Chapter 1

General Gaussian Measures

1.1 Isonormal Gaussian processes

Definition 1.1. Let H be a real separable Hilbert space with inner product $\langle .,. \rangle_H$ and norm $\|.\|_H$. An isonormal Gaussian process over H is a centered Gaussian family

$$X = \{X(h) : h \in H\},$$

indexed by H and defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, such that for all $h, g \in H$,

$$\mathbb{E}[X(h)X(g)] = \langle f, g \rangle_H.$$

Theorem 1.2. *For every real separable Hilbert space H, there exists an isonormal Gaussian process over H.*

Proof. Let $\{\xi_i : i \geqslant 1\}$ be a collection of i.i.d. random variables with normal distribution $\mathcal{N}(0,1)$ defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\{e_i : i \geqslant 1\}$ be an ONB of H (it exists because H is separable). Let $h = \sum_{i=1}^{\infty} \langle e_i, h \rangle_H e_i \in H$. For all $N \geqslant 1$, define

$$X_N(h) := \sum_{i=1}^{N} \langle e_i, h \rangle_H \xi_i.$$

Then $X_N(h)$ is a centered Gaussian r.v. as a linear combination of i.i.d. centered Gaussian variables. For all $M < N$, we have

$$\begin{aligned}
\mathbb{E}\left[\left(X_M(h) - X_N(h)\right)^2\right] &= \mathbb{E}\left[\left(\sum_{i=M+1}^{N} \langle e_i, h \rangle_H \xi_i\right)^2\right] \\
&= \sum_{i=M+1}^{N} \langle e_i, h \rangle_H^2 \mathbb{E}[\xi_i^2] = \sum_{i=M+1}^{N} \langle e_i, h \rangle_H^2 \xrightarrow[M,N\to\infty]{} 0
\end{aligned}$$

because $\mathbb{E}[\xi_i^2] = 1$ and $\sum_{i=1}^{\infty} \langle e_i, h \rangle_H^2 = \|h\|_H^2 < +\infty$. This yields that $\{X_N(h) : N \geqslant 1\}$ is a Cauchy sequence in $L^2(\mathbb{P})$. Since $L^2(\mathbb{P})$ is complete, there exists a r.v. $X(h) \in L^2(\mathbb{P})$ such that

$$\mathbb{E}\left[\left(X_N(h) - X(h)\right)^2\right] \xrightarrow[N\to\infty]{} 0.$$

We have:

(i) For all $h \in H$, $\mathbb{E}(X(h)) = 0$.

(ii) For all $h_1, \dots, h_d \in H$, the random vector $(X(h_1), \dots, X(h_d))$ is Gaussian.

(iii) For all $h, g \in H$,

$$\begin{aligned}\mathbb{E}[X(h)X(g)] &= \lim_{N\to\infty} \mathbb{E}[X_N(h)X_N(g)] = \lim_{N\to\infty} \sum_{i=1}^{N}\sum_{j=1}^{N} \langle e_i, h\rangle_H \langle e_j, g\rangle_H \mathbb{E}[\xi_i\xi_j] \\ &= \lim_{N\to\infty} \sum_{i=1}^{N} \langle e_i, h\rangle_H \langle e_i, g\rangle_H = \sum_{i=1}^{\infty} \langle e_i, h\rangle_H \langle e_i, g\rangle_H = \langle h, g\rangle_H. \end{aligned}$$ □

Proposition 1.3. *Let $X = \{X(h) : h \in H\}$ be an isonormal Gaussian process. Then the following assertions are satisfied:*

(i) For all $h, g \in H$, $X(h) \perp\!\!\!\perp X(g) \iff \langle h, g\rangle_H = 0$.

(ii) For all $h, g \in H$, $X(h+g) = X(h) + X(g)$ a.s.

(iii) For all $h \in H$ and $\alpha \in \mathbb{R}$, $X(\alpha h) = \alpha X(h)$.

(iv) Let $G \subset H$ be a subset such that $\operatorname{span} G$ is dense in H. Then for all $h \in H$, there exists a sequence $\{g_n : n \geqslant 1\}$ of elements of $\operatorname{span} G$ such that $X(g_n) \xrightarrow{L^2(\mathbb{P})} X(h)$.

(v) Let H_0 be a closed subspace of H. We define

$$\sigma(H_0) := \sigma\{X(f) : f \in H_0\}.$$

Then for all $h \in H$, we have

$$\mathbb{E}[X(h) \mid \sigma(H_0)] = X(\operatorname{proj}(h|H_0)).$$

Proof. (i) We know that two jointly Gaussian r.v. are independent if and only if their covariance is zero. In addition, $X(h)$ and $X(g)$ are centered, so $\mathbb{E}[X(h)] = 0 = \mathbb{E}[X(g)]$. Therefore we have the following equivalences:

$$\begin{gathered} X(h) \perp\!\!\!\perp X(g) \iff \operatorname{Covar}(X(h), X(g)) = 0 \iff \\ \iff \mathbb{E}[X(h)X(g)] - \mathbb{E}[X(h)]\mathbb{E}[X(g)] = 0 \iff \mathbb{E}[X(h)X(g)] = 0 \iff \langle h, g\rangle_H = 0. \end{gathered}$$

(ii) We compute

$$\begin{aligned} &\mathbb{E}\left[\left\{X(h+g) - (X(h) + X(g))\right\}^2\right] = \\ &\quad = \mathbb{E}\big[X(h+g)^2\big] + \mathbb{E}\big[X(h)^2\big] + \mathbb{E}\big[X(g)^2\big] + 2\mathbb{E}\big[X(h)X(g)\big] - 2\mathbb{E}\big[X(h+g)\{X(h) + X(g)\}\big] \\ &\quad = \|h+g\|^2 + \|h\|^2 + \|g\|^2 + 2\langle h, g\rangle - 2\langle h+g, h\rangle - 2\langle h+g, g\rangle \\ &\quad = \|h+g\|^2 + \|h\|^2 + \|g\|^2 + 2\langle h, g\rangle - 2\|h\|^2 - 2\langle g, h\rangle - 2\langle h, g\rangle - 2\|g\|^2 \\ &\quad = \|h+g\|^2 - \|h\|^2 - \|g\|^2 - 2\langle g, h\rangle = 0. \end{aligned}$$

The last equality holds true because Hilbert space H is supposed to be real.

(iii) Analogous to (ii):

$$\begin{aligned} \mathbb{E}\left[\left\{X(\alpha h) - \alpha X(h)\right\}^2\right] &= \mathbb{E}[X(\alpha h)^2] + \alpha^2\mathbb{E}[X(h)^2] - 2\alpha\mathbb{E}[X(\alpha h)X(h)] \\ &= \|\alpha h\|^2 + \alpha^2\|h\|^2 - 2\alpha\langle \alpha h, h\rangle = 0. \end{aligned}$$

(iv) Let $h \in H$. Since $\operatorname{span} G$ is dense in H, there exists a sequence $\{g_n \mid n \geqslant 1\} \subset \operatorname{span} G$ converging to h in H. Then

$$\mathbb{E}\left[\left\{X(g_n) - X(h)\right\}^2\right] = \mathbb{E}\left[X(g_n - h)^2\right] = \|g_n - h\|_H^2 \xrightarrow{n\to\infty} 0.$$

(v) Since H_0 is a closed subspace of H, any vector $h \in H$ can be written as $h = \mathrm{proj}(h|H_0) + \mathrm{proj}(h|H_0^\perp)$. Then

$$\begin{aligned}\mathbb{E}[X(h) \mid \sigma(H_0)] &= \mathbb{E}[X(\mathrm{proj}(h|H_0)) + X(\mathrm{proj}(h|H_0^\perp)) \mid \sigma(H_0)] \\ &= X(\mathrm{proj}(h|H_0)) + \mathbb{E}[X(\mathrm{proj}(h|H_0^\perp))] \\ &= X(\mathrm{proj}(h|H_0)).\end{aligned}$$

□

1.2 Gaussian measure

We fix a measurable space $(A, \mathcal{A})$, which we assume to be a Polish space[1] with Borel σ-field. In addition, we fix a σ-finite positive measure μ on $(A, \mathcal{A})$ such that $\mu(\{x\}) = 0$ for all $x \in A$.

Definition 1.4. A Gaussian measure over $(A, \mathcal{A})$ with control μ is a centered Gaussian family

$$G = \{G(B) : \mu(B) < +\infty\}$$

such that for all $B, C \in \mathcal{A}$,

$$\mathbb{E}[G(B)G(C)] = \mu(B \cap C).$$

Remarks.

(i) If $B \cap C = \varnothing$, then $G(B) \perp\!\!\!\perp G(C)$.

(ii) $\mathrm{Var}\, G(B) = \mathbb{E}[G(B)^2] - \mathbb{E}[G(B)]^2 = \mu(B)$. Therefore $G(B) \sim \mathcal{N}(0, \mu(B))$.

Proposition 1.5. *Gaussian measures exist.*

Proof. $L^2(\mu)$ is separable, so by Theorem 1.2, there exists an isonormal Gaussian process $X = \{X(f) : f \in L^2(\mu)\}$. Moreover, for all $B \in \mathcal{A}$ with $\mu(B) < +\infty$, we have $\mathbb{1}_B \in L^2(\mu)$. It follows that $G(B) = X(\mathbb{1}_B)$, $\mu(B) < +\infty$, defines a Gaussian measure with control μ, since

$$\mathbb{E}[G(B)G(C)] = \mathbb{E}[X(\mathbb{1}_B)X(\mathbb{1}_C)] = \langle \mathbb{1}_B, \mathbb{1}_C \rangle_{L^2(\mu)} = \int_A \mathbb{1}_B(x)\mathbb{1}_C(x)\mu(\,\mathrm{d}x) = \mu(B \cap C).$$

□

Proposition 1.6. *For any Gaussian measure G with control μ, the following properties are satisfied:*

(i) G is σ-additive, i.e. for any sequence $\{B_i : i \geqslant 1\}$ of disjoint measurable sets such that $\mu(\bigcup_{i=1}^\infty B_i) < +\infty$, we have

$$G\left(\bigcup_{i=1}^\infty B_i\right) = \sum_{i=1}^\infty G(B_i) \quad \text{in } L^2(\mathbb{P}).$$

(ii) For any $B, C \in \mathcal{A}$, $G(B \cup C) = G(B) + G(C) - G(B \cap C)$.

(iii) For any $x \in A$, $G(\{x\}) = 0$ a.s.

1.3 Gaussian measures are not usual measures

Remark. Let $N \sim \mathcal{N}(0, \sigma^2)$. Then $\mathbb{E}(N^4) = 3\sigma^4$.

[1] A Polish space is a separable completely metrisable topological space.

Proof. First, we show that $\mathbb{E}[NP(N)] = \sigma^2\mathbb{E}[P'(N)]$ for any polynomial P. This can be done by integrating by parts:

$$\begin{aligned}
&\int_{\mathbb{R}} xP(x)\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{x^2}{2\sigma^2}\right)\,\mathrm{d}x \\
&\quad = -\frac{1}{\sqrt{2\pi}}\left[\sigma P(x)\exp\left(-\frac{x^2}{2\sigma^2}\right)\right]_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} P'(x)\sigma\exp\left(-\frac{x^2}{2\sigma^2}\right)\,\mathrm{d}x \\
&\quad = \sigma^2\mathbb{E}[P'(N)].
\end{aligned}$$

Now, let $P(x) = x^n$. We obtain $\mathbb{E}[N^{n+1}] = \sigma^2 n\mathbb{E}[N^{n-1}]$. For $n = 3$, this yields $\mathbb{E}[N^4] = \sigma^2 3\mathbb{E}[N^2] = 3\sigma^4$. $\square$

Remark. If G is a Gaussian measure with control μ then $\mathbb{E}[G(B)^4] = 3\mu(B)^2$.

Proof. Apply the preceding remark and use the fact that $\sigma^2(G(B)) = \operatorname{Var} G(B) = \mu(B)$. $\square$

Let B be a measurable set such that $\mu(B) < +\infty$, and let

$$\left\{B_1^{(n)}, \dots, B_{K(n)}^{(n)}\right\}_{n\geqslant 1}$$

be a sequence of measurable partitions of B such that

$$\max_{j=1,\dots,K(n)} \mu\big(B_j^{(n)}\big) \xrightarrow[n\to\infty]{} 0.$$

Then

$$\sum_{j=1}^{K(n)} \mu\big(B_j^{(n)}\big)^2 \xrightarrow[n\to\infty]{} 0,$$

since $\sum_{j=1}^{K(n)} \mu\big(B_j^{(n)}\big)^2 \leqslant \max_{j=1,\dots,K(n)} \mu\big(B_j^{(n)}\big)\mu(B)$.

Proposition 1.7. *Let G be a Gaussian measure with control μ, B a measurable set with finite measure, and $\{B_1^{(n)}, \dots, B_{K(n)}^{(n)}\}_{n\geqslant 1}$ a sequence of measurable partitions of B as above. Then*

$$\sum_{j=1}^{K(n)} G\big(B_j^{(n)}\big)^2 \xrightarrow[L^2(\mathbb{P})]{n\uparrow+\infty} \mu(B).$$

Proof. We compute

$$\begin{aligned}
\mathbb{E}\left[\left(\sum_{j=1}^{K(n)} G(B_j^{(n)})^2 - \mu(B)\right)^2\right] &= \mathbb{E}\left[\left(\sum_{j=1}^{K(n)} G(B_j^{(n)})^2 - \mu(B_j^{(n)})\right)^2\right] \\
&= \sum_{j=1}^{K(n)} \mathbb{E}\left[\left(G(B_j^{(n)})^2 - \mu(B_j^{(n)})\right)^2\right] \\
&= \sum_{j=1}^{K(n)} \left\{3\mu(B_j^{(n)})^2 + \mu(B_j^{(n)})^2 - 2\mu(B_j^{(n)})^2\right\} \\
&= 2\sum_{j=1}^{K(n)} \mu(B_j^{(n)})^2 \xrightarrow{n\uparrow+\infty} 0.
\end{aligned}$$

The passage to the third line follows from the above remark. $\square$

More detailed proof of Proposition 1.7.

$$\begin{aligned}
\mathbb{E}\left[\left(\sum_{j=1}^{K(n)} G(B_j^{(n)})^2-\mu(B)\right)^2\right] &= \mathbb{E}\left[\left(\sum_{j=1}^{K(n)} G(B_j^{(n)})^2-\mu(B_j^{(n)})\right)^2\right]\\
&= \mathbb{E}\left[\left(\sum_{j=1}^{K(n)} G(B_j^{(n)})^2\right)^2 - 2\left(\sum_{j=1}^{K(n)} G(B_j^{(n)})^2\right)\left(\sum_{j=1}^{K(n)} \mu(B_j^{(n)})\right) + \left(\sum_{j=1}^{K(n)} \mu(B_j^{(n)})\right)^2\right]\\
&= \left(\sum_{i,j=1}^{K(n)} \mathbb{E}\left[G(B_i^{(n)})^2 G(B_j^{(n)})^2\right]\right) - 2\left(\sum_{j=1}^{K(n)} \mu(B_j^{(n)})\right)\mu(B)+\mu(B)^2\\
&= \sum_{i=1}^{K(n)} \mathbb{E}\left[G(B_i^{(n)})^4\right] + \sum_{\substack{i,j=1\\ i\neq j}}^{K(n)} \mathbb{E}\left[G(B_i^{(n)})^2 G(B_j^{(n)})^2\right] - \mu(B)^2\\
&= \sum_{i=1}^{K(n)} \mathbb{E}\left[G(B_i^{(n)})^4\right] + \sum_{\substack{i,j=1\\ i\neq j}}^{K(n)} \mathbb{E}\left[G(B_i^{(n)})^2\right]\mathbb{E}\left[G(B_j^{(n)})^2\right] - \mu(B)^2\\
&= \sum_{i=1}^{K(n)} \mathbb{E}\left[G(B_i^{(n)})^4\right] + \sum_{\substack{i,j=1\\ i\neq j}}^{K(n)} \mu(B_i^{(n)})\mu(B_j^{(n)}) - \mu(B)^2\\
&= \sum_{i=1}^{K(n)} 3\mu(B_i^{(n)})^2 + \sum_{\substack{i,j=1\\ i\neq j}}^{K(n)} \mu(B_i^{(n)})\mu(B_j^{(n)}) - \sum_{i,j=1}^{K(n)} \mu(B_i^{(n)})\mu(B_j^{(n)})\\
&= \sum_{i=1}^{K(n)} 3\mu(B_i^{(n)})^2 - \sum_{i=1}^{K(n)} \mu(B_i^{(n)})^2 \xrightarrow{n\uparrow+\infty} 0.
\end{aligned}$$

To pass from the fourth line to the fifth line, we used the fact that $G(B_i^{(n)}) \perp\!\!\!\perp G(B_j^{(n)})$ whenever $i \neq j$. $\square$

1.4 Wiener-Itô integrals for deterministic functions

Let $(A, \mathcal{A}, \mu)$ be a measure space where A is a Polish space, $\mathcal{A}$ the Borel σ-field on A and μ a non-atomic σ-finite measure. We consider a Gaussian measure $G = \{G(B) : \mu(B) < +\infty\}$ with control μ. Our goal is to define the integral "$\int_A f(x)G(\,\mathrm{d}x)$" of a deterministic function f with respect to the Gaussian measure G.

Let us denote by $\mathcal{E}$ the class of simple functions, that is, functions of the type

$$f(x) = \sum_{j=1}^{M} c_j \mathbb{1}_{B_j}(x), \tag{1.1}$$

where the $B_j \in \mathcal{A}$ are measurable sets of finite measure, $c_j \in \mathbb{R}$ and $M \in \{1, 2, \dots\}$. Set $\mathcal{E}$ is dense in $L^2(\mu)$. Indeed, if $g \perp \mathcal{E}$, then $\int_B g(x)\mu(\,\mathrm{d}x) = 0$ for any measurable set B of finite measure, which together with the σ-finiteness of μ implies that $g = 0$ a.e.-μ.

Definition 1.8. For a simple function $f \in \mathcal{E}$ of the form (1.1), we define

$$\int_A f(x)G(\,\mathrm{d}x) := \sum_{j=1}^{M} c_j G(B_j).$$

We may also use the shorthand notation

$$\int f\,\mathrm{d}G = \int_A f(x)G(\,\mathrm{d}x).$$

Proposition 1.9. *(i) For any simple functions $f, g \in \mathcal{E}$, we have* $\mathbb{E}\left[\int f\,\mathrm{d}G \times \int g\,\mathrm{d}G\right] = \langle f, g\rangle_{L^2(\mu)}$.

(ii) If $\{f_n\}_{n=1}^{\infty} \subset \mathcal{E}$ is a Cauchy sequence in $L^2(\mu)$, then $\{\int f_n\,\mathrm{d}G\}_{n=1}^{\infty}$ is a Cauchy sequence in $L^2(\mathbb{P})$.

Proof. (i) Let f and g be simple functions defined by

$$f(x) = \sum_{j=1}^{M_1} c_j^{(1)} \mathbb{1}_{B_j^{(1)}}(x), \qquad g(x) = \sum_{j=1}^{M_2} c_j^{(2)} \mathbb{1}_{B_j^{(2)}}(x).$$

Without loss of generality, we can assume that $B^{(1)} = B^{(2)} =: B_j$, $M_1 = M_2 =: M$ and $B_j \cap B_\ell = \varnothing$ whenever $j \neq \ell$. Then

$$\mathbb{E}\left[\int f\,\mathrm{d}G \times \int g\,\mathrm{d}G\right] = \mathbb{E}\left[\left(\sum_{j=1}^{M} c_j^{(1)} G(B_j)\right)\left(\sum_{j=1}^{M} c_j^{(2)} G(B_j)\right)\right] = \sum_{j=1}^{M} c_j^{(1)} c_j^{(2)} \mu(B_j) = \langle f, g\rangle_{L^2(\mu)}.$$

The second equality follows from the fact that $G(B_j)$ and $G(B_\ell)$ are independent r.v. for $j \neq \ell$ and hence $\mathbb{E}[G(B_j)G(B_\ell)] = \mathbb{E}[G(B_j)]\mathbb{E}[G(B_\ell)] = 0$.

(ii) Let $\{f_n\}_{n=1}^{\infty} \subset \mathcal{E}$ be a Cauchy sequence in $L^2(\mu)$. Since $\mathcal{E}$ is a vector space, $f_n - f_m$ is an element of $\mathcal{E}$ for all $n, m \in \mathbb{N}$. We can therefore apply point (i) to $f_n - f_m$, which yields

$$\mathbb{E}\left[\left(\int f_n\,\mathrm{d}G - \int f_m\,\mathrm{d}G\right)^2\right] = \mathbb{E}\left[\left(\int f_n - f_m\,\mathrm{d}G\right)^2\right] = \|f_n - f_m\|_{L^2(\mu)}^2.$$

Thus $\{\int f_n\,\mathrm{d}G\}_{n=1}^{\infty}$ is a Cauchy sequence in $L^2(\mathbb{P})$. $\square$

Definition 1.10. Let $f \in L^2(\mu)$. Then there exists a sequence $\{f_n\}_{n=1}^{\infty} \subset \mathcal{E}$ of simple functions converging to f in $L^2(\mu)$. So from point (ii) of the preceding proposition, $\{\int f_n\,\mathrm{d}G\}_{n=1}^{\infty}$ is a Cauchy sequence in $L^2(\mathbb{P})$. We therefore *define*

$$\int_A f(x)G(\mathrm{d}x) = \int f\,\mathrm{d}G := \overset{L^2(\mathbb{P})}{\lim_{n\to+\infty}} \int f_n\,\mathrm{d}G.$$

Theorem 1.11. *(i) The definition of $\int f\,\mathrm{d}G$ is well-given.*

(ii) The class $\left\{\int_A f\,\mathrm{d}G : f \in L^2(\mu)\right\}$ is an isonormal Gaussian process over $L^2(\mu)$.

Proof. (i) Let $\{f_n\}_{n=1}^{\infty}, \{f_n'\}_{n=1}^{\infty} \subset \mathcal{E}$ be two sequences of simple functions converging to f in $L^2(\mu)$. Then

$$\mathbb{E}\left[\left(\int f_n\,\mathrm{d}G - \int f_n'\,\mathrm{d}G\right)^2\right]^{1/2} = \|f_n - f_n'\|_{L^2(\mu)} \leqslant \|f_n - f\|_{L^2(\mu)} + \|f_n' - f\|_{L^2(\mu)} \xrightarrow{n\uparrow+\infty} 0.$$

(ii) The class $\left\{\int_A f\,\mathrm{d}G : f \in L^2(\mu)\right\}$ is centered and jointly Gaussian. This follows from the fact that the L^2-limit of Gaussian r.v. is again a Gaussian r.v. Now, let f, g be two functions in $L^2(\mu)$, and let $\{f_n\}_{n=1}^{\infty}, \{g_n\}_{n=1}^{\infty} \subset \mathcal{E}$ be sequences of simple functions such that $f_n \to f$ and $g_n \to g$ in $L^2(\mu)$. Then

$$\mathbb{E}\left[\int f\,\mathrm{d}G \times \int g\,\mathrm{d}G\right] = \lim_{n\to+\infty} \mathbb{E}\left[\int f_n\,\mathrm{d}G \times \int g_n\,\mathrm{d}G\right] = \lim_{n\to+\infty} \langle f_n, g_n\rangle_{L^2(\mu)} = \langle f, g\rangle_{L^2(\mu)}.$$

Note that the interchanges of integration and taking limits in the last line are allowed by continuity of scalar product in any Hilbert space. $\square$

Definition 1.12. Let $f \in L^2(\mu)$. The (Gaussian) random variable

$$\int_A f(x) G(\mathrm{d}x)$$

is the "Wiener-Itô integral of f with respect to G". The relation

$$\mathbb{E}\left[\int f \,\mathrm{d}G \times \int g \,\mathrm{d}G\right] = \langle f, g \rangle_{L^2(\mu)}$$

is the "Wiener-Itô isometry".

Remark. From the properties of isonormal Gaussian processes (see Proposition 1.3), we deduce that

(i) $\int (\alpha f + g) \,\mathrm{d}G = \alpha \int f \,\mathrm{d}G + \int g \,\mathrm{d}G$, for any $\alpha \in \mathbb{R}$ and $f, g \in L^2(\mu)$,
(ii) $\int fg \,\mathrm{d}\mu = 0 \iff \int f \,\mathrm{d}G \perp\!\!\!\perp \int g \,\mathrm{d}G$,
(iii) $\mathbb{E}\left[\int f \,\mathrm{d}G \;\middle|\; \sigma\{\int h \,\mathrm{d}G : h \in H_0\}\right] = \int \mathrm{proj}(f|H_0) \,\mathrm{d}G$, for any closed subspace H_0 of $L^2(\mu)$.

Chapter 2

Brownian Motion

2.1 Definition and first properties

We fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition 2.1. Let I be either $\mathbb{R}_+ = [0, \infty)$ or $[0, T]$ with $T < +\infty$. A standard Brownian motion, or Wiener process, on I is a Gaussian process $W = \{W_t : t \in I\}$ such that

(a) $W_0 = 0$ a.s.-$\mathbb{P}$,
(b) $\mathbb{E}[W_t] = 0$ for all $t \in I$,
(c) $\mathbb{E}[W_s W_t] = s \wedge t$ for all $s, t \in T$,
(d) The mapping $t \mapsto W_t \colon I \to \mathbb{R}$ is continuous with probability 1.

Remark. Property (d) means that there exists a measurable set $D \in \mathcal{F}$ with $\mathbb{P}(D) = 1$ such that the map $t \mapsto W_t(\omega)$ is continuous for all $\omega \in D$.

A Gaussian process satisfying (a), (b) and (c) but not necessarily (d) is called a pre-brownian motion.

Proposition 2.2. *There exists a pre-brownian motion over I.*

Proof. There exists a Gaussian measure G over I with control given by the Lebesgue measure μ. Then $W_t := G([0, t])$ is a pre-brownian motion. Indeed, $W_0 = G(\{0\}) = 0$, $\mathbb{E}[W_t] = 0$ and $\mathbb{E}[W_s W_t] = \mathbb{E}[G([0, s])G([0, t])] = \mu([0, s] \cap [0, t]) = s \wedge t$. $\square$

Remark. It follows from properties (b) and (c) that $W_t \sim \mathcal{N}(0, t)$ for every $t \in I$.

Proposition 2.3. *Let I be a fixed interval as above. The following assertions are equivalent:*

(i) $W = \{W_t : t \in I\}$ is a pre-Brownian motion.
(ii) $W_0 = 0$ a.s. and for all $t > s \geqslant u \geqslant 0$, $W_t - W_s \sim \mathcal{N}(0, t - s)$ and $(W_t - W_s) \perp\!\!\!\perp W_u$.
(iii) $W_0 = 0$ a.s. and for all $0 < t_1 < t_2 < \cdots < t_n$,

$$\left(W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_n} - W_{t_{n-1}}\right) \sim \mathcal{N}\left(0, \begin{pmatrix} t_1 & 0 & \cdots & 0 & 0 \\ 0 & t_2 - t_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & t_{n-1} - t_{n-2} & 0 \\ 0 & 0 & \cdots & 0 & t_n - t_{n-1} \end{pmatrix}\right).$$

(iv) $W_0 = 0$ a.s. and for all $0 < t_1 < t_2 < \cdots < t_n$, the vector

$$\left(W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_n} - W_{t_{n-1}}\right)$$

has a density given by

$$f(x_1,\dots,x_n) = \frac{1}{(2\pi)^{n/2}\sqrt{t_1(t_2-t_1)\dots(t_n-t_{n-1})}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\frac{(x_i-x_{i-1})^2}{t_1-t_{i-1}}\right)$$

with $x_0 \equiv 0$.

Remark. A pre-Brownian motion has "stationary increments", that is, for any $t,h \geqslant 0$, the r.v. $W_{t+h} - W_t$ follows the same law as $W_h - W_0 = W_h \sim \mathcal{N}(0,h)$.

Proposition 2.4. *Let* $\{W_t : t \in \mathbb{R}_+\}$ *be a standard Brownian motion.*
(i) The process $\{-W_t : t \geqslant 0\}$ *is also a standard Brownian motion.*
(ii) For any $c > 0$, *the process* $\{\frac{1}{c}W_{tc^2} : t \geqslant 0\}$ *is also a standard Brownian motion. This means that* W *is a self-similar process (or random fractal).*
(iii) (Weak Markov property) For any $u > 0$, *the process* $\{W_{u+t} - W_u : t \geqslant 0\}$ *is again a standard Brownian motion, independent of* $\sigma\{W_s : s \leqslant u\}$.
(iv) (Time reversal) For any $T > 0$, *the process* $\{W_{T-t} - W_T : t \in [0,T]\}$ *is a Brownian motion on* $[0,T]$.

Proof. The processes in (i), (ii), (iii) and (iv) are all Gaussian, centered and continuous, so all we have to do is to check covariances:
(i) $\mathbb{E}[(-W_s)(-W_t)] = s \wedge t$.
(ii) $\mathbb{E}[(\frac{1}{c}W_{sc^2})(\frac{1}{c}W_{tc^2})] = \frac{1}{c^2}(c^2 s \wedge c^2 t) = s \wedge t$.
(iii) $\mathbb{E}[(W_{u+s} - W_u)(W_{u+t} - W_u)] = u + s \wedge t - u - u + u = s \wedge t$. Moreover, for all $s \leqslant u$, $\mathbb{E}[(W_{u+t} - W_u)W_s] = s - s = 0$, so $\{W_{u+t} - W_u\}_{t \geqslant 0}$ is independent of W_s for all $s \leqslant u$.
(iv) $\mathbb{E}[(W_{T-s} - W_T)(W_{T-t} - W_T)] = T - s \vee t - (T-s) - (T-t) + T = s \wedge t$. ☐

2.2 Construction of Brownian motion (Lévy-Ciesielski, 1957)

In this section, we shall prove the following result:

Theorem 2.5. *On some probability space* $(\Omega, \mathcal{F}, \mathbb{P})$, *there exists a standard Brownian motion* $\{W_t : t \geqslant 0\}$.

To prove this theorem, we need five lemmata.

Lemma 2.6 (Borel-Cantelli). *Let* $\{A_n\}_{n \geqslant 1}$ *be a sequence of events. Set*

$$\limsup_n A_n := \{A_n \text{ infinitely often}\} = \bigcap_{k \geqslant 1} \bigcup_{n \geqslant k} A_n.$$

Then

$$\sum_{n \geqslant 1} \mathbb{P}(A_n) < +\infty \quad \text{implies} \quad \mathbb{P}(\limsup_n A_n) = 0.$$

Proof. For all k, we have $\mathbb{P}(\limsup_n A_n) \leqslant \mathbb{P}(\bigcup_{n \geqslant k} A_n) \leqslant \sum_{n \geqslant k} \mathbb{P}(A_n)$. By assumption, the last expression goes to zero for $k \to \infty$. ☐

Lemma 2.7. *The set* $C_{[0,1]}$ *of continuous functions on* $[0,1]$ *is a Banach space with respect to the supremum norm* $\|f\|_\infty := \sup_{t \in [0,1]} |f(t)|$.

Proof. Omitted. ☐

Lemma 2.8. *Let $\{f_k\}_{k\geqslant 1} \subset C_{[0,1]}$ be a sequence of continuous functions on $[0,1]$, and let $F_n := \sum_{k=1}^{n}$ for all $n \geqslant 1$. Assume that there exists a sequence $\{b_k\}_{k\geqslant 1} \subset \mathbb{R}_+^*$ of strictly positive real numbers with $\sum_{k=1}^{\infty} b_k < \infty$. Then $\lim_{n\to\infty} F_n = F$ in $C_{[0,1]}$ for some $F \in C_{[0,1]}$.*

Proof. For all $N > M$, we have $\|F_N - F_M\|_\infty \leqslant \sum_{k=M+1}^{N} b_k \xrightarrow{M,N\uparrow+\infty} 0$, so that $\{F_n\}_{n\geqslant 1}$ is a Cauchy sequence. □

Lemma 2.9. *If $Z \sim \mathcal{N}(0,1)$, then for all $c > 0$, we have*

$$\mathbb{P}(|Z| > c) \leqslant \frac{2}{\sqrt{2\pi}} \frac{1}{c} e^{-c^2/2}.$$

Proof. We compute

$$\mathbb{P}(|Z| > c) = 2\mathbb{P}(Z > c) = 2\frac{1}{\sqrt{2\pi}} \int_c^\infty e^{-y^2/2}\,\mathrm{d}y \leqslant \sqrt{\frac{2}{\pi}} \frac{1}{c} \int_c^\infty y e^{-y^2/2}\,\mathrm{d}y = \sqrt{\frac{2}{\pi}} \frac{1}{c} e^{-c^2/2}.$$

The inequality in this computation is due to the fact that $\frac{y}{c} \geqslant 1$ for $y \in [c, \infty)$. □

Lemma 2.10. *We introduce the "Haar system"*

$$\mathbb{H} = \{f_0, f_{n,j} : n \geqslant 1,\ j = 1, \dots, 2^{n-1}\}$$

defined as $f_0 \equiv 1$, and setting $k = 2j-1$,

$$f_{n,j}(t) = \begin{cases} 2^{(n-1)/2} & \text{if } t \in [\frac{k-1}{2^n}, \frac{k}{2^n}), \\ -2^{(n-1)/2} & \text{if } t \in (\frac{k}{2^n}, \frac{k+1}{2^n}], \\ 0 & \text{elsewhere.} \end{cases}$$

Then $\mathbb{H}$ is a complete orthonormal system (or ONB) of $L^2([0,1], \mathrm{d}t)$, where $\mathrm{d}t$ stands for the Lebesgue measure.

Proof. See handwritten notes. □

We introduce also the "Schauder functions":

$$\mathbb{S} = \{F_0, F_{j,n} : n \geqslant 1,\ j = 1, \dots, 2^{n-1}\},$$

where

$$F_0(t) := \int_0^t f_0(x)\,\mathrm{d}x = t$$

and, putting $k = 2j-1$,

$$F_{j,n}(t) = \int_0^t f_{n,j}(x)\,\mathrm{d}x = \begin{cases} 2^{(n-1)/2}(t - \frac{k-1}{2^n}) & \text{if } t \in [\frac{k-1}{2^n}, \frac{k}{2^n}), \\ 2^{(n-1)/2}(\frac{k+1}{2^n} - t) & \text{if } t \in (\frac{k}{2^n}, \frac{k+1}{2^n}], \\ 0 & \text{elsewhere.} \end{cases}$$

Proof of Theorem 2.5. See handwritten notes. □

2.3 Hölder continuity and the Kolmogorov-Čentsov criterion

Definition 2.11. Let X and $\tilde{X}$ be two stochastic processes on $I = [0, T]$ or $\mathbb{R}_0$.

(i) We say that X is a modification of $\tilde{X}$ if for all $t \in I$, $\mathbb{P}\{X_t = \tilde{X}_t\} = 1$.

(ii) We say that X and $\tilde{X}$ are indistinguishable if there exists a measurable set D with $\mathbb{P}(D) = 1$ such that $D \subset \{X_t = \tilde{X}_t \text{ for all } t \in I\}$. That is, if a.s.-$\mathbb{P}$, "$X_t = \tilde{X}_t, \forall t \in I$".

Remark. Condition (ii) implies condition (i), but the converse is false in general (because the index set I may be uncountable). For instance, let $I = [0, 1]$, $X_t \equiv 0$, $U \sim \mathbb{U}_{[0,1]}$ and $\tilde{X}_t = \mathbb{1}_{t=U}$ for $t \in [0, 1]$. Then

$$\forall t \in [0,1], \quad \mathbb{P}(X_t = \tilde{X}_t) = \mathbb{P}(U \neq t) = 1,$$

but, since $\{X_t = \tilde{X}_t,\, \forall t \in [0, 1]\} = \{U \notin [0, 1]\}$,

$$\mathbb{P}(X_t = \tilde{X}_t,\, \forall t \in [0,1]) = \mathbb{P}(U \notin [0,1]) = 0.$$

However, we have

Lemma 2.12. *Let X and $\tilde{X}$ have a.s. continuous paths. Then, if X and $\tilde{X}$ are modifications, they are indistinguishable.*

Proof. X and $\tilde{X}$ have a.s. continuous paths, so there exist measurable sets D_1 and D_2 with $\mathbb{P}(D_1) = \mathbb{P}(D_2) = 1$ such that

$$D_1 \subset \{\omega : X(\omega) \text{ is continuous}\}, \qquad D_2 \subset \{\omega : \tilde{X}(\omega) \text{ is continuous}\}.$$

Let $B = \bigcap_{t \in I \cap \mathbb{Q}} \{X_t = \tilde{X}_t\}$, so that $\mathbb{P}(B) = 1$. Writing $D := D_1 \cap D_2 \cap B$, one has $\mathbb{P}(D) = 1$ and, by densitiy of $\mathbb{Q}$ and by continuity,

$$D \subset \{X_t = \tilde{X}_t \text{ for all } t \in I\}. \qquad \square$$

Remark. It is easy to see that

$$\left.\begin{array}{l} X \text{ is a modification of } Y \\ Y \text{ is a modification of } Z \end{array}\right\} \Longrightarrow X \text{ is a modification of } Z$$

and

$$\left.\begin{array}{l} X \text{ is indistinguishable of } Y \\ Y \text{ is indistinguishable of } Z \end{array}\right\} \Longrightarrow X \text{ is indistinguishable of } Z.$$

This means that both of these relations are equivalence relations.

Theorem 2.13 (Kolmogorov-Čentsov)**.** *Let $I = \mathbb{R}_+$ or $[0, T]$. Let $\{X_t : t \in I\}$ be an $\mathbb{R}^d$-valued stochastic process such that for all $[0, S] \subset I$, there exist $C, \alpha, \beta > 0$ (with C possibly depending on S) such that*

$$\mathbb{E}\left[\|X_t - X_s\|_{\mathbb{R}^d}^{\alpha}\right] \leqslant C|t-s|^{1+\beta}, \qquad \forall t, s \in [0, S].$$

Then there exists a modification $\tilde{X}$ of X such that for any $\omega \in \Omega$, the map $t \mapsto \tilde{X}_t(\omega)$ is locally γ-Hölder continuous for all $\gamma < \frac{\beta}{\alpha}$. That is,

$$\forall [0,S] \subset I,\ \forall \gamma < \tfrac{\beta}{\alpha},\ \forall \omega \in \Omega,\ \exists C_{S,\gamma}(\omega) \quad \text{such that} \quad \|\tilde{X}_t(\omega) - \tilde{X}_s(\omega)\|_{\mathbb{R}^d} \leqslant C_{S,\gamma}(\omega)\|t-s\|^{\gamma}.$$

This modification is unique up to indistinguishability.

Proof. Omitted for the moment. $\square$

Corollary 2.14. *Let $\{W_t : t \in \mathbb{R}_+\}$ be a standard Brownian motion. Then the paths of W are almost surely Hölder continuous (locally), for any $\gamma < \frac{1}{2}$.*

Proof. See handwritten notes. □

Example (1). Let $\{W_t : t \geqslant 0\}$ be a standard Brownian motion. Set

$$X_t := \begin{cases} 0 & \text{if } t = 0, \\ tW_{1/t} & \text{if } t > 0. \end{cases}$$

Then X is a standard Brownian motion on $\mathbb{R}_+$. Indeed, for all $t, s \in \mathbb{R}_+$, we have $\mathbb{E}[X_t] = 0$ and $\mathbb{E}[X_t X_s] = ts\mathbb{E}[W_{1/t} W_{1/s}] = \frac{ts}{t \vee s} = t \wedge s$, and the paths of X are almost surely continuous on $(0, \infty)$. We have to prove continuity at 0. The key fact is that

$$\{X_t : t \in (0, \infty) \cap \mathbb{Q}\} \overset{\text{Law}}{=} \{W_t : t \in (0, \infty) \cap \mathbb{Q}\}$$

(this follows from the fact that both sides are countable families of Gaussian r.v. with the same covariances). In particular,

$$\mathbb{P}\{\lim_{\substack{t \downarrow 0 \\ t \in \mathbb{Q}}} X_t = 0\} = \mathbb{P}\{\lim_{\substack{t \downarrow 0 \\ t \in \mathbb{Q}}} W_t = 0\} = 1.$$

By continuity,

$$\lim_{\substack{t \downarrow 0 \\ t \in \mathbb{Q}}} X_t = \lim_{t \downarrow 0} X_t,$$

proving the statement.

Example (2). We can build a pre-brownian motion with discontinuous paths. Let $\{W_t : t \in [0,1]\}$ be a Brownian motion, and let $U \sim \mathbb{U}_{[0,1]}$ be independent of W. Define

$$\tilde{W}_t := W_t + \mathbb{1}_{t=U}.$$

$\tilde{W}$ is not a Brownian motion (since discontinuous), but it is a pre-brownian motion. To show this, we just have to prove that $\tilde{W}$ is a modification of W; indeed, for all t,

$$\mathbb{P}(W_t = \tilde{W}_t) = \mathbb{P}(U \neq t) = 1.$$

Remark. If X and $\tilde{X}$ are modifications then they have the same finite-dimensional distributions, i.e. for any integer $d \geqslant 1$, we have

$$(X_{t_1}, \dots, X_{t_d}) \overset{\text{Law}}{=} (\tilde{X}_{t_1}, \dots, \tilde{X}_{t_d}) \qquad (\forall t_1, \dots, t_d \in I).$$

Example (3). One consequence of Example (1) is that, almost surely,

$$\lim_{t \to \infty} \frac{W_t}{t} = 0$$

(since $\frac{W_t}{t} = X_{1/t}$). This is the strong law of large numbers for Brownian motions. We have in fact

$$\frac{W_N}{N} = \frac{\sum_{i=1}^{N} (W_i - W_{i-1})}{N}.$$

2.4 "Canonical" construction of Brownian motion

Fact $(*)$. *Let $C(\mathbb{R}_+, \mathbb{R})$ be the set of continuous real-valued functions on $\mathbb{R}_+$, endowed with the topology of uniform convergence on compacts. The Borel σ-field of $C(\mathbb{R}_+, \mathbb{R})$, denoted by $\mathcal{C}$, is generated by "cylindrical sets", which are by definition of the form*

$$A = \{\omega = \{\omega(t) : t \geqslant 0\} : \omega(t_1) \in B_1, \dots, \omega(t_n) \in B_n\}, \qquad B_1, \dots, B_n \in \mathcal{B}(\mathbb{R}).$$

[For instance, if $t_1 = 0$, $t_2 = 2\pi$ and $B_1 = B_2 = \{0\}$, then $\sin \in \{\omega : \omega(t_1) \in B_1, \omega(t_2) \in B_2\}$.]

By using e.g. "a monotone class argument" (see the tutorial on the webpage), we can show the following result:

Fact $(**)$. *If $\mathbb{P}$ and $\mathbb{Q}$ are two probability measures on $(C(\mathbb{R}_+, \mathbb{R}), \mathcal{C})$ such that $\mathbb{P}(A) = \mathbb{Q}(A)$ for all cylindrical sets A, then $\mathbb{P}(B) = \mathbb{Q}(B)$ for all $B \in \mathcal{C}$.*

Fact $(***)$. *Let $W = \{W_t : t \geqslant 0\}$ be a standard Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$, then the set function*

$$\mathbb{W}(B) := \mathbb{P}\{W \in B\} \qquad (B \in \mathcal{C})$$

defines a probability measure on $(C(\mathbb{R}_+, \mathbb{R}), \mathcal{C})$ and due to $(**)$ the definition of $\mathbb{W}$ is independent of the choice of the Brownian motion.

Definition 2.15. (i) The measure $\mathbb{W}$ is called the Wiener measure on $(C(\mathbb{R}_+, \mathbb{R}), \mathcal{C})$.

(ii) The probability space $(C(\mathbb{R}_+, \mathbb{R}), \mathcal{C}, \mathbb{W})$ is the canonical space.

(iii) The process $X = \{X_t : t \geqslant 0\}$ defined as

$$X_t(\omega) = \omega(t), \qquad t \geqslant 0, \forall \omega \in C(\mathbb{R}_+, \mathbb{R})$$

is called the canonical process.

Proposition 2.16. *On the canonical space, the canonical process is a standard Brownian motion.*

Proof. Exercise. □

2.5 Donsker theorem and universality (invariance principles)

Recall that if $\{\xi_k : k \geqslant 1\}$ is a sequence of i.i.d. r.v. with $\mathbb{E}[\xi_1] = 0$ and $\mathbb{E}[\xi_1^2] = 1$ and if $S_n = \sum_{k=1}^n \xi_k$, then the central limit theorem says that

$$\frac{1}{\sqrt{n}} S_n \xrightarrow{\text{Law}} Z \sim \mathcal{N}(0,1),$$

which means that for all $t \in \mathbb{R}$,

$$\lim_{n\to\infty} \mathbb{P}(S_n/\sqrt{n} \leqslant t) = \mathbb{P}(Z \leqslant t).$$

Now, for any $n \in \mathbb{N}$ we can consider the interpolated process on $[0,1]$ defined as

$$X_n(t) := \frac{1}{\sqrt{n}} \left\{ S_{[nt]} + (nt - [nt])\xi_{[nt]+1} \right\}.$$

Then $X_n(1) = S_n$, and in general, for all $j = 0, \dots, n$, we have $X_n(\frac{j}{n}) = \frac{1}{\sqrt{n}} S_j$, with $S_0 = 0$.

Theorem 2.17 (Donsker)**.** *For all $n \in \mathbb{N}$, $X_n = \{X_n(t) : t \in [0,1]\}$ is a continuous process, starting from zero. Moreover, X_n converges in law to W, where $W = \{W_t : t \in [0,1]\}$ is a standard Brownian motion. This means that for any continuous and bounded function $\varphi\colon C([0,1], \mathbb{R}) \to \mathbb{R}$, we have*

$$\lim_{n\to\infty} \mathbb{E}[\varphi(X_n)] = \mathbb{E}[\varphi(W)].$$

2.6 Behaviour of Brownian motion around zero

Let $W = \{W_t : t \geqslant 0\}$ be a standard Brownian motion started from 0, on $(\Omega, \mathcal{F}, \mathbb{P})$. We define the filtration of W to be

$$\mathcal{F}_t = \sigma\{W_u : u \leqslant t\} \qquad (\forall t \geqslant 0).$$

Note that $\mathcal{F}_s \subset \mathcal{F}_t$ whenever $s < t$.

Proposition 2.18 (Blumental's 0-1 law, la loi de tout ou de rien)**.** *Consider the σ-field*

$$\mathcal{F}_{0+} := \bigcap_{\varepsilon > 0} \mathcal{F}_\varepsilon.$$

If $A \in \mathcal{F}_{0+}$, then either $\mathbb{P}(A) = 0$ or $\mathbb{P}(A) = 1$.

Proof. Fix $A \in \mathcal{F}_{0+}$. Let $0 < t_1 < t_2 < \cdots < t_n$ and let $f \colon \mathbb{R}^n \to \mathbb{R}$ be continuous and bounded. Then

$$\begin{aligned}
\mathbb{E}[\mathbb{1}_A f(W_{t_1}, \dots, W_{t_n})] &= \lim_{\varepsilon \downarrow 0} \mathbb{E}[\mathbb{1}_A f(W_{t_1} - W_\varepsilon, \dots, W_{t_n} - W_\varepsilon)] \\
&= \lim_{\varepsilon \downarrow 0} \mathbb{P}(A) \mathbb{E}[f(W_{t_1} - W_\varepsilon, \dots, W_{t_n} - W_\varepsilon)] \\
&= \mathbb{P}(A) \mathbb{E}[f(W_{t_1}, \dots, W_{t_n})]
\end{aligned}$$

(the second equality follows from the fact that $W_{t_j} - W_\varepsilon \perp\!\!\!\perp A$ whenever $\varepsilon < t_1$, since $A \in \mathcal{F}_{\varepsilon/2}$). Since r.v.'s of the type $f(W_{t_1}, \dots, W_{t_n})$ generate $\sigma\{W_t : t > 0\}$, we deduce that $\mathcal{F}_{0+}$ is independent of $\sigma\{W_t : t > 0\} \vee \sigma\{W_0\}$, so $\mathcal{F}_{0+} \perp\!\!\!\perp \sigma\{W_t : t \geqslant 0\}$, and therefore $\mathcal{F}_{0+} \perp\!\!\!\perp \mathcal{F}_{0+}$, and the conclusion follows. ☐

Corollary 2.19. *Almost surely, for any $\varepsilon > 0$, we have $\sup_{t \in [0,\varepsilon]} W_t > 0$ and $\inf_{t \in [0,\varepsilon]} W_t < 0$.*

Proof. By continuity, we can focus on $\sup_{t \in [0,\varepsilon] \cap \mathbb{Q}} W_t$ and $\inf_{t \in [0,\varepsilon] \cap \mathbb{Q}} W_t$. Let $\{\varepsilon_p\}_{p=1}^\infty \subset \mathbb{Q}$ be a sequence such that $\varepsilon_p \downarrow \varepsilon$. Then

$$A := \{\sup_{t \in [0,\varepsilon] \cap \mathbb{Q}} W_t > 0\} = \bigcap_{p \geqslant 1} \{\sup_{t \in [0,\varepsilon_p] \cap \mathbb{Q}} W_t > 0\} \in \mathcal{F}_{0+}.$$

Then, using the continuity from above of probability measures, we get

$$\mathbb{P}(A) = \lim_{p \to \infty} \mathbb{P}\big(\sup_{t \in [0,\varepsilon_p] \cap \mathbb{Q}} W_t > 0\big) \geqslant \lim_{p \to \infty} \underbrace{\mathbb{P}(W_{\varepsilon_p} > 0)}_{=\frac{1}{2}} = \frac{1}{2}.$$

By Blumental's 0-1 law, $\mathbb{P}(A) = 1$. To deal with the infimum, replace W with $-W$ (which is again a Brownian motion). ☐

Corollary 2.20. *For $a \in \mathbb{R}$, let $T_a := \inf\{t \geqslant 0 : W_t = a\}$ be the hitting time of a (with the convention $\inf \varnothing = +\infty$). Then for all $a \in \mathbb{R}$, $\mathbb{P}(T_a < +\infty) = 1$.*

Proof. We deal with $a > 0$ (and implicitly we take suprema with rationals). We have to prove that for all $A > 0$, $\mathbb{P}(\sup_{t \geqslant 0} W_t > A) = 1$. We have

$$\begin{aligned}
&\mathbb{P}(\sup_{t \geqslant 0} W_t > A) = \mathbb{P}(\sup_{t \geqslant 0} \frac{1}{A} W_t > 1) = \mathbb{P}(\sup_{t \geqslant 0} \frac{1}{A} W_{tA^2} > 1) \overset{*}{=} \\
&\overset{*}{=} \mathbb{P}(\sup_{t \geqslant 0} W_t > 1) \overset{\circledast}{=} \lim_{\delta \downarrow 0} \mathbb{P}(\sup_{t \in [0, \frac{1}{\delta^2}]} W_t > 1) \overset{*}{=} \lim_{\delta \downarrow 0} \mathbb{P}(\sup_{t \in [0, \frac{1}{\delta^2}]} \frac{1}{\delta} W_{t\delta^2} > 1) = \\
&= \lim_{\delta \downarrow 0} \mathbb{P}(\sup_{t \in [0,1]} W_t > \delta) \overset{\circledast}{=} \mathbb{P}(\sup_{t \in [0,1]} W_t > 0) = 1,
\end{aligned}$$

where equalities marked with $*$ hold by the scaling property of Brownian motion and equalities marked with $\circledast$ are due to the continuity of $\mathbb{P}$. The last equality follows from Corollary 2.19. To deal with $a < 0$, replace W with $-W$. $\square$

Corollary 2.21. *With probability 1,* $\limsup_{T\to+\infty} W_T = +\infty$ *and* $\liminf_{T\to+\infty} W_T = -\infty$.

In other words, "fluctuations become more and more pronounced as time advances".

Corollary 2.22. *With probability 1, a standard Brownian motion W is nowhere monotone, that is: with probability 1, for all $t \in \mathbb{R}_+$ the following property holds:*

$$P(t): \qquad \forall \eta > 0, \text{ the mapping } t \mapsto W_t \text{ is not monotone on } [t, t+\eta].$$

Proof. Fix $\in \mathbb{R}$. By the weak Markov property, $\{W_{t+s} - W_s : s \geqslant 0\}$ is a standard Brownian motion (independent of $\mathcal{F}_t$). So, by Corollary 2.19, with probability 1, we have for all $\eta > 0$

$$\sup_{s\in[0,\eta]} \{W_{s+t} - W_t\} > 0 \quad \text{and} \quad \inf_{s\in[0,\eta]} \{W_{s+t} - W_t\} < 0,$$

which implies that t has property $P(t)$ (with probability 1). So with probability 1, every $t \in \mathbb{Q}\cap\mathbb{R}_+$ has property $P(t)$ (just an intersection of countably many events of probability 1). Hence, by density of $\mathbb{Q}$ in $\mathbb{R}$, with probability 1 every $t \in \mathbb{R}_+$ has property $P(t)$. $\square$

The results given in the remainder of this section provide some information about maxima of a Brownian motion W. To obtain analogous results for minima, replace W with $-W$.

Corollary 2.23. *Given two non-overlapping (i.e. with disjoint interiors) closed intervals $[a_1, b_1]$ and $[a_2, b_2]$, ($a_1 < b_1 \leqslant a_2 < b_2$), then with probability 1,*

$$m_1 := \max_{t\in[a_1,b_1]} W_t \neq \max_{t\in[a_2,b_2]} W_t =: m_2.$$

Proof. By the weak Markov property, $W_{a_2} < m_2$, meaning that $m_2 = \max_{t\in[a_2+1/n,b_2]} W_t$ for some $n \geqslant 1$, so we can assume without loss of generality that $b_1 < a_2$. Again by the weak Markov property,

$$\begin{aligned} &(1) \qquad && W_{a_2} - W_{b_1} \perp\!\!\!\perp W_{b_1} - m_1, \\ &(2) && m_2 - W_{a_2} \perp\!\!\!\perp W_{a_2} - W_{b_1}, \\ &(3) && m_2 - W_{a_2} \perp\!\!\!\perp W_{b_1} - m_1. \end{aligned}$$

In addition, we have the equivalence

$$m_1 = m_2 \iff W_{a_2} - W_{b_1} = W_{a_2} - m_2 + m_1 - W_{b_1}.$$

Let $Y := W_{a_2} - m_2 + m_1 - W_{b_1}$. $W_{a_2} - W_{b_1}$ is a continuous r.v. and by (1) and (2), it is independent of Y. Therefore

$$\mathbb{P}(m_1 = m_2) = \mathbb{P}(W_{a_2} - W_{b_1} = Y) = \int_\Omega \mathbb{P}(W_{a_2} - W_{b_1} = y)\, \mathscr{L}_Y(\mathrm{d}y) = 0.$$

The second equality is due to $W_{a_2} - W_{b_1} \perp\!\!\!\perp Y$ and the last equality follows from the fact that $\mathbb{P}(W_{a_2} - W_{b_1} = y) = 0$ for all $y \in \mathbb{R}$, since $W_{a_2} - W_{b_1}$ is Gaussian. $\square$

Corollary 2.24. *With probability 1, every local maximum of W is strict.*

By saying that x_0 is a local maximum of f, we mean that there exists $\varepsilon > 0$ such that $f(x) < f(x_0)$ for all x with $0 < |x - x_0| < \varepsilon$.

Proof. Due to Corollary 2.23, with probability 1, for all non-overlapping closed intervals I_1, I_2 with rational endpoints, we have $\max_{t\in I_1} W_t \neq \max_{t\in I_2} W_t$, and this property implies that there are no non-strict local maxima. ☐

Corollary 2.25. *With probability 1, the points of local maxima of W are dense in $\mathbb{R}_+$ and countable.*

Proof. Exercise. ☐

Corollary 2.26. *For any $T > 0$, define*

$$S_T := \max_{t\in[0,T]} W_t.$$

Then with probability 1, for any $T > 0$ there exists a unique $t^* \leqslant T$ such that $W_{t^*} = S_T$. We then define

$$t^* =: \arg\max_{t\leqslant T} W_t.$$

Proof. Assume that there exist $t_1 < t_2 \leqslant T$ such that $W_{t_1} = W_{t_2} = S_T$. Then there exist two non-overlapping closed intervals $I_1, I_2 \subset [0,T]$ such that $\max_{t\in I_1} W_t = \max_{t\in I_2} W_t = S_T$. This event has probability 0. ☐

2.7 Some properties related to “variations”

Let $W = \{W_t : t \geqslant 0\}$ be a standard Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$ with $W_0 = 0$.

Proposition 2.27. *Fix $T \in (0, +\infty)$. Let $t^{(n)} = \{0 = t_0^{(n)} < t_1^{(n)} < \cdots < t_{K_n}^{(n)} = T\}$ be a sequence of partitions of $[0,T]$ such that*

$$\operatorname{mesh}\left(t^{(n)}\right) := \max_{j=1,\dots,K_n} \left[t_j^{(n)} - t_{j-1}^{(n)}\right] \xrightarrow[n\uparrow+\infty]{} 0.$$

Then

$$\sum_{j=1}^{K_n} \left(W_{t_j^{(n)}} - W_{t_{j-1}^{(n)}}\right)^2 \xrightarrow[n\uparrow+\infty]{} T, \quad \textit{in } L^2(\mathbb{P}).$$

Proof. The conclusion is equivalent to

$$\mathbb{E}\left[\left\{\sum_{j=1}^{K_n} \left(W_{t_j^{(n)}} - W_{t_{j-1}^{(n)}}\right)^2 - T\right\}^2\right] \xrightarrow[n\uparrow+\infty]{} 0;$$

this property only depends on the finite-dimensional distribution of W, so we can assume that W is (any) pre-brownian motion (like Brownian motion, but not necessarily continuous). Let

$$G = \{G(A) : A \in \mathcal{B}(\mathbb{R}_+),\, \lambda(A) < +\infty\}$$

be a centered Gaussian measure, controled by the Lebesgue measure λ (see Definition 1.4). Then we know that

(1) $t \mapsto G([0,t])$ is a pre-brownian motion (see the proof of Proposition 2.2),

(2) by Proposition 1.7, for any measurable set B with $\lambda(B) < +\infty$ and for every sequence $\{B_1^{(n)} \cup \cdots \cup B_{K_n}^{(n)}\}$ of partitions of B such that

$$\max_{j=1,\ldots,K_n} \lambda(B_j^{(n)}) \xrightarrow[n\uparrow+\infty]{} 0,$$

we have

$$\sum_{j=1}^{K_n} G(B_j^{(n)})^2 \xrightarrow[n\uparrow+\infty]{L^2(\mathbb{P})} \lambda(B).$$

Selecting $B = [0,T]$ and $B_j^{(n)} = (t_{j-1}^{(n)}, t_j^{(n)}]$ for $j \geqslant 2$ and $B_1^{(n)} = [0, t_1^{(n)}]$ gives the result. □

Definition 2.28. Let f be a real-valued function defined on an interval $[a,b] \subset \mathbb{R}$. The total variation of f is by definition

$$V_{[a,b]}(f) := \sup_{t^{(n)}} \sum_{j=1}^{n} \left| f(t_j^{(n)}) - f(t_{j-1}^{(n)}) \right|,$$

where the supremum is taken over all partitions of $[0,T]$. A function f is said to be of bounded variation on $[a,b]$ if $V_{[a,b]}(f) < +\infty$.

Proposition 2.29. *Let $T > 0$ be arbitrary. Then with probability 1,*

$$V_{[0,T]}(W) = +\infty.$$

In other words, a Brownian motion does not have bounded variation (a.s.) on any finite time interval.

Proof. See handwritten notes. □

Remark. Functions of bounded variations have the following basic properties:

(1) The following are equivalent for $f\colon [0,T] \to \mathbb{R}$:

(a) f has bounded variation on $[0,T]$,

(b) $f(t) = F(t) - G(t)$, where F and G are non-decreasing.

(2) The following are equivalent for $f\colon [0,T] \to \mathbb{R}$:

(a) There exists a finite signed measure[1] μ on $[0,T]$ such that $f(x) - f(a) = \mu([a,x])$ for all $0 \leqslant a \leqslant x \leqslant T$,

(b) f has bounded variation on $[0,T]$ and is right continuous on $[0,T]$.

(3) If f is of bounded variation on $[0,T]$, then the following conditions hold:

(i) f is continuous, except at most on a countable set.

(ii) f is λ-a.e. differentiable and $f' \in L^1((0,T),\lambda)$.

2.8 The strong Markov property of Brownian motions

Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that there exists a standard Brownian motion $W = \{W_t : t \geqslant\}$ satisfying the following conditions:[2]

(a) $\mathcal{F} = \sigma(W)$,

(b) the mapping $t \mapsto W_t(\omega)$ is continuous $\underline{\forall \omega \in \Omega}$.

[1] A signed measure μ on a measurable space $(A, \mathcal{A})$ is a map $\mu\colon \mathcal{A} \to [-\infty, +\infty]$ such that $\mu(\varnothing) = 0$ and $\sum_i \mu(A_i) = \mu(A)$ for all disjoint sets A_i such that $A = \bigcup_i A_i$.

[2] If (b) is not satisfied, we can set the Brownian motion to 0 at a set of measure 0.

Let $\mathcal{N}$ denote the $\mathbb{P}$-null sets of Ω. We extend $\mathbb{P}$ to

$$\mathcal{F}_* := \mathcal{F} \vee \sigma(\mathcal{N}).$$

Our reference filtration will be

$$\mathcal{F}_t = \sigma\{W_u : u \leqslant t\} \vee \sigma(\mathcal{N}) \qquad (t \geqslant 0).$$

This filtration is the "canonical augmentation" of the Brownian filtration. Obviously, it satisfies $\mathcal{F}_t \subset \mathcal{F}_*$ for all t. Moreover, one can show that it has the following properties:
(i) $W_t - W_s \perp\!\!\!\perp \mathcal{F}_s$ for all $0 \leqslant s < t$,
(ii) $\{\mathcal{F}_t : t \geqslant 0\}$ satisfies the "usual conditions", i.e.
(a) $\mathcal{F}_0 \supset \sigma(\mathcal{N})$,
(b) $\mathcal{F}_{t+} := \bigcap_{\varepsilon > 0} \mathcal{F}_{t+\varepsilon} = \mathcal{F}_t$.

We know that a Brownian motion W satisfies the weak Markov property, that is, for any deterministic time T, the process $\{W_{T+t} - W_T : t \geqslant 0\}$ is again a Brownian motion. This property is no longer true if T is a random time. For instance, the random time $T := t^* = \arg\max_{t \leqslant S} W_t$ (with S deterministic) provides a counterexample. However, if T is a stopping time, W satisfies what is called the strong Markov property, which will be stated below.

Definition 2.30. (a) A random variable $T \geqslant 0$ (possibly infinite) is called a $\mathcal{F}_t$-stopping time if for all $t \geqslant 0$,

$$\{T \leqslant t\} \in \mathcal{F}_t.$$

(b) Let T be a $\mathcal{F}_t$-stopping time. We define

$$\mathcal{F}_T := \big\{A \in \mathcal{F}_* : A \cap \{T \leqslant t\} \in \mathcal{F}_t,\ \forall t \geqslant 0\big\}.$$

Examples. (a) For any $a \in \mathbb{R}$, the hitting time $T_a := \inf\{s \geqslant 0 : W_s = a\}$ is an $\mathcal{F}_t$-stopping time, since

$$\{T_a \leqslant t\} = \begin{cases} \{\max_{s \leqslant t} W_s \geqslant a\} & \text{if } a > 0, \\ \{\min_{s \leqslant t} W_s \leqslant a\} & \text{if } a \leqslant 0. \end{cases}$$

(b) $t^* = \arg\max_{t \leqslant T} W_t$ with T deterministic is not a stopping time (see Corollary 2.32).
(c) $g = \sup\{t \leqslant 1 : W_t = 0\}$ is not a stopping time (see Corollary 2.32).

Exercise. (1) $\mathcal{F}_T$ is a σ-field.
(2) $T \in \mathcal{F}_T$.
(3) If $T \equiv t$ is deterministic, then T is a stopping time and $\mathcal{F}_T = \mathcal{F}_t$.
Other properties: see handwritten notes.

Theorem 2.31 (Strong Markov property). *Let T be an $\mathcal{F}_t$-stopping time and set*

$$W_t^{(T)} = \{W_{T+t} - W_T\} \mathbb{1}_{\{T < +\infty\}} \qquad (\forall t \geqslant 0).$$

Then, conditionally on $\{T < +\infty\}$, $W^{(T)}$ is a standard Brownian motion issued from 0 and independent of $\mathcal{F}_T$. More precisely,
* *$W^{(T)}$ has continuous paths,*
* *$\forall B \in \mathcal{F}_t$, $\forall d \geqslant 1$, $\forall (\lambda_1, \dots, \lambda_d) \in \mathbb{R}^d$ and $\forall (t_1, \dots, t_d) \in \mathbb{R}_+^d$, we have*

$$\mathbb{E}\left[\mathbb{1}_B \exp\left(i \sum_{j=1}^d \lambda_j W_{t_j}^{(T)}\right) \,\middle|\, T < +\infty\right] = \mathbb{P}(B \mid T < +\infty) \times \mathbb{E}\left[\exp\left(i \sum_{j=1}^d \lambda_j W_{t_j}\right)\right]$$

Proof. See handwritten notes. $\square$

Corollary 2.32. *The random variables $t^* = \arg\max_{t \leqslant T} W_t$ (with T deterministic) and $g = \sup\{t \leqslant 1 : W_t = 0\}$ are not $\mathcal{F}_t$-stopping times.*

Proof. The behaviours around $t = 0$ of $t \mapsto (W_{t^*+t} - W_{t^*})$ and $t \mapsto (W_{g+t} - W_g)$ are not Brownian (they have a constant sign). $\square$

One important application of the strong Markov property is the "reflection principle", associated with the problem: "What is the law of $S_t = \max_{s \leqslant t} W_s$?"

Theorem 2.33 (Reflection principle)**.** *Fix $t > 0$. Then for any $y \geqslant 0$ and $z \geqslant 0$,*

$$\mathbb{P}\{W_t < z - y,\, S_t \geqslant z\} = \mathbb{P}\{W_t > z + y\}$$

(where $S_t = \max_{s \leqslant t} W_s$).

Proof. Let T_z be the hitting time of z, i.e. $T_z = \inf\{s \geqslant 0 : W_s = z\}$. Then

$$\begin{aligned}
\mathbb{P}\{S_t \geqslant z,\, W_t < z - y\} &= \mathbb{P}\{T_z \leqslant t,\, W_{T_z + (t - T_z)} - z < -y\} \\
&\overset{\circledast}{=} \mathbb{P}\{T_z \leqslant t,\, W^{(T_z)}_{t - T_z} < -y\} \\
&\overset{*}{=} \mathbb{P}\{T_z \leqslant t,\, -W^{(T_z)}_{t - T_z} < -y\} \\
&\overset{\circledast}{=} \mathbb{P}\{T_z \leqslant t,\, z - W_t < -y\} \\
&= \mathbb{P}\{W_t > z + y\}.
\end{aligned}$$

The equality labeled with $*$ follows from the strong Markov property applied to the process $W^{(T_z)}$. Equalities labeled with $\circledast$ rely on the fact that $z = W_{T_z}$. $\square$

Corollary 2.34. *For any fixed $t \geqslant 0$, the r.v. S_t has the same law as $|W_t|$.*

Proof. Fix $z \geqslant 0$. Then

$$\mathbb{P}\{S_t \geqslant z\} = \mathbb{P}\{S_t \geqslant z,\, W_t \geqslant z\} + \mathbb{P}\{S_t \geqslant z,\, W_t < z\} \overset{\circledast}{=} 2\mathbb{P}\{W_t \geqslant z\} = \mathbb{P}\{|W_t| \geqslant z\}.$$

The equality labeled with $\circledast$ is obtained by applying the reflection principle with $y = 0$ and using the fact that $\{W_t \geqslant z\} \subset \{S_t \geqslant z\}$. $\square$

Remark. We have $S_t \overset{\text{Law}}{=} |W_t|$ for all $t \geqslant 0$, but it is not an equality in the sense of stochastic processes. For instance, note that S_t is, unlike $|W_t|$, always non-decreasing.

Corollary 2.35. *Fix $z > 0$. Then the law of the hitting time T_z is*

$$T_z \overset{\text{Law}}{=} \frac{z^2}{W_1^2} \sim \frac{z^2}{\mathcal{N}(0,1)^2}.$$

Proof. Using Corollary 2.34 in $*$ and the scaling property of Brownian motions in $\circledast$, we obtain

$$\mathbb{P}\{T_z \leqslant t\} = \mathbb{P}\{S_t \geqslant z\} \overset{*}{=} \mathbb{P}\{|W_t| \geqslant z\} \overset{\circledast}{=} \mathbb{P}\{\sqrt{t}|W_1| \geqslant z\} = \mathbb{P}\left\{\frac{z^2}{W_1^2} \leqslant t\right\}. \qquad \square$$

Corollary 2.36. *Then process $\{T_z : z \geqslant 0\}$ is*

(1) non-decreasing,

(2) with independent increments, i.e. for all $z > z'$, we have $(T_z - T_{z'}) \perp\!\!\!\perp \sigma\{T_u : u \leqslant z'\}$,

(3) the law of $T_z - T_{z'}$ only depends on $z - z'$ (stationary increments).

So $z \mapsto T_z$ is a non-decreasing Lévy process, also called a "subordinator".

Proof. See handwritten notes. □

Corollary 2.37 (Zeroes of a Brownian motion). *Let $Z := \{t \geqslant 0 : W_t = 0\}$ be the set of zeroes of W. The set Z is a.s.*

(a) *closed,*
(b) *with no isolated points,*
(c) *with Lebesgue measure equal to* 0.

Proof. See handwritten notes. □

Exercise. Use the reflection principle to show that for any $t > 0$, (S_t, W_t) has a joint density given by

$$g(x,y) = \frac{2(2x-y)}{\sqrt{2\pi t^3}} \exp\left(-\frac{(2x-y)^2}{2t}\right) \mathbb{1}_{(x>0,\ y\leqslant x)}.$$

Exercise. Use Corollary 2.35 to prove that for any $z > 0$, T_z has a density

$$f(t) = \frac{z}{\sqrt{2\pi t^3}} \exp\left(-\frac{z^2}{2t}\right) \mathbb{1}_{\{t>0\}}.$$

Remark. If $z < 0$, then $T_z \overset{\text{Law}}{=} T_{-z}$ (since $W \overset{\text{Law}}{=} -W$).

2.9 Some further definitions

Definition 2.38. For every $d \geqslant 1$, a d-dimensional Brownian motion issued from 0 is

$$\overline{W}_t = (W_t^{(1)}, \dots, W_t^{(d)})$$

where the $W^{(j)}$'s are independent Brownian motions.

Definition 2.39. Let W be a one-dimensional Brownian motion. Then the process

$$X_t = x + \mu t + \sigma W_t \qquad (x \in \mathbb{R},\ \mu, \sigma \in \mathbb{R})$$

is called a Brownian motion issued from x with drift μ and volatility σ. In particular, $X_t \sim \mathcal{N}(x + \mu t, \sigma^2 t)$.

Definition 2.40. Gien a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we call filtration an increasing collection $\{\mathcal{H}_t : t \geqslant 0\}$ of sub-σ-fields of $\mathcal{F}$. ("Increasing" means that $\mathcal{H}_s \subset \mathcal{H}_t$ whenever $s < t$.)

Definition 2.41. Given a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{H}_t), \mathbb{P})$, we say that $W = \{W_t : t \geqslant 0\}$ is an $\mathcal{H}_t$-Brownian motion if

(i) $W_0 = 0$ a.s.,
(ii) W has a.s. continuous paths,
(iii) W_t is $\mathcal{H}_t$-measurable for every t (i.e. W_t is $\mathcal{H}_t$-adapted),
(iv) $W_t - W_s \perp\!\!\!\perp \mathcal{H}_s$ whenever $t > s$ and $W_t - W_s \sim \mathcal{N}(0, t-s)$.

Examples. (1) Any Brownian motion W is a $\mathcal{H}_t$-Brownian motion with respect to its canonical filtration $\mathcal{H}_t := \sigma\{W_u : u \leqslant t\}$ and with respect to the filtration $\mathcal{F}_t := \sigma\{W_u : u \leqslant t\} \vee \sigma(\mathcal{N})$.

(2) Let Y be a r.v. independent of a Brownian motion W. Then W is a $\mathcal{H}_t$-Brownian motion for $\mathcal{H}_t := \mathcal{F}_t \vee \sigma(Y)$.

Definition 2.42. We say that a filtration $\{\mathcal{H}_t\}$ satisfies the usual conditions if

$$\mathcal{H}_0 \supset \sigma(\mathcal{N}) \qquad \text{and} \qquad \bigcap_{\varepsilon>0} \mathcal{H}_{t+\varepsilon} = \mathcal{H}_t \quad (\forall t \geqslant 0).$$

Recall also the notation

$$\mathcal{H}_{t+} := \bigcap_{\varepsilon>0} \mathcal{H}_{t+\varepsilon}.$$

Definition 2.43. A stochastic process $X = \{X_t : t \geqslant 0\}$ is called a

$$\left.\begin{array}{l} \mathcal{H}_t\text{-martingale} \\ \mathcal{H}_t\text{-submartingale} \\ \mathcal{H}_t\text{-supermartingale} \end{array}\right\} \text{ if } \forall t, \left[\begin{array}{c} X_t \in \mathcal{H}_t \text{ and} \\ X_t \in L^1(\mathbb{P}) \end{array}\right] \text{ and } \forall s < t, \left\{\begin{array}{l} \mathbb{E}[X_t \mid \mathcal{H}_s] = X_s, \\ \mathbb{E}[X_t \mid \mathcal{H}_s] \geqslant X_s, \\ \mathbb{E}[X_t \mid \mathcal{H}_s] \leqslant X_s. \end{array}\right.$$

Chapter 3

Stochastic Integrals

3.1 Construction of the stochastic integral

For this chapter, we fix a complete probabilistic space $(\Omega, \mathcal{F}, \mathbb{P})$ and a filtration $\{\mathcal{H}\}_{t\geqslant 0}$ of $\mathcal{F}$ satisfying the usual conditions.

Definition 3.1. Let $X = \{X_t : t \geqslant 0\}$ be a stochastic process.

(a) X is measurable if $(t, \omega) \mapsto X_t(\omega)$ is $\mathcal{B}(\mathbb{R}) \otimes \mathcal{F}$-measurable.

(b) X is $\mathcal{H}_t$-adapted, if for all t, $X_t \in \mathcal{H}_t$.

(c) X is $\mathcal{H}_t$-progressively measurable if for any fixed deterministic $T > 0$, the mapping $(t, \omega) \mapsto X_t(\omega)$ restricted to $[0, T] \times \Omega$ is $\mathcal{B}([0, T]) \otimes \mathcal{H}_T$-measurable.

Remark. If X is $\mathcal{H}_t$-progressive, then X is $\mathcal{H}_t$-adapted and measurable.

Lemma 3.2. *Suppose that X is a.s. right-continuous or a.s. left continuous. Then, if X is $\mathcal{H}_t$-adapted, it is also progressive with respect to $\mathcal{H}_t$.*

Proof. We just prove the right-continuous case. For all $n \geqslant 1$, let us consider the process

$$X_t^n(\omega) = \sum_{k=0}^{\infty} \mathbb{1}_{\left[\frac{k-1}{n}, \frac{k}{n}\right)}(t) X_{\frac{k}{n}}(\omega)$$

(we discretize X). By right continuity,

$$X_t^n(\omega) \xrightarrow[n\uparrow+\infty]{} X_t(\omega).$$

Moreover, by construction, for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for any $n \geqslant N$, the process X_t^n is $\mathcal{H}_{t+\varepsilon}$-progressive, meaning that X_t is $\mathcal{H}_{t+\varepsilon}$-progressive for all $\varepsilon > 0$. This means that for any $A \in \mathcal{B}(\mathbb{R})$ and for any $T > 0$, we have

$$\big\{(t, \omega) \in [0, T] \times \Omega : X_t(\omega) \in A\big\} \in \bigcap_{\varepsilon>0} \big(\mathcal{B}([0, T]) \otimes \mathcal{H}_{T+\varepsilon}\big) = \mathcal{B}([0, T]) \otimes \mathcal{H}_T,$$

the last equality being true by the right continuity of $\mathcal{H}_t$. So, X is $\mathcal{H}_t$-progressive. □

Counterexample (Total disorder process). On $(\Omega, \mathcal{F}, \mathbb{P})$, let $\{N_t : t \geqslant 0\}$ be a continuous collection of i.i.d. $\mathcal{N}(0, 1)$ random variables. Then the mapping $(t, \omega) \mapsto N_t(\omega)$ cannot be jointly measurable.

Indeed, if it was the case, then for any $s < u$,

$$\mathbb{E}\left[\left(\int_s^u N_t \,\mathrm{d}t\right)^2\right] \overset{\text{Fubini}}{=} \int_s^u \int_s^u \mathbb{E}[N_t N_{t'}] \,\mathrm{d}t \,\mathrm{d}t' = 0.$$

This would imply that

$$\mathbb{P}\left\{\int_s^u N_t \,\mathrm{d}t = 0 \text{ for all } s, u\right\} = 1,$$

whence

$$\mathbb{P}\{N_t = 0 \text{ for almost every } t\} = 1.$$

But this is absurd, since

$$\mathbb{E}\left[\int_0^T N_t^2 \,\mathrm{d}t\right] \overset{\text{Fubini}}{=} \int_0^T \mathbb{E}[N_t^2] \,\mathrm{d}t = \int_0^T 1 \,\mathrm{d}t = T.$$

Definition 3.3. We denote by $\mathrm{Prog}(\mathcal{H}_t)$ or Prog the smallest σ-field on $\mathbb{R}_+ \times \Omega$ such that all $\mathcal{H}_t$-progressive processes X are measurable (in (t, ω)). Then, then space

$$L^2(\mathrm{Prog}) := L^2(\mathbb{R}_+ \times \Omega, \mathrm{Prog}, \,\mathrm{d}t \times \,\mathrm{d}\mathbb{P})$$

of square-integrable progressive processes is a real Hilbert space. It consists of all processes $h_t = h(t, \omega)$ which are $\mathcal{H}_t$-progressive and satisfy

$$\mathbb{E}\left(\int_0^\infty h_t^2 \,\mathrm{d}t\right) < +\infty.$$

Definition 3.4. We denote by $\mathcal{E}$ the space of elementary processes, i.e.

$$\mathcal{E} := \left\{ f(t,\omega) \;\middle|\; f(t,\omega) = \sum_{j=0}^{N-1} F_{t_j}(\omega) \mathbb{1}_{(t_j, t_{j+1}]}(t), \text{ with } \left\{ \begin{array}{l} 0 \leqslant t_0 < t_1 < t_2 < \cdots < t_N < +\infty, \\ F_{t_j} \in \mathcal{H}_{t_j} \text{ bounded} \end{array} \right. \right\}.$$

Theorem 3.5.
(a) $\mathcal{E} \subset L^2(\mathrm{Prog})$,
(b) $\mathcal{E}$ *is a vector space,*
(c) $\mathcal{E}$ *is dense in* $L^2(\mathrm{Prog})$.

Proof. (a) Each element of $\mathcal{E}$ is

$$\left.\begin{array}{l} \left.\begin{array}{l} \text{— adapted to } \mathcal{H}_t \\ \text{— left-continuous} \end{array}\right\} \Longrightarrow \mathcal{H}_t\text{-progressive (by Lemma 3.2)} \\ \text{— bounded} \\ \text{— with compact support in the variable } t \end{array}\right\} \Longrightarrow \mathcal{E} \subset L^2(\mathrm{Prog}).$$

As for (b) and (c), see handwritten notes. □

From now on, $W = \{W_t : t \geqslant 0\}$ is a standard $\mathcal{H}_t$-Brownian motion issued from 0.

Definition 3.6. Let $f(t, \omega) \in \mathcal{E}$ be given by

$$f(t,\omega) = \sum_{k=0}^{N-1} F_{t_k}(\omega) \mathbb{1}_{(t_k, t_{k+1}]}(t)$$

with $F_{t_k} \in \mathcal{H}_{t_k}$ and bounded. We set

$$\int_0^\infty f(t)\,\mathrm{d}W_t = \int_0^\infty f\,\mathrm{d}W := \sum_{k=0}^{N-1} F_{t_k}(W_{t_{k+1}} - W_{t_k}).$$

For every $t \geqslant 0$, we set

$$\int_0^t f(s)\,\mathrm{d}W_s = \int_0^t f\,\mathrm{d}W := \int_0^\infty f\mathbb{1}_{[0,t]}\,\mathrm{d}W = \sum_{k=0}^{N-1} F_{t_k}(W_{t_{k+1}\wedge t} - W_{t_k\wedge t}).$$

In particular,

$$\int_0^0 f\,\mathrm{d}W = 0.$$

Remark. Thus $\int_0^\infty f\,\mathrm{d}W$ is a sum of increments $W_{t_{k+1}} - W_{t_k}$, each of them multiplied by a weight random function F_{t_k}, which is independent of the increments (since $W_{t_{k+1}} - W_{t_k} \perp\!\!\!\perp \mathcal{H}_{t_k}$).

Proposition 3.7.
(1) For any $f, g \in \mathcal{E}$,

$$\int_0^\infty (f+g)\,\mathrm{d}W = \int_0^\infty f\,\mathrm{d}W + \int_0^\infty g\,\mathrm{d}W.$$

(2) For any $f, g \in \mathcal{E}$,

$$\mathbb{E}\left[\int_0^\infty f\,\mathrm{d}W \times \int_0^\infty g\,\mathrm{d}W\right] = \langle f, g\rangle_{L^2(\mathrm{Prog})} = \mathbb{E}\left[\int_0^\infty f(t)g(t)\,\mathrm{d}t\right].$$

(3) For any $f \in \mathcal{E}$,

$$\mathbb{E}\left[\int_0^\infty f\,\mathrm{d}W\right] = 0.$$

Proof. (1) is trivial. (3) comes from

$$\mathbb{E}\left[F_{t_k}(W_{t_{k+1}} - W_{t_k})\right] = \mathbb{E}\left[F_{t_k}\right]\mathbb{E}\left[(W_{t_{k+1}} - W_{t_k})\right] = 0.$$

To show (2), it is sufficient to consider the case $f = g$, since $f \mapsto \int_0^\infty f\,\mathrm{d}W$ is linear by point (1). We have

$$\begin{aligned}
\mathbb{E}\left[\left(\int_0^\infty f(t)\,\mathrm{d}W_t\right)^2\right] &= \mathbb{E}\left[\left(\sum_{k=0}^{N-1} F_{t_k}(W_{t_{k+1}} - W_{t_k})\right)^2\right] \\
&= \sum_{k=0}^{N-1}\sum_{j=0}^{N-1} \mathbb{E}\left[F_{t_k}F_{t_j}(W_{t_{k+1}} - W_{t_k})(W_{t_{j+1}} - W_{t_j})\right] \\
&= \sum_{k=0}^{N-1} \mathbb{E}\left[F_{t_k}^2(W_{t_{k+1}} - W_{t_k})^2\right] \\
&= \sum_{k=0}^{N-1} \mathbb{E}\left[F_{t_k}^2\right](t_{k+1} - t_k).
\end{aligned}$$

So

$$\mathbb{E}\left[\left(\int_0^\infty f\,\mathrm{d}W\right)^2\right] = \mathbb{E}\left[\sum_{k=0}^{N-1} F_{t_k}^2(t_{k+1} - t_k)\right] = \mathbb{E}\left[\int_0^\infty f(t)^2\,\mathrm{d}t\right]. \qquad \square$$

Corollary 3.8. *Let $\{g_n\}_{n\geqslant 1} \subset \mathcal{E}$ be a sequence which is Cauchy in $L^2(\mathrm{Prog})$. Then $\{\int_0^\infty g_n \,\mathrm{d}W : n \geqslant 1\}$ is Cauchy in $L^2(\mathbb{P})$.*

Proof. Follows from the isometry relation

$$\mathbb{E}\left[\left(\int_0^\infty g_n \,\mathrm{d}W - \int_0^\infty g_m \,\mathrm{d}W\right)^2\right] = \|g_n - g_m\|_{L^2(\mathrm{Prog})}. \qquad \square$$

Definition 3.9. Let $g \in L^2(\mathrm{Prog})$. Then by density, there exists a sequence $\{g_n\}_{n\geqslant 1} \subset \mathcal{E}$ converging in $L^2(\mathrm{Prog})$ to g. By Corollary 3.8, $\{\int_0^\infty g_n \,\mathrm{d}W\}_{n\geqslant 1}$ is Cauchy in $L^2(\mathbb{P})$. We define

$$\int_0^\infty g(t)\,\mathrm{d}W_t = \int_0^\infty g\,\mathrm{d}W := \overset{L^2(\mathbb{P})}{\lim_{n\uparrow+\infty}} \int_0^\infty g_n \,\mathrm{d}W.$$

Proposition 3.10. *For all $g \in L^2(\mathrm{Prog})$, the definition of $\int_0^\infty g\,\mathrm{d}W$ is well-given. Moreover, all assertions of Proposition 3.7 remain valid for every $f, g \in L^2(\mathrm{Prog})$. In particular, assertion (2) of Proposition 3.7 says that the mapping $f \mapsto \int_0^\infty f\,\mathrm{d}W$ is an isometry from $L^2(\mathrm{Prog})$ into $L^2(\mathbb{P})$.*

Proof. See handwritten notes. $\square$

Definition 3.11. We define

$$L^2_{\mathrm{loc}}(\mathrm{Prog}) := \left\{f(t,\omega) : \forall T \geqslant 0 \text{ deterministic}, f\mathbb{1}_{[0,T]} \in L^2(\mathrm{Prog})\right\}.$$

For all $f \in L^2_{\mathrm{loc}}(\mathrm{Prog})$, we define

$$\int_0^t f(s)\,\mathrm{d}W_s = \int_0^t f\,\mathrm{d}W = \int_0^\infty f\mathbb{1}_{[0,t]}\,\mathrm{d}W.$$

Remark. We have, of course, $L^2(\mathrm{Prog}) \subset L^2_{\mathrm{loc}}(\mathrm{Prog})$.

Proposition 3.12. *For every $f \in L^2_{\mathrm{loc}}(\mathrm{Prog})$, the process $\{\int_0^t f\,\mathrm{d}W : t \geqslant 0\}$ is a $\mathcal{H}_t$-martingale.*[1]

Proof. By density, we can take $f \in \mathcal{E}$, and by linearity, we can take f of the form $f(t) = F_a\mathbb{1}_{(a,b]}(t)$ with F_a bounded and $\mathcal{H}_a$-measurable. Then

$$\int_0^t f\,\mathrm{d}W = \begin{cases} 0 & \text{if } t \leqslant a, \\ F_a(W_t - W_a) & \text{if } t \in (a,b], \\ F_a(W_b - W_a) & \text{if } t > b. \end{cases}$$

So for all $t \geqslant 0$, $\int_0^t f\,\mathrm{d}W \in \mathcal{H}_t$, and one can prove the martingale property by considering all possible cases, e.g.:

(1) if $s < t \leqslant a$, then there is nothing to show;
(2) if $a \leqslant s < t \leqslant b$, then

$$\begin{aligned}\mathbb{E}\left[\int_0^t f\,\mathrm{d}W \,\middle|\, \mathcal{H}_s\right] &= \mathbb{E}[F_a(W - t - W_a) \mid \mathcal{H}_s] \\ &= F_a\mathbb{E}[W_t \mid \mathcal{H}_s] - F_aW_a \\ &= F_a(W_s - W_a) = \int_0^s f\,\mathrm{d}W;\end{aligned}$$

and so on. $\square$

[1] We will prove later on that it has an a.s. continuous modification.

Example. Let

$$L^2_{\mathrm{det}}(\mathbb{R}_+) = \left\{ f(t) \text{ deterministic} : \forall T > 0, \int_0^T f^2(t)\,\mathrm{d}t < +\infty \right\}.$$

Then $L^2_{\mathrm{det}}(\mathbb{R}_+) \subset L^2_{\mathrm{loc}}(\mathrm{Prog})$. Every function $f\mathbb{1}_{[0,T]}$ with $f \in L^2_{\mathrm{det}}(\mathbb{R}_+)$ can be approximated by step functions of the type

$$\sum_{k=0}^{N-1} C_k \mathbb{1}_{(t_k,t_{k+1}]}$$

where C_k's are (deterministic) real numbers. This means that for any $f \in L^2_{\mathrm{det}}(\mathbb{R}_+)$, $\int_0^T f\,\mathrm{d}W$ is the limit of r.v.'s of the type

$$\sum_{k=0}^{N-1} C_k (W_{t_{k+1}} - W_{t_k})$$

and $\int_0^T f\,\mathrm{d}W$ coincides with the Wiener-Itô integral of Chapter 1 (Definitions 1.8, 1.10 and 1.12, case $A = \mathbb{R}_+$ and $\mu = \mathrm{Leb}$). In particular:

Proposition 3.13. *Let $d \geqslant 1$, and for $j = 1, \ldots, d$, let $f_j \in L^2_{\mathrm{det}}(\mathbb{R}_+)$ and $t_j \in \mathbb{R}_+$. Set*

$$I_j := \int_0^{t_j} f_j\,\mathrm{d}W.$$

Then $(I_1, \ldots, I_d) \sim \mathcal{N}_d(\mathbf{0}, C)$, where $C = \{C(i,j) : i,j = 1, \ldots, d\}$ is a $d \times d$-matrix the entries of which are given by

$$C(i,j) = \int_0^{t_i \wedge t_j} f_i(x) f_j(x)\,\mathrm{d}x.$$

Remark. (i) Of course, $\int_0^t 1\,\mathrm{d}W = W_t$.

(ii) If $f \in L^2_{\mathrm{det}}(\mathbb{R}_+)$, the process $t \mapsto \int_0^t f\,\mathrm{d}W$ has the same law as the "time-changed" Brownian motion $t \mapsto W_{\int_0^t f^2(s)\,\mathrm{d}s}$. Indeed, $\min(\int_0^t f^2(s)\,\mathrm{d}s, \int_0^u f^2(s)\,\mathrm{d}s) = \int_0^{t\wedge u} f^2(s)\,\mathrm{d}s$, so both processes are Gaussian, centered and with the same covariance.

3.2 Stochastic integrals and continuity

First, let us recall Doob's maximal inequality:

Lemma 3.14. *Let $\{X_n : n = 0, 1, 2, \ldots N\}$ be a discrete martingale. Then for all $p \geqslant 1$ and $\lambda > 0$, we have*

$$\lambda^p\,\mathbb{P}\left\{ \sup_{0\leqslant n\leqslant N} |X_n| \geqslant \lambda \right\} \leqslant \mathbb{E}\left[|X_N|^p\right] \leqslant \sup_{0\leqslant n\leqslant N} \mathbb{E}|X_n|^p.$$

Moreover, for all $p > 1$, we have

$$\mathbb{E}\left[|X_N|^p\right] \leqslant \sup_{0\leqslant n\leqslant N} \mathbb{E}\left[|X_n|^p\right] \leqslant \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[|X_N|^p\right].$$

By apporximating general maxima and minima by maxima and minima over finite sets, we can prove Doob's maximal inequality for continuous martingales:

Proposition 3.15. *Let $X = \{X_t : t \geqslant 0\}$ be a continuous martingale. Then for all $\lambda > 0$, $T > 0$ and $p \geqslant 1$, we have*

$$\lambda^p\,\mathbb{P}\left\{ \sup_{0\leqslant t\leqslant T} |X_t| \geqslant \lambda \right\} \leqslant \sup_{0\leqslant t\leqslant T} \mathbb{E}|X_t|^p.$$

Theorem 3.16. *Let $f \in L^2_{\text{loc}}(\text{Prog})$, then the process $\{\int_0^t f\ dW : t \geqslant 0\}$ has an a.s. continuous modification. This means that there exists an a.s. continuous process $J = \{J_t : t \geqslant 0\}$ such that for all $t \geqslant 0$,*

$$\mathbb{P}\left\{J_t = \int_0^t f(s)\,\mathrm{d}W_s\right\} = 1.$$

Proof. See handwritten notes. The idea is to combine Doob's inequality (for continuous martingales), Borel-Cantelli and the fact that $\mathcal{E}$ is dense in $L^2(\text{Prog})$. □

Remark. From now on, we will always implicitly select a continuous version of $t \mapsto \int_0^t f\,\mathrm{d}W$, for $f \in L^2_{\text{loc}}(\text{Prog})$.

Example (Towards Itô formula). For instance, let us compute $\int_0^t W_s\,\mathrm{d}W_s$. First of all, note that $\{W_t : t \geqslant 0\}$ is an element of $L^2_{\text{loc}}(\text{Prog})$. This follows from the fact that the process $W_s \mathbb{1}_{\{s \leqslant t\}}$ is approximated in $L^2(\text{Prog})$ by the sequence

$$\sum_{k=1}^{N_n} W_{t_{k-1}^{(n)}} \mathbb{1}_{\left(t_{k-1}^{(n)}, t_k^{(n)}\right]}(s),$$

where $0 = t_0^{(n)} < t_1^{(n)} < \cdots < t_{N_n}^{(n)} = t$ is a sequence is partitions of $[0,t]$ with $\text{mesh}(t^{(n)}) \xrightarrow{n \uparrow +\infty} 0$. This also shows that, by construction,

$$\int_0^t W_s\,\mathrm{d}W_s = \overset{L^2(\mathbb{P})}{\lim_{n\to\infty}} \sum_{k=1}^{N_n} W_{t_{k-1}^{(n)}} \left(W_{t_k^{(n)}} - W_{t_{k-1}^{(n)}}\right) = \frac{1}{2}W_t^2 - \frac{t}{2}.$$

This gives

$$W_t^2 = 2\int_0^t W_s\,\mathrm{d}W_s + t,$$

which is a particular case of the Itô formula. A crucial fact in the above computation is

$$\overset{L^2(\mathbb{P})}{\lim_{n\to\infty}} \sum_{k=1}^{N_n} \left(W_{t_k^{(n)}} - W_{t_{k-1}^{(n)}}\right)^2 = t.$$

Remark. (1) If $s < t$, we set

$$\int_s^t f(u)\,\mathrm{d}W_u := \int_0^t f\,\mathrm{d}W - \int_0^s f\,\mathrm{d}W.$$

(2) Assume that $s < t$, $f \in L^2_{\text{loc}}(\text{Prog})$ and $H \in L^2(\mathcal{H}_s)$. Then

$$Hf(u)\mathbb{1}_{\{s \leqslant u \leqslant t\}} \in L^2_{\text{loc}}(\text{Prog})$$

and

$$\int_s^t (Hf)\,\mathrm{d}W = H\int_s^t f\,\mathrm{d}W. \tag{$*$}$$

To see this, observe first that $(*)$ is trivial when $f \in \mathcal{E}$. For general f, observe that if $\mathcal{E} \ni q_n \longrightarrow f\mathbb{1}_{(s,t]}$ in $L^2(\text{Prog})$, then $Hq_n\mathbb{1}_{(s,t]} \longrightarrow Hf\mathbb{1}_{(s,t]}$ in $L^2(\text{Prog})$. So that

$$\int_s^t Hq_n\,\mathrm{d}W \xrightarrow{L^2(\mathbb{P})} \int_s^t Hf\,\mathrm{d}W$$

and also

$$\int_s^t Hq_n\,\mathrm{d}W = H\int_s^t q_n\,\mathrm{d}W \xrightarrow{L^2(\mathbb{P})} H\int_s^t f\,\mathrm{d}W.$$

(3) The same rule applies if H is independent (stochastically) of everything.

The next proposition gives some results on stopping times (proofs are sketched or omitted).

Proposition 3.17. *(1) Let T be a $\mathcal{H}_t$-stopping time, and let $h \in L^2_{\text{loc}}(\text{Prog})$. Then*

$$\big(t \mapsto h(t)\mathbb{1}_{\{t\leqslant T\}}\big) \in L^2_{\text{loc}}(\text{Prog}).$$

(2) For any $t > 0$ deterministic,

$$\int_0^t h(s)\mathbb{1}_{\{s\leqslant T\}}\,\mathrm{d}W_s = \int_0^u h(s)\,\mathrm{d}W_s\bigg|_{u=t\wedge T}.$$

(3) If $h \in L^2(\text{Prog})$,

$$\int_0^\infty h(s)\mathbb{1}_{\{s\leqslant T\}}\,\mathrm{d}W_s = \int_0^u h(s)\,\mathrm{d}W_s\bigg|_{u=T}.$$

Proof. (2) and (3) can be proved approximating T by discrete stopping times (taking values in a countable set); see J.-F. Le Gall, or Mörters—Peres.

To prove (1), observe that $t \mapsto h(t)\mathbb{1}_{\{t\leqslant T\}} =: \tilde{h}(t)$ is such that for all t,

$$\mathbb{E}\left[\int_0^t \tilde{h}(s)^2\,\mathrm{d}s\right] \leqslant \mathbb{E}\left[\int_0^t h(s)^2\,\mathrm{d}s\right] < +\infty.$$

We have to show that $\mathbb{1}_{\{\cdot\leqslant T\}}$ is progressively measurable. Since it is left-continuous, we shall only show that it is $\mathcal{H}_t$-adapted (see Lemma 3.2). But

$$\{t \leqslant T\} = \{T < t\}^c \in \mathcal{H}_t,$$

because T is a $\mathcal{H}_t$-stopping time.[2] □

Definition 3.18. We define

$$\mathcal{I}_{\text{loc}}(\text{Prog}) := \left\{ f \;\middle|\; f \text{ is } \mathcal{H}_t\text{-progressive and } \forall T < +\infty, \text{ we have } \int_0^T f^2(t)\,\mathrm{d}t < +\infty, \text{ a.s.} \right\}.$$

Observe that $L^2(\text{Prog}) \subsetneqq L^2_{\text{loc}}(\text{Prog}) \subsetneqq \mathcal{I}_{\text{loc}}(\text{Prog})$. We will extend the stochastic integral to the class $\mathcal{I}_{\text{loc}}(\text{Prog})$.

Theorem 3.19.

I. *There exists a unique mapping*

$$\begin{aligned} \mathscr{J} : \mathcal{I}_{\text{loc}}(\text{Prog}) &\longrightarrow \{\textit{continuous, } \mathcal{H}_t\textit{-adapted process}\} \\ \{g(t) : t \geqslant 0\} &\longmapsto \{J_t(g) : t \geqslant 0\} \end{aligned}$$

such that

(1) $J_t(g) = \int_0^t g\,\mathrm{d}W$ as defined in previous lectures, for all $g \in L^2_{\text{loc}}(\text{Prog})$,

(2) $\mathscr{J}$ is linear,

(3) $\mathscr{J}$ is continuous in the sense that

$$\int_0^T \big(g^n(t)\big)^2\,\mathrm{d}t \xrightarrow[n\to+\infty]{\mathbb{P}} 0 \quad \textit{implies} \quad \sup_{t\leqslant T}|J_t(g^n)| \xrightarrow[n\to+\infty]{\mathbb{P}} 0.$$

II. *The process $\{J_t(g) : t \geqslant 0\}$ is not in general a $\mathcal{H}_t$-martingale, but only a local martingale, that is, there exists a sequence $\{T_n : n \geqslant 1\}$ of $\mathcal{H}_t$-stopping times such that $T_n \uparrow +\infty$, a.s., and $t \mapsto J_{t\wedge T_n}(g)$ is a uniformly integrable $\mathcal{H}_t$-martingale. One says that $\{T_n\}_{n\geqslant 1}$ "localizes" $J_t(g)$.*

Proof. See handwritten notes. □

Remark. A square-integrable martingale is a local martingale (and the opposite is false).

[2] We have indeed $\{T < t\} = \bigcup_{\varepsilon>0}\{T \leqslant t - \varepsilon\}$ and $\{T \leqslant t - \varepsilon\} \in H_{t-\varepsilon} \subset \mathcal{H}_t$.

Summary

$L^2(\mathrm{Prog})$	$\subsetneq$	$L^2_{\mathrm{loc}}(\mathrm{Prog})$	$\subsetneq$	$\mathcal{I}_{\mathrm{loc}}(\mathrm{Prog})$
The map $L^2(\mathrm{Prog}) \ni f \mapsto \int_0^\infty f \,\mathrm{d}W$ is well-defined. In addition, it is an isometry.		$\int_0^\infty f \,\mathrm{d}W$ is not defined for $f \in L^2_{\mathrm{loc}}(\mathrm{Prog})$ in general. The mapping $f\mathbb{1}_{[0,t]} \mapsto \int_0^t g(s) \,\mathrm{d}W_s$ is an isometry.		$\int_0^\infty f \,\mathrm{d}W$ is not defined for $f \in \mathcal{I}_{\mathrm{loc}}(\mathrm{Prog})$ in general.
$\int_0^t f \,\mathrm{d}W$ is a square-integrable continuous $\mathcal{H}_t$-martingale.		$\int_0^t f \,\mathrm{d}W$ is a square-integrable continuous $\mathcal{H}_t$-martingale.		$\int_0^t f \,\mathrm{d}W$ is a continuous $\mathcal{H}_t$-local martingale. It is localized a.s. by the sequence $T_n = \inf\{t \geqslant 0 : \int_0^t (1+g_s^2) \,\mathrm{d}s = n\}$.

Some further properties

Proposition 3.20. *Suppose $g \in \mathcal{I}_{\mathrm{loc}}(\mathrm{Prog})$ is such that $\int_0^t f \,\mathrm{d}W = 0$ for all t a.s. Then $g_t(\omega) = 0$ almost everywhere- $\mathrm{d}t \otimes \mathbb{P}(\,\mathrm{d}\omega)$.*

Proof. See handwritten notes. □

More generally,

Theorem 3.21. *If $g \in \mathcal{I}_{\mathrm{loc}}(\mathrm{Prog})$ is such that $t \mapsto \int_0^t g \,\mathrm{d}W$ is with bounded variation a.s.-$\mathbb{P}$, then necessarily $g \equiv 0$.*

3.3 Itô processes

Lemma 3.22. *Let $\{M_t\}_{t \geqslant 0}$ be a local $\mathcal{H}_t$-martingale with bounded variation (a.s.-$\mathbb{P}$) and such that $M_0 = 0$. Then $M_t = 0$ for all t, a.s.-$\mathbb{P}$.*

Proof. See handwritten notes. □

Definition 3.23. For $p \geqslant 1$, we define

$$\mathcal{L}^p_{\mathrm{loc}}(\mathrm{Prog}) := \left\{ h \;\middle|\; h \text{ is } \mathcal{H}_t\text{-progressively measurable and } \int_0^t |h_s|^p \,\mathrm{d}s < +\infty \text{ for all } t > 0 \right\}.$$

Note that, in particular, $\mathcal{I}_{\mathrm{loc}}(\mathrm{Prog}) = \mathcal{L}^2_{\mathrm{loc}}(\mathrm{Prog})$.

Definition 3.24. An Itô process $X = \{X_t : t \geqslant 0\}$ is a stochastic process of the type

$$X_t = X_0 + \int_0^t g_s \,\mathrm{d}W_s + \int_0^t h_s \,\mathrm{d}s, \tag{$*$}$$

where $X_0 \in \mathcal{H}_0$, $g \in \mathcal{I}_{\mathrm{loc}}(\mathrm{Prog})$ and $h \in \mathcal{L}^1_{\mathrm{loc}}(\mathrm{Prog})$.

Proposition 3.25. *Representation $(*)$ is unique, i.e., if*

$$X_t = X_0 + \int_0^t \hat{g}_s \,\mathrm{d}W_s + \int_0^t \hat{h}_s \,\mathrm{d}s, \tag{$\hat{*}$}$$

with $\hat{g} \in \mathcal{I}_{\mathrm{loc}}(\mathrm{Prog})$ *and* $\hat{h} \in \mathcal{L}^1_{\mathrm{loc}}(\mathrm{Prog})$, *then*

$$\left.\begin{array}{r} g_s(\omega) = \hat{g}_s(\omega) \\ h_s(\omega) = \hat{h}_s(\omega) \end{array}\right\} \text{ a.e.- } \mathrm{d}s \otimes \mathbb{P}(\mathrm{d}\omega).$$

Proof. See handwritten notes. □

Chapter 4

Itô formula and applications

The framework is the same as in previous chapters. We consider a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{H}_t\}, \mathbb{P})$ and a $\mathcal{H}_t$-Brownian motion $\{W_t\}$. We assume that $\{\mathcal{H}_t\}$ satisfies the usual conditions.

4.1 Simple statements, examples, applications

Theorem 4.1. *Let $X_0 \in \mathcal{H}_0$, $g \in \mathcal{I}_{\mathrm{loc}}(\mathrm{Prog})$ and $h, k \in \mathcal{L}^1_{\mathrm{loc}}(\mathrm{Prog})$. Set*

$$X_t := X_0 + \int_0^t g \,\mathrm{d}W + \int_0^t h_s \,\mathrm{d}s \qquad \textit{and} \qquad A_t := \int_0^t k_s \,\mathrm{d}s,$$

and take $f\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $(s,x) \mapsto f(s,x)$ to be of class $C^{1,2}$. Then

$$\begin{aligned} f(A_t, X_t) = f(0, X_0) &+ \int_0^t \frac{\partial}{\partial s} f(A_u, X_u) \,\mathrm{d}A_u \\ &+ \int_0^t \frac{\partial}{\partial x} f(A_u, X_u) \,\mathrm{d}X_u \\ &+ \frac{1}{2} \int_0^t \frac{\partial^2}{\partial x^2} f(A_u, X_u) \,\mathrm{d}\langle X \rangle_t, \end{aligned}$$

where

$$\begin{aligned} \mathrm{d}A_u &= k_u \,\mathrm{d}u, \\ \mathrm{d}X_u &= (g_u \,\mathrm{d}W_u + h_u \,\mathrm{d}u) \\ \mathrm{d}\langle X \rangle_u &= g_u^2 \,\mathrm{d}u \end{aligned}$$

and

$$\langle X \rangle_t = \langle X, X \rangle_t = \int_0^t g_s^2 \,\mathrm{d}s$$

is the quadratic variation of X.

Examples and applications

Example. Let $g \in \mathcal{I}_{\mathrm{loc}}(\mathrm{Prog})$, and let $X_t = \int_0^t g\,\mathrm{d}W$. We will compute X_t^2 using the above theorem with $f(x) := x^2$. We obtain

$$\begin{aligned} X_t^2 = f(X_t) &= 2\int_0^t X_u\,\mathrm{d}X_u + \langle X\rangle_t \\ &= 2\int_0^t X_u g_u\,\mathrm{d}W_u + \int_0^t g_u^2\,\mathrm{d}u. \end{aligned}$$

In particular, for any $g \in \mathcal{I}_{\mathrm{loc}}(\mathrm{Prog})$, the process

$$t \longmapsto \left(\int_0^t g\,\mathrm{d}W\right)^2 - \int_0^t g_s^2\,\mathrm{d}s = X_t^2 - \langle X\rangle_t$$

is a local martingale (with respect to $\mathcal{H}_t$). Also, $g \equiv 1$ gives

$$W_t^2 = 2\int_0^t W_s\,\mathrm{d}W_s + t.$$

Example. Let $X_t = X_0 + \int_0^t g\,\mathrm{d}W + \int_0^t h_s\,\mathrm{d}s$ and $A_t = \int_0^t K_s\,\mathrm{d}s$. We will compute A_tX_t using the function $f(s,x) := sx$. We obtain

$$A_tX_t = f(A_t, X_t) = \underbrace{0}_{=A_0X_0} + \int_0^t A_u\,\mathrm{d}X_u + \int_0^t X_u\,\mathrm{d}A_u.$$

Example. Let $g \in \mathcal{I}_{\mathrm{loc}}(\mathrm{Prog})$, $\lambda \in \mathbb{R}$ and $i = \sqrt{-1}$. Then the process

$$Z_t^\lambda := \exp\left(i\lambda\int_0^t g\,\mathrm{d}W + \frac{1}{2}\lambda^2\int_0^t g_s^2\,\mathrm{d}s\right)$$

is a complex $\mathcal{H}_t$-local martingale such that

$$Z_t^\lambda = 1 + i\lambda\int_0^t Z_u^\lambda g_u\,\mathrm{d}W_u$$

(note that, in this case, $g_u\,\mathrm{d}W_u = \mathrm{d}X_u$). If g is deterministic, then Z^λ is a true $\mathcal{H}_t$-martingale.

4.2 Stochastic Differential Equations and Girasnov's theorem

$(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P}, W_t, \cdots$ Consider a process $\theta \in \mathcal{I}_{\mathrm{loc}}(Prog)$ such that $\mathbb{E}(\exp\frac{1}{2}\int_0^T \theta_s^2 ds) < \infty$ this $\Longrightarrow$ $Z_t^\theta = exp\int_0^t \theta_s dW_s - \frac{1}{2}\int_0^t \theta_s^2 ds$ is a uniform integrable martingale ($\forall[0,T]$) such that $\mathbb{E}(Z_T^\theta) = 1$ $\forall T > 0$

Definition 4.2.

$$\mathbb{P}^\theta(A) := \mathbb{E}\mathbb{1}_A Z_T^\theta \tag{4.1}$$

Let $A \in \mathcal{F}_T$ where T is the deadline time. Note also that T is positive and fixed.

$\mathbb{P}^\theta(\Omega) = \mathbb{E}[Z_T^\theta] = 1$ where $\mathbb{P}^\theta$ is a probability. We need to prove *infinite additivity*[1] If $A_{i i\geqslant 1}$ is a sequence of disjoint elements of $\mathcal{F}_T$

$$\mathbb{P}^\theta(\bigcup_{i=1}^\infty A_i) = \mathbb{E}(\sum_{i=1}^\infty \mathbb{1}_{A_i}.Z_T^\theta) \underset{\substack{\uparrow \\ \text{by dominated convergence}}}{=}$$

[1] We seem to be able to read countably additive here

$$\sum_{i=1}^{\infty}\mathbb{E}(\mathbb{1}_{A_i}Z_T^{\theta}) \overset{\text{by 4.1}}{\overset{\downarrow}{=}} \sum_{i=1}^{\infty}\mathbb{P}^{\theta}(A_i)$$

Remark. $\mathbb{P}^{\theta}$ and $\mathbb{P}$ are equivalent on $\mathcal{F}_T : \forall A \in \mathcal{F}_T$
$\mathbb{P}^{\theta}(A) = 0 \iff \mathbb{P}(A) = 0$

Proof. Indeed: Assume that $\mathbb{P}(A) = 0 \Rightarrow \mathbb{1}_A(\omega) = 0,$ a.e. $\mathbb{P}(d\omega) \Rightarrow \mathbb{P}^{\theta}(A) = \mathbb{E}(\mathbb{1}_A Z_T^{\theta}) = 0$ The other direction: $\mathbb{P}^{\theta}(A) = 0 \Rightarrow \mathbb{1}_A(\omega) = 0$ a.e. $\mathbb{P}^{\theta}(d\omega) \Rightarrow \mathbb{P}(A) = \mathbb{E}^{\theta}(\mathbb{1}_A(Z_T^{\theta})^{-1}) = 0$ □

Remark. The definition of $\mathbb{P}^{\theta}$ imples that, whenever the expectation is well defined $\mathbb{E}^{\theta}(X) = \mathbb{E}[XZ_T^{\theta}]$

Remark. Using the above remark we can write the following equation $\mathbb{P}(A) = \mathbb{E}(\mathbb{1}_A) = \mathbb{E}(\mathbb{1}_A(Z_T^{\theta})^{-1}Z_T^{\theta})$ $\mathbb{E}^{\theta}(\mathbb{1}_A(Z_T^{\theta})^{-1}$

Definition 4.3. $\hat{W}_t = W_t - \int_0^t \theta_s ds$ (Brownian Motion with drift) The integral in the above definition is adapted, with finite variation

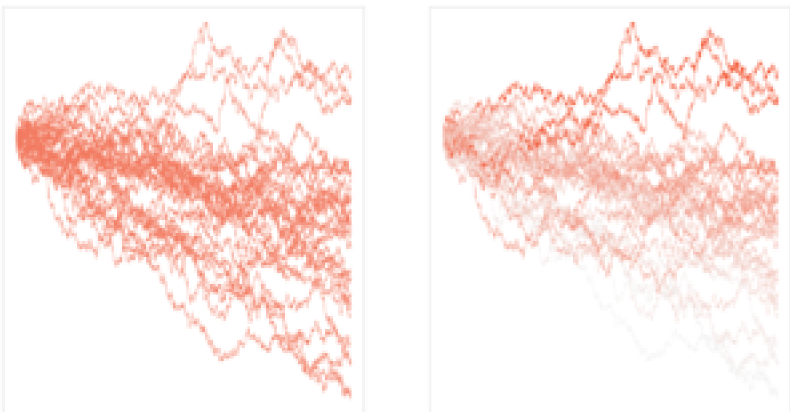
Remark. $\hat{W}_t$ if $\mathcal{F}_t$ - adapted (progressive) but not a Brownian motion under $\mathbb{P}$ (or not an $\mathcal{F}_t$-BM under $\mathbb{P}$),(unless $\theta \equiv 0$) If $\hat{W}$ was a $\mathcal{F}_t$-BM, then $(W_t - \hat{W}_t)$ would be a $\mathcal{F}_t$-martingale. But $\int_0^t \theta_s ds$ has finite variation and this only happens when $\theta \equiv 0$

Proof. The proof was iven in the course when stochastic integrals were introduced □

Example. Take $\theta_s = \theta \times s, \theta \neq 0$ Then: $\lim_{t\hat{}+\infty} \frac{W_t}{t} = 0$, a.s-$d\mathbb{P}$[2] But $\lim_{t\hat{}+\infty} \frac{\hat{W}_t}{t} = \theta \neq 0$a.s-$d\mathbb{P}$ This example shows in particular, that laws of W and $\hat{W}$ under $\mathbb{P}$ (for this particular choice of θ) are **singular**: since $\mathbb{P}\frac{W_t}{t}\ t \overset{\rightarrow}{\rightarrow} \infty\ 0 = 1$, $\mathbb{P}\frac{\hat{W}_t}{t}\ t \overset{\rightarrow}{\rightarrow} \infty\ 0 = 0$ So supported by disjoint sets on $\mathbb{R}_+$

We now come to a very important theorem.

Theorem 4.4 (Girasnov's Theorem).



The graph shows two different measures the one on the right has its measure changed. Let $\theta \in \mathcal{I}_{\text{loc}}(Prog)$ $(T \in (0, +\infty) is fixed)$ We need to verify the Novikov condition. Then, under $\mathbb{P}^{\theta}$, $t \Rightarrow \hat{W}_t = W_t - \int_0^t \theta_s ds$ is a $\mathcal{F}_t - BM$ on $[0, T]$

[2]see Exercise sheet number 1

Remark. A sequence of Girasnov that the laws W and $\hat{W}$ are equivalent on every interval $[0, T]$
Indeed: assume that $\mathbb{P}(W|_{[0,T]} \in A) = 0$
$\mathbb{P}(\hat{W} \in A) = \mathbb{E}^\theta(\mathbb{1}_{(\hat{W} \in A)}(Z_T^\theta)^{-1}) = 0,$
since $\hat{W}$ is a B.M. under $\mathbb{P}^\theta$
If $\mathbb{P}(\hat{W} \in A) = \mathbb{E}^\theta(\mathbb{1}_{\hat{W} \in A}(Z_T^\theta)^{-1})$ since $(Z_T^\theta)^{-1} > 0$ by definition
$\Rightarrow \mathbb{1}_{(\hat{W} \in A)} = 0$ a.e. $d\mathbb{P}^\theta$
$\iff \mathbb{1}_{(W \in A)} = 0$ a.e. $d\mathbb{P}$

Of Girasnov's theorem. We just consider the case "θ_s is deterministic" We know - $\hat{W}_t$ is continuous, $\hat{W}_0 = 0$, $\hat{W}_t$ is $\mathcal{F}_t$-progressive ($\Rightarrow \mathcal{F}_t -$ adapted) Everything we need to show is that

$$\forall 0 \leq t_0 < t_1 < t_2 < \cdots < t_k \leq T$$

$$(\hat{W}_{t_1} - \hat{W}_{t_0}, \cdots, \hat{W}_{t_k} - \hat{W}_{t_k - 1}) \sim \mathcal{N}\left(0, \begin{pmatrix} t_1 - t_0 & 0 & \cdots & 0 & 0 \\ 0 & t_2 - t_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & t_{n-1} - t_{n-2} & 0 \\ 0 & 0 & \cdots & 0 & t_n - t_{n-1} \end{pmatrix}\right).$$

Fix $(\lambda_1, \cdots, \lambda_k) \in \mathbb{R}^k$
$\mathbb{E}^\theta[\exp i \sum_{j=1}^k \lambda_j(\hat{W}_{t_j} - \hat{W}_{t_{j-1}}))](*)$
It is worth pointing out that a Goal of this section is to show that $exp{-}\frac{1}{2}\sum_{j=1}^k \lambda_j^2(t_j - t_{j-1}) = *$
$= \mathbb{E}^\theta[\exp i \sum_{j=1}^k \lambda_j(W_{t_j} - W_{t_{j-1}} - \int_{t_{j-1}}^{t_j}]\theta_s ds = \mathbb{E}(\exp i \sum_{j=1}^k \lambda_j(W_{t_j} - W_{t_{j-1}} - \int_{t_{j-1}}^{t_j} Z_T^\theta)$
(martingale property under $\mathbb{P}$)
$= \mathbb{E}[\exp i \sum_{j=1}^k \lambda_j(W_{t_j} - W_{t_{j-1}} - \int_{t_{j-1}}^{t_j} Z_{t_k}^\theta]$
$=$

$$\mathbb{E}[\exp i \sum_{j=1}^k \lambda_j(W_{t_j} - W_{t_{j-1}} - \int_{t_{j-1}}^{t_j} \theta_s ds) \exp \int_0^{t_k} \theta_s dW_s - \frac{1}{2}\int_0^{t_k} \theta_s^2 ds] \tag{4.2}$$

(we can assume without loss of generatlity that $t_0 = 0$) Now we can write $\int_0^{t_k} \theta_s dW_s$ and $\int_0^{t_k} \theta_s^2 ds = \sum_{j=1}^k \int_{t_{j-1}}^{t_j} \theta_s^2 ds$ So that
4.2$= \mathbb{E}[\exp \sum_{j=1}^k \int_{t_{j-1}}^{t_j(i\lambda_j + \theta_s)dW_s - \sum_{j=1}^k \int_{t_{j-1}}^{t_j} i\lambda_j \theta_s ds - \frac{1}{2}}]$ $\square$