8
Brownian motion and Itô calculus

Brownian motion is a continuous analogue of simple random walk*s* (as described in the previous part), which is very important in many practical applications. This importance has its origin in the universal properties of Brownian motion, which appear as the continuous scaling limit of many simple processes. Moreover, it is also intimately related to martingales and bounded-variation processes in continuous time. Brownian motion is a very rich structure that inherits properties from various fields of mathematics [**à compléter**].

This chapter presents in a first section the main properties of Brownian as well as various constructions. The second section presents Itô calculus for Brownian motion: this construction is only a particular case of stochastic calculus for semi-martingales; we choose however to present it here for two reasons: many applications do not require the general framework and moreover this general framework is abstract enough, so that a short introduction in a simple case may help for the next chapter.

Before going any further in this chapter, the reader is encouraged to read again the reminder of the properties of normal random variables presented in the first part.

8.1 Definition, construction and properties

8.1.1 Definition and structure

Definition 8.1.1 (Brownian motion)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A Brownian motion $(B_t)_{t\in\mathbb{R}_+}$ is a process with values in $\mathbb{R}$ such that:*

1. $B_0 = 0$,
2. *the maps $t \mapsto B_t(\omega)$ are almost surely continuous,*
3. *the increments $B_{t+s} - B_t$ are independent from $\sigma(B_u, u \leq t)$ and are centered normal r.v. with variance s, for all $t, s > 0$.*

Its existence is postponed to the next subsection and we assume now that it exists. The Brownian motion exhibits many interesting features mentioned in the previous chapter, that relates it to other probabilistic tools.

Property 8.1.1. *Let $(B_t)_{t\in\mathbb{R}_+}$ be a Brownian motion. We considered its associated filtration.*

- *the processes $(B_t)_{t\in\mathbb{R}_+}$, $(B_t^2 - t)_{t\in\mathbb{R}_+}$ and $(e^{uB_t - tu^2/2})_{t\in\mathbb{R}_+}$ are martingales.*

- *the process* $(B_t)_{t\in\mathbb{R}_+}$ *is a Markov process with initial law* δ_0 *and transition semi-group* $(Q_t)_{t\in\mathbb{R}_+}$ *given by:*

$$Q_t(x,dy) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(x-y)^2}{2t}\right) dy \tag{8.1}$$

Proof. Integrability of $B_t = B_t - B_0$ is provided by the integrability of normal random variables. A short computation gives:

$$\mathbb{E}\left[B_t|\mathcal{F}_s\right] = \mathbb{E}\left[B_t - B_s|\mathcal{F}_s\right] + \mathbb{E}\left[B_s|\mathcal{F}_s\right] = \mathbb{E}\left[B_t - B_s\right] + B_s = B_s \tag{8.2}$$

for all $s > t$, from the independence of the increments and their centered laws. The same decomposition of B_t as $(B_t - B_s) + B_s$ also gives the two other martingale properties. The Markov transition semi-group follows from the characterization:

$$\begin{aligned}\mathbb{E}\left[f(B_{t+u})g(B_t)\right] &= \mathbb{E}\left[f((B_t-0)+(B_{t+u}-B_t))g(B_t-0)\right] \\ &= \iint f(x+y)g(y)\frac{1}{2\pi\sqrt{tu}}\exp\left(-\frac{x^2}{2t}-\frac{y^2}{2u}\right) \\ &= \iint f(z)g(y)\frac{1}{2\pi\sqrt{tu}}\exp\left(-\frac{(z-y)^2}{2t}-\frac{y^2}{2u}\right) \\ &= \mathbb{E}\left[\left(\int f(z)e^{-(z-B_t)^2/(2u)}\frac{1}{\sqrt{2\pi u}}dz\right)g(B_t)\right]\end{aligned}$$

for all $t, u > 0$ through Fubini's theorem with arbitrary bounded measurable functions f and g. □

We have seen that the strong Markov property is not trivial in general. However, it holds for Brownian motion:

Theorem 8.1.1. *Brownian motion satisfies the weak and strong Markov properties. Let* T *be a stopping time and* $(B_t)_{t\in\mathbb{R}_+}$ *be a Brownian motion; conditionally on* $\{T < \infty\}$*, the process* $(B_{T+t} - B_T)_{t\in\mathbb{R}_+}$ *is a Brownian motion independent of* $\mathcal{F}_T$*.*

Proof. Either we deduce it from general results about Markov processes with càdlàg trajectories and values in a Polish space, or we may check it by hand. We know from the general theory of Markov process with càdlàg trajectories and we need only to prove:

$$\mathbb{E}\left[X\mathbf{1}_{T<\infty}F_n(B_{t_1+T}-B_T,\dots,B_{t_n+T}-B_T)\right] = \mathbb{E}\left[X\mathbf{1}_{T<\infty}\right]\mathbb{E}\left[F_n(B_{t_1},\dots,B_{t_n})\right] \tag{8.3}$$

for arbitrary $\mathcal{F}_T$-measurable integrable r.v. X and arbitrary bounded measurable functions $F_n : \mathbb{R}^n \to \mathbb{R}$.

The proof is very similar to the discrete case, except that the event $\{T < \infty\}$ cannot be written as a countable union of events $\{T = t_k\}$. In order to avoid this difficulty, we introduce approximations of T defined by:

$$T_n = \frac{\lfloor 2^n T\rfloor + 1}{2^n} \tag{8.4}$$

such that T_n is a decreasing sequence of stopping times that converges to T and takes only rational values.

We first prove the strong Markov property for a *rational-valued* stopping time S.

$$\begin{aligned}\mathbb{E}\left[X\mathbf{1}_{S<\infty}F_n(B_{t_1+S}-B_S,\dots,B_{t_n+S}-B_S)\right] &= \sum_{s\in\mathbb{Q}_+}\mathbb{E}\left[X\mathbf{1}_{S=s}F_n(B_{t_1+s}-B_s,\dots,B_{t_n+s}-B_s)\right] \\ &= \sum_{s\in\mathbb{Q}_+}\mathbb{E}\left[X\mathbf{1}_{S=s}\right]\mathbb{E}\left[F_n(B_{t_1},\dots,B_{t_n})\right] \\ &= \mathbb{E}\left[X\mathbf{1}_{S<\infty}\right]\mathbb{E}\left[F_n(B_{t_1},\dots,B_{t_n})\right]\end{aligned}$$

as in discrete time. The property is proved for all the approximations T_n and thus, by convergence of T_n and *right-continuity of Brownian motion*, we obtain the strong Markov property for any stopping time T. □

8.1.2 Marginal laws, scaling properties, some stopping times.

Using the independent increments properties, many marginal laws of Brownian motions can be computed. In particular, we have:

Proposition 8.1.1. *Let $(B_t)_{t\in\mathbb{R}_+}$ be a Brownian motion. The bounded-dimensional marginal laws are given by:*

$$\begin{aligned}&\mathbb{P}\left(B_{t_1}\in A_1, B_{t_2}\in A_2, \dots B_{t_n}\in A_n\right)\\&=\int_{A_1\times A_2\times\ldots\times A_n}\left(\prod_{k=1}^{n}\frac{1}{\sqrt{2\pi(t_k-t_{k-1})}}\right)\exp\left(-\sum_{k=1}^{n}\frac{(x_k-x_{k-1})^2}{2(t_k-t_{k-1})}\right)dx_1dx_2\dots dx_n\end{aligned} \tag{8.5}$$

for any sequence $0=t_0<t_1<t_2<\ldots<t_n$ with the convention $x_0=0$.

Important remark: using this expression of marginal laws and integrating it in order to compute expectation values is at the time a sure way of evaluation but often the worst one! In many cases, using the independent increments property together with expectation values is much more efficient.

Proposition 8.1.2. *Let $(B_t)_{t\in\mathbb{R}_+}$ be a Brownian motion. As a Gaussian process, it is fully characterized by its mean and its covariance. It holds:*

$$\mathbb{E}\left[B_t\right]=0 \tag{8.6}$$

$$\mathbb{E}\left[B_tB_s\right]=\min(s,t) \tag{8.7}$$

Moreover several transformations maps a Brownian motion to another Brownian motion.

Proposition 8.1.3. *Let $(B_t)_{t\in\mathbb{R}_+}$ be a Brownian motion.*

1. time translation invariance: *for all $u>0$, the centered shifted process $(B_{t+u}-B_u)_{t\in\mathbb{R}_+}$ is a Brownian motion.*
2. invariance under scaling: *for all $\alpha>0$, the renormalized process $(\alpha B_{\alpha^{-2}t})_{t\in\mathbb{R}_+}$ is a Brownian motion.*
3. invariance under reflexion: *the process $(-B_t)_{t\in\mathbb{R}_+}$ is a Brownian motion.*
4. invariance under time inversion: *the process $(tB_{1/t})_{t\in\mathbb{R}_+}$ (restricted on the set of probability 1 on which $tB_{1/t}\to 0$ as $t\to 0$) is a Brownian motion (in its own filtration).*

Proof. The scheme of proof is the same for the four invariances: we check the three point of the definition. ☐

As for simple random walks described previously, various hitting times of the Brownian motion are easy to study. The trajectories of the Brownian motion are continuous and thus, for any closed set B of $\mathbb{R}$, the hitting time is a stopping time.

Proposition 8.1.4. *Let $(B_t)_{t\in\mathbb{R}_+}$ be a Brownian motion. Let T_a be the stopping time defined by*

$$T_a=\inf\left\{t\in\mathbb{R}_+; B_t=a\right\} \tag{8.8}$$

and the process $(S_t)_{t\in\mathbb{R}_+}$ defined by

$$S_t=\sup_{0\leq s\leq t}B_s. \tag{8.9}$$

For any $a\geq 0$, it holds

$$\mathbb{P}\left(T_a\leq t\right)=\mathbb{P}\left(S_t\geq a\right)=2\mathbb{P}\left(B_t\geq a\right). \tag{8.10}$$

The r.v. S_t has the same law[1] *as $|B_t|$ and the r.v. T_a has the same law as the r.v. a^2/B_1^2 and admits the density given by*

$$\frac{a}{\sqrt{2\pi t^3}}\exp\left(-\frac{a^2}{2s}\right). \tag{8.11}$$

[1] Be careful, it may not be true as equality in law for the whole process!

Proof. The equality $\{T_a \leq t\} = \{S_t \geq a\}$ is trivial since the Brownian motion has continuous trajectories. This continuity also implies $\{B_t \geq a\} \subset \{T_a \leq t\}$ for $a \geq 0$. Then, we prove the last equality by using the strong Markov and independent increments properties:

$$\begin{aligned}\mathbb{P}(B_t \geq a) &= \mathbb{P}(B_t \geq a, T_a \leq t) = \mathbb{P}((B_t - B_{T_a}) + B_{T_a} \geq a, T_a \leq t)\\ &= \mathbb{P}(B_t - B_{T_a} \geq 0, T_a \leq t) = \mathbb{E}\left[\mathbf{1}_{T_a \leq t}\mathbf{1}_{B_t - B_{T_a} \geq 0}\right]\\ &= \mathbb{E}\left[\mathbf{1}_{T_a \leq t}\mathbf{1}_{B_{t-T_a} \geq 0}\right] = \frac{1}{2}\mathbb{P}(T_a \geq t)\end{aligned}$$

since, for any $s > 0$, $\mathbb{P}(B_s \geq 0) = 1/2$. The densities are deduced directly from the equality. □

8.1.3 Regularity and trajectories

This section is devoted to the fine study of the trajectories of a Brownian motion. We know from the definition that any trajectory $t \mapsto B_t(\omega)$ is continuous and we look for a stronger regularity. To this purpose, we focus on local properties of trajectories and we deduce from the time invariance that a study at $t = 0$ is enough. We start with the following property.

Theorem 8.1.2. *Let $(B_t)_{t\in\mathbb{R}_+}$ be a Brownian motion. It holds:*

$$\limsup_{t\to 0^+} \frac{B_t}{\sqrt{t}} = +\infty, \quad a.s.$$
$$\liminf_{t\to 0^+} \frac{B_t}{\sqrt{t}} = -\infty, \quad a.s.$$

Proof. The random variable $Y = \limsup_{t\to 0+} B_t/\sqrt{t}$ is a typical example of r.v. which is $\mathcal{F}_{0+}$-measurable where $\mathcal{F}_t$ is the filtration associated to the Brownian motion. A $0-1$ law exists for such a process with independent increments.

Lemma 8.1.1. *Let $(B_t)_{t\in\mathbb{R}_+}$ be a Brownian motion and $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ its filtration. Then, for any $B \in \mathcal{F}_{0+}$, $\mathbb{P}(B) = 0$ or 1.*

Proof. We first prove that (B_t) is also a Brownian motion in the filtration $(\mathcal{F}_{t+})_{t\in\mathbb{R}_+}$. We consider a sequence $t < t' < t_1 < t_2 < \ldots < t_n$: the independent increment property implies that $(B_{t'+s} - B_{t'})_s$ are independent of $\mathcal{F}_{t'}$ for any $t' > t$ and $t' < t_1$ and thus are independent of $\mathcal{F}_{t+}$. Moreover, the continuity of the trajectories implies that $B_t = \lim_{t'\downarrow} B_{t'}$ and we obtain the result.

We deduce from third property that $B_t = B_t - B_0$ is independent from $\mathcal{F}_{0+}$ for all $t > 0$ and thus, for all $t > 0$, $\mathcal{F}_t$ is independent from $\mathcal{F}_{0+}$, and $\mathcal{F}_{0+} = \cap_{t>0}\mathcal{F}_t$ is independent from itself, hence the lemma. □

The r.v. Y is thus almost surely constant (eventually infinite). If $Y < a$ for some a finite, then we have $\mathbb{P}\left(B_t/\sqrt{t} > a\right) \to 0$ as $t \to \infty$ (if not, Borel Cantelli lemma implies that Y is above a an infinite number of times). However, this probability is easy to evaluate:

$$\mathbb{P}\left(\frac{B_t}{\sqrt{t}} > a\right) = \int_{a\sqrt{t}}^{\infty} \frac{1}{\sqrt{2\pi t}} e^{-x^2/(2t)} dx = \int_a^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx > 0 \tag{8.12}$$

We thus conclude that $Y = +\infty$. The reflexion invariance gives immediately the second equality. □

This theorem gives a good hint of the irregularity of the Brownian: in a time as short as we wish, any trajectory goes a infinite number of times above $\sqrt{t}$ and below $-\sqrt{t}$. The choice of the square root is also the optimal choice since it is the "last" scaling in (8.12) that gives a non-zero limit.

Corollary 8.1.1. *We deduce the following properties:*

1. *A Brownian trajectory $t \mapsto B_t(\omega)$ is almost surely nowhere derivable.*
2. $\limsup_{t\to\infty} B_t/\sqrt{t} = +\infty$ *and* $\liminf_{t\to\infty} B_t/\sqrt{t} = -\infty$
3. *every point is visited a infinite number of times.*

Proof. The derivability at 0 is in contradiction with the previous theorem. Time translation invariance extends this property to any $t \in \mathbb{R}_+$. The second point is a consequence of the invariance under time inversion $t \mapsto 1/t$ after rescaling. The third point is deduced from the second through the continuity hypothesis of the trajectories. □

We have seen in the previous chapter a new notion of regularity which is the bounded-variation property, between continuity and derivability.

Property 8.1.2. *Let $(B_t)_{t\in\mathbb{R}_+}$ be a Brownian motion. It has a non-zero quadratic variation (equal to t over the interval $[0,t]$) and thus does not have a bounded total variation.*

Proof. We use the characterization with a limit over nested subdivisions. We fix a time t and a subdivision $0 = t_0 < t_1 < \ldots < t_n = t$ and we perform the following L^2 estimation:

$$\begin{aligned}\mathbb{E}\left[\left(\sum_{k=1}^{n}(B_{t_k} - B_{t_{k-1}})^2 - t\right)^2\right] &= \sum_{k=1}^{n} 3(t_k - t_{k-1})^2 + \sum_{k\neq l}(t_k - t_{k-1})(t_l - t_{l-1}) - 2\sum_{k=1}^{n}(t_k - t_{k-1})t + t^2 \\ &= \sum_{k=1}^{n} 2(t_k - t_{k-1})^2\end{aligned}$$

If we consider sequences of subdivisions $(t_k^{(N)})_{1\leq k\leq n_N}$ whose mesh size goes to 0 (i.e. $\sup_{k\in\{1,\ldots,n_N} |t_k^{(N)} - t_{k-1}^{(N)}| \to 0$ as $N \to \infty$, then the previous expectation also goes to zero. We thus have:

$$\sum_{k=1}^{n_N}(B_{t_k^{(N)}} - B_{t_{k-1}^{(N)}})^2 \xrightarrow{L^2, N\to\infty} t \tag{8.13}$$

□

8.1.4 Construction and existence

8.1.5 Numerical simulation

8.2 Itô calculus and stochastic integration

We have already seen that Brownian motion trajectories are not bounded-variation functions; however, it has a bounded quadratic variation and it may help to build quantities such as

$$\int_0^t H_s dB_s \tag{8.14}$$

where H_t is a process with some adequate hypothesis, even if it does not enter directly the construction made in the previous chapter for bounded-variation functions. Understanding how this construction works leads us to Itô calculus, which is one of the main tool to handle quantities such as $\int_0^t h(B_s)dB_s$.

This section has to be considered more as computational tools rather than general variation theory: general stochastic calculus for more general martingales is developed in the next chapter.

8.2.1 Stochastic integral

As for Riemann integral, we first build stochastic integrals for staircase functions.

Definition 8.2.1. *Let $(B_t)_{t\in\mathbb{R}_+}$ be a Brownian motion and let $(H_t)_{t\in\mathbb{R}_+}$ be an adapted process with staircase trajectories, written as:*

$$H_s = \sum_{k=1}^{n} H_{t_k} \mathbf{1}_{[t_k,t_{k+1}[}(s) \tag{8.15}$$

where (t_k) is a subdivision of $\mathbb{R}_+$. We define the stochastic integral as:

$$\int_0^t H_s dB_s = \sum_{k=1}^n H_{t_k}(B_{t_{k+1}\wedge t} - B_{t_k\wedge t}). \tag{8.16}$$

This definition coincides with the usual formula for bounded-variation functions. This definition is linear in (H_s) and we can estimate easily the first moments.

Property 8.2.1. *Let $(B_t)_{t\in\mathbb{R}_+}$ be a Brownian motion and let $(H_t)_{t\in\mathbb{R}_+}$ be an adapted process with staircase trajectories and such that, for all $s \in \mathbb{R}_+$, the r.v. H_s is square-integrable (i.e. is in L^2). Then, it holds:*

$$\mathbb{E}\left[\int_0^t H_s dB_s\right] = 0 \tag{8.17}$$

$$\mathbb{E}\left[\left(\int_0^t H_s dB_s\right)^2\right] = \mathbb{E}\left[\int_0^t H_s^2 ds\right] \tag{8.18}$$

Proof. We use the martingale property of Brownian motion:

$$\begin{aligned}\mathbb{E}\left[\int_0^t H_s dB_s\right] &= \sum_{k=1}^n \mathbb{E}\left[H_{t_k}(B_{t_{k+1}} - B_{t_k})\right] = \sum_{k=1}^n \mathbb{E}\left[\mathbb{E}\left[H_{t_k}(B_{t_{k+1}\wedge t} - B_{t_k\wedge t})\middle|\mathcal{F}_{t_k}\right]\right] \\ &= \sum_{k=1}^n \mathbb{E}\left[H_{t_k}\mathbb{E}\left[(B_{t_{k+1}\wedge t} - B_{t_k\wedge t})\right]\right] = 0\end{aligned}$$

The second computation uses the same type of decomposition:

$$\mathbb{E}\left[\left(\int_0^t H_s dB_s\right)^2\right] = \sum_{k=1}^n\sum_{l=1}^n \mathbb{E}\left[H_{t_k}H_{t_l}(B_{t_{k+1}\wedge t} - B_{t_k\wedge t})(B_{t_{l+1}\wedge t} - B_{t_l\wedge t})\right]$$

Using the independent increments property and introducing the adequate filtrations give:

$$\mathbb{E}\left[\left(\int_0^t H_s dB_s\right)^2\right] = \sum_{k=1}^n\sum_{l=1}^n \mathbb{E}\left[H_{t_k}H_{t_l}\delta_{kl}(t_{k+1}\wedge t - t_k \wedge t)\right] = \int_0^t \mathbb{E}\left[H_s^2\right] ds$$

and we prove the property. □

In order to generalize this stochastic integral, we may use the density of processes with staircase trajectories to more general processes. For example, if $(H_t)_{t\in\mathbb{R}_+}$ is an adapted *bounded* process with continuous trajectories, then the sequence of approximations

$$H_t^{(n)} = \sum_{l=0}^n H_{lT/n}\mathbf{1}_{[kT/n,(k+1)T/n[}(t) \tag{8.19}$$

converges to (H_t) on $[0,T]$ for the following convergence:

$$\forall t \in [0,T], \quad \int_0^t \mathbb{E}\left[(H_s^{(n)} - H_s)^2\right] ds \xrightarrow{n\to\infty} 0 \tag{8.20}$$

This can be proved in a straightforward way using dominated convergence theorems. The hypothesis of *boundedness* can be weakened to the hypothesis

$$\mathbb{E}\left[\int_0^T H_s^2 ds\right] < \infty \tag{8.21}$$

however the corresponding proof is a bit longer and is postponed to the next chapter for a general framework: we use it here without further proof.

Theorem 8.2.1. *Let $(B_t)_{t\in\mathbb{R}}$ be a Brownian motion and $(H_t)_{t\in\mathbb{R}_+}$ be a process adapted to the filtration associated to (B_t), with continuous trajectories and such that*

$$\mathbb{E}\left[\int_0^T H_s^2 ds\right] < \infty \tag{8.22}$$

for some T. Then the stochastic integral $\int_0^t H_s dB_s$ can be defined in a unique way for $t \leq T$ as the limit of $\int_0^t H_s^{(n)} dB_s$ for any approximation $(H_t^{(n)})$ of (H_t) in the sense of (8.20).

The following equality are then valid:

$$\mathbb{E}\left[\int_0^t H_s dB_s\right] = 0, \tag{8.23}$$

$$\mathbb{E}\left[\left(\int_0^t H_s dB_s\right)^2\right] = \mathbb{E}\left[\int_0^t H_s^2 ds\right]. \tag{8.24}$$

Proof. We admit here that such approximations exist. Then, we show that the sequence $\int_0^t H_s^{(n)} dB_s$ is a Cauchy sequence in the L^2 space and thus, by completeness of L^2 spaces, it converges to a limit (and this limit is a.s. unique). The next equalities are obtained from the similar ones proved for staircase processes. □

8.2.2 Variations and Itô calculus

8.2.2.1 Necessity

The previous theorems show the existence of r.v. defined as stochastic integrals $\int_0^t H_s dB_s$. This integral is linear in (H_s); however many standard properties of integrals such as integration by parts or differentiation are not standard anymore in their standard versions. For example, the following intuitive equalities *break down*:

$$B_t^2 \neq \int_0^t 2B_t dB_t \tag{8.25}$$

$$B_t^4 \neq \int_0^t 4B_t^3 dB_t \tag{8.26}$$

$$e^{uB_t} \neq 1 + u\int_0^t e^{uB_t} dB_t \tag{8.27}$$

If there were true, their expectations would give $t = 0$, $3t^2 = 0$ and $e^{u^2 t/2} = 1$ for any $t > 0$. However, we present here without any proof (see below) the correct versions:

$$B_t^2 = \int_0^t 2B_t dB_t + \frac{1}{2}\int_0^t 2dt \tag{8.28}$$

$$B_t^4 = \int_0^t 4B_t^3 dB_t + \frac{1}{2}\int_0^t 12B_t^2 dt \tag{8.29}$$

$$e^{uB_t} = 1 + u\int_0^t e^{uB_t} dB_t + \frac{1}{2}\int_0^t u^2 e^{uB_t} dt \tag{8.30}$$

For example, even if this not a proof, the reader may check easily from the properties of stochastic integrals that the first moments coincide. Some quick further investigations suggest that the additional terms involve *the second derivative* of the function f giving the l.h.s. $f(B_t)$.

Itô calculus gives the rules that govern these additional terms and explains their apparition (and the irrelevance of higher derivatives): this is strongly related the regularity properties of Brownian, and in particular its quadratic variation.

8.2.2.2 Itô's formula

In order to have some L^2 control of the integrals of the type $\int f(B_t)dB_t$, we require the following quadratic variation theorem, which is a generalization of (8.13).

Theorem 8.2.2. *Let $(B_t)_{t\in\mathbb{R}_+}$ be a Brownian motion and let $f:\mathbb{R}\to\mathbb{R}$ be a bounded measurable function. Let $(t_k^{(n)})$ be a sequence of subdivisions $(t_k^{(n)})_{k=0,\dots,N_n}$ (i.e. such that $t_0^{(n)}=0<t_1^{(n)}<\dots<t_{N_n-1}^{(n)}<t_{N_n}^{(n)}=t$) whose mesh size $\sup_k|t_{k+1}^{(n)}-t_k^{(n)}|$ goes to zero as $n\to\infty$. It holds,*

$$\sum_{k=0}^{N_n-1} f(B_{t_k^{(n)}})\left[(B_{t_{k+1}}-B_{t_k})^2-(t_{k+1}^n-t_k^n)\right] \xrightarrow[n\to\infty]{L^2} 0 \tag{8.31}$$

This theorem is useful since it says that every computation of the expectation value of the square (order 2) of a stochastic integral can be reduced (in L^2 to a deterministic term of order 1 in time. This reduction from an apparent second order term to a first order one explains, as we will see, the presence of a second derivative in Itô's formulae.

Proof. We note I_n the l.h.s. of (8.31), expand the square, and use the independent increments property:

$$\begin{aligned}\mathbb{E}\left[I_n^2\right] =& \sum_{k=0}^{N_n-1}\mathbb{E}\left[f(B_{t_k^n})^2\right]\mathbb{E}\left[\left[(B_{t_{k+1}}-B_{t_k})^2-(t_{k+1}^n-t_k^n)\right]^2\right]\\ &+2\sum_{0\le k<lN_n}\mathbb{E}\left[f(B_{t_k^n})f(B_{t_l^n})\left[(B_{t_{k+1}}-B_{t_k})^2-(t_{k+1}^n-t_k^n)\right]\right]\mathbb{E}\left[(B_{t_{l+1}}-B_{t_l})^2-(t_{l+1}^n-t_l^n)\right]\end{aligned}$$

Standard computations of normal r.v. give directly:

$$\begin{aligned}\mathbb{E}\left[(B_{t_{l+1}}-B_{t_l})^2-(t_{l+1}^n-t_l^n)\right]&=0\\ \mathbb{E}\left[\left[(B_{t_{k+1}}-B_{t_k})^2-(t_{k+1}^n-t_k^n)\right]^2\right]&=2(t_{k+1}^n-t_k)^2\end{aligned}$$

and thus

$$\mathbb{E}\left[I_n^2\right]=2\sum_{k=0}^{N_n-1}\mathbb{E}\left[f(B_{t_k^{(n)}})^2\right](t_{k+1}^{(n)}-t_k^{(n)})^2\le 2||f||_\infty\left(\sup_k|t_{k+1}^{(n)}-t_k^{(n)}|\right)t$$

and thus goes to 0: the L^2 convergence is proved. □

We can now state Itô's formula, which is one of the most important result of this section.

Theorem 8.2.3 (Itô's formula I)**.** *Let $(B_t)_{t\in\mathbb{R}}$ be a Brownian motion. Let $f:\mathbb{R}\to\mathbb{R}$ a function that is twice derivable and whose second derivative is continuous (i.e. f is C^2). Then,*

$$f(B_t)=f(B_0)+\int_0^t f'(B_s)dB_s+\frac{1}{2}\int_0^t f''(B_s)ds \tag{8.32}$$

Proof. We first assume that f is bounded and measurable. We expand the increment $f(B_t)-f(B_0)$ as a sum of infinitesimal increments and use, for each increment, Taylor's formula, which gives the existence of quantities $x_k^{(n)}\in[t_k^{(n)},t_{k+1}^{(n)}]$ such that:

$$\begin{aligned}f(B_t)-f(B_0)&=\sum_{k=0}^{N_n-1}\left[f'(B_{t_k^{(n)}})(B_{t_{k+1}^{(n)}}-B_{t_k^{(n)}})+\frac{1}{2}f''(x_k^{(n)})(B_{t_{k+1}^{(n)}}-B_{t_k^{(n)}})^2\right]\\ &=\sum_{k=0}^{N_n-1}\left[f'(B_{t_k^{(n)}})(B_{t_{k+1}^{(n)}}-B_{t_k^{(n)}})+f''(x_k^{(n)})(t_{k+1}^{(n)}-t_k^{(n)})^2\right]\\ &\quad+\frac{1}{2}\sum_{k=0}^{N_n-1}f''(x_k^{(n)})\left[(B_{t_{k+1}}-B_{t_k})^2-(t_{k+1}^n-t_k^n)\right]\end{aligned}$$

The r.h.s. does not depend on the subdivision that is used and thus, along a sequence of subdivisions whose mesh size goes to zero, the limit exists and the previous theorem cancels the second term: Itô's formula is proved.

We then go to general C^2 functions with density theorems that are made precise in the next chapter. □

The control of the variations of functions $f(B_t)$ of the Brownian motion at first order requires the knowledge of the first *two* derivatives of f: this requirement, as the proof shows it, uses the finite quadratic variation of Brownian motion. Another way of understanding this formula is the interpretation of the stochastic integration formula: we made the choice of $H_{t_k}(B_{t_{k+1}} - B_{t_k})$, i.e. we evaluate H_{t_k} at the beginning of the infinitesimal integration interval. However, during this interval, especially if H_{t_k} is a Brownian motion, it may have relevant variations and furthermore these variations are not independent from $(B_{t_{k+1}} - B_{t_k})$; controlling these variations is exactly what we did and it requires the second derivative of f when $H_s = f(B_s)$.

From the beginning, we could have taken other conventions in the approximation scheme used for the stochastic integral: it leads for example to Stratonovitch' calculus, which is an alternative to Itô's calculus, that lead to the same results but with different computations. In particular, the property $\mathbb{E}\left[\int H_s dB_s\right] = 0$ breaks down for other conventions.

The same proof gives the generalization to space-time C^2 functions:

Theorem 8.2.4 (Itô's formula II)**.** *Let $(B_t)_{t\in\mathbb{R}}$ be a Brownian motion. Let $f : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$, $(t,x) \mapsto f(t,x)$ a function that is twice derivable and whose second derivatives are continuous (i.e. f is C^2). Then,*

$$f(B_t) = f(B_0) + \int_0^t \frac{\partial f}{\partial t}(s, B_s)ds + \int_0^t \frac{\partial f}{\partial x}(s, B_s)dB_s + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s, B_s)ds \tag{8.33}$$

8.2.2.3 Applications

We present now several consequences of Itô's formula on the evaluation of various moments related to Brownian motion. We first consider $f(x) = x^n$ for integers $n \geq 2$ and apply Itô's formula:

$$B_t^n = n\int_0^t B_t^{n-1}dB_t + \frac{n(n-1)}{2}\int_0^t B_t^{n-2}dt \tag{8.34}$$

Expectation value maks the first term of the r.h.s. disappear. Writing $a_n(t) = \mathbb{E}\left[B_t^n\right]$ gives:

$$a_n(t) = \frac{n(n-1)}{2}\int_0^t a_{n-2}(t)dt \tag{8.35}$$

Solving this recursion is a second and complicated way of computing the moments of centered reduced normal random variables:

$$\int_{\mathbb{R}} x^n \frac{e^{-x^2/2}}{\sqrt{2\pi}}dx = \begin{cases} 0 & \text{if } n \text{ is odd} \\ \frac{n!}{2^{n/2}} & \text{if } n \text{ is even} \end{cases} \tag{8.36}$$

We consider now $f(x) = e^{ux}$. Itô's formula gives

$$e^{uB_t} - 1 = u\int_0^t e^{uB_s}dB_s + \frac{u^2}{2}\int_0^t e^{uB_t}dt \tag{8.37}$$

and expectation values are given by:

$$\mathbb{E}\left[e^{uB_t}\right] - 1 = \frac{u^2}{2}\int_0^t \mathbb{E}\left[e^{uB_t}\right]dt \tag{8.38}$$

Solving this differential equation gives the well-known result $\mathbb{E}\left[e^{uB_t}\right] = e^{tu^2/2}$.

All these computations can be unified in the following result which is a one of the key results of the bridge between PDE and stochastic calculus.

Theorem 8.2.5. *Let $u : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ be a C^2 solution of the partial differential equation*

$$\partial_t u - \frac{1}{2}\partial_x^2 u = 0 \tag{8.39}$$

such that, for all $\delta > 0$ and for all $t \geq 0$, there exist $C_{\delta,t} > 0$ such that $u(s,x) \leq C \exp \delta x^2$ for $s \in [0,t]$. Then, the quantity $f(t, B_t)$ is a martingale.

Proof. The inequality on u is only present to ensure the integrability of $f(t, B_t)$. Itô's calculus and additivity of the integral gives the conditional expectation requirement for martingales. A detailed proof is given in the next chapter. □

This theorem, together with Feynman-Kac formula, explains how Brownian motion is the probabilistic counterpart of diffusion PDE and how its properties are related to the heat kernel.

8.2.2.4 Feynman-Kac formula

The last important result we present in this section is Feynman-Kac formula, which describes solutions of diffusion PDE as well-chose expectation values over Brownian motion.

Theorem 8.2.6 (Feynman-Kac formula (a particular case))**.** *Let $g : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$, $\phi : \mathbb{R} \to \mathbb{R}$ and $V : \mathbb{R} \to \mathbb{R}_+$ be measurable functions such that there exists a function u that is continuous and has continuous derivatives and that is solution of the PDE:*

$$\partial_t u(t,x) = \frac{1}{2}\partial_x^2 u(t,x) - V(x)u(t,x) + g(t,x) \tag{8.40}$$

with initial conditions $u(0,x) = \phi(x)$. Then, $u(t,x)$ can be represented as an expectation value

$$\begin{aligned} u(t,x) =& \mathbb{E}_x \left[\phi(B_t) e^{-\int_0^t V(B_s)ds} \right] \\ &+ \mathbb{E}_x \left[\int_0^t g(t-s, B_s) e^{-\int_0^s V(B_u)du} ds \right] \end{aligned} \tag{8.41}$$

where $(B_t)_{t \in \mathbb{R}_+}$ is a Brownian motion started[2] at x.

Before the proof, several remarks must be done.

First, we do not give precise hypotheses about the functions g, c and ϕ and invite the reader to refer to any reference about differential equations; in most cases, we have explicit functions and shortcuts can be used to obtain the existence of u easily.

The relation between stochastic evolution of the Brownian motion and the PDE evolution is non-trivial: $u(t,x)$ is the value of $u(t,x)$ at x *after* the evolution time t, whereas x appears in the r.h.s. as the *initial* point of the Brownian motion. Conversely, ϕ appears the initial condition of the PDE, whereas ϕ is functional of B_t (and not B_0). There is an underlying time reversal, which is indeed used in the proof below.

Proof. The main difficulty of this proof is to recognize functions of Brownian motion such that Itô's formula can be applied. We fix $t > 0$ and introduce the time-reversed solution of (8.40):

$$u_*(s,x) = u(t-s,x)$$

The process defined as

$$V_s = e^{-\int_0^s V(B_u)du}$$

does not require any Itô's formula since it does not contain any stochastic integral but only an integral over Brownian motion: we check that V_s has *derivable* trajectories and is a bounded

[2] i.e. it can be written as $x + W_t$ where W_t is a Brownian motion starting at 0, as constructed previously.

variation process: the formalism of stochastic integral is useless and standard variational calculus apply (in particular $V_s - V_0 = \int_0^t V_s' ds$ and $V_s' = V(B_s)V_s$). The interested reader may check easily[3] or read in the next chapter (or prove the very simple case $V = 0$) that Itô's formula can be generalized as:

$$u_*(s,B_s)V_s - u_*(0,x)V_0 = \left(\int_0^t \partial_x u_*(s,B_s)V_s dB_s + \int_0^t \partial_t u_*(s,B_s)V_s ds + \int_0^t u_*(s,B_s)V_s' ds\right) + \frac{1}{2}\int_0^t \partial_x^2 u_*(s,B_s)V_s ds$$

Reordering the terms of r.h.s. gives:

$$u_*(s,B_s)V_s - u_*(0,x)V_0 = \int_0^t \partial_x u_*(s,B_s)V_s dB_s + \int_0^t V_s \left[\partial_t u_*(s,B_s) + \frac{1}{2}\partial_x^2 u_*(s,B_s) + u_*(s,B_s)V(B_s)\right] ds$$

Using the time-reversal and (8.40) gives

$$u_*(s,B_s)V_s - u_*(0,x)V_0 = \int_0^t \partial_x u_*(s,B_s)V_s dB_s - \int_0^t V_s g(t-s,B_s) ds$$

The expectation of the first term of the r.h.s. is zero by construction of the stochastic Itô integral. We have $u_*(0,x)V_0 = u(t,x)$ and thus the final formula:

$$u(t,x) = \mathbb{E}_x\left[u(0,B_s)V_s\right] + \mathbb{E}_x\left[\int_0^t V_s g(t-s,B_s) ds\right]$$

□

An easy example. We look for a solution u of the PDE

$$\begin{cases} \partial_t u(t,x) = \frac{1}{2}\partial_x^2 u(t,x) + x^2 t^5 \\ u(0,x) = 0 \end{cases}$$

that is continuous and with continuous derivatives. We first apply Feynman-Kac formula and represent u as:

$$\begin{aligned} u(t,x) &= \mathbb{E}_x\left[\int_0^t (t-s)^5 B_s^2 ds\right] = \int_0^t (t-s)^5 \mathbb{E}_0\left[(x+W_s)^2\right] ds = \int_0^t (t-s)^5 \mathbb{E}_0\left[(x+W_s)^2\right] ds \\ &= \int_0^t (t-s)^5 (x^2+s) ds = \frac{x^2 t^6 + t^7}{6} \end{aligned}$$

The previous example is very simple and gives a result which could be even guessed! In general, the philosophy behind the use of Feynman-Kac formula is the reduction of a PDE problem to Brownian questions, with the belief that enough things are known about Brownian motion to solve this question. In particular, one could recover many results about PDE from probability theorems.

This Feynman-Kac formula is a particular case of a class of general results relating integro-differential linear equations and probabilistic diffusions: the reader is more invited to understand how this formula was proved, rather than the single result.

The case of PDE with *Dirichlet boundary conditions* can also be integrated to such a formalism and leads to *stopped* processes and the reader is encouraged to refer to specialized books about the subject. The next chapter, which is about general stochastic calculus, does not contain an exploration of such applications.

The case of some non-linear PDE (there are positivity constraints on the non-linearity) can be integrated in this formalism but, in this case, it requires Markov processes indexed by trees and not by $\mathbb{R}_+$ anymore.

---

[3]Exercise: adapt the proof of Itô's formula to a product $f(s,B_s)H_s$ where H_s has continuous derivable trajectories.

Non-trivial application to a PDE upper bound problem. We consider the PDE

$$\partial_t u = \frac{1}{2}\partial_x^2 u - V(x)u \tag{8.42}$$

with $V \geq 0$ such that u is continuous with continuous derivative and such that $(0, x) = 1$. We wish a bound on $u(t, x)$ at large time. To this purpose, we write u as

$$u(t, x) = \mathbb{E}\left[\exp^{-\int_0^t V_x(B_s)ds}\right] \tag{8.43}$$

where $f_x(\cdot) = f(x + \cdot)$. We introduce variables $d(x, t) \geq 0$ and we want to know whether the Brownian motion went out of the domain $[-d(x, t), d(x, t)]$ before time t. We thus have

$$\begin{aligned} u(t, x) =& \mathbb{E}\left[\exp^{-\int_0^t V_x(B_s)ds} \mathbf{1}_{\sup_{0\leq s\leq t}|B_t|>d(x,t)}\right] \\ &+ \mathbb{E}\left[\exp^{-\int_0^t V_x(B_s)ds} \mathbf{1}_{\sup_{0\leq s\leq t}|B_t|\leq d(x,t)}\right] \end{aligned} \tag{8.44}$$

We now control both terms. On the first event, we have no control on the trajectory: we use only $V(x) \geq \inf_{\mathbb{R}} V = V_1$ but wish that this event is rare: it is indeed the case from the maximal inequality

$$\mathbb{P}\left(\sup_{0\leq s\leq t} |B_s| \geq d(x, t)\right) \leq \frac{3}{d(x, t)} \sup_{0\leq s\leq t} \mathbb{E}\left[|B_s|\right] = \frac{C\sqrt{t}}{d(x, t)} \tag{8.45}$$

On the second event, we need only to minimize V over $[x - d(x, t), x + d(x, t)]$ (and call it $m_V(x, d(x, t))$) but this event may not be rare and we bound its probability by 1. We then obtain:

$$u(t, x) \leq \frac{C\sqrt{t}}{d(x, t)} + \exp\left(-t m_V(x, d(x, s))\right) \tag{8.46}$$

for any $d(x, t)$. Depending on the hypothesis on V (behaviour at ∞), one may choose some optimal choice of $d(x, t)$ and obtain a bound about the variation of u.