Stochastic differential equations

Wojciech Dybalski

- There will be lecture notes and exercise sheets on `https://wdybalski.faculty.wmi.amu.edu.pl`. The exercise sheets will be posted about one week before the class at which they will be discussed.
- Office hours: Tuesdays 12-13 (Except Tuesday 7.10) or by appointment.
- Textbook: In general Øksendal [Ok]. But first lecture: Roepstorff [Ro, Ch 1].

First lecture

1 Motivation

1. ODE are prolific in physics. Imagine a particle moving under the influence of an external force, e.g. a pendulum
$$m\frac{d^2x(t)}{dt^2} = -kx(t), \quad k > 0. \tag{1}$$
A solution can be found given the initial conditions $x_0 = x(0), v_0 = \frac{dx(t)}{dt}|_{t=0}$:
$$x(t) = x_0 \cos(\omega t) + \frac{v_0}{\omega}\sin(\omega t), \quad \omega = \sqrt{\frac{k}{m}}. \tag{2}$$
The picture in phase space is an elipse.

2. In a real laboratory situations there are imperfections. For example, the air viscosity will modify the equation to
$$m\frac{d^2x(t)}{dt^2} = -kx(t) - \gamma\frac{dx(t)}{dt}, \quad \gamma > 0. \tag{3}$$
The picture in phase space is a spiral.

3. In addition there are perturbations which act like an external force: Trains passing by, seismic tremors, electromagnetic forces for nearby electric devices (acting on a metal pendulum):
$$m\frac{d^2x(t)}{dt^2} = -kx(t) - \gamma\frac{dx(t)}{dt} + \sigma\frac{dB_t}{dt}, \quad \sigma > 0. \tag{4}$$
The picture in phase space is a rough spiral. **The goal of the lecture is to understand the problem of existence and uniqueness of solutions for such equations.**

   [It may seem hopeless to describe this influence of many factors which we do not control. What saves us is the Central Limit Theorem: If the particle feels the combined effect of a large number of independent kicks,
$$\int_0^t ds\Big\{\text{independent random kicks at time } s\Big\} \tag{5}$$
the net force looks approximately like a Gaussian random variable B_t called Brownian motion. $\frac{dB_t}{dt}$ is called the white noise.]

4. From the practitioner's point of view, random variable is something you can compute expectation E (or average) of and get numbers. For Brownian motion we have

- $E(B_t) = 0$,
- $E(B_t^2) = t$,
- $E((B_0 - B_t)(B_t - B_s)) = 0$ for $s > t > 0$, (independence of increments).
- $B_0 = 0$.

Dealing with $t \mapsto B_t$ naively as if it was a differentiable function is tricky: In fact,

$$1 = \frac{d}{dt}t = \frac{d}{dt}E(B_t^2) = 2E(B_t\frac{dB_t}{dt}) = \lim_{\Delta t \to 0} 2E\big(B_t\frac{B_{t+\Delta t} - B_t}{\Delta t}\big) = 0, \tag{6}$$

which is a contradiction. In reality, one can make sense of $\frac{dB_t^2}{dt} = 2B_t\frac{dB_t}{dt} + 1$, which saves the game. [This example shows, that closer mathematical scrutiny is needed to deal with such objects].

5. Equations involving such random noise are called stochastic differential equations (SDE).
   - Since B_t are random variables, also $x(t)$ are random variables. Concepts like existence and uniqueness of solutions have to be reconsidered.
   - Applications: climate fluctuations, turbulence, neural activity, option pricing.... [It is hard to escape SDE if you want to model real world phenomena].
6. A financial mathematics application:
   - Suppose a person has an asset or resource (e.g. house, stocks, oil...) that they want to sell.
   - To start with, suppose that the value of the asset grows with r (Δt - compounded, annual) interest rate:

$$\frac{\Delta X}{X} = r\Delta t. \tag{7}$$

   - If we take the idealization of a continuous interest rate, this gives

$$\frac{dX_t}{dt} = rX_t, \tag{8}$$

   with a solution $X_t = X_0e^{rt}$ [t measured in years].

   - Since we are talking about a risky asset, the interest rate is fluctuating:

$$\frac{dX_t}{dt} = \left(r + \alpha\frac{dB_t}{dt}\right)X_t, \quad r, \alpha > 0. \tag{9}$$

   Naive solution has the form:

$$X_t = X_0e^{rt+B_t} \tag{10}$$

   One can make a mathematical sense out of it. But one can also make mathematical sense of the solution:

$$X_t = X_0e^{(r-\frac{1}{2})t+B_t}, \tag{11}$$

   depending how exactly one defines stochastic integrals $\int f(t)dB_t$. [This will be important part of the course].

   - **Optimal stopping problem**: The person knows $\{X_s\}_{0\le s\le t}$ and wants to find an optimal time to sale the asset. [If you sell too early, you may loose future gains. If you sell to late the pay-off may shrink].

- **An optimal portfolio problem**: Suppose a person has two possible investments: a risky investment (9) (a stock) and a safe investment (8) (a bond). How to divide the asset $X_t = u_t X_t + (1-u_t)X_t$ into a risky investment and a safe investment, given the selling time T?
- **Pricing of options**: Suppose that at time $t = 0$ the person of the previous question is offered the right (but no obligation) to buy one unit of the risky asset at a specified price K and at a specified future time $t = T$. [European call option]. How much should the person be willing to pay for such an option? Problem solved by Black and Scholes in (1973), related to Nobel Prize in Economics in 1997 (Scholes and Merton).

2 General probability I

[I will explain how B_t emerges from a random walk. We need some basic probabilistic notions needed to describe the random walk and state the CLT. (Picture from the notes of Bertsekas and Tsitsiklis [BT, p.6]).]

1. Def: If Ω is a given set then a σ-algebra $\mathcal{F}$ on Ω is a family of subsets on Ω with the following properties:

   (i) $\emptyset \in \mathcal{F}$.

   (ii) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$, where A^c is the complement of A in Ω.

   (iii) $A_1, A_2, \ldots \in \mathcal{F} \Rightarrow A := \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

   - Example: For $\Omega = \mathbb{R}^d$ the smallest σ-algebra containing all open sets is called the Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$.
   - If (iii) holds only for finite sums, then $\mathcal{F}$ is called an algebra.

2. Def: The pair $(\Omega, \mathcal{F})$ is called a measurable space.

3. Def: A probability measure on a measurable space $(\Omega, \mathcal{F})$ is a function $P : \mathcal{F} \to [0, 1]$ s.t.

   (a) $P(\emptyset) = 0$, $P(\Omega) = 1$.

   (b) If $A_1, A_2, \ldots \in \mathcal{F}$ and $\{A_i\}_{i=1}^{\infty}$ disjoint (i.e., $A_i \cap A_j = \emptyset$ is $i \neq j$) then

   $$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i). \tag{12}$$

   [Question: Why do we need a sigma field and cannot just define the measure on all subsets? Banach-Tarski paradox.]

4. Def: The triple $(\Omega, \mathcal{F}, P)$ is called a probability space.

5. Def. We say that the events A, B, are independent if $P(A \cap B) = P(A)P(B)$. [Note that disjoint sets are dependent. Examples: flipping a coin several times, rolling a die several times]. For a family of events $A_1, \ldots A_n$, we require for any non-empty subset $I \subset \{1, \ldots n\}$

   $$P(\bigcap_{i \in I} A_i) = \prod_{i \in I} P(A_i). \tag{13}$$

6. Def. Conditional probability:

   $$P(A|B) := \frac{P(A \cap B)}{P(B)} \text{ for } P(B) \neq 0, \tag{14}$$

   That is, probability that A occurs provided that B is known to occur.

7. Random variable is a functions $X : \Omega \to \mathbb{R}^d$ which is $\mathcal{F}$-measurable. That is, $X^{-1}(O) \in \mathcal{F}$ for any open $O \subset \mathbb{R}^d$.

8. Def. Every random variable induces a probability measure μ_X on $\mathbb{R}^d$, given by
$$\mu_X(B) = P(X^{-1}(B)) = P(X \in B). \tag{15}$$
It is called the distribution (or law) of X. One also writes $\mathcal{L}(X) := \mu_X$ or $X \sim$ [Name of $\mathcal{L}(X)$].
For example, if $X : \Omega \to \mathbb{R}$ satisfies
$$\mu_X(A) = \int_A \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(x-m)^2}{2\sigma^2}} dx, \tag{16}$$
we say that X is a Gaussian variable with law $N(m, \sigma^2)$ (normal disribution) or $X \sim N(m, \sigma^2)$.

9. Def. For a finite family of random variables $X_1, \dots, X_n$ the joint law is the measure on $\mathbb{R}^{dn}$ given by
$$\mu_{X_1,\dots,X_n}(A) = P((X_1, \dots, X_n) \in A), \quad A \in \mathcal{B}(\mathbb{R}^{dn}). \tag{17}$$

10. Def. If $\int |X(\omega)| dP(\omega) < \infty$ then the number
$$E(X) := \int_\Omega X(\omega) dP(\omega) = \int_\mathbb{R} x \, d\mu_X(x) \tag{18}$$
is called the expectation of X (w.r.t. P). Change of variables formula for measures.
If you are not comfortable with integration w.r.t. an arbitrary measure you may want to have a look at [Ru, Chapter 1].

11. Def. If $\int |X(\omega)|^2 dP(\omega) < \infty$ then the number
$$\mathrm{var}(X) = E(X^2) - E(X)^2 = E[(X - E(X))^2] \tag{19}$$
is called the variance.

12. Def. We say that real-valued random variables $X_1, \dots, X_n$ are independent if
$$\mu_{X_1,\dots,X_n} = \mu_{X_1} \otimes \cdots \otimes \mu_{X_n}. \tag{20}$$

13. Fact: If two random variables X, Y are independent, then they are uncorrelated
$$E[XY] = E[X]E[Y] \tag{21}$$
provided that $E[|X|] < \infty$ and $E[|Y|] < \infty$. (Converse not true in general. But true for Gaussian random variables as we will see).
Idea of proof: We compute for characteristic functions $X = \chi_{F_1}$, $Y = \chi_{F_2}$.
$$E[\chi_{F_1}\chi_{F_2}] = \int \chi_{F_1}(\omega)\chi_{F_2}(\omega) dP(\omega) = P(F_1 \cap F_2) = P(F_1)P(F_2). \tag{22}$$
Then, it follows for step functions. Recalling the definition of the Lebesgue integral, a measurable function is approximated pointwise by step functions.

14. Thm (Central Limit Theorem (CLT)): Let $(X_j)_{j\geq 1}$ be i.i.d. (independent identically distributed) $\mathbb{R}$-valued random variables with $E(X_j) = m$ and $\mathrm{Var}(X_j) = \sigma^2$ with $0 < \sigma^2 < \infty$. Let $S_n = \sum_{j=1}^n X_j$. Let $Y_n = \frac{S_n - nm}{\sigma\sqrt{n}}$. Then Y_n converges in distribution to Y, where $Y \sim N(0, 1)$. That is, for any bounded, continuous function f on $\mathbb{R}$
$$\int f(x) d\mu_{Y_n}(x) \to \int f(x) d\mu_Y(x). \tag{23}$$
(This is also called weak convergence of measures).

3 Example: Bernoulli trials and binomial probabilities

1. Consider one Bernoulli trial with a toss of a coin where we get heads (H) with probability p and tails (T) with probability $1-p$.

   - Sample space: $\Omega_0 = \{H, T\}$
   - σ-algebra: All subsets of Ω_0

   $$\mathcal{F}_0 = \{\emptyset, \{H\}, \{T\}, \{H, T\}\} \tag{24}$$

   - Probability measure: $P_0(\emptyset) = 0$, $P_0(\{H\}) = p$, $P_0(\{T\}) = 1-p$, $P_0(\{H, T\}) = 1$.

2. Consider n independent Bernoulli trials.

   - Sample space $\Omega = \Omega_0^n$. Any $\omega \in \Omega$ has the form $\omega = (\omega_1, \dots, \omega_n)$, $\omega_i \in \{H, T\}$.
   - σ-algebra: All subsets of Ω.
   - Probability measure: By independence of trials,

   $$P(\{\omega\}) = \prod_{i=1}^{n} P_0(\{\omega_i\}) = p^{R_n(\omega)}(1-p)^{n-R_n(\omega)}, \tag{25}$$

   where $R_n(\omega)$ is the number of H in ω. Note that $R_n : \Omega \to \mathbb{R}$ is a random variable.

3. Define random variables $\xi_i : \Omega \to \{-1, 1\}$ s.t.

   $$\xi_i(\omega) = \begin{cases} +1 & \text{if } \omega_i = H, \\ -1 & \text{if } \omega_i = T. \end{cases} \tag{26}$$

4. Then $R_n = \#\{\, i \,:\, \xi_i = +1\}$ is the number of heads in n trials.

5. Fact: We have

   $$P(R_n = \ell) = P(\ell \text{ heads come up in an } n\text{-toss sequence}) = \binom{n}{\ell} p^{\ell}(1-p)^{n-\ell}. \tag{27}$$

   We will write $R_n \sim \text{Binomial}(n, p)$, i.e. the random variable R_n has a binomial probability distribution.

   Proof:

   - Suppose we fix a particular individual outcome, e.g. HTTHTT. There are $\ell = 2$ heads in $n = 6$ trials. By independence of the tosses,

   $$P(\text{HTTHTT}) = P(\text{H})P(\text{H})P(\text{T})P(\text{H})P(\text{T})P(\text{T}) = p^{\ell}(1-p)^{n-\ell}. \tag{28}$$

   - There are many events which have $\ell = 2$ heads in $n = 6$ trials. For example HHTTTT. As this is a disjoint event from HTTHTT we have

   $$P(\text{HHTHTT or HTTHTT}) = P(\text{HHTHTT}) + P(\text{HTTHTT}). \tag{29}$$

   - Now, how many different sequences of n tosses contains exactly ℓ heads? Clearly, $\binom{n}{\ell}$. $\square$

4 From random walks to Brownian motion

1. Particle moves on a lattice $h\mathbb{Z}$, $h > 0$.

2. There occurs one step, to the left or to the right, within the period of time $\tau > 0$.

3. Successive steps are independent: e.g. a fair coin ($p = \frac{1}{2}$) is tossed at each step and the particle moves to the left or to the right depending on the outcome (H or T). We can use the same sample space as above $\Omega = \{H, T\}^n$, with probability P defined by (25).

4. We want to compute probability

$$P(\text{particle ends up at } x = kh \text{ after } n \text{ steps, starting from zero}) =: P(kh, n\tau) =: P[k, n]. \tag{30}$$

   - Define the random variable $S_n = \xi_1 + \ldots + \xi_n$, s.t. $S_n h$ is the position of the particle after n steps.
   - We have $S_n = (+1)R_n + (-1)(n - R_n)$, hence $R_n = (n + S_n)/2$. So the event $\{S_n = k\} := \{\omega \in \Omega : S_n(\omega) = k\}$ is the same as $\{R_n(\omega) = (n + k)/2\} := \{\omega \in \Omega : R_n = (n + k)/2\}$.
   - Then $P[k, n] = P(S_n = k) = P(R_n = (n + k)/2)$. Since $R_n \sim \text{Binomial}(n, 1/2)$, we have

$$P[k, n] := P(R_n = (n + k)/2) = \binom{n}{\frac{1}{2}(n + k)} \frac{1}{2^n}, \tag{31}$$

   provided $\frac{1}{2}(n + k)$ is integer. Otherwise $P[k, n] = 0$.

5. Let us now try to take the continuum limit $h, \tau \to 0$ s.t. $t := n\tau$, $\cancel{x := kh}$, $D := \frac{h^2}{2\tau}$ stay constant. Position of the particle after n steps is $x = hS_n = \sqrt{2\tau D} S_n = \sqrt{\frac{2Dt}{n}} S_n$. We restrict attention to n of the form $n = t\tilde{n}$ for $t \in \mathbb{N}$ and $\tilde{n} \in \mathbb{N}$. Then $x = \sqrt{\frac{2D}{\tilde{n}}} S_{\tilde{n}t}$ is still position of the particle after n steps at time t expressed in terms of $\tilde{n}$ which we will take to ∞. This motivates the definition (setting $D = 1/2$ and renaming $\tilde{n} \to n$):

$$B_t^{(n)} := \frac{1}{\sqrt{n}} S_{\lfloor nt \rfloor}, \tag{32}$$

   with the convention $S_0 = 0$.

6. Fact: By the CLT the limit

$$B_t := \lim_{n \to \infty} B_t^{(n)} \tag{33}$$

   exists in distribution and defines a family of Gaussian random variables $B_t \sim N(0, t)$.

   Def: The family $\{B_t\}_{t \in \mathbb{R}_+}$ is called the Brownian motion.

   Proof of the fact: In our case $X_j = \xi_j$.

   - Let us compute the law of ξ_i. Let $A \in \mathcal{B}(\mathbb{R})$. We note that $\xi_i(\omega) = \tilde{\xi}_i(\omega_i)$.

$$\mu_{\xi_i}(A) := P(\omega : \xi_i(\omega) \in A) = P(\omega : \tilde{\xi}_i(\omega_i) \in A). \tag{34}$$

   - Independence: For $n = 2$

$$\begin{aligned} \mu_{\xi_1, \xi_2}(A_1 \times A_2) &= P(\omega : (\xi_1(\omega), \xi_2(\omega)) \in A_1 \times A_2) \\ &= P(\omega : (\tilde{\xi}_1(\omega_1), \tilde{\xi}_2(\omega_2)) \in A_1 \times A_2) \\ &= P(\omega : \tilde{\xi}_1(\omega_1) \in A_1, \tilde{\xi}_2(\omega_2) \in A_2) \\ &= P(\omega : \tilde{\xi}_1(\omega_1) \in A_1) P(\omega : \tilde{\xi}_2(\omega_2) \in A_2) = \mu_{\xi_1}(A_1)\mu_{\xi_2}(A_2), \end{aligned} \tag{35}$$

   where in the next to the last step we made use of the independence of coin tosses, i.e. (25). For arbitrary n the argument is analogous.

- We compute

$$
\begin{aligned}
m_j &= E(\xi_j) = \int \xi_j(\omega)dP(\omega) = \int \xi_j d\mu_{\xi_j} = (+1)\frac{1}{2} + (-1)\frac{1}{2} = 0, && (36)\\
\sigma_j^2 &= E(\xi_j^2) = \int \xi_j(\omega_j)^2 dP(\omega_j) = \int \xi_j^2 d\mu_{\xi_j} = (+1)\frac{1}{2} + (+1)\frac{1}{2} = 1. && (37)
\end{aligned}
$$

- Thus, by the CLT, $Y_n = \frac{S_n}{\sqrt{n}}$ converges in distribution to a Gaussian random variable $Y = B_{t=1}$ with law $N(0,1)$.
- Similarly, for $t \in \mathbb{N}$, $\frac{S_{nt}}{\sqrt{n}} = \sqrt{t}\frac{S_{nt}}{\sqrt{nt}} = \sqrt{t}Y_{nt} \to \sqrt{t}Y =: B_t \sim N(0,t)$. Its law is

$$
\mu_{\sqrt{t}Y}(A) = P(\sqrt{t}Y \in A) = P(Y \in A/\sqrt{t}) = \int_{A/\sqrt{t}} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy = \int_A \frac{1}{\sqrt{2\pi t}} e^{-y^2/(2t)} dy, \quad (38)
$$

that is $B_t \sim N(0,t)$. (I leave for the reader the extension to $t \in \mathbb{R}_+$). $\square$

7. Remark: We know from the proof that $E(B_t) = 0$, $E(B_t^2) = t$. We would also like to show $E((B_0 - B_t)(B_t - B_s)) = 0$, $s > t$, but so far we do not have a joint probability space for B_t and B_s, $t \neq s$. However, already now, such an outcome is suggested by

$$
E((B_0^{(n)} - B_t^{(n)})(B_t^{(n)} - B_s^{(n)})) = 0. \quad (39)
$$

The two factors have the form

$$
\begin{aligned}
B_t^{(n)} - B_s^{(n)} &= \frac{1}{\sqrt{n}}(\xi_{\lfloor nt \rfloor + 1} + \dots + \xi_{\lfloor ns \rfloor}), && (40)\\
B_0^{(n)} - B_t^{(n)} &= \frac{1}{\sqrt{n}}(\xi_1 + \dots + \xi_{\lfloor nt \rfloor}). && (41)
\end{aligned}
$$

They involve independent coin tosses. For independent random variables we have $E(XY) = E(X)E(Y)$ by (20), so (39) follows.

8. We can also compute $E(B_t^{(n)} B_s^{(n)})$. By (39), for $s > t$,

$$
E(B_t^{(n)} B_s^{(n)}) = E(B_t^{(n)2}) = \frac{1}{n} E(\xi_1^2 + \dots + \xi_{\lfloor nt \rfloor}^2) = \frac{\lfloor nt \rfloor}{n}. \quad (42)
$$

We used $E(\xi_i \xi_j) = E(\xi)E(\xi_j) = 0$ for $i \neq j$. Since the roles of t, s can be exchanged

$$
E(B_t^{(n)} B_s^{(n)}) = \min(\frac{\lfloor nt \rfloor}{n}, \frac{\lfloor ns \rfloor}{n}). \quad (43)
$$

As we will see, in the limit $E(B_t B_s) = \min(t,s)$.

9. Let us consider again the continuum limit $h, \tau \to 0$ s.t. $t := n\tau, x := kh$, $D := \frac{h^2}{2\tau}$ stay constant. There is another way to see (heuristically) that the probability distribution of the position of the particle must be Gaussian:

   - Recall: $\binom{n+1}{m} = \binom{n}{m} + \binom{n}{m-1}$.
   - Hence: $P[k, n+1] = \frac{1}{2}P[k-1, n] + \frac{1}{2}P[k+1, n]$. In fact,

   $$
   P[k, n+1] = \binom{n+1}{\frac{1}{2}(n+1+k)} \frac{1}{2^{n+1}} = \frac{1}{2}\binom{n}{\frac{1}{2}(n+1+k)} \frac{1}{2^n} + \frac{1}{2}\binom{n}{\frac{1}{2}(n+1+k) - 1} \frac{1}{2^n}. \quad (44)
   $$

   Alternatively, we can say that the event { particle at point k after $n+1$ steps } is a union of two disjoint events { particle at point $k-1$ after and n steps, then jump to the right } and { particle at point $k+1$ after n steps, then jump to the left }

- Thus, we have

$$\frac{P(kh,(n+1)\tau)-P(kh,n\tau)}{\tau}=\frac{h^2}{2\tau}\frac{P((k+1)h,n\tau)-2P(kh,n\tau)+P((k-1)h,n\tau)}{h^2} \tag{45}$$

  If the continuum limit $\tilde{p}(x,t):=\lim_{h,\tau\to 0}P(kh,n\tau)$ exists, then $\int_I \tilde{p}(x,t)dx$ is the probability of finding the particle in the interval $I\subset\mathbb{R}$. The probability distribution $\tilde{p}$ satisfies the heat equation

$$\frac{\partial \tilde{p}(x,t)}{\partial t}=D\frac{\partial^2\tilde{p}(x,t)}{\partial x^2}. \tag{46}$$

  This is an example of a Fokker-Planck equation, describing the probability distribution of a stochastic process (in our case $\{B_t\}_{t\in\mathbb{R}_+}$).

- In the continuum limit 'starting' the process at $x=0$ corresponds to the initial condition $\tilde{p}(x,0)=\delta_0(x)$ (Dirac delta) as this means zero probability of finding the particle in any I not containing $\{0\}$. Then the solution of the heat equation is:

$$\tilde{p}(x,t)=\frac{1}{\sqrt{4\pi Dt}}\exp\left(-\frac{x^2}{4Dt}\right). \tag{47}$$

  Setting $D=1/2$ (as above) this is the probability distribution of B_t.

Lecture 2
The goal is now construct a joint probability space for all B_t, $t\in\mathbb{R}_+$.

5 Stochastic processes

1. Def. A stochastic process is a parametrized collection of random variables

$$\{X_t\}_{t\in T} \tag{48}$$

   defined on a probability space $(\Omega,\mathcal{F},P)$ and assuming values in $\mathbb{R}^d$. (T is a set. Typical choices: $\mathbb{R}_+:=[0,\infty)$, $[a,b]$, $\mathbb{N}$...).

2. Remark: For fixed $t\in T$ we have a random variable $\Omega\ni\omega\mapsto X_t(\omega)\in\mathbb{R}^d$.

3. Def: For fixed ω the function $T\ni t\mapsto X_t(\omega)$ is called a path of X_t. (Interpretation: result of an experiment ω at time t. Eg. position of a random walker at time t).

4. Remark: For fixed stochastic process $\{X_t\}_{t\in T}$ we may associate with each ω a function $t\mapsto X_t(\omega)=:\omega(t)$ from T into $\mathbb{R}^d$. (Recall that for random walk $\omega=HHT\dots TH$ was uniquely determining the path).

   - Then, we regard Ω as a subset of $\tilde{\Omega}=(\mathbb{R}^d)^T$. Namely, with any ω we associate a path $\{t\to X_t(\omega)\}$. We denote this embedding by $\iota:\Omega\to\tilde{\Omega}$.
   - The σ-algebra $\tilde{\mathcal{F}}$ on $\tilde{\Omega}$ is generated by 'cylinder sets' of the form

$$C_{t_1,\dots,t_k;F_1,\dots,F_k}:=\{\tilde{\omega}\in(\mathbb{R}^d)^T\ :\ \tilde{\omega}(t_1)\in F_1,\dots,\omega(t_k)\in F_k\},\quad F_i\subset\mathbb{R}^d \text{ Borel}. \tag{49}$$

     Fact: For T countable, $\tilde{\mathcal{F}}$ is the Borel algebra on $\tilde{\Omega}$ provided that $\tilde{\Omega}$ is equipped with the product topology. (The product topology is the smallest topology containing sets (49) with F_i open). [There is a statement in [Ok, p.10] that the same is true for $T=[0,\infty)$, but I have doubts if it really holds. Remains to be checked.]

- We define the measure $\tilde{P}$ on $(\tilde{\Omega}, \mathcal{B})$ as follows

$$\tilde{P}(C) = P(\iota^{-1}(C)) \tag{50}$$

for cylinder sets C.

In view of the above, one can see the stochastic process as a measure $\tilde{P}$ on the space of paths $(\tilde{\Omega}, \tilde{\mathcal{F}})$. We include all possible paths in $\tilde{\Omega} = (\mathbb{R}^d)^T$, but some of them may have probability zero. For example, for the random walk, the path for which after time τ the walker is at position $2h$ has probability zero.

5. Def: Suppose that $\{X_t\}_{t\in T}$ and $\{Y_t\}_{t\in T}$ are stochastic processes on $(\Omega, \mathcal{F}, P)$. Then, we say that $\{X_t\}_{t\in T}$ is a version of (or modification of) $\{Y_t\}_{t\in T}$ if

$$P(\{\omega \in \Omega \,:\, X_t(\omega) = Y_t(\omega)\}) = 1 \text{ for all } t. \tag{51}$$

6. Def: The finite-dimensional distributions of the process $X = \{X_t\}_{t\in T}$ are the measures $\nu_{t_1,\dots,t_k}$ defined on $\mathbb{R}^{dk}$, $k = 1, 2, \dots,$ by

$$\begin{aligned}\nu_{t_1,\dots,t_k}(F_1 \times F_2 \times \cdots \times F_k) &= P[X_{t_1} \in F_1, \dots, X_{t_k} \in F_k] \\ &= \int \chi(X_{t_1}(\omega) \in F_1) \dots \chi(X_{t_k}(\omega) \in F_k) dP(\omega),\end{aligned} \tag{52}$$

where $t_1, \dots, t_k \in T$ and $F_1, \dots F_k$ denote Borel sets in $\mathbb{R}^d$.

7. Under certain conditions, it is possible to reconstruct the stochastic process $\{X_t\}_{t\in T}$ from the distributions:

Theorem 5.1. *(Kolmogorov's extension theorem). For all $t_1, \dots, t_k \in T$ let $\nu_{t_1,\dots,t_k}$ be probability measures on $\mathbb{R}^{dk}$ s.t.*

$$\nu_{t_{\sigma(1)},\cdots,t_{\sigma(k)}}(F_1 \times \cdots \times F_k) = \nu_{t_1,\cdots,t_k}(F_{\sigma^{-1}(1)} \times \cdots \times F_{\sigma^{-1}(k)}) \tag{K1}$$

for all permutations σ on $\{1, 2, \dots, k\}$ and

$$\nu_{t_1,\cdots,t_k}(F_1 \times \cdots \times F_k) = \nu_{t_1,\dots,t_k,t_{k+1},\dots,t_{k+m}}(F_1 \times \cdots \times F_k \times \mathbb{R}^d \times \cdots \times \mathbb{R}^d). \tag{K2}$$

Then, there exists a probability space $(\Omega, \mathcal{F}, P)$ and a stochastic process $\{X_t\}_{t\in T}$ on Ω s.t.

$$\nu_{t_1,\dots,t_k}(F_1 \times \cdots \times F_k) = P(X_{t_1} \in F_1, \cdots, X_{t_k} \in F_k), \tag{53}$$

for all $t_i \in T$, $k \in \mathbb{N}$ and all Borel sets F_i.

6 Brownian motion

1. To construct $\{B_t\}_{t\geq 0}$ it suffices to specify a family of probability measures $\{\nu_{t_1,\dots t_k}\}_{t_1,\dots t_k \geq 0}$ satisfying the conditions from the Kolmogorov theorem.

   - For fixed $x \in \mathbb{R}^d$ define

$$p(t, x, y) = (2\pi t)^{-d/2} \exp\big(-\frac{|x-y|^2}{2t}\big) \quad \text{for } y \in \mathbb{R}^d, t > 0. \tag{54}$$

   - If $0 \leq t_1 \leq t_2 \leq \cdots \leq t_k$ define a measure $\nu_{t_1,\dots t_k}$ on $\mathbb{R}^{dk}$ by

$$\begin{aligned}&\nu_{t_1,\dots,t_k}(F_1 \times \cdots \times F_k) \\ &= \int_{F_1\times\cdots\times F_k} p(t_1, x, x_1) p(t_2 - t_1, x_1, x_2) \cdots p(t_k - t_{k-1}, x_{k-1}, x_k) dx_1 \dots dx_k,\end{aligned} \tag{55}$$

   where we use the convention $p(0, x, y)dy = \delta_x(y)dy$ (unit point mass at x in physicist's notation).

- Extend to all finite sequences using (K1). Since $\int p(t,x,y)dy = 1$ for all $t \geq 0$, also (K2) holds.

2. Remark: Note that for $x = 0$, $d = 1$, we have

$$\nu_{t_1}(F_1) = \int_{F_1} p(t_1, 0, x_1)dx_1 = \int_{F_1} \tilde{p}(t_1, x_1)dx_1 \tag{56}$$

where $\tilde{p}$ appeared in (47) as a limit of random walk probability distributions, satisfying $\frac{\partial \tilde{p}(x,t)}{\partial t} = D\frac{\partial^2 \tilde{p}(x,t)}{\partial x^2}$.

3. Consequently, there exists a probability space $(\Omega, \mathcal{F}, P^x)$ and a stochastic process $\{B_t\}_{t\geq 0}$ on Ω s.t. the finite-dimensional distributions B_t are given by:

$$P^x(B_{t_1} \in F_1, \cdots, B_{t_k} \in F_k) = \int_{F_1\times\cdots\times F_k} p(t_1, x, x_1)\cdots p(t_k - t_{k-1}, x_{k-1}, x_k)dx_1\cdots dx_k. \tag{57}$$

This formula determines the law of a family of random variables in the sense of (17). In the notation of (17) it reads

$$P^x(B_{t_1} \in F_1, \cdots, B_{t_k} \in F_k) = \mu_{B_{t_1},\dots,B_{t_k}}(F_1 \times \ldots \times F_k). \tag{58}$$

Hence, the measure on the l.h.s. of (57) is absolutely continuous w.r.t. the Lebesgue measure on the r.h.s. with density given by $p(t_1, x, x_1)\cdots p(t_k - t_{k-1}, x_{k-1}, x_k)$. Thus, for any Borel measurable function f on $\mathbb{R}^d$ we have

$$\begin{aligned}
&\int f(B_{t_1}, \dots, B_{t_k})P^x(B_{t_1} \in dx_1, \cdots, B_{t_k} \in dx_k) \\
&:= \int f(x_1, \dots, x_k)\mu_{B_{t_1},\dots,B_{t_k}}(x_1, \dots, x_k) \\
&= \int f(x_1, \dots, x_k)\{p(t_1, x, x_1)\cdots p(t_k - t_{k-1}, x_{k-1}, x_k)\}dx_1\cdots dx_k.
\end{aligned} \tag{59}$$

The notation $P^x(B_{t_1} \in dx_1, \cdots, B_{t_k} \in dx_k)$, popular in probability theory, is defined by the above relations.

4. Def: The process constructed above is called (a version of) a Brownian motion starting at x.

5. Properties of the Brownian motion:

   (i) $\{B_t\}_{t\in\mathbb{R}_+}$ is a Gaussian process, i.e., for all $0 \leq t_1 \leq \cdots \leq t_k$ the random variable $Z = (B_{t_1}, \dots, B_{t_k}) \in \mathbb{R}^{dk}$ has a (multi-)normal distribution. (To be defined below.)

   (ii) B_t has independent increments, i.e.,

   $$B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_k} - B_{t_{k-1}}, \tag{60}$$

   are independent for all $0 \leq t_1 < t_2 \cdots < t_k$.

   (iii) Brownian motion has a version with continuous paths. By considerations above (49), Brownian motion can be seen as the space $C([0,\infty), \mathbb{R}^d)$ with probability measures P^x.

These properties will be proven after preparatory Section 7.

7 Gaussian random variables

Definition 7.1. *Let $(\Omega, \mathcal{F}, P)$ be a probability space. A random variable $X : \Omega \to \mathbb{R}$ is normal if the distribution of X has a density of the form*

$$p_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(- \frac{(x-m)^2}{2\sigma^2} \right) \tag{61}$$

where $\sigma > 0$ and $m \in \mathbb{R}$. That is

$$P(X \in G) = \int_G p_X(x)dx \quad \textit{for} \quad G \subset \mathbb{R} \textit{ Borel.} \tag{62}$$

By Gaussian integration:

$$E(X) = \int_\Omega X dP = \int_\mathbb{R} x p_X(x)dx = \int_\mathbb{R} (x+m)\frac{1}{\sigma\sqrt{2\pi}} \exp\left(- \frac{x^2}{2\sigma^2} \right) dx = m, \tag{63}$$

$$\mathrm{var}(X) = E[(X-m)^2] = \int_\mathbb{R} (x-m)^2 p_X(x)dx = \sigma^2. \tag{64}$$

References

[Ok] B. Øksendal. *Stochastic Differential Equations. An Introduction with Applications.* Springer-Verlag, 2000.

[Ro] G. Roepstorff. *Path Integral Approach to Quantum Physics. An Introduction.* Springer-Verlag, 1991.

[BT] D.P. Bertsekas and J.N. Tsitsiklis. *Introduction to Probability.* Lecture Notes, M.I.T. Fall, 2000.

[Ru] W. Rudin. *Real and complex analysis.* McGraw-Hill Inc. 1970