

$$\frac{\partial}{\partial t}W(x,t) = -\frac{\partial}{\partial x}\left[f(x,t)W(x,t) - \frac{\partial}{\partial x}(W(x,t)D(x,t))\right]$$

3.7 The role of temperature

From physical observations, we expect the amplitude of stochastic oscillations in Brownian motion to be dependent on temperature - as it is a direct effect of collisions with molecules in thermal equilibrium. So, we want to derive an explicit relation between the diffusion parameter D and T.
We start by assuming that, for $t \to \infty$, the particle will be *at equilibrium*, meaning that its distribution will be given by the Maxwell-Boltzmann:

$$W(x,t) \xrightarrow[t\to\infty]{} P_{\mathrm{eq}}(x) = \frac{e^{-\beta V(x)}}{Z} \qquad Z = \int_{\mathbb{R}} \mathrm{d}x\, e^{-\beta V(x)}; \qquad \beta = \frac{1}{k_B T}$$

Recall the Fokker-Planck equation:

$$\frac{\partial}{\partial t}W(x,t) = -\frac{\partial}{\partial x}\left[f(x,t)W(x,t) - \frac{\partial}{\partial x}(D(x,t)W(x,t))\right]$$

From the Langevin equivalence, and some physical reasoning, we found that:

$$f(x,t) = \frac{F_{\mathrm{ext}}}{\gamma} = -\frac{1}{\gamma}\frac{\partial V(x)}{\partial x} \qquad \gamma = 6\pi\eta a$$

Where F_{ext} is an external conservative force with potential $V(x)$ acting on the Brownian particle, assumed to be a sphere of radius a moving through a medium of viscosity η. Assuming $D(x,t) \equiv D$ for simplicity, the Fokker-Planck equation becomes:

$$\frac{\partial W^*}{\partial t} = \frac{\partial}{\partial x}\left[\frac{W^*}{\gamma}\frac{\partial V}{\partial x} + D\frac{\partial W^*}{\partial x}\right]$$

Here we are interested in the *particular solution* $W^*(x)$ that will be reached at the equilibrium, as it *does not* depend on time. So:

$$\frac{\partial W^*}{\partial t} \stackrel{!}{=} 0$$

Meaning that:

$$\left[\frac{W^*(x)}{\gamma}\frac{\partial V}{\partial x} + D\frac{\partial W^*}{\partial x}\right] = \text{constant} \qquad \forall x \tag{3.37}$$

As this relation holds for *any* x, we can examine it in the limit $x \to \infty$ to find the value of the constant. In fact, as $W^*(x)$ is a normalized pdf, we expect:

$$W^*, \frac{\partial W^*}{\partial x} \xrightarrow[x\to\infty]{} 0$$

And so the constant in (3.37) must be 0, leading to:

$$\frac{\partial W^*}{\partial x} = -\frac{1}{\gamma D}W^*\frac{\partial V}{\partial x} \Rightarrow \frac{1}{W^*}\frac{\partial W^*}{\partial x} = \frac{\partial \ln\left(W^*\right)}{\partial x} = -\frac{1}{\gamma D}\frac{\partial V}{\partial x}$$

Integrating, we find:

$$\ln W^*(x) = -\frac{1}{\gamma D}V(x) + c \Rightarrow W^*(x) = K\exp\left(-\frac{1}{\gamma D}V(x)\right) \stackrel{!}{=} \frac{1}{Z}\exp\left(-\beta V(x)\right)$$

And by comparing the two functions we obtain the desired relation:

$$\beta = \frac{1}{\gamma D} = \frac{1}{k_B T} \Rightarrow D = \frac{k_B T}{\gamma} = \frac{k_B T}{6\pi\eta a}$$

This is indeed the same relation that Einstein found when examining Brownian motion (*fluctuation-dissipation relationship*, 1905). As $D(x,t) \propto T$, the amplitude of stochastic oscillations (from Langevin equation) is proportional $\sqrt{2D} \propto \sqrt{T}$.

3.8 Harmonic overdamped oscillator

(Lesson 9 of 14/11/19) Compiled: October 13, 2020

Using the framework developed in the previous sections, we now tackle a more general setting, that of a particle moving in a *harmonic potential* and subject to thermal noise. This will be useful to model the local behaviour about the minima of *any* potential - as they are approximately harmonic.
So, consider a particle of mass m moving in one dimension through a *viscous* medium and immersed in a *harmonic* potential. To model the random collisions

with the other (much smaller) particles in the fluid we add a *stochastic term* $\sqrt{2D}\gamma\xi$. The equation of motion becomes:

$$m\ddot{x} = -\gamma\dot{x} - m\omega^2 x + \sqrt{2D}\gamma\xi \tag{3.38}$$

As m/γ is much smaller than the timescale we are interested in, we can neglect it, reaching the *overdamped limit*:

$$\dot{x} = -\underbrace{\frac{m\omega^2}{\gamma}}_{k} x + \sqrt{2D}\xi$$

And multiplying by $\mathrm{d}t$:

$$\mathrm{d}x(t) = -kx(t)\,\mathrm{d}t + \sqrt{2D}\,\mathrm{d}B(t) \tag{3.39}$$

As usual, we introduce a time discretization $\{t_j\}_{j=1,\dots,n}$. Letting:

$$x(t_i) \equiv x_i; \quad \Delta x_i \equiv x_i - x_{i-1}; \qquad B(t_i) \equiv B_i; \qquad \Delta t_i = t_i - t_{i-1}$$

we arrive to:

$$\Delta x_i = -kx_{i-1}\Delta t_i + \sqrt{2D}\Delta B_i \tag{3.40}$$

Note that we evaluated the potential term $-kx(\tau)$ at the *left extremum* of the discretized interval $[t_{i-1}, t_i]$, following Ito's prescription.
To solve (3.39) the plan will be the following:

1. Use the discretization to find the *infinitesimal probability* $\mathbb{P}(\{\Delta x_i\}_{i=1,\dots,n})$ of a *discretized path*, i.e. of a path traversing all *gates* $[x_i, x_i + \mathrm{d}x_i]$ at successive instants $0 \equiv t_1 < \dots < t_n \equiv t$.

2. Find the probability for a continuous path $\mathrm{d}P \equiv \mathbb{P}(\{x(\tau)_{\tau\in[0,t]}\})$ by taking the limit $n \to \infty$.

3. Find the transition probabilities that solve (3.39) by using a *path integral* to evaluate:

$$W(x_t, t; x_0, 0) = \langle \delta(x_t - x(\tau)) \rangle_W \equiv \int_{\mathbb{R}^T} \delta(x_t - x(\tau))\,\mathrm{d}P$$

   In other words, this is the *fraction* of paths (from the set $\mathbb{R}^T$ of all continuous paths happening in the timeframe $[0, t]$) that start in x_0 at instant 0, and reach x_t at instant t.

To find $\mathbb{P}(\{\Delta x_i\}_{i=1,\dots,n})$ we start from the joint pdf $\mathbb{P}(\{\Delta B_i\}_{i=1,\dots,n})$ that we already know, and perform a change of random variables according to (3.40).
In practice, start from:

$$\mathbb{P}(\Delta B_1, \dots, \Delta B_n) = \prod_{i=1}^n \frac{\mathrm{d}\Delta B_i}{\sqrt{2\pi\Delta t_i}} \exp\left(-\sum_{i=1}^n \frac{\Delta B_i^2}{2\Delta t_i}\right)$$

Then insert ΔB_i in terms of Δx_i from (3.40):

$$\Delta B_i = \frac{\Delta x_i + k x_{i-1} \Delta t_i}{\sqrt{2D}}$$

and then multiply by the determinant J of the jacobian of the change of variables to find the desired new pdf:

$$\begin{aligned}
\mathbb{P}(x_1, x_2, \dots, x_n) &= \mathbb{P}(\Delta x_1)\mathbb{P}(\Delta x_2|\Delta x_1)\mathbb{P}(\Delta x_3|\Delta x_1, \Delta x_2)\cdots = \\
&= \prod_{i=1}^n \frac{\mathrm{d}\Delta x_i}{\sqrt{2\pi \Delta t_i}} \exp\left(-\sum_{i=1}^n \frac{1}{2\Delta t_i}\left(\frac{\Delta x_i + k x_{i-1}\Delta t_i}{\sqrt{2D}}\right)^2\right) J
\end{aligned}$$

$$J = \det\left|\frac{\partial(\Delta B_1, \dots, \Delta B_n)}{\partial(\Delta x_1, \dots, \Delta x_n)}\right| = \det\left|\frac{\partial(\Delta x_1, \dots, \Delta x_n)}{\partial(\Delta B_1, \dots, \Delta B_n)}\right|^{-1} = \begin{vmatrix} \sqrt{2D} & 0 & \cdots & 0 \\ * & \sqrt{2D} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ * & \dots & * & \sqrt{2D} \end{vmatrix}_{n\times n}^{-1} = (2D)^{-n/2}$$

The elements under the diagonal are, in general, non-zero derivatives. However, as the matrix is lower triangular, its determinant is just the product of the diagonal elements. Substituting back:

$$\mathbb{P}(\Delta x_1, \dots, \Delta x_n) = \prod_{i=1}^n \left(\frac{\mathrm{d}\Delta x_i}{\sqrt{4\pi D \Delta t_i}}\right) \exp\left(-\sum_{i=1}^n \frac{1}{2\Delta t_i}\left(\frac{\Delta x_i + k x_{i-1}\Delta t_i}{\sqrt{2D}}\right)^2\right) \tag{3.41}$$

Taking the limit $n \to \infty$:

$$\mathrm{d}P \equiv \mathbb{P}(x(\tau)) = \left(\prod_{\tau=0^+}^t \frac{\mathrm{d}x(\tau)}{\sqrt{4\pi D\, \mathrm{d}\tau}}\right) \exp\left(-\frac{1}{4D}\int_0^t (\dot{x} + kx)^2 \,\mathrm{d}\tau\right)$$

where we used:

$$\frac{1}{\Delta t_i}(\Delta x_i + k x_{i-1}\Delta t_i)^2 = \frac{\Delta t_i^2}{\Delta t_i}\left(\frac{\Delta x_i}{\Delta t_i} + k x_{i-1}\frac{\Delta t_i}{\Delta t_i}\right)^2 \xrightarrow[n\to\infty]{} (\dot{x} + kx)^2 \,\mathrm{d}t$$

Expanding the square in (3.41):

$$\mathrm{d}P = \underbrace{\prod_{i=1}^n \frac{\mathrm{d}\Delta x_i}{\sqrt{4\pi D \Delta t_i}} \exp\left(-\sum_{i=1}^n \frac{\Delta x_i^2}{4D\Delta t_i}\right)}_{\text{Wiener measure } (\mathrm{d}x_W)} \underbrace{\exp\left(-\frac{k}{2D}\sum_{i=1}^n x_{i-1}\Delta x_i\right)}_{\text{stochastic integral}} \underbrace{\exp\left(-\frac{k^2}{4D}\sum_{i=1}^n \Delta t_i x_{i-1}^2\right)}_{\text{normal integral}} \tag{3.42}$$

Let's focus on the stochastic integral. We already know that, for Ito's integrals, the usual rules of calculus do not apply. In particular, we can't just do:

$$\sum_{i=1}^n x_{i-1}\Delta x_i \xrightarrow[n\to\infty]{} \int_0^t x(\tau)\,\mathrm{d}x(\tau) \neq \frac{x^2(t) - x^2(0)}{2}$$

So, more in general for a differentiable function $h(x)$:

$$\int_0^t h'(\tau)\,\mathrm{d}x(\tau) \neq h(x(t)) - h(x(0)) \tag{3.43}$$

The idea is now to start from the right side and use Ito's rules to *correct* the left side, so to have a usable identity for integration. As always, we start by discretizing time $\{t_i\}_{i=1,\dots,n}$:

$$h(x(t)) - h(x(0)) = \sum_{i=1}^{n} [h(x(t_i)) - h(x(t_{i-1}))] \equiv \sum_{i=1}^{n} \Delta h_i$$

In the limit, $t_i = t_{i-1} + \mathrm{d}t$, and so the Δh_i are *differentials* of h:

$$\Delta h_i = \frac{\mathrm{d}h}{\mathrm{d}x_i}\Delta x_i + \frac{1}{2}\frac{\mathrm{d}^2 h}{\mathrm{d}x_i^2}\Delta x_i^2 + O(\Delta x_i^3)$$

Now:

$$\Delta x_i = \frac{\mathrm{d}\Delta B_i}{\mathrm{d}\Delta x_i}\Delta B_i + O(\Delta B_i^2) \approx \sqrt{2D}\Delta B_i$$

And by Ito's rules, $\Delta B_i^2 = \Delta t_i$ and $\Delta B_i^n = 0$ for $n \geq 3$. So:

$$\Delta h_i = h' \Delta x_i + \frac{1}{2} h'' \underbrace{\Delta x_i^2}_{2D\Delta t_i}$$

And substituting back in (3.43) leads to:

$$h(x(t)) - h(x(0)) = \sum_{i=1}^{n} (h_i' \Delta x_i + h'' D \Delta t_i)$$

Rearranging:

$$\sum_{i=1}^{n} h_i' \Delta x_i = h(x(t)) - h(x(0)) - D \sum_{i=1}^{n} h'' \Delta t_i$$

In the limit $n \to \infty$, the sums become integrals:

$$\int_0^t h' \,\mathrm{d}x(\tau) = h(x(t)) - h(x(0)) - D \int_0^t h'' \,\mathrm{d}\tau \tag{3.44}$$

We can finally apply the result (3.44) to our case, by setting $h'(x(\tau)) = x(\tau)$, so that:

$$h(x(t)) = \int x(\tau) = \frac{x(t)^2}{2}; \qquad h''(x(\tau)) = 1$$

Substituting in (3.44) leads to:

$$\sum_{i=1}^{n} x_{i-1}\Delta x_i \xrightarrow[n\to\infty]{} \int_0^t x(\tau)\,\mathrm{d}x(\tau) = \frac{x^2(t) - x^2(0)}{2} - D \underbrace{\int_0^t \mathrm{d}\tau}_{t} = \frac{x^2(t) - x^2(0)}{2} - Dt$$

And substituting this result back in (3.42) leads to:

$$\mathrm{d}P \underset{n\to\infty}{=} \mathrm{d}x_W \exp\left(-\frac{k}{2D}\left[\frac{x_t^2 - x_0^2}{2} - Dt\right]\right) \exp\left(-\frac{k^2}{4D}\int_0^t x^2(\tau)\,\mathrm{d}\tau\right)$$

From this expression we can compute *transition probabilities*. Let $T = [0,t]$ and $\mathbb{R}^T$ be the space of continuous functions $T \to \mathbb{R}$, then:

$$
\begin{aligned}
W(x_t, t|x_0, 0) = \langle \delta(x_t - x) \rangle_W = \int_{\mathbb{R}^T} \delta(x_t - x) \, \mathrm{d}P = \\
&= \int_{\mathbb{R}^T} \mathrm{d}x_W \, \delta(x(t) - x) \exp\left(-\frac{k}{2D}\left[\frac{x_t^2 - x_0^2}{2} - Dt\right]\right) \exp\left(-\frac{k^2}{4D}\int_0^t x^2(\tau)\,\mathrm{d}\tau\right) = \\
&= \exp\left(-\frac{k}{2D}\left[\frac{x_t^2 - x_0^2}{2} - Dt\right]\right) \underbrace{\int_{\mathbb{R}^T} \mathrm{d}x_W \, \delta(x(t) - x) \exp\left(-\frac{k^2}{4D}\int_0^t x^2(\tau)\,\mathrm{d}\tau\right)}_{\text{CFR } I_4 \text{ on } 28/10} = \\
&= \exp\left(-\frac{k}{2D}\left[\frac{x_t^2 - x_0^2}{2} - Dt\right]\right) \sqrt{\frac{k}{4\pi D \sinh(kt)}} \exp\left(-\frac{k x_t^2}{4D} \coth(kt)\right)
\end{aligned}
\tag{3.45}
$$

Exercise 3.8.1 (Some more integrals)**:**

> Check that:
>
> $$W(x, 0|x_0, 0) = \delta(x - x_0)$$
>
> **Hint**. Start from the case $x_0 = 0$. Using (3.45), after some algebra:
>
> $$W(x, t|0, 0) = \sqrt{\frac{k}{2\pi D(1 - e^{-2kt})}} \exp\left(-\frac{k}{2D}\frac{x^2}{1 - e^{-2kt}}\right) \tag{3.46}$$
>
> And then show $W(x, t|0, 0) \xrightarrow[t \to 0]{} \delta(x)$. The general case follows by translating that solution.

Alternative derivation The same result can be found solving the Fokker-Planck equation for the transition probabilities $W(x, t|x_0, 0)$:

$$\dot{W}(x, t|x_0, 0) = \frac{\partial}{\partial x}\left(kxW + D\frac{\partial}{\partial x}W\right) \tag{3.47}$$

A quick way to solve this differential equation is to note that $\{\Delta B_i\}$ are all i.i.d. gaussian variables, and so x, which is a sum of ΔB_i must have a *gaussian* pdf. So we can make an *ansatz* for the solution:

$$W(x, t|x_0, 0) = \frac{1}{Z(t)} \exp\left(-a(t)x^2 + b(t)x\right) \tag{3.48}$$

Where $a(t)$ and $b(t)$ are the gaussian parameters, and $Z(t)$ the normalization factor. All that's left is to substitute (3.48) in (3.47) and solve for a, b, Z.

3.8.1 Equilibrium distribution

As before, we expect the equilibrium distribution to follow Maxwell-Boltzmann formula:

$$W_{\text{eq}}(x) = \frac{1}{Z}\exp(-\beta V(x)) = \frac{1}{Z}\exp\left(-\frac{m\omega^2 x^2}{2k_B T}\right) \qquad Z = \int_{\mathbb{R}} \exp(-\beta V(x)) \tag{3.49}$$

Starting from (3.46) and taking the limit $t \to \infty$:

$$\lim_{t\to\infty} W(x,t|0,0) = \sqrt{\frac{k}{2\pi D}}\exp\left(-\frac{k}{2D}x^2\right) \tag{3.50}$$

Comparing (3.49) with (3.50) we find:

$$\frac{m\omega^2}{2k_B T} = \frac{k}{2D} = \frac{m\omega^2}{2\gamma D} \Rightarrow k_B T = \gamma D$$

So we obtain the same relation between D and T that we found in the general case.

3.8.2 High dimensional generalization

We can generalize the previous results to the case where $\boldsymbol{\Delta B_i} = (\Delta B_i^1, \dots, \Delta B_i^d)^T$ are d-dimensional vectors, following a *multivariate gaussian distribution*:

$$\mathbb{P}(\boldsymbol{\Delta B_1}, \dots, \boldsymbol{\Delta B_n}) = \prod_{i=1}^n \prod_{\alpha=1}^d \frac{\mathrm{d}B_i^\alpha}{\sqrt{2\pi\Delta t_i}} \exp\left(-\frac{\Delta B_i^\alpha}{2\Delta t_i}\right)$$

As different components of the same $\boldsymbol{\Delta B_i}$ are independent, by Ito's rules of integration:

$$\mathrm{d}B_i^\alpha\,\mathrm{d}B_i^\beta = \delta_{\alpha\beta}\,\mathrm{d}t_i \qquad \mathrm{d}B_i^\alpha\,\mathrm{d}B_i^\beta\,\mathrm{d}B_i^\gamma = 0$$

We then need to write d different Langevin equations, one for each component:

$$\mathrm{d}x^\alpha(t) = f^\alpha(x(t),t)\,\mathrm{d}t + \sqrt{2D_\alpha(x(t),t)}\,\mathrm{d}B^\alpha(t)$$

More in general, the stochastic term could be:

$$\sum_{\beta=1}^d g_{\alpha\beta}(x(t),t)\,\mathrm{d}B^\beta(t)$$

and in our case $g_{\alpha\beta} = 2\sqrt{2D_\alpha}\delta_{\alpha\beta}$.

The Fokker-Planck equation then becomes:

$$\dot{W}(\boldsymbol{x},t) = \sum_{\alpha=1}^d \frac{\partial}{\partial x^\alpha}\left(-f_\alpha(\boldsymbol{x},t)W(\boldsymbol{x},t) + \frac{\partial}{\partial x^\alpha}D_\alpha(\boldsymbol{x},t)W(\boldsymbol{x},t)\right)$$

And the joint probability for a *discretized* path:

$$\mathbb{P}(\boldsymbol{\Delta x_1}, \dots, \boldsymbol{\Delta x_n}) = \prod_{i=1}^n \prod_{\alpha=1}^d \frac{\mathrm{d}\Delta x_i^\alpha}{\sqrt{4\pi D_\alpha \Delta t_i}} \exp\left(-\sum_{i=1}^n \sum_{\alpha=1}^d \frac{(\Delta x_i^\alpha - f_{i-1}^\alpha \Delta t_i)^2}{4D_\alpha \Delta t_i}\right)$$

And taking the limit $n \to \infty$:

$$\mathbb{P}(\boldsymbol{x}(\tau)) = \prod_{\tau=0^+}^t \left(\frac{\mathrm{d}^d\boldsymbol{x}(\tau)}{\sqrt{4\pi\,\mathrm{d}\tau}\prod_{\alpha=1}^d\sqrt{D_\alpha}}\right)\exp\left(-\sum_{\alpha=1}^d \frac{1}{4D_\alpha}\int_0^t (\dot{x}^\alpha - f^\alpha)^2\,\mathrm{d}\tau\right)$$

3.8.3 Underdamped Harmonic Oscillator

If we do not ignore the inertia term in (3.38) we are left with:

$$m\ddot{\boldsymbol{x}} = m\dot{\boldsymbol{v}} = -\gamma\dot{\boldsymbol{x}} + \boldsymbol{F}(\boldsymbol{x}) + \sqrt{2D}\boldsymbol{\xi}$$

This second order (stochastic) differential equation can be written as a system of two first order equations:

$$\begin{cases} \mathrm{d}\boldsymbol{x} = \boldsymbol{v}\,\mathrm{d}t \\ \mathrm{d}\boldsymbol{v} = \left(-\frac{\gamma}{m}\boldsymbol{v} + \frac{\boldsymbol{F}(\boldsymbol{x})}{m}\right)\mathrm{d}t + \frac{\sqrt{2D}}{m}\,\mathrm{d}\boldsymbol{B} \end{cases}$$

This leads to a *generalization* of the Fokker-Planck equation, named **Kramer equation**:

$$\dot{W}(\boldsymbol{x},\boldsymbol{v},t) = \boldsymbol{\nabla}_{\boldsymbol{v}}\left[\left(\frac{\gamma\boldsymbol{v}}{m} - \frac{\boldsymbol{F}}{m}\right)W(\boldsymbol{x},\boldsymbol{v},t) + \frac{\gamma^2 D}{m^2}\boldsymbol{\nabla}_{\boldsymbol{v}}W(\boldsymbol{x},\boldsymbol{v},t)\right] + \boldsymbol{\nabla}_{\boldsymbol{x}}(-\boldsymbol{v}W(\boldsymbol{x},\boldsymbol{v},t))$$

In the limit $t \to \infty$, the distribution at equilibrium will be:

$$W(\boldsymbol{x},\boldsymbol{v}) = \frac{1}{Z}\exp\left(-\beta\left[\frac{m\|\boldsymbol{v}\|^2}{2} + V(\boldsymbol{x})\right]\right) \qquad D = \frac{k_B T}{\gamma}$$

(Lesson 14 of 18/11/19) Compiled: October 13, 2020

3.9 Particle in a conservative force-field

In last section, we examined a particle of radius a immersed in a harmonic potential $U(x) = m\omega^2 x^2/2$, moving through a medium with viscosity η and subject to thermal fluctuations of amplitude proportional to $\sqrt{2D}$, so that its dynamics are described by the following *stochastic differential equation*:

$$\mathrm{d}x = -kx\,\mathrm{d}t + \sqrt{2D}\,\mathrm{d}B \qquad k = \frac{m\omega^2}{\gamma} \qquad \gamma = 6\pi\eta a$$

The solution, expressed as the *transition probability* between any two given points, is a path integral:

$$W(x_t, t|x_0, 0) = \exp\left(-\frac{x_t^2 - x_0^2}{4D}k + \frac{kt}{2}\right)\left\langle \exp\left(-\int_0^t V(x(\tau))\,\mathrm{d}\tau\right)\delta(x(t) - x_t)\right\rangle_W \tag{3.51}$$

with $V(x(\tau)) = k^2x^2(\tau)/(4D)$. The average is computed with the Wiener measure:

$$\langle f(x(\tau))\rangle_W \equiv \int_{\mathbb{R}^T}\left(\prod_{\tau=0^+}^{t}\frac{\mathrm{d}x(\tau)}{\sqrt{4\pi D\,\mathrm{d}\tau}}\right)\exp\left(-\frac{1}{4D}\int_0^t \dot{x}^2(\tau)\,\mathrm{d}\tau\right)f(x(\tau))$$

with $\mathbb{R}^T$ being the set of continuous functions $T \to \mathbb{R}$, and $T = [0, t]$.

We want now to show that, in the more general case of a particle immersed in a generic potential $U(x)$, a path integral similar to the highlighted term in (3.51) will appear. Of course, the function $V(x(\tau))$ will be different, but it will be proportional to $U(x)$ - as it is evident in the harmonic case.

So, let's consider a particle in a 3D space $\boldsymbol{r} = (x_1, x_2, x_3)^T$, immersed in a conservative force-field $\boldsymbol{F}(\boldsymbol{r}) = -\boldsymbol{\nabla} U(\boldsymbol{r})$ with potential $U(\boldsymbol{r})$, and subject to thermal noise. The Langevin equation becomes:

$$\mathrm{d}\boldsymbol{r} = \boldsymbol{f}(\boldsymbol{r})\,\mathrm{d}t + \sqrt{2D}\,\mathrm{d}\boldsymbol{B} \qquad \boldsymbol{f}(\boldsymbol{r}) = \frac{\boldsymbol{F}(\boldsymbol{r})}{\gamma} \qquad \gamma = 6\pi\eta a \tag{3.52}$$

with $\boldsymbol{B} = (B_1, B_2, B_3)^T$ being a $d = 3$ vector with gaussian components:

$$\Delta B_\alpha \sim \frac{1}{\sqrt{2\pi\Delta t}} \exp\left(-\frac{\Delta B_\alpha^2}{2\Delta t}\right) \qquad \alpha = 1, 2, 3 \tag{3.53}$$

As different components are independent, the joint pdf for the vector $\boldsymbol{\Delta B}$ is just the product of the three terms in (3.53):

$$\boldsymbol{\Delta B} \sim \frac{1}{(2\pi\Delta t)^{3/2}} \exp\left(-\frac{\|\boldsymbol{\Delta B}\|^2}{2\Delta t}\right)$$

As before, we introduce a time discretization $\{t_j\}_{j=0,\dots,n}$ with $t_0 \equiv 0$ and $t_n \equiv t$ fixed, so that (3.52) becomes:

$$\boldsymbol{\Delta r_i} = \boldsymbol{r}(t_i) - \boldsymbol{r}(t_{i-1}) = \boldsymbol{f_{i-1}}\Delta t_i + \sqrt{2D}\boldsymbol{\Delta B_i}$$

where the force $\boldsymbol{f}(\boldsymbol{r})$ is evaluated at the left side t_{i-1} of each discrete interval $[t_{i-1}, t_i]$, following Ito's prescription.
Then, starting from the joint pdf of the $\{\boldsymbol{\Delta B_i}\}$:

$$\mathrm{d}P\left(\boldsymbol{\Delta B}_1, \dots, \boldsymbol{\Delta B}_n\right) = \prod_{i=1}^n \frac{\mathrm{d}^3\boldsymbol{\Delta B_i}}{(2\pi\Delta t)^{3/2}} \exp\left(-\sum_{i=1}^n \frac{\|\boldsymbol{\Delta B_i}\|^2}{2\Delta t_i}\right)$$

we perform a change of variables by inverting (3.52):

$$\boldsymbol{\Delta B_i} = \frac{\boldsymbol{\Delta r_i} - \boldsymbol{f_{i-1}}\Delta t_i}{\sqrt{2D}} \Rightarrow \left|\frac{\partial\{\Delta B_i^\alpha\}}{\partial\{\Delta r_j^\beta\}}\right| = \left|\frac{\partial\{\Delta r_j^\beta\}}{\partial\{\Delta B_i^\alpha\}}\right|^{-1} = (2D)^{3/2}$$

This leads to the joint pdf for the increments $\{\boldsymbol{\Delta r_i}\}$:

$$\mathrm{d}P\left(\boldsymbol{\Delta r_1}, \dots, \boldsymbol{\Delta r_n}\right) = \left(\prod_{i=1}^n \frac{\mathrm{d}^3\boldsymbol{\Delta r_i}}{(4\pi D\Delta t_i)^{3/2}}\right) \exp\left[-\frac{1}{4D}\sum_{i=1}^n \frac{\left\|\boldsymbol{\Delta r_i} - \boldsymbol{f_{i-1}}\Delta t_i\right\|^2}{\Delta t_i}\right] \tag{3.54}$$

Expanding the square in the exponential:

$$-\frac{1}{4D}\sum_{i=1}^n \left[\frac{\|\boldsymbol{\Delta r_i}\|^2}{\Delta t_i} + \left\|\boldsymbol{f_{i-1}}\right\|^2 \Delta t_i - 2\boldsymbol{\Delta r_i}\cdot\boldsymbol{f}_{i-1}\right]$$

allows to recognize the $d=3$ Wiener measure in (3.54):

$$\mathrm{d}P\left(\{\boldsymbol{\Delta r_i}\}\right) = \Bigg(\underbrace{\prod_{i=1}^{n}\frac{\mathrm{d}^3\boldsymbol{\Delta r_i}}{(4\pi D\Delta t_i)^{3/2}}\exp\left[-\frac{1}{4D}\sum_{i=1}^{n}\frac{\|\boldsymbol{\Delta r_i}\|^2}{\Delta t_i}\right]}_{\mathrm{d}^3_W\boldsymbol{r}}\Bigg)\cdot$$
$$\cdot\exp\Bigg(-\frac{1}{4D}\underbrace{\sum_{i=1}^{n}\left\|\boldsymbol{f_{i-1}}\right\|^2\Delta t_i}_{\int_0^t\|\boldsymbol{f}(\boldsymbol{r}(\tau))\|^2\,\mathrm{d}\tau}+\frac{1}{2D}\underbrace{\sum_{i=1}^{n}\boldsymbol{f_{i-1}}\cdot\boldsymbol{\Delta r_i}}_{\int_0^t\boldsymbol{f}(\boldsymbol{r}(\tau))\cdot\mathrm{d}\boldsymbol{r}(\tau)}\Bigg) \tag{3.55}$$

Let's focus on the *stochastic integral* (the one in $\mathrm{d}\boldsymbol{r}(\tau)$). For this we need to generalize to $d=3$ the integration formula we found in the previous section. Consider a multi-variable scalar function $h(\boldsymbol{r})\colon\mathbb{R}^3\to\mathbb{R}$, $\boldsymbol{r}\mapsto h(\boldsymbol{r})$. As before, we start from the difference:

$$\begin{aligned}h(\boldsymbol{r}_n)-h(\boldsymbol{r}_0) &= h(\boldsymbol{r}_n)-h(\boldsymbol{r}_{n-1})+h(\boldsymbol{r}_{n-1})-h(\boldsymbol{r}_{n-2})+\cdots+h(\boldsymbol{r}_1)-h(\boldsymbol{r}_0)=\\ &=\sum_{i=1}^{n}(h(\boldsymbol{r}_i)-h(\boldsymbol{r}_{i-1}))=\sum_{i=1}^{n}\Delta h_i\end{aligned} \tag{3.56}$$

In the discretization, $\boldsymbol{r}_i=\boldsymbol{r}_{i-1}+\boldsymbol{\Delta x}$, with $\boldsymbol{\Delta x}=(\Delta x_i^1,\Delta x_i^2,\Delta x_i^3)$. Each differential Δh_i is then:

$$\begin{aligned}\Delta h_i &= h(\boldsymbol{r}_i)-h(\boldsymbol{r}_{i-1})\equiv h_i-h_{i-1}=\\ &\underset{(a)}{=}\cancel{h(\boldsymbol{r}_{i-1})}+\sum_{\alpha=1}^{3}\left[\frac{\partial}{\partial x^\alpha}h(\boldsymbol{r}_{i-1})\right]\Delta x_i^\alpha+\frac{1}{2}\sum_{\alpha,\beta=1}^{3}\left[\frac{\partial^2}{\partial x^\alpha\partial x^\beta}h(\boldsymbol{r}_{i-1})\right]\Delta x_i^\alpha\Delta x_i^\beta+\cdots-\cancel{h(\boldsymbol{r}_{i-1})}=\\ &\underset{(b)}{=}\sum_{\alpha=1}^{3}\left[\frac{\partial}{\partial x^\alpha}h(\boldsymbol{r}_{i-1})\right]\Delta x_i^\alpha+D\sum_{\alpha,\beta=1}^{3}\left[\frac{\partial^2}{\partial x^\alpha\partial x^\beta}h(\boldsymbol{r}_{i-1})\right]\Delta t_i\end{aligned} \tag{3.57}$$

where in (a) we expanded the first term in Taylor series about $\boldsymbol{r}_{i-1}$, and in (b) we used Ito's rules, and in particular the fact that:

$$\Delta x_i^\alpha\Delta x_i^\beta=\Delta t_i 2D\delta_{\alpha\beta}$$

Substituting (3.57) back in (3.56) leads to:

$$h(\boldsymbol{r}_n)-h(\boldsymbol{r}_0)=\sum_{i=1}^{n}\Delta h_i=\sum_{i=1}^{n}\sum_{\alpha=1}^{3}\frac{\partial}{\partial x^\alpha}h_{i-1}\Delta x_i^\alpha+D\sum_{\alpha=1}^{3}\frac{\partial^2}{\partial x^{\alpha 2}}h_{i-1}\Delta t_i$$

and then, in the continuum limit:

$$h(\boldsymbol{r}(t))-h(\boldsymbol{r}(0))=\int_0^t\boldsymbol{\nabla}h(\boldsymbol{r})\cdot\mathrm{d}^3\boldsymbol{r}+D\int_0^t\nabla^2 h(\boldsymbol{r}(\tau))\,\mathrm{d}\tau$$

Rearranging we arrive at the desired formula for integration:

$$\int_0^t\boldsymbol{\nabla}h(\boldsymbol{r})\cdot\mathrm{d}^3\boldsymbol{r}=h(\boldsymbol{r}(t))-h(\boldsymbol{r}(0))-D\int_0^t\nabla^2 h(\boldsymbol{r}(\tau))\,\mathrm{d}\tau \tag{3.58}$$

Thanks to (3.58) we can solve the stochastic integral in (3.55):

$$\int_0^t \boldsymbol{f}(\boldsymbol{r}(\tau)) \cdot \mathrm{d}^3\boldsymbol{r}(\tau)$$

Inserting $\boldsymbol{f}(\boldsymbol{r}) = -\boldsymbol{\nabla} U(\boldsymbol{r})/\gamma$ and applying the formula leads to:

$$\int_0^t \boldsymbol{f}(\boldsymbol{r}(\tau)) \cdot \mathrm{d}^3\boldsymbol{r}(\tau) = -\frac{1}{\gamma}\int_0^t \boldsymbol{\nabla} U(\boldsymbol{r}(\tau)) \cdot \mathrm{d}^3\boldsymbol{r}(\tau) = -\frac{1}{\gamma}\left[U(\boldsymbol{r}(t)) - U(\boldsymbol{r}(0)) - D\int_0^t \nabla^2 U(\boldsymbol{r}(\tau))\,\mathrm{d}\tau\right]$$

Substituting back in (3.55):

$$\begin{aligned}\mathrm{d}P &= \mathrm{d}^3_W\boldsymbol{r}\exp\left(-\frac{1}{4D}\int_0^t \|\boldsymbol{f}\|^2\,\mathrm{d}\tau + \frac{1}{2D}\left[-\frac{1}{\gamma}[U(\boldsymbol{r}(t)) - U(\boldsymbol{r}(0))] + \frac{D}{\gamma}\int_0^t \nabla^2 U(\boldsymbol{r}(\tau))\,\mathrm{d}\tau\right]\right) = \\ &= \mathrm{d}^3_W\boldsymbol{r}\exp\left(-\frac{1}{4D}\int_0^t \mathrm{d}\tau\,\underbrace{\left[\|\boldsymbol{f}\|^2 - \frac{2D}{\gamma}\nabla^2 U\right]}_{V(\boldsymbol{r})}\right)\exp\left(-\frac{1}{2D\gamma}[U(\boldsymbol{r}(t)) - U(\boldsymbol{r}(0))]\right)\end{aligned} \tag{3.59}$$

where:

$$V = \boldsymbol{f}^2 - \frac{2D}{\gamma}\nabla^2 U = \boldsymbol{f}^2 - 2D\boldsymbol{\nabla}\cdot\boldsymbol{f}$$

Using the just found measure $\mathrm{d}P$ we can compute path integrals, and in particular *transition probabilities*:

$$\begin{aligned}W(\boldsymbol{r}, t|\boldsymbol{r}_0, 0) &= \int_{\mathbb{R}^T} \mathrm{d}P\,\delta(\boldsymbol{r}(t) - \boldsymbol{r}) \equiv \langle\delta(\boldsymbol{r}(t) - \boldsymbol{r}(0))\rangle_W = \\ &= \int_{\mathbb{R}^T} \mathrm{d}^3_W\boldsymbol{r}\exp\left(-\frac{1}{4D}\int_0^t V(\boldsymbol{r}(\tau))\,\mathrm{d}\tau\right)\delta(\boldsymbol{r}(t) - \boldsymbol{r})\exp\left(-\frac{1}{2D\gamma}(U(\boldsymbol{r}) - U(\boldsymbol{r}_0))\right) = \\ &= \langle\exp\left(-\frac{1}{4D}\int_0^t V(\boldsymbol{r}(\tau))\,\mathrm{d}\tau\right)\delta(\boldsymbol{r}(t) - \boldsymbol{r})\rangle_W\exp\left(-\frac{1}{2D\gamma}(U(\boldsymbol{r}) - U(\boldsymbol{r}_0))\right)\end{aligned}$$

This expression is indeed similar to that derived in the specific case of the harmonic oscillator (3.51), meaning that the techniques we used to evaluate previous path integrals can be useful in much more general cases.

This observation has indeed a deeper meaning, as we found a way to describe the dynamics of conservative systems with a path integral. We already know that the behaviour of these systems can be also described with partial differential equations (e.g. the Fokker-Planck equation). So, there should be a link between path integrals and PDEs, that will be explored in the next section.

3.10 Feynman-Kac formula

It is possible to use the machinery of stochastic processes and path integrals to solve certain partial differential equations, which - as we will see - are of fundamental importance in Quantum Mechanics.
In this regard, a very important result is offered by the **Feynman-Kac formula**. The main idea is to use a Brownian process to *simulate many paths*,

and express the solution of the differential equation as the *average* of a certain functional computed over all these paths.
More precisely, consider the following *partial differential equation* (Bloch's equation):

$$\partial_t W_B(x,t) = D\partial_x^2 W_B(x,t) - V(x)W_B(x,t) \qquad D \in \mathbb{R}, V\colon \mathbb{R} \to \mathbb{R} \tag{3.60}$$

The **Feynman-Kac** formula states that the function $W_B(x,t)$ that solves (3.60) can be found by computing a Wiener path integral:

$$W_B(x,t) \equiv \langle \exp\left(-\int_0^t V(x(\tau))\,\mathrm{d}\tau\right)\delta(x(t)-x)\rangle_W \tag{3.61}$$

Note that this result can be generalized to more dimensions - but we will limit ourselves to the $d=1$ case for simplicity.

Proof. We now show that (3.61) indeed satisfies (3.60). As usual, we start by defining a time discretization, $\{t_i\}_{i=0,\dots,n+1}$, so that $t_0 \equiv 0$ and $t_{n+1} \equiv \bar{t}$ is the instant at which we wish to evaluate the solution $W_B(x,t)$. Then (3.61) at that instant will be obtained by the *continuum limit* of the discretized average $\psi_{n+1}(x)$:

$$\begin{aligned} W_B(x,\bar{t}) &= \lim_{n\to\infty} \psi_{n+1}(x) \\ \psi_{n+1}(x) &= \int_{\mathbb{R}^{n+1}} \left(\prod_{i=1}^{n+1} \frac{\mathrm{d}x_i}{\sqrt{4\pi D\Delta t_i}}\right) \exp\left(-\sum_{i=1}^{n+1} \frac{(x_i - x_{i-1})^2}{4D\Delta t_i} - \sum_{i=1}^{n+1} \Delta t_i V(x_i)\right)\delta(x_{n+1}-x) \end{aligned} \tag{3.62}$$

Note that $\psi_{n+1}(x)$ is the average of a functional over *all paths* that arrive in x at the instant $\bar{t}$, making exactly $n+1$ steps from their starting point 0. In the following, the intuition is to see these paths as *being generated*, i.e. evolving *step* after *step* from 0 to x. For example, suppose we want to approximate (3.62). We would start by choosing an *ensemble* of paths arriving to x after $n+1$ timesteps, compute the functional on each of them, and average the results. However, we could also do it in another - a bit stranger - way. Consider the same ensemble of paths we already (supposedly) generated. From each of them, remove the last step. We now have a set of paths that arrive *close* to x, and will arrive exactly there if we let another timestep pass. However, we decide to compute the functional on each of these paths and then average the results, before letting them arrive at their destination. So, in a certain sense, we will estimate the value of the functional "a timestep in the past". Of course, we can repeat this process, removing more and more timesteps at every iteration. At the end, we will have a sequence of numbers detailing the "evolution" of the functional from the start to the end. Turns out that the *rule* for such an evolution is exactly (3.60). So, to prove Feynman-Kac, we just have to find that rule - meaning how to relate ψ_{n+1} to its "past" ψ_n (in the continuum limit $n \to \infty$).
This is just an informal intuition, that will only be useful as a guide for the rest of the proof. So, let's go on.

For simplicity, we choose the time discretization as **uniform**, so that $\Delta t_i \equiv \epsilon$ $\forall i$, and:

$$t_{n+1} = (n+1)\epsilon \equiv \bar{t} \Rightarrow \epsilon = \frac{\bar{t}}{n+1}$$

We then rewrite (3.62), highlight the last term (the one with the x_{n+1}) and integrate to remove the δ:

$$\begin{aligned}
\psi_{n+1}(x) &= \int_{\mathbb{R}^n} \left(\prod_{i=1}^{n} \frac{\mathrm{d}x_i}{\sqrt{4\pi D\epsilon}} \right) \exp\left(-\sum_{i=1}^{n} \frac{(x_i - x_{i-1})^2}{4D\epsilon} - \sum_{i=1}^{n} \epsilon V(x_i) \right) \cdot \\
&\quad \cdot \int_{\mathbb{R}} \frac{\mathrm{d}x_{n+1}}{\sqrt{4\pi D\epsilon}} \exp\left(-\frac{(x_{n+1} - x_n)^2}{4D\epsilon} - \epsilon V(x_{n+1}) \right) \delta(x_{n+1} - x) = \\
&= \int_{\mathbb{R}^n} \left(\prod_{i=1}^{n} \frac{\mathrm{d}x_i}{\sqrt{4\pi D\epsilon}} \right) \exp\left(-\sum_{i=1}^{n} \frac{(x_i - x_{i-1})^2}{4D\epsilon} - \sum_{i=1}^{n} \epsilon V(x_i) \right) \cdot \\
&\quad \cdot \frac{1}{\sqrt{4\pi D\epsilon}} \exp\left(-\frac{(x - x_n)^2}{4D\epsilon} - \epsilon V(x) \right)
\end{aligned} \tag{3.63}$$

Now, with some algebra, we recognize in these integrals a term $\psi_n(x_n)$, indicating the expected value of the functional over all paths reaching x_n (which is close to the end-point x) at timestep $t_n = t - \epsilon$. We start by rearranging, putting the integration over $\mathrm{d}x_n$ at the front:

$$\begin{aligned}
(3.63) &= \int_{\mathbb{R}} \frac{\mathrm{d}x_n}{\sqrt{4\pi D\epsilon}} \exp\left(-\frac{(x - x_n)^2}{4D\epsilon} - \epsilon V(x) \right) \cdot \\
&\quad \cdot \frac{1}{\sqrt{4\pi D\epsilon}} \int_{\mathbb{R}^{n-1}} \left(\prod_{i=1}^{n-1} \frac{\mathrm{d}x_i}{\sqrt{4\pi D\epsilon}} \right) \exp\left(-\sum_{i=1}^{n} \frac{(x_i - x_{i-1})^2}{4D\epsilon} - \epsilon \sum_{i=1}^{n} V(x_i) \right)
\end{aligned} \tag{3.64}$$

Now we *change* all the x_i in the second line to y_i, and then add a δ (with its integral) to *connect* y_n to x_n, which appears in the integral in the first line. In this way we will highlight the desired $\psi_n(x_n)$:

$$\begin{aligned}
(3.64) &= \int_{\mathbb{R}} \frac{\mathrm{d}x_n}{\sqrt{4\pi D\epsilon}} \exp\left(-\frac{(x - x_n)^2}{4D\epsilon} - \epsilon V(x) \right) \cdot \\
&\quad \underbrace{\cdot \int_{\mathbb{R}^n} \left(\prod_{i=1}^{n} \frac{\mathrm{d}y_i}{\sqrt{4\pi D\epsilon}} \right) \exp\left(-\sum_{i=1}^{n} \frac{(y_i - y_{i-1})^2}{4D\epsilon} - \epsilon \sum_{i=1}^{n} V(y_i) \right) \delta(x_n - y_n)}_{\psi_n(x_n)}
\end{aligned} \tag{3.65}$$

And so:

$$\psi_{n+1}(x) = e^{-\epsilon V(x)} \int_{\mathbb{R}} \frac{\mathrm{d}x_n}{\sqrt{4\pi D\epsilon}} \exp\left(-\frac{(x - x_n)^2}{4D\epsilon} \right) \psi_n(x_n) \tag{3.66}$$

This is relation between $\psi_{n+1}(x)$ and its "past" $\psi_n(x)$ that we were searching for. Now, to retrieve (3.60), all that's left is to put (3.66) in differential form.

We start by simplifying the integral, making it similar to a gaussian with a change of variables:

$$-\frac{(x-x_n)^2}{4D\epsilon} \stackrel{!}{=} -\frac{z^2}{2} \Rightarrow z = -\frac{x-x_n}{\sqrt{2D\epsilon}}; \quad x_n = x + \sqrt{2D\epsilon}z; \quad \mathrm{d}x_n = \sqrt{2D\epsilon}\,\mathrm{d}z$$

which leads to:

$$\psi_{n+1}(x) = e^{-\epsilon V(x)} \int_{-\infty}^{+\infty} \frac{\mathrm{d}z}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \psi_n(x_n + z\sqrt{2D\epsilon}) \tag{3.67}$$

As $n \to \infty$, $z \to 0$. We will not prove this, but note that:

$$z = -\frac{x-x_n}{\epsilon}\frac{\sqrt{\epsilon}}{\sqrt{2D}}$$

where the first factor is a *velocity*, which must be physically limited, and so $n \to \infty \Rightarrow \epsilon \to 0 \Rightarrow z \to 0$.
This means that we can expand ψ_n in Taylor series about x:

$$\psi_n(x_n + z\sqrt{2D\epsilon}) = \psi_n(x) + z\sqrt{2D\epsilon}\psi_n'(x) + z^2 D\epsilon\psi_n''(x) + O(z^3\epsilon^{3/2})$$

Substituting back in (3.67):

$$\psi_{n+1}(x) = e^{-\epsilon V(x)} \Bigg[\psi_n(x) \underbrace{\int_{\mathbb{R}} \frac{\mathrm{d}z}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)}_{1} + \sqrt{2D\epsilon}\psi_n'(x) \underbrace{\int_{\mathbb{R}} \frac{\mathrm{d}z}{\sqrt{2\pi}} z \exp\left(-\frac{z^2}{2}\right)}_{0} + \\ + D\epsilon\psi_n''(x) \underbrace{\int_{\mathbb{R}} \frac{\mathrm{d}z}{\sqrt{2\pi}} z^2 \exp\left(-\frac{z^2}{2}\right)}_{1} + O(\epsilon^2)\Bigg] \tag{3.68}$$

as the integrand is just a standard gaussian ($\mu = 0$, $\sigma = 1$). Note how the error term is of order ϵ^2, as the first non-null integral in the series will be that with z^4:

$$\int_{\mathbb{R}} \frac{\mathrm{d}z}{\sqrt{2\pi}} z^k \frac{(2D\epsilon)^{k/2}}{k!} \psi_n^{(k)} = \begin{cases} 0 & k \text{ odd} \\ 1^k(k-1)!! & k \text{ even} \end{cases}$$

Expanding also the $e^{-\epsilon V(x)}$ term:

$$e^{-\epsilon V(x)} = 1 - \epsilon V(x) + \frac{\epsilon^2 V^2(x)}{2} + O(\epsilon^3)$$

Finally, substituting back in (3.68), expanding the product and ignoring all terms of order 2 or higher in ϵ:

$$\psi_{n+1}(x) = \psi_n(x) + D\epsilon\psi_n''(x) - \epsilon V(x)\psi_n(x) + O(\epsilon^2)$$

Rearranging:

$$\frac{\psi_{n+1} - \psi_n}{\epsilon} = D\psi_n'' - V\psi_n$$

And when $\epsilon \to 0$ the first term becomes a time derivative, leading to Bloch's equation, and proving Feynman-Kac formula:

$$\partial_t W_B(x,t) = D\partial_x^2 W_B(x,t) - V(x) W_B(x,t)$$

3.10.1 Application to Quantum Mechanics

It is possible to map the Schrödinger equation to the Bloch equation (3.60), and then use Feynman-Kac formula to solve it.
Recall the time-dependent Schrödinger equation for a particle immersed in a $d = 1$ potential $v(x)$ and described by a wavefunction $\psi(x,t)$:

$$i\hbar\partial_t\psi(x,t) = -\frac{\hbar^2}{2m}\partial_x^2\psi(x,t) + v(x)\psi(x,t)$$

This is already similar to (3.60), except for the presence of a complex coefficient i. We can remove it with a change of variable $t \mapsto it$, leading to:

$$i\hbar(i)\frac{\partial}{\partial t}\psi(x,it) = -\hbar\frac{\partial}{\partial t}\psi(x,it) = -\frac{\hbar^2}{2m}\partial_x^2\psi(x,it) + v(x)\psi(x,it)$$

Defining $\psi(x,it) \equiv \hat{\psi}(x,t)$ and multiplying both sides by $-\hbar^{-1}$ leads to:

$$\frac{\partial}{\partial t}\hat{\psi}(x,t) = \frac{\hbar}{2m}\frac{\partial^2}{\partial x^2}\hat{\psi}(x,t) - \frac{v(x)}{\hbar}\hat{\psi}(x,t)$$

which has the form of Bloch's equation:

$$\frac{\partial}{\partial t}\hat{\psi}(x,t) = D\frac{\partial^2}{\partial x^2}\hat{\psi}(x,t) - V(x)\hat{\psi}(x,t) \qquad D = \frac{\hbar}{2m}; \quad V(x) = \frac{v(x)}{\hbar}$$

3.11 Variational methods

Consider a particle subject to an external conservative force $\boldsymbol{F}(\boldsymbol{r}) = -\boldsymbol{\nabla}U(\boldsymbol{r})$, moving through a viscous medium and subject to thermal noise. The probability density for a path $x(\tau)$ can be derived from (3.59), after "dividing by the volume element" and taking the limit $n \to \infty$:

$$\begin{aligned}\Omega[\boldsymbol{r}(\tau)] \equiv \frac{\mathrm{d}P}{\mathrm{d}V} &= \\ &= \exp\left(-\frac{1}{4D}\int_0^t \mathrm{d}\tau\, \dot{\boldsymbol{r}}^2(\tau) + \int_0^t \mathrm{d}\tau \underbrace{\left(-\frac{1}{4D}\right)\left[\|\boldsymbol{f}\|^2 - \frac{2D}{\gamma}\nabla^2 U\right]}_{V(\boldsymbol{r})} - \frac{1}{2D\gamma}[U(\boldsymbol{r}(t)) - U(\boldsymbol{r}(0))]\right) = \\ &= \exp\left(-\frac{1}{4D}\int_0^t \mathrm{d}\tau\, \dot{\boldsymbol{r}}^2(\tau) - \frac{1}{4D}\int_0^t \mathrm{d}\tau \underbrace{\left[\left\|\frac{\boldsymbol{F}}{\gamma}\right\|^2 + \frac{2D}{\gamma}\boldsymbol{\nabla}\cdot\boldsymbol{F}\right]}_{V(\boldsymbol{r})} + \frac{1}{2D\gamma}\int_{\boldsymbol{r}(0)}^{\boldsymbol{r}(t)} \mathrm{d}^3\boldsymbol{r}\cdot\boldsymbol{F}(\boldsymbol{r})\right)\end{aligned} \tag{3.69}$$

with $\boldsymbol{f}(\boldsymbol{r}) = \boldsymbol{F}/\gamma$, and $\gamma = 6\pi\eta a$ (η being the medium viscosity, and a the particle radius). If we change variables in the last integral:

$$\int_{\boldsymbol{r}(0)}^{\boldsymbol{r}(t)} \mathrm{d}^3\boldsymbol{r}\cdot\boldsymbol{F}(\boldsymbol{r}) = \int_0^t \mathrm{d}\tau\,\boldsymbol{F}(\boldsymbol{r})\cdot\frac{\mathrm{d}\boldsymbol{r}(\tau)}{\mathrm{d}\tau}$$

we can rewrite (3.69) as a single integral:

$$\Omega[\boldsymbol{r}(\tau)] = \exp\left(-\frac{1}{4D}\int_0^t \mathrm{d}\tau\, L(\boldsymbol{r}(\tau))\right) \tag{3.70}$$

with the function $L\colon \mathbb{R} \to \mathbb{R}$, $\boldsymbol{r}(\tau) \mapsto L(\boldsymbol{r}(\tau))$ defined as:

$$L(\boldsymbol{r}(\tau)) \equiv \boldsymbol{r}^2(\tau) + \left\| \frac{\boldsymbol{F}(\boldsymbol{r}(\tau))}{\gamma} \right\|^2 + \frac{2D}{\gamma} \boldsymbol{\nabla} \cdot \boldsymbol{F}(\boldsymbol{r}(\tau)) - \frac{2}{\gamma} \boldsymbol{F}(\boldsymbol{r}(\tau)) \cdot \dot{\boldsymbol{r}}(\tau) =$$
$$= \left\| \dot{\boldsymbol{r}}(\tau) - \frac{\boldsymbol{F}(\boldsymbol{r}(\tau))}{\gamma} \right\|^2 + \frac{2D}{\gamma} \boldsymbol{\nabla} \cdot \boldsymbol{F}(\boldsymbol{r}(\tau)) \tag{3.71}$$

Classical path.
Consider the **classical limit** $D \to 0$. Then the $\boldsymbol{\nabla} \cdot \boldsymbol{F}$ term in (3.71) vanishes, and as $-1/(4D) \to \infty$, there will be only *one path* with *non-zero* probability, i.e. the one $\boldsymbol{r_c}(\tau)$ for which the functional vanishes:

$$\int_0^t \mathrm{d}\tau \, L(\boldsymbol{r_c}(\tau)) = 0 \Rightarrow \int_0^t \mathrm{d}\tau \left\| \dot{\boldsymbol{r}}_c(\tau) - \frac{\boldsymbol{F}(\boldsymbol{r}_c(\tau))}{\gamma} \right\|^2 = 0$$

We can then compute that path. As the integrand is a non-negative function, for its integral to be 0 it must be 0 $\forall t$, leading to:

$$\frac{\mathrm{d}\boldsymbol{r_c}}{\mathrm{d}\tau} = \frac{\boldsymbol{F}(\boldsymbol{r_c})}{\gamma}$$

which is just the equation of motion from classical mechanics.

Let's now use the form (3.70) to compute path integrals. For example, consider a transition probability:

$$W(\boldsymbol{r}_t, t|\boldsymbol{r}_0, 0) = \int_{\mathcal{C}\{\boldsymbol{r}_0, 0; \boldsymbol{r}_t, t\}} \mathrm{d}\boldsymbol{r}(\tau) \underbrace{\exp\left(-\frac{1}{4D} \int_0^t \mathrm{d}\tau \, L(\boldsymbol{r}(\tau)) \right)}_{\Omega[\boldsymbol{r}(\tau)]} \tag{3.72}$$

Let's define the functional:

$$S[\boldsymbol{r}(\tau)] = \int_0^t \mathrm{d}\tau \, L(\boldsymbol{r}(\tau))$$

Given the form of (3.72), the path $\boldsymbol{r_c}(\tau)$ that *minimizes* $S[\boldsymbol{r}(\tau)]$ will give the greatest contribution to the path integral. The parameter D *modulates* the relative contributions of paths. If $D \to 0$, $\boldsymbol{r}_c(\tau)$ will be the only contributing path, but if $D \gg 1$, many different paths will have a significant contribution. Suppose that $\boldsymbol{r_c}(\tau)$ is indeed important, meaning that D is sufficiently small (more precisely, that $S[\boldsymbol{r}_c(\tau)]])/D \gg 1$). Then, we write any generic path $\boldsymbol{x}(\tau)$ as the most important one $\boldsymbol{x_c}(\tau)$ plus a "deviation" $\boldsymbol{y}(\tau)$:

$$\boldsymbol{x}(\tau) = \boldsymbol{x_c}(\tau) + \underbrace{(\boldsymbol{x}(\tau) - \boldsymbol{x_c}(\tau))}_{\boldsymbol{y}(\tau)}$$

Note that as the end-points of every path are fixed, $\boldsymbol{y}(0) = \boldsymbol{y}(t) = 0$. Then, we expand in series the functional:

$$S[\boldsymbol{x}(\tau)] = S[\boldsymbol{x_c}(\tau) + y(\tau)] = S[\boldsymbol{x_c}] + \delta S[\boldsymbol{x_c}, \boldsymbol{y}] + \frac{1}{2!} \delta^2 S[\boldsymbol{x_c}, \boldsymbol{y}] + \dots$$

where the δ terms are the *variations* of the functional[2]. For example, the first variation $\delta S[\boldsymbol{x_c}, \boldsymbol{y}]$ is given by, measures how much S *varies* to first order when changing $\boldsymbol{y}(\tau)$. As $\boldsymbol{x_c}$ is a minimum of S, it is also a *stationary point*, meaning that paths close to $\boldsymbol{x_c}$ do not change the value $S[\boldsymbol{x}_c]$ to first order. Applying the definition of the first variation, this leads to the Euler-Lagrange equations for determining $\boldsymbol{x_c}$:

$$\delta S[\boldsymbol{x_c}, \boldsymbol{y}] = \frac{\partial L(\boldsymbol{r}(\tau))}{\partial x_i} - \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\partial L(\boldsymbol{r}(\tau))}{\partial \dot{x}_i} \stackrel{!}{=} 0 \qquad i = 1, 2, 3$$

Then, note how all other terms of the series involve integrals of $\boldsymbol{y}(\tau)$, which do not depend on x - as $\boldsymbol{y}(\tau)$ starts from 0 and returns to 0 at t. So:

$$S[\boldsymbol{x}(\tau)] = S[\boldsymbol{x_c}] + \underbrace{\frac{1}{2!}\delta^2 S[\boldsymbol{x_c}, \boldsymbol{y}] + \dots}_{h(t)}$$

Substituting back in (3.72):

$$W(\boldsymbol{r}_t, t|\boldsymbol{r}_0, 0) = \underbrace{-\frac{1}{4D}\exp(h(t))}_{\Phi(t)} \exp\left(-\frac{1}{4D}S[\boldsymbol{x_c}]\right) = \Phi(t) \exp\left(-\frac{1}{4D}\int_0^t \mathrm{d}\tau\, L[\boldsymbol{r_c}(\tau)]\right)$$

The function $\Phi(t)$ is called *fluctuation factor*, and its computation is not trivial in the general case. However, if we are dealing with transition probabilities, we can use the *normalization condition* to find it:

$$\int_{\mathbb{R}^3} \mathrm{d}^3\boldsymbol{r}\, W(\boldsymbol{r}, t|\boldsymbol{r}, 0) \equiv 1$$

Example 12 (Simple integral with variational methods):

An example will hopefully clarify the essence of the variational method. Let's start with a already known integral, in the $d = 1$ case:

$$W(x, t|x_0, 0) = \int_{\mathbb{R}^T} \left(\prod_{\tau=0^+}^{t} \frac{\mathrm{d}x(\tau)}{\sqrt{4\pi D\,\mathrm{d}\tau}}\right) \exp\left(-\frac{1}{4D}\int_0^t \dot{x}^2(\tau)\,\mathrm{d}\tau\right)\delta(x - x(t)) = \tag{3.73}$$

$$= \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{(x - x_0)^2}{4Dt}\right)$$

Let's compute it again, this time using variations. In this case we are interested in the functional:

$$S[x(\tau)] = \int_0^t \dot{x}^2(\tau)\,\mathrm{d}\tau \tag{3.74}$$

To minimize it, we solve the Euler-Lagrange equations:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\partial S(x_c)}{\partial \dot{x}} - \frac{\partial S(x_c)}{\partial x} = 0 \Rightarrow 0 - 2\ddot{x}_c = 0 \Rightarrow \ddot{x}_c(\tau) = 0$$

[2]∧See `www2.math.uconn.edu/~gordina/NelsonAaronHonorsThesis2012.pdf` for a refresher

Integrating two times:

$$\dot{x}_c(\tau) = a \Rightarrow x_c(\tau) = a\tau + b$$

The boundary conditions are path's two extrema:

$$x_c(0) = b \stackrel{!}{=} 0; \quad x_c(t) = at + x_0 \stackrel{!}{=} x \Rightarrow a = \frac{x - x_0}{t}$$

leading to:

$$x_c(\tau) = x_0 + \frac{x - x_0}{t}\tau$$

So the path minimizing S is just the straight line joining x_0 to x. We can now express any path $x(\tau)$ as a *deviation* from the $x_c(\tau)$:

$$x(\tau) = x_c(\tau) + y(\tau) \qquad y(0) = y(t) = 0 \tag{3.75}$$

This is a change of variables for (3.73), from $x(\tau)$ to $y(\tau)$ ($x_c(\tau)$ is a *fixed* path). As this is just a translation, $\mathrm{d}x(\tau) = \mathrm{d}y(\tau)$ and the path integral becomes:

$$\begin{aligned} W(x,t|x_0,0) &= \int_{\mathbb{R}^T} \left(\prod_{\tau=0^+}^{t} \frac{\mathrm{d}y(\tau)}{\sqrt{4\pi D\,\mathrm{d}\tau}} \right) \delta(y(t) - 0) S[x_c(\tau) + y(\tau)] = \\ &= \int_{\mathcal{C}\{0,0;0,t\}} \left(\prod_{\tau=0^+}^{t} \frac{\mathrm{d}y(\tau)}{\sqrt{4\pi D\,\mathrm{d}\tau}} \right) S[x_c(\tau) + y(\tau)] \end{aligned} \tag{3.76}$$

To compute S, first we differentiate (3.75):

$$\dot{x}(\tau) = \dot{x}_c(\tau) + \dot{y}(\tau)$$

and substitute in (3.74), leading to:

$$S[x(\tau)] = \int_0^t \mathrm{d}\tau\, (\dot{x}_c + \dot{y})^2 = \int_0^t \dot{x}_c^2(\tau)\,\mathrm{d}\tau + 2\int_0^t \dot{x}_c(\tau)\dot{y}(\tau)\,\mathrm{d}\tau + \int_0^t \dot{y}^2(\tau)\,\mathrm{d}\tau$$

Note that the middle term vanishes. We can see it by integrating by parts:

$$\int_0^t \dot{x}_c(\tau)\dot{y}(\tau)\,\mathrm{d}\tau = \dot{x}_c(\tau)y(\tau)\Big|_0^t - \int_0^t \ddot{x}_c(\tau)y(\tau)\,\mathrm{d}\tau = 0$$

as $y(0) = y(t) = 0$ and $\ddot{x}_c(\tau) \equiv 0$. Going back to (3.76):

$$W(x,t|x_0,0) = \int_{\mathbb{R}^T} \left(\prod_{\tau=0^+}^{t} \frac{\mathrm{d}y(\tau)}{\sqrt{4\pi D\,\mathrm{d}\tau}} \right) \delta(y(t) - 0) \exp\left(-\frac{1}{4D} \left[\int_0^t \dot{x}_c^2(\tau)\,\mathrm{d}\tau + \int_0^t \dot{y}^2(\tau)\,\mathrm{d}\tau \right] \right)$$

As $x_c(\tau)$ is fixed, we can bring it outside the integral:

$$= \exp\left(-\int_0^t \dot{x}_c^2(\tau)\,\mathrm{d}\tau \right) \underbrace{\int_{\mathbb{R}^T} \left(\prod_{\tau=0^+}^{t} \frac{\mathrm{d}y(\tau)}{\sqrt{4\pi D\,\mathrm{d}\tau}} \right) \delta(y(t) - 0) \exp\left(-\frac{1}{4D} \int_0^t \dot{y}^2(\tau)\,\mathrm{d}\tau \right)}_{\Phi(t)}$$

We recognize the remaining path integral as a function $\Phi(t)$ of time only, and finally:

$$W(x,t|x_0,0) = \Phi(t)\int_0^t \mathrm{d}\tau\, \dot{x}_c^2(\tau) = \Phi(t)\exp\left[-\frac{1}{4D}\left(\frac{x^2-x_0^2}{t^2}\right)\underbrace{\int_0^t \mathrm{d}\tau}_{t}\right] =$$
$$= \Phi(t)\exp\left(-\frac{(x-x_0)^2}{4Dt}\right)$$

To find the remaining $\Phi(t)$ we can now use the normalization condition:

$$\int_{-\infty}^{+\infty}\mathrm{d}x\, W(x,t|x_0,0) \overset{!}{=} 1$$

In this case, this is just a gaussian integral:

$$\int_{\mathbb{R}}\mathrm{d}x\,\Phi(t)\exp\left(-\frac{(x-x_0)^2}{4Dt}\right) = \Phi(t)\sqrt{4\pi Dt} \overset{!}{=} 1 \Rightarrow \Phi(t) = \frac{1}{\sqrt{4\pi Dt}}$$

And so we retrieve the correct result:

$$W(x,t|x_0,0) = \frac{1}{\sqrt{4\pi Dt}}\exp\left(-\frac{(x-x_0)^2}{4Dt}\right)$$

Gaussian integrals. There is another, more specific, way to interpret the results we discussed in this section. Instead of working in the continuum, we could use a discretization, and see path integrals as (3.73) as integrals of a highly dimensional gaussian. For example, in the case just examined, we have:

$$W(x,t|x_0,0) = \text{“}\lim_{n\to\infty}\text{”} I_N$$
$$I_n = \int_{\mathbb{R}^n}\left(\prod_{i=1}^n \frac{\mathrm{d}x_i}{\sqrt{4\pi D\Delta t_i}}\right)\exp\left(-\sum_{i=1}^n \frac{(x_i-x_{i-1})^2}{4D\Delta t_i}\right)\delta(x-x_n)$$

Performing the integration over the $\mathrm{d}x_n$ we can remove the δ, leaving only a multivariate gaussian:

$$I_n = \int_{\mathbb{R}^{n-1}}\frac{1}{\sqrt{4\pi D\Delta t_i}}\left(\prod_{i=1}^n \frac{\mathrm{d}x_i}{\sqrt{4\pi D\Delta t_i}}\right)\exp\left(-\sum_{i=1}^n \frac{(x_i-x_{i-1})^2}{4D\Delta t_i}\right)\Big|_{x_n=x}$$

This is a gaussian in the form of:

$$\int_{\mathbb{R}^n}\mathrm{d}^n\boldsymbol{x}\exp\left(-\frac{1}{2}\boldsymbol{x}^T A\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{x}\right) \tag{3.77}$$

Note that removing the δ inserts a linear term in the exponential, here highlighted:

$$\sum_{i=1}^{n} \frac{(x_i - x_{i-1})^2}{4D\Delta t_i}\Big|_{x=x_n} = \sum_{i=1}^{n-1} \frac{(x_i - x_{i-1})^2}{4D\Delta t_i} + \frac{x_n^2 + 2x_n x_{n-1} + x_{n-1}^2}{4D\Delta t_n}$$

and so $\boldsymbol{b} \neq \boldsymbol{0}$.

Recall that to solve (3.77) we proceeded with a change of variables, $\boldsymbol{x} = \boldsymbol{x_c} + \boldsymbol{y}$, where $\boldsymbol{x_c}$ is the *minimum* of the gaussian (see 10/10 notes). This leads to a result that is proportional to the exponential evaluated at $\boldsymbol{x_c}$:

$$\begin{aligned} \int_{\mathbb{R}^n} \mathrm{d}^n\boldsymbol{x} \exp\left(-\frac{1}{2}\boldsymbol{x}^T A\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{x}\right) &= \frac{(2\pi)^{n/2}}{\sqrt{\det(A)}} \exp\left(\frac{1}{2}\boldsymbol{b}^T A^{-1}\boldsymbol{b}\right) = \\ &= \frac{(2\pi)^{n/2}}{\sqrt{\det(A)}} \exp\left(\operatorname{Stat}_{\boldsymbol{x}}\left[-\frac{1}{2}\boldsymbol{x}^T A\boldsymbol{x} + \boldsymbol{b}\cdot\boldsymbol{x}\right]\right) \end{aligned}$$

$$\operatorname{Stat}_{\boldsymbol{x}} F(\boldsymbol{x}) = F(\boldsymbol{x}_c); \qquad \boldsymbol{x_c} \text{ such that } \frac{\partial F(\boldsymbol{x})}{\partial x_i}\Big|_{\boldsymbol{x}=\boldsymbol{x}_c} = 0 \quad \forall i = 1, \dots, n$$

So, in the discrete case, the same *variational result* just derives from choosing the best set of coordinates to describe the multivariate gaussian.

Chapter 4

Variational Methods for Path Integrals

4.1 Variational methods

(Lesson 15 of 21/11/19) Compiled: October 13, 2020

Example 13 (Overdamped harmonic oscillator with variational methods)**:**

Consider a particle immersed in a harmonic potential $U(x) = m\omega^2 x^2/2$ and subject to thermal noise, moving in a viscous medium. In the **overdamped limit** $m/\gamma \to 0$ (where $\gamma = 6\pi\eta a$, with η the medium's viscosity and a the particle's radius), the equation of motion becomes:

$$\mathrm{d}x(t) = -kx(t)\,\mathrm{d}t + \sqrt{2D}\,\mathrm{d}B(t) \qquad k = \frac{m\omega^2}{\gamma}$$

A path $\{x(\tau)\}$ solving that equation has a *infinitesimal* probability given by:

$$\mathrm{d}P = \left(\prod_{\tau=0^+}^{t} \frac{\mathrm{d}x(\tau)}{\sqrt{4\pi D\,\mathrm{d}\tau}}\right) \exp\left(-\frac{1}{4D}\int_0^t (\dot{x}+kx)^2\,\mathrm{d}\tau\right)$$

as we already derived. We are now interested in computing the transition probabilities:

$$W(x,t|x_0,0) = \int_{\mathbb{R}^T} \delta(x(t)-x)\,\mathrm{d}P$$

Following the variational method, we arrive to:

$$W(x,t|x_0,0) = \Phi(t)\exp\left(-\frac{1}{4D}S[x_c(\tau)]\right) \tag{4.1}$$

where S is the *action functional* for the harmonic potential:

$$S[x(\tau)] = \int_0^t L(\dot{x},x)\,\mathrm{d}\tau \qquad L(\dot{x},x) = (\dot{x}+kx)^2$$

and $x_c(\tau)$ is the path that *stationarizes* $S[x(\tau)]$, meaning that $\delta S[x_c(\tau)] = 0$ and so it satisfies the Euler-Lagrange equation:

$$0 \overset{!}{=} \frac{\partial L}{\partial x}\Big|_{x_c} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}}\Big|_{x_c} = 2k(\dot{x}_c + kx_c) - 2(\ddot{x}_c + k\dot{x}_c) = 2(k^2 x_c - \ddot{x}_c)$$

as:

$$\frac{\partial L}{\partial x} = 2k(\dot{x} + kx)$$
$$\frac{\partial L}{\partial \dot{x}} = 2(\dot{x} + kx)$$

So, to find $x_c(\tau)$ we need to solve:

$$\begin{cases} \ddot{x}_c = k^2 x_c \\ x_c(0) = x_0 \\ x_c(t) = x \end{cases}$$

This is the second order ordinary differential equation for an *harmonic repulsor*, which has the following general integral:

$$x_c(\tau) = Ae^{k\tau} + Be^{-k\tau}$$

Imposing the boundary conditions leads to:

$$\begin{cases} x_0 \stackrel{!}{=} A + B \\ x \stackrel{!}{=} Ae^{kt} + Be^{-kt} \end{cases} \Rightarrow \begin{cases} B = x_0 - A \\ xe^{kt} = Ae^{2kt} + B \end{cases} \Rightarrow xe^{kt} - x_0 = A[e^{2kt} - 1]$$

$$\Rightarrow A = \frac{xe^{kt} - x_0}{e^{2kt} - 1} \frac{e^{-kt}}{e^{-kt}} = \frac{(xe^{kt} - x_0)e^{-kt}}{\frac{e^{kt} - e^{-kt}}{2} 2} = \frac{x - x_0 e^{-kt}}{2\sinh(kt)}$$

$$B = x_0 - A = -\frac{x - x_0 e^{kt}}{2\sinh(kt)}$$

Then we evaluate the action at the stationary path $x_c(\tau)$:

$$S[x_c(\tau)] = \int_0^t (\dot{x} + kx)^2 \, \mathrm{d}\tau = \int_0^t [2kAe^{k\tau}]^2 \, \mathrm{d}\tau = 4k^2 A^2 \frac{1}{2k} e^{2k\tau}\Big|_0^t =$$

$$= 4kA^2 \frac{e^{2kt} - 1}{2} \frac{e^{-kt}}{e^{-kt}} = 4kA^2 \sinh(kt)\, e^{kt} =$$

$$= \cancel{4}k \frac{(x - x_0 e^{-kt})^2}{\cancel{4}\sinh(kt)} e^{kt} = \frac{k(x - x_0 e^{-kt})^2}{e^{kt} - e^{-kt}} \frac{2}{e^{-kt}} = \frac{2k(x - x_0 e^{-kt})^2}{1 - e^{-2kt}}$$

Substituting back in (4.1):

$$W(x, t|x_0, 0) = \Phi(t) \exp\left(-\frac{k}{2D(1 - e^{-2kt})}[x - x_0 e^{-kt}]^2\right)$$

All that's left to find $\Phi(t)$ is to use the normalization condition:

$$1 \stackrel{!}{=} \int_{\mathbb{R}} \mathrm{d}x \, W(x, t|x_0, 0) = \Phi(t) \int_{\mathbb{R}} \mathrm{d}x \exp\left(-\overbrace{\frac{k}{2D(1 - e^{-2kt})}}^{\alpha}[x - x_0 e^{-kt}]^2\right) =$$

$$= \Phi(t)\sqrt{\frac{\pi}{\alpha}} = \Phi(t)\sqrt{\frac{2\pi D(1 - e^{-2kt})}{k}} \Rightarrow \Phi(t) = \sqrt{\frac{k}{2\pi D}} \frac{1}{\sqrt{1 - e^{-2kt}}}$$

And so the full solution is:

$$W(x,t|x_0,0) = \sqrt{\frac{k}{2\pi D}} \frac{1}{\sqrt{1-e^{-2kt}}} \exp\left(-\frac{k}{2D}\frac{(x-x_0e^{-kt})^2}{(1-e^{-2kt})}\right) \xrightarrow[t\to\infty]{} \sqrt{\frac{k}{2\pi D}} \exp\left(-\frac{k}{2Dx^2}\right)$$

As before, we can compute the $t \to \infty$ with a Maxwell-Boltzmann distribution $e^{-\beta U(x)}$, obtaining:

$$\frac{1}{2}\beta m\omega^2 x^2 = \frac{k}{2D}x^2 \Rightarrow D = \frac{k}{\beta m \omega^2} = \frac{1}{\beta\gamma} = \frac{k_B T}{\gamma} \Rightarrow D\gamma = k_B T$$

as we previously derived.

If we do not consider the overdamped limit, however, the equation of motion is given by:

$$m\ddot{x} = -\gamma\dot{x} - m\omega^2 x + \sqrt{2D}\gamma\xi$$

This can be rewritten as a system of two first order (stochastic) differential equations:

$$\begin{cases} \mathrm{d}x\,(\tau) = v(\tau)\,\mathrm{d}\tau \\ \mathrm{d}v\,(\tau) = -\frac{\gamma}{m}v(\tau)\,\mathrm{d}\tau + \frac{\gamma\sqrt{2D}}{m}\,\mathrm{d}B \end{cases}$$

It is convenient to "symmetrize" the system, by adding a stochastic term also in the first equation:

$$\begin{cases} \mathrm{d}x\,(\tau) = v(\tau)\,\mathrm{d}\tau + 2\hat{D}\sqrt{\mathrm{d}\hat{B}} \\ \mathrm{d}v\,(\tau) = -\frac{\gamma}{m}v(\tau)\,\mathrm{d}\tau + \frac{\gamma\sqrt{2D}}{m}\,\mathrm{d}B \end{cases}$$

and then we'll consider the limit $\hat{D} \to 0$.
First, as usual, we discretize, with $\{t_i\}_{i=0,\dots,n}$ and $t_0 \equiv 0$, $t_n \equiv t$, arriving to:

$$\begin{cases} \Delta x_i = v_{i-1}\Delta t_i + \sqrt{2\hat{D}}\Delta\hat{B}_i \\ \Delta v_i = -\frac{\gamma}{m}v_{i-1}\Delta t_i + \frac{\gamma}{m}\sqrt{2D}\Delta B_i \end{cases}$$

Where the velocity is evaluated at t_{i-1} as per Ito's prescription. As ΔB_i and $\Delta\hat{B}_i$ are **independent** gaussian increments, their joint distribution is just a product:

$$\mathrm{d}P\,(\Delta B_1, \Delta\hat{B}_1, \dots, \Delta B_n, \Delta\hat{B}_n) = \left(\prod_{i=1}^n \frac{\mathrm{d}\Delta B_i}{\sqrt{2\pi\Delta t_i}}\frac{\mathrm{d}\Delta\hat{B}_i}{\sqrt{2\pi\Delta t_i}}\right) \exp\left(-\frac{1}{2}\sum_{i=1}^n \frac{\Delta B_i^2}{\Delta t_i} - \frac{1}{2}\sum_{i=1}^n \frac{\Delta\hat{B}_i^2}{\Delta t_i}\right)$$

As done previously (see 14/11 notes), to get the distribution for Δx_i and Δv_i we make a change of random variables:

$$\begin{aligned}
\Delta\hat{B}_i &= \frac{\Delta x_i - v_{i-1}\Delta t_i}{\sqrt{2\hat{D}}}\\
\Delta B_i &= \left(\Delta v_i + \frac{\gamma}{m} v_{i-1}\Delta t_i\right)\frac{m}{\gamma\sqrt{2D}}
\end{aligned}$$

with jacobian:

$$\begin{aligned}
&\det\left|\frac{\partial\{\Delta\hat{B}_i\}}{\partial\{\Delta x_i\}}\right| = (2\hat{D})^{-n/2}\\
&\det\left|\frac{\partial\{\Delta B_i\}}{\partial\{\Delta x_i\}}\right| = \det\left|\frac{\partial\{\Delta x_i\}}{\partial\{\Delta B_i\}}\right|^{-1} = \left(\frac{\gamma}{m}\sqrt{2D}\right)^{-n} = \left(\frac{\gamma^2}{m^2}2D\right)^{-n/2}
\end{aligned}$$

leading to:

$$\begin{aligned}
\mathrm{d}P\left(\{\Delta x_i\},\{\Delta v_i\}\right) = &\left(\prod_{i=1}^n \frac{\mathrm{d}\Delta x_i}{\sqrt{4\pi D\Delta t_i}}\frac{\mathrm{d}\Delta v_i}{\sqrt{4\pi D\Delta t_i\gamma^2/m^2}}\right)\cdot\\
&\cdot\exp\left(-\frac{1}{2}\sum_{i=1}^n\frac{m^2}{2\gamma^2 D}\left[\left(\frac{\Delta v_i + \gamma/m v_{i-1}\Delta t_i}{\Delta t_i}\right)^2\Delta t_i\right]\right)\cdot\\
&\cdot\exp\left(-\frac{1}{2}\sum_{i=1}^n\frac{1}{2\hat{D}}\left[\left(\frac{\Delta x_i - v_{i-1}\Delta t_i}{\Delta t_i}\right)^2\Delta t_i\right]\right) =\\
= &\left(\prod_{i=1}^n \frac{\mathrm{d}\Delta x_i}{\sqrt{4\pi D\Delta t_i}}\frac{\mathrm{d}\Delta v_i}{\sqrt{4\pi D\Delta t_i\gamma^2/m^2}}\right)\cdot\\
&\cdot\exp\left(-\frac{m^2}{4D\gamma^2}\sum_{i=1}^n\left[\left(\frac{\Delta v_i}{\Delta t_i} + \frac{\gamma}{m}v_{i-1}\right)^2\Delta t_i\right]\right)\cdot\\
&\cdot\exp\left(-\frac{1}{4\hat{D}}\sum_{i=1}^n\left[\left(\frac{\Delta x_i}{\Delta t_i} - v_{i-1}\right)^2\Delta t_i\right]\right)
\end{aligned}\tag{4.2}$$

Taking the continuum limit $n \to \infty$ leads to:

$$\begin{aligned}
\mathrm{d}P\left(\{x(\tau), v(\tau)\}\right) = &\left(\prod_{\tau=0^+}^t \frac{\mathrm{d}x(\tau)}{\sqrt{4\pi\hat{D}\,\mathrm{d}\tau}}\frac{\mathrm{d}v(\tau)}{\sqrt{4\pi D\,\mathrm{d}\tau\,\gamma^2/m^2}}\right)\cdot\\
&\cdot\exp\left(-\frac{m^2}{4D\gamma^2}\int_0^t \mathrm{d}\tau\left[\dot{v}(\tau) + \frac{\gamma}{m}v(\tau)\right]^2 - \frac{1}{4\hat{D}}\int_0^t \mathrm{d}\tau\left[\dot{x}(\tau) - v(\tau)\right]^2\right)
\end{aligned}$$

In the limit $\hat{D} \to 0^+$, $1/(4\hat{D}) \to +\infty$, and so the gaussian pdf for the $\Delta\hat{B}_i$ becomes *infinitely thin*, and the only path with a non-vanishing probability will be the one where:

$$\int_0^t \mathrm{d}\tau\left[\dot{x} - v(\tau)\right]^2 = 0$$

As any > 0 value will lead to $\exp(-\infty) = 0$. In particular, the i-th factor of the discretization becomes:

$$\frac{1}{\sqrt{4\pi\hat{D}\Delta t_i}}\exp\left[-\frac{1}{4\hat{D}}\left(\frac{\Delta x_i}{\Delta t_i} - v_0^2\right)\Delta t_i\right] =$$

$$= \frac{1}{\sqrt{4\pi\hat{D}\Delta t_i}} \exp\left(-\frac{1}{4\hat{D}\Delta t_i}(\Delta x_i - v_{i-1}\Delta t_i)^2\right) \xrightarrow[\hat{D}\to 0]{} \delta(\Delta x_i - v_{i-1}\Delta t_i)$$

where we used a limit definition for the δ:

$$\lim_{\epsilon\to 0} \frac{1}{\sqrt{4\pi\epsilon}} \exp\left(-\frac{x^2}{4\epsilon}\right) = \delta(x)$$

with $\epsilon = \hat{D}\Delta t_i$ and $x = \Delta x_i - v_{i-1}\Delta t_i$.
Substituting back in (4.2):

$$\mathrm{d}P\left(\{\Delta x_i\},\{\Delta v_i\}\right) = \left(\prod_{i=1}^{n} \mathrm{d}\Delta x_i\, \delta(\Delta x_i - v_{i-1}\Delta t_i)\frac{\mathrm{d}\Delta v_i}{\sqrt{4\pi D\Delta t_i \gamma^2/m^2}}\right)\cdot$$
$$\cdot \exp\left(-\frac{m^2}{4D\gamma^2}\sum_{i=1}^{n}\left[\left(\frac{\Delta v_i}{\Delta t_i} + \frac{\gamma}{m}v_{i-1}\right)^2 \Delta t_i\right]\right)$$

Now consider the discretized transition probability:

$$\begin{aligned} W(x_n, v_n, t_n|x_0, v_0, 0) &= \int_{\mathbb{R}^n\times\mathbb{R}^n} \mathrm{d}P\left(\{x_i, v_i\}\right)\delta(x_n - x)\delta(v_n - x) = \\ &= \int_{\mathbb{R}^n\times\mathbb{R}^n} \left(\prod_{i=1}^{n} \mathrm{d}\Delta x_i\, \delta(\Delta x_i - v_{i-1}\Delta t_i)\frac{\mathrm{d}\Delta v_i}{\sqrt{4\pi D\Delta t_i \gamma^2/m^2}}\right)\cdot \\ &\cdot \exp\left(-\frac{m^2}{4D\gamma^2}\sum_{i=1}^{n}\left[\left(\frac{\Delta v_i}{\Delta t_i} + \frac{\gamma}{m}v_{i-1}\right)^2 \Delta t_i\right]\right)\delta(v_n - v)\delta(x_n - x) \end{aligned} \tag{4.3}$$

Let's focus on the integrations over x_i:

$$\int_{\mathbb{R}^n} \left(\prod_{i=1}^{n} \mathrm{d}\Delta x_i\, \delta(\Delta x_i - v_{i-1}\Delta t_i)\right)\delta(x_n - x) =$$
$$= \int_{\mathbb{R}^n} \mathrm{d}\Delta x_1 \dots \mathrm{d}\Delta x_n\, \delta(\Delta x_1 - v_0\Delta t_1)\dots\delta(\Delta x_n - v_{n-1}\Delta t_n)\delta(x_n - x)$$

We then perform the change of variables $\Delta x_1 = x_1 - x_0$, with x_0 constant, so that $\mathrm{d}\Delta x_1 = \mathrm{d}x_1$. Then we integrate over $\mathrm{d}x_1$, eliminating the first δ and setting $x_1 = x_0 - v_0\Delta t_1$:

$$\int_{\mathbb{R}^n} \mathrm{d}x_1\, \mathrm{d}\Delta x_2 \dots \mathrm{d}\Delta x_n\, \delta(x_1 - x_0 - v_0\Delta t_1)\delta(\Delta x_2 - v_1\Delta t_2)\dots\delta(\Delta x_n - v_{n-1}\Delta t_n)\delta(x_n - x) =$$
$$\int_{\mathbb{R}^{n-1}} \mathrm{d}\Delta x_2 \dots \mathrm{d}\Delta x_n\, \delta(x_2 - x_0 - v_0\Delta t_1 - v_1\Delta t_2)\dots\delta(\Delta x_n - v_{n-1}\Delta t_n)\delta(x_n - x)$$

Repeating these steps for all the other variables except the last one, we arrive to:

$$= \int_{\mathbb{R}} \mathrm{d}x_n\, \delta\left(x_n - x_0 - \sum_{i=1}^{n} v_{i-1}\Delta t_i\right)\delta(x_n - x) = \delta\left(x - x_0 - \sum_{i=1}^{n} v_{i-1}\Delta t_i\right)$$

In the continuum limit, this becomes:

$$\delta\left(x - x_0 - \int_0^t v(\tau)\,\mathrm{d}\tau\right)$$

Substituting back in (4.3) and finally taking the limit $n \to \infty$:

$$W(x,v,t|x_0,v_0,0) = \int_{\mathbb{R}^T} \left(\prod_{\tau=0^+}^{t} \frac{\mathrm{d}v(\tau)}{\sqrt{4\pi D \,\mathrm{d}\tau\, \gamma/m^2}} \right) \exp\left(-\frac{m^2}{4D\gamma} \int_0^t \left(\dot{v}(\tau) + \frac{\gamma}{m} v(\tau) \right)^2 \mathrm{d}\tau \right)$$
$$\cdot\, \delta(v(t)-v)\delta\left(x - x_0 - \int_0^t v(\tau)\,\mathrm{d}\tau \right)$$

We can now use the variational method to compute that integral. So, let $v_c(\tau)$ be the path, starting at $v(0) = v_0$ that *stationarizes* the action functional:

$$S[v(\tau)] = \int_0^t \left(\dot{v}(\tau) + \frac{\gamma}{m} v(\tau) \right)^2 \mathrm{d}\tau$$

so that $\delta S[v_c(\tau)] = 0$, and also satisfies the constraints imposed by the δ:

$$v(t) \stackrel{!}{=} v \qquad x - x_0 \stackrel{!}{=} \int_0^t v(\tau)\,\mathrm{d}\tau$$

Then, the path integral is given by:

$$W(x,v,t|x_0,v_0,0) = \Phi(t) \exp\left(-\frac{m^2}{4D\gamma} \int_0^t \left(\dot{v}_c(\tau) + \frac{\gamma}{m} v_c(\tau) \right)^2 \mathrm{d}\tau \right) \qquad (4.4)$$

All that's left is to compute $v_c(\tau)$ and evaluate the integral. This is a problem of *constrained optimization*, for which we use the method of Lagrange multipliers.

Brief refresher of Lagrange multipliers. Suppose we have two functions $F, g\colon \mathbb{R}^2 \to \mathbb{R}$, with $F(x,y)$ being the function to maximize, and $g(x,y) = c \in \mathbb{R}$ a constraint. A stationary point (x_0, y_0) of F subject to the constraint $g(x,y) = c$ is such that if we move slightly from (x_0, y_0) along the contour $g(x,y) = c$, the value of $F(x,y)$ does not change (to first order). This happens if the contour of F passing through the stationary point $F(x,y) = F(x_0,y_0)$ is parallel at (x_0,y_0) to that of $g(x,y) = c$, meaning that at (x_0,y_0) the gradients of F and g are parallel:

$$\nabla_{x,y} F = \lambda \nabla_{x,y} g \qquad \lambda \in \mathbb{R}$$

(Here we assume that $\nabla_{x,y} g(x_0,y_0) \neq \mathbf{0}$). Rearranging:

$$\nabla_{x,y}(F(x,y) - \lambda g(x,y)) = \mathbf{0}$$

Together with the constraint equation $g(x,y) = c$, we have now 3 equations in 3 unknowns (x, y, λ) that can be solve to yield the desired stationary point (x_0, y_0).

In this case, we have *functionals* instead of functions, and *functionals derivatives* (i.e. variations) instead of derivatives. So, to find the stationary points of:

$$\int_0^t \left(\dot{v}(\tau) + \frac{\gamma}{m} v(\tau) \right)^2 \mathrm{d}\tau \qquad (a)$$

subject to the constraint:

$$\int_0^t v(\tau)\,\mathrm{d}\tau = x - x_0 \qquad (b)$$

we need to solve:

$$\delta \int_0^t \underbrace{\left[\left(\dot{v}(\tau) + \frac{\gamma}{m} v(\tau)\right)^2 - \lambda v(\tau)\right]}_{L(v,\dot{v})} \mathrm{d}\tau = 0$$

And applying the definition of first variation (the δ above) leads to solving the Euler-Lagrange equations:

$$\frac{\partial L}{\partial v} - \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\partial L}{\partial \dot{v}}\Big|_{v=v_c} \stackrel{!}{=} 0$$

Expanding the computations:

$$2\left(\dot{v}_c + \frac{\gamma}{m}v_c\right)\frac{\gamma}{m} - \lambda - \frac{\mathrm{d}}{\mathrm{d}\tau}\left[2\left(\dot{v}_c + \frac{\gamma}{m}v_c\right)\right] = 0 \Rightarrow \ddot{v}_c(\tau) = v_c(\tau)\left(\frac{\gamma}{m}\right)^2 - \frac{\lambda}{2}$$

The homogeneous solution is again a combination of exponentials:

$$v_c(\tau) = A\exp\left(-\frac{\gamma}{m}\tau\right) + B\exp\left(\frac{\gamma}{m}\tau\right)$$

And for the inhomogeneous general integral we just need to add a particular solution, for example the one with constant velocity $\dot{v}(\tau) = \text{const} \Rightarrow \ddot{v}_c(\tau) = 0$, given by:

$$v_c(\tau) = \frac{\lambda}{2}\left(\frac{m}{\gamma}\right)^2$$

Then, we need to impose the boundary conditions:

$$v_c(0) = v_0 \qquad v_c(t) = v \qquad \int_0^t v(\tau)\,\mathrm{d}\tau = (x - x_0)$$

So we have 3 parameters (the two constants of integration A, B and λ) and 3 equations. After finding all of them, we just need to evaluate the integral (4.4) (computations omitted).

4.2 Diffusion with obstacles

(Lesson 15 of 21/11/19) Compiled: October 13, 2020

Consider a particle in a potential $U(x)$ (fig. 4.1), with a local minimum separated by a *barrier*. In the classical case, if the particle's energy is sufficiently low, it can become *forever* trapped inside the minimum. However, in the presence of *thermal fluctuations* there may be a possibility of escape - a sort of classical tunnelling.

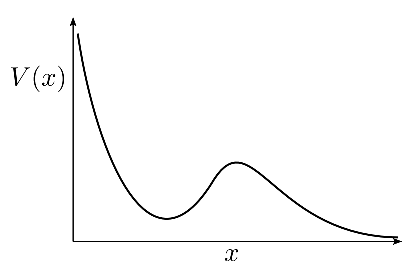


Figure (4.1) – Potential graph

We first consider an easier problem, that of the diffusion process on a compact domain $[a, b]$, representing the *boundaries* of the potential well of fig. 4.1. We then suppose that the particle cannot escape from the left side a, but it can do so - and always does - from the right one b. This means that a is a "reflecting" boundary - i.e. if the particle hits $x = a$ it "bounces back"), while $x = b$ is an *absorbing* boundary, that is a particle reaching b can be "absorbed by the environment" and disappear from the system. In the more general case, the probability of reflection at $x = a$ or absorption at $x = b$ will not be certain, but will depend on the particle's energy.

Recall the Langevin equation:

$$\mathrm{d}x(t) = \underbrace{\frac{F(x,t)}{\gamma}}_{f(x,t)}\,\mathrm{d}t + \sqrt{2D(x,t)}\,\mathrm{d}B \qquad F(x) = -U'(x);\ x \in [a, b] \tag{4.5}$$

This is equivalent to the Fokker-Planck equation:

$$\begin{aligned}\frac{\partial}{\partial t}W(x,t|x_0,0) &= -\frac{\partial}{\partial x}\left[f(x,t)W(x,t|x_0,0) - \frac{\partial}{\partial x}(D(x,t)W(x,t|x_0,0))\right] = \\ &= -\frac{\partial}{\partial x}\overbrace{\left[-\underbrace{\frac{U'(x)}{\gamma}}_{A(x)}W(x,t|x_0,0) - \frac{\partial}{\partial x}\Big(\underbrace{\frac{k_BT}{\gamma}}_{D}W(x,t|x_0,0)\Big)\right]}^{J(x,t)} = \end{aligned} \tag{4.6}$$

$$= -\partial_x[A(x)W(x,t|x_0,0)] + \partial_x^2[D(x)W(x,t|x_0,0)] \tag{4.7}$$

where we inserted $D(x,t) \equiv D = k_BT/\gamma$ (derived from the equilibrium limit). $J(x,t)$ is the probability flux coming out from x at instant t.
To solve (4.6) we need a precise mathematical description for the *reflecting* and *absorbing* boundaries:

- In $x = a$, the *reflecting* boundary condition means that:

$$J(a,t) = A(a)W(a,t|x_0,0) - [\partial_x D(x)W(x,t|x_0,0)]|_{x=a} \overset{!}{=} 0 \qquad \forall t \tag{4.8}$$

  As every particle that goes in a immediately comes out after being reflected, the *inward* flux and *outward* one are the same, and so their sum is 0.

- In b, however, the *absorbing* boundary condition means that the probability to find the particle here is exactly 0:

$$W(b,t|x_0,0) \overset{!}{=} 0 \tag{4.9}$$

As $x \in [a, b]$, the domain of equation (4.6) is not isotropic anymore - meaning that the solution $W(x,t|x_0,0)$ will depend on x_0, making the problem much

difficult. The idea is then to translate the problem from finding the full transition probability $W(x,t|x_0,0)$ to finding a simpler, but still interesting, function, that depends on less parameters.
One possible choice is given by the **survival probability**, i.e. the probability that a particle starting at a given point x will still be inside the interval $[a,b]$ at a later time t:

$$G(x,t) = \int_a^b \mathrm{d}y\, W(y,t|x,0)$$

Note that we keep the starting time fixed at 0, and integrate over all the possible *destinations* of the particle - reducing the number of variables from 4 to 2.
Note that generally $G(x,t) \neq 1$, as the boundary in b offers a possibility of escape, leading to a *violation* of the conservation of probability. In fact the condition (4.9) $W(b,t|x_0,t_0) = 0$ does not mean that the flux here is null. Recalling the definition of $J(x,t)$ from (4.6):

$$\begin{aligned} J(b,t) &= \cancel{A(b)W(b,t|x_0,t_0)} - \partial_x(D(x)W(x,t|x_0,t_0))|_{x=b} = \\ &= \cancel{-(\partial_x D)W(b,t|x_0,t_0)} - D(b)\partial_x W(x,t|x_0,t_0)|_{x=b} \neq 0 \end{aligned}$$

Now, we need to translate (4.6) to a differential equation for $G(x,t)$. We can start by evaluating the time derivative of $G(x,t)$:

$$\frac{\partial}{\partial t}G(x,t) = \int_a^b \mathrm{d}x' \,\frac{\partial}{\partial t}W(x',t|x,0) \tag{4.10}$$

We could use (4.7) to expand the $\partial_t W(x',t|x,0)$ term - but this does not really work:

$$\frac{\partial}{\partial t}G(x,t) = \int_a^b \mathrm{d}x' \left[-\partial_{x'}(A(x')W(x',t|x,0)) + \partial^2_{x'}(D(x')W(x',t|x,0))\right]$$

To reconstruct derivatives of $G(x,t)$ in the right side, we would need to bring the $\partial_{x'}$ out of the integrals - but this is not possible, as x' is the variable of integration. One way to solve this would be to somehow move the derivative from $\partial_{x'}$ to ∂_x.
To do this, we start from the ESCK relation:

$$\int_a^b \mathrm{d}x_1\, W(x_2,t_2|x_1,t_1)W(x_1,t_1|x_0,t_0) = W(x_2,t_2|x_0,t_0) \qquad t_0 < t_1 < t_2$$

Differentiating with respect to the middle time t_1:

$$\int_a^b \mathrm{d}x_1 \left[W(x_1,t_1|x_0,t_0)\partial_{t_1}W(x_2,t_2|x_1,t_1) + W(x_2,t_2|x_1,t_1)\,\partial_{t_1}W(x_1,t_1|x_0,t_0)\right] = 0$$

We then use (4.7) to expand the highlighted term:

$$\begin{aligned} &\int_a^b \mathrm{d}x_1\, W(x_1,t_1|x_0,t_0)\partial_{t_1}W(x_2,t_2|x_1,t_1) + \\ &+ \int_a^b \mathrm{d}x_1\, W(x_2,t_2|x_1,t_1)[-\partial_{x_1}A(x_1)W(x_1,t_1|x_0,t_0) + \partial^2_{x_1}D(x_1)W(x_1,t_1|x_0,t_0)] = 0 \end{aligned}$$

And then we integrate by parts the second term, to move the ∂_{x_1} and $\partial^2_{x_1}$ derivatives:

$$\int_a^b \mathrm{d}x_1\, W(x_1,t_1|x_0,t_0)\partial_{t_1} W(x_2,t_2|x_1,t_1) +$$
$$-A(x_1)W(x_1,t_1|x_0,t_0)W(x_2,t_2|x_1,t_1)\Big|_{x_1=a}^{x_1=b} + W(x_2,t_2|x_1,t_1)[\partial_{x_1} D(x_1)W(x_1,t_1|x_0,t_0)]\Big|_{x_1=a}^{x_1=b}$$
$$-D(x_1)W(x_1,t_1|x_0,t_0)[\partial_{x_1} W(x_2,t_2|x_1,t_1)]\Big|_{x_1=a}^{x_1=b}$$
$$+\int_a^b \mathrm{d}x_1\,[A(x_1)W(x_1,t_1|x_0,t_0)\partial_{x_1} W(x_2,t_2|x_1,t_1) + D(x_1)W(x_1,t_1|x_0,t_0)]\partial^2_{x_1} W(x_2,t_2|x_1,t_1) = 0$$

In the limit $t_1 \to 0$, $W(x_1,t_1|x_0,t_0) = \delta(x_1 - x_0)\delta(t_1 - t_0)$. This makes all the boundary terms vanish (given that $x_0 \neq a, b$), and allows to compute the other integrals (with $x_1 = x_0$ and $t_1 = t_0$), leading to:

$$\frac{\partial}{\partial t_0} W(x_2,t_2|x_0,t_0) + A(x_0)\frac{\partial}{\partial x_0} W(x_2,t_2|x_0,t_0) + D(x_0)\frac{\partial^2}{\partial x_0^2} W(x_2,t_2|x_0,t_0) = 0$$

Rearranging, and dropping some subscripts:

$$\partial_{t_0} W(x,t|x_0,t_0) = -A(x_0)\partial_{x_0} W(x,t|x_0,t_0) - D(x_0)\partial^2_{x_0} W(x,t|x_0,t_0) \quad (4.11)$$

This is the **backward Fokker-Planck equation**, as all derivatives are with respect to the starting time or position - meaning that it can be use to "retrodict" the past given the future. This could be used for computing $\partial_t G(x,t)$ - but first we need to express the derivative ∂_{t_0} in terms of the derivative ∂_t that appears in $\partial_t G(x,t)$.
Supposing that $A(x)$ and $D(x)$ are time-independent (as we implicitly did in the previous notation), then (4.7) is an *autonomous* differential equation, meaning that the solution does not change after a time translation:

$$W(x,t|x_0,t_0) = W(x,t-t_0|x_0,0)$$

Differentiating with respect to t_0:

$$\partial_{t_0} W(x,t|x_0,t_0) = \partial_{t'} W(x,t'|x_0,0)|_{t'=t-t_0}\partial_{t_0}(t-t_0) = -\partial_t W(x,t-t_0|x_0,0) = -\partial_t W(x,t|x_0,t_0)$$

Substituting this relation in (4.11) we get:

$$\partial_t W(x,t|x_0,t_0) = A(x_0)\partial_{x_0} W(x,t|x_0,t_0) + D(x_0)\partial^2_{x_0} W(x,t|x_0,t_0) \quad (4.12)$$

Finally, we can use (4.12) in (4.10):

$$\frac{\partial}{\partial t} G(x,t) = \int_a^b \mathrm{d}x'\, \partial_t W(x',t|x,0) = \quad (4.13)$$
$$\underset{(4.12)}{=} \int_a^b \mathrm{d}x'\,[A(x)\partial_x W(x',t|x,0) + D(x)\partial^2_x W(x',t|x,0)] =$$
$$= A(x)\partial_x \underbrace{\int_a^b \mathrm{d}x'\, W(x',t|x,0)}_{G(x,t)} + D(x)\partial^2_x \underbrace{\int_a^b \mathrm{d}x'\, W(x',t|x,0)}_{G(x,t)} =$$

$$= A(x)\partial_x G(x,t) + D(x)\partial_x^2 G(x,t) \tag{4.14}$$

We have now a differential equation for $G(x,t)$, and we need to translate the appropriate boundary conditions (4.8) and (4.9). The latter is immediate:

$$W(b,t|x_0,0) = 0 \qquad \forall t\, \forall x_0 \in [a,b] \Rightarrow G(x,t)|_{x=b} = 0 \tag{4.15}$$

However, the analogous of (4.8) requires a bit more work. So we start again from the ESCK relation, and differentiate with respect to the mid-time:

$$\partial_\tau \int_a^b \mathrm{d}y\, W(x',t|y,\tau)W(y,\tau|x,0) = \partial_\tau W(x',t|x,0) = 0$$

Expanding the left side:

$$\int_a^b \mathrm{d}y\, [W(y,\tau|x,0)\, \partial_\tau W(x',t|y,\tau) + W(x',t|y,\tau)\, \partial_\tau W(y,\tau|x,0)\,] = 0$$

We can now use (4.11) for the term highlighted in yellow, and (4.7) (also called **forward Fokker-Planck equation**) for the term in green, leading to:

$$\int_a^b \mathrm{d}y\, [-A(y)\partial_y W(x',t|y,\tau) - D(y)\partial_y^2 W(x',t|y,\tau)]W(y,\tau|x,0) +$$
$$\int_a^b \mathrm{d}y\, [-\partial_y A(y) W(y,\tau|X,0) + \partial_y^2 D(y) W(y,\tau|x,0)]W(x',t|y,\tau)$$

We now integrate by parts the first term, moving the ∂_y and ∂_y^2 derivatives away from $W(x',t|y,\tau)$:

$$-A(y)W(x',t|y,\tau)W(y,\tau|x,0)\Big|_{y=a}^{y=b} + \int_a^b \mathrm{d}y\, [\partial_y A(y) W(y,\tau|x,0)]W(x',t|y,\tau) +$$
$$-D(y)W(y,\tau|x,0)[\partial_y W(x',t|y,\tau)]\Big|_{y=a}^{y=b} + W(x',t|y,\tau)[\partial_y D(y)W(y,\tau|x,0)]\Big|_{y=a}^{y=b} +$$
$$-\int_a^b \mathrm{d}y\, [\partial_y^2 D(y) W(y,\tau|x,0)]W(x',t|y,\tau) - \int_a^b \mathrm{d}y\, \partial_y [A(y)W(y,\tau|x,0)]W(x',t|y,\tau) +$$
$$+\int_a^b \mathrm{d}y\, \partial_y^2 [D(y)W(y,\tau|x,0)]W(x',t|y,\tau) = 0$$

The highlighted terms cancel out, leaving only boundaries:

$$-A(y)W(x',t|y,\tau)W(y,\tau|x,0)\Big|_{y=a} - D(y)W(y,\tau|x,0)[\partial_y W(x',t|y,\tau)]\Big|_{y=a}^{y=b} +$$
$$+W(x',t|y,\tau)[\partial_y D(y)W(y,\tau|x,0)]\Big|_{y=a}^{y=b} = 0$$

Now $W(b,t|x_0,0) = 0$ (4.9), and also $W(x',t|b,\tau) = 0$, as a particle starting in b escapes immediately from $[a,b]$. This makes all the boundary terms vanish at $y=b$, leaving only:

$$+A(a)W(x',t|a,\tau)W(a,\tau|x,0) + D(a)W(a,\tau|x,0)[\partial_y W(x',t|y,\tau)]|_{y=a} +$$
$$-W(x',t|a,\tau)[\partial_y D(y)W(y,\tau|x,0)]|_{y=a} = 0$$

Collecting $W(x', t|a, \tau)$ allows to recognize a $J(x, t)$ term:

$$\begin{aligned}&D(a)W(a,\tau|x,0)[\partial_y W(x',t|y,\tau)]|_{y=a}+\\+&W(x',t|a,\tau)\Big[A(a)W(a,\tau|x,0)-\underbrace{[\partial_y D(y)W(y,\tau|x,0)]|_{y=a}}_{J(a,\tau)}\Big]=0\end{aligned}$$

But recall that $J(a, \tau) = 0\, \forall \tau$ as per (4.9). So only a term remains:

$$D(a)W(a,\tau|x,0)[\partial_y W(x',t|y,\tau)]|_{y=a} = 0 \Rightarrow W(a,\tau|x,0) = 0 \vee \partial_y W(x',t|y,\tau)|_{y=a} = 0 \quad \forall\tau$$

Finally, by integrating the second term:

$$\int_a^b \mathrm{d}x'\, \partial_y W(x',t|y,\tau) = \partial_y \int_a^b \mathrm{d}x'\, W(x',t|y,\tau) = \partial_y G(y,\tau)$$

And evaluating at $y = a$ leads to:

$$\partial_x G(x,t)|_{x=a} = 0 \tag{4.16}$$

which is the last boundary condition we needed for $G(x, t)$.
So, the problem now becomes:

$$\begin{cases}\partial_t G(x,t) = A(x)\partial_x G(x,t) + D(x)\partial_x^2 G(x,t)\\ \partial_x G(x,t)|_{x=a} = 0\\ G(x,t)|_{x=b} = 0\end{cases}$$

We can make one last simplification by *removing* the time coordinate. Let's introduce $T(x)$ as being the *lifetime* of a particle starting at x - meaning the amount of time needed for that particle to "disappear" by reaching b (so, in this case, $T(x)$ coincides with $T_{\text{ftv}}(b, x)$, i.e. the *time to the first visit* of b). The exact value of $T(x)$ will depend on the particle's path, making $T(x)$ a random variable. Note that:

$$G(x,t) = \mathbb{P}(T(x) > t)$$

That is, the survival probability is the probability that the particle *has not yet reached* b during the time interval $[0, t]$, which is equivalent to saying that its lifetime is greater than t. Denoting with $\mathbb{P}_{\text{ftv}}(T_b)\, \mathrm{d}T_b$ the probability that a particle will visit b in the time range $[T_b, T_b + \mathrm{d}T_b]$, we have:

$$G(x,t) = \mathbb{P}(T(x) > t) = \int_t^{+\infty} \mathbb{P}_{\text{ftv}}(T_b)\, \mathrm{d}T_b = -\int_{+\infty}^t \mathbb{P}_{\text{ftv}}(T_b)\, \mathrm{d}T_b$$

Differentiating with respect to t:

$$\partial_t G(x,t) = -\mathbb{P}_{\text{fvt}}(t)$$

As we need a function, and $T(x)$ is a random variable, we consider its *average*, i.e. the *mean time of arrival at* b $T_b(x)$:

$$T_b(x) \equiv \langle T(x) \rangle \equiv \int_0^{+\infty} t\mathbb{P}_{\text{fvt}}(t)\, \mathrm{d}t = -\int_0^{+\infty} t\partial_t G(x,t)\, \mathrm{d}t =$$

$$= -tG(x,t)\Big|_{t=0}^{t=+\infty} + \int_0^{+\infty} G(x,t)\,\mathrm{d}t \underset{(a)}{=} \langle G(x)\rangle \tag{4.17}$$

In (a) we used that $tG(x,t)$ vanishes at $t=0$ and also at $t=+\infty$, because the particle will eventually reach $x=b$ if given infinite time to do so. It is not clear if $G(x,t) \xrightarrow[t\to\infty]{} 0$ *faster* than $t\to\infty$, so that $tG(x,t) \xrightarrow[t\to\infty]{} 0$. Here, we will just assume it, as it is physically reasonable.

Then, we need to translate once again everything to expressions involving $T_b(x)$. Fortunately, this time it is much quicker. To get the differential equation, we just integrate (4.14):

$$\int_0^{+\infty} \mathrm{d}t\,\partial_t G(x,t) = A(x)\partial_x \int_0^{+\infty} G(x,t)\,\mathrm{d}t + D(x)\partial_x^2 \int_0^{+\infty} G(x,t)\,\mathrm{d}t$$

And applying (4.17) we get:

$$G(x,t)\Big|_{t=0}^{t=+\infty} = G(x,+\infty) - G(x,0) = -1 = A(x)\partial_x T_b(x) + D(x)\partial_x^2 T_b(x)$$

as $G(x,+\infty)=0$ (no particle lives eternally) and $G(x,0)=0$ (as a particle does not "disappear" immediately for $x\neq b$). Similarly, integrating (4.16) and (4.15) leads to:

$$\begin{cases} A(x)\partial_x T_b(x) + D(x)\partial_x^2 T_b(x) = -1 \\ T_b(x)|_{x=b} = 0 \\ \partial_x T_b(x)|_{x=a} = 0 \end{cases}$$

This is a linear ordinary differential equation. We start by letting $f(x) = \partial_x T_b(x)$, leading to:

$$f'(x) = -\frac{A(x)}{D(x)} f(x) - \frac{1}{D(x)} \qquad f(a) = 0$$

First consider the *homogeneous* equation:

$$A(x)\Phi(x) + D(x)\Phi'(x) = 0$$

This can be solved by separation of variables:

$$A\Phi + D\frac{\mathrm{d}\Phi}{\mathrm{d}x} = 0 \Rightarrow \frac{\mathrm{d}\Phi}{\Phi} = -\frac{A}{D}\,\mathrm{d}x \Rightarrow \ln|\Phi(x)| = -\int_{x_0}^x \frac{A(y)}{D(y)}\,\mathrm{d}y + c$$

where x_0 is a fixed point $\in [a,b]$ (it does not matter which one). Exponentiating:

$$\Phi(x) = \exp\left(-\int_{x_0}^x \frac{A(y)}{D(y)}\,\mathrm{d}y\right) k$$

Where $k = e^c$ will be fixed by the boundary condition $f(a)=0$. First, we need to find the general integral of the inhomogeneous equation - for example by using the method of **variation of parameters**.

Refresher of variation of parameters. Consider the following Cauchy problem:

$$\begin{cases} y' = A(t)y + b(t) \\ y(t_0) = y_0 \end{cases}$$

Suppose we know a solution $\Phi(t)$ of the homogeneous equation $y' = A(t)y$. Then $\Phi' = A\Phi$. We search for a particular solution for the full equation in the form $\tilde{\varphi}(t) = \Phi(t)c(t)$. Substituting in the equation:

$$\Phi' c + c'\Phi = A\Phi c + c'\Phi = A\Phi c + b \Rightarrow c' = \Phi^{-1} b$$

This can be integrated to find c, and then $\tilde{\varphi}$. Then, the general integral will be the sum of the homogeneous solution $\Phi(t)$ and the particular one $\tilde{\varphi}$. Imposing the boundary condition will lead to the general integral:

$$\varphi(t) = \Phi(t)\Phi(t_0)^{-1} y_0 + \Phi(t)\int_{t_0}^t \Phi(\tau)^{-1} b(\tau)\,\mathrm{d}\tau \tag{4.18}$$

Applying formula (4.18) leads to the desired $f(x)$:

$$\begin{aligned}
f(x) &= \Phi(x)\Phi(a)\cdot 0 + \Phi(x)\int_a^x \mathrm{d}z\,\Phi(z)^{-1}\left[-\frac{1}{D(z)}\right] = \\
&= \exp\left(-\int_{x_0}^x \frac{A(y)}{D(y)}\,\mathrm{d}y\right)\int_a^x -\frac{\mathrm{d}z}{D(z)}\exp\left(+\int_{x_0}^z \frac{A(y)}{D(y)}\,\mathrm{d}y\right) = \\
&= -\int_a^x \frac{\mathrm{d}z}{D(z)}\exp\left(+\int_x^z \frac{A(y)}{D(y)}\,\mathrm{d}y\right)
\end{aligned}$$

Recall that $f(x) = \partial_x T_b(x)$, with $T_b(b) = 0$. So, to find $T_b(x)$ we need one last integration:

$$T_b(x) = \int_{x_0}^x \mathrm{d}y\,f(y) + c$$

Imposing $T_b(b) = 0$ leads to:

$$T_b(b) = \int_{x_0}^b \mathrm{d}y\,f(y) + c \overset{!}{=} 0 \Rightarrow c = -\int_{x_0}^b \mathrm{d}y\,f(y)$$

Leading to:

$$T_b(x) = \int_b^x \mathrm{d}y\,f(y) = \int_x^b \mathrm{d}y\int_a^y \frac{\mathrm{d}z}{D(z)}\exp\left(-\int_z^y \mathrm{d}v\,\frac{A(v)}{D(v)}\right) \tag{4.19}$$

4.2.1 Escape from a potential well

Let's now use (4.19) to solve the problem we started from. So, suppose to have a potential $U(x)$ with a local minimum at $x = c$, and a local maximum at $x = d$, with $c < d$. Consider a particle starting at $x = c$. We wish to compute the average first visit time of d, denoted with $\langle T(c \to d)\rangle$. This can be done by redefining the system as the half-line $[-\infty, d]$, with $x = -\infty$ being a *reflective* boundary, and $x = d$ an *absorbing* one. We can do this because we are not interested in the behaviour *after* passing d, but just in the mean arrival times. So $A(x) = -\partial_x U(x)/\gamma$. Supposing to be at equilibrium, $D(x) \equiv D = 1/(\gamma B)$. Letting $a = -\infty$ and $b = d$ leads to:

$$
\begin{aligned}
T_d(x) &= \int_x^d \mathrm{d}y \int_{-\infty}^{y} \beta\gamma \, \mathrm{d}z \exp\left(-\int_z^y - \mathrm{d}v \, \frac{\partial_v U(v)}{\gamma} \gamma\beta \right) = \\
&= \beta\gamma \int_x^d \mathrm{d}y \int_{-\infty}^{y} \mathrm{d}z \exp(\beta[U(y) - U(z)]) = \\
&= \beta\gamma \int_x^d \mathrm{d}y \, e^{\beta U(y)} \underbrace{\int_{-\infty}^{y} \mathrm{d}z \, e^{-\beta U(z)}}_{e^{F(y)}} = \beta\gamma \int_x^d \mathrm{d}y \, e^{\beta U(y) + F(y)}
\end{aligned}
$$

It is not possible to evaluate this integral in the general case. However, in the limit $\beta \to \infty$ ($T \to 0$) we can use the saddle-point approximation.
Recall Laplace's formula:

$$
\int_a^b e^{Mf(x)} \, \mathrm{d}x \underset{M \to +\infty}{\approx} \sqrt{\frac{2\pi}{M|f''(x_0)|}} e^{Mf(x_0)}
$$

where $f'(x_0) = 0$ and $f''(x_0) < 0$.
For the integral in $\mathrm{d}z$, $f(z) = -U(z)$. We search for a maximum of $f(z)$, i.e. a minimum of $U(z)$, which is $z = c$. So:

$$
\int_{-\infty}^{y} e^{-\beta U(z)} \, \mathrm{d}z = \sqrt{\frac{2\pi}{\beta U''(c)}} e^{-\beta U(c)}
$$

This is a constant, and can be brought outside the integral over $\mathrm{d}y$. Then, by applying Laplace's formula once again:

$$
\int_c^d \mathrm{d}y \, e^{\beta U(y)} = \sqrt{\frac{2\pi}{\beta |U''(d)|}} e^{\beta U(d)}
$$

as now $f(y) = U(y)$, and U has a local maximum in $y = d$. Finally, this leads to:

$$
T_d(c) \underset{T \to 0}{\approx} \frac{2\pi\gamma}{\sqrt{U''(c)|U''(d)|}} \exp\left(\beta[U(d) - U(c)]\right) \tag{4.20}
$$

Note that the mean transition time from c to d diverges exponentially as the barrier's height $U(d) - U(c)$ rises. Equivalently, the *escape transition rate* $1/T_d(c) \to 0$.

4.3 Feynman Path Integral

We finish our discussion about the diffusion formalism noting several correspondences with quantum processes.

Recall the Schödinger equation:

$$
\begin{aligned}
i\hbar\frac{\partial}{\partial t}\psi(x,t) &= -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x,t) + V(x)\psi(x,t) = \\
&= H(x,\partial_x^2,t)\psi(x,t)
\end{aligned}
$$

where H is the *Hamiltonian* operator:

$$
H(x,\partial_x^2,t) \equiv -\frac{\hbar^2}{2m}\partial_x^2 + V(x,t)
$$

If we consider a free particle ($V(x,t) \equiv 0$), the Schrödinger equation becomes:

$$
\partial_t\psi = i\frac{\hbar}{2m}\partial_x^2\psi \qquad \psi(x,0) = \delta(x-x_0) \tag{4.21}
$$

which is very similar to the diffusion equation:

$$
\partial_t W(x,t) = D\partial_x W(x,t) \qquad W(x,t|x_0,0)\Big|_{t=0} = \delta(x-x_0) \tag{4.22}
$$

In fact, we can map (6.24) to (4.22) by defining a *quantum diffusion coefficient* $D_{QM} = i\hbar/(2m)$.

Does this mean that all properties of the diffusion equation - and its solution - can be mapped to the quantum case? Unfortunately, the answer is a bit complex.
Recall that the solution of (4.22) for a particle initially starting in x_0 at t_0 is:

$$
W(x,t|x_0,t_0) = \frac{1}{\sqrt{4\pi D(t-t_0)}}\exp\left(-\frac{(x-x_0)^2}{4D(t-t_0)}\right) \tag{4.23}
$$

By substituting $D \leftrightarrow D_{\text{QM}}$ we can construct the analogous *quantum* solution:

$$
\psi(x,t) = \sqrt{\frac{2m}{4\pi(t-t_0)i\hbar}}\exp\left(i\frac{m}{2\hbar}\frac{(x-x_0)^2}{t-t_0}\right) \tag{4.24}
$$

Note that now the exponential argument is *complex*, making basic properties of (4.23) not-trivial. For example, if $t \to t_0$, the exponential in (4.23) tends to a δ:

$$
\lim_{t\to t_0} W(x,t|x_0,t_0) = \delta(x-x_0)
$$

giving back the starting distribution, as expected.
The same, however, does not happen for (4.24), given the presence of the i. Nonetheless, it is true that in the limit $t \to t_0$, (4.24) is a *infinitely oscillating function*, meaning that it is 0 *almost* everywhere. This can be proven by using more sophisticated techniques, such as the *stationary phase approximation*.

What about path integrals? If we start with the usual definition and make the substitution $D \leftrightarrow D_{\text{QM}}$ we get:

$$
\psi(x,t) = \langle\delta(x(t)-x)\rangle_W =
$$

$$= \int_{\mathbb{R}^T} \prod_{\tau=0^+}^{t} \frac{\mathrm{d}x(\tau)}{\sqrt{4\pi D_{\mathrm{QM}}\,\mathrm{d}\tau}} \exp\left(-\frac{1}{4D_{\mathrm{QM}}}\int_0^t \left[\frac{\mathrm{d}x(\tau)}{\mathrm{d}\tau}\right]^2 \mathrm{d}\tau\right)\delta(x(t)-x) =$$

$$= \int_{\mathbb{R}^T} \prod_{\tau=0^+}^{t} \frac{\mathrm{d}x(\tau)}{\sqrt{4\pi D_{QM}\,\mathrm{d}t}} \exp\left(\frac{i}{\hbar}\frac{1}{2}m\int_0^t \left[\frac{\mathrm{d}x(\tau)}{\mathrm{d}\tau}\right]^2 \mathrm{d}\tau\right)\delta(x(t)-x)$$

Note that now *trajectories* are weighted by a *complex number*. This means that they *are not probabilities* - and in particular, we cannot use Kolmogorov extension theorem to prove the existence of such a measure as the *continuum limit* of a measure defined on *discretized paths*.
However, we note that in the limit $\hbar \to 0$, the integral can be approximated with the saddle-point method, which returns the *classical trajectory* - the one where the *phases oscillate slowly*.
In fact, it can be proven that *QM* cannot be derived by statistical mechanics alone: quantum "noise" is very much different from thermal "noise"!

Consider now the more general case of non-zero potential:

$$\frac{\partial}{\partial t}\psi(x,t) = i\frac{\hbar}{2m}\partial_x^2\psi(x,t) - \frac{iV(x)}{\hbar}\psi(x,t)$$

which is just the quantum evaluated version of the Fokker-Planck equation:

$$\partial_t W(x,t) = D\partial_x^2 W(x,t) - V(x)W(x,t)$$

with the substitutions:

$$\begin{aligned} D &\to D_{QM} = \frac{i\hbar}{2m} \\ V &\to \frac{i}{\hbar}V \end{aligned} \tag{4.25}$$

The solution we obtained from discussing the diffusion process is:

$$\begin{aligned} W(x,t|x_0,t_0) &= \langle \exp\left(-\int_0^t V(x(\tau))\,\mathrm{d}\tau\right)\delta(x(t)-x)\rangle_W = \\ &= \int_{\mathbb{R}^T} \prod_{\tau=0^+}^{t} \frac{\mathrm{d}x(\tau)}{\sqrt{4D\pi\,\mathrm{d}\tau}} \exp\left(-\frac{1}{4D}\int_0^t \dot{x}^2(\tau)\,\mathrm{d}\tau - \int_0^t V(x(\tau))\,\mathrm{d}\tau\right)\delta(x(t)-x) \end{aligned}$$

Applying (4.25) we arrive to the **Feynman path integral**:

$$\psi(x,t) = \int_{\mathbb{R}^T} \prod_{\tau=0^+}^{t} \frac{\mathrm{d}x(\tau)}{\sqrt{4\pi D_{QM}\,\mathrm{d}\tau}} \exp\left(\frac{i}{\hbar}\int_0^t \mathrm{d}\tau \underbrace{\left[\frac{\dot{x}^2(\tau)}{2} - V(x(\tau))\right]}_{L(\dot{x},x)}\right)\delta(x(t)-x) \tag{4.26}$$

To compute it we can resort to variational methods. We define the action functional S as:

$$S \equiv \int_0^t \mathrm{d}\tau\, L(\dot{x}(\tau), x(\tau))$$

Note that the Feynman path integral *weights* every trajectory with the following quantity:

$$\exp\left(\frac{i}{\hbar}S\left(\{x(\tau)\}_{\tau\in[0,t]}\right)\right)$$

Then, according to the variational method, we can approximate $\psi(x,t)$ by evaluating it only for the *most contributing trajectory*, i.e. the one that *stationarizes* S: $\delta S = 0$, implying:

$$x_c\colon \frac{\partial L}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}}\Big|_{x_c} = 0$$

Part II

Baiesi's Lectures

Chapter 5

Gaussian integrals

5.1 Moments and Generating Functions

(Lesson 3 of 10/10/19) Compiled: October 13, 2020

Consider a continuous function $f\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto f(x)$. The n-th **moment** of f about a point $c \in \mathbb{R}$ is defined as the integral:

$$\mu_n = \int_{-\infty}^{\infty} (x-c)^n f(x) \, \mathrm{d}x$$

Moments provide a way to quantify, in a certain sense, the *shape* of f. For example, if $f(x)$ is a linear density ($[\mathrm{kg\,m^{-1}}]$), then the 0-th moment is the total mass, the first one (with $c = 0$) is the center of mass, and the second is the *moment of inertia*.

Moments are especially useful if $f(x)$ is a probability density function (pdf), i.e. a non-negative normalized function. In this case the first moment about 0 is the **mean**:

$$\mu_1 \equiv \int_{-\infty}^{\infty} x f(x) \, \mathrm{d}x = \mathrm{E}[X] \equiv \mu; \qquad X \sim f$$

where X is a random variable sampled from f. Note that, if not specified, a moment is intended to be centered around $c = 0$ (it is a *raw moment* or *crude moment*).

The *central second moment*, that is μ_2 with $c = \mu$ is the **variance**:

$$\int_{-\infty}^{\infty} (x-\mu)^2 f(x) \, \mathrm{d}x \equiv \mathrm{E}[(X-\mu)^2] = \mathrm{Var}[X]$$

A **moment-generating function** of a real-valued random variable is a certain function $f\colon \mathbb{R}^n \to \mathbb{R}$, $\boldsymbol{x} \mapsto f(\boldsymbol{x})$ that can be used to *compute* the moment of the distribution where X comes from.
More precisely, for a random variable X, the moment-generating function M_X is defined as:

$$M_X(t) \equiv \mathrm{E}[e^{tX}], \quad t \in \mathbb{R} \tag{5.1}$$

In fact, recall that:

$$e^{tX} = 1 + tX + \frac{t^2X^2}{2!} + \dots$$

Hence, as the *expected value* is a linear operator:

$$\begin{aligned} M_X(t) = \mathrm{E}[e^{tX}] &= 1 + t\,\mathrm{E}[X] + \frac{t^2\,\mathrm{E}[X^2]}{2!} + \dots = \\ &= 1 + t\mu_1 + \frac{t^2\mu_2}{2!} + \dots \end{aligned}$$

Note that the distribution's moments are the coefficients of the power series that defines $M_X(t)$.

In fact, the more general definition of a **generating function** is that of a power-series with "hand-picked" coefficients a_n, such that by simply knowing the function one can compute a_n in an iterative way.

To recover a certain μ_n we start by differentiating M_X n times with respect to t, such that the first $n-1$ terms vanish:

$$\frac{\mathrm{d}^n}{\mathrm{d}t^n} M_X(t) = \underbrace{\frac{n(n-1)\dots 1}{n!}}_{=1}\mu_n + \frac{(n+1)n\dots 2}{(n+1)!} t\mu_{n+1} + \dots$$

Then, by setting $t = 0$, all μ_r with $r > n$ vanish, leaving only the desired μ_n:

$$\frac{\mathrm{d}^n}{\mathrm{d}t^n} M_X(t)\Big|_{t=0} = \mu_n$$

Finally, we note that a moment-generating function can be constructed even for a multi-dimensional vector $\boldsymbol{X} = (X_1, \dots, X_n)^T$ of random variables, by simply taking a scalar product in the exponential:

$$M_{\boldsymbol{X}}(\boldsymbol{t}) \equiv \mathrm{E}\left(e^{\boldsymbol{t}^T\boldsymbol{X}}\right) \qquad \boldsymbol{t} \in \mathbb{R}^n$$

5.2 Multivariate Gaussian

Consider now a normal pdf in $d = 1$:

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

We denote a random variable sampled from $f(x; \mu, \sigma)$ as $X \sim \mathcal{N}(\mu, \sigma)$.
Suppose that we have multiple random variables $\{X_i\}_{i=1,\dots,n}$, each normally distributed ($X_i \sim \mathcal{N}(\mu_i, \sigma_i)$), with covariance matrix $\Sigma \in \mathbb{R}^{n\times n}$ defined as:

$$\Sigma_{ij} = \mathrm{E}[(X_i - \mu_i)(X_j - \mu_j)]$$

Their joint pdf is given by a **multivariate normal distribution** :

$$f(x_1, \dots, x_n; \boldsymbol{\mu}, \Sigma) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\boldsymbol{x} - \boldsymbol{\mu})\right)$$

5.3 Moments and Gaussians

We want now to compute the moment generating function for a multivariate gaussian, that is the value of the integral:

$$M_{\boldsymbol{X}}(\boldsymbol{t}) = \int_{\mathbb{R}^n} e^{\boldsymbol{t}\cdot\boldsymbol{x}} f(\boldsymbol{x}; \boldsymbol{\mu}, \Sigma)\, \mathrm{d}^n x \tag{5.2}$$

Let's start from the easiest case, and work our way out to the most general one.

Recall that the **gaussian integral**, i.e. the 0-th moment of a normal univariate distribution is:

$$\int_{-\infty}^{\infty} \exp\left(-\frac{a}{2}x^2\right) \mathrm{dx} = \sqrt{\frac{2}{a}\pi}$$

Proof. The integral as is can't be computed in terms of elementary functions. However, its square can be calculated:

$$\left(\int_{-\infty}^{\infty} \mathrm{d}x \exp\left(-\frac{a}{2}x^2\right)\right)^2 = \int_{-\infty}^{\infty} \mathrm{d}x \int_{-\infty}^{\infty} \mathrm{d}y \exp\left(-\frac{2}{2}(x^2+y^2)\right)$$

Transforming to polar coordinates:

$$= \int_0^{2\pi} \mathrm{d}\theta \int_0^{\infty} \mathrm{d}r \exp\left(-\frac{a}{2}r^2\right) r = -\frac{2\pi}{a} \exp\left(-\frac{a}{2}r^2\right)\Big|_0^{\infty} = \frac{2\pi}{a}$$

and we arrive at the desired result by simply taking the square root.

Consider now the integral of the multivariate case, with $\boldsymbol{\mu} = \mathbf{0}$ (meaning we applied a translation from the general case):

$$Z(\Sigma) = \int_{\mathbb{R}^n} \mathrm{d}^n\boldsymbol{x} \exp\left(-\frac{1}{2}\boldsymbol{x}^T\Sigma^{-1}\boldsymbol{x}\right)$$

Notice that the inverse of the covariance matrix $\Sigma^{-1} \equiv A$ is a symmetric positive-definite matrix, thus can be used to define a quadratic form:

$$\mathbb{A}(\boldsymbol{x}) = \sum_{i,j=1}^{n} x_i A_{ij} x_j$$

The integral can be computed by applying a change of variables, *rotating* $\boldsymbol{x}$ such that A becomes diagonal:

$$\boldsymbol{y} = O\boldsymbol{x}; \qquad O \in \mathbb{R}^{n\times n}; \quad O^T = O^{-1},\ \det(O) = 1$$

where O is an orthogonal matrix, with a set of orthogonal eigenvectors of A as columns, such that:

$$OAO^{-1} = \mathrm{diag}(a_1, \ldots, a_n)$$

with a_i being the eigenvalues of A.
Note that, as $\det(O) = 1$, the volume element in the integral does not change. So, by substituting:

$$\boldsymbol{x} = O^{-1}\boldsymbol{y}; \qquad \boldsymbol{x}^T = \boldsymbol{y}^T(O^{-1})^T = \boldsymbol{y}^T O$$

in the integral, we get:

$$\begin{aligned} Z(A) &= \int_{\mathbb{R}^n} \mathrm{d}^n\boldsymbol{y} \exp\left(-\frac{1}{2}\boldsymbol{y}^T OAO^T \boldsymbol{y}\right) = \int_{\mathbb{R}^n} \mathrm{d}^n\boldsymbol{y} \exp\left(-\frac{1}{2}\sum_{i=1}^n a_i y_i^2\right) = \\ &= \prod_{i=1}^n \int_{\mathbb{R}} \mathrm{d}y_i \exp\left(-\frac{1}{2}a_i y_i^2\right) = (2\pi)^{n/2}\prod_{i=1}^n a_i^{-1/2} \underset{(a)}{=} (2\pi)^{n/2}(\det(A))^{-1/2} \end{aligned} \tag{5.3}$$

where in (a) we noted that the determinant of a matrix is the product of its eigenvalues.

We are now ready to consider the more general case of (5.2), by simply adding a linear term in the exponential of $Z(A)$:

$$Z(A, \boldsymbol{b}) \equiv \int_{-\infty}^{\infty} \mathrm{d}^n\boldsymbol{x} \exp\left(-\frac{1}{2}\mathbb{A}(\boldsymbol{x}) + \boldsymbol{b}\cdot\boldsymbol{x}\right) \qquad \boldsymbol{b}\cdot\boldsymbol{x} = \sum_{i=1}^n b_i x_i \tag{5.4}$$

To compute this integral, a trick is to translate the maximum of the exponential to the origin. So we start by differentiating:

$$\frac{\partial}{\partial x_i}\left(\frac{1}{2}\mathbb{A}(\boldsymbol{x}) - \boldsymbol{b}\cdot\boldsymbol{x}\right) \overset{!}{=} 0 \quad \forall i \tag{5.5}$$

Note that:

$$\begin{aligned} \frac{\partial}{\partial x_i}\mathbb{A}(\boldsymbol{x}) &= \frac{\partial}{\partial x_i}\sum_{ab} x_a A_{ab} x_b = \sum_{ab}\delta_{ai}A_{ab}x_b + \sum_{ab} x_a A_{ab}\delta_{bi} = \\ &= \sum_b A_{ib}x_b + \sum_a x_a A_{ai} \end{aligned}$$

By renaming the first summation variable to a, we get:

$$= \sum_a (A_{ia} + A_{ai})x_a \underset{(b)}{=} 2\sum_a A_{ia}x_a = 2A\boldsymbol{x}$$

where in (b) we used the fact that A is symmetrical ($A_{ij} = A_{ji}$).
Substituting in (5.5):

$$\frac{1}{\cancel{2}}\cancel{2}\sum_j A_{ij}x_j = b_i \quad \forall i \underset{(c)}{\Leftrightarrow} A^T\boldsymbol{x} = \boldsymbol{b} \underset{(d)}{\Leftrightarrow} \boldsymbol{x}^* = A^{-1}\boldsymbol{b}$$

In (c) we noted that b_i is the scalar product between the i-th column of A and $\boldsymbol{x}$, leading to the transpose in the matrix notation. Of course, as $A = A^T$, in (d) we simply dropped the transpose.

We can now apply the coordinate change:

$$\boldsymbol{x} = \boldsymbol{x}^* + \boldsymbol{y}$$

Substituting in the exponential argument:

$$-\frac{\mathbb{A}(\boldsymbol{x})}{2} + \boldsymbol{b}\cdot\boldsymbol{x} = -\frac{1}{2}\boldsymbol{x}^T A\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{b} = -\frac{1}{2}(\boldsymbol{x}^* + \boldsymbol{y})^T A(\boldsymbol{x}^* + \boldsymbol{y}) + (\boldsymbol{x}^* + \boldsymbol{y})^T\boldsymbol{b} =$$
$$= -\frac{1}{2}\left[\boldsymbol{x}^{*T}A\boldsymbol{x}^* + \boldsymbol{y}^T A\boldsymbol{y} + \cancel{\boldsymbol{x}^{*T}A\boldsymbol{y}} + \cancel{\boldsymbol{y}^T A\boldsymbol{x}^*}\right] + \boldsymbol{x}^{*T}A\boldsymbol{x}^* + \cancel{\boldsymbol{y}^T A\boldsymbol{x}^*} \tag{5.6}$$

Note, in fact, that $\boldsymbol{y}^T A\boldsymbol{x}^* = (\boldsymbol{x}^{*T}A^T\boldsymbol{y})^T = (\boldsymbol{x}^{*T}A\boldsymbol{y})^T$ because A is symmetric, and then $(\boldsymbol{x}^{*T}A\boldsymbol{y})^T = \boldsymbol{x}^{*T}A\boldsymbol{y}$ because they are scalars.
Then:

$$\boldsymbol{x}^{*T}A\boldsymbol{x}^* = (A^{-1}\boldsymbol{b})^T AA^{-1}\boldsymbol{b} = \boldsymbol{b}^T(A^{-1})^T\boldsymbol{b} = \boldsymbol{b}^T A^{-1}\boldsymbol{b} = \boldsymbol{b}\cdot\boldsymbol{x}^*$$

And substituting in (5.6):

$$-\frac{\mathbb{A}(\boldsymbol{x})}{2} + \boldsymbol{b}\cdot\boldsymbol{x} = -\frac{1}{2}\boldsymbol{y}^T A\boldsymbol{y} + \underbrace{\frac{1}{2}\boldsymbol{b}\cdot\boldsymbol{x}^*}_{\omega_2(\boldsymbol{b})}$$

To simplify notation, let's define:

$$w_2(\boldsymbol{b}) = \frac{1}{2}\sum_{i,j=1}^n b_i(A^{-1})_{ij}b_i = \frac{1}{2}\boldsymbol{b}\cdot\boldsymbol{x}^* \tag{5.7}$$

As the change of variables involves only a translation by a constant value, the volume element in the integral does not change, leading to:

$$Z(A,\boldsymbol{b}) = \int_{-\infty}^{\infty} \mathrm{d}^n\boldsymbol{y}\exp\left(-\frac{\mathbb{A}(\boldsymbol{y})}{2} + \omega_2(\boldsymbol{b})\right)$$

Note that $\omega_2(\boldsymbol{b})$ is constant, thus can be extracted from the integral:

$$= e^{\omega_2(\boldsymbol{b})}\int_{-\infty}^{\infty} \mathrm{d}^n\boldsymbol{y}\exp\left(-\frac{\mathbb{A}(\boldsymbol{y})}{2}\right) \underset{(5.3)}{=} e^{\omega_2(\boldsymbol{b})}(2\pi)^{n/2}(\det A)^{-1/2} \tag{5.8}$$

Another way to solve the integral for $Z(A,\boldsymbol{b})$ is by using the matrix equivalent of "completing the square". We start by considering the argument of the exponential in (5.4):

$$-\frac{1}{2}(\boldsymbol{x}^T A\boldsymbol{x} - 2\boldsymbol{b}^T\boldsymbol{x})$$

$\boldsymbol{x}^T A\boldsymbol{x}$ has the role of the square, and $-2\boldsymbol{b}^T\boldsymbol{x}$ that of the double product.
We can then sum and subtract a constant vector $\boldsymbol{c}$ in order to rewrite:

$$\boldsymbol{x}^T A\boldsymbol{x} - 2\boldsymbol{b}^T\boldsymbol{x} + \boldsymbol{c} - \boldsymbol{c} = \boldsymbol{y}^T A\boldsymbol{y} - \boldsymbol{c}$$

for some $\boldsymbol{y}\in\mathbb{R}^n$.
Comparing to a generic square:

$$(\boldsymbol{a}+\boldsymbol{b})^T A(\boldsymbol{a}+\boldsymbol{b}) = \boldsymbol{a}^T A\boldsymbol{a} + \boldsymbol{b}^T A\boldsymbol{b} + 2\boldsymbol{a}^T A\boldsymbol{b}$$

we note that $\boldsymbol{a} = \boldsymbol{x}$ and $\boldsymbol{b} = -A^{-1}\boldsymbol{b}$, leading to:

$$\boldsymbol{x}^T A\boldsymbol{x} - 2\boldsymbol{b}^T\boldsymbol{x} = (\boldsymbol{x} - A^{-1}\boldsymbol{b})^T A(\boldsymbol{x} - A^{-1}\boldsymbol{b}) - \boldsymbol{b}^T A^{-1}\boldsymbol{b}$$

Defining $A^{-1}\boldsymbol{b} \equiv \boldsymbol{x}^*$ and $\boldsymbol{y} = \boldsymbol{x} - \boldsymbol{x}^*$ then leads to the same calculations as before.

Exercise 5.3.1 (Multivariate Gaussian Integral)**:**

Compute $Z(A)$ and $Z(A, \vec{b})$ with:

$$A = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}; \quad b = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Note that $\det A = 8$, and:

$$A^{-1} = \frac{1}{8}\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$$

So, by simply using (5.3) and (5.8):

$$\begin{aligned} Z(A, 0) &= \frac{(2\pi)^{2/2}}{\sqrt{8}} = \frac{\pi}{\sqrt{2}} \\ \frac{1}{2}\begin{pmatrix} 1 & 0 \end{pmatrix}\frac{1}{8}&\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{3}{16} \\ Z(A, \boldsymbol{b}) &= \frac{\pi}{\sqrt{2}}\exp\left(\frac{3}{16}\right) \end{aligned}$$

5.3.1 Gaussian expectation values

The result in (5.8) is exactly what we need to compute the moment generating function for the multivariate normal (5.2).

So, we can finally compute moments:

$$\langle x_{k_1} x_{k_2} \dots x_{k_l} \rangle \equiv \frac{1}{Z(A)} \int \mathrm{d}^n\boldsymbol{x}\, x_{k_1} x_{k_2} \dots x_{k_l} \exp\left(-\frac{1}{2}\mathbb{A}(\boldsymbol{x})\right)$$

by simply deriving the generating function $Z(A, \boldsymbol{b})$ with respect to certain variables in $\boldsymbol{b}$. For example:

$$\langle x_k \rangle = \frac{1}{Z(A)}\frac{\partial}{\partial b_k} Z(A, \vec{b}) = \frac{1}{Z(A)} \int \mathrm{d}^n\boldsymbol{x}\, x_k \exp\left(-\frac{\mathbb{A}(\boldsymbol{x})}{2} + \boldsymbol{b}^T\boldsymbol{x}\right)$$

For the general case:

$$\langle x_{k_1} x_{k_2} \dots x_{k_l} \rangle = (2\pi)^{-n/2}(\det A)^{-1/2}\left[\frac{\partial}{\partial b_{k_1}}\frac{\partial}{\partial b_{k_2}} \dots \frac{\partial}{\partial b_{k_l}} Z(A, \boldsymbol{b})\right]_{\boldsymbol{b}=\boldsymbol{0}} =$$

$$= \frac{\partial}{\partial b_{k_1}} \frac{\partial}{\partial b_{k_2}} \cdots \frac{\partial}{\partial b_{k_l}} e^{w_2(\boldsymbol{b})} \Big|_{\boldsymbol{b}=\boldsymbol{0}}$$

In physics, we say that b_k is "coupled" to x_k, and that $Z(A, \boldsymbol{b})$ is used as "generating function" for $\boldsymbol{x}$.

5.3.2 Wick's Theorem

From the previous formula we know that:

$$\frac{\partial}{\partial b_i} \text{ pulls down a } b_i$$

Explicitly, recall that:

$$\omega_2(\boldsymbol{b}) = \frac{1}{2} \boldsymbol{b}^T A^{-1} \boldsymbol{b}$$

and so:

$$\frac{\partial}{\partial b_i} e^{\omega_2(\boldsymbol{b})} = \frac{1}{2} e^{\omega_2 \boldsymbol{b}} \frac{\partial}{\partial b_i} \sum_{tk} b_t A^{-1}_{tk} b_k = e^{\omega_2 \boldsymbol{b}} \sum_k A^{-1}_{ik} b_k$$

If we now set $\boldsymbol{b} = \boldsymbol{0}$, the result will be 0, meaning that:

$$\langle x_i \rangle = \frac{\partial}{\partial b_i} \frac{Z(A, \boldsymbol{b})}{Z(A)} = 0$$

This result is expected, as in $Z(A, \boldsymbol{b})$ all random variables are centered in 0. However, note that if we derive one more time, with respect to some b_l:

$$\frac{\partial}{\partial b_i} \frac{\partial}{\partial b_l} e^{\omega_2(\boldsymbol{b})} = e^{\omega_2 \boldsymbol{b}} \sum_s A^{-1}_{ls} b_s \sum_k A^{-1}_{ik} b_k + e^{\omega_2 \boldsymbol{b}} A^{-1}_{il}$$

And now, if we set $\boldsymbol{b} = \boldsymbol{0}$, the result may be $\neq 0$.
Note that if we derive one more time we return to the previous situation - and the result will be also 0.

In general, every moment of odd-order is 0, due to the symmetry of the gaussian, we have:

$$\langle x_i x_j x_k \rangle = 0$$

So the expectation value of the product of different random variables, sampled from the same gaussian distribution centered on 0, is only non-zero for an even number of variables. This result is known as the **Wick's theorem** (also known in literature as the **Isserlis theorem**).

By extending this argument, one can find a way to compute the even-order moments, leading to the following formula (which we will not prove):

$$\langle x_{k_1} x_{k_2} \dots x_{k_l} \rangle = \sum_{P \in \sigma(K)} A^{-1}_{k_{P_1} k_{P_2}} A^{-1}_{k_{P_3} k_{P_4}} \cdots A^{-1}_{k_{P_{l-1}} k_{P_l}} = \sum_{P \in \sigma(K)} \langle x_{k_{P_1}} x_{k_{P_2}} \rangle \dots \langle x_{k_{P_{l-1}}} x_{k_{P_l}} \rangle$$

where (k_p, k_q) are a pair of indices from $K = \{k_1, \dots, k_l\}$, and P is a permutation of K, so that (k_{P_1}, k_{P_2}) is the first pair of indices after the permutation P. The sum is over all the distinct ways of partitioning $l = 2s$ variables in pairs to obtain distinct products of s groups.
So, the total number of terms to be added will be $(2s)!/(2^s s!)$ - that is the total number of permutation of $2s$ elements, where the order within couples does not matter (2^s) and neither the order of the couples themselves $(s!)$.
Note that:

$$\frac{(2s)!}{2^s s!} = (2s-1)!! = (2s-1)(2s-3)(2s-5)\dots$$

Where !! denotes the *double factorial*, not to be confused with the factorial of a factorial (which requires brackets: $(a!)!$).

Exercise 5.3.2 (Wick's theorem)**:**

Consider a univariate normal distribution:

$$f(x) = \frac{1}{Z(A)} \exp\left(-\frac{a}{2}x^2\right)$$

Show that:

$$\langle x^2 \rangle = \frac{1}{a}$$
$$\langle x^4 \rangle = \frac{3}{a^2} = 3(\langle x^2 \rangle)^2$$

Here the A matrix is just the scalar $a = \sigma^{-2}$. As the pdf is univariate, there is only one index possible $K = \{1\}$. As $(2-1)!! = 1!! = 1$, there is only one term in the summation, thus:

$$\langle x^2 \rangle = A_{11}^{-1} = \frac{1}{a}$$

For the 4-th order, however, we have more combinations: $(4-1)!! = 3!! = 3 \cdot 1 = 3$. Again, there is only one possible index, so all terms will be the same:

$$\langle x^4 \rangle = A_{11}^{-1}A_{11}^{-1} + A_{11}^{-1}A_{11}^{-1} + A_{11}^{-1}A_{11}^{-1} = \frac{3}{a^2} = 3(\langle x^2 \rangle)^2$$

5.4 Steepest Descent Integrals

It is possible to use gaussian integrals to solve a more general set of integrals, thanks to the *Steepest Descent approximation*.
We start with an integral of the form:

$$I(\lambda) \equiv \int_S \mathrm{d}^n x \exp\left(-\frac{F(\boldsymbol{x})}{\lambda}\right) \tag{5.9}$$

where λ is a small parameter (the approximation is more and more accurate as $\lambda \to 0$), $F(\boldsymbol{x})$ has a global minimum in $\boldsymbol{x_0} \in (a, b)$ and $S \subseteq \mathbb{R}^n$ is a sufficiently large region.
Note that, if λ is lowered, the integral is dominated by the neighborhood of the minimum $\boldsymbol{x_0}$. In fact:

$$h(\boldsymbol{x}) \equiv \exp\left(-\frac{F(\boldsymbol{x})}{\lambda}\right); \quad \frac{h(\boldsymbol{x_0})}{h(\boldsymbol{x})} = \exp\left(-\frac{1}{\lambda}(F(\boldsymbol{x_0}) - F(\boldsymbol{x}))\right)$$

As $F(\boldsymbol{x_0}) - F(\boldsymbol{x}) < 0$, the ratio becomes exponentially higher if $\lambda \to 0$. Basically, for $\lambda \to 0$, the integrand function becomes "more and more similar to a gaussian".

To compute the integral, then, we translate the coordinates about $\boldsymbol{x_0}$:

$$\boldsymbol{x} = \boldsymbol{x_0} + \sqrt{\lambda}\boldsymbol{y} \qquad \mathrm{d}^n\boldsymbol{x} = \lambda^{n/2}\,\mathrm{d}^n\boldsymbol{y}$$

Then we perform a second order Taylor expansion about $\lambda = 0$ and $\boldsymbol{x} = \boldsymbol{x_0}$:

$$\frac{1}{\lambda}F(\boldsymbol{x}) = \frac{1}{\lambda}F(\boldsymbol{x_0}) + \cancel{\frac{1}{\lambda}\sum_i \partial_{x_i} F(\boldsymbol{x_0}) y_i \sqrt{\lambda}} + \frac{1}{\cancel{\lambda}}\frac{1}{2!}\sum_{ij}\partial^2_{x_i x_j} F(\boldsymbol{x_0}) y_i y_j \cancel{\lambda} + O(\lambda^{1/2})$$

where we cancelled the first derivative, as $\boldsymbol{x_0}$ is a stationary point for F.

Substituting back in the integral we get:

$$I(\lambda) = \lambda^{n/2}\exp\left(-\frac{F(\boldsymbol{x_0})}{\lambda}\right)\int_{S'} \mathrm{d}^n\boldsymbol{y}\exp\left[-\frac{1}{2}\sum_{ij}\partial^2_{x_i x_j}F(\boldsymbol{x_0})y_i y_j - R(\boldsymbol{y})\right]$$

This is a gaussian integral $Z(A)$, with A being the Hessian of F evaluated at the minimum $\boldsymbol{x_0}$ (or, equivalently, at the maximum of $-F(\boldsymbol{x})$).
Now, for λ sufficiently small, we can ignore $R(\boldsymbol{y})$ and compute the integral with (5.3), leading to the approximation:

$$I(\lambda) \underset{\lambda \to 0}{\approx} (2\pi\lambda)^{n/2}[\det \partial^2_{x_i x_i} F(\boldsymbol{x_0})]^{-1/2}\exp\left(-\frac{F(\boldsymbol{x_0})}{\lambda}\right) \tag{5.10}$$

Doing this, we implicitly integrated over the entire $\mathbb{R}^n$. This is fine because, for $\lambda \to 0$, the gaussian is "peaked" in a small region around $\boldsymbol{x_0}$, and vanishes exponentially moving further away.

The Steepest Descent approximation generalizes Laplace's method for calculating integrals, which has a much simpler expression for the limited case of univariate integrals:

$$I(s) = \int g(z)e^{sf(z)}dz \underset{s \to \infty}{\approx} \frac{(2\pi)^{1/2} g(z_c) e^{sf(z_c)}}{|sf''(z_c)|^{1/2}} \tag{5.11}$$

with $f, g \in \mathbb{R}$, and z_c is the maximum of f, i.e. $f(z_c) \geq f(z)\,\forall z \in (a, b)$.
This formula is useful in physics: s can model the system's size, and $s \to \infty$ is then the limit for a large system.

Example 14 (Stirling approximation)**:**

We can use the Steepest Descent approximation to derive the formula for the Stirling approximation of factorials.
Recall that a factorial is merely the Γ function evaluated on $\mathbb{N}$:

$$s! = \int_0^\infty x^s e^{-x} dx$$

We then perform a change of variables:

$$x = zs$$

so that:

$$s! = s^{s+1} \int_0^\infty e^{s(\ln z - z)} dz$$

This is an integral in the form:

$$\int \exp\left(-\frac{F(x)}{\lambda}\right)$$

if we let $\lambda = 1/s$ and $F(z) = z - \ln z$. So we need to find the minimum of $F(z)$:

$$\begin{aligned} F'(z) &= \frac{\mathrm{d}}{\mathrm{d}z}(z - \ln z) = 1 - \frac{1}{z} \overset{!}{=} 0 \Rightarrow z_c = 1 \\ F''(z) &= \frac{1}{z^2} \Rightarrow F''(z_c) = 1 > 0 \end{aligned}$$

We can now apply (5.10), leading to:

$$s! \underset{s\to\infty}{\approx} \left(\frac{2\pi}{s}\right)^{1/2} (1)^{1/2} e^{-s} = \sqrt{2\pi} s^{s+\frac{1}{2}} e^{-s}$$

Taking the ln we arrive at the usual form for the Stirling approximation:

$$\ln s! \approx s \ln ns - s + \frac{1}{2}\ln(2\pi s) + O\left(\frac{1}{s}\right)$$

Note that the same result can be obtained by using the much simpler (5.11), with $g(z) \equiv 1$ and $f(z) = \ln z - z$.

Exercise 5.4.1 (Steepest Descent Approximation)**:**

Compute the Steepest Descent Approximation for the following integral (for $s \to \infty$):

$$I(s) = \int_{-\infty}^{\infty} e^{sx - \cosh x} dx$$

By collecting a s in the exponential argument:

$$I(s) = \int_{-\infty}^{\infty} \exp\left(s\left(x - \frac{\cosh x}{s}\right)\right)$$

we can bring back to the form of (5.9) with $F(x) = \cosh x/s - x$ and $\lambda = s^{-1}$.
We find the minimum of $F(x)$ by differentiating:

$$\begin{aligned} F'(x) &= \frac{\sinh x}{s} - 1 \stackrel{!}{=} 0 \Rightarrow x_0 = \sinh^{-1} s \\ F''(x) &= \frac{\cosh x}{s} \Rightarrow F''(x_0) = \frac{\cosh \sinh^{-1} s}{s} = \frac{\sqrt{1+s^2}}{s} > 0 \end{aligned}$$

Finally, by applying (5.10) we obtain the result:

$$\begin{aligned} I(s) \underset{s\to\infty}{\approx} \sqrt{\frac{2\pi}{s}} \sqrt{\frac{s}{\sqrt{1+s^2}}} \exp\left(\frac{\sqrt{1+s^2}}{s} - \sinh^{-1} s\right) = \\ = \frac{\sqrt{2\pi}}{(1+s^2)^{1/4}} \exp\left(\frac{\sqrt{1+s^2}}{s} - \sinh^{-1} s\right) \end{aligned}$$

Note that, for this peculiar case, the simple 1D formula does not work (why?) - and so one should proceed with the general method (full steps: find maximum, second derivative...).

Exercise 5.4.2 (Laplace's formula)**:**

Compute:

$$I(N) = \int_0^{\infty} \cos(x) \exp\left(-N\left[\left(x - \frac{\pi}{3}\right)^2 + \left(x - \frac{\pi}{3}\right)^4\right]\right) dx$$

in the limit $N \to \infty$.

For this exercise we can use Laplace's formula (5.11) with:

$$g(x) = \cos(x) \qquad f(x) = -\left[\left(x - \frac{\pi}{3}\right)^2 + \left(x - \frac{\pi}{3}\right)^4\right]$$

By looking at $f(x)$ one can see directly that it has a global maximum in $x_0 = \pi/3$. In fact:

$$f'(x) = -\left[2\left(x - \frac{\pi}{3}\right) + 4\left(x - \frac{\pi}{3}\right)^3\right] \stackrel{!}{=} 0 \Leftrightarrow x_0 = \frac{\pi}{3}$$

$$f''(x) = -\left[2 + 12\left(x - \frac{\pi}{3}\right)^2\right] \Rightarrow f''(x_0) = -2 < 0$$

And so we arrive at:

$$I(N) \underset{N\to\infty}{\approx} \frac{(2\pi)^{1/2}\cos(\pi/3)e^{N\cdot 0}}{|N(-2)|^{1/2}} = \frac{1}{2}\sqrt{\frac{\pi}{N}}$$

Chapter 6

Integrals of complex variables

In this chapter we discuss several techniques for computing integrals on the complex plane.

6.1 Fourier Transform

One of the most frequent kind of complex integral is given by the *Fourier Transform* (FT). Let $f(x) \in L_2(\mathbb{R})$ be a square-integrable function. Then the Fourier transform maps $f(x)$ to another function $\tilde{f}(k)$ defined as follows:

Fourier transform

$$\mathcal{F}[f(x)](k) = \tilde{f}(k) \equiv \int_{\mathbb{R}} e^{-ikx} f(x)\,\mathrm{d}x \qquad f \in L_2(\mathbb{R}) \tag{6.1}$$

Similarly, it is possible to define the *inverse Fourier transform*, linking $\tilde{f}(k)$ back to $f(x)$:

Inverse Fourier transform

$$\mathcal{F}^{-1}[\tilde{f}(k)](x) = f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ikx} \tilde{f}(k)\,\mathrm{d}k$$

The 2π factor is needed for normalization, so that:

$$\mathcal{F}^{-1}[\mathcal{F}[f(x)](k)](x) = f(x) \tag{6.2}$$

Conventions

As long as (6.2) is satisfied, any different definition of the Fourier transforms is acceptable. For example, it is possible to *switch* the signs in the e^{ikx}, or split differently the normalization factor between $\mathcal{F}$ and $\mathcal{F}^{-1}$.

6.1.1 Refresher on functional analysis

The definition (6.1) is quite limited, as several interesting functions are not in $L_2(\mathbb{R})$ - for example $\sin(x)$, $\cos(x)$, $\theta(x)$. Fortunately, it is possible to extend the Fourier transform by considering *generalized functions* (**distributions**).

Schwartz space

We start by defining a space $\mathcal{S}(\mathbb{R})$ (Schwartz space) containing all functions $\varphi \in C^\infty(\mathbb{R})$ that are *rapidly decreasing*, i.e. such that $\sup_{x\in\mathbb{R}} |x^\alpha \varphi^{(\beta)}(x)| < \infty$ $\forall \alpha, \beta \in \mathbb{N}$. These are also called *test functions*.

Tempered distributions

Then a **tempered distribution** T is a **continuous linear** mapping $\mathcal{S}(\mathbb{R}) \to$

$\mathbb{R}$. So it is possible to "apply" a distribution T to any test function $\varphi \in \mathcal{S}(\mathbb{R})$, resulting in a real number, denoted with $\langle T, \varphi \rangle$.

The choice of $\mathcal{S}$ is made expressly so that the Fourier transform is a linear and invertible operator on $\mathcal{S}$. However, other choices can be made for the space of test functions. For example, one can take the set $\mathcal{D}$ of all functions with *compact support*, i.e. that vanish (along with all their derivatives) outside a compact region.

We can now see that distributions *generalize* the concept of function. We start by noting that any **locally integrable** function $f\colon \mathbb{R} \to \mathbb{R}$ can be used to define a distribution, by considering its inner product with a test function:

$$\langle T_f, \varphi \rangle \equiv \int_{\mathbb{R}} \mathrm{d}x\, f(x)\varphi(x) \qquad \forall \varphi \in \mathcal{S}(\mathbb{R}) \tag{6.3}$$

Distributions that can be defined like this are called **regular**.

In the **complex** case, where $f\colon \mathbb{R} \to \mathbb{C}$, we instead use the Hermitian inner product:

$$\langle T_f, \varphi \rangle = \int_{\mathbb{R}} \mathrm{d}x\, f(x)^* \varphi(x)$$

where $f(x)^*$ is the complex conjugate of $f(x)$. The choice of the *position* of this conjugate (on the first or second entry) is a convention. Physicists tend to use the first position (due to Dirac notation), while mathematicians the second one.

Not all distributions are regular: in general, it is not possible to find a function $f(x)$ for a generic distribution T such that (6.3) is satisfied. The distributions for which this is not possible are called **singular**.

The simplest (and most important) singular distribution is the **Dirac Delta** $\delta(x)$, defined as follows:

Dirac Delta

$$\langle \delta, \varphi \rangle \equiv \varphi(0) \qquad \varphi \in S(\mathbb{R})$$

In other words, applying the δ to any test function φ returns the value of φ at 0.
In practice, we often write *formally*:

$$\langle \delta, \varphi \rangle = \int_{\mathbb{R}} \delta(x)\varphi(x)\, \mathrm{d}x$$

as if $\delta(x)$ were a function (but keep in mind that it isn't). This expression is often just a *shortcut* for quickly reaching useful results, as we will see in the following.

The point of defining *distributions* is that they provide a way to extend rigorously may operations that cannot be done on normal functions. One such example is differentiation. Given a distribution T, its **distributional derivative** is defined as:

Distributional derivative

$$\langle T', \varphi \rangle \equiv -\langle T, \varphi' \rangle \qquad \forall \varphi \in S(\mathbb{R}) \tag{6.4}$$

This is done so that, for a *regular* distribution T_f, that result comes from integration by parts:

$$\langle T_f', \varphi \rangle = \int_{\mathbb{R}} f'(x)\varphi(x)\,\mathrm{d}x = \cancel{f(x)\varphi(x)\Big|_{-\infty}^{+\infty}} - \int_{\mathbb{R}} f(x)\varphi'(x) = -\langle T_f, \varphi' \rangle \tag{6.5}$$

For a singular distribution we use directly the definition (6.4), as the construction in (6.5) has no meaning (but still, sometimes we will write it nonetheless, as a merely *formal* expression).

In the distributional sense, it is possible to differentiate the **Heaviside function** $\theta(x)$:

Heaviside step function

$$\theta(x) \equiv \begin{cases} 1 & x > 0 \\ \frac{1}{2} & x = 0 \\ 0 & x < 0 \end{cases} \tag{6.6}$$

As $\theta(x)$ is locally integrable, we can define a corresponding distribution - that we denote with the same symbol θ. Then:

$$\begin{aligned} \langle \theta', \varphi \rangle = -\langle \theta, \varphi' \rangle = -\int_{\mathbb{R}} \theta(x)\varphi'(x)\,\mathrm{d}x = -\int_0^{+\infty} \varphi'(x)\,\mathrm{d}x = -[\cancel{\varphi(+\infty)} - \varphi(0)] = \\ = \varphi(0) = \langle \delta | \varphi \rangle \end{aligned} \tag{6.7}$$

So $\theta' = \delta$ *in the distributional sense* - i.e. applying θ' or δ to any test function φ leads to the same result.

6.1.2 Fourier transform of distributions

We are finally ready to extend the **Fourier Transform** to tempered distributions. In fact, $S(\mathbb{R})$ has been chosen[1] such that any $\varphi(x) \in S(\mathbb{R})$ has a well-defined transform $\tilde{\varphi}(k)$. Then we define the Fourier transform of a distribution as follows:

Fourier Transform of distributions

$$\langle \mathcal{F}[T], \varphi \rangle \equiv 2\pi \langle T, \mathcal{F}^{-1}[\varphi] \rangle$$

Again, this comes from the expression for regular distributions:

$$\begin{aligned} \langle \mathcal{F}[T_f], \varphi \rangle &= \int_{\mathbb{R}} \mathrm{d}k\, \{\mathcal{F}[f(x)](k)\}^* \varphi(k) = \int_{\mathbb{R}} \mathrm{d}k \int_{\mathbb{R}} \mathrm{d}x \left[e^{-ikx} f(x)\right]^* \varphi(x) = \\ &= \int_{\mathbb{R}} \mathrm{d}x\, f(x) \int_{\mathbb{R}} \mathrm{d}k\, e^{ikx} \varphi(k) = \int_{\mathbb{R}} 2\pi f(x) \mathcal{F}^{-1}[\varphi(k)](x) = 2\pi \langle T, \mathcal{F}^{-1}[\varphi] \rangle \end{aligned}$$

Note that:

$$\langle \mathcal{F}[T], \mathcal{F}[\varphi] \rangle = 2\pi \langle T, \mathcal{F}^{-1}\mathcal{F}[\varphi] \rangle = 2\pi \langle T, \varphi \rangle \tag{6.8}$$

[1] ∧More precisely, the Fourier transform is an *automorphism* of $\mathcal{S}$, i.e. it is linear and invertible

Delta transform

Finally, we can use all this machinery to compute Fourier transforms of some *generalized functions*. We start with the δ:

$$\langle \mathcal{F}[\delta], \varphi \rangle = 2\pi \langle \delta, \mathcal{F}^{-1}[\varphi] \rangle = 2\pi \mathcal{F}^{-1}[\varphi(x)](0)$$

where:

$$\mathcal{F}^{-1}[\varphi(x)](k) = \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{d}x\, e^{ikx} \varphi(x) \Rightarrow 2\pi \mathcal{F}^{-1}[\varphi(x)](0) = \int_{\mathbb{R}} \mathrm{d}x\, \varphi(x) = \langle 1, \varphi \rangle$$

And so $\mathcal{F}[\delta] = 1$.

Note that the same result could be obtained in a simpler way by treating δ as a "formal function":

$$\mathcal{F}[\delta](k) = \int_{\mathbb{R}} e^{-ikx} \delta(x) = e^{-ik0} = 1$$

This leads to an equivalent definition for the δ "function":

$$\delta(x) = \mathcal{F}^{-1}[1](x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ikx}\, \mathrm{d}k$$

Also, note that:

$$\mathcal{F}[1](k) = \int_{\mathbb{R}} e^{-ikx}\, \mathrm{d}x = \int_{\mathbb{R}} e^{ikx}\, \mathrm{d}x = 2\pi \left(\frac{1}{2\pi} \int_{\mathbb{R}} e^{ikx} \right) = 2\pi\delta(k) \tag{6.9}$$

Heaviside transform

We can use the result for the δ to aid the computation of $\mathcal{F}[\theta]$, where $\theta(x)$ is the regular distribution defined from (6.6). We have already seen in (6.7) that $\theta' = \delta$. So, we can use the formula for the Fourier transform of a derivative (which naturally generalizes to distributions):

Fourier transform of a derivative

$$\mathcal{F}[T'] = ik\tilde{T} \tag{6.10}$$

In our case:

$$\mathcal{F}[\theta'] \underset{(6.7)}{=} \mathcal{F}[\delta] = 1 = ik\tilde{\theta} \tag{6.11}$$

However, (6.11) cannot be used to reconstruct $\tilde{\theta}$ by itself, that is we cannot just "solve by $\tilde{\theta}$" and write:

$$\tilde{\theta}(k) = \frac{1}{ik} \tag{6.12}$$

In fact, consider a different $\theta^*(x) \equiv \theta(x) + c$, with $c \in \mathbb{R}$ constant. Their derivatives coincide, and so formula (6.11) would give the same result for both of them. However:

$$\mathcal{F}[\theta^*(x)](k) = \mathcal{F}[\theta(x)](k) + \mathcal{F}[c](k) = \tilde{\theta}(k) + c\delta(k) \neq \tilde{\theta}(k)$$

So we are missing a δ term, meaning that the correct Fourier transform should be:

$$\tilde{\theta}(k) = \mathcal{P}\left(\frac{1}{ik}\right) + c\delta(k) \tag{6.13}$$

Inversion formula

for some constant c. $\mathcal{P}$ denotes the Cauchy principal value, which needs to be used to "fix" the singularity at $k = 0$ (see the following green boxes for the details).

There are several ways to fix c in (6.13). One of the quickest is to reason *with symmetries.*

1. Fix c (symmetries)

Let f be an even function (i.e. a gaussian). Symmetry is preserved by the Fourier transform, and so:

$$\langle \tilde{\theta}, \tilde{f} \rangle = \mathcal{P} \int_{\mathbb{R}} \frac{1}{ik} \tilde{f}(k) \,\mathrm{d}k + c\langle \delta, \tilde{f} \rangle = c\tilde{f}(0) = c \int_{\mathbb{R}} f(x) \,\mathrm{d}x \tag{6.14}$$

The principal value vanishes because $\tilde{f}$ is even (as f is even). The corresponding scalar product without the Fourier transforms is:

$$\langle \theta, f \rangle = \int_0^{+\infty} f(x) \,\mathrm{d}x \underset{(a)}{=} \frac{1}{2} \int_{\mathbb{R}} f(x) \,\mathrm{d}x \tag{6.15}$$

where in (a) we again used the symmetry of f. Then, recalling (6.8), we have:

$$\langle \tilde{\theta}, \tilde{f} \rangle = 2\pi \langle \theta, f \rangle \Rightarrow c \int_{\mathbb{R}} f(x) \,\mathrm{d}x = \frac{2\pi}{2} \int_{\mathbb{R}} f(x) \,\mathrm{d}x \Rightarrow c = \pi$$

(Note that c depends on the choice we made for the normalization in the Fourier transforms).

2. Fix c with symmetries and $\mathrm{sgn}(x)$

A similar argument can be made noting that $\theta(x)$ is just a scaled and shifted sgn function, which is odd:

$$\theta(x) = \frac{1}{2} + \frac{1}{2}\,\mathrm{sgn}(x) \qquad \mathrm{sgn}(x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases}$$

By linearity we have:

$$\tilde{\theta}(k) = \mathcal{F}\left(\frac{1}{2}\right) + \frac{1}{2}\mathcal{F}[\mathrm{sgn}(x)](k) \tag{6.16}$$

Noting that $\mathrm{sgn}' = 2\delta$ and using (6.10) leads to:

$$2 = ik\mathcal{F}[\mathrm{sgn}](k)$$

Inverting with (6.13), we have:

$$\mathcal{F}[\mathrm{sgn}](k) = \mathcal{P}\left(\frac{2}{ik}\right) + c\delta(k) = \mathcal{P}\left(\frac{2}{ik}\right)$$

As this time c must be 0, otherwise $\mathcal{F}[\mathrm{sgn}](k)$ wouldn't be odd (the δ is *even*). Substituting in (6.16) we have:

$$\tilde{\theta}(k) = \frac{1}{2}\underbrace{\mathcal{F}[1]}_{2\pi} + \frac{1}{\cancel{2}}\mathcal{P}\left(\frac{\cancel{2}}{ik}\right) = \mathcal{P}\left(\frac{1}{ik}\right) + \pi\delta(k)$$

Why is (6.12) wrong? There are two main reasons:

- $1/(ik)$ is not locally integrable (as it diverges for $k = 0$), so it cannot be used to define a distribution, such as $\tilde{\theta}$. This can be solved by using the *principal part* of $1/(ik)$ instead.

- The most general solution to the equation $xT = 1$, where T is a tempered distribution, is not just $T = \mathcal{P}(1/x)$, but:

$$T = \mathcal{P}\left(\frac{1}{x}\right) + c\delta$$

  for some constant $c \in \mathbb{R}$.

First, to be precise, the product of a function, such as $f(x) = x$, with a distribution T is *defined* as the following distribution:

$$\langle f(x)T, \varphi\rangle \equiv \langle T, f(x)\varphi\rangle \tag{6.17}$$

where $f(x)$ must be such that $f(x)\varphi \in \mathcal{S}\ \forall\varphi \in \mathcal{S}$, which is indeed the case for any polynomial.
Now consider the *distributional* equation $xT = 1$. If we apply *both sides* to some test function φ, we have:

$$\langle T, x\varphi\rangle = \langle 1, \varphi\rangle = \int_{\mathbb{R}} \varphi(x)\,\mathrm{d}x \tag{6.18}$$

The problem of *finding* T satisfying (6.18) is called the (distributional) **division problem**. To solve it, we want to reduce the equation to something in the form of $xT' = 0$, that can then be solved. So we rewrite the rhs as follows:

$$\int_{\mathbb{R}} \varphi(x)\,\mathrm{d}x = \lim_{\epsilon\to 0^+} \int_{\mathbb{R}\setminus[-\epsilon,\epsilon]} \varphi(x)\,\mathrm{d}x = \lim_{\epsilon\to 0^+} \int_{\mathbb{R}\setminus[-\epsilon,\epsilon]} \frac{x\varphi(x)}{x}\,\mathrm{d}x$$

Then we define the **principal value distribution** $\mathcal{P}(1/x)$ as:

$$\langle \mathcal{P}\left(\frac{1}{x}\right), \varphi\rangle = \lim_{\epsilon\to 0^+} \int_{\mathbb{R}\setminus[-\epsilon,\epsilon]} \frac{\varphi(x)}{x}\,\mathrm{d}x$$

so that:

$$\int_{\mathbb{R}} \varphi(x)\,\mathrm{d}x = \langle \mathcal{P}\left(\frac{1}{x}\right), x\varphi\rangle$$

Substituting back in (6.18) and rearranging we get:

$$\langle T, x\varphi\rangle = \langle \mathcal{P}\left(\frac{1}{x}\right), x\varphi\rangle \Rightarrow \langle T - \mathcal{P}\left(\frac{1}{x}\right), x\varphi\rangle = 0 \underset{(6.17)}{\Rightarrow} x\left[T - \mathcal{P}\left(\frac{1}{x}\right)\right] = 0$$

All that's left is to solve:

$$xT' = 0 \tag{6.19}$$

with $T' = T - \mathcal{P}(1/x)$. We will now see that the general solution of (6.19) is $T = c\delta$, for some constant c. This leads to:

$$T' = T - \mathcal{P}\left(\frac{1}{x}\right) = c\delta \Rightarrow T = \mathcal{P}\left(\frac{1}{x}\right) + c\delta$$

which indeed confirms (6.13).

So, let's see why $T' = c\delta$. In the following, we drop the $'$ for simplicity.
First, we note that any test function $\varphi(x)$ can be written as:

$$\varphi(x) = \varphi(0) + x\psi(x)$$

for some $\psi(x) \in \mathcal{S}(\mathbb{R})$. Explicitly:

$$\begin{aligned}\varphi(x) &= \varphi(0) + \int_0^x \varphi'(t)\,\mathrm{d}t \underset{u=\frac{t}{x}}{=} \varphi(0) + \int_0^1 x\varphi'(xu)\,\mathrm{d}u = \\ &= \varphi(0) + x\underbrace{\int_0^1 \varphi'(xu)\,\mathrm{d}u}_{\psi(x)} = \varphi(0) + x\psi(x)\end{aligned} \tag{6.20}$$

Note that if $\varphi(0) = 0$, then $\varphi(x) = x\psi(x)$.

Now, $xT = 0$ means that:

$$\langle xT, \varphi\rangle = 0 \qquad \forall\varphi \in \mathcal{S}(\mathbb{R}) \tag{6.21}$$

To see what T is, we evaluate it on a test function $\varphi(x)$. The idea is to write $\varphi(x)$ as a sum of two test functions $a(x)$ and $b(x)$, choosing $b(x)$ so that it vanishes at 0, meaning that we can factor a x from it (6.20), and then use $\langle T, xb\rangle = \langle xT, b\rangle = 0$ (6.21).
Note that we can't just directly use (6.20), because while $x\psi(x)$ is indeed a test function, $\varphi(0) \notin \mathcal{S}(\mathbb{R})$ (it is a constant value, so it doesn't vanish for $x \to \infty$). So, the following is ill-defined:

$$\langle T, \varphi\rangle = \underbrace{\langle T, \varphi(0)\rangle}_{?} + \underbrace{\langle T, x\psi(x)\rangle}_{0}$$

as $\langle T, \varphi(0)\rangle$ can't be done, because distributions act *only* on elements of $\mathcal{S}(\mathbb{R})$.

The idea is to *convert* $\varphi(0)$ to a test function by multiplying it with another test function $\chi(x) \in \mathcal{S}(\mathbb{R})$, that we choose (for simplicity) so that $\chi(0) = 1$. Then we write $\varphi(x)$ as:

$$\begin{aligned}\varphi(x) &= \varphi(x) + \varphi(0)\chi(x) - \varphi(0)\chi(x) = \\ &= \underbrace{\varphi(0)\chi(x)}_{a(x)} + \underbrace{[\varphi(x) - \varphi(0)\chi(x)]}_{b(x)}\end{aligned}$$

Note that now $a(x) \in \mathcal{S}(\mathbb{R})$, meaning that $\langle T, a \rangle$ is properly defined. Moreover, as we chose $\chi(0) = 1$, $b(x)$ is a test function that vanishes at 0:

$$b(0) = \varphi(0) - \varphi(0)\chi(0) = \varphi(0) - \varphi(0) = 0$$

And so we can use (6.20) to write $b(x) = x\psi(x)$ for some $\psi(x) \in \mathcal{S}(\mathbb{R})$.
Finally, we are able to apply T to $\varphi(x)$:

$$\begin{aligned}\langle T, \varphi \rangle &= \langle T, \varphi(0)\chi + x\psi \rangle = \\ &= \varphi(0) \underbrace{\langle T, \chi \rangle}_{c} + \underbrace{\langle xT, \psi \rangle}_{0} = \\ &= c\varphi(0) = \langle c\delta, \varphi \rangle\end{aligned}$$

where we denoted with c the result of $\langle T, \chi \rangle$. This proves that the general solution is indeed $T = c\delta$.

Some references on these derivations can be found in:

- https://see.stanford.edu/materials/lsoftaee261/book-fall-07.pdf
- https://math.stackexchange.com/questions/678457/distribution-solution-to-xt-0-in-schwartz-space
- https://math.stackexchange.com/questions/2962209/solve-the-distribution-equation-xt-1

Explicit computation. It is also possible to compute $\tilde{\theta}$ *directly*, at the cost of a longer derivation. The idea is to use a *limit representation* $\theta_\epsilon(x)$ for $\theta(x)$, so that $\theta_\epsilon(x)$ has the same discontinuity of $\theta(x)$ at $x = 0$, and $\lim_{\epsilon \to 0^+} \theta_\epsilon(x) = \theta(x)$. One possible choice is:

$$\theta_\epsilon(x) = \begin{cases} e^{-\epsilon x} & x > 0 \\ 0 & x < 0 \end{cases}$$

When $\epsilon \to 0^+$, $e^{-\epsilon x} \to 1$, reconstructing the Heaviside function. So:

$$\begin{aligned}\tilde{\theta}(k) &= \int_{\mathbb{R}} \theta(x) e^{-ikx} \,\mathrm{d}x = \lim_{\epsilon \to 0^+} \int_0^{+\infty} e^{-\epsilon x} e^{-ikx} \,\mathrm{d}x = \lim_{\epsilon \to 0^+} -\frac{1}{\epsilon + ik}[e^{-\infty} - 1] = \\ &= \lim_{\epsilon \to 0^+} \frac{1}{\epsilon + ik} \frac{-i^2}{-i^2} = \lim_{\epsilon \to 0^+} \frac{-i}{k - i\epsilon}\end{aligned}$$

To manipulate this expression we need to treat it in the context of distributions, meaning that we need to apply it to a test function $\varphi(x)$ and see what happens:

$$\begin{aligned}\langle \tilde{\theta}, \varphi \rangle &= \int_{\mathbb{R}} \tilde{\theta}(k) \varphi(k) \,\mathrm{d}k = \lim_{\epsilon \to 0^+} \int_{\mathbb{R}} \frac{-i}{k - i\epsilon} \frac{k + i\epsilon}{k + i\epsilon} \varphi(k) \,\mathrm{d}k = \\ &= -i \lim_{\epsilon \to 0^+} \int_{\mathbb{R}} \frac{k + i\epsilon}{k^2 + \epsilon^2} \varphi(k) \,\mathrm{d}k =\end{aligned}$$

$$\underset{(a)}{=} -i\left[\lim_{\epsilon\to 0^+} \underbrace{\int_{\mathbb{R}} \frac{k}{k^2+\epsilon^2}\varphi(k)\,\mathrm{d}k}_{A(\epsilon)} + i\lim_{\epsilon\to 0^+} \underbrace{\int_{\mathbb{R}} \frac{\epsilon}{k^2+\epsilon^2}\varphi(k)\,\mathrm{d}k}_{B(\epsilon)}\right]$$

where in (a) we split the real and imaginary part. We then examine each of them separately:

$$\begin{aligned}
A(\epsilon) &= \int_{\mathbb{R}} \frac{k}{k^2+\epsilon^2}\varphi(k)\,\mathrm{d}k = \int_{\mathbb{R}} \left(\frac{\mathrm{d}}{\mathrm{d}k}\frac{1}{2}\ln\!\left(k^2+\epsilon^2\right)\right)\varphi(k)\,\mathrm{d}k = \\
&\underset{(b)}{=} \cancel{a\varphi\Big|_{\mathbb{R}}} - \frac{1}{2}\int_{\mathbb{R}} \ln\!\left(k^2+\epsilon^2\right)\varphi'(k)\,\mathrm{d}k \\
&\xrightarrow[\epsilon\to 0^+]{} -\frac{1}{2}\int_{\mathbb{R}} \underbrace{\ln\!\left(k^2\right)}_{2\ln|k|}\varphi'(k)\,\mathrm{d}k = -\int_{\mathbb{R}} \ln|k|\varphi'(k)\,\mathrm{d}k
\end{aligned}$$

$$\begin{aligned}
B(\epsilon) &= \int_{\mathbb{R}} \frac{\epsilon}{k^2+\epsilon^2}\varphi(k)\,\mathrm{d}k = \int_{\mathbb{R}} \frac{1}{\epsilon}\frac{1}{1+\frac{k^2}{\epsilon^2}}\varphi(k)\,\mathrm{d}k = \\
&= \int_{\mathbb{R}} \left[\frac{\mathrm{d}}{\mathrm{d}k}\arctan\left(\frac{k}{\epsilon}\right)\right]\varphi(k)\,\mathrm{d}k = \\
&\underset{(c)}{=} \cancel{b\varphi\Big|_{\mathbb{R}}} - \int_{\mathbb{R}} \arctan\left(\frac{k}{\epsilon}\right)\varphi'(k)\,\mathrm{d}k \\
&\xrightarrow[\epsilon\to 0^+]{} -\int_0^{+\infty} \frac{\pi}{2}\varphi'(k)\,\mathrm{d}k - \int_{-\infty}^0 \left(-\frac{\pi}{2}\right)\varphi'(k)\,\mathrm{d}k = \\
&= -\frac{\pi}{2}\int_{\mathbb{R}} \mathrm{sgn}(k)\varphi'(k)\,\mathrm{d}k \underset{(d)}{=} \frac{\pi}{\cancel{2}}\int_{\mathbb{R}} \underbrace{\mathrm{sgn}'(k)}_{\cancel{2}\delta(k)}\varphi(k)\,\mathrm{d}k
\end{aligned}$$

where in (b), (c) and (d) we performed integrations by parts. Then we note that:

$$\begin{aligned}
\lim_{\epsilon\to 0^+} \langle B(\epsilon), \varphi\rangle &= \pi\langle \delta, \varphi\rangle \\
\lim_{\epsilon\to 0^+} A(\epsilon) &= -\int_{\mathbb{R}} \ln|k|\varphi'(k)\,\mathrm{d}k \underset{(e)}{=} \mathcal{P}\int_{\mathbb{R}} \frac{1}{k}\varphi(k)\,\mathrm{d}k
\end{aligned}$$

with a final integration by parts in (e). Putting it all together we arrive at the desired result:

$$\tilde{\theta}(k) = -i\mathcal{P}(\frac{1}{k}) + \pi\delta(k) = \mathcal{P}\left(\frac{1}{ik}\right) + \pi\delta(k)$$

Reference: `https://math.stackexchange.com/questions/269809/heaviside-step-function-fourier-transform-and-principal-values`

6.2 Fresnel integral

An important complex integral, appearing for example in the Schrödinger equation, is the Fresnel integral:

$$I(a,b) \equiv \int_{-\infty}^{+\infty} \frac{\mathrm{d}k}{2\pi} \exp\Big(-iak^2 - ibk\Big) = \frac{1}{\sqrt{4\pi ai}} \exp\left(\frac{ib^2}{4a}\right)$$

It is similar to a Gaussian integral, but with complex mean and variance.

To compute it, the idea is to *rotate it* so that it is not entirely along the imaginary axis. Explicitly, we rewrite the i multiplying the a in the exponential argument as:

$$i = \exp\left(i\frac{\pi}{2}\right)$$

And then we subtract an angle ϵ, and consider the limit $\epsilon \to 0^+$:

$$i = \lim_{\epsilon \to 0^+} \exp\left[i\left(\frac{\pi}{2} - \epsilon\right)\right]$$

Then, we evaluate the integral over one segment $[-R, R]$ of the real line, and take the limit $R \to \infty$:

"Regularized" Fresnel integral

$$\begin{aligned} I(a,b) &= \lim_{\epsilon \to 0^+} I_\epsilon(a,b) \\ I_\epsilon(a,b) &= \lim_{R\to\infty} \int_{-R}^{+R} \frac{\mathrm{d}k}{2\pi} \exp\Bigg(-a\underbrace{k^2 \exp\left[i\left(\frac{\pi}{2} - \epsilon\right)\right]}_{z^2} - ibk\Bigg) \qquad a, b \in \mathbb{R} \end{aligned}$$

Suppose that $a > 0$. We make the change of variables:

1. Change of variables

$$z^2 \equiv k^2 \exp\left[i\left(\frac{\pi}{2} - \epsilon\right)\right] \Rightarrow z = k \exp\left[i\underbrace{\left(\frac{\pi}{4} - \frac{\epsilon}{2}\right)}_{\phi_\epsilon}\right] = ke^{i\phi_\epsilon} \Rightarrow k = ze^{-i\phi_\epsilon}$$

And $\mathrm{d}k = \mathrm{d}z\, e^{-i\phi_\epsilon}$. Note that:

$$\phi_\epsilon < \frac{\pi}{4} \tag{6.22}$$

definitely when $\epsilon \to 0^+$.

This change of variables has removed the i multiplying the z^2, meaning that now we have a "standard" Gaussian integral. However, the integration path is now $\gamma_R = \{|z| \leq R, \arg z = \phi_\epsilon\}$, i.e. a segment of length $2R$, centred at the origin and forming an angle ϕ_ϵ with the real line. So the integral becomes:

$$\begin{aligned} I_\epsilon(a,b) &= \lim_{R\to\infty} \int_{\gamma_R} \frac{\mathrm{d}z}{2\pi} e^{-i\phi_\epsilon} \exp\Big(-az^2 - iz\underbrace{be^{-i\phi_\epsilon}}_{b'}\Big) \qquad b' = be^{-i\phi_\epsilon} \\ &= \lim_{R\to\infty} \int_{\gamma_R} \frac{\mathrm{d}z}{2\pi} e^{-i\phi_\epsilon} \exp\Big(-az^2 - ib'z\Big) \end{aligned}$$

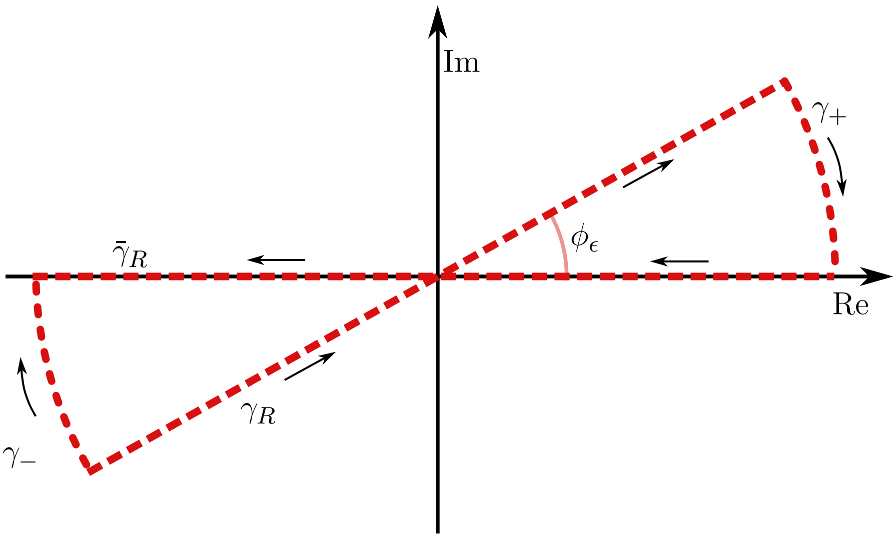


Figure (6.1) – Integration path for the Fresnel integral

We want to *relate* this integral to its version *on the real line*, that we know how to compute. To do this, as always, we *close* the path of integration and use the Cauchy integral theorem, following the schema in fig. 6.1.
Explicitly, consider the closed curve Γ_R defined by:

2. Contour integration

$$\Gamma_R = \gamma_R + \gamma_+ + \bar{\gamma}_R + \gamma_-$$

where:

$$\begin{aligned}
\gamma_+ &= \{z = Re^{i\theta} \colon \theta \in [0, \phi_\epsilon]\} \\
\gamma_- &= \{z = Re^{i\theta} \colon \theta \in [\pi, \pi + \phi_\epsilon]\} \\
\gamma_R &= \{|z| \leq R, \arg z = \phi_\epsilon\} \\
\bar{\gamma}_R &= [-R, R]
\end{aligned}$$

As the integrand has no poles inside Γ_R, we have:

$$\lim_{R\to\infty} \int_{\Gamma_R} \frac{\mathrm{d}z}{2\pi} e^{-i\phi_\epsilon} \exp\left(-az^2 - ib'z\right) = 0$$

Moreover, the integral over γ_+ and γ_- vanish. We show this explicitly only for the γ_+ case:

3. Integrals over $\gamma_\pm$ vanish

$$\left| \int_{\gamma_+} \frac{\mathrm{d}z}{2\pi} e^{-i\phi_\epsilon} \exp\left(-az^2 - ibze^{-i\phi_\epsilon}\right) \right| \tag{6.23}$$

We use the parameterization of γ_+ to change variables:

$$z = Re^{i\theta} \Rightarrow \mathrm{d}z = iRe^{i\theta}\,\mathrm{d}\theta$$

leading to:

$$(6.23) = \left| \int_0^{\phi_\epsilon} \frac{\mathrm{d}\theta}{2\pi} iRe^{i\theta} e^{-i\phi_\epsilon} \exp\left(-aR^2 e^{2i\theta} - ibRe^{i\theta} e^{-i\phi_\epsilon}\right) \right| =$$

$$= \underbrace{\left|\frac{iR}{2\pi}e^{-i\phi_\epsilon}\right|}_{R/(2\pi)} \left|\int_0^{\phi_\epsilon} \mathrm{d}\theta\, e^{i\theta} \exp\Big(-aR^2 e^{2i\theta} - ibRe^{i(\theta-\phi_\epsilon)}\Big)\right| \leq$$

$$\leq \frac{R}{2\pi}\int_0^{\phi_\epsilon} \mathrm{d}\theta \left|\exp\Big(i\theta - aR^2 e^{2i\theta} - ibRe^{i(\theta-\phi_\epsilon)}\Big)\right| =$$

$$= \frac{R}{2\pi}\int_0^{\phi_\epsilon} \mathrm{d}\theta\, \underbrace{|e^{i\theta}|}_{1} |e^{-aR^2(\cos 2\theta + i\sin 2\theta)}||e^{-ibR(\cos(\theta-\phi_\epsilon)+i\sin(\theta-\phi_\epsilon))}| =$$

$$= \frac{R}{2\pi}\int_0^{\phi_\epsilon} \mathrm{d}\theta\, e^{-aR^2\cos 2\theta + Rb\sin(\theta-\phi_\epsilon)} \xrightarrow[R\to\infty]{} 0$$

As the integral is over θ in $[0, \phi_\epsilon]$, we have:

$$0 < \theta < \phi_\epsilon \underbrace{<}_{(6.22)} \frac{\pi}{4} \Rightarrow 0 < 2\theta < \frac{\pi}{2} \Rightarrow \cos(2\theta) > 0$$

So, as we assumed $a > 0$, the integrand decays exponentially fast when $R \to \infty$, making the integral vanish.

Finally, as the integral over γ_+ and γ_- vanish, then:

$$I_{\gamma_R} + I_{\bar{\gamma}_R} = 0 \Rightarrow I_{\gamma_R} = -I_{\bar{\gamma}_R}$$

where $I_{\bar{\gamma}_R}$ is the integral over the real line, that we can compute:

4. Integral over the real line

$$I_{\gamma_R} = -\int_{-R}^{R} \frac{\mathrm{d}z}{2\pi} e^{-i\phi_\epsilon} \exp\Big(-az^2 - ib'z\Big) \xrightarrow[R\to\infty]{} \frac{e^{-i\phi_\epsilon}}{2\pi}\sqrt{\frac{\pi}{a}} \exp\left(-\frac{(b')^2}{4a}\right) =$$

$$= \frac{1}{\sqrt{4\pi a}} e^{-i\phi_\epsilon} \exp\left(-\frac{(b')^2}{4a}\right)$$

Inserting back $b' = be^{-i\phi_\epsilon}$, and taking the limit $\epsilon \to 0^+$, we have:

$$\phi_\epsilon \xrightarrow[\epsilon\to 0^+]{} \frac{\pi}{4} \Rightarrow e^{-i\phi_\epsilon} \xrightarrow[\epsilon\to 0^+]{} \frac{1}{\sqrt{i}} \Rightarrow b' \xrightarrow[\epsilon\to 0^+]{} \frac{b}{\sqrt{i}}$$

and $(b')^2 \to -ib^2$, so that:

$$I(a,b) = \frac{1}{\sqrt{4\pi a i}} \exp\left(\frac{ib^2}{4a}\right)$$

which is the desired result.

For $a < 0$, observe that $I(a,b) = I^*(-a,-b)$, with $-ia = (ia)^*$ and $b^2 = (b^2)^*$, and the same result follows.

6.2.1 Schrödinger Equation

A possible application of the Fresnel integration is solving the Schrödinger equation for a *free* particle:

Example of application

$$i\hbar\partial_t\psi(x,t) = -\frac{\hbar^2}{2m}\partial_x^2\psi(x,t) \tag{6.24}$$

In the following, we will take $\hbar = 1$ for simplicity. Note that (6.24) is very similar to the diffusion equation, and in fact we can solve it in the same way, by applying a Fourier transform to both sides:

$$\begin{aligned} i\partial_t \tilde{\psi}(p,t) &\underset{(a)}{=} -\frac{1}{2m}\int_{\mathbb{R}} \mathrm{d}x\, \partial_x^2 \psi(x,t) e^{-ixp} = \\ &= \frac{p^2}{2m} \underbrace{\int_{\mathbb{R}} \mathrm{d}x\, \psi(x,t) e^{-ipx}}_{\tilde{\psi}(p,t)} = \frac{p^2}{2m}\tilde{\psi}(p,t) \end{aligned}$$

where in (a) we performed two integrations by parts, using the fact that $\psi(x,t)$ vanishes at infinity to remove the boundary terms.

We are left with a first order ODE that can be solved by separation of variables:

$$i\partial_t \tilde{\psi} = \frac{p^2}{2m}\tilde{\psi} \Rightarrow \frac{\mathrm{d}\tilde{\psi}}{\tilde{\psi}} = -i^2 \frac{p^2}{2mi}\,\mathrm{d}t = -\frac{ip^2}{2m}\,\mathrm{d}t \Rightarrow \tilde{\psi}(p,t) = \tilde{\psi}(p,0)\exp\left(-\frac{ip^2 t}{2m}\right)$$

If we assume the particle to be initially localized at $x = 0$, meaning that $\psi(x,0) = \delta(x)$, we have $\tilde{\psi}(p,0) = 1$, and so:

$$\tilde{\psi}(p,t) = \exp\left(-i\frac{p^2 t}{2m}\right)$$

All that's left is to "go back to *position* space" with an inverse Fourier transform, which involves a Fresnel integral:

$$\psi(x,t) = \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{d}p \exp\left(-\frac{ip^2 t}{2m}\right) e^{ipx} = \frac{1}{\sqrt{4\pi a i}} \exp\left(\frac{ib^2}{4a}\right)$$

with $a = t/(2m)$ and $b = -x$, leading to:

$$\psi(x,t) = \sqrt{\frac{m}{2\pi i t}} \exp\left(-\frac{mx^2}{2it}\right)$$

To reinsert $\hbar$ we substitute $t \to t\hbar$:

$$\psi(x,t) = \sqrt{\frac{m}{2\pi \hbar i t}} \exp\left(-\frac{mx^2}{2\hbar i t}\right)$$

This is the Schrödinger **propagator** for a one-dimensional free particle.

6.3 Indented Integrals

Sometimes it is needed to compute integrals with *singularities* on the path of integration. Note that this integrals *do not exist*, meaning that there is not a unique way to compute them. Nonetheless, there are several *rules* (or *prescriptions*) that can be used to assign some result (possibly of physical significance) to these integrals.

Compute integrals that do not exist

Consider, for example, an analytic function $f(z)$, and the following integral:

$$I = \int_{\mathbb{R}} \mathrm{d}x \, \frac{f(x)}{x - x_0}$$

1. Symmetrical integration

The integrand has a pole at x_0, which lies in the path of integration. So I does not exist. However, we could integrate "symmetrically", hoping that the diverging term from one side "cancels" with the one from the other. This is the gist of the **Cauchy Principal Value**:

Cauchy Principal Value

$$\mathcal{P}\int_{\mathbb{R}} \mathrm{d}x \, \frac{f(x)}{x - x_0} = \lim_{\delta \to 0} \left[\int_{-\infty}^{x_0 - \delta} \frac{f(x)}{x - x_0} \, \mathrm{d}x + \int_{x_0 + \delta}^{\infty} \frac{f(x)}{x - x_0} \, \mathrm{d}x \right]$$

For example, this works for $f(x) = 1/x^2$ and $x_0 = 0$:

$$\mathcal{P}\int_{\mathbb{R}} \frac{1}{x^3} \, \mathrm{d}x = \lim_{\delta \to 0} \left[\int_{-\infty}^{-\delta} \frac{1}{x^3} + \int_{\delta}^{\infty} \frac{1}{x^3} \right] \underset{(a)}{=} \lim_{\delta \to 0} 0 = 0$$

where in (a) we used the *symmetry* of $1/x^3$ to cancel the two integrals.

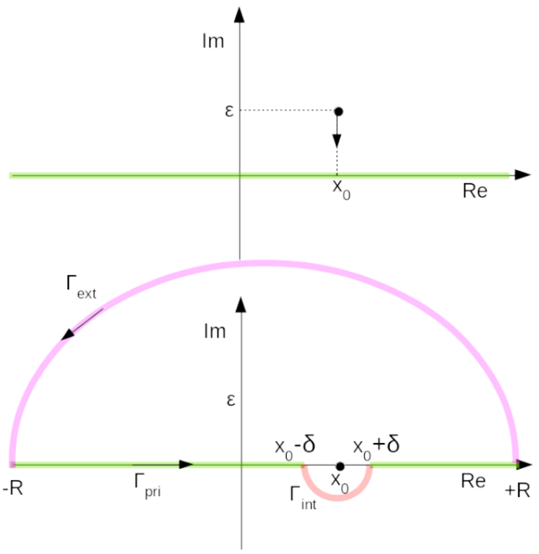


Figure (6.2) – Integration path for an indented integral

2. Path deformation

Another possibility is to *deform* the integration path from the real line to a curve γ_ϵ that avoids the singularity, as can be seen in the bottom half of fig. 6.2. Doing so produces a *different* result from the one of the Cauchy Principal Value, because now we are accounting for half a small circle $C_\epsilon = \{z = x_0 + \epsilon e^{i\theta} \colon \theta \in [-\pi, 0]\}$ around the singularity:

$$\lim_{\epsilon \to 0} \int_{\gamma_\epsilon} \frac{f(x)}{x - x_0} \, \mathrm{d}x = \mathcal{P}\int_{\mathbb{R}} \frac{f(x)}{x - x_0} \, \mathrm{d}x + \lim_{\epsilon \to 0} \int_{C_\epsilon} \frac{f(x)}{x - x_0} \, \mathrm{d}x \qquad (6.25)$$

And the difference amounts to:

$$\lim_{\epsilon \to 0} \int_{C_\epsilon} \mathrm{d}z \, \frac{f(z)}{z - a} \underset{(a)}{=} \lim_{\epsilon \to 0} \int_{-\pi}^{0} \mathrm{d}\theta \, (i\epsilon e^{i\theta}) \frac{f(x_0 + \epsilon e^{i\theta})}{\epsilon e^{i\theta}} = i \lim_{\epsilon \to 0} \mathrm{d}\theta \, f(x_0 + \epsilon e^{i\theta}) = i\pi f(x_0)$$

where in (a) we changed variables using the parameterization of C_ϵ.

Integrating over γ_ϵ that passes *to the right* of the singularity is equivalent to not deforming at all the integration path and moving the singularity "up" instead, as can be seen in fig. 6.2. This is the idea of the **prescription $\pm i\epsilon$**:

3. Moving the singularity

$$\int_{\mathbb{R}} \mathrm{d}x \, \frac{f(x)}{x - (x_0 + i\epsilon)} = \int_{\gamma_\epsilon} \mathrm{d}x \, \frac{f(x)}{x - x_0} = \mathcal{P} \int_{\mathbb{R}} \frac{f(x)}{x - x_0} \, \mathrm{d}x + i\pi f(x_0)$$

Equivalently, it is possible to show that integrating over a path γ_ϵ^- that passes *to the left* of the singularity equates to moving the singularity "down":

$$\int_{\mathbb{R}} \mathrm{d}x \, \frac{f(x)}{x - (x_0 - i\epsilon)} = \int_{\gamma_\epsilon^-} \mathrm{d}x \, \frac{f(x)}{x - x_0} = \mathcal{P} \int_{\mathbb{R}} \frac{f(x)}{x - x_0} \, \mathrm{d}x - i\pi f(x_0)$$

We can summarize these facts as an equation between *operators*:

$$\lim_{\epsilon \to 0} \frac{1}{x - x_0 \mp i\epsilon} = \mathcal{P} \frac{1}{x - x_0} \mp i\pi\delta(x - x_0)$$

Chapter 7

Limit Distributions

In our previous discussion of Brownian motion, we concluded that the sum of many **independent** *gaussian* increments converges, in distribution, to a gaussian.
But what would happen if we consider increments that are still independent and identically distributed, but not gaussian? How would the distribution for their sum change, in the limit of *many* steps? Does it even have a unique form?

In this lesson, we will see that the sum of a general class of i.i.d. random variables - the ones for which it makes sense to compute mean and variance -, after some proper normalization, tends to a normal distribution. This is the gist of the **Central Limit Theorem** (CLT).
Moreover, even the distributions without finite mean or variance, for which the CLT does not apply, can still produce sums that converge to some *distribution* (not gaussian), which we call a **stable distribution**. This observation will allow us to study generalizations of Brownian motion, and in particular the phenomena of subdiffusion and superdiffusion, which have interesting physical applications.

So, we will start by proving the CLT, and then generalize the diffusion equation and study it in some particular cases.

7.1 Characteristic functions

To prove the CLT, we first need a way to efficiently compute the pdf of a sum of i.i.d. random variables.

Let's start with the case of just two **independent** variables X' and X'', with distributions $p_1(x')$ and $p_2(x'')$. Let $X = X' + X'' = f(X, X')$ be their sum, with distribution $p(x)$.
Applying the rule for a change of random variables, we get:

Sum of 2 independent random variables

$$p(x) = \langle \delta(x - f(x', x'')) \rangle_{p_1,p_2} = \int_{\mathbb{R}} \mathrm{d}x' \int_{\mathbb{R}} \mathrm{d}x''\, p_1(x') p_2(x'') \delta(x - x' - x'') \tag{7.1}$$

where we used the independence of X' and X'' to factorize their joint pdf. By symmetry, $\delta(x - x' - x'') = \delta(x' + x'' - x) = \delta(x'' - (x - x'))$. Then, integrating over x'' to remove the δ, we get:

$$= \int_{\mathbb{R}} \mathrm{d}x' \int_{\mathbb{R}} \mathrm{d}x''\, p_1(x') p_2(x'') \delta(x'' - (x - x')) = \int_{\mathbb{R}} \mathrm{d}x'\, p_1(x') p_2(x - x')$$

which is the **convolution** of the distributions p_1 and p_2.

Convolutions are best computed in the Fourier domain, where they reduce to multiplications:

$$\mathcal{F}\left[\int_{\mathbb{R}} \mathrm{d}x'\, p_1(x') p_2(x - x')\right](k) = \mathcal{F}[p_1](k) \cdot \mathcal{F}[p_2](k) \tag{7.2}$$

The Fourier transform[1] of a pdf $p(x)$ is called the **characteristic function** of the corresponding random variable X, and denoted with $\varphi(k)$:

Characteristic function

$$\varphi(k) \equiv \mathcal{F}[p(x)](k) = \int_{\mathbb{R}} \mathrm{d}x\, e^{ikx} p(x) = \langle e^{ikx} \rangle_{p(x)}$$

Note that $\varphi(k)$ is the *moment-generating function* M_X of X, evaluated at a complex argument:

$$M_X(k) = \langle e^{kx} \rangle \Rightarrow \varphi(k) = M_X(ik)$$

This means that we can use $\varphi(k)$ to **compute moments** of X. Note that:

Moments from characteristic functions

$$e^{ikx} = 1 + ikx - \frac{1}{2}k^2x^2 + \cdots = \sum_{n=0}^{+\infty} \frac{(ikx)^n}{n!}$$

And so:

$$\varphi(k) = \langle e^{ikx} \rangle = \sum_{n=0}^{\infty} \frac{i^n k^n}{n!} \langle x^n \rangle \tag{7.3}$$

Then, by differentiating n times and evaluating at 0, all terms of order $\neq n$ vanish, leaving only a multiple of $\langle x^n \rangle$:

$$\frac{\partial \varphi(k)}{\partial k^n} = \underbrace{0}_{\text{First n terms}} + i^n \frac{n!}{n!} \langle x^n \rangle + \sum_{j=n+1}^{+\infty} \frac{n!}{j!} i^j k^{j-n} \langle x^j \rangle =$$

$$\left.\frac{\partial \varphi(k)}{\partial k^n}\right|_{k=0} = i^n \langle x^n \rangle \Rightarrow \langle x^n \rangle = \frac{1}{i^n} (-i^2)^n \frac{\partial}{\partial k^n} \varphi(k)\Big|_{k=0} = (-i)^n \frac{\partial}{\partial k^n} \varphi(k)\Big|_{k=0}$$

Proof of convolution property. Start from the left side of (7.2). By repeating backwards the steps from (7.1) we have:

$$\mathcal{F}\left[\int_{\mathbb{R}} \mathrm{d}x'\, p_1(x') p_2(x - x')\right](k) = \mathcal{F}\left[\int_{\mathbb{R}} \mathrm{d}x' \int_{\mathbb{R}} \mathrm{d}x''\, \delta(x - x' - x'')\right](k) =$$

$$= \int_{\mathbb{R}} \mathrm{d}x\, e^{ikx} \int_{\mathbb{R}} \mathrm{d}x' \int_{\mathbb{R}} \mathrm{d}x''\, p_1(x') p_2(x'') \delta(x - x' - x'') =$$

[1] ∧Here we are using a slightly different convention for the Fourier transform compared to sec. 6.1, where both the $-$ sign and $(2\pi)^{-1}$ normalization factor are contained in the *inverse* transform.

$$= \int_{\mathbb{R}} \mathrm{d}x' \int_{\mathbb{R}} \mathrm{d}x'' \, e^{ik(x'+x'')} p_1(x') p_2(x'') = \int_{\mathbb{R}} \mathrm{d}x' \, e^{ikx'} \int_{\mathbb{R}} \mathrm{d}x'' \, e^{ikx''} =$$
$$= \mathcal{F}[p_1](k) \cdot \mathcal{F}[p_2](k)$$

7.2 Central Limit Theorem

We are finally ready to prove the full Central Limit Theorem.

Consider a set of n **independent** and **identically distributed** (i.i.d.) random variables $\boldsymbol{X} = \{X_1, \dots, X_n\}$, each according to a distribution $f(x)$ with **finite** mean μ and variance σ^2. We want to prove that their sum $S_n = \sum_{i=1}^n x_i$, when *properly translated/scaled*, converges in distribution to a gaussian.

More precisely, the "proper translation/scaling" means considering the random variable Y_n defined by:

$$Y_n \equiv \frac{S_n - n\mu}{\sqrt{n}\sigma} \tag{7.4}$$

Note that, by additivity of mean and variance:

$$\begin{aligned} \langle S_n \rangle &= \langle x_1 \rangle + \dots + \langle x_n \rangle = n\mu \\ \operatorname{Var}(S_n) &= \operatorname{Var}(x_1) + \dots + \operatorname{Var}(x_n) = n\sigma^2 \end{aligned}$$

And so:

$$\langle Y_n \rangle = \frac{\langle S_n \rangle - n\mu}{\sqrt{n}\sigma} = 0 \qquad \operatorname{Var}(Y_n) = \frac{\operatorname{Var}(S_n)}{n\sigma^2} = \frac{\cancel{n\sigma^2}}{\cancel{n\sigma^2}} = 1$$

where we used $\operatorname{Var}(x+a) = \operatorname{Var}(x)$ and $\operatorname{Var}(bx) = b^2 \operatorname{Var}(x)$ where $b \in \mathbb{R}$ is a constant. So, we expect Y_n to converge in distribution to a standard gaussian (0 mean and unit variance).

To compute the distribution of Y_n we apply the rule for changing random variables:

$$Y_n \sim g(y) = \mathbb{P}(Y_n(\boldsymbol{x}) = y | x_i \sim f(x) \, \forall i) = \langle \delta(y - Y_n(x)) \rangle_{\boldsymbol{x} \sim [f(x)]^n} =$$

We rewrite the δ as a Fourier transform $\delta(x) = \mathcal{F}^{-1}[1] = (2\pi)^{-1} \int_{\mathbb{R}} \mathrm{d}k \, e^{-ikx}$, and then insert the definition for Y_n:

$$= \langle \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{d}\alpha \, e^{-i\alpha(y - Y_n(\boldsymbol{x}))} \rangle \underset{(7.4)}{=} \langle \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{d}\alpha \exp\left[-i\alpha y + i\alpha \left(\frac{\sum_{i=1}^n x_i - n\mu}{\sqrt{n}\sigma} \right) \right] \rangle =$$

By linearity we can bring the average inside the integral, which is then factorized as the X_i are **independent**:

$$= \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{d}\alpha \exp(-i\alpha y) \prod_{i=1}^n \langle \exp\left(\frac{i\alpha x_i}{\sqrt{n}\sigma} \right) \rangle \exp\left(-\frac{i\alpha n\mu}{\sqrt{n}\sigma} \right) =$$

Finally we write explicitly the average:

$$= \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{d}\alpha \exp\left(-i\alpha \left[y + \frac{n\mu}{\sqrt{n}\sigma} \right] \right) \prod_{i=1}^n \int_{\mathbb{R}} \mathrm{d}x_i \exp\left(\frac{i\alpha x_i}{\sqrt{n}\sigma} \right) p(x_i) =$$

And then, as the X_i are **identically** distributed, the product becomes the power of the characteristic function of *any* of the n variables:

$$= \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{d}\alpha \exp\left(-i\alpha \left[y + \frac{n\mu}{\sqrt{n}\sigma}\right]\right) \Bigg[\underbrace{\int_{\mathbb{R}} \mathrm{d}x_1\, p(x_1) \exp\left(\frac{i\alpha x_1}{\sqrt{n}\sigma}\right)}_{\varphi_1\left(\frac{\alpha}{\sqrt{n}\sigma}\right)} \Bigg]^n \tag{7.5}$$

As all the n variables are effectively the same, we will drop the subscript in the following steps.

The idea is now to expand φ as in (7.3), bring all the terms inside the same exponential, and show that it reduces to a gaussian after integration. So:

$$\varphi\left(k = \frac{\alpha}{\sqrt{n}\sigma}\right) = 1 + i \underbrace{\langle x \rangle}_{\mu} \frac{\alpha}{\sqrt{n}\sigma} - \langle x^2 \rangle \frac{\alpha^2}{2n\sigma^2} + o(n^{-3/2})$$

This expansion only makes sense if μ and σ are finite. Actually, we need to require only σ to be finite, as then μ is finite by consequence of the Cauchy Schwarz Inequality. To proceed, recall that $\sigma^2 = \langle x^2 \rangle - \langle x \rangle^2 = \langle x^2 \rangle - \mu^2 \Rightarrow \langle x^2 \rangle = \sigma^2 + \mu^2$. Substituting in the previous expression:

$$= 1 + i\mu\frac{\alpha}{\sqrt{n}\sigma} - \frac{\alpha^2}{2n} - \frac{\alpha^2\mu^2}{2n\sigma^2} + o(n^{-3/2}) = \tag{7.6}$$

If we ignore all the higher order terms (as in the limit $n \to \infty$), (7.6) is the expansion of the following exponential, as the only non-negligible terms are the three highlighted above:

$$= \exp\left(\frac{i\alpha\mu}{\sqrt{n}\sigma} - \frac{\alpha^2}{2n} + o(n^{-3/2})\right) \tag{7.7}$$

We then substitute (7.7) in (7.5) and compute the n-th power:

$$\begin{aligned}
\mathbb{P}(Y_n(\boldsymbol{x}) = y) &= \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{d}\alpha \exp\left(-i\alpha\left[y + \frac{n\mu}{\sqrt{n}\sigma}\right]\right) \exp\left(\frac{i\alpha\mu n}{\sqrt{n}\sigma} - \frac{\alpha^2 n}{2n} + o(n^{-1/2})\right) = \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{d}\alpha \exp\left(-i\alpha y - \cancel{\frac{i\alpha n\mu}{\sqrt{n}\sigma}} + \cancel{\frac{i\alpha\mu n}{\sqrt{n}\sigma}} - \frac{\alpha^2}{2} + o(n^{-1/2})\right) = \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{d}\alpha \exp\left(-\frac{\alpha^2}{2} - i\alpha y + o(n^{-1/2})\right) =
\end{aligned}$$

This is a gaussian integral, which evaluates, in the large n limit, to:

$$\underset{n\to\infty}{=} \frac{1}{2\pi}\sqrt{\frac{\pi}{a}} \exp\left(\frac{b^2}{4a}\right) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right)$$

with $a = 1/2$ and $b = -iy$. The final result is the standard gaussian, as desired.

So, we showed that if $X_i \sim p(x)$ with finite variance σ, then the sum of n i.i.d. random variables X_i converges in distribution to a gaussian:

$$\lim_{n\to\infty} Y_n \sim \mathcal{N}(0, 1)$$

By undoing the normalization, we have:

$$\lim_{n\to\infty} S_n \sim \mathcal{N}(n\mu, n\sigma^2)$$

In particular, the sample mean distributes normally around the distribution mean μ:

$$\lim_{n\to\infty} \frac{1}{n} S_n \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

7.3 Subdiffusion and superdiffusion

Recall that, for Brownian motion, the final distribution for a particle starting in $x_0 = 0$ at $t_0 = 0$ is:

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

Its variance, which physically represents how quickly the initial distribution "spreads", is linear in time:

$$\langle x^2(t)\rangle = 2Dt$$

This is indeed a good model for many physical phenomena. However, there are cases of **anomalous diffusion**, in which the "spreading velocity" scales *differently* - as can be seen in fig. 7.1. For example:

- **Subdiffusion**. Sometimes particles tend to persist in the same state for extended periods of time - meaning that the waiting time between *jumps* has a distribution with a "long tail", such as $t^{-1-\alpha}$ with $\alpha \in (0,1)$. This happens, for example, in the transport of charge carriers in semiconductors, and monomers in polymer diffusion. Their paths satisfy:

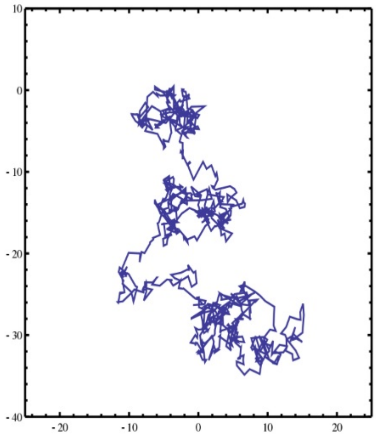
$$\langle x^2(t)\rangle = 2D_\zeta t^\zeta \qquad 0 < \zeta < 1$$

- **Superdiffusion**. Here particles make jumps of large size with non-negligible frequency, meaning that the distribution of displacements Δx has a "long tail", proportional to $|\Delta x|^{-1-\mu}$ for Δx sufficiently large, with $\mu \in (0,2)$. If large jumps happen almost *instantaneously*, we talk about "flights", while if they happen with a fixed maximum velocity, they are "walks". In this case we have:

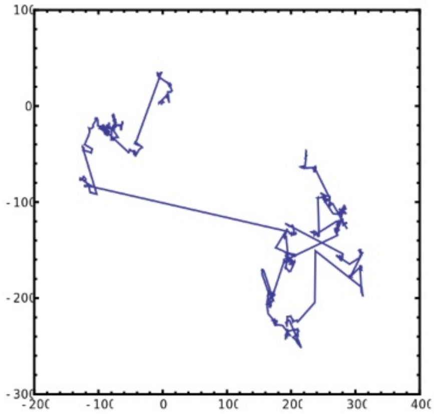
$$\langle x^2(t)\rangle = 2D_\zeta t^\zeta \qquad \zeta > 1$$

7.4 Levy Flights

Anomalous diffusion can be even more complicated, involving *memory* and *long-range correlations*. In our discussion, we will limit ourselves to a case of



(a) – Brownian motion.



(b) – Levy flight (superdiffusion). Note how jumps are frequently of very large size.

Figure (7.1) – Comparison between normal diffusion (a) and anomalous diffusion (b).

superdiffusion - the Levy flights - that can be described with a **generalized diffusion equation**:

$$\begin{cases} \partial_t W(x,t) = D_\mu \frac{\partial}{\partial |x|^\mu} W(x,t) \\ W(x,0) = \rho(x) \end{cases} \qquad 0 < \mu < 2$$

The meaning of the *fractional* derivative can be understood in Fourier space, as a generalization of the transform for a derivative:

$$\partial_t \tilde{W}(k,t)) = -D_\mu |k|^\mu \tilde{W}(k,t)$$

Note that this is equivalent to:

$$\partial_t [\underbrace{e^{D_\mu |k|^\mu t} \tilde{W}(k,t)}_{\tilde{f}(k)}] = D_\mu |k|^\mu \cancel{e^{D_\mu |k|^\mu t}} \tilde{W}(k,t) + \cancel{e^{D_\mu |k|^\mu t}} \partial_t \tilde{W}(k,t) = 0$$

Meaning that the function $\tilde{f}(k)$ is constant in time. Rearranging:

$$\tilde{f}(k) \equiv \exp\left(D_\mu |k|^\mu t\right) \tilde{W}(k,t) \Rightarrow \tilde{W}(k,t) = \tilde{f}(k) e^{-D_\mu |k|^\mu t}$$

As $\tilde{f}(k)$ does not depend on time, we can compute it at any instant, for example at $t = 0$, where $\tilde{f}(k) = \tilde{W}(k,0) = \tilde{\rho}(k)$, and so:

$$\tilde{W}(k,t) = \tilde{\rho}(k) \underbrace{e^{-D_\mu |k|^\mu t}}_{\tilde{W}(k,t|k_0,0)}$$

We interpret the exponential as the Fourier transform of a **propagator**. Multiplication in the Fourier domain corresponds to convolution in the space domain, and so we recover the usual form for the solution of the diffusion problem:

$$W(x,t) = \rho(x_0) * W(x,t|x_0,0)$$

And for $\mu = 2$ we know that:

$$W(x,t) = \int_{\mathbb{R}} \mathrm{d}x_0\, \rho(x_0) \underbrace{\frac{1}{4Dt} \exp\left(-\frac{(x-x_0)^2}{4Dt}\right)}_{W(x,t|x_0,0)}$$

For a general $\mu \in (0,2)$, however, it is difficult to find analytically $W(x,t|x_0,0)$, except for a few cases. One of them is for $\mu = 1$, where $W(x,t)$ becomes a Cauchy distribution, and we talk about **Cauchy random flights**:

Cauchy Random Flights

$$\begin{aligned}\tilde{W}_C(k,t) = \tilde{\rho}(k) e^{-D_1|k|t} \Rightarrow W_C(x,t|0,0) &= \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{d}k \exp\left(-x^*(t)|k| + ikx\right) = \\ &= \frac{1}{\pi} \int_0^\infty \mathrm{d}k\, e^{-x^*(t)k} \cos(kx) = \\ &= \frac{1}{\pi x^*(t)} \frac{1}{1 + \left(\frac{x}{x^*(t)}\right)^2}\end{aligned}$$

where we set $\rho(x) = \delta(x)$, and $x^*(t) = D_1 t$, representing the *typical length scale*. See the exercises for a full derivation.

Note that, in the case of Levy flights, the displacements *do not follow* a distribution with finite variance, as it has a "long tail". Thus, the CLT theorem does not apply, and in fact the sum of many displacements is not normally distributed - for example, in the $\mu = 1$ case it is a Cauchy pdf.

However, the Cauchy pdf has a key property in common with the gaussian: it is a **stable distribution**. This means that a sum of two Cauchy random variables follows again a Cauchy pdf, up to scaling and translation.

Stable distributions

We argue (omitting the proof) that this property holds for all the distributions in the general case $\mu \in (0,2)$, which are called **Lévy alpha-stable distributions**. In particular, these stable distributions behave like "attractors" for the sums of i.i.d. random variables with certain distributions, exactly like the gaussian behaves for all random variables with finite variance. This leads to a **generalization** of the central limit theorem, for which the sum of a number of random variables with symmetric distributions having power-law tails (Paretian tails), decreasing as $|x|^{-\alpha-1}$ for large x, with $\alpha \in (0,2]$ (and therefore with infinite variance), will tend to a Lévy stable distribution as the number of summands grows[2].

Generalized CLT theorem

[2] ∧B.V. Gnedenko, A.N. Kolmogorov. Limit distributions for sums of independent random variables, Cambridge, Addison-Wesley 1954 `https://books.google.com/books/about/Limit_distributions_for_sums_of_independ.html?id=rYsZAQAAIAAJ&redir_esc=y`
See Theorem 5 in Chapter 7, Section 35, page 181.

Part III

Gradenigo's Lectures

Analytical index