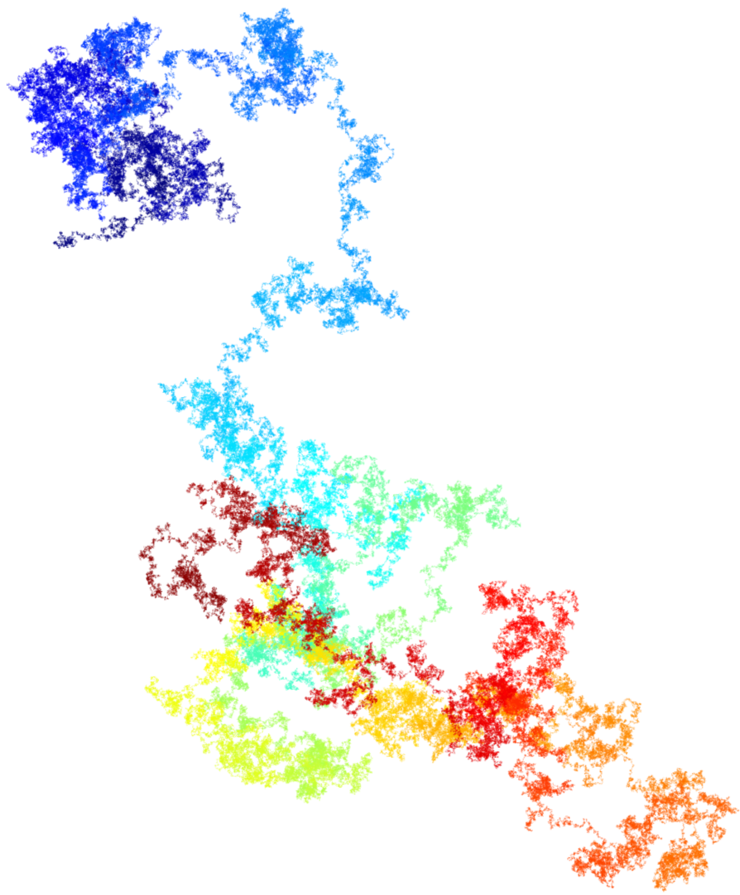
Tsinghua University 2018–2019

Advanced Probability

part III : Brownian Motion

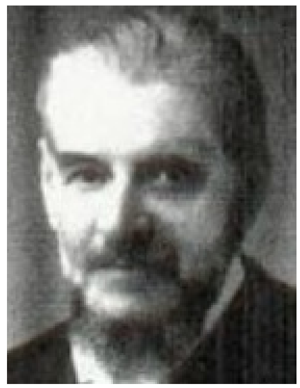


Singh Arvind[1]

Adapted from original lecture notes by **Shi Zhan**

[1]contact: arvind.singh@math.u-psud.fr

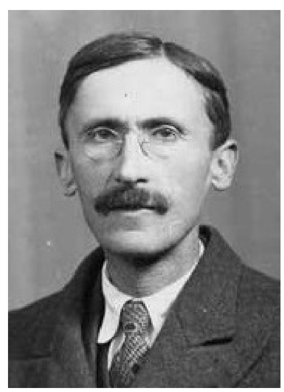
(a) **Robert Brown** (1773 – 1858)



(b) **Louis Bachelier** (1870 – 1946)

(c) **Albert Einstein** (1879 – 1955)

(d) **Norbert Wiener** (1894 – 1964)



(e) **Paul Levy** (1886 – 1971)

Some pioneers of the theory of Brownian motion.

Contents

Chapter 4

Construction of Brownian motion

> Observe what happens when sunbeams are admitted into a building and shed light on its shadowy places. You will see a multitude of tiny particles mingling in a multitude of ways... their dancing is an actual indication of underlying movements of matter that are hidden from our sight... It originates with the atoms which move of themselves. Then those small compound bodies that are least removed from the impetus of the atoms are set in motion by the impact of their invisible blows and in turn cannon against slightly larger bodies. So the movement mounts up from the atoms and gradually emerges to the level of our senses, so that those bodies are in motion that we see in sunbeams, moved by blows that remain invisible.
>
> Lucretius, *De rerum natura* (c. 60 BC).

Brownian motion lies in the intersection of several important families of random processes (martingales, Markov processes, Gaussian processes), and is the fundamental example in each theory. These notes give a brief introduction to Brownian motion, providing an account of its basic properties.

The expression "Brownian motion" originates from the highly irregular movement of pollen grains on the surface of water, observed by the Scottish botanist Robert Brown in 1828. Subsequently, Bachelier (1900) and Einstein (1905) studied quantitatively this irregular movement, in finance and in physics respectively. It is Wiener, in 1923, who established the mathematical model of Brownian motion which we are going to study in these notes. Many deep properties of Brownian motion were later discovered by Paul Lévy (1939, 1948). Brownian motion has since become ubiquitous throughout Probability theory and is a key object encountered in several areas of sciences. For example, in 1973, Black and Scholes used Brownian motion to model options prices in financial mathematics and create the formula now bearing their name, and for which Scholes was awarded the Nobel Prize in Economics in 1997.

4.1. Warm-up: Gaussian law and Gaussian vectors

From now on, we let $(\Omega, \mathscr{F}, \mathbb{P})$ denote the ambient probability space on which all our random variables/processes shall be defined except when explicitly specified otherwise. The main distribution related to Brownian motion is the Gaussian law. We start by recalling well known properties of the Gaussian law and its multidimensional counterparts, the Gaussian vectors.

Let $\mu \in \mathbb{R}$ and $\sigma > 0$. We say that ξ is a **Gaussian random variable** $\mathscr{N}(\mu, \sigma^2)$ if it has density

$$\mathbb{P}(\xi \in \mathrm{d}x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\Big(-\frac{(x-\mu)^2}{2\sigma^2}\Big)\,\mathrm{d}x, \qquad x \in \mathbb{R}.$$

The characteristic function of ξ is given by the formula:

$$\mathbb{E}[\mathrm{e}^{it\xi}] = \mathrm{e}^{\mu it - \frac{\sigma^2 t^2}{2}}, \qquad t \in \mathbb{R}.$$

Clearly, η is Gaussian $\mathscr{N}(\mu, \sigma^2)$ if and only if $\eta = \sigma\xi + \mu$, where ξ is $\mathscr{N}(0, 1)$. Also, if $\tilde{\eta}$ is Gaussian $\mathscr{N}(\tilde{\mu}, \tilde{\sigma}^2)$ independent of η, then $\eta + \eta'$ is Gaussian $\mathscr{N}(\mu + \mu', \sigma^2 + \tilde{\sigma}^2)$.

Theorem 4.1. (Gaussian tail). *If ξ is Gaussian $\mathscr{N}(0, 1)$, then for any $x > 0$,*

$$\frac{1}{\sqrt{2\pi}}\Big(\frac{1}{x} - \frac{1}{x^3}\Big)\mathrm{e}^{-x^2/2} \leq \mathbb{P}(\xi > x) \leq \frac{1}{\sqrt{2\pi}}\frac{1}{x}\mathrm{e}^{-x^2/2},$$
$$\mathbb{P}(\xi > x) \leq \mathrm{e}^{-x^2/2}.$$

Proof. Left as an exercise □

Remark 4.2. 1. We have $\mathbb{P}(\xi > x) \sim \frac{1}{\sqrt{2\pi}}\frac{1}{x}\mathrm{e}^{-x^2/2}$ as $x \to \infty$.

2. The upper bound $\mathrm{e}^{-x^2/2}$ is less precise than $\frac{1}{(2\pi)^{1/2}}\frac{1}{x}\mathrm{e}^{-x^2/2}$ but has the advantage of being bounded on $\mathbb{R}_+$. It is useful when we do not need much precision.

Proposition 4.3 (Convergence of sequences of Gaussian random variable). *Let (ξ_n) be a sequence of random variables such that for any n, ξ_n is Gaussian $\mathscr{N}(\mu_n, \sigma_n^2)$.*

1. *The convergence in distribution $\xi_n \to \xi$ holds if and only if $\mu := \lim_{n\to\infty} \mu_n$ and $\sigma^2 := \lim_{n\to\infty} \sigma_n^2$ both exist. Moreover, the limit ξ is Gaussian $\mathscr{N}(\mu, \sigma^2)$.*

2. *If $\xi_n \to \xi$ in probability, then the convergence also holds in L^p for any $p \in [1, \infty)$.*

Proof. Left as an exercise □

Remark 4.4. The limit in Proposition 4.3 may be degenerated (*i.e.* $\sigma^2 = 0$). In the following, we consider the Dirac measure δ_μ at $\mu \in \mathbb{R}$ as a Gaussian distribution $\mathscr{N}(\mu, 0)$.

Definition 4.5 (Gaussian random vectors). *We say that an n-dimensional random variable $\xi := (\xi_1, \cdots, \xi_n)$ is a* ***Gaussian vector*** *if any linear combination of its components is Gaussian. We define its* ***mean vector***

$$\mu^\xi := (\mathbb{E}[\xi_1], \ldots, \mathbb{E}[\xi_n]) \in \mathbb{R}^n$$

and it ***covariance matrix*** *Σ^ξ as the $n \times n$ real matrix with coefficients*

$$\Sigma^\xi_{i,j} := \mathrm{Cov}(\xi_i, \xi_j) = \mathbb{E}[(\xi_i - \mathbb{E}[\xi_i])(\xi_j - \mathbb{E}[\xi_j])].$$

Remark 4.6. If $(\xi_1, \cdots, \xi_n)$ is a Gaussian random vector, then each of its component is a Gaussian random variable but the converse is wrong! Consider $\xi_1 \sim \mathscr{N}(0,1)$ and B a Rademacher random variable $\mathbb{P}(B=1) = \mathbb{P}(B=-1) = \frac{1}{2}$ independent of ξ. Then ξ_1 and $\xi_2 := B\xi_1$ are both Gaussians but (ξ_1, ξ_2) is not a Gaussian vector because $\xi_1 + \xi_2$ is not Gaussian (there is an atom of weight 1/2 at 0).

Theorem 4.7. *The law of a Gaussian vector ξ is uniquely determined by its mean vector and covariance matrix. Its characteristic function is given by*

$$\mathbb{E}\left[\exp\left(i\left\langle x, \xi\right\rangle\right)\right] = \exp\left(i\left\langle x, \mu^\xi\right\rangle - \frac{1}{2}\left\langle x, \Sigma^\xi x\right\rangle\right) \qquad x \in \mathbb{R}^n. \tag{4.1}$$

where $\langle\cdot,\cdot\rangle$ denote the usual inner product in $\mathbb{R}^n$.

Proof. the random variable $\langle x, \xi\rangle$ is a linear combination of the components of ξ so it is, by definition, a Gaussian random variable $\mathscr{N}(\tilde{\mu}, \tilde{\sigma}^2)$ hence $\mathbb{E}\left[\exp\left(i\left\langle x, \xi\right\rangle\right)\right] = \exp\left(i\tilde{\mu} - \frac{\tilde{\sigma}^2}{2}\right)$. Straightforward computations show that $\tilde{\mu} := \mathbb{E}[\langle x, \xi\rangle] = \langle x, \mu^\xi\rangle$ and $\sigma^2 := \mathrm{Var}(\langle x, \xi\rangle) = \langle x, \Sigma^\xi x\rangle$. This proves (4.1) and shows at the same time that the characteristic function (hence the law) of ξ is uniquely determined by the mean vector and covariance matrix. □

Corollary 4.8. *(i) Let $\xi := (\xi_1, \ldots, \xi_n)$ be a Gaussian random vector. Then $\xi_1, \ldots, \xi_n$ are independent if and only if the covariance matrix of ξ is diagonal.*

(ii) Let $(\xi_1, \ldots, \xi_n, \eta_1, \ldots, \eta_m, \theta_1, \cdots, \theta_\ell)$ be a Gaussian vector. Then the vectors $(\xi_1, \ldots, \xi_n)$ and $(\eta_1, \ldots, \eta_m)$ are independent if and only if $\mathrm{Cov}(\xi_i, \eta_j) = 0$, $\forall i \leq n,\ j \leq m$.

4.2. Definition and construction of Brownian motion

A stochastic process is simply a collection of random variable $X := (X_t,\ t \in \mathbf{T})$ where t is called the time index and $\mathbf{T}$ is the set of times for which the process is defined. In previous chapters, while studying Markov chain and Martingales, we have encountered **discrete-time processes** where time is indexed by integers. By contrast, Brownian motion is a **continuous time process** and the time index t will now range in (a subset of) $\mathbb{R}_+ = [0, \infty)$.

Definition 4.9. *A real-valued process $B = (B_t,\ t \geq 0)$ is said to be* **Brownian motion** *if it satisfies the following properties:*

(i) For any n and any $0 \leq t_1 \leq t_2 \leq \cdots \leq t_n$, the random variables $(B_{t_n} - B_{t_{n-1}})$, $\cdots$, $(B_{t_2} - B_{t_1})$, B_{t_1} are independent.

(ii) For any $t \geq s \geq 0$, the random variable $B_t - B_s$ is Gaussian $\mathcal{N}(0, t-s)$.

(iii) $t \mapsto B_t$ is continuous a.s.

(iv) $B_0 = 0$ a.s.

Remark 4.10. A Brownian motion has independent and stationary increments. Stochastic processes with this property are called **Lévy processes**.

Brownian motion is the most famous member of a class of processes called Gaussian processes.

Definition 4.11. *A stochastic process $X := (X_t,\ t \in \mathbf{T})$ is a* **Gaussian process** *if*

$$\forall n,\ \forall (t_1, \cdots, t_n) \in \mathbf{T}^n,\ (X_{t_1}, \cdots, X_{t_n}) \text{ is a Gaussian vector.}$$

We say that X is ***centered*** *if $\mathbb{E}[X_t] = 0$ for all $t \in \mathbf{T}$.*

Proposition 4.12. *Let $X = (X_t, t \geq 0)$ be a process with continuous paths and $X_0 = 0$ a.s. The following assertions are equivalent:*

1. *X is a Brownian motion*

2. *X is centered Gaussian process with covariance*

$$\mathbb{E}[X_s X_t] = \min\{s, t\} =: s \wedge t, \qquad \text{for all } s, t \geq 0.$$

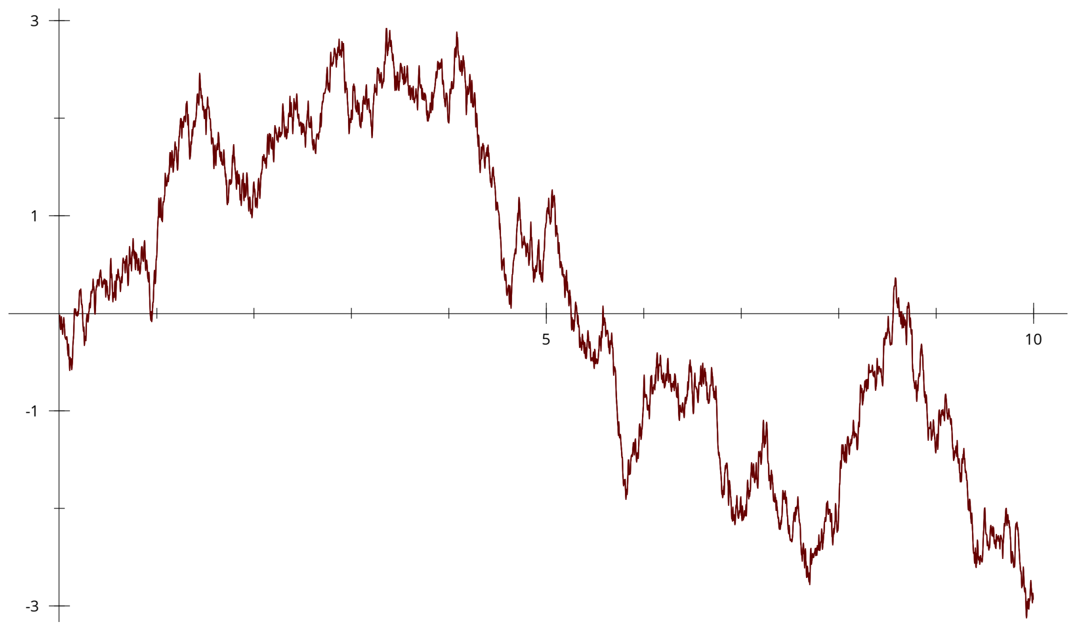


Figure 4.1: Sample path of a standard Brownian motion.

Proof.

"⇒" Let $0 \le t_1 \le t_2 \le \cdots \le t_n$. By assumption, $X_{t_n} - X_{t_{n-1}}, \cdots, X_{t_2} - X_{t_1}, X_{t_1} = X_{t_1} - X_0$ are independent Gaussian hence $(X_{t_1}, X_{t_2} - X_{t_1}, \cdots, X_{t_n} - X_{t_{n-1}})$ is a Gaussian random vector, and so is $(X_{t_1}, X_{t_2}, \cdots, X_{t_n})$. Thus X is a Gaussian process which is obviously centered.

Let us check the covariance of X. Let $t \ge s \ge 0$. We have $\mathbb{E}(X_s X_t) = \mathbb{E}(X_s(X_t - X_s)) + \mathbb{E}(X_s^2)$. Since X_s and $X_t - X_s$ are independent, we have $\mathbb{E}(X_s(X_t - X_s)) = 0$, whereas by Property (ii) of Brownian motion, $\mathbb{E}(X_s^2) = s$. Hence $\mathbb{E}(X_s X_t) = s$.

"⇐" Let $t \ge s \ge 0$. By assumption, $X_t - X_s$ is a centered Gaussian, with variance $\mathbb{E}(X_t - X_s)^2 = \mathbb{E}(X_t^2) + \mathbb{E}(X_s^2) - 2\mathbb{E}(X_s X_t) = t + s - 2s = t - s$. Thus, Property (ii) holds.

Let $0 \le t_1 \le t_2 \le \cdots \le t_n$. We know that $(X_{t_n} - X_{t_{n-1}}, \cdots, X_{t_2} - X_{t_1}, X_{t_1})$ is a Gaussian random vector, whose covariance matrix is diagonal: indeed, for $j > i$, $\mathbb{E}[(X_{t_j} - X_{t_{j-1}})(X_{t_i} - X_{t_{i-1}})] = \mathbb{E}(X_{t_j} X_{t_i}) - \mathbb{E}(X_{t_{j-1}} X_{t_i}) - \mathbb{E}(X_{t_j} X_{t_{i-1}}) + \mathbb{E}(X_{t_{j-1}} X_{t_{i-1}}) = t_i - t_i - t_{i-1} + t_{i-1} = 0$. By Corollary 4.8 (i), the components of this Gaussian random vector are independent hence Property (i) holds.

□

Remark 4.13. Brownian motion, as defined by Definition 4.9 starts from $B_0 = 0$ a.s. and is usually called **standard Brownian motion**. More generally, it is useful to define Brownian motion starting from an arbitrary position $x \in \mathbb{R}$ by considering the process $B + x$. This is still a Gaussian process (but not centered anymore) and which still satisfies (i), (ii) and (iii) of Definition 4.9.

The first thing we must check is that Definition 4.9 is not empty and that such a process exists, which is not obvious.

Theorem 4.14 (**Wiener, 1923**). *Brownian motion does exists.*

Proof. (Lévy 1948). We start by constructing a Brownian motion defined for $t \in [0, 1]$ as a uniform limit of piece-wise linear Gaussian processes.

Let $(\xi_{k,n},\, 0 \le k \le 2^n,\, n \ge 0)$ be a family of i.i.d. Gaussian $\mathscr{N}(0,\, 1)$ random variables. Let $(X_n(t),\, t \in [0, 1],\, n \ge 0)$ be the sequence of random processes defined by:

- for any $n \ge 0$, $t \mapsto X_n(t)$ is affine on each interval of type $[\frac{k}{2^n},\, \frac{k+1}{2^n}]$;
- $X_0(0) = 0$, $X_0(1) = \xi_{0,0}$;
- $X_n(\frac{2j}{2^n}) = X_{n-1}(\frac{2j}{2^n})$, $X_n(\frac{2j+1}{2^n}) = X_{n-1}(\frac{2j+1}{2^n}) + \frac{\xi_{2j+1,n}}{2^{(n+1)/2}}$.

See Figure 4.2 for an illustration of this construction. It is easy (and left as an exercise) to check that for any $n \ge 0$, the family $(X_n(\frac{k}{2^n}),\, 0 \le k \le 2^n)$ is a centered Gaussian vector with covariance matrix $\mathbb{E}[X_n(\frac{k}{2^n})X_n(\frac{\ell}{2^n})] = \frac{k}{2^n} \wedge \frac{\ell}{2^n}$.

Let $n \ge 0$. We see that $(X_n(t),\, t \in [0, 1])$ is a centered Gaussian process because any linear combination $\sum_{i=1}^m a_i X_n(t_i)$ is also a linear combination of $(X_n(\frac{k}{2^n}),\, 0 \le k \le 2^n)$.

Consider the event $A_n := \{\sup_{t\in[0,1]} |X_n(t) - X_{n-1}(t)| > 2^{-n/4}\}$. We have

$$\begin{aligned}\mathbb{P}(A_n) &= \mathbb{P}\Big(\bigcup_{j=0}^{2^{n-1}-1} \Big\{ \sup_{t\in[\frac{2j}{2^n},\frac{2j+2}{2^n}]} |X_n(t) - X_{n-1}(t)| > 2^{-n/4} \Big\}\Big) \\ &= \mathbb{P}\Big(\bigcup_{j=0}^{2^{n-1}-1} \Big\{ \frac{|\xi_{2j+1,n}|}{2^{(n+1)/2}} > 2^{-n/4} \Big\}\Big) \le \sum_{j=0}^{2^{n-1}-1} \mathbb{P}\Big(|\xi_{2j+1,n}| > 2^{(n+2)/4} \Big).\end{aligned}$$

By symmetry and Theorem 4.1, $\mathbb{P}(|\xi_{2j+1,n}| > 2^{(n+2)/4}) \le 2\exp(-2^{n/2})$; so we have $\mathbb{P}(A_n) \le 2^{n-1}\exp(-2^{n/2})$, which implies that $\sum_n \mathbb{P}(A_n) < \infty$. By the Borel-Cantelli lemma, there exists an $E \in \mathscr{F}$ with $\mathbb{P}(E) = 1$, such that for all $\omega \in E$, $\sup_{t\in[0,1]} |X_n(t,\omega) - X_{n-1}(t,\omega)| \le 2^{-n/4}$, $\forall\, n \ge n_0(\omega)$. [For $\omega \notin E$, we can define, for example, $X(t,\, \omega) := 0$, $\forall t \in [0,\, 1]$.] In

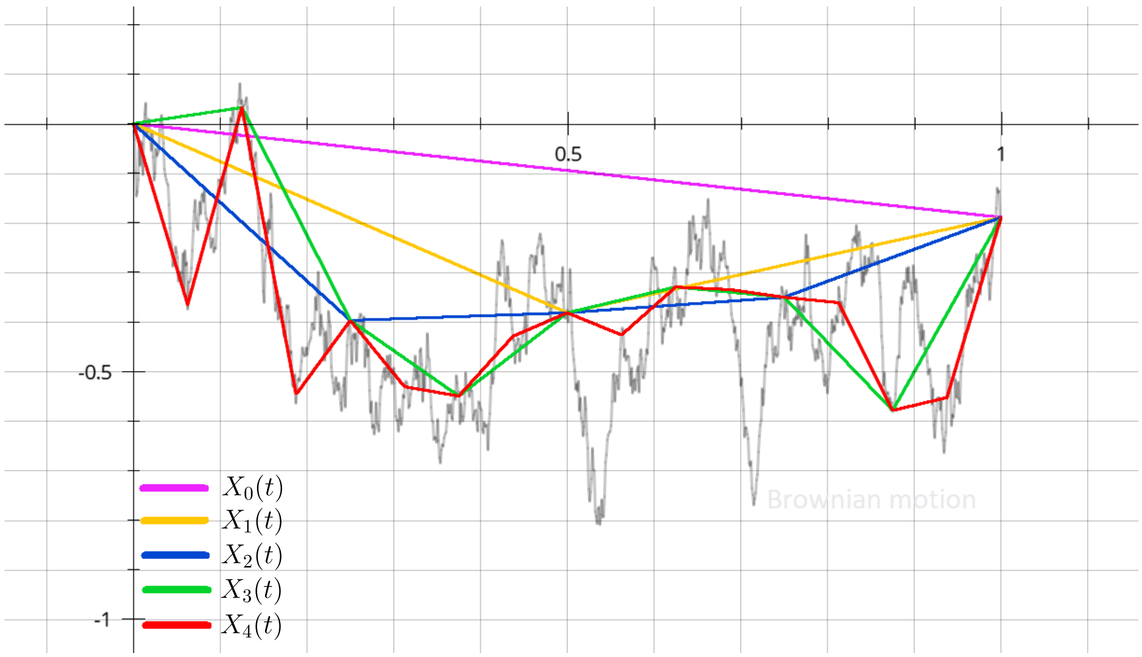


Figure 4.2: Construction of Brownian motion by uniform approximation via linear interpolation on dyadic intervals.

particular, ω-a.s., the function $X_n(\bullet, \omega)$ converges uniformly on $[0, 1]$ to a continuous limit denoted by $X(\bullet, \omega)$. By Proposition 4.3, the process $X = (X(t), t \in [0, 1])$ is a centered Gaussian process with continuous paths and with $X(0) = 0$ a.s.

Let us check the covariance matrix of X. Let $0 \leq s \leq t \leq 1$ and $n \geq 0$. There exists a pair (k, ℓ) with $k \leq \ell$ such that $s \in [\frac{k}{2^n}, \frac{k+1}{2^n}]$ and $t \in [\frac{\ell}{2^n}, \frac{\ell+1}{2^n}]$. Since X_n is affine on $[\frac{k}{2^n}, \frac{k+1}{2^n}]$, we have $X_n(s) = \alpha X_n(\frac{k}{2^n}) + (1-\alpha) X_n(\frac{k+1}{2^n})$, where $\alpha := k + 1 - 2^n s \in [0, 1]$. Similarly, $X_n(t) = \beta X_n(\frac{\ell}{2^n}) + (1-\beta) X_n(\frac{\ell+1}{2^n})$, with $\beta := \ell + 1 - 2^n t \in [0, 1]$. It follows that

$$\mathbb{E}[X_n(s) X_n(t)] = \frac{\alpha\beta k}{2^n} + \frac{(1-\alpha)\beta((k+1) \wedge \ell)}{2^n} + \frac{\alpha(1-\beta)k}{2^n} + \frac{(1-\alpha)(1-\beta)(k+1)}{2^n}.$$

The expression on the right-hand side is, for $n \to \infty$,

$$\frac{\alpha\beta k}{2^n} + \frac{(1-\alpha)\beta k}{2^n} + \frac{\alpha(1-\beta)k}{2^n} + \frac{(1-\alpha)(1-\beta)k}{2^n} + \mathcal{O}(\frac{1}{2^n}),$$

which is $\frac{k}{2^n} + \mathcal{O}(\frac{1}{2^n}) = s + \mathcal{O}(\frac{1}{2^n})$. Letting $n \to \infty$, and by Proposition 4.3 again, we obtain $\mathbb{E}[X(s)X(t)] = s = s \wedge t$. Consequently, $(X(t), t \in [0, 1])$ is a Brownian motion defined on $[0, 1]$.

To conclude, we extend our construction of Brownian motion from $[0, 1]$ to $[0, \infty)$. To do so, we consider a sequence $(B_t^m, t \in [0, 1])$, $m \geq 0$ of independent Brownian motions on

$[0,1]$ and we define

$$B_t := B^{\lfloor t \rfloor}_{t-\lfloor t \rfloor} + \sum_{0 \le m < \lfloor t \rfloor} B^m_1 \quad \text{for } t \ge 0,$$

where $\lfloor t \rfloor$ denote the integer part of t. It is straightforward to check that B is a centered Gaussian process with the required covariance hence it is a Brownian motion on $[0,\infty)$. □

Remark 4.15. One can also define Brownian motion with time indexed by $\mathbb{R}$ as a centered Gaussian process $(B_t, t \in \mathbb{R})$ with $B_0 = 0$ such that $B_s - B_t$ is $\mathscr{N}(0, |t-s|)$ for any s, t. Such a process may be constructed from two centered independent Brownian motions B^- and B^+ on $[0,\infty)$ by setting:

$$B_t := \begin{cases} B^-_{-t} & \text{if } t < 0, \\ B^+_t & \text{if } t \ge 0. \end{cases}$$

4.3. Regularization of sample paths

A probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is said to be **complete** if $\mathscr{F}$ contains all $\mathbb{P}$-negligible set, meaning that any set included in a measurable set of null probability is itself measurable. Given a probability space, we can always enrich its σ-field $\mathscr{F}$ to make it a complete probability space. Doing so does not change law of random variables defined on it while it helps to prevent technical problems about measurability of sets.

From now on, we assume that our probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is complete.

In the definition of Brownian motion, we required the trajectories of the process to be continuous. However, one can ask if this assumption is necessary and whether Properties (i) and (ii) of Definition 4.9 already characterize Brownian motion. To answer this question, we need to introduce the notion of modification and indistinguishability of processes.

Definition 4.16. *Let $(X_t,\, t \in \mathbf{T})$ and $(\widetilde{X}_t,\, t \in \mathbf{T})$ be processes indexed by the same set $\mathbf{T}$. We say that $\widetilde{X}$ is a* ***modification*** *(or version) of X if*

$$\forall\, t \in \mathbf{T}, \qquad \mathbb{P}[X_t = \widetilde{X}_t] = 1.$$

Since an (at most countable) intersection of set of full measure also has full measure, it follows that, for any $t_1, t_2, \cdots, t_n$, the random vectors $(\widetilde{X}_{t_1}, \cdots, \widetilde{X}_{t_n})$ and $(X_{t_1}, \cdots, X_{t_n})$ have the same distribution. In particular, if X is Brownian motion, then so is $\widetilde{X}$. On the other hand, the trajectories of $\widetilde{X}$ may have a totally different behavior from those of X. It can happen that the trajectories of $\widetilde{X}$ are all continuous while those of X are all discontinuous:

Example 4.17. Let Z be a random variable which has no atom (for example uniform on $[0,1]$). Define the processes $(X_t,\ t\geq 0)$ and $(\widetilde{X}_t,\ t\geq 0)$ by

$$X_t = \begin{cases} 1 & \text{if } t-Z\in\mathbb{Q}, \\ 0 & \text{otherwise} \end{cases} \qquad \text{and} \qquad \widetilde{X}_t = 0 \qquad \text{for } t\geq 0.$$

Then $\widetilde{X}$ is a modification of X which is a.s. continuous on $[0,\infty)$ while X is a.s. nowhere continuous!

Definition 4.18. *Two processes X and $\widetilde{X}$ are **indistinguishable** if*

$$\mathbb{P}[\forall\, t\in\mathbf{T},\ \ X_t=\widetilde{X}_t]=1.$$

Notice that, a priori, we know nothing about the measurability $G:=\{\forall\, t\in\mathbf{T},\ \ X_t=\widetilde{X}_t\}$ so the definition above requires it implicitly. But since our probability space is assumed to be complete, this is equivalent to requiring that G contains a set of probability 1.

If X and $\widetilde{X}$ are indistinguishable, then $\widetilde{X}$ is obviously a modification of X. The notion of indistinguishability, however, is stronger: two indistinguishable processes almost surely have the same trajectories. For instance, assume that $\mathbf{T}=I$ is an interval of $\mathbb{R}$, and that the trajectories of X and $\widetilde{X}$ are a.s. continuous[1], then $\widetilde{X}$ is a modification of X if and only if X and $\widetilde{X}$ are indistinguishable: indeed, if $\widetilde{X}$ is a modification of X, then a.s. for all $t\in I\cap\mathbb{Q}$, $X_t=\widetilde{X}_t$. By continuity, a.s. for all $t\in I$, $X_t=\widetilde{X}_t$, which means that X and $\widetilde{X}$ are indistinguishable

Theorem 4.19. (Kolmogorov's criterion). *Let $X=(X_t,\,t\in I)$ be a process indexed by an interval $I\subset\mathbb{R}$, taking values in a complete metric space (E,d). Suppose there exist $p>0$, $\varepsilon>0$ and $C>0$ such that*

$$\mathbb{E}[\,d(X_s,\,X_t)^p]\leq C\,|t-s|^{1+\varepsilon}, \qquad \forall\, s,t\in I.$$

Then there exists a modification $\widetilde{X}$ of X whose trajectories are locally Hölder continuous for exponent α, for any $\alpha\in(0,\,\frac{\varepsilon}{p})$, i.e., for all $T>0$ and $\alpha\in(0,\,\frac{\varepsilon}{p})$, there exists $C_\alpha(T,\omega)>0$ such that

$$d(\widetilde{X}_s(\omega),\,\widetilde{X}_t(\omega))\leq C_\alpha(T,\omega)\,|t-s|^\alpha, \qquad \forall\, s,t\in I,\ s,t\leq T.$$

In particular, there exists a continuous modification of X, which is unique in the sense of indistinguishability

[1] One can relax the continuity assumption and assume, for instance ,that the trajectories are right continuous and with left limits (*càdlàg*). This is the standard assumption when working with non-continuous stochastic processes.

Proof. The uniqueness is clear from the discussions in the previous paragraph. For simplicity, we assume $I = [0,\, 1]$. Let $\mathcal{D} = \{\frac{k}{2^n},\, n \in \mathbb{N}$ and $0 \le k < 2^n\}$ denote the set of dyadic numbers. This set is dense in $[0,1]$. We shall first prove that X is almost surely uniformly continuous on $\mathcal{D}$. First, using Markov's inequality, for $a > 0$ and $s,t \in [0,1]$, we have

$$\mathbb{P}\{d(X_s,\, X_t) \ge a\} \le \frac{\mathbb{E}[d(X_s,\, X_t)^p]}{a^p} \le \frac{C\,|t-s|^{1+\varepsilon}}{a^p}. \tag{4.2}$$

Applying this inequality to $s = \frac{i-1}{2^n}$ and $t = \frac{i}{2^n}$ and $a = 2^{-n\alpha}$ gives

$$\mathbb{P}\{d(X_{(i-1)/2^n},\, X_{i/2^n}) \ge 2^{-n\alpha}\} \le \frac{C}{2^{(1+\varepsilon-p\alpha)n}}, \qquad i = 1,\, 2,\, \cdots,\, 2^n.$$

Thus, by union bound,

$$\mathbb{P}\{\exists i \le 2^n :\ d(X_{(i-1)/2^n},\, X_{i/2^n}) \ge 2^{-n\alpha}\} \le \frac{C}{2^{(\varepsilon-p\alpha)n}},$$

which is summable in n since $p\alpha < \varepsilon$. Let

$$A := \{\exists n_0,\ \ \forall n \ge n_0,\ \ \forall 1 \le i \le 2^n\ \ d(X_{(i-1)/2^n},\, X_{i/2^n}) < 2^{-n\alpha}\}$$

By the Borel-Cantelli lemma, we have $P(A) = 1$. Consider now two dyadic numbers $s, t \in \mathcal{D}$ with $s \le t$. Let $q \ge 0$ be the largest integer satisfying $t - s \le 2^{-q}$. Let $k := \lfloor 2^q s \rfloor$ (so that $k \le \lfloor 2^q t \rfloor \le k+1$). We can find integers $\ell \ge 0$ and $m \ge 0$, such that

$$\begin{aligned} s &= \frac{k}{2^q} + \frac{\varepsilon_{q+1}}{2^{q+1}} + \cdots + \frac{\varepsilon_{q+\ell}}{2^{q+\ell}}, \\ t &= \frac{k}{2^q} + \frac{\widetilde{\varepsilon}_q}{2^q} + \frac{\widetilde{\varepsilon}_{q+1}}{2^{q+1}} + \cdots + \frac{\widetilde{\varepsilon}_{q+m}}{2^{q+m}}, \end{aligned}$$

where $\varepsilon_j,\, \widetilde{\varepsilon}_j \in \{0,1\}$. If we write

$$\begin{aligned} s_i &= \frac{k}{2^q} + \frac{\varepsilon_{q+1}}{2^{q+1}} + \cdots + \frac{\varepsilon_{q+i}}{2^{q+i}}, && 0 \le i \le \ell, \\ t_j &= \frac{k}{2^q} + \frac{\widetilde{\varepsilon}_q}{2^q} + \frac{\widetilde{\varepsilon}_{q+1}}{2^{q+1}} + \cdots + \frac{\widetilde{\varepsilon}_{q+j}}{2^{q+j}}, && 0 \le j \le m, \end{aligned}$$

then for $\omega \in A$,

$$\begin{aligned} d(X_s,\, X_t) &= d(X_{s_\ell},\, X_{t_m}) \\ &\le d(X_{s_0},\, X_{t_0}) + \sum_{i=1}^{\ell} d(X_{s_{i-1}},\, X_{s_i}) + \sum_{j=1}^{m} d(X_{t_{j-1}},\, X_{t_j}) \\ &\le K_\alpha(\omega)\, 2^{-q\alpha} + \sum_{i=1}^{\ell} K_\alpha(\omega)\, 2^{-(q+i)\alpha} + \sum_{j=1}^{m} K_\alpha(\omega)\, 2^{-(q+j)\alpha}, \end{aligned}$$

where

$$K_\alpha(\omega) := \sup_{n\geq 1} \max_{1\leq i\leq 2^n} \frac{d(X_{(i-1)/2^n}, X_{i/2^n})}{2^{-n\alpha}}$$

is finite by definition of A. hence, for $\omega \in A$, we find that

$$d(X_s,\, X_t) \leq 2K_\alpha(\omega) \sum_{i=0}^{\infty} 2^{-(q+i)\alpha} = \frac{2K_\alpha(\omega)2^{-q\alpha}}{1-2^{-\alpha}} \leq \frac{2^{1+\alpha}K_\alpha(\omega)}{1-2^{-\alpha}}\,(t-s)^\alpha,$$

because $2^{-(q+1)} < t-s$. Thus, a.s. the function $t \mapsto X_t(\omega)$ is Hölder continuous on $\mathcal{D}$ and *a fortiori* uniformly continuous on $\mathcal{D}$. Since $(E,\, d)$ is complete, this function a.s. admits an unique continuous extension to $I = [0,\, 1]$, and this extension is also Hölder continuous with exponent α. More precisely, pick $x_0 \in E$ and define, for all $t \in [0, 1]$,

$$\widetilde{X}_t(\omega) := \begin{cases} \lim_{s\to t,\, s\in\mathcal{D}} X_s(\omega) & \text{if } \omega \in A, \\ x_0 & \text{if } \omega \notin A. \end{cases}$$

The trajectories of $\widetilde{X}$ are Hölder continuous for exponent α. It remains to check that $\widetilde{X}$ is a modification of X. Let $t \in [0,\, 1]$. In view of (4.2), we have

$$\lim_{s\to t} X_s = X_t, \qquad \text{in probability.}$$

On the other hand, by construction, $\widetilde{X}_t$ is the a.s. limit of X_s when $s \to t$ and $s \in \mathcal{D}$. By uniqueness of the limit, we conclude that $\widetilde{X}_t = X_t$ a.s. □

Corollary 4.20. *Let $B = (B_t,\, t \geq 0)$ be a process satisfying (i), (ii) and (iv) of Definition 4.9. Then B satisfy the assumption of Kolmogorov's criterion. Thus, there exist a continuous modification of B which is a Brownian motion and it is unique up to indistinguishability. Moreover, any Brownian motion has trajectories that are locally Hölder for exponent $\frac{1}{2} - \varepsilon$, for all $\varepsilon \in (0,\, \frac{1}{2})$.*

Proof. Fix $\varepsilon \in (0,\, \frac{1}{2})$. Let $t,\, s \geq 0$. Since $B_t - B_s$ is Gaussian $\mathscr{N}(0,\, |t-s|)$, we have, for all $p > 0$, $\mathbb{E}[\,|B_t - B_s|^p] = C_p\,(t-s)^{p/2}$, where $C_p := \mathbb{E}[\,|\mathscr{N}(0,1)|^p] < \infty$. It suffices to take p sufficiently large such that $\frac{1}{2} - \varepsilon < \frac{(p/2)-1}{p}$ to see that B admits a modification whose trajectories are locally Hölder continuous for exponent $\frac{1}{2} - \varepsilon$. □

Remark 4.21. Among Levy processes (*i.e.* processes with stationary and independent increments), linear transforms of Brownian motions are the only random processes with continuous trajectories: any Levy process X starting from 0 and with continuous paths can be written in the form

$$X_t = \sigma B_t + \mu t \qquad \text{for all } t$$

where B is a standard Brownian motion and μ is called the drift of the process.

It is natural to ask whether the trajectories of Brownian motion can be locally Hölder continuous for exponent $\frac{1}{2}$. The answer is negative, which we will see later on. For the moment, we prove the following result which will give us an interesting corollary.

Proposition 4.22. *For any $\gamma > 1/2$, we have*

$$\mathbb{P}\Big[\forall\, t \geq 0 : \limsup_{h\to 0+} \frac{|B_{t+h} - B_t|}{h^\gamma} = \infty\Big] = 1\,.$$

Remark 4.23. It is possible to strengthen Proposition 4.22. As a matter of fact, Dvoretzky (1963) proved the existence of a $c > 0$ such that a.s.,

$$\forall\, t \geq 0,\ \limsup_{h\to 0+} \frac{|B_{t+h} - B_t|}{\sqrt{h}} \geq c.$$

Proof of Proposition 4.22. Let $\gamma > 1/2$. Since

$$\Big\{\exists t \geq 0 : \limsup_{h\to 0+} \frac{|B_{t+h} - B_t|}{h^\gamma} < \infty\Big\} \subset \bigcup_{m=1}^\infty \bigcup_{\ell=1}^\infty \bigcup_{k=1}^\infty \Big\{\exists t \in [0, m] : |B_{t+h} - B_t| \leq \ell h^\gamma,\ \forall h \in \big(0, \frac{1}{k}\big]\ \Big\}$$

it suffices to prove that for $m \geq 1$, $\ell \geq 1$ and $\delta > 0$,

$$\mathbb{P}\Big\{\ \exists\, t \in [0, m] : |B_{t+h} - B_t| \leq \ell\, h^\gamma,\ \forall\, h \in (0, \delta]\ \Big\} = 0\,.$$

Consider $A_{i,n} := \{\exists\, t \in [\frac{i}{n}, \frac{i+1}{n}] : |B_{t+h} - B_t| \leq \ell\, h^\gamma, \forall\, h \in (0, \delta]\ \}$. It suffices to check that for all $m \geq 1$, $\ell \geq 1$ and $\delta > 0$, we have $\sum_{i=0}^{nm-1} \mathbb{P}(A_{i,n}) \to 0$, $n \to \infty$.

Let $K > 2$ be an integer with $(K-2)(\gamma - \frac{1}{2}) > 1$. Let $n > n_0 := \lfloor K/\delta \rfloor$. If $\omega \in A_{i,n}$, and let t be as in the definition of $A_{i,n}$ (attention: t depends on ω), then $|B_{\frac{i+j}{n}} - B_t| \leq \ell(\frac{i+j}{n} - t)^\gamma \leq \ell\,(\frac{j}{n})^\gamma$ as long as $0 < \frac{i+j}{n} - t \leq \delta$ (a fortiori, if $2 \leq j \leq K$) ; this implies $|B_{\frac{i+j}{n}} - B_{\frac{i+j-1}{n}}| \leq 2\ell\,(\frac{K}{n})^\gamma$ for $3 \leq j \leq K$. Accordingly,

$$A_{i,n} \subset \bigcap_{j=3}^K \Big\{|B_{\frac{i+j}{n}} - B_{\frac{i+j-1}{n}}| \leq 2\ell\,(\frac{K}{n})^\gamma\Big\}.$$

The events on the right-hand side being independent, we obtain:

$$\mathbb{P}(A_{i,n}) \leq \prod_{j=3}^K \mathbb{P}\Big\{|B_{\frac{i+j}{n}} - B_{\frac{i+j-1}{n}}| \leq 2\ell\,(\frac{K}{n})^\gamma\Big\}.$$

Since $B_{\frac{i+j}{n}} - B_{\frac{i+j-1}{n}}$ is Gaussian $\mathscr{N}(0, \frac{1}{n})$, and since[2] $\mathbb{P}(|\mathscr{N}(0, 1)| < x) \leq (\frac{2}{\pi})^{1/2}\, x, \forall x > 0$, we have $\mathbb{P}(A_{i,n}) \leq \prod_{j=3}^K (\frac{2}{\pi})^{1/2}\, \frac{2\ell\, K^\gamma}{n^{\gamma-1/2}} = \frac{c}{n^{(\gamma-1/2)(K-2)}}$, where $c := [(\frac{2}{\pi})^{1/2}\, 2\ell\, K^\gamma]^{K-2}$. Consequently, $\sum_{i=0}^{nm-1} \mathbb{P}(A_{i,n}) \leq m\, \frac{c}{n^{(\gamma-1/2)(K-2)-1}} \to 0$, $n \to \infty$, as $(K-2)(\gamma - \frac{1}{2}) > 1$. □

[2]The density of $\mathscr{N}(0, 1)$ is bounded by $\frac{1}{(2\pi)^{1/2}}$.

Corollary 4.24 ((Paley, Wiener and Zygmund 1933)). *Almost surely, the function $t \mapsto B_t$ is nowhere differentiable.*

Given an interval $[a, b]$, a function $f : [a, b] \to \mathbb{R}$ is said to have *bounded variation* if

$$V(f) := \sup_{a=t_0<\ldots t_n=b} \sum_{i=1}^{n} |f(t_i) - f(t_{i-1})| < \infty$$

It is not difficult to show that a function f has bounded variation if and only it can be decomposed as $f = g - h$ where g and h are non-decreasing functions (exercise). Since a non-decreasing function is differentiable almost everywhere, the same also holds true for any function with bounded variation. Thus, we deduce from Corollary 4.24 that

Corollary 4.25. *Almost surely, $t \mapsto B_t$ is not of finite variation on any interval $[a, b]$.*

We will give a refinement of this result later on and prove that Brownian motion has finite (non-zero) *quadratic* variation.

4.4. The canonical process and the Wiener measure

In the theory of Markov chain, it is useful to consider the "canonical construction" of the chain given by the process of canonical projections from the space of sequences E^N, embedded into a probability space with the σ-field generated by the cylinder sets. In this way, a Markov chain can be though as a "random variable taking value in the set of sequence E^N". We now do the same thing with Brownian motion which can be seen as random variable taking value in the set of continuous functions.

Consider $C(\mathbb{R}_+, \mathbb{R})$, the space of all real-valued functions on $\mathbb{R}_+$, endowed with the topology of uniform convergence on compacts:

$$d(\mathrm{w}, \widetilde{\mathrm{w}}) = \sum_{n=1}^{\infty} \frac{d_n(\mathrm{w}, \widetilde{\mathrm{w}}) \wedge 1}{2^n},$$

where $d_n(\mathrm{w}, \widetilde{\mathrm{w}}) := \sup_{t\in[0,n]} |\mathrm{w}(t) - \widetilde{\mathrm{w}}(t)|$. Then $(C(\mathbb{R}_+, \mathbb{R}), d)$ is a complete and separable metric space. Let $(X_t,\, t \geq 0)$ denote the "canonical" process of projections:

$$X_t(\mathrm{w}) := \mathrm{w}(t), \qquad \mathrm{w} \in C(\mathbb{R}_+, \mathbb{R}).$$

The next result identifies the σ-field $\sigma(X_t,\, t \geq 0)$ generated by these projections (i.e., the smallest σ-field making all X_t measurable) with the Borel σ-field $\mathscr{C}(\mathbb{R}_+,\, \mathbb{R})$ (i.e., the smallest σ-field containing the topology induced by d). It solves measurability questions regarding continuous random processes seen as "random function".

Lemma 4.26. *We have* $\sigma(X_t,\, t\geq 0) = \mathscr{C}(\mathbb{R}_+,\, \mathbb{R})$.

Proof. For every t, the projection $\mathrm{w} \to X_t(\mathrm{w})$ is continuous hence measurable for $\mathscr{C}(\mathbb{R}_+,\, \mathbb{R})$. This shows that $\sigma(X_t) \subset \mathscr{C}(\mathbb{R}_+,\, \mathbb{R})$ and therefore $\sigma(X_t,\, t \geq 0) \subset \mathscr{C}(\mathbb{R}_+,\, \mathbb{R})$.

Conversely, we can write $d_n(\mathrm{w}, \widetilde{\mathrm{w}}) = \sup_{t\in[0,n]\cap\mathbb{Q}} |\mathrm{w}(t) - \widetilde{\mathrm{w}}(t)|$ which shows that $d_n(\mathrm{w}, \cdot)$ is $\sigma(X_t,\, t \geq 0)$-measurable and so is $d(\mathrm{w}, \cdot)$ as well. Let F be a closed subset of $C(\mathbb{R}_+, \mathbb{R})$ and let (w_n) be a sequence that is dense in F (which exists since $C(\mathbb{R}_+, \mathbb{R})$ is separable). Then,

$$F = \{\mathrm{w} \in C(\mathbb{R}_+, \mathbb{R}) \,:\, d(F, \mathrm{w}) = 0\} = \{\mathrm{w \in C(\mathbb{R}_+, \mathbb{R}) \,:\, d(w_n, w) = 0}\}$$

which shows that F is an element of $\sigma(X_t,\, t \geq 0)$. Hence $\mathscr{C}(\mathbb{R}_+,\, \mathbb{R}) \subset \sigma(X_t,\, t \geq 0)$. □

Corollary 4.27. *A mapping* $\varphi : (\Omega, \mathcal{F}) \mapsto (C(\mathbb{R}_+, \mathbb{R}), \mathscr{C}(\mathbb{R}_+,\, \mathbb{R}))$ *is measurable if and only if* $X_t \circ \varphi$ *is measurable for every* t.

Let $Z = (Z_t,\, t \in \mathbb{R}_+)$ be a continuous process defined on $(\Omega,\, \mathscr{F},\, \mathbb{P})$. Consider the mapping φ

$$\begin{array}{ccc} \Omega & \longrightarrow & C(\mathbb{R}_+, \mathbb{R}) \\ \omega & \longmapsto & \varphi(\omega) = (t \mapsto Z_t(\omega)), \end{array}$$

which is measurable according to the previous corollary. We call the **law** (or: distribution) of Z, the image-measure of $\mathbb{P}$ by φ. By the π-λ theorem, the law of Z is determined by the finite-dimensional distributions $(Z_{t_1}, \cdots, Z_{t_n})$: indeed, two measures on $C(\mathbb{R}_+, \mathbb{R})$ are identical if they attribute same value to sets of type $(X_{t_1}(\mathrm{w}), \cdots, X_{t_n}(\mathrm{w})) \in A$, $A \in \mathscr{B}(\mathbb{R}^n)$ (Borel σ-field of $\mathbb{R}^n$).

In the special case where Z is Brownian motion, this particular image-measure of Z is denoted by $\mathbb{W}$. It is a probability measure on $C(\mathbb{R}_+, \mathbb{R})$ such that $\mathbb{W}\{\mathrm{w} :\, \mathrm{w}(0) = 0\} = 1$, and that for all $n \geq 1$, $0 = t_0 < t_1 < t_2 < \cdots < t_n$ and $A \in \mathscr{B}(\mathbb{R}^n)$,

$$\begin{aligned} &\mathbb{W}\Big\{\mathrm{w} :\, (X_{t_1}(\mathrm{w}), \cdots, X_{t_n}(\mathrm{w})) \in A\Big\} \\ = &\int_A \frac{1}{(2\pi)^{n/2}\sqrt{t_1(t_2 - t_1)\cdots(t_n - t_{n-1})}} \exp\Big(-\frac{1}{2}\sum_{k=1}^{n} \frac{(x_k - x_{k-1})^2}{t_k - t_{k-1}}\Big)\, \mathrm{d}x_1 \cdots \mathrm{d}x_n\,, \end{aligned}$$

with $x_0 := 0$, because the integrand is the density function of the Gaussian random vector $(Z_{t_1},\, Z_{t_2},\, \cdots,\, Z_{t_n})$. This formula characterizes the probability measure $\mathbb{W}$, and does not depend on the choice of Brownian motion in the construction. We call $\mathbb{W}$ the **Wiener measure**, and the process of projections $(X_t,\, t \geq 0)$ is the **canonical process** (of Brownian motion). Summarizing, we have

Theorem 4.28. *There exists a unique probability measure (which is the Wiener measure) on* $C(\mathbb{R}_+, \mathbb{R})$ *under which the process of projections* $(X_t,\, t \geq 0)$ *is a Brownian motion.*

The canonical process is useful for answering measurability questions about functionals of Brownian motion. For example, if B is a Brownian motion, then, the following quantities are clearly random variables: $\sup_{0\leq s\leq t} B_s$, $\int_0^t B_s^2 \,\mathrm{d}s$, $\inf\{s \leq t : B_s = B_t\}, \ldots$

Let $x \in \mathbb{R}$. Let $\mathbb{W}_x$ be the image-measure of $\mathbb{W}$ by the mapping $\mathrm{w} \mapsto \mathrm{w} + x$. Clearly, $\mathbb{W}_x\{\mathrm{w} : \mathrm{w}(0) = x\} = 1$. The process of projections $(X_t,\, t \geq 0)$ under $\mathbb{W}_x$ is called Brownian motion starting at $X_0 = x$. It is a Lévy process whose trajectories are a.s. continuous, $X_0 = x$, a.s., such that $\forall\, t \geq s$, $X_t - X_s$ is Gaussian $\mathscr{N}(0,\, t-s)$. It coincides with the notion of Brownian motion starting from x in Remark 4.13.

Chapter 5

The Markov property for Brownian motion

In the previous chapter, we mainly studied Brownian motion as a Gaussian process. We now study the Markov property of Brownian motion we will use it to obtain basic results concerning the fluctuations of the trajectories. We leave out the martingale property which will be studied in detail in the next chapter.

5.1. Elementary properties

In everything that follows, $B = (B_t,\, t \geq 0)$ denotes a standard Brownian motion.

Proposition 5.1. *The following processes are Brownian motions:*

1. $X_t = -B_t$ ***(symmetry)****.*
2. $X_t = tB_{1/t}$*,* $X_0 = 0$ ***(time inversion)****.*
3. $X_t = \frac{1}{a^{1/2}} B_{at}$ *for* $a > 0$ *fixed* ***(scaling)****.*
4. $X_t = B_T - B_{T-t}$*,* $t \in [0, T]$ *for* $T > 0$ *fixed* ***(time reversal)****.*

Proof. It suffices to check, for each of the processes above, that X is a centered Gaussian process with covariance $s \wedge t$. Only 2. needs some special care because the trajectories are not necessarily continuous at 0. However, this does not cause any trouble because X is, according to Kolmogorov's criterion, indistinguishable to Brownian motion. □

Example 5.2 (**Brownian bridge**)**.** Define $b_t = B_t - tB_1$, $t \in [0, 1]$. It is a centered Gaussian process with a.s. continuous trajectories and with covariance $(s \wedge t) - st$. We call b a Brownian bridge. The following properties hold true

1. The process $(b_t,\, t \in [0,\, 1])$ is independent of the random variable B_1.

2. If b is a Brownian bridge, then so is $(b_{1-t},\, t \in [0,\, 1])$.

3. If b is a Brownian bridge, then $B_t = (1+t)b_{t/(1+t)}$, $t \geq 0$, is Brownian motion. Note that $b_t = (1-t)B_{t/(1-t)}$.

Example 5.3 (Law of large numbers). By continuity of the trajectories of Brownian motion, we have $\lim_{t\to 0+} B_t = 0$, a.s., which, by time inversion, leads to:

$$\lim_{t\to\infty} \frac{B_t}{t} = 0, \qquad \text{a.s.}$$

Notice that the weaker statement $\frac{B_n}{n} \underset{n\to\infty}{\to} 0$ a.s. for integer n's is a direct application of the usual strong law of large number (for centered Gaussian variables).

5.2. The simple Markov property

Recall that we are working on a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Similarly to the discrete time setting, we call **filtration** a family $(\mathscr{F}_t, t \geq 0)$ or sub σ-fields of $\mathscr{F}$ such that

$$\mathscr{F}_s \subset \mathscr{F}_t \qquad \text{for any } s \leq t.$$

The process $B = (B_t,\, t \geq 0)$ is said to be **adapted** to the filtration $(\mathscr{F}_t, t \geq 0)$ if B_t is $\mathscr{F}_t$-measurable for all $t \geq 0$. Of course, B is adapted to its canonical filtration $\sigma(B_s\ s \leq t)$. However, in the continuous-time setting, it is useful to consider the **completed filtration** obtained by making $\mathscr{F}_t$ complete for all $t \geq 0$ *i.e.*[1].

$$\mathscr{F}_t := \sigma(B_s,\, s \leq t) \vee \mathcal{N}$$

where $\mathcal{N}$ represents the null sets of $\mathscr{F}$ for $\mathbb{P}$. Of course, B is still $(\mathscr{F}_t)$-adapted. We also define

$$\mathscr{F}_\infty := \bigvee_{t\geq 0} \mathscr{F}_t.$$

Theorem 5.4 (Simple Markov property). *Let $s \geq 0$. The process $(\widetilde{B}_t := B_{t+s} - B_s)_{t\geq 0}$ is a Brownian motion independent of $\mathscr{F}_s$.*

[1]The notation $\mathcal{F} \vee \mathcal{G}$ in the equation below denotes the smallest σ-field containing both $\mathcal{F}$ and $\mathcal{G}$

Proof. It is clear that $\widetilde{B}$ is a centered Gaussian process with a.s. continuous trajectories and with covariance

$$\begin{aligned}\mathbb{E}(\widetilde{B}_t\widetilde{B}_{t'}) &= \mathbb{E}(B_{t+s}B_{t'+s}) - \mathbb{E}(B_{t+s}B_s) - \mathbb{E}(B_{t'+s}B_s) + \mathbb{E}(B_sB_s)\\ &= (t+s)\wedge(t'+s) - s - s + s = t\wedge t'.\end{aligned}$$

Thus it is Brownian motion. We prove that it is independent of $\mathscr{F}_s$. It suffices to show that for $0 \le t_1 < \cdots < t_n$ and $0 < s_1 < \cdots < s_m \le s$, the random vectors $(\widetilde{B}_{t_1}, \cdots, \widetilde{B}_{t_n})$ and $(B_{s_1}, \cdots, B_{s_m})$ are independent. Since $\mathrm{Cov}(\widetilde{B}_{t_i}, B_{s_j}) = \mathbb{E}[(B_{s+t_i} - B_s)B_{s_j}] = 0$ (because $s \ge s_j$), and since $(\widetilde{B}_{t_1}, \cdots, \widetilde{B}_{t_n}, B_{s_1}, \cdots, B_{s_m})$ is a Gaussian random vector, we see that $(\widetilde{B}_{t_1}, \cdots, \widetilde{B}_{t_n})$ and $(B_{s_1}, \cdots, B_{s_m})$ are independent. □

Given the filtration $(\mathscr{F}_t)_{t\ge0}$, we define another filtration $(\mathscr{F}_{t^+})_{t\ge0}$ by setting

$$\mathscr{F}_{t^+} := \bigcap_{s>t} \mathscr{F}_s.$$

Remark 5.5.

1. We have $\mathscr{F}_t \subset \mathscr{F}_{t^+}$. However, $\mathscr{F}_{t^+}$ gives access, not only to information about the process at time t, but also enables an infinitesimally small peek into the future. Thus, in general $\mathscr{F}_{t^+}$ may be strictly larger than $\mathscr{F}_t$. For example, the event $\{B$ has a right derivative at $t\}$ is in $\mathscr{F}_{t^+}$ but not necessarily in $\mathscr{F}_{t^+}$ (although it is the case for Brownian motion because this event has null probability and our filtrations are complete).

2. A filtration that satisfies $\mathscr{F}_t = \mathscr{F}_{t^+}$ for all t is said to be **right-continuous**. Of course, the filtration $(\mathscr{F}_{t^+})$ is right continuous since $\mathscr{F}_{t^{++}} = \mathscr{F}_{t^+}$. A filtration that is both right-continuous and complete is said to **satisfy the usual conditions**.

We have the seemingly strengthening of the Markov property.

Proposition 5.6. *Let $s \ge 0$. The process $(\widetilde{B}_t := B_{t+s} - B_s,)_{t\ge0}$ is independent of $\mathscr{F}_{s^+}$.*

Proof of Proposition 5.6. It suffices to check that for $A \in \mathscr{F}_{s^+}$ and $0 \le t_1 < t_2 < \cdots < t_n$ and any continuous and bounded function $F : \mathbb{R}^n \to \mathbb{R}$, we have

$$\mathbb{E}\Big[\mathbf{1}_A\, F(\widetilde{B}_{t_1}, \cdots, \widetilde{B}_{t_n})\Big] = \mathbb{P}(A)\,\mathbb{E}\Big[F(B_{t_1}, \cdots, B_{t_n})\Big]. \tag{5.1}$$

Let $\varepsilon > 0$. By the Markov property, $t \mapsto B_{t+s+\varepsilon} - B_{s+\varepsilon}$ is independent of $\mathscr{F}_{s+\varepsilon}$, and is, a fortiori, independent of $\mathscr{F}_{s^+}$. Hence

$$\mathbb{E}\Big[\mathbf{1}_A\, F(B_{t_1+s+\varepsilon} - B_{s+\varepsilon}, \cdots, B_{t_n+s+\varepsilon} - B_{s+\varepsilon})\Big] = \mathbb{P}(A)\,\mathbb{E}\Big[F(B_{t_1}, \cdots, B_{t_n})\Big].$$

Letting $\varepsilon \to 0$, and by continuity of trajectories and the dominated convergence theorem, we obtain (5.1). □

Theorem 5.7 (Blumenthal 0–1 law). *The σ-field $\mathscr{F}_{0^+}$ is trivial, in the sense that*

$$\forall A \in \mathscr{F}_{0^+},\ \mathbb{P}(A) \in \{0, 1\}$$

Proof. By Proposition 5.6, $\mathscr{F}_{0^+}$ is independent of $\sigma(B_t,\, t \geq 0)$, and thus also independent of its completion $\mathscr{F}_\infty$. Let $A \in \mathscr{F}_{0^+} = \bigcap_{s>0} \mathscr{F}_s$. Then, $A \in \mathscr{F}_\infty$ so it is independent of itself. Thus, $\mathbb{P}(A) = \mathbb{P}(A \cap A) = \mathbb{P}(A)\mathbb{P}(A)$ which implies $\mathbb{P}(A) \in \{0, 1\}$. □

Remark 5.8. Since $\mathscr{F}_0$ is complete, Blumenthal's 0–1 law state that $\mathscr{F}_{0^+} = \mathscr{F}_0 = \sigma(\mathcal{N})$. More generally, it follows from the Markov property that $\mathscr{F}_{t^+} = \mathscr{F}_t$ for all $t \geq 0$ hence Theorem 5.4 and Proposition 5.6 are, in fact, identical. In other words: **the completed filtration $(\mathscr{F}_t)$ of Brownian motion satisfies the usual conditions (*i.e.* it is right-continuous)**.

Blumenthal 0–1 law allows us to derive important information about the typical trajectories of Brownian motion.

- Let $\tau^+ := \inf\{t > 0 : B_t > 0\}$. Then we have $\tau^+ = 0$, a.s. To prove this, we note that,

$$\{\tau^+ = 0\} = \bigcap_n \{ \sup_{0 \leq u \leq 1/n} B_u > 0\} \in \mathscr{F}_{0^+}.$$

Therefore, we have $\mathbb{P}(\tau^+ = 0) \in \{0, 1\}$. On the other hand, for $t > 0$, $\mathbb{P}(\tau^+ \leq t) \geq \mathbb{P}(B_t > 0) = \frac{1}{2}$ so that $\mathbb{P}(\tau = 0) = \lim_{t \to 0^+} \mathbb{P}(\tau^+ \leq t) \geq \frac{1}{2}$. Necessarily, $\mathbb{P}(\tau^+ = 0) = 1$. By symmetry, we also have $\tau^- := \inf\{t > 0 : B_t > 0\} = 0$ a.s.

Fact. Brownian motion oscillate at 0: it takes positive and negative values in any neighborhood of the origin. By continuity of the trajectories, the set $\{s : B_s = 0\}$ has an accumulation point at 0.

Using the Markov property, it follows that, Brownian motion oscillate around its position at time t for any fixed $t > 0$.

- Using inversion of time (item (ii) of Proposition 5.1), we find that $\{t > 0 : B_t = 0\}$ is a.s. unbounded toward infinity. Let $x > 0$, from the scaling property of Brownian motion, we get that

$$\mathbb{P}\Big(\sup_{s \in [0,\, t]} B_s > x \Big) = \mathbb{P}\Big(\sup_{s \in [0,\, 1]} B_s > \frac{x}{t^{1/2}} \Big) \underset{t \to \infty}{\longrightarrow} \mathbb{P}\Big(\sup_{s \in [0,\, 1]} B_s > 0 \Big) \geq \mathbb{P}(\tau^+ = 0) = 1$$

This means that $\sup_{s\geq 0} B_s = +\infty$, a.s. By symmetry, $\inf_{s\geq 0} B_s = -\infty$, a.s. (from which we can recover that $\{t > 0 : B_t = 0\}$ is a.s. unbounded).

Fact. Brownian motion is recurrent: Almost surely, for any $x \in \mathbb{R}$ and any $t \geq 0$, there exists $s > t$ with $B_s = x$.

- Let (t_i) be a decreasing sequence converging to 0. The events

$$A_+ := \{B_{t_i} > 0 \text{ for infinitely many } i\}, \qquad A_+ := \{B_{t_i} < 0 \text{ for infinitely many } i\}$$

both belong to $\mathscr{F}_0^+$ and have the same probability by symmetry. Furthermore $\mathbb{P}(A_+ \cup A_-) = 1$ because $\mathbb{P}(B_{t_i} = 0) = 0$ for all t_i. We conclude that $\mathbb{P}(A_+) = \mathbb{P}(A_-) = 1$. By inversion of time, the same result holds for the sequence $(1/t_i)$.

Fact. Almost surely, Brownian motion take both positive and negative values along any deterministic sequence that is either unbounded or has an accumulation point a 0.

- Fix a constant $K > 0$, and let $E_n := \{\sqrt{n}\, B_{1/n} > K\}$. Using Fatou's Lemma, we have $\mathbb{P}(\limsup_n E_n) \geq \limsup_{n\to\infty} \mathbb{P}(E_n) = \mathbb{P}(B_1 > K) > 0$, so by Blumenthal's 0–1 law, $\mathbb{P}(\limsup_n E_n) = 1$. This holds for any K so we obtain that

$$\limsup_{t\to 0^+} \frac{B_t}{t^{1/2}} = \infty, \qquad \liminf_{t\to 0^+} \frac{B_t}{t^{1/2}} = -\infty, \qquad \text{a.s.}$$

(this result shows, in particular that the trajectories of B are not $\frac{1}{2}$-Hölder). By inversion of time, a similar result hold holds true at infinity:

$$\limsup_{t\to \infty} \frac{B_t}{t^{1/2}} = \infty, \qquad \liminf_{t\to \infty} \frac{B_t}{t^{1/2}} = -\infty, \qquad \text{a.s.}$$

We will obtain more accurate results at the end of this chapter when we prove the *law of the iterated logarithm* for Brownian motion.

5.3. The strong Markov property

Just as it is the case for Markov chains, we can strengthen the Markov property stated in the previous section so that it can be used with random times. To do so, we need to introduce the notion of continuous-time stopping times which, once again, is similar to the definition in the discrete setting.

Definition 5.9. *A random variable $T : \Omega \to \mathbb{R}_+ \cup \{\infty\}$ is called a stopping time with respect to a filtration $(\mathscr{F}_t)_{\geq 0}$ if*

$$\{T \leq t\} \in \mathscr{F}_t \qquad \text{for any } t \geq 0.$$

The σ-field $\mathscr{F}_T$ of "past events of T" is defined as

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

Remark 5.10.

1. It is easily checked that $\mathscr{F}_T$ defined as above is indeed a σ-field.
2. $\{T \leq t\} \in \mathscr{F}_t$ means that: *If T occurs before time t, then it must be known by time t.* The σ-field $\mathscr{F}_T$ contains all the events which are known at the time the stopping occurs.
3. The constant time $T = t$ is a stopping time.
4. Let $a \in \mathbb{R}$. The return time $T_a := \inf\{t > 0 : B_t = a\}$ is a stopping time since $\{T_a \leq t\} = \{\inf_{s \in]0,t] \cap \mathbb{Q}} |B_s - a| = 0\} \in \mathscr{F}_t$.
5. If S, T are stopping times, then $S \wedge T$, $S \vee T$ and $S + T$ are also stopping times.
6. If $S \leq T$, then $\mathscr{F}_S \subset \mathscr{F}_T$.
7. The random variables T and $B_T \, \mathbf{1}_{\{T<\infty\}}$ are $\mathscr{F}_T$-measurable. Indeed, for the latter, it suffices to observe that a.s.,

$$B_T \, \mathbf{1}_{\{T<\infty\}} = \lim_{n\to\infty} \sum_{i=0}^{\infty} \mathbf{1}_{\{\frac{i}{2^n} < T \leq \frac{i+1}{2^n}\}} B_{\frac{i}{2^n}},$$

and recall that $B_s \, \mathbf{1}_{\{s<T\}}$ and $\mathbf{1}_{\{T \leq t\}}$ are $\mathscr{F}_T$-measurable.

Recall the construction of Brownian motion as the canonical projections from the set of continuous functions embedded with the Wiener measure describe in Section 4.4. This makes it possible to define a process B together with a collection of probability measures $(\mathbb{P}_x, x \in \mathbb{R})$ such that B under $\mathbb{P}_x$ is a Brownian motion starting from x. In accordance with our previous notation, we write $\mathbb{P} = \mathbb{P}_0$. We also denote by $\mathbb{E}_x$ the expectation under $\mathbb{P}_x$ (with $\mathbb{E} = \mathbb{E}_0$).

Theorem 5.11 (Strong Markov property). *Let T be a stopping time. Let $x \in \mathbb{R}$. Under $\mathbb{P}_x$, conditionally on $\{T < \infty\}$, the process $\widetilde{B} := (B_{T+t} - B_T, \, t \geq 0)$ is Brownian motion starting at 0, independent of $\mathscr{F}_T$.*[2]

[2] We have only defined the process $\widetilde{B}$ on $\{T < \infty\}$. Its values outside of this event do not matter so we can set, for instance $\widetilde{B} := 0$ on $\{T = \infty\}$.

Proof. Suppose first $T < \infty$, $\mathbb{P}_x$-a.s. We are going to prove, for $A \in \mathscr{F}_T$, $0 \le t_1 < \cdots < t_n$ and $F : \mathbb{R}^n \to \mathbb{R}_+$ continuous and bounded, that

$$\mathbb{E}_x\Big[\mathbf{1}_A\, F(\widetilde{B}_{t_1}, \cdots, \widetilde{B}_{t_n})\Big] = \mathbb{P}_x(A)\, \mathbb{E}\Big[F(B_{t_1}, \cdots, B_{t_n})\Big]. \tag{5.2}$$

This will yield that $\widetilde{B}$ is Brownian motion starting at 0 (by taking $A = \Omega$), and that it is independent of $\mathscr{F}_T$. First, observe that, by continuity of F and of the paths of B,

$$\sum_{k=0}^{\infty} \mathbf{1}_{\{\frac{k-1}{2^m} < T \le \frac{k}{2^m}\}} F(B_{\frac{k}{2^m}+t_1} - B_{\frac{k}{2^m}}, \cdots, B_{\frac{k}{2^m}+t_n} - B_{\frac{k}{2^m}})$$

converges a.s. (when $m \to \infty$) to $F(\widetilde{B}_{t_1}, \cdots, \widetilde{B}_{t_n})$. By dominated convergence,

$$\begin{aligned} &\mathbb{E}_x\Big[\mathbf{1}_A\, F(\widetilde{B}_{t_1}, \cdots, \widetilde{B}_{t_n})\Big] \\ =\; &\lim_{m\to\infty} \sum_{k=0}^{\infty} \mathbb{E}_x\Big[\mathbf{1}_A \mathbf{1}_{\{\frac{k-1}{2^m} < T \le \frac{k}{2^m}\}} F(B_{\frac{k}{2^m}+t_1} - B_{\frac{k}{2^m}}, \cdots, B_{\frac{k}{2^m}+t_n} - B_{\frac{k}{2^m}})\Big]. \end{aligned}$$

For each k, $A \cap \{\frac{k-1}{2^m} < T \le \frac{k}{2^m}\} \in \mathscr{F}_{\frac{k}{2^m}}$. By the Markov property,

$$\begin{aligned} \mathbb{E}_x\Big[\mathbf{1}_A\, F(\widetilde{B}_{t_1}, \cdots, \widetilde{B}_{t_n})\Big] &= \lim_{m\to\infty} \sum_{k=0}^{\infty} \mathbb{E}_x\Big[\mathbf{1}_A \mathbf{1}_{\{\frac{k-1}{2^m} < T \le \frac{k}{2^m}\}}\Big]\, \mathbb{E}\Big[F(B_{t_1}, \cdots, B_{t_n})\Big]. \\ &= \mathbb{P}_x(A)\, \mathbb{E}\Big[F(B_{t_1}, \cdots, B_{t_n})\Big], \end{aligned}$$

from which (5.2) follows. When $\mathbb{P}_x(T = \infty) > 0$, the same argument gives

$$\mathbb{E}_x\Big[\mathbf{1}_{A\cap\{T<\infty\}}\, F(\widetilde{B}_{t_1}, \cdots, \widetilde{B}_{t_n})\Big] = \mathbb{P}_x(A \cap \{T < \infty\})\, \mathbb{E}\Big[F(B_{t_1}, \cdots, B_{t_n})\Big],$$

and the desired result follows again. □

Recall the definition of the **return/hitting**[3] **time** to a:

$$T_a := \inf\{t > 0 : \; B_t = a\}. \tag{5.3}$$

We also define the **running supremum process** $(S_t,\; t \ge 0)$ by

$$S_t := \sup_{s\in[0,\,t]} B_s \qquad t \ge 0. \tag{5.4}$$

The processes $(T_a, a \ge 0)$ and $(S_t, t \ge 0)$ are inverse of each other in the sense that

$$\{S_t \le a\} = \{\exists s \le t,\; B_s = a\} = \{T_a \le t\} \qquad \text{for any } a, t \ge 0.$$

An exact formula for the probabilities of the event above can be obtained using the so-called reflection principle which, in turn, is a simple consequence of the strong Markov property.

[3]Since Brownian motion is oscillating, the return and hitting are equal almost surely for any a.

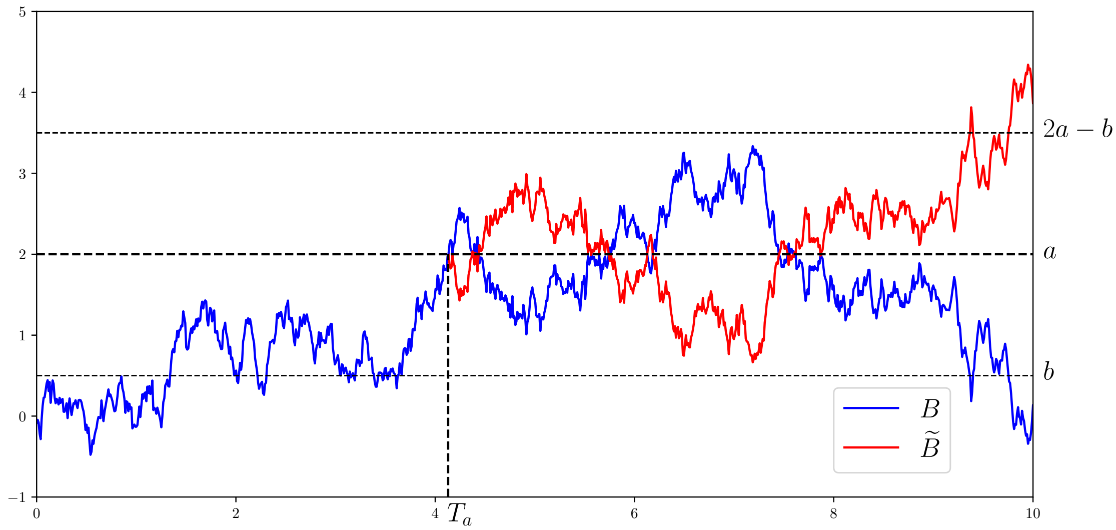


Figure 5.1: Illustration of the reflection principle: the blue Brownian motion is reflected (in red) against the line $y = a$ after time T_a.

Theorem 5.12 (Reflection principle). *We have, for any $a \geq 0$ and any $b \leq a$,*

$$\mathbb{P}(S_t \geq a,\ B_t \leq b) = \mathbb{P}(B_t \geq 2a - b). \tag{5.5}$$

The density of the pair (S_t, B_t) is

$$\mathbb{P}(S_t \in da, B_t \in db) = \frac{2(2a-b)}{\sqrt{2\pi t^3}} \exp\left(-\frac{(2a-b)^2}{2t}\right) \mathbf{1}_{\{a>0,b<a\}} da db.$$

Proof. The idea of the proof it to mirror the trajectory of B after hitting level a and observe that the resulting process is still a Brownian motion (c.f. Figure 5.1 for an illustration). More precisely, recall that $T_a < \infty$ a.s. because Brownian motion is recurrent. We have

$$\mathbb{P}(S_t \geq a,\ B_t \leq b) = \mathbb{P}(T_a \leq t,\ B_t \leq b) = \mathbb{P}(T_a \leq t,\ \widetilde{B}_{t-T_a} \leq b - a),$$

where $\widetilde{B}_s := B_{s+T_a} - B_{T_a} = B_{s+T_a} - a$. So $\mathbb{P}(S_t \geq a,\ B_t \leq b) = \mathbb{P}\{(T_a,\ \widetilde{B}) \in A_t\}$, where $A_t := \{(s,\ \mathrm{w}) \in \mathbb{R}_+ \times C(\mathbb{R}_+,\ \mathbb{R}) :\ s \leq t,\ \mathrm{w}(t-s) \leq b - a\}$ is measurable with respect to $\mathscr{B}(\mathbb{R}_+) \otimes \mathscr{C}(\mathbb{R}_+,\ \mathbb{R})$.

By the strong Markov property, $\widetilde{B}$ is Brownian motion, independent of $\mathscr{F}_{T_a}$, a fortiori of T_a. In particular, $(T_a,\ -\widetilde{B})$ has the same distribution as $(T_a,\ \widetilde{B})$. Therefore, we conclude that

$$\mathbb{P}(S_t \geq a,\ B_t \leq b) = \mathbb{P}(T_a \leq t,\ -\widetilde{B}_{t-T_a} \leq b-a) = \mathbb{P}(T_a \leq t,\ B_t \geq 2a-b) = \mathbb{P}(B_t \geq 2a-b)$$

because $\{B_t \geq 2a-b\} \subset \{T_a \leq t\}$ whenever $b \leq a$. The formula for the density of (B_t, S_t) follows by differentiating both sides of (5.5). □

Corollary 5.13. *1. For each $t > 0$, the random variables S_t and $|B_t|$ have the same law.*

2. For each $a > 0$, the random variables T_a and $(\frac{a}{B_1})^2$ have the same law. The density is given by the formula

$$\mathbb{P}(T_a \in \mathrm{d}t) = \frac{a}{\sqrt{2\pi t^3}} \exp\left(-\frac{a^2}{2t}\right) \mathbf{1}_{t>0}\,\mathrm{d}t \qquad \textit{(Levy's law).} \tag{5.6}$$

Proof. To prove Part 1., we simply observe that

$$\mathbb{P}(S_t \geq a) = \mathbb{P}(S_t \geq a,\, B_t \geq a) + \mathbb{P}(S_t \geq a,\, B_t \leq a) = 2\mathbb{P}(B_t \geq a) = \mathbb{P}(|B_t| \geq a).$$

For Part 2., we combine 1. and the scaling property of Brownian motion to get that

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

This proves that T_a and $(\frac{a}{B_1})^2$ have the same law. The formula for the density follows from that of the Gaussian law. □

Remark 5.14. 1. The identity in law between S_t and $|B_t|$ is valid only for fixed t. The processes $(S_t,\, t \geq 0)$ and $(|B_t|,\, t \geq 0)$ have very different behaviors (for example, the former is non-decreasing while it is certainly not the case for the latter). Still, let us mention a beautiful theorem of Levy which states that the processes $(S_t - B_t, t \geq 0)$ and $(|B_t|,\, t \geq 0)$ have the same law (but the proof is out of the scope of these notes).

2. If T'_a is an independent copy of T_a, it follows the strong Markov property and Brownian scaling that $T_a + T'_a \overset{law}{=} T_{2a} \overset{law}{=} 4T_a$. We say that the Levy distribution is **strictly stable with index $\frac{1}{2}$**. The tail distribution of T_a satisfies

$$\mathbb{P}(T_a > t) \underset{t\to\infty}{\sim} \frac{\sqrt{2}a}{\sqrt{\pi t}}$$

which shows, in particular, that $\mathbb{E}[T_a] = +\infty$ for any $a > 0$.

3. The process $(T_a,\ a \geq 0)$ is Levy process (*i.e.* with independent and stationary increments) which is non-decreasing. Such a process is called a **subordinator**.

5.4. Additional paths properties

Proposition 5.15 (Uniqueness of local extrema). *for any $s < t$, Brownian motion attain its supremum on $[s,t]$ at exactly once a.s.*

Proof. Since Brownian motion is continuous, its supremum is attained on any compact set $[s,t]$ so we must prove that it is attained only once. By applying the simple Markov property at time s, we can assume that $s = 0$. Now, if the maxima is reached (at least) twice on $[0,t]$ with positive probability, then there must exist $r < t$ such that $\mathbb{P}(\sup_{[0,r]} B = \sup_{[r,t]} B) > 0$ because

$$\Big\{\text{maxima of } B \text{ on } [0,t] \text{ reached twice }\Big\} \subset \bigcup_{r\in]0,t[\cap\mathbb{Q}} \Big\{ \sup_{[0,r]} B = \sup_{[r,t]} B\Big\}$$

Now, $\widetilde{B} := (B_{s+r} - B_r,\ s \le t-r)$ and $\widehat{B} := (B_{r-s} - B_r,\ s \le r)$ are independent Brownian motions thanks to the Markov property (combined with time reversal for $\widehat{B}$). But then, using Corollary 5.13 and the independence of $\widetilde{B}$ and $\widehat{B}$, we get

$$\mathbb{P}\Big(\sup_{[0,r]} B = \sup_{[r,t]} B\Big) = \mathbb{P}\Big(\sup_{[0,r]} \widehat{B} = \sup_{[0,t-r]} \widetilde{B}\Big) = \mathbb{P}\Big(|\widehat{B}_r| = |\widetilde{B}_{t-r}|\Big) = 0$$

because the Gaussian law has no atom. Contradiction. □

A random variable Z in $[0,1]$ follows the **arcsine law** if $\mathbb{P}(Z < x) = \frac{2}{\pi}\arcsin\sqrt{x}$ or equivalently, if its density is given by

$$\mathbb{P}(Z \in \,\mathrm{d}x) = \frac{1}{\pi\sqrt{x(1-x)}}\mathbf{1}_{\{x\in]0,1[\}}.$$

This distribution appears naturally in connection functionals of Brownian motion.

Theorem 5.16 (arcsine law for the supremum M). *Let M denote the position where B reaches its maximum over $[0,1]$ (which is well defined thanks to Proposition 5.15). M follows the arcsine laws.*

Theorem 5.17 (arcsine law for the last zero L). *Let L denote the position of the last zero of B on the interval $[0,1]$. L follows the arcsine laws.*

Remark 5.18. 1. There is a third famous arcsine law: the total time spend on the positive half line by B before time 1 (*i.e* $\int_0^1 \mathbf{1}_{\{B_s>0\}}\,\mathrm{d}s$) also follows the arcsine law.

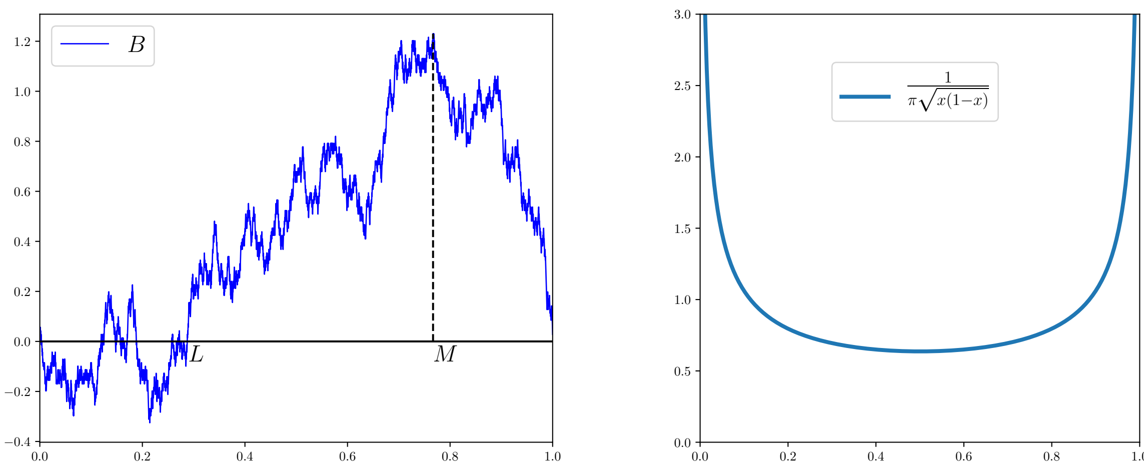


Figure 5.2: Left: last zero L and maximum M of a Brownian motion in $[0,1]$. Right: density of the arcsine law.

2. The arcsine density is minimum at $\frac{1}{2}$ and explodes at 0 and 1. This tells us that, typically (and somewhat surprisingly at first sight), L and M are more likely to be close to 0 or 1 than to the middle of the interval. Also, L and M have symmetric distribution on $[0,1]$ with is not obvious from their definition[4].

Proof of the arcsine law for L. Fix $t \in]0,1[$ and set $\widetilde{B}_s = B_{s+t} - B_t$, which, as usual, is a Brownian motion independent of B_t hence, with obvious notations

$$\begin{aligned}
\mathbb{P}(L < t) &= \mathbb{P}(B_t + \widetilde{B}_s \neq 0 \text{ for all } s \in [0, 1-t]) \\
&= \mathbb{E}\big[\mathbb{P}\big[\widetilde{B}_s \neq -B_t \text{ for all } s \in [0,1-t] \,\big|\, B_t\big]\big] \\
&= \mathbb{E}\big[\mathbb{P}\big[\widetilde{B}_s \neq |B_t| \text{ for all } s \in [0,1-t] \,\big|\, B_t\big]\big] && \text{[symmetry]} \\
&= \mathbb{E}\big[\mathbb{P}\big[\widetilde{T}_{B_t} > 1-t \,\big|\, B_t\big]\big] \\
&= \mathbb{P}\big(B_t^2/\widetilde{B}_1^2 > 1-t\big) && \text{[Corollary 5.13]} \quad (5.7) \\
&= \mathbb{P}\big(B_1^2/\widetilde{B}_1^2 > (1-t)/t\big) && \text{[Brownian scaling]} \\
&= \mathbb{P}\Big(\frac{\widetilde{B}_1^2}{\widetilde{B}_1^2 + B_1^2} < t\Big)
\end{aligned}$$

[4] The symmetry of the law of M can be proved directly using time reversal and symmetry of Brownian motion. What about L ?

Since B_1 and $\widetilde{B}_1$ are independent, the Gaussian vector $(B_1, \widetilde{B}_1)$ has density $\frac{1}{2\pi}\mathrm{e}^{-(x^2+y^2)/2}\,\mathrm{d}x\,\mathrm{d}y$. In polar coordinates $(B_1, \widetilde{B}_1) = (R\cos\Theta, R\sin\Theta)$, this density becomes $\frac{1}{2\pi}\mathrm{e}^{-r^2/2}\mathbf{1}_{\theta\in[0,2\pi[,r>0}\,\mathrm{d}x\,\mathrm{d}y$. In particular, θ is uniform in $[0, 2\pi[$ therefore we conclude that

$$\mathbb{P}(L < t) = \mathbb{P}(\sin^2\Theta < t) = \mathbb{P}(|\sin\Theta| < \sqrt{t}) = \frac{2}{\pi}\arcsin(\sqrt{t}).$$

□

Proof of the arcsine law for M. Fix $t \in]0, 1[$. Let $\widetilde{B}_s := B_{t+s} - B_t$ and $\widehat{B}_s := B_{t-s} - B_t$. Combining the Markov property and time reversal for Brownian motion, we see that $(\widetilde{B}_s,\ s \leq 1-t)$ and $(\widehat{B}_s,\ s \leq t)$ are independent Brownian motions starting from 0, thus

$$\begin{aligned}
\mathbb{P}(M < t) &= \mathbb{P}\big(\sup_{s\in[0,t]} B_s > \sup_{s\in[t,1]} B_s\big) \\
&= \mathbb{P}\big(\sup_{s\in[0,t]} \widehat{B}_s > \sup_{s\in[0,1-t]} \widetilde{B}_s\big) \\
&= \mathbb{P}\big(|\widehat{B}_t| > |\widetilde{B}_{1-t}|\big) && \text{[independence \& Corollary 5.13]} \\
&= \mathbb{P}\Big(|\widehat{B}_1|/|\widetilde{B}_1| > \sqrt{1-t}/\sqrt{t}\Big) && \text{[Brownian scaling]}
\end{aligned}$$

which is the same probability as in (5.7), hence M also follows the arcsine laws. □

Theorem 5.19. *Let $\mathcal{Z} = \{t \geq 0,\ B_t = 0\}$ denote the set of zeros of B. Then $\mathcal{Z}$ is a.s. a perfect set (*i.e *closed and without isolated point) with zero Lebesgue measure.*

Remark 5.20.
1. The usual example of (deterministic) perfect set is the cantor set obtained by trisecting the unit interval recursively while removing at each step the center interval.
2. A perfect set is necessarily uncountable (exercise) so the zeros of Brownian motion are a.s. uncountable yet with null Lebesgue measure.

Proof of Theorem 5.19. The set $\mathcal{Z}$ is closed a.s. since it is the preimage of 0 by B which is continuous. Let $u \geq 0$ and let $\tau_u := \inf\{t \geq u : B_t = 0\}$ denote the first hitting time of 0 after time u. By continuity of Brownian paths, we have $B_{\tau_u} = 0$. Moreover, the random variable τ_u is a stopping time so, by the strong Markov property, $(B_{t+\tau_q} = B_{t+\tau_q} - B_{\tau_q}, t \geq 0)$ is a Brownian motion. Since we already know that the zeros of Brownian motion have an accumulated point at 0, it follows that τ_q is not an isolated zero of B a.s. Thus, we can write

$$\begin{aligned}
\mathbb{P}(\mathcal{Z} \text{ has an isolated zero}) &= \mathbb{P}(\exists (q,r) \in \mathbb{Q}_+^2 \text{ s.t. } \mathcal{Z} \cap]q, r[= \{\tau_q\}) \\
&\leq \sum_{q\in\mathbb{Q}_+} \mathbb{P}(\tau_q \text{ is an isolated zero}) = 0
\end{aligned}$$

Finally, we compute using Fubini's Theorem

$$\mathbb{E}[\text{Lebesgue}(\mathcal{Z})] = \mathbb{E}\left[\int_O^\infty \mathbf{1}_{\{B_s=0\}}\,\mathrm{d}s\right] = \int_0^\infty \mathbb{P}(B_s = 0)\,\mathrm{d}s = 0.$$

which proves that the Lebesgue measure of $\mathcal{Z}$ is a.s. zero. □

Theorem 5.21 (**Law of the iterated logarithm**). *Recall the notation $S_t := \sup_{s\le t} B_s$. We have*

$$\limsup_{t\to\infty} \frac{B_t}{\sqrt{2t\log\log t}} = \limsup_{t\to\infty} \frac{S_t}{\sqrt{2t\log\log t}} = 1 \quad a.s.$$

Proof. The idea of the proof is to sample the Brownian motion at exponential times $t_n := \alpha^n$ and observe that, as $\alpha \to \infty$, the value of B at these times become asymptotically independent which will allow us to use the Borel-Cantelli lemma to conclude.

Upper bound. Let $h(x) := \sqrt{2t\log\log t}$. Let $\varepsilon > 0$ and $\alpha > 1$ an consider the events

$$E_n := \{S_{t_n} \ge (1+\varepsilon)h(t_n)\}.$$

Recalling that S_{t_n} and $\sqrt{t_n}|B_1|$ have the same law and using the Gaussian tail upper bound of Theorem 4.1, we get that

$$\mathbb{P}(E_n) \le 2\mathrm{e}^{-\frac{(1+\varepsilon)^2 h(t_n)^2}{2t_n}} = 2\mathrm{e}^{(1+\varepsilon)^2\log\log\alpha^n} = C(\alpha)n^{-(1+\varepsilon)^2}.$$

Thus, $\sum_n \mathbb{P}(E_n) < \infty$ so the Borel-Cantelli lemmas entails that

$$\limsup_{n\to\infty} \frac{S_{t_n}}{h(t_n)} \le 1+\varepsilon \qquad \text{a.s.}$$

Now, if $t_{n-1} \le t < t_n$, because h and S are non-decreasing, we have

$$\frac{S_t}{h(t)} \le \frac{S_{t_n}}{h(t_{n-1})} \le \frac{S_{t_n}}{h(t_n)}\frac{h(t_n)}{h(t_{n-1})}.$$

Since $h(t_n)/h(t_{n-1}) \underset{n\to\infty}{\to} \sqrt{\alpha}$, we get, taking the limsup in the inequality above,

$$\limsup_{t\to\infty} \frac{S_t}{h(t)} \le (1+\varepsilon)\sqrt{\alpha} \qquad \text{a.s.}$$

and the desired upper bound for S_t (and B_t) follow by taking $\varepsilon \to 0$ and $\alpha \to 1$ (along a countable sequence so that we can swap the limit with the a.s.). By symmetry, we obtain a lower bound for the liminf of $-B_n$. Combining those results yields

$$\limsup_{t\to\infty} \frac{|B_t|}{h(t)} \le 1 \qquad \text{a.s.} \tag{5.8}$$

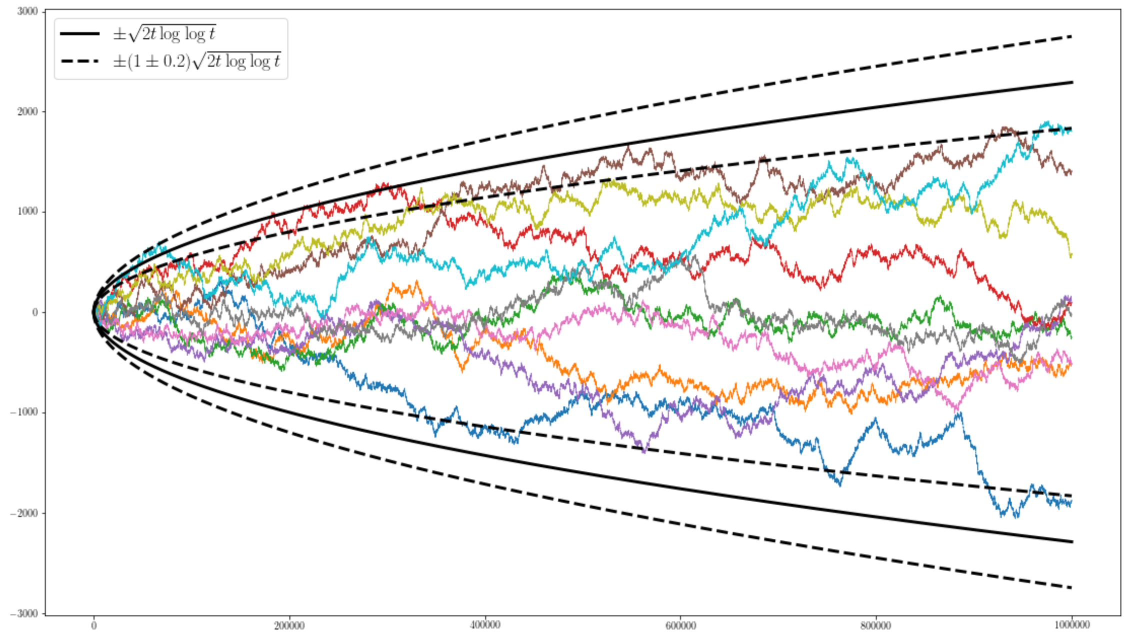


Figure 5.3: Illustration of the law of the iterated logarithm. Simulation of 10 independent Brownian motions up to time 10^6.

which we will need to prove the lower bound.

Lower bound. Consider now the events

$$F_n := \{B_{t_n} - B_{t_{n-1}} > (1-\varepsilon)h(t_n)\}$$

Since Brownian motion has independent increments, the events F_n are independent. Let us compute:

$$\begin{aligned}\mathbb{P}(F_n) &= \mathbb{P}(B_{t_n} - B_{t_{n-1}} > (1-\varepsilon)h(t_n)) \\ &= \mathbb{P}(\sqrt{t_n - t_{n-1}}B_1 > (1-\varepsilon)h(t_n)) \\ &= \mathbb{P}\Big(B_1 > \frac{\sqrt{2\alpha}(1-\varepsilon)}{\sqrt{\alpha-1}}\sqrt{\log\log\alpha^n}\Big)\end{aligned}$$

Setting $\delta := \frac{\sqrt{\alpha}(1-\varepsilon)}{\sqrt{\alpha-1}}$ and using the lower bound on the Gaussian tail (Theorem 4.1), it is easy to check that $\mathbb{P}(F_n) \geq n^{-\delta^2+o(1)}$. Thus, for any $\varepsilon > 0$ fixed, we have $\sum_n \mathbb{P}(F_n) = \infty$ provided α is chosen large enough. Since the events F_n are independent, the second Borel-Cantelli Lemma then asserts that F_n occurs for an infinite number of indexes n a.s. Moreover, on

F_n, we have

$$\frac{B_{t^n}}{h(t_n)} \geq (1-\varepsilon) - \frac{|B_{t^{n-1}}|}{h(t_n)} = (1-\varepsilon) - \frac{|B_{t^{n-1}}|}{h(t_{n-1})}\frac{h(t_{n-1})}{h(t_n)}.$$

We have $h(t_{n-1})/h(t_n) \to 1/\sqrt{\alpha}$. This limit together with (5.8) shows that

$$\limsup_{t\to\infty} \frac{B_t}{h(t)} \geq \limsup_{n\to\infty} \frac{B_{t^n}}{h(t_n)} \geq (1-\varepsilon) - \frac{1}{\sqrt{\alpha}} \qquad \text{a.s.}$$

We conclude the proof by letting first $\alpha \to \infty$ then $\varepsilon \to 0$. □

Remark 5.22. 1. By symmetry of Brownian motion, the previous theorem shows that

$$\liminf_{t\to\infty} \frac{B_t}{\sqrt{2t\log\log t}} = \liminf_{t\to\infty} \frac{\inf_{s\leq t} B_s}{\sqrt{2t\log\log t}} = -1 \quad \text{a.s.}$$

which, combined with the result for the lim sup directly imply that

$$\limsup_{t\to\infty} \frac{|B_t|}{\sqrt{2t\log\log t}} = \limsup_{t\to\infty} \frac{\sup_{s\leq t} |B_s|}{\sqrt{2t\log\log t}} = 1 \quad \text{a.s.}$$

2. Using time inversion, we also obtain the law of the iterated logarithm for $t \to 0^+$. For instance, we have

$$\limsup_{t\to 0^+} \frac{B_t}{\sqrt{2t\log\log\frac{1}{t}}} = 1 \quad \text{a.s.}$$

3. The law of the iterated logarithm is not as precise as it looks like. For example, it does not tell us about the probability that the curve $(2t\log\log t)^{1/2}$ is crossed by Brownian motion at arbitrarily large time a.s. This question was answered by the following integral criterion of Kolmogorov (also referred to sometimes as the Erdős–Feller–Kolmogorov–Petrovski or EFKP integral test): if $h : \mathbb{R} \to \mathbb{R}_+$ is non-decreasing, then

$$\mathbb{P}\Big(B_t \geq \sqrt{t}h(t), \text{ infinitely often}\Big) = \begin{cases} 0 \\ 1 \end{cases} \iff \int_1^\infty \frac{h(t)}{t}\, \mathrm{e}^{-\frac{1}{2}h^2(t)}\, \mathrm{d}t \begin{cases} < +\infty \\ = +\infty \end{cases}$$

In particular, we see that the curve $\sqrt{2t\log\log t}$ is crossed by Brownian motion infinitely often.

4. Another question is about the lower limits of Brownian motion. Chung (1948) proved that

$$\liminf_{t\to\infty} \frac{\sqrt{\log\log t}}{\sqrt{t}} \sup_{s\in[0,\,t]} |B_s| = \frac{\pi}{2\sqrt{2}} \qquad \text{a.s.}$$

Subsequently, Feller (1951) obtained a similar result for the range of Brownian motion:

$$\liminf_{t\to\infty} \frac{\sqrt{\log\log t}}{\sqrt{t}} \Big(\sup_{s\in[0,\,t]} B_s - \inf_{s\in[0,\,t]} B_s \Big) = \frac{\pi}{\sqrt{2}} \qquad \text{a.s.}$$

What happens if we replace $\sup_{s\in[0,\,t]} |B_s|$ by $\sup_{s\in[0,\,t]} B_s$? Intuitively, as far as the lower limits are concerned, it is clear that $\sup_{s\in[0,\,t]} B_s$ can be far smaller than $\sup_{s\in[0,\,t]} |B_s|$: this was confirmed by Hirsch (1965) who established the following criterion:

$$\liminf_{t\to\infty} \frac{(\log t)^a}{\sqrt{t}} \sup_{s\in[0,\,t]} B_s = \begin{cases} 0 & \text{if } a \le 1, \\ \infty & \text{if } a > 1, \end{cases} \qquad \text{a.s.}$$

5.5. Semi-group of Brownian motion

Recall that $\mathbb{P}_x$ denotes the probability under which B is Brownian motion starting from x with the convention $\mathbb{P} = \mathbb{P}_0$[5]. The simple Markov property says that under $\mathbb{P}_x$, the process $\widetilde{B} = (\widetilde{B}_t := B_{t+s} - B_s,\, t \ge 0)$ is Brownian motion independent of $\mathscr{F}_s$. We can state the Markov property of Brownian motion in the following, more familiar way: conditionally on $\mathscr{F}_s$, the process $\widehat{B} = (\widehat{B}_t := B_{t+s},\, t \ge 0)$ is Brownian motion starting at the (random) position $y = B_s$:

$$\begin{aligned} \mathbb{E}_x\Big[F(\widehat{B}_{t_1}, \cdots, \widehat{B}_{t_n}) \,|\, \mathscr{F}_s \Big] &= \mathbb{E}_x\Big[F(\widetilde{B}_{t_1} + y, \cdots, \widetilde{B}_{t_n} + y) \,|\, \mathscr{F}_s \Big] \\ &= \mathbb{E}_0\Big[F(B_{t_1} + y, \cdots, B_{t_n} + y) \Big] \\ &= \mathbb{E}_y\Big[F(B_{t_1}, \cdots, B_{t_n}) \Big], \end{aligned}$$

with $y := B_s$. For $t \ge 0$, we define the operator $P_t : \mathcal{B}(\mathbb{R}, \mathbb{R}_+) \mapsto \mathcal{B}(\mathbb{R}, \mathbb{R}_+)$ from the set of non-negative Borel function to itself by

$$P_t f(x) := \mathbb{E}_x[f(B_t)] = \int_{\mathbb{R}} \frac{1}{(2\pi t)^{1/2}} \exp\Big(- \frac{(x-y)^2}{2t} \Big) f(y)\, \mathrm{d}y. \qquad \text{for } x \in \mathbb{R}.$$

[5]As already mentioned, we can work on the canonical space of Brownian motion, and define $\mathbb{W}_x$ to be the image-measure of the Wiener measure $\mathbb{W}$ by the mapping $\mathrm{w} \mapsto \mathrm{w} + x$. As such, the process of projections $(X_t,\, t \ge 0)$ under $\mathbb{W}_x$ is Brownian motion with $\mathbb{W}_x(X_0 = x) = 1$.

We call $(P_t,\, t \geq 0)$ the **semi-group of** B. Notice that,for any $s, t \geq 0$, we have

$$\begin{aligned} P_{t+s}f(x) &= \mathbb{E}_x[f(B_{t+s})] \\ &= \mathbb{E}_x[\mathbb{E}_x[f(B_{t+s}) \,|\, \mathscr{F}_t]] \\ &= \mathbb{E}_x[\mathbb{E}_{B_s}[f(B_t)]] && \text{[Markov property at time } s\text{]} \\ &= \mathbb{E}_x[P_s f(B_t)] \\ &= P_t(P_s f)(x) \end{aligned}$$

which establishes the **semi-group property**:

$$P_t \circ P_s = P_s \circ P_t = P_{t+s} \qquad \text{for all } s, t \geq 0.$$

One can associate a semi-group to any Markov process. However, the semi-group of Brownian motion enjoys several specials properties (which are left as exercises):

- **Feller property**: if $f \in C_0$ (continuous, with $\lim_{|x|\to\infty} f(x) = 0$), then $P_t f \in C_0$, and $\lim_{t\downarrow 0} P_t f = f$ uniformly on $\mathbb{R}$.
- **Generator**: if $f \in C_c^2$ (class C^2 of compact support), then $\lim_{t\downarrow 0} \frac{P_t f(x) - f(x)}{t} = \frac{1}{2} f''(x)$.
- **Relation with the heat equation**: let $u(t, x) := P_t f(x)$. We have $u(0, x) = f(x)$. If f is a bounded Borel function, then

$$\frac{\partial u}{\partial t} = \frac{1}{2} \frac{\partial^2 u}{\partial x^2}.$$

Chapter 6

Brownian motion and martingales

In this chapter, we study the martingale property of Brownian motion. To this end, we will first define and give general results about continuous-time martingales which we will subsequently apply to Brownian motion.

6.1. Continuous-time martingales

Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$ denote a filtered probability space where, as usual, every σ-field is assumed to be complete. We also assume in this section that:

- **$(\mathscr{F}_t)$ satisfies the usual conditions (*i.e* is right-continuous).**

In practice, we will consider the augmented canonical filtration of Brownian motion which fulfills this assumption (*c.f* Remark 5.8). Continuous-time martingales are defined similarly as in the discrete time setting.

Definition 6.1. *A process $(X_t,\, t \geq 0)$ is a* ***martingale*** *[resp.* ***supermartingale; sub-martingale****] with respect to $(\mathscr{F}_t)_{t\geq 0}$ if*

(i) X is adapted to $(\mathscr{F}_t)_{t\geq 0}$.

(ii) $\forall\, t \geq 0$, $\mathbb{E}(|X_t|) < \infty$.

(iii) $\forall\, s < t$, $\mathbb{E}(X_t \,|\, \mathscr{F}_s) = X_s$, a.s. *[resp., $\leq X_s$; $\geq X_s$].*

The following important example shows that several martingales may be constructed from a process with independent increments.

Example 6.2. Suppose $X = (X_t, t \geq 0)$ is $(\mathscr{F}_t)$-adapted and that $X_{t+s} - X_s$ is independent of $\mathscr{F}_s$ for all $s, t \geq 0$.

1. If for all t, $\mathbb{E}(|X_t|) < \infty$, then $\widetilde{X}_t := X_t - \mathbb{E}(X_t)$ is a martingale.
2. If for all t, $\mathbb{E}(X_t^2) < \infty$, then $Y_t := \widetilde{X}_t^2 - \mathbb{E}(\widetilde{X}_t^2)$ is a martingale.
3. Let $\theta \in \mathbb{R}$. If $\mathbb{E}(\mathrm{e}^{\theta X_t}) < \infty$ for all $t \geq 0$, then $Z_t := \frac{\mathrm{e}^{\theta X_t}}{\mathbb{E}[\mathrm{e}^{\theta X_t}]}$ is a martingale.

(the proof of these statements is left as exercises).

Most results obtained for discrete-time martingales in the first part of the course remain valid for continuous-time martingales. In fact, we just need to assume some regularity of the trajectories of the process (typically right continuity[1]) to leverage the results obtained in the discrete setting and transfer them in continuous time via a discretization procedure.

Theorem 6.3 (Doob's L^p inequality). *Let $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$. Let (X_s) be a right-continuous martingale. Then for any $t \geq 0$,*

$$\left\| \sup_{s\in[0,t]} |X_s| \right\|_p \leq q \, \|X_t\|_p \,, \tag{6.1}$$

As a consequence,

$$\left\| \sup_{s\geq 0} |X_s| \right\|_p \leq q \sup_{s\geq 0} \|X_s\|_p \,. \tag{6.2}$$

Proof. Let $0 \leq t_1 < t_2 < \cdots < t_k = t$. Then $Y_n := X_{t_{n\wedge k}}$ is a discrete-time martingale. By Doob's L^p inequality for discrete-time martingales, we have

$$\left\| \max_{0\leq i\leq k} |X_{t_i}| \right\|_p \leq q \, \|X_t\|_p \,.$$

Let $D \subset [0, t]$ be a countable set containing t. Par the monotone convergence theorem, taking increasingly larger finite sets that exhaust D,

$$\left\| \sup_{s\in D} |X_s| \right\|_p \leq q \, \|X_t\|_p \,.$$

Since the trajectories of (X_s) is a.s. right-continuous, we have $\sup_{s\in[0,t]\cap D} |X_s| = \sup_{s\in[0,t]} |X_s|$ a.s., which yields (6.1) and (6.2) again with the use of the monotone convergence theorem. □

Theorem 6.4 (a.s. convergence). *Let $(X_t,\ t \geq 0)$ be a right-continuous submartingale such that*

$$\sup_{t\geq 0} \mathbb{E}(X_t^+) < \infty.$$

Then, $X_\infty := \lim_{t\to\infty} X_t$ exists a.s. and $\mathbb{E}(|X_\infty|) < \infty$.

[1] In fact, this regularity assumption is not restrictive because it can be proved that, under mild assumptions, every (super/sub)martingale admits a modification that is càdlàg (and therefore unique up to indistinguishability).

Proof. (i) Let $D \subset \mathbb{R}_+$ be countable and dense. Let $a < b$ and let $H_t(a,b)$ denote the number of up-crossings of $(X_s,\ s \in [0,\ t] \cap D)$ over $[a,\ b]$. Then, just as in the previous proof, using Doob's up-crossing lemma for discrete martingales combined with the monotone convergence theorem, we find that

$$\mathbb{E}\left[H_t(a,b)\right] \le \frac{\mathbb{E}[(X_t - a)^+]}{b-a} \le \frac{1}{b-a}\Big(\sup_{s\ge 0}\mathbb{E}(X_s^+) + |a|\Big).$$

Letting $t \to \infty$ and using again the monotone convergence theorem shows $E[H_\infty(a,b)] < \infty$. In particular, we get

$$\mathbb{P}(\text{for any } a,b \in \mathbb{Q},\ H_\infty(a,b) < \infty) = 1.$$

Just as in the discrete case, this fact implies that $X_\infty := \lim_{t\to\infty,\, t\in D} X_t$ exists a.s. But because the paths of X are assumed to be right continuous and because D is dense, we must also have $X_\infty = \lim_{t\to\infty} X_t$. Finally, by Fatou's lemma,

$$\mathbb{E}(|X_\infty|) \le \liminf_{t\to\infty,\, t\in D} \mathbb{E}(|X_t|) \le \sup_{s\ge 0}\mathbb{E}(|X_s|) \le 2\sup_{s\ge 0}\mathbb{E}(X_s^+) - \mathbb{E}(X_0) < \infty$$

where we used that $\mathbb{E}(|X_s|) = 2\mathbb{E}(X_s^+) - \mathbb{E}(X_s) \le 2\mathbb{E}(X_s^+) - \mathbb{E}(X_0)$ because X is a sub-martingale. □

Corollary 6.5. *Let $(X_t,\ t \ge 0)$ be a non-negative, right-continuous supermartingale. Then $X_\infty := \lim_{t\to\infty} X_t$ exists* a.s., *and $\mathbb{E}(X_\infty) \le \mathbb{E}(X_0)$.*

Theorem 6.6 (L^p **convergence**). *Let $p > 1$. If $(X_t,\ t \ge 0)$ is a right-continuous martingale satisfying*

$$\sup_{t\ge 0}\mathbb{E}(|X_t|^p) < \infty,$$

then $X_t \to X_\infty$ a.s. *and in L^p.*

Proof. By Jensen's inequality, $\sup_{t\ge 0}\mathbb{E}(|X_t|) < \infty$. Therefore Theorem 6.4 tells us that $X_t \to X_\infty$ a.s. By Doob's inequality, we also have $\mathbb{E}(\sup_{t\ge 0}|X_t|^p) < \infty$ so the dominated convergence theorem implies that the convergence also holds in L^P. □

Theorem 6.7 (L^1 **convergence**). *Let $(X_t,\ t \ge 0)$ be a uniformly integrable and right-continuous martingale. Then*

(i) $X_t \to X_\infty$ a.s. and in L^1.

(ii) $X_t = \mathbb{E}(X_\infty \,|\, \mathscr{F}_t)$, a.s.

Proof. *(i)* The uniform integrability implies that $\sup_{t\geq 0}\mathbb{E}(|X_t|) < \infty$, which, thanks to Theorem 6.4 shows that X_t converges a.s. to X_∞ with $\mathbb{E}(X_\infty) < \infty$. Furthermore, the combination of uniform integrability and convergence a.s. (convergence in probability would be sufficient) implies convergence in L^1.

(ii) Let $t \geq s$ and $A \in \mathscr{F}_s$. Then $\mathbb{E}(X_t\,\mathbf{1}_A) = \mathbb{E}(X_s\,\mathbf{1}_A)$. We let $t \to \infty$. The L^1 convergence implies that $\mathbb{E}(X_\infty\,\mathbf{1}_A) = \mathbb{E}(X_s\,\mathbf{1}_A)$ for all $A \in \mathscr{F}_s$. This exactly means that $X_s = \mathbb{E}(X_\infty \,|\, \mathscr{F}_s)$, a.s. □

Theorem 6.8 (Doob's optional sampling theorem). *Let $(X_t,\ t \geq 0)$ be a uniformly integrable right-continuous martingale and let $S \leq T$ be stopping times. Then, we have*

$$\mathbb{E}(X_T \,|\, \mathscr{F}_S) = X_S \qquad \text{a.s.} \tag{6.3}$$

In particular, $\mathbb{E}(X_T) = \mathbb{E}(X_0) = \mathbb{E}(X_\infty)$.

Proof. We define

$$\begin{aligned} S_n &:= \sum_{k=0}^{\infty} \frac{k}{2^n}\mathbf{1}_{\{\frac{k-1}{2^n} < S \leq \frac{k}{2^n}\}} + (+\infty)\,\mathbf{1}_{\{S=\infty\}} \\ T_n &:= \sum_{k=0}^{\infty} \frac{k}{2^n}\mathbf{1}_{\{\frac{k-1}{2^n} < T \leq \frac{k}{2^n}\}} + (+\infty)\,\mathbf{1}_{\{T=\infty\}}. \end{aligned}$$

(S_n) and (T_n) are non-increasing sequences of stopping times converging almost surely to S and T. For any n, $(X_{i/2^n},\ i \geq 0)$ is a discrete-time $(\mathscr{F}_{i/2^n})$-martingale and is uniformly integrable. Moreover, S_n and T_n are stopping time for this filtrations and $S_n \leq T_n$ so the optional sampling theorem for discrete time martingales shows that $\mathbb{E}(X_{T_n} \,|\, \mathscr{F}_{S_n}) = X_{S_n}$

Let $A \in \mathscr{F}_S$, we have $S \leq S_n$ so $A \in \mathscr{F}_{S_n}$ and therefore

$$\mathbb{E}(\mathbf{1}_A X_{T_n}) = \mathbb{E}(\mathbf{1}_A X_{S_n}). \tag{6.4}$$

Now, since X has right continuous trajectories, we have $X_{S_n} \to X_S$ and $X_{T_n} \to X_t$ a.s. Furthermore, (X_{S_n}) is uniformly integrable because $X_{S_n} = \mathbb{E}(X_\infty \,|\, \mathscr{F}_{S_n})$ thanks to the optional sampling theorem for discrete time martingales. Thus, by the improved dominated convergence theorem, we get $\mathbb{E}(\mathbf{1}_A X_{S_n}) \to \mathbb{E}(\mathbf{1}_A X_S)$. For the same reasons, $\mathbb{E}(\mathbf{1}_A X_{T_n}) \to \mathbb{E}(\mathbf{1}_A X_T)$. Combining these limits with (6.4) shows that

$$\mathbb{E}(\mathbf{1}_A X_T) = \mathbb{E}(\mathbf{1}_A X_S) \qquad \text{for all } A \in \mathscr{F}_S,$$

which exactly means $\mathbb{E}(X_T \,|\, \mathscr{F}_S) = X_S$. □

Theorem 6.9 (Doob's optional sampling theorem 2). *Let $(X_t,\ t \geq 0)$ be a right-continuous martingale and let $S \leq T$ be stopping times. If T is bounded, then*

$$\mathbb{E}(X_T \mid \mathscr{F}_S) = X_S \qquad \text{a.s.}$$

In particular, $\mathbb{E}(X_T) = \mathbb{E}(X_0)$.

Proof. Let $a > 0$ be such that $S \leq T \leq a$. Then, the martingale $(X_{t\wedge a}, t \geq 0)$ is uniformly integrable because $X_{t\wedge a} = \mathbb{E}[X_a \mid \mathscr{F}_{t\wedge a}]$ and we can apply the previous theorem to this martingale. □

6.2. Brownian motion as a martingale

We now apply the results of the previous section to Brownian motion. Let $(B_t, t \geq 0)$ denote a Brownian motion and let $(\mathscr{F}_t, t \geq 0)$ denote it natural (completed) filtration. Since B is a process with independent increments, Example 6.2 provides three $\mathscr{F}_t$-martingales that are associated with B:

1. $(B_t, t \geq 0)$ itself is a continuous martingale [**natural martingale**].
2. $(B_t^2 - t,\ t \geq 0)$ is a continuous martingale [**quadratic martingale**].
3. $(\mathrm{e}^{\theta B_t - \frac{\theta^2}{2}t},\ t \geq 0)$ for any $\theta \in \mathbb{R}$, is a continuous martingale [**exponential martingale**].

With the help of these martingales, we can compute several functionals of Brownian motion and obtain new insight on the properties of its trajectories. Recall the notation

$$T_a := \inf\{t > 0,\ B_t = a\} \qquad \text{for } a \in \mathbb{R}.$$

Proposition 6.10. *Let $-a < 0 < b$. We have*

$$\mathbb{P}(T_{-a} < T_b) = \frac{b}{a+b} \tag{6.5}$$

and

$$\mathbb{E}(T_{-a} \wedge T_b) = ab. \tag{6.6}$$

Proof. Let $\tau := T_{-a} \wedge T_b$. This is a stopping time and the martingale $(B_{t\wedge\tau}, t \geq 0)$ is uniformly integrable because its absolute value is bounded by $(-a) \vee b$. In particular this martingale converges a.s. which implies that $\tau < \infty$ a.s. because Brownian motion does not converge a.s. (or simply use that B is a.s. unbounded). Applying the optional sampling theorem, we get

$$0 = \mathbb{E}(B_0) = \mathbb{E}(B_\tau) = -a\mathbb{P}(T_{-a} < T_b) + b\mathbb{P}(T_b < T_a)$$

where we used that $B_{T_{-a}} = -a$ and $B_{T_b} = b$ because B has continuous paths. Formula (6.5) follows from the equality $\mathbb{P}(T_b < T_{-a}) = 1 - \mathbb{P}(T_{-a} < T_b)$.

For the second equality, we use the quadratic martingale $(B_t^2 - t,\, t \geq 0)$. Let $\tau_n = \tau \wedge n$ which is a bounded stopping time so the optional sampling theorem (version 2) yields

$$0 = \mathbb{E}(B_0 - 0) = \mathbb{E}(B_{\tau_n}^2 - \tau_n^2) = \mathbb{E}(B_{\tau_n}^2) - \mathbb{E}(\tau_n).$$

We have $\tau_n \to \tau$ a.s. On the one hand, $B_{\tau_n}^2 \leq a^2 \vee b^2$ so the dominated convergence theorem gives $\mathbb{E}(B_{\tau_n}^2) \to \mathbb{E}(B_\tau^2)$. On the other hand, the monotone convergence theorem gives $\mathbb{E}(\tau_n) \to \mathbb{E}(\tau)$ so we conclude that

$$\mathbb{E}(\tau) = \mathbb{E}(B_\tau^2) = a^2\mathbb{P}(T_{-a} < T_b) + b^2\mathbb{P}(T_b < T_{-a}) = \frac{a^2 b}{a+b} + \frac{b^2 a}{a+b} = ab$$

as desired. □

Corollary 6.11. *The random variable* $\sup_{s \leq T_{-1}} B_s$ *has the same distribution as* $\frac{1-U}{U}$ *where* U *is a uniform random variable in* $]0,1[$.

Proof. We simply notice that, according to the previous Proposition, for any $b > 0$, we have $\mathbb{P}(\sup_{s \leq T_{-1}} B_s > b) = \mathbb{P}(T_b < T_{-1}) = \frac{1}{1+b} = \mathbb{P}((1-U)/U > b)$. □

Proposition 6.12. *Let* $-a < 0 < b$. *We have, for all* $\lambda \geq 0$,

$$\mathbb{E}\left[e^{-\lambda T_b}\right] = e^{-b\sqrt{2\lambda}} \tag{6.7}$$

and

$$\mathbb{E}\left[e^{-\lambda(T_{-a} \wedge T_b)}\right] = \frac{\cosh\left(\frac{a-b}{2}\sqrt{2\lambda}\right)}{\cosh\left(\frac{a+b}{2}\sqrt{2\lambda}\right)}$$

Proof. Consider the exponential martingale $M_t := e^{\theta B_t - \frac{\theta^2}{2}t}$. We have $\theta B_{t\wedge T_b} - \frac{\theta^2}{2}(t \wedge T_b) \leq \theta b$, so $M_{t\wedge T_b}$ is a continuous and bounded martingale, thus uniformly integrable, and $M_\infty = e^{\theta a - \frac{\theta^2}{2}T_b}$ since $T_b < \infty$ a.s. By the optional sampling theorem, we deduce that

$$1 = \mathbb{E}[M_0] = \mathbb{E}[M_{T_b}] = \mathbb{E}\left[e^{\theta b - \frac{\theta^2}{2}T_b}\right] = e^{\theta b}\mathbb{E}[\, e^{-\frac{\theta^2}{2}T_a}\,]$$

and the formula for the Laplace transform of T_b follows from the change of variable $\lambda = \theta^2/2$.

We now compute the Laplace transform of $\tau := T_{-a} \wedge T_b$. Let $\theta \in \mathbb{R}$, and define

$$M_t := \sinh(\theta(B_t + a))\, \mathrm{e}^{-\frac{\theta^2}{2} t}\,.$$

It is easy to check that $(M_t, t \geq 0)$ is again a continuous martingale. Also, $M_{t\wedge\tau}$ is bounded hence uniformly integrable. Applying the optional sampling theorem, we get

$$\sinh(\theta a) \;=\; \mathbb{E}\Big[\sinh(\theta(B_\tau + a))\, \mathrm{e}^{-\frac{\theta^2}{2}\tau}\Big] \;=\; \sinh(\theta(a+b))\mathbb{E}\Big[\mathrm{e}^{-\frac{\theta^2}{2}T_b}\, \mathbf{1}_{\{T_b < T_{-a}\}}\Big].$$

Hence,

$$\mathbb{E}\Big[\mathrm{e}^{-\frac{\theta^2}{2}T_b}\, \mathbf{1}_{\{T_b < T_{-a}\}}\Big] = \frac{\sinh(\theta a)}{\sinh(\theta(a+b))}.$$

Exchanging the roles of a and b (which amounts to replacing B by $-B$), we also obtain:

$$\mathbb{E}\Big[\,\mathrm{e}^{-\frac{\theta^2}{2}T_{-a}}\, \mathbf{1}_{\{T_b > T_{-a}\}}\Big] = \frac{\sinh(\theta b)}{\sinh(\theta(a+b))}.$$

So

$$\mathbb{E}[\mathrm{e}^{-\frac{\theta^2}{2}T_{a,b}}] = \mathbb{E}\Big[\,\mathrm{e}^{-\frac{\theta^2}{2}T_b}\, \mathbf{1}_{\{T_b < T_{-a}\}}\Big] + \mathbb{E}\Big[\,\mathrm{e}^{-\frac{\theta^2}{2}T_{-a}}\, \mathbf{1}_{\{T_b > T_{-a}\}}\Big] = \frac{\sinh(\theta a) + \sinh(\theta b)}{\sinh(\theta(a+b))} = \frac{\cosh(\frac{\theta(a-b)}{2})}{\cosh(\frac{\theta(a+b)}{2})},$$

and the result follows again from the change of variable $\lambda = \theta^2/2$. □

Remark 6.13. We already computed the law of T_a in Corollary 5.13. It is easy to check that (6.7) is indeed the Laplace transform of the density given by (5.6).

Chapter 7

Further properties of Brownian motion

7.1. Donsker's theorem and KMT's invariance principle

The Central Limit Theorem tells us that a sum of i.i.d random variable, once centered and correctly re-scaled converges to a Gaussian law provided these random variables have a finite second moment. Thus, the Gaussian law appears as the "universal limit" for a large class of random variables. Similarly, the same holds true for Brownian motion which may be thought as a universal limit for random walks (with finite second moment).

Let $(\xi_i,\, i \geq 1)$ be a sequence of i.i.d. random variables with

$$c := \mathbb{E}(\xi_1) \in \mathbb{R} \quad \text{and} \quad \sigma^2 := \mathrm{Var}(\xi) \in (0, \infty).$$

Let $S_0 := 0$ and $S_n := \xi_1 + \cdots + \xi_n$ for $n \geq 1$. We extend S into a continuous time process $(S_t,\, t \geq 0)$ by linear interpolation:

$$S(t) := S_{\lfloor t \rfloor} + (t - \lfloor t \rfloor)\xi_{\lfloor t \rfloor + 1}$$

For any n, we define the re-scaled process

$$S^n_t := \frac{S_{nt} - cnt}{\sigma\sqrt{n}} \qquad \text{for } t \in [0,1]. \tag{7.1}$$

Thus $(S^n,\, n \geq 1)$ defines a sequence of continuous stochastic processes. We consider the metric space

$$(\mathcal{C}([0,1], \mathbb{R}), \|\cdot\|_\infty)$$

of real-valued continuous functions, endowed with the topology of uniform convergence and the corresponding Borel σ-field. According to Lemma 4.26, any continuous random process is measurable with respect to this σ-field so we can see the processes S^n as random functions taking values in $\mathcal{C}([0,1],\mathbb{R})$.

Theorem 7.1 (Donsker's Theorem). *The sequence of processes $(S^n, n \geq 1)$ converges in distribution, in $(\mathcal{C}([0,1],\mathbb{R}), \|\cdot\|_\infty)$ toward a Brownian motion $B = (B_t, t \in [0,1])$*[1].

This results implies, in particular, that S^n_1 converge in law toward B_1 because the projection $f \to f(1)$ is continuous for $\|\cdot\|_\infty$. Thus, we recover the Central Limit Theorem. More generally, Donsker's theorem implies the convergence of finite-dimensional marginals:

$$(S^n_{t_1}, \ldots, S^n_{t_k}) \xrightarrow[n\to\infty]{\text{law}} (B_{t_1}, \ldots, B_{t_n}).$$

However, Donsker's theorem says much more than that: it is a **functional CLT** and it allows us to obtain results out of reach of the usual CLT because we can now consider functionals of the whole trajectory. To see this, let us first recall the so-called Portmanteau Theorem.[2]

Theorem 7.2 (Portmanteau). *Let (X_n) and X be random variables taking values in a metric space (E,d). The following statements are equivalent and characterize the convergence in distribution of X_n to X.*

1. $\mathbb{E}[F(X_n)] \to \mathbb{E}[F(X)]$ *for every bounded continuous function f.*
2. $\mathbb{E}[F(X_n)] \to \mathbb{E}[F(X)]$ *for every bounded Lipschitz function f.*
3. $\mathbb{E}[F(X_n)] \to \mathbb{E}[F(X)]$ *for every bounded measurable function f such that*

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

4. $\mathbb{P}(X_n \in A) \to \mathbb{P}(X \in A)$ *for any measurable A such that* $\mathbb{P}(X \in \partial A) = 0$.
5. $\limsup \mathbb{P}(X_n \in C) \leq \mathbb{P}(X \in C)$ *for any closed set C.*
6. $\liminf \mathbb{P}(X_n \in O) \geq \mathbb{P}(X \in O)$ *for any open set O.*

In practice, functionals of trajectories are not always continuous for the topology of uniform convergence. But thanks to Portmanteau theorem, they only need to be continuous with probability 1 with respect to the limit law.

[1] Equivalently, the image law of S^n on $\mathcal{C}([0,1],\mathbb{R})$ converges weakly to the Wiener measure.
[2] A proof of Donsker's Theorem as well as Portmanteau's Theorem can be found in Billingsley, "*Convergence of probability Measures*" Wiley, 1999.

Corollary 7.3. *If the sequence (X_n) converges in distribution to X and if F is almost surely continuous at X, then $F(X_n)$ converges in distribution to $F(X)$.*

Proof. Fix a continuous bounded function G defined on the image space of F. Then $G \circ F$ is a.s. continuous at X so item 3. of Portmanteau Theorem tells us that $\mathbb{E}[G(F(X_n))] \to \mathbb{E}[G(F(X))]$. Since this holds for any bounded continuous G, item 1. implies that $F(X_n)$ converges in distribution to $F(X)$. □

Combining Donsker's Theorem together with Portmanteau's Theorem, we can leverage results obtained for Brownian motion and transpose them to random walks. We give here some examples.

Proposition 7.4. *Let $S_n = \xi_1 + \ldots \xi_n$ denote a random walk with i.i.d. increments ξ_n which are centered $\mathbb{E}(\xi_i) = 0$ and with finite second moment $\mathbb{E}(\xi^2) = \sigma^2 > 0$. We have*

$$\frac{1}{\sigma\sqrt{n}} \sup_{k \le n} S_k \xrightarrow[n\to\infty]{law} |\mathcal{N}(0,1)| \tag{7.2}$$

$$\frac{1}{\sigma n\sqrt{n}} \sum_{k=0}^{n} S_k \xrightarrow[n\to\infty]{law} \mathcal{N}(0,1/3) \tag{7.3}$$

$$\frac{1}{n} \inf\left\{0 \le k \le n \,:\, S_k = \sup_{[0,n]} S\right\} \xrightarrow[n\to\infty]{law} \textit{Arsine distribution.} \tag{7.4}$$

Proof.

- We first prove the convergence for the supremum. Recalling notation (7.1), we can write $\frac{1}{\sigma\sqrt{n}} \sup_{k\le n} S_k = F(S^n)$ where

$$F \,:\, \begin{array}{rcl} \mathcal{C}([0,1],\mathbb{R}) & \to & \mathbb{R} \\ g & \mapsto & \sup_{[0,1]} g. \end{array}$$

  Clearly, F is continuous for $\|\cdot\|_\infty$ so, according to Donsker's theorem, $F(S^n)$ converges in distribution to $\sup_{t\in[0,1]} B_t$ which, as we have already seen, as the same law as $|B_1|$. This proves (7.2).

- We prove the convergence for the sum. We observe that, since S^n is the re-scaled linear interpolation of the random walk S, we have

$$\frac{1}{\sigma n\sqrt{n}} \sum_{k=0}^{n} S_k = \int_0^1 S^n_s \,\mathrm{d}s \;+\; \frac{S_0 + S_n}{2\sigma n\sqrt{n}}$$

  The second term on the r.h.s of the equation above converges to 0 a.s. by the law of large number so we just need to find the limit in law of the integral. Using that

$f \to \int_0^1 f$ is continuous for $\|\cdot\|_\infty$, Donsker's theorem tells us that $\int_0^1 S^n_s \,\mathrm{d}s$ converges in law to $Z := \int_0^1 B_s \,\mathrm{d}s$. On the other hand, using a Riemann sum, we can write $Z = \lim_n \frac{1}{n}\sum_i^n B_{i/n}$ which shows that Z is a centered Gaussian as a limit of centered Gaussians. Moreover,

$$\mathbb{E}(Z^2) = \lim_{n\to\infty} \mathbb{E}\Big[\Big(\frac{1}{n}\sum_{i=0}^{n} B_{\frac{i}{n}}\Big)^2\Big] = \lim_{n\to\infty} \mathbb{E}\Big[\Big(\sum_{i=0}^{n-1} \frac{n-i}{n}(B_{\frac{i+1}{n}} - B_{\frac{i}{n}})\Big)^2\Big]$$

$$= \lim_{n\to\infty} \sum_{i=0}^{n-1} \frac{(n-i)^2}{n^2}\mathbb{E}[(B_{\frac{i+1}{n}} - B_{\frac{i}{n}})^2] = \lim_{n\to\infty} \sum_{i=0}^{n-1} \frac{(n-i)^2}{n^3} = \frac{1}{3},$$

and (7.3) follows.

- We prove the arcsine law. We now consider the functional

$$F(f) := \inf\Big\{t \in [0,1] \,:\, f(t) = \sup_{[0,1]} f\Big\}.$$

We have

$$\frac{1}{n}\inf\Big\{0 \le k \le n \,:\, S_k = \sup_{[0,n]} S\Big\} \;=\; F(S^n).$$

However, this time F is not continuous for $\|\cdot\|_\infty$. Consider for example $f_n(x) = x/n$, then $f_n \to 0$ uniformly on $[0,1]$ yet $F(f_n) = 1$ and $F(0) = 0$. Still, it is not difficult to show (and it is left as an exercice) that F is continuous at any $f \in \mathcal{C}([0,1],\mathbb{R})$ such that f attains it maximum at a single point over $[0,1]$. According to Proposition 5.15, this is a.s. the case for Brownian motion, so we can use Corollary 7.3 to conclude that $F(S^n)$ converges in distribution to $F(B)$ and the result follows from the arcsine law for the maximum of Brownian motion (Theorem 5.16).

□

Interestingly enough, it is also possible to use Donsker's Theorem the other way around *i.e.* prove a result for a particular random walk and then transfer it to Brownian motion. For example, If S is the simple random walk on $\mathbb{Z}$, it is easy to see that the process $Y_n := \sup_{k\le n} S_k - S_n$ is a Markov chain that behave like $|S|$ except at 0 where it sticks to 0 with probability 1/2. However, the time spent at 0 is negligible because S is null recurrent. Using Donsker's theorem, we obtain a proof of a celebrated result from Lévy:

Theorem 7.5 (Lévy). *If B is a Brownian motion, then the processes $(|B_t|,\, t \ge 0)$ and $(\sup_{s\le t} B_s - B_t,\, t \ge 0)$ have the same law.*

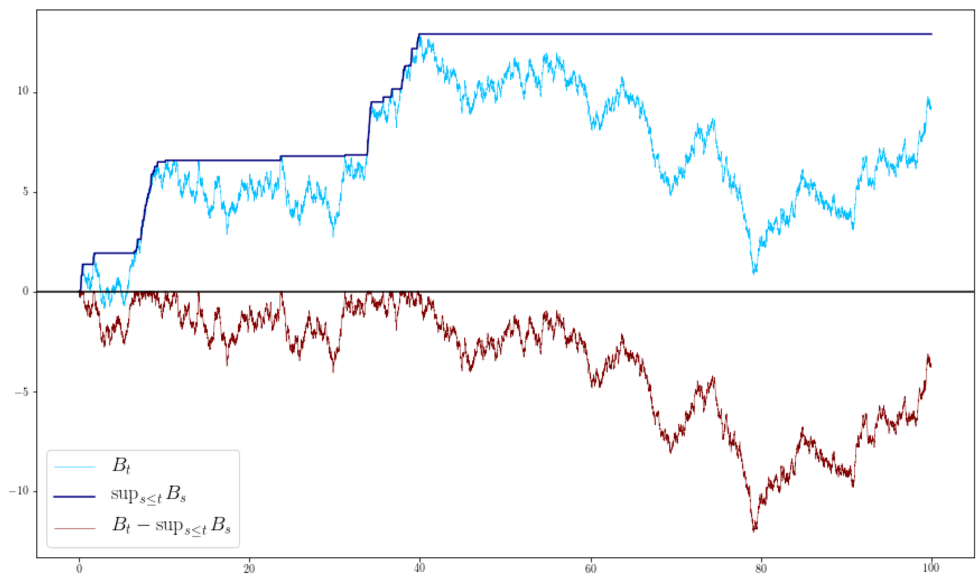


Figure 7.1: Illustration of Theorem 7.5: the red curve has the same law as $-|B|$.

It is possible to strengthen the Donsker's Theorem into an almost-sure invariance principle, if we are allowed to redefine the random variables.

Theorem 7.6 (Komlós, Major and Tusnády (KMT), 1975). *Let ξ be a random real-valued random variable with $\mathbb{E}(\xi) = 0$ and $\mathbb{E}(\xi^2) = 1$. There exist a probability space $(\widetilde{\Omega}, \widetilde{\mathscr{F}}, \widetilde{\mathbb{P}})$ on which we can construct an i.i.d. sequence $(\widetilde{\xi}_i,\, i \ge 1)$ distributed as ξ and a Brownian motion $(\widetilde{B}_t,\, t \in [0, 1])$, such that, as $n \to \infty$,*

$$\sup_{0\le s\le t} |\widetilde{S}_s - \widetilde{B}_s| = O(\log t), \qquad \widetilde{\mathbb{P}}\text{-a.s.}.$$

where $\widetilde{S}_i := \widetilde{\xi}_1 + \ldots + \widetilde{\xi}_i$ is extended by linear interpolation for non-integer times.

The KMT strong invariance principle directly implies Donsker's Theorem. It also enables to prove finer results. For example, we deduce from the law of the iterated logarithm for Brownian motion (Theorem 5.21) that, if S is a centered random walk with unit variance, then

$$\limsup_{n\to\infty} \frac{S_n}{(2n \log\log n)^{1/2}} = 1, \qquad \text{a.s.},$$

This result is known as the Hartman–Wintner law of the iterated logarithm.

7.2. Variations of Brownian motion

In Corollary 4.25, we saw that the trajectories of Brownian motion are a.s. not of bounded variation on any interval. We now show that they have, however, deterministic **quadratic variation**.

Theorem 7.7 (Lévy). *Fix $t > 0$. Let $\Pi := \{0 = t_0 < t_1 < \cdots < t_p = t\}$ be a sequence of subdivisions of $[0, t]$. Write $\|\Pi\| := \max_{1\le i\le p}(t_i - t_{i-1})$. Then*

$$\lim_{\|\Pi\|\to 0} \sum_{i=1}^{p} (B_{t_i} - B_{t_{i-1}})^2 = t, \qquad \text{in } L^2(\mathbb{P}).$$

Moreover, if the limit above is taken over a sequence satisfying $\Pi_1 \subset \Pi_2 \subset \cdots$, then we also have almost sure convergence.

Proof. Let us first prove L^2 convergence. Define

$$Y_i := (B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1}), \qquad 1 \le i \le p\,.$$

Then $(Y_i,\, 1 \le i \le p)$ are i.i.d. centered, with

$$\mathbb{E}(Y_i^2) = (t_i - t_{i-1})^2\, \mathbb{E}[(B_1^2 - 1)^2] = 2(t_i - t_{i-1})^2.$$

Accordingly, we get,

$$\begin{aligned} \mathbb{E}\Big[\Big(\sum_{i=1}^{p}(B_{t_i} - B_{t_{i-1}})^2 - t\Big)^2\Big] &= \mathbb{E}\Big[\Big(\sum_{i=1}^{p} Y_i\Big)^2\Big] \\ &= 2\sum_{i=1}^{p}(t_i - t_{i-1})^2 \\ &\le 2t\|\Pi\| \longrightarrow 0. \end{aligned}$$

This yields the $L^2(\mathbb{P})$ convergence.

We prove a.s. convergence only[3] for the special case $t_i = t_i^n := \frac{i}{2^n}$, $0 \le i \le 2^n$. We have seen that

$$\mathbb{E}\Big[\Big(\sum_{i=1}^{2^n} Y_i\Big)^2\Big] \le 2\sum_{i=1}^{2^n}(t_i^n - t_{i-1}^n)^2 = \frac{1}{2^{n-1}}.$$

[3]The proof of a.s. convergence in the general case is more technical. We refer to Proposition II.2.12 in the book of Revuz and Yor, "*Continuous Martingales and Brownian Motion*" (third edition), Springer, 1999.

By Tchebychev's inequality,

$$\mathbb{P}\Big(\Big|\sum_{i=1}^{2^n} Y_i\Big| > \frac{1}{n}\Big) \leq \frac{n^2}{2^{n-1}},$$

which is summable in $n \geq 1$. By the Borel-Cantelli lemma, it follows that, ω-a.s., there exists $n_0 = n_0(\omega) < \infty$ such that

$$\Big|\sum_{i=1}^{2^n} Y_i\Big| \leq \frac{1}{n}, \qquad \forall\, n \geq n_0,$$

from which convergence a.s. follows. □

Remark 7.8. The statement that "Brownian motion has quadratic variation" is formally incorrect because it would require the almost-sure convergence stated in Theorem 7.7 to hold true for any subdivision as the mesh goes to 0. However, this it is false without the assumption that $\Pi_1 \subset \Pi_2 \subset \cdots$. In fact, Lévy also showed that[4]

$$\limsup_{\|\Pi\|\to 0} \sum_{i=1}^{p} (B_{t_i} - B_{t_{i-1}})^2 = \infty, \qquad \text{a.s.},$$

On the other hand, explosion can be avoided by taking a function that is slightly smaller than x^2: Taylor (1972) proved that

$$\limsup_{\|\Pi\|\to 0} \sum_{i=1}^{p} g(B_{t_i} - B_{t_{i-1}}) = 2, \qquad \text{a.s.},$$

where $g(x) := \frac{x^2}{\log^* \log^* \frac{1}{|x|}}$ with $\log^* x := \max\{1, \log x\}$.

Remark 7.9 (Ito calculus). Given a function f with bounded variation, we can define the associated Stieltjes integral by

$$\int_0^t h \,\mathrm{d}f \;:=\; \lim_{\|\Pi\|\to 0} \sum_{i=1}^{p} h(t_{i-1})\big(f(t_i) - f(t_{i-1})\big).$$

which is well-defined for any continuous function h. In particular, for $f(x) = x$, we recover the Riemann integral[5]. We would like to do the same thing with a Brownian motion[6].

[4]See p. 48 of the book of D. Freedman, "*Brownian Motion and Diffusion*", Holden-Day, 1971.

[5]More generally, if f is differentiable, then $\int h \,\mathrm{d}f = \int h(x) f'(x) \,\mathrm{d}x$.

[6]There are several motivations for defining a notion of "integration along a Brownian". One of them is to provide a framework for studying (stochastic) differential equation with an added white noise which, informally, may be thought as the "derivative of Brownian motion".

Unfortunately, we cannot directly replace f by B since the limit will not exist in general. However, one can prove that, under mild assumptions we have convergence in L^2 (and therefore also in probability). For instance, if H is an adapted, continuous process, then we can define

$$(7.5)\quad \int_0^t H(s)\,\mathrm{d}B(s) := \lim_{\|\Pi\|\to 0} \sum_{i=1}^{p} H(t_{i-1})\big(B(t_i) - B(t_{i-1})\big) \qquad \text{for } L^2\text{-convergence of r.v.}$$

This definition is called **Ito's integral of Brownian motion** and it is the building block of **Stochastic calculus**. This notion of integration with respect to Brownian motion has many nice properties. For instance, the integrated process is a Martingale: this is easily understood from (7.5) by observing that, for each fixed subdivision Π, the right end side is a discrete martingale (in p) which translates, in the limit, to a martingale (in t) for the integral on the left hand side. Moreover, we have an isometry of L^2 spaces:

$$\mathbb{E}\left[\left(\int_0^t H(s)\,\mathrm{d}B(s)\right)^2\right] = \mathbb{E}\left[\int_0^t H^2(s)\,\mathrm{d}s\right].$$

The reader interested in a rigorous exposition of Stochastic calculus can refer the Le Gall's excellent book: "*Brownian motion, Martingale and Stochastic Calculus*" Springer, 2016.

7.3. Multidimensional Brownian motion

In these lecture notes, we defined and studied one-dimensional Brownian motion but one can also define Brownian motion in higher dimension:

Definition 7.10. *Let $d \geq 1$ and Let $B := (B_t = (B_t^{(1)}, \ldots, B_t^{(d)}), t \geq 0)$ be a process taking values in $\mathbb{R}^d$. We say that B is a* ***d-dimensional Brownian motion*** *if the 1-dimensional processes $B^{(1)}, \ldots, B^{(d)}$ are independent Brownian motions.*

At first, it seems that the definition of Brownian motion depend on the choice of the basis in $\mathbb{R}^d$ but this is not the case and the following proposition shows that d-dimensional Brownian motion is **isotropic**.

Proposition 7.11. *If U is an orthogonal matrix (*i.e. *such that $UU^T = I$), then UB is also a d-dimensional Brownian motion.*

Proof. The process UB is a Gaussian process so we just need to check the covariance matrix: $\mathbb{E}[(UB)_s^{(i)}(UB)_t^{(j)}] = (s \wedge t)\mathbf{1}_{i=j}$ which follows from the orthogonality of U. □

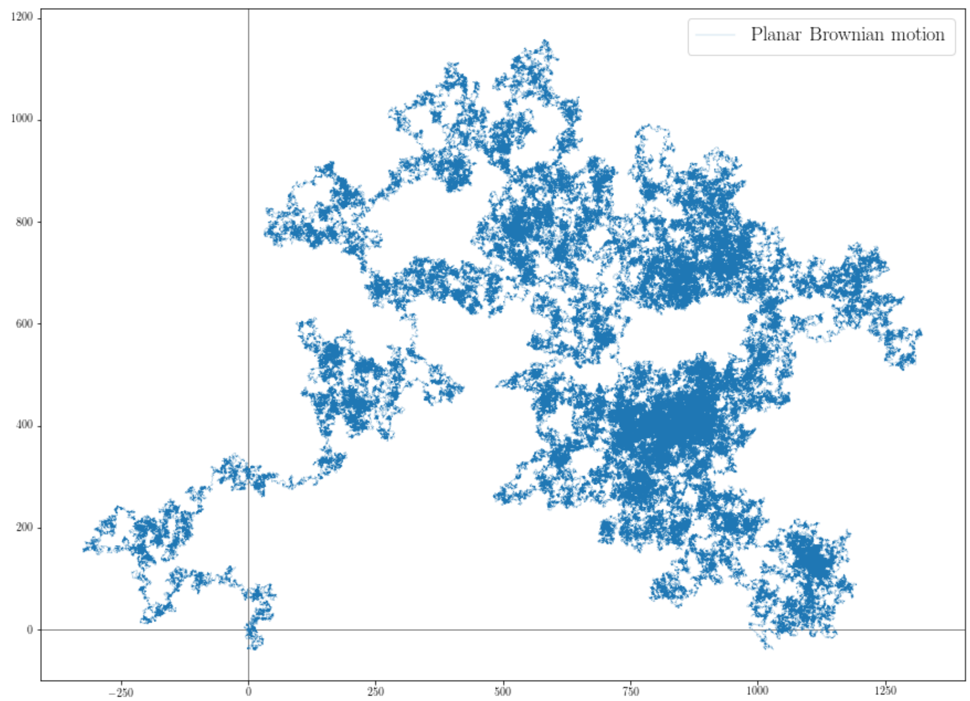


Figure 7.2: Trace of a Brownian motion in $\mathbb{R}^2$ up to time 10^6.

Remark 7.12. Brownian motion in dimension 2 is said to be **conformally invariant**: if B is a planar Brownian motion and f is an holomorphic map, then $f(B)$ is a time-change of a Brownian motion. This result may be seen as a local extension of Proposition 7.11 and can be proved using stochastic calculus.

Multidimensional Brownian motion inherits many properties of 1-dimensional Brownian motion. For instance, the strong Markov property still holds true: **If T is a finite stopping time, then the process $\tilde{B} = B_{T+\cdot} - B_T$ is a d-dimensional Brownian motion independent of the (completed) σ-field $\mathscr{F}_T$.**

Some path properties are also similar to that of one dimensional Brownian motion. For example, the law of the iterated logarithm still holds true:

$$\limsup_{t\to\infty} \frac{\|B_t\|}{(2t\log\log t)^{1/2}} = 1\,, \qquad \text{a.s.}$$

where $\|\cdot\|$ denotes the usual euclidean norm. However, multidimensional Brownian motion has some special properties that are not shared with one-dimensional Brownian motion. For example, it is known that in dimension $d \geq 2$, all points are polar, in the sense that

$$\mathbb{P}(B_t = x,\ \text{for some } t\ > 0) = 0 \qquad \text{for all } x \in \mathbb{R}^d.$$

This means that, Brownian motion is not recurrent in dimension $d \geq 2$. However, planar Brownian motion is still **neighbourhood-recurrent**: any ball $B(x, r)$ is visited at arbitrarily large time a.s., for any $x \in \mathbb{R}^2$ and any $r > 0$. So the trajectory of 2-dimensional Brownian motion is dense in the plane. On the other hand, in dimension $d \geq 3$ and higher, Brownian motion is transient

$$\|B_t\| \underset{t\to\infty}{\longrightarrow} +\infty \quad \text{a.s.} \tag{7.6}$$

This phase transition for the behaviour of Brownian motion between dimension 2 and 3 is analogous to Pólya's phase transition for recurrence and transience of the simple random walk in $\mathbb{Z}^d$.

Let us prove (7.6) in dimension $d \geq 3$. Fix $a > 0$. We have, using the triangular inequality,

$$\begin{aligned}\mathbb{P}\Big(\inf_{s\in[n,n+1]} \|B_s\| < n^a\Big) &\leq \mathbb{P}\Big(\|B_n\| < 2n^a\Big) + \mathbb{P}\Big(\sup_{s\in[n,n+1]} \|B_s - B_n\| > n^a\Big)\\ &= \mathbb{P}\Big(\|B_1\| < 2n^{a-\frac{1}{2}}\Big) + \mathbb{P}\Big(\sup_{s\in[0,1]} \|B_s\| > n^a\Big)\end{aligned}$$

where we used scaling and the Markov property for the last line. On the one hand,

$$\mathbb{P}\Big(\|B_1\| < 2n^{a-\frac{1}{2}}\Big) \leq Cn^{(a-\frac{1}{2})d}$$

for some constant C, because the density of the density of the d-dimensional Gaussian vector B_1 is bounded hence we obtain an upper bound by considering the volume of the ball $B(0, 2n^{a-\frac{1}{2}})$. In particular, we see that, for $a \in]0, \frac{1}{2} - \frac{1}{d}[$ (which is possible for $d \geq 3$), then $\sum_n \mathbb{P}\big(\|B_1\| < 2n^{a-\frac{1}{2}}\big) < \infty$. On the other hand, we have

$$\begin{aligned}\mathbb{P}\Big(\sup_{s\in[0,1]} \|B_s\| > n^a\Big) &= \mathbb{P}\Big(\sum_{i=1}^{d} \sup_{s\in[0,1]} (B_s^{(i)})^2 > n^{2a}\Big)\\ &\leq d\,\mathbb{P}\Big(\sup_{s\in[0,1]} |B_s^{(1)}| > \frac{n^a}{\sqrt{d}}\Big)\\ &\leq 2d\,\mathbb{P}\Big(B_s^{(1)} > \frac{n^a}{\sqrt{d}}\Big)\end{aligned}$$

where we used that $|B_1|$ and $\sup_{[0,1]} B$ have the same law for the last inequality. Thus, it is now clear, using estimates on Gaussian tails that $\sum_n \mathbb{P}\big(\sup_{s\in[0,1]} \|B_s\| > n^a\big) < \infty$ for any $a > 0$. Putting everything together, we conclude that, for $d \geq 3$ and for $a \in]0, \frac{1}{2} - \frac{1}{d}[$, by the Borel-Cantelli Lemma, we have (7.6) and, furthermore

$$\lim_{t\to\infty} \frac{\|B_t\|}{n^a} = +\infty \qquad \text{a.s.} \tag{7.7}$$

It is possible to improve this argument to show that (7.7) holds, in fact, for any $a \in (0, \frac{1}{2})$. More precisely, Dvoretzky and Erdős (1951) proved that, for any $b > 1/(d-2)$, we have

$$\lim_{t\to\infty} \frac{(\log t)^b \|B_t\|}{\sqrt{t}} = \infty \qquad \text{a.s.} \tag{7.8}$$

Since $\|B_t\|$ diverges to infinity, it makes sense to define the process of future infima: $J_t := \inf_{s\geq t} \|B_s\|$, $t \geq 0$. The lower limits of J_t are identical to those of $\|B_t\|$: you can replace $\|B_t\|$ by J_t in (7.8). Concerning the upper limits of J_t, Erdős and Taylor (1962) proved that

$$\limsup_{t\to\infty} \frac{J_t}{\sqrt{2t \log\log t}} = 1 \qquad \text{a.s.}$$

which is the same as the law of the iterated logarithm for $\|B_t\|$.

7.4. For further reading

These lecture notes aim to provide an introduction to Brownian motion. For further reading, the reader is advised to have a look at the following authoritative books on the subject:

- D. Revuz and M. Yor, "*Continuous Martingales and Brownian Motion*" (third edition), Springer, 1999.
- P. Mörters and Y. Peres, "*Brownian Motion*", Cambridge, 2010.
- J.-F. Le Gall, "Brownian motion, Martingales and Stochastic calculus", Springer, 2016.