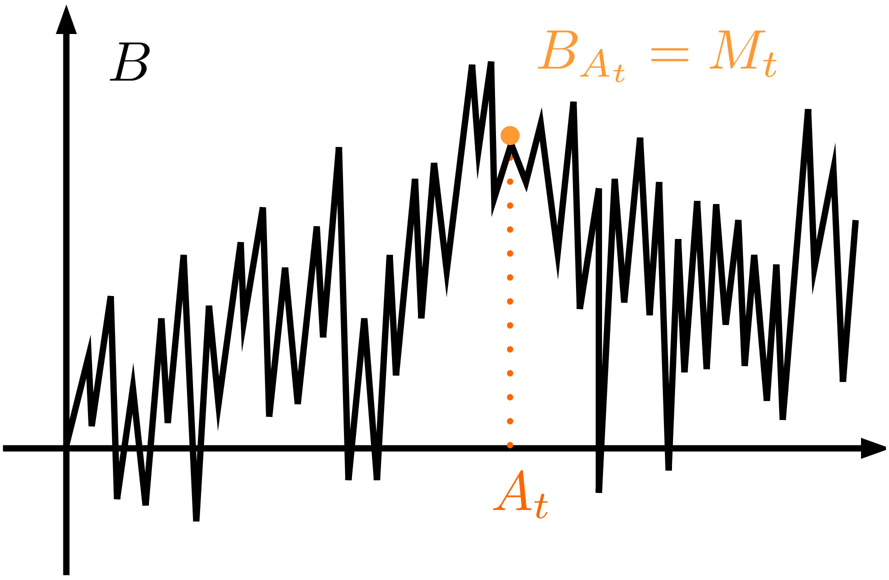
University of Zurich



Brownian Motion

Marco Bertenghi

Based on the lecture notes of
Prof. Wendelin Werner (ETHZ)

November 16, 2018

Contents

1 Construction of a Brownian motion

1.1 Goal

We want to construct a one-dimensional Brownian Motion, that is, we want to find a probability space and a collection of random variables $(B_t)_{t\geq 0}$ (t is indexed over $\mathbb{R}_+$) such that:

1. $B_0 = 0$ almost surely.
2. B has independent and stationary increments.
   (a) Stationary increments: For all $t, h \geq 0$, we have $B(t+h) - B(t) \sim B(h)$.
   (b) Independent increments: For all positive integers k and times $0 \leq t_1 < t_2 < \cdots < t_k$ in $\mathbb{R}_+$, the k increments $B(t_1) - B(0), B(t_2) - B(t_1), \ldots, B(t_k) - B(t_{k-1})$ are independent random variables.
3. $B_t \sim \mathcal{N}(0, t)$ for all $t \in \mathbb{R}_+$.
4. There exists a set of probability 1, such that on this set $t \mapsto B_t$ is a continuous function. *(Hardest part in the proof)*

1.2 Preliminaries

Surprisingly, we only require a bare minimum of prerequisites in order to perform the construction of a Brownian motion. Namely:

1. There exists a probability space on which we can define a countable family of independent centered normal variables, i.e. a collection $N_j \sim \mathcal{N}(0,1)$ for all $j \in \mathbb{N}$.
2. If $N \sim \mathcal{N}(0,1)$, then $aN \sim \mathcal{N}(0, a^2)$ for all constants a.
3. $X \sim \mathcal{N}(0, \sigma_X^2)$, $Y \sim \mathcal{N}(0, \sigma_Y^2)$ and if X and Y are independent, then $X+Y \sim \mathcal{N}(0, \sigma_X^2 + \sigma_Y^2)$.
4. If $N \sim \mathcal{N}(0, \sigma^2)$ and $N' \sim \mathcal{N}(0, \sigma^2)$ and N, N' are independent, then the two random variables

$$\frac{N + N'}{2} \text{ and } \frac{N - N'}{2}$$

   are independent with common law $\mathcal{N}(0, \sigma^2/2)$, notice that their sum adds up to N. This can be seen as a *decomposition law.*

1.3 Construction

1.3.1 Notation and warm-up

For every $n \in \mathbb{N}$ we set $\mathcal{I}_n := \{[j2^{-n}, (j+1)2^{-n}] : j \in \mathbb{N}\}$ and we define $\mathcal{I} := \cup_{n\in\mathbb{N}}\mathcal{I}_n$ (everything nice and countable so far). An element of $\mathcal{I}$ is therefore an interval $I_{j,n} = [j2^{-n}, (j+1)2^{-n}]$ for some n, j. For $I \in \mathcal{I}$ we set $r(I)$ and $l(I)$ to be closed right half, respectively the closed left half of I.

For clarity purposes, here are two examples: Let $I_{j,n} = [j2^{-n}, (j+1)2^{-n}]$ as above, then

$$\begin{aligned} l(I_{j,n}) &= I_{2j,n+1} = [j2^{-n}, (2j+1)2^{-n-1}] \\ r(I_{j,n}) &= I_{2j+1,n+1} = [(2j+1)2^{-n-1}, (j+1)2^{-n}] \end{aligned}$$

For the rest of the construction we consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ that contains a countable collection of independent centered Gaussian random variables, with

$$\left.\begin{aligned} N_j &\sim \mathcal{N}(0,1) \text{ for all } j \in \mathbb{N}. \\ N_{j,n} &\sim \mathcal{N}(0,2^{-n}) \text{ for all } j,n \in \mathbb{N}. \end{aligned}\right\} \text{Independent}$$

1.3.2 Construction on integer times

We first construct the values of the Brownian Motion at **integer times**. This is easy, we simply define the values of the process A (that we want to be a Brownian motion) as follows: $A(0) = 0$, $A(1) = N_1$,$A(2) = A(1) + N_2 = N_1 + N_2$ and $A(j) = A(j-1) + N_j = N_1 + \cdots + N_j$. In particular we have $A(j) \sim \mathcal{N}(0, j)$, for every increment $A(j) - A(j-1) = N_j \sim \mathcal{N}(0,1) \sim A(1)$ and naturally all increments are independent because the N_j's are, as desired.

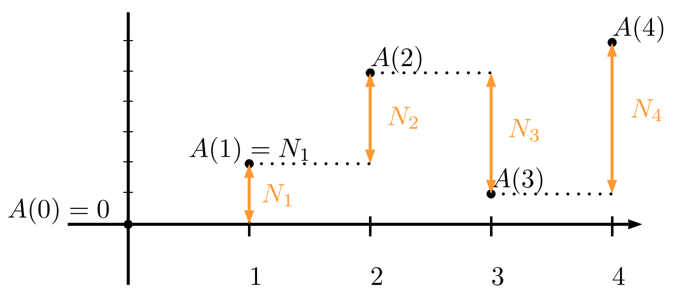


Figure 1: Construction on integer values with increments N_i displayed.

1.3.3 Construction on half integers times

Next we construct the values of the Brownian motion at **half integers times**. We introduce the notation Δ_I, which will denote what will be the increment of the Brownian Motion on the interval I.

We look at each interval $I_{j,0} = [j, j+1]$ separately, we already know the values of the Brownian Motion on such intervals, we introduce the more convenient notation $\Delta_{j,0} := \Delta_{I_{j,0}}$ and we already know that

$$\Delta_{j,0} = A(j+1) - A(j) = N_j \sim \mathcal{N}(0,1).$$

For $N_{j,0} \sim \mathcal{N}(0, 2^{-0}) = \mathcal{N}(0,1)$ we know that

$$\frac{\Delta_{j,0} + N_{j,0}}{2} \text{ and } \frac{\Delta_{j,0} - N_{j,0}}{2}$$

are two independent RV with common law $\mathcal{N}(0, 2^{-1})$ whose sum is $\Delta_{j,0}$.

We now define:

$$A\left(j + \frac{1}{2}\right) := A(j) + \left(\frac{\Delta_{j,0} + N_{j,0}}{2}\right) \sim \mathcal{N}\left(0, j + \frac{1}{2}\right). \tag{*}$$

This definition guarantees now that for the left respectively right half of $[j, j+1]$ we get:

$$\begin{aligned}
\Delta_{[j,j+1/2]} &= A(j+1/2) - A(j) = \frac{\Delta_{j,0} + N_{j,0}}{2} \sim \mathcal{N}(0, 2^{-1}), \\
\Delta_{[j+1/2,j+1]} &= A(j+1) - A(j+1/2) \\
&= A(j+1) - A(j) - \frac{\Delta_{j,0} + N_{j,0}}{2} \\
&= \Delta_{j,0} - \frac{\Delta_{j,0} + N_{j,0}}{2} = \frac{\Delta_{j,0} - N_{j,0}}{2} \sim \mathcal{N}(0, 2^{-1}).
\end{aligned}$$

So we have again $A(j+1/2) \sim \mathcal{N}(0, j+1/2)$, all increments satisfy moreover that $A(j+1/2) - A(j) \sim \mathcal{N}(0, 1/2) \sim A(1/2)$ are therefore stationary and by the above calculations of the decomposition we also have that the increments are independent.

Remark: Notice that the equation (*) above can be rewritten as

$$A\left(j + \frac{1}{2}\right) = \frac{A(j) + A(j+1)}{2} + \frac{N_{j,0}}{2}$$

this equation gives rise to a more geometric interpretation of our construction, see figure 2.

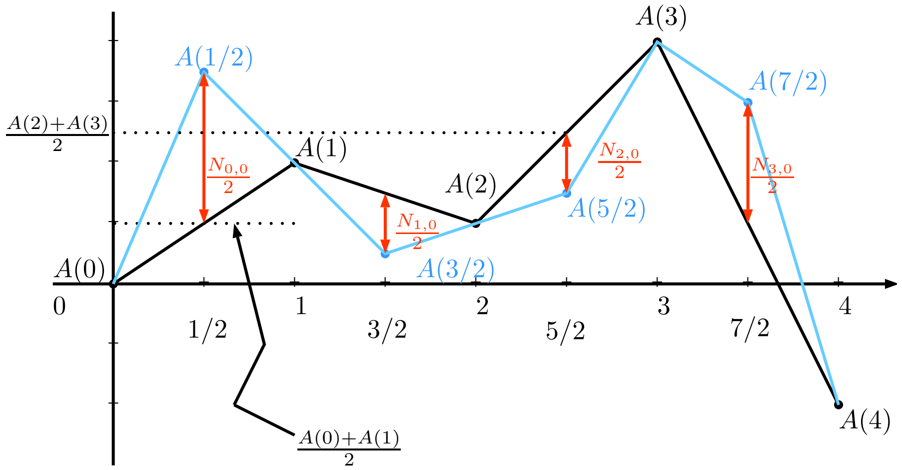


Figure 2: Geometric interpretation of our construction on the half integers

Notice that the blue and the black lines are merely for illustrative purposes. Moreover it's important to remark that $\frac{1}{2}N_{i,0} \sim \mathcal{N}(0, 1/4)$ for all $i \in \mathbb{N}$ and also $((A(i) + A(i+1))/2 \sim \mathcal{N}(0, i + 1/4)$.

1.4 Inductive construction on dyadics

We iterate over $n \in \mathbb{N}$. Let us assume that we know (for a fixed $n \in \mathbb{N}$) the values of $A(j2^{-n})$ for all $j \in \mathbb{N}$. In particular, we know the increments $\Delta_{j,n}$ for all $j \in \mathbb{N}$ and all $(\Delta_{j,n})_{j\geq 0}$ are i.i.d. with common law $\mathcal{N}(0, 2^{-n})$.

We then take again $(N_{j,n})_{j\geq 0}$ i.i.d. $\mathcal{N}(0, 2^{-n})$ and independent of $(\Delta_{j,n})_{j\geq 0}$. We then define for all intervals $I = I_{j,n} = [j2^{-n}, (j+1)2^{-n}]$ the increments $\Delta_{2j,n+1}$, $\Delta_{2j+1,n+1}$ on $l(I)$ respectively on $r(I)$ by:

$$\Delta_{2j,n+1} := \frac{\Delta_{j,n} + N_{j,n}}{2}, \ \Delta_{2j+1,n+1} := \frac{\Delta_{j,n} - N_{j,n}}{2}.$$

The previous arguments (same as on half integers) shows that this time $(\Delta_{j,n+1})_{j\geq 0}$ are i.i.d. $\mathcal{N}(0, 2^{-(n+1)})$. Also, just as in the previous step we define

$$A(\text{middle}(I_{j,n})) := A(j2^{-n}) + \Delta_{2j,n+1} \sim \mathcal{N}(0, (j+1/2)2^{-n}).$$

With this construction we have now a collection of random variables $(A(q))_{q\in\mathcal{D}}$ where $\mathcal{D} = \{j2^{-n}, j \geq 0, n \geq 0\}$ denotes the set of the dyadics. We have the properties $A(0) = 0$ almost surely, $A(q) \sim \mathcal{N}(0, q)$ on $\mathcal{D}$ and it has stationary independent increments on $\mathcal{D}$.

1.5 Continuous extension to $\mathbb{R}_+$

We will see that almost surely the function $\mathcal{D} \ni q \mapsto A(q)$ can be extended into a continuous function $t \mapsto \tilde{A}(t)$ on $\mathbb{R}_+$ and we will show that indeed $t \mapsto \tilde{A}(t)$ is a Brownian motion.

We define for all $n \in \mathbb{N}$, the function $f_n(t)$ where $t \geq 0$ to be the linear interpolation of $A(0), A(1 \cdot 2^{-n}), A(2 \cdot 2^{-n}), \ldots, A(j2^{-n}), \ldots$ (i.e. linear on each $I_{j,n}$).

Our goal is to show that f_n converges uniformly to some continuous function f on any compact interval $[0, K]$, where K is an integer.
In order to achieve this goal, we notice that the function $f_{n+1} - f_n$ is

- a continuous linear function on each interval $I_{j,n+1}$.
- equal to 0 at each $j/2^n$ for all j.
- equal to $N_{j,n}/2$ at each middle point of $I_{j,n}$.

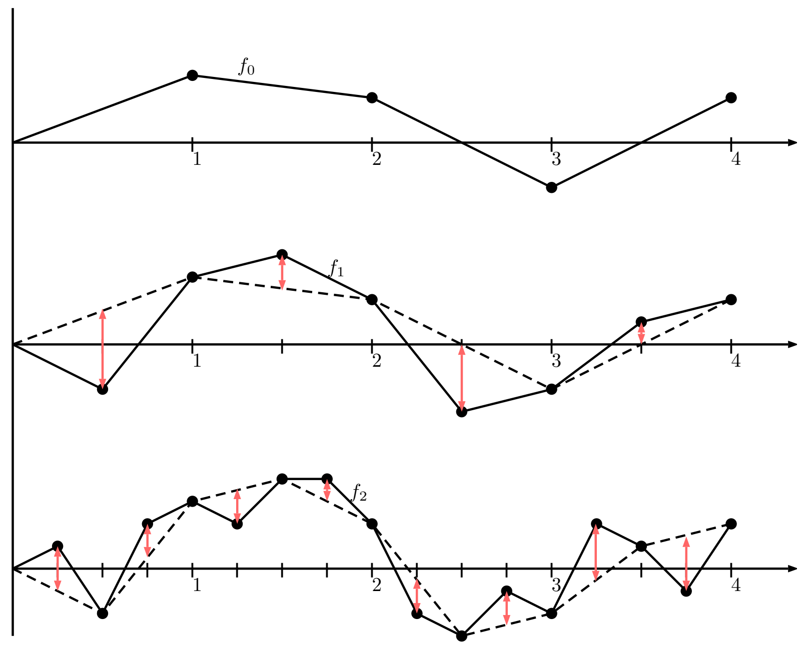


Figure 3: Illustration of the iterative construction: First on the integers, then on the half-integers and so on. The increment of f_n on the interval $I_{j,n} = [j2^{-n}, (j+1)2^{-n}]$ is $\Delta_{j,n}$. The difference $f_{n+1} - f_n$ at the middle point of each $I_{j,n}$ is $N_{j,n}/2$ (red), and at each end point of $I_{j,n}$ is zero.

With figure 3 in mind we obtain for all integers K and $x \geq 1$,

$$\begin{aligned}\mathbb{P}(\max_{[0,K]} |f_{n+1} - f_n| \geq x2^{-n/2}) &\overset{1)}{=} \mathbb{P}(\exists j \leq K2^n - 1 : |N_{j,n}| \geq 2x2^{-n/2}) \\ &\overset{2)}{\leq} \sum_{j \leq K2^n - 1} \mathbb{P}(|N_{j,n}| \geq 2x2^{-n/2}) \overset{3)}{=} K2^n \mathbb{P}(|N_1| \geq 2x) \overset{4)}{\leq} K2^n e^{-2x^2}.\end{aligned}$$

Where we used:

1. If the maximum is attained, then it's attained at the mid point of the intervals (see figure 3) where its value is $N_{j,n}/2$ (it cannot be at the start/end points because there $f_{n+1} - f_n \equiv 0$). Moreover we translate the interval $[0, K]$ for $I_{j,n} = [j2^{-n}, (j+1)2^{-n}]$ i.e. j must run from 0 to $K2^n - 1$.

2. Union bound.

3. We have $N_{j,n}$ i.i.d. $\mathcal{N}(0, 2^{-n})$ which implies that $2^{n/2}N_{j,n} \sim \mathcal{N}(0,1)$ and still i.i.d. so we just take $N_1 \sim \mathcal{N}(0,1)$ (independent of j).

4. Very crude upper bound for a standard Gaussian, easy to show.

If we now choose $x = n \in \mathbb{N}_{\geq 1}$ in our estimate above, we notice that the upper bound is summable, i.e. $\sum_{n \geq 1} \mathbb{P}(\max_{[0,K]} |f_{n+1} - f_n| \geq n2^{-n/2}) < \infty$. By the Borel-Cantelli Lemma, we can conclude that almust surely, there exists a (random) $n_0 = n_0(\omega)$ such that for all $n \geq n_0$ we have

$$|f_{n+1} - f_n| \leq n2^{-n/2} \text{ on } [0, K].$$

Since $\sum n2^{-n/2} < \infty$, it is very easy to conclude from the above that $(f_n)_{n \geq 1}$ is uniformly Cauchy on $[0, K]$ (since the sum over all increments $|f_{n+1} - f_n|$ converges, we must have for large enough p, n that $|f_p - f_n|$ converges to 0).

Since $(f_n)_{n \geq 1}$ is uniformly Cauchy on $[0, K]$ we know that it must converge uniformly on the interval $[0, K]$ as $n \to \infty$ to some function f, and that this limit f is almost surely a continuous function on $[0, K]$.

We remark that this is true for each integer K, so we can interchange "almost surely" and "for all K" (recall that $\forall k, \mathbb{P}(A_k)$ is 'weaker' than $\mathbb{P}(\forall k, A_k) = \mathbb{P}(\cap_k A_k)$) to conclude that: *Almost surely, the function f_n converges uniformly on any compact subset of $\mathbb{R}_+$ to a limiting function f, and this limiting function f is continuous on $\mathbb{R}_+$.* We used that $\forall k \geq 0$ (integer i.e. countable) $\mathbb{P}(A_k^c) = 0$ then $\mathbb{P}(\cup_{k \in \mathbb{N}} A_k^c) \leq \sum_{k \in \mathbb{N}} \mathbb{P}(A_k^c) = 0$ i.e. $\mathbb{P}(\cap_{k \in \mathbb{N}} A_k) = 1$.

1.5.1 The continuous extension is a Brownian motion

We will now prove that the law of the stochastic process $(f(t))_{t\geq 0}$ that has been constructed in the previous section is indeed that of a Brownian motion. We do already know that it is continuous and that $f(0) = 0$ almost surely. It remains to check that for any $0 = t_0 < t_1 < \cdots < t_k$, the random variables

$$f(t_1) - f(t_0), f(t_2) - f(t_1), \ldots, f(t_k) - f(t_{k-1})$$

are independent, and that the law of $f(t+h) - f(t)$ is a centered Gaussian variable with variance h.

To establish this, we first choose for each time t_l where $l = 1, \ldots, k$, a sequence $t_l(n)$ of dyadic times that converges to t_l. (For instance, we take $t_l(n)$ to be the smallest multiple of 2^{-n} that is larger than t_l). Because of the continuity of f, we know that almost surely, $f(t_l(n)) \to f(t_l)$ as $n \to \infty$. This readily implies that the random variables

$$f(t_1) - f(t_0), f(t_2) - f(t_1), \ldots, f(t_k) - f(t_{k-1})$$

are independent as the almost sure limits of the independent random variables

$$f(t_1(n)), f(t_2(n)) - f(t_1(n)), \ldots, f(t_k(n)) - f(t_{k-1}(n)).$$

Moreover we know that $f(t_l(n)) - f(t_{l-1}(n))$ is a centered Gaussian with variance $t_l(n) - t_{l-1}(n)$. Furthermore, $t_l(n) - t_{l-1}(n)$ converges to $t_l - t_{l-1}$ as $n \to \infty$. Therefore we get for all $\lambda \in \mathbb{R}$, that

$$\begin{aligned} \mathbb{E}(\exp(i\lambda(f(t_l) - f(t_{l-1}))) &= \lim_{n\to\infty} \mathbb{E}(\exp(i\lambda(f(t_l(n)) - f(t_{l-1}(n)))) \\ &= \lim_{n\to\infty} \exp(-\lambda^2/(2(t_l(n) - t_{l-1}(n)))) = \exp(-\lambda^2/(2(t_l - t_{l-1}))). \end{aligned}$$

This establishes that $f(t_l) - f(t_{l-1})$ is indeed a centered Gaussian random variable with variance $t_l - t_{l-1}$. This concludes the construction and shows that the process f is indeed a Brownian motion.

Remark 1.1. We see that this last part of the proof is in essence nothing more than using that the set of dyadic times $\mathcal{D}$ is dense in $\mathbb{R}_+$, i.e. we can approximate times $t \in \mathbb{R}_+$ by times $\tilde{t}$ in $\mathcal{D}$.

1.6 Kolmogorov's continuity criterion

Since it matches the current presentation quite well, we will also briefly discuss Kolmogorov's continuity criterion. The underlying question one should have in mind is *if continuity can be read off from the law of a stochastic process or not.* Certainly the answer is no, because it is quite easy to construct two stochastic processes which have the same law, one being continuous (always equal to 0) and the other having jumps.

However, we do add quite a bit more subtlety to this question if we consider *continuous modifications* of a stochastic process. Assume we are given the law of a stochastic process $(X_t)_{t\in I}$. We can then check whether it is possible to find some probability space, and some random process $(Y_t)_{t\in I}$ with this given law, such that there exists a measurable set with probability 1, such that for all ω in this set, $t \mapsto Y_t$ is continuous on I. It is possible to show that for all $t \in I$ we have $X_t = Y_t$ almost surely (i.e. $\forall t \in I,\ \mathbb{P}(X_t = Y_t) = 1$), we then say that the process Y is a continuous modification of X.

Theorem 1.1 (Kolmogorov's continuity criterion). *If $(X_t)_{t\geq 0}$ is a stochastic process such that for all $T > 0$, there exists $\epsilon > 0, \alpha > 0, C > 0$ such that for all $t, s \in [0,T]$ we have*

$$\mathbb{E}(|X_t - X_s|^\alpha) \leq C|t-s|^{1+\epsilon}, \tag{ς}$$

then X admits a continuous modification.

Remark 1.2. Before we write down the proof, it is important to mention that we follow very similar arguments as in the construction of a Brownian motion to establish the proof.

Proof. Let us assume that $Z_q = X_q$ for all dyadics $q \in \mathcal{D}$. Let us try to bound $|Z_{(j+1)2^{-n}} - Z_{j2^{-n}}|$ for all $j \leq 2^n T - 1$.

$$\mathbb{P}(\exists j \leq T2^n - 1 : |Z_{(j+1)2^{-n}} - Z_{j2^{-n}}| \geq x_n) \leq \sum_{j\leq T2^n-1} \mathbb{P}(|X_{(j+1)2^{-n}} - X_{j2^{-n}}|^\alpha \geq x_n^\alpha)$$

$$\overset{\text{Markov}}{\leq} \sum_{j\leq T2^n-1} \frac{1}{x_n^\alpha}\mathbb{E}(|X_{(j+1)2^{-n}} - X_{j2^{-n}}|^\alpha) \overset{\varsigma}{\leq} T2^n C\frac{(2^{-n})^{1+\epsilon}}{x_n^\alpha} = TC\frac{2^{-n\epsilon}}{x_n^\alpha} = TC2^{-n\epsilon/2}$$

Where in the last step we specialized on $x_n = 2^{-n\epsilon/(2\alpha)}$. Since $\sum 2^{-n\epsilon/2} < \infty$, we have by the Borel-Cantelli Lemma that almost surely, there exists $n_0(\omega)$ such that for all $n \geq n_0(\omega), j \leq T2^n - 1$ we have

$$|X_{(j+1)2^{-n}} - X_{j2^{-n}}| < 2^{-n\epsilon/(2\alpha)}$$

We now define f_n to be the linear interpolation of the values of X on the dyadics $\mathcal{D}$ (multiples of 2^{-n}).

$\max(|f_{n+1} - f_n|)$ on $[0, T]$ is attained at one of the middle points of $I_{j,n}$ and is therefore bounded by the maximum over all $j \leq 2^{(n+1)}T - 1$, thus almost surely there exists $n_0(\omega)$, for all $n \geq n_0$ such that

$$\max_{[0,T]} |f_{n+1} - f_n| \leq 2^{-\frac{(n+1)\epsilon}{2\alpha}}.$$

Since $\sum 2^{-n\epsilon/(2\alpha)} < \infty$, we get that the series $\sum |f_{n+1} - f_n|$ is uniformly summable and thus $(f_n)_{n\in\mathbb{N}}$ is uniformly Cauchy on $[0, T]$. Therefore there exists a continuous function f on $[0, T]$ such that $f_n \to f$ uniformly on $[0, T]$.

We now claim that for all fixed $t > 0$ we have $f(t) = X(t)$ almost surely, that is f is a continuous modification of X. To this extent let $q_n(t)$ be a sequence of dyadics that converges to t. Since f is almost surely continuous at t we have (almost surely)

$$f(t) = \lim_{n\to\infty} f(q_n(t)) = \lim_{n\to\infty} X(q_n(t)). \tag{*}$$

On the other hand, by (ς) we also have for all $\delta > 0$

$$\begin{aligned}\mathbb{P}(|X(t) - X(q_n(t))| > \delta) &\leq \delta^{-\alpha}\mathbb{E}(|X(t) - X(q_n(t))|^\alpha) \\ &\leq C\delta^{-\alpha}|t - q_n(t)|^{1+\epsilon} \overset{n\to\infty}{\longrightarrow} 0,\end{aligned}$$

so we have $X(q_n(t)) \to X(t)$ in Probability as $n \to \infty$. Since the almost sure convergence in (*) above implies convergence is Probability and the limit of convergence in probability is unique we must have that $f(t) = X(t)$ almost surely. $\square$

Lemma 1.1. *Kolmogorov's continuity criterion is satisfied by the law of Brownian motion.*

Proof. We know that $B_t - B_s \sim \mathcal{N}(0, t - s)$ for all $t > 0, s < t$. We first notice that $\mathbb{E}((B_t - B_s)^2) = |t - s|^1$, so we have no chance for this exponent to satisfy Kolmogorov's continuity criterion. However, we can easily see that $\sqrt{t - s}B_1 \sim \mathcal{N}(0, t - s)$ and consequently

$$\mathbb{E}((B_t - B_s)^4) = \mathbb{E}((\sqrt{t - s}B_1)^4) = |t - s|^2 \underbrace{\mathbb{E}(B_1^4)}_{<\infty}$$

So we choose $\epsilon = 1, \alpha = 4, C = \mathbb{E}(B_1)^4$ and conclude by Kolmogorv's continuity criterion. $\square$

Remark 1.3. It is possible to adapt the proof of Kolmogorov's continuity criterion to show that if the conditions are satisfied, then the modification will not only be continuous, but also γ-Hölder continuous for any $\gamma < \epsilon/\alpha$ in the sense that for all $T > 0$, there almost surely exists $C = C(T, \gamma)$ such that for all $0 \leq s < t \leq T$ we have

$$|B_t - B_s| \leq C|t-s|^\gamma.$$

Since for Brownian motion (using the same argument as in the proof of the lemma before), one can take $\alpha = 2k$ and $\epsilon = k - 1$ this shows that Brownian motion is almost surely Hölder continuous of exponent γ for all $\gamma < 1/2$. We will later establish that Brownian motion is not Hölder of exponent $1/2$.

1.7 Brownian motion as a Gaussian process, and consequences

Before we give another approach to constructing a Brownian motion, which will be more from a Functional Analytic point of view, it will be fruitful to introduce Gaussian processes in order to describe Brownian motion.

Definition 1.1. *A random vector $(X_1, \ldots, X_n) \in \mathbb{R}^n$ is a centered Gaussian vector if any linear combination of the X_i's is a centered Gaussian random variable.*

We can make some useful remarks (that should be reminders):

- If $N_1, \ldots, N_k$ are i.i.d. centered Gaussian random variables, then any vector whose entries are (fixed) linear combinations of the $N_1, \ldots, N_k$ is a centered Gaussian vector. In other words, if there exists Q a $n \times n$ Matrix such that $(X_1, \ldots, X_n) = (N_1, \ldots, N_d) \cdot Q$, then $(X_1, \ldots, X_n)$ is a Gaussian vector.
- A nice useful property is that the law of a centered Gaussian vector X is completely characterized by its covariance matrix $\sum_X = (\mathbb{E}(X_i X_j))_{1 \leq i,j \leq n}$.
- If $(X_1, \ldots, X_n)$ is a centered Gaussian vector and if for some $n_0 < n$, one has $\mathbb{E}(X_i, X_j) = 0$ for all $1 \leq i \leq n_o < j \leq n$, then the vectors $(X_1, \ldots, X_{n_0})$ and $(X_{n_0+1}, \ldots, X_n)$ are independent.

Definition 1.2. *A stochastic process $(X_t)_{t \in I}$ is said to be a centered Gaussian process for all $n \in \mathbb{N}$ and for all $t_1, \ldots, t_k \in I$ the finite dimensional vector $(X_{t_1}, \ldots, X_{t_n})$ is a Gaussian vector. The covariance function of a stochastic process $X = (X_t)_{t \in I}$ is the function $\sum_X$ defined on $I \times I$ by $\sum_X(s,t) = \mathbb{E}(X_s X_t)$.*

Remark 1.4. We know that the finite-dimensional distributions of X are completely characterized by $\sum_X$, so that the law of the whole process is also characterized by the covariance function $\sum_X$. *(Recall that the law of a stochastic process is determined by its finite-dimensional distributions).*

The important statement in this subsection is the following:

Proposition 1.1. *A Brownian Motion $(B_t)_{t \geq 0}$ is a (centered) Gaussian process with covariance function given by*

$$\Sigma_B(s,t) = \mathbb{E}(B_t B_s) = \min(s,t).$$

Proof. Let us choose $t_1 < t_2 < \cdots < t_k$ for $k \in \mathbb{N}$ and $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$. We want to show that $\lambda_1 B_{t_1} + \cdots + \lambda_k B_{t_k}$ is a centered Gaussian variable. This is easy because we can write this as a linear combination of the independent centered Gaussian variables

$$B_{t_1}, (B_{t_2} - B_{t_1}), \ldots, (B_{t_k} - B_{t_k - 1})$$

using telescopic sums for example. Let us now assume that $s < t$, then

$$\begin{aligned}\mathbb{E}(B_t B_s) &= \mathbb{E}((B_s + B_t - B_s)B_s) = \mathbb{E}(B_s^2) + \mathbb{E}((B_t - B_s)B_s) \\ &= s + \mathbb{E}((B_t - B_s)(B_s - B_0)) = s + \mathbb{E}(B_t - B_s)\mathbb{E}(B_s) = s\end{aligned}$$

where we used that for $s < t$, $B_t - B_s$ is independent of $B_s - B_0 = B_s$. □

Corollary 1.1 (Characterization)**.** *$(B_t)_{t \geq 0}$ is a Brownian motion if and only if $t \mapsto B_t$ is continuous on an event of probability 1 **and** $(B_t)_{t \geq 0}$ is a Gaussian process with covariance function $\mathbb{E}(B_t B_s) = \min(s,t)$.*

1.7.1 Some invariance properties of Brownian motion

The following facts are immediate consequences of the previous description of a Brownian motion as a centered Gaussian process with covariance function $\mathbb{E}(B_s B_t) = s \wedge t$.

Proposition 1.2. *Let $(B_t)_{t \geq 0}$ be a one-dimensional Brownian motion, then:*

- *(Scaling invariance). For every $a > 0$, the process $(a^{-1} B_{a^2 t})_{t \geq 0}$ is a Brownian motion.*
- *(Inversion invariance). The process $(tB_{1/t})_{t > 0}$ is distributed like $(B_t)_{t > 0}$.*
- *(Invariance under time reversal) The process $(B_{1-t} - B_1)_{t \in [0,1]}$ is distributed like $(B_t)_{t \in [0,1]}$.*

Proof. We already know that all three processes involved are continuous (resp. on $[0, \infty), (0, \infty), [0,1]$). Since B is a centered Gaussian process, it follows that the same is true for all three processes. Hence it only remains to check the covariance conditions: For the first one

$$\mathbb{E}(B_{a^2 t} B_{a^2(t+h)}/a^2) = \frac{1}{a^2} \min(a^2 t, a^2(t+h)) = \frac{1}{a^2} a^2 t = t.$$

For the second one:

$$\mathbb{E}(tB_{1/t}(t+h)B_{1/(t+h)}) = t(t+h)/(t+h) = t$$

For the last one: Let $s, t \in [0,1]$ such that $s < t$, then

$$\begin{aligned}\mathbb{E}((B_{1-t} - B_1)(B_{1-s} - B_1)) &= \mathbb{E}(B_{1-t}B_{1-s}) - \mathbb{E}(B_1 B_{1-t}) - \mathbb{E}(B_1 B_{1-s}) + \mathbb{E}(B_1^2) \\ &= (1-t) - (1-t) - (1-s) + 1 = s\end{aligned}$$

which concludes the proof. $\square$

1.7.2 The Brownian bridge

The Gaussian processes framework can be useful to describe processes that are derived from a Brownian motion via some linear operations.

Definition 1.3. *Let $(B_t)_{t \geq 0}$ be a one-dimensional Brownian motion, let us define the process $\beta = (\beta_t)_{t \in [0,1]}$ via*

$$\beta_t = B_t - tB_1.$$

The process β is called a (standard) Brownian bridge.

Proposition 1.3. *The Browindian bridge ($\beta_t = B_t - tB_1, t \in [0,1]$) is a Gaussian process, with covariance function given by (when $0 \leq t \leq s \leq 1$) $\sum_\beta = t(1-s)$ and it is independent of B_1.*

Proof. β is clearly a centered Gaussian process, because B is a centered Gaussian process, thus $((\beta_t)_{t \leq 1}, B_1)$ is a Gaussian process and we have for all $t \leq 1$

$$\mathbb{E}(\beta_t B_1) = \mathbb{E}(B_t B_1 - tB_1^2) = \mathbb{E}(B_t B_1) - t\mathbb{E}(B_1^2) = t - t \cdot 1 = 0,$$

which readily implies that β is independent of B_1. Moreover we have for all $0 \leq t \leq s \leq 1$,

$$\begin{aligned}\mathbb{E}(\beta_t \beta_s) &= \mathbb{E}((B_t - tB_1)(B_s - sB_1)) = \mathbb{E}(B_t B_s - tB_1 B_s - sB_1 B_t + tsB_1^2) \\ &= t \wedge s - ts - st + ts = t - ts = t(1-s).\end{aligned}$$

which concludes the proof. $\square$

1.8 L^2 considerations of constructing a Brownian Motion

1.8.1 Preliminaries

In this subsection we summarize the tools we require.

We work in the space $L^2([0,1])$, more precisely we consider the space of L^2 functions on $[0,1]$. We know that L^2 is the only space amongst the L^p spaces that is a Hilbert space, i.e. that has an inner product (geometry) given by

$$\langle f, g\rangle = \int_0^1 f(t)g(t)dt.$$

Moreover we will use Parseval's identity:

Theorem 1.2 (Parseval's Identity)**.** *Let V be a pre-Hilbert space and $S \subset V$ be an orthonormalsystem, i.e. all elements of S are orthogonal with respect to one another and have Norm (induced by the inner product on V) of* 1*. Then S is a complete orthonormalbasis of V if for all $v \in V$ parseval's identity holds:*

$$\|v\|^2 = \langle v, v\rangle = \sum_{s\in S} |\langle v, s\rangle|^2.$$

More generally, Parseval's identity holds for all $x, y \in V$

$$\langle x, y\rangle = \sum_{s\in S} \langle x, s\rangle\langle y, s\rangle.$$

Last let us recall that if we have a sequence X_n of centered Gaussian random variables such that X_n converges in probability to some finite random variable X, then X is also a centered Gaussian random variable. Indeed, convergence in probability implies the convergence in law, so that for all $\lambda \in \mathbb{R}$ we have

$$\mathbb{E}(\exp(i\lambda X)) = \lim_{n\to\infty} \exp\left(-\frac{\lambda^2}{2\sigma^2_{X_n}}\right).$$

For the right hand side to convergence to some non-zero limit (which has to be the case when λ is small), it is necessary that $\sigma^2_{X_n}$ converges to some finite limit σ^2, and it follows that the law of X is indeed $\mathcal{N}(0, \sigma^2)$.

1.8.2 Construction

Let us take an ONB of $L^2([0,1])$: $(\varphi_n)_{n\in\mathbb{N}}$. Let us define for all $t > 0$ and for all $n \in \mathbb{N}$

$$\Psi_n(t) := \int_0^t \varphi_n(s)ds = \langle 1_{[0,t]}, \varphi_n \rangle.$$

Let us now consider a sequence of i.i.d. centered standard Gaussians $(N_n)_{n\in\mathbb{N}}$ (i.e. $N_n \sim \mathcal{N}(0,1)$ for all $n \in \mathbb{N}$), and let us define the functions

$$S_m(t) = \sum_{n=0}^{m} N_n \Psi_n(t).$$

We notice that the above defines a sum of independent Gaussian random variables with mean 0, in particular it defines a Gaussian process. Moreover we have

$$\begin{aligned}\mathbb{E}(S_m(t)^2) = \sum_{n=0}^{m} \Psi_n(t)^2 \mathbb{E}(N_n^2) = \sum_{n=0}^{m} \Psi_n(t)^2 = \sum_{n=0}^{m} \langle 1_{[0,t]}, \varphi_n \rangle^2 \\ \leq \sum_{n=0}^{\infty} \langle 1_{[0,t]}, \varphi_n \rangle^2 \overset{P.I.}{=} \langle 1_{[0,t]}, 1_{[0,t]} \rangle = t.\end{aligned}$$

Where in the above P.I. stands for Parseval's Identity. Since $S_m(t)$ defines a sum of independent random variables where the sum of variances converges (as $m \to \infty$) by the above, we know (easy to show) that for each t, the series $S_m(t)$ converges almost surely and in L^2 to a random variable $S(t)$ *(alternatively, S_m defines a martingale bounded in L^2 and thus must converge a.s. and in L^2).* Since each $S_m(\cdot)$ is a centered Gaussian process, the same is true (by our preliminaries) for the limiting process $(S(t))_{t\in[0,1]}$.

Furthermore using the convergence in L^2 of $S_m(t)$ to $S(t)$, we get that

$$\begin{aligned}\mathbb{E}(S(t)S(s)) &= \lim_{m\to\infty} \mathbb{E}(S_m(t)S_m(s)) = \lim_{m\to\infty} \sum_{n=0}^{m} \Psi_n(t)\Psi_n(s) \\ &= \sum_{n=0}^{\infty} \Psi_n(t)\Psi_n(s) = \sum_{n=0}^{\infty} \langle 1_{[0,t]}, \varphi_n \rangle \langle 1_{[0,s]}, \varphi_n \rangle \overset{P.I.}{=} \langle 1_{[0,t]}, 1_{[0,s]} \rangle \\ &= \int_0^1 1_{[0,t]}(u) 1_{[0,s]}(u) du = s \wedge t.\end{aligned}$$

So, we see that the law of the process $(S(t))_{t\geq 0}$ is that of a Brownian motion.

Remark 1.5. Here is a heuristic interpretation of what we just derived. The "derivative" of a Brownian motion on $[0,1]$ (recall that a Brownian motion is nowhere differentiable) is given by

$$\sum_{n\geq 0} N_n \varphi_n, \tag{$\mho$}$$

because we have seen that

$$S(t) := \sum_{n\geq 0} N_n \Psi_n(t) = \sum_{n\geq 0} N_n \int_0^t \varphi_n(s)ds$$

has the law of a Brownian Motion. In particular we can read from ($\mho$) that the "derivative" of a BM satisfies that for each orthonormal basis of L^2, the coordinates are just given by i.i.d. standard Gaussians. We stress that this does not make sense, because the sum given in ($\mho$) does not converge, but we guess that there is something there.

This "derivative" of Brownian motion, given at ($\mho$), is sometimes called **white noise**.

Let us now consider two concrete examples of orthonormal basis of $L^2([0,1])$.

First we want to consider the **Haar basis**. For each dyadic interval $I_{j,n} \subset [0,1]$, we define $\varphi_0 = 1$ and we index the others by the dyadic intervals $I_{j,n} = [j2^{-n}, (j+1)2^n]_{n\geq 0}$ for $j \leq 2^n - 1$ as $\varphi_{j,n} = 2^{n/2}$ on the left-half of $I_{j,n}$ and as $\varphi_{j,n} = -2^{n/2}$ on the right-half of $I_{j,n}$.

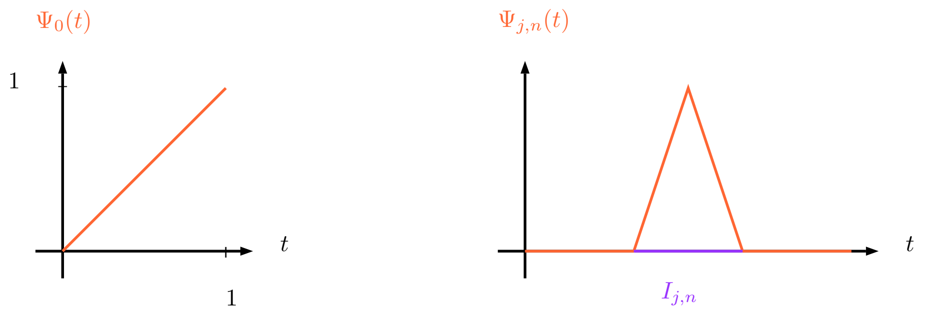


Figure 4: Depiction of Ψ_0 and $\Psi_{j,n}$ for the Haar basis.

We then realize that $S(t) = \sum_j N_{j,n}\Psi_{j,t}(t) = f_{n+1}(t) - f_n(t)$ as we have defined them in the previous construction. So this construction of S leads exactly to our previous construction of Brownian Motion in the first chapter.

1.8.3 The Fourier series of $(B_t)_{t\in[0,1]}$ and the Fourier Basis

In order to motivate the second example of ONB, namely the Fourier basis, we first start with a few considerations. Let us consider a Brownian motion $(B_t)_{t\in[0,1]}$ defined on the time-interval $[0,1]$. We define the Brownian bridge $\beta_t := B_t - tB_1$ (recall that is also a centered Gaussian process) and we know that β is independent of B_1). Since β is almost surely a continuous function with $\beta_0 = \beta_1$, we can almost surely decompose it into a Fourier series (in the sense of L^2 functions). More precisely, we have for all $t \in [0,1]$, the sequence of functions

$$S_m(t) := \sum_{n=1}^{m} b_n \sin(\pi n t), \text{ where } b_n = \int_0^1 \beta_t \sin(\pi n t) dt.$$

We then have that this sequence S_m converges in $L^2([0,1])$ to the function $t \mapsto \beta_t$ as $m \to \infty$. So, if we set $\tilde{b_0} = B_1$ and $\tilde{b_n} = \pi n b_n/\sqrt{2}$, then we obtain

$$B_t = \tilde{b_0} t + \lim_{m\to\infty} \sum_{n=1}^{m} \frac{\tilde{b_n}\sqrt{2}}{\pi n} \sin(\pi n t), \tag{$\clubsuit$}$$

where the limit is in the sense of L^2 limits of functions.

We now use the Fourier basis. The ONB of $L^2([0,1])$ is in this case given by $\varphi_0(t) = 1$ and for all $n \geq 1$ we set $\varphi_n(t) = \sqrt{2}\cos(\pi t n)$. Consequently,

$$\Psi_0(t) = t, \ \Psi_n(t) = \frac{\sqrt{2}}{\pi n} \sin(\pi n t) \text{ for all } n \geq 1.$$

We then know that $S(t) :=$ "$\sum N_n \Psi_n(t)$" has the law of a Brownian motion on $[0,1]$ and

$$S(t) = \sum_{n=1}^{\infty} \frac{N_n\sqrt{2}}{\pi n} \sin(\pi n t) + N_0 t, \tag{$\spadesuit$}$$

since the Fourier coefficients are unique, by comparison of ($\clubsuit$) with ($\spadesuit$) we can identify the Fourier coefficients to be independent centered standard Gaussian random variables.

2 The weak and strong Markov properties for Brownian Motion and first consequences

2.1 The weak Markov property

In what follows, B will denote a real-valued Brownian motion in dimension 1, but the statements will also hold in higher dimensions.

Definition 2.1. *For every $t \geq 0$ we denote by $\mathcal{F}_t$ the σ-algebra generated by the random variables $\{B_r : 0 \leq r \leq t\}$. The Brownian filtration $(\mathcal{F}_t)_{t\geq 0}$ is then just the collection of these σ-fields (in particular we have $\mathcal{F}_t \subset \mathcal{F}_{t+h}$ for all $t, h \geq 0$). Moreover, for each finite time $T \geq 0$, we define the process $B^{(T)} = (B_t^{(T)})_{t\geq 0}$ by*

$$B_t^{(T)} := B_{T+t} - B_T.$$

The definition of Brownian motion (stationary independent increments) ensures immediately that:

Lemma 2.1 (Weak Markov property)**.** *When T is a fixed, deterministic time, then $B^{(T)}$ is a Brownian motion that is independent of $\mathcal{F}_T$.*

2.1.1 Blumenthal's $0-1$ Law and consequences

We also define for each $t \geq 0$, the σ-Field

$$\mathcal{F}_{t+} := \bigcap_{h>0} \mathcal{F}_{t+h},$$

which seems to contain some additional infinitesimal look into the future.

Remark 2.1. We can also think about the use of the above defined σ-field if we consider for instance a local maximum on $[0, t]$, for example a world record, and we wonder if our world record lasts at least for an infinitesimal time into the future $[0, t + h]$.

Proposition 2.1 (Blumenthal's $0-1$ law)**.** *For the Brownian filtration, the σ-field $\mathcal{F}_{0+}$ is trivial in the sense that for all events $A \in \mathcal{F}_{0+}$ we have either $\mathbb{P}(A) = 0$ or $\mathbb{P}(A) = 1$.*

Proof. Let us take $A \in \mathcal{F}_{0+}$, we then have for all $h > 0$ also that $A \in \mathcal{F}_h$. Let us prove that A is necessarily independent of $(B_{t_1}, \ldots, B_{t_p})$ for all fixed times $0 < t_1 < \cdots < t_p$. Recall that for this, it is enough to prove that for any bounded continuous function $f : \mathbb{R}^p \to \mathbb{R}$ we have

$$\mathbb{E}(1_A f(B_{t_1}, \ldots, B_{t_p}) = \mathbb{P}(A)\mathbb{E}(f(B_{t_1}, \ldots, B_{t_p})).$$

Let $h > 0$. By dominated convergence and the continuity of B we easily establish that

$$\begin{aligned}\mathbb{E}(1_A f(B_{t_1}, \ldots, B_{t_p})) &= \lim_{h\to 0} \mathbb{E}(1_A f(B_{t_1+h} - B_h, \ldots, B_{t_p+h} - B_h)) \\ = \lim_{h\to 0} \mathbb{E}(\overbrace{1_A}^{\in \mathcal{F}_h} \underbrace{f(B^{(h)}_{t_1}, \ldots, B^{(h)}_{t_p})}_{\text{indep. of } \mathcal{F}_h}) &\overset{\text{M.P.}}{=} \lim_{h\to 0} \mathbb{P}(A)\mathbb{E}(f(B^{(h)}_{t_1}, \ldots, B^{(h)}_{t_p})) \\ &= \mathbb{P}(A)\mathbb{E}(f(B_{t_1}, \ldots, B_{t_p}))\end{aligned}$$

By a monotone class argument this establishes that A is independent of $\mathcal{F}_\infty$, but then since $A \in \mathcal{F}_h \subset \mathcal{F}_\infty$ we conclude that A is independent of itself, in particular we have

$$\mathbb{P}(A) = \mathbb{P}(A \cap A) = \mathbb{P}(A)^2 \implies \mathbb{P}(A) = 0 \text{ or } 1.$$

$\square$

Here is a nice and easy consequence of Blumenthal's $0 - 1$ law.

Proposition 2.2. *Almost surely,*

$$\limsup_{t\to 0} \frac{B_t}{\sqrt{t}} = \infty \text{ and } \liminf_{t\to 0} \frac{B_t}{\sqrt{t}} = -\infty.$$

Remark 2.2. Note that this implies in particular that for all $\epsilon > 0$, there exists infinitely many times $t \in (0, \epsilon)$ at which $B_t = 0$. Moreover, the Proposition implies that almost surely, Brownian motion is not Hölder continuous with exponent $1/2$ (nor with any exponent greater than $1/2$) and thus Brownian motion is, as expected, nowhere differentiable.

In order to see the second remark, let $0 \leq s < t$ be arbitrary, we then have almost surely

$$\limsup_{t\to 0} \frac{B_t}{\sqrt{t}} = \infty = \limsup_{t\to s} \frac{B_{t-s}}{\sqrt{t-s}} = \limsup_{t\to s} \frac{B_t - B_s}{\sqrt{t-s}}$$

in particular we cannot have $|B_t - B_s| \leq C\sqrt{t-s}$ for some $C \in \mathbb{R}_{>0}$.

Proof. We define for each $\epsilon > 0$, the random variable

$$V_\epsilon := \sup_{s\in(0,\epsilon]} \frac{B_s}{\sqrt{s}}.$$

Thanks to the continuity of Brownian motion we can write $V_\epsilon = \sup\{B_s/\sqrt{s} : s \in (0,\epsilon] \cap \mathbb{Q}\}$, and we notice that V_ϵ is $\mathcal{F}_\epsilon$ measurable. Moreover, obviously $t \mapsto V_t$ is increasing and thus we can define

$$V_{0+} := \lim_{\epsilon\to 0} V_\epsilon = \inf_{\epsilon>0} V_\epsilon.$$

So we see that the random variable V_{0+} is $\mathcal{F}_\epsilon$-measurable for all $\epsilon > 0$ and therefore $\mathcal{F}_{0+}$ measurable. Hence, by Blumenthal's $0-1$ law, we have that it is constant either 0 or 1.

Recall that the scaling property of a Brownian motion tells us that for all $a > 0$ we have $\frac{1}{a}B_{a^2t}$ is also a Brownian motion, in particular $\frac{1}{a}B_{a^2t} \sim B_t$. We note by applying this scaling property, that the law of V_ϵ does in fact not depend on ϵ. Indeed, let us fix $M \geq 1$ integer, then we have almost surely

$$\begin{aligned}\left\{\sup_{s\in(0,\epsilon]} \frac{B_s}{\sqrt{s}} \geq M\right\} \overset{t:=s/\epsilon}{=} \left\{\sup_{t\in(0,1]} \frac{B_{\epsilon t}}{\sqrt{\epsilon t}} \geq M\right\} &= \left\{\sup_{t\in(0,1]} \frac{\frac{1}{\sqrt{\epsilon}}B_{\epsilon t}}{\sqrt{t}} \geq M\right\} \\ &= \left\{\sup_{t\in(0,1]} \frac{B_t}{\sqrt{t}} \geq M\right\}\end{aligned}$$

Where in the last step the scaling property with $a = \sqrt{\epsilon}$ was used. So we have shown that $\mathbb{P}(V_\epsilon \geq M) = \mathbb{P}(V_1 \geq M)$, thus we finally get,

$$\begin{aligned}\mathbb{P}(V_{0+} \geq M) = \lim_{\epsilon\to 0} \mathbb{P}(V_\epsilon \geq M) = \mathbb{P}(V_1 \geq M) &= \mathbb{P}\left(\sup_{t\in(0,1]} \frac{B_t}{\sqrt{t}} \geq M\right) \\ &\geq \mathbb{P}(\sup_{t\in(0,1]} B_t \geq M) \geq \mathbb{P}(B_1 \geq M) > 0.\end{aligned}$$

So by Blumenthal's $0-1$ law we get that for all $M \in \mathbb{N}_{\geq 1}$, $\mathbb{P}(V_{0+} \geq M) = 1$. Thus we can exchange almost surely with for all M integer to get

$$\mathbb{P}(\forall M \geq 1, V_{0+} \geq M) = 1,$$

i.e. $V_{0+} = \infty$ almost surely. We can apply this also to the Brownian motion $-B$, to get the second statement. $\square$

2.2 Stopping times and the strong Markov property

The goal of this section is to extend the weak Markov property to the case where T is replaced by special random times.

Definition 2.2. *A random variable $T \in \mathbb{R}_+ \cup \{\infty\}$ is a stopping time for a filtration $(\mathcal{F}_t)_{t\geq 0}$ if for every $t \geq 0$, the event $\{T \leq t\}$ is in $\mathcal{F}_t$. We also define the σ-field of the past before T as*

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

It intuitively corresponds to all the information (about everything) that has happened up to random time T.

Remark 2.3. A good intuitive way to think about the definition of stopping times is that when T has occurred, then one actually knows it. Here is a neat example to have in mind of a stopping time: You're driving on the highway, when a red car passes you, you take the next exit. Here is an example of a non-stopping time: One year before the next big earthquake.

Proposition 2.3 (Strong Markov property)**.** *Let T be a stopping time (with respect to the Brownian filtration) such that $T < \infty$ almost surely. Then the process $B^{(T)} := (B_t^{(T)} = B_{T+t} - B_T)_{t\geq 0}$ is a Brownian motion independent of $\mathcal{F}_T$.*

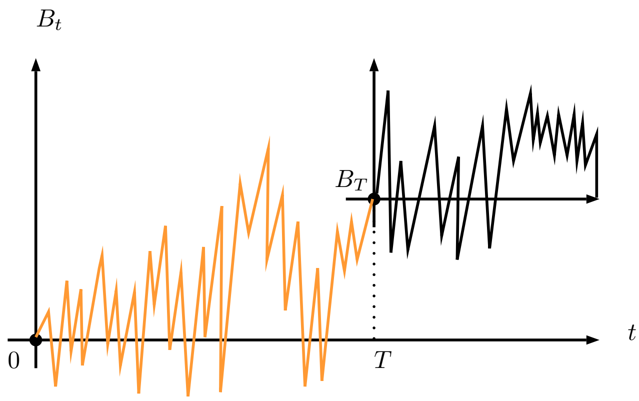


Figure 5: Qualitative picture of the strong Markov property. If we "reboot" the Brownian Motion at random stopping time T (notice the shift of axis, $B^{(T)}$ starts at 0), then what we observe is again a Brownian motion that is independent of its past $\mathcal{F}_T$ (orange).

Proof. Let $A \in \mathcal{F}_T$, we want to prove that for all $t_1 < \cdots < t_p$ the event A is independent of $(B^{(T)}_{t_1}, \dots, B^{(T)}_{t_p})$ and that $(B^{(T)}_{t_1}, \dots, B^{(T)}_{t_p})$ has the same law as $(B_{t_1}, \dots, B_{t_p})$. In order to prove these two statements, it is enough to check that for all $f : \mathbb{R}^p \to \mathbb{R}$ continuous and bounded, we have

$$\mathbb{E}(1_A f(B^{(T)}_{t_1}, \dots, B^{(T)}_{t_p})) = \mathbb{P}(A)\mathbb{E}(f(B^{(T)}_{t_1}, \dots, B^{(T)}_{t_p})),$$

and

$$\mathbb{E}(f(B^{(T)}_{t_1}, \dots, B^{(T)}_{t_p})) = \mathbb{E}(f(B_{t_1}, \dots, B_{t_p}))$$

We work with the discrete approximation of the stopping T. I.e. we define T_n to be the smallest multiple of 2^{-n} such that $T \leq T_n$ (T_n = smallest $j2^{-n}$ with $j2^{-n} \geq T$). It is then an easy exercise to see that T_n are also a stopping times w.r.t. the Brownian filtration and that $\{T_n = j2^{-n}\} \in \mathcal{F}_{j2^{-n}}$.

We then define for each given n, $A_j = A \cap \{T_n = j2^{-n}\}$. Then, each A_j is in $\mathcal{F}_{j2^{-n}}$ and $A = \cup_j A_j$ as a disjoint union of the A_j's. So we get:

$$\begin{aligned} \mathbb{E}(1_A f(B^{(T_n)}_{t_1}, \dots, B^{(T_n)}_{t_p})) &= \sum_{j=0}^{\infty} \mathbb{E}(1_{A_j} f(B^{(T_n)}_{t_1}, \dots, B^{(T_n)}_{t_p})) \\ = \sum_{j=0}^{\infty} \mathbb{E}(1_{A_j} f(B^{(j2^{-n})}_{t_1}, \dots, B^{(j2^{-n})}_{t_p})) &\overset{\text{W.M.P.}}{=} \sum_{j=0}^{\infty} \mathbb{P}(A_j)\mathbb{E}(f(B^{(j2^{-n})}_{t_1}, \dots, B^{(j2^{-n})}_{t_p})) \\ &= \sum_{j=0}^{\infty} \mathbb{P}(A_j)\mathbb{E}(f(B_{t_1}, \dots, B_{t_p})) = \mathbb{P}(A)\mathbb{E}(f(B_{t_1}, \dots, B_{t_p})) \end{aligned}$$

But by dominated convergence and the continuity of Brownian motion we also have $\mathbb{E}(1_A f(B^{(T_n)}_{t_1}, \dots, B^{(T_n)}_{t_p})) \to \mathbb{E}(1_A f(B^{(T)}_{t_1}, \dots, B^{(T)}_{t_p}))$ as $n \to \infty$ almost surely. So finally we get

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

□

Remark 2.4. An equivalent way of phrasing the strong Markov property will be that (under the same conditions, T being a finite stopping time) for all $f : \mathbb{R}^{\times \mathbb{R}_+} \to \mathbb{R}$ bounded and continuous (measurable would be enough) we have for all $x \in \mathbb{R}$,

$$\mathbb{E}_x[f((B_{T+t})_{t \geq 0}) \mid \mathcal{F}_T] = \mathbb{E}_{B_T}[f((B_t)_{t \geq 0})]$$

i.e. conditionally on $\mathcal{F}_T$ the process $(B_{T+t})_{t \geq 0}$ is again a Brownian motion, it is independent of $\mathcal{F}_T$ and has the same law as B started from B_T.

2.2.1 Reflection principle and consequences

Suppose that T is a stopping time for the Brownian filtration and assume that T is almost surely finite. We now construct a new process $\tilde{B}$ as follows: for all $t \geq 0$

$$\tilde{B}_t := \begin{cases} B_t, & t \leq T \\ B_T - (B_t - B_T), & t \geq T \end{cases}$$

In other words, the increments of $\tilde{B}$ after the stopping time T are the opposite of those of B.

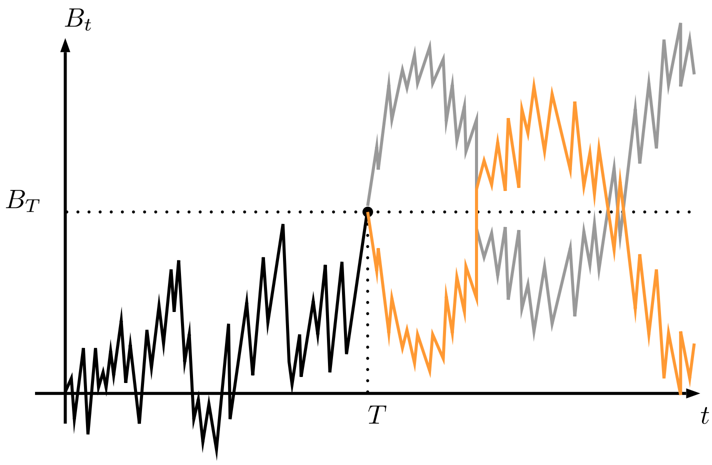


Figure 6: Depiction of the process $\tilde{B}$. We see the Brownian motion B_t (black + grey) and the reflection after stopping time T (depicted in orange) and the process $\tilde{B}$ (black + orange).

Proposition 2.4 (Reflection principle)**.** *The process $\tilde{B}$ is also a Brownian motion.*

Proof. The strong Markov property says that the process $B^{(T)}$ is a Brownian motion and independent of $\mathcal{F}_T$. Hence, the process $-B^{(T)}$ is also a Brownian motion independent of $\mathcal{F}_T$. But we can reconstruct B from the pair $(B_t, t \leq T)$ and $B^{(T)} = (B_{T+t} - B_T)_{t\geq 0}$ in exactly the same way in which $\tilde{B}$ is reconstructed from the pair $(B_t, t \leq T)$ and $-B^{(T)}$, which implies that B and $\tilde{B}$ have the same law. $\square$

Corollary 2.1. *Let B be a Brownian motion in dimension 1. For every $t > 0$, define the running maximum $S_t := \max_{s \le t} B_s$. For every $a \ge 0$ and $h \ge 0$ we have*

$$\mathbb{P}(S_t \ge a, B_t \le a - h) = \mathbb{P}(B_t \ge a + h).$$

Moreover, for each given t, the variable S_t has the same distribution as $|B_t|$.

Proof. Let $T_a = \inf\{t \ge 0 : B_t = a\}$. We know that (best seen in a picture) $\{S_t \ge a\} = \{T_a \le t\}$. We obtain:

$$\begin{aligned}
\mathbb{P}(S_t \ge a, B_t \le a - h) &= \mathbb{P}(T_a \le t, B_t - B_{T_a} \le -h) = \mathbb{P}(T_a \le t, B^{(T_a)}_{t-T_a} \le -h) \\
&\overset{1)}{=} \mathbb{P}(T_a \le t, -\tilde{B}^{(T_a)}_{t-T_a} \le -h) = \mathbb{P}(T_a \le t, \tilde{B}^{(T_a)}_{t-T_a} \ge h) = \mathbb{P}(T_a \le t, \tilde{B}_t - \tilde{B}_{T_a} \ge h) \\
&\overset{2)}{=} \mathbb{P}(T_a \le t, \tilde{B}_t - B_{T_a} \ge h) = \mathbb{P}(T_a \le t, \tilde{B}_t \ge a + h) \overset{\text{R.P.}}{=} \mathbb{P}(T_a \le t, B_t \ge a + h) \\
&\overset{3)}{=} \mathbb{P}(B_t \ge a + h).
\end{aligned}$$

Where we used:

1. It is easily seen that on $t \ge T_a$ we have $\tilde{B}^{(T_a)}_{t-T_a} = \tilde{B}_t - \tilde{B}_{T_a} \overset{\text{def}}{=} B_{T_a} - (B_t - B_{T_a}) - a = B_{T_a} - B_t$. Geometrically this is obvious when we look at the previous figure.
2. By definition, as already used in 1. above, we have $\tilde{B}_{T_a} = B_{T_a} = a$.
3. We have $\{B_t \ge a + h\} \subset \{T_a \le t\}$, indeed if for some fixed $t \ge 0$ we have $B_t \ge a + h$ then necessarily the first time when B_t meets the height a must occur before time t, i.e. $T_a \le t$.

And R.P. stands for Reflection Principle.

To establish the second claim we choose $h = 0$ to see that

$$\begin{aligned}
\mathbb{P}(S_t \ge a) &= \mathbb{P}(S_t \ge a, B_t \le a) + \mathbb{P}(S_t \ge a, B_t \ge a) \\
&\overset{h=0}{=} \mathbb{P}(B_t \ge a) + \mathbb{P}(S_t \ge a, B_t \ge a) = 2\mathbb{P}(B_t \ge a) = \mathbb{P}(|B_t| \ge a).
\end{aligned}$$

Where obviously $\{B_t \ge a\} \subset \{S_t \ge a\}$. ∎

2.2.2 The zero-set of a Brownian motion

We now state and prove another property of a Brownian motion which just further illustrates that a Brownian motion is quite a strange continuous curve. Let us define the zero-set of a Brownian motion as

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

Proposition 2.5. *Almost surely, the set $\mathcal{Z}$ is a perfect set (i.e. it is an non-empty closed set with no isolated points).*

Remark 2.5.

1. Recall that a point $t \in Z$ is called isolated by definition if for all $\epsilon > 0$ we have $Z \cap ((t-\epsilon, t+\epsilon) \setminus \{t\}) \neq \emptyset$.
2. It is an elementary exercise to show that a non-empty closed subset of $\mathbb{R}$ with no isolated points has the same cardinality as $\mathbb{R}$.

Proof. The set $\mathcal{Z}$ is closed almost surely thanks to the continuity of B. *(A subset D of a metric space is closed iff it contains all limits of seq. in D).*

For $q \in \mathbb{Q}_+ (= \mathbb{Q} \cap \mathbb{R}_+)$ we define the stopping time $\tau_q = \inf\{t \geq q : B_t = 0\}$. Clearly if we take $T \in \mathcal{Z}$ (i.e. $B_T = 0$) and if we assume that T is isolated from the left (i.e. $\exists \epsilon > 0$ such that $(T-\epsilon, T) \cap \mathcal{Z} = \emptyset$) then $\exists q \in \mathbb{Q}_+$ such that $\tau_q = T$.

On the other hand, we know that for all such q, $B^{(\tau_q)}$ is distributed like a Brownian motion, and we have seen ($\limsup_{t \to 0} B_t/\sqrt{t} = \infty$) that almost surely $B^{(\tau_q)}$ has infinitely many zeros in any interval $(0, \epsilon)$, in particular τ_q is not isolated from the right in $\mathcal{Z}$.

Consequently, we have almost surely for all $q \in \mathbb{Q}_+$, that τ_q is not isolated in $\mathcal{Z}$ from the right. But then almost surely for all $t \in \mathcal{Z}$ which are isolated from the left we have

$$t \in \bigcup_{q \in \mathbb{Q}} \{\tau_q\},$$

so t is not isolated from the right and therefore $\mathcal{Z}$ has no isolated points. $\square$

Remark 2.6. The Lebesgue measure $\lambda(\mathcal{Z})$ of $\mathcal{Z}$ is almost surely equal to zero, indeed, by Fubini's theorem

$$\mathbb{E}(\lambda(\mathcal{Z})) = \mathbb{E}\left(\int_0^\infty 1_{B_t=0} dt\right) = \int_0^\infty \mathbb{P}(B_t = 0) dt = 0$$

which entails that $\lambda(\mathcal{Z}) = 0$ almost surely.

2.3 Analogous results for multidimensional Brownian motion

Let us briefly list which results in the previous section can be immediately generalized to the case where one considers a Brownian motion B in d-dimensional space with $d \geq 2$ instead of $d = 1$.

- The weak Markov property.
- The strong Markov property.
- Blumenthal's $0-1$ law.

The statements and proofs are exactly the same as in the one-dimensional case.

2.3.1 One extension of some 1D results/ideas

We start with an application of Blumenthal's $0-1$ law in higher dimensions. Let $(B_t)_{t\geq 0}$ be a Brownian motion in $\mathbb{R}^d, d \geq 2$, started at the origin 0. Let $\mathcal{C}$ be an open subset of $\mathbb{R}^d \setminus \{0\}$ such for some fixed $r > 0$, it contains a union of balls $\mathcal{B}(x_n, r|x_n|)$, where x_n is some sequence in $\mathbb{R}^d \setminus \{0\}$ with $x_n \to 0$ (we can always assume that $|x_n|$ decreasing with n). That is, $\exists r > 0$ and there exists a sequence $x_n \to 0$ in $\mathbb{R}^d \setminus \{0\}$ such that

$$\bigcup_{n\in\mathbb{N}} \mathcal{B}(x_n, r|x_n|) \subset \mathcal{C}.$$

A good example to have in mind of such a set if when $\mathcal{C}$ is a cone with apex at 0.

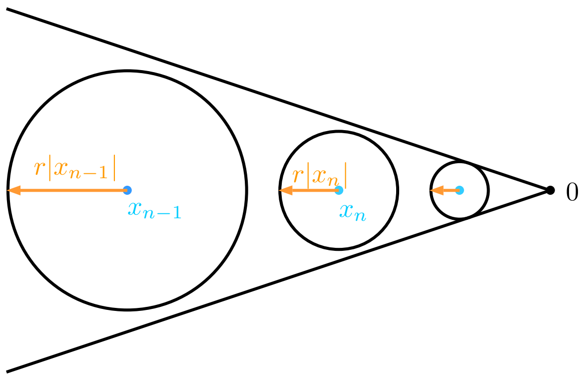


Figure 7: $\mathcal{C}$ is a cone with apex at 0.

Proposition 2.6. *Almost surely, for all $\epsilon > 0$ there exists $t \in (0, \epsilon)$ such that $B_t \in \mathcal{C}$.*

Proof. Let us define for $n_0 \in \mathbb{N}$

$$V_{n_0} := \bigcup_{n \geq n_0} \{B_{|x_n|^2} \in \mathcal{B}(x_n, r|x_n|)\}.$$

Then V_{n_0} is measurable with respect to

$$\mathcal{F}_{\max_{n \geq n_0} |x_n|^2}.$$

We have

$$V_\infty := \bigcap_{n_0 \geq 0} V_{n_0} = \{\exists \text{ infinitely many } n's : B_{|x_n|^2} \in \mathcal{B}(x_n, r|x_n|)\}.$$

Since also $V_{n_0+1} \subset V_{n_0}$ we have for all $\tilde{n}$ that

$$V_\infty = \bigcap_{n_0 > \tilde{n}} V_{n_0},$$

in particular V_∞ is in F_{0+} and therefore we have thanks to Blumenthal's $0-1$ law that $\mathbb{P}(V_\infty) = 0$ or 1. We want to show that this probability is 1.

$$\begin{aligned}
\mathbb{P}(V_\infty) &= \lim_{n_0 \to \infty} \mathbb{P}(V_{n_0}) \geq \lim_{n_0 \to \infty} \mathbb{P}(B_{|x_{n_0}|^2} \in \mathcal{B}(x_{n_0}, r|x_{n_0}|) \\
&= \lim_{n_0 \to \infty} \mathbb{P}\left[\frac{1}{|x_{n_0}|} B_{|x_{n_0}|^2} \in \mathcal{B}\left(\frac{x_{n_0}}{|x_{n_0}|}, r\right)\right] \overset{1)}{=} \lim_{n_0 \to \infty} \mathbb{P}(B_1 \in \mathcal{B}(x_{n_0}/|x_{n_0}|, r)) \\
&\overset{2)}{=} \lim_{n_0 \to \infty} \mathbb{P}(B_1 \in \mathcal{B}(1., r)) = \mathbb{P}(B_1 \in \mathcal{B}(1., r)) > 0.
\end{aligned}$$

Where we used:

1. the scaling invariance at time $t = 1$ $\frac{1}{a} B_{a^2 t} \sim B_t$.
2. The isotropy property of a Brownian motion, it just states that for all linear isometries ϕ we have $\Phi(B)$ is still a BM in $\mathbb{R}^d$ and $1. = (1, 0, 0 \ldots, 0) \in \mathbb{R}^d$.

□

3 The Dirichlet Problem via Brownian Motion

Before we discuss the connection between Brownian motion and the Dirichlet problem, we will briefly talk about coupling techniques which will come in useful.

3.1 Coupling

3.1.1 Coupling in 1D

Suppose that one considers x and $\hat{x}$ in $\mathbb{R}$, such that $x \neq \hat{x}$. It is our goal to define on the same probability space two Brownian motions B and $\hat{B}$ started from x and $\hat{x}$ respectively in such a way that *"after they meet"*, they *"stick together"*. Obviously these two Brownian motions would not be independent of one another. We discuss two ideas:

3.1.2 First idea: B and $\hat{B}$ move independently, until they meet

Let B and $\tilde{B}$ be two independent Brownian motions, started from x respectively $\hat{x}$ and define $T = \inf\{t > 0 : B_t = \tilde{B}_t\}$. We then define $\hat{B}$ by:

$$\hat{B}_t := \begin{cases} \tilde{B}_t, & t \leq T \\ B_t, & t > T \end{cases}$$

one can then check that the law of $(\hat{B}_t)_{t \geq 0}$ is still that of a Brownian motion.

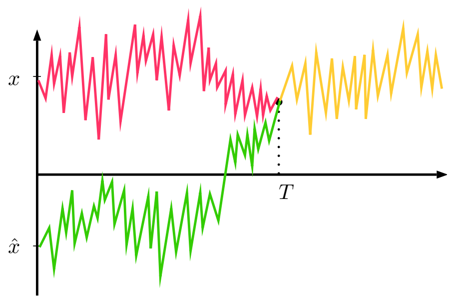


Figure 8: We see the sample of a Brownian motion B (red + orange) and the sample of $\tilde{B}$ independent BM (green) and the new BM $\hat{B}$ (green + orange).

3.1.3 Second idea: Mirror coupling

Let assume that $x = -\hat{x} > 0$ (for convenience purposes, otherwise we shift everything by $(x + \hat{x})/2$). Let us sample the Brownian motion $(B_t)_{t \geq 0}$ started from x and then define $T = \inf\{t > 0 : B_t = 0\}$. We then define $\hat{B}$ by

$$\hat{B}_t := \begin{cases} -B_t, & t \leq T \\ B_t, & t > T \end{cases}$$

it is then again easy to check that $\hat{B}$ is a Brownian motion. $\hat{B}$ is called the mirror coupling of B started from $-x$.

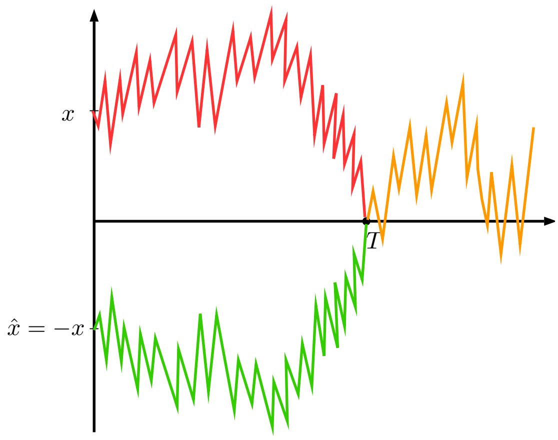


Figure 9: Mirror coupling: We see the sample of the Brownian motion B (red + orange) and it's mirror $-B_t$ (green) and the BM $\hat{B}$ (green + orange).

3.1.4 Mirror coupling in $\mathbb{R}^d$

Let us consider $x \neq 0$ in $\mathbb{R}^d$ and $\hat{x} = -x$. Let $\mathcal{H}$ be the hyperlane of points that are equidistant to x and $\hat{x}$ (i.e. $\mathcal{H}$ is the median hyperplane between x and $-x$). Let now $(B_t)_{t \geq 0}$ be a Brownian motion started from x.

Let $T = \inf\{t > 0 : B_t \in \mathcal{H}\}$ and let us denote by σ the symmetry with respect to the hyperplane $\mathcal{H}$. We then define $\hat{B}$ as

$$\hat{B}_t := \begin{cases} \sigma(B_t), & t \leq T \\ B_t, & t > T \end{cases}$$

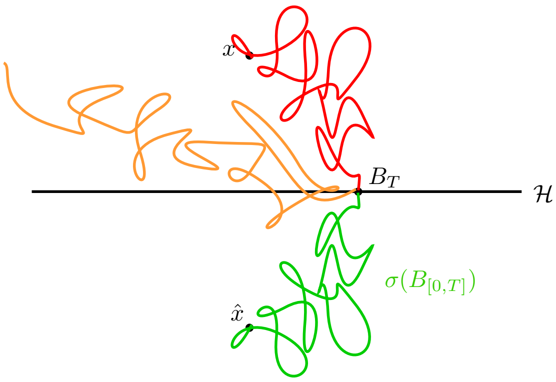


Figure 10: Mirror coupling in $\mathbb{R}^d$: A BM in $\mathbb{R}^d$ started from x (red + orange) and its mirror (w.r.t. $\mathcal{H}$) started $\hat{x}$ (green) and it's mirror coupling $\hat{B}$ which is a BM started from $\hat{x}$ (green + orange).

3.2 Towards the Dirichlet Problem

We use our multidimensional ideas to develop a more substantial consequence.

Definition 3.1. *We say that a bounded open subset D of $\mathbb{R}^d$ satisfies property $(\mathcal{P})$ if, for every $x \in \partial D$, there exists $r > 0$ and a sequence $x_n \to x$ such that for all n we have*

$$\mathcal{B}(x_n, r\|x_n - x\|) \cap D = \emptyset.$$

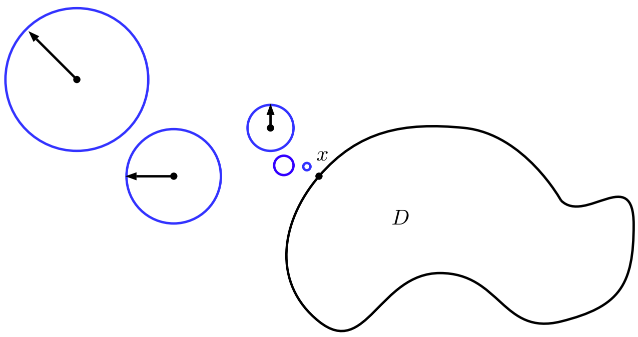


Figure 11: A depiction of property $(\mathcal{P})$.

Remark 3.1. Obviously $(\mathcal{P})$ is satisfied as soon as for each $x \in \partial D$, there exists an open cone C with apex at x such that $C \cap \overline{D} = \emptyset$. This is often referred to as the *outside cone condition*.

Suppose that D is a bounded open subset of $\mathbb{R}^d$. Suppose that we are given a continuous, real-valued function f that is defined on ∂D.

We are now going to define a (very nice) function U on $\overline{D}$ using Brownian motion. For each given $x \in \overline{D}$, we consider a Brownian motion $(B_t)_{t\geq 0}$ started from x. Let T be the first time where B hits the boundary of D, i.e.

$$T = \inf\{t > 0 : B_t \in \partial D\},$$

(note that T is finite because D is bounded). We then define for all $x \in \overline{D}$

$$U(x) := \mathbb{E}_x(f(B_T)).$$

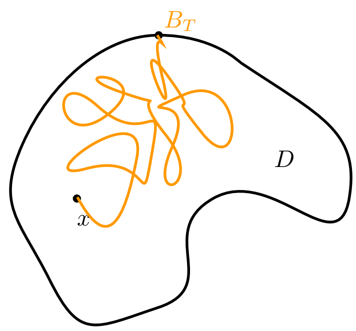


Figure 12: We start a BM at $x \in \overline{D}$ and let it evolve until it hits the boundary of D (equivalently: until it escapes the set D).

Proposition 3.1. *When D is a bounded open subset of $\mathbb{R}^d$, and f is a continuous, real-valued function defined on ∂D, then the function U satisfies the following properties:*

- *U is continuous in D.*
- *U is equal to f on ∂D.*
- *U satisfies the "mean-value property" in D: For all $x \in D$ and for all $r < d(x, \partial D)$, the mean value of U on the sphere of radius r around x is equal to $U(x)$, or in a formula:*

$$U(x) = \int_{S(x,r)} U(z) d\sigma_{x,r}(z),$$

where $\sigma_{x,r}$ is the (unique) uniform probability measure on the sphere $S(x,r)$ of radius r around x.

Remark 3.2. Notice that we do not state here that the function U is continuous at the boundary points of D. However, we will later on give a condition such that this will be the case.

Proof. We first give a rough sketch before we provide more detail. The second statement is obvious, if we have $x \in \partial D$, then obviously $T = 0$ and thus $B_T = B_0 = x$, in particular we have $U(x) = \mathbb{E}_x(f(x)) = f(x)$. The first statement will be derived using mirror coupling. The final statement will follow from the strong Markov property at the hitting time of the sphere $S(x, r)$ by B.

Continuity: Let us fix $x \in D$ and let $r_0 = d(x, \partial D)$. Choose y with $d(x, y) < r_0/8$. Set $x_0 = (x + y)/2$, then x_0 is equidistant from x and y and we can consider the hyperplane $\mathcal{H}$ through x_0. We now consider the two coupled Brownian motions:

$$\begin{cases} B, & \text{Brownian motion started from } x \\ B', & \text{"mirror coupled BM" started from } y \end{cases}$$

We then have for fixed r_1, the probability that B and B' do not couple before B hits $\mathcal{B}(x_0, r_1)$ goes to 0 as $d(x, y) \to 0$.

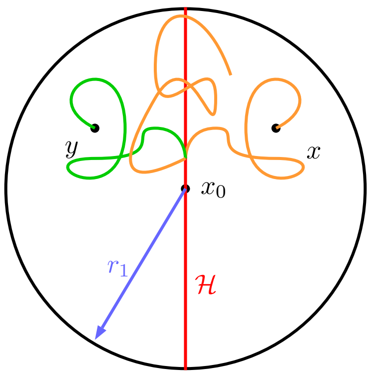


Figure 13: For a fixed r_1, the probability that the BM (orange) does not couple with its coupled mirror B' (green) before B exits $\mathcal{B}(x_0, r_1)$ goes to 0 when the distance between x, y is small. In Particular $B_T = B'_T$ with high probability when y is close to x.

$$\begin{aligned} |\mathbb{E}(f(B_T) - f(B'_T))| &\leq \mathbb{E}(|f(B_T) - f(B'_T)|) \\ \leq 2\|f\|_\infty \mathbb{P}(B \text{ and } B' \text{ do not couple before reaching } \mathcal{B}(x_0, r_1)) &\xrightarrow{y \to x} 0 \end{aligned}$$

We conclude that $|U(y) - U(x)| \to 0$ as $y \to x$.

Mean-value property: We take $x \in D$ and let $r < d(x, \partial D)$. We consider a Brownian motion B started from x, and we let τ be the first time at which it hits the sphere $S(x,r) = \partial \mathcal{B}(x,r)$ around x of radius r, i.e.

$$\tau = \inf\{t > 0 : d(B_t, x) = r\}.$$

We have

$$U(x) = \mathbb{E}_x(\mathbb{E}(f(B_T) \mid \mathcal{F}_\tau))$$

and we can apply the strong Markov property at time τ, it states that conditionally on $\mathcal{F}_\tau$, the process $(B_{\tau+t)})_{t\geq 0}$ is again a Brownian motion started from τ, hence we have

$$\mathbb{E}(f(B_T) \mid \mathcal{F}_\tau) = \mathbb{E}_{B_\tau}(f(B_T)) = U(B_\tau),$$

but we know that the law of B_τ is uniformly distributed on the sphere $\partial\mathcal{B}(x,r)$ which then proves the claim. Recall that a Brownian motion (in distribution) is invariant under rotations (in fact, by the law of Isotropy it is invariant under linear isometries). The only probability measure on the sphere $S(x,r)$ that is invariant under rotations is the uniform measure (see for instance Exercise 5.1).

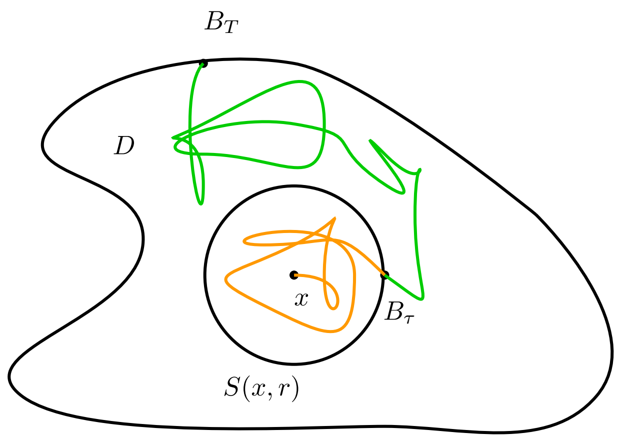


Figure 14: Conditionally on $\mathcal{F}_\tau$, we observe another Brownian motion $(B_{\tau+t})_{t\geq 0}$ started from τ (and independent of $\mathcal{F}_\tau$)

$\square$

Let us now give a condition on D that ensures that U will also be continuous at boundary points (i.e. on ∂D).

Definition 3.2. *We say that $x \in \partial D$ is a regular boundary point of D, if when one considers a Brownian motion B started from x, then almost surely*

$$\inf\{t > 0 : B_t \notin D\} = 0,$$

i.e. a BM started from $x \in \partial D$ exits immediately $\overline{D}$. We say that D has a regular boundary for Brownian motion (or in short, a regular boundary) if every boundary point is regular.

Remark 3.3. We know that the exterior cone condition implies property $(\mathcal{P})$, but then it also has a regular boundary (because we know that almost surely, for all $\epsilon > 0$ there exists $t \in (0, \epsilon)$ such that $B_t \in \mathcal{C}$).

Proposition 3.2. *Let us consider D and U just as in Proposition 3.1. If we also assume that D has a regular boundary for Brownian motion, then the function U is continuous on $\overline{D}$.*

The idea in order to prove this Proposition goes as follows: Let $x_0 \in \partial D$ be a regular boundary point and $y \in D$ such that $d(y, x)$ is very small, then with high probability, a Brownian motion started from y will exit D near x_0, so in particular with high probability the value of f at the exit point is close to $f(x_0)$. This is made more precisely by the following Lemma:

Lemma 3.1. *Let x_0 be a regular boundary point of the bounded domain D. For all $\epsilon >$ and $\delta > 0$, there exists $r > 0$, such that for all $x \in \mathcal{B}(x_0, r) \cap D$, we have*

$$\mathbb{P}_x(d(B_t, x_0) > \delta) < \epsilon.$$

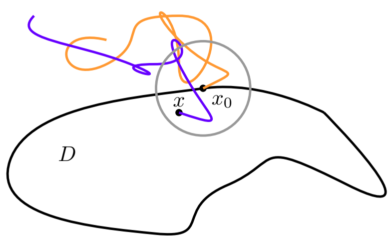


Figure 15: x_0 is a regular boundary point (BM exits D immediately started from there), if we choose x according to the Lemma above and start a BM from there, then with high probability we exit D near x_0.

Proof of Proposition. Let $x_0 \in \partial D$. Our goal is to prove that the function U is continuous at x_0. We notice that $U(x_0) = f(x_0)$ by definition. We know that f is continuous at x_0 and $|f|$ is bounded by $M := \|f\|_\infty < \infty$. Thus by the continuity of f at x_0 we have for all $\epsilon > 0$, the existence of $\delta > 0$ such that for all $y \in \mathcal{B}(x_0, \delta) \cap \partial D : |f(y) - f(x_0)| < \epsilon$. (*)

For this choice of ϵ and δ we use the previous Lemma. It gives us the existence of a radius $r > 0$ such that for all $x \in D \cap \mathcal{B}(x_0, r)$, when B is a Brownian motion started from x, with probability at least $1 - \epsilon$ (very high), one has $d(B_T, x_0) < \delta$, hence we have $B_T \in \mathcal{B}(x_0, \delta) \cap \partial D$ which immediately implies (making use of (*)) that $|f(x_0) - f(B_T)| < \epsilon$.

In the event that $d(B_T, x_0) \geq \delta$ (which happens only with very small probability $\leq \epsilon$), we anyway have that $|f(x_0) - f(B_T)| \leq 2M$. We can then conclude that

$$\begin{aligned}|U(x) - f(x_0)| \leq \mathbb{E}_x(|f(B_T) - f(x_0)|) &\leq \epsilon \overbrace{\mathbb{P}_x(d(B_T, x_0) < \delta)}^{\leq 1} + 2M\mathbb{P}_x(d(B_T, x_0) \geq \delta) \\ &\leq \epsilon + 2M \underbrace{\mathbb{P}_x(d(B_T, x_0) \geq \delta)}_{\leq \epsilon} \leq \epsilon(1 + 2M),\end{aligned}$$

where we again used the previous Lemma. This proves that U is continuous at x_0. □

3.3 Harmonic Functions, Brownian Motion and the Dirichlet Problem

Throughout this section, Ω will denote any (not necessarily bounded) open subset of $\mathbb{R}^d$. There are two natural definitions of harmonic functions: Via the mean value property, or via the Laplacian. As we shall see, these two definitions are in fact equivalent.

Definition 3.3. *We say that the function $h : \Omega \to \mathbb{R}$ is harmonic in Ω, if it is continuous in Ω and if it satisfies the following mean-value property: For each $x \in \Omega$ and each $r < d(x, \partial\Omega)$,*

$$h(x) = \int_{S(x,r)} h(z) d\sigma_{x,r}(z),$$

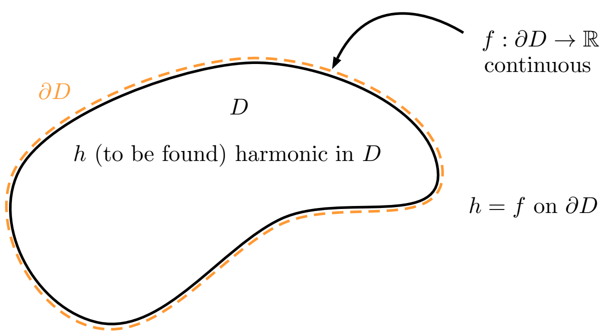
where $\sigma_{x,r}$ is the uniform probability measure on the sphere $S(x,r) = \partial\mathcal{B}(x,r)$.

Remark 3.4.
1. We can read the mean-value property as $h(x) =$ mean-value of h on $S(x,r) = \partial\mathcal{B}(x,r)$.
2. We shall later prove that harmonic functions are in fact necessarily smooth.

We suppose from now on that D denotes a *bounded* open subset of $\mathbb{R}^d$. Suppose we are given a continuous function $f : \partial D \to \mathbb{R}$. Given that ∂D is compact, such a function f is necessarily bounded. We then define:

Definition 3.4. *A function $h : \overline{D} \to \mathbb{R}$ is said to be a solution to the Dirichlet problem in D with boundary values f if it satisfies:*

- *h is harmonic in D.*
- *h is continuous in $\overline{D}$ and $h = f$ on ∂D.*



Here is a simple classical result that we will use

Lemma 3.2 (Maximum principle I). *Suppose that H is a continuous function on $\overline{D}$ that is harmonic in D. Then $\sup_{\overline{D}} H = \max_{\partial D} H$, i.e. such a function takes its maximum value on the boundary of D.*

Corollary 3.1. *There exists at most one solution to the Dirichlet problem.*

Proof. If we have two solutions h_1 and h_2, then $H := h_1 - h_2$ and $-H$ are harmonic in D, continuous in $\overline{D}$ and equal to 0 on ∂D (because they both satisfy the same boundary condition). Applying the maximum principle I to H and $-H$,

$$\begin{cases} \max_{\overline{D}} H = \max_{\partial D} H = \max_{\partial D}(h_1 - h_2) = 0 \\ \max_{\overline{D}}(h_2 - h_1) = 0 \end{cases}$$

from which we conclude that $H = 0$ on $\overline{D}$ and thus, i.e. $h_1 = h_2$ on $\overline{D}$. □

At the end of the previous section, we have used Brownian to construct a concrete candidate for the solution of the Dirichlet problem, and we did in fact show that under some conditions on D (on its boundary), it was a solution to the Dirichlet problem.

Let us repeat the construction here: We suppose that D and f are given as introduced in this section. Then for each $x \in \overline{D}$, we define a Brownian motion B started from x under the probability measure $\mathbb{P}_x$ (we write $\mathbb{E}_x$ for the expectation with respect to this probability measure). We also let T be the first time at which the Brownian motion B hits ∂D. Then we define for each $x \in \overline{D}$,

$$U(x) := \mathbb{E}_x(f(B_T)).$$

If we now combine Proposition 3.1, Proposition 3.2 together with the uniqueness statement of the above corollary we obtain

Theorem 3.1. *If all the boundary points of the bounded domain D are regular, then there exists a unique solution to the Dirichlet problem, and this solution is equal to the function U.*

So, it remains to see what can happen when some boundary points of D are not regular. We will state (and prove) a general fact that is valid for any bounded domain D. Notice that in the next proposition we don't impose any conditions on the boundary of D (i.e. regularity).

Proposition 3.3. *If the solution to the Dirichlet problem in bounded D with boundary values f exists, then it is necessarily equal to the function U.*

As an important consequence to the above Proposition (and the previous Theorem) there are just two options:

- Either U is a solution to the Dirichlet problem and then a solution exists and it is unique,
- or U is not a solution to the Dirichlet problem and then it means that there is no solution to the Dirichlet problem at all (because if it would exist, it would have to be equal to U according to the above Proposition).

Proposition 3.2 shows that the only thing that can prevent U from being a solution to the Dirichlet problem is the possible lack of continuity of U at the boundary points. As we will see a little bit later, it can indeed happen that U has discontinuities at non-regular boundary points.

Proof of Proposition. Suppose that h is a solution to the Dirichlet problem in D with boundary values f. Let us fix $x \in D$ and we want to show that $h(x) = U(x)$. Let B be a Brownian motion started from x, and we define iteratively the following stopping times: $T_0 = 0$ and for each $n \geq 0$, we set $r_n = d(B_{T_n}, \partial D)$ and

$$T_{n+1} = \inf\{t > T_n : d(B_t, B_{T_n}) = r_n/2\}.$$

We notice that $r_0 = d(B_0, \partial D) = d(x, \partial D)$ and $T_1 = \inf\{t > 0 : d(B_t, x) = r_0/2\}$ is the first time we hit the sphere $S(x, r_0/2)$.

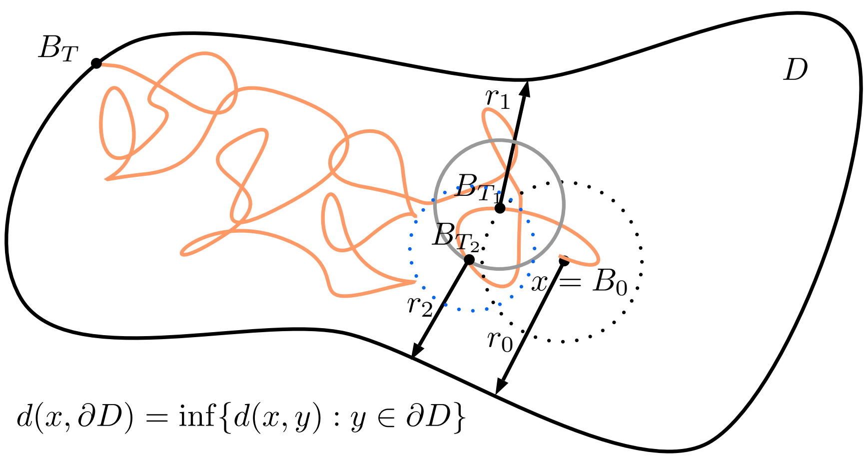


Figure 16: The sequence of the stopping times T_n.

Since we know that almost surely, B is continuous and will exit D almost surely, we can deduce iteratively that $T_n < T < \infty$ almost surely and that $r_n \geq r_0/2^n > 0$. The continuity of Brownian motion also implies that T_n is increasing, thus it converges almost surely to some T' with $T' \leq T$. To see that $T' \geq T$ assume for contradiction that $r' := d(B_{T'}, \partial D) > 0$, then for $n \in \mathbb{N}$ sufficiently large enough we get that

$$d(B_{T_{n+1}}, B_{T_n}) = \frac{r_n}{2} > \frac{r'}{4} > 0.$$

But the continuity of B ensures that B_{T_n} converges to $B_{T'}$ and thus the above is a contradiction, we conclude that $r' = d(B_{T'}, \partial D) = 0$, i.e. $T' \in \partial D$ and since by definition T is the first hitting time of the boundary of D we conclude that $T' \geq T$.

The continuity of h and B now ensures that,

$$h(B_{T_n}) \xrightarrow{a.s.} h(B_T) = f(B_T).$$

On the other hand, using the strong Markov property at time T_n and the harmonicity of h we get that for all $n \in \mathbb{N}_{\geq 1}$, almost surely

$$h(x) = \mathbb{E}_x(h(B_{T_n})).$$

Indeed, for $n = 1$ we have by the harmonicity of h and by the definition of T_1 as the hitting time of sphere with radius $r_0/2$ around x that

$$h(x) = \text{mean value of } h \text{ on } S(x, r_0/2) = \mathbb{E}_x(h(B_{T_1}))$$

Assume now that we have $h(x) = \mathbb{E}_x(h(B_{T_n}))$. We then have by the strong Markov property,

$$\begin{aligned}\mathbb{E}_x(h(B_{T_{n+1}})) = \mathbb{E}_x(\mathbb{E}(h(B_{T_{n+1}}) \mid \mathcal{F}_{T_n})) \overset{1)}{=} \mathbb{E}_x(\mathbb{E}_{B_{T_n}}(h(B_{T_{n+1}}))) &\overset{2)}{=} \mathbb{E}_x(h(B_{T_n})) \\ &= h(x),\end{aligned}$$

where in 1) we used the strong Markov property at time T_n and in 2) the harmonicity of h. Hence, by dominated convergence (h is bounded on $\overline{D}$, because $\overline{D}$ is compact and h is continuous on $\overline{D}$) we get that

$$h(x) = \mathbb{E}_x(h(B_{T_n})) \to \mathbb{E}_x(f(B_T)) = U(x)$$

which concludes the proof. □

Remark 3.5.

1. In class Prof. Werner mentioned that he likes this proof.

3.3.1 Harmonic Functions

Let us come back to the other possible definition of harmonicity:

Proposition 3.4. *Let Ω be an open set in $\mathbb{R}^d$. A function $h : \Omega \to \mathbb{R}$ is harmonic in Ω if and only if h is C^2 in Ω and $\Delta h = 0$ in Ω. Moreover, a harmonic function is necessarily smooth (i.e. C^∞).*

Remark 3.6. Hence, the Dirichlet problem (as we defined it) is equivalent to the more standard phrasing: Find a function h that is continuous in $\overline{D}$, equal to f on ∂D, is C^2 in D and such that $\Delta h = 0$ in D.

Proof very sketchy. "$\Longrightarrow$" is a convolution argument which shows that $h \in C^\infty$, in particular in C^2 and we can use the Taylor expansion and the harmonicity of h to derive $\Delta h = 0$.
"$\Longleftarrow$" Uses the Maximum principle II below: For $r < d(x, \Omega)$ we work on $D = \mathcal{B}(x, r)$ (this domain satisfies property $(\mathcal{P})$) and we know that the D.P. with boundary values given by H has a unique solution given by U. But then U is C^∞ and its Laplacian vanishes in D, by 2MP $U = H$ in D. But then $U(x) = H(x)$ is the mean value of H on $\mathcal{S}(x, r)$.

□

Lemma 3.3 (Maximum principle II)**.** *If h is a C^2 function in D that is continuous on $\overline{D}$ such that $\Delta h \geq 0$ in D, then*

$$\max_{\overline{D}} h = \max_{\partial D} h.$$

3.3.2 Some consequences for multidimensional Brownian motion

In the previous sections, we have used Brownian motion in order to construct the unique possible solution to the Dirichlet problem. In this section, we will use their identification in terms of Brownian motion expectations, in order to deduce some explicit formulas for Brownian motion hitting probabilities.

Let us first look at the case $d = 2$. Let B be a Brownian motion in dimension 2 started from x. Let $H(x) = \log(|x|)$ be defined on $\mathbb{R}^2 \setminus \{0\}$, we can then easily see that H is harmonic because its Laplacian is 0. This means, by the uniqueness of the solution to the Dirichlet problem, that for every bounded domain D with regular boundary and such that $\overline{D} \subset \mathbb{R}^2 \setminus \{0\}$ we have

$$\log(|x|) = \mathbb{E}_x(\log |B_T|), \text{ for all } x \in D,$$

where T denotes the exit time of B from D.

Let us now apply this to a more concrete example.

Example 3.1. *Let us fix $0 < r < R$ and take $D = \mathcal{B}(0,R) \setminus \mathcal{B}(0,r)$, then D has a regular boundary because it satisfies property $(\mathcal{P})$. Let us start a Brownian motion inside $D \subset \mathbb{R}^2 \setminus \{0\}$ and consider the hitting times*

$$\begin{aligned} T_r &= \inf\{t > 0 : |B_t| = r\} \\ T_R &= \inf\{t > 0 : |B_t| = R\} \\ T &= T_r \wedge T_R \text{ is the exit time of } D. \end{aligned}$$

We know that for all $x \in D$ we have

$$\log(|x|) = \mathbb{E}_x(\log|B_T|).$$

Thus we obtain

$$\begin{aligned} \log(|x|) &= \mathbb{E}_x(\log|B_T|1_{T_r<T_R}) + \mathbb{E}_x(\log|B_T|1_{T_R<T_r}) \\ &= \mathbb{E}_x(\log|B_{T_r}|1_{T_r<T_R}) + \mathbb{E}_x(\log(|B_{T_R}|1_{T_R<T_r}) \\ &= \log|r|\mathbb{P}_x(T_r < T_R) + \log|R|\mathbb{P}_x(T_R < T_r) \\ &= \log(r)(1 - \mathbb{P}_x(T_R < T_r)) + \log(R)\mathbb{P}_x(T_R < T_r). \end{aligned}$$

The above yields:

$$\mathbb{P}_x(T_R < T_r) = \frac{\log(|x|) - \log(r)}{\log(R) - \log(r)} \tag{*}$$

If we now let $R \to \infty$, then we get almost surely that $T_R \to \infty$ (by the continuity of Brownian motion) and thus

$$\mathbb{P}_x(T_r = \infty) = \lim_{R\to\infty} \mathbb{P}_x(T_r > T_R) = 0$$

for some fixed $r > 0$ and x with $|x| > r$. Thus we have shown:

Proposition 3.5. *Almost surely, for all fixed $r > 0$, a two dimensional Brownian motion started from x with $|x| > r$ will visit $\mathcal{B}(0,r)$.*

From this proposition we can immediately get:

Theorem 3.2. *For two-dimensional Brownian motion, the set $\{B_t : t \geq 0\}$ is almost surely dense in the plane, i.e. $\overline{B[0,\infty)} = \mathbb{R}^2$.*

Proof. The previous proposition equivalently states that for B Brownian motion in $d = 2$ started from 0 we have for all $x \in \mathbb{R}^2 \setminus \{0\}$ and for all $r < d(x,0)$ almost surely that B will visit the disk $\mathcal{D}(x,r)$. This statement is true almost surely simultaneously for all $x \in \mathbb{Q}^2 \setminus \{0\}$ and for all $r \in \mathbb{Q}_+$, but the set of such rational couples (x,r) (simultaneously) is countable. □

If we now fix x and R in (*) and let instead $r \to 0$, we have $T_r \to 0$ (continuity of Brownian motion), we then easily see that

$$\begin{aligned}\mathbb{P}_x(B \text{ hits the point } 0 \text{ before } T_R) &= \lim_{r\to 0} \mathbb{P}_x(T_r < T_R) \\ &= 1 - \lim_{r\to 0} \mathbb{P}_x(T_R < T_r) = 1 - 1 = 0.\end{aligned}$$

We have thus proven

Proposition 3.6. *Almost surely, a two-dimensional Brownian motion that starts away from the origin will never visit the origin.*

Let us now consider the 3 dimensional case. In this case we know that $H(x) = 1/\|x\|$ defined on $\mathbb{R}^3 \setminus \{0\}$ is Harmonic because its Laplacian is 0.

Example 3.2. *Let us again fix $0 < r < R$, D, T_r, T_R, T all as in the previous example. Let B be a 3 dimensional Brownian motion started inside $D \subset \mathbb{R}^3 \setminus \{0\}$. We then have again for all $x \in D$ that*

$$|x|^{-1} = \mathbb{E}_x(1/|B_T|) = r^{-1}\mathbb{P}_x(T_r < T_R) + R^{-1}(1 - \mathbb{P}_x(T_r < T_R))$$

which implies that

$$\mathbb{P}_x(T_r < T_R) = \frac{|x|^{-1} - R^{-1}}{r^{-1} - R^{-1}}.$$

If we let $R \to \infty$ we get that $T_R \to \infty$ (because BM is continuous) and thus

$$0 < \mathbb{P}_x(T_r < \infty) = \frac{r}{|x|} < 1 \tag{**}$$

where $|x| > r$ (because we start B in D). If we now let $r \to 0$ in (**) we obtain

$$\mathbb{P}_x(B \text{ hits the point } 0 \text{ before it escapes to } \infty) = \lim_{r\to 0} \mathbb{P}_x(T_r < \infty) = 0$$

This (and the fact that when $d \geq 3$, the first three coordinates of a d-dimensional Brownian motion form a 3-dimensional Brownian motion) in turn implies that:

Proposition 3.7. *If B is a Brownian motion in dimension $d \geq 3$, then almost surely it escapes to infinity, in particular $\|B_t\| \to \infty$ as $t \to \infty$.*

Remark 3.7. Recall that for a simple random walk in $\mathbb{Z}^d$, it is recurrent when $d = 2$ and transient when $d \geq 3$. So, we see that there is a little difference in the way in which two-dimensional Brownian motion is recurrent (it almost surely visits all open sets infinitely often) compared to the recurrence of simple random walk (that almost surely visits every point infinitely often).

3.4 Dirichlet Problem on non-regular boundary points

Let us now briefly come back to the discussion about the existence of bounded domains D for which the solution to the Dirichlet problem might not exist.

Example 3.3. *Arguably the simplest example, is the case of the planar domain $D = \mathcal{B}(0,1) \setminus \{0\}$, i.e. the unit disk minus its center. Then, it is clear that the origin is not a regular boundary point. Indeed, we know that a Brownian motion started from the origin will almost surely never return to the origin. So if we define*

$$f(x) = \begin{cases} f(x) = 1, & \text{on } \partial B(0,1) \\ f(0) = 0 \end{cases}$$

then $f : \partial D \to \mathbb{R}$ is continuous on the boundary of D (i.e. on $\partial\mathcal{B}(0,1) \cup \{0\}$).

We then have for all $x \in D$

$$U(x) = \mathbb{E}_x(f(B_T)) = 1,$$

because Brownian motion started from $x \in D$ will almost surely exit D on the unit circle. But

$$U(x) \overset{x\to 0}{\longrightarrow} 1 \neq U(0) = f(0) = 0.$$

Hence U is not continuous at $x = 0$ and consequently 0 is not a regular boundary point. So, there is no solution to the Dirichlet problem for this choice of D and f.

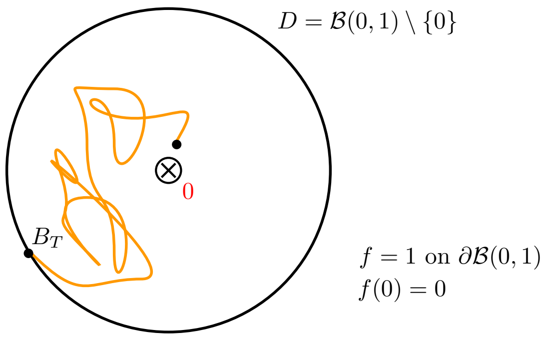


Figure 17: The above example suits to give a non regular boundary point $x = 0$.

Remark 3.8. We highlight that the most important argument used in this example is that a 2-dimensional Brownian motion started away from the origin will never visit the origin again.

Example 3.4. *Another example is the following three-dimensional counterpart to the previous example. Let us consider the unit ball $\mathcal{B}(0,1)$ in $\mathbb{R}^3$. Let $u = (0,0,1) \in \mathbb{R}^3$ be the north pole of the unit sphere and let us denote by I the segment that joins the origin to the point u, i.e. $I = [0,u]$. We then choose $D = \mathcal{B}(0,1) \setminus I$.*

We then obtain a continuous function $f : \partial D = \mathcal{S}(0,1) \cup I \to \mathbb{R}$ by setting

$$f(x) = \begin{cases} f(x) = 1, & \text{on } \mathcal{S}(0,1) \\ f(x) = 0, & \text{on } I \end{cases}$$

We know that almost surely a two-diemsnional Brownian never hits the origin at positive times. Hence, the Brownian motion B started in $x \in D$ will not hit I at positive times (because the projection of such a BM onto the horizontal plane is a 2d BM that will a.s. never visit the origin at positive times).

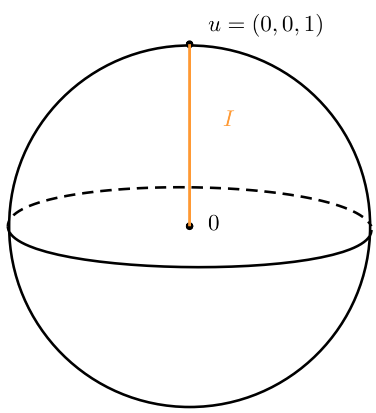
Thus we have $B_T \in \mathcal{S}(0,1)$ whenever $B_0 \notin I$. Thus we obtain that every $x \in [0,u)$ (notice right open, u is regular because of the outside cone condition) is not a regular boundary point. Indeed we have for all $x \in D$

$$U(x) = \mathbb{E}_x(f(B_T)) = 1$$

in particular we obtain that

$$U(x) \xrightarrow{x \to 0} 1 \neq U(0) = f(0) = 0$$

This shows that for all $x \in [0,u)$ that U is not continuous at such a x. In particular all said x are not regular boundary points for D and there exists no solution to the Dirichlet problem for said choice of D and f.



Example 3.5. *One may wonder whether the result we obtained from the previous example is only due to the fact that I is "one-dimensional" and has no interior. Here is another, more involved example, that shows that this is not the case. Recall that a Brownian motion B started from 0 in $\mathbb{R}^3$ does almost never visit the line that goes through the origin and $u = (0, 0, 1)$ at positive times and that $\|B_t\| \to \infty$ (which implies that the set of points visited by B is closed).*

Let $(r_n)_{n\in\mathbb{N}}$ be a sequence of radii such that $r_n \to 0$ as $n \to \infty$. Let c_n be the closed segment of radius r_n between the points $(0, 0, 2^{-n})$ and $(0, 0, 2^{-(n+1)})$.

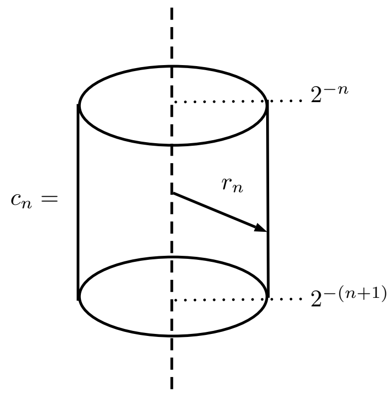


Figure 18: The c_n as depicted as above are the "drills" we're going to apply to sphere, see figure further below.

We then define: $D = \mathcal{B}(0, 1) \setminus \left(\overline{\bigcup_{n\geq 0} c_n}\right)$

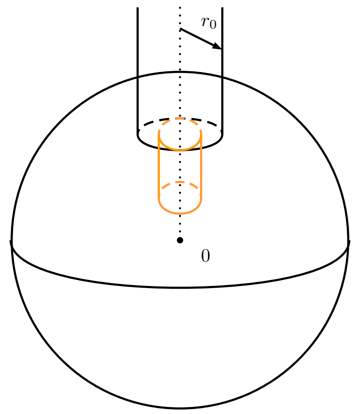


Figure 19: We apply gradually finer and finer drills (c_n above) to $\mathcal{B}(0, 1)$ until we reach 0.

Let us define $I = [u/2, u] \subset \mathbb{R}^3$. *This is a compact set and thus the distance between the closed set* $\{B_t, t \geq 0\}$ *is strictly positive. I.e. if we take a Brownian motion* B *started from* 0 *then almost surely* $d(\{B_t, t \geq 0\}, I) > 0$ *which gives that*

$$\mathbb{P}(d(\{B_t, t \geq 0\}, I) < r) \to 0 \text{ as } r \to 0.$$

Let us choose r'_n *such that*

$$\mathbb{P}(d(B, I) < r'_n) \leq \frac{1}{2^{n+2}}.$$

By scaling invariance of Brownian motion we obtain that

$$\mathbb{P}(d(B, 2^{-n}I) < 2^{-n}r'_n) = \mathbb{P}(d(B, I) < r'_n) \leq \frac{1}{2^{n+2}},$$

therefore we get

$$\mathbb{P}(\exists n \geq 0 : d(B, 2^{-n}I) < 2^{-n}r'_n) \leq \sum_{n \geq 0} \frac{1}{2^{n+2}} = \frac{1}{2}$$

and consequently

$$\mathbb{P}\left(B \text{ never intersects } \bigcup_{n \geq 0} \{z : d(z, 2^{-n}I) < 2^{-n}r'_n\}\right) \geq \frac{1}{2}.$$

So if we define $r_n := 2^{-n}r'_n$, *then the above yields that a Brownian motion started from* 0 *has a probability of at least* $1/2$ *to exit* D *through* $\partial B(0,1)$. *In other words* 0 *is not a regular boundary point of* D.

We can then define the function $f(x) = \|x\|$ *on* ∂D, *and we can show that there exists a sequence* x_n *in* D *such that* $x_n \to 0$ *but* $U(x_n) \not\to 0 = U(0) = f(0) = 0$. *(Will be discussed in more detail in the exercise sheets).*

3.4.1 Some final remarks on the Dirichlet problem

We want to give some concluding remarks on the Dirichlet problem. In particular one should pay attention to the second remark presented here, where we discuss how odd things can happen once unbounded domains D are considered (we can for instance lose uniqueness of solutions).

Remark 3.9. It is worthwhile to stress that while for most domains (for instance bounded domains that satisfy the exterior cone condition), it is possible to prove existence and uniqueness of the solution to the Dirichlet problem without the use of probability theory and Brownian motion.

However, Brownian motion appears to be the right tool when one tries to describe exactly which bounded domains have the property that for all continuous functions f, the solution to the Dirichlet problem exists. Indeed, the previous ideas (theorems and examples) suggest that a **necessary and sufficient condition** is exactly that all boundary points are **regular**. This will be derived in the exercise sheets.

Remark 3.10. When the domain D is not bounded and $d \geq 3$, one has to be careful because it could happen that with positive probability, a Brownian motion started in D does in fact stay in D forever.

For instance, if one considers $D = \mathbb{R}^d \setminus \overline{\mathcal{B}(0,1)}$, we have seen that with positive probability, a Brownian motion started form $x \in D$ does wander away to infinity without hitting $\mathcal{S}(0,1)$ (see (**) before Proposition 3.7), i.e. the exit time T of D is infinite with positive probability.

So we see that for all $c \in \mathbb{R}$, the functions

$$H(x) := c\mathbb{P}_x(T = \infty)$$

are all harmonic in D, continuous on $\overline{D}$ and equal to 0 on $\mathcal{S}(0,1)$. In particular the solution to the Dirichlet problem in D for $f = 0$ is therefore not unique.

3.5 Story time

In this section we briefly discuss some further features of Brownian motion. The content of this section is more for *scientific culture* and should be viewed more as a story telling session without full proofs, where only the main ideas of the proofs will be hinted at.

3.5.1 Polar sets for planar Brownian motion

Suppose we consider a compact set K in the plane and a planar Brownian motion started at $x \notin K$. We can ask which sets K do have the property that there will almost surely never be hit by B.

Definition 3.5. *If $\mathbb{P}(\exists t \geq 0 : B_t \in K) = 0$, then we say that K is a polar set for our planar Brownian motion.*

We want to study which compact sets in $\mathbb{R}^2$ a Brownian motion does not hit. So what do we know?

1. If K is a singleton, then it is polar (because planar Brownian motion almost surely does not visit a given point).
   (a) Consequently, if K is countable, then it is polar (because the union of countably many events with probability 0 has probability 0).
2. If K is a line segment $[a, b]$, then it is not polar (i.e. $\mathbb{P}(B \text{ hits } K) = 1$).
3. If $\overset{\circ}{K} \neq \emptyset$, then K is not polar.

The question we are going to address here is what happens to sets K such as the classical Cantor set, which is a subset of an interval with length zero. Recall briefly the construction of the Cantor set:

We start with $I_0 = [0, 1] \subset \mathbb{C} \sim \mathbb{R}^2$ (unit length). We then define iteratively

$$I_1 = [0, 1/3] \cup [2/3, 1]$$

I_2 is obtained by cutting out the (open) middle third of each of the 2 closed intervals of I_1.
I_{k+1} is obtained by cutting out the (open) middle third of each of the 2^k closed intervals (of length 3^{-k}) that I_k is the union of. It is then well known that $K := \cap_k I_k$ is a compact set, with zero length, but it is uncountable and has the same cardinality as $\mathbb{R}$.

So we ask us the question, does Brownian motion hit the Cantor set K?

Proposition 3.8. *The Cantor set K is not a polar set for planar Brownian motion (i.e. B hits K almost surely).*

This shows that while two-dimensional Brownian motion does not hit points, it nevertheless hits fairly small sets.

3.5.2 Two cousins of the Dirichlet problem

We are now going to quickly discuss how problems related to the Dirichlet problem can also be interpreted and solved using Brownian motion.

Just as in the Dirichlet problem, we are going to suppose that D is an open bounded subset of $\mathbb{R}^d$ with a regular boundary (all boundary points are regular boundary points), and we are given a real valued continuous function f defined on ∂D. Let us also assume that we are given a constant $\alpha \geq 0$.

Goal: Find a function h such that:

- h is C^2 in D and in D satisfies $\Delta h = \alpha h$.
- h is continuous on $\overline{D}$ and $h = f$ on ∂D.

Remark 3.11. We notice that in the special case of $\alpha = 0$ this is just classical Dirichlet problem that we have already studied.

Proposition 3.9. *The solution to this new problem exists and is unique, and it is equal to the function*

$$U(x) = \mathbb{E}_x(f(B_T)e^{-\alpha T}), \text{ for all } x \in \overline{D}.$$

Remark 3.12. We notice again that for $\alpha = 0$ we just obtain again the classical solution to the Dirichlet problem.

We will prove this result later with the help of stochastic calculus.

Yet another variant of the Dirichlet problem is the following: We still consider a bounded domain D in $\mathbb{R}^d$, and we are given a constant β. This time we look for a solution to the following problem: Find a function H that is continuous in $\overline{D}$, equal to 0 on ∂D, that is C^2 in D and that satisfies $\Delta H = -2\beta$ in D.

Proposition 3.10. *There exists a unique solution to this problem, and it is given by the function*

$$U(x) = \mathbb{E}_x(\beta T), \text{ for all } x \in \overline{D}.$$

4 Donsker's Theorem

We will now briefly outline two proofs of Donsker's theorem.

4.1 Formulation of the theorem

- We work on the metric space (E, d) where $E = \mathcal{C}([0, 1] \to \mathbb{R})$ denotes the space of continuous function from $[0, 1]$ to $\mathbb{R}$, endowed with the metric $d(f, g) = \max_{[0,1]} |f - g|$.
- The Wiener measure is then the law of $(B_t, t \in [0, 1])$ viewed as an element of said metric space (E, d).
- On the same metric space we consider the *"interpolated, rescaled random walk"* defined as follows:
  1. We start with a standard random walk $S_n := X_1 + \cdots + X_n$, i.e. $S_0 := 0$ and the X_i are i.i.d. coin flips.
  2. We extend this into a continuous function $(S_t, t \geq 0)$ by linearly interpolating the previous sequence on each interval $[n, n + 1]$.
  3. For each $\epsilon > 0$, we define the rescaled Random Walk (notice that we both scale time and our "height") by

     $$S_t^{(\epsilon)} := \epsilon S_{\epsilon^{-2}t}.$$

     We then look at $(S_t^{(\epsilon)}, t \leq 1)$ as a random element in (E, d).

Theorem 4.1 (Donsker's invariance theorem)**.** *The law of* $(S_t^{(\epsilon)}, t \leq 1)$ *(viewed as a measure on* (E, d)*) converges weakly as* $\epsilon \to 0$ *to the law of* B *(viewed as a law in* (E, d) *i.e. Wiener measure).*

Remark 4.1. This means that for all continuous functions $F : (E, d) \to \mathbb{R}$ we have

$$\mathbb{E}(F(S^{(\epsilon)})) \xrightarrow{\epsilon \to 0} \mathbb{E}(F(B))$$

Also it is worth to mention that if we work with $F(f) := f(1)$ (which is continuous with respect to d) then this can be used to deduce the regular Central Limit Theorem (substitute $u = \frac{1}{\epsilon^2}$ so as $\epsilon \to 0$ we have $u \to \infty$ and $\epsilon = \frac{1}{\sqrt{u}}$)

4.2 Warm-up: Doob's inequality for discrete martingales

A general useful tool to have information about *"regularity of entire trajectories"* is Doob's inequality for discrete martingales. We provide here a quick recap of the various results.

We write

$$M_n^* := \max_{0 \leq j \leq n} |M_j|.$$

Trivially we have $M_n^* \geq |M_n|$. The purpose of Doob's inequalities is to conversely provide upper bounds for M_n^* in terms of $|M_n|$:

- **Doob's maximal inequality:** If $(M_n)_{n\geq 0}$ is a martingale with respect to the filtration $(\mathcal{F}_n)_{n\geq 0}$ such that $M_0 = 0$ almost surely, then for all $\lambda > 0$ and all $N \geq 0$:

$$\lambda \mathbb{P}(M_N^* \geq \lambda) \leq \mathbb{E}(|M_N| 1_{M_N^* \geq \lambda}) \leq \mathbb{E}(|M_N|).$$

- **Doob's L^2 inequality:** If furthermore the random variables M_n are in L^2, then one always has

$$\mathbb{E}((M_N^*)^2) \leq 4\mathbb{E}(M_N^2).$$

  This is very useful as it allows to control the whole trajectory of M until time N via the law of M_N only.

- **Doob's L^p inequality:** Similarly, if the random variables M_n are in L^p for some $p > 1$, then

$$\mathbb{E}((M_N^*)^p) \leq (p/(p-1))^p \mathbb{E}(|M_N|^p).$$

4.3 Proof via coupling

Proof of Donsker's theorem. First we roughly describe the idea of the proof. We start with a Brownian motion $(B_t)_{t\geq 0}$. Then we define deterministically (as a function of B) for each given $\epsilon > 0$ a rescaled Random Walk $S^{(\epsilon)}$ that is going to be very close to B.

Let us define for each $\epsilon > 0$ a sequence of T_n^ϵ of stopping times by: $T_0^\epsilon = 0$

$$\begin{aligned} T_1^\epsilon &= \inf\{t > 0 : |B_t| = \epsilon\} \\ T_n^\epsilon &= \min\{t > T_{n-1}^\epsilon : |B_t - B_{T_{n-1}^\epsilon}| = \epsilon\}. \end{aligned}$$

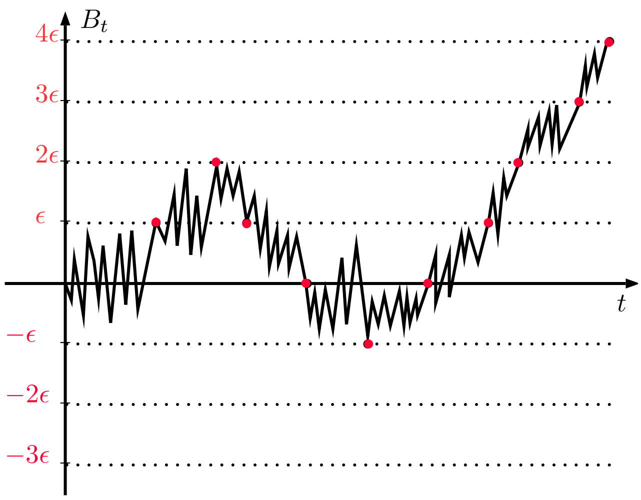


Figure 20: The stopping times T_n^ϵ indicated with red dots.

We then define

$$S_n^\epsilon := \frac{B_{T_n^\epsilon}}{\epsilon},$$

we notice that in our realisation as in the figure above we have $S_1^\epsilon = 1$ whereas $S_5^\epsilon = -1$. We see that $(S_n^\epsilon)_{n\geq 0}$ is a simple random walk. We will later discuss this in more detail.

Next we define $(S_t^\epsilon)_{t\geq 0}$ by linear interpolation on each $[n, n+1]$. Then the rescaled random walk

$$(S_t^{(\epsilon)} = \epsilon S_{\epsilon^{-2}t}^\epsilon)_{t\geq 0}.$$

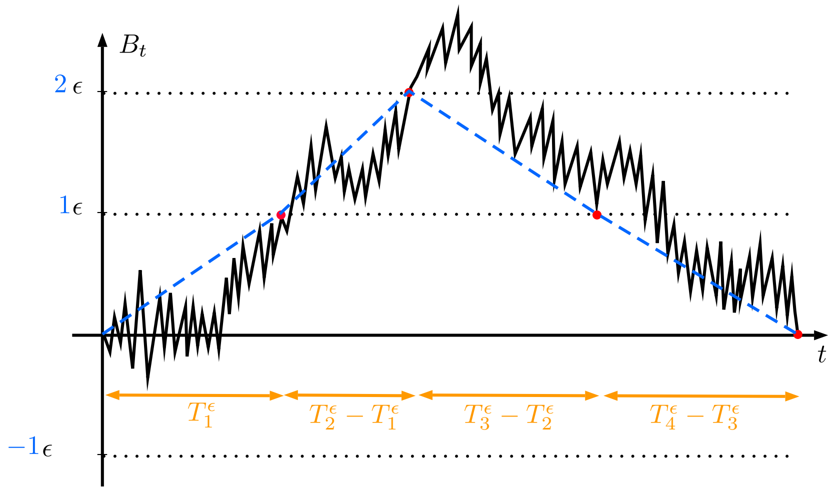


Figure 21: We notice ($\approx$blue line) that $S_1^\epsilon = 1$, $S_2^\epsilon = 2$, $S_3^\epsilon = 1$, $S_4^\epsilon = 0$. So our increments are ± 1.

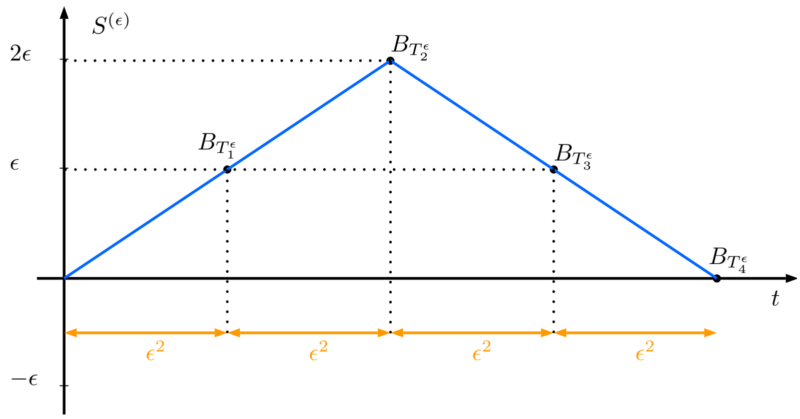


Figure 22: The graph of $S^{(\epsilon)}$, notice that for each $t = n\epsilon^2$, the value of $S^{(\epsilon)}$ is exactly $B_{T_n^\epsilon}$.

We observe:

- $(T_{n+1}^\epsilon - T_n^\epsilon)_{n\geq 0}$ is a sequence of i.i.d. RV (strong Markov property).
- By scaling: The law of $\tau_n := \dfrac{T_{n+1}^\epsilon - T_n^\epsilon}{\epsilon^2}$ does not depend on ϵ.
- $\mathbb{E}(\tau_1) = 1$ (using $B_t^2 - t$ is martingale) $\rightsquigarrow \mathbb{E}(T_n^\epsilon) = n\epsilon^2$ (by independence).
- The random variables $(X_n^\epsilon := \epsilon^{-1}(B_{T_n^\epsilon} - B_{T_{n-1}^\epsilon})$ are i.i.d. coin flips.

Goal: We want to show that $(S_s^{(\epsilon)})_{s\leq 1}$ converges in law/weakly in (E,d) to $(B_s, s \leq 1)$ as $\epsilon \to 0$.

We will show a stronger statement, i.e. that $S_s^{(\epsilon)}$ is actually very close to B in the sup norm. More precisely, we want to show that

$$\forall \eta > 0, \forall \alpha > 0, \exists \epsilon_0 : \forall \epsilon \text{ with } \epsilon < \epsilon_0,\ \mathbb{P}(\sup_{s\in[0,1]} |B_s - S_s^{(\epsilon)}| > \alpha) < \eta. \qquad (*)$$

In order to establish (*) we will derive another inequality: For each $\epsilon > 0$ and each $s \leq 1$, let us define $t_\epsilon(s)$ to be the largest multiple of ϵ^2 such that $t_\epsilon(s) \leq s$, we then get:

$$\begin{aligned}
\sup_{s\in[0,1]} |B_s - S_s^{(\epsilon)}| &= \sup_{s\in[0,1]} |B_s - B_{t_\epsilon(s)} + B_{t_\epsilon(s)} - S^{(\epsilon)}_{t_\epsilon(s)} + S^{(\epsilon)}_{t_\epsilon(s)} - S_s^{(\epsilon)}| \\
&\leq \sup_{s\in[0,1]} |B_s - B_{t_\epsilon(s)}| + \sup_{s\leq 1} |S_s^{(\epsilon)} - S^{(\epsilon)}_{t_\epsilon(s)}| + \sup_{s\leq 1} |B_{t_\epsilon(s)} - S^{(\epsilon)}_{t_\epsilon(s)}| \\
&\leq \sup_{s\in[0,1]} |B_s - B_{t_\epsilon(s)}| + \sup_{s\leq 1} |S_s^{(\epsilon)} - S^{(\epsilon)}_{t_\epsilon(s)}| + \max_{j\leq\epsilon^{-2}} |B_{j\epsilon^2} - S^{(\epsilon)}_{j\epsilon^2}| \\
&= \sup_{s\in[0,1]} |B_s - B_{t_\epsilon(s)}| + \sup_{s\leq 1} |S_s^{(\epsilon)} - S^{(\epsilon)}_{t_\epsilon(s)}| + \max_{j\leq\epsilon^{-2}} |B_{j\epsilon^2} - B_{T_j^\epsilon}|
\end{aligned}$$

We immediately notice that by definition we have

$$|S^{(\epsilon)}_{t_\epsilon(s)} - S_s^{(\epsilon)}| \leq \epsilon.$$

So we will choose $\epsilon_0 \leq \alpha/4$ such that finally we want to get an upper bound of $\sup |B_s - S_s^{(\epsilon)}| \leq 3\alpha/4$.

On the other hand, we know that almost surely B is continuous on $[0,2]$, so that almost surely, for all given $\alpha > 0$, there exists a (random) $\delta > 0$ such that for all $s, t \leq 2$ with $|s-t| \leq \delta$ we have

$$|B_s - B_t| \leq \frac{\alpha}{4}. \qquad (**)$$

For all $\eta > 0$, we can always choose $\delta_0 = \delta_0(\alpha, \eta)$ small enough, such that the random $\delta > 0$ above is larger than δ_0 with probability at least $1 - \eta/4$.

Hence, on this event of probability at least $1 - \eta/4$, by (**) in order to ensure that

$$\sup_{s\in[0,1]} |B_s - B_{t_\epsilon(s)}| + \max_{j\leq\epsilon^{-2}} |B_{j\epsilon^2} - B_{T_j^\epsilon}| \leq \frac{\alpha}{2}$$

it is enough to force that

$$\sup_{s\leq 1} |s - t_\epsilon(s)| \leq \delta_0 \text{ and } \max_{j\leq\epsilon^{-2}} |j\epsilon^2 - T_j^\epsilon| \leq \delta_0.$$

We first notice that $|s - t_\epsilon(s)| \leq \epsilon^2$, so we choose $\epsilon_0^2 \leq \delta_0$.

In order to establish the second bound, we consider

$$M_n := T_n^\epsilon - \mathbb{E}(T_n^\epsilon) = (T_1^\epsilon - \mathbb{E}(T_1^\epsilon)) + (T_2^\epsilon - T_1^\epsilon - \mathbb{E}(T_1^\epsilon)) + \cdots + (T_n^\epsilon - T_{n-1}^\epsilon - \mathbb{E}(T_1^\epsilon))$$

Since M_n can be written as the sum of i.i.d. Random Variables with mean 0 we can conclude that $(M_n)_{n\geq 0}$ is a martingale. Moreover it has finite bounded variance, in particular we can use Doob's L^2-inequality (note that $M_n = T_n^\epsilon - n\epsilon^2$). We denote by u_ϵ the integer part of ϵ^{-2}:

$$\begin{aligned}
\mathbb{E}((\max_{j\leq u_\epsilon}(T_j^\epsilon - j\epsilon^2)^2) &\leq 4\mathbb{E}((T_{u_\epsilon}^\epsilon - u_\epsilon\epsilon^2)^2) = 4\mathrm{var}(T_{u_\epsilon}^\epsilon) \\
&= 4\mathrm{var}(T_1^\epsilon + (T_2^\epsilon - T_1^\epsilon) + \cdots + (T_{u_\epsilon}^\epsilon - T_{u_\epsilon - 1}^\epsilon)) \\
&= 4u_\epsilon \mathrm{var}(T_1^\epsilon) \leq 4\epsilon^{-2}\mathrm{var}(T_1^\epsilon) \leq 4\epsilon^{-2}\epsilon^4\mathrm{var}(T_1^1) \\
&= 4\epsilon^2\mathrm{var}(T_1^1)
\end{aligned}$$

where the last inequality is just due to scaling. Hence, by Markov's inequality we obtain,

$$\mathbb{P}(\max_{j\leq u_\epsilon}|T_j^\epsilon - j\epsilon^2| \geq \sqrt{\epsilon}) \leq \epsilon^{-1}\mathbb{E}((\max_{j\leq u_\epsilon}(T_j^\epsilon - j\epsilon^2))^2) \leq 4\epsilon\mathrm{var}(T_1^1) \overset{!}{<} \frac{\eta}{4},$$

hence, if we choose $\epsilon_0 \leq \delta_0^2$ and $\epsilon_0 \leq \eta/(16\mathrm{var}(T_1^1))$, we get as soon as $\epsilon \leq \epsilon_0$, with probability at least $1 - \eta/4$,

$$\max_{j\leq u_\epsilon}|T_j^\epsilon - j\epsilon^2| \leq \delta_0.$$

In conclusion, for all $\alpha > 0$ and $\eta > 0$, if we choose
$\epsilon_0 \leq \min(\delta_0^2, \eta/(16\mathrm{var}(T_1^1)), \sqrt{\delta_0}, \alpha/4)$ for $\delta_0 = \delta_0(\eta, \alpha)$, we get indeed that for all $\epsilon \leq \epsilon_0$,

$$\mathbb{P}(\sup_{s\in[0,1]}|B_s - S_s^{(\epsilon)}| > \alpha) < \eta,$$

which concludes the proof of Donsker's theorem. □

Remark 4.2. We stress the fact, that the approach we choose to prove Donsker's theorem here was that of a natural embedding of a random walk within a Brownian motion. This is an idea that will resurface again when we will study general continuous martingales.

4.3.1 Some consequences

Donsker's theorem allows for instance to derive results for the limiting distribution of *continuous functionals* of simple random walks in terms of the corresponding continuous functionals of Brownian motion. More precisely, if F is a continuous function from the space of continuous function on $[0,1]$ (endowed with the sup norm) onto some other metric space, then Donsker's theorem shows that the law of $F((S_t^{(\epsilon)}, t \in [0,1]))$ converges weakly to that of $F((B_t, t \in [0,1]))$ as $\epsilon \to 0$.

Example 4.1. *Let f^* denote the function that associates to a continuous function f its maximum on $[0,1]$, i.e. $f^* : (\mathcal{C}, d) \to \mathbb{R}$ given by*

$$f^*(f) := \max_{t \in [0,1]} f(t).$$

By Donsker's theorem we get that $\max_{[0,1]} S^{(\epsilon)}$ converges to the law of $\max_{[0,1]} B$ as $\epsilon \to 0$.

Example 4.2. *We define for each $f \in \mathcal{C}$:*

$$f^*(t) := \sup_{s \le t} f(s)$$
$$f^{\#}(t) := \inf_{s \le t} f(s)$$

then $f^*, f^{\#}$ are again in $\mathcal{C}$, we then define

$$F : \begin{cases} E & \longrightarrow E^3 \\ f & \longmapsto (f, f^*, f^{\#}) \end{cases}$$

Donsker's theorem then implies that weak convergence of the triples

$$(S^{(\epsilon)}, (S^{(\epsilon)})^*, (S^{(\epsilon)})^{\#}) \overset{\epsilon \to 0}{\Longrightarrow} (B, B^*, B^{\#})$$

It will be shown in the exercise sheet that one can use Donsker's theorem to prove the following result:

Proposition 4.1 (Paul Lévy - one of his many results on BM)**.**
Let $(B_t, t \ge 0)$ be a one dimensional Brownian motion started from the origin and define $\overline{B}_t := \max_{[0,t]} B$. Then the process $(\overline{B}_t - B_t, t \ge 0)$ has the same law as the process $(|B_t|, t \ge 0)$.

Remark 4.3. Recall that it is possible to see as a rather consequence of the reflection principle that for a *fixed* time t, the law of $\overline{B}_t - B_t$ is the same as that of $|B_t|$. The present statement is much stronger because it shows the identity in the law of the entire processes. This result will be derived (with two approaches, one using Donsker's thm) in the exercise sheet 6.

4.4 Proof via compactness arguments

We briefly discuss an alternate way how once could prove Donsker's theorem. We will only quickly browse through the arguments here, since the ideas presented wont be used again later on.

We start with a general fact:

- If $(P_n)_{n\in\mathbb{N}}$ is a sequence of probability measures on a compact metric space (K,d), then there exists a subsequence (n_k) such that P_{n_k} converges weakly to some probability measure P on K.

This result (and its proof) is a rather direct extension of the usual diagonal extraction trick.

Definition 4.1. *If $(P_n)_{n\in\mathbb{N}}$ is a sequence of probability measures in a separable metric space, then the sequence is said to be* ***tight*** *if for all $\epsilon > 0$ there exists a compact set K, such that for all $n \in \mathbb{N}$*

$$P_n(K) \geq 1-\epsilon.$$

Theorem 4.2 (Prokhorov Theorem 1/2)**.** *If $(P_n)_{n\in\mathbb{N}}$ is a tight sequence of probability measures in a separable metric space (E,d), then there exists a subsequence (n_k) such that P_{n_k} converges weakly to some probability measure $\mathbb{P}$ on E.*

We will apply this as follows in order to conclude Donsker's theorem:

- We consider (E,d) the metric space of continuous functions on $[0,1]$ to $\mathbb{R}$ endowed with the sup norm, which is indeed a separable metric space.
- Let $(P_n)_{n\in\mathbb{N}}$ be a sequence of probability measures given by the law of $S^{(\epsilon_n)}$ for some $\epsilon_n \to 0$ as $n \to \infty$.
- We ask if the sequence $(P_n)_{n\in\mathbb{N}}$ is tight?

If we can prove that $(P_n)_{n\in\mathbb{N}}$ is tight, then by Prokhorov's theorem there exists n_k and a probability measure $\mathbb{P}$ on (E,d) such that $P_{n_k} \implies \mathbb{P}$. In particular we have for all $0 \leq t_1 < \cdots < t_p \leq 1$, the function

$$F : \begin{cases} (E,d) & \longrightarrow \mathbb{R}^p \\ f & \longmapsto (f(t_1),\ldots,f(t_p)) \end{cases}$$

then F is continuous and thus we have $F(S^{(\epsilon_{n_k})}) \implies F(f)$ where $f \sim \mathbb{P}$. But we also know by the usual CLT that $(S^{(\epsilon_{n_k})}(t_1),\ldots,S^{(\epsilon_{n_k})}(t_p)) \implies (B_{t_1},\ldots,B_{t_p})$ where B is a BM, we conclude that $\mathbb{P}$ must be the Wiener measure.

So, in order to conclude Donsker's invariance theorem with this approach, it suffices to show that $(P_n)_{n\in\mathbb{N}}$ is indeed a tight sequence. This is part of the exercise sheet 7.

For this, we can for instance use the following compact subset of E: For each $M > 0$ let us define

$$K_M := \{f \in E : f(0) = 0 \text{ and } \forall s, t \in [0,1],\ |f(t) - f(s)| \leq M|t-s|^{1/8}\}.$$

The set K_M is then a compact subset of E by standard Arzelà-Ascoli considerations. On the other hand, it is not very difficult to show that for every $\eta > 0$, there exists M_η such that for all $\epsilon \leq 1$ we have

$$\mathbb{P}(S^{(\epsilon)} \in K_M) \geq 1 - \eta.$$

which shows that the sequence P_n is tight and provides another proof of Donsker's invariance principle.

5 Continuous martingales

5.1 Continuous martingales, convergence and optional stopping theorems

5.1.1 Basic defintions, Doob's inequalities

We work on a **filtered probability space**, i.e. a space $(\Omega, (\mathcal{F})_{t\geq 0}, \mathcal{F}, \mathbb{P})$ where $(\mathcal{F}_t)_{t\geq 0}$ is a given filtration (meaning that for all $s \leq t$ we have $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ and $\mathcal{F}_t$ is a σ-field). The way to think about it is that t represents time and that $\mathcal{F}_t$ corresponds to what one can observe at time t.
Typically one defines $\mathcal{F}_\infty := \sigma(\bigcup_{t\geq 0} \mathcal{F}_t)$ and in general, one does not always have $\mathcal{F}_\infty = \mathcal{F}$.

It can be convenient (in order to simplify some minor measurability issues) to consider a filtered probability space that satisfies the so-called **usual conditions**. These are two conditions we want our filtration to satisfy, namely:

1. The filtration is **complete**: The σ-field $\mathcal{F}_0$ (and therefore all $\mathcal{F}_t$) does contain all sets $A \subset \Omega$ (not necessarily measurable) that are contained in a measurable set $A' \in \mathcal{F}$ of probability zero.

2. The filtration is **right-continuous**: For each $t \geq 0$ we have

$$\mathcal{F}_t = \bigcap_{\epsilon > 0} \mathcal{F}_{t+\epsilon},$$

   i.e. at time t we already have the maximum amount of information we can have, we don't need to look into an "ϵ-future".

Definition 5.1. *A process $(X_t)_{t\geq 0}$ is said to be **adapted** to the filtration $(\mathcal{F}_t)_{t\geq 0}$ if for each $t \geq 0$, the random variable X_t is $\mathcal{F}_t$-measurable. We say that the process is in L^1 if for each $t \geq 0$ the random variable X_t is in L^1.*

Definition 5.2. *A process $(M_t)_{t\geq 0}$ defined in this filtered probability space is said to be a continuous martingale with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$ if it is an L^1 process that is adapted to this filtration and such that*

- *For all $s \leq t$, $M_s = \mathbb{E}(M_t \mid \mathcal{F}_s)$ almost surely.*
- *There exists an event of probability 1 such that on this event, $t \mapsto M_t$ is a continuous function on $\mathbb{R}_+$.*

Remark 5.1.

1. Note that the first property in the definition above can equivalently be rewritten as for all $t, s \geq 0$ we have $\mathbb{E}(M_{t+s} \mid \mathcal{F}_s) = M_s$.
2. Of course an example of a continuous martingale is Brownian motion, indeed let $s, t \geq 0$, by the weak Markov property we know that $\tilde{B}_t = B_{t+s} - B_s$ is a Brownian motion which is independent of $\mathcal{F}_s$, hence:
$$\begin{aligned}\mathbb{E}(B_{t+s} \mid \mathcal{F}_s) &= \mathbb{E}(\tilde{B}_t + B_s \mid \mathcal{F}_s) = B_s + \mathbb{E}(\tilde{B}_t \mid \mathcal{F}_s) \\ &= B_s + \mathbb{E}(\tilde{B}_t) = B_s.\end{aligned}$$
3. As a more general remark, it is apparent from the definition of a martingale, that it depends on the filtration $(\mathcal{F}_t)_{t\geq 0}$, it is thus good practice to actually call it an $\mathcal{F}_t$-martingale, to highlight this dependency. However in this course we usually deal with a given filtered probability space or we work with a canonical filtration (for example the Brownian filtration). Hence we use the terminology of a continuous martingale.

Very loosely speaking the outcome of our study will be that a continuous martingale can (always) be obtained as follows: One has on the one hand a Brownian motion $(B_s)_{s\geq 0}$ and also some other available randomness (at request). The time s of a Brownian motion is a priori not the same as the time of the continuous martingale $(M_t)_{t\geq 0}$. Then, one explores (starting from B_0) the profile of the function B by moving to the right in a continuous way. In other words, one has a (possibly random) continuous non-decreasing function $t \mapsto s(t)$ and then M_t is equal to $B_{s(t)}$.

We note that when $(M_t)_{t\geq 0}$ is a continuous martingale with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$ then for all $j \geq 0$ and $u \geq 0$, the process $(M_{ju}, j \geq 0)$ is a discrete martingale with respect to the discrete filtration $(\mathcal{F}_{ju})_{j\geq 0}$. This, together with the continuity of $t \mapsto M_t$ allows to deduce some results for continuous martingales from the theory of discrete martingales.

For instance, we have already mentioned Doob's L^2 inequality, for discrete martingales that implies (when applies to the previous discrete martingale for $u = t2^{-n}$) that if $(M_t)_{t\geq 0}$ is a continuous martingale and $M_t \in L^2$ for some given t, then

$$\mathbb{E}((\max_{j\leq 2^n} M_{jt2^{-n}})^2) \leq 4\mathbb{E}(M_t^2).$$

On the other hand the almost sure continuity of the paths shows that almost surely,

$$(\max_{j\leq 2^n} M_{jt2^{-n}})^2 \overset{n\to\infty}{\longrightarrow} (\max_{s\in[0,t]} M_s)^2$$

so that by monotone convergence we get:

Proposition 5.1 (Doob's L^2 inequality for continuous martingales). *If $(M_t)_{t\geq 0}$ is a continuous martingale and $M_t \in L^2$ for some given t, then we have*

$$\mathbb{E}((\sup_{s\leq t} M_s)^2) \leq 4\mathbb{E}(M_t^2).$$

In the same way, one obtains the L^p inequalities or Doob's maximal inequality for continuous martingales as well

Proposition 5.2 (Doob's L^p inequality for continuous martingales). *If $(M_t)_{t\geq 0}$ is a continuous martingale and $M_t \in L^p$ for some given $t \geq 0$ and $p > 1$, then*

$$\mathbb{E}((\sup_{s\leq t} M_s)^p) \leq (p/(p-1))^p \mathbb{E}(|M_t|^p).$$

Proposition 5.3 (Doob's maximal inequality for continuous martingales). *If $(M_t)_{t\geq 0}$ is a continuous martingale, then for all $\lambda > 0$*

$$\mathbb{P}(\sup_{[0,t]} |M_s| > \lambda) \leq \frac{1}{\lambda}\mathbb{E}(|M_t| 1_{\sup_{[0,t]} |M_s| > \lambda}) \leq \frac{1}{\lambda}\mathbb{E}(|M_t|).$$

5.1.2 Convergence theorems

The convergence results for continuous martingales are also basically the same as those for discrete martingales. The story is again that if one knows the proof in the discrete case well, then said proof can be extended to the continuous case (using the continuity of the paths).

Recall that we say that a martingale $(M_t)_{t\geq 0}$ is:

- in L^p for $p > 1$, if for all $t \geq 0$, $M_t \in L^p$.
- Bounded in L^1 if $\sup_{t\geq 0} \mathbb{E}(|M_t|) < \infty$.
- Bounded in L^p (for a given $p \geq 1$) if $\sup_{t\geq 0} \mathbb{E}(|M_t|^p) < \infty$
- Uniformly integrable (UI) if $\lim\limits_{A\to\infty} (\sup\limits_{t\geq 0} \mathbb{E}(|M_t| 1_{|M_t|\geq A})) = 0$.

Remark 5.2. Recall that a sequence of random variables $(X_n)_{n\in\mathbb{N}}$ converges in L^1 if and only if it converges in probability and is uniformly integrable. Also, we have the following chain of implications:

$$\text{Bounded in } L^p \text{ (for } p > 1) \implies \text{Uniformly integrable} \implies \text{Bounded in } L^1.$$

Remark 5.3. Let us consider a family $(X_i)_{i\in I}$ of L^1 random variables. Recall that any single one of the following three criteria implies that $(X_i)_{i\in I}$ is uniformly integrable:

- There exists $p > 1$ and $C > 0$ such that for all $i \in I$, $\mathbb{E}(|X_i|^p) \leq C$.
- There exists a RV X in L^1 such that for all $i \in I$, $|X_i| \leq |X|$ a.s.
- There exists a random variable X and a collection of σ-fields $\mathcal{G}_i$ such that for all $i \in I$, we have $X_i = \mathbb{E}(X \mid \mathcal{G}_i)$ almost surely.

Proposition 5.4 (Convergence criteria for continuous martingales)**.** *Let $(M_t)_{t\geq 0}$ be a continuous martingale, then we have*

- *If $(M_t)_{t\geq 0}$ is bounded in L^1, then M_t converges almost surely as $t \to \infty$ to some random variable $M_\infty \in L^1(\mathcal{F}_\infty)$.*
- *If $(M_t)_{t\geq 0}$ is UI, then M_t converges almost surely and in L^1 to some random variable $M_\infty \in L^1(\mathcal{F}_\infty)$. ($\rightsquigarrow \mathbb{E}(M_\infty) = \lim_{t\to\infty} \mathbb{E}(M_t) = \mathbb{E}(M_0)$).*
- *If $(M_t)_{t\geq 0}$ is bounded in L^p for some given $p > 1$, then M_t converges almost surely and in L^p to some random variable $M_\infty \in L^p(\mathcal{F}_\infty)$.*

5.1.3 Optional stopping theorems

One can also define, as for discrete filtrations, the notion of stopping times. It is then our goal in this section to find conditions on T (finite) stopping time, such that we have $\mathbb{E}(M_T) = \mathbb{E}(M_0)$ for a continuous martingale $(M_t)_{t\geq 0}$.

Definition 5.3. *The random variable T with values in $[0,\infty]$ is said to be a stopping time for the filtration $(\mathcal{F}_t)_{t\geq 0}$ if for any positive t, the event $\{T \leq t\}$ is in $\mathcal{F}_t$.*

Remark 5.4. M_T is well-defined, because $t \mapsto M_t$ is continuous.

Let us start with the following easiest version of the optional stopping theorem, together with a useful characterization of continuous martingales:

Lemma 5.1 (Optional stopping, bounded stopping times)**.** *Suppose that the continuous process $(M_t)_{t\geq 0}$ is an adapted L^1 process in some filtered probability space. Then the following two statements are equivalent:*

1. *The process $(M_t)_{t\geq 0}$ is a (continuous) martingale.*
2. *For any bounded stopping time T, one has $\mathbb{E}(M_T) = \mathbb{E}(M_0)$.*

Corollary 5.1. *If $(M_t)_{t\geq 0}$ is a continuous martingale, and S is a stopping time (w.r.t. the same filtration), then the stopped process $M^S := (M_t^S := M_{t\wedge S})_{t\geq 0}$ is also a (continuous) martingale (w.r.t. the same filtration).*

Proof. We first note that for any fixed $t_0 \geq 0$, $\min(t_0, S)$ is a bounded stopping time, and thus $M_{t_0}^S$ is in L^1 by the previous Lemma. Let now T be any bounded stopping time, then $\min(S,T)$ is also a bounded stopping time, which implies that (also by the previous Lemma) $\mathbb{E}(_{S\wedge T}) = \mathbb{E}(M_0)$ which can be rewritten as $\mathbb{E}(M_T^S) = \mathbb{E}(M_0^S)$, from which we conclude (using that 2. implies 1. in the previous Lemma) that M^S is a martingale. □

We can now apply these ideas to obtain a workable version of the optional stopping theorem. Given a continuous martingale $(M_t)_{t\geq 0}$ and a finite stopping time T, then we know that M^T is again a continuous martingale, in particular we have $M_t^T = M_{t\wedge T} \to M_T$ almost surely as $t \to \infty$. So if the stopped martingale $(M_t^T)_{t\geq 0}$ is UI, then $M_t^T \to M_T$ also in L^1 as $t \to \infty$ and therefore

$$\mathbb{E}(M_T) = \mathbb{E}(M_\infty^T) = \lim_{t\to\infty}(M_t^T) = \lim_{t\to\infty}\mathbb{E}(M_0^T) = \mathbb{E}(M_0).$$

We have thus shown:

Theorem 5.1 (Optional stopping theorem, workable version)**.** *If $(M_t)_{t\geq 0}$ is a continuous martingale and if T is a finite stopping time such that the stopped martingale M^T is uniformly integrable, then $\mathbb{E}(M_T) = \mathbb{E}(M_0)$.*

Example 5.1 (**Warning:** Important counterexample!). *Here we shall talk about the enemy in optional stopping theorems. Just as in the discrete martingale case, in order to apply the optional stopping theorem, one has to* ***carefully*** *check that the stopped martingale is* ***uniformly integrable*** *(for example it suffices to check for a $p > 1$ that the martingale is bounded in L^p).*

Let us take for example the continuous martingale $(B_t)_{t\geq 0}$ a 1-dimensional Brownian motion associated with its natural filtration. Let $T := \inf\{t > 0 : B_t = -1\}$, then T is a well-defined and finite stopping time, moreover we have $B_T = -1$ almost surely, but

$$\mathbb{E}(B_T) = -1 \neq \mathbb{E}(B_0) = 0.$$

In particular we see that the stopped martingale B^T is not uniformly integrable and thus the optional stopping time cannot be applied.

5.1.4 Application of the optional stopping theorem

The simple (workable version) of the optional stopping theorem provides us already with non-trivial results for the case of the two most natural Brownian martingales:

Lemma 5.2. *Consider a one-dimensional Brownian motion $(B_t)_{t\geq 0}$ and suppose for instance that $(\mathcal{F}_t)_{t\geq 0}$ is the Brownian filtration.*

- ***The quadratic martingale****: The process $(M_t := B_t^2 - t)_{t\geq 0}$ is a martingale with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$.*
- ***The exponential martingales****: For any $\lambda \in \mathbb{R}$, the processes $\mathcal{E}_\lambda(t) := M_t^\lambda := \exp(\lambda B_t - \lambda^2 t/2)$ are martingales with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$.*

Proof. The proofs are immediate consequences of the fact that B has stationary independent increments, in other words we make use of the weak Markov property, namely that $B_{t+h} - B_t$ is independent of $\mathcal{F}_t$. Let $t, h \geq 0$:

$$\begin{aligned}
\mathbb{E}(B_{t+h}^2 \mid \mathcal{F}_t) &= \mathbb{E}((B_{t+h} - B_t + B_t)^2 \mid \mathcal{F}_t) \\
&= \mathbb{E}(B_t^2 + (B_{t+h} - B_t)^2 + 2B_t(B_{t+h} - B_t) \mid \mathcal{F}_t) \\
&= B_t^2 + 2B_t\mathbb{E}((B_{t+h} - B_t) \mid \mathcal{F}_t) + \mathbb{E}((B_{t+h} - B_t)^2 \mid \mathcal{F}_t) \\
&= B_t^2 + 2B_t\mathbb{E}(B_{t+h} - B_t) + \mathbb{E}((B_{t+h} - B_t)^2) \\
&= B_t^2 + 0 + h = B_t^2 + h
\end{aligned}$$

Which proves the claim for the quadratic martingale.

For the exponential martingale, one notes that almost surely

$$\begin{aligned}\mathbb{E}(\exp(\lambda B_{t+h}) \mid \mathcal{F}_t) &= \mathbb{E}(\exp(\lambda B_t)\exp(\lambda(B_{t+h}-B_t)) \mid \mathcal{F}_t)\\ &= \exp(\lambda B_t)\mathbb{E}(\exp(\lambda(B_{t+h}-B_t)) \mid \mathcal{F}_t)\\ &= \exp(\lambda B_t)\mathbb{E}(\exp(\lambda(B_{t+h}-B_t)))\\ &= \exp(\lambda B_t)\exp(\lambda^2 h/2).\end{aligned}$$

which concludes the proof of the lemma. □

We now study some simple examples:

Example 5.2. *Let $\tau = \inf\{t > 0 : B_t \notin (-1,1)\}$ be the exit time from $[-1,1]$. The quadratic martingale $M_t = B_t^2 - t$ stopped at τ is uniformly integrable because for all $t \geq 0$ we have*

$$|M_t^\tau| = |B_{t\wedge\tau}^2 - (t\wedge\tau)| \leq |B_{t\wedge\tau}^2| + \tau \leq 1 + \tau.$$

But we know that $\mathbb{E}(\tau^2) < \infty$, so that the stopped martingale is bounded in L^2 and we know that this implies that it is uniformly integrable. Hence, we can apply the optional stopping theorem to obtain

$$\mathbb{E}(M_\tau) = \mathbb{E}(M_0) = 0 \iff \mathbb{E}(\tau) = \mathbb{E}(B_\tau^2) = 1.$$

Example 5.3. *Suppose now that $a < 0 < b$ and that $T := T_{a,b} := \inf\{t > 0 : B_t \notin (a,b)\}$ is the first time at which the Brownian motion B exits the interval (a,b). If we first work with the martingale $(M_t = B_t)_{t\geq 0}$ then the stopped martingale is obviously bounded because*

$$|B_t^T| \leq |a| + |b|$$

and since bounded implies in particular uniformly integrable we can again use the optional stopping theorem to obtain just like for random walks

$$\begin{aligned}&\mathbb{E}(B_T) = 0 = a\mathbb{P}(B_T = a) + b(1 - \mathbb{P}(B_T = a))\\ &\implies \mathbb{P}(B_T = a) = \frac{b}{b-a},\ \mathbb{P}(B_T = b) = \frac{-a}{b-a}.\end{aligned}$$

Similarly, if we consider the quadratic martingale $M_t = B_t^2 - t$, then the stopped martingale M^T is bounded in L^2 because

$$|M_t| \leq |a|^2 + |b|^2 + T$$

and again T is in L^2, thus M^T is UI and we can apply the optional stopping theorem to obtain

$$\mathbb{E}(T) = a^2\mathbb{P}(B_T = a) + b^2\mathbb{P}(B_T = b) = |ab|.$$

Example 5.4. *Let us now consider a final example in which we make use of the exponential martingale. By considering the sum* $\mathcal{E}_\lambda(t) + \mathcal{E}_{-\lambda}(t)$ *and the difference* $\mathcal{E}_\lambda(t) - \mathcal{E}_{-\lambda}(t)$*, we get that*

$$\tilde{\mathcal{E}}_\lambda(t) := e^{-\lambda^2 t/2} \cosh(\lambda B_t) \text{ and } \hat{\mathcal{E}}_\lambda(t) := e^{-\lambda^2 t/2} \sinh(\lambda B_t)$$

are also continuous martingales. Let τ *be again the exit time of a Brownian motion of the interval* $[-1, 1]$*. If we consider the stopped martingale* $\tilde{\mathcal{E}}_\lambda^\tau$ *we easily see that*

$$|\tilde{\mathcal{E}}_\lambda^\tau| \leq \cosh(\lambda),$$

so the stopped martingale is bounded and especially uniformly integrable. The optional stopping theorem then shows that

$$\begin{aligned}
\mathbb{E}(\tilde{\mathcal{E}}_\lambda(\tau)) &= \mathbb{E}(\tilde{\mathcal{E}}_\lambda(0)) = \cosh(0) = 1 \\
\iff \mathbb{E}(e^{-\lambda^2 \tau/2} \cosh(\lambda B_\tau)) &= \mathbb{E}(e^{-\lambda^2 \tau/2} \cosh(\lambda)) = 1 \\
&\implies \mathbb{E}(e^{-\lambda^2 \tau/2}) = 1/\cosh(\lambda).
\end{aligned}$$

If we substitute $x = \frac{\lambda^2}{2} \geq 0$ *in the above, we see that for all* $x \geq 0$ *we get*

$$\mathbb{E}(e^{-x\tau}) = 1/\cosh(\sqrt{2x}).$$

This gives the formula for the Laplace transform of the positive random variable τ*, which we know does characterize the law of* τ*. So, we see that the collection of formulas given by the optional stopping theorem for the collection of exponential martingales at time* τ *does provide a full analytical description of the law of* τ*.*

6 Quadratic variation

6.1 Quadratic variation of continuous martingales

6.1.1 Warm-up via embedded random walks

Let us revisit what we did in the proof of Donsker's theorem: We considered a Brownian motion started from 0 and we defined for all $\epsilon > 0$ iteratively the stopping times

$$T_0^\epsilon := 0,\ T_{n+1}^\epsilon = \inf\{t > T_n^\epsilon : |B_t - B_{T_n^\epsilon}| = \epsilon\}.$$

The walk given by $Z_n := B_{T_n^\epsilon}$ is then a fair random walk that moves up or down by $\pm\epsilon$ at each step. If we define $Z^p := (Z_n^p)_{n\geq 0}$ to be the walk corresponding to the specialized $\epsilon_p = 2^{-p}$, we then note that we can recover the walk Z^p from the knowledge of Z^{p+1}, we say that the sequence of walks Z^p is nested.

To summarize our ideas above: Given a Brownian motion, then the Brownian motion defines a nested family of walks Z^p (for $p \geq 0$) of random walks on finer and finer meshes.

One feature that we established during the course of the proof of Donsker's theorem is that the knowledge of all the nested walks Z^p allows to recover B. Indeed, we have seen that for each given $t \geq 0$, we have that

$$T^\epsilon_{[t\epsilon^{-2}]} \xrightarrow{\mathbb{P}} t,\ \text{as } \epsilon \to 0.$$

In particular, by Donsker's theorem, this implies that

$$B_t = \overset{\mathbb{P}}{\lim_{p\to\infty}} Z^p_{[t\epsilon_p^{-2}]} = \overset{\mathbb{P}}{\lim_{p\to\infty}} Z^p_{[t4^p]},$$

which shows that one can recover all the trajectory and time-parametrization of B from the knowledge of the nested family of random walks Z^p.

We will now apply the same idea to continuous martingales. Suppose that $(M_t)_{t\geq 0}$ is a continuous martingale with respect to some filtration $(\mathcal{F}_t)_{t\geq 0}$ with $M_0 = 0$ almost surely, let us also assume (for reasons of simplicity) that $\limsup_{t\to\infty} M_t = \infty$.

We can then define in exactly the same fashion as we defined T_n^ϵ for B the sequence of stopping times $S_0^\epsilon = 0$ and for all $n \geq 1$:

$$S_n^\epsilon := \inf\{t > S_{n-1}^\epsilon : |M_t - M_{S_{n-1}^\epsilon}| = \epsilon\}.$$

The condition about the limsup of M ensures that all these stopping times are finite, and the continuity of M ensures also that for each $\epsilon > 0$ we have $S_n^\epsilon \to \infty$ as $n \to \infty$ (as otherwise, $\limsup_{n\to\infty} M_{S_n^\epsilon} = \infty$).

Let us look at the case $n = 1$, then the stopped martingale $M^{S_1^\epsilon}$ is bounded (it's bounded by ϵ) and thus in particular uniformly integrable, by the optional stopping theorem we obtain

$$\mathbb{E}(M_{S_1^\epsilon}) = \mathbb{E}(M_0) = 0 \implies \mathbb{P}(M_{S_1^\epsilon} = \epsilon) = \mathbb{P}(M_{S_1^\epsilon} = -\epsilon) = \frac{1}{2}.$$

But also by the optional stopping time (more elaborate version) we get that

$$\begin{aligned}\mathbb{E}(M_{S_{n+1}^\epsilon} \mid \mathcal{F}_{S_n^\epsilon}) = M_{S_n^\epsilon} &\iff \mathbb{E}(M_{S_{n+1}^\epsilon} - M_{S_n^\epsilon} \mid \mathcal{F}_{S_n^\epsilon}) = 0 \\ &\implies \mathbb{P}(M_{S_{n+1}^\epsilon} - M_{S_n^\epsilon} = \pm\epsilon \mid \mathcal{F}_{S_n^\epsilon}) = \frac{1}{2}\end{aligned}$$

This shows that the walk $Y_n := M_{S_n^\epsilon}$ is a fair random walk that moves up or down $\pm\epsilon$ at each step. So, if we define Y^p to be the walk corresponding to $\epsilon_p = 2^{-p}$, we get that the walks Y^p for $p \geq 1$ are nested exactly as the walks Z^p were. In other words, the nested sequence of random walks $(Y_n^p, \geq 0)_{p\geq 1}$ and $(Z_n^p, n \geq 0)_{p\geq 1}$ have the same law.

In particular, this shows that starting from the trajectory of a continuous martingale M, one can first define the nested family Y^p, then define Z^p to be equal to Y^p and then one can construct a Brownian motion B from this nested family Z^p using the procedure we have described on the previous page (using Donsker's invariance theorem). In other words, starting from a continuous martingale M, one can define a Brownian motion B which is some deterministic function of M.

Recap: *Starting with a continuous martingale M, one can define a Brownian motion B which is some deterministic function of M.*

Lets make this more concrete: for each time t, we can define $N_\epsilon(t)$ to be the largest integer $n \in \mathbb{N}$ such that $S_n^\epsilon \leq t$. In other words, $N_\epsilon(t)$ is the number of ϵ-steps performed by M up to time t, in particular we have

$$|M_t - Y_{N_\epsilon(t)}| \leq \epsilon, \text{ (where } Y_n = M_{S_n^\epsilon})$$

We therefore conclude that (see picture in proof of Donsker's theorem), as $\epsilon \to 0$, $\epsilon^2 N_\epsilon(t)$ converges in probability to the time A_t such that $B_{A_t} = M_t$. It suggests therefore that if one defines

$$A_t := \overset{\mathbb{P}}{\lim_{\epsilon \to 0}} \epsilon^2 N_\epsilon(t),$$

then it will be possible to define a continuous modification of the process $(A_t)_{t \geq 0}$ such that for all $t \geq 0$ we have

$$M_t = B_{A_t}. \tag{*}$$

This continuous process $(A_t)_{t\geq 0}$ is what we will call the **quadratic variation** of the martingale M. One often refers to (*), as the fact that M is a time-changed Brownian motion.

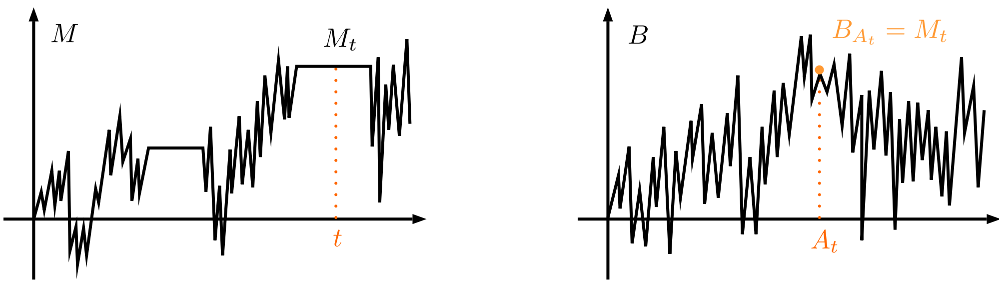


Figure 23: Notice that as $\epsilon \to 0$, according to the definition of $N_\epsilon(t)$ we count "0"-steps. The picture then displays the time change of a Brownian motion.

In the next section, we will in fact construct this quadratic variation by other means, but then check that indeed it corresponds to the process $A_t = \lim_{\epsilon\to 0} \epsilon^2 N_\epsilon(t)$.

6.1.2 Quadratic variation of bounded martingales

Idea: We know that $B_t^2 - t$ is a continuous martingale, this suggests that the time-changed process

$$(B_{A_t}^2 - A_t)_{t\geq 0} = (M_t^2 - A_t)_{t\geq 0}$$

will be a (sort of) martingale (or rather what we will call a local martingale). We will in fact show that this feature, that $A = (A_t)_{t\geq 0}$ is a continuous non-decreasing process such that $((M_t)^2 - A_t)_{t\geq 0}$ is a local martingale does characterize A uniquely and provides a neat way to construct it.

Let us now suppose that $(M_t)_{t\geq 0}$ is a continuous martingale that is started from 0 such that:

- $|M|$ is almost surely bounded by some deterministic constant C (i.e., almost surely, for all $t \geq 0$, $|M_t| \leq C$).
- For some deterministic integer K, M is almost surely constant on $[K, \infty)$ (i.e., almost surely, for all $t \geq K$, $M_t = M_K$).

These two conditions look at first pretty restrictive, but we shall see that it will be easy to deduce the general results for continuous martingales from the results that we derive with these two restrictions. In other words, everything happens here, all the core arguments of the construction of the quadratic variation of continuous martingales. We also already mention here that the quadratic variation will be at the core of our definitions of stochastic integrals.

The key result (under the *'restrictions'* above) of this section will be the following:

Proposition 6.1. *There exists a unique continuous non-decreasing process $(A_t)_{t\geq 0}$ with $A_0 = 0$ (A_t is $\mathcal{F}_t$-measurable and in L^2) such that $(M_t^2 - A_t)_{t\geq 0}$ is an L^2-martingale.*

Definition 6.1. *This unique process $(A_t)_{t\geq 0}$ will be called the quadratic variation of the martingale M.*

The next definitions are crucial:

Definition 6.2. *When $t \geq 0$, we say that Δ_n is a nested sequence of subdivisions of $[0,t]$ such that $|\Delta_n| \to 0$ as $n \to \infty$ if:*

- *Each Δ_n consists of a finite family $(t_i^n)_{i \leq m_n}$ such that $0 = t_0^n < t_1^n < \cdots < t_{m_n}^n = t$ (i.e. it is a subdivision of the interval $[0,t]$).*
- *The mesh-size $|\Delta_n| := \max_{0 \leq i \leq m_n}(t_{i+1}^n - t_i^n)$ tends to 0 as $n \to \infty$.*
- *The sequence $(\Delta_n)_{n \geq 0}$ of subdivisions is nested, meaning that for each $n \leq n'$ and each $i \leq m_n$, there exists $i' \leq m_{n'}$ such that $t_i^n = t_{i'}^{n'}$.*

Definition 6.3. *When Δ_n is a subdivision of $[0,t]$, we then define*

$$V_{\Delta_n} := \sum_{i=0}^{m_n-1} (M_{t_{i+1}^n} - M_{t_i^n})^2$$

to be the sum of the squares of the increments of M along the subdivision Δ_n. We will refer to V_{Δ_n} as the ***discrete quadratic variation*** *of M along the subdivision Δ_n.*

Proposition 6.2. *If Δ_n is a nested sequence of subdivisions of $[0,t]$ with $|\Delta_n| \to 0$ (mesh sizes tends to 0), then the sequence V_{Δ_n} converges in L^2 to A_t.*

We will prove both propositions at once:

Proof. **Uniqueness:** Let us suppose that there were two such process A and A'. Note that A and A' are then necessarily in L^2, since they are the difference between two processes that are in L^2 (namely the bounded process $(M_t^2)_{t \geq 0}$ and the L^2 martingales given by $M^2 - A$ respectively $M^2 - A'$). Let us define $(A''_t := A_t - A'_t)_{t \geq 0}$ which is an L^2 martingale.

Let us consider a nested sequence Δ_n of subdivisions on $[0,K]$ with $|\Delta_n| \to 0$.

$$\begin{aligned}
\mathbb{E}((A''_K)^2) &= \sum_{i=0}^{m_n-1} \mathbb{E}((A''_{t_{i+1}^n} - A''_{t_i^n})^2) \\
&\leq \mathbb{E}(\max_{i \leq m_n-1} |A''_{t_{i+1}^n} - A''_{t_i^n}| \times \sum_{i=0}^{m_n-1} |A''_{t_{i+1}^n} - A''_{t_i^n}|) \\
&\overset{\Delta}{\leq} \mathbb{E}(\max_{i \leq m_n-1} |A''_{t_{i+1}^n} - A''_{t_i^n}| \times \sum_{i=0}^{m_n-1} (|A_{t_{i+1}^n} - A_{t_i^n}| + |A'_{t_{i+1}^n} - A'_{t_i^n}|)) \\
&\leq \mathbb{E}(\max_{i \leq m_n-1} |A''_{t_{i+1}^n} - A''_{t_i^n}| \times (A_K + A'_K)).
\end{aligned}$$

Where in the last inequality we used the fact that A and A' are non-decreasing.

But by the continuity of A and A' we have that almost surely

$$\max_{i \leq m_n - 1} |A''_{t^n_{i+1}} - A''_{t^n_i}| \to 0 \text{ as } n \to \infty.$$

On the other hand,

$$\begin{aligned} &\max_{i \leq m_n - 1} |A''_{t^n_{i+1}} - A''_{t^n_i}| \times (A_K + A'_K) \\ \overset{\Delta}{\leq} &\max_{i \leq m_n - 1} (|A_{t^n_{i+1}} - A_{t^n_i}| + |A'_{t^n_{i+1}} - A'_{t^n_i}|) \times (A_K + A'_K) \leq (A_K + A'_K)^2 \end{aligned}$$

where we used again that A and A' are non-decreasing, since we also know that $A_t + A'_t$ is in L^2, we can apply dominated convergence for $n \to \infty$ to obtain that

$$\mathbb{E}((A''_K)^2) = 0.$$

Since $(A''_t)_{t \geq 0}$ is an L^2-martingale (as the difference of L^2 martingales), we can apply Doob's L^2 inequality to establish

$$\mathbb{E}(\sup_{t \leq K} (A''_t)^2) \leq 4\mathbb{E}((A''_K)^2) = 0$$

from which we conclude that almost surely $A''_t = 0$ for all $t \leq K$, and thus $A = A'$ as processes.

Existence: *Only idea of the proof.*

Instead of trying to approximate the process $(A_t)_{t \geq 0}$, we first try to find the right approximation of the martingale $2X_t := M_t^2 - A_t$. We take a nested sequence Δ_n of subdivisions of $[0, K]$ such that the mesh-size $|\Delta_n|$ tends to 0 as $n \to \infty$ and then for each $n \in \mathbb{N}$ we define

$$X^n_t := \sum_{i=0}^{m_n - 1} M_{t^n_i} (M_{t \wedge t^n_{i+1}} - M_{t \wedge t^n_i}),$$

we notice that this is a bounded continuous martingale such that on each interval $[t^n_i, t^n_{i+1}]$, the increments of this martingale are exactly $M_{t^n_i}$ times the increments of M. In other words, for all $t \in [t^n_i, t^n_{i+1}]$ we have

$$X^n_t - X^n_{t^n_i} = M_{t^n_i} \times (M_t - M_{t^n_i}).$$

The idea is that $2X_t^n$ is a martingale that will approximate $M_t^2 - A_t$ well. Indeed we have

$$V_{\Delta_n} = \sum_{i=0}^{m_n-1} (M_{t_{i+1}^n} - M_{t_i^n})^2 \text{ and } 2X_K^n = \sum_{i=0}^{m_n-1} 2M_{t_i^n}(M_{t_{i+1}^n} - M_{t_i^n})$$

from which we can conclude that

$$\begin{aligned} V_{\Delta_n} + 2X_K^n &= \sum_{i=0}^{m_n-1} (M_{t_{i+1}^n} - M_{t_i^n} + 2M_{t_i^n})(M_{t_{i+1}^n} - M_{t_i^n}) \\ &= \sum_{i=0}^{m_n-1} (M_{t_{i+1}^n}^2 - M_{t_i^n}^2) = M_K^2 - M_0^2 = M_K^2 \end{aligned}$$

So, if we can show that X_K^n converges in L^2 to some X_K as $n \to \infty$, we conclude that V_{Δ_n} does converge in L^2 to $M_K^2 - 2X_K$.

We now briefly outline the rest of the proof:

- **Step 1: Lemma:** X_K^n converges in L^2 to some random variable X_K as $n \to \infty$. Establishing this Lemma is quite technical and will be skipped.
- **Step 2:** There exists a sequence $n_k \to \infty$, such that almost surely X^{n_k} converges uniformly on $[0, K]$ to a continuous function $(X_t)_{t \le K}$ ($X_t = X_K$ for $t \ge K$) and we check that $(X_t)_{t \ge 0}$ is an L^2-martingale.
- **Step 3:** Finally we check that $A_t := M_t^2 - 2X_t$ is a continuous, adapted L^2 process which is non-decreasing. (The non-decreasing part is again a bit technical).

□

Remark 6.1. Let us comment on the martingale $(M_t^2 - A_t)_{t \ge 0}$ in our construction. What we have seen is that when $M_0 = 0$ almost surely, then for all K, if one chooses a nested sequence $(\Delta_n)_{n \ge 0}$ of subdivisions of $[0, K]$ with $|\Delta_n| \to 0$, then for all $t \le K$, $M_t^2 - A_t$ is "twice" the limit in L^2 of

$$\sum_{i=0}^{m_n-1} M_{t_i^n}(M_{t \wedge t_{i+1}^n} - M_{t \wedge t_i^n}).$$

This limiting martingale is what we will call the stochastic integral $\int_0^t M_s dM_s$ when we will define stochastic integrals. So we can write

$$(M_t)^2 = 2\int_0^t M_s dM_s + A_t.$$

6.1.3 Quadratic variation of general L^2 martingales

In the previous section, we have constructed and characterized the quadratic variation of a martingale $(M_t)_{t\geq 0}$ under the assumption that $|M|$ was bounded by some deterministic constant C and that it was almost surely constant after some deterministic time K. As we will explain now, neither assumption is actually needed to show the existence and uniqueness of the quadratic variation of a martingale. We remark that the condition that M is in L^2 is actually needed in order to ensure that the process $(M_t^2 - A_t)_{t\geq 0}$ is in L^1, and can therefore be a martingale.

Theorem 6.1 (Quadratic variation for L^2 martingales)**.** *There exists a unique adapted continuous non-decreasing process $(A_t)_{t\geq 0}$ with $A_0 = 0$ such that $(M_t^2 - A_t)_{t\geq 0}$ is a martingale. Furthermore, for all $t \geq 0$, and for all choice of nested sequences Δ_n of subdivisions of $[0,t]$ with $|\Delta_n| \to 0$, V_{Δ_n} converges in probability to A_t as $n \to \infty$.*

6.1.4 Quadratic variation of local martingales

Motivation: Assume we are given a Brownian motion $(B_t)_{t\geq 0}$ started from 0 and a random variable Z that is independent of B. We then can define

$$M_t := Z + B_t,$$

i.e. a *"Brownian motion started from a random point Z"*. We would like to say that M is like a martingale and that its quadratic variation at time t is given by t. However, if $Z \notin L^1$, we cannot say this.

Definition 6.4. *A continuous adapted process $(M_t)_{t\geq 0}$ is said to be a local martingale started from 0 (in some filtered probability space) if:*

- *$M_0 = 0$ almost surely.*
- *There exists a sequence of non-decreasing stopping times such that $T_n \to \infty$ almost surely and $(M_t^{T_n})_{t\geq 0}$ is a continuous martingale started from 0.*

Equivalent to this is the following definition:

Definition 6.5. *A continuous adapted process $(M_t)_{t\geq 0}$ is said to be a local martingale if M_0 is $\mathcal{F}_0$-measurable and there exists a sequence of non-decreasing stopping times such that $T_n \to \infty$ almost surely and $(M_t^{T_n} - M_0)_{t\geq 0}$ is a continuous martingale.*

Remark 6.2. According to our definition, local martingales are always continuous.

Lemma 6.1. *$(M_t)_{t\geq 0}$ is a local martingale started from 0 if and only if:*

- *$M_0 = 0$ almost surely.*
- *M is almost surely continuous.*
- *If we define $\sigma_n = \inf\{t > 0 : |M_t| = n\}$ for all $n \in \mathbb{N}$, then M^{σ_n} is a martingale.*

Proof. Easy exercise (from exercise sheets). $\square$

Lemma 6.2. *If a local martingale started from 0 has finite variation (meaning that it can be written as the difference between two adapted continuous non-decreasing processes), then it is almost surely equal to 0 for all times.*

Proof. Very similar to the uniqueness proof we have shown in the previous Proposition. $\square$

It's quite evident from the definition of local martingales, that every continuous martingale (started from 0) is a local martingale (because we can take any sequence of increasing stopping times that diverges to ∞ and already know that the stopped martingale is a continuous martingale). We ask us when is a local martingale a (true) martingale?

Lemma 6.3. *Every bounded local martingale is a continuous martingale.*

Proof. Let $(M_t)_{t\geq 0}$ be a local martingale that is bounded by some constant C, i.e. we have $|M_t| \leq C$ for all $t \geq 0$. Let $(\tau_n)_{n\in\mathbb{N}}$ be a localizing sequence for M, i.e. it is a non-decreasing sequence of stopping times such that $(M_t^{\tau_n})_{t\geq 0}$ is a martingale for any $n \in \mathbb{N}$ and $\tau_n \nearrow \infty$ almost surely. Fix $s \leq t$, by the martingale property we have

$$\mathbb{E}(M_{t\wedge\tau_n} \mid \mathcal{F}_s) = M_{s\wedge\tau_n}.$$

By dominated convergence theorem, which we can apply by the uniform boundedness of M, we get that (since $\tau_n \nearrow \infty$)

$$\mathbb{E}(M_t \mid \mathcal{F}_s) = \lim_{n\to\infty} \mathbb{E}(M_{t\wedge\tau_n} \mid \mathcal{F}_s) = \lim_{n\to\infty} M_{s\wedge\tau_n} = M_s$$

which shows that indeed M is a (true) martingale. $\square$

Remark 6.3. The proof of the Lemma above suggests that it is sufficient to check that $(M_t^{\tau_n}) \to M_t$ in L^1 in order to conclude that a local martingale is a martingale. Notice that the previous limit is indeed taken with respect to n as $n \to \infty$.

We can now state the main results about the quadratic variation of local martingales.

Theorem 6.2 (Quadratic variation of local martingales)**.** *If $(M_t)_{t\geq 0}$ is a local martingale started from 0, then there exists a unique adapted continuous non-decreasing process $(A_t)_{t\geq 0}$ with $A_0 = 0$ almost surely and such that $(M_t^2 - A_t)_{t\geq 0}$ is a local martingale. Furthermore, for all $t \geq 0$ and for all choice of nested sequences Δ_n of subdivisions of $[0,t]$ with $|\Delta_n| \to 0$*

$$V_{\Delta_n} = \sum_{i=0}^{m_n-1} (M_{t_{i+1}^n} - M_{t_i^n})^2 \xrightarrow{\mathbb{P}} A_t, \ \text{as } n \to \infty.$$

6.1.5 Crossvariation between local martingales

It is convenient to denote the quadratic variation of a local martingale M by $(\langle M \rangle_t)_{t\geq 0}$. It is natural and useful to also define the *cross-variation* between two local martingales (defined w.r.t. the same filtration).

Definition 6.6 (Crossvariation)**.** *If M and N are local martingales, then one defines the process $(\langle M, N\rangle_t)_{t\geq 0}$ and calls this the cross-variation between M and N, by*

$$\langle M, N\rangle_t := \frac{1}{4}(\langle M+N\rangle_t - \langle M-N\rangle_t).$$

By our previous results we can say that:

- For all $t \geq 0$, and for all nested sequence Δ_n of subdivisions of $[0,t]$ with $|\Delta_n| \to 0$,

$$\sum_{i=0}^{m_n-1} (M_{t_{i+1}^n} - M_{t_i^n})(N_{t_{i+1}^n} - N_{t_i^n}) \xrightarrow{\mathbb{P}} \langle M, N\rangle_t \text{ as } n \to \infty.$$

  - If M and N are in L^2, then this convergence takes place in L^2.

- Furthermore, $\langle M, N\rangle$ is the only continuous process with bounded variation (i.e., that can be written as the difference of two non-decreasing adapted processes) such that $MN - \langle M, N\rangle$ is a local martingale.

Remark 6.4. With this notation, we set $\langle M\rangle_t = \langle M, M\rangle_t$.

7 Stochastic integrals

7.1 Warm-up

We are now ready to define *stochastic integrals*. Before doing this rigorously, it is useful to informally explain what this means. On the one hand, we are given a continuous martingale $(M_t)_{t\geq 0}$ (or a local martingale) in some filtered probability space. For the time being, we can first think of this continuous martingale to be a Brownian motion $(B_t)_{t\geq 0}$. Then, suppose that $(H_t)_{t\geq 0}$ is a continuous real-valued process that is adapted to the same filtration. We will define a new continuous (local) martingale $(I_t)_{t\geq 0}$, denoted by

$$I_t = \int_0^t H_s dB_s, \tag{Notation}$$

which is called a stochastic integral. It can be understood as follows: Locally, at time t, it follows the increments of B but it *"decided"* to multiply those increments by the local factor H_t.

In particular, we see that a good approximation of the value of I_t is obainted by considering a nested sequence of subdivisions Δ_n of $[0,t]$ with $|\Delta_n| \to 0$ and to consider the sum

$$I_n := \sum_{j=0}^{m_n-1} H_{t_j^n}(B_{t_{j+1}^n} - B_{t_j^n}).$$

Indeed, we shall see that I_n does converge in probability to our stochastic integral I_t as $n \to \infty$.

Remark 7.1. As we have already discussed, an example of stochastic integral is provided by the process $B_t^2 - t$ that can be viewed as the stochastic integral $\int_0^t 2B_s dB_s$ and we have seen that indeed for a given t, this quantity $B_t^2 - t$ is the limit (in L^2) of

$$I_n = \sum_{j=0}^{m_n-1} 2B_{t_j^n}(B_{t_{j+1}^n} - B_{t_j^n}).$$

It is very important to notice that in this discrete approximation of I_n, one chooses the magnification factor $H_{t_j^n} = B_{t_j^n}$ *just before* the increment $(B_{t_{j+1}^n} - B_{t_j^n})$ occurs (i.e. we choose $B_{t_j^n}$ instead of $B_{t_{j+1}^n}$). So even if the notation $H_s dB_s$ in a stochastic integral describes things at infinitesimal levels, it still means at that infinitesimal level, one *chooses* H_s (infinitesimally) before the increments dB_s occurs. As opposed to the Lebesgue integral, where "orientation" of the time-axis will play an important role in stochastic integrals.

7.2 The L^2 theory of stochastic integrals with respect to Brownian motion

Let us now first describe how to properly define stochastic integrals with respect to a one-dimensional Brownian motion $B := (B_t)_{t\geq 0}$.

7.2.1 Integral of elementary processes, and strategy of construction

Definition 7.1. *The process $(K_t)_{t\geq 0}$ is said to be an elementary process if it is of the type*

$$K_t = \sum_{j=0}^{m-1} Y_{a_j} 1_{t\in(a_j,a_{j+1}]}, \tag{*}$$

where $m \in \mathbb{N}$, $0 \leq a_0 < \cdots < a_m$ are deterministic times and each Y_{a_j} is an $\mathcal{F}_{a_j}$-measurable random variable in L^2. We denote by $\mathcal{E}_B$ the set of all elementary processes such that K_t is in L^2 (that is exactly the fact if the Y_{a_j} are in L^2).

Remark 7.2. $\mathcal{E}_B$ is a vector space.

Definition 7.2 (Stochastic integral for $\mathcal{E}_B$)**.** *For each elementary process K_t as in the form (*) in $\mathcal{E}_B$, we define the stochastic integral to be the continuous process $I(K) = (I(K)_t)_{t\geq 0}$, such that for all $t \geq 0$*

$$I_t^K := \sum_{j=0}^{m-1} Y_{a_j}(B_{t\wedge a_{j+1}} - B_{t\wedge a_j}).$$

We also denote this by

$$I_t^K = \int_0^t K_s dB_s.$$

Let us give a couple of remarks that follow immediately by these definitions:

- $t \mapsto I_t^K$ is continuous and on each interval $[a_j, a_{j+1}]$ its increments are Y_{a_j} times the increments of B.

- $(I_t^K)_{t\geq 0}$ is indeed a continuous martingale in L^2 (difference of L^2 martingales). Moreover it is constant for all $t \geq a_m$, and that if we denote the constant value $I_{a_m}^K$ by I_∞^K, then (using that BM has independent increments, and $B_{a_{j+1}} - B_{a_j}$ is independent of $\mathcal{F}_{a_j} \ni Y_{a_j}$)

$$\begin{aligned}
\mathbb{E}((I_\infty^K)^2) &= \mathbb{E}\left[\left(\sum_{j=0}^{m-1} Y_{a_j}(B_{a_j+1} - B_{a_j})\right)^2\right] \\
&= \sum_{j=0}^{m-1} \mathbb{E}(Y_{a_j}^2)(a_{j+1} - a_j) = \sum_{k=0}^{m-1} \mathbb{E}(Y_{a_j}^2) \int_0^\infty 1_{t\in(a_j,a_{j+1}]} dt \\
&= \mathbb{E}\left(\int_0^\infty \sum_{j=0}^{m-1} Y_{a_j}^2 1_{t\in(a_j,a_{j+1}]} dt\right) = \mathbb{E}\left(\int_0^\infty K_t^2 dt\right).
\end{aligned}$$

  Which shows that

$$\mathbb{E}\left[\left(\int_0^\infty K_s dB_s\right)^2\right] = \mathbb{E}\left[\int_0^\infty K_s^2 ds\right].$$

  So, the martingale I^K is bounded in L^2 if and only if $K \in \mathcal{E}_B$.

- The quadratic variation of I^K is given by the process

$$t \mapsto \int_0^t K_s^2 ds$$

- Similarly, as in the item described above, simply because the difference between two elementary processes K and K' is again an elementary process, and because I^K is clearly a linear function in K, we see that for all $K, K' \in \mathcal{E}_B$,

$$\mathbb{E}\left[\left(\int_0^\infty (K_s - K_s')dB_s\right)^2\right] = \mathbb{E}[(I(K-K')_\infty)^2] = \mathbb{E}\left[\int_0^\infty (K_s - K_s')^2 ds\right]$$

  Hence, one can measure the distance between $I(K)$ and $I(K')$ in terms of some distance of K and K'. This identity implies also that the crossvariation between $I(K)$ and $I(K')$ is given by

$$\int_0^t K_s K_s' ds.$$

7.2.2 Plan of the construction

Let us briefly outline the strategy of the construction of the stochastic integral $I_t^K = \int_0^t K_s dB_s$ for more general processes K:

1. We have already constructed $\int_0^t K_s dB_s$ when K is an elementary process $K \in \mathcal{E}_B$.
2. We will consider a larger class of processes $(K_t)_{t\geq 0}$. Heuristically speaking will it be a class of processes that can be *"nicely"* approximated by elementary processes. Typically K will be in this class if there exists a sequence K^n of elementary processes such that
$$\mathbb{E}\left(\int_0^\infty (K_s - K_s^n)^2 ds\right) \to 0, \text{ as } n \to \infty.$$
3. When K is in that class, we are going to show that the sequence of martingales $(I_t^{K_n})_{t\geq 0}$ will converge to a limiting process $(I_t)_{t\geq 0}$ that we will define to be $(I_t^K)_{t\geq 0}$ and call it the stochastic integral of K with respect to B.
4. In the end, we will see that I^K is a continuous martingale, that its quadratic variation is given by $\int_0^t K_s^2 ds$ and if K is a continuous process in that class and Δ_n is a nested sequence of subdivisions of $[0,t]$ with mesh-size going to zero, then we have
$$I_t^K = \int_0^t K_s dB_s = \overset{\mathbb{P}}{\lim_{n\to\infty}} \sum_{i=0}^{m_n-1} K_{t_i^n}(B_{t_{i+1}^n} - B_{t_i^n}).$$

7.2.3 The space of continuous martingales that are bounded in L^2

In order to implement this strategy, let us first describe what notion of convergence of sequences of continuous martingales we will use: We consider the space $\mathcal{M}^2$ of continuous martingales that are started from 0 and bounded in L^2 (i.e. $\sup_{t\geq 0}\mathbb{E}(M_t^2) < \infty$). We then endow this vector space with the scalar product $(M,N) := \mathbb{E}(M_\infty N_\infty)$ for all $M, N \in \mathcal{M}^2$.

Remark 7.3. For $M \in \mathcal{M}^2$, we always have $M \xrightarrow{t\to\infty} M_\infty$ almost surely and in L^2 (because M is bounded in L^2). moreover, all of $(M_t)_{t\geq 0}$ can be recovered from M_∞ by
$$M_t = \mathbb{E}(M_\infty \mid \mathcal{F}_t).$$
Thus if $(M,M) = 0$, then $M_\infty = 0$ almost surely, and therefore $\forall t \geq 0$, $M_t = \mathbb{E}(M_\infty \mid \mathcal{F}_t) = 0$ almost surely, which implies that almost surely $\forall t \geq 0$, $M_t = 0$. Hence (M,N) is indeed a scalar product.

Lemma 7.1. *The space $\mathcal{M}^2$ is a Hilbert space.*

Proof. We only need to prove that $\mathcal{M}^2$ is complete. Suppose that $(M^n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{M}^2$, i.e. we have

$$\mathbb{E}((M^n_\infty - M^l_\infty)^2) \to 0, \text{ as } n,l \to \infty. \tag{*}$$

(*) above shows that the sequence $(M^n_\infty)_{n\in\mathbb{N}}$ is Cauchy in L^2 but we know that L^2 is complete, thus there exists $M_\infty \in L^2$ such that $M^n_\infty \to M_\infty$ in L^2 as $n \to \infty$. By properties of conditional expectations this gives

$$M^n_t = \mathbb{E}(M^n_\infty \mid \mathcal{F}_t) \xrightarrow[n\to\infty]{L^2} \mathbb{E}(M_\infty \mid \mathcal{F}_t) =: M_t$$

We have to show that $M := (M_t)_{t\geq 0}$ is in $\mathcal{M}^2$. Certainly M is a martingale (because its written as a closed martingale), that is bounded in L^2, because

$$M_t^2 \leq \mathbb{E}(M^2_\infty \mid \mathcal{F}_t) \implies \mathbb{E}(M_t^2) \leq \mathbb{E}(M^2_\infty) < \infty.$$

We thus need to check if $(M_t)_{t\geq 0}$ is continuous. By Doob's L^2 inequality we get

$$\mathbb{E}[(\sup_{t\geq 0}(M^n_t - M^l_t))^2] \leq 4\mathbb{E}((M^n_\infty - M^l_\infty)^2) \xrightarrow{n,l\to\infty} 0.$$

Hence there exists a deterministic sequence $n_k \to \infty$ such that

$$\mathbb{E}[(\sup_{t\geq 0}(M^{n_k}_t - M^{n_{k+1}}_t))^2] \leq \frac{1}{8^k},$$

and consequently by Markov's inequality

$$\mathbb{P}[(\sup_{t\geq 0}(M^{n_k}_t - M^{n_{k+1}}_t))^2 \geq 4^{-k}] \leq \frac{1}{2^k}.$$

Since the above is summable, we get by the Borel-Cantelli lemma that almost surely, for all but finitely many k's, we have

$$\sup_{t\geq 0}|M^{n_{k+1}}_t - M^{n_k}_t| \leq 2^{-k} = \sqrt{4^{-k}}.$$

This implies that $(M^{n_k}_t)_{t\geq 0}$ converges uniformly on $\mathbb{R}_+$ to some continuous function that we call $t \mapsto \tilde{M}_t$. But, since we already know that $M^n_t \to M_t$ in L^2, we must have for all $t \geq 0$, $M_t = \tilde{M}_t$ almost surely. Hence, we can conclude that $(M_t)_{t\geq 0}$ is a continuous martingale that is bounded in L^2, and we see that M^n converges to M in $\mathcal{M}^2$ as $n \to \infty$ because:

$$(M - M^n, M - M^n) = \mathbb{E}((M_\infty - M^n_\infty)^2) \to 0$$

by the convergence of M^n_∞ to M_∞ in L^2. $\square$

7.2.4 Progressively measurable processes

Let us now look more closely at the class of processes that can be nicely approximated by processes in $\mathcal{E}_B$. Note that we want to look at processes $(K_t)_{t\geq 0}$ such that the integral $\int_0^t K_s^2 ds$ makes sense, so that some measurability with respect to time is required to be sure that this integral is a well-defined random variable.

Definition 7.3. *The process $(K_t)_{t\geq 0}$ is said to be progressively measurable with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$ if there exists a measurable set E of probability 1, such that for all $t \geq 0$, the map $(s,\omega) \mapsto K_s(\omega)$ defined on $[0,t]\times\Omega$ is measurable with respect to the product σ-field $\mathcal{B}_{[0,t]} \otimes \mathcal{F}_t$.*

Remark 7.4. In essence this (cryptic) definition means that at each time t, one can look at once at the whole function $s \mapsto K_s$ defined on $[0,t]$ as a random measurable function on $[0,t]$ and this function is $\mathcal{F}_t$ measurable.

It is easy to see that most adapted processes with some regularity will be progressively measurable as the next Lemma suggests (see exercise sheets):

Lemma 7.2. *An adapted process that has right-continuous paths with probability 1 is progressively measurable. An adapted process that has left-continuous paths with probability 1 is progressively measurable.*

Proof. Exercise sheets. □

Definition 7.4. *The set of progressively measurable processes will be denoted by $\mathcal{P}$. We can then define the subset $\mathcal{P}_B$ of $\mathcal{P}$ consisting of progressively measurable processes such that*

$$\mathbb{E}\left(\int_0^\infty K_s^2 ds\right) < \infty.$$

This set is naturally endowed with the scalar product

$$(K,K')_B := \mathbb{E}\left(\int_0^\infty K_s K_s' ds\right).$$

Remark 7.5. In Exercise 10.1 we will derive that $\mathcal{P}_B$ is the L^2 space on $\Omega\times\mathbb{R}_+$ equipped with the progressive σ-algebra (10.1.a) and product measure $\mathbb{P}\otimes d\lambda$ (where λ denotes the Lebesgue measure), so that in particular $\mathcal{P}_B$ is a Hilbert space.

We recall from Functional Analysis the following important Proposition:

Proposition 7.1. *A subspace M of a Hilbert space H is dense in H if and only if $M^{\perp} = \{0\}$ (i.e. trivial orthogonal complement).*

Lemma 7.3. *The subspace $\mathcal{E}_B$ (of $\mathcal{P}_B$) is dense in the Hilbert space $\mathcal{P}_B$. In other words, for any progressively measurable process in $\mathcal{P}_B$, one can find a sequence of elementary processes K^n in $\mathcal{E}_B$ such that*

$$\lim_{n\to\infty} \mathbb{E}\left(\int_0^\infty (K_s - K_s^n)^2 ds\right) = 0.$$

Proof. It suffices to show that a process K in $\mathcal{P}_B$ that is orthogonal to $\mathcal{E}_B$ is necessarily equal to 0, i.e. for all $\tilde{K} \in \mathcal{E}_B$ we have

$$(K, \tilde{K})_B = 0. \tag{*}$$

Let us first note that by Cauchy-Schwarz inequality, for each $t \geq 0$, the random variable $X_t := \int_0^t K_s ds$ is in L^1, thus we can choose $t \mapsto X_t$ to be continuous almost surely, moreover it then has finite variation because X is the difference of the non-decreasing integral of $K1_{K>0}$ and of $-K1_{K<0}$.

From the orthogonality condition (*) we get that for all $a < b$ and for all $Y_a \in L^2(\mathcal{F}_a)$, the process $\tilde{K}_t := (Y_a 1_{t\in(a,b]})_{t\geq 0}$ is in $\mathcal{E}_B$ and thus orthogonal to $K \in \mathcal{P}_B$, which by (*) translates to

$$\begin{aligned} (K, \tilde{K})_B = \mathbb{E}(Y_a \int_a^b K_s ds) = \mathbb{E}(Y_a(X_b - X_a)) &\overset{(*)}{=} 0 \\ \implies \mathbb{E}(X_b \mid \mathcal{F}_a) &= X_a \end{aligned}$$

So $(X_t)_{t\geq 0}$ is a continuous martingale started from 0 and its of bounded variation, we know that this implies that $t \mapsto X_t$ is identically 0. Hence, K itself is equal to 0. ∎

7.2.5 Definition of the stochastic integral

With the previous Lemma at hand we are now ready to define the stochastic integral $I(K)$ of a progressively measurable process $K \in \mathcal{P}_B$ with respect to B, where B is a Brownian motion.

- We first consider any sequence K^n of elementary processes in $\mathcal{E}_B$ that converges to K in $\mathcal{P}_B$ (this is guaranteed to work by the previous Lemma, i.e. $\mathcal{E}_B$ is dense in $\mathcal{P}_B$).

- The sequence $(K^n)_{n\in\mathbb{N}}$ is then Cauchy with respect to the norm of $\mathcal{P}_B$ (norm induced by scalar product), because every convergent sequence is Cauchy, i.e. we have
$$\mathbb{E}\left(\int_0^\infty (K_s^n - K_s^l)^2 ds\right) \xrightarrow{n,l\to\infty} 0.$$
This implies that the sequence $I(K^n)$ is Cauchy in $\mathcal{M}^2$ because we know that the norm of $I(K^n) - I(K^l)$ in $\mathcal{M}^2$ is equal to the distance between K^n and K^l in $\mathcal{P}_B$. Recall that we have shown this isometry:
$$\mathbb{E}((I_\infty^{K_n} - I_\infty^{K_l})^2) = \mathbb{E}((I_\infty^{K_n - K_l})^2) = \mathbb{E}\left(\int_0^\infty (K_s^n - K_s^l)^2 ds\right) \xrightarrow{n,l\to\infty} 0.$$
  - Notice that since $(K^n)_{n\in\mathbb{N}}$ is a sequence of elementary processes in $\mathcal{E}_B$ we know already how to define the stochastic integral for such processes $(I_t^{K_n} = \int_0^t K_s^n dB_s)_{t\geq 0}$ and I^{K_n} is a continuous martingale.

- The space $\mathcal{M}^2$ is complete (it is in fact a Hilbert space), so there exists a continuous martingale $I(K)$ in $\mathcal{M}^2$ such that $I(K^n) \to I(K)$ in this space.

- We note that the continuous martingale $I(K)$ does not depend on our choice of sequence $(K^n)_{n\in\mathbb{N}}$ we did begin with in our first step to approximate K in $\mathcal{P}_B$. Indeed, if $(\tilde{K}^n)_{n\in\mathbb{N}}$ is another such sequence, then $I(K^n - \tilde{K}^n)$ does converge to 0 in $\mathcal{M}^2$ by using again the aforementioned isometry.

Definition 7.5. *We call this process $I(K)$ the stochastic integral of K with respect to B and we will denote it by*
$$I_t^K := \int_0^t K_s dB_s.$$

Remark 7.6. A more compact and equivalent way to summarize this construction is to say that $K \mapsto I(K)$ is an isometry from the set $\mathcal{E}_B$ into its image in $\mathcal{M}^2$, i.e. $\|I(K)\|_{\mathcal{M}^2} = \|K\|_B$. The extension of this map to $\mathcal{P}_B = \overline{\mathcal{E}}_B$ is the map $K \mapsto I(K)$.

7.3 Basic properties of stochastic integrals

We now list some of the basic properties of the stochastic integral with respect to Brownian motion. The proofs of these facts often proceeds as follows: First one checks directly that it holds in the case where the process K is an elementary process, and one then extends the result to the case $K \in \mathcal{P}_B$ by continuity.

- **An isometry:** Let $K \in \mathcal{P}_B$ and K^n be a sequence of elementary processes that approximates K. We know, since K^n is elementary that we have

$$\mathbb{E}[(I(K^n)_\infty)^2] \stackrel{\text{def}}{=} \mathbb{E}\left[\left(\int_0^\infty K_s^n dB_s\right)^2\right] = \mathbb{E}\left(\int_0^\infty (K_s^n)^2 ds\right). \qquad (*)$$

  But we also know that $I(K^n)_\infty \to I(K)_\infty$ in L^2 (because $I(K^n) \to I(K)$ in $\mathcal{M}^2$). But on the other hand, by the very definition of convergence of K^n to K in $\mathcal{P}_B$, we know that

$$\int_0^\infty (K_s^n)^2 ds \xrightarrow[n\to\infty]{L^1} \int_0^\infty K_s^2 ds.$$

  Thus by taking the limit as $n \to \infty$ in (*) above we obtain

$$\mathbb{E}\left[\left(\int_0^\infty K_s dB_s\right)^2\right] = \mathbb{E}\left(\int_0^\infty K_s^2 ds\right).$$

- **The quadratic variation of $I(K)$:** The quadratic variation of the continuous martingale $I(K)$ is the process $t \mapsto \int_0^t K_s^2 ds$. Indeed, one checks it first directly in the case for elementary processes. Then, using the same approximation methods as above we see that

$$(I(K^n)_\infty)^2 - \int_0^\infty (K_s^n)^2 ds \xrightarrow[n\to\infty]{L^1} (I(K)_\infty)^2 - \int_0^\infty K_s^2 ds$$

  The conditional expectations with respect to $\mathcal{F}_t$ do therefore converge as well. But the conditional expectation of the LHS (LHS we know is a martingale) is equal to $(I(K^n)_t)^2 - \int_0^t (K_s^n)^2 ds$. With the help of Doob's L^2-inequality

$$\begin{aligned}\sup_{t\geq 0} \mathbb{E}\left[(I(K^n)_t - I(K)_t)^2\right] &\leq \mathbb{E}\left[\sup_{t\geq 0}(I(K^n)_t - I(K)_t)^2\right] \\ &\leq 4\mathbb{E}\left[(I(K^n)_\infty - I(K)_\infty)^2\right] \xrightarrow{n\to\infty} 0,\end{aligned}$$

  which shows that $(I(K^n)_t)^2 \to (I(K)_t)^2$ in L^1 as $n \to \infty$. Similarly one shows that $\int_0^t (K_s^n)^2 ds \to \int_0^t K_s^2 ds$ in L^1. Hence we have shown that

$$(I(K^n)_t)^2 - \int_0^t (K_s^n)^2 ds \xrightarrow[n\to\infty]{L^1} (I(K)_t)^2 - \int_0^t K_s^2 ds =: M_t$$

  which shows that M_t is indeed a continuous martingale.

- **Approximations of stochastic integrals:** Suppose that K is a **continuous** process in $\mathcal{P}_B$. We then claim that for all $t \geq 0$, if Δ_n is a nested sequence of subdivisions of $[0,t]$ with $|\Delta_n| \to 0$, then

$$\sum_{i=0}^{m_n-1} K_{t_i^n}(B_{t_{i+1}^n} - B_{t_i^n}) \xrightarrow[n\to\infty]{L^2} \int_0^t K_s dB_s$$

Proof. Let us suppose first that $|K_t| \leq C$ almost surely for all $t \geq 0$. We define $\tilde{K}_s := K_s 1_{s\leq t}$, then we know that $\tilde{K}$ is still in $\mathcal{P}_B$ and we have

$$\int_0^\infty \tilde{K}_s dB_s = \int_0^t K_s dB_s.$$

We also define for all $n \in \mathbb{N}$

$$\tilde{K}_s^n := \sum_{i=0}^{m_n-1} K_{t_i^n} 1_{s\in(t_i^n, t_{i+1}^n]}$$

and we observe that

$$\int_0^\infty \tilde{K}_s^n dB_s = \int_0^t \tilde{K}_s^n dB_s = \sum_{i=0}^{m_n-1} K_{t_i^n}(B_{t_{i+1}^n} - B_{t_i^n}),$$

so in order to establish the claim it suffices to show that $\tilde{K}^n \to \tilde{K}$ in $\mathcal{P}_B$, since we know that this implies in particular the L^2 convergence of the integrals. We notice that by the continuity of K we get that

$$|\tilde{K}_s^n - \tilde{K}_s| \xrightarrow[n\to\infty]{a.s} 0$$

but also

$$|\tilde{K}_s^n - \tilde{K}_s|^2 \leq (2C)^2 1_{s\leq t},$$

thus by dominated convergence we get that

$$\mathbb{E}\left(\int_0^\infty |\tilde{K}_s^n - \tilde{K}_s|^2 ds\right) \overset{n\to\infty}{\to} 0$$

which establishes that $\tilde{K}^n \to \tilde{K}$ in $\mathcal{P}_B$. □

Remark 7.7. It is possible to remove the condition of boundedness in the previous proof, however, we will skip that here.

7.4 Generalization to stochastic integrals with respect to local martingales

7.4.1 The L^2 theory of stochastic integrals with respect to a continuous martingale

Let us now consider a continuous martingale $(M_t)_{t\geq 0}$ in our filtered probability space. We want to define the stochastic integral of processes K with respect to M, i.e., we want to define the continuous martingale $t \mapsto \int_0^t K_s dM_s$.

In some sense, what we are now going to do is to use the very same idea, modulo the fact that time plays now a different role, because of the time-change that is due to the quadratic variation A of M. To start with, we introduce a new set $\mathcal{E}_M$ of elementary processes as follows: These are elementary processes in $\mathcal{E}$, i.e. of the form

$$K_t = \sum_{j=0}^{m-1} K_{a_j} 1_{t\in(a_j,a_{j+1}]},$$

where $0 \leq a_0 < \cdots < a_m$ and each K_{a_j} is an $\mathcal{F}_{a_j}$-measurable random variable such that for each j

$$\mathbb{E}((K_{a_j})^2(A_{a_{j+1}} - A_{a_j})) < \infty.$$

Note that when M is a Brownian motion, then this is exactly the class of elementary processes $\mathcal{E}_B$, because if $M = B$, then $A = t$ and $t_{a_{j+1}} - t_{a_j} = a_{j+1} - a_j$.

Next, for each $K \in \mathcal{E}_M$, one can define the process

$$I(K)_t := \int_0^t K_s dM_s := \sum_{j=0}^{m-1} K_{a_j}(M_{t\wedge a_{j+1}} - M_{t\wedge a_j}),$$

the definition of $\mathcal{E}_M$ then guarantees us that the above is an L^2 martingale. We then define the set $\mathcal{P}_M$ of progressively measurable processes such that

$$\mathbb{E}\left(\int_0^\infty K_s^2 dA_s\right).$$

This space can be viewed as an L^2 space, when endowed with the scalar product

$$(K, K')_M := \mathbb{E}\left(\int_0^\infty K_s K'_s dA_s\right).$$

Now everything is exactly the same as in the construction we've already done. We keep the same space $\mathcal{M}^2$ of continuous martingales started from 0 that are bounded in L^2, we get

- The mapping $K \mapsto I(K)$ is an isometry from $\mathcal{E}_M$ into its image in $\mathcal{M}^2$.
  - In other words, $\|I(K)\|_{\mathcal{M}^2} = \|K\|_M$ or written out

$$\mathbb{E}\left[\left(\int_0^\infty K_s dM_s\right)^2\right] = \mathbb{E}\left(\int_0^\infty K_s^2 dA_s\right).$$

- The set $\mathcal{E}_M$ is dense in the Hilbert space $\mathcal{P}_M$.
- The mapping $K \mapsto I(K)$ can therefore be extended in a unique way into an isometry from $\mathcal{P}_M$ into the space $\mathcal{M}^2$.

Let us now list the basic properties of these stochastic integrals:

Proposition 7.2. *Suppose that M is a continuous martingale and that $K \in \mathcal{P}_M$ and $K' \in \mathcal{P}_M$, then:*

1. *The quadratic variation of the martingale $(\int_0^t K_s dM_s)_{t\geq 0}$ is the process $\int_0^t K_s^2 d\langle M\rangle_s$.*
2. *The cross-variation of the two martingales $(\int_0^t K_s dM_s)_{t\geq 0}$ and $(\int_0^t K'_s dM_s)_{t\geq 0}$ is the process $\int_0^t K_s K'_s d\langle M\rangle_s$.*
3. *If T denotes a stopping time, then if $\tilde{K}_t = K_t 1_{t<T}$, one has for all $t \geq 0$,*

$$\int_0^t \tilde{K}_s dM_s = \int_0^{t\wedge T} K_s dM_s = \int_0^t K_s dM_s^T.$$

4. *If K has almost surely left-continuous paths, and if $(\Delta_n)_{n\geq 0}$ denotes a nested sequence of subdivisions of $[0,t]$ with $|\Delta_n| \to 0$, then*

$$\sum_{i=0}^{m_n-1} K_{t_i^n}(M_{t_{i+1}^n} - M_{t_i^n}) \xrightarrow[n\to\infty]{\mathbb{P}} \int_0^t K_s dM_s$$

Proof. The proofs are essentially word for word the same as for Brownian motion. $\square$

7.4.2 Stochastic integrals with respect to a local martingale

Item (2) in the previous proposition enables us to directly generalize the definition of stochastic integrals to the following setting:

- The process M is a local martingale (and one denotes its quadratic variation by $(A_t)_{t\geq 0}$).
- The process K is a progressively measurable process such that almost surely, for all $t \geq 0$, $\int_0^t K_s^2 dA_s < \infty$.

Indeed, under these conditions, one can define $(\tilde{M}_t = M_t - M_0)_{t\geq 0}$ which will be a local martingale started from 0 and for all k, we define the sequence of stopping times

$$T_k := \inf\{t > 0 : t = k \text{ or } \int_0^t K_s^2 dA_s = k \text{ or } |\tilde{M}_t| = k\}.$$

One notes that almost surely, $T_k \nearrow \infty$ as $k \to \infty$, because $\tilde{M}$ is continuous and $\int_0^t K_s^2 dA_s$ is non-decreasing and finite. Thus the stopped process $\tilde{M}^{T_k}$ is then a proper martingale and the process K is in $\mathcal{P}_{\tilde{M}^{T_k}}$.

Then, one can define the stochastic integral of K with respect to the martingale $\tilde{M}^{T_k}$ and check that for each given t

$$\int_0^t K_s d\tilde{M}_s^{T_k}$$

is almost surely the same for all integer values of k such that $t \leq T_k$. We define this value to be $\int_0^t K_s dM_s$, and one can then check that this process is a local martingale started from 0, because for each k, this process stopped at T_k is a martingale.

Finally, one can check that all the items in the previous stated proposition still remains valid in this most general setting.

8 Itô's formula

8.1 Semimartingales

It will be useful to consider the class of processes that can be written as the sum of a local martingale and of an adapted continuous process with bounded variation:

Definition 8.1. *A process $(Z_t)_{t\geq 0}$ is a continuous semimartingale in the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ if it can be written as $Z_t = M_t + V_t$ where:*

- *M is a local martingale in this filtered probability space.*
- *V is the difference between two adapted continuous non-decreasing process V^+ and V^- started from 0 (equivalently: V is a continuous adapted bounded variation process started from 0).*

Remark 8.1. If Z is a continuous semimartingale, then the decomposition $Z = M + V$ is unique because if $Z = M + V = M' + V'$, then $C := M - M' = V' - V$, so C is a local martingale (LHS is difference of two local martingales) and it is of bounded variation (because RHS is the difference between two bounded variation processes), we know that this implies that $C = 0$.

This is a nice class of processes because if Z is a continuous semimartingale, we will see that $F(Z)$ is also a semimartingale when F is C^2, and we will describe the decomposition of $F(Z)$ in terms of $Z = M + V$.

Definition 8.2. *The quadratic variation of a continuous semimartingale is the quadratic variation of its local martingale part.*

This definition makes sense because

Lemma 8.1. *Suppose that Z is a continuous semimartingale and that Y is a continuous adapted process. For each $t \geq 0$, for all nested sequences $(\Delta_n)_{n\in\mathbb{N}}$ of subdivisions of $[0,t]$ with $|\Delta_n| \to 0$ one has*

$$\sum_{i=0}^{m_n-1} Y_{t_i^n}(Z_{t_{i+1}^n} - Z_{t_i^n})^2 \xrightarrow[n\to\infty]{\mathbb{P}} \int_0^t Y_s d\langle Z\rangle_s$$

Definition 8.3. *When $Z = M + V$ and $Z' = M' + V'$ are two continuous semimartingales (in the same filtered probability space), we define the cross-variation $\langle Z, Z'\rangle_t$ to be the crossvariation of $\langle M, M'\rangle_t$ of their two local martingale parts.*

Corollary 8.1. *Suppose that Z and Z' are two continuous semimartingales and that Y is an adapted continuous process. For each $t \geq 0$, for all nested sequences $(\Delta_n)_{n\geq 0}$ of subdivisions of $[0,t]$ with $|\Delta_n| \to 0$, one has*

$$\sum_{i=0}^{m_n-1} Y_{t_i^n}(Z_{t_{i+1}^n} - Z_{t_i^n})(Z'_{t_{i+1}^n} - Z'_{t_i^n}) \xrightarrow[n\to\infty]{\mathbb{P}} \int_0^t Y_s d\langle Z, Z'\rangle_s$$

8.1.1 Stochastic integrals with respect to a continuous semimartingale

We can then note that it is no problem to define the integral of a progressively measurable stochastic process K with respect to an adapted continuous bounded variation process V as soon as

$$\int_0^t |K_s||dV_s| < \infty$$

(see for instance exercise sheet for details about how to define $|dV_s|$). In the sequel, we will in fact almost only deal with the case where the process K_s is almost surely continuous, in which case the integral $\int_0^t K_s dV_s$ is always well-defined, as a Lebesgue-Stieltjes integral.

Definition 8.4. *Let $(Z_t)_{t\geq 0}$ be a continuous semimartingale, K a progressively measurable stochastic process, then we define*

$$\int_0^t K_s dZ_s := \int_0^t K_s dM_s + \int_0^t K_s dV_s$$

which is well-defined as soon $\int_0^t K_s^2 d\langle M\rangle_s$ and $\int_0^t |K_s||dV_s|$ are finite. The integral with respect to M is the previously defined stochastic integral and the integral with respect to V is the Stieltjes integral.

Remark 8.2. When K is an adapted and continuous process, one can always define $\int_0^t K_s dZ_s$, because $\int_0^t K_s dV_s$ is always well-defined.

Lemma 8.2. *When K is adapted and continuous, we have that for all $t \geq 0$, for all given sequences $(\Delta_n)_{n\in\mathbb{N}}$ of subdivisions of $[0,t]$ with $|\Delta_n| \to 0$*

$$\sum_{i<m_n} K_{t_i^n}(Z_{t_{i+1}^n} - Z_{t_i^n}) \xrightarrow[n\to\infty]{\mathbb{P}} \int_0^t K_s dZ_s.$$

8.2 Itô's formula

We are now ready to state Itô's formula in the case of one-dimensional continuous semimartingales:

Theorem 8.1 (Itô's formula, one-dimensional case)**.** *Suppose that Z is a continuous semimartingale (with decomposition $Z = M + V$) and that $F : D \to \mathbb{R}$ is a C^2 function from some open set $D \subset \mathbb{R}$ into $\mathbb{R}$ and that almost surely $Z_t \in D$ for all $t \geq 0$. Then, the process $(F(Z_t))_{t\geq 0}$ is also a continuous semimartingale, and almost surely for each $t \geq 0$ we have*

$$F(Z_t) = F(Z_0) + \int_0^t F'(Z_s)dZ_s + \frac{1}{2}\int_0^t F''(Z_s)d\langle Z\rangle_s.$$

Remark 8.3.

- We stress that in this formula, the integral $\int_0^t F(Z_s)dZ_s$ is the sum of the two integrals corresponding to the martingale part and to the bounded variation part, so that

$$F(Z_t) = F(Z_0) + \int_0^t F'(Z_s)dM_s + \int_0^t F'(Z_s)dV_s + \frac{1}{2}\int_0^t F''(Z_s)d\langle M\rangle_s.$$

- We then see, that the local martingale part of the continuous semimartingale $(F(Z_t))_{t\geq 0}$ is given by

$$F(Z_0) + \int_0^t F'(Z_s)dM_s,$$

  whereas the bounded variation part is the sum of the latter two terms, i.e.

$$\int_0^t F'(Z_s)dV_s + \frac{1}{2}\int_0^t F''(Z_s)d\langle M\rangle_s.$$

  In particular we see that Itô's formula gives us the unique decomposition of the semimartingale $F(Z_t)$.

- If $Z_t = B_t$ is a Brownian motion in 1 dimension (every continuous martingale is of course a semimartingale with $V = 0$) then

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

Proof of Itô's formula. The proof will be a direct consequence of the combination of Taylor's expansion for F with the approximation results of stochastic integrals and of quadratic variations.

We notice that it will suffice to prove for each fixed t, the identity asserted by Itô's formula for $F(Z_t)$ almost surely, this is enough because both sides are almost surely continuous.

Moreover, for each k let D_k denote the set of points that are at a distance at least $1/k$ of $\mathbb{R} \setminus D$. We can define the stopping time

$$T_k = \inf\{t > 0, |M_t| \geq k \text{ or } |V_t| = k \text{ or } |Z_t| = k \text{ or } Z_t \notin D_k\}$$

and if we know that the result holds for all the continuous semimartingales Z^{T_k}, then we can deduce that it will hold for Z as well, by letting $k \to \infty$.

Hence, we conclude that it suffices to show the identity for a fixed time t and in the case where the semi-martingale Z does almost surely not exit $[-k, k] \cap D_k$.

Note first that F being C^2 in a $1/2k$ neighborhood of $[-k, k] \cap D_k$, there exists a constant such that for all $z \in [-k - (1/2k), k + (1/2k)] \cap D_k$, we have $\max(|F(z)|, F'(z)|, |F''(z)|) \leq C$. In particular Taylor's formula for F can simply be written as

$$F(z + h) = F(z) + hF'(z) + h^2 F''(z)/2 + h^2 R(z, h), \tag{*}$$

where $\sup_{z\in[-k,k]} R(z, h)$ tends to 0 as $h \to 0$.

We now choose Δ_n to be a subdivision of $[0, t]$ into 2^n intervals of length $t2^{-n}$, we denote the increments $h_i^n := Z_{t_{i+1}^n} - Z_{t_i^n}$, we then simply get

$$\begin{aligned} F(Z_t) - F(Z_0) &= \sum_{i<2^n} (F(Z_{t_{i+1}^n}) - F(Z_{t_i^n})) = \sum_{i<2^n} (F(Z_{t_i^n} + h_i^n) - F(Z_{t_i^n})) \\ &\overset{(*)}{=} \sum_{i<2^n} (F'(Z_{t_i^n}) \times h_i^n) + \frac{1}{2} \sum_{i<2^n} (F''(Z_{t_i^n}) \times (h_i^n)^2) \\ &\quad + \sum_{i<2^n} (R(Z_{t_i^n}, h_i^n) \times (h_i^n)^2). \end{aligned}$$

By Lemma 8.1 and Lemma 8.2 we know that for the first two sums

$$\begin{aligned} \sum_{i<2^n} (F'(Z_{t_i^n}) \times h_i^n) + \frac{1}{2} \sum_{i<2^n} (F''(Z_{t_i^n}) \times (h_i^n)^2) \xrightarrow[n\to\infty]{\mathbb{P}} &\int_0^t F'(Z_s) dZ_s \\ &+ \frac{1}{2} \int_0^t F''(Z_s) d\langle Z \rangle_s \end{aligned}$$

So it remains to check that the last sum goes to 0 in probability as $n \to \infty$. But we do already know, also by Lemma 8.1 that

$$\sum_{i<2^n} (h_i^n)^2 \xrightarrow[n\to\infty]{\mathbb{P}} \langle Z \rangle_t.$$

Since Z is almost surely continuous we get that

$$\lim_{n\to\infty} \sup_{i<2^n} |R(Z_{t_i^n}, h_i^n)| = 0.$$

We therefore conclude that

$$\left| \sum_{i<2^n} R(Z_{t_i^n}, h_i^n) \times (h_i^n)^2 \right| \leq \sup_{i<2^n} |R(Z_{t_i^n}, h_i^n)| \sum_{i<2^n} (h_i^n)^2 \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

which concludes the proof of Itô's formula. □

8.2.1 Itô's formula in higher dimensions

It is in fact very useful to consider continuous semimartingales with values in $\mathbb{R}^d$ for $d \geq 2$.

Definition 8.5. *A process $(Z_t = (Z_t^1, \ldots, Z_t^d))_{t\geq 0}$ with values in $\mathbb{R}^d$ is a continuous (d-dimensional) semimartingales, if for reach $1 \leq j \leq d$, the process Z^j is a (one-dimensional) continuous semimartingale.*

Theorem 8.2 (Itô's formula in higher dimensions)**.** *Suppose that Z is a d-dimensional continuous semimartingale (with decomposition $Z = M + V$) and that F is a C^2 function, from some open domains $D \subset \mathbb{R}^d$ to $\mathbb{R}$. Suppose that almost surely, $Z_t \in D$ for all $t \geq 0$. Then, the process $(F(Z_t))_{t\geq 0}$ is a continuous one-dimensional semimartingale, and almost surely, for each $t \geq 0$,*

$$F(Z_t) = F(Z_0) + \sum_{j=1}^{d} \int_0^t \frac{\partial F}{\partial z_j}(Z_s) dZ_s^j + \frac{1}{2} \sum_{1\leq i,j\leq d} \int_0^t \frac{\partial^2 F}{\partial z_i \partial z_j}(Z_s) d\langle Z^i, Z^j \rangle_s.$$

Proof. Is exactly as the one dimensional case. Just use Corollary 8.1 to control the crossvariation terms. □

Remark 8.4. One simple case is when $Z = (X, Y)$ and where $F(x, y) = xy$. It shows that the product of two continuous semimartingales is a continuous semimartingale and that (integration by parts formula for cont. semimg.)

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \langle X, Y \rangle_t$$

8.3 Applications of Itô's formula

A first important consequence of Itô's formula is the fact that, just as in the case of Brownian motion, when one starts with a local martingale M, it is possible to define an infinite family of exponential local martingales $\mathcal{E}(\lambda M)$.

Proposition 8.1. *When M is a local martingale in a given filtered probability space. We then define for all $t \geq 0$,*

$$\mathcal{E}_t^\lambda = \mathcal{E}(\lambda M)_t := \exp\left(\lambda M_t - \frac{\lambda^2}{2}\langle M\rangle_t\right).$$

The process $\mathcal{E} = (\mathcal{E}_t^\lambda)_{t\geq 0}$ is then a local martingale. Moreover, for $\lambda = 1$ and for all $t \geq 0$,

$$\begin{aligned}\mathcal{E}_t &= \mathcal{E}_0 + \int_0^t \mathcal{E}_s dM_s, \ \textit{and}\\ M_t &= M_0 + \int_0^t \frac{d\mathcal{E}_s}{\mathcal{E}_s}.\end{aligned}$$

Proof. Let us consider $Z_t = (M_t, \langle M\rangle_t)$, which is a continuous semimartingale with values in $\mathbb{R}^2$. Let us also take the function $F(a,b) = \exp(\lambda a - (\lambda^2/2)b)$, which is a C^2 function from $\mathbb{R}^2$ to $\mathbb{R}$.

Before we apply Itô's formula we notice that the quadratic variation terms involving the derivatives with respect to the second variable vanish, because $\langle M\rangle_t$ is a continuous bounded variation process with no quadratic variation. Thus by Itô's formula we obtain that

$$\begin{aligned}F(Z_t) = F(Z_0) &+ \int_0^t \lambda f(Z_s) dM_s + \int_0^t \overbrace{-\frac{\lambda^2}{2}F(Z_s)}^{=\partial_a F}\, d\langle M\rangle_s\\ &+ \frac{1}{2}\int_0^t \underbrace{\lambda^2 F(Z_s)}_{=\partial_a^2 F}\, d\langle M\rangle_s\\ = F(Z_0) &+ \int_0^t \lambda f(Z_s) dM_s \rightsquigarrow \text{local martingale.}\end{aligned}$$

We thus see we have indeed for $\lambda = 1$,

$$\mathcal{E}_t = \mathcal{E}_0 + \int_0^t \mathcal{E}_s dM_s.$$

In particular, the decomposition

$$\mathcal{E}_t = \mathcal{E}_0 + \int_0^t \mathcal{E}_s dM_s.$$

yields us

$$\langle \mathcal{E}_t \rangle_t = \langle \mathcal{E}_0 \rangle_t + \langle \int_0^t \mathcal{E}_s dM_s \rangle = 0 + \int_0^t \mathcal{E}_s^2 d\langle M \rangle_s.$$

or equivalently:

$$d\langle \mathcal{E} \rangle_t = \mathcal{E}_t^2 d\langle M \rangle_t \tag{*}$$

Conversely, since we know now that $(\mathcal{E}_t)_{t\geq 0}$ is a local martingale, and moreover since it is always positive, we can apply Itô's formula to obtain the local martingale/bounded variation process decomposition of $\log \mathcal{E}_t$, we get

$$\begin{aligned}\log \mathcal{E}_t &= \log \mathcal{E}_0 + \int_0^t \frac{d\mathcal{E}_s}{\mathcal{E}_s} - \frac{1}{2}\int_0^t (\mathcal{E}_s)^{-2} d\langle \mathcal{E} \rangle_s \\ &\overset{(*)}{=} M_0 + \int_0^t \frac{d\mathcal{E}_s}{\mathcal{E}_s} - \frac{1}{2}\langle M \rangle_t\end{aligned}$$

So that indeed we get (by solving for M_t in $\log \mathcal{E}_t$) that

$$M_t = M_0 + \int_0^t \frac{d\mathcal{E}_s}{\mathcal{E}_s}$$

□

Remark 8.5. The proof above shows that if F is any C^2 function on $\mathbb{R}^2$, such that

$$\partial_a F(a,b) + \frac{1}{2}\partial_a^2 F(a,b) = 0$$

then $(F(M_t, \langle M \rangle_t))_{t\geq 0}$ is a local martingale as soon M is a local martingale. In particular, this shows that for any $z \in \mathbb{C}$, the real part and the imaginary parts of $\exp(zM_t - z^2\langle M \rangle_t/2)$ are local martingales.

Especially, this shows that for any $z \in \mathbb{C}$,

$$(\exp(zM_t - z^2\langle M \rangle_t/2))_{t\geq 0}$$

is a complex-valued local martingale.

Corollary 8.2 (Lévy's characterization of a Brownian motion in one dimension)**.** *Suppose that M is a local martingale started from the origin, such that $(M_t^2 - t)_{t\geq 0}$ is also a local martingale (equivalently $\langle M\rangle_t = t$ for all $t \geq 0$). Then M is a Brownian motion.*

Proof. Let us first check that for a given $t \geq 0$, the law of M_t is that of a Brownian motion at time t. Let us look at

$$\tilde{\mathcal{E}}_t := \mathcal{E}(i\lambda M)_t = \exp(i\lambda M_t + \lambda^2\langle M\rangle_t/2) = \exp(i\lambda M_t + \lambda^2 t/2).$$

Since $|\tilde{\mathcal{E}}_t| = \exp(\lambda^2 t/2)$ (which is deterministic and thus in L^1). Let us then take a sequence of non-decreasing stopping times T_n that diverge to ∞. Then $(\tilde{\mathcal{E}}_{t\wedge T_n})_{t\geq 0}$ is a martingale and $|\tilde{\mathcal{E}}_{t\wedge T_n}| \leq \exp(\lambda^2 t/2)$ (integrable), thus by dominated convergence (taking $n \to \infty$) we get that $(\tilde{\mathcal{E}}_t)_{t\geq 0}$ is a complex martingale, thus

$$\mathbb{E}(\exp(i\lambda M_t + \lambda^2 t/2)) = \mathbb{E}(\tilde{\mathcal{E}}_t) = \mathbb{E}(\tilde{\mathcal{E}}_0) = 1,$$

so that indeed the characteristic function of M_t is given by

$$\mathbb{E}(\exp(i\lambda M_t)) = \exp(-\lambda^2 t/2) \implies M_t \sim \mathcal{N}(0,t).$$

More generally, we have for all $s \leq t$

$$\mathbb{E}(\exp(i\lambda M_{t+s} + \frac{\lambda^2}{2}(t+s)) \mid \mathcal{F}_s) = \exp(i\lambda M_s + \frac{\lambda^2}{2}s)$$
$$\implies \mathbb{E}(e^{i\lambda(M_{s+t}-M_s)} \mid \mathcal{F}_s) = e^{-\frac{\lambda^2}{2}t}$$

Hence for all $t \geq 0$ and for all $s \leq t$, $M_{t+s} - M_s$ is a distributed as $\mathcal{N}(0,t)$ and is independent of $\mathcal{F}_s$, which shows that M has stationary independent increments. Thus M is indeed a Brownian motion. □

Corollary 8.3 (Lévy's characterization of a Brownian motion in $\mathbb{R}^d$)**.** *Suppose that $M = (M^1, \ldots, M^d)$ is a d-dimensional local martingale started from the origin such that:*

- *For all $k \neq j$, $\langle M^k, M^j\rangle_t = 0$.*
- *For all $1 \leq j \leq d$, $\langle M^j\rangle_t = t$.*

Then M is a Brownian motion in $\mathbb{R}^d$.

Corollary 8.4. *Suppose that $(M_t)_{t\geq 0}$ is a local martingale started from the origin, such that $\langle M\rangle_\infty = \infty$ almost surely. Define for all $u \geq 0$, $\tau_u := \inf\{t > 0, \langle M\rangle_t > u\}$. Then the process $(M_{\tau_u})_{u\geq 0}$ is a Brownian motion.*

8.4 More applications of Itô's formula

Itô's formula is an incredible powerful tool which we will demonstrate with the next couple of examples. We will start by revisiting some results that we already know (related to the Dirichlet Problem).

Proposition 8.2. *Let $D \subset \mathbb{R}^d$ be a bounded open domain, if $H : \overline{D} \to \mathbb{R}$ is such that*

- *H is continuous in $\overline{D}$.*
- *H is C^2 in D and $\Delta H = 0$ in D. (i.e. Harmonic in D).*

Then $H(x) = \mathbb{E}_x(H(B_\tau))$ for all $x \in \overline{D}$ where B denotes a d-dimensional Brownian motion started from x and $\tau = \inf\{t > 0 : B_t \in \partial D\}$ is the first exit time of B from D.

Proof. Let us fix $x \in D$ and start a Brownian motion B from x. For $\epsilon > 0$ define $T_\epsilon := \inf\{t > 0 : d(B_t, \partial D) = \epsilon\}$ and we notice that $T_\epsilon \to \tau$ as $\epsilon \to 0$. We need this stopping time because it guarantees that a.s. $B_{t\wedge T_\epsilon} \in D$ for all $t \geq 0$. For a d-dimension Brownian motion we have for the crossvariations

$$\begin{cases} \langle B^i, B^j \rangle_t = 0, & \text{if } i \neq j \\ \langle B^i, B^i \rangle_t = t, & \text{if } i = j \end{cases}$$

(in order to show the first case in the statements above one actually needs to use the fact that B^i, B^j are independent for $i \neq j$).
Since $H \in C^2$ in D, we can apply Itô's formula to the martingale B^{T_ϵ} to obtain

$$\begin{aligned} H(B_{t\wedge T_\epsilon}) &= H(B_0 = x) + \sum_{j=1}^d \int_0^{t\wedge T_\epsilon} \frac{\partial H}{\partial z_j}(B_s) dB_s^j + \frac{1}{2}\int_0^{t\wedge T_\epsilon} \Delta H(B_s) ds \\ &= H(x) + \sum_{j=1}^d \int_0^{t\wedge T_\epsilon} \frac{\partial H}{\partial z_j}(B_s) dB_s^j \end{aligned}$$

hence $(H(B_{t\wedge T_\epsilon}))_{t\geq 0}$ is a local martingale, but H is also bounded (because it is continuous on the compact set $\overline{D}$) and thus $(H(B_{t\wedge T_\epsilon}))_{t\geq 0}$ is in fact a (bounded) continuous martingale. Since it is bounded it is also UI and thus by the optional stopping theorem we get that

$$\mathbb{E}(H(B_{T\epsilon})) = H(x)$$

Letting $\epsilon \to 0$ we finally get by dominated convergence theorem (H is bounded), that

$$\mathbb{E}_x(H(B_\tau)) = H(x)$$

which proves the claim. □

Let us now come back to one of the cousins of the Dirichlet problem, which is in fact a generalisation of the Dirichlet problem.

Proposition 8.3. *Suppose that D is an open bounded subset of $\mathbb{R}^d$ and we are given $\alpha : D \to \mathbb{R}$ a continuous, bounded, non-negative function. Let $H : \overline{D} \to \mathbb{R}$ is continuous, C^2 in D and such that*

$$\Delta H(x) = 2\alpha(x)H(x), \text{ for all } x \in D,$$

then

$$H(x) = \mathbb{E}_x \left[H(B_\tau) \exp\left(- \int_0^\tau \alpha(B_s) ds \right) \right].$$

Proof. Let us define for all $t \leq T_\epsilon$

$$Z_t := H(B_t) e^{-\int_0^t \alpha(B_s)ds}$$

We then notice that $Z_t = F(B_t, Y_t)$ where $F(b, y) = H(b)e^{-y}$ is a C^2 function from $\mathbb{R}^d \times \mathbb{R} = \mathbb{R}^{d+1}$ to $\mathbb{R}$ (because H is C^2) where $Y_t = \int_0^t \alpha(B_s)ds$ is clearly of bounded variation, in particular (B_t, Y_t) is indeed a semimartingale in dimension $d + 1$.

We also notice that the crossvariation of the any Brownian motion B^j for $1 \leq j \leq d$ with the process Y_t always gives 0, because Y_t is of bounded variation, in particular $\langle Y_t, Y_t \rangle_t = 0$. Itô's formula thus yields:

$$\begin{aligned} Z_t = F(B_t, Y_t) &= Z_0 + \sum_{j=1}^d \int_0^t \frac{\partial H}{\partial z_j}(B_s) e^{-Y_s} dB_s^j - \int_0^t H(B_s) e^{-Y_s} \underbrace{dY_s}_{=\alpha(B_s)ds} \\ &\quad + \frac{1}{2} \int_0^t \underbrace{\Delta H(B_s)}_{=2\alpha(B_s)H(B_s)} e^{-Y_s} ds \\ &= Z_0 + \sum_{j=1}^d \int_0^t \frac{\partial H}{\partial z_j}(B_s) e^{-Y_s} dB_s^j \end{aligned}$$

We thus conclude that $(F(B_{t\wedge T_\epsilon}, Y_{t\wedge T_\epsilon}))_{t\geq 0}$ is a local martingale, since F is also bounded (because H and $\exp^{-y}$ is) so in particular it is a bounded continuous martingale (and thus also UI), by optional stopping we get that

$$\mathbb{E}_x(Z_{T_\epsilon}) = H(x) = \mathbb{E}_x \left[H(B_{T_\epsilon}) \exp\left(- \int_0^{T_\epsilon} \alpha(B_s) ds \right) \right].$$

Letting $\epsilon \to 0$ we obtain the claim by dominated convergence. $\square$

8.4.1 Brownian motion and the heat equation

The same approach as used in the previous applications shows that Brownian motion is directly related to the *heat equation*.

Setup: Suppose that we are given a bounded open domain $D \subset \mathbb{R}^d$ with a regular boundary, and that $f : \overline{D} \to \mathbb{R}$ is continuous on $\overline{D}$ that is equal to 0 on ∂D.

We say that the continuous function F defined on $\overline{D} \times [0, \infty)$ solves the heat equation with 0 boundary conditions on ∂D and initial value f if:

- The function $F : (x,t) \mapsto F(x,t)$ is C^2 on $D \times (0,\infty)$, and it satisfies $\partial_t F(t,x) = \frac{1}{2}\Delta_x F(t,x)$.
- For each $t \geq 0$, $F(x,t) = 0$ when $x \in \partial D$.
- For each $x \in D$, $F(x,0) = f(x)$.

Proposition 8.4. *If the solution to the heat equation exists, then it is unique and it is equal to*

$$F(x,t) = \mathbb{E}_x(f(B_t)1_{t<T})$$

where B is a Brownian motion started from x and T denotes its exit time from D.

Proof. Suppose that F is a solution to the heat equation. Let T_ϵ be as before the first time at which B is at distance less than ϵ from ∂D. Let us fix $t_0 >$ and let us consider for all $t \leq T_\epsilon \wedge t_0$ the process

$$Z_t = F(B_t, t_0 - t)$$

(we set $Z_t = Z_{t_0 \wedge T_\epsilon}$ for $t \geq t_0 \wedge T_\epsilon$). Then $(B_t, t_0 - t)$ is again a $d+1$ dimensional semi-martingale. Itô's formula yields

$$\begin{aligned} Z_t &= Z_0 + \sum_{j=1}^{d} \int_0^t \partial_{x_j} F(B_s, t_0 - s) dB_s^j - \int_0^t \partial_t F(B_s, t_0 - s) ds \\ &\quad + \frac{1}{2} \int_0^t \Delta_x F(B_s, t_0 - s) ds \\ &= Z_0 + \sum_{j=1}^{d} \int_0^t \partial_{x_j} F(B_s, t_0 - s) dB_s^j \end{aligned}$$

So we obtain once again that $(Z_t)_{t\geq 0}$ is a local martingale, it is bounded and thus a (true) bounded martingale, thus we get

$$\mathbb{E}(Z_{t_0\wedge T_\epsilon}) = \mathbb{E}(Z_0) = F(B_0, t_0) = F(x, t_0)$$

Letting $\epsilon \to 0$ (dominated convergence) gives us for the left hand side

$$\begin{aligned}
\mathbb{E}(Z_{t_0\wedge T}) &= \mathbb{E}(F(B_{t_0\wedge T}, t_0 - (t_0 \wedge T))) \\
&= \mathbb{E}(F(B_{t_0}, 0)1_{t_0<T}) + \mathbb{E}(F(B_T, t_0 - T)1_{t_0>T}) \\
&= \mathbb{E}(f(B_{t_0})1_{t_0<T}) + 0 = \mathbb{E}(f(B_{t_0})1_{t_0<T})
\end{aligned}$$

where we used that $F(B_T, t_0-T) = 0$ because $B_T \in \partial D$. This completes the proof of the statement. $\square$

9 Stochastic Differential Equations

9.1 Motivation and general remarks

Let us first informally describe stochastic differential equations (SDE), first in 1 dimension. We are given two function $\sigma : \mathbb{R} \to \mathbb{R}$ and $b : \mathbb{R} \to \mathbb{R}$ and a starting point $x \in \mathbb{R}$. It is then our goal to find a process $(X_t)_{t\geq 0}$ which is a semimartingale and such that for all $t \geq 0$

$$X_t = x + \int_0^t \sigma(X_s) dB_s + \int_0^t b(X_s) ds$$

where B is a 1-dimensional Brownian motion.

More generally in dimension d, we are given $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ (i.e. σ takes values in the space of $d \times d$ matrices), and $b : \mathbb{R}^d \to \mathbb{R}^d$ and $x \in \mathbb{R}^d$, then the SDE can be formally written as

$$dX_t^i = \sum_{j=1}^d \sigma_{i,j}(X_s) dB_s^j + b_i(X_s) ds, \; 1 \leq i \leq d$$

or more compactly

$$\begin{aligned} X_t &= x + \int_0^t \sigma(X_s) dB_s + \int_0^t b(X_s) ds \\ &\iff dX_t = \sigma(X_t) dB_t + b(X_t) dt \end{aligned}$$

where B is a d-dimensional Brownian motion.

Remark 9.1. We notice that when $\sigma = 0$, an SDE is an ODE.

As opposed to the case of ordinary differential equations (ODE), it turns out that there are different conceptual notions of solutions of such stochastic differential equations (we will call them **weak** and **strong** - not to be confused with weak solutions in PDEs). These notions correspond to the fact that one may be allowed to choose the probability space that one defines the process X and B in, or that one is already given the probability space and the Brownian motion B and we have to find X in said probability space.

Let us briefly recall two main families of results for ODEs:

- There is the Picard-Lindelölf theorem. This theorem shows existence and uniqueness of the solutions to an ODE $y' = b(y)$ with $y(0) = y_0$ when b is a bounded Lipschitz function.
  - Existence part is derived via Picard's iteration which provides a constructive existence proof.
  - Uniqueness part follows from a fixed point argument.

  When b is not bounded and only locally Lipschitz, we still have existence and uniqueness of local solutions.

- There is the Peano Theorem (or the stronger Carathéodory Theorem). It shows existence (but **not uniqueness**) of solutions to an ODE $y' = b(y)$ under weaker assumptions for b, i.e. boundedness and continuity of b ensures existence of solutions. Moreoever if b is not bounded we still have local existence of solutions.

These two types of theorems have analogues for Stochastic Differential Equations. The analogue of the Picard-Lindelöf approach will give rise to the existence and uniqueness of **strong** solutions, i.e. of solutions that are fully determined by the Brownian motion B in the equation and no other randomness is involved.

The analog of the Peano approach will use tightness in the space of probability measures on continuous functions (like when discussed Donsker's Theorem). It therefore constructs a probability space (with a probability measure $\mathbb{P}$) on which there exists a solution, by defining $\mathbb{P}$ to be the weak limit of probability measures. This will give rise to existence of so called **weak** solutions.

9.2 Existence and uniqueness of strong solutions

Definition 9.1. *We say that a function f defined on $\mathbb{R}^d$ (or on $\mathbb{R}^{d\times d}$) is K-Lipschitz (continuous) if for all x, y*

$$\|f(x) - f(y)\| \leq K\|x - y\|$$

where $K \geq 0$ is some constant and $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^d$ or on $\mathbb{R}^{d\times d}$. We also say that if a function f is K-Lipschitz continuous just that it is Lipschitz.

Theorem 9.1 (Globally Lipschitz: Existence and uniqueness of strong solutions)**.** *Let us suppose that we are given a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ and a d-dimensional Brownian motion $(B_t)_{t\geq 0}$ in this filtration. Suppose that the functions σ and b are Lipschitz functions with values in $\mathbb{R}^{d\times d}$ and $\mathbb{R}^d$ respectively. Then:*

- *For all $x_0 \in \mathbb{R}^d$, there exists a continuous semimartingale X such that for all $t \geq 0$,*

$$X_t = x_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dB_s.$$

- *This process $(X_t)_{t\geq 0}$ is the unique one with these properties.*

The proof of the theorem will be reminiscent of the proof of the Picard-Lindelöf Theorem for ODEs. Let us briefly recall the following two easy lemmas which will be the building blocks in that proof:

Lemma 9.1 (Gronwall's lemma)**.** *If f is a bounded real-valued measurable function on some interval $[0, T]$ such that there exists c such that for all $t \leq T$, $|f(t)| \leq c\int_0^t |f(s)|ds$, then $f = 0$ on $[0, T]$.*

Proof. We iterate the condition $|f(t)| \leq c\int_0^t |f(s)|ds$ to obtain

$$|f(t)| \leq c\max(|f|)t \implies |f(t)| \leq c^2 \max(|f|)\frac{t^2}{2} \overset{\dots}{\implies} |f(t)| \leq \max(|f|)\frac{t^n c^n}{n!}$$

The last quantity converges to 0 as $n \to \infty$. □

Lemma 9.2 (Fixed point of a contraction, slight variation)**.** *Suppose that F is a continuous function from some complete metric space (E, d) into itself, such that there exists $n > 0$ such that for all $x, y \in E$, $d(F^n(x), F^n(y)) \leq d(x, y)/2$. Then F has a unique fixed point z in E i.e. $F(z) = z$. (Here F^n denotes the n-th iterate of F, i.e. $F^n = F \circ F^{n-1}$).*

Proof. Choose some $x \in E$ and consider the sequence $x_k = F^{n_k}(x)$. We then easily obtain by the contraction assumption that

$$d(x_k, x_{k+1}) \leq 2^{-k} d(x_0, x_1) \text{ and } d(x_k, F(x_k)) \leq 2^{-k} d(x, F(x)),$$

So we see that $(x_k)_{k\geq 0}$ is a Cauchy sequence in E and since E is complete $x_k \to x_\infty \in E$, but we also have since F is continuous that $(F(x_k))_{k\to 0}$ converges to the same point, i.e. $\lim_{k\to\infty} F(x_k) = F(x_\infty) = x_\infty$.

Assume now that $x' \in E$ is another fixed point of F, then we have

$$d(x', x_\infty) = d(F^n(x'), F^n(x_\infty)) \leq \frac{d(x', x_\infty)}{2}$$

which can only hold when $d(x', x_\infty) = 0$, i.e. $x' = x_\infty$. □

Proof of theorem. We will write out the proof in the case of $d = 1$, but the proof for general d is exactly the same. We also notice that for $T \geq 0$ fixed, it suffices to prove existence and uniqueness of a process $(X_t)_{t\in[0,T]}$ that solves the SDE on the time-interval $[0, T]$.

Uniqueness (Gronwall): Suppose that X and Y are two solutions, let τ_n denote the first time at which either $|X|$ or $|Y|$ reach n, and let us prove that the stopped process X^{τ_n} and Y^{τ_n} defined on $[0, T]$ are almost surely equal, if we prove this for all n then this clearly implies that $X = Y$ on $[0, T]$. For $t \leq T$ let us denote $t_n = \min(t, \tau_n)$ (just to make notation easier).
Recall that the convexity of $t \mapsto t^p$ for $p \geq 1$ implies $\|f+g\|_p^p \leq 2^{p-1}(\|f\|_p^p + \|g\|_p^p)$, or equivalently for RV $\mathbb{E}((X+Y)^2) \leq 2(\mathbb{E}(X^2) + \mathbb{E}(Y^2))$. $(p = 2)$

$$\begin{aligned}
\mathbb{E}[(X_t^{\tau_n} - Y_t^{\tau_n})^2] &= \mathbb{E}\left[\left(\int_0^{t_n} (\sigma(X_s) - \sigma(Y_s))dB_s + \int_0^{t_n} (b(X_s) - b(Y_s))ds\right)^2\right] \\
&\leq 2\mathbb{E}\left[\left(\int_0^{t_n} (\sigma(X_s) - \sigma(Y_s))dB_s\right)^2\right] + 2\mathbb{E}\left[\left(\int_0^{t_n} (b(X_s) - b(Y_s))ds\right)^2\right] \\
&\overset{\text{Isometry}}{\leq} 2\mathbb{E}\left[\int_0^{t_n} (\sigma(X_s) - \sigma(Y_s))^2 ds\right] + 2\mathbb{E}\left[\left(\int_0^{t_n} |b(X_s) - b(Y_s)|ds\right)^2\right] \\
&\overset{\text{Lipschitz}}{\leq} 2K^2\mathbb{E}\left[\int_0^{t_n} (X_s - Y_s)^2 ds\right] + 2K^2\mathbb{E}\left[\left(\int_0^{t_n} |X_s - Y_s|ds\right)^2\right] \\
&\overset{\text{Hölder}}{\leq} 2K^2(1+T)\mathbb{E}\left[\int_0^{t_n} (X_s - Y_s)^2 ds\right] \\
&\leq 2K^2(1+T)\mathbb{E}\left[\int_0^t (X_s^{\tau_n} - Y_s^{\tau_n})^2 ds\right] = C\int_0^t \mathbb{E}[(X_s^{\tau_n} - Y_s^{\tau_n})^2]ds
\end{aligned}$$

Hence if we write $f(t) = \mathbb{E}((X_t^{\tau_n} - Y_t^{\tau_n})^2)$, then we can apply Gronwall's lemma to conclude that $f = 0$ which concludes the uniqueness part of the proof.

Existence (Picard's iteration, fixed point of contraction): We work on the set $E := \{$continuous adapted process $Y : [0, T] \to \mathbb{R}^d$, with $Y(0) = x_0$ and $\mathbb{E}(\sup_{[0,T]} |Y_t|^2) < \infty\}$. We equip this space with the metric $d(X, Y) := \mathbb{E}(\sup_{s\in[0,T]}(X_s - Y_s)^2)$, then (E, d) is a complete metric space (seen similar arguments in class already to prove this).

Let us define $F : E \to E$ by

$$F(Y)_t := x_0 + \int_0^t \sigma(Y_s)dB_s + \int_0^t b(Y_s)ds.$$

Also for $t_0 \leq T$, we denote $d_{t_0}(X, Y) = \mathbb{E}(\sup_{s\leq t_0}(X_s - Y_s)^2)$. Almost the same arguments as used in the uniqueness part (except that one uses also Doob's inequality) shows that when X, Y are in E, then

$$\begin{aligned}
d_{t_0}(F(X), F(Y)) &= \mathbb{E}[\sup_{t\leq t_0}(F(X)_t - F(Y)_t)^2] \\
&\leq 2\mathbb{E}\left[\left(\sup_{t\leq t_0}\int_0^t (\sigma(X_s) - \sigma(Y_s))dB_s\right)^2\right] + 2\mathbb{E}\left[\left(\sup_{t\leq t_0}\int_0^t (b(X_s) - b(Y_s))ds\right)^2\right] \\
&\overset{\text{Doob}}{\leq} 2\cdot 4\mathbb{E}\left[\left(\int_0^{t_0} (\sigma(X_s) - \sigma(Y_s))dB_s\right)^2\right] + 2\mathbb{E}\left[\left(\int_0^{t_0} |b(X_s) - b(Y_s)|ds\right)^2\right] \\
&\leq 8K^2\mathbb{E}\left[\int_0^{t_0} (X_s - Y_s)^2 ds\right] + 2K^2\mathbb{E}\left[\left(\int_0^{t_0} |X_s - Y_s|ds\right)^2\right] \\
&\leq 2K^2(4 + t_0)\mathbb{E}\left[\int_0^{t_0} (X_s - Y_s)^2 ds\right] \\
&\overset{\text{Tonelli}}{\leq} 2K^2(4 + T)\int_0^{t_0} d_s(X, Y)ds \leq C\int_0^{t_0} d_T(X, Y)ds \\
&\leq Ct_0 d_T(X, Y), \text{ where } C := 2K^2(4 + T)
\end{aligned}$$

So we have shown that for all $t \leq t_0$, we have $d_t(F(X), F(Y)) \leq Ctd_T(X, Y)$, and then iteratively (just like in proof of Gronwall's Lemma) we get that for all $n \geq 1$

$$d_t(F^n(X), F^n(Y)) \leq Ct^n d_T(X, Y)/n!.$$

So for all n large enough we have $d_T(F^n(X), F^n(Y)) \leq d_T(X, Y)/2$. We conclude by using the fixed-point lemma, that there exists a unique continuous adapted process $(X_t)_{t\in[0,T]}$ such that $F(X) = X$. □

9.3 Survey of some other results on SDEs

9.3.1 Other results obtained via Picard's iteration method

Let us now suppose that the conditions on σ and b are relaxed to local Lipschitz conditions, i.e. for all $R > 0$, there exists $K_R \geq 0$, such that for all $x, y \in \mathbb{R}^d$ with $\|x\|, \|y\| \leq R$ we have

$$\|\sigma(x) - \sigma(y)\| \leq K_R \|x - y\|,$$
$$\|b(x) - b(y)\| \leq K_R \|x - y\|.$$

The following result can be derived using the same ideas as for the globally Lipschitz case, in particular this theorem establishes existence and uniqueness of strong solutions up to a explosion time:

Theorem 9.2 (Locally Lipschitz: Existence and uniqueness of strong solutions)**.** *Suppose that the functions σ and b are locally Lipschitz functions with values in $\mathbb{R}^{d\times d}$ and $\mathbb{R}^d$ respectively, then:*

- *For all $x_0 \in \mathbb{R}^d$, there exists a stopping time $T \in [0, \infty]$ and an adapted continuous process X such that for all $t \leq T$,*

$$X_t = x_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dB_s,$$

and such that almost surely on the event $\{T < \infty\}$, the set $\{X_t, t < T\}$ is not bounded.

- *If T' and X' is another stopping time and another process with these properties, then $T = T'$ almost surely, and almost surely $X_t = X'_t$ for all $t \leq T$.*

When T is finite, T is called the explosion time of the solution to the SDE.

It is also possible to get the uniqueness statement under somewhat weaker assumptions on σ:

Theorem 9.3 (Yamade-Watanabe's pathwise uniqueness)**.** *Suppose that b is Lipschitz, and that there exists an increasing function $h : [0, 1] \to \mathbb{R}_+$ with $h(0) = 0$ and $\int_0^1 h(x)^{-2}dx = \infty$, such that for all x and y,*

$$\|\sigma(x) - \sigma(y)\| \leq h(\|x - y\|).$$

Then, if X and Y are two processes defined on the same filtered probability space (on which one is given a Brownian motion B), such that for all $t \geq 0$,

$$X_t = x_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dB_s = Y_t$$

then $X = Y$ almost surely.

9.3.2 Tanaka's example

Let us describe an instructive example that illustrates what sort of surprising things can happen when σ is not continuous.

Let us look at the easy case when $d = 1$, $b = 0$, $x_0 = 0$ and $\sigma(x) = \mathrm{sgn}(x) = 1_{x\geq 0} - 1_{x<0}$. It is our goal to find an adapted process $(X_t)_{t\geq 0}$ started from the origin, such that for all $t \geq 0$

$$X_t = \int_0^t \mathrm{sgn}(X_s) dB_s,$$

where B is a 1-dimensional Brownian motion.

Let us make some comments about this situation:

- Assume that such a process X would exists, then it is a local martingale (since it is a stochastic integral), but we would also have

$$\langle X\rangle_t = \int_0^t (\mathrm{sgn}(X_s))^2 ds = \int_0^t 1 ds = t$$

  so by Lévy's characterisation, necessarily X is in fact a Brownian motion (in this filtered probability space).

- We recall that the zero-set of a Brownian motion X has almost surely zero Lebesgue measure, this is easy to prove because by Fubini we have

$$\mathbb{E}\left(\int_0^\infty 1_{X_s=0} ds\right) = \int_0^\infty \overbrace{\mathbb{P}(X_s = 0)}^{=0} ds = 0 \implies \int_0^\infty 1_{X_s=0} ds = 0.$$

  This implies in particular that the local martingale $(\int_0^t 1_{X_s=0} dB_s)_{t\geq 0}$ is constant and equal to 0, indeed

$$\mathbb{E}\left[\left(\int_0^\infty 1_{X_s=0} dB_s\right)^2\right] = \mathbb{E}\left(\int_0^\infty (1_{X_s=0})^2 ds\right) = \mathbb{E}\left(\int_0^\infty 1_{X_s=0} ds\right) = 0.$$

  So we have $\int_0^\infty 1_{X_s=0} dB_s =$ and we get the values of $\int_0^t 1_{X_s=0} dB_s$ by using martingales arguments (conditioning on $\mathcal{F}_t$).

- If such a process X exists and if we define $Y = -X$, then this process Y would also satisfy the SDE, because ($-\mathrm{sgn}(x) = \mathrm{sgn}(-x) - 2 \times 1_{x=0}$)

$$\begin{aligned} Y_t &= -\int_0^t \mathrm{sgn}(X_s) dB_s = \int_0^t \mathrm{sgn}(-X_s) dB_s - 2\int_0^t 1_{X_s=0} dB_s \\ &= \int_0^t \mathrm{sgn}(Y_s) dB_s - 0 = \int_0^t \mathrm{sgn}(Y_s) dB_s \end{aligned}$$

The last item shows that for this SDE, one can not have a unique strong solution as in the previous Lipschitz case. It is in fact possible to show that there is no strong solution at all. On the other hand the first item shows that any solution has a specific law (as it is necessarily a Brownian motion).

Let us now suppose that we are given a Brownian motion X on some filtered probability space, then one can define the process B by

$$B_t := \int_0^t \mathrm{sgn}(X_s) dX_s.$$

This process B is therefore a Brownian motion in this space (because it is a local martingale with $\langle B \rangle_t = t$), and for all $t \geq 0$, we have

$$\int_0^t \mathrm{sgn}(X_s) dB_s = \int_0^t (\mathrm{sgn}(X_s))^2 dX_s = \int_0^t dX_s = X_t$$

Hence we have constructed a filtered probability space and a pair of processes (X, B) on this space such that B is a Brownian motion and such that for all $t \geq 0$,

$$X_t = \int_0^t \mathrm{sgn}(X_s) dB_s.$$

This type of construction is described as the existence of **weak solutions** to this SDE.

9.3.3 Existence of weak solutions

Let us now state without proof the existence result for weak solutions, that is the SDE analog of Peano's theorem for ODEs.

Theorem 9.4 (Bounded & continuous coefficients, existence of weak solutions)**.** *Suppose that σ and b are bounded continuous functions (from $\mathbb{R}^d$ into $\mathbb{R}^{d \times d}$ and into $\mathbb{R}^d$ respectively). Suppose that $x \in \mathbb{R}^d$. Then it is possible to construct a filtered probability space and two continuous adapted processes X and B in this probability space, such that B is a d-dimensional Brownian motion for this filtration, and such that for all $t \geq 0$,*

$$X_t = x + \int_0^t b(X_s) ds + \int_0^t \sigma(X_s) dB_s.$$

Remark 9.2. Again, if we remove the boundedness constraint, one can show the existence of weak solutions up to some explosion time. As we have already mentioned, this type of result is based on compactness/tightness argument (but there is more to it in the proof).

9.3.4 Combining weak existence and pathwise uniqueness

Suppose that we are in the framework of the Yamade-Watanabe theorem, and that σ and b are continuous and also bounded. In other words we assume two things:

- σ and b are continuous and also bounded (see previous theorem on weak existence).
- σ satisfies moreover the condition of Yamade-Watanabe's theorem.

One then has the existence of weak solutions on some probability space, and by the Yamade-Watanabe Theorem, in that probability space, there is no other solution to the SDE. It is then in fact possible to deduce that this solution on that probability space is even a strong solution (i.e. our weak solution is a strong solution). Essentially because if the solution was not a deterministic function of a Brownian motion in that space, it would be possible to construct another one.

To summarize, in the framework as introduced above, it is possible to show that a weak solution, is actually a strong solution.

An application of this is the construction of the squared Bessel processes (and then of the Bessel processes) that we now give (pay special attention to the exercise sheets!):

Corollary 9.1 (Definition of Bessel processes)**.** *The previous remarks show that for any non-negative real d, one has existence and uniqueness of the solution to the SDE*

$$dZ_t = 2\sqrt{|Z_t|}dB_t + d \times dt.$$

Definition 9.2. *The process $Y = \sqrt{Z}$ where Z is a squared Bessel process of dimension d, is called a Bessel process of dimension d.*

Itô's formula then shows that for all $t \geq 0$, and as long as $t < \inf\{s \geq 0 : Y_s = 0\}$,

$$Y_t = Y_0 + B_t + \frac{d-1}{2}\int_0^t \frac{ds}{Y_s}.$$

10 Absolute continuity between laws of processes

In this final section, we will describe a more probabilistic idea that relates martingales to change of probability measures. This will also provide tools to construct weak solutions of SDEs, and results that have no real analog in the ODE setting, and that are making use of the randomness in an SDE.

10.1 Warm-up

Recall the definition of absolute continuity

Definition 10.1 (Absolute continuity of measures). *Given two measures μ and ν on the same measurable space $(\Omega, \mathcal{A})$, then μ is said to be absolutely continuous with respect to ν (or dominated by ν) denoted by $\mu \ll \nu$ if for all $A \in \mathcal{A}$ with $\nu(A) = 0$ implies $\mu(A) = 0$.*

The relevant theorem that goes with this theorem is that of the Radon-Nikodým derivative:

Theorem 10.1 (Radon-Nikodým). *Let $(\Omega, \mathcal{A})$ be a measurable space, on which two σ-finite measure μ and ν are defined. Assume that $\nu \ll \mu$ (i.e. ν is absolutely continuous with respect to μ), then there exists a measurable function $f : \Omega \to [0, \infty)$ such that for every measurable set $A \in \mathcal{A}$*

$$\nu(A) = \int_A f d\mu.$$

Furthermore, f is unique μ-almost everywhere. The function f is called the Radon-Nikodým derivative of ν with respect to μ and is denoted by

$$f = \frac{d\nu}{d\mu}$$

Remark 10.1. Since we work in a probabilistic settings, i.e. on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which we are given two probability measures, say $\mathbb{P}$ and $\mathbb{Q}$, then these measures are always finite (because they assign a unit value to the measure of the whole space Ω), in particular they are σ-finite.

A relevant question we're going to discuss here is the following: Given a Brownian motion $(B_t)_{t\leq 1}$ and a deterministic continuous function $(h(t))_{t\geq 1}$ we wonder, under what conditions on h is it true that the law of $(B_t + h(t))_{t\leq 1}$ is absolutely continuous with respect to the law of $(B_t)_{t\leq 1}$.

10.1.1 Change of probability measures and martingales

We will first discuss why uniformly integrable non-negative martingales with mean 1 can be interpreted as a Radon-Nikodým derivative between two probability measures in a given filtered probability space.

Suppose that we are given a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ (let us assume for convenience that the filtration is right-continuous) and lets assume that we are given another probability measure $\mathbb{Q}$ in this space that is absolutely continuous with respect to $\mathbb{P}$.

By the theorem of Radon-Nikodým there exists $D \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ (non-negative) such that for all $A \in \mathcal{F}$ we have

$$\mathbb{Q}(A) = \mathbb{E}_{\mathbb{P}}(1_A D).$$

Let us define for all $t \in [0, \infty]$, $D_t := \mathbb{E}(D \mid \mathcal{F}_t)$. We notice that for all $t \in [0, \infty]$ we have $\mathbb{E}(D_t) = \mathbb{E}(D) = \mathbb{E}(\mathbb{Q}(\Omega)) = 1$. We also define $D_\infty = \mathbb{E}(D \mid \mathcal{F}_\infty)$ on $\mathcal{F}_\infty$.

Proposition 10.1. *This process $(D_t)_{t\geq 0}$ is a uniformly integrable, non-negative martingale (in the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$) that converges almost surely and in L^1 to the non-negative random variable D_∞.*

Proof. The martingale property is trivial. But also since it's in closed form ($X_i = \mathbb{E}(X \mid \mathcal{G}_i)$ for a sequence of σ-fields), we also know that it is UI. The convergence is then a consequence of the martingale convergence thm. □

Conversely, let us start with a non-negative uniformly integrable martingale $(\mathcal{D}_t)_{t\geq 0}$ in $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ with $\mathbb{E}(\mathcal{D}_t) = 1$ for all $t \geq 0$. Since $\mathcal{D}$ is UI we know that $\mathcal{D}_t \to \mathcal{D}_\infty$ in L^1 and almost surely, thus in particular $\mathbb{E}(\mathcal{D}_\infty) = \mathbb{E}(\mathcal{D}_t) = 1$. We can then use this to define another probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ by

$$\mathbb{Q}(A) := \mathbb{E}(1_A \mathcal{D}_\infty), \text{ for all } A \in \mathcal{F}.$$

It is easy to check that $\mathbb{Q}$ is a probability measure, using that $\mathbb{E}(\mathcal{D}_\infty) = 1$ and $(\mathcal{D})_{t\geq 0}$ is non-negative. Then we see that $(\mathcal{D}_t)_{t\geq 0}$ is exactly the process $(D_t)_{t\geq 0}$ defined as in the proposition above. That is we have

$$\mathcal{D}_t = \mathbb{E}(\mathcal{D}_\infty \mid \mathcal{F}_t).$$

We also note that in this setup, $\mathbb{Q}$ is indeed absolutely continuous with respect to $\mathbb{P}$, because if $\mathbb{P}(A) = \mathbb{E}(1_A) = 0$, then in particular $\mathbb{E}(1_A \mathcal{D}_\infty) = \mathbb{Q}(A) = 0$.

10.2 Girsanov's theorem

10.2.1 Cameron-Martin space

Let us restate our question: Let h be a continuous deterministic, real-valued function defined on $[0,1]$. Under what conditions on h is it true that the law of $B_t+h(t)$ is absolutely continuous with respect to the law of B_t on the time-interval $[0,1]$? For those functions h for which we do have absolute continuity, what is the Radon-Nikodým derivative of the law of $(B_t+h(t))_{t\leq 1}$ with respect to that of $(B_t)_{t\leq 1}$.

Clearly, for some continuous functions h, we will not have absolute continuity. If we take for example $h(t)=t^{1/3}$, then we know that almost surely

$$\limsup_{t\to 0+}\frac{B_t+h(t)}{t^{1/3}}=1 \text{ and } \limsup_{t\to 0+}\frac{B_t}{t^{1/3}}=0,$$

so that the laws of $(B_t)_{t\in[0,1]}$ and of $(B_t+h(t))_{t\in[0,1]}$ are singular with respect to each other.

The answer to our question is given by:

Proposition 10.2 (Cameron-Martin space)**.** *Define the class $\mathcal{C}$ of functions*

$$\mathcal{C}:=\{h:[0,1]\to\mathbb{R} \textit{ continuous } \mid \exists v\in L^2[0,1]: h(t)=\int_0^t v(s)ds\}.$$

Then, for $h\in\mathcal{C}$, the law of $(B_t+h(t))_{t\leq 1}$ is absolutely continuous with respect to that of $(B_t)_{t\leq 1}$. Moreover, the Radon-Nikodým derivative of the law of $(B_t+h(t))_{t\leq 1}$ with respect to that of $(B_t)_{t\leq 1}$ is given by

$$\exp\left(\int_0^1 v(s)dB_s-\frac{1}{2}\int_0^1 v(s)^2ds\right)$$

We wont prove this Proposition, however we will state and prove the stronger Girsanov Theorem which will in particular imply the Cameron-Martin space.

10.2.2 Girsanov's theorem

Let us consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F})_{t\geq 0}, \mathbb{P})$. Let N denote a continuous martingale started from $N_0 = 0$. We then define the exponential local martingale

$$\mathcal{E}_t := \mathcal{E}(N)_t = \exp(N_t - \langle N \rangle_t / 2).$$

We assume that this exponential local martingale is UI and denote its limit by $\mathcal{E}_\infty$, we then have $\mathcal{E}_\infty \in L^1$ with $\mathbb{E}(\mathcal{E}_\infty) = \mathbb{E}(\mathcal{E}_0) = 1$. Our previous considerations allow us to define another probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ by

$$\mathbb{Q}(A) := \mathbb{E}(1_A \mathcal{E}_\infty), \text{ for all } A \in \mathcal{F}.$$

We also notice that $\mathcal{E}_t = \mathbb{E}(\mathcal{E}_\infty \mid \mathcal{F}_t)$, which implies that for all $A \in \mathcal{F}_t$, $\mathbb{Q}(A) = \mathbb{E}(1_A \mathcal{E}_t)$. So, one can view $\mathcal{E}_t$ as the Radon-Nikodým derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$ when viewed on $(\Omega, \mathcal{F}_t)$.

Theorem 10.2 (Girsanov)**.** *Under the above conditions (i.e. when the local martingale $\mathcal{E}(N)$ is uniformly integrable), if $M = (M_t)_{t\geq 0}$ is a local martingale (under the probability $\mathbb{P}$), then in the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{Q})$ (i.e. under the probability measure $\mathbb{Q}$), the process*

$$\tilde{M}_t := M_t - \langle M, N \rangle_t$$

is a local martingale. Moreover we have ($\mathbb{Q}$ and $\mathbb{P}$)-a.s. $\langle \tilde{M} \rangle_t = \langle M \rangle_t$.

Let us summarize what Girsanov's theorem tells us exactly:

- Under $\mathbb{P}$: M is a local martingale and $\tilde{M}_t = M_t - \langle M, N \rangle_t$ is a semimartingale (since it is given in its decomposition of a local martingale M and a bounded variation term $\langle M, N \rangle$).
- Under $\mathbb{Q}$: $\tilde{M}$ is now a local martingale and $M_t = \tilde{M}_t + \langle M, N \rangle_t = \tilde{M}_t + \langle \tilde{M}, N \rangle_t$ is a semimartingale (again written in its unique decomposition).

Let us also write down the special case for Brownian motion:

Theorem 10.3 (Girsanov, Brownian case)**.** *Under the above conditions (i.e., when the local martingale $\mathcal{E}(N)$ is uniformly integrable), if B denotes a Brownian motion (under the probability measure $\mathbb{P}$), then under the probability measure $\mathbb{Q}$, the process $\tilde{B}_t = B_t - \langle B, N \rangle_t$ is a Brownian motion.*

This is indeed an immediate consequence from the first theorem because the local martingale has quadratic variation $\langle \tilde{B}_t \rangle = \langle B_t \rangle = t$ and one concludes with Lévy's characterisation.

Before we prove Girsanov's theorem, we want to illustrate how one can use it to deduce the previously mentioned absolute continuity result (Cameron-Martin space).

Consider a Brownian motion B and some deterministic function h defined on $[0,1]$ in the space $\mathcal{C}$. We can then extend h and v by letting $h(t) = h(1)$ and $v(t) = 0$ on $(1,\infty)$. We define

$$N_t := \int_0^t v(s)dB_s.$$

This is of course a continuous local martingale, and we note that

$$\mathcal{E}(N)_t = \exp\left(\int_0^t v(s)dB_s - \frac{1}{2}\int_0^t v(s)^2 ds\right),$$

and is easy to check (because N is the time-change of a Brownian motion running until the deterministic finite time $\int_0^1 v(s)^2 ds$), that $\mathcal{E}(N)$ is uniformly integrable. Then by Girsanov's theorem we know that under the probability measure $\mathbb{Q}$ defined as

$$\mathbb{Q}(A) = \mathbb{E}_{\mathbb{P}}(1_A \mathcal{E}_\infty) = \mathbb{E}_{\mathbb{P}}\left[1_A \exp\left(\int_0^1 v(s)dB_s - \frac{1}{2}\int_0^1 v(s)^2 ds\right)\right],$$

the process $\tilde{B}_t = B_t - \langle B, N\rangle_t = B_t - \int_0^t v(s)ds = B_t - h(t)$ is a Brownian motion. In particular we see that the Radon-Nikodým derivative of the law of $(B_t = \tilde{B}_t + h(t))_{t\leq 1}$ with respect to the law of $(B_t)_{t\leq 1}$ is given by

$$\exp\left(\int_0^1 v(s)dB_s - \frac{1}{2}\int_0^1 v(s)^2 ds\right).$$

Proof of Girsanov's theorem. Let $T_n := \inf\{t > 0 : |\tilde{M}_t| = n\}$ (enough to work with this sequence of stopping times thanks to exercise 9.1.). It is our goal to check that for all $t, s \geq 0$

$$\mathbb{E}_{\mathbb{Q}}(\tilde{M}_{t+s}^{T_n} \mid \mathcal{F}_t) = \tilde{M}_t^{T_n} \text{ almost surely.}$$

In other words for all $A \in \mathcal{F}_t$ we have $\mathbb{E}_{\mathbb{Q}}(\tilde{M}_{t+s}^{T_n} 1_A) = \mathbb{E}_{\mathbb{Q}}(\tilde{M}_t^{T_n} 1_A)$. Or equivalently to this (only rephrased, see definition of $\mathbb{Q}$) we want to show that for all $A \in \mathcal{F}_t$

$$\mathbb{E}_{\mathbb{P}}(\mathcal{E}_\infty \tilde{M}_{t+s}^{T_n} 1_A) = \mathbb{E}_{\mathbb{P}}(\mathcal{E}_\infty \tilde{M}_t^{T_n} 1_A). \tag{*}$$

In order to verify (*), we will show that $(\mathcal{E}_t \tilde{M}_t^{T_n})_{t\geq 0}$ is a local martingale in $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$.

Recall that $\mathcal{E}_t = \mathcal{E}(N)_t$ is a local martingale and we have already shown with the help of Itô's formula that $d\mathcal{E}_t = \mathcal{E}_t dN_t$. Also we have derived the integration by parts formula for the product of continuous semimartingales ($dX_tY_t = X_t dY_t + Y_t dX_t + d\langle X, Y\rangle_t$). Since $(\tilde{M}_t = M_t - \langle M, N\rangle_t)_{t\geq 0}$ is a semimartingale, we can apply Itô's formula to the product $\mathcal{E}_t\tilde{M}_t$.

$$\begin{aligned} d(\mathcal{E}_t\tilde{M}_t) &\overset{\text{Itô}}{=} \mathcal{E}_t d\tilde{M}_t + \tilde{M}_t d\mathcal{E}_t + d\langle \tilde{M}_t, \mathcal{E}_t\rangle_t \\ &= \mathcal{E}_t dM_t - \mathcal{E}_t d\langle M, N\rangle_t + \tilde{M}_t d\mathcal{E}_t + d\langle M, \mathcal{E}\rangle_t \\ &\overset{!}{=} \mathcal{E}_t dM_t - \mathcal{E}_t d\langle M, N\rangle_t + \tilde{M}_t d\mathcal{E}_t + \mathcal{E}_t d\langle M, N\rangle_t \\ &= \mathcal{E}_t dM_t + \tilde{M}_t d\mathcal{E}_t \end{aligned}$$

Hence $(\mathcal{E}_t\tilde{M}_t)_{t\geq 0}$ is a local martingale (careful, at the step we highlighted with a ! we used an exercise, namely 11.3.e, generalization of crossvariation). Similarly, the process $(\mathcal{E}_t\tilde{M}_t^{T_n})_{t\geq 0}$ is a local martingale, because $d(\mathcal{E}_t\tilde{M}_t^{T_n}) = 1_{t<T_n} d(\mathcal{E}_t\tilde{M}_t) + 1_{t>T_n}\tilde{M}_{T_n} d\mathcal{E}_t$.

But we also know that $(\mathcal{E}_t)_{t\geq 0}$ is by assumption uniformly integrable, and $|\mathcal{E}_t\tilde{M}_t^{T_n}| \leq n\mathcal{E}_t$, this implies that for all fixed $n \geq 1$, $(\mathcal{E}_t\tilde{M}_t^{T_n})_{t\geq 0}$ is also uniformly integrable and in particular it is a true continuous martingale (for fixed $n \geq 1$).

Let $\mathcal{E}_\infty \in L^1$ denote the limit of the UI martingale $(\mathcal{E}_t)_{t\geq 0}$. We then have $\mathcal{E}_{t+s} = \mathbb{E}(\mathcal{E}_\infty \mid \mathcal{F}_{t+s})$ and thus:

$$\mathbb{E}(\mathcal{E}_\infty\tilde{M}_{t+s}^{T_n}1_A) = \mathbb{E}(\mathcal{E}_{t+s}\tilde{M}_{t+s}^{T_n}1_A) = \mathbb{E}(\mathcal{E}_t\tilde{M}_t^{T_n}1_A) = \mathbb{E}(\mathcal{E}_\infty\tilde{M}_t^{T_n}1_A)$$

Where we used the shorthand notation $\mathbb{E} = \mathbb{E}_{\mathbb{P}}$, which is exactly statement (*). Since this is true for all $n \in \mathbb{N}$ this means that $\tilde{M}$ is a local martingale under the probability measure $\mathbb{Q}$ and thus concludes our proof of Girsanov's theorem. □

Of course, in order to apply Girsanov's theorem, it is crucial to be able to check that $\mathcal{E}(N)$ is a uniformly integrable martingale. There exist some useful criteria for this (proof will be omitted here):

Proposition 10.3 (Kawazaki's and Novikov's criteria)**.** *Suppose that N is a local martingale with $N_0 = 0$ almost surely, then:*

- *If $\mathbb{E}(\exp(\langle N\rangle_\infty/2)) < \infty$, then $\mathcal{E}(N)$ is uniformly integrable.*
- *If N is a uniformly integrable martingale and if $\mathbb{E}(\exp(N_\infty/2)) < \infty$, then is $\mathcal{E}(N)$ uniformly integrable.*

10.3 Some consequences for SDEs

Let us now outline how Girsanov's theorem can be used to obtain existence results for (weak) solutions of SDEs in one dimensions under fairly weak regularity assumptions on the functions σ and b.

Let us first focus on the case where σ is constant and equal to 1. In other words, one is looking for a real-valued process X such that

$$X_t = x_0 + B_t + \int_0^t b(X_s)ds. \tag{*}$$

Strategy of construction: Start with a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ on which one has defined a Brownian motion started from x_0, denoted by $(X_t)_{t\geq 0}$.

Idea: Find another probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0})$ such that under $\mathbb{Q}$

$$X_t - x_0 - \int_0^t b(X_s)ds$$

is a Brownian motion and then call this process $(B_t)_{t\geq 0}$. Consequently in the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{Q})$ one has 2 processes, namely X and B where B is a Brownian motion and

$$X_t = x_0 + B_t + \int_0^t b(X_s)ds \text{ almost surely.}$$

In particular we have constructed a weak solution to the SDE (*).

Lets make our idea more formal and explicit, i.e. we wonder how we can define such a $\mathbb{Q}$:

- We start with a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ on which the process X is a Brownian motion started from x_0.
- We then define for each T, the probability measure $\mathbb{Q}^T$ on $(\Omega, \mathcal{F}_T)$ (denote $\mathbb{P}^T = \mathbb{P}$ restricted to $\mathcal{F}_T$) via

  $$\frac{d\mathbb{Q}^T}{d\mathbb{P}^T} = \exp\left(\int_0^T b(X_s)dX_s - \frac{1}{2}\int_0^T b(X_s)^2 ds\right)$$

  Recall that Girsanov's theorem tells us that given a Brownian motion B under $\mathbb{P}$, and $d\mathbb{Q}/d\mathbb{P} = \exp(\int_0^t v(B_s)dB_s - \frac{1}{2}\int_0^t v(s)^2 ds)$, then under $\mathbb{Q}$ the process $\tilde{B}_t = B_t - \int_0^t v(B_s)ds$ is a Brownian motion.

    - This is for instance possible for all b that are bounded and measurable.

- We then define $\mathbb{Q}$ on $(\Omega, \mathcal{F}_\infty)$ such that $\mathbb{Q} = \mathbb{Q}^T$ on $\mathcal{F}_T$ for each T.
    - Under this $\mathbb{Q}$:

$$\left(B_t := X_t - x_0 - \int_0^t b(X_s)ds \right)_{t \leq T}$$

      is a Brownian motion up to time T. Where we used again that from Girsanov's theorem we know that $B_t = X_t - x_0 - \langle X, N \rangle_t$ is a Brownian motion.

- Then we conclude that in the probability space $(\Omega, \mathcal{F}_\infty, (\mathcal{F}_t)_{t \geq 0}, \mathbb{Q})$, B is a Brownian motion and that almost surely, for all $t \geq 0$

$$X_t = x_0 + B_t + \int_0^t b(X_s)ds.$$

This way to construct a weak solution to this SDE can easily be generalized to higher dimensions. This time, the constraint is that σ is constant times the identity matrix. We thus have derived the following result:

Proposition 10.4. *When b is a bounded measurable function from $\mathbb{R}^d$ into $\mathbb{R}^d$, then for each $x_0 \in \mathbb{R}^d$, one can construct a weak solution to the SDE (in $\mathbb{R}^d$)*

$$dX_t = dB_t + b(X_t)dt$$

started from $X_0 = x_0$.

Remark 10.2. It is noticeable that there is no continuity-type assumption on b in this statement. This is in contrast with the case of ODEs, where some regularity is required to obtain existence of solutions. In some sense, the presence of the Brownian motion term has a regularising effect.

Let us now consider the case when σ is not constant, it is then still possible to use the previous idea in order to construct a (weak) solution to the SDE

$$dX_t = \sigma(X_t)dB_t + b(X_t)dt$$

in the **one dimensional case** (we will not generalize to higher dimensions here). The difference is that instead of starting with a probability space on which X is a Brownian motion (and then apply Girsanov's theorem), one has to first find a way to define a local martingale that has the right quadratic variation structure, before applying Girsanov's theorem.

Assumption: It turns out that it is useful to assume σ is a measurable function and that there exists $0 < \epsilon < C$ such that for all $x \in \mathbb{R}$, we have $|\sigma(x)| \in [\epsilon, C]$.

- First we need to construct a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ on which X is a local martingale started from x_0 with $\langle X \rangle_t = \int_0^t \sigma(X_s)^2 ds$.
  - This can be achieved by starting with a Brownian motion $(\beta_u)_{u\geq 0}$ with respect to some filtration $(\mathcal{G}_u)_{u\geq 0}$ and by appropriately time-changing it. More precisely, let

    $$t(u) := \int_0^u \frac{dv}{\sigma(\beta_v)^2}.$$

    Our conditions on σ guarantee that $u \mapsto t(u)$ is an increasing continuous bijection from $[0, \infty)$ into itself. We can then define its inverse function $u(t) := t^{-1}(u)$ and let $X_t := \beta_{u(t)}$. The process $(X_t)_{t\geq 0}$ is then a local martingale for the filtration $\mathcal{F}_t := \mathcal{G}_{u(t)}$ and it's quadratic variation is indeed

    $$\langle X \rangle_t = u(t) = \int_0^t \sigma(X_s)^2 ds.$$

- Then one applies Girsanov's theorem: One chooses

  $$N_t := \int_0^t \frac{b(X_s)}{\sigma(X_s)^2} dX_s.$$

  This is a continuous martingale and we have chosen it in such a way that $\langle X, N \rangle_t = \int_0^t b(X_s)ds$. We construct another probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0})$ such that for all T, the Radon-Nikodým derivative between $\mathbb{Q}$ and $\mathbb{P}$ on $\mathcal{F}_T$ is given by

  $$\frac{d\mathbb{Q}^T}{d\mathbb{P}^T} = \exp(N_T - \langle N \rangle_T / 2)$$

- Girsanov's theorem establishes that under the probability measure $\mathbb{Q}$, the process

  $$\tilde{X}_t := X_t - x_0 - \langle X, N \rangle_t = X_t - x_0 - \int_0^t b(X_s) ds$$

  is a local martingale and it is such that that $d\langle \tilde{X} \rangle_t = d\langle X \rangle_t = \sigma(X_t)^2 dt$. Let us then define

  $$B_t := \int_0^t \sigma(X_s)^{-1} d\tilde{X}_s.$$

  We easily see that the process B is a local martingale under the probability measure $\mathbb{Q}$ and we have

  $$\langle B_t \rangle_t = \int_0^t \sigma(X_s)^{-2} d\langle \tilde{X} \rangle_s = \int_0^t \sigma(X_s)^{-2} \sigma(X_s)^2 ds = \int_0^t ds = t$$

  so by Lévy's characterization theorem, we know that B is indeed a Brownian motion under the probability measure $\mathbb{Q}$. But then under this probability measure $\mathbb{Q}$

  $$\int_0^t \sigma(X_s) dB_s = \int_0^t \sigma(X_s)\sigma(X_s)^{-1} d\tilde{X}_s = \int_0^t d\tilde{X}_s = \tilde{X}_t$$

  so that finally

  $$X_t = x_0 + \tilde{X}_t + \int_0^t b(X_s) ds = x_0 + \int_0^t \sigma(X_s) dB_s + \int_0^t b(X_s) ds.$$

We have thus derived:

Proposition 10.5 (Existence for one-dimensional SDE via Girsanov)**.** *If there exists $\epsilon > 0$ and C such that for all $x \in \mathbb{R}$, $|b(x)| \leq C$ and $|\sigma(x)| > \epsilon$, and if $|\sigma|$ is bounded on any compact interval, then there exists a weak solution to the SDE $dX_t = \sigma(X_t)dB_t + b(X_t)dt$ stated from $x_0 \in \mathbb{R}$.*

Remark 10.3. The construction of solutions via Girsanov's theorem does actually provide more than just existence of weak solutions. It does in fact also show that any two weak solutions have necessarily the same law. (The idea is to go our construction above backwards and obtain that the law must be the same, will be skipped here).