Introduction to Brownian motion

October 31, 2013

Lecture notes for the course given at Tsinghua university in May 2013.
Please send an e-mail to nicolas.curien@gmail.com for any error/typo found.

Historic introduction

From wikipedia : Brownian motion is the random moving of particles suspended in a fluid (a liquid or a gas) resulting from their bombardment by the fast-moving atoms or molecules in the gas or liquid. In 1827, the botanist Robert Brown, looking through a microscope at particles found in pollen grains in water, noted that the particles moved through the water but was not able to determine the mechanisms that caused this motion. Atoms and molecules had long been theorized as the constituents of matter, and many decades later, Albert Einstein published a paper in 1905 that explained in precise detail how the motion that Brown had observed was a result of the pollen being moved by individual water molecules. This explanation of Brownian motion served as definitive confirmation that atoms and molecules actually exist, and was further verified experimentally by Jean Perrin in 1908. Perrin was awarded the Nobel Prize in Physics in 1926 "for his work on the discontinuous structure of matter" (Einstein had received the award five years earlier "for his services to theoretical physics" with specific citation of different research). The direction of the force of atomic bombardment is constantly changing, and at different times the particle is hit more on one side than another, leading to the seemingly random nature of the motion. This transport phenomenon is named after Robert Brown. The first mathematical rigorous construction of Brownian motion is due to Wiener in 1923 (that is why Brownian motion is sometimes called Wiener process).

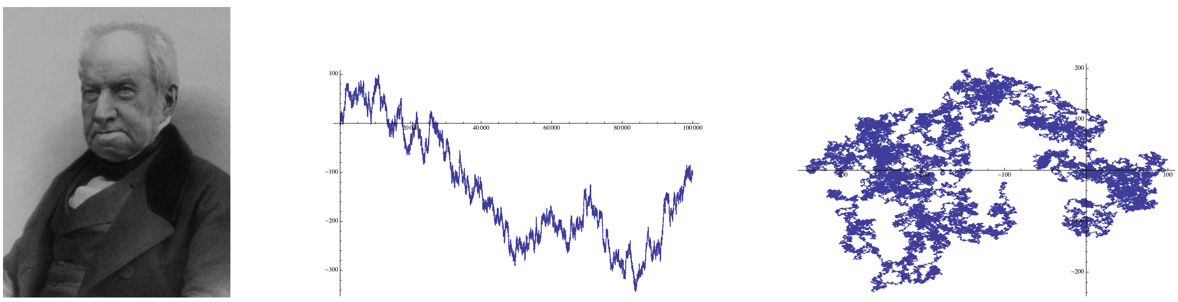

Figure 1: Robert Brown and Brownian motions in 1 and 2 dimensions.

A discrete model

A possible model of the above motion of a particle in d dimension can be as follows. We consider that the particle is moving as a random walk $S_n = Y_1 + \ldots + Y_n$ with Y_i i.i.d. uniform over $\{-1,1\}^d$ (at each step, the d coordinates are updated by an independent fair ± 1). We then look at the motion of the particle seen from far away. Since $n^{-1}S_n \to 0$, the central limit theorem tells us that the good renormalization in space is the square root of the time, and we thus consider the random function

$$S_n^*(t) = \frac{1}{\sqrt{n}} S_{[nt]},$$

where $[x]$ is the largest integer less than x. For every $0 = t_0 < t_1 < t_2 < \ldots < t_p$ the vectors $(S_n^*(t_i) - S_n^*(t_{i-1})), i \in \{1, \ldots, p\}$ are independent and converge in distribution by the central limit theorem towards centered Gaussian vectors of covariance $(t_i - t_{i-1})\mathrm{Id}$. Obviously this remains true (up to a multiplicative factor) if we change the distribution of the steps as long as they remain centered and have finite variance. This leads us to set up a definition for the (candidate) continuous limiting process:

Definition 1. *A d-dimensional Brownian motion (starting from 0) is a family of $\mathbb{R}^d$-valued random variables $(B_t : t \geq 0)$ living on a probability space $(\Omega, \mathcal{F}, P)$ such that*

- $B_0 = 0$ *almost surely,*
- *for every $0 = t_0 < t_1 < t_2 < \ldots < t_p$ the variables $B_{t_i} - B_{t_{i-1}}$ for $i \in \{1, \ldots, p\}$ are independent and*

$$B_{t_i} - B_{t_{i-1}} = \mathcal{N}\big(0, \mathrm{Id}(t_i - t_{i-1})\big).$$

- *the function $t \mapsto B_t$ is almost surely continuous.*

The last point of the definition could seem trivial from a physics point of view but turns out to be mathematically essential. We will see that existence of Brownian motion is not trivial. We will come back later to the fact that Brownian motion is the universal limit of scaled random walks. Before constructing Brownian motion, let us quickly dive into the realm of stochastic processes.

Remark. Brownian motion thus has *stationary* and *independent* increments. Meaning that $B_{t_i} - B_{t_{i-1}}$ for $i \in \{1, \ldots, p\}$ are independent and $B_{t+s} - B_t = B_{t'+s} - B_{t'} = B_s$ in distribution for every $t, t' \geq 0$. Among the class of stochastic processes satisfying these assumptions (The Lévy processes) Brownian motion is the only continuous one. Do you know another (non-continuous) process $(X_t : t \geq 0)$ which has stationary and independent increments?

1 Generality on stochastic processes

Let us call *stochastic process* a family $(X_t : t \in \mathbb{R}_+)$ of random variables with values in $\mathbb{R}$ endowed with its Borel σ-field $\mathcal{B}$ (the vectorial case is similar) defined on a probability space $(\Omega, \mathcal{F}, P)$. Hence, for every $\omega \in \Omega$ we can speak about the "trajectory" $t \mapsto X_t(\omega)$ and interpret it as a "random function". Where does this object live? How to characterize it? What are its properties?

1.1 First approach

We could consider the random process $(X_t)_{t\geq 0}$ as taking values in $\mathbb{R}^{\mathbb{R}_+}$. As a product space, $\mathbb{R}^{\mathbb{R}_+}$ is endowed with the product σ-field $\mathcal{B}^{\mathbb{R}_+}$ (where $\mathcal{B}$ is the Borel σ-field on $\mathbb{R}$) generated by the coordinate mappings

$$(X_t)_{t\geq 0} \in \mathbb{R}^{\mathbb{R}_+} \mapsto X_s \in (\mathbb{R}, \mathcal{B}).$$

Since $\omega \mapsto X_s(\omega) \in (\mathbb{R}, \mathcal{B})$ is measurable for every $s \geq 0$, it follows that $\omega \mapsto (X_t(\omega), t \geq 0) \in (\mathbb{R}^{\mathbb{R}_+}, \mathcal{B}^{\mathbb{R}_+})$ is measurable as well. The law of such a random process is thus a probability measure on $(\mathbb{R}^{\mathbb{R}_+}, \mathcal{B}^{\mathbb{R}_+})$. The *finite-dimensional* distributions (or marginals) of $X = (X_t : t \geq 0)$ are the laws of all the finite vectors $(X_{t_1}, X_{t_2}, \ldots, X_{t_p})$ for every $t_1 \leq t_2 \leq \ldots \leq t_p \in \mathbb{R}_+$.

Proposition 1. *The finite-dimensional distributions of X characterize its law over* $(\mathbb{R}^{\mathbb{R}_+}, \mathcal{B}^{\mathbb{R}_+})$.

Proof. Denote ν the law of (X) over $(\mathbb{R}^{\mathbb{R}_+}, \mathcal{B}^{\mathbb{R}_+})$. The finite-dimensional distributions of X fix the ν-probability of events of the form $\{f \in \mathbb{R}^{\mathbb{R}_+} : f(t_1) \in A_1, \ldots, f(t_p) \in A_p\}$ for $A_i \in \mathcal{B}$. Such events are called "cylinders". It is easy to see that cylinder events is a collection that is closed under finite intersection and which generates $\mathcal{B}^{\mathbb{R}_+}$ as a σ-field. Hence the monotone class theorem entails that ν is characterized by its values on cylinder events. □

Exercise 1. Show that items 1 and 2 of Definition 1 characterize the finite-dimensional distribution of Brownian motion.

Exercise 2 (One-dimensional does not suffices)**.** Find two stochastic processes $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ which have the same one-dimensional distributions i.e. such that $X_t = Y_t$ in distribution for every $t \geq 0$, but (X) and (Y) have different laws. * Same question with two-dimensional distributions.

This approach enjoys an abstract criterion to construct stochastic processes:

Theorem 1 (Kolmogorov extension theorem)**.** *Suppose that we are given a family of finite-dimensional laws $\pi_{t_1,\ldots,t_n}$ for every $t_1, \ldots, t_n \in \mathbb{R}_+$ which are coherent in the sense that if $(X_i)_{i=t_1,\ldots,t_n,t_{n+1}}$ has law $\pi_{t_1,\ldots,t_n,t_{n+1}}$ then $(X_i)_{i=t_1,\ldots,t_n}$ has law $\pi_{t_1,\ldots,t_n,}$. Then there exists a probability π on $(\mathbb{R}^{\mathbb{R}_+}, \mathcal{B}^{\mathbb{R}_+})$ such that $(X_t)_{t\geq 0}$ has law π and $(X_i)_{i=t_1,\ldots,t_n}$ has law $\pi_{t_1,\ldots,t_n}$.*

Proof. Omitted. □

Exercise 3. Check that the finite-dimensional distributions of Brownian motion are coherent and deduce that there exists a stochastic process verifying conditions 1 and 2 of Definition 1.

The problem with the σ-field $\mathcal{B}^{\mathbb{R}_+}$ is that it is "trajectorially" very poor.

Exercise 4. *

1. Show that the event $\{t \mapsto X_t \text{ is continuous}\}$ is not measurable for the σ-field $\mathcal{B}^{\mathbb{R}_+}$.

   Hint : Construction two stochastic processes X, X' which have the same finite-dimensional distributions and such that X is continuous almost surely but X' is not.

2. Show that the σ-field $\mathcal{B}^{\mathbb{R}_+}$ is made of all sets that can be written $A \times \mathbb{R}^{\mathbb{R}_+\setminus D}$, where D is a countable set of indices of $\mathbb{R}_+$ and $A \in \mathcal{B}^D$.

1.2 Second approach

We now suppose that our stochastic process (X) is continuous and thus lives in $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$. We thus have to endow this set with a σ-field. The first try would be to use the $\|\|_\infty$ norm but that would yield a non-separable space which causes many problems, see Exercise 6. We instead use the topology of uniform convergence over every compact sets which is metrizable by

$$d(f,g) = \sum_{i=1}^{+\infty} 2^{-i}(1 \wedge \sup_{[0,i]} |f-g|).$$

Exercise 5. Check that d is a distance on $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$ which makes it separable and complete.

The Borel σ-field associated to this topology is denoted by $\mathcal{B}_c$. We also denote $\mathcal{C}$ the trace of the σ-field $\mathcal{B}^{\mathbb{R}_+}$ on $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$.

Proposition 2. *We have $\mathcal{B}_c = \mathcal{C}$.*

Proof. The inclusion $\mathcal{C} \subset \mathcal{B}_c$ follows from the fact that the coordinate mappings $f \mapsto f(t)$ are continuous for d hence measurable. For the other inclusion, since $(\mathcal{C}(\mathbb{R}_+, \mathbb{R}), d)$ is separable, it suffices to show that balls for d are measurable. But by continuity of the functions considered we can give an equivalent definition of d using rational values:

$$\sum_{i=1}^{+\infty} 2^{-i}(1 \wedge \sup_{s\in[0,i]} |f(s)-g(s)|) = \sum_{i=1}^{+\infty} 2^{-i}(1 \wedge \sup_{s\in[0,i]\cap\mathbb{Q}} |f(s)-g(s)|).$$

It easily follows that balls for d are $\mathcal{C}$-measurable. $\square$

Exercise 6 (Nightmares[1]***)**.** Let ℓ^∞ be the space of bounded real sequences. We endow this space with two σ-fields: the (trace of the) product σ-field $\mathcal{B}^{\mathbb{N}}$ and the Borel σ-field from the norm $\|\|_\infty$ over ℓ^∞ where $\|u\|_\infty = \sup_{i\geq 0} |u_i|$.

1. Show that $\mathcal{B}^{\mathbb{N}} \subset \mathcal{B}_{\|\|_\infty}$.

Consider a sequence A_n of intervals of $[0,1]$ such that $|A_n| \to 0$ when $n \to \infty$ and such that for every $t \in [0,1]$, the set $B_t = \{n \geq 0 : t \in A_n\}$ is infinite. Finally, define the mapping

$$\xi : t \in [0,1] \mapsto (\mathbf{1}_{t\in A_1}, \mathbf{1}_{t\in A_2}, \mathbf{1}_{t\in A_3}, \ldots).$$

2. Show that ξ is measurable from $([0,1], \mathcal{B}_{[0,1]})$ into $(\ell^\infty, \mathcal{B}^{\mathbb{N}})$.

3. Let $A \subset [0,1]$. Show that $\xi(A)$ is closed in $(\ell^\infty, \|\|_\infty)$, and that $\xi^{-1}(\xi(A)) = A$.

4. Assume the axiom of choice, and recall that there are subsets of $[0,1]$ which are non Lebesgue measurable. Conclude that ξ is not measurable from $([0,1], \mathcal{B}_{[0,1]})$ into $(\ell^\infty, \mathcal{B}_{\|\|_\infty})$ and that $\mathcal{B}^{\mathbb{N}} \neq \mathcal{B}_{\|\|_\infty}$.

[1]Taken from "Probability Distributions on Banach Spaces" by Vakhania et al, p. 23 - 24.

Intermezzo: Gaussian process

A random vector $(X_1, \ldots, X_p) \in \mathbb{R}^p$ is a *Gaussian vector* if any linear combinaison of the X_i is Gaussian. If $N_1, \ldots, N_k$ are independent Gaussian random variables then any vector whose entries are (fixed) linear combinaisons of the $N_1, \ldots, N_k$ is a Gaussian vector (actually any Gaussian vector can be represented as such). Let us give an example of a vector with Gaussian entries which is not a Gaussian vector. Consider a standard normal variable N and let ϵ be an independent fair $\pm$ Bernoulli coin flip. It is easy to see that the vector $(X, \epsilon X)$ has both entries distributed as $\mathcal{N}(0,1)$ but this vector is not Gaussian because $X + \epsilon X$ is not Gaussian.

The marvelous property is that the law of a Gaussian vector is completely characterized by the means $E[X_i]$ and the covariance matrix $(E[X_i X_j])_{1 \le i,j \le n}$. Indeed, for any $\lambda_1, \ldots \lambda_n \in \mathbb{R}$, the characteristic function $E[\exp(\mathrm{i} \sum \lambda_i X_i)]$ is the Fourier transform of a Gaussian variable $\sum \lambda_i X_i$ and thus equal $\exp(\mathrm{i}m - \sigma^2/2)$ where m is its mean and σ^2 its variance. Both m and σ^2 are computable in terms of the λ_i, of the covariance matrix and the means of the X_i. In particular if $(X_1, \ldots, X_n, Y_1, \ldots, Y_m)$ is a Gaussian vector and such that $\mathrm{Cov}[X_i Y_j] = 0$ for every $i \in \{1, \ldots, n\}$ and every $j \in \{1, \ldots, m\}$ then $(X_1, \ldots, X_n)$ is independent of $(Y_1, \ldots, Y_m)$.

A stochastic process $(X_t)_{t \ge 0}$ is a Gaussian process if any finite-dimensional vector $(X_{t_1}, \ldots, X_{t_p})$ is a Gaussian vector. By the above remark, the finite-dimensional distributions of (X) are completely characterized by the means of X_t and the covariance function $s, t \mapsto E[X_s X_t]$.

Example. Brownian motion is a Gaussian process. Indeed, for any $t_1, \ldots, t_p$ the vector $(B_{t_1}, \ldots, B_{t_p})$ is a linear combinaison of independent Gaussian variables $(B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_p - t_{p-1}})$ and is thus a Gaussian vector. Furthermore, the process is centered meaning that $E[B_t] = 0$ for all $t \ge 0$ and the covariance function is

$$E[B_s B_t] = \min(s,t) \qquad \text{(exercise)}.$$

The last line is thus an alternative (and much simpler to manipulate) to the points 1 and 2 of Definition 1.

Exercise 7. Let X_n be a sequence of Gaussian vectors.

1. Suppose that $b = \lim E[X_n]$ (the mean vector) and $\Sigma = \lim \mathrm{Cov}(X_n)$ (the covariance matrix) exist. Show that $X_n \to X$ in distribution where X is Gaussian with mean b and covariance matrix Σ.

2. * Suppose that $X_n \to X$ almost surely. Show that $X_n \to X$ in $\mathbb{L}^2$ and deduce that $b = \lim E[X_n]$ (the mean vector) and $\Sigma = \lim \mathrm{Cov}(X_n)$ (the covariance matrix) exist.

We finish this section by stating a bound on the tail of the normal distribution which will be used a several places in this course.

Lemma 1 (Tail bounds for the standard normal)**.** *Let N be a standard normal $\mathcal{N}(0,1)$ random variable. Then we have*

$$P(N > x) \sim \frac{1}{x} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad \text{as } x \to \infty.$$

Proof. On the one hand we have

$$\int_x^\infty \frac{ds}{\sqrt{2\pi}} e^{-s^2/2} \le \int_x^\infty \frac{ds}{\sqrt{2\pi}} \frac{s}{x} e^{-s^2/2} = \frac{1}{x} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

On the other hand, after performing an integration by parts

$$\int_x^\infty \frac{ds}{\sqrt{2\pi}} e^{-s^2/2} = \left[-\frac{1}{\sqrt{2\pi}} \frac{1}{s} e^{-s^2/2} \right]_x^\infty - \int_x^\infty \frac{1}{s^2} \frac{ds}{\sqrt{2\pi}} e^{-s^2/2}.$$

The second term is easily seen to be negligible compared to the first one as $x \to \infty$. □

2 Construction of Brownian motion

Theorem 2. *Brownian motion exists*[2].

Proof. The following proof is due to Paul Lévy. We first construct the process for $t \in [0, 1]$ and in dimension 1. Let

$$\mathcal{D}_n = \{k2^{-n} : 0 \le k \le 2^n\}$$

the set of dyadic points of order n and $\mathcal{D} = \cup \mathcal{D}_n$ the set of all dyadic points of $[0, 1]$. Let $(\Omega, \mathcal{F}, P)$ a probability space where we have a collection $(Z_i : i \in \mathcal{D})$ of independent identically distributed standard $\mathcal{N}(0, 1)$ variables. We put $B_0 = 0$ and $B_1 = Z_1$ and then construct inductively B on the set $\mathcal{D}_n$ so that B has the right finite-dimensional marginals on $\mathcal{D}_n$. This is done for $\mathcal{D}_0$. Suppose that we succeeded in doing it at step $n - 1$. Then for every $t \in \mathcal{D}_n \backslash \mathcal{D}_{n-1}$ we update the interpolated value of B_t by adding $2^{-(n+1)/2} Z_t$, formally we put

$$B_t = \frac{B_{t+2^{-n}} + B_{t-2^{-n}}}{2} + \frac{Z_t}{2^{(n+1)/2}}.$$

Note that after step n the values of $B_t \in \mathcal{D}_n$ are fixed and are independent of $\{Z_t : t \notin \mathcal{D}_n\}$. It is easy to see that after step $n \ge 0$, the variables $B_{(k+1)2^{-n}} - B_{k2^{-n}}$ for $k \in \{0, \dots, 2^n - 1\}$ are independent and distributed as $\mathcal{N}(0, 2^{-n})$. Indeed, it is easy to see by induction that the covariance matrix of the Gaussian vector $(B_{i2^{-n}})_{0 \le i \le 2^n}$ is equal to

$$E[B_t B_s] = (s \wedge t) \mathbf{1}_{t \in \mathcal{D}_n} \mathbf{1}_{s \in \mathcal{D}_n}.$$

This implies that B has the right finite-dimensional distributions over $\mathcal{D}_n$. We denote by $B^{(n)}$ the function which interpoles linearly $\{B_t : t \in \mathcal{D}_n\}$.

Lemma 2. *The random functions $B^{(n)}$ almost surely converge towards a random continuous function B (with a slight abuse of notation) for the uniform norm $\|\|$ over $[0, 1]$.*

Proof. Fix $n \ge 1$ and let us upper bound $\|B^{(n)} - B^{(n-1)}\|$. To go from the function $B^{(n-1)}$ to $B^{(n)}$ the slope of $B^{(n-1)}$ over each dyadic interval of length $2^{-(n-1)}$ is updated in the middle by adding an independent $2^{-(n+1)/2} \mathcal{N}(0, 1)$. Hence we have

$$\|B^{(n)} - B^{(n-1)}\| \le 2^{-(n+1)/2} \max_{t \in \mathcal{D}_n \backslash \mathcal{D}_{n-1}} |Z_t|.$$

Fix $\alpha \in (1/\sqrt{2}, 1)$. We thus have

$$\begin{aligned} P(\|B^{(n)} - B^{(n-1)}\| \ge \alpha^n) &= P(2^{-(n+1)/2} \max_{t \in \mathcal{D}_n \backslash \mathcal{D}_{n-1}} |Z_t| \ge \alpha^n) \\ &\le 2^n P(|Z| \ge \alpha^n 2^{(n+1)/2}) \\ &\le C 2^n \exp(-(\alpha^2 2)^n) \end{aligned}$$

[2]That is, we can construct a probability space which supports a Brownian motion

for some $C > 0$ using Lemma 1. Since $2\alpha^2 > 1$, the last probabilities are summable for $n = \{1, 2, \ldots\}$ hence by the Borel–Cantelli lemma we have almost surely $\|B^{(n)} - B^{(n-1)}\| \leq \alpha^n$ eventually. Consequently, the functions $B^{(n)}$ almost surely converge (at exponential speed!) for the $\|\|$ norm towards a random continuous function B. $\square$

It remains to show that B has the desired finite-dimensional marginals. Fix $t_1 \leq \ldots \leq t_p \in [0, 1]$ and let $t_1^{(n)} \leq \ldots \leq t_p^{(n)}$ be approximations of the latter in $\mathcal{D}_n$. We already know that the variables $B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}}$ are independent and are distributed as $\mathcal{N}(0, t_{i+1}^{(n)} - t_i^{(n)})$. By continuity of B these variables converge towards $B_{t_{i+1}} - B_{t_i}$. By Exercice 7 the limit has the desired distribution.

To construct (one-dimensional) Brownian motion on the whole of $\mathbb{R}_+$ we just concatenate independent copies of Brownian motion over the time interval $[0, 1]$. The Brownian motion in dimension d is constructed as $(B_1(t), \ldots, B_d(t))$ where $B_1, \ldots, B_d$ are independent standard Brownian motions in one dimension. It is easy to see that these processes have the desired properties. $\square$

Canonical representation. Let (B) be a Brownian motion defined over $(\Omega, \mathcal{F}, P)$. The push-forward of P by the measurable mapping $\omega \mapsto B_t(\omega) \in (\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d), \mathcal{B}_c)$ is called the Wiener measure and is denoted by $\mathbb{P}_0$. For every $t \geq 0$, we put for $x \in \mathbb{R}^d$

$$p_t(x) = \frac{1}{\sqrt{2\pi t}^d} \exp\left(-\frac{\|x\|^2}{2t}\right).$$

The Wiener measure is thus characterized on $(\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d), \mathcal{B}_c)$ by

$$\begin{aligned}
&\mathbb{P}_0\Big(\{\omega \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d) : \omega(t_1) \in A_1, \ldots, \omega(t_p) \in A_p\}\Big) \\
&= \int_{A_1 \times A_2 \ldots \times A_p} dy_1 \ldots dy_p \; p_{t_1}(y_1) p_{t_2 - t_1}(y_2 - y_1) \ldots p_{t_p - t_{p-1}}(y_p - y_{p-1}),
\end{aligned}$$

for every Borel sets $A_i \in \mathcal{B}$. In the remaining of this course, we will often use the canonical representation of BM (Brownian motion) by setting $(\Omega, \mathcal{F}, P) = (\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d), \mathcal{B}_c, \mathbb{P}_0)$ and $B_t(\omega) = \omega(t)$. For every $x \in \mathbb{R}^d$, under the probability $\mathbb{P}_x$, the Brownian motion starts at x.

Exercise 8. * Let $f : [0, 1] \to \mathbb{R}$ be a continuous function such that $f(0) = 0$. Prove that for every $\varepsilon > 0$ we have

$$P\left(\sup_{t \in [0,1]} |B_t - f(t)| < \varepsilon\right) > 0.$$

2.1 Simple properties

Proposition 3. *Let (B) be a standard Brownian motion in dimension $d \geq 1$.*

- *Isometry. If ϕ is a linear isometry of $\mathbb{R}^d$ then $\phi(B_t)$ is a Brownian motion.*
- *Translation. For every $s \geq 0$, the process $B_t^{(s)} = B_{s+t} - B_s$ is a Brownian motion.*
- *Time reversal. The process $(B_1 - B_{1-t})_{t \in [0,1]}$ is distributed as $(B_t)_{t \in [0,1]}$.*
- *Scale invariance. For every $a > 0$, the process $(\frac{1}{a} B_{a^2 t} : t \geq 0)$ is a Brownian motion.*

Proof. All the processes considered are continuous, Gaussian, centered (mean zero) and the covariance functions are easily seen to coincide with that of Brownian motion. □

Exercise 9 (A first use of Brownian scaling)**.** For $b \geq 0$, let $\tau_{[-b,b]} = \inf\{t \geq 0 : B_t \notin [-b,b]\}$. Show that $E[\tau_{[-b,b]}] = b^2 E[\tau_{[-1,1]}]$.

Exercise 10 (Time inversion)**.** * The process $(\mathbf{1}_{t>0} t B_{1/t} : t \geq 0)$ is a Brownian motion.

Exercise 11 (Ornstein-Uhlenbeck diffusion)**.** For $t \in \mathbb{R}$, let $X_t = e^{-t} B_{e^{2t}}$.

1. Show that for every $t \in \mathbb{R}$, X_t is distributed as $\mathcal{N}(0,1)$.
2. Show that $(X)_{t\in\mathbb{R}}$ is a centered Gaussian process and compute its covariance function $E[X_s X_t]$ for $s, t \in \mathbb{R}$.
3. Show that for every $s \in \mathbb{R}$ and $a \geq 0$, the processes $(X_{-t})_{t\in\mathbb{R}}$ and $(X_{t+s})_{t\in\mathbb{R}}$ have the same law as $(X)_{t\in\mathbb{R}}$ on $\mathcal{C}(\mathbb{R},\mathbb{R})$ endowed with the Borel σ-field of uniform convergence over every compact.

Exercise 12 (Quadratic variation)**.** * Let $0 \leq a < b$. For every $n \geq 0$, set

$$X_n = \sum_{k=1}^{2^n} \left(B_{a+k(b-a)2^{-n}} - B_{a+(k-1)(b-a)2^{-n}} \right)^2 .$$

Compute the mean and the variance of X_n and find the almost sure limit of X_n. Deduce that a.s. BM has no finite variation on any interval –we recall that a function $f : \mathbb{R}_+ \longrightarrow \mathbb{R}$ has finite variation on $[a,b]$ if there exists $M > 0$ such that for every subdivision $a = t_0 < t_1 < \ldots < t_p = b$ we have

$$\sum_{i=1}^{p} |f(t_i) - f(t_{i-1})| < M.$$

3 The Markov property

Unless explicitly mentioned, B is a one-dimensional Brownian motion starting from 0.

3.1 Blumenthal 0-1 law and consequences

Filtrations. For every $s \geq 0$ we denote by $\mathcal{F}_s$ the filtration generated by the random variables $\{B_r : 0 \leq r \leq s\}$. We suppose that this filtration is *complete* in the sense that $\mathcal{F}_s$ contains all the P-negligible sets[3]. In words, $\mathcal{F}_s$ contains all the information before time s. We also define

$$\mathcal{F}_{s^+} = \bigcap_{t>s} \mathcal{F}_t,$$

which contains an additional infinitesimal look into the future. Finally, $\mathcal{F}_\infty$ is the σ-field generated by all the $B_t, t \in \mathbb{R}_+$. Let us start to strengthen item 2 of Proposition 3. With the notation introduced there, for every $0 < r_1 < r_2 < \ldots < r_q \leq s$ and every $t_1, \ldots, t_p > 0$

$$\text{the vector } (B^{(s)}_{t_1}, \ldots, B^{(s)}_{t_p}) \text{ is independent of } (B_{r_1}, \ldots, B_{r_q}),$$

[3] this assumption is usually used to restrict oneself to the set of trajectories where BM is continuous

this can be seen from the definition of Brownian motion or using the covariance structure. By the monotone class theorem, it follows that $(B^{(s)}_{t_1}, \dots, B^{(s)}_{t_p})$ is independent of $\sigma(B_u : u \leq s) = \mathcal{F}_s$. Applying once again the monotone class theorem it comes that

$$\text{the Brownian motion } (B^{(s)}_t)_{t\geq 0} \text{ is independent of } \mathcal{F}_s.$$

This reinforcement of Proposition 3 item 2 is called *the simple Markov property*. Very soon, we will extend this property when the fixed time s is replaced by a random one. Before that we prove a zero-one law for "infinitesimal" events.

Theorem 3 (Blumenthal 0-1 law)**.** *The σ-field $\mathcal{F}_{0+}$ is trivial in the sense that every event which is $\mathcal{F}_{0+}$ measurable has probability* 0 *or* 1*.*

Exercise 13. Do you know others 0-1 laws?

Proof. Fix $A \in \mathcal{F}_{0+}$. Let $t_1, \dots, t_p > 0$ and f be a bounded continuous function on $\mathbb{R}^p$. For every $\varepsilon > 0$ the Brownian motion $B^{(\varepsilon)}$ is independent of $\mathcal{F}_\varepsilon$ and so a fortiori of $\mathcal{F}_{0+}$. Thus we have

$$E[1_A f(B^{(\varepsilon)}_{t_1}, \dots, B^{(\varepsilon)}_{t_p})] = P(A)E[f(B^{(\varepsilon)}_{t_1}, \dots, B^{(\varepsilon)}_{t_p})] = P(A)E[f(B_{t_1}, \dots, B_{t_p})].$$

On the other hand, by (almost sure) continuity of the paths of Brownian motion we have

$$E[1_A f(B_{t_1}, \dots, B_{t_p})] = \lim_{\varepsilon\to 0} E[1_A f(B^{(\varepsilon)}_{t_1}, \dots, B^{(\varepsilon)}_{t_p})].$$

Combining the two displays we find that A is independent of $(B_{t_1}, \dots, B_{t_p})$. By the monotone class theorem, A is independent of the σ-field generated by all B_s for $s > 0$ which is $\mathcal{F}_\infty$. The event A is in particular independent of itself, hence $P(A) = 0$ or 1. □

Exercise 14. * Prove that for every $s \geq 0$ we have $\mathcal{F}_s = \mathcal{F}_{s+}$.

Applications. Consider the variables $\overline{B}_\varepsilon = \sup\{B_s : 0 \leq s \leq \varepsilon\}$ and $\underline{B}_\varepsilon = \inf\{B_s : 0 \leq s \leq \varepsilon\}$. A priori it is no clear that these are measurable with respect to $\mathcal{F}_\varepsilon$. To see this, we use the (almost sure) continuity of the trajectories and write

$$\overline{B}_\varepsilon = \sup_{t\in[0,\varepsilon]\cap\mathbb{Q}} B_t.$$

This clearly entails that $\overline{B}_\varepsilon$ is $\mathcal{F}_\varepsilon$-measurable. The same reasoning holds for $\underline{B}_\varepsilon$. Then the (measurable) events $A_+ = \{\forall n > 0 : \overline{B}_{1/n} > 0\}$ and $A_- = \{\forall n > 0 : \underline{B}_{1/n} < 0\}$ belong to $\mathcal{F}_{0+}$. By the previous theorem they have P-probability 0 or 1. But $P(\overline{B}_\varepsilon > 0) \geq P(B_\varepsilon > 0) = 1/2$ which entails that $P(A_+) \geq 1/2$ and so $P(A_+) = 1$ and similarly $P(A_-) = 1$. In words, when starting from 0, Brownian motion immediately takes positive and negative values. The picture to keep in mind is the function $t \mapsto t\sin(t^{-1})$ which exhibits the same property:

Proposition 4. *For every $a \in \mathbb{R}$, let $T_a = \inf\{t \geq 0 : B_t = a\}$ with the convention that* $\inf\varnothing = \infty$*. Then for every $a \in \mathbb{R}$ we have $T_a < \infty$ a.s.*

Proof. By the discussion before the Proposition we have

$$1 = P(\overline{B}_1 > 0) = \lim_{\delta\to 0} P(\overline{B}_1 > \delta).$$

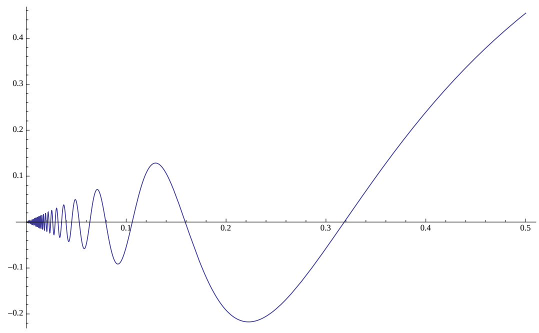

Figure 2: The function $t \mapsto t\sin(t^{-1})$ takes negative and positive values immediately after 0.

A scaling argument (Proposition 3) shows that

$$P(\overline{B}_1 > \delta) = P(\sup_{[0,1]} B > \delta) = P(\sup_{[0,\delta^{-2}]} B > 1).$$

Letting $\delta \to 0$, we deduce that $P(\sup_{s\geq 0} B > 1) = 1$. Another scaling argument shows that $P(\sup_{s\geq 0} B > A) = 1$ for every $A > 0$. By symmetry the proposition holds. $\square$

Exercise 15. Show that in the last proposition we can exchange a.s. and $\forall$ meaning that we have a.s. for every $a \in \mathbb{R}$, $T_a < \infty$. Deduce that $\sup B = +\infty$ and $\inf B = -\infty$ almost surely.

Exercise 16.

1. Show that the event $\limsup_{t\to 0} t^{-1/2} B_t = +\infty$ is $\mathcal{F}_{0+}$-measurable.
2. Deduce that a.s. $\limsup_{t\to 0} t^{-1/2} B_t = +\infty$ and $\liminf_{t\to 0} t^{-1/2} B_t = -\infty$.
3. Using (or admitting) Exercise 10, show that the statements of 2. hold when $t \to \infty$.

Exercise 17. Show that Brownian motion is a.s. non monotone on every interval.

Exercise 18. * Show that for every $t \geq 0$, time t is not a local maximum for B almost surely. Can we exchange a.s. and $\forall$ in the last sentence ?

Exercise 19. Show the following convergence in distribution

$$\left(\int_0^t e^{B_s} ds\right)^{1/\sqrt{t}} \xrightarrow[t\to\infty]{} e^{\overline{B}_1}.$$

Hint : Use scaling.

3.2 The strong Markov property

Recall that with the notation of Proposition 3, for every $s \geq 0$ the Brownian motion $B^{(s)}$ is independent of $\mathcal{F}_s$. The goal of this section is to extend the previous statement when s is replaced by a random time... but not any random time!

Definition 2. *A random variable $T \in \mathbb{R}_+ \cup \{\infty\}$ is a stopping time if for every $t \geq 0$, the event $\{T \leq t\}$ is $\mathcal{F}_t$ measurable. In words, "if you must stop before time t you should know it by time t". The σ-field of the past before T is defined as*

$$\mathcal{F}_T = \{A \in \mathcal{F}_\infty : \forall t \geq 0 : A \cap \{T \leq t\} \in \mathcal{F}_t\}.$$

Example. For every $a \in \mathbb{R}$, T_a is a stopping time. Indeed,

$$\{T_a \le t\} = \left\{ \inf_{r \in \mathbb{Q} \cap [0,t]} |B_r - a| = 0 \right\} \in \mathcal{F}_t.$$

Exercise 20. Let T be stopping time.

- Check that $\mathcal{F}_T$ is a σ-field.
- If S, T are stopping times then $\min(S,T)$, $\max(S,T)$, $S+T$ are stopping times.
- If $S \le T$ are stopping times show that and $\mathcal{F}_S \subset \mathcal{F}_T$.

It is trivial that the random variable T is $\mathcal{F}_T$-measurable. Also $B_T \mathbf{1}_{T<\infty}$ is $\mathcal{F}_T$-measurable. To see this, notice that by the almost sure continuity of Brownian motion we have the almost sure limit

$$B_T = \lim_{n\to\infty} \sum_{k=0}^{\infty} \mathbf{1}_{T \in [k/n,(k+1)/n)} B_{k/n},$$

and remark that each member in the sum is $\mathcal{F}_T$-measurable. We can now state one of the most useful theorem about Brownian motion.

Theorem 4 (strong Markov property)**.** *Let T be a stopping time such that $P(T<\infty) > 0$. For every $t \ge 0$ we put*

$$B_t^{(T)} = \mathbf{1}_{T<\infty}\big(B_{t+T} - B_T\big).$$

Then under $P(.\mid T<\infty)$ the process $B^{(T)}$ is a Brownian motion independent of $\mathcal{F}_T$.

Proof. We first suppose that $T < \infty$ a.s. Let $A \in \mathcal{F}_T$, pick a bounded continuous function $F : \mathbb{R}^p \to \mathbb{R}$ and fix $0 < t_1 < \ldots < t_p$. We thus aim at showing that

$$E[\mathbf{1}_A f(B_{t_1}^{(T)}, \ldots, B_{t_p}^{(T)})] = P(A) E[f(B_{t_1}, \ldots, B_{t_p})].$$

For every integer $n \ge 1$ denote by $[T]_n$ the smallest number of the form $k2^{-n}$ larger than or equal to T. Since we have $f(B_{t_1}^{([T]_n)}, \ldots, B_{t_p}^{([T]_n)}) \to f(B_{t_1}^{(T)}, \ldots, B_{t_p}^{(T)})$ a.s., then by the dominated convergence theorem we have

$$\begin{aligned} E[\mathbf{1}_A f(B_{t_1}^{(T)}, \ldots, B_{t_p}^{(T)})] & \\ &= \lim_{n\to\infty} E[\mathbf{1}_A f(B_{t_1}^{([T]_n)}, \ldots, B_{t_p}^{([T]_n)})] \\ &= \lim_{n\to\infty} \sum_{k=0}^{\infty} E\left[\mathbf{1}_A \mathbf{1}_{(k-1)2^{-n} < T \le k2^{-n}} f(B_{t_1}^{(k2^{-n})}, \ldots, B_{t_p}^{(k2^{-n})})\right] \end{aligned}$$

Now, notice that $A \cap \{(k-1)2^{-n} < T \le k2^{-n}\} \in \mathcal{F}_{k2^{-n}}$. Since by the simple Markov property $B^{(k2^{-n})}$ is independent of $\mathcal{F}_{k2^{-n}}$ we can split the expectation is the sum and get

$$E[\mathbf{1}_A f(B_{t_1}^{(T)}, \ldots, B_{t_p}^{(T)})] = \lim_{n\to\infty} \sum_{k=0}^{\infty} E\left[\mathbf{1}_A \mathbf{1}_{(k-1)2^{-n} < T \le k2^{-n}}\right] E\left[f(B_{t_1}, \ldots, B_{t_p})\right].$$

By re-summing we get the desired result. The case when $P(T=\infty) > 0$ is similar. $\square$

In the following, except mentioned all the stopping times considered are almost surely finite.

Let us give a useful reformulation of the strong Markov property. Suppose that $T < \infty$ a.s. and let X be a random variable which is $\mathcal{F}_T$ measurable: it could for example be a function of T, $B_T \mathbf{1}_{T<\infty}$ of the path before T... and can take value in an arbitrary measurable space $(E, \mathcal{A})$. Then for every positive measurable $\phi : E \times \mathcal{C}(\mathbb{R}_+, \mathbb{R})$ we have

$$E\left[\phi\big(X, B^{(T)}\big)\right] = E\left[\int \mathbb{P}_0(d(W))\phi\big(X, (W_t)_{t\geq 0}\big)\right], \tag{1}$$

where $\mathbb{P}_0$ is the Wiener measure. To see this, just remark that X is independent of $B^{(T)}$ by the above theorem and thus we can write the law $P_{(X,B^{(T)})}$ as a product $P_X \otimes \mathbb{P}_0$. Let us see an application on the canonical space to remind you of the strong Markov property in the case of Markov chains. We put $\Omega = \mathcal{C}(\mathbb{R}_+, \mathbb{R})$ endowed with the product σ-field and $B_t(\omega) = \omega_t$. Recall also the notation $\mathbb{P}_x$. The translation operator is

$$\theta_s(\omega) = (\omega_{s+t})_{t\geq 0}, \qquad \text{for } s \geq 0.$$

Then the Markov property is equivalently described as follows. Let T by an a.s. finite stopping time. For every measurable positive functions F, G on Ω such that F is $\mathcal{F}_T$-measurable then

$$\mathbb{E}_x\left[F \cdot G \circ \theta_T\right] = \mathbb{E}_x\left[F \cdot G(B^{(T)}_{\cdot} + B_T)\right] = \mathbb{E}_x\left[F \int \mathbb{P}_0(dW)G(W + B_T)\right],$$

by (1). But by definition of the measures $\mathbb{P}_x$, we have

$$\mathbb{E}_x\left[F \cdot G \circ \theta_T\right] = \mathbb{E}_x\left[F \int \mathbb{P}_{B_T}(dW)G(W)\right] = \mathbb{E}_x[F\ \mathbb{E}_{B_T}[G]].$$

Exercise 21. Show that for every $a > 0$, almost surely

$$\inf\{t \geq 0 \,|\, B_t = a\} \ = \ \inf\{t \geq 0 \,|\, B_t > a\}\,.$$

Exercise 22. Let $a \geq 0$ and recall that $T_a = \inf\{s \geq 0 : B_s = a\}$.

1. Show that $T_a = a^2 T_1$ in distribution.
2. For $0 \leq a \leq b < \infty$, prove that $T_b - T_a$ is distributed as T_{b-a} and is independent of $(\mathcal{F}_{T_s})_{0\leq s\leq a}$.

Remark: the process $a \mapsto T_a$ is thus another process with stationary and independent increments.

Exercise 23. Show that if S, T are stopping times then ST is not necessarily a stopping time.

4 Paths properties

4.1 Reflection principle

Theorem 5. *Let B be a Brownian motion in dimension 1. For every $t > 0$, recall that $\overline{B}_t = \sup_{s\leq t} B_s$. For every $a \geq 0$ and $b \leq a$ we have*

$$P(\overline{B}_t \geq a,\ B_t \leq b) = P(B_t \geq 2a - b).$$

Proof. The idea is to apply the strong Markov property at the stopping time T_a and to reflect the path after time a around the horizontal line $y = a$. We have already seen that $T_a < \infty$ a.s. Then we can apply the strong Markov property (1) and get that

$$\begin{aligned} P(\overline{B}_t \geq a,\ B_t \leq b) &= P(T_a \leq t,\ B^{(T_a)}_{t-T_a} \leq b - a) \\ &= E\left[\mathbf{1}_{T_a \leq t} F_{t-T_a}([-\infty, b-a])\right], \end{aligned}$$

where for $s \geq 0$ and $A \subset \mathbb{R}$

$$F_s(A) = \int \mathbb{P}(d\omega)\mathbf{1}_{\omega_s \in A}.$$

By symmetry of the Wiener measure we have $F_s(A) = F_s(-A)$ so that we have

$$\begin{aligned} P(\overline{B}_t \geq a,\ B_t \leq b) &= E\left[\mathbf{1}_{T_a \leq t} F_{t-T_a}([a-b, \infty])\right] \\ &= P(T_a \leq t,\ B^{(T_a)}_{t-T_a} \geq a - b) \\ &= P(\overline{B}_t \geq a,\ B_t \geq 2a - b) \\ &= P(B_t \geq 2a - b). \end{aligned}$$

□

Corollary 1. *With the notation of the previous theorem we have*

1. *The variable $\overline{B}_t$ has the same distribution as $|B_t|$.*
2. *The variable T_a is distributed as $(\frac{a}{B_1})^2$ for every $a \in \mathbb{R}$.*

Proof. For (i) we have

$$\begin{aligned} P(\overline{B}_t \geq a) &= P(\overline{B}_t \geq a,\ B_t \leq a) + P(\overline{B}_t \geq a,\ B_t > a) \\ &= P(B_t \geq 2a - a) + P(B_t > a) \\ &= 2P(B_t \geq a). \end{aligned}$$

For (ii), if $a \geq 0$ using (i) and the scaling property of Brownian motion we have

$$P(T_a \leq t) = P(\overline{B}_t \geq a) = P(|B_t| \geq a) = P(|\sqrt{t}B_1| \geq a) = P(\frac{a^2}{B_1^2} \leq t).$$

When $a \leq 0$, we use the symmetry of Brownian motion to say that $T_a = T_{-a}$ in law. □

Exercise 24 (Densities)**.** Deduce from the last corollary that

1. The pair $(\overline{B}_t, B_t)$ has density
$$\frac{2(2a-b)}{\sqrt{2\pi t^3}} \exp\left(-\frac{(2a-b)^2}{2t}\right)\mathbf{1}_{a>0, b<a}.$$
2. For every $a \in \mathbb{R}$, the variable T_a has density
$$\frac{a}{\sqrt{2\pi t^3}} \exp\left(-\frac{a^2}{2t}\right)\mathbf{1}_{t>0}.$$

Exercise 25. Try Exercise 10 again.

Exercise 26 (local maxima)**.** * Let p, q, r, $s \in \mathbb{Q}_+$ such that $p < q < r < s$. Show that

$$P\left(\sup_{p \leq t \leq q} B_t = \sup_{r \leq t \leq s} B_t\right) = 0.$$

Deduce that local extrema of Brownian motion are almost surely distinct.

4.2 Zeros of Brownian motion

In this section, we study the zero set of a standard one-dimensional Brownian motion that is

$$\mathcal{Z} = \{t \geq 0 : B_t = 0\}.$$

Theorem 6. *Almost surely, the set $\mathcal{Z}$ is a perfect set (that is closed with no isolated points) of zero Lebesgue measure.*

Proof. Since B is an a.s. continuous function the set $\mathcal{Z}$ is closed a.s. It follows from the discussion after the Blumenthal law that $T_0^+ = \inf\{t>0 : B_t = 0\}$ satisfies $T_0^+ = 0$ almost surely. Using this, we will prove that $\mathcal{Z}$ has no isolated point. Fix $q \in [0, \infty)$ a rational number and consider the first return to zero after q:

$$\tau_q = \inf\{t \geq q : B_t = 0\}.$$

It is easy to see that τ_q is an almost surely finite stopping time. Note also that $B_{\tau_q} = 0$ by continuity. We can thus apply the strong Markov property at time τ_q and get

$$P(\inf\{z \in \mathcal{Z} : z > q\} = q) = P(\inf\{t > 0 : B^{(\tau_q)}\} = 0) = 1.$$

We deduce that $\forall q \in \mathbb{Q}_+$, almost surely, the zero τ_q is not isolated from the right. Since the rationals numbers are countable, we can exchange $\forall$ and *a.s.* in the last sentence. This easy implies that all the zero of $\mathcal{Z}$ are no isolated, indeed if we argue by contradiction and find $z \in \mathcal{Z}$ alone in a neighborhood then $z = \tau_q$ for some $q \in \mathbb{Q}_+$ and we reach a contradiction.

To show that the Lebesgue measure of $\mathcal{Z}$ is zero, we apply Fubini's theorem

$$E[\mathsf{Leb}(\mathcal{Z})] = E\left[\int_0^\infty dt\ \mathbf{1}_{B_t=0}\right] = \int_0^\infty dt\ P(B_t = 0) = 0,$$

which proves the desired result... However to justify the use of Fubini's theorem we have to show that the function $(\omega, t) \mapsto \mathbf{1}_{B_t=0}$ is measurable with respect to $\mathcal{F}_\infty \otimes \mathcal{B}$. This is left as an measurability exercise. □

Exercise 27. Show that the function $(\omega, t) \mapsto \mathbf{1}_{B_t=0}$ is measurable with respect to $\mathcal{F}_\infty \otimes \mathcal{B}$.

Exercise 28. Show that for every $s \geq 0$, almost surely, the set $\mathcal{L}_s = \{t \geq 0 : B_t = s\}$ is also a perfect set of zero Lebesgue measure. Can we exchange $\forall$ and *a.s.* ?

Exercise 29. Show that Brownian motion has isolated zeros from the left and from the right.

Exercise 30 (Classic exercise about perfect sets)**.** Show that a perfect set of $\mathbb{R}$ has the same cardinal as $\mathbb{R}$. Do you know of other perfect sets of $\mathbb{R}$ of Lebesgue measure 0? (If you do not, then learn the classical construction of Cantor sets).

4.3 Arcsine laws

A random variable $X \in [0, 1]$ follows the arcsine law if

$$P(X \in dx) = \frac{dx}{\pi\sqrt{x(1-x)}}\mathbf{1}_{x\in(0,1)}.$$

The name comes from the fact that the cumulative distribution function of X is then $P(X \leq t) = \frac{2}{\pi}\arcsin(\sqrt{x})$. The arcsine law pops-up naturally when studying basic properties of Brownian motion.

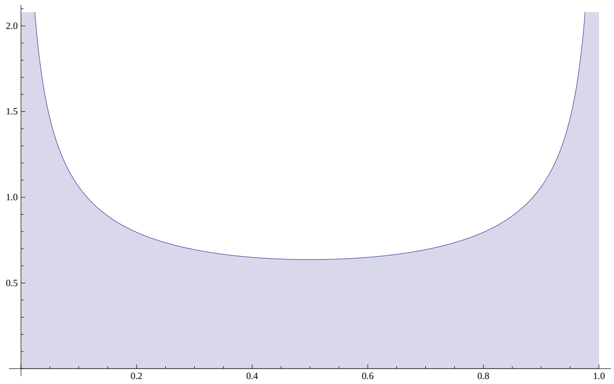

Figure 3: The arcsine distribution

Theorem 7 (Arcsine laws)**.** *Let $L = \sup\{t \in [0,1] : B_t = 0\}$ be the last zero of Brownian motion before time $t = 1$. Let also M be the almost surely unique random time in $[0,1]$ such that $B_M = \overline{B}_1$. Then L and M are both arcsine distributed.*

Exercise 26 shows that M is almost surely unique and thus well-defined.

We will see later that there is yet another arcsine law for Brownian motion: the time spent by B in positive (or negative) regions before time 1 is again distributed according to the arcsine distribution.

Exercise 31. Show that L and M are *not* stopping times.

Proof. We start with the case of L. Fix $t \in (0,1)$. By applying the (simple) Markov property at (the stopping) time t we deduce that

$$\begin{aligned} P(L < t) &= E\left[\mathbf{1}\{B_s^{(t)} + B_t \neq 0,\ \forall s \in [0, 1-t]\}\right] \\ &= E\left[\mathbb{E}_{B_t}[\omega_s > 0,\ \forall s \in [0, 1-t]\}]\right] \\ &= E\left[\mathbb{P}_{B_t}(T_0 > 1-t)\right] \\ &= E\left[\mathbb{P}_0(T_{B_t} > 1-t)\right]. \end{aligned}$$

We can now use Corollary 1 and get

$$\begin{aligned} P(L < t) &= E[P((B_t/N')^2 > 1-t)] \\ &= P((N/N')^2 > \frac{1-t}{t}), \end{aligned}$$

where N, N' are independent standard normal $\mathcal{N}(0,1)$ independent of B. At this point, either we compute directly the integral or we use a trick: After a simple manipulation the last probability is also equal to

$$P\left(\frac{|N'|}{\sqrt{N^2 + N'^2}} \leq \sqrt{t}\right).$$

We write $(N, N') = (R\sin(\theta), R\cos(\theta))$ in polar coordinates, where $R \geq 0$ and $\theta \in (0, 2\pi]$. The density of (N, N') is $dxdy\, \frac{1}{2\pi} e^{-(x^2+y^2)/2}$ which becomes $d\theta\, \frac{1}{2\pi}\, dr\, re^{-r^2/2}$ in polar. The point is that θ is uniformly distributed over $(0, 2\pi]$. It follows that the last probability is equal to $P(|\sin(\theta)| \leq \sqrt{t}) = \frac{2}{\pi}\arcsin(\sqrt{t})$ as desired.

The case of the variable M is similar in spirit. Fix $t \in (0,1)$ and write

$$\begin{aligned} P(M \le t) &= P\left(\sup_{[0,t]} B \ge \sup_{[t,1]} B\right) \\ &= P\left(\sup_{u\in[0,t]} B_{t-u} - B_t \ge \sup_{u\in[0,1-t]} B_u^{(t)}\right). \end{aligned}$$

By Proposition 3, the process $\tilde{B}_u = B_{t-u} - B_t$ is a Brownian motion over $[0,t]$. Hence, applying the (simple) Markov property at time t we get that

$$\begin{aligned} P(M \le t) &= P(\overline{\tilde{B}}_t \ge \overline{B^{(t)}}_{1-t}) \\ &= P(\sqrt{t}|N| \ge \sqrt{1-t}|N'|) \end{aligned}$$

by Theorem 5 and with the same notation as above. The rest of the proof is the same. $\square$

Exercise 32 (Another proof using time inversion)**.** Let $R = \inf\{t \ge 1 : B_t = 0\}$.

1. Using the (simple) Markov property at time 1, show that $R = 1 + (N/N')^2$ in distribution where N and N' are two independent standard normal.
2. Use Exercise 10 to show that $R = L^{-1}$ in distribution and recover the arcsine law for L.

4.4 Law of iterated logarithm

Theorem 8. *Let B be a standard one-dimensional Brownian motion. Then we have*

$$\limsup_{t\to\infty} \frac{B_t}{\sqrt{2t \log\log(t)}} = 1.$$

Proof. The main idea of the proof is to look at values of Brownian motion at exponential times α^n for some $\alpha > 1$. Indeed, when $\alpha \to \infty$ these values becomes more and more independent and we can use the Borel–Cantelli lemmas. Let us even imagine in the proof that we do not know the scale function $\sqrt{2t \log\log t}$ and let us try to recover it. Notice that it should be large compared to $\sqrt{t}$ since the latter is the typical behavior of Brownian motion at time t.

Upper bound. Fix $\alpha > 1$ and look at the events

$$A_n = \{\overline{B}_{\alpha^n} \ge \sqrt{\alpha^n}\psi(\alpha^n)\}.$$

The goal here is to apply the first Borel–Cantelli lemma and to see that for a suitable function ψ we have $\sum P(A_n) < \infty$ which will imply the upper bound. By the scaling property of Brownian motion and Theorem 5 we have

$$P(A_n) = P(\overline{B}_1 \ge \psi(\alpha^n)) = P(|B_1| \ge \psi(\alpha^n)).$$

As we said, it should be clear that $\psi(t) \to \infty$ as $t \to \infty$. Using Lemma 1 the last probability is thus of order

$$\frac{1}{\psi(\alpha^n)} \exp(-\psi(\alpha^n)^2/2).$$

Since $\psi(\alpha^n) \to \infty$, the important factor in the last display is the exponential one. We thus want to see if the series $\sum P(A_n)$ converges and thus compare $\exp(-\psi(\alpha^n)^2/2)$ with $1/n$. Doing so, we clearly see that the threshold function is $\psi(t) = \sqrt{2\log\log t}$. Indeed for every $\varepsilon > 0$ if $\psi(t) < (1-\varepsilon)\sqrt{2\log\log t}$ the series $\sum P(A_n)$ diverges and if $\psi(t) = (1+\varepsilon)\sqrt{2\log\log t}$ then $\sum P(A_n)$ converges. In this case, by the first Borel–Cantelli lemma, A_n happens finitely many often. For every $t > 1$, we now interpolate and find $n \geq 1$ such that $\alpha^{n-1} < t \leq \alpha^n$ and so

$$\frac{B_t}{\sqrt{2t\log\log t}} \leq \frac{\overline{B}_{\alpha^n}}{(1+\varepsilon)\sqrt{2\alpha^n\log\log\alpha^n}} \frac{(1+\varepsilon)\sqrt{2\alpha^n\log\log\alpha^n}}{\sqrt{2t\log\log t}}$$

Taking $\limsup$ we get that

$$\limsup_{t\to\infty} \frac{B_t}{\sqrt{2t\log\log t}} \leq (1+\varepsilon)\sqrt{\alpha}.$$

Now let $\varepsilon \ \to 0$ and $\alpha \to 1$ to get the upper bound. We can generalize slightly the upper bound to get that (this will be used in the lower bound)

$$\limsup_{t\to\infty} \frac{|B_t|}{\sqrt{2t\log\log t}} = 1. \tag{2}$$

Lower bound. For the lower bound we will apply the second Borel–Cantelli lemma and thus need independent events. Fix $\alpha > 1$ and $\varepsilon > 0$ and consider the events

$$A'_n = \{B_{\alpha^n} - B_{\alpha^{n-1}} > (1-\varepsilon)\sqrt{2\alpha^n\log\log\alpha^n}\}, \qquad \text{for } n \geq 1$$

Remark that $B_{\alpha^n} - B_{\alpha^{n-1}} = B^{(\alpha^{n-1})}_{\alpha^n-\alpha^{n-1}}$ so that the Markov property applied at (the stopping time) α^{n-1} yields that A'_n is independent of $A'_1, A'_2, \dots, A'_{n-1}$. It follows that all the A'_n are independent. On the other hand by Lemma 1 we have

$$\begin{aligned} P(A'_n) &= P(B_{\alpha^n-\alpha^{n-1}} > (1-\varepsilon)\sqrt{2\alpha^n\log\log\alpha^n}) \\ &= P\left(B_1 > \frac{(1-\varepsilon)\sqrt{2\alpha^n\log\log\alpha^n}}{\sqrt{\alpha^n-\alpha^{n-1}}}\right) \\ &\geq P\left(B_1 > \frac{\sqrt{2\alpha}(1-\varepsilon)}{\sqrt{\alpha-1}}\sqrt{\log\log\alpha^n}\right). \end{aligned}$$

Using Lemma 1 the last quantity is of order

$$n^{-\gamma^2/2+o(1)}, \qquad \text{where } \gamma = \frac{\sqrt{2\alpha}(1-\varepsilon)}{\sqrt{\alpha-1}}.$$

For ε fixed we thus pick α large enough so that $\gamma^2/2 < 1$. If we do so, the series $\sum P(A'_n)$ is infinite and by the second Borel–Cantelli lemma, A'_n happens infinitely often. When A'_n happens we have

$$\frac{B_{\alpha^n}}{\sqrt{2\alpha^n\log\log\alpha^n}} \geq (1-\varepsilon) - \frac{|B_{\alpha^{n-1}}|}{\sqrt{2\alpha^n\log\log\alpha^n}}.$$

We now use the upper bound (2) for the second term of the right-hand side to deduce that the limsup of the right-hand side is at least $(1-\varepsilon) - \frac{1}{\sqrt{\alpha}}$. We can now let $\alpha \to \infty$ followed by $\varepsilon \to 0$ and get the desired result. □

Remark. By symmetry it follows that $\liminf B_t/\sqrt{2t\log\log(t)} = -1$ and by time inversion (see Exercise 10) we also have

$$\limsup_{t\to 0} \frac{B_t}{\sqrt{2t\log|\log t|}} = 1.$$

In particular, a.s. Brownian motion is not differentiable at time $t = 0$. Applying the simple Markov property it follows that, for every $t \geq 0$, almost surely B is not differentiable at time t. The purpose of the following theorem is to exchange $\forall$ and *a.s.* and show that Brownian motion is a.s. nowhere differentiable.

Theorem 9 (Paley, Wiener and Zygmung 1933)**.** *Almost surely, the function $t \mapsto B_t$ is nowhere differentiable.*

Proof. We can restrict ourself to $[0,1]$. If there exists $t_0 \in [0,1)$ such that B is differentiable at t_0 that implies that

$$\sup_{h\in[0,1]} \frac{|B_{t_0+h} - B_{t_0}|}{h} \leq M, \tag{3}$$

for some (random) $M > 0$. We split the interval $[0,1]$ into the 2^n intervals $[(k-1)2^{-n}, k2^{-n}]$ for $k \in \{1,\dots,2^n\}$. If t_0 falls into some interval $[(k-1)2^{-n}, k2^{-n}]$ that means that for every $1 \leq j \leq 2^n - k$ we have using (3)

$$\begin{aligned}|B_{(k+j)2^{-n}} - B_{(k+j-1)2^{-n}}| &\leq |B_{(k+j)2^{-n}} - B_{t_0}| + |B_{(k+j-1)2^{-n}} - B_{t_0}| \\ &\leq M\frac{2j+1}{2^n}.\end{aligned}$$

In fact, we will only use the first values of $j = 1,2,3$ for our purposes. We are thus led to consider the events

$$\Omega_{n,k}^M := \{|B_{(k+j)2^{-n}} - B_{(k+j-1)2^{-n}}| \leq 7M2^{-n}, \text{ for } j = 1,2,3\}.$$

By the independence of the increments and the scaling property of Brownian motion we have

$$\begin{aligned}P(\Omega_{n,k}^M) &= P(|B_{2^{-n}}| \leq 7M2^{-n})^3 \\ &= P(|B_1| \leq 7M2^{-n/2})^3 \\ &\leq (7M)^3 2^{-3n/2},\end{aligned}$$

where for the last line we used the fact that the density of the normal distribution is bounded by 1. By the union bound we have

$$P\left(\bigcup_{k=1}^{2^n} \Omega_{n,k}^M\right) \leq (7M)^3 2^{-n/2}.$$

Now, if B is differentiable at some $t_0 \in [0,1)$ by (3) it implies that for some M, the event $\cup_k \Omega_{n,k}^M$ holds for n large enough and so for infinitely many n's. But by the first Borel–Cantelli lemma we deduce that

$$\begin{aligned}
&P(\exists t_0 \in [0,1] : B \text{ differentiable at } t_0) \\
&\leq P\left(\bigcup_{M=1}^{\infty} \left\{ \text{the event } \bigcup_{k=1}^{2^n} \Omega_{n,k}^M \text{ holds i.o.} \right\} \right) \\
&= \sum_{M=1}^{\infty} 0 = 0.
\end{aligned}$$

□

Exercise 33. * Let $f \in \mathcal{C}([0,1], \mathbb{R})$ be any fixed continuous function. Adapt the proof of Theorem 9 and show that, almost surely, the function $(B_t + f(t) : t \in [0,1])$ is nowhere differentiable.

We end this section by mentioning Lévy's modulus of continuity theorem (without giving the proof) which shares striking similarities with the law of the iterated logarithm (except the iterated logarithm!).

Theorem 10 (Lévy 1937). *Almost surely,*

$$\limsup_{h \to 0} \sup_{0 \leq t \leq 1-h} \frac{|B_{t+h} - B_t|}{\sqrt{2h \log(1/h)}} = 1.$$

Exercise 34. * Prove the lower bound in Lévy's modulus of continuity theorem.

5 Brownian motion and martingales

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Suppose that we have a filtration $(\mathcal{F}_t)_{t \geq 0}$. For the purposes of this course we will suppose that $\mathcal{F}_t$ is the filtration generated by the Brownian motion B. We will need to consider martingales in continuous time which are defined similarly as in the discrete setting.

Definition 3. *A family $(M_t)_{t \geq 0}$ is an $\mathcal{F}_t$-martingale if for every $t \geq 0$, M_t is $\mathcal{F}_t$-measurable (the process is said to be adapted to the filtration), $E[|M_t|] < \infty$ and if for every $s \leq t$ we have*

$$E[M_t \mid \mathcal{F}_s] = M_s.$$

The process is a submartingale or supermartingale if the $=$ of the last display is replaced by $\geq$ respectively $\leq$.

As in the discrete setting, martingales represent fair games where the current value of the process gives the best prediction for the future of the game. Let us give a few examples of very important continuous martingales based on Brownian motion.

Proposition 5. *The processes $(B_t : t \geq 0)$, $(B_t^2 - t : t \geq 0)$ and $(e^{\lambda B_t - t\lambda^2/2} : t \geq 0)$ for all $\lambda \in \mathbb{R}$ are all continuous martingales for the Brownian filtration.*

Proof. All the processes considered are integrable, adapted to the filtration $(\mathcal{F}_t)_{t \geq 0}$ and continuous. Let us check the martingale property, for $t \geq s$ we have

$$E[B_t \mid \mathcal{F}_s] = E[B_s + (B_t - B_s) \mid \mathcal{F}_s] = E[B_s],$$

because $B_t - B_s$ is independent of $\mathcal{F}_s$ by the (simple) Markov property. The other cases are similar (for the exponential martingale, recall that if N is a standard Gaussian r.v. then $E[\exp(zN)] = \exp(z^2/2)$ for every $z \in \mathbb{C}$). □

Exercise 35. Find $\alpha \in \mathbb{R}$ such that $B_t^3 - \alpha t B_t$ is an $(\mathcal{F}_t)$-martingale.

Exercise 36. Show that the process

$$t \mapsto tB_t - \int_0^t ds\ B_s,$$

is an $(\mathcal{F}_t)$-martingale.

We will now extend some of the results of the theory of discrete-time martingales to our setup. The proofs always go through an approximation by discrete martingales after discretizing the time and then passing to the limit.

Proposition 6. *Let T be a bounded stopping time and let $(M_t)_{t\geq 0}$ be a continuous martingale adapted to $\mathcal{F}_t$. Then we have*

$$E[M_T] = E[M_0].$$

Proof. Let $K > 0$ such that $T < K$ a.s. For $n \geq 0$, we divide the time-scale $[0, K]$ into n steps by putting $\mathcal{F}_i^{(n)} = \mathcal{F}_{iK/n}$, defining $T^{(n)}$ to be the smallest i such that iK/n larger than T and setting $M_i^{(n)} = M_{iK/n}$. It is then clear that $M^{(n)}$ is an $(\mathcal{F}_i^{(n)})_{i\geq 0}$-martingale and $T^{(n)}$ is a (discrete) stopping time for $(\mathcal{F}_i^{(n)})_{i\geq 0}$. By continuity of M we clearly have

$$M^{(n)}_{T^{(n)}} \xrightarrow[n\to\infty]{a.s.} M_T.$$

For every $n \geq 0$, by the discrete theory we have

$$E[M^{(n)}_{T^{(n)}}] = E[M_0],$$

besides

$$M^{(n)}_{T^{(n)}} = E[M_K \mid \mathcal{F}^{(n)}_{T^{(n)}}].$$

Since the family $\{E[M_K \mid \mathcal{G}] : \mathcal{G} \subset \mathcal{F} \text{ subfiltration}\}$ is uniformly integrable it follows that $M^{(n)}_{T^{(n)}}$ is also uniformly integrable. Consequently, by the enhanced version of dominated convergence theorem we have

$$E[M^{(n)}_{T^{(n)}}] \to E[M_T] = E[M_0] \quad \text{as } n \to \infty. \qquad \square$$

Theorem 11. *For every $a < 0 < b$ we have*

1. $P(T_a < T_b) = \dfrac{b}{b-a}$,
2. $E[\min(T_a, T_b)] = -ab$,

Proof. Fix $n \geq 1$. The stopping time $\tau_n = \min(n, T_a, T_b)$ is bounded by n. We use the continuous martingale B_t and we get by the last proposition that

$$E[B_{\tau_n}] = 0.$$

Since $T_a, T_b < \infty$, we can let $n \to \infty$ to get $B_{\tau_n} \to B_{\min(T_a,T_b)}$ a.s. Since furthermore B_{τ_n} is bounded by $\max(|a|, |b|)$ the dominated convergence theorem shows that

$$0 = E[B_{\min(T_a,T_b)}] = aP(T_a < T_b) + bP(T_b < T_a).$$

Combining this with the obvious fact $P(T_a < T_b) + P(T_b < T_a) = 1$ yields the first point of the theorem.
The second point is similar but we now work with the continuous martingale $(B_t^2 - t)_{t\geq 0}$. Using the same stopping time $\tau_n = \min(n, T_a, T_b)$ we get by the same proposition that

$$0 = E[B_{\tau_n}^2 - \tau_n] = E[B_{\tau_n}^2] - E[\tau_n].$$

On the one hand, $E[B_{\tau_n}^2] \to E[B_{\min(T_a,T_b)}^2]$ by dominated convergence. On the other hand $E[\tau_n] \uparrow E[\min(T_a, T_b)]$ by monotone convergence. We thus obtain using the first point of the theorem that

$$E[\min(T_a, T_b)] = E[B_{\min(T_a,T_b)}^2] = a^2 \frac{b}{b-a} + b^2 \frac{-a}{b-a} = -ab. \quad \square$$

Thanks to the last theorem, the expectation of the stopping time $\min(T_a, T_b)$ for $a < 0 < b$ is thus seen to be finite. Notice that it is very important to constrain the Brownian motion in a finite strip to get a finite expectation for the exit time. Indeed for every $a \in \mathbb{R}^*$ it is easy to see using the exact density of T_a (see Exercise 24) that

$$E[T_a] = \infty.$$

Exercise 37. Use Theorem 11 and let $b \to \infty$ to get another proof of the identity $E[T_a] = \infty$.

Exercise 38. Using the martingale $\exp(\lambda B_t - t\lambda^2/2)$, show that for every $\lambda \in \mathbb{R}_+$ we have

$$E\left[\exp(-\lambda \min(T_a, T_{-a}))\right] = \cosh(a\sqrt{2\lambda})^{-1}.$$

Exercise 39. Show that if $|B_{t+1} - B_t| > 2$ for some $t \geq 0$ then we have $\min(T_1, T_{-1}) \leq t + 1$. Deduce that $P(\min(T_1, T_{-1}) \geq n) \leq \alpha^n$ for some $\alpha \in (0, 1)$.

Exercise 40 (Doob's maximal inequality for submartingales)**.** Let $(X_t)_{t\geq 0}$ be a continuous submartingale for the filtration $(\mathcal{F}_t)_{t\geq 0}$. Let $p > 1$ and $t \geq 0$, show using the discrete version of Doob's maximal inequality that

$$E\left[\left(\sup_{s\leq t} |X_t|\right)^p\right] \leq \left(\frac{p}{p-1}\right)^p E\left[|X_t|^p\right].$$

Cultural remark. A fascinating theorem, due to Dubins and Schwarz (not the same guy as the Cauchy-Schwarz inequality...) roughly says that all the *continuous* martingales are variations of Brownian motion (more precisely time-changed Brownian motion). Hence, in a certain sense there is only "one" continuous martingale which is the Brownian motion. Of course if we remove the statement "continuous" that is another story...

6 Donsker's theorem

Reminder on the convergence in distribution. Let X_n and X be random variables with values in a metric space (E, d) endowed with its Borel σ-field. We say that $X_n \to X$ in distribution (or in law) if for every bounded continuous function $F : E \to \mathbb{R}$ we have

$$E[F(X_n)] \longrightarrow E[F(X)], \qquad \text{as } n \to \infty.$$

The following proposition gathers most of the properties which are commonly used:

Proposition 7 (Portmanteau)**.** *Let X_n and X be random variables with values in an arbitrary metric space (E, d). The the following propositions are equivalent:*

- *X_n converges in law towards X*
- *$E[F(X_n)] \to E[F(X)]$ for any bounded continuous function F*
- *$E[F(X_n)] \to E[F(X)]$ for any bounded Lipschitz function F*
- *$\limsup P(X_n \in C) \leq P(X \in C)$ for any $C \subset E$ closed*
- *$\liminf P(X_n \in O) \geq P(X \in O)$ for any $O \subset E$ open*
- *$P(X_n \in A) \to P(X \in A)$ for any $A \subset E$ measurable such that $P(X \in \partial A) = 0$*
- *$E[F(X_n)] \to E[F(X)]$ for any bounded measurable function $F : E \to \mathbb{R}$ with*
$$P(F \text{ is discontinuous at } X) = 0.$$

Proof. Can be found in any textbook on advanced probability. For example "Convergence of probability measures" by Billingsley. □

Proposition 8 (Continuous mapping theorem)**.** *Suppose that X_n converges in distribution towards X and let F be a continuous on (E, d). Then $F(X_n)$ converges in law towards $F(X)$. This remains true if F is only measurable but such that*

$$P(F \text{ is discontinuous at } X) = 0.$$

Proof. Suppose that is measurable and that $P(F \text{ is discontinuous at } X) = 0$. Let G be a bounded continuous function (defined on the image space of F) with values in $\mathbb{R}$. Obviously, $G \circ F$ is a bounded measurable function such that $P(G \circ F \text{ is discontinuous at } X) = 0$. We can thus apply the last item of Proposition 7 to deduce that

$$E[G(F(X_n))] \longrightarrow E[G(F(X))],$$

as $n \to \infty$ which entails that $F(X_n) \to F(X)$ in law as desired. □

Exercise 41. * Let X_n and X be random variables with values in a metric space (E, d).

1. Suppose that $X_n \to X$ almost surely. Show that $X_n \to X$ in distribution.
2. Suppose now that $X_n \to X$ in probability, i.e. $\forall \varepsilon > 0$ we have $P(d(X_n, X) > \varepsilon) \to 0$.
   (a) Show that there exists a subsequence n_k such that $X_{n_k} \to X$ almost surely.
   (b) Deduce that $X_n \to X$ in distribution.

We will focus in this chapter on the case

$$(E, d) = \Big(\mathcal{C}\big([0, K], \mathbb{R}\big), \|\|_\infty\Big),$$

the space of continuous functions over $[0, K]$ for $K > 0$ endowed with the topology of uniform convergence. Although not useful in the sequel, recall that this space is separable and complete.

The random variables that we will consider are rescaled random walks and their limit is always Brownian motion.

Let $X_1, \ldots, X_n, \ldots$ be independent identically distributed real random variables with mean $c \in \mathbb{R}$ and finite variance σ^2. We form the cumulative sum $S_0 = 0$ and $S_n = X_1 + \ldots + X_n$ for $n \geq 1$ which we interpolate linearly between integer values by setting

$$S(t) = S_{[t]} + (t - [t])(S_{[t]+1} - S_{[t]})$$

for every $t \geq 0$. We now define a sequence (S_n^*) of random continuous functions by putting

$$S_n^*(t) = \frac{S(nt) - cnt}{\sigma\sqrt{n}}.$$

We already know from the central limit theorem that $S_n^*(1) \to B_1$ in distribution and more generally that for every $0 \leq t_1 \leq t_2 \leq \ldots \leq t_k$, we have

$$(S_n^*(t_1), \ldots, S_n^*(t_k)) \xrightarrow[n\to\infty]{(d)} (B_{t_1}, \ldots, B_{t_k})$$

in distribution as $n \to \infty$ (if you do not see why, prove it!). In words, the finite dimensional marginals of S_n^* converge towards that of Brownian motion. The following theorem (sometimes called the functional central limit theorem) shows that the convergence actually holds in distribution for the uniform norm over $[0, K]$ for any $K > 0$. This is much stronger (as a comparison, the uniform convergence over every compact in much stronger that the point wise convergence of functions) and enables to control the function entirely over $[0, K]$. For example, this theorem will imply the convergence in distribution of all $F(S_n^*)$ towards $F(B)$ where F is a continuous function for the uniform norm over $[0, 1]$, see Section 6.1 for applications.

Theorem 12 (Donsker's invariance principle)**.** *For every $K \geq 0$ the random continuous functions S_n^* converge in distribution for the supremum norm over $[0, K]$ towards B.*

One of the advantage of Donsker's theorem is its universality in the sense that Brownian motion appears as the limit of any rescaled random walk as soon as the step distribution has a finite variance. However, for the sake of clarity (and for timing reason) we will prove the theorem in the case when X_i is a sequence of iid fair Bernoulli variables, that is

$$P(X_1 = 1) = P(X_1 = -1) = \frac{1}{2}.$$

There are many proofs of the Donsker's theorem. The one we undertake is based on the idea of coupling: we will construct, on the same probability space the random walk S_n as well as a Brownian motion B so that they are very close to each other. To do so, we will embed the walk into the Brownian motion B.

Proof. Consider the sequence of random times defined recursively as follows: $\tau_0 = 0$ and for $i \geq 0$

$$\tau_{i+1} = \inf\{t \geq \tau_i : |B_t - B_{\tau_i}| = 1\}.$$

This forms a sequence of increasing stopping times (exercise) such that we have $B_{\tau_{i+1}} = B_{\tau_i} \pm 1$. By the strong Markov property applied at time τ_i, we deduce that $(\tau_{i+1} - \tau_i, B_{\tau_{i+1}} - B_{\tau_i})$ is independent of $\mathcal{F}_{\tau_i}$ and is distributed as (τ_1, B_{τ_1}). Since the $(\tau_{j+1} - \tau_j, B_{\tau_{j+1}} - B_{\tau_j})$ for $j < i$ are $\mathcal{F}_{\tau_i}$-measurable, it follows that

$$(\tau_{i+1}-\tau_i, B_{\tau_{i+1}}-B_{\tau_i})_{i\geq 0} \text{ are iid and distributed as } (\tau_1, B_{\tau_1}).$$

By Theorem 11, $E[\tau_1] = 1$ and B_{τ_1} is a fair Bernoulli variable. Hence, the discrete process $(S_n)_{n\geq 0} = (B_{\tau_n})_{n\geq 0}$ is a simple symmetric random walk on $\mathbb{Z}$. This is the coupling we were looking for. We set $B_n^*(t) = n^{-1/2}B_{nt}$ and we will show that for every $K > 0$ we have

$$\sup_{t\in[0,K]} |B_n^*(t) - S_n^*(t)| \to 0, \quad \text{in probability as } n \to \infty. \tag{4}$$

This will imply the theorem. Indeed, let $F : \mathcal{C}([0,K],\mathbb{R}) \to \mathbb{R}$ be a bounded continuous function for the uniform norm over $[0,K]$. Since for every $n \geq 0$, B_n^* is distributed as a standard Brownian motion we have

$$\big|E[F(S_n^*)] - E[F(B)]\big| = \big|E[F(S_n^*) - E[F(B_n^*)]\big| \leq E\Big[\big|F(S_n^*) - F(B_n^*)\big|\Big].$$

We then proceed as in the proof of the fact that convergence in probability implies convergence in distribution (Exercise 41). Namely, we argue by contradiction and suppose that we can find infinitely many n's so that the right-hand side of the last display is larger than some $\varepsilon > 0$. Then by (4), we can extract from this sequence a subsequence n_k such that $S_{n_k}^* - B_{n_k}^* \to 0$ almost surely for the uniform norm over $[0,K]$. We can then use the dominated convergence theorem to get that

$$E\Big[\big|F(S_{n_k}^*) - F(B_{n_k}^*)\big|\Big] \to 0, \qquad \text{as } k \to \infty,$$

which yields a contradiction.

It thus suffices to prove (4) to show the theorem. We fix $K = 1$ for simplicity (the general case is similar). Recall the linear interpolation of S. For every $t \in [0,1]$ we have

$$|B_n^*(t) - S_n^*(t)| = \frac{|B_{nt} - S(nt)|}{\sqrt{n}}.$$

Since S interpolates between $S([nt])$ and $S([nt]+1) = S([nt]) \pm 1$ we have $|S([nt]) - S(nt)| \leq 1$. Recall also that with our coupling we have $S([nt]) = B_{\tau_{[nt]}}$ and so

$$\begin{aligned}|B_n^*(t) - S_n^*(t)| &\leq \frac{|B_{nt} - S([nt])| + |S(nt) - S([nt])|}{\sqrt{n}} \\ &\leq \frac{1}{\sqrt{n}} + \Big|B_n^*(t) - B_n^*(\tau_{[nt]}/n)\Big|.\end{aligned}$$

Taking the sup over $[0,1]$ we deduce using the notation Mod_f for the modulus of continuity of the continuous function f over the interval $[0,2]$ that

$$\sup_{t\in[0,1]} |B_n^*(t) - S_n^*(t)| \leq \frac{1}{\sqrt{n}} + \mathsf{Mod}_{B_n^*}\Big(\sup_{t\in[0,1]}\Big|t - \frac{\tau_{[nt]}}{n}\Big|\Big).$$

Since for every $n \geq 0$, B_n^* is distributed as a standard Brownian motion B we have the equality in distribution $\mathsf{Mod}_{B_n^*}(\eta) = \mathsf{Mod}_B(\eta)$ for every $n \geq 0$ and every $\eta > 0$. However, since B is almost surely continuous we have $\mathsf{Mod}_B(\eta) \to 0$ in probability as $\eta \to 0$ (exercise). We are thus reduced to prove that

$$\sup_{t\in[0,1]}\Big|t - \frac{\tau_{[nt]}}{n}\Big| \to 0, \qquad \text{in probability.} \tag{5}$$

Recall that $\tau_i = (\tau_1 - \tau_0) + (\tau_2 - \tau_1) + \ldots + (\tau_i - \tau_{i-1})$ is a sum of iid random variables distributed as $\min(T_1, T_{-1})$ and thus of mean 1 by Theorem 11. It follows from the law of large numbers that

$$\frac{\tau_n}{n} \xrightarrow[n\to\infty]{a.s.} 1.$$

Thanks to Exercise 42 this implies (5) and completes the proof of the theorem. $\square$

Exercise 42 (functional law of large numbers)**.** Let a_n be a sequence of real numbers so that $\lim n^{-1} a_n \to 1$. Prove that

$$\lim_{n\to\infty} \sup_{0\le k\le n} \left| \frac{a_k}{n} - \frac{k}{n} \right| = 0.$$

Deduce that if $(X_i)_{i\ge 0}$ is a sequence of i.i.d. random variables such that $E[|X_0|] < \infty$ and $E[X_0] = 1$ then the random functions

$$\left(\frac{X_1 + \ldots + X_{[nt]}}{n} \right)_{t\ge 0},$$

converge almost surely on every compact of $\mathbb{R}_+$ towards the deterministic map $t \mapsto t$.

6.1 From the continuous to the discrete

In this section we use the functional central limit theorem to deduce limiting laws for random walks. In the following S_n is a random walk with independent increments having a finite variance $\sigma^2 \in (0, \infty)$ and zero mean (centered). Also S is interpolated linearly between integers and S_n^* denotes the same function as in the last section.

Corollary 2. *We have the following convergence in distribution*

$$\frac{1}{\sqrt{n}} \sup_{0\le k\le n} S_k \xrightarrow[n\to\infty]{(d)} \sigma \cdot |N|,$$

where N is a standard normal distribution.

Proof. By Donsker's theorem we have $S_n^* \to B$ in distribution for the topology of uniform convergence over $[0,1]$. Let F be the function $F : f \in \mathcal{C}([0,1], \mathbb{R}) \mapsto \sup_{[0,1]} f$. Using this notation we have

$$\frac{1}{\sqrt{n}} \sup_{0\le k\le n} S_k = \sigma \cdot F\big(S_n^*\big).$$

Since F is continuous for the uniform topology over $[0,1]$, we deduce by the continuous mapping theorem that $F(S_n^*)$ converges in distribution towards $F(B)$. The statement of the corollary then follows from the identity $\overline{B}_1 = |N|$ in law, see Theorem 5. $\square$

Corollary 3 (Arcsine law of the maximum)**.** *We have the following convergence in distribution*

$$\frac{1}{n} \inf\left\{0 \le k \le n : S_k = \sup_{0\le i\le n} S_i\right\} \xrightarrow[n\to\infty]{(d)} \textit{Arcsine distribution.}$$

Proof. If f is a continuous function over $[0,1]$ we set

$$G(f) = \inf \left\{ t \leq 1 : f(t) = \sup_{[0,1]} f \right\}$$

so that the left-hand side in the corollary is nothing but $G(S_n^*)$. Since by Donsker's theorem $S_n^* \to B$ in distribution for the uniform norm over $[0,1]$, we are willing to deduce that $G(S_n^*) \to G(B)$ in law as $n \to \infty$. The latter distribution being distributed as the arcsine distribution by Theorem 7, that would finish the job...

However, the function G is *not* continuous for the uniform norm over $[0,1]$ (exercise: find a counterexample). But it turns out that G is continuous at every function f such that the maxima of f over $[0,1]$ is uniquely attained (exercise). Since the local maxima of Brownian motion are distinct (Exercise 26) we deduce that a.s. G is continuous at B (and measurable). We then use the enhanced version of the continuous mapping theorem to deduce that $G(S_n^*) \to G(B)$ in distribution as $n \to \infty$. $\square$

Exercise 43 (Arcsine law for the last zero)**.** Let $L_n = \sup\{k \in \{0,1,2,...,n-1\} : S_k S_{k+1} \leq 0\}$. Show that $n^{-1} L_n$ converges in distribution towards the arcsine distribution.

Exercise 44. * Show that we have the following convergence in distribution

$$\frac{1}{n^2} \inf\{k \geq 0 : S_k \geq n\} \xrightarrow[n \to \infty]{(d)} \frac{1}{\sigma^2 |N|^2}.$$

Exercise 45. * Let S_n by the simple symmetric random walk on $\mathbb{Z}$. Use the coupling with the Brownian motion of the proof of Theorem 12 to show that

$$P(S_i \geq 0 : i \leq n) \sim P(T_{-1} \geq n) \sim \frac{\sqrt{2}}{\sqrt{\pi n}}, \qquad \text{as } n \to \infty$$

Exercise 46. ** Think about transferring the law of iterated logarithm for the Brownian motion onto the simple symmetric random walk via the coupling of the proof of Theorem 12.

6.2 From the discrete to the continuous

We now use Donsker's theorem in the other direction, meaning that we will prove results on Brownian motion by first establishing the corresponding result for *one particular* random walk and then transfer it to Brownian motion using Theorem 12. Obviously, in this section we will focus on the most simple random walk (S_n) which is the simple symmetric random walk on $\mathbb{Z}$.

6.2.1 Third arcsine law

Theorem 13. *The random variable* $\int_0^1 dt\, \mathbf{1}_{B_t > 0}$ *is arcsine distributed.*

There is a proof of this theorem directly in the continuous setting. But this is not within the reach of this course. Rather, we first prove a discrete statement which directly implies the theorem.

Lemma 3 (Richard)**.** *Let S_n be the simple symmetric random walk on $\mathbb{Z}$. Then*

$$\#\big\{k \in \{1,\dots,n\} : S_k > 0\big\} \overset{(d)}{=} \min\big\{k \in \{0,1,\dots,n\} : S_k = \max_{0 \leq j \leq n} S_j\big\}.$$

Proof of Theorem 13 with Lemma 3. If f is a continuous function over $[0,1]$ we let

$$H(f) = \int_0^1 dt\ \mathbf{1}_{f(t)>0}.$$

Using this notation, it is easy to see that

$$\frac{1}{n}\#\{k \in \{1,\dots,n\} : S_k > 0\} + \frac{\#\{k \le n : S_{k-1} = 1 \text{ and } S_k = 0\}}{n} = H(S_n^*).$$

It is easy to show that the second term of the left-hand side is small, indeed

$$E[\#\{k \le n : S_k = 0\}] = \sum_{k=0}^{n} P(S_k = 0) \sim C\sqrt{n}$$

so that $n^{-1}\#\{k \le n : S_k = 0\}$ converges in probability to 0 as $n \to \infty$. Combining this observation with Lemma 3 and Corollary 3 we deduce that $H(S_n^*)$ converges in distribution to the arcsine law.

On the other hand, by Donsker's theorem we have $S_n^* \to B$ uniformly over $[0,1]$. We are tempted to say that this implies $H(S_n^*) \to H(B)$ in distribution... which would imply the theorem. But the function H is not continuous. However, an exercise shows that the function H is continuous at functions f where

$$\lim_{\varepsilon \to 0} \int_0^1 dt\ \mathbf{1}_{f(t)\in[0,\varepsilon]} = 0,$$

or equivalently at functions f so that

$$\int_0^1 dt\ \mathbf{1}_{f(t)=0} = 0,$$

which is almost surely the case for the Brownian motion as seen in Theorem 6. The statement thus follows from an application of the enhanced continuous mapping theorem. $\square$

Proof of Richard's lemma. We see a random walk as a list $[+,+,-,+\dots]$ ordered from tail to head. The standard random walk S just grows by adding $\pm$ to the head. We define a new random walk $\tilde{S}$ by re-arranging the steps of S. Inductively we apply the following procedure: Initially the walk is empty list $[]$, and then for every $k \ge 1$, if $S_k > 0$ then the step $S_k - S_{k-1}$ is inserted at the tail of the list $\tilde{S}$ otherwise it is inserted at the head of $\tilde{S}$. Here is an example of the walk S and the associated walk $\tilde{S}$ at times goes on:

$$[] \leftrightarrow [] \qquad [-] \leftrightarrow [-] \qquad [-,+] \leftrightarrow [-,+] \qquad [-,+,+] \leftrightarrow [+,-,+] \qquad [-,+,+,-] \leftrightarrow [+,-,+,-]$$

It is easy to check (exercise) by induction that

$$\#\{k \in \{1,\dots,n\} : S_k > 0\} = \min\{k \in \{0,1,\dots,n\} : \tilde{S}_k = \max_{0\le j\le n} \tilde{S}_j\}.$$

To prove the lemma, we will now show that for every fixed $n \ge 0$ we have $S_n = \tilde{S}_n$ in distribution. We here also use induction. The statement is clear for $n = 0$ and $n = 1$. Let $i_1, \dots, i_{n+1} \in$

$\{+,-\}^{n+1}$ and let us compute $P(\tilde{S}_n = [i_1, \ldots, i_{n+1}])$. By construction, if I_n is the nth increment of S we have

$$
\begin{aligned}
&P(\tilde{S}_{n+1} = [i_1, \ldots, i_{n+1}]) \\
&= P\left(\tilde{S}_{n+1} = [i_1, \ldots, i_{n+1}]\right) \mathbf{1}_{i_1+\ldots+i_{n+1}>0} + P\left(\tilde{S}_{n+1} = [i_1, \ldots, i_{n+1}]\right) \mathbf{1}_{i_1+\ldots+i_{n+1}\leq 0} \\
&= P\left(\tilde{S}_n = [i_2, \ldots, i_{n+1}]\mathbf{1}_{I_{n+1}=i_1}\right) \mathbf{1}_{i_1+i_2+\ldots+i_{n+1}>0} + P\left(\tilde{S}_n = [i_1, \ldots, i_n]\mathbf{1}_{I_{n+1}=i_{n+1}}\right) \mathbf{1}_{i_1+i_2+\ldots+i_n+i_{n+1}\leq 0} \\
&= 2^{-n}\frac{1}{2}\left(\mathbf{1}_{i_1+i_2+\ldots+i_n+i_{n+1}>0} + \mathbf{1}_{i_1+i_2+\ldots+i_n+i_{n+1}\leq 0}\right) = 2^{-n-1},
\end{aligned}
$$

where we used the independence of the steps of S to go from the third to the last line: note that $\tilde{S}_n$ only depends on the first n steps of S and is thus independent of I_{n+1}. $\square$

6.2.2 Lévy's $\overline{B} - B$

Theorem 14 (Lévy)**.** *The process $(\overline{B}_t - B_t)_{t\geq 0}$ has the same law as $(|B|_t)_{t\geq 0}$.*

Sketch of the proof. Let (S_n) be the simple symmetric random walk and recall the notation S^*_n. Fix $K > 0$. We introduce the supremum process $\overline{S}^*_n(t) = \sup_{s\leq t} S^*_n$. By Donsker's theorem we have

$$
\overline{S}^*_n \to \overline{B} \quad \text{and } |S^*_n| \to |B| \tag{6}
$$

in distribution uniformly over $[0, K]$. On the other hand, the process $X_n = \sup_{k\leq n} S_k - S_n$ is a Markov chain whose probabilities transition are exactly the same as the Markov chain $|S|_n$ except at 0 where $P(X_n = 0) = 1/2$. In others words, X_n can be seen as the Markov chain $|S|_n$ where each time $|S|_n$ touches 0 it stays there for a independent geometric variable of parameter 1/2. We can thus couple the chain (X_n) with the chain $(|S|_n)$ such that

$$
X_{n+\xi_n} = |S|_n,
$$

where $\xi_n = \sum_{i=0}^n \mathbf{1}_{|S|_i=0} G_i$ with i.i.d. G_i geometric of parameter 1/2. Using similar estimates as in the proof of Theorem 13 we have $n^{-1}\sup_{i\leq n} \xi_i = n^{-1}\xi_n \to 0$ in probability. We thus deduce using (6) that (with an abuse of notation) we have the following convergences in distribution uniformly over $[0, K]$,

$$
\overline{B} - B \quad \longleftarrow \quad X^*_n \approx |S^*_n| \quad \longrightarrow \quad |B|. \quad \square
$$

This theorem sheds a new light on Theorem 7. Indeed, it is clear now that M and L have the same distribution since the last zero of $|B|$ (or B) before 1 coincides with the distribution of the last zero before 1 of $\overline{B} - B$ which is nothing but the last time B reaches its maximum over $[0, 1]$.