Processus Stochastiques

MAT 4514

2025

Sandro Franceschi

Wojciech Pieczynski

Sholom Schechtman

Table des matières

Avertissement

Ce cours s'inspire des références cités dans la bibliographie, la principale étant le livre de Jean François Le Gall, mais aussi le cours de Randal Douc (qui reprend la majeure partie de la structure du livre "Calcul Stochastique appliqué à la finance" de Lamberton et Lapeyre). De nombreux exercices sont repris des TD du cours de calcul stochastique de l'université Paris Diderot donnés par Benoît Laslier.

N'hésitez pas à nous communiquer par mail les erreurs qu'il pourrait rester ainsi qu'à proposer toute sorte de suggestions visant à améliorer ce cours : `sandro.franceschi@telecom-sudparis.eu` pour les martingales à temps continu et la partie sur le mouvement brownien et le calcul stochastique, `sholom.schechtman@telecom-sudparis.eu` sur les martingales à temps discret et les chaînes de Markov et `jules.flin@telecom-sudparis.eu` pour les TP.

Programme prévisionnel

Séance	Date	Heure	Type	Thématique	Intervenant
1	30/04	3H	Cours/TD	Martingales discrètes	Sholom Schechtman
2	07/05	3H	Cours/TD	Martingales discrètes	Sholom Schechtman
3	09/05	3H	Cours/TD	Martingales discrètes	Sholom Schechtman
4	12/05	3H	TP	Martingales	Jules Flin
5	13/05	3H	Cours/TD	Martingales continues	Sandro Franceschi
6	14/05	3H	Cours/TD	Mouvement Brownien	Sandro Franceschi
7	20/05	3H	Cours/TD	Intégrale stochastique	Sandro Franceschi
8	21/05	3H	Cours/TD	Formules d'Itô	Sandro Franceschi
9	22/05	3H	Cours/TD	Équation Différentielle Stochastique	Sandro Franceschi
10	26/05	3H	TP	Mouvement brownien	Jules Flin
11	27/05	3H	Cours/TD	Chaines de Markov	Sholom Schechtman
12	28/05	3H	Cours/TD	Chaines de Markov	Sholom Schechtman
13	04/06	3H	TP	Chaines de Markov	Jules Flin
14	12/06	2H	CF	Contrôle final	

Notation

— Une partie sur les martingales discrètes et les chaînes de Markov, notée sur 6 corrigée par Sholom Schechtman.

— Une partie sur le mouvement brownien et le calcul stochastiques, notée sur 6 corrigée par Sandro Franceschi.

— Les 3 TP sont notés sur 3 points chacun et corrigés par Jules Flin.

— Cela donne une note sur 21, laissée sur 20.

Consignes pour le contrôle final

— La partie de Sandro Franceschi et Sholom Schechtman sont à rédiger sur deux copies séparées.

— Indiquez de manière lisible la question traitée et justifiez vos réponses.

— La qualité de la rédaction sera prise en compte. Soulignez ou encadrez vos résultats.

— Tous les documents papier sont autorisés.

— L'ordinateur, les calculatrices et les téléphones portables sont interdit.

Première partie

Martingales

Chapitre 1

Martingales à temps discret

D'après le Larousse une martingale est « *Un système de jeu qui prétend, selon des principes fondés sur le calcul des probabilités, assurer un bénéfice certain dans les jeux de hasard.* »

Mathématiquement une martingale est un processus stochastique dont la moyenne n'évolue pas au cours du temps. Contrairement à la phrase citée il est bien sûr impossible d'établir une stratégie qui gagne contre un casino ; nous pouvons le démontrer mathématiquement à l'aide des martingales. Plus important, ces dernières sont un outil puissant pour démontrer des théorèmes probabilistes et apparaissent naturellement dans un grand nombre d'applications.

1.1 Premières définitions

Dans ce chapitre on fixe $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Toutes les variables aléatoires (v.a.) seront définies sur cet espace et seront à valeurs réelles.

Définition 1.1. On dit que $(\mathcal{F}_k)_{k\in\mathbb{N}}$ est une filtration si $(\mathcal{F}_k)$ est une suite croissante de tribus : $\mathcal{F}_m \subset \mathcal{F}_l$ si $m \leqslant l$.

On interprète souvent $\mathcal{F}_k$ comme l'ensemble d'informations disponibles au temps $k \in \mathbb{N}$. Une filtration est alors un flux d'informations. Le fait que $m \leqslant l \implies \mathcal{F}_m \subset \mathcal{F}_l$ reflète le fait que l'ensemble d'informations disponibles croît avec le temps.

Définition 1.2. Une suite de v.a. (X_k) est dite adaptée à $(\mathcal{F}_k)$ si pour tout $k \in \mathbb{N}$, X_k est $\mathcal{F}_k$-mesurable.

En d'autres mots, une suite adaptée (X_k) est telle que $\mathscr{F}_k$ (les informations disponibles au temps $k \in \mathbb{N}$) contienne toutes les informations nécessaires pour caractériser X_k.

Dans la suite, nous fixerons une filtration $(\mathcal{F}_k)$ et on parlera juste de suite adaptée (sous-entendu à $(\mathcal{F}_k)$).

Définition 1.3. Une suite de v.a. (X_k) adaptée et intégrable (c-à-d $\mathbb{E}[|X_k|] < +\infty$) est une :

1. Martingale si pour tout $k \in \mathbb{N}$, $X_k = \mathbb{E}[X_{k+1}|\mathcal{F}_k]$.
2. Sous-martingale si pour tout $k \in \mathbb{N}$, $X_k \leqslant \mathbb{E}[X_{k+1}|\mathcal{F}_k]$.
3. Sur-martingale si pour tout $k \in \mathbb{N}$, $X_k \geqslant \mathbb{E}[X_{k+1}|\mathcal{F}_k]$.

Remarque 1.4. Notons que pour une martingale (resp. sous/sur martingale) on a immédiatement $\mathbb{E}[X_k] = \mathbb{E}[X_{k+1}]$ (resp. $\mathbb{E}[X_k] \leqslant \mathbb{E}[X_{k+1}]$/ $\mathbb{E}[X_k] \geqslant \mathbb{E}[X_{k+1}]$).

On interprète souvent la suite (X_k) comme la fortune d'un joueur jouant à un jeu selon une stratégie. Dans cette interprétation une situation favorable est serait que $(\mathbb{E}[X_k])$ est une suite croissante : c'est le cas pour une sous-martingale. Bien sûr, en jouant dans un casino (X_k) est plutôt une sur-martingale (on est donc en moyenne perdant).

Exemple 1.5 (Marche aléatoire)**.** Soit (ξ_k) une suite de variables i.i.d. de moyennes finies. On définit $\mathcal{F}_k = \sigma(\xi_0, \dots, \xi_k)$ et

$$X_{k+1} = \sum_{i=0}^{k+1} \xi_i = X_k + \xi_{k+1} .$$

Alors (X_k) est une martingale si $\mathbb{E}[\xi_0] = 0$, sur-martingale si $\mathbb{E}[\xi_0] \leqslant 0$ et sous-martingale si $\mathbb{E}[\xi_0] \geqslant 0$.

Exemple 1.6 (Martingale formée par un produit)**.** Soit (ξ_k) une suite de variable i.i.d., positives et de moyennes finies. On définit $\mathcal{F}_k = \sigma(\xi_0, \dots, \xi_k)$ et

$$X_{k+1} = \prod_{i=0}^{k+1} \xi_i = X_k \xi_{k+1} .$$

On a alors que (X_k) est une martingale si $\mathbb{E}[\xi_0] = 1$, une sur-martingale si $\mathbb{E}[\xi_0] < 1$ et une sous-martingale si $\mathbb{E}[\xi_0] > 1$.

Exemple 1.7 (Martingale de Doob [1])**.** Soit X une v.a. intégrable, soit $(\mathcal{F}_k)$ une filtration, la suite (X_k) définie par

$$X_k = \mathbb{E}[X | \mathcal{F}_k]$$

est une martingale. Une telle martingale est aussi appelée fermée.

Les propositions suivantes montrent comment construire une grande classe de (sous-)martingales.

Proposition 1.8. *Soit (X_k) une martingale et $g : \mathbb{R} \to \mathbb{R}$ une fonction convexe t.q. $\mathbb{E}[|g(X_k)|] < +\infty$ pour tout k. Alors le processus $(g(X_k))$ est une sous-martingale.*

Démonstration. Par hypothèse, pour tout $k \in \mathbb{N}$, $g(X_k)$ est intégrable. De plus, par l'inégalité de Jensen :

$$g(X_k) = g(\mathbb{E}[X_{k+1} | \mathcal{F}_k]) \leqslant \mathbb{E}[g(X_{k+1}) | \mathcal{F}_k] .$$

□

Proposition 1.9. *Soit (X_k) une martingale et (B_k) un processus* ***borné****, adapté à $(\mathcal{F}_k)$ alors*

$$Y_k := \sum_{i=0}^{k} B_{i-1}(X_i - X_{i-1})$$

est une martingale.

On peut interpréter (Y_k) comme le gain d'un joueur qui mise à chaque tour B_k et gagne (pour une unité mise) $X_{k+1} - X_k$.

Démonstration. Pour tout k, Y_k est bien intégrable comme somme de v.a. intégrables. De plus, on a bien

$$\begin{aligned}
\mathbb{E}[Y_{k+1} | \mathcal{F}_k] &= \mathbb{E}[\sum_{i=0}^{k+1} B_{i-1}(X_i - X_{i-1}) | \mathcal{F}_k] \\
&= \sum_{i=0}^{k} B_{i-1}(X_i - X_{i-1}) + B_k \mathbb{E}[X_{k+1} - X_k | \mathcal{F}_k] \\
&= \sum_{i=0}^{k} B_{i-1}(X_i - X_{i-1}) = Y_k ,
\end{aligned}$$

où à la dernière ligne on a utilisé $\mathbb{E}[X_{k+1} | \mathcal{F}_k] = X_k$. □

1. Joseph Leo Doob (1910-2004) est un mathématicien américain et un des fondateurs de la théorie des martingales.

1.2 Temps d'arrêt et martingales arrêtées

Nous rappelons que se place $(\Omega, \mathcal{F}, \mathbb{P})$ est un espace probabilisé.

Définition 1.10. Soit $(\mathcal{F}_k)$ une filtration. On dit que $\tau : \Omega \to \mathbb{N}$, une v.a. à valeurs dans $\mathbb{N}$, est un temps d'arrêt si pour tout $k \in \mathbb{N}$, on a $[\tau \leqslant k] \in \mathcal{F}_k$.

Naturellement, on interprète τ comme un temps. Ainsi, τ est un temps d'arrêt si à l'instant k on est capable de certifier si τ a déjà eu lieu : l'évènement $[\tau \leqslant k]$.

Lemme 1.11. *Soit τ un temps d'arrêt adapté à $\mathcal{F}_k$, alors pour tout $k \in \mathbb{N}$ et $l < k$ on a :*

$$[\tau \leqslant l] \in \mathcal{F}_k \,.$$

En particulier, $[\tau = k]$ et $[\tau > k]$ sont des éléments de $\mathcal{F}_k$.

Démonstration. En effet, $[\tau \leqslant l] \in \mathcal{F}_l \subset \mathcal{F}_k$. De plus, $[\tau = k] = [\tau \leqslant k] \cap [\tau \leqslant k-1]^c \in \mathcal{F}_k$ et $[\tau > k] = [\tau \leqslant k]^c \in \mathcal{F}_k$. □

Exemple 1.12 (Exemple : premier passage)**.** Soit (X_k) une suite de v.a. à valeurs dans $\mathbb{R}$ et $\mathcal{F}_k := \sigma(X_0, \ldots, X_k)$. Soit $A \subset \mathbb{R}$ un sous-ensemble mesurable. Alors $\tau_A = \inf\{k \in \mathbb{N} : X_k \in A\}$ est un temps d'arrêt mesurant le premier passage de (X_k) par A.

Exemple 1.13 (Contre-exemple : dernier passage)**.** Dans le cadre de l'exemple précédent $\tau_A = \sup\{k \in \mathbb{N} : X_k \in A\}$ est en principe pas un temps d'arrêt. En effet, pour $k \in \mathbb{N}$,

$$[\tau_A \leqslant k] = [X_k \in A] \cap [X_{k+1} \notin A, X_{k+2} \notin A, \ldots]$$

est un événement qui dépend du futur et en principe n'appartient pas à $\mathcal{F}_k$.

Exemple 1.14 (Exemple : deux temps d'arrêts)**.** Soit T, S deux temps d'arrêts. Alors les variables $T + S$, $\max(T, S)$ et $\min(T, S)$ sont également des temps d'arrêts.

Démonstration. Il suffit d'écrire les évènements correspondants. Par exemple

$$[T + S \leqslant k] = \bigcup_{0 \leqslant i \leqslant k} [T \leqslant i] \cap [S \leqslant k - i] \,,$$

et chaque élément de l'union (et donc l'union aussi) est un élément de $\mathcal{F}_k$. De même

$$[\min(T, S) \leqslant k] = [T \leqslant k] \cup [S \leqslant k] \,.$$

Finalement,

$$[\max(T, S) \leqslant k] = [T \leqslant k] \cap [S \leqslant k] \,.$$

□

Étant donné un temps d'arrêt τ on peut définir la tribu $\mathcal{F}_\tau$ qui contient toutes les informations disponibles au temps τ (insistons sur le fait que τ est une variable aléatoire)/

Définition 1.15. Soit τ un temps d'arrêt. On définit $\mathcal{F}_\tau$ comme l'ensemble d'évènements $A \in \mathcal{F}$ t.q.

$$\forall k \in \mathbb{N} \,, \quad A \cap [\tau \leqslant k] \in \mathcal{F}_k \,.$$

Définition 1.16 (Processus arrêté)**.** Soit (X_k) une suite de v.a. et τ un temps d'arrêt. Le processus arrêté (X_k^τ) est défini comme

$$X_k^\tau = X_{k \wedge \tau} \,,$$

où $l \wedge k = \min(l, k)$.

On a la proposition suivante.

Proposition 1.17. *Soit (X_k) une martingale (resp. sous-martingale, sur-martingale) et τ un temps d'arrêt, la suite (X_k^τ) est alors également une martingale (resp. sous-martingale, sur-martingale).*

Démonstration. 1. (X_k^τ) est un processus adapté. Remarquons que

$$\begin{aligned} X_k^\tau &= X_k \mathbb{1}_{k\leqslant\tau} + X_\tau \mathbb{1}_{\tau<k} \\ &= X_k \mathbb{1}_{k\leqslant T} + \sum_{i=0}^{k-1} X_i \mathbb{1}_{\tau=i} \,. \end{aligned}$$

Tous les événements dans les indicatrices étant dans $\mathcal{F}_k$ on trouve que (X_k^τ) est adapté à $(\mathcal{F}_k)$.
2. Pour tout k, X_k^τ est intégrable. En effet, on a

$$\mathbb{E}[|X_k^\tau|] \leqslant \mathbb{E}[|X_k|] + \mathbb{E}[|X_\tau \mathbb{1}_{\tau\leqslant k}|] \leqslant \mathbb{E}[|X_k|] + \sum_{i=0}^{k} \mathbb{E}[|X_i|] < +\infty \,.$$

Égalité de l'espérance conditionnelle. On va prouver seulement le cas où (X_k) est une martingale, les deux autres cas se prouvent similairement.

Remarquons que $[k+1 \leqslant \tau] = [\tau \leqslant k]^c \in \mathcal{F}_k$. Ainsi,

$$\begin{aligned} \mathbb{E}[X_{k+1}^\tau|\mathcal{F}_k] &= \mathbb{1}_{k+1\leqslant\tau}\mathbb{E}[X_{k+1}|\mathcal{F}_k] + \mathbb{E}[X_\tau \mathbb{1}_{\tau\leqslant k}|\mathcal{F}_k] \\ &= \mathbb{1}_{k+1\leqslant\tau} X_k + \mathbb{E}\left[\sum_{i=0}^{k} X_k \mathbb{1}_{\tau=i}|\mathcal{F}_k\right] \\ &= \mathbb{1}_{k+1\leqslant\tau} X_k + X_\tau \mathbb{1}_{\tau\leqslant k} \\ &= \mathbb{1}_{k<\tau} X_k + X_\tau \mathbb{1}_{\tau\leqslant k} = X_k^\tau \,. \end{aligned}$$

□

Dans certains cas la relation entre les espérances conditionnelles d'une (sur)-martingale peut s'étendre au cas où on remplace l'indice k par un temps d'arrêt.

Proposition 1.18. *Soit (X_k) une martingale (respectivement sur-martingale, respectivement sous-martingale), $N \in \mathbb{N}$ et T, S deux temps d'arrêts t.q. $T \leqslant S < N$. Alors*

$$\mathbb{E}[X_S|\mathcal{F}_T] = X_T \qquad \textit{(respectivement } \leqslant X_T,\ \geqslant X_T\textit{)}.$$

Démonstration. Montrons la proposition dans le cas des martingales (le cas des sur-martingale se traite de manière similaire).
1. Montrons d'abord que pour tout temps d'arrêt $T \leqslant N$ on a $\mathbb{E}[X_N|\mathcal{F}_T] = X_T$. En effet, soit $A \in \mathcal{F}_T$ on peut écrire :

$$\begin{aligned} \mathbb{E}[X_N \mathbb{1}_A] &= \mathbb{E}\left[\sum_{i=0}^{N} X_N \mathbb{1}_{A\cap[T=i]}\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\sum_{i=0}^{N} X_N \mathbb{1}_{A\cap[T=i]}|\mathcal{F}_i\right]\right] \\ &= \mathbb{E}\left[\mathbb{1}_{A\cap[T=i]}\mathbb{E}[\sum_{i=0}^{N} X_N|\mathcal{F}_i]\right] \\ &= \mathbb{E}\left[\sum_{i=0}^{N} X_i \mathbb{1}_{A\cap[T=i]}\right] = X_T \,. \end{aligned}$$

2. Cas général. D'après la proposition 1.17 le processus arrêté $(X^S) = (X_{S\wedge k})$ est une martingale. Donc en utilisant la première partie de la preuve on trouve

$$\mathbb{E}[X_N^S|\mathcal{F}_T] = X_T^S = X_T \,.$$

Or comme $S \leqslant N$ on a bien $X_N^S = X_{S\wedge N} = X_S$. Le résultat est prouvé. □

Remarque 1.19. Dans la proposition précédente la condition de bornitude est importante. En effet, considérons la marche aléatoire sur $\mathbb{Z}$: $X_k = \sum_{i=1}^k \xi_i$, avec (ξ_k) i.i.d., intégrables, et de moyenne nulle et $\xi_0 \neq 0$. Soit $m \in \mathbb{Z}$, et $T = \inf_{k\in\mathbb{N}}\{X_k : X_k = m\}$, on peut par ailleurs montrer que $T < +\infty$. Cependant, on a bien sûr $\mathbb{E}[X_T] = 1 \neq \mathbb{E}[X_0] = 0$. Ceci montre que le temps d'arrêt T n'est pas borné par une constante de manière déterministe. En passant par l'analogie des jeux du hasard, si notre stratégie est de jouer jusqu'au moment où on gagne m euros alors en moyenne on gagnera m euros. On présuppose cependant d'avoir un capital infini à investir dans le jeu !

1.3 Inégalités de Doob

Rappelons que pour une variable aléatoire X, disons intégrable autant de fois qu'il le faut, on a toujours $\mathbb{P}(|X| \geqslant \varepsilon) \leqslant \frac{\mathbb{E}[|X|^p]}{\varepsilon^p}$. Dans le cas où (X_k) est une (sous)-martingale on peut renforcer cette inégalité. Ce sont les inégalités dites de Doob.

Proposition 1.20 (Inégalité de Doob)**.** *Soit (X_k) une **sous-martingale**, $\varepsilon > 0$ et notons $X_k^* = \sup_{l\leqslant k} X_l$ alors on a*

$$\mathbb{P}(X_k^* \geqslant \varepsilon) \leqslant \frac{\mathbb{E}\left[\mathbb{1}_{X_k^*\geqslant\varepsilon} X_k\right]}{\varepsilon} \leqslant \frac{\mathbb{E}\left[\mathbb{1}_{X_k^*\geqslant\varepsilon} |X_k|\right]}{\varepsilon}$$

Démonstration. La deuxième inégalité est triviale, montrons la première. Définissons le temps d'arrêt $\tau = \inf\{k \in \mathbb{N} : X_k \geqslant \varepsilon\}$. Remarquons que $[X_k^* \geqslant \varepsilon] = [\tau \leqslant k]$. On a alors

$$\varepsilon\mathbb{P}(X_k^* \geqslant \varepsilon) = \mathbb{E}[\varepsilon\mathbb{1}_{X_k^*\geqslant\varepsilon}] = \mathbb{E}[\varepsilon\mathbb{1}_{\tau\leqslant k}] \leqslant \mathbb{E}[X_\tau\mathbb{1}_{\tau\leqslant k}] = \mathbb{E}[X_{\tau\wedge k}\mathbb{1}_{\tau\leqslant k}]$$

De plus, si on note $T = \tau \wedge k$, T est un temps d'arrêt (voir l'exemple 1.14) et $T \leqslant k$. On peut donc appliquer la proposition 1.18 :

$$X_{\tau\wedge k} \leqslant \mathbb{E}[X_k|\mathcal{F}_{\tau\wedge k}] .$$

En combinant les deux équations on obtient

$$\varepsilon\mathbb{P}(X_k^* \geqslant \varepsilon) \leqslant \mathbb{E}[X_{\tau\wedge k}\mathbb{1}_{\tau\leqslant k}] \leqslant \mathbb{E}[\mathbb{1}_{\tau\leqslant k}\mathbb{E}[X_k|\mathcal{F}_{\tau\wedge k}]] .$$

Finalement, comme $[\tau \leqslant k] \in \mathcal{F}_{\tau\wedge k}$ et $[\tau \leqslant k] = [X_k^* \geqslant \varepsilon]$ on obtient

$$\varepsilon\mathbb{P}(X_k^* \geqslant \varepsilon) \leqslant \mathbb{E}[\mathbb{1}_{\tau\leqslant k} X_k] = \mathbb{E}\left[\mathbb{1}_{X_k^*\geqslant\varepsilon} X_k\right] .$$

□

On peut améliorer ce résultat pour les (sous)-martingales L^p-intégrables.

Proposition 1.21 (Inégalité de Doob, L^p)**.** *Soit (X_k) une martingale ou une **sous-martingale positive**. Soit $p \geqslant 1$, t.q. pour tout $k \in \mathbb{N}$, $\mathbb{E}[|X_k|^p] < +\infty$. Notons $|X|_k^* = \sup\{|X_l| : l \leqslant k\}$.*

1. *Pour tout $\varepsilon > 0$,*

$$\varepsilon^p\mathbb{P}(|X|_k^* \geqslant \varepsilon) \leqslant \mathbb{E}[|X_k|^p] .$$

Démonstration. Il suffit de remarquer que la fonction $x \mapsto |x|^p$ est convexe, donc par la proposition 1.8 la suite $(|X_k|^p)$ est une sous-martingale. On applique alors la proposition 1.20. □

1.4 Décomposition de Doob d'un processus stochastique

La décomposition suivante montre que tout processus stochastique a une composante « martingale ». D'où l'importance de l'étude de ces dernières !

Avant de l'énoncer rappelons qu'un processus (B_k) est dit prévisible par rapport à une filtration $(\mathcal{F}_k)$ si pour tout k, B_k est $\mathcal{F}_{k-1}$ mesurable.

Proposition 1.22. *Soit (X_k) une suite de v.a. adaptée à la filtration $(\mathcal{F}_k)$ et intégrable. Alors il existe un processus (B_k), prévisible, et (M_k) une martingale t.q. $0 = M_0 = B_0$ et pour tout $k \in \mathbb{N}$:*

$$X_k = X_0 + B_k + M_k \,.$$

De plus cette décomposition est unique.

Démonstration. Dans cette preuve on notera $\Delta X_k = X_k - X_{k-1}$ (et de même pour ΔB_k et ΔM_k).

Notons que si une telle décomposition existe on a forcément $\Delta B_k = \Delta X_k - \Delta M_k$ et comme B_k est prévisible et M_k est une martingale on trouve $\mathbb{E}[\Delta B_k|\mathcal{F}_k] = \Delta B_k = \mathbb{E}[\Delta X_k|\mathcal{F}_k]$. Ceci donne l'idée de la décomposition :

$$X_k = X_0 + \sum_{i=0}^{k} \mathbb{E}[\Delta X_i|\mathcal{F}_{i-1}] + \sum_{i=0}^{k} (\Delta X_i - \mathbb{E}[\Delta X_i|\mathcal{F}_{i-1}]) \ ,$$

et cette décomposition convient.

Supposons maintenant qu'il existe une autre décomposition $X_k = X_0 + B'_k + M'_k$. On trouve alors, pour tout $k \in \mathbb{N}$,

$$B'_k - B_k = M'_k - M_k \,.$$

Donc le processus $(M'_k - M_k)$ est une martingale prévisible, ce qui implique pour tout $k > 0$,

$$M'_k - M_k = \mathbb{E}[M'_k - M_k|\mathcal{F}_{k-1}] = M'_{k-1} - M_{k-1} \,.$$

En réitérante le procédé on trouve que pour tout $k \in \mathbb{N}$, $M'_k = M_k$ et donc $B_k - B'_k$. □

Remarque 1.23. Avec la décomposition de Doob on voit que (X_k) est une martingale ssi $B_k \equiv 0$ est une sous-martingale si (B_k) est croissant et une sur-martingale si (B_k) est décroissant.

1.5 Convergence des martingales

Nous allons enfin aborder les résultats fondamentaux de la théorie des martingales — sous quelles conditions (et sous quel type de convergence) nous pouvons affirmer que X_k converge vers une v.a. X_∞.

Théorème 1.24. *Soit (X_k) une sur-martingale (ou une sur-martingale) t.q.* $\sup \mathbb{E}[|X_k|] < +\infty$*. Il existe une variable X_∞ t.q.*

$$X_k \to X_\infty \textit{ presque sûrement.}$$

Théorème 1.25. *Soit $p > 1$ et (X_k) une martingale t.q.* $\sup \mathbb{E}[|X_k|^p] < +\infty$*. Alors, il existe une variable aléatoire X_∞ t.q.*

$$X_k \to X_\infty \textit{ presque sûrement } \textbf{et} \textit{ en norme } L^p \,.$$

Remarque 1.26. Attention Le théorème 1.25 s'applique au cas où $p > 1$. Le cas $p = 1$ est un peu compliqué comme le montre l'exercice I.7.

Nous allons démontrer le théorème 1.25 dans le cas plus simple où la martingale est de carré intégrable : pour tout $k \in \mathbb{N}$, $\mathbb{E}[X_k^2] < +\infty$.

Remarque 1.27 (Rappel : L^2 est un espace de Hilbert)**.** Notons L^2 l'ensemble des v.a. de carré intégrable : $Z \in L^2 \Leftrightarrow \mathbb{E}[Z^2] < +\infty$. Alors $X, Y \mapsto \mathbb{E}[XY]$ est un produit scalaire pour cet espace. Ce produit scalaire induit une norme $\|Z\|_{L^2} = \mathbb{E}[|Z|^2]$ et L^2 muni de cette norme est complet : toute suite de Cauchy [2] converge. En résumé : L^2 muni de ce produit scalaire est un espace de Hilbert.

2. (Z_k) est une suite de Cauchy, si pour tout $\varepsilon > 0$, il existe $k_0 \in \mathbb{N}$, t.q. pour tout $l, k \geqslant k_0$, $\|Z_k - Z_l\| \leqslant \varepsilon$.

L'idée principale pour démontrer le théorème 1.25 est alors d'écrire X_k comme

$$X_k = X_0 + \sum_{i=0}^{k} \Delta X_i \,,$$

avec $\Delta X_i := X_i - X_{i-1}$ et par convention $\Delta X_0 = X_0$.

On peut alors remarquer que pour $j \neq i$, ΔX_j est orthogonale à ΔX_i (dans L^2 muni du produit scalaire énoncé ci-dessous). En effet, sans perte de généralité supposons que $j < i$

$$\mathbb{E}[\Delta X_i \Delta X_j] = \mathbb{E}[\mathbb{E}[\Delta X_i \Delta X_j | \mathcal{F}_{i-1}]] = \mathbb{E}[\Delta X_j \mathbb{E}[\Delta X_i | \mathcal{F}_i]] = 0 \,.$$

En particulier, par lé théorème de Pythagore, on trouve :

$$\mathbb{E}[|X_k|^2] = \sum_{i=0}^{k} \mathbb{E}[(\Delta X_i)^2] \,, \tag{1.1}$$

et $\mathbb{E}[|X_k|^2]$ est une suite croissante. En particulier, cette dernière converge si elle est bornée.

Preuve de la convergence L^2. Comme $\sup_k \mathbb{E}[|X_k|^2] < +\infty$ on trouve par l'équation (1.1), que $(\mathbb{E}[|X_k|^2])$ converge et est en particulier de Cauchy. Fixons, $k \in \mathbb{N}$, on a alors pour tout $l > 0$:

$$\mathbb{E}[|X_{k+l} - X_k|^2] = \mathbb{E}\left[\left|\sum_{i=k+1}^{k+l} \Delta X_i\right|^2\right] = \sum_{i=k+1}^{k+l} \mathbb{E}[(\Delta X_i)^2] = \mathbb{E}[X_{k+l}^2] - \mathbb{E}[X_k^2] \to_{k\to+\infty} 0 \,,$$

où la deuxième égalité vient de l'orthogonalité des (ΔX_k) et la convergence vient du fait que $(\mathbb{E}[|X_k|^2])$ est de Cauchy. Donc (X_k) est de Cauchy dans L^2 qui est complet ce qui montre que (X_k) converge. ☐

Avant de prouver la convergence presque sûre établissons un lemme préliminaire.

Lemme 1.28. *Soient les variables aléatoires (X_k) et X_∞. Alors $X_k \to X_\infty$ p.s. ssi $(\sup_{l \geqslant k} |X_l - X_\infty|)_k$ converge vers 0 en probabilité. En d'autres mots, ssi pour tout $\varepsilon > 0$, on a :*

$$\mathbb{P}(\sup_{l \geqslant k} |X_l - X_\infty| \geqslant \varepsilon) \to 0.$$

Démonstration. Remarquons que l'événement $A = [X_k \to X_\infty]$ s'écrit comme

$$A = \bigcap_{m>1} \bigcup_{k\in\mathbb{N}} [\sup_{l\geqslant k} |X_l - X_\infty| \leqslant 1/m] := \bigcap_{m>1} A_m \,.$$

La suite d'événements (A_m) est une suite décroissante et donc $\mathbb{P}(A) = \lim_{m\to\infty} \mathbb{P}(A_m)$. Donc $\mathbb{P}(A) = 1 \Leftrightarrow \lim_{m\to+\infty} \mathbb{P}(A_m) = 1$ et cette dernière condition est équivalent au fait que $\mathbb{P}(A_m) = 1$ pour tout $m > 1$ (car A_m est une suite décroissante).

Or, $A_m = \bigcup_k A_{m,k}$ avec $A_{m,k} := [\sup_{l\geqslant k} |X_l - X_\infty| \leqslant 1/m]$ et $(A_{m,k})_k$ est une suite croissante. Donc pour tout $m > 1$,

$$\mathbb{P}(A) = 1 \Leftrightarrow \forall m \geqslant 1 \,, \lim_{k\to+\infty} \mathbb{P}(A_{m,k}) = 1 \,.$$

La dernière condition est exactement celle de la convergence de X_k vers X_∞ en probabilité. ☐

Preuve de la convergence presque sûre. Fixons $k \in \mathbb{N}$, X_∞ la limite L^2 de (X_k) et notons que par l'inégalité de Tchebychev on a :

$$\begin{aligned}\mathbb{P}(\sup_{l\geqslant k} |X_l - X_\infty| \geqslant \varepsilon) &\leqslant \frac{1}{\varepsilon^2}\mathbb{E}[\sup_{l\geqslant k} |X_l - X_\infty|^2] \leqslant \frac{2}{\varepsilon^2}\mathbb{E}[|X_k - X_\infty^2] + \frac{2}{\varepsilon^2}\mathbb{E}[\sup_{l\geqslant k} |X_l - X_k|^2] \\ &= \frac{2}{\varepsilon^2}\mathbb{E}[|X_k - X_\infty^2] + \frac{2}{\varepsilon^2} \lim_{N\to+\infty} \mathbb{E}[\sup_{k\leqslant l\leqslant N} |X_l - X_k|^2] \,,\end{aligned} \tag{1.2}$$

où la dernière égalité est obtenue par le théorème de la convergence monotone. Or d'une part on a :

$$\frac{1}{\varepsilon^2}\mathbb{E}[|X_k - X_\infty^2] \to_{k\to+\infty} 0 \tag{1.3}$$

par le fait que $X_k \to X_\infty$ dans L^2. D'autre part, notons que $(|X_l - X_k|)_{l\geqslant k}$ est une sous-martingale par rapport à la filtration $(\mathcal{F}_l)_{l\geqslant k}$. Donc d'après la proposition 1.21, appliquée avec $p = 2$, on trouve :

$$\lim_{N\to+\infty} \mathbb{E}[\sup_{k\leqslant l\leqslant N} |X_l - X_k|^2] \leqslant 4 \lim_{N\to+\infty} \mathbb{E}[|X_N - X_k|^2] = 0\,.$$

Ainsi,

$$\mathbb{P}(\sup_{l\geqslant k} |X_l - X_\infty| \geqslant \varepsilon) \leqslant \frac{2}{\varepsilon^2}\mathbb{E}[|X_k - X_\infty^2]\,,$$

et la dernière expression tends vers 0 par l'équation 1.3. Donc on peut appliquer le lemme 1.28 pour aboutir au résultat. □

Nous admettrons les démonstrations des théorèmes 1.24 et 1.25 dans le cadre général.

1.6 Optionnel : démonstration du théorème 1.24

L'idée de la preuve se comprend en interprétant (X_k) comme le prix d'une action. Dans ce cas-là, une stratégie raisonnable pour gagner de l'argent serait de fixer $a < b$ et d'acheter cette action quand son prix est inférieur à a puis d'en vendre quand son prix est supérieur à b. Ainsi, si le prix de l'action fluctue infiniment autour de l'intervalle $[a, b]$ on gagnera forcément de l'argent. Nous allons voir que sous les hypothèses du théorème 1.24 ce n'est justement pas possible.

Du point de vue mathématique la preuve du théorème se sert de la caractérisation suivante d'une suite convergente.

Lemme 1.29. *Soit $(x_k) \in \mathbb{R}^{\mathbb{N}}$, alors (x_k) converge ssi pour tout $a, b \in \mathbb{Q}$, avec $a < b$ on a que*

$$u_{a,b} := \textit{le nombre maximal d'intervalles } [l_1, l_2] \textit{ t.q. } l_1 < l_2 \textit{ et } x_{l_1} \leqslant a \textit{ et } x_{l_2} \geqslant b$$

est fini.

Pour chaque a, b, $u_{a,b}$ représente le nombre de passages croissants d'un niveau inférieur à a à un niveau supérieur à b.

Démonstration. En effet, supposons que $x_k \to x$ et soit $a < b$. Alors, il existe $\varepsilon > 0$ t.q. $[x-\varepsilon, x+\varepsilon] \not\subset [a, b]$ et pour tout $k_0 \in \mathbb{N}$, $x \in [x - \varepsilon, x + \varepsilon]$. On voit donc bien que $u_{a,b}$ est fini.

Inversement, soit $a = \underline{\lim} x_k$ et $b = \overline{\lim} x_k$ [3]. Si $a < b$ alors il existe ε t.q. $a' = a + \varepsilon < b' = b - \varepsilon$, avec $a', b' \in \mathbb{Q}$, et par hypothèse $u_{a',b'}$ est fini. Ceci qui veut dire qu'à partir d'un certain rang k_0 ou bien $\inf_{k\geqslant k_0} x_k \geqslant a' > a$ ou bien $\sup_{k\geqslant k_0} x_k \leqslant b' < b$. Les deux possibilités contredisent le fait que $a = \underline{\lim} x_k < b = \overline{\lim} x_k$. Ainsi, $a = b$ ce qui veut dire que x_k converge vers un unique point. □

Soit (X_k) une sur-martingale et fixons $a < b$ deux réels. Définissons la suite des temps d'arrêts $\tau_0 = 0$, $\sigma_{k+1} = \inf\{l > \tau_k : X_l \leqslant a\}$, $\tau_{k+1} = \inf\{l \geqslant \sigma_{k+1} : X_l \geqslant b\}$. Alors pour tout k, la variable aléatoire

$$U_k^{a,b} := \max\{l \in \mathbb{N} : \tau_l \leqslant k\}$$

représente le nombre de passages croissants de l'intervalle $[a, b]$ avant l'itération k. L'idée de la preuve du théorème 1.24 est d'alors de borner cette quantité, uniformément en k et d'appliquer le lemme 1.29.

Proposition 1.30. *Soit (X_k) une sur-martingale et $a < b$ deux réels. Alors*

$$\mathbb{E}[U_k^{a,b}] \leqslant \frac{\mathbb{E}[\max(a - X_k, 0)]}{b - a}\,.$$

3. $\underline{\lim} x_k = \lim_{k\to+\infty} \inf_{l\geqslant k} x_k$ est la plus petite valeur d'adhérence de (x_k). De même, $\overline{\lim} x_k = \lim_{k\to+\infty} \sup_{l\geqslant k} x_k$ est la plus grande valeur d'adhérence de (x_k).

Démonstration. Écrivons formellement la stratégie d'achat d'une action décrite au début de cette section. Soit (H_k) défini comme suit.

$$H_k = \begin{cases} 1 & \text{si } \sigma_l + 1 \leqslant k \leqslant \tau_l, \text{ pour un certain } l \in \mathbb{N}, \\ 0 & \text{sinon.} \end{cases}$$

Remarquons que pour chaque k, H_k est $\mathcal{F}_{k-1}$ mesurable, car

$$[H_k = 1] = \bigcup_{l \in \mathbb{N}} [\sigma_l \leqslant k-1] \cap [\tau_l > k-1] \,,$$

et chaque événement dans l'union est $\mathcal{F}_{k-1}$-mesurable.

Définissons alors $Z_k = \sum_{i=0}^{k} H_i(X_i - X_{i-1})$, cette variable aléatoire représente exactement le profit réalisé par la stratégie décrite en début de la section. On peut vérifier (ou appliquer directement la proposition 1.9) que (Z_k) est une sur-martingale. En notant $l = U_k^{a,b}$ on trouve :

$$Z_k = \sum_{i=1}^{l} (X_{\tau_i} - X_{\sigma_i}) + (X_k - X_{\sigma_l}) \geqslant l(b-a) + (X_k - a) \,.$$

Or (Z_k) étant une sur-martingale on trouve $\mathbb{E}[Z_k] \leqslant \mathbb{E}[Z_0] = 0$, ce qui montre que

$$\mathbb{E}[U_k^{a,b}] \leqslant \frac{\max(a - X_k, 0)}{b-a} \,.$$

□

Nous pouvons maintenant terminer la preuve du théorème. En effet, pour tout $a, b \in \mathbb{R}$, $a < b$ on a $U_k^{a,b} \nearrow_{k \to \infty} U^{a,b}$, où $U^{a,b}$ est le nombre de passages croissants de (X_k) par l'intervalle $[a, b]$. En utilisant la proposition précédente on a aussi

$$\mathbb{E}[U_k^{a,b}] \leqslant \frac{\sup_{k \in \mathbb{N}} \mathbb{E}[X_k] + a}{b-a} < +\infty \,.$$

Donc par le théorème de la convergence monotone on a :

$$\mathbb{E}[U^{a,b}] < +\infty \,.$$

Donc $U^{a,b}$ est fini presque sûrement et donc (X_k) converge presque sûrement.

Chapitre 2

Martingales à temps continu

2.1 Processus à temps continu, filtrations et mesurabilité

On se donne $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilité.

Définition 2.1 (Processus aléatoire)**.** Un *processus aléatoire* indexé par un ensemble d'indices T et à valeurs dans un espace mesurable $(E, \mathcal{E})$ est une famille $(X_t)_{t\in T}$ de variables aléatoires définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ et à valeurs dans $(E, \mathcal{E})$, c'est à dire, pour tout $t \in T$ l'application

$$X_t : \begin{array}{rcl} (\Omega, \mathcal{A}, \mathbb{P}) & \longrightarrow & (E, \mathcal{E}) \\ \omega & \longmapsto & X_t(\omega) \end{array}$$

est mesurable. On parle de

- *processus discret* si $T = \mathbb{N}$ (martingales discrètes, chaînes de Markov, ...) ;
- *processus à temps continu* si $T = \mathbb{R}_+$ (mouvement Brownien, procéssus de Markov à temps continu, processus de Lévy, ...) ;
- *champs aléatoire* si $T = \mathbb{R}^2$ (champs libre Gaussien, ...).

On s'interessera dans cette section aux procéssus aléatoires à temps continu et à valeurs réelles, et on prendra donc $T = \mathbb{R}_+$ et $(E, \mathcal{E}) = (\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$. Pour un $\omega \in \Omega$ fixé on dit que la fonction $t \mapsto X_t(\omega)$ est une *trajectoire* du processus. Un processus $(X_t)_{t\in\mathbb{R}}$ est dit *continu* si ses trajectoires sont continues (idem pour continu à droite et continu à gauche).

Définition 2.2 (Filtration)**.** On dit que $\mathcal{F} = (\mathcal{F}_t)_{t\in\mathbb{R}}$ est une *filtration* si $\mathcal{F}$ est une famille croissante (i.e. $\mathcal{F}_t \subset \mathcal{F}_{t'}$ pour $t \leqslant t'$) de sous tribus de $\mathcal{A}$.

Il faut comprendre $(\mathcal{F}_t)$ comme "l'information disponible au temps t" : plus le temps croît ($s \leqslant t$), plus on a d'informations ($\mathcal{F}_s \subset \mathcal{F}_t$).

Définition 2.3 (Mesurabilité)**.** On dit que le processus $(X_t)_{t\in\mathbb{R}_+}$ est

- *mesurable* si $(t, \omega) \mapsto X_t(\omega)$ est $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{A}$-mesurable ;
- *adapté* à la filtration $(\mathcal{F}_t)$ si pour tout $t \in \mathbb{R}_+$, $\omega \mapsto X_t(\omega)$ est $\mathcal{F}_t$-mesurable ;
- *progressif* ou *progressivement mesurable* si $[0, t] \times \Omega \ni (t, \omega) \mapsto X_t(\omega)$ est $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$-mesurable.

Proposition 2.4.
- *Un processus continu à droite (ou à gauche) est mesurable.*
- *Tout processus progressif est adapté et mesurable.*
- *Tout processus adapté et continu à droite (ou continu à gauche) est progressif.*

Démonstration. Admis. Voir [1, §3.1]. □

Remarque 2.5 (Filtration canonique augmentée)**.** On dit que $\mathcal{F} = (\mathcal{F}_t)$ est la *filtration canonique* ou *naturelle* associée au processus (X_t) si

$$\mathcal{F}_t = \sigma(X_s, s \leqslant t).$$

On note $\mathcal{F}_t^+ = \cap_{t<s}\mathcal{F}_s$. Et en rêgle générale, on considèrera la filtration $(\widetilde{\mathcal{F}}_t)$ définie par

$$\widetilde{\mathcal{F}}_t = \mathcal{F}_t^+ \vee \{A \in \mathcal{A}, \mathbb{P}(A) = 0\}$$

il s'agit de l'*augmentation habituelle* de la filtration naturelle qui est complèté par les ensembles négligeables et continue à droite ($\widetilde{\mathcal{F}}_t = \widetilde{\mathcal{F}}_t^+$). Par abus de notation on notera $\mathcal{F}_t$ au lieu de $\widetilde{\mathcal{F}}_t$. Muni de cette tribu un processus est évidemment adapté. Si il est de plus coninu (à droite ou à gauche) il est donc aussi progressif pas la Proposition 2.4.

Dans ce qui suit, tous les processus à temps continu que nous considérerons seront progressivement mesurables sur un espace de probabilité filtré

$$(\Omega, \mathcal{A}, (\mathcal{F}_t, t \geqslant 0), \mathbb{P}).$$

Cela permettra entre autres de considérer des événements s'exprimant à l'aide de toute la trajectoire du processus.

2.2 Temps d'arrêt

Définition 2.6 (Temps d'arrêt et tribu associée)**.** La variable aléatoire $\tau : \Omega \to [0, \infty]$ est un *temps d'arrêt* pour la filtration $(\mathcal{F}_t)$ si $\forall t \geqslant 0$, $\{\tau \leqslant t\} \in \mathcal{F}_t$. La tribu associée à ce temps d'arrêt est définie par

$$\mathcal{F}_\tau = \{A \in \mathcal{A}; \forall t \geqslant 0,\ A \cap \{\tau \leqslant t\} \in \mathcal{F}_t\}.$$

L'essentiel des propriétés élémentaires que nous utiliserons sur les temps d'arrêt se résume à cette proposition.

Proposition 2.7. *1. Si τ est un temps d'arrêt, τ est $\mathcal{F}_\tau$-mesurable.*

2. *Si τ est un temps d'arrêt fini $\mathbb{P}$-p.s. et X $\mathcal{F}$-adapté alors X_τ est $\mathcal{F}_\tau$-mesurable.*
3. *Si τ et τ' sont deux temps d'arrêt tels que $\tau \leqslant \tau'$ $\mathbb{P}$-p.s., alors $\mathcal{F}_\tau \subset \mathcal{F}_{\tau'}$.*
4. *Si τ et τ' sont deux temps d'arrêt, alors, $\tau \wedge \tau'$ l'est aussi. En particulier pour tout T déterministe, $\tau \wedge T$ est un temps d'arrêt.*

Démonstration. Admis. Voir [1, §3.2]. □

2.3 Martingale, convergence et théorème d'arrêt

Les résultats pour les martingales en temps continu sont les mêmes qu'en temps discret. Les preuves sont cependant plus complexes à cause des questions de mesurabilités et seront donc la plupart du temps admises. Les hypothèses de continuité faite sur les filtrations et les processus, permettent un certain nombre de simplifications.

Définition 2.8 (Martingales, sous-martingales et sur-martingales)**.** Un processus (M_t) est une *martingale* si il est adapté, si $M_t \in L^1$ (i.e. $\forall t \geqslant 0$, $\mathbb{E}[|M_t|] < \infty$) et si

$$\forall t \geqslant s, \quad M_s = \mathbb{E}(M_t|\mathcal{F}_s) \tag{2.1}$$

La définition est équivalente pour les *sous et sur-martingales* en remplaçant (2.1) par

— $\forall t \geqslant s, \quad M_s \leqslant \mathbb{E}(M_t|\mathcal{F}_s)$ pour les sous-martingales,

— $\forall t \geqslant s, \quad M_s \geqslant \mathbb{E}(M_t|\mathcal{F}_s)$ pour les sur-martingales.

En particulier, pour les martingales l'espérance est constante :

$$\mathbb{E}(M_t) = \mathbb{E}(M_0).$$

Théorème 2.9 (Convergence des surmartingales bornées dans L^1)**.** *Soit M une surmartingale à trajectoires continues à droite et bornée dans L^1. Alors, il existe une variable aléatoire $M_\infty \in L^1$ telle que*

$$\lim_{t\to\infty} M_t = M_\infty, \; p.s.$$

Démonstration. Admis. Voir [1, §3.4]. □

Définition 2.10 (Martingale fermée)**.** Une martingale (M_t) est *fermée* si il existe $Z \in L^1$ tel que pour tout $t \geqslant 0$,

$$M_t = E(Z|\mathcal{F}_t).$$

L'ensemble du processus d'une martingale fermée est ainsi déterminé par une valeur terminale à l'horizon infini.

Avant d'énoncer le théorème suivant, on rappelle le définition d'uniforme intégrabilité.

Définition 2.11 (Uniforme intégrabilité)**.** Une famille $(M_t)_{t\geqslant}$ de variables aléatoires est dite uniformément intégrable si :

$$\lim_{K\to\infty} \sup_{t\geqslant 0} \mathbb{E}\left(|M_t|\, 1_{|M_t|>K}\right) = 0.$$

Une telle famille est nécessairement bornée dans L^1.

Théorème 2.12 (Théorème de convergence)**.** *Soit M une martingale continue. Les assertions suivantes sont équivalentes :*

1. *M est une martingale fermée par M_∞ ;*
2. *M converge p.s. et dans L^1 vers M_∞ ;*
3. *M est uniformément intégrable.*

Démonstration. Admis. Voir [1, §3.4]. □

Il existe des martingales qui ne sont pas fermé ou pas uniformément intégrables, autrement dit qui ne convergent pas. Un exemple typique de martingale non convergente est le mouvement brownien, que nous verrons au chapitre suivant.

Théorème 2.13 (Théorème d'arrêt de Doob)**.** *Soit M une continue et S et T deux temps d'arrêt tels que $S \leqslant T$ p.s. Supposons que M soit uniformément intégrable ou que les temps d'arrêt soient bornés par une constante finie déterministe. Alors M_T est une v. a. intégrable et*

$$M_S = E(M_T|\mathcal{F}_S).$$

Démonstration. Admis. Voir [1, §3.4]. □

On a le même type de résultat avec des sous-martingales mais dans ce cas, l'égalité $=$ doit être remplacée par l'inégalité $\leqslant$.

Le corrolaire suivant nous dit qu'une martingale arrété est encore une martingale.

Corolaire 2.14 (Martingale arrété)**.** *Si M_t est une martingale continue à droite et T un temps d'arrêt alors $M_{t\wedge T}$ est encore une martingale. Si M_t est uniformément intégrable on a de plus*

$$M_{t\wedge T} = E(M_T|\mathcal{F}_t).$$

Démonstration. Admis. Voir [1, §3.4]. □

2.4 Variation quadratique

Le variation quadratique est une notion utilisée dans la construction de l'intégrale stochastique par rapport à une martingale. On commence par la définir simplement dans le cas discret.

Proposition 2.15 (Variation quadratique, martingale discrète)**.** *Soit (M_n) une martingale discrète de carré intégrable, c'est-à-dire que pour tout n, $\mathbb{E}(M_n^2)$ est fini. On définit la variation quadratique de M pour tout $n \geqslant 0$,*

$$\langle M \rangle_n := \sum_{k=1}^{n} \mathbb{E}\left[(M_k - M_{k-1})^2 \middle| \mathcal{F}_{k-1} \right].$$

Alors, $\langle M \rangle_n$ est l'unique processus croissant et prévisible (i.e. M_n est $\mathcal{F}_{n-1}$-mesurable), tel que $\langle M \rangle_0 = 0$ et tel que

$$M_n^2 - \langle M \rangle_n$$

est une martingale.

Démonstration. La croissance du processus est évidente et le fait qu'il soit prévisible. On cherche maintenant à vérifier que $(M_n^2 - \langle M \rangle_n)$ est bien une martingale. Le processus $\langle M \rangle_n$ est $\mathcal{F}_n$ mesurable. On a

$$\begin{aligned}
\mathbb{E}(M_{n+1}^2 - \langle M \rangle_{n+1} | \mathcal{F}_n) &= \mathbb{E}((M_n + (M_{n+1} - M_n))^2 - \langle M \rangle_n - (M_{n+1} - M_n)^2 | \mathcal{F}_n) \\
&= \mathbb{E}(M_n^2 + 2(M_{n+1} - M_n) M_n - \langle M \rangle_n | \mathcal{F}_n) \\
&= M_n^2 - \langle M \rangle_n + 2 M_n \mathbb{E}(M_{n+1} - M_n | \mathcal{F}_n) \\
&= M_n^2 - \langle M \rangle_n.
\end{aligned}$$

L'unicité découle du fait que si A_n et B_n sont prévisibles partant de 0 et tels que $M_n^2 - A_n$ et $M_n^2 - B_n$ sont des martingales, alors $A_n - B_n$ est une martingale prévisible et est donc constante (car $A_n - B_n = \mathbb{E}(A_n - B_n | \mathcal{F}_{n-1}) = A_{n-1} - B_{n-1}$) et égale à $A_0 - B_0 = 0$. Notons que l'existence et l'unicité peuvent aussi être déduite du théorème de décomposition de Doob appliqué à M_n^2 qui est une sous-martingale par l'inégalité de Jensen. □

Théorème et définition 2.16 (Variation quadratique, martingale continue)**.** *On se donne une martingale $(M_t, \mathcal{F}_t)$ à trajectoires continues et de carré intégrable, c'est-à-dire que pour tout $t \geqslant 0$, $\mathbb{E}(M_t^2)$ est fini (on dit aussi que M est L^2). On peut montrer qu'il existe un unique processus croissant noté $\langle M \rangle$ et appelé variation quadratique de M tel que $\langle M \rangle_0 = 0$ et tel que*

$$M_t^2 - \langle M \rangle_t$$

soit une martingale. On a la limite en probabilité suivante,

$$\langle M \rangle_t = \lim_{n \to \infty} \sum_{k=1}^{2^n} \left(M_{\frac{k}{2^n} t} - M_{\frac{k-1}{2^n} t} \right)^2.$$

Démonstration. L'existence de la variation quadratique d'une martingale en temps continu se base sur la discrétisation du temps. Voir [1, §4.3]. □

La variation quadratique $(\langle M \rangle_t)$ est un processus mesurable pour la filtration $(\mathcal{F}_t)$ et croissant.

Remarque 2.17 (Norme L^2 et variation quadratique)**.** On a en particulier, pour une martingale de carré intégrable (i.e. $M \in L^2$),

$$\mathbb{E}(M_t^2) = \mathbb{E}(\langle M \rangle_t).$$

On en déduit que si $M_0 = 0$ et $\langle M \rangle_t = 0$ pour tout $t \geqslant 0$ alors $(M_t) = 0$ presque surement.

Théorème et définition 2.18 (Covariation quadratique)**.** *On se donne une martingale* (M_t) *et* (N_t) *de carré intégrable. On définit la covariation quadratique par*

$$\langle M, N\rangle_t = \frac{1}{2}\left(\langle M+N\rangle_t - \langle M\rangle_t - \langle N\rangle_t\right).$$

Alors $\langle M, N\rangle_t$ *est l'unique processus à variation fini tel que*

$$M_t N_t - \langle M, N\rangle_t$$

soit une martingale. On a la limite en probabilité suivante,

$$\langle M, N\rangle_t = \lim_{n\to\infty} \sum_{k=1}^{2^n} \left(M_{\frac{k}{2^n}t} - M_{\frac{k-1}{2^n}t}\right)\left(N_{\frac{k}{2^n}t} - N_{\frac{k-1}{2^n}t}\right).$$

Démonstration. On se contente de montrer que $(M_t N_t - \langle M, N\rangle_t)$ est une martingale. On a

$$\begin{aligned} M_t N_t - \langle M, N\rangle_t &= \frac{1}{2}((M_t+N_t)^2 - M_t^2 - N_t^2) - \frac{1}{2}\left(\langle M+N\rangle_t - \langle M\rangle_t - \langle N\rangle_t\right) \\ &= \frac{1}{2}\Big(((M_t+N_t)^2 - \langle M+N\rangle_t) - (M_t^2 - \langle M\rangle_t) - (N_t^2 - \langle N\rangle_t)\Big) \end{aligned}$$

qui est donc la somme de trois martingales par le Théorème 2.16 et est donc une martingale. □

En remarquant que $\langle 2M\rangle_t = 4\langle M\rangle_t$ on déduit que

$$\langle M\rangle_t = \langle M, M\rangle_t.$$

Remarque 2.19 (Forme bilinéaire symétrique)**.** On peut aussi remarque que la covariation quadratique est une forme bilinéaire symétrique, en particulier

— $\langle M, N\rangle_t = \langle N, M\rangle_t$,

— $\langle aM, bN\rangle_t = ab\langle M, N\rangle_t$ pour tout a et $b \in \mathbb{R}$,

— $\langle M_1 + M_2, N\rangle_t = \langle M_1, N\rangle_t + \langle M_2, N\rangle_t$.

On en déduit que

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

Remarque 2.20 (Covariation et indépendance)**.** Le produit de deux martingales indépendante de carré intégrable M et N est encore une martingale, ce qui implique que $\langle M, N\rangle = 0$.

Travaux dirigés

Martingales discrètes

Exercice I.1. Soit (ξ_k) une suite i.i.d. t.q. $\mathbb{P}(\xi_1 = -1) = \mathbb{P}(\xi_1 = 1) = 1/2$ et posons $X_0 = 0$ et $X_k = \sum_{i=1}^k \xi_i$. Soient $a, b \in \mathbb{N}^*$ et $T = \inf\{k \in \mathbb{N} : X_k \leqslant -a \text{ ou } X_k \geqslant b\}$.

1. Montrer que T est un temps d'arrêt par rapport à la filtration $\mathscr{F}_k := \sigma(X_0, \xi_1, \ldots, \xi_k)$ et que (X_k) est une martingale.
2. Donner la valeur $\mathbb{E}[X_{k\wedge T}]$.
3. En admettant que $T < +\infty$ presque sûrement donner une relation entre $\mathbb{E}[X_T]$ et $p = \mathbb{P}(X_T = -a)$.
4. En admettant que $T < +\infty$ presque sûrement, en déduire la valeur de p.

Exercice I.2 (Première inégalité de Wald)**.** Soit $(X_k)_{k\geqslant 1}$ une suite i.i.d. de variables intégrables, $\mathscr{F}_k := \sigma(X_1, \ldots, X_k)$ et $\mu = \mathbb{E}[X_1]$. Soit T un temps d'arrêt t.q. $\mathbb{E}[T] < +\infty$ et $S_T = \sum_{i=1}^T X_i$. Le but de cet exercice est de montrer que $\mathbb{E}[S_T] = \mathbb{E}[T]\mathbb{E}[X_1]$.

1. Montrer que $S_k - k\mu$ est une martingale.
2. Montrer que $\mathbb{E}[S_{T\wedge k}] = \mathbb{E}[T \wedge k]\mu$
3. Si tous les X_k sont positifs p.s. montrer que $\mathbb{E}[S_T] = \mathbb{E}[T]\mathbb{E}[X_1]$.
4. Montrer le résultat dans le cas général.

Exercice I.3 (Deuxième égalité de Wald)**.** Dans le contexte de l'exercice I.2 supposons de plus que $\mu = 0$ et $\mathbb{E}[X_1^2] = \sigma^2 < +\infty$. Le but de cet exercice est de montrer que $\mathbb{E}[S_T^2] = \mathbb{E}[T]\sigma^2$.

1. Pour $k \in \mathbb{N}$, calculer $\mathrm{Var}(S_k)$.
2. Montrer que $Z_k := S_k^2 - k\sigma^2$ est une martingale.
3. Quelle est la valeur de $\mathbb{E}[S_{k\wedge T}^2]$?
4. Conclure.

Exercice I.4 (Troisième identité de Wald)**.** Soit $(X_k)_{k\geqslant 1}$ une suite de variable i.i.d. et $\theta \in \mathbb{R}$ t.q. $\mathbb{E}[e^{\theta X_1}] = \phi(\theta) < +\infty$. Soit T un temps d'arrêt et $S_k := \sum_{i=1}^k X_i$.

1. Montrer que $Y_k := \frac{e^{\theta S_k}}{\phi(\theta)^k}$ est une martingale.
2. Quelle est la valeur de $\mathbb{E}[Y_{T\wedge k}]$?
3. S'il existe $N \in \mathbb{N}$, t.q. $T < N$ montrer $\mathbb{E}[\phi(\theta)^{-T} e^{S_T}] = 1$.

Exercice I.5 (Une autre inégalité de Doob)**.** Soient $p > 1$, $\varepsilon > 0$ et (X_k) une martingale t.q. pour tout $k \in \mathbb{N}$, $\mathbb{E}[|X_k|^p] < +\infty$. Notons $|X|_k^* = \sup_{l\leqslant k} |X_l|$. Le but de l'exercice est de montrer l'inégalité suivante :

$$\mathbb{E}[(|X|_k^*)^p] \leqslant \left(\frac{p}{p-1}\right)^p \mathbb{E}[|X_k|^p] \,.$$

1. Pour $x \geqslant 0$, montrer que
$$x^p = \int_0^x p\varepsilon^{p-1}\, \mathrm{d}(\varepsilon) \,.$$

2. Pour $K > 0$, montrer
$$\mathbb{E}[(|X|_k^* \wedge K)^p] = \int_0^K p\varepsilon^{p-1}\mathbb{P}(|X|_k^* > \varepsilon)\,\mathrm{d}\varepsilon\,.$$
3. En déduire
$$\mathbb{E}[(|X|_k^* \wedge K)^p] \leqslant \mathbb{E}\left[|X_k| \int_0^{K\wedge |X|_k^*} p\varepsilon^{p-2}\,\mathrm{d}\varepsilon\right]\,.$$
et puis
$$\mathbb{E}[(|X|_k^* \wedge K)^p] \leqslant \frac{p}{p-1}\mathbb{E}\left[|X_k|(|X|_k^* \wedge K)^{p-1}\right]\,.$$
4. En appliquant l'inégalité de Hölder, montrer
$$\mathbb{E}[(|X|_k^* \wedge K)^p] \leqslant \frac{p}{p-1}\mathbb{E}\left[|X_k|^p\right]$$
et conclure.

Exercice I.6 (Loi des grands nombres)**.** Soit (M_k) une martingale t.q. $\sup_{k\in\mathbb{N}} \mathbb{E}[|M_k|^2] < +\infty$ et
$$\sum_{i=1}^{+\infty} \frac{\mathbb{E}[|M_i - M_{i-1}|^2]}{i^2} < +\infty\,.$$

1. Notons $X_k = \sum_{i=1}^k \frac{1}{i}(M_i - M_{i-1})$. Montrer que X_k est une martingale t.q. $\sup_k \mathbb{E}[|X_k|^2] < +\infty$.
2. Montrer
$$\frac{1}{k}M_k = X_k - \frac{1}{k}\sum_{i=1}^k X_{i-1}\,.$$
En déduire la limite de $\frac{1}{k}M_k$.

Soit maintenant (Y_k) une suite i.i.d. de v.a. intégrables.

3. Posons
$$M_k = \sum_{i=1}^k \left(X_i \mathbb{1}_{|X_i|\leqslant i} - \mathbb{E}[X_i \mathbb{1}_{|X_i|\leqslant i}]\right)\,.$$
Montrer que (M_k) est une martingale qui vérifie les conditions du début de l'exercice.
4. En déduire la limite de
$$\frac{1}{k}\sum_{i=1}^k X_i \mathbb{1}_{|X_i|\leqslant i}\,.$$
5. Montrer que p.s. il existe $N \in \mathbb{N}$ (une variable aléatoire), t.q. l'évènement $[X_k \geqslant N]$ se produise un nombre fini de fois. *Indication : penser à utiliser le lemme de Borel-Cantelli.*
6. Conclure sur la convergence de
$$\frac{1}{k}\sum_{i=1}^k X_i\,.$$

Exercice I.7 (Convergence p.s mais pas L^1)**.** Soit (ξ_k) une suite i.i.d. de loi de Bernoulli de paramètre p. On définit
$$X_k = e^{-k+2\sum_{i=1}^k \xi_i}\,.$$
Montrer les résultats suivants.

1. (X_k) est une martingale ssi $p = (1+e)^{-1}$.
2. $|X_k|^{1/k}$ converge vers une constante $s < 1$.
3. $\mathbb{E}[X_k] \not\to 0$ et donc $X_k \not\to 0$ en norme L^1.

Exercice I.8 (Urnes de Polya)**.** A l'instant 0 on dispose d'une urne avec $R_0 = r_0$ boules rouges et $B_0 = b_0$ boules bleues. On notera R_k (resp. B_k) le nombre de boules rouges (resp. bleues) à l'instant k et $X_k = \frac{B_k}{B_k+R_k}$ la proportion de boules bleues à l'instant k. La dynamique des proportions est évolue comme suit. À chaque instant k une variable U_{k+1} de loi uniforme dans $[0,1]$ est tirée de manière indépendante. Si $U_{k+1} \leqslant X_k$ alors on rajoute une boule bleue, sinon on rajoute une boule rouge.

1. Notons $\mathcal{F}_k = \sigma(U_1, \ldots, U_k)$. Montrer que $(X_k)_{k\geqslant 1}$ est une martingale.
2. Est-ce que (X_k) converge presque sûrement ? Et dans L^p pour $p \geqslant 1$?

Exercice I.9 (Marche aléatoire sur un graphe)**.** Un graphe non orienté G est représenté par le couple (V,E) où V est l'ensemble de sommets de G et E est l'ensemble de ses arêtes. Pour $i \in E$ on peut définir le degré de i comme le nombre de voisins de i. Une marche aléatoire sur un tel graphe est définie récursivement par $X_0 = x_0$ et

$$\mathbb{P}(X_{k+1} = j | X_k = i) = \begin{cases} 0 & \text{si } \{i,j\} \notin E\,, \\ \frac{1}{\deg i} & \text{sinon}\,. \end{cases}$$

Une fonction $f : V \to \mathbb{R}$ est dite harmonique si

$$f(i) = \frac{1}{\deg i} \sum_{j:\{i,j\}\in E} f(j)\,.$$

1. Montrer que $f(X_k)$ est une martingale.

Martingales à temps continu

Exercice I.10 (Propriété des Martingales de carré intégrable)**.** Soit $(M_t)_{t\geqslant 0}$ une martingale de carré intégrable $\left(M_t \in L^2 \text{ pour tout } t \geqslant 0\right)$. Soient $0 \leqslant s < t$ et soit $s = t_0 < t_1 < \cdots < t_p = t$ une subdivision de l'intervalle $[s,t]$. Montrer que

$$\mathbb{E}\left[\sum_{i=1}^{p}\left(M_{t_i} - M_{t_{i-1}}\right)^2 \mid \mathscr{F}_s\right] = \mathbb{E}\left[M_t^2 - M_s^2 \mid \mathscr{F}_s\right] = \mathbb{E}\left[(M_t - M_s)^2 \mid \mathscr{F}_s\right].$$

En déduire qu'en particulier on a,

$$\mathbb{E}\left[\sum_{i=1}^{p}\left(M_{t_i} - M_{t_{i-1}}\right)^2\right] = \mathbb{E}\left[M_t^2 - M_s^2\right] = \mathbb{E}\left[(M_t - M_s)^2\right].$$

Exercice I.11 (PAI)**.** On dit qu'un processus $(Z_t)_{t\geqslant 0}$ (à valeurs réelles) est un processus à accroissements indépendants (PAI) par rapport à la filtration $(\mathscr{F}_t)$ si Z est adapté et si, pour tous $0 \leqslant s < t, Z_t - Z_s$ est indépendant de la tribu $\mathscr{F}_s$ (par exemple un mouvement brownien est un PAI par rapport à sa filtration canonique, complétée ou non). Si Z est un PAI par rapport à $(\mathscr{F}_t)$, montrer que

(i) si $Z_t \in L^1$ pour tout $t \geqslant 0, \widetilde{Z}_t = Z_t - \mathbb{E}[Z_t]$ est une martingale ;

(ii) si $Z_t \in L^2$ pour tout $t \geqslant 0, X_t = \widetilde{Z}_t^2 - \mathbb{E}\left[\widetilde{Z}_t^2\right]$ est une martingale ;

(iii) s'il existe $\theta \in \mathbb{R}$ tel que $E\left[\mathrm{e}^{\theta Z_t}\right] < \infty$ pour tout $t \geqslant 0$,

$$X_t = \frac{\mathrm{e}^{\theta Z_t}}{\mathbb{E}\left[\mathrm{e}^{\theta Z_t}\right]}$$

est une martingale.

Exercice I.12 (Martingale à variation finie)**.** Une fonction $f : \mathbb{R}_+ \to \mathbb{R}$ est à variation finie si

$$V_t(f) := \sup\left\{\sum_{i=0}^{n-1}\left|f(t_{i+1}) - f(t_i)\right| : n \in \mathbb{N}^*, 0 = t_0 \leqslant \cdots \leqslant t_n = t\right\} < +\infty$$

pour tout $t \geqslant 0$. On rappelle que si f est continue, alors $t \mapsto V_t(f)$ l'est aussi. Soit $M = \{M_t\}_{t\geqslant 0}$ une martingale dont les trajectoires sont continues et à variation finie. Montrer que p.s., les trajectoires de M sont constantes. Indication : on pourra supposer que $V_t(M)$ est borné.

Deuxième partie

Mouvement Brownien et calcul stochastique

Chapitre 3

Mouvement Brownien

Un peu d'histoire Avant d'être un objet mathématique rigoureux, le mouvement brownien a été étudié en Botanique, en Finance, et en Physique. Le botaniste R. Brown observe d'abord vers 1828 le mouvement irrégulier de particules de pollen en suspension dans l'eau. En 1877, Delsaux explique les changements incessants de direction de trajectoire par les chocs entre les particules de pollen et les molécules d'eau. Un mouvement de ce type est alors appelé mouvement au hasard. En 1900, L. Bachelier, en vue d'étudier les cours de la Bourse de Paris dans sa thèse, met en évidence le caractère markovien du mouvement brownien : la position d'une particule à l'instant t+s dépend de sa position en t, et ne dépend pas de sa position avant t. Peu après, vers 1905, A. Einstein détermine la densité de transition du Brownien par l'intermédiaire de l'équation de la chaleur. La même année, Smoluchowski décrit le mouvement brownien comme une limite de promenades aléatoires. La première étude mathématique rigoureuse du Brownien est faite par N. Wiener (1923), qui construit une mesure de probabilités sur l'espace des fonctions continues sous laquelle le processus canonique est un mouvement Brownien. Des recherches d'une influence considérable ont ensuite été menées par P. Lévy (1948), lequel s'est intéressé aux propriétés fines des trajectoires du Brownien. Ces objets ont été developpés par les potentialistes américains à la suite de J. L. Doob, puis systématisés par les spécialistes de la "Théorie Générale des Processus" de l'école de Strasbourg, autour de P.-A. Meyer. [2]

3.1 Définitions du mouvement brownien

Dans toute la suite, on considerera un espace de probabilité muni de la filtration canonique

$$(\Omega, \mathcal{A}, (\mathcal{F}_t, t \geqslant 0), \mathbb{P}).$$

Définition 3.1 (Mouvement Brownien standard)**.** On appelle *mouvement brownien standard* un processus réel (W_t) qui est

i) à *trajectoires continues* (i.e. $t \mapsto W_t$ est continue) ;

ii) à *accroissements indépendants* (i.e. pour tout $n \geqslant 2$ et pour tout $0 \leqslant t_1 < \cdots < t_n$, la famille $(W_{t_1}, W_{t_2} - W_{t_1}, \cdots, W_{t_n} - W_{t_{n-1}})$ est indépendante, i.e. $\forall s < t$, $W_t - W_s$ indépendant de $\mathcal{F}_s$) ;

iii) à *accroissements stationnaires* (i.e. $W_t - W_s$ a même loi que W_{t-s} pour tout $s \leqslant t$) ;

iv) et tel que pour tout $t \geqslant 0$, W_t suit la loi gaussienne $\mathcal{N}(0, t)$.

On remarque que les propriétés iii) et iv) sont équivalentes au fait d'avoir un processus

iii') issu de 0 (i.e. $W_0 = 0$) ;

iv') et tel que pour tout $0 \leqslant s \leqslant t$, $W_t - W_s$ suit la loi gaussienne $\mathcal{N}(0, t - s)$.

Remarque 3.2 (Mouvement brownien, diffusion, dérive et point de départ)**.** On dit que (B_t) est un mouvement brownien de *coefficient de diffusion* (ou *volatilité*) σ^2, de *dérive* (*drift* en anglais) μ, issu de x, si $W_t = (B_t - \mu_t - x)/\sigma$ est un mouvement brownien standard, i.e. si il existe W_t un mouvement brownien standard tel quel $B_t = x + \mu t + \sigma W_t$. La propriété iv) est remplacée par $B_t - x \sim \mathcal{N}(\mu t, \sigma^2 t)$.

Remarque 3.3 (Processus de Lévy (PAIS))**.** Les processus à accroissement stationnaires et indépendants (i.e. qui vérifient ii) et iii)) sont appelés processus de Lévy. On peut montrer (mais c'est un résultat difficile qui découle de ce qu'on appelle la représentation de Lévy-Khintchine) que les seuls processus de Lévy continus sont des mouvements Brownien avec dérive, c'est à dire des processus gaussiens qui vérifient $W_t - W_0 \sim \mathcal{N}(\mu t, \sigma^2 t)$ où μ, σ sont des constantes réelles appelées la dérive et le coefficient de diffusion.

Remarque 3.4 (Continuité)**.** Dans la définition du mouvement brownien, on aurait en fait pu se passer de la continuité, et montrer que (W_t) admet une modification $(\widetilde{W}_t)$ (i.e. pour $t \geqslant 0$, $\mathbb{P}(W_t = \widetilde{W}_t) = 1$) continue qui est encore un mouvement brownien. Voir [1, §2.2].

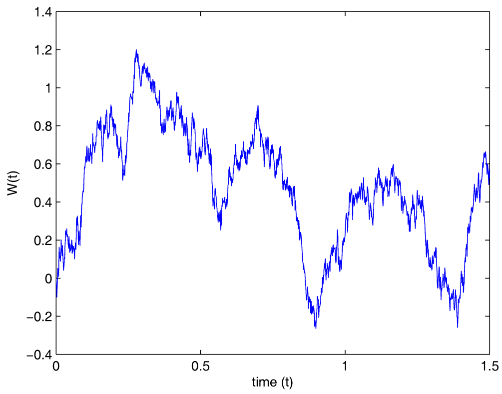


FIGURE 3.1 – Simulation d'une trajectoire d'un mouvement Brownien

Le théorème suivant fourni une définition équivalente du mouvement Brownien.

Théorème 3.5 (Processus gaussien de covariance $s \wedge t$)**.** *Soit (W_t) un processus continu. Alors (W_t) est un mouvement brownien standard si et seulement si c'est un processus gaussien centré de covariance $\mathbb{E}(W_s W_t) = s \wedge t$, c'est à dire un processus tel que pour tout $0 \leqslant t_1 \leqslant \cdots \leqslant t_n$, $(W_{t_1}, \ldots, W_{t_n})$ est un vecteur gaussien centré de matrice de variance-covariance $[t_i \wedge t_j]_{1 \leqslant i,j \leqslant n}$.*

Démonstration. $\Rightarrow$. Par indépendance et stationnarité des accroissements, pour $t_1 \leqslant t_2 \leqslant \ldots \leqslant t_n$, $(W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_n} - W_{t_{n-1}})$ est un vecteur gaussien à composantes indépendantes. Par transformation linéaire, on en déduit donc que $(W_{t_1}, \ldots, W_{t_n})$ est un vecteur gaussien (centré). Il reste à calculer $\mathbb{C}\mathrm{ov}(W_s, W_t)$. Prenons par exemple, $s \leqslant t$, alors par indépendance de W_t et $W_t - W_s$, on a

$$\mathbb{C}\mathrm{ov}(W_s, W_t) = \mathbb{E}(W_s W_t) - \underbrace{\mathbb{E}(W_s)\mathbb{E}(W_t)}_{0} = \underbrace{\mathbb{E}(W_s(W_t - W_s))}_{\mathbb{E}(W_s)\mathbb{E}(W_t - W_s) = 0\times 0} + \underbrace{\mathbb{E}(W_s^2)}_{s} = s$$

$\Leftarrow$. Pour tout $t > s \geqslant 0$, $W_t - W_s$ est une gaussienne centré de variance

$$\mathbb{V}\mathrm{ar}(W_t - W_s) = \mathbb{V}\mathrm{ar}(W_t) + \mathbb{V}\mathrm{ar}(W_s) - 2\mathbb{C}\mathrm{ov}(W_t, W_s) = t + s - 2s = t - s.$$

Donc $W_t - W_s \sim \mathcal{N}(0, t-s)$ qui est une loi qui ne dépend pas de t, les accroissements sont stationnaires. De plus, le vecteur des accroissements $(W_{t_2} - W_{t_1}, W_{t_3} - W_{t_2} \ldots, W_{t_n} - W_{t_{n-1}})$ est obtenu par transformation linéaire du vecteur $(W_{t_1}, \ldots, W_{t_n})$ qui est un vecteur gaussien, donc le vecteur des accroissements est aussi gaussien et on sait que pour un vecteur *gaussien*, ses composantes sont indépendantes ssi leurs covariances sont nulles. On achève la preuve en montrant que pour $i < j$,

$$\begin{aligned}\mathbb{C}\mathrm{ov}(W_{t_i} - W_{t_{i-1}}, W_{t_j} - W_{t_{j-1}}) = \mathbb{C}\mathrm{ov}(W_{t_i}, W_{t_j}) - \mathbb{C}\mathrm{ov}(W_{t_i}, W_{t_{j-1}}) - \mathbb{C}\mathrm{ov}(W_{t_{i-1}}, W_{t_j}) + \mathbb{C}\mathrm{ov}(W_{t_{i-1}}, W_{t_{j-1}}) \\ = t_i - t_i - t_{i-1} + t_{i-1} = 0\end{aligned}$$

□

Théorème 3.6 (Existence)**.** *Le mouvement brownien existe. (admis)*

On considère $\mathscr{C}(\mathbb{R}_+, \mathbb{R})$ l'espace des fonctions continues de $\mathbb{R}_+$ dans $\mathbb{R}$ munit de la tribu $\mathcal{C}$ qui est la plus petite tribu rendant mesurables les applications coordonnées $w \mapsto w(t)$ de $\mathscr{C}(\mathbb{R}_+, \mathbb{R}) \to \mathbb{R}$.

Théorème et définition 3.7 (Mesure de Wiener)**.** *On considère un mouvement brownien W_t. L'application*

$$\begin{aligned} \Phi : \quad (\Omega, \mathcal{F}, \mathbb{P}) &\longrightarrow (\mathscr{C}(\mathbb{R}_+, \mathbb{R}), \mathcal{C}) \\ \omega &\longmapsto (t \mapsto W_t(\omega)) \end{aligned}$$

est mesurable. On définit la mesure de Wiener $\mathbb{W}$ *comme la mesure image de $\mathbb{P}$ par Φ, i.e. pour tout $A \subset \mathscr{C}(\mathbb{R}_+, \mathbb{R})$,*

$$\mathbb{W}(A) := \mathbb{P}(\Phi^{-1}(A)).$$

Cette mesure ne dépend pas du mouvement brownien (W_t) choisi.

Démonstration. Admis. □

La mesure de Wiener est une probabilité sur $(\mathscr{C}(\mathbb{R}_+, \mathbb{R}), \mathcal{C})$. Il s'agit de la *loi du mouvement brownien.* Le théorème précédent montre que la loi du mouvement brownien est unique.

Lorsqu'on considère un mouvement Brownien, on prend désormais l'espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ avec

$$\Omega = \mathscr{C}(\mathbb{R}_+, \mathbb{R}), \quad \mathcal{F} = \mathcal{C}, \quad \mathbb{P} = \mathbb{W}.$$

Le processus canonique défini pour $w \in \Omega$ par $X_t(w) = w(t)$ est un mouvement brownien.

On conclu cette section en donnant quatre mouvements brownien que l'on peut déduire d'un mouvement brownien (W_t).

Proposition 3.8 (Symétrie, changement d'échelle et inversion et retournement du temps)**.** *Les processus suivant sont des mouvements brownien :*

1. $(-W_t)$ *(symétrie) ;*
2. $(c^{-1}W_{tc^2})$ *pour tout $c > 0$ (changement d'échelle) ;*
3. $(tW_{1/t})$ *(inversion du temps) ;*
4. $(W_T - W_{T-t})$*, pour tout $T > 0$ et $t \in [0, T]$ (retournement du temps).*

Démonstration. Utiliser le Théorème 3.5 et vérifier qu'on a des processus gaussien centrés de covariance $s \wedge t$, à trajectoires p.s. continues. Voir TD. (Seule difficulté admise, la continuité en 0 de $tW_{1/t}$). □

3.2 Approximation via une marche aléatoire

Il est possible de construire le Brownien comme limite de marches aléatoires renormalisées. Cette approximation peut être utilisée pour simuler un mouvement Brownien, voir le TP 2.

On prend $X_1, \cdots, X_n$ une suite de variables aléatoires, indépendantes et de même loi, de type Bernoulli de paramètre 1/2, telles que $\mathbb{P}(X_k = 1) = \mathbb{P}(X_k = -1) = 1/2$. On considère la marche symétrique simple

$$S_n = X_1 + \cdots + X_n$$

qui vérifie $\mathbb{E}(S_n) = 0$ et $\mathbb{V}\mathrm{ar}(S_n) = n$. On interpole alors cette marche aléatoire pour la transformer en fonction continue, on définit $S_0 = 0$ et pour tout $t > 0$,

$$S_t := S_{\lfloor t \rfloor} + (t - \lfloor t \rfloor) X_{\lfloor t \rfloor + 1}.$$

On procède alors à une <u>double renormalisation</u> :

- en *temps* (avec un facteur $1/n$),
- en *espace* (avec un facteur $1/\sqrt{n}$),

et on définit pour tout $t \in [0, 1]$ le processus

$$S_t^{(n)} := \frac{S_{nt}}{\sqrt{n}} = \frac{1}{\sqrt{n}} \left(S_{\lfloor nt \rfloor} + (nt - \lfloor nt \rfloor) X_{\lfloor nt \rfloor + 1} \right).$$

Théorème 3.9 (Donsker)**.** *La suite de processus* $(S_t^{(n)})$ *converge en loi vers un mouvement Brownien standard* (W_t) *dans* $\mathscr{C}([0,1],\mathbb{R})$.

Idée de preuve et heuristique. On cherche à étudier les propriétés i), ii), iii') et iv') de la définition 3.1 sur le processus $S_t^{(n)}$ ou sa limite.

i) Le processus $S_t^{(n)}$ est continu.

ii) On peut montrer que la partie qui sert à interpoler, $(nt-\lfloor nt\rfloor)X_{\lfloor nt\rfloor+1}/\sqrt{n}$ converge en probabilité vers 0. De plus, il est évident que $\frac{S_{\lfloor nt\rfloor}}{\sqrt{n}}$ a des accroissements indépendants en effet pour tout $0 \leqslant t_1 \leqslant \cdots \leqslant t_k$,

$$\begin{aligned}&\left(S_{\lfloor nt_1\rfloor}, S_{\lfloor nt_2\rfloor}-S_{\lfloor nt_1\rfloor}, \cdots, S_{\lfloor nt_k\rfloor}-S_{\lfloor nt_{k-1}\rfloor}\right)\\ &\qquad= (X_1+\cdots+X_{\lfloor nt_1\rfloor}, X_{\lfloor nt_1\rfloor+1}+\cdots+X_{\lfloor nt_2\rfloor}, \cdots, X_{\lfloor nt_{k-1}\rfloor+1}+\cdots+X_{\lfloor nt_k\rfloor})\end{aligned}$$

qui est une famille indépendante car les X_i sont indépendants.

iii') Le processus est issu de 0, $S_0^{(n)}=0$.

iv') En faisant tendre n vers l'infini on a la convergence en loi de $S_t^{(n)}-S_s^{(n)}$ vers $\mathcal{N}(0,t-s)$. En effet, comme $\mathbb{E}(X_k)=0$ et $\mathbb{V}\mathrm{ar}(X_k)=1$, le théorème central limite implique

$$\frac{S_{\lfloor ns\rfloor}}{\sqrt{n}}-\frac{S_{\lfloor nt\rfloor}}{\sqrt{n}} = \frac{\sqrt{\lfloor nt\rfloor-\lfloor ns\rfloor}}{\sqrt{n}}\frac{X_{\lfloor ns\rfloor+1}+\cdots+X_{\lfloor nt\rfloor}}{\sqrt{\lfloor nt\rfloor-\lfloor ns\rfloor}} \xrightarrow[n\to\infty]{\text{(loi)}} \sqrt{t-s}\mathcal{N}(0,1)=\mathcal{N}(0,t-s).$$

Pour rendre la preuve rigoureuse il faudrait montrer la convergence des lois fini-dimensionnelles et utiliser de la notion de tension, voir la remarque suivante. □

Remarque 3.10 (Loi d'un processus et convergence)**.** On considère (X_t) un processus réel continu. Tout comme on a défini la mesure de Wiener pour le mouvement Brownien, on considère

$$\begin{aligned}\Phi:\quad (\Omega,\mathcal{F},\mathbb{P}) &\longrightarrow (\mathscr{C}(\mathbb{R}_+,\mathbb{R}),\mathcal{C})\\ \omega &\longmapsto (t\mapsto X_t(\omega))\end{aligned}$$

et on définit $\mathbb{P}_X$ la loi du processus (X_t) comme la mesure image de $\mathbb{P}$ par Φ, i.e. pour tout $A\subset\mathscr{C}(\mathbb{R}_+,\mathbb{R})$,

$$\mathbb{P}_X(A) := \mathbb{P}(\Phi^{-1}(A)).$$

La loi d'un processus est caractérisée par ses *lois fini-dimensionnelles*, c'est à dire les lois des vecteurs aléatoires $(X_{t_1},\cdots,X_{t_n})$ pour tout $t_1<t_2<\cdots<t_n$. Cela signifie que si deux processus (X_t) et (Y_t) ont les mêmes lois fini-dimensionnelles si et seulement si ils ont la même loi $\mathbb{P}_X=\mathbb{P}_Y$.

La convergence en loi d'un processus peut se montrer par la convergence des lois fini-dimensionnelles et par un critère dit de *tension* qui est un concept subtil qui dépasse largement le cadre de ce cours.

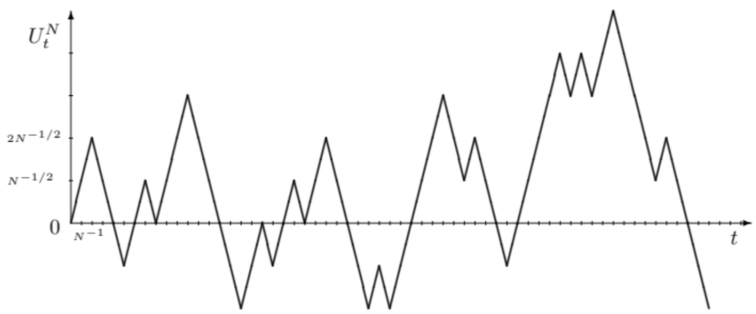


FIGURE 3.2 – Illustration du théorème de Donsker

3.3 Propriété de Markov

Heuristique *En langage courant, dire qu'un processus vérifie la propriété de Markov signifie que la loi du processus dans le futur étant donnés les états passés et l'état présent, ne dépend en fait que de l'état présent et non pas des états passés. On parle d'absence de mémoire.*

La condition d'indépendance des accroissement ii) de la définition 3.1 est équivalente au fait que $W_t - W_s$ est indépendant de $\mathcal{F}_s$ pour tout $s \leqslant t$. On peut ainsi montrer que si $(W_t)_{t \geqslant 0}$ est un mouvement brownien pour la filtration $(\mathcal{F}_t)$ alors pour tout $s \geqslant 0$, $(W_{t+s} - W_s)_{t \geqslant 0}$ est un mouvement brownien indépendant de $\mathcal{F}_s$ pour la filtration $(\mathcal{F}_{t+s})_{t \geqslant 0}$. On en déduit ainsi la propriété de Markov simple.

Théorème 3.11 (Propriété de Markov simple)**.** *Pour tout $s \geqslant 0$, $(W_{t+s} - W_s)_{t \geqslant 0}$ est un mouvement brownien indépendant de $\mathcal{F}_s$ et pour toute fonction f mesurable bornée, pour tout $t \geqslant 0$,*

$$\mathbb{E}(f(W_{t+s})|\mathcal{F}_s) = \mathbb{E}(f(W_{t+s})|W_s) = \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} f(x + W_s) e^{-\frac{x^2}{2t}} \mathrm{d}x.$$

Une autre façon de décrire cette propriété est de dire que conditionnellement $\mathcal{F}_s$, la v.a. W_{t+s} est de loi gaussienne d'espérance W_s et de variance t et que le processus (W_{t+s}) est un mouvement Brownien issu de W_s.

Le théorème suivant (admis) étend la propriété de Markov à des temps d'arrêt.

Théorème 3.12 (Propriété de Markov forte)**.** *Pour tout T temps d'arrêt à valeurs finies, $(W_{t+T} - W_T)_{t \geqslant 0}$ est un mouvement brownien indépendant de $\mathcal{F}_T$ et pour toute fonction f mesurable bornée, pour tout $t \geqslant 0$,*

$$\mathbb{E}(f(W_{t+T})|\mathcal{F}_T) = \mathbb{E}(f(W_{t+T})|W_T) = \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} f(x + W_T) e^{-\frac{x^2}{2t}} \mathrm{d}x.$$

La propriété de Markov forte est l'une des propriétés les plus utiles du Brownien.

Dans le théorème suivant on considère la filtration canonique augmentée de la remarque 2.5 que l'on notera $(\mathcal{F}_t)$ et telle que $\mathcal{F}_t^+ = \mathcal{F}_t$ pour tout $t \geqslant 0$. Le théorème suivant s'appelle la loi du tout ou rien.

Théorème 3.13 (Loi du 0-1 de Blumenthal)**.** *La tribu $\mathcal{F}_0^+$ est triviale au sens où pour tout $A \in \mathcal{F}_0^+$, $\mathbb{P}(A) = 0$ ou 1.*

Démonstration. Cela découle directement de la propriété de Markov simple, qui implique que pour tout $t \geqslant 0$, W_t est indépendante de $\mathcal{F}_0$, ce qui implique que pour tout t, $\mathcal{F}_t$ est indépendante de $\mathcal{F}_0$ et comme $\mathcal{F}_0 \subset \mathcal{F}_t$, $\mathcal{F}_0$ est indépendante d'elle même. Donc, tout ses éléments sont de probabilité 0 ou 1. □

On en déduit quelques propriétés de trajectoire du mouvement Brownien. On notera $\overline{\lim}$ (resp. $\underline{\lim}$) la limite supérieure (resp. inférieure). On rappelle par exemple que pour une fonction f, $\overline{\lim}_\infty f = \lim_{t \to \infty} \sup\{f(s) : s \geqslant t\}$ et $\overline{\lim}_0 f = \lim_{t \to 0} \sup\{f(s) : |s| \leqslant t\}$ et idem pour la limite inférieure.

Proposition 3.14 (Comportement des trajectoires Browniennes)**.** *Presque surement,*

1. $\overline{\lim}_{t \to \infty} \dfrac{W_t}{\sqrt{t}} = \overline{\lim}_{t \to 0} \dfrac{W_t}{\sqrt{t}} = +\infty$;
2. $\underline{\lim}_{t \to \infty} \dfrac{W_t}{\sqrt{t}} = \underline{\lim}_{t \to 0} \dfrac{W_t}{\sqrt{t}} = -\infty$;
3. $\lim\limits_{t \to \infty} \dfrac{W_t}{t} = 0$.

Démonstration. C'est une conséquence de la loi du tout ou rien. Pour 1, on considère la v.a.

$$R = \overline{\lim}_{t \to +\infty} t^{-1/2} W_t = \overline{\lim}_{t \to +\infty} t^{-1/2} (W_t - W_s)$$

pour tout $s \geqslant 0$. Par indépendance des accroissements de W, R est indépendante de $\sigma\{B_u, u \leqslant s\}$ pour tout $s \geqslant 0$, et donc de $\sigma\{B_u, u \geqslant 0\}$ en faisant tendre $s \to \infty$. Or R est clairement mesurable par rapport à $\sigma\{W_u, u \geqslant 0\}$, de sorte que R est indépendante d'elle-même, et ceci n'est évidemment possible que si R est une constante déterministe (loi du tout ou rien). Supposons que R soit finie. Alors $\mathbb{P}\left[t^{-1/2}W_t \geqslant R+1\right] \to 0$ quand $t \to +\infty$ (car $\overline{\lim}\mathbb{E}(\mathbb{1}_{t^{-1/2}W_t \geqslant R+1}) \leqslant \mathbb{E}(\overline{\lim}\mathbb{1}_{t^{-1/2}W_t \geqslant R+1}) = 0$). Mais par scaling,

$$\mathbb{P}\left[t^{-1/2}W_t \geqslant R+1\right] = \mathbb{P}\left[W_1 \geqslant R+1\right] \neq 0$$

pour tout $t \geqslant 0$, de sorte que nécessairement $R = +\infty$. De même, on montre que $\overline{\lim}_{t\to 0} t^{-1/2}W_t = +\infty$ p.s.

Le point 2 est une conséquence immédiate de la symétrie du Brownien, et le point 3 découle de la loi des grands nombres, on peut montrer que $\lim \frac{W_t}{t} = \lim \frac{W_1+(W_2-W_1)+\cdots+(W_{\lfloor t\rfloor}-W_{\lfloor t\rfloor-1})}{\lfloor t\rfloor} = \mathbb{E}(W_1) = 0$. □

Cette proposition permet de mesurer grossièrement deux quantités reliées à la courbe brownienne, d'une part sa vitesse locale en 0 plus grande que $\sqrt{t}$ (et idem en tout point t_0, plus grande que $\sqrt{|t-t_0|}$), et d'autre part son comportement en l'infini qui oscille au delà des courbes $\pm\sqrt{t}$ mais en dessous des courbes $\pm t$. En fait, le mouvement Brownien se comporte presque comme $\sqrt{t}$ en 0 et l'∞.

Le théorème suivant dû à Khintchine est bien plus précis. Il donne la vitesse locale et les taux de croissance exacts du Brownien.

Théorème 3.15 (Loi du logarithme itéré)**.** *Presque surement,*

1. $\overline{\lim}_{t\to 0} \dfrac{W_t}{\sqrt{2t\log\log(1/t)}} = 1$ *et* $\underline{\lim}_{t\to 0} \frac{W_t}{\sqrt{2t\log\log(1/t)}} = -1$ *;*
2. $\overline{\lim}_{t\to \infty} \dfrac{W_t}{\sqrt{2t\log\log(t)}} = 1$ *et* $\underline{\lim}_{t\to \infty} \frac{W_t}{\sqrt{2t\log\log(t)}} = -1$*.*

Démonstration. Voir TD. □

Remarque 3.16 (Zéros du Brownien)**.** La proposition précédente implique directement par continuité de $t \mapsto W_t$ que le mouvement brownien a une infinité de zéros au voisinage de l'infini et au voisinage de 0.

Remarque 3.17 (Non-dérivabilité)**.** Il est facile de voir que pour tout $t \geqslant 0$ fixé, (W_t) n'est presque surement pas dérivable en t. Heuristiquement, $(W_{t+h} - W_t)/h \sim \mathcal{N}(0, 1/h)$ est une gaussienne de variance très grande lorsque $h \to 0$ et ne converge donc pas. La non-dérivabilité en 0 découle directement du fait que p.s. $\overline{\lim}_{t\to 0} \frac{W_t}{\sqrt{t}} = +\infty$ et la propriété de Markov faible permet de généraliser à tous les t.

Avec la propriété de changement d'échelle et d'indépendance des accroissement on peut montrer une propriété encore plus forte (admise). Presque surement, $t \mapsto W_t$ est non dérivable en tout point. (On a échangé le p.s. et le $\forall$ point.)

Remarque 3.18 (Caractère Hölderien)**.** De la loi du logarithme itéré on déduit facilement que en 0 (ou en t fixé), le mouvement Brownien est presque surement α*-Hölderienne* pour tout $\alpha \in]0, 1/2[$ et pas pour $\alpha \in [1/2, 1]$. On pourrait montrer que presque surement ses trajectoires sont localement α*-Hölderienne* en tout point, pour tout $\alpha \in]0, 1/2[$ et pas pour $\alpha \in [1/2, 1]$.

Pour rappel, soit f une fonction définie sur un intervalle I de $\mathbb{R}$. On dit que f est α-höldérienne, où $\alpha \in]0, 1]$, s'il existe une constante $C_\alpha > 0$ telle que, pour tous $x, y \in I$,

$$|f(x) - f(y)| \leqslant C_\alpha |x-y|^\alpha.$$

En particulier, si $\alpha = 1$, on retrouve la définition des fonctions lipschitziennes. Pour tout $\alpha \in]0, 1]$, toute fonction α-höldérienne est continue, et même uniformément continue. Si I est un segment, toute fonction dérivable sur I est α höldérienne, pour tout $\alpha \in]0, 1]$. Cette condition de régularité, qui est donc intermédiaire entre "être continue" et "être dérivable".

Équation de la chaleur L'équation de la chaleur est une équation aux dérivées partielles parabolique, pour décrire le phénomène physique de conduction thermique. Une solution fondamentale de cette équation de diffusion s'obtient en considérant la densité d'un mouvement brownien. Voir TD.

3.4 Martingales à temps continu

Voici des exemples de martingales construites à partir du mouvement brownien standard W_t.

Proposition 3.19. *Les processus suivants sont des martingales :*

1. (W_t)
2. $(W_t^2 - t)$
3. $(\exp(\lambda W_t - (\lambda^2/2)t))$

Démonstration. Dans tout ce qui suit, $s \leqslant t$ et on utilise le fait que $W_t - W_s$ est indépendant de $\mathcal{F}_s$.

1. $\mathbb{E}(W_t - W_s|\mathcal{F}_s) = \mathbb{E}(W_t - W_s) = 0$ donc $\mathbb{E}(W_t|\mathcal{F}_s) = W_s$.
2.

$$\begin{aligned}\mathbb{E}(W_t^2 - t|\mathcal{F}_s) &= \mathbb{E}(((W_t - W_s) + W_s)^2|\mathcal{F}_s) - t \\ &= \mathbb{E}((W_t - W_s)^2|\mathcal{F}_s) + \mathbb{E}(W_s^2|\mathcal{F}_s) + 2\mathbb{E}((W_t - W_s)W_s|\mathcal{F}_s) - t \\ &= \mathbb{E}((W_t - W_s)^2) + W_s^2 + 2W_s\mathbb{E}((W_t - W_s)|\mathcal{F}_s) - t \\ &= t - s + W_s^2 + 2W_s \times 0 - t = W_s^2 - s\end{aligned}$$

3. Enfin, comme la loi de $W_t - W_s$ conditionnellement à $\mathcal{F}_s$ est une loi normale $\mathcal{N}(0, t-s)$, la valeur de la fonction caractéristique des variables gaussienne donne $\mathbb{E}(\exp[\lambda(W_t - W_s)]|\mathcal{F}_s) = \mathrm{e}^{\frac{\lambda^2}{2}(t-s)}$. Ce qui assure que $\mathbb{E}(\exp(\lambda W_t - (\lambda^2/2)t)|\mathcal{F}_s) = \exp(\lambda W_s - (\lambda^2/2)s)$.

□

La réciproque est aussi vraie !

Théorème 3.20 (Théorème de caractérisation de P. Lévy)**.** *Soit M une martingale (locale) à trajectoires continues.*

— *Si $(M_t^2 - t)$ est une martingale alors (M_t) est un mouvement brownien.*

— *Si $(\exp(\lambda M_t - (\lambda^2/2)t)$ est une martingale alors (M_t) est un mouvement brownien.*

On en déduit que si $\langle M \rangle_t = t$ alors M est un mouvement brownien.

Démonstration. Si $(\exp(\lambda M_t - (\lambda^2/2)t)$ est une martingale, alors

$$\mathbb{E}\left[\exp[i(\lambda \cdot (M_t - M_s)] \,|\, \mathcal{F}_s\right] = \exp\left(-\frac{\lambda^2}{2}(t-s)\right).$$

C'est la fonction caractéristique d'une loi normale de variance $t - s$. Ainsi, la loi conditionnelle de $(M_t - M_s)$ sachant $\mathcal{F}_s$ est la loi $\mathcal{N}(0, (t-s))$, et on conclut que M est un mouvement brownien : en effet il est continu, adapté, et pour $0 < s < t$, $(M_t - M_s)$ est indépendant de $\mathcal{F}_s$ et de loi $\mathcal{N}(0, (t-s))$.

On verra plus tard (avec la formule d'Itô) que si M est une martingale alors $(\exp(\lambda M_t - (\lambda^2/2)\langle M \rangle_t)$ est aussi une martingale. Ainsi, si $\langle M \rangle_t = t$ alors M est un mouvement brownien.

De plus, le fait que $(M_t^2 - t)$ soit une martingale implique par unicité de la variation quadratique que $\langle M \rangle_t = t$ et donc que M est un mouvement brownien. □

Le Théorème d'arrêt 2.13 permet de calculer la loi du temps d'atteinte d'un seuil pour un brownien.

Proposition 3.21 (Temps d'atteinte)**.** *Soit (W_t) un mouvement brownien. Soit $a \in \mathbb{R}$, posons $T_a = \inf\{s \geqslant 0; W_s = a\}$ avec la convention $\inf \emptyset = \infty$. Alors T_a est un temps d'arrêt fini presque sûrement, tel que $\mathbb{E}(T_a) = \infty$ et pour tout $\lambda \geqslant 0$, on obtient la transformée de Laplace de la densité de T_a et par inversion sa densité f_{T_a},*

$$\mathbb{E}\left(e^{-\lambda T_a}\right) = e^{-\sqrt{2\lambda}|a|} \quad et \quad f_{T_a}(t) = \frac{|a|}{\sqrt{2\pi t^3}} e^{-\frac{a^2}{2t}}.$$

Démonstration. L'idée est de prouver que $\mathbb{E}\left[e^{\sqrt{2\lambda}W_{T_a} - \lambda T_a}\right] = 1$. Le temps d'arrêt T_a n'étant pas borné, on considérera $T_a \wedge n$ et on appliquera le théorème d'arrêt 2.13. On fait ensuite tendre n vers l'infini et on remarque que $T_a \wedge n \to T_a$ cat T_a fini p.s. et $W_{T_a \wedge n} \to W_{T_a} = a$ p.s. par continuité. On conclu par convergence dominée car $\left|e^{\sqrt{2\lambda}W_{T_a \wedge n} - \lambda T_a \wedge n}\right| \leqslant e^{\sqrt{2\lambda}a}$. □

Variation quadratique du mouvement brownien Rappelons la définition de la variation quadratique d'une martingale énoncé dans le Théorème 2.16 (et qui avait été admis). La variation quadratique du mouvement Brownien est l'unique processus croissant $\langle W\rangle_t$ tel que $W_t^2 - \langle W\rangle_t$ soit une martingale. La Proposition 3.19 nous dit que $W_t^2 - t$ est une martingale et on peut donc conclure que

$$\langle W\rangle_t = t.$$

Dans la proposition suivante que l'on va démontrer, on retrouve par calcul direct la valeur de la variation quadratique du mouvement brownien. Cette propriété met une une nouvelle fois en évidence le caractère "quadratique" de la courbe brownienne.

Proposition 3.22 (Variation quadratique du mouvement brownien)**.** *On a la limite presque sure et dans L^2 suivante :*

$$\langle W\rangle_t := \lim_{n\to\infty} \sum_{k=1}^{2^n} \left(W_{\frac{k}{2^n}t} - W_{\frac{k-1}{2^n}t}\right)^2 = t.$$

Démonstration. On pose $Z_t^n = \sum_{k=1}^{2^n} \left(W_{\frac{k}{2^n}t} - W_{\frac{k-1}{2^n}t}\right)^2$. On remarque que $W_{\frac{k}{2^n}t} - W_{\frac{k-1}{2^n}t}$ suit une loi $\mathcal{N}(0, t/2^n)$ et on déduit donc que $\mathbb{E}(Z_t^n) = \sum_{k=1}^{2^n} t/2^n = t$. Pour la convergence dans L^2 on cherche à montrer que $\mathbb{E}((Z_t^n - t)^2) = \mathbb{V}\mathrm{ar}(Z_t^n)$ tend vers 0. En se souvenant que les accroissement $W_{\frac{k}{2^n}t} - W_{\frac{k-1}{2^n}t}$ sont indépendants et grâce au fait que pour $X \sim \mathcal{N}(0,\sigma^2)$ on a $\mathbb{V}\mathrm{ar}X^2 = 2\sigma^4$, on obtient

$$\mathbb{V}\mathrm{ar}(Z_t^n) = \sum_{k=1}^{2^n} \mathbb{V}\mathrm{ar}\left(W_{\frac{k}{2^n}t} - W_{\frac{k-1}{2^n}t}\right)^2 = \sum_{k=1}^{2^n} 2\left(\frac{t}{2^n}\right)^2 = t^2/2^{n-1} \underset{n\to\infty}{\longrightarrow} 0.$$

On a montré la convergence L^2. Pour la convergence presque sure on remarque que

$$\mathbb{E}\left(\sum_{n=1}^{\infty} (Z_t^n - t)^2\right) = \sum_{n=1}^{\infty} t^2/2^{n-1} < \infty.$$

On en déduit donc que presque surement $\sum_{k=1}^{\infty}((Z_t^n - t)^2) < \infty$ et que le terme général de cette série $(Z_t^n - t)^2$ converge presque surement vers 0. $\square$

La variation quadratique du mouvement brownien (ou crochet de W) est utilisée dans la construction de l'intégrale stochastique.

Corolaire 3.23 (Variation infinie du mouvement brownien)**.** *Le mouvement brownien a presque surement une variation infini sur tout intervalle. C'est à dire, presque surement, pour tout $a < b$,*

$$V_{[a,b]}(W) := \sup_{a=t_0<\ldots<t_N=b} \sum_{i=1}^{N} |W_{t_i} - W_{t_{i-1}}| = \infty.$$

Démonstration. On se contente de montrer que presque surement, le mouvement brownien a une variation infinie sur $[0,1]$. On a vu dans la proposition précédente que

$$\lim_{n\to\infty} \sum_{k=1}^{2^n} \left(W_{\frac{k}{2^n}} - W_{\frac{k-1}{2^n}}\right)^2 = 1.$$

Nous avons de plus,

$$\sum_{k=1}^{2^n} \left(W_{\frac{k}{2^n}} - W_{\frac{k-1}{2^n}}\right)^2 \leqslant V_{[a,b]}(B) \times \sup_{i=1,\cdots,2^n} |W_{\frac{k}{2^n}} - W_{\frac{k-1}{2^n}}|.$$

Comme W est continu sur le segment $[0,1]$ on a $\sup_{i=1,\cdots,2^n} |W_{\frac{k}{2^n}} - W_{\frac{k-1}{2^n}}|$ tend vers 0 p.s. quand $n \to \infty$. Comme la limite du membre de gauche de l'inégalité est 1 on en déduit que p.s. $V_{[a,b]}(B) = \infty$. $\square$

Cela signifie que la trajectoire d'un brownien est presque sûrement de longueur infinie.

Avec une démonstration similaire on peut montrer le résultat suivant.

Proposition 3.24 (Martingale à variation finie)**.** *Soit M une martingale à variation finie et à trajectoires continues. Alors p.s., pour tout $t \geqslant 0$, $M_t = M_0$.*

Démonstration. Voir l'exercice I.12 du TD sur les martingales à temps continu. □

De la même manière on peut montrer qu'un processus à variation totale finie a une variation quadratique nulle.

Chapitre 4

Intégrale stochastique

L'objectif de ce chapitre est d'expliquer la construction de l'intégrale stochastique qui permet de donner sens à la quantité
$$(H\cdot W)_t := \int_0^t H_s \mathrm{d}W_s$$
pour un processus adapté (H_t) et un mouvement brownien (W_t). Sous certaines conditions d'intégrabilité de H, le processus $t \mapsto (H\cdot W)_t$ sera une martingale (locale).

Dans la Section 4.1 on fait des rappels sur l'intégrale déterministe de Stieljes, qui permet d'intégrer par rapport à un processus à variations finies A et donner sens à des quantités telles que $\int_0^t H_s \mathrm{d}A_s$ en les définissant pour chaque trajectoire (i.e. pour chaque ω) comme la limite des sommes de Riemann $\int_0^t H_s \mathrm{d}A_s = \lim_{N\to\infty}\sum_{i=1}^N H_{t_{i-1}}(A_{t_i} - A_{t_{i-1}})$.

Le mouvement brownien n'étant pas à variation fini, cette approche ne pourra pas fonctionner directement, mais l'idée sera conservée en regardant cette fois ci une limite dans L^2. L'idée pour définir l'intégrale stochastique sera, non pas de définir l'intégrale trajectoire par trajectoire mais plutôt, de regarder directement les limites L^2 des sommes de Riemann. On aura
$$\lim_{N\to\infty}\mathbb{E}\left[\left(\sum_{i=0}^N H_{t_{i-1}}(W_{t_i} - W_{t_{i-1}}) - \int_0^t H_s \mathrm{d}W_s\right)^2\right] = 0.$$

Dans la Section 4.2 on construira l'intégrale stochastique dite de Wiener, c'est à dire du type $\int_0^t h(s)\mathrm{d}W_s$ pour $H_s = h(s)$ une fonction continue déterministe. Dans ce cas là, le processus $(H\cdot W)_t$ sera un processus gaussien et une martingale.

Dans le Section 4.3, on généralisera cette intégrale en donnant sens aux intégrales d'Itô du type $(H\cdot W)_t := \int_0^t H_s \mathrm{d}W_s$ pour un processus stochastique H_s tel que $\mathbb{E}\left(\int_0^T H_s^2 \mathrm{d}s\right) < \infty$. Dans ce cas là, le processus $(H\cdot W)_t$ sera une martingale.

Enfin, dans la Section 4.4, on étendra encore cette définition de l'intégrale stochastique à tout processus tel que $\int_0^T H_s^2 \mathrm{d}s < \infty$ presque surement. Dans ce cas là, le processus $(H\cdot W)_t$ sera une martingale locale (mais pas nécessairement une martingale).

4.1 Intégrale de Stieljes

L'intégrale de Riemann est définie comme une limite de sommes dites de Riemann. Si $(t_i)_{0\leqslant i\leqslant N}$ désigne une subdivision de $[0,t]$ (i.e. $0 = t_0 < t_1 < \ldots < t_N = t$) dont le pas, $\sup_{i=1,\ldots,N}|t_i - t_{i-1}|$, tend vers 0 lorsque N tend vers l'infini, alors pour toute fonction h Riemann-intégrable on a
$$\int_0^t h(s)\mathrm{d}s = \lim_{N\to\infty}\sum_{i=1}^N h(t_{i-1})(t_i - t_{i-1}).$$

L'intégrale de Lebesgue étend l'intégrale de Riemann grâce à la théorie de la mesure en considérant la mesure de Lebesgue λ définie par
$$\lambda(]a,b]) = b - a$$
et on a $\int_0^t h(s)\mathrm{d}\lambda(s) = \int_0^t h(s)\mathrm{d}s$.

Définition 4.1 (Mesure et intégrale de Stieljes)**.** Pour une fonction $F : [0, T] \to \mathbb{R}$ croissante et continue (continue à droite pourrait suffire), on peut associer une unique mesure $\mathrm{d}F$, appelée *mesure de Stieljes* de F, telle que

$$\mathrm{d}F(]a, b]) = F(b) - F(a).$$

L'intégrale par rapport à cette mesure, est appelée *intégrale de Stieljes*. Cette intégrale peut être obtenue comme la limite de sommes de Riemann. Si $(t_i)_{0 \leqslant i \leqslant N}$ désigne une subdivision de $[0, t]$ dont le pas tend vers 0, alors

$$\int_0^t h(s)\mathrm{d}F(s) = \lim_{N \to \infty} \sum_{i=1}^{N} h(t_{i-1})(F(t_i) - F(t_{i-1})). \tag{4.1}$$

En probabilité, cette intégrale est souvent utilisée en prenant F une fonction de répartition.

Définition 4.2 (Variation totale)**.** La *variation totale* d'une fonction F sur $[0, t]$ notée $V_t(F)$ est définie par

$$V_t(F) := \sup_{0=t_0<t_1<\ldots<t_N=t} \sum_{i=1}^{N} |F(t_i) - F(t_{i-1})|.$$

On dit que F est à variation finie sur $[0, t]$ si $V_t(F) < \infty$. On notera $V_{[a,b]}(F)$ la variation sur $[a, b]$.

Toute fonction continûment différentiable vérifie

$$V_t(f) = \int_0^t |F'(s)|\, ds < \infty,$$

et toute fonction monotone vérifie

$$V_t(F) = |F(t) - F(0)| < \infty.$$

Proposition 4.3 (Fonction à variation finie)**.** *Une fonction F est a variation finie si et seulement si il existe deux fonctions croissantes F_1 et F_2 telles que $F = F_1 - F_2$.*

Démonstration. Si $F = F_1 - F_2$ avec F_1, F_2 croissantes alors

$$\sum_k |F(t_{k+1}) - F(t_k)| \leqslant \sum_k |F_1(t_{k+1}) - F_1(t_k)| + \sum_k |F_2(t_{k+1}) - F_2(t_k)|,$$

de sorte que $V_t(F) \leqslant V_t(F_1) + V_t(F_2) < \infty$ et ainsi F est à variation finie. Réciproquement, si F est à variation finie alors les fonctions $F_1 = V_t(F)$ et $F_2 = V_t(F) - F(t)$ sont croissantes. En effet $V_{t+s}(F) - V_t(F) = V_{[t,t+s]}(F) \geqslant 0$, et $V_{t+s}(F) - V_t(F) - (F(t+s) - F(t)) = V_{[t,t+s]}(F) - (F(t+s) - F(t)) \geqslant 0$ car $\{t, t+s\}$ forme une subdivision de l'intervalle $[t, t+s]$. ☐

Si F est à variation borné, on peut alors simplement étendre la notion d'intégrale de Stieljes et définir

$$\int_0^t h(s)\mathrm{d}F(s) := \int_0^t h(s)\mathrm{d}F_1(s) - \int_0^t h(s)\mathrm{d}F_2(s).$$

et l'intégrale peut encore être obtenue comme limite de sommes de Riemann par l'équation (4.1). On dit que h est intégrable par rapport à F si h est intégrable par rapport à F_1 et F_2 (i.e. pour tout t, $\int_0^t |h|\mathrm{d}F_i(s) < \infty$ pour $i = 1, 2$.)

On dit qu'un processus A est à variation fini si toutes ses trajectoires sont à variation finie.

Proposition 4.4 (Intégrale de Stieljes d'un processus stochastique)**.** *Soit H un processus stochastique progressif et A un processus à variation finie, si $H(\omega)$ est bien intégrable par rapport à $A(\omega)$ pour chaque ω, alors le processus*

$$(H \cdot A)_t := \int_0^t H_s \mathrm{d}A_s$$

est un processus adapté à variation finie.

Voir [1, Prop. 4.2] pour la démonstration. Remarquons que ce processus est bien défini sans soucis pour chaque trajectoire (i.e. pour chaque ω) via l'intégrale de Stieljes d'une fonction déterministe.

En particulier, pour toute fonction h continue, $t \mapsto \int_0^t h(W_s)\mathrm{d}s$ est (évidemment) bien défini pour chaque ω et est un processus adapté à variation finie.

4.2 Intégrale de Wiener

Heuristique L'idée première pour construire une intégrale sur le brownien $\int_0^t h(s)\mathrm{d}W_s$ serait de la définir "trajectoire par trajectoire" comme une intégrale de Sieljes par la limite d'une somme de Riemann. Malheureusement, quand on veut contrôler les approximations de Riemann, la variation totale du mouvement brownien intervient et on a vu dans le Corolaire 3.23 qu'elle était infini presque surement, i.e.

$$V_t(B) := \sup_{0=t_0<\ldots<t_N=t} \sum_{i=1}^{N} |W_{t_i} - W_{t_i-1}| = \infty.$$

Ainsi l'intégrale stochastique ne pourra être définie comme limite presque sûre, trajectoire par trajectoire. Comment faire alors ?

Espaces de Hilbert L^2 On se donne un horizon fini $T > 0$ (éventuellement $T = \infty$). On définit l'espace des fonctions de carré intégrable

$$L^2([0,T]) := \left\{ h : \mathbb{R}_+ \to [0,T] \,:\, \int_0^T |h(s)|^2 \,\mathrm{d}s < \infty \right\}.$$

muni de la norme $\|h\|^2_{L^2([0,T])} = \int_0^T |h(s)|^2 \,\mathrm{d}s$ et du produit scalaire $\langle h, g\rangle_{L^2([0,T])} = \int_0^T h(s)g(s)\mathrm{d}s$. Il s'agit d'un espace de Hilbert (espace complet muni d'un produit scalaire). De même on définit l'espace des variables aléatoires de carré intégrable par

$$L^2(\Omega) := \left\{ X : \Omega \to \mathbb{R} \,:\, \mathbb{E}(X^2) < \infty \right\}.$$

Il s'agit aussi d'un espace de Hilbert muni du produit scalaire $\langle X, Y\rangle_{L^2(\Omega)} = \mathbb{E}(XY)$ et de la norme $\|X\|^2_{L^2(\Omega)} = \mathbb{E}(X^2)$.

Construction sur les fonctions étagées On définit l'ensemble des fonctions étagées sur $\mathbb{R}_+$ par

$$\mathcal{E} := \left\{ \sum_{k=1}^{n} a_{k-1} \mathbf{1}_{[t_{k-1},t_k[},\ a_0, \ldots, a_n \in \mathbb{R},\ 0 = t_0 < \cdots < t_n \leqslant T \right\} \subset L^2([0,T]).$$

Définition 4.5 (Intégrale de Wiener sur $\mathcal{E}$)**.** L'*intégrale de Wiener* sur $\mathcal{E}$ est l'application

$$\begin{array}{rccc} I_T : & \mathcal{E} & \longrightarrow & L^2(\Omega) \\ & h & \longmapsto & I_T(h) \end{array}$$

telle que pour tout $h = \sum_{k=1}^n a_{k-1}\mathbf{1}_{[t_{k-1},t_k[} \in \mathcal{E}$, on a

$$I_T(h) := \sum_{k=1}^{n} a_{k-1}(W_{t_k} - W_{t_{k-1}}).$$

En fonction du contexte on trouvera les notations suivantes $I_T(f) = (h \cdot W)_T = \int_0^T h(s)\mathrm{d}W_s$.

Proposition 4.6 (Intégrale gaussienne)**.** *Soit $h \in \mathcal{E}$. L'intégrale de Wiener $I_T(f)$ est une variable aléatoire gaussienne centrée en* 0 *et de variance $\|h\|^2_{L^2([0,T])}$, c'est à dire*

$$\int_0^T h(s)\mathrm{d}W_s \sim \mathcal{N}\left(0\,, \int_0^T h^2(s)\mathrm{d}s\right). \tag{4.2}$$

Démonstration. Remarquons que par le caractère gaussien du Brownien et l'indépendance de ses accroissements, la variable aléatoire $I_T\,(h)$ est une variable gaussienne d'espérance nulle et de variance

$$\mathbb{V}\mathrm{ar}\,[I_T\,(h)] = \sum_{k=1}^{n} a_{k-1}^2 \mathbb{V}\mathrm{ar}(W_{t_k} - W_{t_{k-1}}) = \sum_{k=1}^{n} a_{k-1}^2 (t_k - t_{k-1}) = \int_0^T h^2(s)\mathrm{d}s.$$

□

Proposition 4.7 (Isométrie)**.** *La fonction $I_T : \mathcal{E} \to L^2(\Omega)$ est un isométrie, c'est à dire une application qui conserve la norme, $\|I(h)\|_{L^2(\Omega)} = \|h\|_{L^2([0,T])}$, i.e.*

$$\mathbb{E}\left[\left(\int_0^T h(s)\mathrm{d}W_s\right)^2\right] = \int_0^T h^2(s)\mathrm{d}s. \tag{4.3}$$

Le produit scalaire est donc aussi conservé et on a

$$\mathbb{E}\left[\left(\int_0^T h(s)\mathrm{d}W_s\right)\left(\int_0^T g(s)\mathrm{d}W_s\right)\right] = \int_0^T h(s)g(s)\mathrm{d}s. \tag{4.4}$$

Démonstration. On remarque que $f \mapsto I_T(h)$ est une fonction linéaire au sens où $I_T(ah + bg) = aI_T(h) + bI_T(g)$ pour toutes fonctions h, g en escalier et tous $a, b \in \mathbb{R}$. Enfin, si f et g sont deux fonctions en escalier, on a

$$\begin{aligned}
\mathbb{E}\left[I_T(h)I_T(g)\right] &= \frac{1}{2}\left(\mathbb{V}\mathrm{ar}\left[I_T(h) + I_T(g)\right] - \mathbb{V}\mathrm{ar}\left[I_T(h)\right] + \mathbb{V}\mathrm{ar}\left[I_T(g)\right]\right) \\
&= \frac{1}{2}\left(\int_0^T (f+g)^2(s)\mathrm{d}s - \int_0^T h^2(s)\mathrm{d}s - \int_0^T g^2(s)\mathrm{d}s\right) \\
&= \int_0^T h(s)g(s)\mathrm{d}s.
\end{aligned}$$

□

Extension à $L^2(\mathbb{R}_+)$ Soit $f \in L^2([0,T])$, il existe une suite $(f_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{E}$ telle que

$$\lim_{n\to\infty} \|f_n - f\|_{L^2([0,T])} = 0. \tag{4.5}$$

On admet ce résultat qui permet de prolonger l'intégrale, définie précédemment sur $\mathcal{E}$, à $L^2([0,T])$.

Proposition 4.8 (Intégrale de Wiener sur $L^2([0,T])$)**.** *Il existe une application $L^2([0,T]) \to L^2(\Omega)$ qui prolonge l'application I_T de la Définition 4.5, on la notera encore I_T. L'application*

$$\begin{array}{rccc}
I_T : & L^2([0,T]) & \longrightarrow & L^2(\Omega) \\
& h & \longmapsto & I_T(h) = \int_0^T h(s)\mathrm{d}W_s
\end{array}$$

est encore une isométrie et $I_T(h)$ est encore une gaussienne centrée en 0 et de variance $\|h\|^2_{L^2([0,T])}$. Les formules (4.2), (4.3) et (4.4) et restent donc valide.

Démonstration. Admis. □

Pour tout $t \in [0,T]$ et $f \in L^2([0,T])$, on note

$$I_T(h\mathbf{1}_{[0,t]}) = I_t(h) = \int_0^t h(s)dW_s.$$

Proposition 4.9 (Martingale intégrale)**.** *L'extension de l'intégrale de Wiener à $L^2([0,T])$ vérifie les propriétés suivantes :*

1. *$t \mapsto \int_0^t h(s)\mathrm{d}W_s(\omega)$ est continue sur $\mathbb{R}_+$ pour presque tout $\omega \in \Omega$.*
2. *$\left(\int_0^t h(s)\mathrm{d}W_s\right)_{t\in[0,T]}$ est une martingale de carré intégrable.*

Démonstration. Admis. □

On peut déduire de ce qui précède, et en particulier (4.2) et (4.4) la proposition suivante.

Proposition 4.10 (Processus gaussien)**.** *L'intégrale $I_t(h) = \int_0^t h(s)dW_s$ est un processus gaussien, centré, à accroissement indépendants et de fonction de covariance*

$$\mathbb{C}ov(I_t(h), I_s(h)) = \mathbb{E}\left(\int_0^t h(u)\mathrm{d}W_u \int_0^s h(u)\mathrm{d}W_u\right) = \int_0^{t\wedge s} h^2(u)\mathrm{d}u.$$

4.3 Intégrale d'Itô

Notons l'ensemble

$$\mathcal{H} := \{(H_t) \text{ processus adapté}, \mathbb{E}\left(\int_0^T H_s^2 \mathrm{d}s\right) < \infty\}$$

et munissons d'un norme et d'un produit scalaire

$$\|H\|_{\mathcal{H}} = \sqrt{\mathbb{E}\left(\int_0^T H_s^2 \mathrm{d}s\right)}, \quad \langle H, \widetilde{H}\rangle_{\mathcal{H}} = \mathbb{E}\left(\int_0^T H_s \widetilde{H}_s \mathrm{d}s\right).$$

On admet qu'il s'agit d'un espace vectoriel normé complet (toute suite de Cauchy converge), et comme la norme est issue du produit scalaire on dit alors que c'est un *espace de Hilbert*.

Cela va avoir plusieurs conséquences. La première est, comme on l'a fait pour l'intégrale de Wiener, dans la construction de l'intégrale stochastique : on la définit pour des fonctions H constantes par morceaux qui forment un sous ensemble dense de $\mathcal{H}$. Par une propriété d'isométrie, on montre ensuite que les intégrales stochastiques des fonctions constantes par morceaux forment une suite de Cauchy dans L^2 et on conclut en remarquant que l'espace est complet.

Une autre conséquence du caractère hilbertien de $\mathcal{H}$ apparaîtra dans la résolution des Equations Différentielles Stochastiques où l'existence de solution découlera d'un théorème de point fixe.

Passons maintenant aux propriétés essentielles des intégrales stochastiques définies sur des fonctions de $\mathcal{H}$ et qui généralisent l'intégrale de Wiener vue à la section précédente.

Théorème 4.11 (Intégrale d'Itô dans $\mathcal{H}$)**.** *Pour tout H dans $\mathcal{H}$, il existe un processus noté $(\int_0^t H_s \mathrm{d}W_s)_{t\geqslant 0}$ continue en la variable t et vérifiant de plus,*

1. *Si $(t_i)_{0\leqslant i\leqslant N}$ désigne une subdivision de $[0,t]$ (i.e. $0 = t_0 < t_1 < \ldots < t_N = t$) dont le pas, $\sup_{i=1,\ldots,N} |t_i - t_{i-1}|$, tend vers 0 lorsque N tend vers l'infini, alors on a la convergence L^2 suivante,*
$$\sum_{i=1}^N H_{t_{i-1}}(W_{t_i} - W_{t_{i-1}}) \xrightarrow[N\to\infty]{L^2(\Omega)} \int_0^t H_s \mathrm{d}W_s.$$
2. *$(\int_0^t H_s \mathrm{d}W_s)_{t\geqslant 0}$ est une martingale de carré intégrable, i.e. une martingale L^2.*
3. *L'application $\mathcal{H} \to L^2(\Omega)$ telle que $H \mapsto \int_0^T H_s \mathrm{d}W_s$ est une isométrie, i.e. $\| \int_0^T H_s \mathrm{d}W_s\|_{L^2(\Omega)} = \|H\|_{\mathcal{H}}$. Plus généralement, pour tout $(G,H) \in \mathcal{H}\times\mathcal{H}$, et pour tout t,*
$$\mathbb{E}\left[\left(\int_0^t G_s \mathrm{d}W_s\right)\left(\int_0^t H_s \mathrm{d}W_s\right)\right] = \mathbb{E}\left[\int_0^t G_s H_s \mathrm{d}s\right].$$

Remarque 4.12 (Convergence dans L^2)**.** On notera que la convergence dans i) est dans $L^2(\Omega)$ et non presque sure, cela signifie en d'autres termes que

$$\lim_{N\to\infty} \mathbb{E}\left[\left(\sum_{i=0}^N H_{t_{i-1}}(W_{t_i} - W_{t_{i-1}}) - \int_0^t H_s \mathrm{d}W_s\right)^2\right] = 0.$$

D'autre part, la propriété iii) est une propriété d'isométrie. La propriété iv) vient de l'inégalité maximale pour les martingales continues L^2 et du fait que $(\int_0^t H_s \mathrm{d}W_s)_{t\geqslant 0}$ est une $\mathcal{F}$-martingale (voir ii)).

Remarque 4.13 (Intégrale de Wiener)**.** Un cas particulier important est celui où $H_s = h(s)$ où h est une fonction continue. Alors, $\int_0^t h(s)\mathrm{d}W_s$ est obtenue comme limite L^2 de v.a. gaussiennes donc est une v.a. gaussienne. Ainsi, lorsque $H_s = h(s)$ est une fonction déterministe , $\int_0^t h(s)\mathrm{d}W_s$ suit une loi normale, ce qui n'est pas le cas en général (i.e pour un H quelconque de $\mathcal{H}$).

Remarque 4.14 (Martingales sur $\mathbb{R}$)**.** Dans les définitions précédentes on peut prendre $T = \infty$. Mais, même lorsque $\mathbb{E}\left(\int_0^\infty H_s^2 \mathrm{d}s\right) = \infty$, si $\mathbb{E}\left(\int_0^T H_s^2 \mathrm{d}s\right) < \infty$ pour tout $T > 0$ alors $\int_0^t H_s \mathrm{d}W_s$ est aussi une martingale pour $t \in \mathbb{R}$.

4.4 Extension de l'intégrale stochastique

On pourrait se contenter de définir l'intégrale stochastique sur $\mathcal{H}$. Malheureusement, (et notamment nous le verrons lorsque nous parlerons de la formule d'Itô), la plupart des fonctions que nous aurons à manipuler ne seront pas dans $\mathcal{H}$ mais plutôt dans l'ensemble plus large de fonctions $\widetilde{\mathcal{H}}$ que nous définissons maintenant. Notons

$$\widetilde{\mathcal{H}} = \{(H_t) \text{ processus adapté}, \int_0^T H_s^2 \mathrm{d}s < \infty \ p.s.\}.$$

Évidemment, $\mathcal{H} \subset \widetilde{\mathcal{H}}$.

Il est possible de l'intégrale stochastique sur ce nouvel ensemble de fonctions. Pour tout H dans $\widetilde{\mathcal{H}}$, il existe un processus $(\int_0^t H_s \mathrm{d}W_s)$, continue en la variable t, prolongeant l'intégrale stochastique sur $\mathcal{H}$. Définir une intégrale stochastique sur un ensemble de fonctions plus large appelle quelques questions : notamment on peut se demander quelles sont les propriétés sur $\mathcal{H}$ qui sont conservées dans $\widetilde{\mathcal{H}}$? La réponse est brutale : aucune des propriétés 1., 2., 3. du Théorème 4.11 ne sont maintenant vérifiées pour des fonctions de $\widetilde{\mathcal{H}}$. On peut en effet remarquer que dans la définition de $\widetilde{\mathcal{H}}$, aucun moment d'ordre 2 par rapport à la probabilité $\mathbb{P}$ n'est supposée et c'est donc naturel que les résultats exprimés en moments d'ordre 2, du Théorème 4.11 valables dans $\mathcal{H}$, ne le soient plus dans $\widetilde{\mathcal{H}}$. Mais des propriétés similaires restent encore satisfaites.

1. La limite de la propriété 1 du Théorème 4.11 reste valable mais avec une convergence en probabilité (et non plus une convergence L^2) [1, Prop 5.5].
2. Pour la propriété 2 du Théorème 4.11, lorsque $H \in \widetilde{\mathcal{H}}$, le processus $(\int_0^t H_s \mathrm{d}W_s)_{t \geqslant 0}$ n'est plus forcément une martingale mais il n'en est pas loin. Pour tout H dans $\widetilde{\mathcal{H}}$, le processus $(\int_0^t H_s \mathrm{d}W_s)$ est une *martingale locale*. C'est à dire qu'il existe une suite de temps d'arrêt, (τ_n) vérifiant les deux propriétés suivantes :
   i) $\lim \uparrow_{n \to \infty} \tau_n = \infty$, $\mathbb{P}$-p.s. i.e. (τ_n) est une suite croissante de temps d'arrêt tendant vers l'infini.
   ii) $(\int_0^{t \wedge \tau_n} H_s \mathrm{d}W_s)_{t \geqslant 0}$ est une martingale.
3. La propriété 3 du Théorème 4.11 est remplacé par une égalité sur la (co-)variation quadratique. On rappelle les Définitions 2.16 et 2.18. Il est possible d'étendre la notion de variation quadratique aux martingales locales. Si M est une martingale locale, il existe un unique processus croissant $\langle M \rangle$ tel que $(M_t^2 - \langle M \rangle_T)$ soit une martingale locale.
   Pour tout $(G, H) \in \widetilde{\mathcal{H}} \times \widetilde{\mathcal{H}}$, on a alors $\langle (\int_0^{\cdot} G_s \mathrm{d}W_s), (\int_0^{\cdot} H_s \mathrm{d}W_s) \rangle_t = \int_0^t G_s H_s \mathrm{d}s$.

On rassemble ces trois propriétés dans le théorème suivant (admis) qui fait écho au Théorème 4.11.

Théorème 4.15 (Intégrale stochastique dans $\widetilde{\mathcal{H}}$)**.** *Pour tout H dans $\widetilde{\mathcal{H}}$, il existe un processus noté $(\int_0^t H_s \mathrm{d}W_s)_{t \geqslant 0}$, continue en la variable t, prolongeant l'intégrale stochastique sur $\mathcal{H}$, et vérifiant de plus,*

1. *Si $(t_i)_{0 \leqslant i \leqslant N}$ désigne une subdivision de $[0, t]$ (i.e. $0 = t_0 < t_1 < \ldots < t_N = t$) dont le pas, $\sup_{i=1,\ldots,N} |t_i - t_{i-1}|$, tend vers 0 lorsque N tend vers l'infini, alors on a la convergence en probabilité suivante,*
$$\sum_{i=1}^N H_{t_{i-1}} (W_{t_i} - W_{t_{i-1}}) \xrightarrow[N \to \infty]{\mathbb{P}} \int_0^t H_s \mathrm{d}W_s.$$
2. *$(\int_0^t H_s \mathrm{d}W_s)_{t \geqslant 0}$ est une martingale locale.*
3. *Pour tout $(G, H) \in \widetilde{\mathcal{H}} \times \widetilde{\mathcal{H}}$,*
$$\left\langle \left(\int_0^{\cdot} G_s \mathrm{d}W_s \right), \left(\int_0^{\cdot} H_s \mathrm{d}W_s \right) \right\rangle_t = \int_0^t G_s H_s \mathrm{d}s.$$

Remarque 4.16 (Martingales locales sur $\mathbb{R}$)**.** Dans les définitions précédentes on peut prendre $T = \infty$. Mais, même lorsque $\int_0^\infty H_s^2 \mathrm{d}s = \infty$, si $\int_0^T H_s^2 \mathrm{d}s < \infty$ p.s. pour tout $T > 0$ alors $\int_0^t H_s \mathrm{d}W_s$ est aussi une martingale locale pour $t \in \mathbb{R}$.

Une remarque importante pour la suite est donc que si $H \in \mathcal{H}$ alors le processus $(\int_0^t H_s \mathrm{d}W_s)_{t \geqslant 0}$ est une martingale. Par contre si $H \in \widetilde{\mathcal{H}}$, alors on ne peut rien dire quand au caractère martingale ou pas du processus $(\int_0^t H_s \mathrm{d}W_s)_{t \geqslant 0}$.

Si une propriété sur $\mathcal{H}$ est directement issue de la "martingalité" de $(\int_0^t H_s \mathrm{d}W_s)$, on peut donc tenter de la prolonger sur $\widetilde{\mathcal{H}}$ par l'intermédiaire de la suite de temps d'arrêt τ_n en arguant que $(\int_0^{t \wedge \tau_n} H_s \mathrm{d}W_s)$ est une martingale puis on fait tendre n vers l'infini.

4.5 Intégrale par rapport à une martingale

Il est en fait possible de généraliser tout cette théorie pour définir l'intégrale d'un processus H par rapport à une martingale M (ou même une semi-martingale qui est la somme d'une martingale locale et d'un processus à variation finie)

$$(H \cdot M)_t := \int_0^t H_s \mathrm{d}M_t.$$

Le processus H doit vérifier, soit

$$\mathbb{E}\left(\int_0^T H_s^2 \mathrm{d}\langle M \rangle_s\right) < \infty$$

c'est à dire l'équivalent de ce qu'on a noté $\mathcal{H}$ lorsque $M = W$ est un mouvement brownien, et dans ce cas la l'intégrale $H \cdot M$ est une martingale, soit

$$\left(\int_0^T H_s^2 \mathrm{d}\langle M \rangle_s\right) \infty \ p.s.$$

c'est à dire l'équivalent de ce qu'on a noté $\widetilde{\mathcal{H}}$ lorsque $M = W$ est un mouvement brownien, et dans ce cas la l'intégrale $H \cdot M$ est une martingale locale.

On a naturellement la formule de la variation quadratique

$$\left\langle \int_0^\cdot H_s \mathrm{d}M_s \right\rangle_t = \int_0^t H_s^2 \mathrm{d}\langle M \rangle_s$$

et pour la covariation quadratique

$$\left\langle \int_0^\cdot H_s \mathrm{d}M_s, \int_0^\cdot G_s \mathrm{d}N_s \right\rangle_t = \int_0^t H_s G_s \mathrm{d}\langle M, N \rangle_s.$$

TABLE 4.1 – Propriétés des intégrales

Intégrale	Stieljes	Wiener	Itô	Stochastique
par rapport à	A_t à variation finie	W_t mouvement Brownien		
Écriture	$\int_0^t H_s \mathrm{d}A_s$	$\int_0^t h(s)\mathrm{d}W_s$	$\int_0^t H_s \mathrm{d}W_s$	
Somme de Riemann	Limite p.s.	Limite L^2		Limite en proba
Ensemble de définition	Processus intégrables $\int_0^t \lvert H_s\rvert\lvert \mathrm{d}A_s\rvert < \infty$	$L^2([0,T])$ $\int_0^T \lvert h(s)\rvert^2 \,\mathrm{d}s < \infty$	$\mathcal{H}$ $\mathbb{E}\left(\int_0^T H_s^2 \mathrm{d}s\right) < \infty$	$\widetilde{\mathcal{H}}$ $\int_0^T H_s^2 \mathrm{d}s < \infty$ p.s.
Type de processus	Variation finie	Martingale		Martingale locale
		Processus gaussien		
Variation quadratique	$\langle \int_0^\cdot H_s \mathrm{d}A_s \rangle_t = 0$	$\langle \int_0^\cdot h(s)\mathrm{d}W_s \rangle_t = \int_0^t h(s)^2 \mathrm{d}s$	$\langle \int_0^\cdot H_s \mathrm{d}W_s \rangle_t = \int_0^t H_s^2 \mathrm{d}s$	
Isométrie		$\mathbb{E}\left[\left(\int_0^T h(s)\mathrm{d}W_s\right)^2\right] = \int_0^T h^2(s)\mathrm{d}s$	$\mathbb{E}\left[\left(\int_0^T H_s \mathrm{d}W_s\right)^2\right] = \mathbb{E}\left[\int_0^T H_s^2 \mathrm{d}s\right]$	

Les propriétés de ce tableau restent vérifiées en remplaçant le mouvement brownien W par une martingale (locale) M et tous les ds par des d$\langle M\rangle_s$.

Chapitre 5

Calcul d'Itô

Ce chapitre a pour objectif de déterminer les nouvelles règles de calcul qui permettent d'étudier et de calculer des intégrales stochastique par rapport à un mouvement brownien et plus généralement par rapport à des processus d'Itô. Avec une intégrale stochastique, il ne suffit pas de calculer la primitive comme c'est le cas pour les intégrales classiques. Il faut aller au deuxième ordre du développement de Taylor, ce qui rajoute un terme. Ainsi $\int_0^t W_s \mathrm{d}W_s$ ne sera pas égale à $W_t^2/2$ mais à $W_t^2/2 - t/2$. La formule principale de ce chapitre est la formule d'Itô, qui sera énoncé en toute généralité dans les sections suivantes, et qui donne dans sa forme la plus simple la formule

$$f(W_t) = f(0) + \int_0^t f'(W_s)\mathrm{d}W_s + \frac{1}{2}\int_0^t f''(W_s)\mathrm{d}s.$$

5.1 Processus d'Itô et crochets

Définition 5.1 (Processus d'Itô)**.** On suppose les 4 conditions suivantes réunies :

1. X_0 est $\mathcal{F}_0$-mesurable.
2. $\mu = (\mu_s)$ un processus adapté et $\int_0^T |\mu_s|\mathrm{d}s < \infty$ p.s.
3. $\sigma = (\sigma_s)$ un processus adapté et $\int_0^T \sigma_s^2 \mathrm{d}s < \infty$ p.s., i.e. $\sigma \in \widetilde{\mathcal{H}}$.
4. (W) est un mouvement brownien standard.

On dit que le processus $(X_t)_{0 \leqslant t \leqslant T}$ est un *processus d'Itô* si

$$\text{p.s.,} \quad \forall t \leqslant T, \quad X_t = X_0 + \int_0^t \mu_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s \quad \text{(Ecriture intégrale),}$$

ce qu'on note de façon infinitésimale,

$$\forall t \leqslant T, \quad \mathrm{d}X_t = \mu_t \mathrm{d}t + \sigma_t \mathrm{d}W_t \quad \text{(Ecriture différentielle).}$$

Dans ce cas, le processus (μ_s) sera appelé *dérive* (en français) ou *drift* (en anglais) de X_t et (σ_s) son coefficient de *diffusion* (ou encore, terme de *volatilité*). De plus, la décomposition de (X_t) à l'aide du triplet $(X_0, (\mu), (\sigma))$ vérifiant 1, 2, 3 est **unique**.

On notera que les hypothèses de la définition précédente nous permettent bien de définir proprement les intégrales $\int_0^t \mu_s \mathrm{d}s$ (à variation finie) et $\int_0^t \sigma_s \mathrm{d}W_s$ (martingale locale) et d'affirmer que X est adapté.

Remarque 5.2 (Intégrale stochastique par rapport à un processus d'Itô)**.** Il est naturel de définir l'intégrale d'un processus H, adapté par rapport au processus d'Itô X, par :

$$\int_0^T H_s \mathrm{d}X_s := \int_0^T H_s\left(\mu_t \mathrm{d}t + \sigma_t \mathrm{d}W_t\right) = \int_0^T H_s \mu_s \mathrm{d}s + \int_0^T H_s \sigma_s \mathrm{d}W_s$$

C'est effectivement possible à condition de supposer que les deux termes à droite de l'égalité soient bien définies, il faut donc s'assurer que les deux conditions supplémentaires soient remplies :

1. $\int_0^T |H_s\mu_s|\mathrm{d}s < \infty$, p.s.
2. $\int_0^T (H_s\sigma_s)^2\mathrm{d}s < \infty$, p.s., i.e. $H\sigma \in \widetilde{\mathcal{H}}$.

Proposition 5.3. *Soit X est un processus d'Itô tel que $\sigma \in \mathcal{H}$. Alors, X est une martingale si et seulement si μ_t est le processus nul.*

Démonstration. Evidemment, si μ_t est le processus nul, alors $\mathrm{d}X_t = \sigma_t\mathrm{d}W_t$ et comme $\sigma \in \mathcal{H}$ il vient directement que X est une martingale. Il reste à montrer l'implication en sens inverse. Pour cela, supposons que X est une martingale. On ne démontrera cette implication que pour un processus (μ_t) tel que $\mathbb{E}\int_0^t |\mu_s|\mathrm{d}s < \infty$. Alors, comme $\int_0^t \mu_s\mathrm{d}s = \int_0^t \mathrm{d}X_s - \int_0^t \sigma_s\mathrm{d}W_s$ est différence de deux martingales, c'est encore une martingale. Pour tout $t \geqslant u$, $\mathbb{E}(\int_u^t \mu_s\mathrm{d}s|\mathcal{F}_u) = 0 = \int_u^t \mathbb{E}(\mu_s|\mathcal{F}_u)\mathrm{d}s$ ce qui donne par dérivation par rapport à t : $\mathbb{E}(\mu_t|\mathcal{F}_u) = 0$. En prenant $t = u$, il vient puisque μ étant $\mathcal{F}$-adapté, (donc μ_t est $\mathcal{F}_t$-mesurable) $\mathbb{E}(\mu_t|\mathcal{F}_t) = \mu_t = 0$. On aurait aussi pu utiliser la Proposition 3.24 qui dit qu'une martingale à variation finie est constante et deduire que comme $\int_0^t \mu_s\mathrm{d}s$ est à variation finie, si c'est une martingale c'est alors la constante nulle et donc $(\mu) = 0$. □

Le résultat précédent permet de montrer simplement l'unicité de la décomposition du processus d'Itô.

Nous allons maintenant définir le crochet ou variation quadratique d'un processus d'Itô comme la variation quadratique de la partie martingale de sa décomposition. La partie à variation finie n'est pas prise en compte dans la variation quadratique.

Définition 5.4 (Covariation quadratique ou crochet d'un processus d'Itô)**.** Soient deux processus d'Itô (X^1) et (X^2) définis sur le même brownien (W).

$$\mathrm{d}X_t^i = \mu_t^i\mathrm{d}t + \sigma_t^i\mathrm{d}W_t, \quad (i \in \{1,2\}).$$

On appelle *covariation quadratique ou crochet* entre X^1 et X^2 défini par

$$\langle X^1, X^2\rangle_t = \langle \sigma^1 W, \sigma^2 W\rangle_t = \int_0^t \sigma_s^1\sigma_s^2\mathrm{d}s$$

et on note alors $\mathrm{d}\langle X^1, X^2\rangle_t = \sigma_s^1\sigma_s^2\mathrm{d}s$. Si (X) est un processus d'Itô telle que $\mathrm{d}X_t = \mu_t\mathrm{d}t + \sigma_t\mathrm{d}W_t$, la *variation quadratique* ou *crochet* de X est donc défini par $\langle X\rangle_t = \langle X, X\rangle_t$ et $\mathrm{d}\langle X\rangle_t = \mathrm{d}\langle X, X\rangle_t = \sigma_t^2\mathrm{d}t$.

Contrairement à la Proposition 3.24 qui dit qu'une martingale de variation quadratique nulle est constante, si $\langle X, X\rangle = 0$, le processus d'Itô X n'est pas forcément constant. Voir les exemples ci dessous.

Exemple 5.5. $\langle W\rangle_t = t, \quad \langle \int_0^\cdot \mu_s\mathrm{d}s\rangle_t = 0, \quad \langle \int_0^\cdot \mu_s\mathrm{d}s, X\rangle_t = 0.$

Règles de calcul (peu rigoureuses, mais pratiques !)

$\langle\cdot,\cdot\rangle$	$\mathrm{d}t$	$\mathrm{d}W_t$
$\mathrm{d}t$	0	0
$\mathrm{d}W_t$	0	$\mathrm{d}t$

Remarque 5.6 (Propriétés du crochet)**.** Les propriétés suivantes vrais pour les martingales restent aussi valable pour (X) et (Y) deux processus d'Itô.

1. $\langle X\rangle$ est un processus croissant positif.
2. $\langle X, Y\rangle_t$ est bilinéaire, symétrique. En particulier, $\langle aX, aX\rangle_t = a^2\langle X, X\rangle_t$ pour tout $a \in \mathbb{R}$.
3. $\langle X, Y\rangle = \frac{\langle X+Y\rangle - \langle X-Y\rangle}{4}$
4. Si X est un processus d'Itô satisfaisant $X_0 = 0$, $\mu = 0$ et $\sigma \in \mathcal{H}$ alors X_t est une martingale et par le Théorème 2.16 le processus

$$(X_t^2 - \langle X\rangle_t)_{t\geqslant 0} = \left(\left(\int_0^t \sigma_s\mathrm{d}W_s\right)^2 - \int_0^t \sigma_s^2\mathrm{d}s\right)_{t\geqslant 0}$$

est une martingale.

5.2 Formules d'Itô pour les processus d'Itô

Si X_t était un processus dérivable (ce qui n'est en général pas le cas !) on aurait la formule $(f \circ X)' = (f' \circ X)X'$ et par la formule classique d'intégration on aurait

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)X_s'\mathrm{d}s = f(X_0) + \int_0^t f'(X_s)\mathrm{d}X_s.$$

Mais cette formule est FAUSSE dans le cas général d'une diffusion. A cause du caractère quadratique de la partie martingale de X, on a besoin d'aller au second terme du développement de Taylor pour obtenir une formule correcte. C'est ce que fait la formule d'Itô !

Théorème 5.7 (Première formule d'Itô)**.** *Soient (X) un processus d'Itô et f une fonction de classe $\mathcal{C}^2$. Alors,*

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)\mathrm{d}X_s + \frac{1}{2}\int_0^t f''(X_s)\mathrm{d}\langle X\rangle_s.$$

Sous la forme différentielle la formule d'Itô s'écrit,

$$\mathrm{d}f(X_t) = f'(X_t)\mathrm{d}X_t + \frac{1}{2}f''(X_t)\mathrm{d}\langle X\rangle_t.$$

Le fait que X soit un processus d'Itô et que f soit $\mathcal{C}^2$ suffit à garantir que les intégrales sont bien définies.

Comme $\mathrm{d}X_s = \mu_s\mathrm{d}s + \sigma_s\mathrm{d}W_s$ et $\mathrm{d}\langle X\rangle_s = \sigma_s^2\mathrm{d}s$ la formule d'Itô se réécrit comme

$$\begin{aligned}
f(X_t) &= f(X_0) + \int_0^t f'(X_s)(\mu_s\mathrm{d}s + \sigma_s\mathrm{d}W_s) + \frac{1}{2}\int_0^t f''(X_s)\sigma_s^2\mathrm{d}s \\
&= f(X_0) + \int_0^t \left(\mu_s f'(X_s) + \frac{1}{2}\sigma_s^2 f''(X_s)\right)\mathrm{d}s + \int_0^t \sigma_s f'(X_s)\mathrm{d}W_s.
\end{aligned}$$

Ainsi, $f(X_t)$ est encore un processus d'Itô de dérive $\mu_s f'(X_s) + \frac{1}{2}\sigma_s^2 f''(X_s)$ et de diffusion $\sigma_s f'(X_s)$. La première intégrale avec la terme de dérive est la partie à variation finie, la deuxième intégrale avec le terme de diffusion est la partie martingale locale (ou martingale si $\sigma_s f'(X_s) \in \mathcal{H}$).

Pour un mouvement brownien standard W la formule d'Itô donne donc

$$f(W_t) = f(0) + \int_0^t f'(W_s)\mathrm{d}W_s + \frac{1}{2}\int_0^t f''(W_s)\mathrm{d}s.$$

Démonstration de la formule d'Itô. On se contentera ici de donner les lignes directrices de la démonstration en se restreignant au cas où l'on considère W un mouvement brownien standard et f une fonction $\mathcal{C}^2$ à support compact. La preuve est repose sur l'écriture suivante :

$$f(W_t) - f(0) = \sum_{i=1}^n f(W_{t_i}) - f(W_{t_{1-1}})$$

où $(t_i)_{i=0,\dots,n}$ une subdivision de $[0,t]$ définie par $t_i = \frac{it}{n}$. Il suffit ensuite d'utiliser une formule de Taylor sur les différences $f(W_{t_1}) - f(W_{t_{1-1}})$ pour arriver à nos fins. Rappelons ainsi que,

$$f(y) - f(x) = (y-x)f'(x) + \frac{1}{2}(y-x)^2 f''(x) + r(x,y)$$

avec le reste intégral $r(x,y)$ qui vérifie $|r(x,y)| \leqslant (x-y)^2 h(x,y)$ pour une fonction h continue bornée telle que $h(x,x) = 0$. On a alors $f(W_t) - f(0) = A_n + B_n + C_n$ avec :

1. $A_n = \sum_{i=1}^n f'\left(W_{t_{t-1}}\right)\left(W_{t_1} - W_{t_{i-1}}\right)$
2. $B_n = \frac{1}{2}\sum_{i=1}^n f''\left(W_{t_{i-1}}\right)\left(W_{t_1} - W_{t_{i-1}}\right)^2$
3. et enfin le reste C_n qui vérifie l'inégalité suivante $|C_n| \leqslant \sum_{i=1}^n \left(W_{t_1} - W_{t_{t-1}}\right)^2 h\left(W_{t_{1-1}}, W_{t_1}\right)$.

Il est facile de traiter le premier terme en utilisant les sommes de Riemann. En effet, nous savons (voir Théorème 4.15) que

$$\lim_{n\to+\infty} A_n \overset{\mathbb{P}}{=} \int_0^t f'(W_s)\, dW_s.$$

Traitons maintenant le deuxième terme en écrivant B_n sous la forme suivante

$$B_n = \frac{1}{2}\sum_{i=1}^{n} f''(W_{t_{i-1}})(t_i - t_{i-1}) + \frac{1}{2}\sum_{i=1}^{n} f''(W_{t_{i-1}})\left[(W_{t_i} - W_{t_{i-1}})^2 - (t_i - t_{i-1})\right].$$

Puisque f'' est continue, le premier terme de la somme ci-dessus converge presque sûrement (comme une somme classique de Riemann) vers

$$\frac{1}{2}\int_0^t f''(W_s)\, ds.$$

Il reste à montrer que le 2ème terme de la somme converge vers zéro. Pour cela, il est possible de faire une démonstration qui s'inspire de la preuve de la Proposition 3.22 qui établie la variation quadratique du mouvement brownien, à savoir $\sum_{i=1}^n \left[(W_{t_i} - W_{t_{i-1}})^2 - (t_i - t_{i-1})\right] = \sum_{i=1}^n (W_{t_i} - W_{t_{i-1}})^2 - t \to 0$. Il ne resterait ensuite plus qu'à montrer que C_n converge vers zéro en probabilité. □

La deuxième formule d'Itô fait intervenir le temps en première variable.

Théorème 5.8 (Deuxième formule d'Itô)**.** *Soient (X) un processus d'Itô et $f : (t,x) \mapsto f(t,x)$ une fonction de classe de classe $\mathcal{C}^1$ par rapport à t et de classe $\mathcal{C}^2$ par rapport à x. Alors,*

$$f(t, X_t) = f(0, X_0) + \int_0^t f'_t(s, X_s)\mathrm{d}s + \int_0^t f'_x(s, X_s)\mathrm{d}X_s + \frac{1}{2}\int_0^t f''_{xx}(X_s)\mathrm{d}\langle X\rangle_s$$

ce qui donne sous la forme différentielle,

$$\mathrm{d}f(t, X_t) = f'_t(t, X_t)\mathrm{d}t + f'_x(t, X_t)\mathrm{d}X_t + \frac{1}{2}f''_{x,x}(t, X_t)\mathrm{d}\langle X\rangle_t$$

avec les notations classiques $f'_t = \frac{\partial f}{\partial t}$, $f'_x = \frac{\partial f}{\partial x}$ et $f''_{x,x} = \frac{\partial^2 f}{\partial x^2}$.

5.3 Formule d'Itô multidimensionnelle

On va définir ici des processus d'Itô vectoriel, (X_t) où X_t **est un vecteur de** $\mathbb{R}^n$ définis sur le brownien vectoriel (W_t) où W_t **est un vecteur de** $\mathbb{R}^p$. On notera que n et p ne sont pas nécessairement égales.

Définition 5.9 (Mouvement brownien multidimensionnel)**.** Un processus (W) à valeurs dans $\mathbb{R}^p$ est un mouvement brownien p-dimensionnel s'il s'écrit sous la forme $W_t = (W_t^1, \cdots, W_t^p)$ où les (W_t^ℓ) sont des mouvements browniens standards indépendants.

Par indépendance des coordonnées, on a via Remarque 2.20 que

$$\langle W^i, W^j\rangle_t = \mathbb{1}_{\{i=j\}} \times t \quad \text{ou encore} \quad \mathrm{d}\langle W^i, W^j\rangle_t = \mathbb{1}_{\{i=j\}}\, \mathrm{d}t \tag{5.1}$$

Définition 5.10 (Processus d'Itô vectoriel)**.** On définit $(X_t) = (X_t^1, \cdots, X_t^p)$ un processus d'Itô vectoriel par

$$\forall i \in \{1, \ldots, n\}, \quad \mathrm{d}X_t^i = \mu_t^i \mathrm{d}t + \sum_{\ell=1}^{p} \sigma_s^{i,\ell}\mathrm{d}W_s^\ell$$

où les coefficients μ^i et $\sigma^{i,\ell}$ vérifient les conditions d'intégrabilités de la Définition 5.1.

Nous pouvons alors obtenir la covariation quadratique entre X^i et X^j :

$$\mathrm{d}\langle X^i, X^j\rangle_t = \sum_{\ell=1}^{p} \sigma_t^{i,\ell}\sigma_t^{j,\ell}\mathrm{d}t. \tag{5.2}$$

Théorème 5.11 (Formule d'Itô vectorielle)**.** *Soient* (X_t) *un processus d'Itô vectoriel et* $(t,x) \mapsto f(t,x)$ *une fonction à valeurs réelles de classe de classe* $\mathcal{C}^1$ *par rapport à* t *et de classe* $\mathcal{C}^2$ *par rapport à* $x \in \mathbb{R}^n$. *Alors,*

$$f(t, X_t) = f(0, X_0) + \int_0^t f_t'(s, X_s)\mathrm{d}s + \sum_{i=1}^n \int_0^t f_{x_i}'(s, X_s)\mathrm{d}X_s^i + \frac{1}{2} \sum_{i,j=1}^n \int_0^t f_{x_i x_j}''(X_s)\mathrm{d}\langle X^i, X^j \rangle_s$$

ce qui donne sous la forme différentielle,

$$\mathrm{d}f(t, X_t) = f_t^{'}(t, X_t)\mathrm{d}t + \sum_{i=1}^n f_{x_i}^{'}(t, X_t^i)\mathrm{d}X_t + \frac{1}{2} \sum_{i,j=1}^n f_{x_i,x_j}^{''}(t, X_t)\mathrm{d}\langle X^i, X^j \rangle_t.$$

Un conséquence quasi immédiate de la formule d'Itô vectorielle est l'intégration par parties.

Théorème 5.12 (Intégration par partie stochastique)**.** *Soient* (X) *et* (Y) *deux processus d'Itô réels (définis sur le même brownien p-dimensionnel). Alors*

$$\int_0^t X_s \mathrm{d}Y_s = X_t Y_t - X_0 Y_0 - \int_0^t Y_s \mathrm{d}X_s - \langle X, Y \rangle_t.$$

Démonstration. On pose $f(x, y) = xy$ et on applique la formule d'Itô à f pour le processus (X_s, Y_s) :

$$\begin{aligned}\mathrm{d}f(X_t, Y_t) = &\frac{\partial f(X_t, Y_t)}{\partial x}\mathrm{d}X_t + \frac{\partial f(X_t, Y_t)}{\partial y}\mathrm{d}Y_t \\ &+ \frac{1}{2}\left(\frac{\partial^2 f(X_t, Y_t)}{\partial x^2}\mathrm{d}\langle X \rangle_t + 2\frac{\partial^2 f(X_t, Y_t)}{\partial x \partial y}\mathrm{d}\langle X, Y \rangle_t + \frac{\partial^2 f(X_t, Y_t)}{\partial y^2}\mathrm{d}\langle Y \rangle_t\right)\end{aligned}$$

Après un calcul rapide des dérivées de f, il reste : $\mathrm{d}(X_t Y_t) = Y_t \mathrm{d}X_t + X_t \mathrm{d}Y_t + d\langle X, Y \rangle_t$ $\square$

Chapitre 6

Équations différentielles stochastiques

Une équation différentielle stochastique (EDS) s'écrit sous la forme :

$$X_0 = x, \quad \mathrm{d}X_t = \mu(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}W_t \tag{6.1}$$

Définition 6.1 (Équation différentielle stochastique)**.** Un processus X est *solution de l'EDS* (6.1) sur $[0, T]$ si X est un processus adapté satisfaisant

$$\int_0^T |\mu(t, X_t)|\mathrm{d}t + \int_0^T \sigma^2(t, X_t)\mathrm{d}t < \infty, \quad \text{p.s.} \tag{6.2}$$

et

$$\forall t \leqslant T, \quad X_t = x + \int_0^t \mu(s, X_s)\mathrm{d}s + \int_0^t \sigma(s, X_s)\mathrm{d}W_s.$$

La condition (6.2) permet l'intégrabilité des intégrales.

6.1 Exemples d'EDS

Mouvement brownien géométrique

Cherchons à résoudre l'équation différentielle stochastique dite de *Black et Scholes*,

$$dX_t = \mu X_t dt + \sigma X_t dB_t \quad \text{avec} \quad X_0 = x_0$$

où $\mu \in \mathbb{R}$ et $\sigma > 0$ sont des constantes. Le plus naturel semble de supposer qu'une solution de cette E.D.S. soit de la forme $X_t = f(t, B_t)$. Ainsi, en utilisant la formule d'Itô nous savons que

$$dX_t = \left[\partial_t f(t, B_t) + \frac{1}{2}\partial^2_{xx}(t, B_t)\right] dt + \partial_x f(t, B_t)\, dB_t$$

Il est alors raisonnable d'identifier les coefficients apparaissant dans ces deux expressions de dX_t. C'est-à-dire, nous devons trouver une fonction $f(t, x)$ telle que

$$\mu f(t, x) = \partial_t f(t, x) + \frac{1}{2}\partial^2_{xx} f(t, x) \quad \text{et} \quad \sigma f(t, x) = \partial_x f(t, x)$$

Puisque la deuxième partie s'écrit sous la forme $\frac{\partial_x f}{f} = \sigma$ nous savons alors que $f(t, x) = \exp(\sigma x + g(t))$ avec une fonction g à déterminer. En remplaçant cette expression de f dans la première partie, nous trouvons que $g'(t) = \mu - \frac{1}{2}\sigma^2$. En conclusion, nous avons donc obtenu

$$X_t = x_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2 t\right) + \sigma B_t\right)$$

Ce processus, très utilisé en finance dans le modèle de Black et Scholes, porte le nom de mouvement brownien géométrique.

Le processus d'Ornstein-Uhlenbeck

La solution de l'E.D.S. suivante

$$dX_t = -\alpha X_t dt + \sigma dB_t \quad \text{avec} \quad X_0 = x_0$$

où $\alpha > 0$ et $\sigma > 0$ correspond à un processus gaussien très important. Voyons de quelle manière nous pouvons obtenir une expression de celui-ci.

Supposons que la solution de l'E.D.S. soit de la forme $X_t = a(t)\left(x_0 + \int_0^t b(s)dB_s\right)$ avec $a(\cdot)$ et $b(\cdot)$ des fonctions dérivables à déterminer. En utilisant la règle de dérivation stochastique d'un produit, nous obtenons que

$$dX_t = a'(t)\left(x_0 + \int_0^t b(s)dB_s\right)dt + a(t)b(t)dB_t$$

En d'autres termes, si nous supposons que $a(0) = 1$ et $a(t) > 0$ pour tout $t > 0$, le processus X_t est solution de l'équation

$$dX_t = \frac{a'(t)}{a(t)} X_t dt + a(t)b(t)dB_t, \quad \text{avec} \quad X_0 = x_0$$

Par identification, nous devons alors résoudre

$$\frac{a'(t)}{a(t)} = -\alpha \quad \text{et} \quad a(t)b(t) = \sigma$$

L'unique solution de la première équation différentielle, sous la contrainte $a(0) = 1$ est $a(t) = e^{-\alpha t}$ et en conséquent $b(t) = \sigma e^{\alpha t}$. Autrement dit

$$X_t = e^{-\alpha t}\left(x_0 + \sigma \int_0^t e^{\alpha s} dB_s\right) = x_0 e^{-\alpha t} + \sigma \int_0^t e^{-\alpha(t-s)} dB_s$$

Cette écriture permet de constater plusieurs choses. L'influence du point de départ x_0 diminue exponentiellement avec le temps lorsque $t \to +\infty$. Puisque l'intégrande apparaissant dans l'intégrale stochastique est déterministe nous avons affaire à un processus gaussien à accroissement indépendant d'après la proposition 4.10. De plus

$$\mathbb{E}[X_t] = x_0 e^{-\alpha t} \quad \text{et} \quad \operatorname{Var}(X_t) = \sigma^2 \int_0^t e^{-2\alpha(t-s)} ds = \frac{\sigma^2}{2\alpha}\left(1 - e^{-2\alpha t}\right).$$

6.2 Théorème de Cauchy-Lipschitz stochastique

Le théorème suivant révèle des conditions suffisantes sous lesquelles l'EDS (6.1) admet une unique solution dans $\mathcal{H}$.

Théorème 6.2 (Cauchy-Lipschitz stochastique)**.** *Si les fonctions μ et σ satisfont les conditions*

i) Il existe $K > 0$ tel que pour tout $(t, x, y) \in [0, T] \times \mathbb{R}^2$, on ait :

$$|\mu(t, x) - \mu(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leqslant K|x - y|$$

ii) Il existe $L > 0$ tel que pour tout $(t, x) \in [0, T] \times \mathbb{R}$, on ait :

$$|\mu(t, x)| + |\sigma(t, x)| \leqslant L(1 + |x|)$$

alors l'EDS (6.1) *admet une unique solution dans $\mathcal{H}$.*

La démonstration de ce théorème s'appuie sur un principe de point fixe et s'approche donc de la manière dont est montré l'existence et l'unicité de solutions d'équations différentielles.

Démonstration. Commençons par l'existence. Nous allons munir $\mathcal{H}$ de la norme $\|\cdot\|_c$ définie par : $\|X\|_c = \sqrt{\mathbb{E}(\int_0^T \mathrm{e}^{-ct}X_t^2\mathrm{d}t)}$ pour tout $X \in \mathcal{H}$, où c est un réel positif ou nul. Rappelons nous que $(\mathcal{H}, \|\cdot\|)$ est un espace de Hilbert où $\|X\| = \sqrt{\mathbb{E}(\int_0^T X_t^2\mathrm{d}t)}$. $(\mathcal{H}, \|\cdot\|_c)$ sera alors aussi un espace de Hilbert pourvu que les deux normes $\|\cdot\|$ et $\|\cdot\|_c$ soient équivalentes. C'est effectivement le cas puisqu'un calcul élémentaire montre que

$$\mathrm{e}^{-cT/2}\|X\| \leqslant \|X\|_c \leqslant \|X\|$$

Définissons maintenant la fonction Φ de source dans $\mathcal{H}$ par

$$\forall t \leqslant T, \quad \Phi(X)_t = x + \int_0^t \mu(s, X_s)\mathrm{d}s + \int_0^t \sigma(s, X_s)\mathrm{d}W_s \quad (X \in \mathcal{H})$$

La condition ii) permet de montrer que Φ est bien définie (c'est à dire que les intégrales ont bien un sens) et que Φ laisse stable $\mathcal{H}$, i.e. $\Phi(X) \in \mathcal{H}$ pour $X \in \mathcal{H}$. Nous laissons cette vérification au lecteur pour nous concentrer à présent sur la propriété de contraction de Φ qu'on montrera grâce à la condition i). Plus précisément, on va vérifier que la fonction $\Phi : \mathcal{H} \to \mathcal{H}$ est une application contractante pour la norme $\|\cdot\|_c$ en choisissant c suffisamment grand. On a en utilisant $(a+b)^2 \leqslant 2(a^2+b^2)$ et la condition i) du Théorème 6.2,

$$\begin{aligned}
\|\Phi(X) - \Phi(Y)\|_c &= \mathbb{E}\left(\int_0^T \mathrm{e}^{-ct}[\Phi(X)_t - \Phi(Y)_t]^2\mathrm{d}t\right) \\
&\leqslant 2\int_0^T \mathrm{e}^{-ct}\mathbb{E}\left(\left[\int_0^t (\mu(s, X_s) - \mu(s, Y_s))\mathrm{d}s\right]^2\right)\mathrm{d}t \\
&\quad +2\int_0^T \mathrm{e}^{-ct}\mathbb{E}\left(\left[\int_0^t (\sigma(s, X_s) - \sigma(s, Y_s))\mathrm{d}W_s\right]^2\right)\mathrm{d}t \\
&\leqslant 2K^2(T+1)\int_0^T \mathrm{e}^{-ct}\mathbb{E}\left(\int_0^t (X_s - Y_s)^2\mathrm{d}s\right)\mathrm{d}t \\
&\leqslant 2K^2(T+1)\mathbb{E}\left(\int_0^T (X_s - Y_s)^2\left(\int_s^T \mathrm{e}^{-ct}\mathrm{d}t\right)\mathrm{d}s\right) \\
&\leqslant 2K^2(T+1)\mathbb{E}\left(\int_0^T (X_s - Y_s)^2\mathrm{e}^{-cs}\left(\frac{1 - \mathrm{e}^{-c(T-s)}}{c}\right)\mathrm{d}s\right) \\
&\leqslant \frac{2K^2(T+1)}{c}\|X - Y\|_c
\end{aligned}$$

En choisissant c suffisamment grand, Φ est une application α-lipschitzienne avec $\alpha < 1$ (on dit alors que Φ est une application contractante) et le théorème du point fixe dans un espace de Hilbert assure alors de l'existence d'une solution, i.e. d'un point fixe tel que $\Phi(X) = X$. L'unicité est encore une conséquence de la contraction puisque si (X) et (Y) sont deux solutions de l'EDS, ce sont des points fixes de Φ et l'inégalité

$$\|X - Y\|_c = \|\Phi(X) - \Phi(Y)\|_c \leqslant \alpha\|X - Y\|_c \quad (\alpha < 1)$$

montre que $X = Y$. $\square$

Travaux dirigés

Mouvement Brownien

Exercice II.1 (Quizz)**.** Soit $B = \{B_t\}_{t \geqslant 0}$ un processus gaussien à trajectoires continues, tel que $B_t \sim \mathcal{N}(0, t)$ pour tout $t \geqslant 0$. Peut-on conclure que B est un mouvement brownien ? Si oui montrez le sinon donnez un contre exemple.

Exercice II.2 (Densité et mesure de Wiener d'un cylindre)**.** Soient W un mouvement brownien, $0 < t_1 < \cdots < t_n$ et $A_1, \cdots, A_n \in \mathcal{B}(\mathbb{R})$. Montrer que

$$\mathbb{P}(B_{t_1} \in A_1, \cdots, B_{t_n} \in A_n) = \int_{A_1 \times \cdots \times A_n} \frac{\mathrm{d}x_1 \cdots \mathrm{d}x_n}{(2\pi)^{n/2} \sqrt{t_1(t_2 - t_1) \cdots (t_n - t_{n-1})}} \exp\left(-\frac{1}{2} \sum_{i=1}^{n} \frac{(x_i - x_{i-1})^2}{t_i - t_{i-1}} \right).$$

où on prend $x_0 = 0$ et en déduire la densité $p(x_1, \cdots, x_n)$ du vecteur aléatoire $(B_{t_1}, \cdots, B_{t_n})$.

Exercice II.3 (Transformations)**.** Soit $\{B_t\}_{t \geqslant 0}$ un mouvement brownien et $a > 0$.

1. Montrer que $\{-B_t\}_{t \geqslant 0}$ est un mouvement brownien.
2. Montrer que $\{B_{a+t} - B_a\}_{t \geqslant 0}$ est un mouvement brownien indépendant de $\{B_t\}_{0 \leqslant t \leqslant a}$.
3. Montrer que $\left\{\frac{B_{at}}{\sqrt{a}}\right\}_{t \geqslant 0}$ est un mouvement brownien.
4. Montrer que $\left\{(1+t) B_{\frac{t}{1+t}} - tB_1\right\}_{t \geqslant 0}$ est un mouvement brownien.
5. Montrer que $\left\{tB_{\frac{1}{t}} 1_{(t>0)}\right\}_{t \geqslant 0}$ est un mouvement brownien.

Indication : On pourra dans un premier temps montrer qu'il suffit de contrôler la covariance entre deux temps (ou la variance d'un incrément) et la continuité.

Exercice II.4 (Pont brownien)**.** Un pont brownien est un processus gaussien $\{Z_t\}_{0 \leqslant t \leqslant 1}$ centré, à trajectoires continues et de fonction de covariance $\Gamma(s, t) = s \wedge t - st$.

1. Vérifier que la loi d'un pont brownien est invariante par retournement du temps : si $\{Z_t\}_{0 \leqslant t \leqslant 1}$ est un pont brownien, alors $\{Z_{1-t}\}_{0 \leqslant t \leqslant 1}$ aussi.
2. Soit $\{Z_t\}_{0 \leqslant t \leqslant 1}$ un pont brownien. Que dire du processus $\left\{(1+t) Z_{\frac{t}{1+t}}\right\}_{t \geqslant 0}$?
3. Soit $B = \{B_t\}_{t \geqslant 0}$ un mouvement brownien. Pour $0 \leqslant t \leqslant 1$ on pose $Z_t := B_t - tB_1$. Montrer que $Z = \{Z_t\}_{0 \leqslant t \leqslant 1}$ est un pont brownien indépendant de B_1.
4. On note $\mathcal{C}$ l'espace des fonctions continues de $[0, 1]$ dans $\mathbb{R}$, muni de la norme $\| \cdot \|_\infty$. Montrer que pour toute fonction continue bornée $G : \mathcal{C} \to \mathbb{R}$,

$$\mathbb{E}\left[G(B) | |B_1| \leqslant \varepsilon\right] \underset{\varepsilon \to 0}{\longrightarrow} \mathbb{E}[G(Z)]$$

Exercice II.5 (Principe de réflexion)**.** Soit $\{W_t\}_{t \geqslant 0}$ un mouvement brownien, et soit $t > 0$. On pose

$$S_t := \sup_{0 \leqslant s \leqslant t} W_s$$

1. Soit $x \geqslant 0$ et $y \leqslant a$. A l'aide de la propriété de Markov forte, établir l'identité (Théorème de Désiré André)
$$\mathbb{P}\left(S_t \geqslant x, W_t \leqslant y\right) = \mathbb{P}\left(W_t \geqslant 2x - y\right)$$
2. En déduire que le couple (S_t, W_t) admet une densité que l'on explicitera.
3. Montrer que S_t a même loi que $|W_t|$.
4. Montrer que $T_x = \inf\{t \geqslant 0 : W_t = a\}$ a même loi que $(x/B_1)^2$, puis en déduire sa densité.

Exercice II.6 (Quelques martingales du mouvement brownien)**.** Soit $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P}\right)$ un espace de probabilité filtré et soit $\{B_t\}_{t\geqslant 0}$ un $\{\mathcal{F}_t\}_{t\geqslant 0}$-mouvement brownien.

1. Montrer que $\left\{B_t^2 - t\right\}_{t\geqslant 0}$ est une martingale.
2. Construire une martingale à partir du processus $\{B_t^3\}_{t\geqslant 0}$.
3. Construire une martingale à partir du processus $\{B_t^4\}_{t\geqslant 0}$.
4. Soit $\lambda \in \mathbb{C}$. Montrer que le processus $\left\{e^{\lambda B_t - \frac{\lambda^2 t}{2}}\right\}_{t\geqslant 0}$ est une martingale.
5. Construire une martingale à partir du processus $\{\cosh(\lambda B_t)\}_{t\geqslant 0}$.

Exercice II.7 (Loi de temps d'atteinte)**.** Soit $\{B_t\}_{t\geqslant 0}$ un mouvement brownien et $a > 0$.

1. À l'aide de la martingale $\left\{B_t^2 - t\right\}_{t\geqslant 0}$, calculer l'espérance de $T_a^\star := \inf\{t \geqslant 0 : |B_t| = a\}$.
2. À l'aide d'une martingale bien choisie, calculer la variance de $T_a^\star$.
2. À l'aide d'une martingale bien choisie, calculer la transformée de Laplace de $T_a^\star$.
3. Calculer la transformée de Laplace de $T_a := \inf\{t \geqslant 0 : B_t = a\}$ et retrouver le fait que T_a a même loi que $(a/B_1)^2$. Que vaut $\mathbb{E}[T_a]$?

Exercice II.8 (Maximum du mouvement brownien avec dérive)**.** Soit $\{B_t\}_{t\geqslant 0}$ un mouvement brownien. On fixe $a, b > 0$ et on pose $\tau := \inf\{t \geqslant 0 : B_t - bt = a\}$ un temps d'arrêt (admis).

1. À l'aide d'une martingale bien choisie, calculer la transformée de Laplace de τ.
2. En déduire la probabilité que la courbe du mouvement brownien soit au dessous de la demi-droite $t \mapsto a + bt$. Pouvait-on prévoir que la réponse ne dépendrait que de ab ?
3. Quelle est la loi de la variable aléatoire $U := \sup_{t\geqslant 0} B_t - bt$?

Exercice II.9 (Maximum du pont brownien)**.** Soit $\{Z_t\}_{0\leqslant t\leqslant 1}$ un pont brownien.

1. Pour $t \geqslant 0$ on pose $B_t = (1+t) Z_{\frac{t}{1+t}}$. Vérifier que $\{B_t\}_{t\geqslant 0}$ est un mouvement brownien.
2. En utilisant l'exercice précédant, déterminer la loi de la variable $V := \sup_{0\leqslant t\leqslant 1} Z_t$.

Intégrale stochastiques

Exercice II.10 (Intégrale d'un processus gaussien). Soit $X = \{X_t\}_{t \geqslant 0}$ un processus gaussien à trajectoires continues. Montrer que le processus $Y = \{Y_t\}_{t \geqslant 0}$ défini par

$$Y_t(\omega) := \int_0^t X_u(\omega) \mathrm{d}u$$

est encore un processus gaussien dont on précisera les fonctions de moyenne et de covariance, en fonction de celles de X. Qu'obtient-on dans le cas où X est un mouvement brownien ?

Exercice II.11 (Intégrales gaussiennes et Martingales). Soit $f \in L^2(\mathbb{R}_+)$. Pour $t \geqslant 0$ on pose

$$M_t := \int_0^t f(s) \mathrm{d}B_s$$

1. Le processus $(M_t)_{t \geqslant 0}$ est-il un processus gaussien à accroissements indépendants ?
2. Le processus $(M_t)_{t \geqslant 0}$ est-il une martingale ?
3. Construire une martingale à partir de $(M_t^2)_{t \geqslant 0}$.
4. Soit $\theta \in \mathbb{C}$. Construire une martingale à partir de $\left(e^{\theta M_t}\right)_{t \geqslant 0}$.
5. Pour quels choix de f le processus $(M_t)_{t \geqslant 0}$ est-il un mouvement brownien ?
6. Soit $g \in L^2(\mathbb{R}_+)$ et $N_t := \int_0^t g(s) \mathrm{d}B_s$. A quelle condition $(M_t)_{t \geqslant 0}$ et $(N_t)_{t \geqslant 0}$ sont-ils indépendants ?

Exercice II.12 (Une intégrale brownienne). Que dire du processus $(X_t)_{t \geqslant 0}$ défini par la formule suivante ?

$$X_t := \int_0^{\sqrt{t}} \sqrt{2s}\, \mathrm{d}B_s$$

Exercice II.13 (Pont brownien). Soit $(a, b) \in \mathbb{R}^2$. On considère le processus $(Z_t)_{0 \leqslant t < 1}$ défini par

$$Z_t := a(1-t) + bt + (1-t) \int_0^t \frac{1}{1-s}\, \mathrm{d}B_s \quad (0 \leqslant t < 1)$$

1. Montrer que $(Z_t)_{0 \leqslant t < 1}$ est un processus gaussien dont on explicitera les paramètres.
2. Que dire de ce processus dans le cas $a = b = 0$?
3. Montrer que lorsque $t \to 1$, on a $Z_t \to b$ au sens de la convergence L^2.

Exercice II.14 (Processus d'Ornstein-Uhlenbeck). Soit V_0 une variable aléatoire réelle indépendante d'un brownien $(B_t)_{t \geqslant 0}$, et $b, \sigma > 0$. On définit un processus $(V_t)_{t \geqslant 0}$ par

$$V_t := e^{-bt} \left(V_0 + \sigma \int_0^t e^{bs}\, \mathrm{d}B_s \right)$$

1. Montrer que V_t converge en loi lorsque $t \to \infty$, et déterminer la limite.
2. On suppose désormais que $V_0 \sim \mathcal{N}\left(0, \frac{\sigma^2}{2b}\right)$. Montrer que $(V_t)_{t \geqslant 0}$ est un processus gaussien stationnaire dont on précisera les paramètres.
3. Que dire du processus $(X_t)_{t \in \mathbb{R}}$ défini ci-dessous ?

$$X_t := \frac{\sigma}{\sqrt{2b}} e^{-bt} B_{e^{2bt}}$$

Formule d'Itô

Exercice II.15 (Un exemple facile)**.** Calculer $\int_0^t B_s \mathrm{d}B_s$ en fonction de B_t et de t.

Exercice II.16 (Intégration stochastique par partie)**.** Soient $(X_t)_{t\geqslant 0}$ et $(Y_t)_{t\geqslant 0}$ deux processus d'Itô. Montrer que $(X_tY_t)_{t\geqslant 0}$ est un processus d'Itô et calculer sa différentielle stochastique.

Exercice II.17 (Formule d'Itô)**.** Dans chacun des cas suivants, montrer que $\{Z_t\}_{t\geqslant 0}$ est un processus d'Itô, calculer sa différentielle stochastique, et déterminer si $\{Z_t\}_{t\geqslant 0}$ est une martingale.

1. $Z_t = B_t + 4t$
2. $Z_t = B_t^2 - t$
3. $Z_t = t^2 B_t - 2\int_0^t sB_s \ \mathrm{d}s$
4. $Z_t = B_t^3 - 3tB_t$
5. $Z_t = B_t^2 \left(B_t^2 - 6t\right)$
6. $Z_t = \exp\left(\mu t + \sigma B_t\right)$, avec $(\mu, \sigma) \in \mathbb{R}^2$.
7. $Z_t = (\cos B_t, \sin B_t)$

Exercice II.18 (Polarisation)**.** Soient ϕ_1, ϕ_2 des fonctions L^2. Montrer que

$$\left\{\int_0^t \phi_1(s)\mathrm{d}B_s \int_0^t \phi_2(s)\mathrm{d}B_s - \int_0^t \phi_1(s)\phi_2(s)\mathrm{d}s\right\}_{t\geqslant 0}$$

est une martingale.

Exercice II.19 (Fonction du brownien)**.** Soit $f \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$. On suppose que pour tout $t \geqslant 0$,

$$\mathbb{E}\left[\int_0^t \left(f'\left(B_s\right)\right)^2 \ \mathrm{d}s\right] < \infty \quad \text{et} \quad \mathbb{E}\left[\int_0^t \left|f''\left(B_s\right)\right| \mathrm{d}s\right] < \infty$$

1. Établir l'identité suivante, valable pour tout $t \geqslant 0$:

$$\mathbb{E}\left[f\left(B_t\right)\right] = f(0) + \frac{1}{2}\mathbb{E}\left[\int_0^t f''\left(B_s\right) \mathrm{d}s\right]$$

2. Retrouver la formule donnant les moments de la loi gaussienne.

Exercice II.20 (EDP)**.** Soit $(t, x) \mapsto f(t, x)$ une fonction de classe $\mathcal{C}^{1,2}\left(\mathbb{R}_+ \times \mathbb{R}, \mathbb{R}\right)$ solution de l'EDP

$$\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2} = 0$$

1. Montrer que le processus $\{f\left(t, B_t\right)\}_{t\geqslant 0}$ est une martingale locale, et donner une condition suffisante sur f pour que ce soit une vraie martingale.
2. Que dire de $\{B_t^3 - 3tB_t\}_{t\geqslant 0}$? et $\{B_t^4 - 6tB_t^2 + 3t^2\}_{t\geqslant 0}$?

Exercice II.21 (Cas vectoriel)**.** Soit $\left\{\left(B_t, \widetilde{B}_t\right)\right\}_{t\geqslant 0}$ un mouvement brownien plan. Pour $t \geqslant 0$ on pose :

$$X_t := \exp\left(B_t\right)\cos\left(\widetilde{B}_t\right) \quad \text{et} \quad Y_t := \exp\left(B_t\right)\sin\left(\widetilde{B}_t\right)$$

1. Montrer que $\{X_t\}_{t\geqslant 0}$ et $\{Y_t\}_{t\geqslant 0}$ sont des martingales de carré intégrable.
2. Le produit $\{X_tY_t\}_{t\geqslant 0}$ est-il une martingale ?
3. Calculer la différentielle stochastique du processus $\{Z_t\}_{t\geqslant 0}$ défini par

$$Z_t := \left(X_t - 1\right)^2 + \left(Y_t\right)^2$$

Exercice II.22 (Loi de l'arcsinus)**.** (Loi de l'arcsinus). Soit $\{B_t\}_{t\geqslant 0}$ un mouvement brownien réel. Il est facile de voir que la formule d'Itô pour $\{f(B_t)\}_{t\geqslant 0}$ reste valable si f est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ et de classe $\mathcal{C}^2$ partout sauf en un nombre fini de points au voisinage desquels f'' est bornée. Voici une application. Soit

$$H_t := \int_0^t \mathbf{1}_{(B_s\geqslant 0)}\mathrm{d}s \quad (t \geqslant 0)$$

Il s'agit de montrer que la variable aléatoire $\frac{H_t}{t}$ admet pour densité $x \mapsto \frac{1}{\pi\sqrt{x(1-x)}}$ sur $(0,1)$.

1. Étant donnés $\alpha, \beta > 0$, trouver $a, b \in \mathbb{R}$ pour que f satisfasse les conditions ci-dessus, avec

$$f(x) := \begin{cases} a\exp(-\sqrt{2(\alpha+\beta)}x) + \frac{1}{\alpha+\beta} & \text{si } x \geqslant 0 \\ b\exp(\sqrt{2\alpha}x) + \frac{1}{\alpha} & \text{si } x < 0 \end{cases}$$

2. Construire une martingale bornée à partir de $\left\{f(B_t)\, e^{-(\alpha t+\beta H_t)}\right\}_{t>0}$ et en déduire que

$$\int_0^\infty e^{-\alpha t}\mathbb{E}\left[e^{-\beta H_t}\right]\mathrm{d}t = \frac{1}{\sqrt{\alpha(\alpha+\beta)}}$$

3. On admet que $\int_0^\infty \int_0^1 e^{-\alpha t}\frac{e^{-\beta t x}}{\pi\sqrt{x(x-1)}}\mathrm{d}x\ \mathrm{d}t = \frac{1}{\sqrt{\alpha(\alpha+\beta)}}$ pour tout α et β, conclure.

Exercice II.23 (Temps local)**.** Soit $\{B_t\}_{t\geqslant 0}$ un mouvement brownien réel. La formule d'Itô

$$f(B_t) = f(0) + \int_0^t f'(B_s)\,\mathrm{d}B_s + \frac{1}{2}\int_0^t f''(B_s)\,\mathrm{d}s$$

s'étend aisément à toute fonction $f \in \mathcal{C}^1(\mathbb{R},\mathbb{R})$ de classe $\mathcal{C}^2$ partout sauf en un nombre fini de points au voisinage desquels f'' reste bornée. Voici un exemple où cette extension est utile.

1. Soit $\varepsilon > 0$. Calculer la différentielle stochastique du processus $\{f_\varepsilon(B_t)\}_{t\geqslant 0}$, où

$$f_\varepsilon(x) = \begin{cases} |x| & \text{si } |x| \geqslant \varepsilon \\ \frac{1}{2}\left(\varepsilon + \frac{x^2}{\varepsilon}\right) & \text{si } |x| < \varepsilon \end{cases}$$

2. On note Λ la mesure de Lebesgue sur $\mathbb{R}$. En déduire que pour tout $t \geqslant 0$,

$$\frac{1}{2\varepsilon}\Lambda\{s \in [0,t] : |B_s| < \varepsilon\} \xrightarrow[\varepsilon\to 0^+]{L^2} |B_t| - \int_0^t \operatorname{sgn}(B_s)\,\mathrm{d}B_s$$

Exercice II.24 (Récurrence ou transience du mouvement brownien)**.**

1. Trouver une fonction harmonique non-triviale $u : \mathbb{R}^2\backslash\{0\} \to \mathbb{R}$ à symétrie sphérique, i.e.

$$\|z\| = \|z'\| \Longrightarrow u(z) = u(z')$$

2. On fixe $0 < r < R < \infty$ et $z \in \mathbb{R}^2$ tel que $r < \|z\| < R$. Soit $\{B_t\}_{t\geqslant 0}$ un mouvement brownien plan issu de z, et

$$\tau := \inf\{t \geqslant 0 : \|B_t\| \leqslant r \text{ ou } \|B_t\| \geqslant R\}$$

Que dire du processus $\{u(B_{t\wedge\tau})\}_{t>0}$? En déduire la probabilité que $\{B_t\}_{t\geqslant 0}$ atteigne le cercle de centre 0 et de rayon r avant d'avoir touché celui de rayon R.

3. Montrer qu'un point donné distinct du point de départ ne sera presque-sûrement jamais visité par le mouvement brownien plan, mais que presque-sûrement, chaque ouvert de $\mathbb{R}^2$ est visité infiniment souvent.
4. On se place désormais en dimension $d \geqslant 3$. Pour $r < \|z\| < R$, déterminer la probabilité qu'un mouvement brownien issu de z touche la sphère de rayon r avant celle de rayon R.

5. Quelle est la probabilité qu'une fois sorti de la boule de rayon n^3, le mouvement brownien ne revienne plus jamais dans la boule de rayon n ?
6. En déduire qu'en dimension $d \geqslant 3$, le mouvement brownien est transient : $\|B_t\| \to \infty$ p.s.

Equations différentielles stochastiques

Exercice II.25 (Pont brownien)**.** Soit $(a, b) \in \mathbb{R}^2$. On considère le processus $\{Z_t\}_{0 \leqslant t<1}$ défini par

$$Z_t = a(1-t) + bt + (1-t)\int_0^t \frac{1}{1-s}\, \mathrm{d}B_s$$

Montrer que $\{Z_t\}_{0 \leqslant t<1}$ est solution de l'équation différentielle stochastique suivante :

$$\mathrm{d}Z_t = \frac{b - Z_t}{1-t}\, \mathrm{d}t + \mathrm{d}B_t$$

Exercice II.26 (Une EDS)**.** On considère l'équation différentielle stochastique

$$X_0 = 0, \quad dX_t = \frac{1}{2\left(1+X_t^2\right)} dt + \frac{1}{\sqrt{1+X_t^2}} dB_t$$

où B désigne un mouvement brownien réel. Le but est d'étudier les variables aléatoires

$$X_\star := \inf\{X_t : t \geqslant 0\} \in [-\infty, 0] \quad \text{ et } \quad X^\star := \sup\{X_t : t \geqslant 0\} \in [0, +\infty]$$

1. Justifier que cette EDS admet une unique solution forte.
2. Trouver $F \in \mathcal{C}^2$ avec $F(0) = 1, F(\infty) = 0$, telle que $(F(X_t))_{t \geqslant 0}$ soit une martingale locale.
3. Soient $a, b > 0$. On pose $\tau := \tau_{-a} \wedge \tau_b$, où pour tout $r \in \mathbb{R}, \tau_r := \inf\{t \geqslant 0 : X_t = r\}$.

(a) Montrer que $\left(e^{-X_{t\wedge\tau}}\right)_{t \geqslant 0}$ est une martingale de carré intégrable, et que pour $t \geqslant 0$,

$$\mathbb{E}\left[e^{-2X_{t\wedge\tau}}\right] = 1 + \mathbb{E}\left[\int_0^{t\wedge\tau} \frac{e^{-2X_s}}{1+X_s^2} ds\right]$$

(b) En déduire une constante $C(a, b)$ telle que pour tout $t \geqslant 0, \mathbb{E}[t \wedge \tau] \leqslant C(a, b)$.

4. Déduire des deux questions précédentes que

$$\mathbb{P}(\tau_{-a} < \tau_b) = \frac{1-e^{-b}}{e^a - e^{-b}} \quad \text{ et } \quad \mathbb{P}(\tau_{-a} > \tau_b) = \frac{e^a - 1}{e^a - e^{-b}}$$

5. Conclure que $X^\star = \infty$ p.s. et que $-X_\star$ suit une loi exponentielle.

Exercice II.27 (Monotonie)**.** Soient $b, \sigma : \mathbb{R} \to \mathbb{R}$ des fonctions lipschitziennes, et $x \leqslant y$ des réels. On note $(X_t)_{t \geqslant 0}$ et $(Y_t)_{t \geqslant 0}$ les solutions de $dZ_t = b(Z_t)\, dt + \sigma(Z_t)\, dB_t$ avec $X_0 = x$ et $Y_0 = y$.

1. Justifier que le processus $(U_t)_{t \geqslant 0}$ défini par la formule suivante a bien un sens :

$$U_t := \int_0^t \frac{\sigma(X_s) - \sigma(Y_s)}{X_s - Y_s} \mathbf{1}_{X_s \neq Y_s} dB_s + \int_0^t \frac{b(X_s) - b(Y_s)}{X_s - Y_s} \mathbf{1}_{X_s \neq Y_s} ds$$

2. Établir l'identité suivante, valable pour tout $t \geqslant 0$:

$$X_t - Y_t = (x-y) \exp\left\{U_t - \frac{1}{2}\int_0^t \left(\frac{\sigma(X_s) - \sigma(Y_s)}{X_s - Y_s}\right)^2 \mathbf{1}_{X_s \neq Y_s} ds\right\}$$

3. En déduire que presque-sûrement : $\forall t \geqslant 0, X_t \leqslant Y_t$.

Exercice II.28 (Changement de variable)**.** Établir l'existence d'une unique solution forte à l'EDS

$$dX_t = \left(\sqrt{1+X_t^2} + \frac{X_t}{2}\right) dt + \sqrt{1+X_t^2} dB_t, \quad X_0 = x$$

puis la déterminer explicitement en effectuant le changement de variable $Y_t = \operatorname{argsinh}(X_t)$.

Exercice II.29 (Brownien géométrique)**.** Le but de cet exercice est de résoudre l'EDS suivante :

$$dX_t = \{r(t)X_t + f(t)\}\, dt + \{v(t)X_t + g(t)\}\, dB_t, \quad X_0 = \zeta$$

où $v, r, f, g : \mathbb{R}_+ \to \mathbb{R}$ sont boréliennes bornées et $\zeta \in L^2$ est indépendante de $(B_t)_{t\geqslant 0}$.

1. Montrer que cette équation admet une unique solution forte.
2. Trouver une solution dans le cas où $f = g = 0$ et où r et v sont des fonctions constantes.
3. En déduire une solution dans le cas où $f = g = 0$ mais r et v ne sont pas constantes.
4. Résoudre le cas général.

Indication : On pourra s'inspirer de la variation de la constante pour passer de la solution d'une équation différentielle homogène à une équation avec second membre.

Exercice II.30 (Sinh du brownien)**.** Soit $((B_t, C_t))_{t\geqslant 0}$ un brownien plan. Pour $t \geqslant 0$, on pose

$$Y_t := \int_0^t e^{C_t - C_s} dB_s \quad \text{et} \quad W_t := \int_0^t \frac{1}{\sqrt{1+Y_s^2}} dB_s + \int_0^t \frac{Y_s}{\sqrt{1+Y_s^2}} dC_s$$

1. Montrer que $(W_t)_{t\geqslant 0}$ est un mouvement brownien réel.
2. Vérifier que $(Y_t)_{t\geqslant 0}$ est solution de l'équation différentielle stochastique

$$dY_t = \sqrt{1+Y_t^2} dW_t + \frac{Y_t}{2} dt; \quad Y_0 = 0$$

3. Pour $t \geqslant 0$, on pose $X_t := \sinh(W_t)$. Quelle équation différentielle stochastique vérifie le processus $(X_t)_{t\geqslant 0}$? Que peut-on en conclure ?
4. Pour $0 \leqslant s \leqslant t$, on pose $\widetilde{B}_s^t := B_t - B_{t-s}$. On rappelle que $\left(\widetilde{B}_s^t\right)_{0\leqslant s\leqslant t}$ est un mouvement brownien sur $[0, t]$. Prouver l'identité suivante, valable pour tout $f \in L^2([0, t])$:

$$\int_0^t f(s) dB_s = \int_0^t f(t-s) d\widetilde{B}_s^t$$

5. Pour $t \geqslant 0$, on pose $Z_t := \int_0^t e^{C_s} dB_s$. Montrer que pour tout $t \geqslant 0, Z_t$ a même loi que Y_t.
6. Montrer que $(Y_t)_{t\geqslant 0}$ n'est pas une martingale, mais que $(Z_t)_{t\geqslant 0}$ en est une.

Troisième partie

Processus de Markov

Chapitre 7

Chaînes de Markov

Introduction

Une chaîne de Markov est un des exemples les plus simples d'une suite de variables aléatoires. Ce sont des suites dites « sans mémoire », dans le sens qu'à l'instant présent, la loi du futur ne dépend pas du passé. Elles ont été introduites par Andreï Markov au début du XX^e siècle, entre autres choses, comme un exemple de suites non i.i.d, mais vérifiant la loi des grands nombres.

Le but de ce cours, qui suit de très prêt les notes de cours de Jean-Jacques Ruch et Marie-Line Chabanol[1], est d'établir quelques propriétés de ces chaînes et en particulier d'étudier leurs propriétés asymptotiques.

7.1 Définitions et premières propriétés

Une chaîne de Markov à états au plus dénombrable est une suite de variables aléatoires (X_n) à valeurs dans E, un espace au plus dénombrable, vérifiant la propriété suivante :

$$\mathbb{P}(X_{k+1} = x_{k+1} | X_k = x_k, \dots, X_0 = x_0) = \mathbb{P}(X_{k+1} = x_{k+1} | X_k = x_k)\,,$$

pour tout $k \in \mathbb{N}$ et $(x_0, \dots, x_{k+1}) \in E^k$.

En d'autres mots la loi de X_{k+1} conditionnée au passé ne dépend que de X_k - l'instant présent.

Exemple 7.1 (Marche aléatoire sur $\mathbb{Z}$)**.** Soit (ξ_n) une suite de variables aléatoires indépendantes de lois Bernoullis de paramètres (p_n), $X_0 = 0$ et pour tout $k \geqslant 0$,

$$X_{k+1} = X_k + (2\xi_{k+1} - 1) = \sum_{i=0}^{k} (2\xi_{i+1} - 1)\,.$$

On appelle cet exemple marche aléatoire : à chaque instant on lance une pièce (peut-être truquée) et on va à droite si c'est pile et à gauche si c'est face. Bien sûr, notre position à l'instant suivant ne dépend que de notre position à l'instant présent.

On peut remarquer que si dans le cas où dans l'exemple précédent (p_n) est une suite constante $p_n \equiv p$ alors la loi de $X_{k+1} | X_k$ est exactement la loi de $X_1 | X_0$. On parle dans ce cas-là d'une chaîne de Markov homogène.

Définition 7.2. On dit qu'une chaine de Markov (X_k) est homogène si pour tout $k \in \mathbb{N}$ et tout $x_0, x_1 \in E$,

$$\mathbb{P}(X_{k+1} = x_1 | X_k = x_0) = \mathbb{P}(X_1 = x_1 | X_0 = x_0)\,.$$

Proposition 7.3. *Soit (X_k) une chaîne de Markov homogène. Pour tout $m, n \in \mathbb{N}$ et $x, y \in E$, on a*

$$\mathbb{P}(X_{m+n} = y | X_n = x) = \mathbb{P}(X_m = y | X_0 = x)\,.$$

En particulier, pour tout $m \in \mathbb{N}$, $(X_{km})_{k \in \mathbb{N}}$ est également une chaîne de Markov homogène.

1. Disponibles sur ce lien.

Démonstration. Fixons $n \in \mathbb{N}$ et démontrons cette propriété par récurrence sur m. Pour $m = 0$ le résultat est évident, pour $m = 1$ il découle de la propriété d'homogénéité. Supposons que le résultat soit vrai pour $m \in \mathbb{N}$, alors on a

$$\begin{aligned}
\mathbb{P}(X_{n+(m+1)} = y | X_n = x) &= \sum_{z \in E} \mathbb{P}(X_{n+(m+1)} = y | X_{n+m} = z, X_n = x)\mathbb{P}(X_{n+m} = z | X_n = x) \\
&= \sum_{z \in E} \mathbb{P}(X_{n+(m+1)} = y | X_{n+m} = z)\mathbb{P}(X_m = z | X_0 = x) \\
&= \sum_{z \in E} \mathbb{P}(X_{m+1} = y | X_m = z)\mathbb{P}(X_m = z | X_0 = x) \\
&= \sum_{z \in E} \mathbb{P}(X_{m+1} = y | X_m = z, X_0 = x)\mathbb{P}(X_m = z | X_0 = x) \\
&= \mathbb{P}(X_{m+1} = y | X_0 = x) \,,
\end{aligned}$$

où la deuxième égalité se déduit de l'hypothèse de récurrence et la troisième de l'homogénéité de la chaîne. □

Dans la suite, nous allons considérer seulement les chaînes de Markov homogènes.

Définition 7.4. On appelle probabilité de transition pour aller d'un état $x \in E$ à un état $y \in E$ la probabilité

$$p_{x,y} = \mathbb{P}(X_{k+1} = y | X_k = x) = \mathbb{P}(X_1 = y | X_0 = x) \,.$$

Comme E est au plus dénombrable on écrira par la suite $E = \{x_k \in E : k \in \mathbb{N}\}$.

Proposition 7.5. *Soit ν la loi de X_0, alors pour tout $(x_0, \ldots, x_k) \in E^k$ on a :*

$$\mathbb{P}(X_0 = x_0, \ldots, X_k = x_k) = \nu(x_0) \prod_{i=0}^{k-1} p_{x_k, x_{k+1}} \,.$$

Démonstration. Immédiat par conditionnements successifs. □

Définition 7.6. On appelle *matrice de transition* la matrice $\mathcal{P} = (p_{x_i,x_j})_{i,j \in \mathbb{N}}$:

$$\mathcal{P} = \begin{pmatrix} p_{x_0,x_0} & p_{x_0,x_1} & p_{x_0,x_2} & \cdots \\ p_{x_1,x_0} & p_{x_1,x_1} & p_{x_1,x_2} & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} .$$

C'est une matrice finie si E est fini et dénombrable sinon.

Notons quelques propriétés simples de $\mathcal{P}$.

Lemme 7.7. *Toute matrice de transition vérifie les propriétés suivantes.*

1. *Pour tout $i, j \in \mathbb{N}$, $0 \leqslant p_{x_i,x_j} \leqslant 1$.*
2. *Pour tout $i \in \mathbb{N}$, on a $\sum_{j \in \mathbb{N}} p_{x_i,x_j} = 1$.*

Démonstration. Le premier point est immédiat, car $p_{x_i,x_j} = \mathbb{P}(X_1 = x_j | X_0 = x_i)$. Pour le second on écrit

$$\sum_{j \in \mathbb{N}} p_{x_i,x_j} = \sum_{j \in \mathbb{N}} \mathbb{P}(X_1 = x_j | X_0 = x_i) = 1 \,.$$

□

On pourra aussi remarquer (et ça sera important pour la suite) la propriété suivante.

Lemme 7.8. *La matrice de transition $\mathcal{P}$ admet 1 comme valeur propre. Le vecteur propre associé est le vecteur qui ne contient que des 1.*

Comme chaque X_k est à valeurs dans $(x_0, \ldots, x_k) \in E^k$, on identifiera ν^k la loi de X_k avec le vecteur ligne $(\nu^k(0), \ldots, \nu^k(n), \ldots)$, où $\nu^k(i) = \mathbb{P}(X_k = x_i)$. La matrice de transition $\mathcal{P}$ permet alors d'exprimer simplement ν^k à partir de μ^0.

Proposition 7.9. *Soit ν^0 la loi de X_0 alors ν^1 la loi de X_1 s'écrit comme*

$$\nu^1 = \nu^0 \mathcal{P} .$$

De plus pour $k \in \mathbb{N}$, on a que ν^{k+1} la loi de X_{k+1} s'écrit comme

$$\nu^k = \nu^0 \mathcal{P}^k ,$$

où $\mathcal{P}^k$ correspond au produit matriciel de $\mathcal{P}$ avec elle-même k fois.

En particulier, $\mathbb{P}(X_k = x_j | X_0 = x_i) = (\mathcal{P}^k)_{i,j}$.

Démonstration. Pour le premier point on écrit

$$\nu^1(i) = \mathbb{P}(X_1 = x_i) = \sum_{j \in \mathbb{N}} \mathbb{P}(X_1 = x_i | X_0 = x_j) \nu^0(x_j) = (\nu^0 \mathcal{P})_i .$$

Prouvons le cas général par récurrence. Le cas $k = 1$ venant d'être prouvé (le cas $k = 0$ est évident), supposons que la propriété soit vraie pour un $k \geqslant 1$. Alors

$$\nu^{k+1}(i) = \sum_{j \in \mathbb{N}} \mathbb{P}(X_{k+1} = x_i | X_k = x_j) \nu^k(x_j) = (\nu^k \mathcal{P})_i = (\nu^0 \mathcal{P}^{k+1})_i ,$$

où la deuxième égalité provient de l'homogénéité de la chaîne. □

7.1.1 Représentation par graphe

À toute chaîne de Markov homogène on peut naturellement associer un graphe dirigé. Les états seront représenter par des sommets s'il existe une arête $x_i \to x_j$ si $p_{x_i, x_j} > 0$. La chaîne peut alors s'interpréter comme une marche aléatoire sur ce graphe.

À partir de ce graphe, nous pouvons facilement déduire des propriétés importantes de la chaîne. Par exemple dans le cas où il existe un chemin entre x_i et x_j ça veut dire que si notre chaîne de Markov est initialisé en x_i il existe une probabilité non nulle de passer à x_j.

Plus généralement, nous allons voir dans la suite, que l'existence, le nombre et la nature des composantes connexes jouent un rôle important dans l'étude du comportement d'une chaîne de Markov.

Nous rappelons qu'une composante fortement connexe d'un graphe orienté est un sous-graphe maximal t.q. pour tout couple (u, v) de ses sommets, il existe un chemin contenu dans le sous-graphe allant de u à v.

Exemple 7.10 (La chaîne à deux états)**.** Son graphe est représenté sur la figure 7.1. La matrice de transition de cette chaîne s'écrit comme suit :

$$\mathcal{P} = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix} .$$

Si $\alpha, \beta > 0$, alors c'est un graphe fortement connexe, la chaîne va continuer de visiter tous les états peu importe l'initialisation. Si $\alpha = 0$ (respectivement $\beta = 0$) alors la chaîne s'arrêtera au premier (respectivement deuxième) état dès le premier passage. On dit alors que cet état est absorbant.

Exemple 7.11 (Marche aléatoire homogène)**.** Nous reprenons l'exemple de la marche aléatoire sur $\mathbb{Z}$ homogène. Son graphe est représenté sur la figure 7.2. Si $p \notin \{0, 1\}$, c'est un graphe fortement connexe. Sa matrice de transition s'écrit $\mathcal{P} = (p_{i,j})_{i,j \in \mathbb{Z}}$ avec

$$p_{i,j} = \begin{cases} p \text{ si } j = i+1 \\ 1-p \text{ si } j = i-1 \\ 0 \text{ sinon} \end{cases} .$$

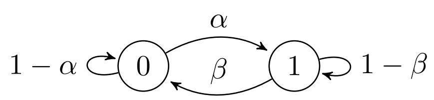


FIGURE 7.1 – Chaîne à deux états

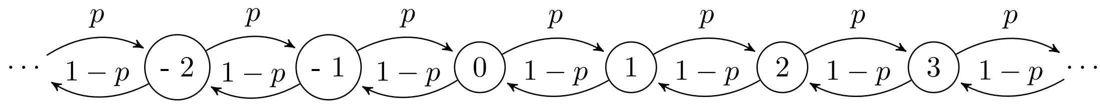


FIGURE 7.2 – Marche aléatoire sur $\mathbb{Z}$

Exemple 7.12 (Ruine d'un joueur)**.** Deux joueurs A, B partagent entre eux n euros. À chaque tour, ils lancent une pièce (truquée ou non), si c'est pile le joueur A donne 1 euro au joueur B et inversement. Le jeu s'arrête lorsqu'un des joueurs est ruiné. Si (X_k) représente la fortune du joueur A après k lancés, alors (X_k) est une chaîne de Markov dans l'espace des états $E = \{0, \dots, n\}$. Son graphe est représenté sur la figure 7.3. Sa matrice de transition $\mathcal{P} = (p_{i,j})_{i,j \in E^2}$ est la suivante :

$$p_{i,j} = \begin{cases} 1 \text{ si } i = j = 0 \text{ ou } i = j = n \\ p \text{ si } i \neq 0 \text{ et } j = i + 1 \\ 1 - p \text{ si } i \neq n \text{ et } j = i - 1 \\ 0 \text{ sinon } . \end{cases}$$

Dans cet exemple-là, si $p \notin \{0, 1\}$, le graphe possède 3 composantes fortement connexes et 2 états absorbants.

Exemple 7.13 (Modèles des urnes d'Ehrenfest[2])**.** Nous supposons être en présence de N molécules disposés dans deux urnes A, B. À chaque instant une des molécules est choisie aléatoirement et est déposée dans l'autre urne. Ainsi si l'on note (X_k) le nombre de molécules dans l'urne A après k tirages, la matrice de transition s'écrit comme $\mathcal{P} = (p_{i,j})_{0 \leqslant i,j \leqslant N}$ avec

$$p_{i,j} = \begin{cases} 1 - \frac{i}{N} \text{ si } j = i + 1 \\ \frac{i}{N} \text{ si } j = i - 1 \\ 0 \text{ sinon } . \end{cases}$$

Notons que le graphe correspondant n'à qu'une seule composante fortement connexe.

7.1.2 Relation de Chapman-Kolmogorov

Cette relation lie la probabilité de passage entre deux états aux probabilités de passage par les états intermédiaires en n étapes.

Proposition 7.14 (Relation de Chapman-Kolomogorov)**.** *Pour tout $m, n \in \mathbb{N}$, et $y, x \in E$, on a*

$$\mathbb{P}(X_{m+n} = y | X_0 = x) = \sum_{z \in E} \mathbb{P}(X_m = z | X_0 = x) \mathbb{P}(X_n = y | X_0 = z) \,.$$

2. Modèle introduit en 1911 par Paul Ehrenfest (1880-1933) et sa femme Tatiana Afanassieva (1876-1963) pour illustrer certains paradoxes de la physique statistique.

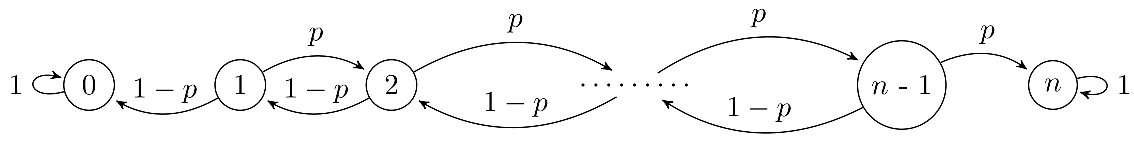


FIGURE 7.3 – Ruine d'un joueur

Démonstration. En utilisant la proposition 7.9, cette relation traduit tout simplement l'expression du produit matriciel.

$$(\mathcal{P}^{n+m})_{i,j} = (\mathcal{P}^m\mathcal{P}^n)_{i,j} = \sum_{l\in\mathbb{N}} (\mathcal{P}^m)_{i,l}(\mathcal{P}^n)_{l,j} \,.$$

□

Proposition 7.15. *Soit $n \geqslant 0$ et $m \geqslant 1$, on a*

$$\begin{aligned} &\mathbb{P}(X_{m+n} = j_{n+m}, \ldots, X_{n+1} = j_{n+1} | X_n = j_n, \ldots, X_0 = j_0) \\ &= \mathbb{P}(X_{m+n} = j_{n+m}, \ldots, X_{n+1} = j_{n+1} | X_n = j_n) \\ &= \mathbb{P}(X_m = j_{n+m}, \ldots, X_1 = j_{n+1} | X_0 = j_n) \,. \end{aligned}$$

Démonstration. Par conditionnements successifs nous avons que la première expression est égale à :

$$\prod_{l=1}^{m} \mathbb{P}(X_{n+l} = j_{n+l} | X_{n+l-1} = j_{n+l-1}, \ldots, X_0 = j_0) = \prod_{l=1}^{m} \mathbb{P}(X_{n+l} = j_{n+l} | X_n = j_{n+l-1}) \,,$$

et la dernière expression est bien égale à

$$\mathbb{P}(X_{m+n} = j_{n+m}, \ldots, X_{n+1} = j_{n+1} | X_n = j_n) \,.$$

□

L'idée principale qu'il faut avoir en tête est que le futur « sachant le passé » ne dépend que du présent.

7.1.3 Filtration naturelle

Plus généralement, définissons, pour chaque $k \in \mathbb{N}$, la tribu qui contient toutes les informations sur la chaîne de Markov jusqu'au temps k : $\mathcal{F}_k = \sigma(X_0, \ldots, X_k)$. La suite $(\mathcal{F}_k)$ est alors ce qu'on appelle une filtration : c'est une tribu croissante au sens que, pour $n \leqslant m$, $\mathcal{F}_n \subset \mathcal{F}_m$.

Notons en plus $\mathcal{F}_k^+ = \sigma(X_{k+1}, X_{k+2}, \ldots)$ la tribu du futur. Notons alors le théorème.

Théorème 7.16. *Soit $n \in \mathbb{N}$, $A^+ \in \mathcal{F}_n^+$, $A^- \in \mathcal{F}_n$ et $j_n \in E$ t.q. $\mathbb{P}(A^-, X_n = j_n) > 0$:*

$$\mathbb{P}(A^+ | X_n = j_n, A^-) = \mathbb{P}(A^+ | X_n = j_n) \,.$$

Démonstration. On admettra le résultat pour $A^+ \in \mathcal{F}_n^+$ général, mais on notera que si $A^+ \in \mathcal{F}_{n+m}$, avec $m \geqslant 1$ alors il peut s'écrire comme une réunion dénombrable d'évènements du type $[X_{m+n} = j_{n+m}, \ldots, X_{n+1} = j_{n+1}]$. De même, l'évènement $A^- \cap [X_n = j_n]$ s'écrit comme une réunion dénombrable d'événements du type $[X_n = j_n, \ldots, X_0 = j_0]$. Dans ce cas-là il suffit d'appliquer la proposition 7.15.

Le cas général où A^+ ne s'écrit pas comme une telle union est admis. □

7.2 Classification des états

Étant donné une chaîne de Markov un des points qui nous intéresse est de caractériser son comportement asymptotique. Idéalement, on voudrait affirmer que X_k, quand $k \to +\infty$, converge en loi vers une loi de probabilité, et ce, indépendamment de l'initalisation !

Il se trouve que pour certifier un tel comportement de la chaîne, il suffit d'étudier son graphe. Cette section est donc dédiée au diverses propriétés de ce dernier.

En résumé :

— *Irréductibilité.* Les composantes connexes du graphe sont appelés les classes d'irréductibilité. Il existe toujours un chemin entre deux états dans la même classe. Une chaîne est dite **irréductible** si son graphe possède **une seule composante connexe**.

— *Périodicité.* Pour un état $i \in E$, sa période, noté $d(i)$ est telle que pour $n \in \mathbb{N}$ assez grand il existe toujours un chemin de i à i de longueur de taille exactement $nd(i)$ (voir la section 7.2.2 pour la définition exacte). Une chaîne est dite **apériodique** si elle est **irréductible** et **tous ses états sont de période** 1.

— *État récurrent.* Un état i est **récurrent** si la chaîne débutant en i revient à i **une infinité de fois**.

— *État transitoire.* Un état i est **transitoire** si il n'est pas récurrent.

Enfin, en s'avançant un peu aux résultats de la section 7.3, on sait qu'une chaîne de Markov à états finis qui est irréductible et apériodique converge vers son unique mesure invariante.

7.2.1 Irréductibilité

Comme évoqué plus haut il nous sera intéressant d'étudier les composantes fortement connexes du graphe associé à la chaîne.

Définition 7.17. On dit qu'un état $j \in E$ est accessible à partir de $i \in E$, noté $i \rightsquigarrow j$, s'il existe un entier $n \geqslant 0$ tel que $\mathbb{P}(X_n = j | X_0 = i) > 0$.

En d'autres mots, $i \rightsquigarrow j$ s'il existe un chemin orienté de i à j sur le graphe associé à la chaîne.

Définition 7.18. On dit que deux états $i, j \in E$ communiquent si $i \rightsquigarrow j$ et $j \rightsquigarrow i$. On notera $i \leftrightsquigarrow j$.

Ainsi, si $i \leftrightsquigarrow j$ ssi ils se situent dans la même composant fortement connexe du graphe associé à la chaîne.

Proposition 7.19. *La relation $i \leftrightsquigarrow j$ est une relation d'équivalence.*

Démonstration. En effet, si $i \rightsquigarrow j$ et $j \rightsquigarrow l$, alors il existe $n, m \in \mathbb{N}$, t.q. $\mathbb{P}(X_n = j | X_0 = i) > 0$ et $\mathbb{P}(X_m = l | X_0 = j) > 0$. Donc

$$\begin{aligned}
\mathbb{P}(X_{n+m} = l | X_0 = i) &\geqslant \mathbb{P}(X_{m+n} = l | X_n = j)\mathbb{P}(X_n = j | X_0 = i) \\
&= \mathbb{P}(X_n = j | X_0 = i)\mathbb{P}(X_n = j | X_0 = i) \\
&> 0\,,
\end{aligned}$$

ce qui veut dire que $\rightsquigarrow$ est transitive et donc de même pour $\leftrightsquigarrow$. De plus $\leftrightsquigarrow$ est symétrique par définition et on a toujours $i \rightsquigarrow i$, car $\mathbb{P}(X_0 = i | X_0 = i) = 1$, ce qui veut dire que la relation est réflexive. □

Les classes d'équivalence pour cette relation sont exactement les composantes fortement connexes du graphe associé à la chaîne de Markov. On les appelle les *classes irréductibles*.

Si $\mathcal{C}_1$, $\mathcal{C}_2$ sont deux classes d'équivalences distinctes alors on peut éventuellement passer d'un état de $\mathcal{C}_1$ à un état de $\mathcal{C}_2$, mais dans ce cas-là on ne pourra pas passer de $\mathcal{C}_2$ à $\mathcal{C}_1$.

Définition 7.20. On dit que la chaîne de Markov est irréductible, s'il existe qu'une seule classe irréductible.

7.2.2 Périodicité

Dans la section précédente, nous avons étudié les états accessibles à partir d'un état i. Dans cette section nous allons nous intéresser aux propriétés de retour de la chaîne à son état initial.

Définition 7.21. Soit $i \in E$, on appelle la période de i, notée $d(i)$, l'entier suivant :

$$d(i) = \mathrm{pgcd}\{k \geqslant 1 : \mathbb{P}(X_k = i|X_0 = i) > 0\} .$$

Par convention, on dit que $\mathrm{pgcd}(\emptyset) = +\infty$.

Si $2 \leqslant d(i) < +\infty$, on dira que l'état i est périodique de période $d(i)$. Si $d(i) = 1$, on dira que l'état i est apériodique.

La proposition suivante montre que la période est en fait une propriété de classe.

Proposition 7.22. *Si $d(i)$ la période de i est finie et $i \leftrightsquigarrow j$ alors j est périodique de même période.*

Démonstration. Comme $i \leftrightsquigarrow j$, il existe $m, n \in \mathbb{N}$ t.q. $\mathbb{P}(X_m = i|X_0 = j) > 0$ et $\mathbb{P}(X_n = j|X_0 = i) > 0$. Soit $l \in \mathbb{N}$ t.q. $\mathbb{P}(X_l = i|X_0 = i) > 0$. Alors

$$\begin{aligned}\mathbb{P}(X_{l+(n+m)} = j|X_0 = j) &\geqslant \mathbb{P}(X_n = j|X_0 = i)\mathbb{P}(X_m = i|X_0 = i)\mathbb{P}(X_l = i|X_0 = j)\\ &> 0 .\end{aligned}$$

Donc $d(j)$ divise $l + (n+m)$. Le même raisonnement prouve que $d(j)$ divise $2l + (n+m)$. Finalement, on a que $d(j)$ divise l et donc $d(j)$ divise $d(i)$. Par symétrie on a aussi que $d(i)$ divise $d(j)$. Donc $d(i) = d(j)$. □

7.2.3 Temps d'arrêts et propriété de Markov forte

La prochaîne étape de classification des états d'une chaîne de Markov consiste à établir qu'il y a des états qui vont être visité une infinité de fois, on les appelle les *états récurrents*, et ceux qui sont visité qu'un nombre fini de fois - les *états transitoires*.

Pour établir un tel résultat, nous aurons besoin d'introduire quelques notions mathématiques supplémentaires.

Rappelons que la filtration naturelle de la chaîne de Markov est la suite de tribus $(\mathcal{F}_k)$, où $\mathcal{F}_k = \sigma(X_0, \ldots, X_k)$ est la tribu qui contient toute l'information sur la chaîne jusqu'au temps k. Les définitions qui suivent sont identiques à celles vu dans les chapitres sur les martingales.

Définition 7.23. Soit $(\mathcal{F}_k)$ une filtration. On appelle un temps d'arrêt, une variable aléatoire à valeur dans $\mathbb{N}$, t.q. pour tout $k \in \mathbb{N}$,

$$[T \leqslant k] \in \mathcal{F}_k .$$

Exemple 7.24 (premier passage)**.** Soit $j \in E$, $T_j = \arg\min\{k \in \mathbb{N} : X_k = j\}$ est un temps d'arrêt mesurant le premier passage de la chaîne par l'état j.

Exemple 7.25 (Dernier passage)**.** Soit $j \in E$, $T_j = \arg\max\{k \in \mathbb{N} : X_k = j\}$ n'est en principe pas un temps d'arrêt. En effet, pour $k \in \mathbb{N}$,

$$[T_j = k] = [X_k = j] \cap [X_{k+1} \neq j, X_{k+2} \neq j, \ldots]$$

est un événement qui dépend du futur et en principe n'appartient pas à $\mathcal{F}_k$.

Exemple 7.26 (Deux temps d'arrêts)**.** Soit T, S deux temps d'arrêts. Alors les variables $T + S$, $\max(T, S)$ et $\min(T, S)$ sont également des temps d'arrêts.

On peut alors définir la tribu $\mathcal{F}_T$ qui contient toutes les informations sur la chaîne de Markov jusqu'au temps T (qui on remarque est une variable aléatoire).

Définition 7.27. Soit T un temps d'arrêt. On définit $\mathcal{F}_T$ comme l'ensemble d'évènements $A \in \mathcal{F}_\infty := \sigma(\bigcup_{k\in\mathbb{N}} \mathcal{F}_k)$ t.q.

$$\forall k \in \mathbb{N}, \quad A \cap [T = k] \in \mathcal{F}_k .$$

Proposition 7.28. *La variable aléatoire $\mathbb{1}_{T<+\infty} X_T$ est $\mathcal{F}_T$ mesurable.*

Démonstration. On doit prouver que pour $A \subset E$, on a $[\mathbb{1}_{T<+\infty} X_T \in A] \in \mathcal{F}_T$. Soit $j \in E$ et $k \in \mathbb{N}$, on a

$$[X_T = j] \cap [T = k] = [X_k = j] \in \mathcal{F}_k\,.$$

Comme $[\mathbb{1}_{T<+\infty} X_T \in A]$ est une réunion au plus dénombrable d'éléments de tels événements le résultat est prouvé. □

On peut alors généraliser le théorème 7.16 comme suit.

Théorème 7.29 (Propriété de Markov forte)**.** *Soit $i \in E$, T un temps d'arrêt et $A \in \mathcal{F}_T$, t.q. $\mathbb{P}(T < +\infty, A, X_T = i) > 0$ on a alors :*

$$\mathbb{P}(X_{T+1} = j_1, \ldots, X_{T+n} = j_n | T < +\infty, A, X_T = i) = \mathbb{P}(X_1 = j_1, \ldots, X_n = j_n | X_0 = i)\,.$$

Démonstration. On doit prouver que

$$\mathbb{P}(X_{T+1} = j_1, \ldots, X_{T+n} = j_n, T < +\infty, A, X_T = i) \tag{7.1}$$

est égal à

$$\mathbb{P}(X_1 = j_1, \ldots, X_n = j_n | X_0 = i)\mathbb{P}(T < +\infty, A, X_T = i)\,.$$

En effet, l'équation (7.1) peut s'écrire comme

$$\begin{aligned}
&\sum_{k\in\mathbb{N}} \mathbb{P}(X_{k+1} = j_1 \ldots X_{k+n} = j_n, T = k, A, X_k = i)\\
&= \sum_{k\in\mathbb{N}} \mathbb{P}(X_{k+1} = j_1 \ldots X_{k+n} = j_n | A, T = k, X_k = i)\mathbb{P}(T = k, A, X_k = i)\\
&= \sum_{k\in\mathbb{N}} \mathbb{P}(X_{k+1} = j_1 \ldots X_{k+n} = j_n | X_k = i)\mathbb{P}(T = k, A, X_k = i) \qquad ,\\
&= \mathbb{P}(X_1 = j_1 \ldots X_n = j_n | X_0 = i) \sum_{k\in\mathbb{N}} \mathbb{P}(T = k, A, X_k = i)\\
&= \mathbb{P}(X_1 = j_1 \ldots X_n = j_n | X_0 = i)\mathbb{P}(T < +\infty, A, X_T = i)\,.
\end{aligned}$$

où pour passer à la troisième égalité on a utilisé le fait que $A \cap [T = k] \in \mathcal{F}_k$ et le théorème 7.16. □

7.2.4 États récurrents et transitoires

Soit $i \in E$ un état de la chaîne. La variable

$$N_i := \sum_{k\in\mathbb{N}} \mathbb{1}_{X_k=i}$$

représente le nombre de passage de la chaîne par l'état i.

Définition 7.30. On dit qu'un état i est **récurrent** si

$$\mathbb{P}(N_i = +\infty | X_0 = i) = 1\,.$$

Sinon, on dit que i est **transitoire** et

$$\mathbb{P}(N_i = +\infty | X_0 = i) = 0\,.$$

Il nous faudra démontrer la dernière affirmation de la définition précédente. Pour cela nous allons utiliser dans la proposition suivante une autre caractérisation de ces états faisant intervenir le temps d'arrêt $T_i = \arg\min\{k \in \mathbb{N} : X_k = i\}$.

Proposition 7.31. *Un état est récurrent si*

$$\mathbb{P}(T_i < +\infty | X_0 = i) = 1\,.$$

Un état est transitoire si

$$\mathbb{P}(T_i < +\infty | X_0 = i) < 1\,,$$

de plus pour un état transitoire on a

$$\mathbb{E}[N_i | X_0 = i] < +\infty$$

et en particulier $\mathbb{P}(N_i = +\infty | X_0 = i) = 0$.

Cette proposition est assez intuitive, en effet si on revient presque sûrement à l'état i une première fois alors on y reviendra presque sûrement une deuxième fois et ainsi de suite. La démonstration repose sur la propriété de Markov forte.

Démonstration. Notons pour tout $n \in \mathbb{N}$, $T_i^0 = T_i$ et $T_i^{n+1} = \arg\min\{k \in \mathbb{N} : k > T_i^n, X_k = i\}$ les passages successifs par l'état i. On peut écrire

$$\begin{aligned}
\mathbb{P}(N_i \geqslant n | X_0 = i) &= \mathbb{P}(T_i^n < +\infty | X_0 = i) \\
&= \mathbb{P}(T_i^n < +\infty, T_i^{n-1} < +\infty | X_0 = i) \\
&= \mathbb{P}(T_i^n < +\infty | T_i^{n-1} < +\infty, X_0 = i)\mathbb{P}(T_i^{n-1} < +\infty | X_0 = i) \\
&= \mathbb{P}(T_i^n < +\infty | X_{T_i^{n-1}} = i, T_i^{n-1} < +\infty)\mathbb{P}(N_i \geqslant n-1 | X_0 = i)\,.
\end{aligned} \tag{7.2}$$

De plus, on peut décomposer l'évènement $T_i^n < +\infty$ en évènements disjoints A_l avec

$$A_l = [X_{T_i^n + l} = i, X_{T_i^n + l - 1} \neq i, \dots X_{T_i^n + 1} \neq i]\,.$$

En définissant les évènements $B_l := [X_l = i, X_{l-1} \neq i, \dots X_1 \neq i]$ on a alors par la propriété de Markov fort.

$$\begin{aligned}
\mathbb{P}(T_i^n < +\infty | X_{T_i^{n-1}} = i, T_i^{n-1} < +\infty) &= \sum_{l \in \mathbb{N}} \mathbb{P}(A_l | X_{T_i^{n-1}} = i, T_i^{n-1} < +\infty) \\
&= \sum_{l \in \mathbb{N}} \mathbb{P}(B_l | X_0 = i) \\
&= \mathbb{P}(T_i < +\infty | X_0 = i)\,.
\end{aligned} \tag{7.3}$$

En combinant les équations (7.2) et (7.3) on trouve ainsi que la suite $\mathbb{P}(N_i \geqslant n | X_0 = i)$ est géométrique de facteur $\mathbb{P}(T_i < +\infty | X_0 = i) = q$. On a donc :

$$\begin{aligned}
\mathbb{P}(N_i = +\infty | X_0 = i) &= \mathbb{P}\left(\bigcap_{n \in \mathbb{N}} [N_i \geqslant n] | X_0 = i\right) \\
&= \lim_{n \to \infty} \mathbb{P}(N_i \geqslant n | X_0 = i) \\
&= \lim_{n \to \infty} q^n\,.
\end{aligned}$$

Cette limite vaut 1 ssi $q = 1$ et vaut 0 sinon. De plus,

$$\begin{aligned}
\mathbb{E}[N_i | X_0 = i] &= \sum_{n \in \mathbb{N}} n\mathbb{P}(N_i = n | X_0 = i) \\
&= \sum_{n \in \mathbb{N}} n\mathbb{P}(N_i \geqslant n | X_0 = i) - \sum_{n \in \mathbb{N}} n\mathbb{P}(N_i \geqslant n+1 | X_0 = i) \\
&= \sum_{n \geqslant 1} n\mathbb{P}(N_i \geqslant n | X_0 = i) - \sum_{n \geqslant 1} (n-1)\mathbb{P}(N_i \geqslant n | X_0 = i) \\
&= \sum_{n \geqslant 1} \mathbb{P}(N_i \geqslant n | X_0 = i) = \sum_{n \geqslant 1} q^{n-1}\,.
\end{aligned}$$

Cette série converge ssi $q < 1$.

□

Proposition 7.32. *Soit $j \in E$ un état transitoire et $i \in E$ un autre état. Alors*

$$\mathbb{P}(N_j < +\infty | X_0 = i) = 1$$

et

$$\mathbb{E}[N_j | X_0 = i] < +\infty\,.$$

Démonstration. Comme j est transitoire on a

$$\mathbb{P}(T_j < +\infty | X_0 = j) < 1\,.$$

Donc, pour $n \in \mathbb{N}$,

$$\begin{aligned}\mathbb{P}(N_j \geqslant n | X_0 = i) &= \mathbb{P}(N_j \geqslant n | X_0 = i, T_j < +\infty)\mathbb{P}(T_j < +\infty | X_0 = i)\\ &= \mathbb{P}(\sum_{k \geqslant T_j} \mathbb{1}_j(X_k) \geqslant n | X_{T_j} = j, T_j < +\infty, X_0 = i)\mathbb{P}(T_j < +\infty | X_0 = i)\\ &= \mathbb{P}(N_j \geqslant n | X_0 = j)\mathbb{P}(T_j < +\infty | X_0 = i)\\ &\leqslant \mathbb{P}(N_j \geqslant n | X_0 = j)\,.\end{aligned}$$

Donc $\mathbb{P}(N_j = +\infty | X_0 = i) \leqslant \mathbb{P}(N_j = +\infty | X_0 = j) = 0$ et

$$\mathbb{E}[N_j | X_0 = i] \leqslant \mathbb{E}_[N_j | X_0 = j] < +\infty\,.$$

□

Corolaire 7.33. *Soit (X_k) une chaîne de Markov de matrice de transition $\mathcal{P}$. Alors $i \in E$ est un état récurrent ssi*

$$\sum_{n \in \mathbb{N}} (\mathcal{P}^n)_{i,i} = +\infty\,.$$

De même i est un état transitoire ssi

$$\sum_{n \in \mathbb{N}} (\mathcal{P}^n)_{i,i} < +\infty\,.$$

De plus, si j est transitoire et $i \in E$ un état quelconque alors

$$\sum_{n \in \mathbb{N}} (\mathcal{P}^n)_{i,j} < +\infty\,.$$

Démonstration. Un état $i \in E$ est récurrent ssi $\mathbb{E}[N_i | X_0 = i] = +\infty$. Or

$$\begin{aligned}\mathbb{E}[N_i | X_0 = i] = \mathbb{E}[\sum_{n \in \mathbb{N}} \mathbb{1}_i(X_n) | X_0 = i] &= \sum_{n \in \mathbb{N}} \mathbb{E}[\mathbb{1}_i(X_n) | X_0 = i]\\ &= \sum_{n \in \mathbb{N}} (\mathcal{P}^n)_{i,i}\,.\end{aligned}$$

De même si j est transitoire on a toujours $\mathbb{E}[N_j | X_0 = i] < +\infty$ et cette espérance vaut

$$\sum_{n \in \mathbb{N}} (\mathcal{P}^n)_{i,j}$$

□

On en déduit que les propriétés de récurrence et de transience sont **une propriété de classe**.

Proposition 7.34. *Soit $i, j \in E$ t.q. $i \leftrightsquigarrow j$, alors i est récurrent ssi j l'est et i est transitoire ssi j l'est.*

Démonstration. Si $i \leftrightsquigarrow j$ il existe $l_1, l_2 \in \mathbb{N}$ t.q. $(\mathcal{P}^{l_1})_{i,j} > 0$ et $(\mathcal{P}^{l_2})_{j,i} > 0$. Donc

$$\sum_{n \in \mathbb{N}} (\mathcal{P}^n)_{i,i} \geqslant \sum_{m \in \mathbb{N}} (\mathcal{P}^{l_1})_{i,j}(\mathcal{P}^m)_{j,j}(\mathcal{P}^{l_2})_{j,i} = (\mathcal{P}^{l_1})_{i,j}(\mathcal{P}^{l_2})_{j,i} \sum_{m \in \mathbb{N}} (\mathcal{P}^m)_{j,j}\,.$$

Ainsi si j est récurrent i l'est et si i est transitoire j l'est. Par symétrie de la relation $\leftrightsquigarrow$ l'inverse est vrai aussi. □

7.3 Comportement asymptotique

Dans cette partie nous allons enfin caractériser le comportement asymptotique d'une chaîne de Markov.

7.3.1 Mesures invariantes

La notion importante est celle de la mesure de probabilité invariante. Comme précédemment on identifiera une probabilité sur E au vecteur ligne $\nu = (\nu(0), \nu(1), \ldots, \nu(n), \ldots)$ avec $\nu(n) = \nu(x_n)$.

Définition 7.35. Soit (X_k) une chaîne de Markov de matrice de transition $\mathcal{P}$. On dit qu'une probabilité ν est invariante (ou stationnaire) pour cette chaîne si $\nu = \nu\mathcal{P}$. Ceci s'écrit de manière équivalente comme :

$$\forall i \in E\,, \quad \nu(i) = \sum_{j \in E} \nu_j p_{j,i}\,.$$

Remarquons que ν est donc tout simplement un vecteur propre à gauche de $\mathcal{P}$ associé à la valeur propre 1 (de manière équivalente $\nu^\top$ est un vecteur propre à droite de la matrice $\mathcal{P}^\top$).

Ainsi, si $X_0 \sim \nu$, alors $X_1 \sim \nu\mathcal{P} = \nu$. De même pour tout $k \in \mathbb{N}$,

$$X_k \sim \nu\mathcal{P}^k = \nu\,.$$

La chaîne de Markov devient donc un processus stationnaire.

De plus, supposons que X_k convergence en loi vers une mesure de probabilité μ. Alors, pour tout $i \in E$

$$\mu(i) = \lim_{k\to\infty} (\nu^0\mathcal{P}^k)(i)\,,$$

et

$$\mu(i) = \lim_{k\to\infty} (\nu^0\mathcal{P}^{k+1})(i) = \lim_{k\to\infty} \sum_{j\in E} (\nu^0\mathcal{P}^k)(j)\mathcal{P}_{j,i}\,.$$

Ainsi, si on peut intervertir la limite et la somme (ce qui est par exemple le cas à partir du moment où E est fini), on obtient

$$\mu(i) = \sum_{j\in E} \lim_{k\to\infty} (\nu^0\mathcal{P}^k)(j)P_{j,i} = \sum_{j\in E} \mu(j)P_{j,i} = (\mu P)(i)\,.$$

Et donc μ est une probabilité invariante de la Chaîne de Markov. Les mesures invariantes caractérisent donc le comportement asymptotique de (X_k).

On peut déjà voir, qu'une mesure invariante associe une masse nulle aux états transitoires.

Proposition 7.36. *Soit $i \in E$ un état transitoire et ν une mesure invariante, alors $\nu(i) = 0$.*

Démonstration. En effet, on doit avoir pour un tel i et pour tout $k \in \mathbb{N}$,

$$\nu(i) = \sum_{j\in E} \nu(j)(\mathcal{P}^k)_{j,i}\,.$$

Or, pour tout $j \in E$, $(\mathcal{P}^k)_{j,i} \to_{k\to\infty} 0$ par le corollaire 7.33. Ceci implique $\nu(i) = 0$. □

7.3.2 Convergence

Dans cette partie, car c'est plus simple, nous nous plaçons dans le cas où l'espace des états E est **fini**. Les deux théorèmes importants sont les suivants.

Théorème 7.37. *Soit (X_k) une chaîne Markov à espace d'états fini.*

1. *(X_k) admet toujours une mesure de probabilité invariante.*

2. *Si de plus (X_k) est une chaîne de Markov **irréductible**, alors cette mesure invariante, qu'on notera ν est unique. Elle vérifie, pour tout $i \in E$,*

$$\nu(i) = \frac{1}{\mathbb{E}[T_i | X_0 = i]} .$$

3. *Enfin si (X_k) est **irréductible** et **apériodique**, alors X_k converge en loi vers ν la mesure invariante associée.*

Théorème 7.38. *Soit (X_k) une chaîne de Markov **irreductible** à espace d'états E fini et soit ν sa loi de probabilité invariante. Alors pour tout $d \in \mathbb{N}$ et toute fonction $f : E \to \mathbb{R}^d$, on a :*

$$\frac{1}{N} \sum_{i=0}^{N-1} f(X_i) \longrightarrow_{N\to\infty} \mathbb{E}_{X\sim\nu}[f(X)] \quad \textit{presque sûrement.}$$

De plus, on a que

$$\sqrt{N} \left(\frac{1}{N} \sum_{i=0}^{N-1} f(X_i) - \mathbb{E}_{X\sim\nu}[f(X)] \right)$$

converge en loi vers une gaussienne centrée.

7.3.3 Optionnel : preuve du théorème 7.37

Nous allons admettre la forme de la mesure invariante : $\nu(i) = \frac{1}{\mathbb{E}[T_i|X_0=i]}$, mais démontrer les autres résultats. Les premiers points proviennent des simples manipulations d'algèbre linéaire.

Preuve de l'existence de la mesure invariante. Notons n le cardinal de E. Remarquons que le vecteur $(1, \ldots, 1)$ est un vecteur propre, associé à la valeur propre 1, de $\mathcal{P}$. Ceci implique que $\mathcal{P}^\top$ admet aussi un vecteur propre $v = (v_1, \ldots, v_n)^\top$. Si toutes les coordonnées de ce vecteur sont positives alors $\nu^\top = \frac{v^\top}{\sum_{i=0}^n v_i}$ défini une probabilité invariante car :

$$\nu\mathcal{P} = (\mathcal{P}^\top \nu^\top)^\top = (\nu^\top)^\top = \nu .$$

Montrons que le vecteur $u = (|v_1|, |v_2|, \ldots, |v_n|)$ est un vecteur propre de $\mathcal{P}^\top$. En effet, on a pour tout $1 \leqslant i \leqslant n$:

$$|u_i| = \left| \sum_{j=1}^n v_j \mathcal{P}_{j,i} \right| \leqslant \sum_{j=1}^n |u_j| \mathcal{P}_{j,i} . \tag{7.4}$$

En particulier,

$$\sum_{i=1}^n |u_i| \leqslant \sum_{i=1}^n \sum_{j=1}^n |u_j| \mathcal{P}_{j,i} = \sum_{j=1}^n |u_j| .$$

Comme l'expression à gauche est égale à l'expression de droite, ceci signifie que toutes les inégalités d'auparavant étaient des égalités. En particulier, l'équation (7.4) implique que u est un vecteur propre, à coordonnées positives, associé à la valeur propre 1 □

Preuve de l'unicité de la mesure invariante. On va maintenant montrer que si (X_k) est irréductible alors il existe qu'un seul vecteur propre à droite de $\mathcal{P}$ associé à la valeur propre 1 (et donc de même pour les vecteurs propres à gauche). En effet, on sait que $(1, \ldots, 1)^\top$ est un vecteur propre de $\mathcal{P}$, supposons qu'il en existe un autre u, t.q. toutes ces coordonnées ne soient pas égales. Sans perte de généralité, supposons que $u_1 = \arg\max_{1\leqslant i \leqslant n} |u_i|$ et $u_2 \neq u_1$ (si ce n'est pas le cas il suffit de permuter la base canonique). Comme (X_k) est irréductible, tous les états communiquent et il existe $l \in \mathbb{N}$ t.q. $\mathcal{P}^l_{1,2} = \mathbb{P}(X_l = 2 | X_1 = 1) > 0$.

Donc

$$u(1) = (\mathcal{P}^l u)(1) = \sum_{j=1}^n \mathcal{P}^l_{1,j} u(j) = u(2)\mathcal{P}^l_{1,2} + \sum_{j\neq 2} \mathcal{P}^l_{1,j} u(j) < u(1) \sum_{j=1}^n \mathcal{P}^l_{1,j} = u(1) .$$

On trouve $u(1) < u(1)$, ce qui est absurde. Donc la dimension de l'espace engendré par les vecteurs propres de $\mathcal{P}$ associé à la valeur 1 est de 1. Il en est donc de même pour $\mathcal{P}^\top$.

□

Pour montrer le résultat sur l'apériodicité nous aurons besoin d'un résultat intermédiaire.

Proposition 7.39. *Soit (X_k) une chaîne irréductible, apériodique, à espace d'états fini. Alors, il existe $k_0 \in \mathbb{N}$, t.q. pour tout $i, j \in E$, et $k \geqslant k_0$, $\mathbb{P}(X_k = j | X_0 = i) > 0$.*

Démonstration. Pour $i \in E$, notons $A(i) = \{k \geqslant 1 : \mathbb{P}(X_k = i | X_0 = i) > 0\}$. La preuve se passe en plusieurs étapes.

1) Il existe $k(i) \in A(i)$ t.q. $k(i) + 1 \in A(i)$. En effet, comme la chaîne est apériodique, il existe $k_1, \ldots, k_l \in A(i)$ t.q. leur pgcd est égal à 1. Alors, par le théorème de Bézout, il existe $a_1, \ldots, a_l \in \mathbb{Z}$ t.q.

$$\sum_{j=1}^{l} a_j k_j = 1 \,. \tag{7.5}$$

Définissons

$$k(i) = \sum_{1 \leqslant j \leqslant l, a_j < 0} |a_j| k_j \,.$$

Si $k - i) = 0$, ça veut dire qu'un des entiers parmi $n_1, \ldots, n_l$ est égal à 1 et le résultat est prouvé. Sinon, $k(i) \in A(i)$ est $k(i) + 1 \in A(i)$, car d'après l'équation (7.5),

$$k(i) + 1 = \sum_{1 \leqslant j \leqslant l, a_j > 0} a_j k_j \,.$$

2) Pour tout $k \geqslant k^2(i)$, on a $\mathbb{P}(X_k = i | X_0 = i) > 0$. En effet, pour un tel k on peut, par l'algorithme de la division euclidienne, trouver un $m \in \mathbb{N}$ et $0 \leqslant l < k(i)$

$$k = (k(i) + m)k(i) + l = l(k(i) + 1) + (m + k(i) - l)k(i) \,.$$

Comme $l \geqslant 0$ et $m + n(i) - l \geqslant 0$ le résultat est prouvé.

3) Pour chaque $j \in E$, il existe $k_{j,i}$, t.q. pour tout $k \geqslant k_{j,i}$, $\mathbb{P}(X_k = i | X_0 = j) > 0$. Comme la chaîne est irréductible il existe $l \in \mathbb{N}$, t.q. $\mathbb{P}(X_l = i | X_0 = j) > 0$. Il suffit donc prendre $k_{j,i} = l + k^2(i)$.

4) Fin de la preuve. La preuve est alors finie en prenant $k_0 = \max_{i,j \in E} k_{i,j}$. □

Enfin, on aura besoin du théorème suivant.

Théorème 7.40 (Théorème de Perron-Frobenius)**.** *Soit $A \in \mathbb{R}^{n \times n}$ une matrice réelle dont tous les coefficients sont strictement positifs. Alors la plus grande, en module, des valeurs propres complexes de A est réelle et simple.*

Preuve de la convergence en loi d'une chaîne apériodique. Par le résultat de la proposition précédente, il existe un $k \in \mathbb{N}$, t.q. $\mathcal{P}^k$ est une matrice dont toutes les coordonnées sont strictement positives. On peut alors appliquer le théorème de Perron-Frobenius pour trouver que la plus grande valeur propre de $\mathcal{P}^k$, λ est réelle et est une valeur propre simple.

1)Montrons que $\lambda = 1$. En effet, on sait que $(1, \ldots, 1)$ est un vecteur propre de $\mathcal{P}$ (et donc de $\mathcal{P}^k$) associé à la valeur propre 1. On a donc que $\lambda \geqslant 1$. De plus, soit u le vecteur propre associé à λ et prenons i_0 l'indice de la coordonnée maximale (en valeur absolue) de u.

$$|(\mathcal{P}^k u)_{i_0}| = \lambda |u_{i_0}| = \left| \sum_{j \in E} (\mathcal{P}^k)_{i,j} u_j \right| \leqslant u_{i_0} \sum_{j \in E} (\mathcal{P}^k)_{i,j} = |u_{i_0}| \,.$$

Donc $\lambda = 1$ et c'est une valeur propre simple.

2) Montrons que 1 est une valeur propre simple de $\mathcal{P}$ et toutes les autres valeurs propres de $\mathcal{P}$ sont de module strictement inférieur à 1. En effet, si u est un vecteur propre associé à la valeur propre λ, alors $\mathcal{P}^k u = \lambda^k u$ et u est donc un vecteur propre de $\mathcal{P}^k$ de valeur propre λ^k. Le résultat se déduit alors du point *1)* et du théorème de Perron-Frobenius.

La décomposition de Jordan nous donne alors l'existence d'une matrice $A \in \mathbb{R}^{(n-1)\times(n-1)}$ dont la plus grande valeur propre est de module strictement inférieur à 1 et $Q \in \mathbb{R}^{n\times n}$, inversible, vérifiants

$$\mathcal{P} = Q \begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix} Q^{-1} \tag{7.6}$$

3) Montrons que la première colonne de Q est proportionnelle au vecteur $(1,\dots,1)^\top$. En effet, on sait que $(1,\dots,1)^\top$ est le vecteur propre (à droite) de $\mathcal{P}$ associé à la valeur propre 1. L'équation (7.6) implique alors que la première colonne de Q est un vecteur propre associé à 1 et est donc proportionnel à $(1,\dots,1)^\top$.

4) Montrons que la première ligne de Q^{-1} est proportionnelle à ν, la probabilité invariante de (X_k). Comme la chaîne est irréductible, le seul, à constante de proportionnalité près, vecteur propre (à gauche) de $\mathcal{P}$ associé à la valeur propre 1 est ν. L'équation (7.6) implique alors que la première ligne de Q^{-1} est ce vecteur propre.

5) Fin de la preuve.

Finalement,

$$\mathcal{P}^l = Q \begin{pmatrix} 1 & 0 \\ 0 & A^l \end{pmatrix} Q^{-1} \to_{l\to+\infty} = Q \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} Q^{-1} \propto \begin{pmatrix} \nu(1) & \cdots & \nu(n) \\ \nu(1) & \cdots & \nu(n) \\ \vdots & \cdots & \vdots \\ \nu(1) & \cdots & \nu(n) \end{pmatrix},$$

où on a utilisé le fait que toutes les valeurs propres de A sont de module strictement inférieur à 1 et donc $A^l \to 0$.

En particulier, si $X_0 \sim \nu^0$ on a pour tout $1 \leqslant i \leqslant n$,

$$\mathbb{P}(X_l = i) = (\nu^0 \mathcal{P}^l)(i) \to \nu(i) \sum_{j=1}^n \nu^0(j) = \nu(i)\,.$$

$\square$

Travaux Dirigés

Chaînes de Markov

Ci-dessous deux exercices pour vous entraîner. Vous pouvez trouver des exercices supplémentaires (avec correction) dans le polycopié de Béatrice de Tilière sur ce lien `https://www.ceremade.dauphine.fr/~detiliere/Cours/polycop_Proc_Stoch.pdf`. Ça vous permettra d'être à l'aise sur les notions de matrice de transition, récurrence, périodicité...

Exercice III.1 (Transmission d'un signal informatique)**.** Soit $(Y_k)_{k\geqslant 1}$ une suite i.i.d. avec $\mathbb{P}(Y_1 = 1) = p$ et $\mathbb{P}(Y_1 = -1)$. Au temps 0 on dispose d'un signal $X_0 \in \{-1, 1\}$ un signal. Au temps n, le signal est altéré par $X_{n+1} = Y_{n+1}X_n$.

1. Dessiner le graphe et écrire $\mathcal{P}$ la matrice de transition de la chaîne.
2. Trouver les vecteurs et les valeurs propres de $\mathcal{P}$.
3. Calculer $\mathbb{P}(X_n = X_0)$. Quelle est sa limite quand $n \to +\infty$?

Exercice III.2 (Joueur de Casino)**.** Tous les jours un joueur joue dans un des trois casinos de la ville. Chaque jour il choisit un des deux casinos où il n'est pas allé la veille avec probabilité 1/2. On note X_n le numéro du casino visité le jour n.

1. Dessiner le graphe et écrire $\mathcal{P}$ la matrice de transition de la chaîne.
2. Calculer $\mathcal{P}^n$. On pourra remarquer que

$$2\mathcal{P} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

3. Soit μ la loi de X_0, calculer $\lim_{n\to+\infty} \mathbb{P}(X_n = j)$ pour $j \in \{1, 2, 3\}$.

Exercice III.3 (Chaîne à deux états)**.** Soit (X_k) une chaîne de Markov à deux états $\{0, 1\}$, où $\mathbb{P}(X_1 = 1|X_0 = 0) = \alpha$ et $\mathbb{P}(X_1 = 0|X_0 = 1) = \beta$.

1. Dessiner le graphe et écrire la matrice de transition de la chaîne.
2. Décrire, selon les valeurs de α, β les classes d'irréductibilité de la chaîne et leur périodicité.
3. Décrire, selon les valeurs de α, β, les états récurrents et transitoires.
4. Pour $\alpha, \beta \in (0, 1)$, trouver la probabilité invariante de la chaîne.
5. Dans le cadre de la question précédente, quelle est la valeur de $\mathbb{E}[T_0|X_0 = 0]$, où $T_0 = \inf\{k > 0 : X_k = 0\}$ est le temps du premier passage en 0.
6. Quelle est la valeur de $\mathbb{E}[T_0|X_0 = 1]$?

Exercice III.4 (Joueur de Casino (suite))**.** On se place dans le cadre de l'exercice III.2.

1. Déterminer les classes d'irréductibilité de la chaîne et leur périodicité.
2. Trouver une probabilité invariante de la chaîne. Est-elle unique ?
3. Soit $T = \arg\min\{n \geqslant 1, X_n = 1\}$. Calculer $\mathbb{E}[T|X_0 = 1]$.
4. Comment pouvait-on procéder à la dernière question de l'exercice III.2 pour éviter les calculs ?

Exercice III.5 (Chaîne à 5 états). Soit (X_k) une chaîne de Markov avec la matrice de transition

$$\mathcal{P} = \begin{pmatrix} 1/2 & 0 & 1/2 & 0 & 0 \\ 1/4 & 1/2 & 1/4 & 0 & 0 \\ 1/2 & 1/4 & 1/4 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1/2 & 1/2 \end{pmatrix}.$$

1. Dessiner le graphe correspondant.
2. Classer les états de la chaîne (irréductibilité, transitivité, périodicité).
3. Trouver 2 probabilités invariantes de la chaîne à support disjoint.
4. Montrer que la chaîne possède une infinité de mesures invariantes.
5. Soit $T_2 = \inf\{k \geqslant 1, X_k = 2\}$. Calculer $\mathbb{E}[T_2|X_0 = 2]$.
6. Si ν^0 la loi de X_0 est à support dans $\{1, 2, 3\}$ (respectivement dans $\{4, 5\}$) est-ce que (X_k) converge en loi ? Préciser la limite au cas échéant.
7. Discuter de la convergence en loi dans le cas général.
8. Soit $f : E \to \mathbb{R}^d$ une fonction. Supposons que X_0 soit distribué selon ν. Quelle est la limite de
$$\frac{1}{N}\sum_{i=1}^{N} f(X_i)\,,$$
en fonction de ν^0 la loi de X_0 quand N tend vers l'infini.
9. Que changerait dans les questions précédentes si on met $\mathcal{P}_{3,3} = 1/4 - \varepsilon$, $\mathcal{P}_{4,3} = \mathcal{P}_{3,4} = \varepsilon$ et $\mathcal{P}_{4,4} = 1/2 - \varepsilon$?

Exercice III.6 (Chaîne à 4 états). Soit (X_k) une chaîne de Markov de matrice de transition

$$\mathcal{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

1. Dessiner le graphe correspondant.
2. Classer les états de la chaîne (irréductibilité, transitivité, périodicité).
3. Soit $T = \inf\{k \in \mathbb{N} : X_k = 4\}$. Exprimer $\mathbb{P}(T < +\infty|X_0 = 2)$ en fonction de $\mathbb{P}(T < +\infty|X_0 = 3)$.
4. Calculer la probabilité que la chaîne soit absorbée par l'état 4 en étant issue de 2.
5. Soit $T' = \inf\{k \in \mathbb{N} : X_k \in \{1, 4\}\}$, calculer $\mathbb{E}[T'|X_0 = 2]$.

Exercice III.7 (Le modèle de l'urne d'Ehrenfest [3]). Nous supposons être en présence de N molécules disposés dans deux urnes A, B. À chaque instant une des molécules, parmi les N, est choisie aléatoirement et est changée d'urne. Pour $k \in \mathbb{N}$, on notera X_k le nombre de molécules dans l'urne A à l'instant k

1. Écrire la matrice de transition de la chaîne.

3. C'est un modèle introduit en 1911 par Paul Ehrenfest et sa femme Tatiana Afanassieva pour illustrer certains paradoxes de la physique statistique. En effet, nous savons que l'entropie d'un tel système doit augmenter et donc tendre vers l'état parfaitement mélangé où les molécules sont parfaitement reparties entre les deux urnes.
Au premier abord, la théorie des chaînes de Markov nous dit que ce n'est pas vrai, puisque l'état 0 (contraire au mélange parfait) est toujours accessible et sera même visité une infinité de fois. Cependant, le résultat de la dernière question nous montre que le temps d'en arriver à cet état est exponentiellement grand — 2^N en nombre de molécules. Ainsi, cet état est en pratique non observable par aucune de nos expériences.

2. Déterminer les classes d'irréductibilité de la chaîne.
3. Déterminer la période des états de la chaîne.
4. Quels sont les états récurrents, transitoires ?
5. Donner les N équations que doit vérifier la mesure invariante $\nu = (\nu_1, \ldots, \nu_N)$.
6. Montrer par récurrence que pour tout $k \in \{0, \ldots, N\}$:

$$\nu_k = \nu_0 \binom{N}{k} .$$

7. Trouver ν_0 et en déduire la forme de la mesure invariante.
8. Est-ce que la chaîne est garantie de converger en loi vers cette mesure invariante ?
9. Soit $\varepsilon \in (0, 1)$, on considère maintenant qu'à chaque étape, il y a une probabilité ε qu'on ne déplace pas la molécule choisie. Écrire la nouvelle matrice de transition de la chaîne.
10. Qu'est ce qui change (périodicité, récurrence, mesure invariante ...) par rapport au modèle précédent ?
11. À quoi est égal $\mathbb{E}[T_i|X_i = i]$? Supposons que N est pair, $N = 2l$, en utilisant la formule de Stirling, comparer $\mathbb{E}[T_0|X_0 = 0]$ et $\mathbb{E}[T_l|X_0 = l]$ quand $N \to \infty$.

Quatrième partie

Travaux Pratiques

TP1 Martingales

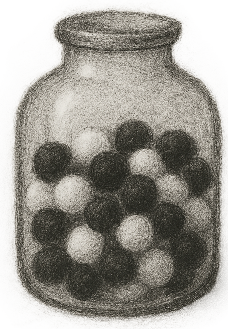
Jules Flin

Exercice 1. (*la ruine du joueur et théorème d'arrêt*) On considère la suite de variables aléatoires $(X_n)_n$ définie par $X_0 = 0$ et $X_{n+1} = X_n + \xi_{n+1}$ où les ξ_n sont *i.i.d.* de loi commune Rad. On se donne deux entiers $a > 0$ et $b > 0$, et on s'intéresse à la variable aléatoire

$$T = \inf\{n, X_n \notin (-a, b)\}$$

1. Écrire une fonction `partie()` qui renvoie une trajectoire de $(X_n)_{n \leqslant T}$, et en tracer plusieurs.
2. Montrer que $(X_n)_n$ est une martingale par rapport la filtration $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$.
3. En appliquant le théorème d'arrêt, calculer $\mathbb{P}(X_T = b)$.
4. Illustrer ce résultat par une méthode de Monte-Carlo.
5. Montrer que $Z_n := X_n^2 - n$ est une $\mathcal{F}_n$-martingale.
6. En déduire $\mathbb{E}[T]$.
7. Illustrer ce résultat par une méthode de Monte-Carlo.

Exercice 2. (*l'urne de Pólya et convergence de martingales*) On s'intéresse à une urne contenant des boules rouges et bleues. On suppose qu'initialement l'urne contient A boules rouges et B boules bleues. A chaque étape, on tire uniformément au hasard une boule, et on la remet dans l'urne avec une boule de même couleur. On s'intéresse à la fraction de boules rouges au cours du temps.



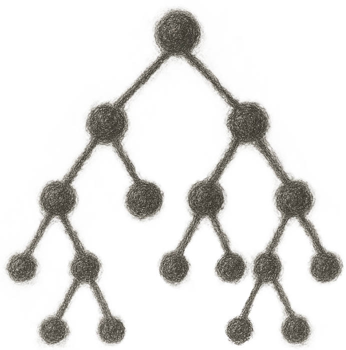
1. On note X_k la proportion de boules rouges à la k-ième étape. Montrer que $(X_k)_k$ est une martingale, et qu'elle converge presque sûrement.
2. Écrire une fonction `polya(A,B,n)` qui renvoie une réalisation de la suite $(X_k)_{k \leqslant n}$ quand l'urne est initialisée avec `A` boules rouges et `B` boules bleues.
3. Illustrer la convergence presque sûre en traçant plusieurs trajectoires de X en faisant varier les paramètres.
4. On peut montrer que (X_n) converge en loi vers une variable aléatoire $X \sim \beta(A, B)$. Simuler un K-échantillon de X_n pour $K = 100$ et $n = 1000$, et superposer un histogramme de cet échantillon avec la densité d'une loi beta de paramètres `A` et `B`.
5. On s'intéresse au cas particulier $A = B = 1$, et à la loi limite. Simuler un K-échantillon dans ce cas là, et proposer une expression plus simple de la loi $\beta(1, 1)$.
6. Prouver le résultat conjecturé à la question précédente.

Exercice 3. (*application des martingales dans l'étude de l'arbre de Galton-Watson surcritique*) Dans ce dernier exercice, on s'intéresse à un modèle issu de la biologie : l'arbre de Galton-Watson. On en donne ici une définition naïve. Soit $(\xi_{i,j})_{(i,j)}$ un ensemble de variables aléatoire *i.i.d.* à valeurs dans $\mathbb{N}$ et intégrables d'espérance μ. Considérons la suite $(Z_n)_n$ définie par récurrence par $Z_0 = 1$ et

$$Z_{n+1} = \sum_{i=1}^{Z_n} \xi_{n+1,i}$$

avec la convention $\sum_\emptyset = 0$. On peut interpréter Z_n comme la taille d'une population à la génération n : $\xi_{n+1,i}$ est alors le nombre de descendants du i individu de la génération n.

1. Rédiger un court programme qui simule un trajectoire de $(Z_n)_n$. Tester ce programme avec plusieurs lois/paramètres et conjecturer le comportement limite de $(Z_n)_n$ quand $\mu < 1$ et $\mu > 1$.
2. On suppose que $\mu > 1$. Montrer que $(Z_n/\mu^n)_n$ est une martingale.
3. En déduire $\mathbb{E}[Z_n]$ et l'illustrer par des méthodes de Monte-Carlo.
4. On suppose que $\mathrm{Var}(\xi_{0,0}) = \sigma^2 < +\infty$. Montrer que $(Z_n/\mu^n)_n$ converge `p.s.` et dans $\mathcal{L}^2$ vers une variable aléatoire Y_∞.
5. Tracer la distribution empirique (approchée) de Y_∞, conditionnellement à $Y_\infty \neq 0$.



TP2 Mouvement brownien

Jules Flin

Exercice 1. (*simulation du mouvement brownien*) Dans cet exercice, on cherche à produire des simulations du mouvement brownien. On rappelle que le mouvement brownien standard est un processus à accroissements indépendants, stationnaires et gaussiens, et dont les trajectoires sont *p.s.* continues.

1. Écrire une fonction `bm(T,...)` qui trace une trajectoire du mouvement brownien standard sur $[0,T]$. On pourra ajouter d'autres arguments.
2. Adapter la fonction précédente pour tracer une trajectoire du mouvement brownien, avec drift μ, variance σ^2 et partant de $x \in \mathbb{R}$.
3. Adapter les fonctions précédentes pour tracer une trajectoire d'un mouvement brownien bidimensionnel.
4. Adapter les fonctions précédentes pour tracer une trajectoire d'un mouvement brownien tridimensionnel.

Exercice 2. (*quelques propriétés du mouvement brownien standard*)

1. On rappelle que le mouvement brownien standard vérifie le principe de réflexion : pour tout $a > 0$ et tout $t > 0$,
$$\mathbb{P}(\sup_{s \leqslant t} B_s \geqslant a) = 2\mathbb{P}(B_t \geqslant a) = \mathbb{P}(|B_t| \geqslant a)$$
Illustrer cette propriété à l'aide d'une méthode de Monte-Carlo.
2. Dans cette question, on s'intéresse à la distribution de
$$Y := \operatorname{argmax}\{B_s : s \in [0,1]\}$$
Écrire un programme qui génère un K-échantillon de Y, trace un histogramme. Comparer cet histogramme avec la densité de la loi arc sinus
$$f(x) = \frac{1}{\pi\sqrt{x(1-x)}}$$
3. On se donne $a > 0$ et $b > 0$. Notons $T_{-a} := \inf\{t > 0 : B_t = -a\}$, $T_b := \inf\{t > 0 : B_t = b\}$ et $T := T_{-a} \wedge T_b$. Comme pour la marche aléatoire symétrique simple sur $\mathbb{Z}$, on a
$$\mathbb{P}(T = T_b) = \frac{a}{a+b}$$
Vérifier cette affirmation numériquement.

Exercice 3. (*autour du Théorème de Donsker*) Dans cet exercice, on illustre un résultat incontournable à propos du mouvement brownien : c'est la limite d'échelle d'une marche aléatoire discrète ! Plus précisément, le théorème de Donsker peut s'énoncer de la manière suivante : soit $(X_n)_n$ une suite de variables aléatoires *i.i.d.* de carrés intégrables, centrées et réduite. On note $(S_n)_n$ la marche aléatoire associée
$$S_n := \sum_{k=1}^{n} X_k$$

On interpole et normalise les trajectoires de cette marche aléatoire

$$X_n(t) := \frac{1}{\sqrt{n}}\left(S_{\lfloor nt \rfloor} + \{nt\}X_{\lfloor nt \rfloor + 1}\right)$$

$(X_n(t))_t$ est un variable aléatoire à valeurs dans $[0,1]$, et elle converge en loi vers un mouvement brownien standard.

1. Écrire un programme qui trace les trajectoires de $(S_n)_n$, jusqu'à un instant N.
2. Écrire un programme qui trace le graphe de X_n pour un n fixé.
3. Illustrer le résultat de Donsker en simulant X_n pour n suffisamment grand.

Exercice 4. (*des variantes du mouvement brownien*) Dans ce dernier exercice, on simule quelques variantes du mouvement brownien, fréquemment rencontrées dans des applications.

1. (*mouvement brownien géométrique*) le mouvement brownien géométrique représente, dans le modèle de Black-Scholes, le prix d'une action en bourse. On se donne pour valeur initiale S_0 (variable aléatoire positive), $(B_t)_t$ un mouvement brownien drifté par μ et de coefficient de diffusion σ^2, et on pose
$$S_t := S_0 \exp\left(\mu t - \frac{\sigma^2}{2}t + \sigma B_t\right)$$
Tracer des trajectoires de ce processus.
2. (*pont brownien*) le pont brownien est le processus obtenu en conditionnant un mouvement brownien à valeur 0 aux temps 0 et 1. Tracer des trajectoires de ce processus. On ne peut évidemment pas appliquer de méthodes de rejet.
3. (*excursion brownienne*) l'excursion brownienne est un pont brownien conditionné à rester positif. On va simuler ce processus à l'aide de la transformée de Verwaat. On note
$$T := \operatorname{argmin}\{W_t : t \in [0,1]\}$$
où $(W_t)_t$ est un pont brownien sur $[0,1]$. Notons Z_t le processus défini par
$$Z_t := \begin{cases} W_{T+t} - W_T \text{ si } 0 \leqslant t \leqslant 1-T \\ W_{T+t-1} - W_T \text{ sinon.} \end{cases}$$
Simuler des trajectoires de ce dernier processus.

TP3 Chaînes de Markov

Jules Flin

Exercice 1. (*paradoxe du singe savant*)

On s'intéresse à l'expérience suivante : un singe tape des lettres au hasard sur une machine à écrire. A chaque étape, une nouvelle lettre est tirée uniformément au hasard parmi les lettres d'un alphabet $\mathcal{A}$ (indépendamment de toutes les lettres tapées auparavant), de sorte à obtenir un mot aléatoire infini $M = a_1 a_2 \ldots$ sur $\mathcal{A}$. On se fixe un texte de référence M_0 de longueur ℓ, et on s'intéresse au temps aléatoire T_0 (à valeurs dans $\mathbb{N} \cup \{+\infty\}$ défini par

$$T_0 = \inf\{n : a_{n-\ell+1} \ldots a_n = M_0\}$$

autrement dit, T_0 est le nombre de frappes avant d'observer une première occurrence du texte de référence. On peut par exemple prendre l'alphabet latin pour $\mathcal{A}$ et *Hamlet* de Shakespeare pour M_0. Toutefois, et afin de simplifier le code, on choisira $\mathcal{A} := \{1, \ldots, N\}$ (en faisant correspondre chaque lettre à sa position dans l'alphabet).

1. Montrer que $T_0 < +\infty$ presque sûrement (*paradoxe du singe savant*). Dans le reste de l'exercice, on cherche à estimer précisément ce temps.
2. Soit X_n la taille du plus long suffixe de $a_1 \ldots a_n$ qui soit un préfixe de M_0. Montrer que $(X_n)_n$ est une chaîne de Markov. Comment interpréter T_0 dans ce contexte, si on fixe $\mathbb{P}(X_{n+1} = 0 \,|\, X_n = \ell) = 1$?
3. Écrire un programme qui prend en entrée le texte de référence M_0 et renvoie la matrice de transition de cette chaîne de Markov. On pourra tester cette fonction sur des mots courts.
4. Écrire un programme qui à la matrice de transition associe le temps moyen $\mathbb{E}[T_0]$ (*indice : on pourra écrire une fonction auxiliaire qui renvoie un vecteur propre à gauche de la matrice de transition pour la valeur propre* 1*...*).
5. La loi de T_0 ne dépend en réalité que de la *structure* de M_0 (autrement dit, permuter les lettres de $\mathcal{A}$ laisse invariant la loi de T_0). Dénombrer, à ℓ fixé, ces structures. Pour chaque structure, dénombrer les mots de longueur ℓ sur $\mathcal{A}$ qui admettent cette structure.
6. Écrire un court programme qui prend en entrée ℓ et renvoie la liste de toutes les structures de longueur ℓ.
7. On peut à présent se placer dans un contexte plus général : dorénavant le mot M_0 est tirée uniformément au hasard parmi les mots de longueur ℓ sur $\mathcal{A}$. Écrire un programme qui calcule, à partir de la longueur ℓ le temps moyen avant d'observer le mot de référence (aléatoire) $\mathbb{E}[T_0]$.

Exercice 2. (*flaques aléatoires*) Dans ce dernier exercice, on considère une suite aléatoire de sous-ensembles $(A_n)_n$ du réseau $\mathbb{Z}^d$ construits de la manière suivante : $A_0 = \{\mathbf{0}\}$ ($\mathbf{0}$ est l'origine de $\mathbb{Z}^d$), et pour tout n, $A_{n+1} = A_n \cup \{X_n\}$ où X_n est le premier état de $\overline{A}_n$ visité par une marche aléatoire symétrique sur $\mathbb{Z}^d$, initialisée en $\mathbf{0}$ (toutes ces marches sont supposées indépendantes). Pour $d = 2$, on peut (par exemple) interpréter $(A_n)_n$ comme une flaque se formant sous un réservoir qui fuit au goutte à goutte.

1. Écrire un programme qui simule (et représente séquentiellement) une trajectoire de la chaîne de Markov $(A_n)_n$.
2. On s'intéresse au rayon $(R_n)_n$ de l'amas A_n :
$$R_n := \sup\{|a|,\ a \in A_n\}$$
La suite $(R_n)_n$ est-elle une chaîne de Markov ?
3. À l'aide de méthodes de Monte Carlo, conjecturer un équivalent $f_d(n)$ pour $(\mathbb{E}[R_n])_n$ (dépend de la dimension d).
4. En simulant un grand amas (avec $d = 2$), conjecturer une limite pour $(A_n/f_d(n))_n$.

Annexes

Annexe A

Rappels de probabilités et de théorie de la mesure

A.1 Tribus

Soit Ω un ensemble.

Définition A.1. Une tribu $\mathcal{F}$ est un ensemble de parties de Ω vérifiant les propriétés suivantes.

- $\emptyset \in \mathcal{F}$.
- Stabilité par passage au complémentaire : pour $A \in \Omega$, $A^c \in \Omega$.
- Stabilité par union dénombrable : pour $(A_n) \in \Omega^{\mathbb{N}}$, on a $\bigcup_{n\in\mathbb{N}} A_n \in \Omega$.

Exemple A.2. Exemples

- La tribu $\{\emptyset, \Omega\}$ est dite la tribu grossière.
- La tribu $\mathcal{P}(\Omega)$ est dite la tribu discrète.
- Si $\Omega = \mathbb{R}^d$, alors $\mathcal{B}(\mathbb{R}^d)$, la tribu des boréliens, est la tribu engendrée par les ouverts de $\mathbb{R}^d$.

Définition A.3. Soit $\mathcal{C}$ un ensemble de partie de Ω, on appelle la tribu engendrée par $\mathcal{C}$, noté $\sigma(\mathcal{C})$ la plus petite tribu sur Ω qui contient $\mathcal{C}$.

Si $\mathcal{F}$ est une tribu sur Ω on parle de l'espace mesurable $(\Omega, \mathcal{F})$. La définition de mesure est alors assez intuitive, on veut que le volume du vide soit nul et que la mesure de deux ensembles disjoints soit la somme de leurs mesure.

Définition A.4. On dit que $\mu : \mathcal{F} \to \mathbb{R}_+$ est une mesure si elle vérifie les propriétés suivantes.

- $\mu(\emptyset) = 0$.
- Pour $A_n \in \Omega^{\mathbb{N}}$ disjoints on a

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n)$$

Le triplet $(\Omega, \mathcal{F}, \mu)$ est alors nommé un espace mesuré.

Définition A.5. On dit qu'une mesure μ est une mesure de probabilité si $\mu(\Omega) = 1$. Le triplet $(\Omega, \mathcal{F}, \mu)$ est alors nommé un espace probabilisé.

Si μ est une mesure de probabilité on la notera souvent $\mathbb{P}$

Remarque A.6 (Motivation)**.** L'idée derrière la notion de tribu est de fournir un ensemble de parties qu'on peut mesurer. Bien sûr, $\mathcal{P}(\Omega)$, l'ensemble des parties de Ω forme toujours une tribu. Cependant, elle est, souvent, trop grande. Par exemple, on peut montrer, qu'il est possible de découper une boule dans $\mathbb{R}^3$ en un nombre fini de morceaux, puis de les déplacer et les réassembler afin de produire deux copies identiques de la boule initiale. Ce paradoxe, appelé le paradoxe de Banach-Tarski, est contraire à notre intuition physique. Mathématiquement, il provient du fait que sur la tribu $\mathcal{P}(\mathbb{R}^3)$ il n'est pas possible de définir une mesure de volume. Ainsi, quand $\Omega = \mathbb{R}^d$ on se restreint généralement à $\mathcal{B}(\mathbb{R}^d)$, la tribu des boréliens qui est la tribu engendrée par les ouverts de $\mathbb{R}^d$.

A.2 Variables aléatoires

Définition A.7. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et $(E, \mathcal{E})$ un espace mesuré. On dit que $X : \Omega \to E$ est une variable aléatoire, si pour tout $B \in \mathcal{E}$, $X^{-1}(B) \in \Omega$.

Si E est un espace au plus dénombrable on prendra généralement $\mathcal{E} = \mathcal{P}(E)$, on dit dans ce cas que X est une *variable aléatoire discrète*. Si $(E, \mathcal{E}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ on dira que X est une variable aléatoire réelle. Si $(E, \mathcal{E}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ on dit que X est un vecteur aléatoire.

Définition A.8. Soit $X : \Omega \to E$ une variable aléatoire, on appelle $\sigma(X)$ la plus petite tribu sur Ω t.q. pour tout $B \in \mathcal{E}$, $X^{-1}(B) \in \sigma(X)$.

$\sigma(X)$ est en fait juste $\sigma(\{X^{-1}(B) : B \in \mathcal{E}\})$. Ainsi, X peut être vu comme une variable aléatoire allant de $(\Omega, \sigma(X), \mathbb{P})$ dans $(E, \mathcal{E})$.

De même pour une famille quelconque de variables aléatoires $(X_i)_{i \in I}$, on définit

$$\sigma(X_i, i \in I) = \sigma(\{X_i^{-1}(B) : i \in I, B \in \mathcal{E}\}) \ .$$

C'est la plus petite tribu qui rend toutes les variables $(X_i)_{i \in I}$ mesurables **simultanément**.

Une intuition qui peut être utile est de se dire que la tribu $\sigma(X)$ contient toutes les informations nécessaires à décrire la variable aléatoire X. En particulier, pour tout $A \in \mathcal{E}$, l'évènement $[X \in A] = \{w \in \Omega : X(w) \in A\}$ est un élément de la tribu $\sigma(X)$. Ainsi, pour parler de $\mathbb{P}(X \in A)$, nous n'avons pas besoin de définir $\mathbb{P}$ sur tout $\mathcal{F}$, mais seulement sur $\sigma(X)$.

A.3 Espérance conditionnelle

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace mesuré. Soit $\mathcal{G} \subset \mathcal{F}$ une sous-tribu.

Définition A.9. Soit $X : \Omega \to \mathbb{R}$ une variable aléatoire intégrable : $\mathbb{E}[|X|] < +\infty$. On note $\mathbb{E}[X|\mathcal{G}]$ l'espérance conditionnelle de X sachant $\mathcal{G}$. C'est l'unique variable aléatoire $Y = \mathbb{E}[X|\mathcal{G}]$ qui vérifie les deux points suivant.

- Y est $\mathcal{G}$-mesurable.
- Pour tout $A \in \mathcal{G}$,
$$\mathbb{E}[X \mathbb{1}_A] = \mathbb{E}[Y \mathbb{1}_A] \, .$$

Attention. Alors que $\mathbb{E}[X]$ est un réel, $\mathbb{E}[X|\mathcal{G}]$ est bien une variable aléatoire !

Remarque A.10 (Quelques notations)**.**
- Si $Z : \Omega \to \mathbb{R}$ est une autre variable aléatoire on note $\mathbb{E}[X|Z] = \mathbb{E}[X|\sigma(Z)]$.
- De même si $Z_1, \dots, Z_k$ sont des v.a. alors $\mathbb{E}[X|Z_1, \dots, Z_k] = \mathbb{E}[X|\sigma(Z_1, \dots, Z_k)]$.

Dans le cas où X est de carré intégrable, alors $\mathbb{E}[X|\mathcal{G}]$ peut s'interpréter comme la projection de X sur l'espace des v.a. mesurables par rapport à $\mathcal{G}$.

Proposition A.11. *Soit X une v.a. t.q. $\mathbb{E}[|X|^2] < +\infty$. Alors*

$$\mathbb{E}[X|\mathcal{G}] = \arg\min\{\mathbb{E}[|X - Y|^2] : Y \text{ est une v.a. } \mathcal{G}\text{-mesurable}\} \, .$$

A.4 Filtrations

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé. On dit que $(\mathcal{F}_k)_{k \in \mathbb{N}}$ est une filtration si pour tout $k \in \mathbb{N}$, $\mathcal{F}_k$ est une tribu sur Ω et de plus :

$$\forall m \leqslant l \, , \quad \mathcal{F}_m \subset \mathcal{F}_l \, .$$

L'idée derrière cette notion est de modéliser le flux d'information. Dans cette interprétation $\mathcal{F}_m$ est toute l'information (qu'on peut mesurer/observer) au temps m, qui est incluse, dans toute l'information au temps l.

On dit que (X_k), une suite de variables aléatoires de $(\Omega, \mathcal{F}, \mathbb{P})$ dans $(E, \mathcal{E})$, est adapté à une filtration $(\mathcal{F}_k)$ si pour tout k, X_k est mesurable par rapport à $\mathcal{F}_k$: $\sigma(X_k) \subset \mathcal{F}_k$.

La filtration naturelle est alors, $\mathcal{F}_k := \sigma(X_0, \ldots, X_k)$ (on notera que X_k est $\mathcal{F}_k$ mesurable par construction). C'est la plus petite tribu qui contient toutes les informations sur les (X_k) jusqu'à l'instant k.

Bibliographie

[1] Le Gall, J. Mouvement brownien, martingales et calcul stochastique. *Mathématiques Et Applications.* (2013) `http://dx.doi.org/10.1007/978-3-642-31898-6`

[2] Monique Jeanblanc. Cours de Calcul stochastique *Master 2IF EVRY.* Septembre 2006.

[3] Randal Douc. Eléments de calcul stochastique et applications à la finance *MAT 5501. Télécom SudParis.*

[4] Wojciech Pieczynski Processus Stochastiques *MAT 4514. Télécom SudParis.* Avril 2022.