Calcul stochastique et modèles de diffusion

Cyril Labbé

Année universitaire 2023/2024
M2MO : Modélisation Aléatoire, Finance et Data Science
Université Paris Cité

Table des matières

Chapitre I

Mouvement Brownien

Approx. 2 séances de cours

1 Processus à temps continu

Définition I.1

Soit $(E, \mathcal{E})$ un ensemble mesurable, $\mathbb{T}$ un ensemble quelconque et $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité. On dit que $(X_t, t \in \mathbb{T})$ est un processus stochastique sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans $(E, \mathcal{E})$ si pour tout $t \in \mathbb{T}$, X_t est une variable aléatoire de $(\Omega, \mathcal{F})$ dans $(E, \mathcal{E})$.

Lorsque $\mathbb{T} = \mathbb{N}$, on parle de processus à temps discret. Lorsque $\mathbb{T}$ est égal à $\mathbb{R}_+$ ou à un intervalle quelconque de $\mathbb{R}$, on parle de processus à temps continu. L'essentiel de ce cours porte sur les processus stochastiques à temps continu et à valeurs réelles. De façon générique, on supposera dans la suite que $\mathbb{T} = \mathbb{R}_+$, et que $E = \mathbb{R}$ muni de la tribu Borélienne $\mathcal{E} = \mathcal{B}(\mathbb{R})$ (mais toutes les définitions présentées font encore sens si $\mathbb{T}$ est un intervalle borné, etc.).

Etant donné un processus stochastique $(X_t, t \in \mathbb{R}_+)$ et un n-uplet $t_1, \ldots, t_n \in \mathbb{R}_+$, le vecteur $(X_{t_1}, \ldots, X_{t_n})$ forme une variable aléatoire à valeurs dans $\mathbb{R}^n$. Les variables X_t sont appelées marginales uni-dimensionnelles, et les vecteurs $(X_{t_1}, \ldots, X_{t_n})$ marginales finies-dimensionnelles. Ces variables aléatoires peuvent être vues comme des projections d'un objet plus gros : l'application

$$\begin{aligned} X : \Omega &\to \mathbb{R}^{\mathbb{R}_+} \\ \omega &\mapsto (X_t(\omega), t \in \mathbb{R}_+) \end{aligned}$$

qui prend ses valeurs dans un espace de *trajectoires*.

De quelle tribu peut-on munir cet espace de trajectoires ? La réponse la plus simple est : la tribu produit $\mathcal{B}(\mathbb{R})^{\otimes \mathbb{R}_+}$, c'est-à-dire, la plus petite tribu sur $\mathbb{R}^{\mathbb{R}_+}$ rendant mesurables les applications coordonnées $t \mapsto x_t$ pour tout $t \in \mathbb{R}_+$. On notera qu'il s'agit également de la plus petite tribu contenant les cylindres de dimension finie, c'est-à-dire, les sous-ensembles de $\mathbb{R}^{\mathbb{R}_+}$ de la forme

$$\prod_{t \in \mathbb{R}_+} A_t \ , \qquad \text{avec } A_t \in \mathcal{B}(\mathbb{R}_+) \ , \text{ et } \qquad \#\{t \in \mathbb{R}_+ : A_t \neq \mathbb{R}\} < \infty \ .$$

Lemme I.2

Si $(X_t, t \in \mathbb{R}_+)$ est un processus stochastique, alors l'application X définie ci-dessus est mesurable de $(\Omega, \mathcal{F})$ dans $(\mathbb{R}^{\mathbb{R}_+}, \mathcal{B}(\mathbb{R})^{\otimes \mathbb{R}_+})$.

Démonstration Si l'on note $\mathcal{C}$ la classe des cylindres de dimension finie, alors $X^{-1}(\mathcal{C}) \subset \mathcal{F}$ par définition d'un processus stochastique ; or $\sigma(\mathcal{C})$ n'est rien d'autre que la tribu produit et une propriété générale de théorie de la mesure assure que $X^{-1}(\sigma(\mathcal{C})) = \sigma(X^{-1}(\mathcal{C}))$ et ainsi X est mesurable. □

Ainsi un processus stochastique $(X_t, t \in \mathbb{R}_+)$ peut être vu comme *une variable aléatoire* à valeurs dans un espace de trajectoires. La loi $\mathbf{P}$ d'un tel processus stochastique est la mesure image de $\mathbb{P}$ par l'application X. Comme la classe des cylindres est stable par intersection finie, un corollaire du théorème de classe monotone assure que la loi d'un processus stochastique est entièrement caractérisée par la donnée de $\mathbf{P}(C)$ pour tout $C \in \mathcal{C}$. Or si l'on note $t_1, \ldots, t_n$ les indices pour lesquels $A_t \neq \mathbb{R}$ on observe que

$$\mathbf{P}(C) = \mathbb{P}((X_{t_1}, \ldots, X_{t_n}) \in A_{t_1} \times \cdots \times A_{t_n}) \ .$$

Ainsi la loi d'un processus stochastique est entièrement caractérisée par les lois de ses marginales fini-dimensionnelles : si Y est un autre processus stochastique et si pour tout n et tous $t_1, \ldots, t_n \in \mathbb{R}_+$ les vecteurs $(X_{t_1}, \ldots, X_{t_n})$ et $(Y_{t_1}, \ldots, Y_{t_n})$ ont même loi, alors X et Y ont même loi.

A ce stade, il nous faut formuler une observation subtile mais importante. Si $(X_t, t \in \mathbb{R}_+)$ est un processus stochastique, rien n'assure que les ensembles

$$\{\omega : \forall t \in [0,1] : X_t(\omega) \leqslant 10\} \ , \quad \{\omega : t \mapsto X_t(\omega) \text{ est continu}\} \ ,$$

soient mesurables en général. En effet, ces ensembles dépendent d'une collection *non-dénombrables* d'indices $t \in \mathbb{R}_+$, et ne peuvent pas s'écrire a priori comme union/intersection dénombrables d'événements ne faisant intervenir que les marginales uni-dimensionnelles. Autrement dit, les ensembles

$$\{\forall t \in [0,1] : x_t \leqslant 10\} \ , \quad \{t \mapsto x_t \text{ est continu}\} \ ,$$

n'appartiennent pas à la tribu produit et rien ne dit que X reste mesurable si l'on munit l'ensemble des trajectoires d'une tribu contenant ces ensembles.
Nous reviendrons sur la continuité des trajectoires d'un processus stochastique un peu plus loin dans ce chapitre.

La tribu engendrée par un processus stochastique X, parfois notée $\sigma(X)$, est la plus petite tribu rendant mesurables les v.a. X_t, $t \in \mathbb{R}_+$. On dira qu'un processus stochastique X est indépendant d'une tribu $\mathcal{G}$ (sur Ω) si les tribus $\sigma(X)$ et $\mathcal{G}$ sont indépendantes. On peut vérifier que cela est vrai si et seulement si toutes les marginales finies dimensionnelles de X sont indépendantes de la tribu $\mathcal{G}$.

Citons à présent un résultat important d'existence de processus stochastiques. Sa preuve repose sur des arguments délicats de construction de mesures et est omise.

Théorème I.3 (*Théorème d'extension de Kolmogorov*)

Pour toute partie finie I de $\mathbb{R}_+$, soit P_I une mesure de probabilité sur $\mathbb{R}^I$. On suppose que les *mesures $(P_I)_I$ sont compatibles au sens suivant :*

$$(P_J)_{|I} = P_I \ , \quad I \subset J \subset \mathbb{R}_+ \ , \quad J \text{ fini} \ .$$

Alors il existe une unique mesure de probabilité $\mathbf{P}$ sur la tribu produit de $\mathbb{R}^{\mathbb{R}_+}$ telle que $\mathbf{P}_{|I} = P_I$ pour toute partie finie $I \subset \mathbb{R}_+$.

Nous allons appliquer ce théorème pour constuire des processus gaussiens.

Définition I.4

On dit qu'un processus stochastique $(X_t, t \in \mathbb{R}_+)$ est gaussien si toutes ses marginales finies-dimensionnelles forment des vecteurs gaussiens.

Autrement dit, un processus stochastique $(X_t, t \in \mathbb{R}_+)$ est gaussien si pour tous $t_1, \ldots, t_n \in \mathbb{R}_+$ et tous $\lambda_1, \cdots, \lambda_n \in \mathbb{R}$, la variable aléatoire $\sum_{i=1}^n \lambda_i X_{t_i}$ est gaussienne.

Proposition I.5

La loi d'un processus gaussien $(X_t, t \in \mathbb{R}_+)$ est complétement caractérisée par la donnée de sa fonction moyenne et de sa fonction de covariance :

$$t \mapsto \mathbb{E}[X_t] \ , \quad (s,t) \mapsto \mathrm{Cov}(X_s, X_t) \ .$$

Démonstration Il est bien connu que la loi d'un vecteur gaussien $(X_{t_1}, \ldots, X_{t_n})$ est caractérisée par son vecteur moyenne et sa matrice de covariance. Or ces deux quantités sont complétement décrites par les fonctions introduites dans l'énoncé. Pour conclure, on rappelle que la loi d'un processus stochastique est complétement caractérisée par les lois de ses marginales finies-dimensionnelles. □

Réciproquement, on se donne une fonction $t \mapsto m_t$ de $\mathbb{R}_+$ dans $\mathbb{R}$ ainsi qu'une fonction $\Gamma : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ vérifiant :

$$\Gamma(s,t) = \Gamma(t,s) \ , \quad \sum_{i,j=1}^n \lambda_i \lambda_j \Gamma(t_i, t_j) \geqslant 0 \ , \quad \forall n \geqslant 1 \ , \lambda_i \in \mathbb{R} \ , t_i \in \mathbb{R}_+ \ .$$

Le théorème d'extension de Kolmogorov assure qu'il existe un processus gaussien de fonction moyenne égal à m et de fonction de covariance donnée par Γ. En effet, pour tout ensemble fini $I \subset \mathbb{R}_+$ il est bien connu que la loi normale $\mathcal{N}(m_{|I}, \Gamma_{|I})$ existe, où $m_{|I}$ et $\Gamma_{|I}$ sont les restrictions à I et $I \times I$ des fonctions m et Γ. Par ailleurs, il est facile de vérifier la propriété de consistence : la restriction de la loi normale $\mathcal{N}(m_{|J}, \Gamma_{|J})$ aux indices I est la loi normale $\mathcal{N}(m_{|I}, \Gamma_{|I})$, pour tout $I \subset J \subset \mathbb{R}_+$ avec J fini.

On notera que si X est un processus gaussien de moyenne m et de fonction de covariance Γ, alors $(X_t - m_t, t \in \mathbb{R}_+)$ est un processus gaussien centré de même fonction de covariance.

2 Définition du mouvement Brownien

Définition I.6

Un processus stochastique $(B_t, t \in \mathbb{R}_+)$ est appelé mouvement Brownien (issu de 0) si :

1. *B est un processus gaussien centré de fonction de covariance*

$$\mathrm{Cov}(B_s, B_t) = \mathbb{E}[B_s B_t] = \min(s,t) \ , \quad \forall s,t \in \mathbb{R}_+ \ ,$$

2. *les trajectoires de B sont continues.*

Quelques remarques s'imposent. Tout d'abord, un tel processus vérifie $B_0 = 0$ p.s. car toute variable gaussienne de variance nulle est égale p.s. à sa moyenne. Par ailleurs, par la Proposition I.5 la loi d'un processus vérifiant cette définition est unique. On pourra parler de la loi du mouvement Brownien sans ambiguité.
L'existence d'un processus vérifiant la première condition a déjà été démontrée : en effet, il n'est pas difficile de vérifier que la fonction de covariance de la définition est bien symétrique positive et l'on peut alors appliquer le théorème d'extension de Kolmogorov comme expliqué à la section précédente. En revanche la continuité d'un tel processus n'est pas acquise : ce faisant, l'existence du mouvement Brownien reste à établir. Ce sera l'objet des deux sections suivantes.
Notons enfin qu'il existe des processus gaussiens qui ne sont pas continus. Par exemple, si X est un processus gaussien centré de fonction de covariance $\Gamma(s,t) = \min(s,t)\mathbf{1}_{\{s,t \in \mathbb{Q}_+\}}$, alors X est constant égal à 0 aux temps irrationnels et a une variance non-triviale aux temps rationnels : s'il était continu, il serait constant égal à 0 partout.

Commençons par manipuler les propriétés présentées dans la définition.

Proposition I.7

Soit $(B_t, t \in \mathbb{R}_+)$ un processus stochastique. Il y a équivalence entre :

(i) *B est un processus gaussien centré de fonction de covariance* $\mathrm{Cov}(B_s, B_t) = \min(s,t)$,

(ii) *B vérifie :*

 (a) *$B_0 = 0$ p.s.,*

 (b) *pour tout $n \geqslant 2$, pour tous $0 \leqslant t_1 < t_2 < \ldots < t_n$, les variables aléatoires B_{t_1}, $B_{t_2} - B_{t_1}$, $\ldots$, $B_{t_n} - B_{t_{n-1}}$ sont indépendantes,*

 (c) *pour tous $0 \leqslant s \leqslant t$, la variable $B_t - B_s$ suit la loi gaussienne $\mathcal{N}(0, t-s)$*

Ainsi tout processus B dont les trajectoires sont continues et qui satisfait les propriétés (a), (b) et (c) est un mouvement Brownien.

Démonstration Supposons que B vérifie (i). On a déjà vu plus haut que $B_0 = 0$ p.s, d'où le point (a). Par ailleurs la variable $B_t - B_s$ est gaussienne centrée de variance

$$\mathrm{Var}(B_t - B_s) = \mathrm{Var}(B_t, B_t) - 2\,\mathrm{Cov}(B_s, B_t) + \mathrm{Var}(B_s, B_s) = t - s \ ,$$

ce qui prouve le point (c). En ce qui concerne les point (b), le vecteur $(B_{t_1}, B_{t_2} - B_{t_1}, \cdots, B_{t_n} - B_{t_{n-1}})$ étant gaussien, il suffit de vérifier que sa matrice de covariance est diagonale (exercice).
Supposons que B vérifie (ii). Les points (b) et (c) assurent que pour tout $n \geqslant 2$ et pour tous $0 \leqslant t_1 < t_2 < \ldots < t_n$, les variables $B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}}$ sont indépendantes et gaussiennes centrées de variances t_1, $t_2 - t_1$, $\cdots$, $t_n - t_{n-1}$; ces variables forment donc un vecteur gaussien de matrice de covariance diagonale avec pour entrées sur la diagonale les variances précédentes. La stabilité sous transformation linéaire des vecteurs gaussiens assure que $(B_{t_1}, \cdots, B_{t_n})$ forme un

vecteur gaussien centré tel que pour tous $1 \leqslant i \leqslant j \leqslant n$

$$\begin{aligned}
\operatorname{Cov}(B_{t_i}, B_{t_j}) &= \sum_{k=1}^{i} \sum_{\ell=1}^{j} \operatorname{Cov}(B_{t_k} - B_{t_{k-1}}, B_{t_\ell} - B_{t_{\ell-1}}) \\
&= \sum_{k=1}^{i} \operatorname{Var}(B_{t_k} - B_{t_{k-1}}) \\
&= t_i = \min(t_i, t_j) \ .
\end{aligned}$$

On a donc prouvé que B est un processus gaussien centré de fonction de covariance $\operatorname{Cov}(B_s, B_t) = \min(s, t)$. $\square$

Corollaire I.8

Soit B un mouvement Brownien. Pour tout $n \geqslant 1$ et tous $0 \leqslant t_1 < \cdots < t_n$, la loi du vecteur $(B_{t_1}, B_{t_2}, \ldots, B_{t_n})$ admet une densité sur $\mathbb{R}^n$ donnée par

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

avec $x_0 = 0$ et $(x_1, \ldots, x_n) \in \mathbb{R}^n$.

Démonstration Comme on l'a vu dans la preuve précédente, le vecteur $(B_{t_1}, B_{t_2} - B_{t_1}, \cdots, B_{t_n} - B_{t_{n-1}})$ est composé de v.a. gaussiennes centrées indépendantes, il est donc gaussien et admet pour densité

$$p(y_1, \ldots, y_n) = \frac{1}{(2\pi)^{n/2}\sqrt{t_1(t_2 - t_1)\cdots(t_n - t_{n-1})}} \exp\left(- \sum_{i=1}^{n} \frac{(y_i - y_{i-1})^2}{2(t_i - t_{i-1})} \right) \ .$$

En appliquant le changement de variables $x_i = y_1 + \cdots + y_i$ (de Jacobien égal à 1) on obtient le résultat voulu. $\square$

Pour l'instant, nous n'avons pas établi l'existence du mouvement Brownien : la difficulté réside dans la continuité des trajectoires, que nous examinons à présent.

3 Continuité des trajectoires

Afin de prouver l'existence du mouvement Brownien, nous allons nous appuyer sur un résultat général dû à Kolmogorov qui affirme que, sous une hypothèse sur les moments d'un processus stochastique, on peut "modifier" le processus pour le rendre continu. Commençons par introduire cette notion de modification :

Définition I.9

Soient deux processus stochastiques $(X_t, t \in I)$ et $(\tilde{X}_t, t \in I)$ indexés par un ensemble $I \subset \mathbb{R}$, et définis sur un même espace de probabilité. On dit que $\tilde{X}$ est une modification de X si

$$\forall t \in I, \quad \mathbb{P}(X_t = \tilde{X}_t) = 1 \ .$$

On notera que si $\tilde{X}$ est une modification de X, alors les marginales de dimension finie des deux processus coïncident et ainsi les deux processus ont même loi. En revanche, il se peut que les trajectoires des deux processus soient très différentes ! En effet, l'égalité de la définition ne fournit aucun contrôle sur des collections *non-dénombrables* d'indices $t \in I$. Pour cela, il faut imposer une condition plus forte, ce qui donne lieu à la notion suivante.

Définition I.10

Soient deux processus stochastiques $(X_t, t \in I)$ et $(\tilde{X}_t, t \in I)$ indexés par un ensemble $I \subset \mathbb{R}$, et définis sur un même espace de probabilité. On dit que X et $\tilde{X}$ sont indistinguables si

$$\mathbb{P}(\forall t \in I : X_t = \tilde{X}_t) = 1 \ .$$

Remarque I.11

Une formulation plus rigoureuse est : On dit que X et $\tilde{X}$ sont indistinguables s'il existe un sous-ensemble $N \subset \Omega$ qui est négligable, c'est-à-dire, N est inclus dans un événement de mesure nulle et tel que

$$\forall \omega \in \Omega \backslash N, \forall t \in I, \quad X_t(\omega) = \tilde{X}_t(\omega) \ .$$

En supposant N négligeable, on relâche la contrainte de mesurabilité de l'ensemble des réalisations où les deux processus diffèrent. Il s'agit là d'un point de détail.

Il est facile de vérifier que si deux processus sont indistinguables, alors l'un est une modification de l'autre.

Théorème I.12 (*Théorème de continuité de Kolmogorov*)

Soit $(X_t, t \in I)$ un processus stochastique à valeurs réelles indicé par un intervalle borné $I \subset \mathbb{R}$. *Supposons qu'il existe des réels $q, \epsilon, C > 0$ tels que pour tous $s, t \in I$*

$$\mathbb{E}[|X_s - X_t|^q] \leqslant C|t-s|^{1+\epsilon} \ .$$

Alors il existe une modification $\tilde{X}$ de X dont les trajectoires sont Hölderiennes d'exposant α pour tout $\omega \in \Omega$ et tout $\alpha \in]0, \epsilon/q[$, c'est-à-dire, pour tout $\alpha \in]0, \epsilon/q[$ il existe une variable aléatoire positive C_α telle que pour tous $s, t \in I$ et tout $\omega \in \Omega$

$$|\tilde{X}_s(\omega) - \tilde{X}_t(\omega)| \leqslant C_\alpha(\omega)|t-s|^\alpha \ .$$

Dans la preuve, nous aurons besoin du lemme technique suivant. Soit D l'ensemble dénombrable des nombres dyadiques de l'intervalle $[0, 1]$.

Lemme I.13

Soit $f : [0,1] \to \mathbb{R}$. On suppose qu'il existe $\alpha, K > 0$ tels que pour tout entier $n \geqslant 1$ et tout $i \in \{1, \ldots, 2^n\}$

$$|f((i-1)2^{-n}) - f(i2^{-n})| \leqslant K2^{-n\alpha} \ .$$

Alors pour tous $s, t \in D$

$$|f(s) - f(t)| \leqslant \frac{2K}{1-2^{-\alpha}}|t-s|^\alpha \ .$$

Démonstration Soient $s, t \in D$ tels que $s < t$. On note $p \geqslant 0$ l'unique entier tel que $2^{-p} \leqslant t-s < 2^{-(p-1)}$. Nécessairement il existe k, ℓ, m entiers tels que

$$\begin{aligned}
s &= k2^{-p} - \epsilon_{p+1}2^{-p-1} - \cdots - \epsilon_{p+\ell}2^{-p-\ell} \;,\\
t &= k2^{-p} + \epsilon'_p 2^{-p} + \epsilon'_{p+1}2^{-p-1} + \cdots + \epsilon'_{p+m}2^{-p-m} \;.
\end{aligned}$$

où les ϵ_i, ϵ'_j valent 0 ou 1. On pose

$$\begin{aligned}
s_i &= k2^{-p} - \epsilon_{p+1}2^{-p-1} - \cdots - \epsilon_{p+i}2^{-p-i} \;, \quad i \in \{0, \ldots, \ell\}\\
t_j &= k2^{-p} + \epsilon'_p 2^{-p} + \epsilon'_{p+1}2^{-p-1} + \cdots + \epsilon'_{p+j}2^{-p-j} \;, \quad j \in \{0, \ldots, m\} \;.
\end{aligned}$$

On observe que $s_\ell = s$ et $t_m = t$. Ainsi l'hypothèse du lemme assure que

$$\begin{aligned}
|f(s) - f(t)| &= |f(s_\ell) - f(t_m)|\\
&\leqslant |f(s_0) - f(t_0)| + \sum_{i=1}^{\ell} |f(s_{i-1}) - f(s_i)| + \sum_{j=1}^{m} |f(t_{j-1}) - f(t_j)|\\
&= K\Big(2^{-p\alpha} + \sum_{i=1}^{\ell} 2^{-(p+i)\alpha} + \sum_{j=1}^{m} 2^{-(p+j)\alpha}\Big)\\
&\leqslant 2K2^{-p\alpha}(1-2^{-\alpha})^{-1} \leqslant 2K(t-s)^\alpha(1-2^{-\alpha})^{-1} \;.
\end{aligned}$$

□

Présentons à présent la preuve du théorème. On va se restreindre aux indices de temps dyadiques et montrer que le processus X, restreint aux dyadiques, est continu p.s. (cela fait sens car on ne manipule qu'un nombre dénombrable de v.a.). Puis on étendra cette définition à tous les indices de temps par densité des dyadiques et continuité. Il ne restera plus qu'à vérifier que l'objet ainsi construit est bien une modification du processus de départ.

Démonstration [du Théorème de continuité de Kolmogorov] Sans perte de généralité, on peut supposer que $I = [0,1]$. On fixe $\alpha \in]0, \epsilon/q[$. L'hypothèse de l'énoncé combinée à l'inégalité de Markov assure que pour tout $n \geqslant 1$ et pour tout $i \in \{1, \ldots, 2^n\}$

$$\mathbb{P}(|X_{(i-1)2^{-n}} - X_{i2^{-n}}| > 2^{-n\alpha}) \leqslant 2^{n\alpha q}\mathbb{E}[|X_{(i-1)2^{-n}} - X_{i2^{-n}}|^q] \leqslant C2^{-n(1+\epsilon)}2^{n\alpha q} \;.$$

Ainsi

$$\mathbb{P}\left(\bigcup_{i=1}^{2^n}\{|X_{(i-1)2^{-n}} - X_{i2^{-n}}| > 2^{-n\alpha}\}\right) \leqslant C2^{-n\epsilon}2^{n\alpha q} \;.$$

Comme $\epsilon - \alpha q > 0$ on obtient

$$\sum_{n\geqslant 1} \mathbb{P}\left(\bigcup_{i=1}^{2^n}\{|X_{(i-1)2^{-n}} - X_{i2^{-n}}| > 2^{-n\alpha}\}\right) < \infty \;.$$

Le lemme de Borel-Cantelli assure donc que pour $\mathbb{P}$-presque tout $\omega \in \Omega$, il existe un entier $n_0 = n_0(\omega)$ tel que

$$\forall n \geqslant n_0, \forall i \in \{1, \ldots, 2^n\}, \quad |X_{(i-1)2^{-n}} - X_{i2^{-n}}| \leqslant 2^{-n\alpha} \;.$$

Ceci assure que la variable aléatoire

$$K_\alpha := \sup_{n\geqslant 1} \sup_{i\in\{1,\dots,2^n\}} \frac{|X_{(i-1)2^{-n}} - X_{i2^{-n}}|}{2^{-n\alpha}} ,$$

est finie presque sûrement. (En effet pour $n \geqslant n_0(\omega)$ cette quantité est majorée par 1, tandis que le supremum sur $n < n_0$ porte sur un nombre fini de termes.)
Sur l'événement $\{K_\alpha < \infty\}$, qui est de probabilité 1, on déduit du lemme précédent que pour tous $s, t \in D$

$$|X_s(\omega) - X_t(\omega)| \leqslant C_\alpha(\omega)|t-s|^\alpha ,$$

avec $C_\alpha(\omega) = 2K_\alpha(\omega)/(1-2^{-\alpha})$. On a donc montré que sur l'événement $\{K_\alpha < \infty\}$, les trajectoires de X restreintes à l'ensemble D sont α-Höldériennes. On peut alors poser sans ambiguité

$$\tilde{X}_t(\omega) := \begin{cases} \lim_{s\to t, s\in D} X_s(\omega) & \text{si} K_\alpha(\omega) < \infty \\ 0 & \text{si} K_\alpha(\omega) = \infty . \end{cases}$$

Il reste à montrer que $\tilde{X}$ est une modification de X. Si $t \in D$ alors

$$\mathbb{P}(\tilde{X}_t \neq X_t) \leqslant \mathbb{P}(K_\alpha = \infty) = 0 .$$

Supposons à présent que $t \notin D$. L'hypothèse du théorème assure que pour tout $s \in [0,1]$

$$\mathbb{E}[1 \wedge |X_t - X_s|^q] \leqslant \mathbb{E}[|X_t - X_s|^q] \leqslant C|t-s|^{1+\epsilon} .$$

Pour $s \in D$, on sait déjà que $X_s = \tilde{X}_s$ p.s. de sorte que

$$\mathbb{E}[1 \wedge |X_t - \tilde{X}_s|^q] \leqslant C|t-s|^{1+\epsilon} .$$

En faisant tendre $s \to t$ avec $s \in D$, le théorème de convergence dominée assure que

$$\mathbb{E}[1 \wedge |X_t - \tilde{X}_t|^q] = 0 ,$$

et ainsi $X_t = \tilde{X}_t$ p.s. □

Corollaire I.14

Le mouvement Brownien existe !

Démonstration Soit B un processus gaussien centré de fonction de covariance $\mathrm{Cov}(B_s, B_t) = \min(s,t)$. (L'existence d'un tel processus a déjà été discutée plus haut). Afin d'appliquer le théorème de continuité de Kolmogorov, il nous faut évaluer les moments des incréments de ce processus. On remarque que $B_t - B_s$ a même loi que $\sqrt{t-s}\,Y$ où Y suit une loi normale centrée réduite. Ainsi en posant $C_q := \mathbb{E}[|Y|^q]$ qui est fini (les gaussiennes admettent des moments de tous ordres) on obtient

$$\mathbb{E}[|B_t - B_s|^q] = C_q(t-s)^{q/2} .$$

En choisissant $q > 2$ et un intervalle borné $I \subset \mathbb{R}_+$ arbitraire, on déduit du théorème l'existence d'une modification $(\tilde{B}_t, t \in I)$ dont les trajectoires sont Höldériennes donc continues. En appliquant cette construction sur chacun des intervalles $[0,1]$, $[1,2]$, $\cdots$, $[n, n+1]$, $\cdots$ on obtient alors une collection de modifications que l'on peut concaténer afin de former un processus $(\tilde{B}_t, t \in \mathbb{R}_+)$ à trajectoires continues et qui est une modification de $(B_t, t \in \mathbb{R}_+)$. □

Construction du mouvement Brownien par Paul Lévy. Nous venons d'appliquer un théorème général de la théorie des processus afin de construire le mouvement Brownien. Il se trouve que l'on peut donner une construction beaucoup plus géométrique (et élégante !) du mouvement Brownien. Celle-ci est due à Paul Lévy et s'appuie sur l'observation suivante

Lemme I.15

Pour tout $0 \leqslant s < t$, le vecteur $(B_s, B_{(s+t)/2}, B_t)$ a même loi que le vecteur

$$(B_s, \frac{B_s + B_t}{2} + Y, B_t) \ ,$$

où Y est une variable aléatoire $\mathcal{N}(0,(t-s)/4)$ indépendante de B_s et B_t.

Démonstration Exercice. □

On se restreint à l'intervalle $[0,1]$, on note $D_n := \{k2^{-n} : k \in \{0,\ldots,2^n\}\}$ et l'on construit de façon récursive une suite de processus $(B^{(n)}_t, t \in [0,1])$ tels que :

1. Les marginales de dimensions finies de $B^{(n)}$ restreint à D_n coïncident avec celles du Brownien. Plus précisément : les vecteurs $(B^{(n)}_t, t \in D_n)$ et $(B_t, t \in D_n)$ ont même loi.
2. La procédure est consistente sur les dyadiques. Plus précisément : $B^{(n)}_t = B^{(m)}_t$ pour tout $t \in D_n \cap D_m$.
3. Chaque processus $B^{(n)}$ est affine entre deux points consécutifs de D_n. Plus précisément :

$$B^{(n)}_t = 2^n(t - k2^{-n})B^{(n)}_{(k+1)2^{-n}} + 2^n((k+1)2^{-n} - t)B^{(n)}_{k2^{-n}} \ , \quad t \in [k2^{-n}, (k+1)2^{-n}] \ .$$

On commence la procédure en se donnant une variable aléatoire X de loi $\mathcal{N}(0,1)$ et l'on pose

$$B^{(0)}_t := tX \ .$$

On suppose avoir construit le processus $B^{(n)}$ pour un certain entier $n \geqslant 0$. Pour construire $B^{(n+1)}$ il suffit de définir correctement les marginales aux temps $t \in D_{n+1} \backslash D_n$ puis d'interpoler linéairement entre les points de D_{n+1}. Pour ce faire, on se donne une collection de v.a. $X_{k,n}$, $k \in \{0,\ldots,2^n - 1\}$ i.i.d. $\mathcal{N}(0, 2^{-n}/4)$, indépendantes de toutes les v.a. utilisées jusqu'alors dans la construction. Pour chaque $k \in \{0,\ldots,2^n-1\}$, on pose

$$B^{(n+1)}_{(2k+1)2^{-(n+1)}} := B^{(n)}_{(2k+1)2^{-(n+1)}} + X_{k,n} \ ,$$

tandis que pour tout $t \in D_n$, on pose

$$B^{(n+1)}_t = B^{(n)}_t \ .$$

Il est clair que $(B^{(n+1)}_t, t \in D_{n+1})$ est un vecteur gaussien centré. Pour s'assurer qu'il a la bonne loi, il suffit de calculer les covariances entre les entrées de ce vecteur (exercice). On peut ensuite achever la construction de $B^{(n+1)}$ en interpolant de façon affine entre deux points consécutifs de D_{n+1}.

On peut alors montrer qu'avec probabilité 1, la suite de fonctions continues $(B^{(n)}_t, t \in [0,1])$ converge uniformément vers un objet limite que l'on note $(B_t, t \in [0,1])$.

Nécessairement B est continu. Par construction, ses marginales de dimensions finies restreintes à D coïncident en loi avec celle du mouvement Brownien. Pour conclure, on utilise le lemme suivant.

Lemme I.16

Soient $(X_t, t \in [0,1])$ et $(Y_t, t \in [0,1])$ deux processus stochastiques dont les trajectoires sont continues. Supposons qu'il existe un ensemble $J \subset [0,1]$ dense dans $[0,1]$ tel que les marginales de dimensions finies de X et Y, restreintes à J, coïncident en loi. Alors les deux processus $(X_t, t \in [0,1])$ et $(Y_t, t \in [0,1])$ ont même loi.

Démonstration Il suffit de montrer que pour tout $n \geqslant 1$ et tous $0 \leqslant t_1 \leqslant t_2 \leqslant \ldots \leqslant t_n$, les lois des vecteurs $(X_{t_1}, \ldots, X_{t_n})$ et $(Y_{t_1}, \ldots, Y_{t_n})$ coïncident.
On sait que cela est vrai dès que tous les indices de temps appartiennent à J. Pour conclure, il suffit de montrer que le vecteur aléatoire $(X_{s_1}, \ldots, X_{s_n})$ (resp. $(Y_{s_1}, \ldots, Y_{s_n})$) converge en loi vers $(X_{t_1}, \ldots, X_{t_n})$ (resp. $(Y_{t_1}, \ldots, Y_{t_n})$) dès que $s_i \to t_i$ pour tout $i \in \{1, \ldots, n\}$ avec les $s_i \in J$. Or, sous cette hypothèse, la continuité des trajectoires assure que $(X_{s_1}, \ldots, X_{s_n})$ converge presque sûrement vers $(X_{t_1}, \ldots, X_{t_n})$, et ainsi la convergence a également lieu en loi. □

4 Quelques propriétés du mouvement Brownien

Dans cette partie, $B = (B_t, t \in \mathbb{R}_+)$ désignera toujours un mouvement Brownien. Pour tout $t \geqslant 0$, on note $\mathcal{F}_t := \sigma(B_s, s \in [0,t])$: intuitivement, il s'agit de la plus petite tribu contenant toute l'information encodée par la trajectoire Brownienne jusqu'au temps t. On introduit également $\mathcal{F}_\infty := \sigma(B_s, s \geqslant 0)$.

Proposition I.17 (*Propriétés d'invariance*)

1. *$-B$ est un mouvement Brownien [invariance par changement de signe],*
2. *pour tout $\lambda > 0$, le processus B^λ, défini par $B_t^\lambda := \lambda^{-1} B_{t\lambda^2}$, est un mouvement Brownien [invariance par changement d'échelle],*
3. *pour tout $T \geqslant 0$, le processus $B^{(T)}$, défini par $B_t^{(T)} := B_{T+t} - B_T$, est un mouvement Brownien [invariance par translation]*

Démonstration Dans les trois cas, il est immédiat que les processus sont gaussiens centrés : en effet, toute combinaison linéaire de marginales de dimension finie du processus $-B$, resp. B^λ, resp. $B^{(T)}$, est combinaison linéaire de marginales de dimension finie du processus B, et ainsi, est une gaussienne centrée. Par ailleurs, la continuité des trajectoires est préservée dans chaque cas. Il reste à calculer la fonction de covariance dans chaque cas.

$$\begin{aligned}
\mathrm{Cov}(-B_t, -B_s) &= \mathrm{Cov}(B_t, B_s) = \min(s,t) \ , \\
\mathrm{Cov}(B_t^\lambda, B_s^\lambda) &= \lambda^{-2}\,\mathrm{Cov}(B_{t\lambda^2}, B_{s\lambda^2}) = \lambda^{-2} \min(t\lambda^2, s\lambda^2) = \min(t,s) \ , \\
\mathrm{Cov}(B_t^{(T)}, B_s^{(T)}) &= \mathrm{Cov}(B_{T+t} - B_T, B_{T+s} - B_T) \\
&= \mathrm{Cov}(B_{T+t}, B_{T+s}) - \mathrm{Cov}(B_T, B_{T+s}) - \mathrm{Cov}(B_T, B_{T+t}) + \mathrm{Cov}(B_T, B_T) \\
&= \min(T+s, T+t) + T - 2T = \min(s,t) \ .
\end{aligned}$$

□

La propriété d'invariance par translation s'agrémente d'une propriété d'indépendance.

Proposition I.18 (*Propriété de Markov simple*)

Pour tout $T \geqslant 0$, le processus $B^{(T)}$, défini par $B_t^{(T)} := B_{T+t} - B_T$, est un mouvement Brownien *indépendant de la tribu $\mathcal{F}_T$.*

Démonstration Nous avons déjà montré que $B^{(T)}$ était un mouvement Brownien. Il nous faut prouver qu'il est indépendant de $\mathcal{F}_T$. Rappelons que cela signifie que les tribus engendrées par $(B_s, s \in [0,T])$ d'une part et par $(B_{T+t} - B_T, t \in \mathbb{R}_+)$ d'autre part sont indépendantes. Pour prouver cette assertion, il suffit de montrer que pour tous $n, m \geqslant 1$ et tous $0 \leqslant s_1 \leqslant s_2 \leqslant \ldots \leqslant s_n \leqslant T$, $0 \leqslant t_1 \leqslant t_2 \leqslant \ldots \leqslant t_m$ les vecteurs

$$(B_{s_1}, \ldots, B_{s_n}) \ , \quad (B_{T+t_1} - B_T, \ldots, B_{T+t_m} - B_T) \ ,$$

sont indépendants. Comme ces deux vecteurs forment conjointement un vecteur gaussien dans $\mathbb{R}^{n+m}$, l'indépendance des deux vecteurs est équivalente aux identités

$$\mathrm{Cov}(B_{s_i}, B_{T+t_j} - B_T) = 0 \ , \quad \forall i \in \{1, \ldots, n\} \ , \forall j \in \{1, \ldots, m\} \ .$$

On vérifie aisément que ces identités sont vraies. □

Le résultat suivant montre que le comportement “juste après” le temps 0 suit la loi du tout ou rien.

Proposition I.19

La tribu

$$\mathcal{F}_{0+} := \bigcap_{s>0} \mathcal{F}_s \ ,$$

est grossière au sens suivant : pour tout $A \in \mathcal{F}_{0+}$, $\mathbb{P}(A) = 0$ ou 1.

Démonstration Il suffit de montrer que $\mathcal{F}_{0+}$ est indépendante de $\mathcal{F}_\infty$. En effet, comme $\mathcal{F}_{0+} \subset \mathcal{F}_\infty$, cela permet de déduire que $\mathcal{F}_{0+}$ est indépendante d'elle-même, et est donc grossière.
Soient $n \geqslant 1$, $0 < t_1 < \ldots < t_n$ et $f : \mathbb{R}^n \to \mathbb{R}$ une fonction continue bornée. La continuité des trajectoires combinée au Théorème de Convergence Dominée assure que pour tout $A \in \mathcal{F}_{0+}$

$$\mathbb{E}[\mathbf{1}_A f(B_{t_1}, \ldots, B_{t_n})] = \lim_{\epsilon \to 0} \mathbb{E}[\mathbf{1}_A f(B_{t_1} - B_\epsilon, \ldots, B_{t_n} - B_\epsilon)] \ .$$

Or $A \in \mathcal{F}_\epsilon$ et la propriété de Markov simple assure que le vecteur $(B_{t_1} - B_\epsilon, \ldots, B_{t_n} - B_\epsilon)$ est indépendant de $\mathcal{F}_\epsilon$. Ainsi

$$\mathbb{E}[\mathbf{1}_A f(B_{t_1}, \ldots, B_{t_n})] = \lim_{\epsilon \to 0} \mathbb{P}(A) \mathbb{E}[f(B_{t_1} - B_\epsilon, \ldots, B_{t_n} - B_\epsilon)] = \mathbb{P}(A) \mathbb{E}[f(B_{t_1}, \ldots, B_{t_n})] \ .$$

On a donc montré que $\mathcal{F}_{0+}$ est indépendante de $\sigma(B_t, t > 0)$. Par ailleurs, B_0 étant déterministe, il est mesurable par rapport à n'importe quelle tribu et ainsi $\sigma(B_t, t > 0) = \sigma(B_t, t \geqslant 0) = \mathcal{F}_\infty$, ce qui achève la preuve. □

On déduit de cette dernière proposition des propriétés trajectorielles remarquables.

Corollaire I.20

On a presque sûrement

$$\forall \epsilon > 0, \quad \sup_{0 \leqslant s \leqslant \epsilon} B_s > 0 \ , \quad \inf_{0 \leqslant s \leqslant \epsilon} B_s < 0 \ .$$

Par ailleurs pour tout $a \in \mathbb{R}$, si l'on note $T_a := \inf\{t \geqslant 0 : B_t = a\}$ alors presque sûrement

$$\forall a \in \mathbb{R}, \quad T_a < \infty \ .$$

En conséquence presque sûrement

$$\limsup_{t\to\infty} B_t = +\infty \ , \quad \liminf_{t\to\infty} B_t = -\infty \ .$$

Démonstration Soit $(\epsilon_n)_n$ une suite de réels strictement positifs décroissant vers 0. On pose

$$A := \bigcap_n \{ \sup_{0\leqslant s\leqslant \epsilon_n} B_s > 0\} \ .$$

L'événement A est $\mathcal{F}_{0+}$-mesurable et est une intersection d'événements décroissants, de sorte que

$$\mathbb{P}(A) = \lim_{n\to\infty} \mathbb{P}(\sup_{0\leqslant s\leqslant \epsilon_n} B_s > 0) \ .$$

Or

$$\mathbb{P}(\sup_{0\leqslant s\leqslant \epsilon_n} B_s > 0) \geqslant \mathbb{P}(B_{\epsilon_n} > 0) = 1/2 \ ,$$

et ainsi la probabilité de A est supérieure ou égale à 1/2. Cet événement étant trivial, sa probabilité est donc égale à 1. On a donc montré que presque sûrement $\forall \epsilon > 0$, $\sup_{0\leqslant s\leqslant\epsilon} B_s > 0$. L'assertion concernant l'infimum est obtenue en remplaçant B par $-B$.
On prouve à présent la finitude p.s. de tous les T_a. On commence par écrire

$$1 = \mathbb{P}(\sup_{0\leqslant s\leqslant 1} B_s > 0) = \lim_{\delta\downarrow 0} \mathbb{P}(\sup_{0\leqslant s\leqslant 1} B_s > \delta) \ .$$

On fixe $a > 0$. En posant $\lambda = \delta/a$ et en utilisant l'invariance d'échelle du mouvement Brownien on obtient

$$\mathbb{P}(\sup_{0\leqslant s\leqslant 1} B_s > \delta) = \mathbb{P}(\sup_{0\leqslant s\leqslant 1} \lambda^{-1} B_s > \lambda^{-1}\delta) = \mathbb{P}(\sup_{0\leqslant s\leqslant \lambda^{-2}} \lambda^{-1} B_{s\lambda^2} > a) = \mathbb{P}(\sup_{0\leqslant s\leqslant \lambda^{-2}} B_s > a) \ .$$

En prenant la limite $\delta \downarrow 0$, c'est-à-dire, $\lambda^{-2} \uparrow \infty$ on obtient que

$$\mathbb{P}(\sup_{s\geqslant 0} B_s > a) = 1 \ .$$

On a donc montré que $T_a < \infty$ p.s. On en déduit que presque sûrement, pour tout $a \in (0,\infty) \cap \mathbb{Q}$, $T_a < \infty$. Or la continuité des trajectoires assure que pour tout $0 < a < b$, $T_b < \infty \Rightarrow T_a < \infty$. On a donc montré que presque sûrement pour tout $a > 0$, $T_a < \infty$.
En remplaçant B par $-B$, on en déduit que cela reste vrai pour $a < 0$.
On a donc montré que presque sûrement la trajectoire $t \mapsto B_t$ visite toutes les valeurs $a \in \mathbb{R}$. Nécessairement $\limsup_{t\to\infty} B_t = +\infty$ et $\liminf_{t\to\infty} B_t = -\infty$. □

5 Propriété de Markov forte du mouvement Brownien

Nous allons montrer que la propriété de Markov simple reste vraie lorsque le temps T est aléatoire mais n'anticipe pas le futur, c'est-à-dire, dès que T est un temps d'arrêt.

Définition I.21

On dit qu'une variable aléatoire T à valeurs dans $[0,\infty]$ est un temps d'arrêt si pour tout $t \geqslant 0$, *l'événement* $\{T \leqslant t\} \in \mathcal{F}_t$.

En termes intuitifs, cette définition impose que l'information encodée dans la tribu $\mathcal{F}_t$ suffit à déterminer si T s'est réalisé avant le temps t ou pas.

Donnons quelques exemples et contre-exemples de temps d'arrêt. La v.a. constante $T = t_0$ est un temps d'arrêt, ainsi que le premier temps d'atteinte de a par le Brownien

$$T_a := \inf\{t \geqslant 0 : B_t = a\} .$$

En revanche, le dernier temps d'atteinte de 0 avant le temps 1 par le Brownien

$$\sup\{t \leqslant 1 : B_t = 0\} ,$$

ainsi que le premier temps où la trajectoire Brownienne atteint son maximum sur $[0, 1]$

$$\inf\{t \in [0, 1] : B_t = \sup_{s\in[0,1]} B_s\} ,$$

ne sont pas des temps d'arrêt.

Enonçons quelques propriétés élémentaires vérifiées par tout temps d'arrêt.

Lemme I.22

Soit T un temps d'arrêt. Pour tout $t \geqslant 0$, les événements $\{T < t\}$ et $\{T = t\}$ sont dans $\mathcal{F}_t$. Par ailleurs, pour tout $n \geqslant 1$, $T_n := 2^{-n}(\lfloor 2^n T\rfloor + 1)$ est un temps d'arrêt.

Démonstration On note que

$$\{T < t\} = \bigcup_{q\in[0,t[\cap\mathbb{Q}} \{T \leqslant q\} ,$$

de sorte que cet événement est bien dans $\mathcal{F}_t$. En écrivant $\{T = t\} = \{T \leqslant t\}\backslash\{T < t\}$ la $\mathcal{F}_t$ mesurabilité de $\{T = t\}$ s'en suit.
On observe que T_n prend ses valeurs dans l'ensemble $k2^{-n}$, $k \geqslant 0$. Pour tout $t \geqslant 0$, on introduit $k \geqslant 0$ tel que $k2^{-n} \leqslant t < (k + 1)2^{-n}$ et l'on écrit

$$\{T_n \leqslant t\} = \{T_n \leqslant k2^{-n}\} = \{\lfloor 2^n T\rfloor + 1 \leqslant k\} = \{2^n T < k\} = \{T < 2^{-n}k\} .$$

Comme $2^{-n}k \leqslant t$, ce dernier événement est $\mathcal{F}_t$-mesurable. □

Etant donné un temps d'arrêt T, on introduit la tribu du passé avant T en posant

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

(Nous reviendrons sur cette définition dans le chapitre suivant). On pourra vérifier en exercice que $\mathcal{F}_T$ est bien une tribu. On pourra aussi vérifier en exercice que les v.a. T et $\mathbf{1}_{\{T<\infty\}}B_T$ sont $\mathcal{F}_T$-mesurables.

Théorème I.23 (*Propriété de Markov forte*)

Soit T un temps d'arrêt fini presque sûrement. On pose pour tout $t \geqslant 0$

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

Le processus $B^{(T)}$ est un mouvement Brownien indépendant de $\mathcal{F}_T$.

La propriété de Markov forte a de très nombreuses applications dans l'étude du mouvement Brownien, nous en verrons certaines un peu plus loin. Notons qu'elle est fondamentalement due à l'indépendance et à la stationarité des incréments du mouvement Brownien.

Démonstration Nous allons montrer que pour tout $A \in \mathcal{F}_T$ et tout $m \geqslant 1$, tous $0 \leqslant t_1 < t_2 < \ldots < t_m$, et toute fonction continue bornée $f : \mathbb{R}^m \to \mathbb{R}$

$$\mathbb{E}[\mathbf{1}_A f(B^{(T)}_{t_1}, \ldots, B^{(T)}_{t_m})] = \mathbb{P}(A)\mathbb{E}[f(B_{t_1}, \ldots, B_{t_m})] \; . \tag{I.1}$$

Cela suffit pour établir les assertions du théorème. En effet, en prenant $A = \Omega$, on déduit que B^T a même loi qu'un mouvement Brownien : comme ses trajectoires sont également continues, on en déduit alors que c'est un mouvement Brownien. Par ailleurs cette identité assure que le vecteur $(B^{(T)}_{t_1}, \ldots, B^{(T)}_{t_m})$ est indépendant de la tribu $\mathcal{F}_T$, et ce pour tout choix de m et de valeurs $t_1, \ldots, t_m$. L'indépendance entre la tribu engendrée par $B^{(T)}$ et $\mathcal{F}_T$ s'en suit.
Pour établir (I.1), commençons par observer que le cas où T est déterministe est déjà couvert par la propriété de Markov simple. Supposons à présent que T prend ses valeurs dans un ensemble dénombrable $\{r_k, k \geqslant 1\}$ donné. On calcule alors

$$\begin{aligned}
\mathbb{E}[\mathbf{1}_A f(B^{(T)}_{t_1}, \ldots, B^{(T)}_{t_m})] &= \sum_{k\geqslant 1} \mathbb{E}[\mathbf{1}_A f(B^{(T)}_{t_1}, \ldots, B^{(T)}_{t_m})\mathbf{1}_{\{T=r_k\}}] \\
&= \sum_{k\geqslant 1} \mathbb{E}[f(B_{r_k+t_1} - B_{r_k}, \ldots, B_{r_k+t_m} - B_{r_k})\mathbf{1}_{A\cap\{T=r_k\}}]
\end{aligned}$$

Or $A \cap \{T = r_k\} \in \mathcal{F}_{r_k}$. Ainsi la propriété de Markov simple permet de poursuivre le calcul et d'obtenir

$$\begin{aligned}
&= \sum_{k\geqslant 1} \mathbb{E}[f(B_{t_1}, \ldots, B_{t_m})]\mathbb{P}(A \cap \{T = r_k\}) \\
&= \mathbb{E}[f(B_{t_1}, \ldots, B_{t_m})]\mathbb{P}(A) \; .
\end{aligned}$$

Dans le cas général où T prend ses valeurs dans $[0, \infty]$ (et est fini p.s.), on introduit la suite d'approximations $T_n := 2^{-n}(\lfloor 2^n T\rfloor + 1)$, $n \geqslant 1$. On note que $T_n \to T$ presque sûrement quand $n \to \infty$. Ainsi la continuité des trajectoires combinée au Théorème de Convergence Dominée assure que

$$\mathbb{E}[\mathbf{1}_A f(B^{(T)}_{t_1}, \ldots, B^{(T)}_{t_m})] = \lim_{n\to\infty} \mathbb{E}[\mathbf{1}_A f(B^{(T_n)}_{t_1}, \ldots, B^{(T_n)}_{t_m})] \; .$$

Le cas dénombrable déjà établi s'applique à T_n et $B^{(T_n)}$, et l'on obtient (en notant que si $A \in \mathcal{F}_T$ alors $A \in \mathcal{F}_{T_n}$)

$$\mathbb{E}[\mathbf{1}_A f(B^{(T_n)}_{t_1}, \ldots, B^{(T_n)}_{t_m})] = \mathbb{P}(A)\mathbb{E}[f(B_{t_1}, \ldots, B_{t_m})] \; ,$$

de sorte que

$$\mathbb{E}[\mathbf{1}_A f(B^{(T)}_{t_1}, \ldots, B^{(T)}_{t_m})] = \mathbb{P}(A)\mathbb{E}[f(B_{t_1}, \ldots, B_{t_m})] \; .$$

□

Nous présentons une très jolie identité en loi portant sur le supremum courant du mouvement Brownien

$$S_t := \sup_{0\leqslant s\leqslant t} B_s \; , \quad t \geqslant 0 \; .$$

La preuve de cette identité repose sur la propriété de Markov forte du mouvement Brownien.

Théorème I.24 (***Principe de réflexion***)

Pour tout $t > 0$, tout $a \geqslant 0$ et tout $b \leqslant a$ on a

$$\mathbb{P}(S_t \geqslant a, B_t \leqslant b) = \mathbb{P}(B_t \geqslant 2a - b) \ .$$

En particulier, S_t a même loi que $|B_t|$.

La preuve de ce théorème sera vue en TD.

Démonstration On a déjà vu que $T_a := \inf\{t \geqslant 0 : B_t = a\} < \infty$ p.s. On écrit alors

$$\mathbb{P}(S_t \geqslant a, B_t \leqslant b) = \mathbb{P}(T_a \leqslant t, B_t \leqslant b) = \mathbb{P}(T_a \leqslant t, B^{(T_a)}_{t-T_a} \leqslant b - a) \ .$$

Le Théorème I.23 assure que $B^{(T_a)}$ est indépendant de T_a. Comme $B^{(T_a)}$ a même loi que $-B^{(T_a)}$ on en déduit que $(T_a, B^{(T_a)})$ a même loi que $(T_a, -B^{(T_a)})$ et ainsi

$$\mathbb{P}(T_a \leqslant t, B^{(T_a)}_{t-T_a} \leqslant b-a) = \mathbb{P}(T_a \leqslant t, -B^{(T_a)}_{t-T_a} \leqslant b-a) = \mathbb{P}(T_a \leqslant t, B_{T_a} - B_t \leqslant b-a) = \mathbb{P}(T_a \leqslant t, B_t \geqslant 2a-b) \ .$$

Comme $2a - b \geqslant a$, si $B_t \geqslant 2a - b$ alors nécessairement $T_a \leqslant t$ et ainsi

$$\mathbb{P}(S_t \geqslant a, B_t \leqslant b) = \mathbb{P}(B_t \geqslant 2a - b) \ .$$

Pour prouver la deuxième assertion, on écrit

$$\mathbb{P}(S_t \geqslant a) = \mathbb{P}(S_t \geqslant a, B_t \leqslant a) + \mathbb{P}(S_t \geqslant a, B_t > a) \ .$$

On note que si $B_t > a$ alors nécessairement $S_t \geqslant a$. En utilisant le principe de réflexion ainsi que cette dernière observation on obtient

$$\mathbb{P}(S_t \geqslant a) = \mathbb{P}(B_t \geqslant a) + \mathbb{P}(B_t > a) = 2\mathbb{P}(B_t \geqslant a) = \mathbb{P}(|B_t| \geqslant a) \ .$$

$\square$

6 Mesure de Wiener, mouvement Brownien multi-dimensionnel

Soit $(B_t, t \geqslant 0)$ un mouvement Brownien. L'application

$$\begin{aligned} B : \Omega &\to \mathbb{R}^{\mathbb{R}_+} \\ \omega &\mapsto (B_t(\omega), t \in \mathbb{R}_+) \end{aligned}$$

est mesurable par rapport aux tribus $\mathcal{F}$ et $\mathcal{B}(\mathbb{R})^{\otimes \mathbb{R}_+}$. Par définition du mouvement Brownien, nous savons que les trajectoires de B sont continues et ainsi l'application prend ses valeurs dans $C(\mathbb{R}_+, \mathbb{R})$. Ainsi B est une application mesurable de $(\Omega, \mathcal{F})$ dans $C(\mathbb{R}_+, \mathbb{R})$ muni de la tribu produit, c'est-à-dire, de la plus petite tribu rendant mesurables les applications $x \mapsto x_t$ pour tous $t \in \mathbb{R}_+$.

Il est habituel de munir $C(\mathbb{R}_+, \mathbb{R})$ de la métrique suivante :

$$d(x, y) := \sum_{n \geqslant 1} 2^{-n} \frac{\sup_{t \in [0,n]} |x_t - y_t|}{1 + \sup_{t \in [0,n]} |x_t - y_t|} \ .$$

On pourra se convaincre que cette distance métrise la topologie de la convergence uniforme sur tout compact de $\mathbb{R}_+$. On note alors $\mathcal{C}$ la tribu Borélienne associée à cet espace métrique. Il est alors possible de montrer le résultat suivant.

Lemme I.25

La tribu produit sur $C(\mathbb{R}_+, \mathbb{R})$ coïncide avec la tribu $\mathcal{C}$.

La loi W du mouvement Brownien sur $(C(\mathbb{R}_+, \mathbb{R}), \mathcal{C})$ est appelée *mesure de Wiener*.

Terminons ce chapitre en introduisant la version multi-dimensionnelle du mouvement Brownien. Un processus stochastique $(B_t, t \in \mathbb{R}_+) = ((B_t^{(1)}, \ldots, B_t^{(d)}), t \in \mathbb{R}_+)$ à valeurs dans $\mathbb{R}^d$ est un mouvement Brownien en dimension d si les processus $(B_t^{(i)}, t \in \mathbb{R}_+)$, $i \in \{1, \ldots, d\}$ sont des mouvements Browniens indépendants.

Chapitre II

Martingales

Approx. 1,5 séance de cours

1 Martingales à temps continu

Soit $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P})$ un espace de probabilité filtré, c'est-à-dire, un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ muni d'une collection $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ de sous-tribus de $\mathcal{F}$ satisfaisant $s \leqslant t \Rightarrow \mathcal{F}_s \subset \mathcal{F}_t$. Une telle collection est appelée *filtration.*

Définition II.1

Soit $M = (M_t, t \in \mathbb{R}_+)$ un processus stochastique à valeurs dans $\mathbb{R}$. On dit que M est une martingale (resp. sur-martingale, resp. sous-martingale) si :

1. *M est adapté à la filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$, c'est-à-dire, pour tout $t \in \mathbb{R}_+$, M_t est $\mathcal{F}_t$-mesurable,*
2. *pour tout $t \in \mathbb{R}_+$, M_t est intégrable,*
3. *pour tous $0 \leqslant s \leqslant t$, $\mathbb{E}[M_t \mid \mathcal{F}_s] = M_s$ p.s. (resp. $\mathbb{E}[M_t \mid \mathcal{F}_s] \leqslant M_s$ p.s., resp. $\mathbb{E}[M_t \mid \mathcal{F}_s] \geqslant M_s$ p.s.).*

Commençons par énoncer quelques propriétés simples.

Lemme II.2

Soient M une martingale (resp. une sous-martingale) et $f : \mathbb{R} \to \mathbb{R}$ une fonction convexe (resp. convexe croissante). On suppose que pour tout $t \in \mathbb{R}_+$, $f(M_t)$ est intégrable. Alors $f(M)$ est une sous-martingale.

Démonstration L'inégalité de Jensen pour l'espérance conditionnelle assure que

$$f(\mathbb{E}[M_t \mid \mathcal{F}_s]) \leqslant \mathbb{E}[f(M_t) \mid \mathcal{F}_s] .$$

Si M est une martingale alors $f(\mathbb{E}[M_t \mid \mathcal{F}_s]) = f(M_s)$ et l'on peut conclure. Si M est une sous-martingale, la croissance de f assure que $f(\mathbb{E}[M_t \mid \mathcal{F}_s]) \geqslant f(M_s)$ et l'on peut conclure. □

On rappelle qu'une variable aléatoire $T : \Omega \to \mathbb{N} \cup \{+\infty\}$ est un temps d'arrêt si pour tout $t \in \mathbb{R}_+$, $\{T \leqslant t\} \in \mathcal{F}_t$. On notera que si T est un temps d'arrêt, alors $\{T > t\} = \{T \leqslant t\}^\complement \in \mathcal{F}_t$ et $\{T = t\} = \lim_n \downarrow \{T \leqslant t\} \cap \{T \leqslant t - 1/n\}^\complement \in \mathcal{F}_t$.

Définition II.3

Soit T un temps d'arrêt. On appelle tribu du passé avant T la collection $\mathcal{F}_T$ de tous les événements $A \in \mathcal{F}$ tels que pour tout $t \in \mathbb{R}_+$, $A \cap \{T \leqslant t\} \in \mathcal{F}_t$.

On peut vérifier que $\mathcal{F}_T$ est bien une tribu (exercice !).

Remarque II.4

Est-il clair que la tribu $\mathcal{F}_T$ ainsi définie correspond au passé avant T *? Dans le cas où T est déterministe, et vaut disons a, alors $\mathcal{F}_T = \mathcal{F}_a$. Plus généralement, lorsque T prend ses valeurs dans un ensemble dénombrable, disons $\{t_n, n \in \mathbb{N}\}$, alors $\mathcal{F}_T$ n'est rien d'autre que l'ensemble des événements A qui s'écrivent*

$$A = \bigcup_n \left(A_{t_n} \cap \{T = t_n\} \right) ,$$

où chaque $A_{t_n} \in \mathcal{F}_{t_n}$, et l'on peut se convaincre qu'il s'agit là du passé avant T.
Lorsque T est un temps d'arrêt quelconque, il n'existe pas de description aussi simple de $\mathcal{F}_T$. Cependant, si $T < \infty$, on peut prouver que $A \in \mathcal{F}_T$ si et seulement si il existe un processus adapté X à trajectoires continues à droite ainsi qu'un Borélien B tels que $A = \{X_T \in B\}$, voir [cn11, Chap 5 Th 1.14].

Lemme II.5

Soient S, T deux temps d'arrêt. Alors $S \wedge T$ est encore un temps d'arrêt et $\mathcal{F}_{S\wedge T} = \mathcal{F}_S \cap \mathcal{F}_T$.

Démonstration Exercice ! □

A ce stade, notons une difficulté technique propre aux processus à temps continu. Soit $(X_t, t \in \mathbb{R}_+)$ un processus stochastique adapté à la filtration $(\mathcal{F}_t, t \in \mathbb{R}_+)$. Soit T un temps d'arrêt. Il est naturel de vouloir manipuler l'application

$$X_T : \omega \mapsto X_{T(\omega)}(\omega)$$

Si T prend ses valeurs dans un ensemble dénombrable $(t_n)_n$ alors X_T dépend de façon explicite des v.a. $(X_{t_n})_n$ et l'on peut montrer qu'il est mesurable. Si T est quelconque, la mesurabilité n'est pas forcément vraie ! En revanche, on a le résultat suivant :

Lemme II.6

Soit X un processus stochastique à trajectoires continues et soit T un temps d'arrêt fini p.s. Alors X_T est une v.a. $\mathcal{F}_T$ mesurable.
Si X_t admet une limite p.s. quand $t \to \infty$, notée X_∞, alors X_T est une v.a. $\mathcal{F}_T$ mesurable sans condition de finitude sur T.

Remarque II.7

On notera que sur l'événement $\{T = \infty\}$, la définition de X_T n'a de sens que si X_t admet une limite quand $t \to \infty$. Cependant, si $T < \infty$ p.s., l'événement $\{T = \infty\}$ est de mesure nulle et l'on peut y définir X_T de façon arbitraire.

Démonstration On suppose que $T < \infty$ p.s. On a par continuité des trajectoires que p.s.

$$X_T = \lim_{n\to\infty} \sum_{k=0}^{\infty} \mathbf{1}_{k2^{-n} \leqslant T < (k+1)2^{-n}} X_{k2^{-n}} .$$

On note que $\mathbf{1}_{k2^{-n}\leqslant T<(k+1)2^{-n}}X_{k2^{-n}} = \mathbf{1}_{k2^{-n}\leqslant T}X_{k2^{-n}}\mathbf{1}_{T<(k+1)2^{-n}}$. La v.a. $\mathbf{1}_{T<(k+1)2^{-n}}$ est bien $\mathcal{F}_T$ mesurable. Montrons alors que chaque $\mathbf{1}_{k2^{-n}\leqslant T}X_{k2^{-n}}$ est $\mathcal{F}_T$ mesurable. Cela revient à montrer que pour tout $s \in \mathbb{R}_+$, $\mathbf{1}_{s\leqslant T}X_s$ est $\mathcal{F}_T$ mesurable. Soit B un Borélien de $\mathbb{R}$ ne contenant pas 0. Pour tout $t \in \mathbb{R}_+$ on a

$$\{\mathbf{1}_{s\leqslant T}X_s \in B\} \cap \{T \leqslant t\} = \begin{cases} \varnothing & \text{si } t < s \\ \{X_s \in B\} \cap \{s \leqslant T \leqslant t\} & \text{si } s \leqslant t \end{cases}$$

qui est $\mathcal{F}_t$ mesurable dans les deux cas (on notera que $\{s \leqslant T \leqslant t\} = \{T < s\}^{\complement} \cap \{T \leqslant t\}$). Si B est un Borélien de $\mathbb{R}$ contenant 0 alors

$$\{\mathbf{1}_{s\leqslant T}X_s \in B\} \cap \{T \leqslant t\} = \{T \leqslant t\}\backslash\Big(\{\mathbf{1}_{s\leqslant T}X_s \in B^{\complement}\} \cap \{T \leqslant t\}\Big) .$$

□

Dans la suite, on manipulera toujours des martingales à *trajectoires continues* de sorte que les quantités M_T, $\int_0^t M_s ds$ ou $\sup_{0\leqslant s\leqslant t} M_s$ seront bien mesurables.

Les martingales sont des objets qui apparaissent dans de nombreux contextes, et il est utile d'établir des résultats généraux les concernant. Il se trouve que les résultats que nous allons établir sont tout à fait analogues à ceux connus à temps discret : ce n'est pas étonnant car si M est une martingale à temps continu alors pour toute suite croissante $(t_n)_n$, le processus $(M_{t_n}, n \in \mathbb{N})$ est une martingale à temps discret adapté à la filtration $(\mathcal{F}_{t_n}, n \in \mathbb{N})$. Cette "projection" va nous permettre d'établir certains résultats (comme les inégalités maximales) en utilisant leurs analogues discrets et en passant à la limite. Cependant, certains énoncés ne pourront pas être obtenus de façon élémentaire à partir du cas discret, et nécessiteront des détours techniques conséquents : on notera en particulier que le fait qu'une martingale arrêtée est encore une martingale est délicat à établir dans le continu (ce sera notre dernier résultat !) tandis qu'il s'agissait d'une propriété élémentaire dans le discret.

1) Inégalités maximales

Proposition II.8

— *Si M est une sous-martingale à trajectoires continues alors pour tout $a > 0$ et pour tout $t \in \mathbb{R}_+$*

$$a\mathbb{P}(\sup_{0\leqslant s\leqslant t} M_s > a) \leqslant \mathbb{E}[M_t^+] ,$$

— *Si M est une sur-martingale à trajectoires continues alors pour tout $a > 0$ et pour tout $t \in \mathbb{R}_+$*

$$a\mathbb{P}(\sup_{0\leqslant s\leqslant t} M_s > a) \leqslant \mathbb{E}[M_0] + \mathbb{E}[M_t^-] .$$

Démonstration On suppose que M est une sous-martingale. Soit $n \geqslant 1$ un entier et soient $0 = t_0 < t_1^{(n)} < \ldots < t_n^{(n)} = t$. Le processus $(M_{t_k^{(n)}}, 0 \leqslant k \leqslant n)$ est une sous-martingale à temps discret dans la filtration $\mathcal{G}_k := \mathcal{F}_{t_k^{(n)}}$. Ainsi l'inégalité maximale à temps discret (et le fait que $t_n^{(n)} = t$) assure que

$$a\mathbb{P}(\sup_{0\leqslant k\leqslant n} M_{t_k^{(n)}} > a) \leqslant \mathbb{E}[M_t^+] .$$

On peut choisir les $(t_k^{(n)}, 0 \leqslant k \leqslant n)$ de sorte que la suite d'ensembles $\{t_k^{(n)}; 0 \leqslant k \leqslant n\}$, $n \geqslant 1$ soit croissante et converge vers $[0,t] \cap \mathbb{Q}$. Par la continuité croissante des probabilités, on obtient

$$a\mathbb{P}(\sup_{s\in[0,t]\cap\mathbb{Q}} M_s > a) \leqslant \mathbb{E}[M_t^+] \,.$$

Par continuité des trajectoires de M on a

$$\mathbb{P}(\sup_{s\in[0,t]\cap\mathbb{Q}} M_s > a) = \mathbb{P}(\sup_{s\in[0,t]} M_s > a) \,,$$

et la première inégalité est prouvée. On procède de façon analogue pour la seconde. □

Proposition II.9 (*Inégalité maximale de Doob*)

Soit M une martingale à trajectoires continues. Posons $M_t^* := \sup_{0\leqslant s\leqslant t} |M_s|$ pour tout $t \in \mathbb{R}_+$. Pour tout $p > 1$ et tout $t \in \mathbb{R}_+$

$$\mathbb{E}[(M_t^*)^p] \leqslant \Big(\frac{p}{p-1}\Big)^p \mathbb{E}[|M_t|^p] \,.$$

Démonstration On suit la même démarche que dans la preuve précédente : si l'on restreint le supremum à un ensemble fini de marginales, l'inégalité est vraie (cas discret). Par croissance et continuité, on peut en déduire l'inégalité recherchée. □

L'inégalité maximale de Doob a une conséquence tout à fait remarquable.

Corollaire II.10

Soit M une martingale à trajectoires continues et $p > 1$. Si M est bornée dans L^p, c'est-à-dire, $\sup_t \mathbb{E}[|M_t|^p] < \infty$ alors la famille de v.a. $|M_t|^p, t \in \mathbb{R}_+$ est uniformément intégrable.

Démonstration On pose $C := \sup_t \mathbb{E}[|M_t|^p]$. L'inégalité maximale de Doob assure que pour tout $t \geqslant 0$

$$\mathbb{E}[(M_t^*)^p] \leqslant C\Big(\frac{p}{p-1}\Big)^p \,.$$

Or M_t^* est une fonction croissante de t et ainsi admet une limite p.s. que l'on note $M_\infty^* = \sup_{s\geqslant 0} |M_s|$. Le Théorème de Convergence Monotone assure alors que

$$\mathbb{E}[(M_\infty^*)^p] \leqslant C\Big(\frac{p}{p-1}\Big)^p \,.$$

On peut appliquer la Proposition B.4. □

2) Convergence

Nous allons admettre le résultat suivant, qui est tout à fait analogue à celui vu dans le discret (Théorème C.7).

Théorème II.11

Soit M une martingale (resp. sous-martingale, resp. sur-martingale) à trajectoires continues, et bornée dans L^1, c'est-à-dire, telle que $\sup_{t\in\mathbb{R}_+} \mathbb{E}[|M_t|] < \infty$. Alors quand $t \to \infty$, M_t converge *presque sûrement vers une variable aléatoire intégrable M_∞.*

3) Martingales fermées

Définition II.12

On dit qu'une martingale M est fermée s'il existe une v.a. $X \in L^1$ telle que pour tout $t \in \mathbb{R}_+$

$$M_t = \mathbb{E}[X \mid \mathcal{F}_t] \, .$$

Théorème II.13

Soit M une martingale à trajectoires continues. Il y a équivalence entre :

(i) M est fermée

(ii) M est uniformément intégrable.

(iii) M converge p.s. et dans L^1 vers une variable aléatoire M_∞ intégrable.

Si l'une des trois conditions est vérifiée, alors nécessairement M est fermée par sa limite M_∞.

Démonstration (i) $\Rightarrow$ (ii). Il suffit d'appliquer la Proposition B.7.
(ii) $\Rightarrow$ (iii). Etant u.i., la collection de v.a. $(M_t, t \in \mathbb{R}_+)$ est bornée dans L^1. Comme c'est en outre une martingale, le Théorème II.11 assure qu'elle converge p.s. vers une variable aléatoire intégrable M_∞. Le théorème de convergence dominée optimal Théorème B.6 montre alors que M_t converge aussi dans L^1 vers M_∞.
(iii) $\Rightarrow$ (i). Pour tout $A \in \mathcal{F}_t$, la propriété de martingale assure que

$$\mathbb{E}[M_t \mathbf{1}_A] = \mathbb{E}[M_{t+s} \mathbf{1}_A] \, .$$

Or $M_{t+s} \to M_\infty$ dans L^1 quand $s \to \infty$. Ainsi

$$\mathbb{E}[M_{t+s} \mathbf{1}_A] \to \mathbb{E}[M_\infty \mathbf{1}_A] \, ,$$

et l'on en déduit que

$$\mathbb{E}[M_t \mathbf{1}_A] = \mathbb{E}[M_\infty \mathbf{1}_A] \, .$$

Ceci prouve que p.s.

$$M_t = \mathbb{E}[M_\infty \mid \mathcal{F}_t] \, .$$

□

Théorème II.14 (*Théorème d'arrêt*)

Soit M une martingale à trajectoires continues et fermée. Pour tout temps d'arrêt T on a presque sûrement

$$M_T = \mathbb{E}[M_\infty \mid \mathcal{F}_T] \, .$$

Par ailleurs pour tous temps d'arrêt $S \leqslant T$ on a presque sûrement

$$M_S = \mathbb{E}[M_T \mid \mathcal{F}_S] \, .$$

Démonstration La deuxième propriété est une conséquence de la première. En effet

$$\mathbb{E}[M_T \mid \mathcal{F}_S] = \mathbb{E}[\mathbb{E}[M_\infty \mid \mathcal{F}_T] \mid \mathcal{F}_S] = \mathbb{E}[M_\infty \mid \mathcal{F}_S] = M_S \ .$$

La première propriété se prouve comme suit. On introduit pour tout $n \geqslant 1$

$$T^{(n)} := \sum_{k=0}^{\infty} \frac{k+1}{2^n} \mathbf{1}_{\{k2^{-n} < T \leqslant (k+1)2^{-n}\}} + \infty \mathbf{1}_{\{T=\infty\}} \ .$$

Il s'agit d'une suite de temps d'arrêt qui converge de façon décroissante vers T.
Pour tout $n \geqslant 1$, la martingale discrète $(M_{k2^{-n}}, k \in \mathbb{N})$ est adaptée à la filtration $(\mathcal{F}_{k2^{-n}}, k \in \mathbb{N})$, et est fermée par M_∞. Par ailleurs $T^{(n)}$ est un temps d'arrêt dans cette filtration. Le théorème d'arrêt dans le cas discret assure que presque sûrement

$$M_{T^{(n)}} = \mathbb{E}[M_\infty \mid \mathcal{F}_{T^{(n)}}] \ .$$

Une conséquence de cette identité est que la collection de v.a. $(M_{T^{(n)}}, n \geqslant 1)$ est uniformément intégrable.
Soit $A \in \mathcal{F}_T$. Comme $T \leqslant T^{(n)}$, on a $A \in \mathcal{F}_{T^{(n)}}$ et ainsi

$$\mathbb{E}[\mathbf{1}_A M_{T^{(n)}}] = \mathbb{E}[\mathbf{1}_A M_\infty] \ .$$

La suite de v.a. $\mathbf{1}_A M_{T^{(n)}}$, $n \geqslant 1$ converge p.s. vers $\mathbf{1}_A M_T$ et est u.i. : par le Théorème B.6 on en déduit que la convergence a aussi lieu dans L^1 et ainsi

$$\mathbb{E}[\mathbf{1}_A M_T] = \mathbb{E}[\mathbf{1}_A M_\infty] \ .$$

Comme le Lemme 1 assure que M_T est $\mathcal{F}_T$ mesurable, ceci assure que $M_T = \mathbb{E}[M_\infty \mid \mathcal{F}_T]$ et conclut la preuve. □

Ce résultat (difficile) a plusieurs conséquences (a priori simples).

Corollaire II.15 (*Théorème d'arrêt borné*)

Soit M une martingale. Soient $S \leqslant T$ deux temps d'arrêt bornés, *c'est-à-dire, tels qu'il existe $t_0 \geqslant 0$ (déterministe !) tel que $0 \leqslant S \leqslant T \leqslant t_0$ p.s. Alors M_S, M_T sont intégrables et presque sûrement*

$$\mathbb{E}[M_T \mid \mathcal{F}_S] = M_S \ .$$

Démonstration Le processus stochastique $(M_{t\wedge t_0}, t \geqslant 0)$ est à trajectoires continues et est adapté à la filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$. C'est une martingale fermée par M_{t_0} car la propriété de martingale de M assure que

$$\mathbb{E}[M_{t_0} \mid \mathcal{F}_t] = M_{t\wedge t_0} \ , \quad \forall t \in \mathbb{R}_+ \ .$$

On peut donc appliquer le théorème précédent aux temps d'arrêt $S \leqslant T$ et obtenir

$$\mathbb{E}[M_{T\wedge t_0} \mid \mathcal{F}_S] = M_{S\wedge t_0} \ .$$

Comme $T \leqslant t_0$ et $S \leqslant t_0$, on peut conclure. □

Corollaire II.16

Soient M une martingale à trajectoires continues et fermée, et T un temps d'arrêt. On pose $M_t^T := M_{T\wedge t}$ pour tout $t \in \mathbb{R}_+$. Alors M^T est une martingale (dans la filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$) fermée par M_T.

Démonstration Par le théorème d'arrêt appliqué à $t \wedge T$ et T, on sait que $M_{t\wedge T}, M_T \in L^1$ et que

$$\mathbb{E}[M_T \mid \mathcal{F}_{t\wedge T}] = M_{t\wedge T} \; .$$

Si l'on montre que cela reste vrai en remplaçant $\mathcal{F}_{t\wedge T}$ par $\mathcal{F}_t$ alors on en déduira que $(M_{t\wedge T}, t \geqslant 0)$ est fermée par M_T et que c'est une martingale dans la filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$, ce qui conclura la preuve. On sait par le Lemme 1 que $M_T\mathbf{1}_{T<\infty}$ est $\mathcal{F}_T$ mesurable. Vérifions que $M_T\mathbf{1}_{T\leqslant t}$ est $\mathcal{F}_{t\wedge T}$ mesurable. La $\mathcal{F}_T$ mesurabilité est immédiate. Concernant la $\mathcal{F}_t$ mesurabilité, on voit que pour tout Borélien B ne contenant pas 0, on a

$$\{M_T\mathbf{1}_{T\leqslant t} \in B\} = \{M_T \in B\} \cap \{T \leqslant t\} \in \mathcal{F}_t \; ,$$

tandis que si B contient 0 on a

$$\{M_T\mathbf{1}_{T\leqslant t} \in B\} = \{T > t\} \cup \Big(\{T \leqslant t\} \cap \{M_T\mathbf{1}_{T<\infty} \in B\}\Big) \in \mathcal{F}_t \; .$$

Comme $\mathcal{F}_{t\wedge T} = \mathcal{F}_t \cap \mathcal{F}_T$, on en déduit que $M_T\mathbf{1}_{T\leqslant t}$ est $\mathcal{F}_{t\wedge T}$ mesurable. Ainsi

$$\mathbb{E}[M_T\mathbf{1}_{T\leqslant t} \mid \mathcal{F}_{t\wedge T}] = M_T\mathbf{1}_{T\leqslant t} = \mathbb{E}[M_T\mathbf{1}_{T\leqslant t} \mid \mathcal{F}_t] \; .$$

Pour compléter la preuve, nous allons montrer que

$$\mathbb{E}[M_T\mathbf{1}_{T>t} \mid \mathcal{F}_{t\wedge T}] = \mathbb{E}[M_T\mathbf{1}_{T>t} \mid \mathcal{F}_t] \; .$$

Soit $A \in \mathcal{F}_t$. On sait que $\{T > t\} \in \mathcal{F}_t$ et ainsi $A \cap \{T > t\} \in \mathcal{F}_t$. Par ailleurs pour tout $s \geqslant 0$

$$A \cap \{T > t\} \cap \{T \leqslant s\} = \begin{cases} \varnothing \text{ si } s \leqslant t \\ A \cap \{t < T \leqslant s\} \text{ si } s > t \; . \end{cases}$$

Dans les deux cas, cet événement est $\mathcal{F}_s$ mesurable et ainsi $A \cap \{T > t\} \in \mathcal{F}_T$. On a donc montré que $A \cap \{T > t\} \in \mathcal{F}_{t\wedge T}$. On peut alors écrire

$$\mathbb{E}[M_T\mathbf{1}_{T>t}\mathbf{1}_A] = \mathbb{E}[\mathbf{1}_{T>t}\mathbf{1}_A\mathbb{E}[M_T \mid \mathcal{F}_{t\wedge T}]] = \mathbb{E}[\mathbf{1}_A\mathbb{E}[\mathbf{1}_{T>t}M_{t\wedge T} \mid \mathcal{F}_{t\wedge T}]] \; ,$$

ce qui achève la preuve. □

Finalement, on peut établir que toute martingale arrêtée est encore une martingale : cette propriété d'apparence simple repose sur les résultats élaborés précédents. On notera que dans le cas discret, cette propriété est très facile à établir.

Corollaire II.17

Soient $(M_t, t \in \mathbb{R}_+)$ une martingale à trajectoires continues (pas forcément fermée !), et T un temps d'arrêt. Alors M^T est encore une martingale.

Démonstration Soit $a > 0$. Le processus stochastique $(M_{t\wedge a}, t \geqslant 0)$ est une martingale fermée par M_a. On peut donc lui appliquer le corollaire précédent et l'on en déduit que $(M_{t\wedge a\wedge T}, t \geqslant 0)$ est une martingale. En particulier $(M_{t\wedge T}, t \in [0, a])$ est une martingale sur $[0, a]$. Comme a est arbitraire, on peut conclure. □

2 Martingales locales

Il est pratique de relâcher certaines hypothèses dans la définition d'une martingale (nous en verrons de nombreuses applications lorsque nous introduirons l'intégrale stochastique).

Définition II.18

On appelle martingale locale issue de 0 tout processus stochastique $M = (M_t, t \in \mathbb{R}_+)$ adapté, à trajectoires continues, qui vérifie $M_0 = 0$ p.s. et pour lequel il existe une suite croissante de temps d'arrêt $(T_n, n \in \mathbb{N})$ telle que $T_n \uparrow \infty$ et M^{T_n} est une martingale.
Plus généralement, on appelle martingale locale tout processus stochastique $M = (M_t, t \in \mathbb{R}_+)$ tel que M_0 est $\mathcal{F}_0$-mesurable et $N_t = M_t - M_0$, $t \in \mathbb{R}_+$ est une martingale locale issue de 0.

On dira qu'une suite $(T_n, n \in \mathbb{N})$ de temps d'arrêt réduit M si p.s. $T_n \uparrow \infty$ et $M^{T_n} - M_0$ est une martingale.

Proposition II.19

1. *Toute martingale à trajectoires continues est une martingale locale.*
2. *Soient M une martingale locale et T un temps d'arrêt. Alors M^T est encore une martingale locale.*
3. *Soit M une martingale locale. S'il existe une v.a. positive $Z \in L^1$ telle que pour tout $t \geqslant 0$, $|M_t| \leqslant Z$ alors M est une martingale (uniformément intégrable).*
4. *Si M est une martingale locale issue de 0, alors il existe une suite $(T_n, n \in \mathbb{N})$ qui réduit M et telle que M^{T_n} est une martingale uniformément intégrable.*

Remarque II.20

On peut se demander si l'hypothèse de domination s'écrit : p.s., pour tous $t \geqslant 0$, $|M_t| \leqslant Z$ ou alors : pour tout $t \geqslant 0$, p.s. $|M_t| \leqslant Z$. En fait, la continuité des trajectoires assure que ces deux assertions sont équivalentes.

Démonstration

1. Par le Corollaire II.17, n'importe quelle suite croissante de temps d'arrêt qui tend p.s. vers $+\infty$ réduit M.
2. Soit T_n une suite de temps d'arrêt qui réduit M. Le Corollaire II.17 appliqué à la martingale $M^{T_n} - M_0$ assure que $(M^{T_n} - M_0)^T$ est encore une martingale. Or pour tout $t \geqslant 0$, $(M^{T_n} - M_0)^T_t = M_{T_n \wedge T \wedge t} - M_0 = (M^T - M_0)^{T_n}_t$ ce qui assure que $(M^T - M_0)^{T_n}$ est une martingale. On en déduit que $M^T - M_0$ est une martingale locale.
3. Soit $(T_n)_{n \in \mathbb{N}}$ une suite qui réduit M. Alors

$$M_{s \wedge T_n} - M_0 = \mathbb{E}[M_{t \wedge T_n} - M_0 \mid \mathcal{F}_s] .$$

Par hypothèse de domination, on peut ajouter M_0 aux deux membres de l'inégalité et obtenir

$$M_{s \wedge T_n} = \mathbb{E}[M_{t \wedge T_n} \mid \mathcal{F}_s] .$$

Par continuité des trajectoires et domination, on peut passer à la limite sur n et obtenir

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

4. Il suffit de prendre $T_n := \inf\{t \geqslant 0 : |M_t| \geqslant n\}$ car alors M^{T_n} est bornée par n et l'on peut appliquer le point précédent.

$\square$

Remarque II.21

On pourrait penser que si une martingale locale est intégrable à tout temps, alors il s'agit d'une vraie martingale. Il s'avère que cela est faux, voir par exemple [RY99, Exercise V.2.13].

3 Variation finie et martingales

Commençons par rappeler la notion de variation d'une fonction. Soit $f : \mathbb{R}_+ \to \mathbb{R}$, on définit

$$V(f,s,t) = \sup_{(t_k)_k} \sum_k |f(t_{k+1}) - f(t_k)| ,$$

où le supremum porte sur toutes les subdivisions finies $s = t_0 < t_1 < \ldots < t_n = t$ de l'intervalle $[s,t]$.

On dit que f est à variation finie si $V(f,s,t) < \infty$ pour tous $0 \leqslant s \leqslant t$.

Il est facile de vérifier que $V(f,s,t) + V(f,t,u) = V(f,s,u)$ pour tous triplets $s < t < u$. Par ailleurs toute fonction continûment différentiable vérifie

$$V(f,s,t) = \int_s^t |f'(r)| dr < \infty ,$$

et toute fonction monotone vérifie

$$V(f,s,t) = |f(t) - f(s)| .$$

La fonction $f : t \mapsto \mathbf{1}_{[1,\infty[}$ est à variation finie et $V(f,0,\infty) = 1$. En revanche, la fonction $t \mapsto \mathbf{1}_{\mathbb{Q}}(t)$ est à variation infinie. Enfin la fonction

$$[0,1] \ni t \to t\sin(1/t) ,$$

est continue mais sa variation sur l'intervalle $[0,1]$ est infinie. Cela est laissé en exercice (indication : considérer les temps $t = \frac{1}{(2k+1)\pi/2}$ pour $k \in \mathbb{N}$).

Proposition II.22

Une fonction $f : \mathbb{R}_+ \to \mathbb{R}$ est à variation finie si et seulement si c'est la différence de deux fonctions croissantes.

Démonstration Si $f = g - h$ avec g, h croissantes alors

$$\sum_k |f(t_{k+1}) - f(t_k)| \leqslant \sum_k |g(t_{k+1}) - g(t_k)| + \sum_k |h(t_{k+1}) - h(t_k)| ,$$

de sorte que $V(f,s,t) \leqslant V(g,s,t) + V(h,s,t)$ et ainsi f est à variation finie. Réciproquement, si f est à variation finie alors les fonctions $g = V(f,0,t)$ et $h = V(f,0,t) - f(t)$ sont croissantes. En effet $V(f,0,t+s) - V(f,0,t) = V(f,t,t+s) \geqslant 0$, et $V(f,0,t+s) - V(f,0,t) - (f(t+s) - f(t)) = V(f,t,t+s) - (f(t+s) - f(t)) \geqslant 0$ car $\{t, t+s\}$ forme une subdivision de l'intervalle $[t, t+s]$. $\square$

Il se trouve que le Brownien n'est pas à variation finie.

Lemme II.23

Presque sûrement, pour tous $0 \leqslant s < t$, $V(B,s,t) = \infty$.

Démonstration Il suffit de montrer que pour tous $s < t \in \mathbb{Q}_+$, presque sûrement $V(B,s,t) = \infty$. Posons $t_k = s + k(t-s)/n$ pour $k \in \{0,\ldots,n\}$. Alors

$$\sum_{k=0}^{n-1} |B(t_{k+1}) - B(t_k)| \;,$$

est égal en loi à

$$\sqrt{\frac{t-s}{n}}(|X_1| + \ldots + |X_n|) \;,$$

où les v.a. X_i sont i.i.d. $\mathcal{N}(0,1)$. Si l'on note $c = \mathbb{E}[|X_1|] > 0$ alors

$$\sqrt{\frac{t-s}{n}}(|X_1| + \ldots + |X_n|) = \sqrt{n(t-s)}c + \sqrt{\frac{t-s}{n}}((|X_1| - c) + \ldots + (|X_n| - c)) \;.$$

Le théorème central limite assure que le second terme converge en loi vers une gaussienne, tandis que le premier terme converge vers $+\infty$. Ainsi la somme converge en probabilité vers $+\infty$ quand $n \to \infty$, ce qui implique que $V(B,s,t) = \infty$ p.s. □

En fait, il n'existe pas de martingale non-triviale à trajectoires continues qui soit à variation finie.

Théorème II.24

Soit M une martingale locale issue de 0. Si M est un processus à variation finie alors presque sûrement pour tout $t \in \mathbb{R}_+$, $M_t = 0$.

Remarque II.25

On peut s'interroger sur l'importance de la continuité des trajectoires dans un tel résultat. Il se trouve qu'elle est cruciale. En effet, si N est un processus de Poisson d'intensité $\lambda > 0$ alors $N_t - \lambda t$ est une martingale. Par ailleurs, c'est un processus à variation finie car c'est la différence de deux processus croissants : N_t et λt.

Démonstration Posons pour tout $n \in \mathbb{N}$

$$\tau_n := \inf\{t \geqslant 0 : V(M,0,t) \geqslant n\} \;.$$

Comme M est à variation finie, on a $\tau_n \uparrow \infty$. Fixons $n \in \mathbb{N}$ et posons $N := M^{\tau_n}$. Alors $V(N,0,\infty) = V(M,0,\tau_n) = n$ et pour tout $t \in \mathbb{R}_+$, $|N_t| \leqslant V(N,0,t) \leqslant n$. Ainsi N est une martingale locale bornée donc par la Proposition II.19 3. c'est une vraie martingale bornée. On note alors que pour tous $0 \leqslant s \leqslant t$

$$\mathbb{E}[N_s(N_t - N_s)] = \mathbb{E}[\mathbb{E}[N_s(N_t - N_s) \mid \mathcal{F}_s]] = \mathbb{E}[N_s\mathbb{E}[(N_t - N_s) \mid \mathcal{F}_s]] = 0 \;,$$

or $(N_t - N_s)^2 = N_t^2 - N_s^2 - 2N_s(N_t - N_s)$ et ainsi

$$\mathbb{E}[(N_t - N_s)^2] = \mathbb{E}[N_t^2 - N_s^2] \;.$$

On se donne alors une suite (indicée par $p \geqslant 1$) de subdivisions $0 = t_0^{(p)} < \ldots < t_{k_p}^{(p)} = t$ de pas tendant vers 0 et l'on calcule

$$\mathbb{E}[N_t^2] = \sum_{i=1}^{k_p} \mathbb{E}[N^2_{t_i^{(p)}} - N^2_{t_{i-1}^{(p)}}] = \sum_{i=1}^{k_p} \mathbb{E}[(N_{t_i^{(p)}} - N_{t_{i-1}^{(p)}})^2] \leqslant \mathbb{E}[(\sup_{1\leqslant i\leqslant k_p} |N_{t_i^{(p)}} - N_{t_{i-1}^{(p)}}|) \sum_{i=1}^{k_p} |N_{t_i^{(p)}} - N_{t_{i-1}^{(p)}}|]$$
$$\leqslant n\, \mathbb{E}[\sup_{1\leqslant i\leqslant k_p} |N_{t_i^{(p)}} - N_{t_{i-1}^{(p)}}|] \ .$$

Comme le pas tend de la subdivision tend vers 0, la continuité des trajectoires assure que $\sup_{1\leqslant i\leqslant k}(N_{t_i^{(p)}} - N_{t_{i-1}^{(p)}}) \to 0$ p.s. quand $p \to \infty$. Le caractère borné de N permet alors d'appliquer le théorème de convergence dominée pour déduire que

$$\lim_{p\to\infty} \mathbb{E}[\sup_{1\leqslant i\leqslant k_p} (N_{t_i^{(p)}} - N_{t_{i-1}^{(p)}})] = 0 \ .$$

Ainsi $\mathbb{E}[N_t^2] = 0$ et $\mathbb{E}[M^2_{t\wedge\tau_n}] = 0$. La continuité des trajectoires assure que $M_{t\wedge\tau_n} \to M_t$ p.s. quand $n \to \infty$. Par le lemme de Fatou, on en déduit que $\mathbb{E}[M_t^2] \leqslant \liminf_n \mathbb{E}[M^2_{t\wedge\tau_n}] = 0$. Ainsi pour tout $t \in \mathbb{R}_+$, $M_t = 0$ p.s. La continuité des trajectoires assure que l'on peut permuter "p.s." et "pour tout $t \in \mathbb{R}_+$". □

4 Variation quadratique

Dans la section précédente, nous avons vu que la variation d'une martingale non-triviale est toujours infinie. Nous allons voir qu'en revanche toute martingale admet une variation *quadratique* finie.

Théorème II.26

Soit M une martingale locale. Il existe un unique processus à trajectoires continues et croissantes, noté $(\langle M, M\rangle_t, t \geqslant 0)$ et appelé variation quadratique de M, tel que $M_t^2 - \langle M, M\rangle_t$ soit une martingale locale et $\langle M, M\rangle_0 = 0$. De plus, pour tout $t > 0$ et pour toute suite indicée par $n \geqslant 1$ de subdivisions emboîtées $0 = t_0^n < t_1^n < \ldots < t_{p_n}^n = t$ de l'intervalle $[0, t]$ dont le pas tend vers 0 quand $n \to \infty$, la convergence suivante a lieu en probabilité

$$\langle M, M\rangle_t = \lim_{n\to\infty} \sum_{i=1}^{p_n} (M_{t_i^n} - M_{t_{i-1}^n})^2 \ . \tag{II.1}$$

Calculons la variation quadratique du mouvement Brownien. On commence par noter que les v.a. $(B_{t_i^n} - B_{t_{i-1}^n})^2$ sont indépendantes de moyenne $t_i^n - t_{i-1}^n$ et de variance $2(t_i^n - t_{i-1}^n)^2$. Ainsi

$$\mathbb{E}\Big[\sum_{i=1}^{p_n} (B_{t_i^n} - B_{t_{i-1}^n})^2\Big] = \sum_{i=1}^{p_n} t_i^n - t_{i-1}^n = t \ ,$$

et

$$\mathbb{E}\Big[\Big(\sum_{i=1}^{p_n} (B_{t_i^n} - B_{t_{i-1}^n})^2 - t\Big)^2\Big] = \mathrm{Var}\Big(\sum_{i=1}^{p_n} (B_{t_i^n} - B_{t_{i-1}^n})^2\Big) = \sum_{i=1}^{p_n} \mathrm{Var}\Big((B_{t_i^n} - B_{t_{i-1}^n})^2\Big) = \sum_{i=1}^{p_n} 2(t_i^n - t_{i-1}^n)^2$$
$$\leqslant 2\Big(\sup_i (t_i^n - t_{i-1}^n)\Big) t \ ,$$

et cette quantité tend vers 0 quand le pas de la subdivision tend vers 0. Ainsi $\langle B, B\rangle_t = t$ et la convergence (II.1) a lieu dans L^2.

Démonstration [Démonstration du Théorème II.26]

Unicité. Si A et A' sont deux processus croissants tels que $M_t^2 - A_t$ et $M_t^2 - A'_t$ sont des martingales locales, alors $A_t - A'_t$ est une martingale locale de variation finie. Par le Théorème II.24, c'est un processus constant égal à $A_0 = A'_0 = 0$.

Existence. Il s'agit d'un résultat difficile : on renvoie vers [LG13, Th 4.2]. □

Remarque II.27

On peut vérifier que si M est une martingale locale et T un temps d'arrêt alors $\langle M^T, M^T\rangle_t = \langle M, M\rangle_{t\wedge T}$ pour tout $t \in \mathbb{R}_+$.

La variation quadratique d'une martingale locale apporte de nombreuses informations sur celle-ci. Un premier résultat dans ce sens est le suivant.

Proposition II.28

Soit M une martingale locale issue de 0. Il y a équivalence entre :

1. *M est une vraie martingale bornée dans L^2,*
2. *$\mathbb{E}[\langle M, M\rangle_\infty] < \infty$.*

Sous ces hypothèses, $M_t^2 - \langle M, M\rangle_t$ est une martingale uniformément intégrable.

Démonstration On suppose que M est une vraie martingale bornée dans L^2, et ainsi, est uniformément intégrable. On sait alors que M_t converge p.s. et dans L^1 vers une variable M_∞. Par ailleurs, une conséquence de l'inégalité de Doob, voir le Corollaire II.10, assure que

$$\mathbb{E}[\sup_{t\geqslant 0} M_t^2] < \infty \; .$$

Combiné à la Proposition B.4, ceci assure que $(M_t^2, t \geqslant 0)$ est également uniformément intégrable de sorte que M_t converge aussi dans L^2 vers M_∞.

On introduit alors le temps d'arrêt

$$S_n := \inf\{t \geqslant 0 : \langle M, M\rangle_t \geqslant n\} \; ,$$

et l'on a $\langle M, M\rangle_{t\wedge S_n} \leqslant n$. Ainsi la martingale locale $M^2_{t\wedge S_n} - \langle M, M\rangle_{t\wedge S_n}$ est dominée par la v.a. $\sup_{t\geqslant 0} M_t^2 + n$ qui est intégrable. La Proposition II.19 3. assure qu'il s'agit alors d'une vraie martingale (issue de 0) de sorte que

$$\mathbb{E}[M^2_{t\wedge S_n} - \langle M, M\rangle_{t\wedge S_n}] = 0 \; .$$

Du fait de l'intégrabilité de chacun des deux termes, cela se réécrit

$$\mathbb{E}[M^2_{t\wedge S_n}] = \mathbb{E}[\langle M, M\rangle_{t\wedge S_n}] \; .$$

En faisant tendre t vers l'infini, et en appliquant le théorème de convergence dominée pour le terme de gauche et le théorème de convergence monotone pour celui de droite, on obtient

$$\mathbb{E}[M^2_{S_n}] = \mathbb{E}[\langle M, M\rangle_{S_n}] \; .$$

En faisant tendre n vers l'infini, on obtient à l'aide des mêmes arguments

$$\mathbb{E}[M_\infty^2] = \mathbb{E}[\langle M, M\rangle_\infty] \ .$$

On a donc montré que $\mathbb{E}[\langle M, M\rangle_\infty] < \infty$. En outre, la martingale locale $M_t^2 - \langle M, M\rangle_t$ est dominée par la v.a. intégrable $\sup_{t\geqslant 0} M_t^2 + \langle M, M\rangle_\infty$ et il s'agit ainsi d'une vraie martingale uniformément intégrable.
On suppose que $\mathbb{E}[\langle M, M\rangle_\infty] < \infty$. On introduit le temps d'arrêt $T_n := \inf\{t \geqslant 0 : |M_t| \geqslant n\}$ de sorte que M^{T_n} est une vraie martingale bornée (Proposition II.19 3.). La martingale locale $M_{t\wedge T_n}^2 - \langle M, M\rangle_{t\wedge T_n}$ est dominée par la v.a. intégrable $n^2 + \langle M, M\rangle_\infty$, c'est donc une vraie martingale uniformément intégrable (issue de 0). On en déduit que

$$\mathbb{E}[M_{t\wedge T_n}^2] = \mathbb{E}[\langle M, M\rangle_{t\wedge T_n}] \leqslant \mathbb{E}[\langle M, M\rangle_\infty] < \infty \ .$$

L'inégalité de Doob assure donc que

$$\mathbb{E}[\sup_{t\geqslant 0} M_{t\wedge T_n}^2] \leqslant 4 \sup_{t\geqslant 0} \mathbb{E}[M_{t\wedge T_n}^2] \leqslant 4\mathbb{E}[\langle M, M\rangle_\infty] < \infty \ .$$

Par le Lemme de Fatou on obtient

$$\mathbb{E}[\sup_{t\geqslant 0} M_t^2] = \mathbb{E}[\lim_n \sup_{t\geqslant 0} M_{t\wedge T_n}^2] \leqslant \liminf_n \mathbb{E}[\sup_{t\geqslant 0} M_{t\wedge T_n}^2] \leqslant 4\mathbb{E}[\langle M, M\rangle_\infty] < \infty \ .$$

Ainsi la Proposition II.19 3. assure que M est une vraie martingale, qui est de plus bornée dans L^2. □

On peut aussi donner une version plus faible de cette équivalence. On dit que M est de carré intégrable si $\mathbb{E}[M_t^2] < \infty$ pour tout $t \geqslant 0$. On notera que cela n'implique pas que M est bornée dans L^2. En revanche, le Lemme II.2 assure que si M est une martingale de carré intégrable alors M^2 est une sous-martingale et l'on a $\sup_{s\in[0,t]} \mathbb{E}[M_s^2] \leqslant \mathbb{E}[M_t^2] < \infty$. Ainsi toute martingale de carré intégrable est localement bornée dans L^2, c'est-à-dire, $\sup_{s\in[0,t]} \mathbb{E}[M_s^2] < \infty$.

Corollaire II.29

Soit M une martingale locale issue de 0. Il y a équivalence entre :

1. *M est une vraie martingale de carré intégrable,*
2. *$\mathbb{E}[\langle M, M\rangle_t] < \infty$ pour tout $t \geqslant 0$.*

Sous ces hypothèses, $M_t^2 - \langle M, M\rangle_t$ est une martingale.

Démonstration Soit M une martingale locale issue de 0. Pour tout $a > 0$, M^a est encore une martingale locale. Le résultat précédent appliqué à M^a assure qu'il y a équivalence entre

1. M^a est une vraie martingale bornée dans L^2,
2. $\mathbb{E}[\langle M^a, M^a\rangle_\infty] < \infty$.

En utilisant la définition de M^a et le commentaire précédant l'énoncé du corollaire, on peut réécrire cette équivalence sous la forme :

1. M est une vraie martingale sur $[0, a]$ et $\mathbb{E}[M_t^2] < \infty$ pour tout $t \in [0, a]$,
2. $\mathbb{E}[\langle M, M\rangle_t] < \infty$ pour tout $t \in [0, a]$.

Comme a est arbitraire cela suffit à conclure. Par ailleurs, si l'une de ces conditions est remplie alors $M_{t\wedge a}^2 - \langle M, M\rangle_{t\wedge a}$ est une vraie martingale : là encore, comme a est arbitraire, on en déduit que $M_t^2 - \langle M, M\rangle_t$ est une martingale. □

Pour finir, on introduit le crochet de deux martingales locales M, N en posant

$$\langle M, N\rangle_t := \frac{1}{2}\Big(\langle M+N, M+N\rangle_t - \langle M, M\rangle_t - \langle N, N\rangle_t\Big)\ , \quad t \in \mathbb{R}_+\ .$$

Il est possible de montrer que ce processus est l'unique processus à variation finie issu de 0 tel que $M_t N_t - \langle M, N\rangle_t$ est une martingale locale. Il est aussi possible d'établir un résultat d'approximation similaire à ce qui est énoncé dans le Théorème II.26.

Chapitre III

Intégrale stochastique et formule d'Itô

Approx. 1,5 séance de cours
Il existe divers champs d'application (physique, finance, biologie) dans lesquels on est amené à modéliser l'évolution d'une quantité d'intérêt (la trajectoire d'une particule, le cours d'une action, etc.) par une équation différentielle dirigée par un mouvement Brownien :

$$dX_t = \sigma(X_t)dB_t \ ,$$

où $\sigma : \mathbb{R} \to \mathbb{R}$ est une fonction donnée. Pour donner un sens à la solution d'une telle équation, il est nécessaire de développer une théorie de l'intégration contre le mouvement Brownien.

Une approche naturelle serait d'approximer l'intégrale par des sommes de Riemann. Par exemple, pour donner un sens à

$$\int_0^t B_s dB_s \ ,$$

on peut se donner une suite de subdivisions emboîtées $0 = t_0^{(n)} < t_1^{(n)} < \ldots < t_n^{(n)} = t$ dont le pas tend vers 0 quand $n \to \infty$ et considérer la somme finie

$$\sum_{i=0}^{n-1} \tilde{B}_{i,n}(B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}}) \ ,$$

où $\tilde{B}_{i,n}$ est une valeur prise par B sur l'intervalle $[t_i^{(n)}, t_{i+1}^{(n)}]$. L'idée serait alors de définir l'intégrale comme la limite (si elle existe !) de la suite des sommes finies.
Dans la théorie de l'intégration de Riemann, la valeur que l'on choisit sur chaque intervalle de la subdivision (ici $\tilde{B}_{i,n}$) n'a pas d'impact sur la convergence, ni sur la limite, de la suite des sommes finies.
Il se trouve que la convergence de cette suite de sommes finies est *très* instable car elle dépend fortement du choix de $\tilde{B}_{i,n}$. Illustrons cela par un calcul simple. Considérons deux approximations distinctes : la première considère la valeur au début de l'intervalle $[t_i^{(n)}, t_{i+1}^{(n)}]$

$$I_t^{(n)} = \sum_{i=0}^{n-1} B_{t_i^{(n)}}(B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}}) \ ,$$

alors que la seconde choisit la valeur à la fin de l'intervalle $[t_i^{(n)}, t_{i+1}^{(n)}]$

$$S_t^{(n)} = \sum_{i=0}^{n-1} B_{t_{i+1}^{(n)}} (B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}}) \ .$$

On voit alors que

$$S_t^{(n)} - I_t^{(n)} = \sum_{i=0}^{n-1} (B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}})^2 \ .$$

Or le Théorème II.26 montre que cette dernière quantité tend en probabilité vers la variation quadratique du mouvement Brownien au temps t, c'est-à-dire t. Autrement dit les deux suites de sommes finies ne peuvent avoir la même limite !

Cet exemple est très instructif : il illustre non seulement l'instabilité des schémas d'approximation de l'intégrale contre le mouvement Brownien, et il suggère également que la source des difficultés est la *variation quadratique* du mouvement Brownien. On rappelle que la théorie de l'intégration contre des fonctions à variation finie est un cas particulier de la théorie générale de l'intégration (de Lebesgue). Si f est à variation finie, alors $\mu(ds) = df(s)$ est une mesure (signée) contre laquelle on peut intégrer toute fonction mesurable bornée ; en particulier, l'intégrale $\int_{[0,t]} f(s)df(s)$ est bien définie, et peut être obtenue par approximation.
On a vu que la trajectoire Brownienne est à variation infinie p.s., ainsi il n'est pas possible d'appliquer la théorie de la mesure pour définir l'intégrale contre le Brownien. Dans ce chapitre nous allons présenter une théorie de l'intégration contre le Brownien due à Itô (1948), ainsi que sa généralisation contre toute martingale à trajectoires continues due à Kunita et Watanabe (1967). Nous allons voir que la construction par Itô consiste à approximer l'intégrale en choisissant la valeur de l'intégrand *au début* de chaque intervalle de la subdivision : dans l'exemple ci-dessus, il s'agit donc de $I_t^{(n)}$. On notera que de telles approximations préservent la propriété de martingale du Brownien : en particulier, $I_t^{(n)}$ est une martingale (dont on peut calculer explicitement la variation quadratique !). Cela jouera un rôle prépondérant dans la construction de l'intégrale.

1 Intégrale stochastique contre le mouvement Brownien

1) Intégrale stochastique des processus élémentaires

Le point de départ de la construction est de considérer la classe $\mathcal{E}$ des *processus élémentaires*, c'est-à-dire, de tous les processus H de la forme

$$H_t(\omega) = \sum_{i=0}^{p-1} H^{(i)}(\omega) \mathbf{1}_{]t_i, t_{i+1}]}(t) \ , \quad t \in \mathbb{R}_+ \ , \omega \in \Omega \ ,$$

où p est un entier positif ou nul, $0 = t_0 < t_1 < \ldots < t_p$ est une collection de réels positifs et chaque $H^{(i)}$ est une v.a. $\mathcal{F}_{t_i}$-mesurable bornée.

Remarque III.1

On impose ici que chaque intervalle $]t_i, t_{i+1}]$ est fermé à droite et ouvert à gauche. Il se trouve *que ce choix n'a aucun impact sur la suite de la construction : on pourrait considérer par exemple l'intervalle ouvert à droite et fermé à gauche sans changer les résultats à venir. Cela est dû à la continuité de la trajectoire Brownienne.*

Pour tout processus élémentaire H, on définit l'intégrale de H contre B comme la v.a.

$$I(H) := \sum_{i=0}^{p-1} H^{(i)}(\omega)(B_{t_{i+1}} - B_{t_i})(\omega) .$$

Pour définir l'intégrale de H contre B restreinte à chaque intervalle $[0,t]$, on commence par noter que pour tout $t \in \mathbb{R}_+$, $H\mathbf{1}_{[0,t]}$ est encore un processus élémentaire , et que

$$I(H\mathbf{1}_{[0,t]}) = \sum_{i=0}^{p-1} H^{(i)}(\omega)(B_{t_{i+1}\wedge t} - B_{t_i\wedge t})(\omega) .$$

Pour tout processus élémentaire H, on pose alors

$$I_t(H) := I(H\mathbf{1}_{[0,t]}) , \quad t \in \mathbb{R}_+ .$$

Il est facile de voir que le processus $(I_t(H), t \in \mathbb{R}_+)$ est constant égal à $I(H)$ à partir du temps t_p.

Lemme III.2

Pour tout processus élémentaire H, le processus $(I_t(H), t \in \mathbb{R}_+)$ est une martingale continue issue de 0, bornée dans L^2 et de variation quadratique

$$\langle I.(H), I.(H)\rangle_t = \int_0^t H_s^2 ds .$$

Démonstration Tout d'abord, il n'est pas difficile de vérifier que $I_t(H)$ est $\mathcal{F}_t$-mesurable pour tout $t \in \mathbb{R}_+$ et intégrable (il s'agit d'une somme finie de v.a. bornées). Vérifions la propriété de martingale. Pour tout $s \in \mathbb{R}_+$, soit $j \in \{0, \ldots, p\}$ l'unique entier tel que $t_j \leqslant s < t_{j+1}$ (avec pour convention que $t_{p+1} = \infty$). On écrit alors pour tout $t \geqslant s$

$$I_t(H) = \sum_{i=0}^{j-1} H^{(i)}(B_{t_{i+1}} - B_{t_i}) + H^{(j)}(B_{t_{j+1}\wedge t} - B_{t_j}) + \sum_{i=j+1}^{p} H^{(i)}(B_{t_{i+1}\wedge t} - B_{t_i\wedge t}) .$$

Le premier terme est $\mathcal{F}_{t_j} \subset \mathcal{F}_s$-mesurable. Le second terme est la somme du terme $\mathcal{F}_s$-mesurable $H^{(j)}(B_s - B_{t_j})$ et de $H^{(j)}(B_{t_{j+1}\wedge t} - B_s)$ dont l'espérance conditionnelle vaut

$$\mathbb{E}[H^{(j)}(B_{t_{j+1}\wedge t} - B_s) \mid \mathcal{F}_s] = H^{(j)}\mathbb{E}[B_{t_{j+1}\wedge t} - B_s \mid \mathcal{F}_s] = H^{(j)}\mathbb{E}[B_{t_{j+1}\wedge t} - B_s] = 0 .$$

Calculons l'espérance conditionnelle du troisième terme. Soit $i \in \{j+1, \ldots, p\}$. Si $t < t_i$ alors $H^{(i)}(B_{t_{i+1}\wedge t} - B_{t_i\wedge t}) = 0$ et l'espérance conditionnelle s'annule. Si $t \geqslant t_i$ alors

$$\begin{aligned}
\mathbb{E}[H^{(i)}(B_{t_{i+1}\wedge t} - B_{t_i\wedge t}) \mid \mathcal{F}_s] &= \mathbb{E}[\mathbb{E}[H^{(i)}(B_{t_{i+1}\wedge t} - B_{t_i\wedge t}) \mid \mathcal{F}_{t_i}] \mid \mathcal{F}_s] \\
&= \mathbb{E}[H^{(i)}\mathbb{E}[B_{t_{i+1}\wedge t} - B_{t_i\wedge t} \mid \mathcal{F}_{t_i}] \mid \mathcal{F}_s] \\
&= \mathbb{E}[H^{(i)}\mathbb{E}[B_{t_{i+1}\wedge t} - B_{t_i\wedge t}] \mid \mathcal{F}_s] \\
&= 0 .
\end{aligned}$$

On a donc montré que

$$\mathbb{E}[I_t(H) \mid \mathcal{F}_s] = \sum_{i=0}^{j-1} H^{(i)}(B_{t_{i+1}} - B_{t_i}) + H^{(j)}(B_s - B_{t_j}) = \sum_{i=0}^{p-1} H^{(i)}(B_{t_{i+1}\wedge s} - B_{t_i \wedge s}) = I_s(H) .$$

On est donc face à une martingale. Notons que pour tout $t \in \mathbb{R}_+$

$$\mathbb{E}[I_t(H)^2]^{1/2} \leqslant \sum_{i=0}^{p-1} \mathbb{E}[(H^{(i)})^2(B_{t_{i+1}\wedge t} - B_{t_i\wedge t})^2]^{1/2} .$$

Comme H est un processus élémentaire, il existe une constante $C > 0$ telle que pour tout i

$$\mathbb{E}[(H^{(i)})^2(B_{t_{i+1}\wedge t} - B_{t_i\wedge t})^2]^{1/2} \leqslant C\mathbb{E}[(B_{t_{i+1}\wedge t} - B_{t_i\wedge t})^2]^{1/2} \leqslant C(t_{i+1} - t_i)^{1/2} ,$$

de sorte que $I_t(H)$ est bornée dans L^2. La Proposition II.28 et le Théorème II.26 assurent alors que $\langle I.(H), I.(H)\rangle_t$ est l'unique processus adapté et croissant tel que $I_t(H)^2 - \langle I.(H), I.(H)\rangle_t$ est une (vraie) martingale. Comme $\int_0^t H_s^2 ds$ est adapté et croissant, il nous suffit de vérifier que $I_t(H)^2 - \int_0^t H_s^2 ds$ est une martingale. On écrit alors

$$I_t(H)^2 = \sum_{i=0}^{p-1}(H^{(i)})^2(B_{t_{i+1}\wedge t} - B_{t_i\wedge t})^2 + 2\sum_{i<j:i,j=0}^{p-1} H^{(i)}H^{(j)}(B_{t_{i+1}\wedge t} - B_{t_i\wedge t})(B_{t_{j+1}\wedge t} - B_{t_j\wedge t}) .$$

Un calcul, similaire au précédent, montre que l'espérance conditionnelle sachant $\mathcal{F}_s$ du premier terme vaut

$$\sum_{i=0}^{p-1}(H^{(i)})^2(B_{t_{i+1}\wedge s} - B_{t_i\wedge s})^2 + \mathbb{E}[\int_s^t H_r^2 dr \mid \mathcal{F}_s] , \tag{III.1}$$

tandis que celle du second terme vaut

$$2\sum_{i<j:i,j=0}^{p-1} H^{(i)}H^{(j)}(B_{t_{i+1}\wedge s} - B_{t_i\wedge s})(B_{t_{j+1}\wedge s} - B_{t_j\wedge s}) .$$

En regroupant les termes, on obtient alors que

$$\mathbb{E}[I_t(H)^2 - \int_0^t H_r^2 dr \mid \mathcal{F}_s] = I_s(H)^2 - \int_0^s H_r^2 dr .$$

Donnons les détails du calcul du premier terme. En reprenant les notations du début de la preuve, on obtient aisément que

$$\begin{aligned}\mathbb{E}[\sum_{i=0}^{p-1}(H^{(i)})^2(B_{t_{i+1}\wedge t} - B_{t_i\wedge t})^2 \mid \mathcal{F}_s] &= \sum_{i=0}^{j-1}(H^{(i)})^2(B_{t_{i+1}} - B_{t_i})^2 \\ &+ (H^{(j)})^2(B_s - B_{t_j})^2 + \mathbb{E}[(H^{(j)})^2(B_{t_{j+1}\wedge t} - B_s)^2 \mid \mathcal{F}_s] \\ &+ \mathbb{E}[\sum_{i=j+1}^{p} (H^{(i)})^2(B_{t_{i+1}\wedge t} - B_{t_i\wedge t})^2 \mid \mathcal{F}_s]\end{aligned}$$

Or

$$\begin{aligned}\mathbb{E}[(H^{(j)})^2(B_{t_{j+1}\wedge t}-B_s)^2 \mid \mathcal{F}_s] &= (H^{(j)})^2\mathbb{E}[(B_{t_{j+1}\wedge t}-B_s)^2 \mid \mathcal{F}_s]\\ &= (H^{(j)})^2\mathbb{E}[(B_{t_{j+1}\wedge t}-B_s)^2]\\ &= (H^{(j)})^2(t_{j+1}\wedge t-s)\ ,\end{aligned}$$

tandis que

$$\begin{aligned}\mathbb{E}[(H^{(i)})^2(B_{t_{i+1}\wedge t}-B_{t_i\wedge t})^2 \mid \mathcal{F}_s] &= \mathbb{E}[\mathbb{E}[(H^{(i)})^2(B_{t_{i+1}\wedge t}-B_{t_i\wedge t})^2 \mid \mathcal{F}_{t_i\wedge t}] \mid \mathcal{F}_s]\\ &= \mathbb{E}[(H^{(i)})^2\mathbb{E}[(B_{t_{i+1}\wedge t}-B_{t_i\wedge t})^2] \mid \mathcal{F}_s]\\ &= \mathbb{E}[(H^{(i)})^2 \mid \mathcal{F}_s](t_{i+1}\wedge t-t_i\wedge t)\ .\end{aligned}$$

Or, le processus H étant constant par morceaux, on obtient

$$\int_s^t H_r^2 dr = (H^{(j)})^2(t_{j+1}\wedge t-s) + \sum_{i=j+1}^{p}(H^{(i)})^2(t_{i+1}\wedge t-t_i\wedge t)\ ,$$

et

$$\mathbb{E}[\int_s^t H_r^2 dr \mid \mathcal{F}_s] = (H^{(j)})^2(t_{j+1}\wedge t-s) + \sum_{i=j+1}^{p}\mathbb{E}[(H^{(i)})^2 \mid \mathcal{F}_s](t_{i+1}\wedge t-t_i\wedge t)\ .$$

En regroupant les termes, on obtient (III.1). □

On notera que $\mathcal{E}$ est un espace vectoriel et que $H \mapsto I(H)$ est linéaire. Ainsi, par polarisation on calcule que pour tous $H, K \in \mathcal{E}$

$$\langle I_\cdot(H), I_\cdot(K)\rangle_t = \int_0^t H_s K_s ds\ .$$

On notera également que le lemme précédent, combiné au fait que $I_t(H)$ est constant égal à $I(H)$ pour t assez grand, assure que

$$\mathbb{E}[I(H)^2] = \mathbb{E}[\int_0^\infty H_s^2 ds]\ .$$

2) Intégrale stochastique des processus progressifs intégrables

On souhaite étendre la construction précédente à des integrands plus généraux. Par exemple, on veut pouvoir considérer des processus H dont les trajectoires sont continues, voire seulement continues à droites.

Quelle est la "bonne" classe d'intégrands à considérer ?

Jusqu'à présent nous avons construit une application linéaire I sur l'espace vectoriel $\mathcal{E}$ et nous avons vu que pour tout $H \in \mathcal{E}$

$$\mathbb{E}[I(H)^2] = \mathbb{E}[\int_0^\infty H_s^2 ds]\ .$$

L'idée est d'utiliser cette identité comme une propriété d'*isométrie* afin d'étendre par *densité* l'application I.

On considère alors $\mathcal{H}^2$, l'ensemble des processus $(H_s, s \in \mathbb{R}_+)$ tels que :

(i) (Mesurabilité) l'application $(s, \omega) \mapsto H_s(\omega)$ est mesurable de $(\mathbb{R}_+ \times \Omega, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F})$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$,

(ii) (Progressive mesurabilité) pour tout $t \in \mathbb{R}_+$, l'application $(s, \omega) \mapsto H_s(\omega)$ est mesurable de $([0, t] \times \Omega, \mathcal{B}([0, t]) \otimes \mathcal{F}_t)$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$,

(iii) (Intégrabilité) $\mathbb{E}[\int_0^\infty H_s^2 ds] < \infty$

Quelques commentaires s'imposent. L'hypothèse (iii) est nécessaire pour établir la propriété d'isométrie. L'hypothèse (i) assure que pour tout ω, $s \mapsto H_s(\omega)$ est mesurable de sorte que l'on peut manipuler $\int_0^\infty H_s^2 ds$. Elle assure aussi que cette dernière quantité est une variable aléatoire et qu'ainsi on peut en calculer l'espérance, ce qui est nécessaire pour énoncer l'hypothèse (iii). L'hypothèse (ii) est plus subtile : elle assure que les processus élémentaires sont denses dans $\mathcal{H}^2$.

Cet ensemble de processus est compliqué, et on ne cherchera pas à en caractériser tous les éléments. En revanche, il contient les processus que l'on souhaite manipuler :

Lemme III.3

Tout processus stochastique adapté dont les trajectoires sont continues à gauche vérifie les propriétés (i) et (ii). En particulier, $\mathcal{E} \subset \mathcal{H}^2$.

Démonstration Soit H un processus adapté dont les trajectoires sont continues à gauche. Pour tout $n \geq 1$, on définit le processus

$$H_t^{(n)} := H_{k2^{-n}} , \quad k2^{-n} \leq t < (k+1)2^{-n} .$$

Par continuité à gauche des trajectoires, $H_t^{(n)} \to H_t$ pour tout $t \in \mathbb{R}_+$. Il suffit donc d'établir (i) et (ii) pour chaque $H^{(n)}$.
Soit A un Borélien, on a

$$(H^{(n)})^{-1}(A) = \bigcup_k \big[k2^{-n}, (k+1)2^{-n}\big[\times \{\omega : H_{k2^{-n}}(\omega) \in A\} ,$$

et cette union dénombrable d'ensembles produits appartient à $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}$ (ici on utilise le caractère adapté de H). La mesurabilité (i) de $H^{(n)}$ s'en suit.
La progressive mesurabilité (ii) s'établit de façon quasiment identique (on doit restreindre la trajectoire à $[0, t]$). $\square$

Mentionnons enfin que l'espace $\mathcal{H}^2$ est un espace de Hilbert une fois muni du produit scalaire (nous ne démontrons pas cette affirmation)

$$\langle H, K \rangle_{\mathcal{H}^2} := \mathbb{E}[\int_{\mathbb{R}_+} H_s K_s ds] .$$

Venons-en au résultat de densité évoqué précédemment. Etant donné une fonction $f \in L^2(\mathbb{R}_+)$ et un entier $n \geq 1$, on définit la fonction en escalier

$$(P_n f)_t := \sum_{i=1}^{n2^n} \Big[(-n) \vee \Big(2^n \int_{(i-1)2^{-n}}^{i2^{-n}} f_s ds\Big) \wedge n\Big] \mathbf{1}_{]i2^{-n},(i+1)2^{-n}]}(t) , \quad t \in \mathbb{R}_+ .$$

On peut montrer que pour tout $f \in L^2(\mathbb{R}_+)$:

1. $\|P_n f\|_{L^2(\mathbb{R}_+)} \leq \|f\|_{L^2(\mathbb{R}_+)}$,
2. $P_n f \to f$ dans $L^2(\mathbb{R}_+)$.

Lemme III.4

L'ensemble $\mathcal{E}$ est dense dans $\mathcal{H}^2$. Plus précisément, si $H \in \mathcal{H}^2$ alors :

1. *pour tout $n \geqslant 1$, $((P_nH)_t, t \in \mathbb{R}_+)$ est un processus élémentaire,*
2. *$((P_nH)_t, t \in \mathbb{R}_+)$ converge, quand $n \to \infty$, vers H dans $\mathcal{H}^2$.*

Démonstration Pour vérifier le premier point, on observe que $t \mapsto (P_nH)_t$ est une fonction en escalier, qui s'annule hors de $[2^{-n}, n + 2^{-n}]$ et que sa valeur est constante sur tout intervalle de la forme $]i2^{-n}, (i+1)2^{-n}]$ et que cette valeur est donnée par une v.a. $\mathcal{F}_{i2^{-n}}$-mesurable bornée.
Pour vérifier le second point, on commence par observer que

$$\|P_nH - H\|^2_{\mathcal{H}^2} = \mathbb{E}[\|P_nH - H\|^2_{L^2(\mathbb{R}_+)}] \ .$$

Or les propriétés de l'opérateur P_n assurent que pour tout $\omega \in \Omega$

$$\|P_nH - H\|_{L^2(\mathbb{R}_+)} \to 0 \ , \quad n \to \infty \ .$$

et

$$\|P_nH(\omega) - H(\omega)\|_{L^2(\mathbb{R}_+)} \leqslant \|P_nH(\omega)\|_{L^2(\mathbb{R}_+)} + \|H(\omega)\|_{L^2(\mathbb{R}_+)} \leqslant 2\|H(\omega)\|_{L^2(\mathbb{R}_+)} \ .$$

Ainsi le Théorème de Convergence Dominée assure que

$$\|P_nH - H\|_{\mathcal{H}^2} \to 0 \ , \quad n \to \infty \ .$$

□

On peut alors prolonger de manière unique l'application I en une application linéaire continue sur l'espace $\mathcal{H}^2$ tout entier. Cette application vérifie la propriété d'isométrie : pour tout $H \in \mathcal{H}^2$

$$\mathbb{E}[I(H)^2] = \mathbb{E}[\int_0^\infty H_t^2 dt] \ .$$

On peut ensuite introduire l'intégrale fonction de sa borne supérieure en posant

$$I_t(H) := I(H\mathbf{1}_{[0,t]}) \ , \quad t \in \mathbb{R}_+ \ .$$

Proposition III.5

Pour tout $H \in \mathcal{H}^2$, le processus $(I_t(H), t \in \mathbb{R}_+)$ est une martingale continue issue de 0, bornée dans L^2 et de variation quadratique

$$\langle I_\cdot(H), I_\cdot(H)\rangle_t = \int_0^t H_s^2 ds \ .$$

Démonstration Ces propriétés s'obtiennent en approximant H par une suite de processus élémentaires et en passant à la limite. □

3) Intégrale stochastique des processus progressifs localement intégrables

Il est naturel de vouloir définir l'intégrale stochastique d'un processus H adapté et à trajectoires continues ; malheureusement, l'espace $\mathcal{H}^2$ considéré jusqu'à présent est trop restrictif. En effet, un tel processus vérifie les hypothèses (i) et (ii), mais rien ne garantit que l'hypothèse d'intégrabilité

(iii) est vérifiée. Cependant, la continuité des trajectoires assure que pour tout $t \in \mathbb{R}_+$, $\int_0^t H_s^2 ds < \infty$ p.s. Le but de ce paragraphe est d'étendre la construction afin de couvrir de tels processus.

On considère l'espace $\mathcal{H}^2_{\text{loc}}$ de tous les processus H vérifiant :

(i) (Mesurabilité) l'application $(s, \omega) \mapsto H_s(\omega)$ est mesurable de $(\mathbb{R}_+ \times \Omega, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F})$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$,

(ii) (Progressive mesurabilité) pour tout $t \in \mathbb{R}_+$, l'application $(s, \omega) \mapsto H_s(\omega)$ est mesurable de $([0, t] \times \Omega, \mathcal{B}([0, t]) \otimes \mathcal{F}_t)$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$,

(iii) (Intégrabilité locale) pour tout $t \in \mathbb{R}_+$, $\int_0^t H_s^2 ds < \infty$ p.s.

Soit $H \in \mathcal{H}^2_{\text{loc}}$. On considère la suite de temps d'arrêt $T_n := \inf\{t \geqslant 0 : \int_0^t H_s^2 ds \geqslant n\}$. La suite T_n est croissante et converge vers $+\infty$ p.s. Pour tout n, le processus tronqué $H_t^{(n)} := H_t \mathbf{1}_{[0,T_n]}$ appartient à $\mathcal{H}^2$. Il est facile de vérifier que presque sûrement : pour tout $t \geqslant 0$, la suite $I_t(H^{(n)})$, $n \geqslant 1$, est constante à partir du premier rang n où $T_n \geqslant t$. On pose alors

$$I_t(H) = \lim_{n \to \infty} I_t(H^{(n)}) .$$

Lemme III.6

Pour tout $H \in \mathcal{H}^2_{\text{loc}}$, le processus $(I_t(H), t \in \mathbb{R}_+)$ est une martingale locale issue de 0 de variation quadratique

$$\langle I_\cdot(H), I_\cdot(H) \rangle_t = \int_0^t H_s^2 ds .$$

Démonstration La définition de $I_t(H)$ assure que presque sûrement : pour tout $t \geqslant 0$ et tout $n \geqslant 1$

$$I_{t \wedge T_n}(H) = I_t(H^{(n)}) .$$

En effet, la définition assure que si $t \wedge T_n \leqslant T_m$ alors

$$I_{t \wedge T_n}(H) = I_{t \wedge T_n}(H^{(m)}) .$$

Or $t \wedge T_n \leqslant T_n$ et ainsi

$$I_{t \wedge T_n}(H) = I_{t \wedge T_n}(H^{(n)}) .$$

Or

$$I_{t \wedge T_n}(H^{(n)}) = I_{t \wedge T_n}(H \mathbf{1}_{[0,T_n]}) = I(H \mathbf{1}_{[0,T_n]} \mathbf{1}_{[0,t \wedge T_n]}) = I(H \mathbf{1}_{[0,T_n]} \mathbf{1}_{[0,t]}) = I_t(H \mathbf{1}_{[0,T_n]}) = I_t(H^{(n)}) ,$$

d'où le résultat.

Comme la suite de temps d'arrêt T_n, $n \geqslant 1$ tend vers $+\infty$ p.s., on a montré que $I_t(H)$ est bien une martingale locale et que la suite T_n, $n \geqslant 1$ la réduit. Pour identifier sa variation quadratique, on commence par rappeler que (voir Remarque II.27)

$$\langle I_\cdot(H), I_\cdot(H) \rangle_{t \wedge T_n} = \langle I_{\cdot \wedge T_n}(H), I_{\cdot \wedge T_n}(H) \rangle_t .$$

La première identité prouvée dans cette preuve permet d'écrire

$$\langle I_\cdot(H), I_\cdot(H) \rangle_{t \wedge T_n} = \langle I_\cdot(H^{(n)}), I_\cdot(H^{(n)}) \rangle_t = \int_0^t (H_s^{(n)})^2 ds = \int_0^{t \wedge T_n} H_s^2 ds .$$

Toutes ces expressions sont croissantes en n, on peut donc passer à la limite $n \to \infty$ et obtenir

$$\langle I.(H), I.(H)\rangle_t = \int_0^t H_s^2 ds \ .$$

$\square$

Dans la suite, on manipulera l'intégrale stochastique à l'aide de la notation naturelle suivante :

$$I_t(H) = \int_0^t H_s dB_s \ , \quad t \geqslant 0 \ .$$

Remarque III.7

La notion de martingale locale prend tout son sens ici. L'intégrale stochastique d'un processus H adapté et à trajectoires continues n'est, en général, pas une martingale mais seulement une martingale locale. On peut ainsi fournir de nombreux exemples de martingales locales.

4) Un résultat d'approximation

En lien avec la discussion présentée en début de chapitre, nous énonçons un résultat affirmant que l'intégrale d'Itô que l'on vient de construire résulte bien d'une approximation dans laquelle l'integrand est évalué au *début* de chaque intervalle de la subdivision.

Proposition III.8

Soit H un processus continu adapté. Pour tout $t > 0$ et toute suite $0 = t_0^n < \cdots < t_{p_n}^n = t$ de subdivisions de $[0,t]$ de pas tendant vers 0, la convergence suivante a lieu en probabilité

$$\lim_{n\to\infty} \sum_{i=0}^{p_n-1} H_{t_i^n}(B_{t_{i+1}^n} - B_{t_i^n}) = \int_0^t H_s dB_s \ .$$

Démonstration Commençons par supposer qu'il existe $C > 0$ tel que pour tout $s \in [0,t]$ et tout $\omega \in \Omega$, $|H_s(\omega)| < C$. On introduit alors le processus élémentaire

$$K_s^{(n)}(\omega) := \sum_{i=0}^{p_n-1} H_{t_i^n}(\omega)\mathbf{1}_{]t_i^n, t_{i+1}^n]}(s) \ , \quad s \in \mathbb{R}_+ \ .$$

Le théorème de convergence dominée assure que

$$\mathbb{E}\Big[\int_{\mathbb{R}_+} (\mathbf{1}_{[0,t]}(s)H_s - K_s^{(n)})^2 ds\Big] \to 0 \ , \quad n \to \infty \ ,$$

c'est-à-dire $\|\mathbf{1}_{[0,t]}(s)H_s - K_s^{(n)}\|_{\mathcal{H}^2} \to 0$ et ainsi $\int_0^t H_s dB_s - \int_0^t K_s^{(n)} dB_s$ converge vers 0 dans $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Dans le cas général, on introduit le temps d'arrêt $T_m := \inf\{t \geqslant 0 : |H_s| \geqslant m\}$ pour tout entier $m \geqslant 1$. La discussion précédente assure que la convergence suivante a lieu dans L^2 et donc en probabilité

$$\lim_{n\to\infty} \sum_{k=0}^{p_n-1} H_{t_i^n \wedge T_m}(B_{t_{i+1}^n} - B_{t_i^n}) = \int_0^t H_{s\wedge T_m} dB_s \ .$$

Sur l'événement $\{T_m > t\}$, cela assure que

$$\lim_{n\to\infty} \sum_{k=0}^{p_n-1} H_{t_i^n}(B_{t_{i+1}^n} - B_{t_i^n}) = \int_0^t H_s dB_s \ .$$

On conclut en utilisant le fait que $\mathbb{P}(T_m > t) \to 1$ quand $m \to \infty$. □

2 Intégrale stochastique contre une martingale continue

Pourquoi souhaiter généraliser *encore* l'intégrale stochastique ?
On a vu que si $H \in \mathcal{H}^2_{\text{loc}}$ alors

$$\int_0^t H_s dB_s \ , \quad t \geqslant 0 \ ,$$

est une martingale locale. On pourrait souhaiter définir une notion d'intégrale contre cette martingale locale. En fait, la formule d'Itô que nous allons introduire bientôt nécessite d'introduire une telle notion d'intégrale.

Il se trouve que toute la construction présentée précédemment se généralise au cas où B est remplacé par une martingale locale M. La différence principale porte sur la condition d'intégrabilité des integrands : au lieu d'imposer que pour tout $t \in \mathbb{R}_+$

$$\int_0^t H_s^2 ds < \infty \quad \text{p.s.} \ ,$$

on impose

$$\int_0^t H_s^2 d\langle M, M\rangle_s < \infty \quad \text{p.s.} \ .$$

C'est tout-à-fait cohérent car dans le cas où M est un mouvement Brownien $d\langle M, M\rangle_s = ds$.

Le résultat s'énonce comme suit. Soit $\mathcal{H}^2_{\text{loc}}(M)$ l'ensemble des processus H vérifiant :

(i) (Mesurabilité) l'application $(s, \omega) \mapsto H_s(\omega)$ est mesurable de $(\mathbb{R}_+ \times \Omega, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F})$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$,

(ii) (Progressive mesurabilité) pour tout $t \in \mathbb{R}_+$, l'application $(s, \omega) \mapsto H_s(\omega)$ est mesurable de $([0, t] \times \Omega, \mathcal{B}([0, t]) \otimes \mathcal{F}_t)$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$,

(iii) (Intégrabilité locale) pour tout $t \in \mathbb{R}_+$, $\int_0^t H_s^2 d\langle M, M\rangle_s < \infty$ p.s.

Pour tout $H \in \mathcal{H}^2_{\text{loc}}(M)$, il existe un processus $(I_t(H), t \in \mathbb{R}_+)$ qui est une martingale locale issue de 0 de variation quadratique

$$\langle I_\cdot(H), I_\cdot(H)\rangle_t = \int_0^t H_s^2 d\langle M, M\rangle_s \ .$$

Ce processus est caractérisé par la propriété suivante : il s'agit de l'unique martingale locale issue de 0 telle que pour toute martingale locale N

$$\langle I(H), N\rangle_t = \int_0^t H_s d\langle M, N\rangle_s \ , \quad t \in \mathbb{R}_+ \ .$$

Là encore, on utilisera la notation

$$I_t(H) = \int_0^t H_s dM_s \ , \quad t \in \mathbb{R}_+ \ .$$

3 Formule d'Itô

Soit X un processus obtenu comme intégrale stochastique contre un mouvement Brownien :

$$X_t = \int_0^t H_s dB_s \ , \quad t \in \mathbb{R}_+ \ .$$

Si $f : \mathbb{R} \to \mathbb{R}$ est une fonction $\mathcal{C}^1$, alors on s'attend à ce que

$$df(X_t) = f'(X_t) dX_t \ ,$$

c'est-à-dire

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) H_s dB_s \ .$$

Il s'avère que cette identité est *fausse* en général ! Illustrons cela sur un exemple très simple. Prenons $X_t = B_t = \int_0^t dB_s$ et $f(x) = x^2$. Il se trouve que

$$f(B_t) = B_t^2 \neq B_0^2 + 2\int_0^t B_s dB_s = f(B_0) + \int_0^t f'(B_s) dB_s \ .$$

En effet, nous avons vu en tout début de chapitre que

$$\int_0^t B_s dB_s = \lim_n \sum_{i=0}^{n-1} B_{t_i^{(n)}} (B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}}) \ .$$

Or

$$B_t^2 = B_0^2 + \sum_{i=0}^{n-1} (B^2_{t_{i+1}^{(n)}} - B^2_{t_i^{(n)}}) = B_0^2 + 2\sum_{i=0}^{n-1} B_{t_i^{(n)}} (B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}}) + \sum_{i=0}^{n-1} (B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}})^2 \ .$$

Par la Proposition III.8, le premier terme de gauche converge vers $2\int_0^t B_s dB_s$. Par contre, le second terme de gauche converge vers t, par le Théorème II.26. Ainsi

$$B_t^2 = B_0^2 + 2\int_0^t B_s dB_s + t \ .$$

En d'autres termes, l'intégrale stochastique ne vérifie pas les règles usuelles de calcul différentiel. En particulier, la *chain rule* qui affirme que pour une fonction $t \mapsto h_t$ continûment différentiable et $f : \mathbb{R} \to \mathbb{R}$ de classe $\mathcal{C}^1$

$$df(h_t) = f'(h_t) dh_t \ ,$$

n'a plus cours lorsque h est remplacé par un processus mettant en jeu une intégrale stochastique. La formule d'Itô, que l'on introduit à présent, fournit la bonne formule pour une telle composition.

Avant cela, nous devons introduire une classe de processus à laquelle cette formule d'Itô s'appliquera.

Définition III.9

Un processus stochastique $(X_t, t \in \mathbb{R}_+)$ est une semi-martingale s'il existe une martingale locale M

issue de 0, un processus adapté à variation finie A et à trajectoires continues issu de 0 ainsi qu'une v.a. X_0 $\mathcal{F}_0$-mesurable tels que

$$X_t = X_0 + M_t + A_t \ , \quad t \in \mathbb{R}_+ \ .$$

On pose alors $\langle X, X\rangle_t := \langle M, M\rangle_t$ pour tout $t \geqslant 0$.

Par le Théorème II.24 une telle décomposition est nécessairement unique. Un exemple typique de semi-martingale est

$$X_t = X_0 + \int_0^t H_s dB_s + \int_0^t K_s ds \ ,$$

avec $H \in \mathcal{M}^2_{\text{loc}}$ et K un processus progressivement mesurable et localement borné (on parle parfois de processus d'Itô pour désigner de telles semi-martingales).

Remarque III.10

On peut montrer que $\langle X, X\rangle_t$ est la limite en probabilité de la somme des carrés des incréments de X le long d'une suite de subdivisions emboîtées.

Théorème III.11 (*Formule d'Itô en dimension* 1)

Soit X une semi-martingale et soit $f : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ une fonction de classe $\mathcal{C}^{1,2}$. Alors le processus $(f(t, X_t), t \in \mathbb{R}_+)$ satisfait presque sûrement : pour tout $t \in \mathbb{R}_+$

$$f(t, X_t) = f(0, X_0) + \int_0^t \partial_s f(s, X_s) ds + \int_0^t \partial_x f(s, X_s) dX_s + \frac{1}{2}\int_0^t \partial_x^2 f(s, X_s) d\langle X, X\rangle_s \ .$$

Remarque III.12

L'intégrale contre dX_s est la somme de deux termes : une intégrale stochastique contre la partie martingale locale de X

$$\int_0^t \partial_x f(s, X_s) dM_s \ ,$$

et une intégrale classique contre la partie à variation finie de X

$$\int_0^t \partial_x f(s, X_s) dA_s \ .$$

En notation différentielle cette formule s'écrit

$$df(t, X_t) = \partial_t f(t, X_t) + \partial_x f(t, X_t) dX_t + \frac{1}{2}\partial_x^2 f(t, X_t) d\langle X, X\rangle_t \ .$$

Les deux premiers termes apparaissent dans la formule classique (régie par les règles usuelles d'intégration/différentiation), par contre le troisième terme est spécifique à l'intégrale stochastique. On notera d'ailleurs qu'il respose sur la variation quadratique de X et que celle-ci est nulle dans la formule classique.

Démonstration On se contente de présenter la preuve dans le cas où X est un mouvement Brownien B et f ne dépend pas du temps. Le cas général s'appuie sur les mêmes idées. Le gros du travail porte sur le terme mettant en jeu la variation quadratique.

On se donne une suite indicée par $n \geqslant 1$ de subdivisions emboîtées $0 = t_0^n < t_1^n < \ldots < t_{p_n}^n = t$ de l'intervalle $[0, t]$ dont le pas tend vers 0 quand $n \to \infty$. Alors

$$f(B_t) = f(B_0) + \sum_{i=0}^{p_n-1} \left(f(B_{t_{i+1}^n}) - f(B_{t_i^n}) \right) .$$

La formule de Taylor assure que

$$f(B_{t_{i+1}^n}) - f(B_{t_i^n}) = f'(X_{t_i^n})(B_{t_{i+1}^n} - B_{t_i^n}) + \frac{f_{n,i}(\omega)}{2}(B_{t_{i+1}^n} - B_{t_i^n})^2 ,$$

où

$$\inf_{\theta \in [0,1]} f''(B_{t_i^n} + \theta(B_{t_{i+1}^n} - B_{t_i^n})) \leqslant f_{n,i} \leqslant \sup_{\theta \in [0,1]} f''(B_{t_i^n} + \theta(B_{t_{i+1}^n} - B_{t_i^n})); .$$

La Proposition III.8 assure que la convergence suivante a lieu en probabilité

$$\lim_{n \to \infty} \sum_{i=0}^{p_n-1} f'(X_{t_i^n})(B_{t_{i+1}^n} - B_{t_i^n}) = \int_0^t f'(B_s) dB_s .$$

Il nous reste à montrer que la convergence suivante a lieu en probabilité

$$\lim_{n \to \infty} \sum_{i=0}^{p_n-1} f_{n,i}(B_{t_{i+1}^n} - B_{t_i^n})^2 = \int_0^t f''(B_s) ds .$$

L'idée va être d'utiliser deux échelles d'approximation, pour faire apparaître successivement la variation quadratique puis l'approximation de f''. Plus précisément pour tout $m < n$ on introduit la v.a.

$$Z_{m,n} := \sup_{0 \leqslant j \leqslant p_m - 1} \sup_{i : t_j^m \leqslant t_i^n \leqslant t_{j+1}^m} |f_{n,i} - f_{m,j}| .$$

La continuité de f'' assure que $Z_{m,n} \to 0$ presque sûrement, lorsque $m \to \infty$ (on rappelle que $m < n$ et qu'ainsi $n \to \infty$ nécessairement). On note alors que

$$\Big| \sum_{i=0}^{p_n-1} f_{n,i}(B_{t_{i+1}^n} - B_{t_i^n})^2 - \sum_{j=0}^{p_m-1} f_{m,i} \sum_{i : t_j^m \leqslant t_i^n \leqslant t_{j+1}^m} (B_{t_{i+1}^n} - B_{t_i^n})^2 \Big| \leqslant Z_{m,n} \sum_{i=0}^{p_n-1} (B_{t_{i+1}^n} - B_{t_i^n})^2 .$$

Comme la somme apparaissant dans le deuxième terme converge vers t en probabilité, pour tout $\varepsilon > 0$ il existe m tel que pour tout $n > m$

$$\mathbb{P}(Z_{m,n} \sum_{i=0}^{p_n-1} (B_{t_{i+1}^n} - B_{t_i^n})^2 > \varepsilon) < \varepsilon .$$

Pour cette valeur de m fixée, on voit que pour tout $j \in \{0, \ldots, p_m - 1\}$ et au sens de la convergence en probabilité

$$\lim_{n \to \infty} \sum_{i : t_j^m \leqslant t_i^n \leqslant t_{j+1}^m} (B_{t_{i+1}^n} - B_{t_i^n})^2 = t_{j+1}^m - t_j^m .$$

Par linéarité

$$\lim_{n \to \infty} \sum_{j=0}^{p_m-1} f_{m,i} \sum_{i : t_j^m \leqslant t_i^n \leqslant t_{j+1}^m} (B_{t_{i+1}^n} - B_{t_i^n})^2 = \int_0^t h_s^m ds ,$$

où $h^m_s := \sum_{j=0}^{p_m-1} f_{m,i}\mathbf{1}_{[t^m_j,t^m_{j+1}[}(s)$. Quitte à prendre m plus grand, la continuité de f'' assure que

$$\mathbb{P}(\sup_{s\in[0,t]} |h^m_s - f''(B_s)| > \varepsilon) < \varepsilon \ ,$$

et ainsi

$$\mathbb{P}(|\int_0^t h^m_s ds - \int_0^t f''(B_s)ds| \geqslant t\varepsilon) < \varepsilon \ .$$

En regroupant les estimations précédentes, on a montré que

$$\mathbb{P}(\Big|\sum_{i=0}^{p_n-1} f_{n,i}(B_{t^n_{i+1}} - B_{t^n_i})^2 - \int_0^t f''(B_s)ds\Big| \geqslant \varepsilon(2+t)) < 2\varepsilon \ ,$$

ce qui conclut la preuve. □

Enonçons à présent la formule dans le cas multi-dimensionnel.

Théorème III.13 (*Formule d'Itô en dimension quelconque*)

Soient $X^{(1)},\dots,X^{(d)}$ d semi-martingales et soit $f:\mathbb{R}_+\times\mathbb{R}^d\to\mathbb{R}$ une fonction de classe $\mathcal{C}^{1,2}$. Alors le processus $(f(t,X_t), t\in\mathbb{R}_+)$ satisfait l'équation

$$\begin{aligned} f(t,X^{(1)}_t,\dots,X^{(d)}_t) &= f(0,X^{(1)}_0,\dots,X^{(d)}_0) + \int_0^t \partial_s f(s,X^{(1)}_s,\dots,X^{(d)}_s)ds \\ &+ \sum_{i=1}^d \int_0^t \partial_{x^{(i)}} f(s,X^{(1)}_s,\dots,X^{(d)}_s)dX^{(i)}_s \\ &+ \frac{1}{2}\sum_{i,j=1}^d \int_0^t \partial^2_{x^{(i)},x^{(j)}} f(s,X_s)d\langle X^{(i)},X^{(j)}\rangle_s \ . \end{aligned}$$

Cette formule s'appuie sur la covariation entre deux semi-martingales. Si X et Y sont deux semi-martingales, dont les parties martingales locales sont données par M et N respectivement alors

$$\langle X,Y\rangle_t := \langle M,N\rangle_t \ .$$

On peut vérifier la prorpriété de bilinéarité suivante :

$$\langle aX+bY,Z\rangle_t = a\langle X,Z\rangle + b\langle Y,Z\rangle_t \ ,$$

et de symétrie

$$\langle X,Y\rangle_t = \langle Y,X\rangle_t \ .$$

Chapitre IV

Outils en calcul stochastique

Approx. 1 séance de cours

1 Martingales exponentielles

On a déjà observé en TD que le processus $(\exp(B_t - \frac{1}{2}t), t \in \mathbb{R}_+)$ est une martingale. La preuve de ce résultat s'appuie sur l'indépendance des incréments du Brownien et la connaissance de la transformée de Laplace de la gaussienne. Il se trouve que ce résultat n'est pas limité au mouvement Brownien comme le montre la proposition suivante.

Proposition IV.1

Soit L une martingale locale issue de 0. On définit

$$\mathcal{E}(L)_t := \exp(L_t - \frac{1}{2}\langle L, L\rangle_t) \ , \quad t \in \mathbb{R}_+ \ .$$

Le processus $(\mathcal{E}(L)_t, t \in \mathbb{R}_+)$ est une martingale locale issue de 1.

Le processus ci-dessus est parfois appelé *exponentielle stochastique* ou *exponentielle de Doléans-Dade* de M.

Démonstration On va appliquer la formule d'Itô à la paire de semi-martingales $(L_t, \langle L, L\rangle_t)$. On notera que le second processus est à variation finie de sorte que sa variation quadratique est nulle ainsi que sa covariation avec L.

Soit $F(x,y) := \exp(x - \frac{1}{2}y)$. Il s'agit d'une fonction $\mathcal{C}^2$ en ses deux arguments qui vérifie

$$\partial_x F = F \ , \quad \frac{1}{2}\partial_x^2 F + \partial_y F = 0 \ .$$

On applique alors la formule d'Itô à F et à la paire de semi-martingales $(L_t, \langle L, L\rangle_t)$ pour obtenir

$$\begin{aligned} F(L_t, \langle L, L\rangle_t) &= 1 + \int_0^t \partial_x F(L_s, \langle L, L\rangle_s) dL_s \\ &+ \int_0^t \Big(\frac{1}{2}\partial_x^2 F + \partial_y F\Big)(L_s, \langle L, L\rangle_s) d\langle L, L\rangle_s \\ &= 1 + \int_0^t F(L_s, \langle L, L\rangle_s) dL_s \ . \end{aligned}$$

ce qui prouve le résultat. □

Remarque IV.2

Si L est une martingale locale et $z \in \mathbb{C}$ alors le processus $(\mathcal{E}(zL)_t = \exp(zL_t - \frac{z^2}{2}\langle L, L\rangle_t), t \in \mathbb{R}_+)$ est une martingale locale à valeurs complexes. La preuve est identique à celle présentée ci-dessus, il suffit d'appliquer la formule d'Itô séparément aux parties réelles et imaginaires de la fonction $F(x,y) = \exp(zx - z^2\frac{1}{2}y)$.

Dans la suite, il sera utile d'avoir un critère assurant que l'exponentielle stochastique construite ci-dessus est une vraie martingale.

Théorème IV.3 (*Novikov*)

Soit L une martingale locale issue de 1. On fixe $t > 0$. Si

$$\mathbb{E}\Big[\exp\Big(\frac{1}{2}\langle L, L\rangle_t\Big)\Big] < \infty ,$$

alors $(\mathcal{E}(L)_s, s \in [0,t])$ est une martingale.

On rappelle que toute martingale sur un intervalle de temps borné est uniformément intégrable (elle est fermée par sa valeur terminale).

Démonstration Soit $\varepsilon \in (0,1)$. Il est facile de vérifier que

$$\Big(\mathcal{E}((1-\varepsilon)L)_s\Big)^{\frac{1}{1-\varepsilon^2}} = \Big(\mathcal{E}(L)_s\Big)^{\frac{1}{1+\varepsilon}}\Big(e^{\frac{1}{2}\langle L,L\rangle_s}\Big)^{\frac{\varepsilon}{1+\varepsilon}} .$$

On introduit le temps d'arrêt $T_n := \inf\{t \geqslant 0 : |L|_t + \langle L, L\rangle_t \geqslant n\}$. Il s'agit d'une suite qui réduit les martingales locales $\mathcal{E}(L)$ et $\mathcal{E}((1-\varepsilon)L)$. On applique l'égalité précédente au temps $s = t \wedge T_n$ et l'on en prend l'espérance (qui est bien définie car tout est positif) :

$$\mathbb{E}\Big[\Big(\mathcal{E}((1-\varepsilon)L)_{t\wedge T_n}\Big)^{\frac{1}{1-\varepsilon^2}}\Big] = \mathbb{E}\Big[\Big(\mathcal{E}(L)_{t\wedge T_n}\Big)^{\frac{1}{1+\varepsilon}}\Big(e^{\frac{1}{2}\langle L,L\rangle_{t\wedge T_n}}\Big)^{\frac{\varepsilon}{1+\varepsilon}}\Big] .$$

On applique alors l'inégalité de Hölder avec $p = 1+\varepsilon$ et $q = (1+\varepsilon)/\varepsilon$ et l'on obtient

$$\mathbb{E}\Big[\Big(\mathcal{E}((1-\varepsilon)L)_{t\wedge T_n}\Big)^{\frac{1}{1-\varepsilon^2}}\Big] \leqslant \mathbb{E}\Big[\mathcal{E}(L)_{t\wedge T_n}\Big]^{\frac{1}{1+\varepsilon}}\mathbb{E}\Big[e^{\frac{1}{2}\langle L,L\rangle_{t\wedge T_n}}\Big]^{\frac{\varepsilon}{1+\varepsilon}} . \tag{IV.1}$$

Comme T_n réduit $\mathcal{E}(L)$, on sait que $\mathbb{E}\Big[\mathcal{E}(L)_{t\wedge T_n}\Big] = \mathbb{E}\Big[\mathcal{E}(L)_0\Big] = 1$. Par ailleurs $\langle L, L\rangle_{t\wedge T_n} \leqslant \langle L, L\rangle_t$ p.s. et ainsi

$$\mathbb{E}\Big[\Big(\mathcal{E}((1-\varepsilon)L)_{t\wedge T_n}\Big)^{\frac{1}{1-\varepsilon^2}}\Big] \leqslant \mathbb{E}\Big[e^{\frac{1}{2}\langle L,L\rangle_t}\Big]^{\frac{\varepsilon}{1+\varepsilon}} .$$

Ceci assure que la suite $(\mathcal{E}((1-\varepsilon)L)_{t\wedge T_n})_{n\geqslant 1}$ est bornée dans L^p pour $p = \frac{1}{1-\varepsilon^2} > 1$, et est ainsi uniformément intégrable (voir la Proposition B.5). Comme cette suite converge p.s. vers $\mathcal{E}((1-\varepsilon)M)_t$, le Théorème de Convergence Dominée Optimal B.6 assure que

$$\mathbb{E}\Big[\mathcal{E}((1-\varepsilon)L)_t\Big] = \lim_{n\to\infty}\mathbb{E}\Big[\mathcal{E}((1-\varepsilon)L)_{t\wedge T_n}\Big] = 1 .$$

L'inégalité de Jensen assure que

$$\mathbb{E}\Big[\mathcal{E}((1-\varepsilon)L)_t^p\Big] \geqslant \mathbb{E}\Big[\mathcal{E}((1-\varepsilon)L)_t\Big]^p = 1 .$$

On reprend alors l'inégalité (IV.1) mais appliquée au temps t plutôt que $t \wedge T_n$ et l'on obtient

$$1 \leqslant \mathbb{E}\Big[\Big(\mathcal{E}((1-\varepsilon)L)_t\Big)^{\frac{1}{1-\varepsilon^2}}\Big] \leqslant \mathbb{E}\Big[\mathcal{E}(L)_t\Big]^{\frac{1}{1+\varepsilon}} \mathbb{E}\Big[e^{\frac{1}{2}\langle L,L\rangle_t}\Big]^{\frac{\varepsilon}{1+\varepsilon}} .$$

En prenant la limite $\varepsilon \downarrow 0$ on trouve

$$\mathbb{E}\Big[\mathcal{E}(L)_t\Big] \geqslant 1 .$$

Pour conclure, il suffit d'appliquer le lemme ci-dessous. □

Lemme IV.4

Soit N une martingale locale positive issue de 1. Alors N est une super-martingale. Par ailleurs $(N_s, s \in [0,t])$ est une martingale si et seulement si $\mathbb{E}[N_t] \geqslant \mathbb{E}[N_0]$.

Démonstration Soit T_n, $n \geqslant 1$ une suite de temps d'arrêt qui réduit N. Pour tout $0 \leqslant s \leqslant t$

$$N_{s\wedge T_n} = \mathbb{E}[N_{t\wedge T_n} \mid \mathcal{F}_s] .$$

Or $N_{s\wedge T_n}$ et $N_{t\wedge T_n}$ convergent p.s. vers N_s et N_t quand $n \to \infty$. Comme il s'agit de v.a. positives, on peut appliquer le Lemme de Fatou pour l'espérance conditionnelle (qui se prouve en utilisant la définition de l'espérance conditionnelle et le lemme de Fatou classique) :

$$\mathbb{E}[N_t \mid \mathcal{F}_s] \leqslant \liminf_n \mathbb{E}[N_{t\wedge T_n} \mid \mathcal{F}_s] = \liminf_n N_{s\wedge T_n} = N_s ,$$

ce qui assure que N est une super-martingale.
Passons à la preuve de l'équivalence. L'implication directe est immédiate. Prouvons l'implication réciproque. On suppose que $\mathbb{E}[N_t] \geqslant \mathbb{E}[N_0]$. La propriété de super-martingale implique $s \mapsto \mathbb{E}[N_s]$ est décroissante. L'inégalité précédente implique que $\mathbb{E}[N_s] = \mathbb{E}[N_t] = \mathbb{E}[N_0]$. Ainsi $N_s - \mathbb{E}[N_t \mid \mathcal{F}_s]$ est une v.a. positive p.s. et d'espérance nulle : nécessairement cette v.a. est nulle p.s. □

2 Caractérisation de Lévy et théorème de Dubins-Schwarz

Théorème IV.5 (*Caractérisation de Lévy du mouvement Brownien*)

Soit X un processus adapté, à trajectoires continues et issu de 0. Il y a équivalence entre :

(i) X est un mouvement Brownien,

(ii) X est une martingale locale et $\langle X, X\rangle_t = t$ pour tout $t \in \mathbb{R}_+$.

Démonstration L'implication (i) ⇒ (ii) est déjà connue. Supposons (ii). Par la Proposition IV.1 et la Remarque IV.2, le processus

$$\mathcal{E}(i\lambda X)_t = \exp(i\lambda X_t + \frac{\lambda^2}{2}t) , \quad t \in \mathbb{R}_+ ,$$

est une martingale locale. Comme elle est bornée sur tous les intervalles bornés, il s'agit d'une vraie martingale. Ainsi pour tous $0 \leqslant s \leqslant t$

$$\exp(i\lambda X_s + \frac{\lambda^2}{2}s) = \mathbb{E}[\exp(i\lambda X_t + \frac{\lambda^2}{2}t) \mid \mathcal{F}_s] .$$

Ainsi pour tout $A \in \mathcal{F}_s$

$$\mathbb{E}[\mathbf{1}_A \exp(i\lambda(X_t - X_s))] = \mathbb{P}(A) \exp(-\frac{1}{2}\lambda^2(t-s)) \ .$$

En prenant $A = \Omega$, on déduit que $X_t - X_s$ est de loi $\mathcal{N}(0, t-s)$. Par ailleurs, en prenant $A \in \mathcal{F}_s$ tel que $\mathbb{P}(A) > 0$ et en notant $\mathbb{P}_A$ la probabilité conditionnelle sachant A, on voit que

$$\mathbb{E}_A[\exp(i\lambda(X_t - X_s))] = \exp(-\frac{1}{2}\lambda^2(t-s)) \ .$$

Ainsi, sous la probabilité $\mathbb{P}_A$, la v.a. $X_t - X_s$ suit une $\mathcal{N}(0, t-s)$, ce qui implique que pour toute fonction mesurable positive $f : \mathbb{R} \to \mathbb{R}$

$$\mathbb{E}_A[f(X_t - X_s)] = \mathbb{E}[f(X_t - X_s)] \ ,$$

soit encore

$$\mathbb{E}[\mathbf{1}_A f(X_t - X_s)] = \mathbb{P}(A)\mathbb{E}[f(X_t - X_s)] \ .$$

Cela étant vrai pour tout $A \in \mathcal{F}_s$ et toute fonction mesurable positive f, on en déduit que $X_t - X_s$ est indépendante de $\mathcal{F}_s$. On a donc montré que X est un processus à trajectoires continues, dont les accroissements sont indépendants et stationnaires, et tels que $X_t - X_s \sim \mathcal{N}(0, t-s)$. Par la Proposition I.7, on en déduit que X est un mouvement Brownien. □

Enonçons également la version multi-dimensionnelle de ce résultat, sa preuve est tout à fait similaire.

Théorème IV.6 (*Caractérisation de Lévy du mouvement Brownien multi-dimensionnel*)

Soit $X = (X^{(1)}, \ldots, X^{(d)})$ un processus adapté, à trajectoires continues et issu de 0. Il y a équivalence entre :

1. *X est un mouvement Brownien multi-dimensionnel,*
2. *Chaque processus $X^{(i)}$ est une martingale locale et $\langle X^{(i)}, X^{(j)} \rangle_t = t\mathbf{1}_{i=j}$ pour tout $t \in \mathbb{R}_+$ et tous $i, j \in \{1, \ldots, d\}$.*

Enonçons, sans preuve, un autre résultat reliant martingales et mouvement Brownien.

Théorème IV.7 (*Dubins-Schwarz*)

Soit M une martingale locale issue de 0 telle que $\langle M, M \rangle_\infty = \infty$ p.s. Il existe un mouvement Brownien B tel que presque sûrement

$$\forall t \geqslant 0 \ , \quad M_t = B_{\langle M,M \rangle_t} \ .$$

Ce très joli résultat montre que toute martingale locale issue de 0 est un mouvement Brownien changé de temps. Par ailleurs la contrainte $\langle M, M \rangle_\infty = \infty$ p.s. est plutôt d'ordre technique, et peut être levée.

3 Théorème de Girsanov

Le théorème de Girsanov est un outil très utile en calcul stochastique. Il s'appuie sur un changement de mesure : à l'aide d'une martingale exponentielle, on introduit une mesure de probabilité $\mathbb{Q}$ qui est équivalente à $\mathbb{P}$. Le théorème établit alors une correspondance entre martingales locales sous la mesure d'origine $\mathbb{P}$ et martingales locales sous la nouvelle mesure $\mathbb{Q}$.

Commençons par introduire ce changement de mesure : pour éviter les confusions, nous notons à présent $\mathbb{E}_{\mathbb{P}}$ l'espérance par rapport à la probabilité $\mathbb{P}$ d'origine, et $\mathbb{E}_{\mathbb{Q}}$ l'espérance sous une autre probabilité $\mathbb{Q}$.

Proposition IV.8

Soit L une martingale locale issue de 0. On fixe $t \geqslant 0$ et on suppose que $(\mathcal{E}(L)_s, s \in [0,t])$ est une vraie martingale. On pose pour tout $A \in \mathcal{F}_t$

$$\mathbb{Q}(A) := \mathbb{E}_{\mathbb{P}}[\mathbf{1}_A \mathcal{E}(L)_t] .$$

Alors $\mathbb{Q}$ est une mesure de probabilité sur $(\Omega, \mathcal{F}_t)$ et pour tout $s \in [0,t]$ et tout $A \in \mathcal{F}_s$

$$\mathbb{Q}(A) = \mathbb{E}_{\mathbb{P}}[\mathbf{1}_A \mathcal{E}(L)_s] .$$

Démonstration Vérifions que $\mathbb{Q}$ est effectivement une mesure de probabilité sur $(\Omega, \mathcal{F}_t)$. Il est clair que $\mathbb{Q}(\varnothing) = 0$. Par ailleurs si A_n, $n \geqslant 1$ est une suite d'événements deux-à-deux disjoints alors, en utilisant le théorème de Fubini, on obtient

$$\mathbb{Q}(\cup_n A_n) = \mathbb{E}_{\mathbb{P}}[\mathbf{1}_{\cup_n A_n} \mathcal{E}(L)_t] = \mathbb{E}_{\mathbb{P}}[\sum_n \mathbf{1}_{A_n} \mathcal{E}(L)_t] = \sum_n \mathbb{E}_{\mathbb{P}}[\mathbf{1}_{A_n} \mathcal{E}(L)_t] = \sum_n \mathbb{Q}(A_n) .$$

Soit maintenant $s \in [0,t]$ et $A \in \mathcal{F}_s$. En utilisant la propriété de martingale de $\mathcal{E}(L)$ on trouve :

$$\mathbb{Q}(A) = \mathbb{E}_{\mathbb{P}}[\mathbf{1}_A \mathcal{E}(L)_t] = \mathbb{E}_{\mathbb{P}}[\mathbb{E}_{\mathbb{P}}[\mathbf{1}_A \mathcal{E}(L)_t \mid \mathcal{F}_s]] = \mathbb{E}_{\mathbb{P}}[\mathbf{1}_A \mathbb{E}_{\mathbb{P}}[\mathcal{E}(L)_t \mid \mathcal{F}_s]] = \mathbb{E}_{\mathbb{P}}[\mathbf{1}_A \mathcal{E}(L)_s] .$$

□

Pour s'assurer que $(\mathcal{E}(L)_s, s \in [0,t])$ est une vraie martingale, on peut utiliser le critère de Novikov du Théorème IV.3.

Théorème IV.9

Soit L une martingale locale issue de 0. On fixe $t \geqslant 0$ et on suppose que $(\mathcal{E}(L)_s, s \in [0,t])$ est une vraie martingale. On note alors $\mathbb{Q}$ la mesure de probabilité induite par $\mathcal{E}(L)_t$, comme introduit dans le résultat précédent.
Si $(B_s, s \in [0,t])$ est un mouvement Brownien sous $\mathbb{P}$ alors le processus

$$\tilde{B}_s := B_s - \langle B, L \rangle_s \ , s \in [0,t] ,$$

est un mouvement Brownien sous $\mathbb{Q}$.
Plus généralement si $(M_s, s \in [0,t])$ est une martingale locale sous $\mathbb{P}$ alors le processus

$$\tilde{M}_s := M_s - \langle M, L \rangle_s \ , \quad s \in [0,t] ,$$

est une martingale locale sous $\mathbb{Q}$ de même variation quadratique que M.

Démonstration On se contente de prouver l'énoncé qui concerne $\tilde{B}$. On commence par supposer qu'il existe une constante $C > 0$ telle que p.s. $\langle L, L \rangle_t \leqslant C$. On peut alors vérifier que p.s. pour tout

$s \in [0,t]$

$$|\langle L,B\rangle_s| \leqslant \frac{1}{2}\big(\langle L,L\rangle_s + \langle B,B\rangle_s\big) \leqslant \frac{1}{2}\big(C+t\big)\ .$$

Pour la première inégalité, on a utilisé le fait que $\langle L-B,L-B\rangle \geqslant 0$, $\langle L+B,L+B\rangle \geqslant 0$ et ainsi

$$-\langle L,B\rangle \leqslant \frac{1}{2}(\langle L,L\rangle + \langle B,B\rangle)\ ,$$
$$\langle L,B\rangle \leqslant \frac{1}{2}(\langle L,L\rangle + \langle B,B\rangle)\ ,$$

de sorte que $|\langle L,B\rangle_s| \leqslant \frac{1}{2}\big(\langle L,L\rangle_s + \langle B,B\rangle_s\big)$.
Pour tout $\theta \in \mathbb{R}$

$$|\langle L+\theta B, L+\theta B\rangle_t| = |\langle L,L\rangle_t + 2\theta\langle L,B\rangle_t + \theta t| \leqslant C + |\theta|(C+t) + |\theta| t\ .$$

le critère de Novikov assure que la martingale locale $\mathcal{E}(L+\theta B)$ est une vraie martingale sur l'intervalle $[0,t]$. On en déduit que pour tous $0 \leqslant u \leqslant v \leqslant t$

$$\mathbb{E}_{\mathbb{P}}[\mathcal{E}(L+\theta B)_v \mid \mathcal{F}_u] = \mathcal{E}(L+\theta B)_u\ .$$

Un calcul simple permet d'en déduire que

$$\mathbb{E}_{\mathbb{P}}[\mathcal{E}(L)_v \exp(\theta(\tilde{B}_v - \tilde{B}_u)) \mid \mathcal{F}_u] = \mathcal{E}(L)_u \exp(\frac{1}{2}\theta^2(v-u))\ .$$

Soit $A \in \mathcal{F}_u$, on obtient alors

$$\mathbb{E}_{\mathbb{P}}[\mathcal{E}(L)_v \exp(\theta(\tilde{B}_v - \tilde{B}_u))\mathbf{1}_A] = \exp(\frac{1}{2}\theta^2(v-u))\mathbb{E}_{\mathbb{P}}[\mathcal{E}(L)_u\mathbf{1}_A]\ .$$

Cela se ré-écrit

$$\mathbb{E}_{\mathbb{Q}}[\exp(\theta(\tilde{B}_v - \tilde{B}_u))\mathbf{1}_A] = \exp(\frac{1}{2}\theta^2(v-u))\mathbb{Q}(A)\ .$$

En prenant $A = \Omega$, on voit que $\tilde{B}_v - \tilde{B}_u \sim \mathcal{N}(0, v-u)$ sous $\mathbb{Q}$. Par ailleurs, un raisonnement similaire à celui présenté dans la preuve de la caractérisation de Lévy du mouvement Brownien permet de déduire de l'identité précédente que $\tilde{B}_v - \tilde{B}_u$ est indépendante de $\mathcal{F}_u$ sous $\mathbb{Q}$. Le processus $\tilde{B}$ est à trajectoires continues, à accroissements indépendants et pour tous $0 \leqslant u \leqslant v \leqslant t$, $\tilde{B}_v - \tilde{B}_u \sim \mathcal{N}(0, v-u)$: par la Proposition I.7 on en déduit que $(\tilde{B}_s, s \in [0,t])$ est un mouvement Brownien sous $\mathbb{Q}$. Avant de passer au cas général, remarquons que l'on peut déduire de ce qui vient d'être établi que

$$\mathbb{E}_{\mathbb{P}}[\mathcal{E}(L)_v \exp(i\theta(\tilde{B}_v - \tilde{B}_u))\mathbf{1}_A] = \exp(-\frac{1}{2}\theta^2(v-u))\mathbb{E}_{\mathbb{P}}[\mathcal{E}(L)_u\mathbf{1}_A]\ .$$

En effet, comme $\tilde{B}$ est un Brownien sous $\mathbb{Q}$ on a

$$\mathbb{E}_{\mathbb{Q}}[\exp(i\theta(\tilde{B}_v - \tilde{B}_u))\mathbf{1}_A] = \exp(-\frac{1}{2}\theta^2(v-u))\mathbb{E}_{\mathbb{Q}}(A)\ ,$$

or $\mathbb{E}_{\mathbb{Q}}(A) = \mathbb{E}_{\mathbb{P}}[\mathcal{E}(L)_u\mathbf{1}_A]$ et

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}[\exp(i\theta(\tilde{B}_v - \tilde{B}_u))\mathbf{1}_A] &= \mathbb{E}_{\mathbb{P}}[\mathcal{E}(L)_t \exp(i\theta(\tilde{B}_v - \tilde{B}_u))\mathbf{1}_A] \\
&= \mathbb{E}_{\mathbb{P}}[\mathbb{E}_{\mathbb{P}}[\mathcal{E}(L)_t \exp(i\theta(\tilde{B}_v - \tilde{B}_u))\mathbf{1}_A \mid \mathcal{F}_v]] \\
&= \mathbb{E}_{\mathbb{P}}[\exp(i\theta(\tilde{B}_v - \tilde{B}_u))\mathbf{1}_A \mathbb{E}_{\mathbb{P}}[\mathcal{E}(L)_t \mid \mathcal{F}_v]] \\
&= \mathbb{E}_{\mathbb{P}}[\exp(i\theta(\tilde{B}_v - \tilde{B}_u))\mathbf{1}_A \mathcal{E}(L)_v]\ .
\end{aligned}$$

Dans le cas général, on introduit la suite de temps d'arrêt $T_n := \inf\{s \geqslant 0 : \langle L, L\rangle_s \geqslant n\}$ pour $n \geqslant 1$. On considère le processus arrêté L^{T_n}, la martingale $\mathcal{E}(L^{T_n})$ et le processus $\tilde{B}_s^{(n)} := B_s - \langle L^{T_n}, B\rangle_s$. Les arguments précédents assurent que

$$\mathbb{E}_{\mathbb{P}}[\mathcal{E}(L^{T_n})_v \exp(i\theta(\tilde{B}_v^{(n)} - \tilde{B}_u^{(n)}))\mathbf{1}_A] = \exp(-\frac{1}{2}\theta^2(v-u))\mathbb{E}_{\mathbb{P}}[\mathcal{E}(L^{T_n})_u \mathbf{1}_A] .$$

Or pour tout $s \in [0,t]$, $\mathcal{E}(L^{T_n})_s = \mathcal{E}(L)_{s\wedge T_n} \to \mathcal{E}(L)_s$ presque sûrement quand $n \to \infty$. De même $\tilde{B}_v^{(n)} - \tilde{B}_u^{(n)} \to \tilde{B}_v - \tilde{B}_u$. Par ailleurs, les v.a. $\mathcal{E}(L^{T_n})_s$ sont d'espérance (sous $\mathbb{P}$) égale à 1 et sont positives. Le Lemme de Scheffé assure alors $\mathcal{E}(L)_{s\wedge T_n} \to \mathcal{E}(L)_s$ dans L^1 quand $n \to \infty$. On peut alors passer à la limite dans l'identité précédente et obtenir

$$\mathbb{E}_{\mathbb{P}}[\mathcal{E}(L)_v \exp(i\theta(\tilde{B}_v - \tilde{B}_u))\mathbf{1}_A] = \exp(-\frac{1}{2}\theta^2(v-u))\mathbb{E}_{\mathbb{P}}[\mathcal{E}(L)_u \mathbf{1}_A] .$$

On peut alors conclure comme précédemment. □

4 Inégalités de Burkholder-Davis-Gundy

Nous terminons cette section en introduisant des inégalités générales reliant le supremum d'une martingale locale et sa variation quadratique.

Théorème IV.10

Pour tout réel $p > 0$ il existe des constantes $c_p, C_p > 0$ telles que, pour toute martingale locale M issue de 0

$$c_p\mathbb{E}[\langle M, M\rangle_\infty^{p/2}] \leqslant \mathbb{E}[(M_\infty^*)^p] \leqslant C_p\mathbb{E}[\langle M, M\rangle_\infty^{p/2}] ,$$

où $M_t^* = \sup_{s\in[0,t]} |M_s|$ pour tout $t \in [0,\infty]$.

Bien que ces inégalités soient énoncées au temps $+\infty$, on peut facilement les étendre à un temps d'arrêt quelconque. En effet, si T est un temps d'arrêt et M est une martingale locale, alors M^T est encore une martingale locale et l'on a

$$(M_\infty^T)^* = \sup_{t\in[0,T]} |M_t| , \quad \langle M^T, M^T\rangle_\infty = \langle M, M\rangle_T .$$

Chapitre V

Equations différentielles stochastiques et processus de diffusion

Approx. 1 séance de cours
Comme mentionné précédemment, dans de nombreux champs d'application on est amené à modéliser l'évolution d'une quantité d'intérêt X_t par une équation du type

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t \ ,$$

où B est un mouvement Brownien en dimension 1. Une telle équation est appelée équation différentielle stochastique. Le terme $b(t, X_t)dt$ est appelé *terme de drift* : il s'agit d'un terme de nature déterministe ; tandis que le terme $\sigma(t, X_t)dB_t$ est appelé *terme de diffusion* : il s'agit d'un terme de nature aléatoire. En général, le terme aléatoire modélise une incertitude ou une action aléatoire macroscopique due à une multitude d'effets infinitésimaux.

Les processus stochastiques obtenus comme solutions d'équations différentielles stochastiques dirigées par un mouvement Brownien sont communément appelés *processus de diffusion*. Nous verrons dans le chapitre suivant qu'ils tissent des liens étroits avec certains opérateurs différentiels et certaines EDP.

Le but de ce chapitre est d'établir des résultats d'existence et d'unicité pour de telles équations. Commençons par étudier un cas simple.

1 Processus d'Ornstein-Uhlenbeck

Considérons une particule qui se déplace sur la droite réelle et qui est soumise à une force de friction de coefficient $c > 0$. Sa position x_t est alors solution de l'équation différentielle

$$m\frac{d^2}{dt^2}x_t = -c\frac{d}{dt}x_t \ .$$

Si en plus de cette force de friction, la particule subit une multitude de chocs infinitésimaux avec d'autres particules, on peut être amené à modéliser la force résultante par l'incrément d'un mouvement Brownien dB_t. L'équation d'évolution pour la vitesse, notée V_t, de la particule devient alors :

$$dV_t = -bV_tdt + \sigma dB_t \ , \quad t \geqslant 0 \ , \tag{V.1}$$

partant d'une condition initiale $V_0 \in \mathbb{R}$, et où $b, \sigma > 0$ sont des constantes. Cette équation doit être comprise au sens intégral :

$$V_t = V_0 - b\int_0^t V_s ds + \sigma B_t\ , \quad t \geqslant 0\ . \tag{V.2}$$

Proposition V.1

Il existe un unique processus continu V satisfaisant l'équation (V.2). Ce processus, appelé processus d'Ornstein-Uhlenbeck, est donné par

$$V_t = e^{-bt}V_0 + \sigma\int_0^t e^{-b(t-s)}dB_s\ , \quad t \geqslant 0\ .$$

Démonstration L'unicité est facile à vérifier. Si X et Y sont deux solutions alors $Z_t := X_t - Y_t$ satisfait

$$Z_t = -b\int_0^t Z_s ds\ , \quad t \geqslant 0\ ,$$

ainsi que $Z_0 = 0$. Cette équation différentielle ordinaire admet pour unique solution $Z_t = 0$ pour tout $t \geqslant 0$.
Passons à la preuve de l'existence. On procède à un changement d'inconnu $Z_t := V_t - \sigma B_t$. Alors

$$Z_t = V_0 - b\int_0^t Z_s ds + b\sigma\int_0^t B_s ds\ .$$

En d'autres termes, pour tout $\omega \in \Omega$, $t \mapsto Z_t(\omega)$ satisfait une équation différentielle ordinaire :

$$Z_t' + bZ_t = -b\sigma B_t\ .$$

On peut résoudre cette équation explicitement :

$$Z_t = e^{-bt}V_0 - b\sigma\int_0^t e^{-b(t-s)}B_s ds\ ,$$

et ainsi

$$V_t = e^{-bt}V_0 - b\sigma\int_0^t e^{-b(t-s)}B_s ds + \sigma B_t\ , \quad t \geqslant 0\ .$$

Or la formule d'Itô appliquée à $s \mapsto e^{-b(t-s)}B_s$ montre que

$$B_t = \int_0^t e^{-b(t-s)}dB_s + b\int_0^t e^{-b(t-s)}B_s ds\ ,$$

ce qui permet de conclure. (En fait l'argument d'existence permet également de prouver l'unicité).
□

La solution obtenue est composée de deux termes : un terme d'amortissement de la condition initiale $e^{-bt}V_0$ et un terme gaussien $\sigma\int_0^t e^{-b(t-s)}dB_s$ centré de variance

$$\sigma^2\int_0^t e^{-2b(t-s)}ds = \sigma^2\frac{1-e^{-2bt}}{2b}\ .$$

Le terme d'amortissement tend vers 0 quand $t \to \infty$. En revanche le terme gaussien converge en loi vers une v.a. $V_\infty \sim \mathcal{N}(0, \sigma^2/(2b))$.

On pourra noter que dans la construction de la solution de l'équation différentielle stochastique ci-dessus nous n'avons pas fait appel à l'intégrale d'Itô. Ceci est essentiellement dû au fait que le coefficient de diffusion ne dépend pas de la solution. Dans la section suivante, nous allons nous intéresser à une classe plus générale d'EDS pour lesquelles l'intégrale stochastique est nécessaire.

2 Solutions fortes

On se donne deux fonctions $b : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ et $\sigma : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ et l'on considère l'équation

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t \ , \quad t \geqslant 0 \ , \tag{V.3}$$

partant d'une condition initiale $X_0 \in \mathbb{R}^d$ donnée.

Définition V.2

On dit qu'un processus $(X_t, t \in \mathbb{R}_+)$ est solution forte de l'équation (V.3) si :

1. *X est adapté à la filtration du mouvement Brownien B,*
2. *presque sûrement pour tout $t \in \mathbb{R}_+$*

$$\int_0^t |b(s, X_s)|ds < \infty \ , \quad \int_0^t \sigma(s, X_s)^2 ds < \infty \ .$$

3. *presque sûrement pour tout $t \in \mathbb{R}_+$*

$$X_t = X_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s \ .$$

On notera que la condition 2. est nécessaire pour que les intégrales apparaissant dans la condition 3. aient un sens.

Théorème V.3 (*Existence et unicité de solutions fortes*)

On suppose que les fonctions b et σ sont continues et qu'il existe une constante $K > 0$ telle que pour tout $t \in \mathbb{R}_+$ et tous $x, y \in \mathbb{R}$

$$|b(t,0)| + |\sigma(t,0)| \leqslant K \ , \quad |b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leqslant K|x-y| \ .$$

Il existe une unique solution forte $(X_t, t \in \mathbb{R}_+)$ à (V.3).

Avant de présenter la preuve du théorème, rappelons l'énoncé du lemme de Grönwall :

Lemme V.4

Soit $T > 0$ et g une fonction positive mesurable bornée sur l'intervalle $[0, T]$. Supposons que

$$g(t) \leqslant a + b \int_0^t g(s)ds \ , \quad t \in [0, T] \ ,$$

pour deux constantes $a, b \geqslant 0$ données. Alors pour tout $t \in [0,T]$

$$g(t) \leqslant a\exp(bt)$$

La preuve du lemme procède par itération, nous ne la présentons pas. Mentionnons simplement qu'il est facile de retrouver le résultat de ce lemme. En effet, si g satisfait

$$g(t) = a + b\int_0^t g(s)ds\ , \quad t \in [0,T]\ , \tag{V.4}$$

alors la théorie des équations différentielles ordinaires assure que

$$g(t) = a\exp(bt)\ .$$

Si l'égalité dans (V.4) est remplacée par une inégalité, il est raisonnable d'obtenir également une inégalité dans la dernière équation.

Démonstration [unicité - Théorème V.3] Commençons par la preuve de l'unicité. Soient X et Y deux solutions. Pour $M > 0$ fixé on introduit le temps d'arrêt

$$\tau := \inf\{t \geqslant 0 : |X_t| \wedge |Y_t| \geqslant M\}\ .$$

On voit alors que

$$X_{t\wedge\tau} - Y_{t\wedge\tau} = \int_0^{t\wedge\tau}\Big(\sigma(s,X_s) - \sigma(s,Y_s)\Big)dB_s + \int_0^{t\wedge\tau}\Big(b(s,X_s) - b(s,Y_s)\Big)ds\ .$$

Fixons alors un horizon de temps $T > 0$. Pour tout $t \in [0,T]$, en utilisant l'inégalité $(x+y)^2 \leqslant 2x^2 + 2y^2$ on trouve

$$\mathbb{E}[|X_{t\wedge\tau} - Y_{t\wedge\tau}|^2] \leqslant 2\mathbb{E}\Big[\Big(\int_0^{t\wedge\tau}\Big(\sigma(s,X_s) - \sigma(s,Y_s)\Big)dB_s\Big)^2\Big] + 2\mathbb{E}\Big[\Big(\int_0^{t\wedge\tau}\Big(b(s,X_s) - b(s,Y_s)\Big)ds\Big)^2\Big]\ .$$

On note alors que $s \mapsto \mathbf{1}_{[0,t\wedge\tau]}(s)\Big(\sigma(s,X_s) - \sigma(s,Y_s)\Big)$ appartient à $\mathcal{H}^2$ de sorte que l'isométrie d'Itô assure que

$$\begin{aligned}
\mathbb{E}\Big[\Big(\int_0^{t\wedge\tau}\Big(\sigma(s,X_s) - \sigma(s,Y_s)\Big)dB_s\Big)^2\Big] &= \mathbb{E}\Big[\int_0^{t\wedge\tau}\Big(\sigma(s,X_s) - \sigma(s,Y_s)\Big)^2 ds\Big]\\
&\leqslant K^2\mathbb{E}\Big[\int_0^{t\wedge\tau}\Big|X_s - Y_s\Big|^2 ds\Big]\\
&\leqslant K^2\int_0^t \mathbb{E}\Big[\Big|X_{s\wedge\tau} - Y_{s\wedge\tau}\Big|^2\Big]ds\ .
\end{aligned}$$

Par ailleurs, en utilisant l'inégalité de Jensen on trouve

$$\begin{aligned}
\mathbb{E}\Big[\Big(\int_0^{t\wedge\tau}\Big(b(s,X_s) - b(s,Y_s)\Big)ds\Big)^2\Big] &= \mathbb{E}\Big[\Big(\int_0^{t\wedge\tau} T\Big(b(s,X_s) - b(s,Y_s)\Big)\frac{ds}{T}\Big)^2\Big]\\
&\leqslant \mathbb{E}\Big[T\int_0^{t\wedge\tau}\Big(b(s,X_s) - b(s,Y_s)\Big)^2 ds\Big]\\
&\leqslant TK^2\mathbb{E}\Big[\int_0^{t\wedge\tau}\Big|X_s - Y_s\Big|^2 ds\Big]\\
&\leqslant TK^2\int_0^t \mathbb{E}\Big[\Big|X_{s\wedge\tau} - Y_{s\wedge\tau}\Big|^2\Big]ds\ .
\end{aligned}$$

On pose alors

$$g(s) := \mathbb{E}\Big[\Big|X_{s\wedge\tau} - Y_{s\wedge\tau}\Big|^2\Big] ,$$

et l'on observe que

$$g(t) \leqslant 2K^2(1+T)\int_0^t g(s)ds .$$

Le Lemme V.4 assure alors que

$$g(t) \leqslant 0\exp(2K^2(1+T)t) = 0 , \quad t \in [0,T] .$$

La continuité des trajectoires de X et Y assure donc que presque sûrement pour tout $s \in [0, T \wedge \tau]$, $X_s = Y_s$. Or $\tau \to +\infty$ presque sûrement quand $M \to \infty$ et l'on peut conclure que presque sûrement pour tout $s \in [0,T]$, $X_s = Y_s$. □

Démonstration [existence - Théorème V.3] On applique la méthode des itérations de Picard. On pose

$$X_t^{(0)} := x , \quad t \geqslant 0 ,$$

puis par récurrence pour tout $n \geqslant 1$

$$X_t^{(n)} := x + \int_0^t \sigma(s, X_s^{(n)})dB_s + \int_0^t b(s, X_s^{(n)})ds . \tag{V.5}$$

On notera que chaque $X^{(n)}$ est continu et adapté à la filtration Brownienne et qu'ainsi chaque processus $s \mapsto \sigma(s, X_s^{(n)})$ appartient à $\mathcal{H}^2_{\text{loc}}$ de sorte que les intégrales stochastiques sont bien définies.
Fixons $T > 0$. En utilisant des arguments tout à fait similaires à ce qui a été développé pour l'uncité, on peut montrer qu'il existe pour tout $n \geqslant 0$ une constante $C_n \geqslant 0$ telle que

$$\mathbb{E}[(X_t^{(n)})^2] \leqslant C^{(n)} , \quad t \in [0,T] . \tag{V.6}$$

Majorons alors $\mathbb{E}[\sup_{0\leqslant s\leqslant t}(X_s^{(n+1)} - X_s^{(n)})^2]$. Par le même argument que dans la preuve de l'unicité, on obtient

$$\begin{aligned}\mathbb{E}[\sup_{0\leqslant s\leqslant t}(X_s^{(n+1)} - X_s^{(n)})^2] &\leqslant 2\mathbb{E}\Big[\sup_{0\leqslant s\leqslant t}\Big(\int_0^s \Big(\sigma(r, X_r^{(n)}) - \sigma(r, X_r^{(n-1)})\Big)dB_r\Big)^2\Big] \\ &\quad + 2\mathbb{E}\Big[\sup_{0\leqslant s\leqslant t}\Big(\int_0^s \Big(b(r, X_r^{(n)}) - b(r, X_r^{(n-1)})\Big)dr\Big)^2\Big] .\end{aligned}$$

La majoration (V.6) assure que les martingales locales sont de vraies martingales. L'inégalité de Doob puis l'isométrie d'Itô assurent alors que

$$\begin{aligned}\mathbb{E}\Big[\sup_{0\leqslant s\leqslant t}\Big(\int_0^s \Big(\sigma(r, X_r^{(n)}) - \sigma(r, X_r^{(n-1)})\Big)dB_r\Big)^2\Big] &\leqslant 4\mathbb{E}\Big[\Big(\int_0^t \Big(\sigma(r, X_r^{(n)}) - \sigma(r, X_r^{(n-1)})\Big)dB_r\Big)^2\Big] \\ &\leqslant 4\mathbb{E}\Big[\int_0^t \Big(\sigma(r, X_r^{(n)}) - \sigma(r, X_r^{(n-1)})\Big)^2 dr\Big] \\ &\leqslant 4K^2\mathbb{E}\Big[\int_0^t \Big|X_r^{(n)} - X_r^{(n-1)}\Big|^2 dr\Big] \\ &\leqslant 4K^2\mathbb{E}\Big[\int_0^t \sup_{0\leqslant r\leqslant s}\Big|X_r^{(n)} - X_r^{(n-1)}\Big|^2 ds\Big] .\end{aligned}$$

Par ailleurs l'inégalité de Jensen assure que

$$\mathbb{E}\Big[\sup_{0\leqslant s\leqslant t}\Big(\int_0^s\Big(b(r,X_r^{(n)})-b(r,X_r^{(n-1)})\Big)dr\Big)^2\Big]\leqslant \mathbb{E}\Big[\sup_{0\leqslant s\leqslant t} s\int_0^s\Big(b(r,X_r^{(n)})-b(r,X_r^{(n-1)})\Big)^2 dr\Big]$$
$$\leqslant T\mathbb{E}\Big[\int_0^t\Big(b(r,X_r^{(n)})-b(r,X_r^{(n-1)})\Big)^2 dr\Big]$$
$$\leqslant TK^2\mathbb{E}\Big[\int_0^t\Big(X_r^{(n)}-X_r^{(n-1)}\Big)^2 dr\Big]$$
$$\leqslant TK^2\mathbb{E}\Big[\int_0^t\sup_{0\leqslant r\leqslant s}\Big|X_r^{(n)}-X_r^{(n-1)}\Big|^2 ds\Big]\,.$$

En posant $C_T := 2K^2(4+T)$ et $g_n(t)=\mathbb{E}[\sup_{0\leqslant s\leqslant t}(X_s^{(n+1)}-X_s^{(n)})^2]$ on trouve ainsi que

$$g_n(t)\leqslant C_T\int_0^t g_{n-1}(s)ds\,.$$

On note que pour tout $t\in[0,T]$

$$g_1(t)\leqslant 2T(C^{(1)}+C^{(0)}) =: C'_T\,.$$

Ainsi

$$g_1(t)\leqslant C_TC'_Tt\,,\quad t\in[0,T]\,,$$

et par récurrence

$$g_n(t)\leqslant C_T^nC'_T\frac{t^n}{n!}\,,\quad t\in[0,T]\,.$$

Ceci assure que $\sum_{n\geqslant 0}g_n(T)^{1/2}<\infty$, et ainsi par le Théorème de Fubini

$$\mathbb{E}[\sum_{n\geqslant 0}\sup_{0\leqslant s\leqslant t}|X_s^{(n+1)}-X_s^{(n)}|]=\sum_{n\geqslant 0}\mathbb{E}[\sup_{0\leqslant s\leqslant t}|X_s^{(n+1)}-X_s^{(n)}|]$$
$$\leqslant\sum_{n\geqslant 0}\mathbb{E}[\sup_{0\leqslant s\leqslant t}|X_s^{(n+1)}-X_s^{(n)}|^2]^{1/2}=\sum_{n\geqslant 0}g_n(T)^{1/2}<\infty\,.$$

En conséquence presque sûrement

$$\sum_{n\geqslant 0}\sup_{0\leqslant s\leqslant t}|X_s^{(n+1)}-X_s^{(n)}|<\infty\,,$$

et ainsi presque sûrement, la suite de processus $(X_t^{(n)},t\in[0,T])_{n\geqslant 0}$ converge uniformément sur $[0,T]$ vers un processus $(X_t,t\in[0,T])$ qui est adapté (par rapport à la filtration canonique du Brownien) et continu.
Il reste à montrer que X satisfait l'équation voulue. Pour cela on va passer à la limite sur l'équation satisfaite par $X^{(n)}$. Pour tout $n_0\geqslant 0$, l'inégalité de Minkowski (sur la norme $L^2(\Omega,\mathcal{F},\mathbb{P})$) assure que

$$\mathbb{E}[\sup_{0\leqslant s\leqslant T}|X_s^{(n)}-X_s|^2]^{1/2}=\mathbb{E}\Big[\Big(\sum_{n=n_0}^{+\infty}\sup_{0\leqslant s\leqslant T}|X_s^{(n+1)}-X_s^{(n)}|\Big)^2\Big]^{1/2}$$
$$\leqslant\sum_{n=n_0}^{+\infty}\mathbb{E}\Big[\Big(\sup_{0\leqslant s\leqslant T}|X_s^{(n+1)}-X_s^{(n)}|\Big)^2\Big]^{1/2}$$
$$\leqslant\sum_{n=n_0}^{+\infty}g_{n+1}(T)^{1/2}\to 0\,,\quad n_0\to\infty\,.$$

On peut alors en déduire les convergences suivantes dans $L^2(\Omega, \mathcal{F}, \mathbb{P})$

$$\int_0^t \sigma(s, X_s) dB_s = \lim_{n\to\infty} \int_0^t \sigma(s, X_s^{(n)}) dB_s \,,$$
$$\int_0^t b(s, X_s) ds = \lim_{n\to\infty} \int_0^t b(s, X_s^{(n)}) ds \,,$$

En passant à la limite sur (V.5) on obtient l'égalité suivante dans L^2 pour tout $t \in [0, T]$

$$X_t := x + \int_0^t \sigma(s, X_s) dB_s + \int_0^t b(s, X_s) ds \,.$$

La continuité des trajectoires assure que l'égalité a lieu p.s. pour tous les $t \in [0, T]$ simultanément. Comme $T > 0$ est arbitraire, on peut conclure. □

Illustrons le résultat que l'on vient de prouver sur un exemple. Considérons l'EDS

$$dX_t = bX_t dt + \sigma X_t dB_t \,,$$

partant de $X_0 = x \in \mathbb{R}$. Ici $b \in \mathbb{R}$ et $\sigma > 0$. Les hypothèses du théorème sont vérifiées et l'on en déduit qu'il existe une unique solution forte à cette EDS.

La formule d'Itô permet de vérifier que la solution est donnée par

$$X_t = x \exp(\sigma B_t + (b - \frac{\sigma^2}{2})t) \,, \quad t \in \mathbb{R}_+ \,.$$

Dans le cas particulier où $b = 0$, on obtient la martingale exponentielle $x\mathcal{E}(\sigma B)_t$.

3 Solutions fortes - dimensions quelconque

Le résultat précédent reste vrai en dimension supérieure. On travaille dans $\mathbb{R}^d$ pour $d \geqslant 1$ fixé et l'on se donne un espace de probabilité sur lequel est défini un mouvement Brownien $B = (B^{(1)}, \ldots, B^{(k)})$ de dimension $k \geqslant 1$. (Notons que d et k ne sont pas nécessairement égaux). On note $M_{d,k}(\mathbb{R})$ l'espace des matrices réelles de taille $d \times k$.

On se donne deux fonctions continues $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ et $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \to M_{d,k}(\mathbb{R})$ et l'on considère l'équation vectorielle

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t \,, \quad t \geqslant 0 \,, \tag{V.7}$$

partant d'une condition initiale $X_0 \in \mathbb{R}^d$ donnée.

Définition V.5

On dit qu'un processus $(X_t, t \in \mathbb{R}_+)$ est solution forte de l'équation (V.3) si :

1. *X est adapté à la filtration du mouvement Brownien B,*
2. *presque sûrement pour tout $t \in \mathbb{R}_+$*

$$\int_0^t |b(s, X_s)| ds < \infty \,, \quad \int_0^t \sigma(s, X_s)^* \sigma(s, X_s) ds < \infty \,.$$

3. presque sûrement pour tout $t \in \mathbb{R}_+$

$$X_t = X_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s \ .$$

Théorème V.6 (*Existence et unicité de solutions fortes*)

On suppose qu'il existe une constante $K > 0$ *telle que pour tous* $t \in \mathbb{R}_+$ *et tous* $x, y \in \mathbb{R}^d$

$$|b_i(0,x)| + |\sigma_{i,j}(0,x)| \leqslant K \ , \quad |b_i(t,x) - b_i(t,y)| + |\sigma_{i,j}(t,x) - \sigma_{i,j}(t,y)| \leqslant K|x-y| \ .$$

Il existe une unique solution forte $(X_t, t \in \mathbb{R}_+)$ *à* (V.7).

La preuve repose sur les mêmes arguments que précédemment.

4 Solutions faibles

Peut-on s'affranchir des hypothèses utilisées dans les deux théorèmes précédents ? Plus particulièrement, est-il nécessaire que les coefficients b et σ soient Lipschitziens en espace ?

1) Un exemple d'EDS sans solution forte

Concentrons-nous sur une équation particulière pour laquelle l'hypothèse Lipschitziannité n'est pas vérifiée :

$$dX_t = \mathrm{sgn}(X_t)dB_t \ , \tag{V.8}$$

où sgn désigne la fonction “signe” définie par

$$\mathrm{sgn}(u) = \begin{cases} 1 & \text{si } u \geqslant 0 \\ -1 & \text{si } u < 0 \end{cases}$$

.
Tout d'abord, on observe que si $\mathcal{F}$ est une filtration à laquelle X est adapté et dans laquelle B est un mouvement Brownien, alors X est une martingale locale de crochet

$$\langle X, X\rangle_t = \int_0^t \mathrm{sgn}(X_s)^2 ds = t \ .$$

La caractérisation de Lévy, Théorème IV.5, assure alors que X est un mouvement Brownien.

Or $-X$ satisfait

$$-X_t = -\int_0^t \mathrm{sgn}(X_s)dB_s = \int_0^t \mathrm{sgn}(-X_s)dB_s \ .$$

(En effet, ces deux intégrales diffèrent là où X_s s'annulent cependant $\int_0^t \mathbf{1}_{X_s=0}dB_s$ est une martingale locale de crochet égal à $\int_0^t \mathbf{1}_{X_s=0}ds$ et cette quantité est nulle pour un mouvement Brownien). Ainsi $-X$ satisfait également l'EDS de sorte qu'il ne peut y avoir d'unicité forte pour l'équation.

De façon intuitive, le défaut d'unicité de l'EDS vient de la liberté que l'on a, au moment où la solution touche 0, de la faire partir vers le haut ou vers le bas. Si l'on décomposait la trajectoire de X en excursions hors de 0, on pourrait obtenir une autre solution en changeant le signe d'une partie des excursions.

En fait, on peut même montrer qu'il n'y a pas existence forte. Partons d'un mouvement Brownien W et définissons B à partir de W de la manière suivante :

$$B_t := \int_0^t \operatorname{sgn}(W_s) dW_s \ .$$

Il s'agit d'une martingale locale de crochet donné par

$$\langle B, B\rangle_t = \int_0^t \operatorname{sgn}(W_s)^2 ds = t \ .$$

La caractérisation de Lévy, Théorème IV.5, assure alors que B est un mouvement Brownien. Or

$$\int_0^t \operatorname{sgn}(W_s) dB_s = \int_0^t \operatorname{sgn}(W_s)^2 dW_s = \int_0^t dW_s = W_t \ ,$$

de sorte que W est solution de l'EDS (V.8).
Il est alors possible de vérifier que B est adapté à filtration du processus $|W|$: or cette filtration contient strictement moins d'information que celle du processus W de sorte que W *n'est pas* adapté à la filtration de B. Pour vérifier que B est adapté à filtration du processus $|W|$, on peut appliquer la formule d'Itô avec $\varphi_\epsilon(x) = \sqrt{\epsilon + x^2}$:

$$\varphi_\epsilon(W_t) = \epsilon + \int_0^t \varphi_\epsilon'(W_s) dW_s + \frac{1}{2} \int_0^t \varphi_\epsilon''(W_s) ds \ .$$

Par parité des fonctions φ_ϵ et φ_ϵ'', les processus $\varphi_\epsilon(W_t)$ et $\frac{1}{2}\int_0^t \varphi_\epsilon''(W_s)ds$ sont $\mathcal{F}^{|W|}$-adaptés. Nécessairement le processus $\int_0^t \varphi_\epsilon'(W_s)dW_s$ l'est aussi. Par ailleurs on peut montrer que $\int_0^t \varphi_\epsilon'(W_s)dW_s$ converge dans L^2 vers

$$\int_0^t \operatorname{sgn}(W_s) dW_s = B_t \ ,$$

et ainsi la limite est aussi $\mathcal{F}^{|W|}$-adaptée.

2) Un résultat d'existence utilisant le théorème de Girsanov

Il est possible de construire des solutions dites faibles à l'aide du théorème de Girsanov. Considérons l'EDS

$$dX_t = b(t, X_t)dt + dB_t \ , \quad t \in \mathbb{R}_+ \ ,$$

partant de $X_0 = x \in \mathbb{R}$. On suppose que $b : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ est une fonction borélienne bornée. On va construire un espace de probabilité filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{Q})$ sur lequel B est un mouvement Brownien et X est un processus $\mathcal{F}_t$-adapté et satisfait l'équation précédente : c'est ce que l'on appelle une solution faible de l'EDS. Attention : rien n'assurera que X est adapté à la filtration du Brownien (ça n'est pas forcément une solution forte).

On part d'un mouvement Brownien X sur un espace de probabilité filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P})$. On fixe $T > 0$ et l'on définit le changement de mesure

$$\mathbb{Q}(A) := \mathbb{E}^{\mathbb{P}}[\mathbf{1}_A \mathcal{E}(L)_T] , \quad A \in \mathcal{F}_T ,$$

où

$$L_t := \exp(\int_0^t b(s, X_s) dX_s - \frac{1}{2} \int_0^t |b(s, X_s)|^2 ds) , \quad t \geqslant 0 .$$

(On notera que l'intégrale stochastique fait sens car le processus $(b(s, X_s), s \in [0, t])$ est progressivement mesurable et borné). Le Théorème IV.9 assure alors que sous $\mathbb{Q}$

$$B_t := X_t - \langle X, L \rangle_t = X_t - \int_0^t b(s, X_s) ds ,$$

est un mouvement Brownien.

Chapitre VI

Processus de Markov et EDP

Approx. 1 séance de cours

1 Processus de Markov

Commençons par introduire la notion de processus de Markov. Informellement, on dit qu'un processus stochastique X est de Markov si son évolution future ne dépend que de l'état présent du processus et non de l'ensemble de son passé. Pour en donner une définition plus rigoureuse il nous faut introduire un peu de vocabulaire.

Définition VI.1

Une famille $(p_t(x,\cdot))_{t\geqslant 0, x\in\mathbb{R}^d}$ de mesures de probabilités sur $\mathbb{R}^d$ est appelée noyau de transition *si*

1. *$p_0(x,\cdot) = \delta_x$ pour tout $x \in \mathbb{R}^d$,*
2. *pour tout Borélien $A \in \mathcal{B}(\mathbb{R}^d)$, l'application $(t,x) \mapsto p_t(x,A)$ est mesurable,*
3. *pour tout Borélien $A \in \mathcal{B}(\mathbb{R}^d)$ et toute paire $0 \leqslant s \leqslant t$, la relation dite de Chapman-Kolmogorov est satisfaite*

$$p_t(x,A) = \int_{\mathbb{R}^d} p_s(x,dy) p_{t-s}(y,A) \ , \quad \forall x \in \mathbb{R}^d \ .$$

Etant donné un noyau de transition $(p_t(x,\cdot))_{t\geqslant 0, x\in\mathbb{R}^d}$, nous pouvons introduire un semigroupe d'opérateurs $(P_t, t \geqslant 0)$ en posant

$$P_t f(x) := \int_{\mathbb{R}^d} p_t(x,dy) f(y) \ , \quad \forall x \in \mathbb{R}^d \ , \quad \forall t \geqslant 0 \ ,$$

pour toute fonction mesurable bornée $f : \mathbb{R}^d \to \mathbb{R}$.

Remarque VI.2

Il s'agit en fait de la définition d'un noyau de transition homogène en temps. *Il est possible de donner une définition plus générale.*

La relation de Chapman-Kolmogorov implique la propriété de *semigroupe* suivante :

$$P_t = P_{t-s} \circ P_s \ , \quad \forall 0 \leqslant s \leqslant t \ .$$

Nous pouvons à présent définir les processus de Markov.

Définition VI.3

Un processus stochastique X défini sur $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P})$ et à valeurs dans $\mathbb{R}^d$ est dit Markovien *s'il existe un noyau de transtion $(p_t(x,\cdot))_{t\geqslant 0, x\in\mathbb{R}^d}$ tel que pour toute fonction $f : \mathbb{R}^d \to \mathbb{R}$ mesurable bornée presque sûrement*

$$\mathbb{E}[f(X_t) \mid \mathcal{F}_s] = P_{t-s}f(X_s) .$$

Examinons cette dernière identité. Le terme de gauche peut s'interpréter comme : la meilleure approximation de $f(X_t)$ sachant toute la trajectoire jusqu'au temps s. Le terme de droite, lui, ne dépend que de la valeur de la trajectoire au temps s et, à l'aide du semigroupe, *propage* cette quantité vers le temps t.

Proposition VI.4

Le mouvement Brownien est un processus de Markov dont le noyau de transition est donné par le noyau gaussien

$$p_t(x,dy) = \frac{1}{(2\pi t)^{d/2}} \exp(-\frac{|x-y|^2}{2t})dy .$$

Démonstration Il n'est pas difficile de vérifier les deux premières propriétés que doit vérifier un noyau de transition. Concernant la troisième, elle peut se prouver comme suit. L'application

$$A \mapsto \int_{\mathbb{R}^d} p_s(x,dy)p_{t-s}(y,A) ,$$

est la convolée de deux lois de probabilités sur $\mathbb{R}^d$: $\mathcal{N}(x, sId)$ et $\mathcal{N}(0,(t-s)Id)$. Il est bien connu que cette convolée n'est rien d'autre que $\mathcal{N}(x,tId)$ qui n'est rien d'autre que $p_t(x,\cdot)$.
Soit B un mouvement Brownien et $(\mathcal{F}_t)_{t\geqslant 0}$ sa filtration naturelle. Calculons à présent :

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

Or $B_t - B_s$ est indépendant de $\mathcal{F}_s$ par la Proposition I.18 de sorte que les propriétés de l'espérance conditionnelle assurent que si l'on pose pour tout $z \in \mathbb{R}^d$

$$\Phi(z) := \mathbb{E}[f(z + B_t - B_s)] ,$$

alors presque sûrement

$$\mathbb{E}[f(B_s + (B_t - B_s)) \mid \mathcal{F}_s] = \Phi(B_s) .$$

Or

$$\Phi(z) := \mathbb{E}[f(z + B_t - B_s)] = \int_{\mathbb{R}^d} f(z+y)p_{t-s}(0,dy) = \int_{\mathbb{R}^d} f(y)p_{t-s}(z,dy) = P_{t-s}f(z) ,$$

ce qui conclut la preuve □

Cette propriété s'étend à une classe de solutions fortes d'EDS introduites dans le chapitre précédent. On se donne deux fonctions $b : \mathbb{R}^d \to \mathbb{R}^d$ et $\sigma : \mathbb{R}^d \to \mathcal{M}_{d,k}$, et l'on suppose qu'il existe une constante $K > 0$ telle que pour tous $t \in \mathbb{R}_+$ et tous $x, y \in \mathbb{R}^d$

$$|b_i(0)| + |\sigma_{i,j}(0)| \leqslant K , \quad |b_i(x) - b_i(y)| + |\sigma_{i,j}(x) - \sigma_{i,j}(y)| \leqslant K|x-y| .$$

Proposition VI.5

On se place sous les hypothèses du Théorème V.6 en supposant que les coefficients ne dépendent pas du temps, et l'on note X_t^x la solution de l'EDS correspondante partant de $x \in \mathbb{R}^d$. On pose pour tout $x \in \mathbb{R}^d$ et tout $t \geqslant 0$

$$p_t(x, A) := \mathbb{P}(X_t \in A) , \quad A \in \mathcal{B}(\mathbb{R}^d) .$$

Alors pour tout $x \in \mathbb{R}^d$, le processus $(X_t^x, t \geqslant 0)$ est un processus de Markov de noyau de transition $(p_t(y, \cdot))_{t \geqslant 0, y \in \mathbb{R}^d}$.

Ce résultat découle essentiellement de l'unicité des solutions fortes des EDS considérées qui assure le fait suivant : la solution au temps t partant de x coïncide avec la solution au temps $t - s$ partant de la valeur (aléatoire !) de la solution au temps s partant de x.
Il y a cependant des détails pénibles à vérifier, et nous ne présentons donc pas la preuve.

On se place toujours sous les hypothèses du Théorème V.6 en supposant que les coefficients ne dépendent pas du temps.

Théorème VI.6 (*Propriété de Markov forte*)

Soit $x \in \mathbb{R}^d$. Si T est un temps d'arrêt fini p.s. alors pour toute fonction $F : C([0, \infty[, \mathbb{R}) \to \mathbb{R}$ mesurable bornée on a presque sûrement

$$\mathbb{E}[F(X_{T+\cdot}^x) \mid \mathcal{F}_T] = \Phi(X_T^x) ,$$

où

$$\Phi(y) = \mathbb{E}[F(X_\cdot^y)] , \quad y \in \mathbb{R}^d .$$

La fonction Φ associe à un point de départ y l'espérance de F évalué en la trajectoire $X_\cdot^y$. La quantité $\Phi(X_T^x)$ est une variable aléatoire qui, à ω donné, prend la valeur $\Phi(y)$ où $y = X_T^x(\omega)$.

Il convient de comparer cet énoncé avec la propriété de Markov forte vérifiée par le mouvement Brownien. On notera cependant que dans le cas des solutions fortes d'EDS, il n'est pas vrai en général que le processus

$$(X_{T+t}^x - X_T^x, t \geqslant 0) ,$$

est indépendant de $\mathcal{F}_T$ et distribué comme $(X_t^0, t \geqslant 0)$!

2 Liens avec des EDP linéaires

Dans toute cette partie, on se donne deux fonctions continues $b : \mathbb{R}^d \to \mathbb{R}^d$ et $\sigma : \mathbb{R}^d \to \mathcal{M}_{d,k}$, et l'on suppose qu'il existe une constante $K > 0$ telle que pour tous $t \in \mathbb{R}_+$ et tous $x, y \in \mathbb{R}^d$

$$|b_i(x)| + |\sigma_{i,j}(x)| \leqslant K , \quad |b_i(x) - b_i(t, y)| + |\sigma_{i,j}(x) - \sigma_{i,j}(y)| \leqslant K|x - y| .$$

(On notera qu'on a ici supposé que les fonctions σ et b étaient bornées par K en tout point). Le théorème V.6 s'applique sous ces hypothèses et fournit un processus de Markov $(X_t^x, t \geqslant 0)$ pour tout point de départ $x \in \mathbb{R}^d$.

Cette partie esquisse des liens profonds entre d'un côté certains opérateurs différentiels et certaines EDP (équations aux dérivées partielles), et d'un autre côté certains processus de diffusion. Nous recommandons de ne pas s'attarder sur les hypothèses ou les détails techniques, mais plutôt d'apprécier les liens à un niveau formel.

1) Générateur infinitésimal

Théorème VI.7

On introduit l'opérateur différentiel suivant :

$$\mathcal{L}f(x) := \sum_{i=1}^{d} b_i(x)\frac{\partial f}{\partial x_i}(x) + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}(\sigma\sigma^*)_{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x) , \quad x \in \mathbb{R}^d ,$$

agissant sur les fonctions $f : \mathbb{R}^d \to \mathbb{R}$ de classe $\mathcal{C}_c^\infty$. Pour toute fonction $f : \mathbb{R}^d \to \mathbb{R}$ de classe $\mathcal{C}_c^\infty$, le processus

$$f(X_t^x) - f(x) - \int_0^t \mathcal{L}f(X_s^x)ds , \quad t \geqslant 0 ,$$

est une martingale.

Démonstration La formule d'Itô assure que

$$f(X_t) = f(x) + \sum_{i=1}^{d}\int_0^t \partial_{x_i} f(X_s)dX_s^{(i)} + \frac{1}{2}\sum_{i,j=1}^{d}\int_0^t \partial^2_{x_i,x_j} f(s,X_s)d\langle X^{(i)},X^{(j)}\rangle_s .$$

Or

$$dX_s^{(i)} = b_i(X_s)ds + \sum_{\ell=1}^{k}\sigma_{i,\ell}(X_s)dB_s^{(\ell)} ,$$

et

$$d\langle X^{(i)},X^{(j)}\rangle_s = \sum_{\ell=1}^{k}\sigma_{i,\ell}(X_s)\sigma_{j,\ell}(X_s) = (\sigma\sigma^*(X_s))_{ij} .$$

On introduit le processus $M_s \in \mathbb{R}$ en posant

$$M_s := \sum_{i=1}^{d}\sum_{\ell=1}^{k}\int_0^s \partial_{x_i} f(X_r)\sigma_{i,\ell}(X_r)dB_r^{(\ell)} .$$

Il s'agit d'une martingale car l'integrand est continu et borné. On obtient alors

$$\begin{aligned} f(X_t) &= f(x) + \sum_{i=1}^{d}\int_0^t \partial_{x_i} f(X_s)b_i(X_s)ds + \frac{1}{2}\sum_{i,j=1}^{d}\int_0^t \partial^2_{x_i,x_j} f(s,X_s)(\sigma\sigma^*(X_s))_{ij} + M_t \\ &= f(x) + \int_0^t \mathcal{L}f(X_s)ds + M_t . \end{aligned}$$

Le résultat s'en suit. □

Ce résultat permet de relier le semigroupe associé au processus de Markov X_t et l'opérateur différentiel ci-dessus. En effet, si l'on calcule un incrément de la fonction $t \mapsto \mathbb{E}[f(X_t)]$ alors on trouve

$$\mathbb{E}[f(X_{t+s})] - \mathbb{E}[f(X_t)] = \mathbb{E}[\int_t^{t+s} \mathcal{L}f(X_r)dr] \ ,$$

(car le terme de martingale est d'espérance nulle). En passant à la limite $s \downarrow 0$ on obtient alors (sans se soucier des justifications mathématiques)

$$\partial_t \mathbb{E}[f(X_t)] = \mathbb{E}[\mathcal{L}f(X_t)] \ .$$

Or $\mathbb{E}[f(X_t)] = P_t f(x)$ ainsi

$$\partial_t P_t f(x) = P_t \mathcal{L} f(x) \ .$$

L'opérateur $\mathcal{L}$ est appelé *générateur infinitésimal* du semigroupe : il s'agit d'un opérateur qui décrit l'évolution infinitésimale du semigroupe (et le caractérise !). Nous ne donnons aucun détail rigoureux car il s'agit-là d'une théorie mathématique délicate.

2) Equation de Fokker-Planck

On note X_t^x le processus de Markov solution de (V.7) partant de x. On peut souhaiter obtenir une équation d'évolution pour sa loi, en d'autres termes, une équation d'évolution pour $t \mapsto p_t(x,\cdot)$. On introduit l'opérateur différentiel

$$\mathcal{L}^* f = -\sum_{i=1}^d \frac{\partial}{\partial x_i}(b_i f)(x) + \frac{1}{2}\sum_{i=1}^d\sum_{j=1}^d \frac{\partial^2}{\partial x_i \partial x_j}(f(\sigma\sigma^*)_{ij})(x) \ , \quad x \in \mathbb{R}^d \ .$$

Il s'agit de l'adjoint de l'opérateur $\mathcal{L}$ (par rapport à la mesure de Lebesgue). Plus précisément, pour toutes fonctions $f, g \in C_c^\infty$

$$\int_x f(x)\mathcal{L}g(x)dx = \int_x \mathcal{L}^* f(x) g(x) dx \ .$$

Alors $t \mapsto p_t(x,\cdot)$ est solution de l'équation d'évolution dite de Fokker-Planck

$$\partial_t p_t(x,\cdot) = \mathcal{L}^* p_t(x,\cdot) \ ,$$

partant de $p_0(x,\cdot) = \delta_x(\cdot)$.

En effet, pour toutes fonctions $f, g \in \mathcal{C}_c^\infty$

$$\partial_t \int_x f(x) P_t g(x) dx = \int_x f(x) P_t \mathcal{L} g(x) dx = \int_x\int_y f(x) p_t(x,dy)\mathcal{L}g(y)dx = \int_x f(x)\Big(\int_y \mathcal{L}^* p_t(x,dy) g(y)\Big)dx \ .$$

Or

$$\partial_t \int_x f(x) P_t g(x) dx = \partial_t \int_x\int_y f(x) p_t(x,dy) g(y) dx = \int_x f(x)\partial_t\Big(\int_y p_t(x, dyg(y)\Big)dx \ .$$

On a donc obtenu l'égalité

$$\partial_t p_t(x,\cdot) = \mathcal{L}^* p_t(x,\cdot) \ ,$$

au sens des distributions.

Ajoutons une hypothèse supplémentaire dite d'uniforme ellipticité : il existe $c > 0$ tel que pour tout vecteur $\lambda \in \mathbb{R}^d$

$$\sum_{i,j=1}^{d} (\sigma\sigma^*)_{ij}\lambda_i\lambda_j \geqslant c\|\lambda\|^2 .$$

Dans ce cas, il est possible de montrer que la loi $p_t(x, \cdot)$ admet une densité :

$$p_t(x, dy) = p_t(x, y)dy .$$

L'équation de Fokker-Planck s'écrit alors

$$\partial_t p_t(x, y) = \mathcal{L}^* p_t(x, y) .$$

3) EDP parabolique

On se place sous les mêmes hypothèses que dans la partie précédente et l'on note $\mathcal{L}$ l'opérateur différentiel associé. Soit $u : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ une fonction continue bornée telle que

$$u(\cdot, x) \in C^1(]0, +\infty[) , \quad \forall x \in \mathbb{R}^d ,$$

et

$$u(t, \cdot) \in C_b^2(\mathbb{R}^d) , \quad \forall t \in]0, +\infty[.$$

Supposons que u satisfait l'EDP parabolique suivante

$$\begin{cases} \frac{\partial u}{\partial t} = \mathcal{L}u , \quad t > 0 \\ u(0, x) \qquad\qquad = f(x) , \quad x \in \mathbb{R}^d , \end{cases} \tag{VI.1}$$

pour une certaine fonction $f \in C_b(\mathbb{R}^d)$.

Théorème VI.8

Pour tout $x \in \mathbb{R}^d$

$$u(t, x) = \mathbb{E}[f(X_t^x)] , \quad t \geqslant 0 ,$$

où X_t^x est la solution de l'EDS (V.7) partant de x.

Ce résultat assure que l'existence d'une solution à l'EDP parabolique (VI.1) implique l'unicité de la solution.

Démonstration Une application de la formule d'Itô au processus $(u(t - s, X_s^x), s \in [0, t[)$ assure que pour tout $s \in [0, t[$

$$\mathbb{E}[u(t - s, X_s^x)] = \mathbb{E}[u(t, X_0^x)] = u(t, x) .$$

Or la fonction u est continue bornée donc le théorème de convergence dominée assure que lorsque $s \uparrow t$

$$\mathbb{E}[u(t - s, X_s^x)] \to \mathbb{E}[u(0, X_t^x)] = \mathbb{E}[f(X_t^x)] ,$$

et le résultat s'en suit. On notera que l'on n'a pas appliqué la formule d'Itô au temps t mais à un temps s antérieur à t : en effet, la fonction f est de classe $C^{1,2}$ sur $]0, +\infty[\times\mathbb{R}^d$ mais rien n'assure qu'elle le soit aussi au temps 0. $\square$

4) Formule de Feynman-Kac

Soit $V : \mathbb{R}^d \to \mathbb{R}$ une fonction continue et bornée inférieurement. Soit $u : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ une fonction continue bornée telle que

$$u(\cdot, x) \in C^1(]0, +\infty[) \ , \quad \forall x \in \mathbb{R}^d \ ,$$

et

$$u(t, \cdot) \in C_b^2(\mathbb{R}^d) \ , \quad \forall t \in]0, +\infty[\ .$$

On suppose que u satisfait l'EDP

$$\begin{cases} \frac{\partial u}{\partial t} = \mathcal{L}u - Vu \ , \quad t > 0 \\ u(0, x) \qquad\qquad\qquad = f(x) \ , \quad x \in \mathbb{R}^d \ , \end{cases} \tag{VI.2}$$

pour une certaine fonction $f \in C_b(\mathbb{R}^d)$.

Théorème VI.9

Pour tout $x \in \mathbb{R}^d$

$$u(t, x) = \mathbb{E}[f(X_t^x) \exp(-\int_0^t V(X_s^x) ds)] \ , \quad t \geqslant 0 \ .$$

Démonstration Appliquer Itô au processus $(e^{-\int_0^s V(X_u) du} u(t - s, X_s), s \in [0, t[)$. □

Annexe A

Généralités

1 Variables aléatoires

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité. Soit $(E, \mathcal{E})$ un espace mesurable.

Définition A.1

Une application $X : \Omega \to E$ mesurable par rapport aux tribus $\mathcal{F}$ et $\mathcal{E}$ est appelée variable aléatoire.

On rappelle que X est mesurable par rapport aux tribus $\mathcal{F}$ et $\mathcal{E}$ si pour tout $B \in \mathcal{E}$, $X^{-1}(B) \in \mathcal{F}$.

Définition A.2

Soit X un variable aléatoire de $(\Omega, \mathcal{F}, \mathbb{P})$ dans $(E, \mathcal{E})$. L'application de $\mu_X : \mathcal{E} \to [0,1]$ définie par

$$\mu_X(B) := \mathbb{P}(X^{-1}(B)) , \quad B \in \mathcal{E} ,$$

est une mesure de probabilité sur $(E, \mathcal{E})$ que l'on appelle loi de X.

On notera que la loi de X n'est rien d'autre que l'image par X de la mesure de probabilité $\mathbb{P}$.
Quelques exemples :

1. Jet de dés : $\Omega = \{1, \ldots, 6\}^n$ muni de la tribu $\mathcal{F}$ de toutes les parties de Ω et de la mesure uniforme $\mathbb{P}$. L'application $X_k : \omega \mapsto \omega_k$ qui retourne la k-ème coordonnée est une variable aléatoire à valeurs dans $\{1, \ldots, 6\}$ muni de la tribu de toutes ses parties. Sa loi est uniforme. L'application $S : \Omega \to \sum_{k=1}^n X_k$ est une variable aléatoire à valeurs dans $\{n, \ldots, 6n\}$. Sa loi peut être explicitée.
2. Variable gaussienne : $\Omega = E = \mathbb{R}$, $\mathcal{F} = \mathcal{E} = \mathcal{B}(\mathbb{R})$ (la tribu Borélienne) et $\mathbb{P}$ la mesure de probabilté
$$\mathbb{P}(dx) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx .$$
3. Variable constante égale à $a \in E$: $\mu_X = \delta_a$ où l'on définit la masse de Dirac $\delta_a(B) = \mathbf{1}_B(a)$ pour tout $B \in \mathcal{E}$.

Proposition A.3 (*Formule de transfert*)

Soit X une variable aléatoire à valeurs dans un espace $(E, \mathcal{E})$ et $h : E \to \mathbb{R}$ une fonction $(\mathcal{E}, \mathcal{B}(\mathbb{R}))$-

mesurable. Si $h(X)$ est intégrable alors

$$\mathbb{E}[h(X)] = \int_E h \, d\mu_X \, .$$

2 Caractérisation d'une loi

Proposition A.4

Soit E un espace métrique et $\mathcal{E}$ la tribu Borélienne associée (c'est-à-dire, la plus petite tribu contenant les ouverts de E). Soient X et Y deux variables aléatoires à valeurs $(E, \mathcal{E})$ (pas forcément définies sur un même espace de probabilité). On a équivalence entre :

1. $\mu_X = \mu_Y$,
2. $\mu_X(O) = \mu_X(O)$ *pour tout O ouvert borné de E,*
3. $\int_E f d\mu_X = \int_E f d\mu_Y$ *pour toute fonction continue bornée $f : E \to \mathbb{R}$.*

Démonstration 1.⇒3. est clair.
3.⇒2. : Soit O un ouvert borné de E. Soit $f_n : E \to \mathbb{R}$ définie par

$$f_n(x) := (1 - \mathrm{dist}(x, O)/n)_+ \, , \quad x \in E \, .$$

Il s'agit d'une fonction continue bornée. Par ailleurs f_n converge ponctuellement vers l'indicatrice de O quand $n \to \infty$. Ainsi le théorème de convergence dominée assure que

$$\mu_X(O) = \lim_n \int_E f_n d\mu_X \, , \quad \mu_Y(O) = \lim_n \int_E f_n d\mu_Y \, ,$$

et l'on peut conclure.
2.⇒1. : tout d'abord, l'égalité s'étend à tous les ouverts. En effet, si O est un ouvert de E, alors $O_n := O \cap B(x, n)$ est un ouvert borné (ici $B(x, n)$ désigne la boule ouvert de centre x et de rayon n), $O_n \uparrow O$ et l'on peut utiliser la continuité croissante des mesures de probabilité.
La classe

$$\mathcal{M} := \{A \in \mathcal{E} : \mu_X(A) = \mu_Y(A)\} \, ,$$

est une classe monotone (càd contient E, est stable par différence et par union croissante). Elle contient la classe des ouverts. Le lemme de classe monotone assure qu'elle contient la tribu engendrée par cette classe, c'est-à-dire, la tribu Borélienne sur E. □

Dans la suite, on se donne une variable aléatoire X à valeurs dans $\mathbb{R}^d$ (pour un entier $d \geqslant 1$) muni de sa tribu Borélienne.

1) Fonction caractéristique

L'application

$$\begin{aligned} \Phi_X : &\mathbb{R}^d \to \mathbb{R} \\ &t \mapsto \mathbb{E}[e^{i\langle t, X\rangle}] \, , \end{aligned}$$

est appelée fonction caractéristique de la variable aléatoire X. La formule de transfert montre que la fonction caractéristique est complétement déterminée par la loi de X. Il se trouve que la réciproque est vraie :

Théorème A.5

La loi de X est entièrement caractérisée par sa fonction caractéristique.

Démonstration [Idées de la preuve] On introduit une variable aléatoire Z_σ, indépendante de X et de loi $\mathcal{N}(0, \sigma Id)$. On montre alors que :

1. la loi de $X + Z_\sigma$ est complétement caractérisée par Φ_X. Plus précisément, on montre que la loi de $X + Z_\sigma$ admet une densité f_σ qui admet une expression explicite en terme de Φ_X.
2. la v.a. $X + Z_\sigma$ converge en loi vers X quand $\sigma \to 0$. Nous rappellerons plus loin cette notion de convergence, mais on peut dores et déjà noter que cette convergence est très intuitive car Z_σ converge (en loi) vers 0 quand $\sigma \to 0$.

□

Remarque A.6

On différenciera bien la v.a. X de sa loi : deux v.a. ayant même fonction caractéristique ont même loi, mais n'ont aucune raison d'être égales (ni même d'être définies sur le même espace de probabilité !).

2) Fonction de répartition

On suppose que $d = 1$.

Définition A.7

La fonction $F_X : \mathbb{R} \to [0, 1]$ définie par

$$F(t) := \mathbb{P}(X \leqslant t) , \quad t \in \mathbb{R} ,$$

est appelée la fonction de répartition de X.

Comme les ensembles $(-\infty, t]$, $t \in \mathbb{R}$, engendrent la tribu Borélienne, le lemme de classe monotone permet de déduire que F_X caractérise la loi μ_X.

Proposition A.8

L'ensemble des fonctions de répartitions est exactement donné par l'ensemble des fonctions $F : \mathbb{R} \to [0, 1]$, continues à droite, croissantes, qui convergent vers 1 en $+\infty$ et 0 en $-\infty$.

3 Variables gaussiennes

La loi gaussienne uni-dimensionnelle de moyenne $m \in \mathbb{R}$ et de variance $\sigma^2 > 0$ est la mesure de probabilité sur $\mathbb{R}$ de densité

$$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{|x-m|^2}{2\sigma^2}} , \quad x \in \mathbb{R} .$$

Dans le cas $\sigma^2 = 0$, il s'agit de la mesure de probabilité sur $\mathbb{R}$ donnée par la masse de Dirac en m :

$$\delta_m(B) = \mathbf{1}_B(m) , \quad B \in \mathcal{B}(\mathbb{R}) .$$

Définition A.9

On dit qu'une variable aléatoire à valeurs dans $\mathbb{R}^d$ pour $d \geqslant 1$ est gaussienne si toute combinaison linéaire de ses coordonnées est une variable aléatoire gaussienne (uni-dimensionnelle).

Comme pour tout vecteur aléatoire à valeurs dans $\mathbb{R}^d$ (dont la norme euclidienne admet un moment d'ordre 2) on peut définir le vecteur moyenne m et la matrice de covariance Σ d'une v.a. gaussienne X

$$m_k := \mathbb{E}[X_k] \;, \quad \Sigma(k,\ell) = \mathrm{Cov}(X_k, X_\ell) \;.$$

On notera bien que m est un élément de $\mathbb{R}^d$ alors que Σ est une matrice symétrique positive de taille $d \times d$.

Proposition A.10

La fonction caractéristique d'une variable aléatoire gaussienne X est donnée par

$$\Phi_X(t) = \exp(i\langle t, m\rangle - \frac{\langle t, \Sigma t\rangle}{2}) \;, \quad t \in \mathbb{R}^d \;.$$

En conséquence, la loi d'une variable aléatoire gaussienne est entièrement déterminée par son vecteur moyenne et sa matrice de covariance.

Soit $Y = (Y_1, \ldots, Y_d)$ un vecteur aléatoire dont les coordonnées sont des v.a. gaussiennes unidimensionnelles, indépendantes centrées réduites. Il est facile de vérifier que Y est une variable aléatoire gaussienne de moyenne le vecteur nul et de matrice de covariance la matrice identité.

Etant donné un vecteur $m \in \mathbb{R}^n$ et une matrice $n \times n$ symétrique positive Σ, la variable aléatoire $X := m + \Sigma^{1/2} Y$ est gaussienne de moyenne m et de matrice de covariance Σ. (La matrice $\Sigma^{1/2}$ peut être définie comme suit. Soit P une matrice orthogonale telle que $D = P\Sigma P^t$ est diagonale, on définit alors $D^{1/2}$ comme la matrice diagonale dont les entrées sont les racines carrés des entrées de D (celles-ci sont positives car Σ est positive). On pose alors $\Sigma^{1/2} := P^t D^{1/2} P$.) On note $\mathcal{N}(m, \Sigma)$ la loi correspondante.

Proposition A.11

Soit $X \sim \mathcal{N}(m, \Sigma)$. X admet une densité sur $\mathbb{R}^d$ si et seulement si Σ est définie positive, et dans ce cas celle-ci est donnée par

$$\frac{1}{(2\pi)^{d/2}\sqrt{\det(\Sigma)}} \exp\Big(-\frac{\langle \Sigma^{-1}(x-m), x-m\rangle}{2}\Big) \;, \quad x \in \mathbb{R}^d \;.$$

4 Espérance conditionnelle

Dans cette partie, on manipule des variables aléatoires réelles.

Définition A.12 (*Espérance conditionnelle dans L^1*)

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité, $\mathcal{B} \subseteq \mathcal{F}$ une sous-tribu et $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. On définit l'espérance conditionnelle de X sachant $\mathcal{B}$, que l'on note $\mathbb{E}[X \mid \mathcal{B}]$, comme l'unique variable aléatoire

$U \in L^1(\Omega, \mathcal{B}, \mathbb{P})$ qui satisfait

$$\mathbb{E}[XZ] = \mathbb{E}[UZ],$$

pour toute variable aléatoire bornée Z qui est $\mathcal{B}$-mesurable.

On définit alors l'espérance conditionnelle d'une variable aléatoire X sachant une autre variable aléatoire Y comme l'espérance conditionnelle de X sachant $\sigma(Y)$, que l'on notera souvent $\mathbb{E}[X \mid Y]$. Citons un cas particulier. Supposons que Y prenne ses valeurs dans un ensemble dénombrable, c'est-à-dire, on suppose qu'il existe un ensemble dénombrble $\mathcal{Y}$ tel que $\mathbb{P}(Y = y) > 0$ pour tout $y \in \mathcal{Y}$ et $\sum_{y \in \mathcal{Y}} \mathbb{P}(Y = y) = 1$. Alors

$$\mathbb{E}[X \mid Y] = \sum_{y \in \mathcal{Y}} \mathbb{E}[X \mid \{Y = y\}] \mathbf{1}_{\{Y=y\}} .$$

Proposition A.13 (*Propriétés de l'espérance conditionnelle*)

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité, $\mathcal{B} \subseteq \mathcal{F}$ une sous-tribu et $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. Les propriétés suivantes sont vérifiées :

(a) $\forall a, b \in \mathbb{R}$ et pour toute v.a. $Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, on a $\mathbb{E}[aX + bY \mid \mathcal{B}] = a\mathbb{E}[X \mid \mathcal{B}] + b\mathbb{E}[Y \mid \mathcal{B}]$.

(b) Si $X \geqslant 0$, alors $\mathbb{E}[X \mid \mathcal{B}] \geqslant 0$.

(c) Si a est une constante, alors $\mathbb{E}[a \mid \mathcal{B}] = a$.

(d) Si $\mathcal{B} = \{\emptyset, \Omega\}$, alors $\mathbb{E}[X \mid \mathcal{B}] = \mathbb{E}[X]$ et $\mathbb{E}[X \mid \mathcal{F}] = X$.

(e) Soit Z une variable $\mathcal{B}$-mesurable telle que $XZ \in L^1$. Alors $\mathbb{E}[XZ \mid \mathcal{B}] = Z\mathbb{E}[X \mid \mathcal{B}]$. En particulier si X est $\mathcal{B}$-mesurable, $\mathbb{E}[X \mid \mathcal{B}] = X$.

(f) Si $\mathcal{C} \subseteq \mathcal{B}$, alors $\mathbb{E}\big[\mathbb{E}[X \mid \mathcal{B}] \mid \mathcal{C}\big] = \mathbb{E}[X \mid \mathcal{C}]$. En particulier, $\mathbb{E}\big[\mathbb{E}[X \mid \mathcal{B}]\big] = \mathbb{E}[X]$.

(g) Si $\sigma(X)$ et $\mathcal{B}$ sont indépendantes, alors $\mathbb{E}[X \mid \mathcal{B}] = \mathbb{E}[X]$, et plus généralement, pour toute Y $\mathcal{B}$-mesurable

$$\mathbb{E}[g(X, Y) \mid \mathcal{B}] = \int g(x, Y)\, \mathcal{L}_X(dx),$$

pour toute fonction mesurable $g : \mathbb{R}^2 \to \mathbb{R}$ telle que $g(X, Y)$ est intégrable.

(h) (Jensen) Si $\phi : \mathbb{R} \to \mathbb{R}_+$ est convexe et $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, alors

$$\phi(\mathbb{E}[X \mid \mathcal{B}]) \leqslant \mathbb{E}[\phi(X) \mid \mathcal{B}] .$$

5 Modes de convergence

Soit E un espace métrique. Soit $(X_n)_{n \geqslant 1}$ une suite de v.a. à valeurs dans E muni de sa tribu Borélienne. Soit également X une variable aléatoire à valeurs dans E.

Définition A.14

On suppose que toutes les v.a. sont définies sur un même espace de probabilité. On dit que X_n converge en probabilité vers X si pour tout $\delta > 0$

$$\mathbb{P}(dist(X_n, X) > \delta) \to 0 , \quad n \to \infty .$$

Définition A.15

On suppose que toutes les v.a. sont définies sur un même espace de probabilité. On dit que X_n converge presque sûrement vers X si

$$\mathbb{P}(dist(X_n, X) \to 0 , n \to \infty) = 1 .$$

Définition A.16

On suppose que toutes les v.a. sont définies sur un même espace de probabilité. Dans le cas où $E = \mathbb{R}^d$, on dit que X_n converge dans $L^p(\Omega, \mathcal{F}, \mathbb{P})$ vers X si $X_n, X \in L^p$ et si

$$\mathbb{E}[|X_n - X|^p] \to 0 , \quad n \to \infty .$$

On introduit à présent une notion de convergence qui ne nécessite pas que les v.a. soient définies sur un même espace de probabilité : en effet, cette notion de convergence porte sur les lois des v.a.

Définition A.17

On dit que X_n converge en loi vers X si pour toute fonction continue bornée $f : E \to \mathbb{R}$

$$\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)] , \quad n \to \infty .$$

Proposition A.18

1. *La convergence en probabilités implique la convergence en loi.*
2. *La convergence presque sûre (resp. L^p) implique la convergence en probabilités.*
3. *La convergence en probabilités implique la convergence presque sûre le long d'une sous-suite.*

Annexe B

Uniforme intégrabilité

Commençons par rappeler que toute variable aléatoire intégrable Y sur $(\Omega, \mathcal{F}, \mathbb{P})$ satisfait (de par le théorème de convergence dominée)

$$\lim_{a\to\infty} \mathbb{E}[|Y|\mathbf{1}_{|Y|>a}] = 0 \ . \tag{B.1}$$

Plus généralement :

Lemme B.1 (*Uniforme continuité de l'intégrale*)

Pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que pour tout $B \in \mathcal{F}$

$$\mathbb{P}(B) < \delta \Rightarrow \mathbb{E}[|Y|\mathbf{1}_B] < \varepsilon \ .$$

Démonstration Il suffit d'écrire

$$\begin{aligned} \mathbb{E}[|Y|\mathbf{1}_B] &= \mathbb{E}[|Y|\mathbf{1}_B\mathbf{1}_{|Y|\leqslant a}] + \mathbb{E}[|Y|\mathbf{1}_B\mathbf{1}_{|Y|>a}] \\ &\leqslant a\mathbb{P}(B) + \mathbb{E}[|Y|\mathbf{1}_{|Y|>a}] \ . \end{aligned}$$

Il suffit alors de choisir a tel que $\mathbb{E}[|Y|\mathbf{1}_{|Y|>a}] < \varepsilon/2$ puis $\delta = \varepsilon/2a$. □

Nous allons généraliser ces idées à des collections de v.a.

Soit $(X_i, i \in I)$ une collection (pas forcément dénombrable) de variables aléatoires à valeurs dans $\mathbb{R}$ (ou $\mathbb{R}^d$).

Définition B.2

On dit que la famille $(X_i, i \in I)$ est uniformément intégrable si

$$\lim_{a\to\infty} \sup_{i\in I} \mathbb{E}[|X_i|\mathbf{1}_{|X_i|>a}] = 0 \ .$$

La convergence (B.1) montre que toute variable aléatoire intégrable Y forme une famille (à un élément) qui est u.i. Plus généralement, il est facile de vérifier que toute collection finie de v.a. intégrables est uniformément intégrable.

Donnons à présent une caractérisation de l'uniforme intégrabilité, qui se rapproche de la continuité uniforme de l'intégrale.

Proposition B.3

La famille $(X_i, i \in I)$ est uniformément intégrable si et seulement si

1. $\sup_{i\in I} \mathbb{E}[|X_i|] < \infty$
2. *pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que pour tout $B \in \mathcal{F}$*

$$\mathbb{P}(B) < \delta \Rightarrow \sup_{i\in I} \mathbb{E}[|X_i|\mathbf{1}_B] < \varepsilon \ .$$

Démonstration Supposons que la famille est u.i. On a pour tout $B \in \mathcal{F}$

$$\mathbb{E}[|X_i|\mathbf{1}_B] = \mathbb{E}[|X_i|\mathbf{1}_B\mathbf{1}_{|X_i|\leqslant a}] + \mathbb{E}[|X_i|\mathbf{1}_B\mathbf{1}_{|X_i|>a}] \leqslant a\mathbb{P}(B) + \mathbb{E}[|X_i|\mathbf{1}_{|X_i|>a}] \ .$$

Si l'on prend $B = \Omega$, on trouve 1. Si l'on choisit a tel que

$$\sup_{i\in I} \mathbb{E}[|X_i|\mathbf{1}_{|X_i|>a}] < \varepsilon/2 \ ,$$

ainsi que $\delta = \varepsilon/2a$ on trouve 2.
Réciproquement on suppose que 1. et 2. sont vérifiées. Soit $\varepsilon > 0$ fixé, on souhaite montrer que pour tout a assez grand on a

$$\sup_{i\in I} \mathbb{E}[|X_i|\mathbf{1}_{|X_i|>a}] < \varepsilon \ .$$

On pose $M := \sup_{i\in I} \mathbb{E}[|X_i|]$, et $a = M/\delta$. Ainsi pour tout $i \in I$

$$\mathbb{P}(|X_i| > a) \leqslant M/a \leqslant \delta \ ,$$

et 2. assure que

$$\mathbb{E}[|X_i|\mathbf{1}_{|X_i|>a}] \leqslant \varepsilon \ .$$

□

Énonçons à présent deux critères assurant l'uniforme intégrabilité.

Proposition B.4

S'il existe Y dans L^1 telle que pour tout $i \in I$, p.s. $|X_i| \leqslant |Y|$ alors la famille $(X_i, i \in I)$ est uniformément intégrable.

Démonstration Il suffit de vérifier les points 1. et 2. de la proposition précédente. Le point 1. est clair. Concernant le point 2., on sait que

$$\sup_{i\in I} \mathbb{E}[|X_i|\mathbf{1}_B] \leqslant \mathbb{E}[|Y|\mathbf{1}_B] \ ,$$

et l'on peut conclure en utilisant les propriétés rappelées en début de section. □

Proposition B.5

Soit $g : \mathbb{R}_+ \to \mathbb{R}_+$ une fonction telle que

$$\lim_{t\to\infty} \frac{g(t)}{t} = +\infty \ .$$

Si $\sup_{i\in I} \mathbb{E}[g(|X_i|)] < \infty$ alors la famille $(X_i, i \in I)$ est uniformément intégrable. En particulier, si

$(X_i, i \in I)$ est bornée dans L^p pour un certain $p > 1$ alors elle est uniformément intégrable.

Démonstration Soient $M := \sup_{i\in I} \mathbb{E}[g(|X_i|)]$ et $\varepsilon > 0$. Il existe $a > 0$ tel que pour tout $t > a$ on a

$$\frac{g(t)}{t} > \frac{M}{\varepsilon} .$$

On écrit alors pour tout $i \in I$

$$\mathbb{E}[|X_i|\mathbf{1}_{|X_i|>a}] \leqslant \frac{\varepsilon}{M}\mathbb{E}[g(|X_i|)\mathbf{1}_{|X_i|>a}] \leqslant \frac{\varepsilon}{M}\mathbb{E}[g(|X_i|)] \leqslant \varepsilon .$$

□

Théorème B.6 (*Théorème de convergence dominée optimal*)

Soit $(X_n)_n$ une suite de v.a. qui converge en probabilités vers une variable aléatoire X. Si $(X_n, n \geqslant 1)$ est uniformément intégrable alors $X - X_n$ converge vers 0 dans L^1.

L'implication du théorème est en fait une équivalence (exercice).
Démonstration Par une proposition précédente, on sait qu'il existe une sous-suite n_k le long de laquelle X_{n_k} converge p.s. vers X. Par Fatou, on trouve

$$\mathbb{E}[|X|] \leqslant \liminf_{k\to\infty} \mathbb{E}[|X_{n_k}|] \leqslant \sup_n \mathbb{E}[|X_n|] .$$

Par la Proposition B.3, on sait que l'u.i. assure que $(X_n)_n$ est bornée dans L^1, et ainsi $X \in L^1$. On fixe $\varepsilon > 0$. On écrit

$$\begin{aligned}\mathbb{E}[|X_n - X|] &= \mathbb{E}[|X_n - X|\mathbf{1}_{\{|X_n-X|\leqslant\varepsilon\}}] + \mathbb{E}[|X_n - X|\mathbf{1}_{\{|X_n-X|>\varepsilon\}}]\\ &\leqslant \varepsilon + \mathbb{E}[|X_n|\mathbf{1}_{\{|X_n-X|>\varepsilon\}}] + \mathbb{E}[|X|\mathbf{1}_{\{|X_n-X|>\varepsilon\}}]\end{aligned}$$

La convergence en probabilité assure que $\mathbb{P}(|X_n - X| > \varepsilon) \to 0$. Par la Proposition B.3, on sait alors que pour tout n assez grand

$$\mathbb{E}[|X_n|\mathbf{1}_{\{|X_n-X|>\varepsilon\}}] \leqslant \varepsilon .$$

Par ailleurs, l'uniforme continuité de l'intégrale assure que pour tout n assez grand

$$\mathbb{E}[|X|\mathbf{1}_{\{|X_n-X|>\varepsilon\}}] \leqslant \varepsilon .$$

□

On conclut avec un résultat d'uniforme intégrabilité qui sera utilisé à plusieurs reprises dans le cours.

Proposition B.7

Soit X une variable aléatoire intégrable. Soit $(\mathcal{F}_i)_{i\in I}$ une collection de sous-tribus de $\mathcal{F}$ et soit $X_i := \mathbb{E}[X \mid \mathcal{F}_i]$ pour tout $i \in I$. Alors la famille $(X_i, i \in I)$ est uniformément intégrable.

Démonstration Par l'inégalité de Jensen pour l'espérance conditionnelle

$$|X_i| \leqslant \mathbb{E}[|X| \mid \mathcal{F}_i] ,$$

et ainsi
$$\sup_i \mathbb{E}[|X_i|] \leqslant \sup_i \mathbb{E}[\mathbb{E}[|X| \mid \mathcal{F}_i]] = \mathbb{E}[|X|] < \infty \ .$$
Ainsi $(X_i, i \in I)$ est bornée dans L^1. Par ailleurs l'événement $\{|X_i| > a\}$ est $\mathcal{F}_i$ mesurable et ainsi

$$\begin{aligned}
\mathbb{E}[|X_i|\mathbf{1}_{\{|X_i|>a\}}] &= \mathbb{E}[|\mathbb{E}[X \mid \mathcal{F}_i]|\mathbf{1}_{\{|X_i|>a\}}] \\
&= \mathbb{E}[|\mathbb{E}[X\mathbf{1}_{\{|X_i|>a\}} \mid \mathcal{F}_i]|] \\
&\leqslant \mathbb{E}[\mathbb{E}[|X|\mathbf{1}_{\{|X_i|>a\}} \mid \mathcal{F}_i]] \\
&\leqslant \mathbb{E}[|X|\mathbf{1}_{\{|X_i|>a\}}] \ .
\end{aligned}$$

L'uniforme continuité de l'intégrale assure que pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que

$$\mathbb{P}(A) < \delta \Rightarrow \mathbb{E}[|X|\mathbf{1}_A] < \varepsilon \ .$$

Comme
$$\sup_i \mathbb{P}(|X_i| > a) \leqslant \frac{1}{a} \sup_i \mathbb{E}[|X_i|] \leqslant \frac{1}{a}\mathbb{E}[|X|] \ .$$
Il existe $a > 0$ tel que
$$\sup_i \mathbb{P}(|X_i| > a) < \delta \ ,$$
et ainsi
$$\sup_i \mathbb{E}[|X_i|\mathbf{1}_{\{|X_i|>a\}}] \leqslant \varepsilon \ .$$
On en déduit l'u.i. de $(X_i, i \in I)$.

$\square$

Par exemple, soit X une variable aléatoire intégrable et $(\mathcal{F}_i)$ une filtration. Alors, la famille $\{\mathbb{E}[X \mid \mathcal{F}_T] : T \text{ est un temps d'arrêt}\}$ est uniformément intégrable.

Annexe C

Martingales à temps discret

Soit $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\in\mathbb{N}}, \mathbb{P})$ un espace de probabilité filtré, c'est-à-dire, un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ muni d'une collection $(\mathcal{F}_n)_{n\in\mathbb{N}}$ de sous-tribus de $\mathcal{F}$ satisfaisant

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \dots .$$

Une telle collection est appelée *filtration.*

Définition C.1

Soit $M = (M_n)_{n\in\mathbb{N}}$ une collection de variables aléatoires définies sur $(\Omega, \mathcal{F}, \mathbb{P})$ et à valeurs réelles. On dit que M est une martingale (resp. sur-martingale, resp. sous-martingale) si :

1. *M est adapté à la filtration $(\mathcal{F}_n)_n$, c'est-à-dire, pour tout $n \in \mathbb{N}$, M_n est $\mathcal{F}_n$-mesurable,*
2. *pour tout $n \in \mathbb{N}$, M_n est intégrable,*
3. *pour tout $n \in \mathbb{N}$, $\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] = M_n$ p.s. (resp. $\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] \leqslant M_n$ p.s., resp. $\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] \geqslant M_n$ p.s.).*

Quelques exemples : la marche aléatoire simple est une martingale ; la marche aléatoire biaisée (proba p d'augmenter de 1 et $1 - p$ de diminuer de 1) est une sous-martingale lorsque $p \geqslant 1/2$ et une sur-martingale lorsque $p \leqslant 1/2$; pour toute collection de variables aléatoires indépendantes intégrables $(X_k)_{k\geqslant 0}$ le processus $M_n = X_0 + \ldots + X_n$ est une martingale dans la filtration $\mathcal{F}_n := \sigma(X_0, \ldots, X_n)$ pour peu que les $(X_k)_{k\geqslant 0}$ soient d'espérance nulle.
Commençons par rappeler quelques propriétés simples.

Lemme C.2

Soient M une martingale (resp. une sous-martingale) et $f : \mathbb{R} \to \mathbb{R}$ convexe (resp. convexe croissante). On suppose que pour tout $n \in \mathbb{N}$, $f(M_n)$ est intégrable. Alors $f(M)$ est une sous-martingale.

Démonstration Il suffit d'appliquer l'inégalité de Jensen dans l'espérance conditionnelle. □

On rappelle qu'une variable aléatoire $T : \Omega \to \mathbb{N} \cup \{+\infty\}$ est un temps d'arrêt si pour tout $n \in \mathbb{N}$, $\{T \leqslant n\} \in \mathcal{F}_n$. On notera que si T est un temps d'arrêt, alors $\{T > n\} = \{T \leqslant n\}^\complement \in \mathcal{F}_n$ et $\{T = n\} = \{T \leqslant n\} \cap \{T \leqslant n - 1\}^\complement \in \mathcal{F}_n$.

Lemme C.3

Soient M une martingale (resp. sous-martingale, resp. sur-martingale) et T un temps d'arrêt. On pose $M_n^T := M_{T\wedge n}$ pour tout $n \in \mathbb{N}$. Alors M^T est une martingale (resp. sous-martingale, resp. sur-

martingale).

Démonstration Tout d'abord $|M_n^T| \leqslant |M_0| + \ldots + |M_n|$ et ainsi M_n^T est intégrable. Par ailleurs, chaque variable $M_k \mathbf{1}_{T=k}$ est $\mathcal{F}_k$-mesurable et ainsi la somme

$$M_n^T = \sum_{k=0}^{n} M_k \mathbf{1}_{T=k} + M_n \mathbf{1}_{T>n} ,$$

est $\mathcal{F}_n$-mesurable. Enfin

$$\begin{aligned}
\mathbb{E}[M_{n+1}^T \mid \mathcal{F}_n] &= \sum_{k=0}^{n} \mathbb{E}[M_k \mathbf{1}_{T=k} \mid \mathcal{F}_n] + \mathbb{E}[M_{n+1} \mathbf{1}_{T>n} \mid \mathcal{F}_n] \\
&= \sum_{k=0}^{n} M_k \mathbf{1}_{T=k} + \mathbf{1}_{T>n} \mathbb{E}[M_{n+1} \mid \mathcal{F}_n] \\
&= \sum_{k=0}^{n} M_k \mathbf{1}_{T=k} + M_n \mathbf{1}_{T>n} \\
&= M_n^T .
\end{aligned}$$

□

Les martingales apparaissent dans de nombreux contextes, il est donc utile d'établir des résultats généraux les concernant. Nous allons aborder trois thèmes :

1. les théorèmes d'arrêt, qui permettent d'évaluer $\mathbb{E}[M_T]$ où T est un temps d'arrêt,
2. les inégalités maximales, qui fournissent des bornes sur la queue de probabilité du supremum $\sup_{0 \leqslant k \leqslant n} M_n$ d'une martingale (resp. sur-martingale, resp. sous-martingale),
3. les théorèmes de convergence, qui fournissent des informations sur le comportement asymptotique d'une martingale (resp. sur-martingale, resp. sous-martingale).

1 Théorème d'arrêt borné

Etant donné un temps d'arrêt T, on introduit la tribu du passé avant T :

$$\mathcal{F}_T := \{A \in \mathcal{F} : A \cap \{T \leqslant n\} \in \mathcal{F}_n , \quad \forall n \in \mathbb{N}\} .$$

On peut vérifier que $A \in \mathcal{F}_T$ si et seulement si A s'écrit

$$A = \bigcup_n \Big(B_n \cap \{T = n\} \Big) ,$$

où $B_n \in \mathcal{F}_n$. Cela justifie l'appellation "tribu du passé" car les événements ne peuvent pas s'appuyer sur une information postérieure à T.

Théorème C.4 (*Théorème d'arrêt borné*)

Soit M une martingale (resp. une sous-martingale, resp. une sur-martingale). Soient $S \leqslant T$ deux temps d'arrêt bornés, *c'est-à-dire, tels qu'il existe $m \geqslant 0$ tel que $0 \leqslant S \leqslant T \leqslant m$ p.s. Alors M_S, M_T*

sont intégrables et presque sûrement

$$\mathbb{E}[M_T \mid \mathcal{F}_S] = M_S , \quad \textit{resp.} \ \geqslant M_S , \quad \textit{resp.} \ \leqslant M_S .$$

Démonstration On a $|M_T| \leqslant |M_0| + \ldots + |M_m|$ p.s., et ainsi M_T est intégrable. De même pour M_S. Par ailleurs pour tout $A \in \mathcal{F}_S$

$$\mathbb{E}[M_T \mathbf{1}_A] = \mathbb{E}[M_{T\wedge m}\mathbf{1}_A] = \sum_{k=0}^{m} \mathbb{E}[M_{T\wedge m}\mathbf{1}_{A\cap\{S=k\}}] .$$

Or $(M_{T\wedge n}, n \in \mathbb{N})$ est une martingale et $A \cap \{S = k\}$ est $\mathcal{F}_k$ mesurable de sorte que

$$\mathbb{E}[M_{T\wedge m}\mathbf{1}_{A\cap\{S=k\}}] = \mathbb{E}[M_{T\wedge k}\mathbf{1}_{A\cap\{S=k\}}] .$$

Ainsi

$$\mathbb{E}[M_T \mathbf{1}_A] = \sum_{k=0}^{m} \mathbb{E}[M_{T\wedge k}\mathbf{1}_{A\cap\{S=k\}}] = \sum_{k=0}^{m} \mathbb{E}[M_S\mathbf{1}_{A\cap\{S=k\}}] = \mathbb{E}[M_S\mathbf{1}_A] .$$

□

Nous verrons plus loin des conditions sur la martingale M assurant que cette égalité reste vraie pour tous temps d'arrêt $S \leqslant T$.

2 Inégalités maximales

Proposition C.5

- *Si M est une sous-martingale alors pour tout $a > 0$*

$$a\mathbb{P}(\sup_{0\leqslant k\leqslant n} M_k > a) \leqslant \mathbb{E}[M_n\mathbf{1}_{\sup_{0\leqslant k\leqslant n} M_k > a}] \leqslant \mathbb{E}[M_n^+] ,$$

- *Si M est une sur-martingale alors pour tout $a > 0$*

$$a\mathbb{P}(\sup_{0\leqslant k\leqslant n} M_k > a) \leqslant \mathbb{E}[M_0] + \mathbb{E}[M_n^-] .$$

Démonstration On fixe $n \in \mathbb{N}$ et l'on note $T := \inf\{k \in \mathbb{N} : M_k > a\} \wedge n$. On note que l'on a $\{\sup_{0\leqslant k\leqslant n} M_k \leqslant a\} \subset \{T = n\}$ et $\{\sup_{0\leqslant k\leqslant n} M_k > a\} = \{M_T > a\}$. On calcule alors

$$\mathbb{E}[M_T] = \mathbb{E}[M_T\mathbf{1}_{\sup_{0\leqslant k\leqslant n} M_k > a}] + \mathbb{E}[M_T\mathbf{1}_{\sup_{0\leqslant k\leqslant n} M_k \leqslant a}] \geqslant a\mathbb{P}(\sup_{0\leqslant k\leqslant n} M_k > a) + \mathbb{E}[M_n\mathbf{1}_{\sup_{0\leqslant k\leqslant n} M_k \leqslant a}] .$$

- Si M est une sous-martingale alors le théorème d'arrêt borné appliqué aux temps d'arrêt $T \leqslant n$ assure que

$$\mathbb{E}[M_n] \geqslant \mathbb{E}[M_T] \geqslant a\mathbb{P}(\sup_{0\leqslant k\leqslant n} M_k > a) + \mathbb{E}[M_n\mathbf{1}_{\sup_{0\leqslant k\leqslant n} M_k \leqslant a}] ,$$

  et ainsi

$$a\mathbb{P}(\sup_{0\leqslant k\leqslant n} M_k > a) \leqslant \mathbb{E}[M_n\mathbf{1}_{\sup_{0\leqslant k\leqslant n} M_k > a}] \leqslant \mathbb{E}[M_n^+] .$$

— Si M est une sur-martingale alors le théorème d'arrêt borné appliqué aux temps d'arrêt $0 \leqslant T$ assure que
$$\mathbb{E}[M_0] \geqslant \mathbb{E}[M_T] \geqslant a\mathbb{P}(\sup_{0\leqslant k\leqslant n} M_k > a) + \mathbb{E}[M_n \mathbf{1}_{\sup_{0\leqslant k\leqslant n} M_k \leqslant a}] \ ,$$
et ainsi
$$a\mathbb{P}(\sup_{0\leqslant k\leqslant n} M_k > a) \leqslant \mathbb{E}[M_0] - \mathbb{E}[M_n \mathbf{1}_{\sup_{0\leqslant k\leqslant n} M_k \leqslant a}] \leqslant \mathbb{E}[M_0] + \mathbb{E}[M_n^-] \ .$$

□

Proposition C.6 (*Inégalité maximale de Doob*)

Soit M une martingale. Posons $M_n^* := \sup_{0\leqslant k\leqslant n} |M_k|$ pour tout $n \in \mathbb{N}$. Pour tout $p > 1$ et tout $n \in \mathbb{N}$
$$\mathbb{E}[(M_n^*)^p] \leqslant \Big(\frac{p}{p-1}\Big)^p \mathbb{E}[|M_n|^p] \ .$$

Démonstration On va prouver une propriété plus générale. Soit X une sous-martingale positive et soit $\tilde{X}_n := \sup_{0\leqslant k\leqslant n} X_k$. Pour tout $p > 1$ et tout $n \in \mathbb{N}$
$$\mathbb{E}[(\tilde{X}_n)^p] \leqslant \Big(\frac{p}{p-1}\Big)^p \mathbb{E}[X_n^p] \ .$$
Une fois cette propriété établie, il suffit de prendre $X_n = |M_n|$ pour déduire le résultat.
On peut supposer que $\mathbb{E}[X_n^p] < \infty$ car sinon il n'y a rien à montrer. Comme $x \mapsto x^p$ est convexe croissante sur $\mathbb{R}_+$, X^p est toujours une sous-martingale positive et on en déduit que $\mathbb{E}[X_k^p] < \infty$ pour tout $0 \leqslant k \leqslant n$, puis que pour une constante $C_p > 0$
$$\mathbb{E}[(\tilde{X}_n)^p] \leqslant \mathbb{E}[(X_0 + X_1 + \ldots + X_n)^p] \leqslant C_p\Big(\mathbb{E}[X_0^p] + \ldots \mathbb{E}[X_n^p]\Big) < \infty \ .$$
L'inégalité maximale précédemment prouvée montre que
$$a\mathbb{P}(\sup_{0\leqslant k\leqslant n} X_k > a) \leqslant \mathbb{E}[X_n \mathbf{1}_{\sup_{0\leqslant k\leqslant n} X_k > a}] \ .$$
En multipliant chaque membre de cette inégalité par a^{p-2} et en intégrant par rapport à la mesure de Lebesgue sur $[0, \infty[$ on obtient
$$\int_0^\infty a^{p-1}\mathbb{P}(\sup_{0\leqslant k\leqslant n} X_k > a)da = \frac{1}{p}\mathbb{E}[(\tilde{X}_n)^p] \ ,$$
ainsi que
$$\int_0^\infty a^{p-2}\mathbb{E}[X_n \mathbf{1}_{\sup_{0\leqslant k\leqslant n} X_k > a}]da = \frac{1}{p-1}\mathbb{E}[X_n(\tilde{X}_n)^{p-1}] \leqslant \frac{1}{p-1}\mathbb{E}[(X_n)^p]^{1/p}\mathbb{E}[(\tilde{X}_n)^p]^{\frac{p-1}{p}} \ .$$
Ainsi
$$\frac{1}{p}\mathbb{E}[(\tilde{X}_n)^p] \leqslant \frac{1}{p-1}\mathbb{E}[(X_n)^p]^{1/p}\mathbb{E}[(\tilde{X}_n)^p]^{\frac{p-1}{p}} \ ,$$
et l'inégalité voulue en découle. □

3 Convergence de martingales

La preuve du résultat suivant est assez longue, nous ne la rappellerons pas et nous renvoyons vers [LG06, Th 12.3.3].

Théorème C.7

Soit M une martingale (resp. sous-martingale, resp. sur-martingale) bornée dans L^1, c'est-à-dire, telle que $\sup_{n\in\mathbb{N}} \mathbb{E}[|M_n|] < \infty$. Alors M_n converge presque sûrement vers une variable aléatoire intégrable M_∞.

On notera qu'il suffit de prouver le résultat pour une sous-martingale : en effet, si M est une sur-martingale bornée dans L^1 alors $-M$ est une sous-martingale bornée dans L^1.

Il se trouve qu'en général cette convergence n'a pas lieu dans L^1. Considérons l'exemple suivant. Soit S la marche aléatoire simple issue de 1 et $T := \inf\{n \in \mathbb{N} : S_n = 0\}$ son premier temps d'atteinte de 0. On notera que $T < \infty$ presque sûrement mais que T n'est pas borné. Par un résultat précédent, le processus $M_n = S_{T\wedge n}$ est une martingale et l'on a

$$\mathbb{E}[M_n] = \mathbb{E}[M_0] = 1 \; .$$

Etant positive, la martingale arrêtée M est alors bornée dans L^1. On en déduit qu'elle converge p.s. vers une variable aléatoire intégrable M_∞. Or pour tout $k \in \mathbb{N}$ et pour tout $n \geqslant k$

$$M_n \mathbf{1}_{T=k} = S_{T\wedge n} \mathbf{1}_{T=k} = 0 \; .$$

Ainsi $M_\infty = 0$ sur l'événement $\{T = k\}$, et comme $T < \infty$ p.s., on en déduit que $M_\infty = 0$ p.s. Si M_n convergeait vers M_∞ dans L^1 alors $1 = \mathbb{E}[M_n] \to \mathbb{E}[M_\infty] = 0$.

4 Martingales fermées

L'objectif de cette section est double. Tout d'abord, on souhaiterait trouver des conditions sur une martingale sous lesquelles le théorème d'arrêt est encore vrai pour des temps d'arrêt quelconques (pas forcément bornés). Par ailleurs, on souhaiterait trouver des conditiones sous lesquelles la convergence p.s. du théorème précédent a également lieu dans L^1. Il se trouve que dans les deux cas, la notion de martingale fermée joue un rôle prépondérant.

Définition C.8

On dit qu'une martingale M est fermée s'il existe une v.a. $X \in L^1$ telle que pour tout $n \in \mathbb{N}$

$$M_n = \mathbb{E}[X \mid \mathcal{F}_n] \; .$$

Théorème C.9

Il y a équivalence entre :

(i) M est fermée

(ii) M est uniformément intégrable.

(iii) M converge p.s. et dans L^1 vers une variable aléatoire M_∞ intégrable.

Si l'une des trois conditions est vérifiée, alors nécessairement M est fermée par sa limite M_∞.

On notera que la condition de martingale fermée est nécessaire et suffisante pour que la convergence p.s. du théorème vu précédemment soit renforcée en une convergence L^1 !
Démonstration (i) ⇒ (ii). C'est une conséquence de la Proposition B.7.
(ii) ⇒ (iii). Etant u.i., la suite $(M_n, n \in \mathbb{N})$ est bornée dans L^1. Comme c'est en outre une martingale, elle converge p.s. vers une variable aléatoire intégrable M_∞. Le théorème de convergence dominée optimal, Théorème B.6, montre alors que M_n converge aussi dans L^1 vers M_∞.
(iii) ⇒ (i). Pour tout $A \in \mathcal{F}_n$, on a

$$\mathbb{E}[M_n \mathbf{1}_A] = \mathbb{E}[M_{n+k} \mathbf{1}_A] \ .$$

Or $M_{n+k} \to M_\infty$ dans L^1 quand $k \to \infty$. Ainsi

$$\mathbb{E}[M_{n+k} \mathbf{1}_A] \to \mathbb{E}[M_\infty \mathbf{1}_A] \ ,$$

et l'on en déduit que

$$\mathbb{E}[M_n \mathbf{1}_A] = \mathbb{E}[M_\infty \mathbf{1}_A] \ .$$

On a donc montré que $\mathbb{E}[M_\infty \mid \mathcal{F}_n] = M_n$ p.s., pour tout $n \in \mathbb{N}$. □

Soit M une martingale fermée. On vient de voir qu'il existe alors une v.a. M_∞ telle que p.s. $M_n \to M_\infty$. Ainsi pour tout temps T, on peut définir

$$M_T(\omega) = M_{T(\omega)}(\omega) \ .$$

(En particulier, sur l'événement $\{T = \infty\}$ cette définition fait bien sens).

Théorème C.10

Soit M une martingale fermée. Pour tout temps d'arrêt T on a presque sûrement

$$M_T = \mathbb{E}[M_\infty \mid \mathcal{F}_T] \ .$$

Par ailleurs pour tous temps d'arrêt $S \leqslant T$ on a presque sûrement

$$M_S = \mathbb{E}[M_T \mid \mathcal{F}_S] \ .$$

Démonstration La deuxième propriété est une conséquence de la première. En effet

$$\mathbb{E}[M_T \mid \mathcal{F}_S] = \mathbb{E}[\mathbb{E}[M_\infty \mid \mathcal{F}_T] \mid \mathcal{F}_S] = \mathbb{E}[M_\infty \mid \mathcal{F}_S] = M_S \ .$$

La première propriété se prouve comme suit. Supposons que pour tout $n \in \mathbb{N} \cup \{\infty\}$

$$\mathbb{E}[M_\infty \mid \mathcal{F}_T] \mathbf{1}_{\{T=n\}} = \mathbb{E}[M_\infty \mid \mathcal{F}_n] \mathbf{1}_{\{T=n\}} \ . \tag{C.1}$$

Alors p.s.

$$\mathbb{E}[M_\infty \mid \mathcal{F}_T] \mathbf{1}_{\{T=n\}} = \mathbb{E}[M_\infty \mid \mathcal{F}_n] \mathbf{1}_{\{T=n\}} = M_n \mathbf{1}_{\{T=n\}} = M_T \mathbf{1}_{\{T=n\}} \ ,$$

et comme cela est vrai pour tout $n \in \mathbb{N} \cup \{\infty\}$, on en déduit que $M_T = \mathbb{E}[M_\infty \mid \mathcal{F}_T]$.
Il nous faut donc prouver l'identité ci-dessus. Pour cela, on commence par prouver que le produit d'une v.a. X $\mathcal{F}_T$ mesurable par $\mathbf{1}_{\{T=n\}}$ est $\mathcal{F}_n$ mesurable. En effet pour tout Borélien B ne contenant pas 0, du fait que $\{X \in B\} \in \mathcal{F}_T$ on a

$$\{X \mathbf{1}_{\{T=n\}} \in B\} = \{X \in B\} \cap \{T = n\} \in \mathcal{F}_n \ .$$

Par ailleurs pour tout Borélien B contenant 0 on a

$$\{X\mathbf{1}_{\{T=n\}} \in B\} = \{T \neq n\} \cup \Big(\{X \in B^{\complement}\} \cap \{T = n\}\Big) ,$$

qui est également $\mathcal{F}_n$ mesurable.
Les deux membres de (C.1) sont donc $\mathcal{F}_n$ mesurables. Nous allons montrer que leurs espérances contre $\mathbf{1}_A$ coïncident, pour tout $A \in \mathcal{F}_n$. Commençons par observer que $A \cap \{T = n\}$ est $\mathcal{F}_T$ mesurable : en effet, si $k < n$ alors $A \cap \{T = n\} \cap \{T \leqslant k\} = \varnothing$ et si $k \geqslant n$, $A \cap \{T = n\} \cap \{T \leqslant k\} = A \cap \{T = n\}$. Ainsi

$$\begin{aligned}
\mathbb{E}[\mathbb{E}[M_\infty \mid \mathcal{F}_T]\mathbf{1}_{A\cap\{T=n\}}] &= \mathbb{E}[\mathbb{E}[M_\infty \mathbf{1}_{A\cap\{T=n\}} \mid \mathcal{F}_T]] \\
&= \mathbb{E}[M_\infty \mathbf{1}_{A\cap\{T=n\}}] \\
&= \mathbb{E}[\mathbb{E}[M_\infty \mathbf{1}_{A\cap\{T=n\}} \mid \mathcal{F}_n]] \\
&= \mathbb{E}[\mathbb{E}[M_\infty \mid \mathcal{F}_n]\mathbf{1}_{A\cap\{T=n\}}] .
\end{aligned}$$

□

Bibliographie

[cn11] E. Çinlar. *Probability and stochastics*, vol. 261 of *Graduate Texts in Mathematics*. Springer, New York, 2011, xiv+557. doi:10.1007/978-0-387-87859-1.

[ELVE19] W. E, T. Li, and E. Vanden-Eijnden. *Applied stochastic analysis*, vol. 199 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2019, xxi+305. doi:10.1090/gsm/199.

[KS91] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*, vol. 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second ed., 1991, xxiv+470. doi:10.1007/978-1-4612-0949-2.

[LG06] J.-F. Le Gall. Intégration, probabilités et processus aléatoires, 2006. URL https://www.imo.universite-paris-saclay.fr/~jean-francois.le-gall/IPPA2.pdf.

[LG13] J.-F. Le Gall. *Mouvement brownien, martingales et calcul stochastique*, vol. 71 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer, Heidelberg, 2013, viii+176. doi:10.1007/978-3-642-31898-6.

[RY99] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, vol. 293 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third ed., 1999, xiv+602. doi:10.1007/978-3-662-06400-9.

[Sal22] J. Salez. Introduction to stochastic calculus, 2022. URL https://www.ceremade.dauphine.fr/~salez/stoc.pdf.