CALCUL STOCHASTIQUE ET PROCESSUS DE MARKOV

Jean-François LE GALL

Notes de Cours de Master 2, 2010-2011

Université Paris-Sud
Master Probabilités et Statistiques

Septembre 2010

Sommaire

Chapitre 1

Vecteurs et Processus Gaussiens

1.1 Rappels sur les variables gaussiennes en dimension un

Une variable aléatoire X à valeurs dans $\mathbb{R}$ est dite gaussienne (ou normale) centrée réduite si elle admet pour densité

$$p_X(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2}).$$

La transformée de Laplace complexe de X est alors donnée par

$$E[e^{zX}] = e^{z^2/2}, \qquad \forall z \in \mathbb{C}.$$

Pour obtenir cette formule (et aussi voir que la transformée de Laplace complexe est bien définie) on traite d'abord le cas où $z = \lambda \in \mathbb{R}$:

$$E[e^{\lambda X}] = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{\lambda x} \, e^{-x^2/2} \, dx = e^{\lambda^2/2} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-(x-\lambda)^2/2} \, dx = e^{\lambda^2/2}.$$

Ce calcul assure que $E[e^{zX}]$ est bien défini pour tout $z \in \mathbb{C}$ et définit une fonction analytique sur $\mathbb{C}$. L'égalité $E[e^{zX}] = e^{z^2/2}$ étant vraie pour tout $z \in \mathbb{R}$ doit donc l'être aussi pour tout $z \in \mathbb{C}$.

En prenant $z = i\xi$, $\xi \in \mathbb{R}$, on trouve la transformée de Fourier de la loi de X:

$$E[e^{i\xi X}] = e^{-\xi^2/2}.$$

A partir du développement

$$E[e^{i\xi X}] = 1 + i\xi E[X] + \cdots + \frac{(i\xi)^n}{n!} E[X^n] + O(|\xi|^{n+1}),$$

valable quand X est dans tous les espaces L^p, $p < \infty$, ce qui est le cas ici, on calcule facilement

$$E[X] = 0, \quad E[X^2] = 1, \quad E[X^{2n}] = \frac{(2n)!}{2^n n!}.$$

Pour $\sigma > 0$ et $m \in \mathbb{R}$, on dit qu'une variable aléatoire réelle Y suit une loi gaussienne $\mathcal{N}(m, \sigma^2)$ si Y vérifie l'une des trois propriétés suivantes équivalentes:

(i) $Y = \sigma X + m$ où X suit une loi gaussienne centrée réduite (i.e. $\mathcal{N}(0,1)$);

(ii) la densité de Y est

$$p_Y(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp -\frac{(y-m)^2}{2\sigma^2};$$

(iii) la fonction caractéristique de Y est

$$E[e^{i\xi Y}] = \exp(im\xi - \frac{\sigma^2}{2}\,\xi^2).$$

On a alors

$$E[Y] = m, \quad \mathrm{var}(Y) = \sigma^2.$$

Par extension, on dit que Y suit une loi $\mathcal{N}(m, 0)$ si $Y = m$ p.s. (la propriété (iii) reste vraie).

Sommes de variables gaussiennes indépendantes. Supposons que Y suit la loi $\mathcal{N}(m, \sigma^2)$, Y' suit la loi $\mathcal{N}(m', \sigma'^2)$, et Y et Y' sont indépendantes. Alors $Y + Y'$ suit la loi $\mathcal{N}(m + m', \sigma^2 + \sigma'^2)$. C'est une conséquence immédiate de (iii).

Proposition 1.1 *Soit (X_n) une suite de variables aléatoires gaussiennes, telle que X_n soit de loi $\mathcal{N}(m_n, \sigma_n)$. Supposons que X_n converge en loi vers X. Alors*

1) La v.a. X est aussi une v.a. gaussienne $\mathcal{N}(m, \sigma^2)$ et $m = \lim m_n$, $\sigma = \lim \sigma_n$.

2) Si la suite (X_n) converge en probabilité vers X la convergence a lieu dans tous les espaces L^p, $p < \infty$.

Démonstration. 1) La convergence en loi est équivalente à dire que, pour tout $\xi \in \mathbb{R}$,

$$E[e^{i\xi X_n}] = \exp(im_n\xi - \frac{\sigma_n^2}{2}\,\xi^2) \longrightarrow E[e^{i\xi X}].$$

En passant au module, on a aussi

$$\exp(-\frac{\sigma_n^2}{2}\,\xi^2) \longrightarrow |E[e^{i\xi X}]|, \quad \forall \xi \in \mathbb{R},$$

ce qui ne peut se produire que si σ_n^2 converge vers $\sigma^2 \geq 0$ (le cas $\sigma_n^2 \to +\infty$ est à écarter car la fonction limite $\mathbf{1}_{\{\xi=0\}}$ ne serait pas continue). On a ensuite, pour tout $\xi \in \mathbb{R}$,

$$e^{im_n\xi} \longrightarrow e^{\frac{1}{2}\sigma^2\xi^2}\, E[e^{i\xi X}].$$

Montrons que cela entraîne la convergence de la suite (m_n). Si on sait a priori que la suite (m_n) est bornée, c'est facile: si m et m' sont deux valeurs d'adhérence on a $e^{im\xi} = e^{im'\xi}$ pour tout $\xi \in \mathbb{R}$, ce qui entraîne $m = m'$. Supposons la suite (m_n) non bornée et montrons qu'on arrive à une contradiction. On peut extraire une sous-suite (m_{n_k}) qui converge vers $+\infty$ (ou $-\infty$ mais le raisonnement est le même). Alors, pour tout $A > 0$,

$$P[X \geq A] \geq \limsup_{k\to\infty} P[X_{n_k} \geq A] \geq \frac{1}{2},$$

puisque $P[X_{n_k} \geq A] \geq P[X_{n_k} \geq m_{n_k}] \geq 1/2$ pour k assez grand. En faisant tendre A vers $+\infty$ on arrive à une contradiction.

On a donc $m_n \to m$, $\sigma_n \to \sigma$, d'où

$$E[e^{i\xi X}] = \exp(im\xi - \frac{\sigma^2}{2}\,\xi^2),$$

ce qui montre que X est gaussienne $\mathcal{N}(m, \sigma^2)$.

2) Puisque X_n a même loi que $\sigma_n N + m_n$, où N désigne une v.a. de loi $\mathcal{N}(0,1)$, et que les suites (m_n) et (σ_n) sont bornées (d'après 1)), on voit immédiatement que

$$\sup_n E[|X_n|^q] < \infty, \qquad \forall q > 0.$$

Le lemme de Fatou entraîne alors que $E[|X|^q] < \infty$, puis

$$\sup_n E[|X_n - X|^q] < \infty, \qquad \forall q > 0.$$

Soit $p \geq 1$. La suite $Y_n = |X_n - X|^p$ converge en probabilité vers 0 par hypothèse et est uniformément intégrable car bornée dans L^2 d'après ce qui précède (avec $q = 2p$). Elle converge donc dans L^1 vers 0 d'où le résultat recherché. $\square$

1.2 Vecteurs gaussiens

Soit E un espace euclidien de dimension d (E est isomorphe à $\mathbb{R}^d$ et on peut si on le souhaite prendre $E = \mathbb{R}^d$, mais il sera plus commode de travailler avec un espace abstrait). On note $\langle u, v\rangle$ le produit scalaire de E. Une variable aléatoire X à valeurs dans E est un vecteur gaussien si $\forall u \in E$, $\langle u, X\rangle$ est une variable gaussienne. (Par exemple, si $E = \mathbb{R}^d$ et si $X_1, \dots, X_d$ sont des variables gaussiennes indépendantes, la propriété des sommes de v.a. gaussiennes indépendantes montre que le vecteur $X = (X_1, \dots, X_d)$ est un vecteur gaussien.)

Si X est un vecteur gaussien à valeurs dans E il existe $m_X \in E$ et une forme quadratique positive q_X sur E tels que

$$\begin{aligned} E[\langle u, X\rangle] &= \langle u, m_X\rangle, \\ \mathrm{var}(\langle u, X\rangle) &= q_X(u), \end{aligned}$$

En effet, soit $(e_1, \dots, e_d)$ une base orthonormée de E et soit $X = \sum_{i=1}^d X_j\, e_j$ la décomposition de X dans cette base. Remarquons que les variables aléatoires $X_j = \langle e_j, X\rangle$ sont gaussiennes. On vérifie alors immédiatement les formules précédentes avec $m_X = \sum_{i=1}^d E[X_j]\, e_j \overset{\text{(not.)}}{=} E[X]$, et, si $u = \sum_{i=1}^d u_j e_j$,

$$q_X(u) = \sum_{j,k=1}^n u_j u_k \,\mathrm{cov}(X_j, X_k).$$

Comme $\langle u, X\rangle$ est une variable gaussienne $\mathcal{N}(\langle u, m_X\rangle, q_X(u))$, on en déduit la transformée de Fourier

$$E[\exp(i\langle u, X\rangle)] = \exp(i\langle u, m_X\rangle - \frac{1}{2} q_X(u)).$$

Proposition 1.2 *Sous les hypothèses précédentes, les variables aléatoires $X_1, \dots, X_d$ sont indépendantes si et seulement si la matrice de covariance $(\mathrm{cov}(X_j, X_k))$ est diagonale, soit si et seulement si la forme quadratique q_X est diagonale dans la base $(e_1, \dots, e_d)$.*

Démonstration. Il est évident que si les variables aléatoires $X_1, \dots, X_d$ sont indépendantes, la matrice de covariance $(\mathrm{cov}(X_j, X_k))_{j,k=1,\dots d}$ est diagonale. Inversement, si cette matrice est diagonale, on a

$$q_X(u) = \sum_{j=1}^d \lambda_j\, u_j^2, \qquad \text{où } \lambda_j = \mathrm{var}(X_j),$$

et donc

$$E[\exp i \sum_{j=1}^d u_j X_j] = \prod_{j=1}^d \exp(i u_j E[X_j] - \frac{1}{2}\lambda_j u_j^2) = \prod_{j=1}^d E[\exp(i u_j X_j)],$$

ce qui entraîne l'indépendance de $X_1, \dots, X_d$. $\square$

A la forme quadratique q_X on associe l'endomorphisme symétrique positif γ_X de E tel que

$$q_X(u) = \langle u, \gamma_X(u) \rangle$$

(γ_X a pour matrice $(\text{cov}(X_j, X_k))$ dans la base $(e_1, \dots, e_d)$ mais comme le montre la formule ci-dessus la définition de γ_X ne dépend pas de la base choisie).

A partir de maintenant, pour simplifier l'écriture, on se limite à des vecteurs gaussiens centrés, i.e. tels que $m_X = 0$.

Théorème 1.3 1) *Si γ est un endomorphisme symétrique positif de E il existe un vecteur gaussien centré X tel que $\gamma_X = \gamma$.*

2) *Soit X un vecteur gaussien centré. Soit $(\varepsilon_1, \dots, \varepsilon_d)$ une base de E qui diagonalise γ_X: $\gamma_X \varepsilon_j = \lambda_j \varepsilon_j$ avec $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_r > 0 = \lambda_{r+1} = \cdots = \lambda_d$ (r est le rang de γ_X). Alors,*

$$X = \sum_{j=1}^{r} Y_j \varepsilon_j,$$

où les v.a. Y_j sont des variables gaussiennes (centrées) indépendantes, et Y_j est de variance λ_j. En conséquence, si P_X désigne la loi de X, le support topologique de P_X est

$$\text{supp}\, P_X = \text{e.v.}(\varepsilon_1, \dots, \varepsilon_r).$$

La loi P_X est absolument continue par rapport à la mesure de Lebesgue sur E si et seulement si $r = d$, et dans ce cas la densité de X est

$$p_X(x) = \frac{1}{(2\pi)^{d/2}\sqrt{\det \gamma_X}} \exp -\frac{1}{2} \langle x, \gamma_X^{-1}(x) \rangle.$$

Démonstration. 1) Soit $(\varepsilon_1, \dots, \varepsilon_d)$ une base orthonormée de E dans laquelle γ est diagonale, $\gamma(\varepsilon_j) = \lambda_j \varepsilon_j$, et soient $Y_1, \dots, Y_d$ des v.a. gaussiennes centrées indépendantes avec $\text{var}(Y_j) = \lambda_j$. On pose alors

$$X = \sum_{j=1}^{d} Y_j \varepsilon_j,$$

et on vérifie alors facilement que, si $u = \sum u_j \varepsilon_j$,

$$q_X(u) = E[(\sum u_j Y_j)^2] = \sum \lambda_j u_j^2 = \langle u, \gamma u \rangle.$$

2) Si on écrit X dans la base $(\varepsilon_1, \dots, \varepsilon_d)$, la matrice de covariance des coordonnées $Y_1, \dots, Y_d$ est la matrice de γ_X dans cette base, donc une matrice diagonale avec $\lambda_1, \dots, \lambda_d$ sur la diagonale. Pour $j \in \{r+1, \dots, d\}$, on a $E[Y_j^2] = 0$ donc $Y_j = 0$ p.s. De plus, d'après la Proposition 1.2 les variables $Y_1, \dots, Y_r$ sont indépendantes.

Ensuite, puisque $X = \sum_{j=1}^r Y_j \varepsilon_j$ p.s. il est clair que

$$\operatorname{supp} P_X \subset \text{e.v.}\{\varepsilon_1, \ldots, \varepsilon_r\},$$

et inversement, si O est un pavé de la forme

$$O = \{u = \sum_{j=1}^r \alpha_j \varepsilon_j;\ a_j < \alpha_j < b_j\},$$

on a $P[X \in O] = \prod_{j=1}^r P[a_j < Y_j < b_j] > 0$. Cela suffit pour obtenir l'égalité $\operatorname{supp} P_X = \text{e.v.}\{\varepsilon_1, \ldots, \varepsilon_r\}$.

Si $r < d$, puisque e.v.$\{\varepsilon_1, \ldots, \varepsilon_r\}$ est de mesure nulle, la loi de X est étrangère par rapport à la mesure de Lebesgue sur E. Si $r = d$, notons Y le vecteur aléatoire de $\mathbb{R}^d$ défini par $Y = (Y_1, \ldots, Y_d)$ et remarquons que X est l'image de Y par la bijection $\varphi(y_1, \ldots, y_d) = \sum y_j \varepsilon_j$. Alors,

$$\begin{aligned}
E[g(X)] &= E[g(\varphi(Y))] \\
&= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} g(\varphi(y)) \exp\Big(-\frac{1}{2}\sum_{j=1}^d \frac{y_j^2}{\lambda_j}\Big) \frac{dy_1 \ldots dy_d}{\sqrt{\lambda_1 \ldots \lambda_d}} \\
&= \frac{1}{(2\pi)^{d/2}\sqrt{\det \gamma_X}} \int_{\mathbb{R}^d} g(\varphi(y)) \exp(-\frac{1}{2}\langle \varphi(y), \gamma_X^{-1}(\varphi(y))\rangle)\, dy_1 \ldots dy_d \\
&= \frac{1}{(2\pi)^{d/2}\sqrt{\det \gamma_X}} \int_E g(x) \exp(-\frac{1}{2}\langle x, \gamma_X^{-1}(x)\rangle)\, dx,
\end{aligned}$$

puisque la mesure de Lebesgue sur E est par définition l'image de la mesure de Lebesgue sur $\mathbb{R}^d$ par φ (ou par n'importe quelle autre isométrie de $\mathbb{R}^d$ sur E). Pour la deuxième égalité, on a utilisé le fait que $Y_1, \ldots, Y_d$ sont des variables gaussiennes $\mathcal{N}(0, \lambda_j)$ indépendantes, et pour la troisième l'égalité

$$< \varphi(y), \gamma_X^{-1}(\varphi(y))\rangle = \langle \sum y_j \varepsilon_j, \sum \frac{y_j}{\lambda_j} \varepsilon_j \rangle = \sum \frac{y_j^2}{\lambda_j}.$$

1.3 Espaces et processus gaussiens

Définition 1.1 *Un espace gaussien (centré) est un sous-espace fermé de $L^2(\Omega, \mathcal{A}, P)$ formé de variables gaussiennes centrées.*

Par exemple, si $X = (X_1, \ldots, X_d)$ est un vecteur gaussien centré dans $\mathbb{R}^d$, e.v.$\{X_1, \ldots, X_d\}$ est un espace gaussien.

Définition 1.2 *Une famille* $(\{X_t, t \in T\})$ *de v.a. réelles est un processus gaussien (centré) si toute combinaison linéaire finie des v.a.* X_t, $t \in T$ *est gaussienne centrée.*

Proposition 1.4 *Si* $X = (X_t, t \in T)$ *est un processus gaussien, le sous-espace vectoriel fermé de* L^2 *engendré par les v.a.* X_t, $t \in T$ *est un espace gaussien, appelé espace gaussien engendré par le processus* X.

Démonstration. Il suffit de remarquer qu'une limite dans L^2 de variables gaussiennes centrées est encore gaussienne centrée (cf Proposition 1.1). □

Si H est un sous-ensemble de L^2 on note $\sigma(H)$ la tribu engendrée par les v.a. $\xi \in H$.

Théorème 1.5 *Soit* H *un espace gaussien et soit* $(H_i, i \in I)$ *une famille de sous-espaces vectoriels de* H. *Alors les sous-espaces* H_i, $i \in I$ *sont orthogonaux dans* L^2 *si et seulement si les tribus* $\sigma(H_i), i \in I$ *sont indépendantes.*

Remarque. Il est crucial que les espaces H_i soient contenus dans un même espace gaussien. Considérons par exemple une variable X de loi $\mathcal{N}(0,1)$ et une seconde variable ε indépendante de X et telle que $P[\varepsilon = 1] = P[\varepsilon = -1] = 1/2$. Alors $X_1 = X$, $X_2 = \varepsilon X$ sont deux variables $\mathcal{N}(0,1)$. De plus, $E[X_1 X_2] = E[\varepsilon]E[X^2] = 0$. Cependant X_1 et X_2 ne sont évidemment pas indépendantes ($|X_1| = |X_2|$). Dans cet exemple, le couple (X_1, X_2) n'est pas un vecteur gaussien dans $\mathbb{R}^2$ bien que ses coordonnées soient des variables gaussiennes.

Démonstration. Si les tribus $\sigma(H_i)$ sont indépendantes, on a pour $i \neq j$, si $X \in H_i$ et $Y \in H_j$,

$$E[XY] = E[X]E[Y] = 0,$$

et donc les espaces H_i sont deux à deux orthogonaux.

Inversement, supposons les espaces H_i deux à deux orthogonaux. Par définition de l'indépendance d'une famille infinie de tribus, il suffit de montrer que pour $i_1, \ldots, i_p \in I$ distincts les tribus $\sigma(H_{i_1}), \ldots, \sigma(H_{i_p})$ sont indépendantes. Ensuite, il suffit de montrer que, si $\xi_1^1, \ldots, \xi_{n_1}^1 \in H_{i_1}, \ \ldots \ , \xi_1^p, \ldots, \xi_{n_p}^p \in H_{i_p}$, les vecteurs $(\xi_1^1, \ldots, \xi_{n_1}^1), \ldots, (\xi_1^p, \ldots, \xi_{n_p}^p)$ sont indépendants (en effet, pour chaque j les ensembles de la forme $\{\xi_1^j \in A_1, \ldots, \xi_{n_j}^j \in A_{n_j}\}$ forment une classe stable par intersection finie qui engendre la tribu $\sigma(H_j)$, et on peut ensuite utiliser un argument classique de classe monotone). Or, pour chaque $j \in \{1, \ldots, p\}$ on peut trouver $\eta_1^j, \ldots, \eta_{m_j}^j$ base orthonormée de e.v.$\{\xi_1^j, \ldots, \xi_{n_j}^j\}$. La matrice de covariance du vecteur

$$(\eta_1^1, \ldots, \eta_{m_1}^1, \eta_1^2, \ldots, \eta_{m_2}^2, \ldots, \eta_1^p, \ldots, \eta_{m_p}^p)$$

est donc la matrice identité (pour $i \neq j$, $E[\eta_l^i \eta_r^j] = 0$ à cause de l'orthogonalité de H_i et H_j). Ce vecteur est gaussien car ses composantes sont dans H. D'après la

Proposition 1.2, les composantes sont indépendantes. On conclut que les vecteurs $(\eta_1^1, \ldots, \eta_{m_1}^1), \ldots, (\eta_1^p, \ldots, \eta_{m_p}^p)$ sont indépendants. De manière équivalente les vecteurs $(\xi_1^1, \ldots, \xi_{n_1}^1), \ldots, (\xi_1^p, \ldots, \xi_{n_p}^p)$ sont indépendants, ce qui était le résultat recherché. $\square$

Corollaire 1.6 *Soient H un espace gaussien et K un sous-espace vectoriel fermé de K. On note p_K la projection orthogonale sur K. Soit $X \in H$.*

(i) *On a*

$$E[X \mid \sigma(K)] = p_K(X).$$

(ii) *Soit $\sigma^2 = E[(X - p_K(X))^2]$. Alors, pour tout borélien Γ de $\mathbb{R}$,*

$$P[X \in \Gamma \mid \sigma(K)] = Q(\omega, \Gamma),$$

où $Q(\omega, \cdot)$ est la loi $\mathcal{N}(p_K(X)(\omega), \sigma^2)$:

$$Q(\omega, \Gamma) = \frac{1}{\sigma\sqrt{2\pi}} \int_\Gamma dy \, \exp\Big(- \frac{(y - p_K(X))^2}{2\sigma^2}\Big)$$

($Q(\omega, \Gamma) = \mathbf{1}_\Gamma(p_K(X))$ si $\sigma = 0$).

Remarques. a) La partie (ii) de l'énoncé signifie que la loi conditionnelle de X sachant la tribu $\sigma(K)$ est la loi $\mathcal{N}(p_K(X), \sigma^2)$. D'une manière générale, la loi conditionnelle d'une variable aléatoire réelle X sachant une sous-tribu $\mathcal{G}$ est un noyau $Q(\omega, \Gamma)$ $\mathcal{G}$-mesurable, i.e. une application $Q : \Omega \times \mathcal{B}(\mathbb{R}) \to [0, 1]$ telle que

- pour tout ω, $\Gamma \to Q(\omega, \Gamma)$ est une mesure de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$,
- pour tout $\Gamma \in \mathcal{B}(\mathbb{R})$, $\omega \to Q(\omega, \Gamma)$ est $\mathcal{G}$-mesurable,

avec la propriété

$$P[X \in \Gamma \mid \mathcal{G}] = Q(\omega, \Gamma), \qquad \forall \Gamma \in \mathcal{B}(\mathbb{R})$$

(plus généralement $E[f(X) \mid \mathcal{G}] = \int f(y)\, Q(\omega, dy)$).

b) En général, pour une variable aléatoire X dans L^2, on a

$$E[X \mid \sigma(K)] = p_{L^2(\Omega, \sigma(K), P)}(X).$$

L'assertion (i) montre que dans notre cadre gaussien, cette projection orthogonale coïncide avec la projection orthogonale sur l'espace K, qui est bien plus petit que $L^2(\Omega, \sigma(K), P)$.

c) L'assertion (i) donne aussi le principe de la régression linéaire. Par exemple, si (X_1, X_2, X_3) est un vecteur gaussien, la meilleure approximation de X_3 connaisssant

X_1 et X_2 s'écrit $\lambda_1 X_1 + \lambda_2 X_2$ où λ_1 et λ_2 sont déterminés en disant que $X_3 - (\lambda_1 X_1 + \lambda_2 X_2)$ est orthogonal à e.v.(X_1, X_2).

Démonstration. (i) Soit $Y = X - p_K(X)$. Alors Y est orthogonal à K et d'après le Théorème 1.5, Y est indépendante de $\sigma(K)$. Ensuite,

$$E[X \mid \sigma(K)] = E[p_K(X) \mid \sigma(K)] + E[Y \mid \sigma(K)] = p_K(X) + E[Y] = p_K(X).$$

(ii) On écrit, pour toute fonction f mesurable positive sur $\mathbb{R}_+$,

$$E[f(X) \mid \sigma(K)] = E[f(p_K(X) + Y) \mid \sigma(K)] = \int P_Y(dy)\, f(p_K(X) + y),$$

où P_Y est la loi de Y qui est une loi $\mathcal{N}(0, \sigma^2)$ puisque Y est une variable gaussienne (centrée) de variance σ^2. (On a utilisé le fait général suivant: si Z est une variable $\mathcal{G}$-mesurable et si Y est indépendante de $\mathcal{G}$ alors $E[g(Y, Z) \mid \mathcal{G}] = \int g(y, Z)\, P_Y(dy)$.) Le résultat annoncé en (ii) découle aussitôt de la formule précédente. □

Soit $(X_t, t \in T)$ un processus gaussien. La fonction de covariance de X est la fonction $\Gamma : T \times T \to \mathbb{R}$ définie par $\Gamma(s,t) = \mathrm{cov}(X_s, X_t) = E[X_s X_t]$. La loi de X est déterminée par la fonction Γ. En effet, pour toute famille finie $\{t_1, \ldots, t_p\}$ de T, le vecteur $(X_{t_1}, \ldots, X_{t_p})$ est un vecteur gaussien centré de matrice de covariance $(\Gamma(t_i, t_j))$. (Dans le cas non centré, il faudrait en plus se donner la fonction moyenne $m(t) = E[X_t]$.)

On peut se demander si inversement, étant donné une fonction Γ sur $T \times T$, il existe un processus gaussien X dont Γ est la fonction de covariance. La fonction Γ doit être symétrique ($\Gamma(s,t) = \Gamma(t,s)$) et de type positif au sens suivant: si c est une fonction à support fini sur T, alors

$$\sum_{T \times T} c(s) c(t)\, \Gamma(s,t) = E[(\sum_T c(s) X_s)^2] \geq 0.$$

Remarquons que dans le cas où T est fini, le problème d'existence de X (qui est alors simplement un vecteur gaussien) est résolu sous les hypothèses précédentes sur Γ par le Théorème 1.3.

Théorème 1.7 *Soit Γ une fonction symétrique de type positif sur $T \times T$. Il existe alors un processus gaussien dont la fonction de covariance est Γ.*

Démonstration. C'est une application simple du théorème d'extension de Kolmogorov (cf Théorème 7.1 dans le Chapitre 7 ci-dessous). On construit une probabilité P sur l'espace mesurable

$$(\mathbb{R}^T, \mathcal{B}(\mathbb{R})^{\otimes T})$$

de telle sorte que sous P le processus des coordonnées $X_t(\omega) = \omega(t)$ soit un processus gaussien de fonction de covariance Γ. Si $\{t_1, \ldots, t_n\}$ est une partie finie de T, on construit d'abord une probabilité $P_{\{t_1,\ldots,t_n\}}$ sur $\mathbb{R}^{\{t_1,\ldots,t_n\}} \simeq \mathbb{R}^n$ comme la loi du vecteur gaussien de matrice de covariance $(\Gamma(t_i, t_j))$ (qui existe d'après le Théorème 1.3). Il est ensuite immédiat de vérifier que les probabilités $P_{\{t_1,\ldots,t_n\}}$ forment une famille compatible, et le théorème de Kolmogorov donne le résultat voulu.

Exemple. On considère le cas $T = \mathbb{R}$ et on se donne une mesure finie symétrique (i.e. $\mu(-A) = \mu(A)$) sur $\mathbb{R}$. On pose alors

$$\Gamma(s,t) = \int e^{i\xi(t-s)}\,\mu(d\xi).$$

On vérifie alors immédiatement que Γ a les propriétés requises: en particulier,

$$\sum_{T\times T} c(s)c(t)\,\Gamma(s,t) = \int |\sum_T c(s)e^{i\xi s}|^2\,\mu(ds) \geq 0.$$

La fonction Γ possède la propriété supplémentaire de dépendre seulement de la différence $t-s$. On en déduit aussitôt que le processus X associé à Γ par le théorème précédent est stationnaire (au sens strict), au sens où

$$(X_{t_1+t}, \ldots, X_{t_n+t}) \overset{(loi)}{=} (X_{t_1}, \ldots, X_{t_n})$$

pour tout choix de $t_1, \ldots, t_n, t \in \mathbb{R}$. Il est aussi vrai inversement que si $(X_t, t \in \mathbb{R})$ est un processus gaussien stationnaire, la fonction de covariance de X doit être du type précédent (théorème de Bochner).

1.4 Mesures gaussiennes

Définition 1.3 *Soit $(E, \mathcal{E})$ un espace mesurable, et soit μ une mesure σ-finie sur $(E, \mathcal{E})$. Une mesure gaussienne d'intensité μ est une isométrie G de $L^2(E, \mathcal{E}, \mu)$ sur un espace gaussien.*

Donc, si $f \in L^2(E, \mathcal{E}, \mu)$, $G(f)$ est une variable aléatoire gaussienne centrée de variance

$$E[G(f)^2] = \|G(f)\|^2_{L^2(\Omega)} = \|f\|^2_{L^2(\mu)}.$$

En particulier, si $f = \mathbf{1}_A$ avec $\mu(A) < \infty$, $G(\mathbf{1}_A) \overset{(not)}{=} G(A)$ suit la loi $\mathcal{N}(0, \mu(A))$.

Soient $A_1, \ldots, A_n \in E$ disjoints tels que $\mu(A_j) < \infty$. Alors le vecteur

$$(G(A_1), \ldots, G(A_n))$$

est un vecteur gaussien dans $\mathbb{R}^n$ de matrice de covariance diagonale, puisque pour $i \neq j$, la propriété d'isométrie donne

$$E[G(A_i)G(A_j)] = \langle \mathbf{1}_{A_i}, \mathbf{1}_{A_j} \rangle_{L^2(\mu)} = 0.$$

D'après la Proposition 1.2 on en déduit que les variables $G(A_1), \ldots, G(A_n)$ sont indépendantes.

Si A (toujours tel que $\mu(A) < \infty$) s'écrit comme réunion disjointe d'une famille dénombrable $A_1, \ldots, A_n, \ldots$ alors $\mathbf{1}_A = \sum_{j=1}^{\infty} \mathbf{1}_{A_j}$ avec une série convergeant dans $L^2(\mu)$, ce qui encore par la propriété d'isométrie, entraîne que

$$G(A) = \sum_{j=1}^{\infty} G(A_j)$$

avec convergence de la série dans $L^2(P)$ (grâce à l'indépendance des $G(A_j)$ on peut même montrer que la série converge p.s.).

Les propriétés de l'application $A \to G(A)$ "ressemblent" donc à celles d'une mesure (dépendant de ω). Cependant on peut montrer que en général, pour ω fixé, l'application $A \to G(A)(\omega)$ ne définit pas une mesure (nous reviendrons sur ce point plus loin).

Proposition 1.8 *Soient $(E, \mathcal{E})$ un espace mesurable et μ une mesure σ-finie sur $(E, \mathcal{E})$. Il existe alors une mesure gaussienne d'intensité μ.*

Démonstration. Soit $(f_i, i \in I)$ un système orthonormé total de $L^2(E, \mathcal{E}, \mu)$. Pour toute $f \in L^2(E, \mathcal{E}, \mu)$ on a

$$f = \sum_{i \in I} \alpha_i \, f_i$$

où les $\alpha_i = \langle f, f_i \rangle$ sont tels que

$$\sum_{i \in I} \alpha_i^2 = \|f\|^2 < \infty.$$

Sur un espace de probabilité, on se donne alors une famille $(X_i, i \in I)$ de v.a. $\mathcal{N}(0,1)$ indépendantes, et on pose

$$G(f) = \sum_{i \in I} \alpha_i \, X_i$$

(la série converge dans L^2 puisque les X_i forment un système orthonormé dans $L^2(P)$). Il est alors clair que G prend ses valeurs dans l'espace gaussien engendré par les $(X_i, i \in I)$. De plus il est immédiat que G est une isométrie. $\square$

On aurait pu aussi déduire le résultat précédent du Théorème 1.7 en prenant $T = L^2(E, \mathcal{E}, \mu)$ et $\Gamma(f, g) = \langle f, g \rangle_{L^2(\mu)}$. On construit ainsi un processus gaussien $(X_f, f \in L^2(\mu))$ et il suffit de prendre $G(f) = X_f$.

Remarque. Dans la suite du cours, on considérera uniquement le cas où $L^2(E, \mathcal{E}, \mu)$ est séparable. Par exemple, si $E = \mathbb{R}_+$ et μ est la mesure de Lebesgue, on peut pour construire G se donner une suite (ξ_n) de v.a. $\mathcal{N}(0,1)$ indépendantes, une base (φ_n) de $L^2(\mathbb{R}_+, dt)$ et définir G par

$$G(f) = \sum_n \langle f, \varphi_n \rangle \, \xi_n.$$

Proposition 1.9 *Soit G une mesure gaussienne sur $(E, \mathcal{E})$ d'intensité μ. Soit $A \in \mathcal{E}$ tel que $\mu(A) < \infty$. Supposons qu'il existe une suite de partitions de A,*

$$A = A_1^n \cup \ldots \cup A_{k_n}^n$$

de pas tendant vers 0, i.e.

$$\lim_{n \to \infty} \left(\sup_{1 \le j \le k_n} \mu(A_j^n) \right) = 0.$$

Alors,

$$\lim_{n \to \infty} \sum_{j=1}^{k_n} G(A_j^n)^2 = \mu(A)$$

dans L^2.

Démonstration. Pour n fixé les variables $G(A_1^n), \ldots, G(A_{k_n}^n)$ sont indépendantes. De plus, $E[G(A_j^n)^2] = \mu(A_j^n)$. On calcule alors facilement

$$E\left[\left(\sum_{j=1}^{k_n} G(A_j^n)^2 - \mu(A) \right)^2 \right] = \sum_{j=1}^{k_n} \mathrm{var}(G(A_j^n)^2) = 2 \sum_{j=1}^{k_n} \mu(A_j^n)^2,$$

car si X est une variable $\mathcal{N}(0, \sigma^2)$, $\mathrm{var}(X^2) = E(X^4) - \sigma^4 = 3\sigma^4 - \sigma^4 = 2\sigma^4$. Or,

$$\sum_{j=1}^{k_n} \mu(A_j^n)^2 \le \left(\sup_{1 \le j \le k_n} \mu(A_j^n) \right) \mu(A)$$

qui tend vers 0 quand $n \to \infty$ par hypothèse. $\square$

Chapitre 2

Le Mouvement Brownien

2.1 Le pré-mouvement brownien

Définition 2.1 *Soit G une mesure gaussienne sur $\mathbb{R}_+$ d'intensité la mesure de Lebesgue. Le processus $(B_t, t \in \mathbb{R}_+)$ défini par*

$$B_t = G(\mathbf{1}_{[0,t]})$$

est appelé pré-mouvement brownien.

Proposition 2.1 *Le processus $(B_t, t \geq 0)$ est un processus gaussien (centré) de covariance*

$$K(s,t) = \inf\{s,t\} \overset{\text{(not.)}}{=} s \wedge t.$$

Démonstration. Par définition d'une mesure gaussienne, les v.a. B_t appartiennent à un même espace gaussien, et $(B_t, t \geq 0)$ est donc un processus gaussien. De plus,

$$E[B_s B_t] = E[G([0,s])G([0,t])] = \int_0^\infty dr\, \mathbf{1}_{[0,s]}(r)\mathbf{1}_{[0,t]}(r) = s \wedge t.$$

Proposition 2.2 *Un processus $(X_t, t \in \mathbb{R}_+)$ est un pré-mouvement brownien si et seulement si*

(i) $X_0 = 0$ p.s.

(ii) Pour tous $0 \leq s < t$, la v.a. $X_t - X_s$ est indépendante de $\sigma(X_r, r \leq s)$ et suit la loi $\mathcal{N}(0, t-s)$.

Démonstration. Supposons d'abord que X est un pré-mouvement brownien, $X_t = G([0,t])$. Alors $X_0 = G(\{0\})$ suit une loi $\mathcal{N}(0,0)$ et donc $X_0 = 0$ p.s. Ensuite, fixons $s > 0$ et notons H_s l'espace vectoriel engendré par $(X_r, 0 \leq r \leq s)$, $\tilde{H}_s$ l'espace vectoriel engendré par $(X_{s+u}-X_s, u \geq 0)$. Alors H_s et $\tilde{H}_s$ sont orthogonaux puisque pour $r \in [0,s]$, $u \geq 0$

$$E[X_r(X_{s+u} - X_s)] = E[G([0,r])G(]s, s+u])] = 0$$

d'après les propriétés des mesures gaussiennes. Comme H_s et $\tilde{H}_s$ sont aussi contenus dans un même espace gaussien, on déduit du Théorème 1.5 que $\sigma(H_s)$ et $\sigma(\tilde{H}_s)$ sont indépendantes. En particulier $X_t - X_s$ est indépendante de $\sigma(H_s) = \sigma(X_r, r \leq s)$. Enfin, $X_t - X_s = G(]s,t])$ suit la loi $\mathcal{N}(0, t-s)$.

Inversement, supposons (i) et (ii) satisfaites. D'après (ii), pour tout choix de $0 = t_0 < t_1 < \cdots < t_n$ les v.a. $X_{t_i} - X_{t_{i-1}}$, $i \in \{1, \ldots, n\}$ sont indépendantes, et $X_{t_i} - X_{t_{i-1}}$ suit la loi $\mathcal{N}(0, t_i - t_{i-1})$. Il en découle facilement que X est un processus gaussien. Ensuite, si f est une fonction en escalier sur $\mathbb{R}_+$ de la forme $f = \sum_{i=1}^n \lambda_i \mathbf{1}_{]t_{i-1},t_i]}$, on pose

$$G(f) = \sum_{i=1}^{n} \lambda_i \, (X_{t_i} - X_{t_{i-1}})$$

(observer que cette définition ne dépend pas de l'écriture choisie pour f). On vérifie immédiatement que si g est une autre fonction en escalier du même type on a

$$E[G(f)G(g)] = \int_{\mathbb{R}_+} f(t)g(t)dt.$$

Grâce à la densité des fonctions en escalier dans $L^2(\mathbb{R}_+)$, on en déduit que l'application $f \to G(f)$ s'étend en une isométrie de $L^2(\mathbb{R}_+)$ dans l'espace gaussien engendré par X. Enfin $G([0,t]) = X_t - X_0 = X_t$ à cause de (i). $\square$

Remarque. La propriété (ii) (avec seulement le fait que la loi de $X_t - X_s$ ne dépend que de $t-s$) est souvent appelée propriété d'indépendance et de stationnarité des accroissements. Le pré-mouvement brownien est un cas particulier de la classe des processus à accroissements indépendants et stationnaires.

Corollaire 2.3 *Un processus $X = (X_t, t \in \mathbb{R}_+)$ est un pré-mouvement brownien si et seulement si $X_0 = 0$ p.s. et pour tout choix de $0 = t_0 < t_1 < \cdots < t_n$ les v.a. $X_{t_i} - X_{t_{i-1}}$, $i \in \{1, \ldots, n\}$ sont indépendantes, et $X_{t_i} - X_{t_{i-1}}$ suit la loi $\mathcal{N}(0, t_i - t_{i-1})$. En particulier, la loi du vecteur $(X_{t_1}, X_{t_2}, \ldots, X_{t_n})$ a pour densité*

$$\frac{1}{(2\pi)^{n/2}\sqrt{t_1(t_2-t_1)\ldots(t_n-t_{n-1})}} \exp\Big(-\sum_{i=1}^{n} \frac{(y_i - y_{i-1})^2}{2(t_i - t_{i-1})}\Big)$$

(par convention $y_0 = 0$).

Démonstration. La nécessité de la condition a déjà été obtenue dans la preuve de la Proposition 2.2. Inversement, sous la condition du corollaire, pour $0 = s_0 < s_1 < \ldots < s_n = s < t$, la v.a. $X_t - X_s$ est indépendante de $(X_{s_1} - X_{s_0}, \ldots, X_{s_n} - X_{s_{n-1}})$ donc aussi de $(X_{s_1}, \ldots, X_{s_n})$. Par un argument de classe monotone, $X_t - X_s$ est indépendant de $\sigma(X_r, r \leq s)$, et on est ramené à la Proposition 2.2.

Ensuite, les v.a. $X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$ étant indépendantes, le vecteur $(X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}})$ a pour densité

$$\frac{1}{(2\pi)^{n/2}\sqrt{t_1(t_2 - t_1)\ldots(t_n - t_{n-1})}} \exp\Big(-\sum_{i=1}^{n} \frac{x_i^2}{2(t_i - t_{i-1})}\Big),$$

et il suffit de faire le changement de variables $y_i = x_1 + \cdots + x_i$. □

La condition de la Proposition 2.1 (X processus gaussien centré de covariance $K(s,t) = s \wedge t$) est aussi nécessaire et suffisante pour que X soit un pré-mouvement brownien, puisque nous savons déjà que cette condition caractérise la loi de $(X_{t_1}, \ldots, X_{t_n})$ pour tout choix de $(t_1, \ldots, t_n)$, et on applique le corollaire 2.3.

Proposition 2.4 *Soit B un pré-mouvement brownien. Alors,*

(i) $-B$ est aussi un pré-mouvement brownien;

(ii) pour tout $\lambda > 0$, le processus $B_t^\lambda = \frac{1}{\lambda} B_{\lambda^2 t}$ est aussi un pré-mouvement brownien (invariance par changement d'échelle);

(iii) pour tout $s \geq 0$, le processus $B_t^{(s)} = B_{s+t} - B_s$ est un pré-mouvement brownien indépendant de $\sigma(B_r, r \leq s)$ (propriété de Markov simple).

Démonstration. (i) et (ii) sont très faciles. Démontrons (iii). Avec les notations de la preuve de la Proposition 2.2, la tribu engendrée par $B^{(s)}$ est $\sigma(\tilde{H}_s)$, qui est indépendante de $\sigma(H_s) = \sigma(B_r, r \leq s)$. Pour voir que $B^{(s)}$ est un pré-mouvement brownien, il suffit de vérifier la condition du Corollaire 2.3, ce qui est immédiat puisque $B_{t_i}^{(s)} - B_{t_{i-1}}^{(s)} = B_{s+t_i} - B_{s+t_{i-1}}$. □

Si B est un mouvement pré-brownien et G la mesure gaussienne associée, on note souvent pour $f \in L^2(\mathbb{R}_+, dt)$,

$$G(f) = \int_0^\infty f(s)\, dB_s$$

ou bien

$$G(f\mathbf{1}_{[0,t]}) = \int_0^t f(s)\, dB_s.$$

Cette notation est justifiée par le fait que si $u < v$,

$$\int_u^v dB_s = G(]u,v]) = B_v - B_u.$$

L'application $f \to \int_0^\infty f(s)\,dB_s$ (c'est-à-dire la mesure gaussienne G) est alors appelée intégrale de Wiener par rapport au pré-mouvement brownien B. Comme les mesures gaussiennes ne sont pas de "vraies" mesures, il ne s'agit pas d'une "vraie" intégrale dépendant de ω. Rappelons que $\int_0^\infty f(s)dB_s$ suit une loi $\mathcal{N}(0, \int_0^\infty f(s)^2 ds)$. Une partie importante de la suite de ce cours est consacrée à étendre la définition de $\int_0^\infty f(s)dB_s$ à des fonctions f qui peuvent dépendre de ω.

2.2 La continuité des trajectoires

Soit $B = (B_t, t \geq 0)$ un pré-mouvement brownien. Les applications $t \to B_t(\omega)$ pour $\omega \in \Omega$ sont appelées les trajectoires de B. Au stade où nous en sommes, on ne peut rien affirmer au sujet de ces trajectoires : il n'est même pas évident (ni vrai en général) que ces applications soient mesurables. Le but de ce paragraphe est de montrer que, quitte à modifier "un peu" B, on peut faire en sorte que les trajectoires soient continues.

Définition 2.2 *Soient $(X_t, t \in T)$, $(\tilde{X}_t, t \in T)$ deux processus aléatoires (i.e. familles de variables aléatoires) indexés par le même ensemble T. On dit que $\tilde{X}$ est une modification de X si*

$$\forall t \in T, \qquad P[\tilde{X}_t = X_t] = 1.$$

Remarquons que le processus $\tilde{X}$ a alors même loi que X au sens où pour tout choix de $t_1, \ldots, t_n$, le vecteur $(\tilde{X}_{t_1}, \ldots, \tilde{X}_{t_n})$ a même loi que $(X_{t_1}, \ldots, X_{t_n})$. En particulier, si X est un pré–mouvement brownien, $\tilde{X}$ est aussi un pré-mouvement brownien. En revanche, les trajectoires de $\tilde{X}$ peuvent avoir un comportement très différent de celles de X. Il peut arriver par exemple que les trajectoires de $\tilde{X}$ soient toutes continues alors que celles de X sont toutes discontinues.

Définition 2.3 *Les deux processus X et $\tilde{X}$ sont dits indistinguables si*

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

(On admet implicitement que l'événement $\{\forall t \in T,\ X_t = \tilde{X}_t\}$ est mesurable.)

Si deux processus sont indistinguables, l'un est une modification de l'autre. La notion d'indistinguabilité est cependant (beaucoup) plus forte : deux processus indistinguables ont p.s. les mêmes trajectoires.

Si $T = I$ est un intervalle de $\mathbb{R}$, et si X et $\tilde{X}$ sont deux processus dont les trajectoires sont p.s. continues, alors $\tilde{X}$ est une modification de X si et seulement si X et $\tilde{X}$ sont indistinguables. En effet, si $\tilde{X}$ est une modification de X on a p.s. $\forall t \in I \cap \mathbb{Q}$, $X_t = \tilde{X}_t$ (on écarte une réunion *dénombrable* d'ensembles de probabilité nulle) d'où p.s. $\forall t \in I$, $X_t = \tilde{X}_t$ par continuité.

Théorème 2.5 (lemme de Kolmogorov) *Soit* $X = (X_t, t \in I)$ *un processus aléatoire indexé par un intervalle borné* I *de* $\mathbb{R}$, *à valeurs dans un espace métrique complet* (E, d). *Supposons qu'il existe trois réels* $q, \varepsilon, C > 0$ *tels que, pour tous* $s, t \in I$,

$$E[d(X_s, X_t)^q] \leq C\,|t-s|^{1+\varepsilon}.$$

Alors, il existe une modification $\tilde{X}$ *de* X *dont les trajectoires sont höldériennes d'exposant* α *pour tout* $\alpha \in]0, \frac{\varepsilon}{q}[$: *pour tout* $\alpha \in]0, \frac{\varepsilon}{q}[$ *il existe une constante* $C_\alpha(\omega)$ *telle que, pour tous* $s, t \in I$,

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

En particulier, $\tilde{X}$ *est une modification continue de* X *(unique à indistinguabilité près d'après ci-dessus).*

Remarques. (i) Si I est non borné, par exemple si $I = \mathbb{R}_+$, on peut appliquer le Théorème 2.5 à $I = [0,1], [1,2], [2,3]$,etc. et on trouve encore que X a une modification continue, qui est localement höldérienne d'exposant α pour tout $\alpha \in]0, \varepsilon/q[$.

(ii) Il suffit de montrer que pour $\alpha \in]0, \varepsilon/q[$ fixé, X a une modification dont les trajectoires sont höldériennes d'exposant α. En effet, on applique ce résultat à une suite $\alpha_k \uparrow \varepsilon/q$ en observant que les processus obtenus sont alors tous indistinguables, d'après la remarque précédant le théorème.

Démonstration. Pour simplifier l'écriture, on prend $I = [0,1]$. On note D l'ensemble (dénombrable) des nombres dyadiques de l'intervalle $[0,1[$, c'est-à-dire des réels $t \in [0,1]$ qui s'écrivent

$$t = \sum_{k=1}^{p} \varepsilon_k\, 2^{-k}$$

avec $\varepsilon_k = 0$ ou 1. L'étape-clé de la preuve du théorème est le lemme suivant.

Lemme 2.6 *Pour tout* $\alpha \in]0, \varepsilon/q[$ *il existe p.s. une constante* $C_\alpha(\omega)$ *telle que, pour tous* $s, t \in D$,

$$d(X_s, X_t) \leq C_\alpha(\omega)\,|t-s|^\alpha.$$

Démonstration. L'hypothèse du théorème entraîne que, pour $a > 0$ et $s, t \in I$

$$P[d(X_s, X_t) \geq a] \leq a^{-q} E[d(X_s, X_t)^q] \leq C\,a^{-q}\,|t-s|^{1+\varepsilon}.$$

Appliquons cette inégalité à $s = (i-1)2^{-n}$, $t = i2^{-n}$ (pour $i \in \{1, \ldots, 2^n\}$) et $a = 2^{-n\alpha}$:

$$P[d(X_{(i-1)2^{-n}}, X_{i2^{-n}}) \geq 2^{-n\alpha}] \leq C\,2^{nq\alpha} 2^{-(1+\varepsilon)n}.$$

En sommant sur i on trouve

$$P[\bigcup_{i=1}^{2^n}\{d(X_{(i-1)2^{-n}},X_{i2^{-n}})\geq 2^{-n\alpha}\}]\leq 2^n\cdot C\,2^{nq\alpha-(1+\varepsilon)n}=C\,2^{-n(\varepsilon-q\alpha)}.$$

Par hypothèse, $\varepsilon-q\alpha>0$. En sommant maintenant sur n on a donc

$$\sum_{n=1}^{\infty}P[\bigcup_{i=1}^{2^n}\{d(X_{(i-1)2^{-n}},X_{i2^{-n}})\geq 2^{-n\alpha}\}]<\infty,$$

et le lemme de Borel-Cantelli montre que

$$\text{p.s. }\exists n_0(\omega):\ \forall n\geq n_0(\omega),\ \forall i\in\{1,\dots,2^n\},\quad d(X_{(i-1)2^{-n}},X_{i2^{-n}})\leq 2^{-n\alpha}.$$

En conséquence, la constante $K_\alpha(\omega)$ définie par

$$K_\alpha(\omega)=\sup_{n\geq 1}\left(\sup_{1\leq i\leq 2^n}\frac{d(X_{(i-1)2^{-n}},X_{i2^{-n}})}{2^{-n\alpha}}\right)$$

est finie p.s. (Pour $n\geq n_0(\omega)$, le terme entre parenthèses est majoré par 1, et d'autre part, il n'y a qu'un nombre fini de termes avant $n_0(\omega)$.)

Nous affirmons alors que le résultat du lemme est vrai avec $C_\alpha(\omega)=2(1-2^{-\alpha})^{-1}K_\alpha(\omega)$. Pour le voir, considérons $s,t\in D$ avec $s<t$. Soit $p\geq 1$ le plus petit entier tel que $2^{-p}<t-s$. Alors il est facile de voir qu'on peut trouver un entier $k\geq 0$ et deux entiers $l,m\geq 0$ tels que

$$\begin{aligned}s&=k2^{-p}-\varepsilon_{p+1}2^{-p-1}-\ldots-\varepsilon_{p+l}2^{-p-l}\\ t&=k2^{-p}+\varepsilon'_p2^{-p}+\varepsilon'_{p+1}2^{-p-1}+\ldots+\varepsilon'_{p+m}2^{-p-m},\end{aligned}$$

où $\varepsilon_i,\varepsilon'_j=0$ ou 1. Notons

$$\begin{aligned}s_i&=k2^{-p}-\varepsilon_{p+1}2^{-p-1}-\ldots-\varepsilon_{p+i}2^{-p-i}\qquad(\text{pour }0\leq i\leq l)\\ t_j&=k2^{-p}+\varepsilon'_p2^{-p}+\varepsilon'_{p+1}2^{-p-1}+\ldots+\varepsilon'_{p+j}2^{-p-j}\qquad(\text{pour }0\leq j\leq m).\end{aligned}$$

Alors,

$$\begin{aligned}d(X_s,X_t)&=d(X_{s_l},X_{t_m})\\ &\leq d(X_{s_0},X_{t_0})+\sum_{i=1}^{l}d(X_{s_{i-1}},X_{s_i})+\sum_{j=1}^{m}d(X_{t_{j-1}},X_{t_j})\\ &\leq K_\alpha(\omega)2^{-p\alpha}+\sum_{i=1}^{l}K_\alpha(\omega)2^{-(p+i)\alpha}+\sum_{j=1}^{m}K_\alpha(\omega)2^{-(p+j)\alpha}\\ &\leq K_\alpha(\omega)\,2(1-2^{-\alpha})^{-1}\,2^{-p\alpha}\\ &\leq C_\alpha(\omega)\,(t-s)^\alpha,\end{aligned}$$

d'où le résultat voulu. $\square$

Nous pouvons maintenant terminer la preuve du théorème. Le lemme montre que p.s. la fonction $t \to X_t(\omega)$ est höldérienne sur D, donc uniformément continue sur D. Puisque (E, d) est complet, cette fonction a p.s. un unique prolongement continu à $I = [0, 1]$, et ce prolongement est lui aussi höldérien d'exposant α. De manière plus précise, on pose pour tout $t \in [0, 1]$

$$\tilde{X}_t(\omega) = \lim_{s \to t,\, s \in D} X_s(\omega)$$

sur l'ensemble $\{K_\alpha < \infty\}$ des valeurs de ω pour lesquelles la fonction $s \to X_s(\omega)$ est höldérienne d'exposant α sur D, et $\tilde{X}_t(\omega) = x_0$ sur l'ensemble $\{K_\alpha = \infty\}$ qui est de probabilité nulle. (Ici x_0 est un point fixé de E.)

D'après les remarques précédentes, le processus $\tilde{X}$ a des trajectoires höldériennes d'exposant α sur $[0, 1]$. Il reste à voir que $\tilde{X}$ est une modification de X. Or l'hypothèse du théorème entraîne aussitôt que pour tout $t \in I$,

$$\lim_{s \to t} X_s = X_t$$

au sens de la convergence en probabilité. Comme par définition $\tilde{X}_t$ est aussi la limite p.s. de X_s quand $s \to t$, $s \in D$, on conclut que $X_t = \tilde{X}_t$ p.s. $\square$

Corollaire 2.7 *Soit $B = (B_t, t \geq 0)$ un pré-mouvement brownien. Le processus B a une modification dont les trajectoires sont continues, et même localement höldériennes d'exposant $\frac{1}{2} - \delta$ pour tout $\delta \in]0, \frac{1}{2}[$.*

Démonstration. Pour $s < t$, la v.a. $B_t - B_s$ suit une loi $\mathcal{N}(0, t - s)$, et donc $B_t - B_s$ a même loi que $\sqrt{t - s}N$, où N suit une loi $\mathcal{N}(0, 1)$. Donc, pour tout $q > 0$,

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

avec $C_q = E[|N|^q] < \infty$. Dès que $q > 2$, on peut appliquer le théorème avec $\varepsilon = \frac{q}{2} - 1$. On trouve ainsi que B a une modification dont les trajectoires sont continues, et même localement höldériennes d'exposant α pour tout $\alpha < (q-2)/(2q)$. En choisissant q grand on trouve le résultat souhaité. $\square$

Définition 2.4 *Un processus $(B_t, t \geq 0)$ est un mouvement brownien (réel, issu de* 0*) si :*

(i) $(B_t, t \geq 0)$ *est un pré-mouvement brownien.*

(ii) *Les trajectoires de B, c'est-à-dire les applications $t \longrightarrow B_t(\omega)$ pour $\omega \in \Omega$, sont toutes continues.*

L'existence du mouvement brownien découle du corollaire précédent. En effet la modification obtenue dans ce corollaire est encore un pré-mouvement brownien, et ses trajectoires sont continues. Dans la suite on ne parlera plus de pré-mouvement brownien et on s'intéressera uniquement au mouvement brownien.

Il est important de remarquer que l'énoncé de la Proposition 2.4 reste vrai mot pour mot si on remplace partout pré-mouvement brownien par mouvement brownien. En effet, avec les notations de cette proposition, on vérifie immédiatement que les processus $-B, B^\lambda, B^{(s)}$ ont des trajectoires continues si c'est le cas pour B.

Mesure de Wiener. Notons $C(\mathbb{R}_+, \mathbb{R})$ l'espace des fonctions continues de $\mathbb{R}_+$ dans $\mathbb{R}$. La donnée d'un mouvement brownien B fournit donc une application

$$\begin{aligned} \Omega &\longrightarrow C(\mathbb{R}_+, \mathbb{R}) \\ \omega &\longrightarrow (t \longrightarrow B_t(\omega)) \end{aligned}$$

qui est mesurable lorsque $C(\mathbb{R}_+, \mathbb{R})$ est muni de la plus petite tribu rendant mesurables les applications coordonnées $\mathrm{w} \to \mathrm{w}(t)$. Il est facile de voir que cette tribu notée $\mathcal{C}$ coïncide avec la tribu borélienne pour la topologie de la convergence uniforme sur tout compact). La **mesure de Wiener** (loi du mouvement brownien) est par définition la mesure-image de $P(d\omega)$ par cette application. Si $W(d\mathrm{w})$ désigne cette mesure-image, le Corollaire 2.3 montre que, pour $0 = t_0 < t_1 < \cdots < t_n$, et $A_0, A_1, \ldots, A_n \in \mathcal{B}(\mathbb{R})$,

$$\begin{aligned} &W(\{\mathrm{w}; \mathrm{w}(t_0) \in A_0, \mathrm{w}(t_1) \in A_1, \ldots, \mathrm{w}(t_n) \in A_n\}) \\ &\quad = \mathbf{1}_{A_0}(0) \int_{A_1 \times \cdots \times A_n} \frac{dy_1 \ldots dy_n}{(2\pi)^{n/2} \sqrt{t_1(t_2 - t_1) \ldots (t_n - t_{n-1})}} \exp\Big(-\sum_{i=1}^n \frac{(y_i - y_{i-1})^2}{2(t_i - t_{i-1})} \Big). \end{aligned}$$

(avce la convention y_0=0). De plus ces propriétés caractérisent la probabilité W : en effet, la classe des ensembles de la forme $\{\mathrm{w}; \mathrm{w}(t_0) \in A_0, \mathrm{w}(t_1) \in A_1, \ldots, \mathrm{w}(t_n) \in A_n\}$ (les "cylindres") est stable par intersection finie et engendre la tribu $\mathcal{C}$, ce qui, par un argument standard de classe monotone, suffit pour dire qu'une mesure de probabilité sur $\mathcal{C}$ est caractérisée par ses valeurs sur cette classe. En conséquence, la mesure de Wiener ne dépend pas du choix du mouvement brownien B pour la construction. Si B' est un autre mouvement brownien, on a pour tout $A \in \mathcal{C}$,

$$P((B_t)_{t \geq 0} \in A) = W(A) = P((B'_t)_{t \geq 0} \in A).$$

Nous utiliserons fréquemment cette propriété sans plus de commentaires. Voir par exemple la deuxième partie de la preuve du Corollaire 2.9 ci-dessous.

Si l'on prend maintenant comme espace de probabilité

$$\Omega = C(\mathbb{R}_+, \mathbb{R}), \quad \mathcal{F} = \mathcal{C}, \quad P(d\mathrm{w}) = W(d\mathrm{w}),$$

le processus, dit *canonique*,

$$X_t(\mathrm{w}) = \mathrm{w}(t)$$

est un mouvement brownien (d'après le Corollaire 2.3 et les formules ci-dessus). C'est la construction canonique du mouvement brownien.

2.3 Comportement des trajectoires du mouvement brownien

Dans ce paragraphe, nous obtenons quelques informations sur l'allure des trajectoires du mouvement brownien B. Un ingrédient très utile est le résultat suivant, connu sous le nom de loi du tout ou rien.

Théorème 2.8 *Pour tout $t \geq 0$ soit $\mathcal{F}_t$ la tribu définie par*

$$\mathcal{F}_t = \sigma(B_s, s \leq t),$$

et soit

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

La tribu $\mathcal{F}_{0+}$ est grossière, au sens où $\forall A \in \mathcal{F}_{0+}$, $P(A) = 0$ ou 1.

Démonstration. Soient $0 < t_1 < t_2 < \cdots < t_k$ et soit $g : \mathbb{R}^k \longrightarrow \mathbb{R}$ une fonction continue bornée. Soit aussi $A \in \mathcal{F}_{0+}$. Alors, par un argument de continuité,

$$E[\mathbf{1}_A\, g(B_{t_1}, \ldots, B_{t_k})] = \lim_{\varepsilon \to 0} E[\mathbf{1}_A\, g(B_{t_1} - B_\varepsilon, \ldots, B_{t_k} - B_\varepsilon)].$$

Mais, dès que $\varepsilon < t_1$, les v.a. $B_{t_1} - B_\varepsilon, \ldots, B_{t_k} - B_\varepsilon$ sont indépendantes de $\mathcal{F}_\varepsilon$ (par la propriété de Markov simple) et donc aussi de la tribu $\mathcal{F}_{0+}$. Il en découle que

$$\begin{aligned} E[\mathbf{1}_A\, g(B_{t_1}, \ldots, B_{t_k})] &= \lim_{\varepsilon \to 0} P(A)\, E[g(B_{t_1} - B_\varepsilon, \ldots, B_{t_k} - B_\varepsilon)] \\ &= P(A)\, E[g(B_{t_1}, \ldots, B_{t_k})]. \end{aligned}$$

Ainsi on trouve que $\mathcal{F}_{0+}$ est indépendante de $\sigma(B_{t_1}, \ldots, B_{t_k})$. Comme cela est vrai pour toute famille finie $\{t_1, \ldots, t_k\}$ de réels strictement positifs, $\mathcal{F}_{0+}$ est indépendante de $\sigma(B_t, t > 0)$. Finalement $\sigma(B_t, t > 0) = \sigma(B_t, t \geq 0)$ puisque B_0 est la limite simple de B_t quand $t \to 0$. Comme $\mathcal{F}_{0+} \subset \sigma(B_t, t \geq 0)$, on voit que $\mathcal{F}_{0+}$ est indépendante d'elle-même, ce qui entraîne que $\mathcal{F}_{0+}$ est grossière. $\square$

Corollaire 2.9 *On a p.s. pour tout* $\varepsilon > 0$

$$\sup_{0\le s\le \varepsilon} B_s > 0, \qquad \inf_{0\le s\le \varepsilon} B_s < 0.$$

Pour tout $a \in \mathbb{R}$, *soit* $T_a = \inf\{t, B_t = a\}$ ($\inf\emptyset = \infty$). *Alors,*

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

En conséquence, p.s.,

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

Remarque. Il n'est pas a priori évident que la variable $\sup_{0\le s\le\varepsilon} B_s$ soit mesurable: il s'agit d'un supremum non dénombrable de fonctions mesurables. Cependant, parce que nous savons que les trajectoires de B sont continues, on peut se restreindre aux valeurs **rationnelles** de $s \in [0,\varepsilon]$ et on obtient alors un supremum dénombrable de variables aléatoires. Nous utiliserons ce type de remarque implicitement dans la suite.

Démonstration. Soit (ε_p) une suite de réels strictement positifs décroissant vers 0, et soit

$$A = \bigcap_p \{\sup_{0\le s\le \varepsilon_p} B_s > 0\}.$$

Il est clair que l'événement A est $\mathcal{F}_{0+}$-mesurable. D'autre part,

$$P(A) = \lim_{p\to\infty} \downarrow P[\sup_{0\le s\le\varepsilon_p} B_s > 0],$$

et

$$P[\sup_{0\le s\le \varepsilon_p} B_s > 0] \ge P[B_{\varepsilon_p} > 0) = \frac{1}{2},$$

ce qui montre que $P[A] \ge 1/2$. D'après le Théorème 2.8 on a $P[A] = 1$, d'où

$$\text{p.s. } \forall \varepsilon > 0, \quad \sup_{0\le s\le\varepsilon} B_s > 0.$$

L'assertion concernant $\inf_{0\le s\le\varepsilon} B_s$ est obtenue en remplaçant B par $-B$.

Ensuite, on écrit

$$1 = P[\sup_{0\le s\le 1} B_s > 0] = \lim_{\delta\downarrow 0} \uparrow P[\sup_{0\le s\le 1} B_s > \delta],$$

et on remarque en appliquant la propriété d'invariance d'échelle (Proposition 2.4 (ii)) avec $\lambda = 1/\delta$ que

$$P[\sup_{0\le s\le 1} B_s > \delta] = P[\sup_{0\le s\le 1/\delta^2} B_s^{1/\delta} > 1] = P[\sup_{0\le s\le 1/\delta^2} B_s > 1].$$

En faisant tendre δ vers 0, on trouve

$$P[\sup_{s\geq 0} B_s > 1] = 1.$$

A nouveau un argument de changement d'échelle montre que pour tout $A > 0$,

$$P[\sup_{s\geq 0} B_s > A] = 1$$

et en utilisant le changement $B \rightarrow -B$ on a aussi

$$P[\inf_{s\geq 0} B_s < -A] = 1.$$

Les dernières assertions du corollaire en découlent facilement: pour la dernière, on observe qu'une fonction continue $f : \mathbb{R}_+ \longrightarrow \mathbb{R}$ ne peut visiter tous les réels que si $\limsup_{t\to+\infty} f(t) = +\infty$, $\liminf_{t\to+\infty} f(t) = -\infty$. □

En utilisant la propriété de Markov simple, on déduit facilement du corollaire que p.s. la fonction $t \rightarrow B_t$ n'est monotone sur aucun intervalle non-trivial.

Proposition 2.10 *Soit* $0 = t_0^n < t_1^n < \cdots < t_{p_n}^n = t$ *une suite de subdivisions de* $[0, t]$ *de pas tendant vers* 0 *(i.e.* $\sup_i (t_i^n - t_{i-1}^n) \rightarrow 0$*). Alors,*

$$\lim_{n\to\infty} \sum_{i=1}^{p_n} (B_{t_i^n} - B_{t_{i-1}^n})^2 = t,$$

dans L^2*.*

Démonstration. C'est une conséquence presque immédiate de la Proposition 1.9, en écrivant $B_{t_i^n} - B_{t_{i-1}^n} = G(]t_{i-1}^n, t_i^n])$. □

On déduit facilement de la Proposition 2.10 et de la continuité des trajectoires que p.s. la fonction $t \rightarrow B_t$ n'est à variation finie sur aucun intervalle non trivial.

2.4 La propriété de Markov forte

Notre but est d'étendre la propriété de Markov simple (Proposition 2.4 (iii)) au cas où l'instant déterministe s est remplacé par un temps aléatoire T. Nous devons d'abord préciser la classe des temps aléatoires admissibles. On garde la notation $\mathcal{F}_t$ introduite dans le Théorème 2.8 et on note aussi $\mathcal{F}_\infty = \sigma(B_s, s \geq 0) \vee \sigma(\mathcal{N})$.

Définition 2.5 *Une variable aléatoire* T *à valeurs dans* $[0, \infty]$ *est un temps d'arrêt si* $\forall t \geq 0$, $\{T \leq t\} \in \mathcal{F}_t$.

Exemples. Les temps $T = t$ (temps constant) ou $T = T_a$ sont des temps d'arrêt (pour le deuxième cas remarquer que $\{T_a \le t\} = \{\sup_{0\le s\le t} B_s \ge a\}$ pour $a \ge 0$). En revanche, $T = \sup\{s \le 1, B_s = 0\}$ n'est pas un temps d'arrêt (cela découlera par l'absurde de la propriété de Markov forte ci-dessous et de la Proposition 2.9).

Définition 2.6 *Soit T un temps d'arrêt. La tribu des événements antérieurs à T est*

$$\mathcal{F}_T = \{A \in \mathcal{F}_\infty;\ \forall t \ge 0,\ A \cap \{T \le t\} \in \mathcal{F}_t\}.$$

On vérifie facilement que les variables aléatoires T et B_T sont $\mathcal{F}_T$-mesurables (pour B_T, remarquer que p.s.

$$B_T = \lim_{n\to\infty} \sum_{i=0}^{\infty} \mathbf{1}_{\{i2^{-n}<T\le(i+1)2^{-n}\}} B_{i2^{-n}},$$

puis que $B_s \mathbf{1}_{\{s<T\}}$ est $\mathcal{F}_T$ mesurable).

Théorème 2.11 (Propriété de Markov forte) *Soit T un temps d'arrêt. Alors, conditionnellement à $\{T < \infty\}$, le processus $B^{(T)}$ défini par*

$$B^{(T)}_t = B_{T+t} - B_T$$

est un mouvement brownien indépendant de $\mathcal{F}_T$.

Démonstration. Supposons d'abord $T < \infty$ p.s. On va montrer que, pour $A \in \mathcal{F}_T$, $0 \le t_1 < \cdots < t_p$ et F continue bornée sur $\mathbb{R}^p$ on a

$$E[\mathbf{1}_A\, F(B^{(T)}_{t_1}, \ldots, B^{(T)}_{t_p})] = P[A]\, E[F(B_{t_1}, \ldots, B_{t_p})]. \tag{2.1}$$

Cela suffit pour établir les différentes assertions du théorème : le cas $A = \Omega$ montre que $B^{(T)}$ est un mouvement brownien (remarquer que les trajectoires de $B^{(T)}$ sont continues) et d'autre part (2.1) entraîne que pour tout choix de $0 \le t_1 < \cdots < t_p$, le vecteur $(B^{(T)}_{t_1}, \ldots, B^{(T)}_{t_p})$ est indépendant de $\mathcal{F}_T$, d'où il découle par classe monotone que $B^{(T)}$ est indépendant de $\mathcal{F}_T$.

Pour montrer (2.1), on observe d'abord que p.s.

$$\begin{aligned}&F(B^{(T)}_{t_1}, \ldots, B^{(T)}_{t_p})\\&= \lim_{n\to\infty} \sum_{k=0}^{\infty} \mathbf{1}_{\{(k-1)2^{-n}<T\le k2^{-n}\}} F(B_{k2^{-n}+t_1} - B_{k2^{-n}}, \ldots, B_{k2^{-n}+t_p} - B_{k2^{-n}}),\end{aligned}$$

d'où par convergence dominée,

$$\begin{aligned}&E[\mathbf{1}_A\, F(B^{(T)}_{t_1}, \ldots, B^{(T)}_{t_p})]\\&= \lim_{n\to\infty} \sum_{k=0}^{\infty} E[\mathbf{1}_A \mathbf{1}_{\{(k-1)2^{-n}<T\le k2^{-n}\}} F(B_{k2^{-n}+t_1} - B_{k2^{-n}}, \ldots, B_{k2^{-n}+t_p} - B_{k2^{-n}})].\end{aligned}$$

Pour $A \in \mathcal{F}_T$, l'événement $A \cap \{(k-1)2^{-n} < T \leq k2^{-n}\}$ est $\mathcal{F}_{k2^{-n}}$-mesurable, donc coïncide p.s. avec un événement de $\sigma(B_r, r \leq k2^{-n})$. D'après la propriété de Markov simple (Proposition 2.4 (iii)), on a donc

$$\begin{aligned}
&E[\mathbf{1}_{A\cap\{(k-1)2^{-n}<T\leq k2^{-n}\}} F(B_{k2^{-n}+t_1} - B_{k2^{-n}}, \ldots, B_{k2^{-n}+t_p} - B_{k2^{-n}})] \\
&= P[A \cap \{(k-1)2^{-n} < T \leq k2^{-n}\}]\, E[F(B_{t_1}, \ldots, B_{t_p})],
\end{aligned}$$

et il ne reste plus qu'à sommer sur k pour arriver au résultat souhaité.

Lorsque, $P[T = \infty] > 0$, les mêmes arguments conduisent à

$$E[\mathbf{1}_{A\cap\{T<\infty\}}\, F(B^{(T)}_{t_1}, \ldots, B^{(T)}_{t_p})] = P[A \cap \{T < \infty\}]\, E[F(B_{t_1}, \ldots, B_{t_p})]$$

et le résultat recherché en découle à nouveau. □

Une application importante de la propriété de Markov forte est le principe de réflexion illustré dans la preuve du théorème suivant.

Théorème 2.12 *Pour tout $t > 0$, notons $S_t = \sup_{s\leq t} B_s$. Alors, si $a \geq 0$ et $b \leq a$, on a*

$$P[S_t \geq a,\ B_t \leq b] = P[B_t \geq 2a - b].$$

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

Démonstration. On applique la propriété de Markov forte au temps d'arrêt

$$T_a = \inf\{t \geq 0, B_t = a\}.$$

On a déjà vu (Corollaire 2.9) que $T_a < \infty$ p.s. Ensuite,

$$P[S_t \geq a,\ B_t \leq b] = P[T_a \leq t,\ B_t \leq b] = P[T_a \leq t,\ B^{(T_a)}_{t-T_a} \leq b - a],$$

puisque $B^{(T_a)}_{t-T_a} = B_t - B_{T_a} = B_t - a$. Notons $B' = B^{(T_a)}$, de sorte que d'après le théorème 2.11, le processus B' est un mouvement brownien indépendant de $\mathcal{F}_{T_a}$ donc en particulier de T_a. Comme B' a même loi que $-B'$, le couple (T_a, B') a aussi même loi que $(T_a, -B')$. Notons $H = \{(s, \mathrm{w}) \in \mathbb{R}_+ \times C(\mathbb{R}_+, \mathbb{R});\ s \leq t,\ \mathrm{w}(t-s) \leq b - a\}$. La probabilité précédente vaut

$$\begin{aligned}
P[(T_a, B') \in H] &= P[(T_a, -B') \in H] \\
&= P[T_a \leq t,\ -B^{(T_a)}_{t-T_a} \leq b - a] \\
&= P[T_a \leq t,\ B_t \geq 2a - b] \\
&= P[B_t \geq 2a - b]
\end{aligned}$$

parce que l'événement $\{B_t \geq 2a - b\}$ est contenu dans $\{T_a \leq t\}$.

Pour la deuxième assertion on observe que

$$P[S_t \geq a] = P[S_t \geq a, B_t \geq a] + P[S_t \geq a, B_t \leq a] = 2P[B_t \geq a] = P[|B_t] \geq a],$$

d'où le résultat voulu. $\square$

On déduit immédiatement du théorème précédent que la loi du couple (S_t, B_t) a pour densité

$$g(a,b) = \frac{2(2a-b)}{\sqrt{2\pi t^3}} \exp\left(-\frac{(2a-b)^2}{2t}\right) \mathbf{1}_{\{a>0,b<a\}}.$$

Corollaire 2.13 *Pour tout $a > 0$, T_a a même loi que $\dfrac{a^2}{B_1^2}$ et a donc pour densité*

$$f(t) = \frac{a}{\sqrt{2\pi t^3}} \exp\Big(-\frac{a^2}{2t}\Big) \mathbf{1}_{\{t>0\}}.$$

Démonstration. On écrit

$$\begin{aligned} P[T_a \leq t] &= P[S_t \geq a] \\ &= P[|B_t| \geq a] \quad \text{(Th.2.12)} \\ &= P[B_t^2 \geq a^2] \\ &= P[tB_1^2 \geq a^2] \\ &= P[\frac{a^2}{B_1^2} \leq t]. \end{aligned}$$

Ensuite, puisque B_1 suit une loi $\mathcal{N}(0,1)$ on calcule facilement la densité de a^2/B_1^2.$\square$

Généralisations. Soit Z une variable aléatoire réelle. Un processus $(X_t, t \geq 0)$ est appelé mouvement brownien (réel) issu de Z si on peut écrire $X_t = Z + B_t$ où B est un mouvement brownien issu de 0 et *indépendant* de Z.

Un processus $B_t = (B_t^1, \ldots, B_t^d)$ à valeurs dans $\mathbb{R}^d$ est un mouvement brownien (en dimension d) issu de 0 si ses composantes $B^1, \ldots, B^d$ sont des mouvements browniens réels issus de 0 *indépendants.* On vérifie facilement que si Φ est une isométrie vectorielle de $\mathbb{R}^d$, le processus $(\Phi(B_t))$ est encore un mouvement brownien en dimension d.

Enfin, si Z est une variable aléatoire à valeurs dans $\mathbb{R}^d$ et $X_t = (X_t^1, \ldots, X_t^d)$ un processus à valeurs dans $\mathbb{R}^d$, on dit que X est un mouvement brownien en dimension d issu de Z si on peut écrire $X_t = Z + B_t$ où B est un mouvement brownien en dimension d issu de 0 et *indépendant* de Z.

La plupart des résultats qui précèdent peuvent être étendus au mouvement brownien en dimension d. En particulier, la propriété de Markov forte reste vraie, avec exactement la même démonstration.

Chapitre 3

Filtrations et Martingales

Dans ce chapitre, nous introduisons les rudiments de la théorie générale des processus sur un espace de probabilité muni d'une filtration. Cela nous amène à généraliser plusieurs notions introduites dans le chapitre précédent dans le cadre du mouvement brownien. Dans un second temps, nous développons la théorie des martingales à temps continu et nous établissons en particulier les résultats de régularité des trajectoires.

3.1 Filtrations et processus

Définition 3.1 *Soit* $(\Omega, \mathcal{A}, P)$ *un espace de probabilité. Une filtration sur cet espace est une famille croissante* $(\mathcal{F}_t)_{0\le t\le\infty}$ *de sous-tribus de* $\mathcal{A}$.

Exemple. Si B est un mouvement brownien, on peut prendre

$$\mathcal{F}_t = \sigma(B_s, 0 \le s \le t), \quad \mathcal{F}_\infty = \sigma(B_s, s \ge 0).$$

Plus généralement, si $X = (X_t, t \ge 0)$ est un processus indexé par $\mathbb{R}_+$, la *filtration canonique* de X est $\mathcal{F}_t = \sigma(X_s, s \le t)$ (ou par abus de langage l'augmentation habituelle de $(\mathcal{F}_t)$ comme cela sera défini plus loin).

On pose pour tout $t > 0$

$$\mathcal{F}_{t+} = \bigcap_{s>t} \mathcal{F}_s.$$

La famille $(\mathcal{F}_{t+})$ (avec $\mathcal{F}_{\infty+} = \mathcal{F}_\infty$) est aussi une filtration. On dit que la filtration $(\mathcal{F}_t)$ est *continue à droite* si

$$\mathcal{F}_{t+} = \mathcal{F}_t, \qquad \forall t \ge 0.$$

Soit $(\mathcal{F}_t)$ une filtration et soit $\mathcal{N}$ la classe des ensembles P-négligeables de $\mathcal{F}_\infty$ (i.e. $A \in \mathcal{N}$ si $\exists A' \in \mathcal{F}_\infty$ tel que $A \subset A'$ et $P(A') = 0$). La filtration est dite *complète* si $\mathcal{N} \subset \mathcal{F}_0$ (et donc $\mathcal{N} \subset \mathcal{F}_t$ pour tout t).

On dira qu'une filtration $(\mathcal{F}_t)$ satisfait les *conditions habituelles* si elle est à la fois continue à droite et complète. Partant d'une filtration quelconque $(\mathcal{F}_t)$ on peut construire une filtration qui satisfait les conditions habituelles, simplement en ajoutant à chaque tribu $\mathcal{F}_{t+}$ la classe des P-négligeables de $\mathcal{F}_\infty$. C'est l'*augmentation habituelle* de la filtration $(\mathcal{F}_t)$.

Définition 3.2 *Un processus $(X_t, t \geq 0)$ est dit mesurable si l'application*

$$(t, \omega) \longrightarrow X_t(\omega)$$

définie sur $\mathbb{R}_+ \times \Omega$ muni de la tribu produit $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{A}$ est mesurable.

Cette propriété est plus forte que de dire que, pour tout $t \geq 0$, X_t est $\mathcal{A}$-mesurable. Cependant, si l'on impose que les trajectoires de X sont continues, ou seulement continues à droite, il est facile de voir que les deux propriétés sont équivalentes (approcher X par des processus "en escalier" qui sont mesurables).

Dans toute la suite on suppose qu'on s'est donné une filtration $(\mathcal{F}_t)$ sur un espace de probabilité.

Définition 3.3 *Un processus X est dit adapté si pour tout $t \geq 0$, X_t est $\mathcal{F}_t$-mesurable. Le processus est dit progressif si, pour tout $t \geq 0$, l'application*

$$(s, \omega) \longrightarrow X_s(\omega)$$

est mesurable sur $[0, t] \times \Omega$ muni de la tribu $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$.

Remarquons qu'un processus progressif est adapté et mesurable (dire qu'un processus est mesurable est équivalent à dire que pour tout $t \geq 0$, $(s, \omega) \longrightarrow X_s(\omega)$ est mesurable sur $[0, t] \times \Omega$ muni de la tribu $\mathcal{B}([0, t]) \otimes \mathcal{A}$).

Proposition 3.1 *Soit $(X_t, t \geq 0)$ un processus à valeurs dans un espace métrique (E, d). Supposons que X est adapté et à trajectoires continues à droite (i.e. pour tout $\omega \in \Omega$, $t \to X_t(\omega)$ est continue à droite). Alors X est progressif.*

Lemme 3.2 *Pour tout $\varepsilon > 0$, soit $\mathcal{F}_t^\varepsilon = \mathcal{F}_{\varepsilon+t}$. Un processus adapté X est progressif relativement à la filtration $(\mathcal{F}_t)$ si et seulement si, pour tout $\varepsilon > 0$, il est progressif relativement à la filtration $(\mathcal{F}_t^\varepsilon)$.*

Démonstration. Il est évident que si X est progressif relativement à $(\mathcal{F}_t)$, il l'est aussi relativement à $(\mathcal{F}_t^\varepsilon)$. Inversement, fixons $t > 0$ et pour tout $\varepsilon \in]0, t[$ posons

$$X_s^\varepsilon(\omega) = X_s(\omega)\mathbf{1}_{[0,t-\varepsilon]}(s) + X_t(\omega)\mathbf{1}_{\{t\}}(s).$$

Alors X^ε est mesurable pour $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$ (par hypothèse, l'application

$$\begin{array}{rcl} [0,t-\varepsilon] \times \Omega & \longrightarrow & E \\ (s,\omega) & \longrightarrow & X_s(\omega) \end{array}$$

est $\mathcal{B}([0,t-\varepsilon]) \otimes \mathcal{F}_t$-mesurable). De plus, pour $s \leq t$, $X_s = X_s^\varepsilon$ dès que ε est assez petit. Le résultat recherché en découle. $\square$

Démonstration de la Proposition 3.1. Pour tout $n \geq 1$, posons

$$X_t^n = \sum_{k=1}^{\infty} X_{k/n} \, \mathbf{1}_{[(k-1)/n, k/n[}(t).$$

La continuité à droite des trajectoires assure que

$$\forall t \geq 0, \ \forall \omega \in \Omega, \quad X_t(\omega) = \lim_{n\to\infty} X_t^n(\omega).$$

D'autre part, X^n est progressif par rapport à $(\mathcal{F}_t^\varepsilon)$ dès que $n^{-1} < \varepsilon$. On conclut donc que X est progressif par rapport à $(\mathcal{F}_t^\varepsilon)$ et on applique le lemme. $\square$

Remarque. On peut remplacer continu à droite par continu à gauche dans l'énoncé de la proposition. La preuve est même plus facile et n'utilise pas le lemme.

Tribu progressive. La famille des parties $A \in \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{A}$ telles que le processus $X_t(\omega) = \mathbf{1}_A(t,\omega)$ soit progressif forme une tribu sur $\mathbb{R}_+ \times \Omega$, appelée la *tribu progressive.* Il est alors facile de vérifier (exercice !) qu'un processus X est progressif si et seulement si l'application $(t,\omega) \to X_t(\omega)$ est mesurable sur $\mathbb{R}_+ \times \Omega$ muni de la tribu progressive.

3.2 Temps d'arrêt et tribus associées

Définition 3.4 *Soit $(\mathcal{F}_t)$ une filtration sur $(\Omega, \mathcal{A}, P)$. Une v.a. $T : \Omega \longrightarrow \bar{\mathbb{R}}_+ = \mathbb{R}_+ \cup \{+\infty\}$ est un $(\mathcal{F}_t)$-temps d'arrêt si $\forall t \geq 0$, $\{T \leq t\} \in \mathcal{F}_t$.*

On associe à un temps d'arrêt T les tribus suivantes:

$$\begin{aligned} \mathcal{F}_T &= \{A \in \mathcal{F}_\infty, \forall t \geq 0, \ A \cap \{T \leq t\} \in \mathcal{F}_t\} \\ \mathcal{F}_{T+} &= \{A \in \mathcal{F}_\infty, \forall t \geq 0, \ A \cap \{T < t\} \in \mathcal{F}_t\} \\ \mathcal{F}_{T-} &= \sigma(A \cap \{T > t\}; \ t \geq 0, A \in \mathcal{F}_t\}. \end{aligned}$$

Propriétés.

(a) On a toujours $\mathcal{F}_{T-} \subset \mathcal{F}_T \subset \mathcal{F}_{T+}$. Si la filtration $(\mathcal{F}_t)$ est continue à droite, on a $\mathcal{F}_{T+} = \mathcal{F}_T$.

(b) Une v.a. $T : \Omega \to \bar{\mathbb{R}}_+$ est un $(\mathcal{F}_{t+})$-temps d'arrêt si et seulement si $\forall t \geq 0$, $\{T < t\} \in \mathcal{F}_t$. Cela équivaut encore à dire que $T \wedge t$ est $\mathcal{F}_t$-mesurable, pour tout t.

(c) Si $T = t$, $\mathcal{F}_T = \mathcal{F}_t$, $\mathcal{F}_{T+} = \mathcal{F}_{t+}$.

(d) Pour $A \in \mathcal{F}_\infty$, posons

$$T^A(\omega) = \begin{cases} T(\omega) & \text{si } \omega \in A \\ +\infty & \text{si } \omega \notin A \end{cases}$$

Alors $A \in \mathcal{F}_T$ si et seulement si T^A est un temps d'arrêt.

(e) T est $\mathcal{F}_T$-mesurable.

(f) Si S, T sont deux temps d'arrêt et si $S \leq T$, alors $\mathcal{F}_S \subset \mathcal{F}_T$ et $\mathcal{F}_{S+} \subset \mathcal{F}_{T_+}$. En général, $S \vee T$ et $S \wedge T$ sont deux temps d'arrêt et $\mathcal{F}_{S \wedge T} = \mathcal{F}_S \cap \mathcal{F}_T$. De plus, $\{S \leq T\} \in \mathcal{F}_{S \wedge T}$, $\{S = T\} \in \mathcal{F}_{S \wedge T}$.

(g) Si (S_n) est une suite croissante de temps d'arrêt, alors $S = \lim \uparrow S_n$ est aussi un temps d'arrêt, et

$$\mathcal{F}_{S-} = \bigvee_n \mathcal{F}_{S_n-}.$$

(h) Si (S_n) est une suite décroissante de temps d'arrêt, alors $S = \lim \downarrow S_n$ est un temps d'arrêt de la filtration $(\mathcal{F}_{t+})$, et

$$\mathcal{F}_{S+} = \bigcap_n \mathcal{F}_{S_n+}.$$

(i) Si (S_n) est une suite décroissante *stationnaire* de temps d'arrêt (i.e. $\forall \omega$, $\exists N(\omega)$: $\forall n \geq N(\omega)$, $S_n(\omega) = S(\omega)$) alors $S = \lim \downarrow S_n$ est aussi un temps d'arrêt, et

$$\mathcal{F}_S = \bigcap_n \mathcal{F}_{S_n}.$$

La preuve de ces propriétés est facile. Démontrons par exemple les quatre dernières.

(f) Si $S \leq T$ et $A \in \mathcal{F}_S$ alors

$$A \cap \{T \leq t\} = (A \cap \{S \leq t\}) \cap \{T \leq t\} \in \mathcal{F}_t,$$

d'où $A \in \mathcal{F}_T$.

En général,

$$\{S \wedge T \leq t\} = \{S \leq t\} \cup \{T \leq t\} \in \mathcal{F}_t,$$
$$\{S \vee T \leq t\} = \{S \leq t\} \cap \{T \leq t\} \in \mathcal{F}_t.$$

Il est évident d'après ci-dessus que $\mathcal{F}_{S\wedge T} \subset (\mathcal{F}_S \cap \mathcal{F}_T)$. De plus, si $A \in \mathcal{F}_S \cap \mathcal{F}_T$,

$$A \cap \{S \wedge T \leq t\} = (A \cap \{S \leq t\}) \cup (A \cap \{T \leq t\}) \in \mathcal{F}_t,$$

d'où $A \in \mathcal{F}_{S\wedge T}$.

Ensuite, pour $t \geq 0$,

$$\{S \leq T\} \cap \{T \leq t\} = \{S \leq t\} \cap \{T \leq t\} \cap \{S \wedge t \leq T \wedge t\} \in \mathcal{F}_t$$
$$\{S \leq T\} \cap \{S \leq t\} = \{S \wedge t \leq T \wedge t\} \cap \{S \leq t\} \in \mathcal{F}_t,$$

ce qui donne $\{S \leq T\} \in \mathcal{F}_S \cap \mathcal{F}_T = \mathcal{F}_{S\wedge T}$. L'argument pour $\{S = T\}$ est le même.

(g) On a

$$\{S \leq t\} = \bigcap_n \{S_n \leq t\} \in \mathcal{F}_t.$$

De plus, on voit d'abord que $\mathcal{F}_{S_n-} \subset \mathcal{F}_{S-}$ (pour $A \in \mathcal{F}_t$, écrire $A \cap \{S_n > t\} = (A \cap \{S_n > t\}) \cap \{S > t\})$). Donc $\bigvee_n \mathcal{F}_{S_n-} \subset \mathcal{F}_{S-}$. Ensuite, pour $A \in \mathcal{F}_t$,

$$A \cap \{S > t\} = \bigcup_n (A \cap \{S_n > t\}),$$

ce qui montre que $A \cap \{S > t\} \in \bigvee_n \mathcal{F}_{S_n-}$.

(h) On écrit

$$\{S < t\} = \bigcup_n \{S_n < t\} \in \mathcal{F}_t.$$

De plus, il est facile de voir que $\mathcal{F}_{S+} \subset \mathcal{F}_{S_n+}$ pour tout n, et inversement si $A \in \bigcap_n \mathcal{F}_{S_n+}$,

$$A \cap \{S < t\} = \bigcup_n (A \cap \{S_n < t\}) \in \mathcal{F}_t,$$

d'où $A \in \mathcal{F}_{S+}$.

(i) Dans ce cas on a aussi

$$\{S \leq t\} = \bigcup_n \{S_n \leq t\} \in \mathcal{F}_t,$$

et pour $A \in \bigcap_n \mathcal{F}_{S_n}$,

$$A \cap \{S \leq t\} = \bigcup_n (A \cap \{S_n \leq t\}) \in \mathcal{F}_t,$$

d'où $A \in \mathcal{F}_S$. $\square$

Théorème 3.3 *Soit* $(X_t, t \geq 0)$ *un processus progressif et* T *un temps d'arrêt. Alors* $X_T \, \mathbf{1}_{\{T<\infty\}}$ *est* $\mathcal{F}_T$*-mesurable.*

Démonstration. De la définition il découle très facilement qu'une fonction $Y : \Omega \to \mathbb{R}$ nulle sur $\{T = \infty\}$ est $\mathcal{F}_T$-mesurable si et seulement si pour tout $t \geq 0$, $Y \, \mathbf{1}_{\{T \leq t\}}$ est $\mathcal{F}_t$-mesurable (écrire pour $A \in \mathcal{B}(\mathbb{R})$ avec $0 \notin A$, $\{Y \in A\} \cap \{T \leq t\} = \{Y \, \mathbf{1}_{\{T \leq t\}} \in A\}$).

On applique cette observation à $Y = X_T \, \mathbf{1}_{\{T<\infty\}}$, de sorte que $Y \, \mathbf{1}_{\{T \leq t\}} = X_T \, \mathbf{1}_{\{T \leq t\}} = X_{T \wedge t} \, \mathbf{1}_{\{T \leq t\}}$. Or $X_{T \wedge t}$ est la composition des deux applications

$$\begin{array}{ccc} \omega & \longrightarrow & (T(\omega) \wedge t, \omega) \\ \mathcal{F}_t & & \mathcal{B}([0,t]) \otimes \mathcal{F}_t \end{array}$$

et

$$\begin{array}{ccc} (s, \omega) & \longrightarrow & X_s(\omega) \\ \mathcal{B}([0,t]) \otimes \mathcal{F}_t & & \mathcal{B}(\mathbb{R}) \end{array}$$

qui sont toutes les deux mesurables (la deuxième par définition d'un processus progressif). On conclut que $X_{T \wedge t}$, donc aussi $X_{T \wedge t} \, \mathbf{1}_{\{T \leq t\}}$ est $\mathcal{F}_t$-mesurable. □

Proposition 3.4 *Soient* T *un temps d'arrêt et* S *une v.a.* $\mathcal{F}_T$*-mesurable telle que* $S \geq T$*. Alors* S *est aussi un temps d'arrêt.*

En particulier, si T *est un temps d'arrêt,*

$$T_n = \sum_{k=0}^{\infty} \frac{k+1}{2^n} \, \mathbf{1}_{\{k2^{-n} < T \leq (k+1)2^{-n}\}} + \infty \cdot \mathbf{1}_{\{T=\infty\}}$$

est une suite de temps d'arrêt qui décroît vers T*.*

Démonstration. On écrit

$$\{S \leq t\} = \{S \leq t\} \cap \{T \leq t\} \in \mathcal{F}_t$$

puisque $\{S \leq t\}$ est $\mathcal{F}_T$-mesurable. La deuxième assertion en découle puisque T_n est clairement $\mathcal{F}_T$-mesurable et $T_n \geq T$. □

Nous terminons ce paragraphe avec un théorème important qui permet la construction de nombreux temps d'arrêt.

Théorème 3.5 *On suppose que la filtration* $(\mathcal{F}_t)$ *vérifie les conditions habituelles (elle est complète et continue à droite). Soit* $A \subset \mathbb{R}_+ \times \Omega$ *un ensemble progressif (i.e. tel que le processus* $X_t(\omega) = \mathbf{1}_A(t, \omega)$ *soit progressif) et soit* D_A *le début de* A *défini par*

$$D_A(\omega) = \inf\{t \geq 0, \, (t, \omega) \in A\} \qquad (\inf \emptyset = \infty).$$

Alors D_A *est un temps d'arrêt.*

La preuve de ce théorème repose sur le résultat difficile de théorie de la mesure qui suit.
Théorème. *Soit* $(\Lambda, \mathcal{G}, \Pi)$ *un espace de probabilité complet, au sens où la tribu* $\mathcal{G}$ *contient tous les ensembles* Π*-négligeables. Soit* $H \subset \mathbb{R}_+ \times \Lambda$ *un ensemble mesurable pour la tribu produit* $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{G}$*. Alors la projection de* H*,*

$$p(H) = \{\omega \in \Lambda; \exists t \geq 0,\ (t, \omega) \in H\}$$

appartient à la tribu $\mathcal{G}$*.*
(Pour une preuve voir le traité de C. Dellacherie et P.A. Meyer, Probabilités et Potentiel, Vol. I, pages 68 et 92-93.)

Démonstration du Théorème 3.5. On applique le théorème ci-dessus en fixant $t \geq 0$, en prenant $(\Lambda, \mathcal{G}, \Pi) = (\Omega, \mathcal{F}_t, P)$ et $H = ([0, t[\times \Omega) \cap A$. La définition d'un processus progressif montre que H est mesurable pour $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}_t$. En conséquence,

$$p(H) = \{\omega \in \Omega,\ \exists s < t,\ (s, \omega) \in A\} \in \mathcal{F}_t.$$

Or il est clair que l'événement $\{D_A < t\}$ coïncide avec cet ensemble. On a donc obtenu $\{D_A < t\} \in \mathcal{F}_t$ et puisque la filtration est continue à droite, D_A est un temps d'arrêt. □

Remarque. Dans des cas particuliers, qui sont ceux qu'on utilise le plus fréquemment, on peut obtenir le résultat du Théorème 3.5 plus facilement, et sans hypothèse sur la filtration. On considère un processus X à valeurs dans un espace métrique (E, d):

(i) Si $(X_t, t \geq 0)$ est un processus à trajectoires continues à droite et adapté, si O est un ouvert de E, alors

$$T_O = \inf\{t \geq 0,\ X_t \in O\}$$

est un temps d'arrêt de $(\mathcal{F}_{t+})$. En effet,

$$\{T_O < t\} = \bigcup_{s \in [0,t[\cap \mathbb{Q}} \{X_s \in O\}.$$

(ii) Si $(X_t, t \geq 0)$ est un processus à trajectoires continues et adapté, et si F est un fermé, alors

$$T_F = \inf\{t \geq 0,\ X_t \in F\}$$

est un temps d'arrêt. En effet,

$$\{T_F \leq t\} = \{\inf_{0 \leq s \leq t} d(X_s, F) = 0\} = \{\inf_{s \in [0,t] \cap \mathbb{Q}} d(X_s, F) = 0\}.$$

3.3 Martingales et surmartingales à temps continu

On suppose donné un espace de probabilité muni d'une filtration $(\Omega, \mathcal{A}, (\mathcal{F}_t), P)$.

Définition 3.5 *Un processus $(X_t, t \geq 0)$ adapté et tel que, pour tout $t \geq 0$, $X_t \in L^1$ est appelé*

- *martingale si pour $s < t$, $E[X_t \mid \mathcal{F}_s] = X_s$;*
- *surmartingale si pour $s < t$, $E[X_t \mid \mathcal{F}_s] \leq X_s$;*
- *sous-martingale si pour $s < t$, $E[X_t \mid \mathcal{F}_s] \geq X_s$.*

Exemples. On dit qu'un processus $(Z_t, t \geq 0)$ est un processus à accroissements indépendants (PAI) par rapport à la filtration $(\mathcal{F}_t)$ si Z est $(\mathcal{F}_t)$-adapté et si, pour $s < t$, $Z_t - Z_s$ est indépendant de la tribu $\mathcal{F}_s$ (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 $(\mathcal{F}_t)$ alors

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

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

(iii) si pour $\theta \in \mathbb{R}$, $E[e^{\theta Z_t}] < \infty$ pour tout $t \geq 0$,

$$X_t = \frac{e^{\theta Z_t}}{E[e^{\theta Z_t}]}$$

est une martingale.

Les démonstrations sont très faciles. Dans le deuxième cas,

$$\begin{aligned}
E[(\widetilde{Z}_t)^2 \mid \mathcal{F}_s] &= E[(\widetilde{Z}_s + \widetilde{Z}_t - \widetilde{Z}_s)^2 \mid \mathcal{F}_s] \\
&= \widetilde{Z}_s^2 + 2\widetilde{Z}_s E[(\widetilde{Z}_t - \widetilde{Z}_s) \mid \mathcal{F}_s] + E[(\widetilde{Z}_t - \widetilde{Z}_s)^2 \mid \mathcal{F}_s] \\
&= \widetilde{Z}_s^2 + 2\widetilde{Z}_s E[(\widetilde{Z}_t - \widetilde{Z}_s)] + E[(\widetilde{Z}_t - \widetilde{Z}_s)^2] \\
&= \widetilde{Z}_s^2 + E[\widetilde{Z}_t^2] - E[\widetilde{Z}_s^2],
\end{aligned}$$

d'où le résultat voulu. Dans le troisième cas,

$$E[X_t \mid \mathcal{F}_s] = \frac{e^{\theta Z_s}\, E[e^{\theta(Z_t - Z_s)} \mid \mathcal{F}_s]}{E[e^{\theta Z_s}]\, E[e^{\theta(Z_t - Z_s)}]} = \frac{e^{\theta Z_s}}{E[e^{\theta Z_s}]} = X_s.$$

En prenant $Z = B$ (avec sa filtration canonique), on voit que les processus

$$B_t \ , \ B_t^2 - t \ , \ e^{\theta B_t - \frac{\theta^2}{2} t}$$

sont des martingales. On peut aussi prendre pour $f \in L^2_{\rm loc}(\mathbb{R}_+, dt)$,

$$Z_t = \int_0^t f(s)\, dB_s \quad (= G(f\mathbf{1}_{[0,t]})).$$

Les propriétés élémentaires des mesures gaussiennes montrent que Z est un PAI par rapport à la filtration canonique de B, et donc

$$\int_0^t f(s)dB_s \, , \, \Big(\int_0^t f(s)dB_s\Big)^2 - \int_0^t f(s)^2 ds \, , \, \exp\Big(\theta \int_0^t f(s)dB_s - \frac{\theta^2}{2} \int_0^t f(s)^2 ds\Big)$$

sont des martingales. On peut aussi prendre $Z = N$, processus de Poisson standard, et on trouve alors en particulier que

$$N_t - t \ , \ (N_t - t)^2 - t$$

sont des martingales.

Soit (X_t) une martingale (respectivement une sous-martingale) et soit $f : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction convexe (resp. une fonction convexe croissante). Supposons aussi que $f(X_t) \in L^1$ pour tout $t \geq 0$. Alors, $(f(X_t), t \geq 0)$ est une sous-martingale.

En effet, d'après l'inégalité de Jensen, on a pour $s < t$

$$E[f(X_t) \mid \mathcal{F}_s] \geq f(E[X_t \mid \mathcal{F}_s]) \geq f(X_s).$$

On peut appliquer cette observation à la fonction convexe croissante $f(x) = x^+$. On trouve que si $(X_t, t \geq 0)$ est une sous-martingale, $((X_t)^+, t \geq 0)$ est aussi une sous-martingale. En particulier, on a pour tout $s \in [0, t]$,

$$E[(X_s)^+] \leq E[(X_t)^+].$$

D'autre part, puisque X est une sous-martingale on a aussi pour $s \in [0, t]$,

$$E[X_s] \geq E[X_0].$$

En combinant ces deux inégalités, et en remarquant que $|x| = 2x^+ - x$, on trouve

$$\sup_{s\in[0,t]} E[|X_s|] \leq 2\, E[(X_t)^+] - E[X_0] < \infty.$$

Donc, si $(X_t, t \geq 0)$ est une sous-martingale (ou une surmartingale, en changeant X en $-X$), on a pour tout $t \geq 0$, $\sup_{s\in[0,t]} E[|X_s|] < \infty$. Cette remarque simple sera utile dans la suite. Observons aussi que si X est une martingale, pour tout $p \geq 1$, $|X_t|^p$ est une sous-martingale dès que $E[|X_t|^p] < \infty$ pour tout t, et donc $E[|X_t|^p]$ est fonction croissante de t.

Notre objectif est maintenant d'établir des résultats de régularité des trajectoires pour les martingales et les surmartingales à temps continu. Nous commençons par rappeler les résultats connus à temps discret [une bonne source pour ces résultats est le livre de Dellacherie et Meyer "Probabilités et Potentiel. Volume 2. Théorie des martingales"], dont on peut immédiatement déduire des conséquences intéressantes à temps continu.

Inégalités maximales et inégalités de Doob.

Soit $(Y_n, n \in \mathbb{N})$ une surmartingale à temps discret. Alors, pour tous $\lambda > 0$, $k \geq 0$,

$$\lambda\, P[\sup_{n \leq k} |Y_n| \geq \lambda] \leq E[|Y_0|] + 2E[|Y_k|].$$

Si maintenant $(X_s, s \geq 0)$ est une surmartingale à temps continu, on peut appliquer ce résultat à la surmartingale à temps discret $Y_n = X_{t_{n \wedge k}}$, pour toute partie finie $\{t_1 < t_2 < \cdots < t_k = t\}$ de $[0, t]$. Ensuite, si D est un sous-ensemble dénombrable de $\mathbb{R}_+$, en écrivant D comme limite croissante de parties finies, on obtient que

$$\lambda\, P[\sup_{s \in [0,t] \cap D} |X_s| \geq \lambda] \leq E[|X_0|] + 2E[|X_t|].$$

Si on suppose de plus que les trajectoires de X sont continues à droite alors, en prenant D dense et tel que $t \in D$, on a

$$\sup_{s \in [0,t] \cap D} |X_s| = \sup_{s \in [0,t]} |X_s|,$$

ce qui montre que dans l'inégalité précédente on peut prendre $\sup_{s \in [0,t]}$ au lieu de $\sup_{s \in [0,t] \cap D}$.

L'inégalité de Doob affirme que si $(Y_n, n \in \mathbb{N})$ est une martingale à temps discret telle que $Y_n \in L^p$ (pour $p > 1$) alors, pour tout $k \geq 0$,

$$\| \sup_{n \leq k} |Y_n| \|_p \leq q\, \|Y_k\|_p \ ,$$

où $\frac{1}{p} + \frac{1}{q} = 1$. On en déduit que si $(X_t, t \geq 0)$ est une martingale à temps continu dans L^p,

$$\| \sup_{s \in [0,t] \cap D} |X_s| \|_p \leq q \|X_t\|_p,$$

et on peut remplacer $\sup_{s \in [0,t] \cap D}$ par $\sup_{s \in [0,t]}$ si X a des trajectoires continues à droite. Résumons les résultats obtenus.

Proposition 3.6 *Soit* $(X_t, t \geq 0)$ *une surmartingale dont les trajectoires sont continues à droite. Alors, pour tous* $t \in [0, \infty]$, $\lambda > 0$,

$$\lambda\, P[\sup_{s \leq t} |X_s| \geq \lambda] \leq 3 \sup_{s \leq t} E[|X_t|].$$

Si $(X_t, t \geq 0)$ *est une martingale dont les trajectoires sont continues à droite, et si* $X_t \in L^p$ *pour tout* $t \geq 0$, *alors*

$$\| \sup_{s \leq t} |X_s| \|_p \leq q \|X_t\|_p,$$

où $\frac{1}{p} + \frac{1}{q} = 1$.

Remarque. La quantité $\sup_{s \leq t} E[|X_s|]$ est finie d'après une observation précédente.

Nombres de montées.

Si $f : T \longrightarrow \mathbb{R}$ est une fonction définie sur une partie T de $\mathbb{R}_+$, et si $a < b$, le nombre de montées de f le long de $[a, b]$, noté $M_{ab}^f(T)$ est le supremum des entiers k tels que l'on puisse trouver une suite croissante $s_1 < t_1 < \cdots < s_k < t_k$ avec $f(s_i) < a$, $f(t_i) > b$.

On sait que si (Y_n) est une surmartingale (à temps discret),

$$E[M_{ab}^Y([0, n])] \leq \frac{1}{b - a} E[(Y_n - a)_-].$$

On en déduit immédiatement le résultat suivant, en prenant d'abord D fini.

Proposition 3.7 *Soit* $(X_t, t \geq 0)$ *une surmartingale. Si* D *est un sous-ensemble dénombrable de* $\mathbb{R}_+$,

$$E[M_{ab}^X([0, t] \cap D)] \leq \frac{1}{b - a} E[(X_t - a)_-].$$

Théorèmes de convergence.

Si $(Y_n, n \geq 0)$ est une surmartingale bornée dans L^1 (en particulier si elle est positive), alors $Y_n \longrightarrow Y_\infty$ p.s. et $Y_\infty \in L^1$.

Si $(Y_n, n \leq 0)$ est une surmartingale *indexée par les entiers négatifs*, et telle que $\sup_{n \leq 0} E[Y_n] < \infty$, alors la suite (Y_n) est uniformément intégrable, et converge p.s. et dans L^1 quand $n \to -\infty$ vers une variable Z telle que $Z \geq E[X_n \mid \mathcal{F}_{-\infty}]$ (par définition, $\mathcal{F}_{-\infty} = \bigcap_{n \leq 0} \mathcal{F}_n$).

Le théorème suivant applique ces résultats "à temps discret" aux surmartingales à temps continu.

Théorème 3.8 *Soit* $(X_t, t \geq 0)$ *une surmartingale et soit* D *un sous-ensemble dénombrable dense de* $\mathbb{R}_+$.

(i) *Pour presque tout* $\omega \in \Omega$, *l'application* $s \to X_s(\omega)$ *définie sur* D *et à valeurs dans* $\mathbb{R}$ *admet en tout point* t *de* $\mathbb{R}_+$ *une limite à droite* $X_{t+}(\omega)$ *finie et en tout point* t *de* $\mathbb{R}_+\backslash\{0\}$ *une limite à gauche* $X_{t-}(\omega)$ *finie.*

(ii) *Pour tout* $t \in \mathbb{R}_+$, $X_{t+} \in L^1$ *et*

$$X_t \geq E[X_{t+} \mid \mathcal{F}_t]$$

avec égalité si la fonction $t \longrightarrow E[X_t]$ *est continue à droite (en particulier si* X *est une martingale). Le processus* $(X_{t+}, t \geq 0)$ *est une surmartingale par rapport à la filtration* $(\mathcal{F}_{t+})$, *une martingale si* X *en est une.*

Démonstration. (i) Fixons $T > 0$. L'inégalité maximale de Doob entraîne aussitôt que

$$\lim_{\lambda \to \infty} P[\sup_{s \in D \cap [0,T]} |X_s| \geq \lambda] = 0,$$

d'où

$$\sup_{s \in D \cap [0,T]} |X_s| < \infty \ , \quad \text{p.s.}$$

Ensuite, l'inégalité sur les nombres de montées montre que, p.s. pour tous $a < b$ rationnels,

$$M_{ab}^X(D \cap [0,T]) < \infty.$$

Donc p.s. la fonction $s \longrightarrow X_s(\omega)$ est bornée sur $D \cap [0,T]$ et, pour tous $a < b$ rationnels, ne fait qu'un nombre fini de montées le long de $[a,b]$. La propriété énoncée en (i) en découle: en effet si une fonction bornée $f : D \cap [0,T] \to \mathbb{R}$ n'a par exemple pas de limite à gauche finie en $t \in]0,T]$, on peut choisir $a < b$ rationnels de façon que

$$\liminf_{s \uparrow\uparrow t} f(s) < a < b < \limsup_{s \uparrow\uparrow t} f(s),$$

et on voit facilement que le nombre de montées de f le long de $[a,b]$ doit être infini.

(ii) Pour que $X_{t+}(\omega)$ soit défini pour tout ω et pas seulement sur l'ensemble de probabilité un où la limite existe, on peut convenir de prendre $X_{t+}(\omega) = 0$ sur l'ensemble ($\mathcal{F}_{t+}$-mesurable et de probabilité nulle) où la limite à droite en t le long de D n'existe pas. Avec cette convention, $X_{t+}(\omega)$ est bien défini pour tout ω et est $\mathcal{F}_{t+}$-mesurable.

Fixons $t \geq 0$ et choisissons une suite (t_n) dans D avec $t_n \downarrow\downarrow t$. Alors, par construction on a p.s.,

$$X_{t+} = \lim_{n \to -\infty} X_{t_n}.$$

De plus, posons, pour $k \leq 0$, $Y_k = X_{t_{-k}}$. Alors Y est une surmartingale indexée par les entiers négatifs, par rapport à la filtration $\mathcal{G}_k = \mathcal{F}_{t_{-k}}$. Comme il est clair que

$\sup_k E[|Y_k|] < \infty$, on conclut d'après le résultat "discret" rappelé immédiatement avant le théorème que la suite X_{t_n} converge dans L^1 vers X_{t+}. En particulier $X_{t+} \in L^1$.

Grâce à la convergence dans L^1, on peut passer à la limite dans l'inégalité $X_t \geq E[X_{t_n} \mid \mathcal{F}_t]$, et on trouve

$$X_t \geq E[X_{t_+} \mid \mathcal{F}_t].$$

De plus, toujours grâce à la convergence dans L^1, on a $E[X_{t+}] = \lim E[X_{t_n}]$ et donc si la fonction $s \longrightarrow E[X_s]$ est continue à droite on doit avoir $E[X_{t+}] = E[X_t]$. Clairement, l'inégalité précédente n'est alors possible que si $X_t = E[X_{t_+} \mid \mathcal{F}_t]$.

Nous avons déjà remarqué que X_{t+} est $\mathcal{F}_{t+}$-mesurable. Soit ensuite $s < t$ et soit aussi une suite (s_n) dans D telle que $s_n \downarrow\downarrow s$. On peut bien sûr supposer $s_n \leq t_n$. Alors X_{s_n} converge vers X_{s+} dans L^1, et donc, si $A \in \mathcal{F}_{s+}$,

$$E[X_{s+}\mathbf{1}_A] = \lim_{n\to\infty} E[X_{s_n}\mathbf{1}_A] \geq \lim_{n\to\infty} E[X_{t_n}\mathbf{1}_A] = E[X_{t+}\mathbf{1}_A] = E[E[X_{t+} \mid \mathcal{F}_{s+}]\mathbf{1}_A],$$

ce qui entraîne $X_{s+} \geq E[X_{t+} \mid \mathcal{F}_{s+}]$. Enfin, si X est une martingale, on peut remplacer dans le calcul précédent l'inégalité $\geq$ par une égalité. $\square$

Théorème 3.9 *Supposons que la filtration $(\mathcal{F}_t)$ satisfait les conditions habituelles. Si $X = (X_t, t \geq 0)$ est une surmartingale et si la fonction $t \longrightarrow E[X_t]$ est continue à droite, alors X a une modification, qui est aussi une $(\mathcal{F}_t)$-surmartingale, dont les trajectoires sont continues à droite avec des limites à gauche en tout point (càdlàg).*

Démonstration. Fixons un ensemble D comme dans le Théorème 3.8. Soit N l'ensemble négligeable en dehors duquel la fonction $s \longrightarrow X_s(\omega)$ a le long de D des limites à droite et à gauche en tout $t \in \mathbb{R}_+$. Posons alors

$$Y_t(\omega) = \begin{cases} X_{t+}(\omega) & \text{si } \omega \notin N \\ 0 & \text{si } \omega \in N. \end{cases}$$

Les trajectoires de Y sont alors continues à droite: en effet, si $\omega \notin N$ on a pour $t \geq 0$ et $\varepsilon > 0$

$$\sup_{[t,t+\varepsilon[} Y_s(\omega) \leq \sup_{]t,t+\varepsilon[\cap D} X_s(\omega),$$
$$\inf_{[t,t+\varepsilon[} Y_s(\omega) \geq \inf_{]t,t+\varepsilon[\cap D} X_s(\omega),$$

et

$$\lim_{\varepsilon\downarrow 0}\Big(\sup_{]t,t+\varepsilon[\cap D} X_s(\omega)\Big) = \lim_{\varepsilon\downarrow 0}\Big(\inf_{]t,t+\varepsilon[\cap D} X_s(\omega)\Big) = X_{t+}(\omega) = Y_t(\omega).$$

Le même argument montre que les trajectoires de Y ont des limites à gauche. Enfin, comme $\mathcal{F}_{t+} = \mathcal{F}_t$ on a d'après le Théorème 3.8 (ii),

$$X_t = E[X_{t+} \mid \mathcal{F}_t] = E[X_{t+} \mid \mathcal{F}_{t+}] = X_{t+} = Y_t, \quad \text{p.s.}$$

puisque X_{t+} est $\mathcal{F}_{t+}$-mesurable. On voit ainsi que Y est une modification de X. Puisque la filtration $(\mathcal{F}_t)$ est complète, il est clair que Y_t est $\mathcal{F}_t$-mesurable et il est aussi immédiat que Y est une $(\mathcal{F}_t)$-surmartingale. $\square$

Remarque. Si la filtration est quelconque, on peut toujours appliquer le théorème précédent à l'augmentation habituelle de $(\mathcal{F}_t)$ et au processus (X_{t+}) du Théorème 3.8 (remarquer que l'application $t \to E[X_{t+}]$ est toujours continue à droite). On trouve ainsi que (X_{t+}) a une modification càdlàg.

En particulier, si on sait à l'avance que les trajectoires de X sont continues à droite, on en déduit que p.s. elles ont aussi des limites à gauche (une modification continue à droite est unique à indistinguabilité près).

3.4 Théorèmes d'arrêt

Nous commençons par un théorème de convergence analogue au résultat discret rappelé dans la partie précédente.

Théorème 3.10 *Soit X une surmartingale continue à droite et bornée dans L^1 (ce qui équivaut à $\sup_{t\geq 0} E[X_{t-}] < \infty$). Alors, il existe une v.a. $X_{\infty-} \in L^1$ telle que*

$$\lim_{t\to\infty} X_t = X_{\infty-}, \quad \text{p.s.}$$

Démonstration. Soit D un sous-ensemble dénombrable dense de $\mathbb{R}_+$. On déduit de la Proposition 3.7 que, pour tous $a < b$,

$$E[M_{ab}^X(D)] \leq \frac{1}{b-a} \sup_{t\geq 0} E[(X_t - a)_-] < \infty.$$

Donc, p.s. pour tout rationnels $a < b$ on a $M_{ab}^X(D) < \infty$. Comme dans la preuve du Théorème 3.8, on en déduit que

$$\lim_{D\ni t\to\infty} X_t$$

existe p.s. (et cette limite est dans L^1 d'après le lemme de Fatou). La continuité à droite des trajectoires permet de supprimer la restriction $t \in D$ dans la limite précédente. $\square$

Définition 3.6 *Une surmartingale $(X_t, t \geq 0)$ est dite fermée par une v.a. $X_\infty \in L^1$ si pour tout $t \geq 0$,*

$$X_t \geq E[X_\infty \mid \mathcal{F}_t].$$

Une martingale $(X_t, t \geq 0)$ est dite fermée (comme martingale) par $X_\infty \in L^1$ si pour tout $t \geq 0$,

$$X_t = E[X_\infty \mid \mathcal{F}_t].$$

Remarquons qu'une martingale peut être fermée en tant que surmartingale sans l'être en tant que martingale. Par exemple une martingale positive est toujours fermée comme surmartingale par 0, mais n'est pas nécessairement fermée comme martingale. Rappelons qu'une martingale à temps discret $(Y_n, n \in \mathbb{N})$ est fermée (i.e. on peut écrire comme ci-dessus $Y_n = E[Y_\infty \mid \mathcal{G}_n]$ avec $Y_\infty \in L^1$) si et seulement si elle est uniformément intégrable, et cela équivaut encore à dire que Y_n converge p.s. et dans L^1. On a le résultat analogue à temps continu.

Proposition 3.11 *Soit $(X_t, t \geq 0)$ une martingale continue à droite. Alors il y a équivalence entre*

(i) *X est fermée;*

(ii) *X_t converge p.s. et dans L^1 vers $X_{\infty-}$;*

(iii) *la famille $(X_t, t \geq 0)$ est uniformément intégrable.*

Démonstration. L'implication (i)⇒(iii) est facile puisque si $Z \in L^1$, la famille de v.a. $\{E[Z \mid \mathcal{G}], \mathcal{G} \text{ sous} - \text{tribu de } \mathcal{A}\}$ est uniformément intégrable. Si (iii) est vrai, la Proposition 3.10 entraîne que X_t converge p.s. vers $X_{\infty-}$, donc aussi dans L^1 par uniforme intégrabilité. Enfin, si (ii) est vérifié, on peut passer à la limite $t \to \infty$ dans L^1 dans l'égalité $X_s = E[X_t \mid \mathcal{F}_s]$ et on trouve $X_s = E[X_{\infty-} \mid \mathcal{F}_s]$. □

Si une surmartingale continue à droite $(X_t, t \geq 0)$ est fermée par X_∞, alors un argument de convexité donne $E[(X_t)^-] \leq E[(X_\infty)^-]$, et donc, d'après la Proposition 3.10, X_t converge p.s. vers $X_{\infty-}$. On vérifie alors que X est fermée par $X_{\infty-}$ (passer à la limite $t \to \infty$ dans l'inégalité $X_s - E[X_\infty \mid \mathcal{F}_s] \geq E[X_t - E[X_\infty \mid \mathcal{F}_t] \mid \mathcal{F}_s]$ en utilisant le lemme de Fatou).

Le théorème d'arrêt pour une surmartingale discrète $(Y_k, k \in \mathbb{N})$, relativement à une filtration $(\mathcal{G}_k)$, fermée par Y_∞, énonce que si S et T sont deux temps d'arrêt tels que $S \leq T$ on a $Y_S, Y_T \in L^1$, et $Y_S \geq E[Y_T \mid \mathcal{G}_S]$ (avec la convention $Y_T = Y_\infty$ sur $\{T = \infty\}$). L'analogue continu de ce théorème est le résultat le plus important de cette partie.

Théorème 3.12 (a) *Soit X une surmartingale continue à droite et fermée par une v.a. $\mathcal{F}_\infty$-mesurable X_∞. Soient S et T deux temps d'arrêt avec $S \leq T$. Alors X_S et X_T sont dans L^1 et*

$$X_S \geq E[X_T \mid \mathcal{F}_S],$$

avec la convention $X_T = X_\infty$ sur $\{T = \infty\}$.
(b) *Soit X une martingale continue à droite fermée par une v.a. $\mathcal{F}_\infty$-mesurable X_∞. Alors, si S et T sont deux temps d'arrêt avec $S \leq T$ on a*

$$X_S = E[X_T \mid \mathcal{F}_S],$$

avec la même convention que ci-dessus. En particulier, $X_S = E[X_\infty \mid \mathcal{F}_S]$.

Démonstration. (a) Pour tout $n \geq 1$, posons

$$D_n = \{\frac{k}{2^n}, k \in \mathbb{N}\} \cup \{+\infty\},$$

puis $T_n = \inf\{t \in D_n, t > T\}$. D'après la Proposition 3.4, les T_n forment une suite de t.a. qui décroît vers T.

Considérons pour n fixé les deux variables T_n et T_{n+1}, qu'on peut voir comme deux temps d'arrêt de la filtration discrète $(\mathcal{F}_t, t \in D_{n+1})$. Puisque $T_{n+1} \leq T_n$, le théorème d'arrêt appliqué à la surmartingale discrète $(X_t, t \in D_{n+1})$ (qui est fermée par X_∞), dans la filtration discrète $(\mathcal{F}_t, t \in D_{n+1})$, entraîne que

$$X_{T_{n+1}} \geq E[X_{T_n} \mid \mathcal{F}_{T_{n+1}}].$$

On obtient ainsi que $Y_k = X_{T_{-k}}$ pour $k \leq 0$ est une surmartingale indexée par les entiers négatifs relativement à la filtration $(\mathcal{F}_{T_{-k}}, k \in -\mathbb{N})$. De plus on a aussi $\sup_k E[Y_k] = \sup_n E[X_{T_n}] \leq E[X_0]$, toujours d'après le théorème d'arrêt discret. D'après le résultat discret rappelé avant le Théorème 3.8, on conclut que X_{T_n} converge p.s. et dans L^1. La limite est forcément X_T par continuité à droite, et en particulier $X_T \in L^1$.

Au temps d'arrêt S, on associe de même la suite $S_n \downarrow S$, et X_{S_n} converge vers X_S p.s. et dans L^1. Puisque $S_n \leq T_n$ le théorème d'arrêt discret donne

$$X_{S_n} \geq E[X_{T_n} \mid \mathcal{F}_{S_n}]$$

donc, pour $A \in \mathcal{F}_S \subset \mathcal{F}_{S_n}$,

$$E[X_{S_n} \mathbf{1}_A] \geq E[X_{T_n} \mathbf{1}_A].$$

En passant à la limite L^1 quand $n \to \infty$, il vient

$$E[X_S \mathbf{1}_A] \geq E[X_T \mathbf{1}_A],$$

pour tout $A \in \mathcal{F}_S$. Comme X_S est $\mathcal{F}_S$-mesurable ($\mathbf{1}_{\{S<\infty\}} X_S$ l'est d'après le Théorème 3.3 et $\mathbf{1}_{\{S=\infty\}} X_S$ l'est parce que X_∞ est $\mathcal{F}_\infty$-mesurable), cette inégalité suffit pour conclure que $X_S \geq E[X_T \mid \mathcal{F}_S]$.
(b) Il suffit d'appliquer la partie (a) à X et à $-X$. □

Corollaire 3.13 *Soit X une surmartingale (resp. une martingale) continue à droite et soient $S \leq T$ deux temps d'arrêt bornés. Alors*

$$X_S \geq E[X_T \mid \mathcal{F}_S] \qquad (\textit{resp. } X_S = E[X_T \mid \mathcal{F}_S]).$$

Démonstration. Soit $a \geq 0$ tel que $T \leq a$. On applique le théorème précédent à $(X_{t\wedge a}, t \geq 0)$ qui est bien sûr fermée par X_a. □

Corollaire 3.14 *Soit X une martingale continue à droite uniformément intégrable, et soit T un temps d'arrêt. Alors le processus $(X_{t\wedge T}, t \geq 0)$ est aussi une martingale uniformément intégrable, et*

$$X_{t\wedge T} = E[X_T \mid \mathcal{F}_t],$$

avec la convention $X_T = X_{\infty-}$ sur $\{T = \infty\}$.

Démonstration. Il suffit d'établir la dernière assertion. Rappelons que $X_T \in L^1$ d'après le Théorème 3.12. D'après ce théorème, on a aussi

$$X_{t\wedge T} = E[X_T \mid \mathcal{F}_{t\wedge T}].$$

Puisque $X_T \mathbf{1}_{\{T<\infty\}}$ est $\mathcal{F}_T$-mesurable, on sait que $X_T \mathbf{1}_{\{T\leq t\}}$ est $\mathcal{F}_t$-mesurable, et aussi $\mathcal{F}_T$-mesurable, donc $\mathcal{F}_{t\wedge T}$-mesurable. Il en découle que

$$E[X_T \mathbf{1}_{\{T\leq t\}} \mid \mathcal{F}_{t\wedge T}] = X_T \mathbf{1}_{\{T\leq t\}} = E[X_T \mathbf{1}_{\{T\leq t\}} \mid \mathcal{F}_t].$$

Pour compléter la preuve, il reste à voir que

$$E[X_T \mathbf{1}_{\{T>t\}} \mid \mathcal{F}_{t\wedge T}] = E[X_T \mathbf{1}_{\{T>t\}} \mid \mathcal{F}_t]. \tag{3.1}$$

Or si $A \in \mathcal{F}_t$, $A \cap \{T > t\} \in \mathcal{F}_T \cap \mathcal{F}_t = \mathcal{F}_{t\wedge T}$, et donc

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

ce qui entraîne (3.1) et complète la preuve. □

Si on suppose seulement que X est une martingale continue à droite, alors on peut appliquer le corollaire ci-dessus à la martingale u. i. $(X_{t\wedge a}, t \geq 0)$. On trouve que pour tout temps d'arrêt T, $(X_{t\wedge T\wedge a}, t \geq 0)$ est une martingale. Comme on peut prendre a aussi grand qu'on le désire, cela signifie que $(X_{t\wedge T}, t \geq 0)$ est une martingale.

Chapitre 4

Semimartingales Continues

Les semimartingales continues constituent la classe générale de processus à trajectoires continues pour laquelle on peut développer une théorie de l'intégrale stochastique, qui sera traitée dans le chapitre suivant. Par définition, une semimartingale est la somme d'une martingale (locale) et d'un processus à variation finie. Dans ce chapitre nous étudions séparément ces deux classes de processus. En particulier, nous introduisons la notion de variation quadratique d'une martingale, qui jouera plus tard un rôle fondamental.

4.1 Processus à variation finie

4.1.1 Fonctions à variation finie

Définition 4.1 *Soit $T > 0$. Une fonction continue $a : [0,T] \longrightarrow \mathbb{R}$ telle que $a(0) = 0$ est dite à variation finie s'il existe une mesure signée (i.e. différence de deux mesures positives finies) μ telle que $a(t) = \mu([0,t])$ pour tout $t \in [0,T]$.*

La mesure μ est alors déterminée de façon unique. La décomposition de μ comme différence de deux mesures positives finies n'est bien sûr pas unique, mais il existe une seule décomposition $\mu = \mu_+ - \mu_-$ telle que μ_+ et μ_- soient deux mesures positives finies portées par des boréliens disjoints. Pour obtenir l'existence d'une telle décomposition, on peut partir d'une décomposition quelconque $\mu = \mu_1 - \mu_2$, poser $\nu = \mu_1 + \mu_2$ puis utiliser le théorème de Radon-Nikodym pour écrire

$$\mu_1(dt) = h_1(t)\nu(dt), \quad \mu_2(dt) = h_2(t)\nu(dt).$$

Ensuite, si $h(t) = h_1(t) - h_2(t)$ on a

$$\mu(dt) = h(t)\nu(dt) = h(t)^+\nu(dt) - h(t)^-\nu(dt)$$

ce qui donne la décomposition $\mu = \mu_+ - \mu_-$ avec $\mu_+(dt) = h(t)^+\nu(dt)$, $\mu_-(dt) = h(t)^-\nu(dt)$, les mesures μ_+ et μ_- étant portées respectivement par les boréliens disjoints $D_+ = \{t, h(t) > 0\}$ et $D_- = \{t, h(t) < 0\}$. L'unicité de la décomposition $\mu = \mu_+ - \mu_-$ découle du fait que l'on a nécessairement

$$\mu_+(A) = \sup\{\mu(C), C \subset A,\, C \text{ borélien}\}.$$

On note $|\mu|$ la mesure positive $|\mu| = \mu_+ + \mu_-$. Remarquons que la dérivée de Radon-Nikodym de μ par rapport à $|\mu|$ est

$$\frac{d\mu}{d|\mu|} = \mathbf{1}_{D_+} - \mathbf{1}_{D_-}.$$

On a $a(t) = \mu_+([0,t]) - \mu_-([0,t])$, ce qui montre que la fonction a est différence de deux fonctions croissantes continues et nulles en 0 (la continuité de a entraîne que μ n'a pas d'atomes, et il en va alors de même pour μ_+ et μ_-). Inversement une différence de fonctions croissantes (continues et nulles en 0) est aussi à variation finie au sens précédent. En effet, cela découle du fait bien connu que la formule $g(t) = \nu([0,t])$ établit une bijection entre les fonctions croissantes continues à droite $g : [0,T] \longrightarrow \mathbb{R}_+$ et les mesures positives finies sur $[0,T]$.

Soit $f : [0,T] \longrightarrow \mathbb{R}$ une fonction mesurable telle que $\int_{[0,T]} |f(s)|\,|\mu|(ds) < \infty$. On note

$$\int_0^T f(s)\,da(s) = \int_{[0,T]} f(s)\,\mu(ds),$$
$$\int_0^T f(s)\,|da(s)| = \int_{[0,T]} f(s)\,|\mu|(ds).$$

On vérifie facilement l'inégalité

$$\left|\int_0^T f(s)\,da(s)\right| \leq \int_0^T |f(s)|\,|da(s)|.$$

Remarquons de plus que la fonction $t \longrightarrow \int_0^t f(s)\,da(s)$ est aussi à variation finie (la mesure associée est simplement $\mu'(ds) = f(s)\mu(ds)$).

Proposition 4.1 *Pour tout* $t \in [0,T]$,

$$|\mu|([0,t]) = \sup\left\{\sum_{i=1}^{p} |a(t_i) - a(t_{i-1})|\right\},$$

où le supremum porte sur toutes les subdivisions $0 = t_0 < t_1 < \cdots < t_p = t$ *de* $[0,t]$.

Remarque. Dans la présentation habituelle des fonctions à variation finie, on part de la propriété que le supremum ci-dessus est fini.
Démonstration. Il suffit clairement de traiter le cas $t = T$. L'inégalité $\geq$ est très facile puisque

$$|a(t_i) - a(t_{i-1})| = |\mu(]t_{i-1}, t_i])| \leq |\mu|(]t_{i-1}, t_i]).$$

Pour l'autre inégalité, on peut montrer plus précisément que pour toute suite $0 = t_0^n < t_1^n < \cdots < t_{p_n}^n = T$ de subdivisions emboitées de $[0, T]$ de pas tendant vers 0 on a

$$\lim_{n\to\infty} \sum_{i=1}^{p_n} |a(t_i^n) - a(t_{i-1}^n)| = |\mu|([0, T]).$$

Considérons pour simplifier les subdivisions dyadiques $t_i^n = i2^{-n}T$, $0 \leq i \leq 2^n$. On utilise un argument de martingales en introduisant l'espace de probabilité $\Omega = [0, T]$ muni de la tribu borélienne $\mathcal{B} = \mathcal{B}([0, T])$ et de la probabilité $P(ds) = (|\mu([0, T])|^{-1})|\mu|(ds)$. Introduisons sur cet espace la filtration discrète $(\mathcal{B}_n)$ telle que pour chaque n, $\mathcal{B}_n$ soit la tribu engendrée par les intervalles $](i-1)2^{-n}T, i2^{-n}T]$, $1 \leq i \leq 2^n$. Posons enfin

$$X(s) = \frac{d\mu}{d|\mu|}(s) = \mathbf{1}_{D_+}(s) - \mathbf{1}_{D_-}(s),$$

et pour chaque $n \geq 0$

$$X_n = E[X \mid \mathcal{B}_n].$$

Les propriétés de l'espérance conditionnelle montrent que X_n doit être constante sur chaque intervalle $](i-1)2^{-n}T, i2^{-n}T]$ et vaut sur cet intervalle

$$\frac{\mu(](i-1)2^{-n}T, i2^{-n}T])}{|\mu|(](i-1)2^{-n}T, i2^{-n}T])} = \frac{a(i2^{-n}T) - a((i-1)2^{-n}T)}{|\mu|(](i-1)2^{-n}T, i2^{-n}T])}.$$

D'autre part, il est clair que la suite (X_n) est une martingale fermée (relativement à la filtration $(\mathcal{B}_n)$) et parce que X est mesurable par rapport à $\mathcal{B} = \bigvee_n \mathcal{B}_n$ cette martingale converge p.s. et dans L^1 vers X. En particulier,

$$\lim_{n\to\infty} E[|X_n|] = E[|X|] = 1,$$

cette dernière égalité étant claire puisque $|X(s)| = 1$, μ p.p. Le résultat annoncé en découle puisque d'après ci-dessus

$$E[|X_n|] = (|\mu|([0, T])^{-1}) \sum_{i=1}^{2^n} |a(i2^{-n}T) - a((i-1)2^{-n}T)|.$$

Lemme 4.2 *Si* $f : [0,T] \longrightarrow \mathbb{R}$ *est une fonction continue et si* $0 = t_0^n < t_1^n < \cdots < t_{p_n}^n = T$ *est une suite de subdivisions de* $[0,T]$ *de pas tendant vers* 0 *on a*

$$\int_0^T f(s)\,da(s) = \lim_{n\to\infty} \sum_{i=1}^{p_n} f(t_{i-1}^n)\,(a(t_i^n) - a(t_{i-1}^n)).$$

Démonstration. Soit f_n la fonction définie par $f_n(s) = f(t_{i-1}^n)$ si $s \in]t_{i-1}^n, t_i^n]$. Alors,

$$\sum_{i=1}^{p_n} f(t_{i-1}^n)\,(a(t_i^n) - a(t_{i-1}^n)) = \int_{[0,T]} f_n(s)\,\mu(ds),$$

et le résultat voulu en découle par convergence dominée. □

On dira qu'une fonction continue $a : \mathbb{R}_+ \longrightarrow \mathbb{R}$ est à variation finie sur $\mathbb{R}_+$ si la restriction de a à $[0,T]$ est à variation finie, pour tout $T > 0$. Il est alors facile d'étendre les définitions précédentes. En particulier, on peut définir $\int_0^\infty f(s)da(s)$ pour toute fonction f telle que $\int_0^\infty |f(s)||da(s)| = \sup_{T>0} \int_0^T |f(s)||da(s)| < \infty$.

4.1.2 Processus à variation finie

On se place maintenant sur un espace de probabilité filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ satisfaisant les conditions habituelles.

Définition 4.2 *Un processus à variation finie* $A = (A_t, t \geq 0)$ *est un processus adapté dont toutes les trajectoires sont à variation finie au sens de la définition précédente. Le processus* A *est appelé processus croissant si de plus les trajectoires de* A *sont croissantes.*

Remarque. En particulier on a $A_0 = 0$ et les trajectoires de A sont continues.

Proposition 4.3 *Soit* A *un processus à variation finie et soit* H *un processus progressif tel que*

$$\forall t \geq 0\ \forall \omega\ ,\ \int_0^t |H_s(\omega)|\,|dA_s(\omega)| < \infty.$$

Alors le processus $H \cdot A$ *défini par*

$$(H \cdot A)_t = \int_0^t H_s\,dA_s$$

est aussi un processus à variation finie.

Démonstration. D'après des remarques précédentes, il est clair que les trajectoires de $H \cdot A$ sont à variation finie. Il reste donc à montrer que $H \cdot A$ est adapté. Pour cela, il suffit de voir que, si $h : [0,t] \times \Omega \longrightarrow \mathbb{R}$ est mesurable pour la tribu produit $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$ et si $\int_0^t |h(s,\omega)||dA_s(\omega)|$ est fini pour tout ω, alors la variable $\int_0^t h(s,\omega)dA_s(\omega)$ est $\mathcal{F}_t$-mesurable.

Si $h(s,\omega) = \mathbf{1}_{]u,v]}(s)\mathbf{1}_\Gamma(\omega)$ avec $]u,v] \subset [0,t]$ et $\Gamma \in \mathcal{F}_t$, le résultat est évident. On passe ensuite au cas $h = \mathbf{1}_G$, $G \in \mathcal{B}([0,t]) \otimes \mathcal{F}_t$ par un argument de classe monotone. Enfin, dans le cas général, on observe qu'on peut toujours écrire h comme limite ponctuelle d'une suite de fonctions étagées h_n telles que pour tout n on ait $|h_n| \le |h|$, ce qui assure que $\int_0^t h_n(s,\omega)dA_s(\omega) \longrightarrow \int_0^t h(s,\omega)dA_s(\omega)$ par convergence dominée. $\square$

Remarques. (i) Il arrive souvent qu'on ait l'hypothèse plus faible

$$\text{p.s.} \quad \forall t \ge 0, \qquad \int_0^t |H_s|\,|dA_s| < \infty.$$

On peut encore définir $H \cdot A$ en convenant que, sur l'ensemble de probabilité nulle où $\int_0^t |H_s|\,|dA_s|$ devient infini, on prend $(H \cdot A)_t(\omega) = 0$ pour tout t. Le processus $(H \cdot A)$ ainsi défini reste adapté parce que la filtration est supposée complète.

(ii) Sous des hypothèses convenables, on a la propriété d'associativité $K \cdot (H \cdot A) = (KH) \cdot A$.

Un cas particulier important est celui où $A_t = t$. Si H_s est un processus progressif tel que

$$\text{p.s. } \forall t \ge 0 \qquad \int_0^t |H_s|\,ds < \infty,$$

le processus $\int_0^t H_s\,ds$ est un processus à variation finie.

4.2 Martingales locales

Comme dans la partie précédente, nous nous plaçons sur un espace de probabilité filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ satisfaisant les conditions habituelles. Si T est un temps d'arrêt et $X = (X_t, t \ge 0)$ est un processus à trajectoires continues, on note X^T le processus arrêté $X_t^T = X_{t \wedge T}$ pour tout $t \ge 0$.

Définition 4.3 *Un processus adapté à trajectoires continues $M = (M_t, t \ge 0)$ tel que $M_0 = 0$ p.s. est une martingale locale (continue) s'il existe une suite croissante $(T_n, n \in \mathbb{N})$ de temps d'arrêt telle que $T_n \uparrow \infty$ et pour tout n le processus arrêté M^{T_n} est une martingale uniformément intégrable.*

Plus généralement, lorsque $M_0 \neq 0$, on dit que M est une martingale locale (continue) si $M_t = M_0 + N_t$, où le processus N est une martingale locale issue de 0.

Dans tous les cas, on dit que la suite de temps d'arrêt $T_n \uparrow \infty$ réduit M si pour tout n le processus arrêté M^{T_n} est une martingale uniformément intégrable.

Remarque. On n'impose pas dans la définition d'une martingale locale que les variables M_t soient dans L^1 (comparer avec la définition des martingales). En particulier, on voit sur la définition précédente que M_0 peut être n'importe quelle variable $\mathcal{F}_0$-mesurable.

Propriétés.

(a) Une martingale à trajectoires continues est une martingale locale (et la suite $T_n = n$ réduit M).

(b) Dans la définition d'une martingale locale (issue de 0) on peut remplacer "martingale uniformément intégrable" par "martingale" (en effet on peut ensuite remplacer T_n par $T_n \wedge n$).

(c) Si M est une martingale locale, pour tout temps d'arrêt T, M^T est une martingale locale (cf Corollaire 3.14).

(d) Si (T_n) réduit M et si S_n est une suite de temps d'arrêt telle que $S_n \uparrow \infty$, alors la suite $(T_n \wedge S_n)$ réduit aussi M.

(e) L'espace des martingales locales est un espace vectoriel (utiliser la propriété précédente).

Proposition 4.4 (i) *Une martingale locale positive M telle que $M_0 \in L^1$ est une surmartingale.*

(ii) *Une martingale locale M bornée, ou plus généralement telle qu'il existe une variable $Z \in L^1$ telle que, pour tout $t \geq 0$, $|M_t| \leq Z$, est une martingale.*

(iii) *Si M est une martingale locale avec $M_0 = 0$, la suite de temps d'arrêt*

$$T_n = \inf\{t \geq 0, |M_t| = n\}$$

réduit M.

Démonstration. (i) Ecrivons $M_t = M_0 + N_t$ et soit (T_n) une suite de temps d'arrêt qui réduit N. Alors, si $s \leq t$, on a pour tout n,

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

En ajoutant des deux cotés la variable M_0 (qui est $\mathcal{F}_0$-mesurable et dans L^1), on trouve

$$M_{s\wedge T_n} = E[M_{t\wedge T_n} \mid \mathcal{F}_s]. \tag{4.1}$$

Puisque M est à valeurs positives, on peut faire tendre n vers ∞ et appliquer le lemme de Fatou (pour les espérances conditionnelles) qui donne

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

En prenant $s = 0$, on voit que $E[M_t] \leq E[M_0] < \infty$, donc $M_t \in L^1$ pour tout $t \geq 0$. L'inégalité précédente montre alors que M est une surmartingale.

(ii) Si M est bornée (ou plus généralement dominée par une variable intégrable), le même raisonnement que ci-dessus donne pour $s \leq t$

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

Or par convergence dominée la suite $M_{t\wedge T_n}$ converge dans L^1 vers M_t, et donc on peut passer à la limite $n \to \infty$ pour trouver $M_s = E[M_t \mid \mathcal{F}_s]$.

(iii) C'est une conséquence immédiate de (ii) $\square$

Théorème 4.5 *Soit M une martingale locale (continue) issue de 0. Alors si M est un processus à variation finie, M est indistinguable de 0.*

Démonstration. Supposons que M est un processus à variation finie et posons pour tout $n \in \mathbb{N}$,

$$\tau_n = \inf\{t \geq 0, \int_0^t |dM_s| \geq n\}.$$

Les temps τ_n sont des temps d'arrêt d'après le Théorème 3.5 (remarquer que le processus $\int_0^t |dM_s|$ est continu et adapté). Fixons $n \geq 1$ et posons $N = M^{\tau_n}$. Alors N est une martingale locale telle que $\int_0^\infty |dN_s| \leq n$, et donc en particulier $|N_t| \leq n$. D'après la proposition 4.4, N est une (vraie) martingale bornée. Ensuite, soit $t > 0$ et soit $0 = t_0 < t_1 < \cdots < t_p = t$ une subdivision de $[0, t]$. Alors,

$$\begin{aligned}
E[N_t^2] &= \sum_{i=1}^p E[N_{t_i}^2 - N_{t_{i-1}}^2] \\
&= \sum_{i=1}^p E[(N_{t_i} - N_{t_{i-1}})^2] \\
&\leq E\Big[\Big(\sup_{1\leq i\leq p} |N_{t_i} - N_{t_{i-1}}|\Big) \sum_{i=1}^p |N_{t_i} - N_{t_{i-1}}|\Big] \\
&\leq n\, E\Big[\sup_{1\leq i\leq p} |N_{t_i} - N_{t_{i-1}}|\Big]
\end{aligned}$$

en utilisant la proposition 4.1. On applique l'inégalité précédente à une suite $0 = t_0^k < t_1^k < \cdots < t_{p_k}^k = t$ de subdivisions de $[0, t]$ de pas tendant vers 0. En utilisant la continuité des trajectoires, et le fait que N est bornée (pour justifier la convergence dominée), on a

$$\lim_{k\to\infty} E\Big[\sup_{1\leq i\leq p_k} |N_{t_i^k} - N_{t_{i-1}^k}|\Big] = 0.$$

On conclut alors que $E[N_t^2] = 0$, soit $E[M_{t\wedge\tau_n}^2] = 0$. En faisant tendre n vers ∞ on obtient $E[M_t^2] = 0$. $\square$

4.3 Variation quadratique d'une martingale locale

Le théorème ci-dessous joue un rôle très important dans la suite du cours.

Théorème 4.6 *Soit $M = (M_t, t \geq 0)$ une martingale locale (continue). Il existe un processus croissant, noté $(\langle M, M\rangle_t, t \geq 0)$ unique à indistinguabilité près, tel que $M_t^2 - \langle M, M\rangle_t$ soit une martingale locale continue. De plus, pour tout $T > 0$, si $0 = t_0^n < t_1^n < \cdots < t_{p_n}^n = T$ est une suite de subdivisions emboitées de $[0, T]$ de pas tendant vers 0, on a*

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

au sens de la convergence en probabilité. Le processus $\langle M, M\rangle$ est appelé la variation quadratique de M.

Remarques. (i) Si $M = B$ est un mouvement brownien, la Proposition 2.10 montre que $\langle B, B\rangle_t = t$.

(ii) Pour la dernière assertion, il n'est en fait pas nécessaire de supposer que les subdivisions soient emboitées.

Démonstration. L'unicité découle du théorème 4.5. En effet, si A et A' sont deux processus croissants satisfaisant la condition donnée, le processus $A_t - A'_t = (M_t^2 - A'_t) - (M_t^2 - A_t)$ doit être à la fois une martingale locale et un processus à variation finie.

Pour l'existence considérons d'abord le cas où $M_0 = 0$ et M est bornée (donc en particulier est une vraie martingale). Donnons dans ce cas les grandes étapes de la preuve. On se fixe $T > 0$ et $0 = t_0^n < t_1^n < \cdots < t_{p_n}^n = T$ une suite de subdivisions emboitées de $[0, T]$ de pas tendant vers 0. Pour chaque n, le processus X^n défini par

$$X_t^n = \sum_{i=1}^{p_n} M_{t_{i-1}^n}(M_{t_i^n \wedge t} - M_{t_{i-1}^n \wedge t})$$

est une martingale continue. De plus on vérifie par des calculs directs assez fastidieux utilisant la propriété de martingale (exercice!) que

$$\lim_{n,m\to\infty} E[(X_T^n - X_T^m)^2] = 0.$$

Via l'inégalité de Doob (Proposition 3.6), cela entraîne que

$$\lim_{n,m\to\infty} E\Big[\sup_{t\leq T}(X_t^n - X_t^m)^2\Big] = 0.$$

Il est facile d'en déduire qu'une sous-suite X^{n_k} converge p.s. uniformément sur $[0,T]$ vers un processus $Y=(Y_t, 0\le t\le T)$ à trajectoires continues. En passant à la limite L^2 dans l'égalité de martingale pour X^n, on voit que Y est une martingale continue (sur l'intervalle $[0,T]$). Ensuite, un calcul simple montre que pour tout $n\ge 1$, pour tout $j\in\{1,\dots,p_n\}$,

$$M_{t_j^n}^2 - 2X_{t_j^n}^n = \sum_{i=1}^{j}(M_{t_i^n}-M_{t_{i-1}^n})^2, \tag{4.2}$$

Donc le processus $M_t^2-2X_t^n$ est croissant le long de la subdivision $(t_i^n, 0\le i\le p_n)$. En passant à la limite $n\to\infty$, on voit que la fonction $t\longrightarrow M_t^2-2Y_t$ doit être (p.s.) croissante sur $[0,T]$. On pose, pour $t\in[0,T]$, $\langle M,M\rangle_t = M_t^2-2Y_t$ (on peut prendre $\langle M,M\rangle_t\equiv 0$ sur l'ensemble de probabilité nulle où la fonction $t\longrightarrow M_t^2-2Y_t$ n'est pas croissante). Alors, $\langle M,M\rangle$ est un processus croissant et $M_t^2-\langle M,M\rangle_t = 2Y_t$ est une martingale, sur l'intervalle de temps $[0,T]$. Il est facile d'étendre la définition de $\langle M,M\rangle_t$ à tout $t\in\mathbb{R}_+$: on applique ce qui précède avec $T=k$ pour tout entier $k\ge 1$, en remarquant que le processus obtenu avec $T=k$ doit être la restriction à $[0,k]$ de celui obtenu avec $T=k+1$, à cause de l'argument d'unicité. Le processus $\langle M,M\rangle_t$ ainsi étendu satisfait manifestement la première propriété de l'énoncé.

La partie unicité montre aussi que le processus $\langle M,M\rangle_t$ ne dépend pas de la suite de subdivisions choisie pour le construire. On déduit alors de (4.2) (avec $j=p_n$) que pour tout $T>0$, pour n'importe quelle suite de subdivisions emboitées de $[0,T]$ de pas tendant vers 0, on a

$$\lim_{n\to\infty}\sum_{j=1}^{p_n}(M_{t_j^n}-M_{t_{j-1}^n})^2 = \langle M,M\rangle_T$$

dans L^2. Cela achève la preuve du théorème dans le cas borné.

Considérons maintenant le cas général. En écrivant $M_t = M_0+N_t$, donc $M_t^2 = M_0^2+2\,M_0N_t+N_t^2$, et en remarquant que M_0N_t est une martingale locale, on se ramène facilement au cas où $M_0=0$. On pose alors

$$T_n = \inf\{t\ge 0, |M_t|\ge n\}$$

et on peut appliquer ce qui précède aux martingales bornées M^{T_n}. Notons $A^n = \langle M^{T_n}, M^{T_n}\rangle$. Grâce à la partie unicité, on voit facilement que les processus $A^{n+1}_{t\wedge T_n}$ et A^n_t sont indistinguables. On en déduit qu'il existe un processus croissant A tel que, pour tout n, $A_{t\wedge T_n}$ et A_t^n soient indistinguables. Par construction, $M^2_{t\wedge T_n}-A_{t\wedge T_n}$ est une martingale, ce qui entraîne précisément que $M_t^2-A_t$ est une martingale locale. On prend $\langle M,M\rangle_t = A_t$ et cela termine la preuve de la partie existence.

Enfin, la deuxième assertion du théorème est vraie si on remplace M et $\langle M,M\rangle_T$ par M^{T_n} et $\langle M,M\rangle_{T\wedge T_n}$ (même avec convergence L^2). Il suffit alors de faire tendre

n vers ∞ en observant que, pour tout $T > 0$, $P[T \leq T_n]$ converge vers 1 quand $n \to \infty$. $\square$

Propriété. Si T est un temps d'arrêt on a

$$\langle M^T, M^T \rangle_t = \langle M, M \rangle_{t \wedge T}$$

simplement parce que $M^2_{t\wedge T} - \langle M, M \rangle_{t\wedge T}$ est une martingale locale comme martingale locale arrêtée.

Nous énonçons maintenant un théorème qui montre comment les propriétés d'une martingale locale sont liées à celles de sa variation quadratique. Si A est un processus croissant, A_∞ désigne de manière évidente la limite croissante de A_t quand $t \to \infty$.

Théorème 4.7 *Soit M une martingale locale avec $M_0 = 0$.*

(i) *Si M est une (vraie) martingale bornée dans L^2, alors $E[\langle M, M \rangle_\infty] < \infty$, et $M_t^2 - \langle M, M \rangle_t$ est une martingale uniformément intégrable.*

(ii) *Si $E[\langle M, M \rangle_\infty] < \infty$, alors M est une (vraie) martingale bornée dans L^2.*

(iii) *Il y a équivalence entre:*

(a) *M est une (vraie) martingale de carré intégrable ($E[M_t^2] < \infty$).*

(b) *$E[\langle M, M \rangle_t] < \infty$ pour tout $t \geq 0$.*

De plus si ces conditions sont satisfaites, $M_t^2 - \langle M, M \rangle_t$ est une martingale.

Remarque. Dans la partie (i) de l'énoncé, il est essentiel de supposer que M est une martingale, et pas seulement une martingale locale. L'inégalité de Doob utilisée dans la preuve suivante n'est pas valable pour une martingale locale !

Démonstration. (i) Si M est une martingale (continue) bornée dans L^2, l'inégalité de Doob montre que

$$E\Big[\sup_{t\geq 0} M_t^2\Big] \leq 4 \sup_{t\geq 0} E[M_t^2] < \infty.$$

En particulier, M est une martingale uniformément intégrable. On note M_∞ la limite (p.s. et dans L^2) de M_t quand $t \to \infty$. Pour tout entier $n \geq 1$, posons

$$S_n = \inf\{t \geq 0, \langle M, M \rangle_t \geq n\}.$$

Alors S_n est un temps d'arrêt et $\langle M, M \rangle_{t\wedge S_n} \leq n$. On en déduit que la martingale locale $M^2_{t\wedge S_n} - \langle M, M \rangle_{t\wedge S_n}$ est dominée par la variable intégrable $n + \sup_{t\geq 0} M_t^2$. En conséquence cette martingale locale est une vraie martingale (proposition 4.4 (ii)) et on a

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

En faisant tendre t vers ∞ (et en utilisant le théorème de convergence monotone pour le terme de gauche, le théorème de convergence dominée pour celui de droite) on trouve

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

Enfin en faisant tendre n vers ∞ on trouve, avec les mêmes arguments,

$$E[\langle M,M\rangle_\infty] = E[M_\infty^2] < \infty.$$

De plus la martingale locale $M_t^2 - \langle M,M\rangle_t$ est dominée par la variable intégrable $\sup_{t\geq 0} M_t^2 + \langle M,M\rangle_\infty$, et est donc une (vraie) martingale uniformément intégrable.

(ii) On suppose maintenant $E[\langle M,M\rangle_\infty] < \infty$. Notons $T_n = \inf\{t \geq 0, |M_t| \geq n\}$, de sorte que M^{T_n} est une vraie martingale bornée. La martingale locale $M_{t\wedge T_n}^2 - \langle M,M\rangle_{t\wedge T_n}$ est aussi une vraie martingale uniformément intégrable, car elle est dominée par la variable intégrable $n^2 + \langle M,M\rangle_\infty$. Soit S un temps d'arrêt fini p.s. D'après le théorème d'arrêt (théorème 3.12) on a

$$E[M_{S\wedge T_n}^2] = E[\langle M,M\rangle_{S\wedge T_n}],$$

d'où, à l'aide du lemme de Fatou,

$$E[M_S^2] \leq E[\langle M,M\rangle_\infty].$$

La famille $\{M_S,\, S$ temps d'arrêt fini$\}$ est bornée dans L^2 et donc uniformément intégrable. En passant à limite L^1 quand $n \to \infty$ dans l'égalité $E[M_{t\wedge T_n} \mid \mathcal{F}_s] = M_{s\wedge T_n}$ (pour $s \leq t$) on en déduit que M est une vraie martingale. De plus, M est bornée dans L^2.

(iii) Soit $a > 0$. D'après (i) et (ii), on a $E[\langle M,M\rangle_a] < \infty$ si et seulement si $M_{t\wedge a}$ est une vraie martingale de carré intégrable. L'équivalence de (a) et (b) en découle aussitôt. Enfin, si ces conditions sont remplies, (i) montre que $M_{t\wedge a}^2 - \langle M,M\rangle_{t\wedge a}$ est une vraie martingale, d'où la dernière assertion. $\square$

Corollaire 4.8 *Soit M une martingale locale continue telle que $M_0 = 0$. Alors on a $\langle M,M\rangle_t = 0$ p.s. pour tout $t \geq 0$ si et seulement si M est indistinguable de 0.*

Démonstration. Supposons $\langle M,M\rangle_t = 0$ p.s. pour tout $t \geq 0$. D'après les parties (ii) et (i) du théorème ci-dessus, M_t^2 est une martingale uniformément intégrable, d'où $E[M_t^2] = E[M_0^2] = 0$. $\square$

Crochet de deux martingales locales. Si M et N sont deux martingales locales, on pose

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

Proposition 4.9 (i) *$\langle M,N\rangle$ est l'unique (à indistinguabilité près) processus à variation finie tel que $M_tN_t - \langle M,N\rangle_t$ soit une martingale locale.*

(ii) *L'application $(M,N) \longrightarrow \langle M,N\rangle$ est bilinéaire symétrique.*

(iii) *Si* $0 = t_0^n < t_1^n < \cdots < t_{p_n}^n = t$ *est une suite de subdivisions emboitées de* $[0, t]$ *de pas tendant vers* 0*, on a*

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

en probabilité.

(iv) *Pour tout temps d'arrêt* T, $\langle M^T, N^T \rangle_t = \langle M^T, N \rangle_t = \langle M, N \rangle_{t\wedge T}$.

Démonstration. (i) découle de la caractérisation analogue dans le théorème 4.6 (et l'unicité découle du théorème 4.5). (iii) est de même une conséquence de l'assertion analogue dans le théorème 4.6. (ii) découle de (iii). Enfin, on peut voir (iv) comme une conséquence de la propriété (iii), en remarquant que cette propriété entraîne p.s.

$$\begin{aligned} &\langle M^T, N^T \rangle_t = \langle M^T, N \rangle_t = \langle M, N \rangle_t && \text{sur } \{T \geq t\}, \\ &\langle M^T, N^T \rangle_t - \langle M^T, N^T \rangle_s = \langle M^T, N \rangle_t - \langle M^T, N \rangle_s = 0 && \text{sur } \{T \leq s < t\}. \end{aligned}$$

Remarque. Une conséquence de (iv) est que $M^T(N - N^T)$ est une martingale locale, ce qui n'est pas si facile à voir directement.

Définition 4.4 *Deux martingales locales* M *et* N *sont dites orthogonales si* $\langle M, N \rangle = 0$*, ce qui équivaut à dire que le produit* MN *est une martingale locale.*

Remarque. Un $(\mathcal{F}_t)$-mouvement brownien B est un mouvement brownien qui est $(\mathcal{F}_t)$-adapté et à accroissements indépendants par rapport à $(\mathcal{F}_t)$. Un $(\mathcal{F}_t)$-mouvement brownien B est bien sûr une martingale locale, de variation quadratique $\langle B, B \rangle_t = t$. De plus, deux $(\mathcal{F}_t)$-mouvements browniens indépendants B et B' sont des martingales orthogonales. Le plus simple pour le voir est d'observer que le processus $\frac{1}{\sqrt{2}}(B_t + B'_t)$ est encore une $(\mathcal{F}_t)$-martingale locale, et d'autre part il est très facile de vérifier que c'est aussi un mouvement brownien. Donc sa variation quadratique est t et par bilinéarité cela entraîne $\langle B, B' \rangle_t = 0$.

Si M et N sont deux (vraies) martingales bornées dans L^2, on déduit facilement du théorème 4.7 que $M_tN_t - \langle M, N \rangle_t$ est une martingale uniformément intégrable. Si de plus M et N sont orthogonales, on a $E[M_tN_t] = 0$, et même $E[M_SN_S] = 0$ pour tout temps d'arrêt S.

Proposition 4.10 (Inégalité de Kunita-Watanabe) *Soient* M *et* N *deux martingales locales et* H *et* K *deux processus mesurables. Alors*

$$\int_0^\infty |H_s|\,|K_s|\,|d\langle M, N \rangle_s| \leq \Big(\int_0^\infty H_s^2\, d\langle M, M \rangle_s\Big)^{1/2} \Big(\int_0^\infty K_s^2\, d\langle N, N \rangle_s\Big)^{1/2}.$$

Démonstration. Notons $\langle M,N\rangle_s^t = \langle M,N\rangle_t - \langle M,N\rangle_s$ pour $s \le t$. On commence par remarquer que p.s. pour tous $s<t$ rationnels (donc aussi par continuité pour tous $s<t$) on a

$$|\langle M,N\rangle_s^t| \le \sqrt{\langle M,M\rangle_s^t}\sqrt{\langle N,N\rangle_s^t}.$$

En effet, cela découle immédiatement des approximations de $\langle M,M\rangle$ et $\langle M,N\rangle$ données dans le théorème 4.6 et la proposition 4.9 respectivement, ainsi que de l'inégalité de Cauchy-Schwarz. A partir de maintenant, on fixe ω tel que l'inégalité précédente soit vraie pour tous $s<t$, et on raisonne sur cette valeur de ω. On remarque qu'on a aussi

$$\int_s^t |d\langle M,N\rangle_u| \le \sqrt{\langle M,M\rangle_s^t}\sqrt{\langle N,N\rangle_s^t}. \tag{4.3}$$

En effet, il suffit d'utiliser la proposition 4.1 et de majorer pour une subdivision $s=t_0<t_1<\dots<t_p=t$,

$$\begin{aligned}
\sum_{i=1}^p |\langle M,N\rangle_{t_{i-1}}^{t_i}| &\le \sum_{i=1}^p \sqrt{\langle M,M\rangle_{t_{i-1}}^{t_i}}\sqrt{\langle N,N\rangle_{t_{i-1}}^{t_i}} \\
&\le \Big(\sum_{i=1}^p \langle M,M\rangle_{t_{i-1}}^{t_i}\Big)^{1/2}\Big(\sum_{i=1}^p \langle N,N\rangle_{t_{i-1}}^{t_i}\Big)^{1/2} \\
&= \sqrt{\langle M,M\rangle_s^t}\sqrt{\langle N,N\rangle_s^t}
\end{aligned}$$

On peut généraliser et obtenir pour toute partie borélienne bornée A de $\mathbb{R}_+$,

$$\int_A |d\langle M,N\rangle_u| \le \sqrt{\int_A d\langle M,M\rangle_u}\sqrt{\int_A d\langle N,N\rangle_u}.$$

Lorsque $A=[s,t]$, c'est l'inégalité (4.3). Si A est une réunion finie d'intervalles, cela découle de (4.3) et d'une nouvelle application de l'inégalité de Cauchy-Schwarz. Un argument de classe monotone montre alors que cette inégalité est vraie pour toute partie borélienne bornée.

Soient $h=\sum \lambda_i \mathbf{1}_{A_i}$ et $k=\sum \mu_i \mathbf{1}_{A_i}$ deux fonctions étagées positives. Alors,

$$\begin{aligned}
\int h(s)k(s)|d\langle M,N\rangle_s| &= \sum \lambda_i\mu_i \int_{A_i} |d\langle M,N\rangle_s| \\
&\le \Big(\sum \lambda_i^2 \int_{A_i} d\langle M,M\rangle_s\Big)^{1/2}\Big(\sum \mu_i^2 \int_{A_i} d\langle N,N\rangle_s\Big)^{1/2} \\
&= \Big(\int h(s)^2 d\langle M,M\rangle_s\Big)^{1/2}\Big(\int k(s)^2 d\langle N,N\rangle_s\Big)^{1/2},
\end{aligned}$$

ce qui donne l'inégalité voulue pour des fonctions étagées. Il reste à écrire une fonction mesurable positive comme limite croissante de fonctions étagées. □

4.4 Semimartingales continues

Définition 4.5 *Un processus $X = (X_t, t \geq 0)$ est une semimartingale continue s'il s'écrit sous la forme*

$$X_t = X_0 + M_t + A_t,$$

où M est une martingale locale (nulle en $t = 0$) et A est un processus à variation finie.

La décomposition ci-dessus est unique à indistinguabilité près, toujours à cause du théorème 4.5.

Si $Y_t = Y_0 + M'_t + A'_t$ est une autre semimartingale continue on pose par définition

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

En particulier, $\langle X, X \rangle_t = \langle M, M \rangle_t$.

Proposition 4.11 *Soit $0 = t_0^n < t_1^n < \cdots < t_{p_n}^n = t$ une suite de subdivisions emboitées de $[0, t]$ de pas tendant vers 0. Alors,*

$$\lim_{n\to\infty} \sum_{i=1}^{p_n} (X_{t_i^n} - X_{t_{i-1}^n})(Y_{t_i^n} - Y_{t_{i-1}^n}) = \langle X, Y \rangle_t$$

en probabilité.

Démonstration. Pour simplifier, traitons seulement le cas où $X = Y$. Alors,

$$\begin{aligned}\sum_{i=1}^{p_n} (X_{t_i^n} - X_{t_{i-1}^n})^2 &= \sum_{i=1}^{p_n} (M_{t_i^n} - M_{t_{i-1}^n})^2 + \sum_{i=1}^{p_n} (A_{t_i^n} - A_{t_{i-1}^n})^2 \\ &\quad +2 \sum_{i=1}^{p_n} (M_{t_i^n} - M_{t_{i-1}^n})(A_{t_i^n} - A_{t_{i-1}^n}).\end{aligned}$$

On sait déjà (théorème 4.6) que

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

D'autre part,

$$\begin{aligned}\sum_{i=1}^{p_n} (A_{t_i^n} - A_{t_{i-1}^n})^2 &\leq \Big(\sup_{1\leq i\leq p_n} |A_{t_i^n} - A_{t_{i-1}^n}|\Big) \sum_{i=1}^{p_n} |A_{t_i^n} - A_{t_{i-1}^n}| \\ &\leq \Big(\int_0^t |dA_s|\Big) \sup_{1\leq i\leq p_n} |A_{t_i^n} - A_{t_{i-1}^n}|,\end{aligned}$$

qui tend vers 0 p.s. quand $n \to \infty$ par continuité de la fonction $s \to A_s$. Le même raisonnement montre que

$$\Big|\sum_{i=1}^{p_n}(A_{t_i^n} - A_{t_{i-1}^n})(M_{t_i^n} - M_{t_{i-1}^n})\Big| \leq \Big(\int_0^t |dA_s|\Big) \sup_{1\leq i\leq p_n} |M_{t_i^n} - M_{t_{i-1}^n}|$$

tend vers 0 p.s. $\square$

Chapitre 5

Intégrale Stochastique

5.1 Construction de l'intégrale stochastique

Dans tout ce chapitre on se place sur un espace de probabilité $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ muni d'une filtration satisfaisant les conditions habituelles.

Définition 5.1 *On note H^2 l'espace des martingales continues M bornées dans L^2 et telles que $M_0 = 0$.*

Le Théorème 4.7 montre que, si $M \in H^2$, $E[\langle M, M\rangle_\infty] < \infty$. Il en découle que, si $M, N \in H^2$ on a $E[|\langle M, N\rangle_\infty|] < \infty$. En effet, d'après l'inégalité de Kunita-Watanabe (Proposition 4.10),

$$E[|\langle M, N\rangle_\infty|] \le E[\int_0^\infty |d\langle M, N\rangle_s|] \le E[\langle M, M\rangle_\infty]^{1/2} E[\langle N, N\rangle_\infty]^{1/2}.$$

Cela permet de définir un produit scalaire sur H^2 par la formule

$$(M, N)_{H^2} = E[\langle M, N\rangle_\infty].$$

Le Corollaire 4.8 montre aussi que $(M, M)_{H^2} = 0$ si et seulement si $M = 0$ (on identifie des processus indistinguables). La norme sur H^2 associée au produit scalaire est

$$\|M\|_{H^2} = (M, M)_{H^2}^{1/2} = E[\langle M, M\rangle_\infty]^{1/2}.$$

Proposition 5.1 *L'espace H^2 muni du produit scalaire $(M, N)_{H^2}$ est un espace de Hilbert.*

Démonstration. Il faut voir que H^2 est complet pour la norme $\| \; \|_{H^2}$. Soit donc (M^n) une suite de Cauchy pour cette norme: d'après le Théorème 4.7, on a

$$\lim_{m,n\to\infty} E[(M^n_\infty - M^m_\infty)^2] = \lim_{m,n\to\infty} E[\langle M^n - M^m, M^n - M^m\rangle_\infty] = 0.$$

L'inégalité de Doob entraîne qu'on a aussi

$$\lim_{m,n\to\infty} E\Big[\sup_{t\geq 0}(M_t^n - M_t^m)^2\Big] = 0.$$

Il est ensuite facile d'extraire une sous-suite (n_k) telle que

$$E\Big[\sum_{k=1}^{\infty} \sup_{t\geq 0} |M_t^{n_k} - M_t^{n_{k+1}}|\Big] \leq \sum_{k=1}^{\infty} E\Big[\sup_{t\geq 0}(M_t^{n_k} - M_t^{n_{k+1}})^2\Big]^{1/2} < \infty.$$

On en déduit que p.s.

$$\sum_{k=1}^{\infty} \sup_{t\geq 0} |M_t^{n_k} - M_t^{n_{k+1}}| < \infty,$$

et donc p.s. la suite $(M_t^{n_k}, t \geq 0)$ converge uniformément sur $\mathbb{R}_+$ vers une limite notée $(M_t, t \geq 0)$. Clairement le processus limite M a des trajectoires continues (on se débarasse facilement de l'ensemble de probabilité nulle en prenant $M \equiv 0$ sur cet ensemble). Puisque $M_t^{n_k}$ converge aussi dans L^2 vers M_t, pour tout $t \geq 0$ (car la suite $(M_t^{n_k})$ est évidemment de Cauchy dans L^2) on voit immédiatement en passant à la limite sur la propriété de martingale que M est aussi une martingale. Comme les variables M_t^n sont uniformément bornées dans L^2, la martingale M est aussi bornée dans L^2, et on a donc $M \in H^2$. Enfin,

$$\lim_{k\to\infty} E[\langle M^{n_k} - M, M^{n_k} - M\rangle_\infty] = \lim_{k\to\infty} E[(M_\infty^{n_k} - M_\infty)^2] = 0,$$

ce qui montre que la sous-suite (M^{n_k}), donc aussi la suite (M^n) converge vers M dans H^2. □

Rappelons que Prog désigne la tribu progressive sur $\mathbb{R}_+ \times \Omega$.

Définition 5.2 *Pour $M \in H^2$, on note*

$$L^2(M) = L^2(\mathbb{R}_+ \times \Omega, \mathrm{Prog}, dP\, d\langle M, M\rangle_s)$$

l'espace des processus progressifs H tels que

$$E\Big[\int_0^\infty H_s^2\, d\langle M, M\rangle_s\Big] < \infty.$$

Comme n'importe quel espace L^2, l'espace $L^2(M)$ est un espace de Hilbert pour le produit scalaire

$$(H, K)_{L^2(M)} = E\Big[\int_0^\infty H_s K_s\, d\langle M, M\rangle_s\Big].$$

Définition 5.3 *On note $\mathcal{E}$ le sous-espace vectoriel de $L^2(M)$ formé des processus élémentaires, c'est-à-dire des processus H de la forme*

$$H_s(\omega) = \sum_{i=0}^{p-1} H_{(i)}(\omega)\, \mathbf{1}_{]t_i,t_{i+1}]}(s),$$

où $0 \leq t_0 < t_1 < t_2 < \cdots < t_p$ et pour chaque i, $H_{(i)}$ est une variable $\mathcal{F}_{t_i}$-mesurable et bornée.

Proposition 5.2 *Pour tout $M \in H^2$, $\mathcal{E}$ est dense dans $L^2(M)$.*

Démonstration. Il suffit de montrer que si $K \in L^2(M)$ est orthogonal à $\mathcal{E}$ alors $K = 0$. Supposons donc K orthogonal à $\mathcal{E}$. Soient $0 \leq s < t$ et soit F une variable $\mathcal{F}_s$-mesurable bornée. En écrivant $(H, K)_{L^2(M)} = 0$ pour $H = F\,\mathbf{1}_{]s,t]} \in \mathcal{E}$, on trouve

$$E\Big[F \int_s^t K_u\, d\langle M, M\rangle_u\Big] = 0.$$

Posons, pour tout $t \geq 0$,

$$X_t = \int_0^t K_u\, d\langle M, M\rangle_u.$$

Le fait que cette intégrale soit (p.s.) absolument convergente découle facilement de l'inégalité de Cauchy-Schwarz et des propriétés $M \in H^2$, $K \in L^2(M)$. Ce raisonnement montre même que $X_t \in L^1$. De plus, l'identité précédente conduit à $E[F(X_t - X_s)] = 0$ pour tous $s < t$ et toute variable $\mathcal{F}_s$-mesurable F. Cela montre que X est une martingale. D'autre part puisque $X_0 = 0$ et X est aussi un processus à variation finie (Proposition 4.3), cela n'est possible (Théorème 4.5) que si $X = 0$. On a donc

$$\int_0^t K_u\, d\langle M, M\rangle_u = 0 \qquad \forall t \geq 0, \quad \text{p.s.}$$

ce qui entraîne

$$K_u = 0, \qquad d\langle M, M\rangle_u \text{ p.p.}, \quad \text{p.s.}$$

c'est-à-dire $K = 0$ dans $L^2(M)$. □

Théorème 5.3 *Soit $M \in H^2$. Pour tout $H \in \mathcal{E}$, de la forme*

$$H_s(\omega) = \sum_{i=0}^{p-1} H_{(i)}(\omega)\, \mathbf{1}_{]t_i,t_{i+1}]}(s),$$

on définit $H \cdot M \in H^2$ par la formule

$$(H \cdot M)_t = \sum_{i=0}^{p-1} H_{(i)}\, (M_{t_{i+1}\wedge t} - M_{t_i \wedge t}).$$

L'application $H \mapsto H \cdot M$ s'étend à une isométrie de $L^2(M)$ dans H^2. De plus, $H \cdot M$ est caractérisé par la relation

$$\langle H \cdot M, N\rangle = H \cdot \langle M, N\rangle, \qquad \forall N \in H^2. \tag{5.1}$$

Si T est un temps d'arrêt, on a

$$(\mathbf{1}_{[0,T]}H) \cdot M = (H \cdot M)^T = H \cdot M^T. \tag{5.2}$$

On notera très souvent

$$(H \cdot M)_t = \int_0^t H_s \, dM_s.$$

Remarque. L'intégrale $H \cdot \langle M, N\rangle$ qui figure dans le terme de droite de (5.1) est une intégrale par rapport à un processus à variation finie, comme cela a été défini dans le paragraphe 4.1.

Démonstration. On va d'abord vérifier que l'application $M \mapsto H \cdot M$ est une isométrie de $\mathcal{E}$ dans H^2. On vérifie très facilement que pour $H \in \mathcal{E}$, $H \cdot M$ est une martingale continue bornée dans L^2, donc appartient à H^2. De plus l'application $H \mapsto H \cdot M$ est clairement linéaire. Ensuite, on observe que $H \cdot M$ est la somme des martingales

$$M^i_t = H_{(i)} \, (M_{t_{i+1}\wedge t} - M_{t_i \wedge t})$$

qui sont orthogonales et de processus croissants respectifs

$$\langle M^i, M^i\rangle_t = H^2_{(i)}(\langle M, M\rangle_{t_{i+1}\wedge t} - \langle M, M\rangle_{t_i\wedge t})$$

(ces propriétés sont faciles à vérifier, par exemple en utilisant les approximations de $\langle M, N\rangle$). On conclut que

$$\langle H \cdot M, H \cdot M\rangle_t = \sum_{i=0}^{p-1} H^2_{(i)}(\langle M, M\rangle_{t_{i+1}\wedge t} - \langle M, M\rangle_{t_i\wedge t}).$$

En conséquence,

$$\begin{aligned}
\|H \cdot M\|^2_{H^2} &= E\Big[\sum_{i=0}^{p-1} H^2_{(i)}(\langle M, M\rangle_{t_{i+1}} - \langle M, M\rangle_{t_i})\Big] \\
&= E\Big[\int_0^\infty H^2_s \, d\langle M, M\rangle_s\Big] \\
&= \|H\|^2_{L^2(M)}.
\end{aligned}$$

L'application $M \mapsto H \cdot M$ est donc une isométrie de $\mathcal{E}$ dans H^2. Puisque $\mathcal{E}$ est dense dans $L^2(M)$ (Proposition 5.2) et H^2 est un espace de Hilbert (Proposition

5.1), on peut prolonger, de manière unique, cette application en une isométrie de $L^2(M)$ dans H^2.

Vérifions maintenant la propriété (5.1). D'abord, si H est élémentaire comme ci-dessus, on a

$$\langle H \cdot M, N \rangle = \sum_{i=0}^{p-1} \langle M^i, N \rangle$$

et on vérifie facilement que

$$\langle M^i, N \rangle_t = H_{(i)}(\langle M, N \rangle_{t_{i+1} \wedge t} - \langle M, N \rangle_{t_i \wedge t}).$$

On en déduit que

$$\langle H \cdot M, N \rangle_t = \sum_{i=0}^{p-1} H_{(i)}(\langle M, N \rangle_{t_{i+1} \wedge t} - \langle M, N \rangle_{t_i \wedge t}) = \int_0^t H_s \, d\langle M, N \rangle_s$$

ce qui donne la relation (5.1) lorsque $H \in \mathcal{E}$. Ensuite, on remarque que l'application $X \mapsto \langle X, N \rangle_\infty$ est continue de H^2 dans L^1: en effet,

$$E[|\langle X, N \rangle_\infty|] \leq E[\langle X, X \rangle_\infty]^{1/2} E[\langle N, N \rangle_\infty]^{1/2} = E[\langle N, N \rangle_\infty]^{1/2} \|X\|_{H^2}.$$

Si $H^n \in \mathcal{E}$ et $H^n \to H$ dans $L^2(M)$ on a donc

$$\langle H \cdot M, N \rangle_\infty = \lim_{n \to \infty} \langle H^n \cdot M, N \rangle_\infty = \lim_{n \to \infty} (H^n \cdot \langle M, N \rangle)_\infty = (H \cdot \langle M, N \rangle)_\infty,$$

où les convergences ont lieu dans L^1 et la dernière égalité découle de l'inégalité de Kunita-Watanabe:

$$E\Big[\Big| \int_0^\infty (H_s^n - H_s) \, d\langle M, N \rangle_s \Big|\Big] \leq E[\langle N, N \rangle_\infty]^{1/2} \, \|H^n - H\|_{L^2(M)}$$

(le fait que $(H \cdot \langle M, N \rangle)_\infty$ soit bien défini découle du même argument).

En remplaçant N par N^t dans l'égalité $\langle H \cdot M, N \rangle_\infty = (H \cdot \langle M, N \rangle)_\infty$ on trouve $\langle H \cdot M, N \rangle_t = (H \cdot \langle M, N \rangle)_t$, ce qui termine la preuve de (5.1).

Il est facile de voir que la relation (5.1) caractérise $H \cdot M$. En effet, si X est une autre martingale de H^2 qui satisfait la même propriété, on a pour tout $N \in H^2$,

$$\langle H \cdot M - X, N \rangle = 0$$

et en prenant $N = H \cdot M - X$ on trouve que $X = H \cdot M$.

Il reste à vérifier la dernière propriété. En utilisant les propriétés du crochet de deux martingales, on remarque que si $N \in H^2$,

$$\langle (H \cdot M)^T, N \rangle_t = \langle H \cdot M, N \rangle_{t \wedge T} = (H \cdot \langle M, N \rangle)_{t \wedge T} = (H \, \mathbf{1}_{[0,T]} \cdot \langle M, N \rangle)_t$$

ce qui montre que la martingale arrêtée $(H \cdot M)^T$ vérifie la propriété caractéristique de l'intégrale $(\mathbf{1}_{[0,T]}H) \cdot M$. On obtient ainsi la première égalité de (5.2). La preuve de la seconde est analogue, en écrivant

$$\langle H \cdot M^T, N\rangle = H \cdot \langle M^T, N\rangle = H \cdot \langle M, N\rangle^T = H\,\mathbf{1}_{[0,T]} \cdot \langle M, N\rangle.$$

Cela termine la preuve du théorème. □

Remarque. On aurait pu utiliser la relation (5.1) pour *définir* l'intégrale stochastique $H \cdot M$, en observant que l'application $N \mapsto E[(H \cdot \langle M, N\rangle)_\infty]$ définit une forme linéaire continue sur H^2, et donc qu'il existe un élément unique $H \cdot M$ de H^2 tel que

$$E[(H \cdot \langle M, N\rangle)_\infty] = (H \cdot M, N)_{H^2} = E[\langle H \cdot M, N\rangle_\infty].$$

Proposition 5.4 *Si $K \in L^2(M)$ et $H \in L^2(K \cdot M)$ alors $HK \in L^2(M)$ et*

$$(HK) \cdot M = H \cdot (K \cdot M).$$

Démonstration. D'après le Théorème 5.3, on a

$$\langle K \cdot M, K \cdot M\rangle = K \cdot \langle M, K \cdot M\rangle = K^2 \cdot \langle M, M\rangle,$$

et donc

$$\int_0^\infty H_s^2 K_s^2\, d\langle M, M\rangle_s = \int_0^\infty H_s^2\, d\langle K \cdot M, K \cdot M\rangle_s$$

ce qui donne la première assertion. Pour la seconde il suffit de remarquer que si $N \in H^2$,

$$\begin{aligned}\langle (HK) \cdot M, N\rangle = HK \cdot \langle M, N\rangle &= H \cdot (K \cdot \langle M, N\rangle)\\ &= H \cdot \langle K \cdot M, N\rangle\\ &= \langle H \cdot (K \cdot M), N\rangle\end{aligned}$$

d'où le résultat voulu. □

Remarque. De manière informelle, l'égalité de la proposition précédente s'écrit

$$\int_0^t H_s\,(K_s\, dM_s) = \int_0^t H_s K_s\, dM_s.$$

De même la propriété (5.1) s'écrit

$$\langle \int_0^\cdot H_s dM_s, N\rangle_t = \int_0^t H_s\, d\langle M, N\rangle_s.$$

En appliquant deux fois cette relation on a aussi

$$\langle \int_0^\cdot H_s dM_s, \int_0^\cdot K_s dN_s\rangle_t = \int_0^t H_s K_s\, d\langle M, N\rangle_s.$$

En particulier,
$$\langle \int_0^{\cdot} H_s dM_s, \int_0^{\cdot} H_s dM_s \rangle_t = \int_0^t H_s^2 \, d\langle M, M\rangle_s.$$
Remarquons enfin que puisque $H \cdot M$ est une martingale de H^2 on a si $M \in H^2$, $N \in H^2$, $H \in L^2(M)$ et $K \in L^2(N)$, pour tout $t \in [0, \infty]$
$$E\Big[\int_0^t H_s \, dM_s\Big] = 0$$
$$E\Big[\Big(\int_0^t H_s dM_s\Big)\Big(\int_0^t K_s dN_s\Big)\Big] = E\Big[\int_0^t H_s K_s \, d\langle M, N\rangle_s\Big].$$
Il est important de remarquer que ces relations ne seront plus forcément vraies pour les extensions de l'intégrale stochastique qui vont être décrites ci-dessous.

A l'aide des identités (5.2) il est facile d'étendre la définition de $H \cdot M$ à une martingale locale continue quelconque.

Soit M une martingale locale issue de 0. On note $L^2_{\text{loc}}(M)$ (resp. $L^2(M)$) l'espace des processus progressifs H tels que pour tout $t \geq 0$,
$$\int_0^t H_s^2 \, d\langle M, M\rangle_s < \infty, \quad \text{p.s.} \qquad (\text{resp.} \quad E\Big[\int_0^\infty H_s^2 \, d\langle M, M\rangle_s\Big] < \infty).$$
Théorème 5.5 *Soit M une martingale locale issue de* 0*. Pour tout $H \in L^2_{\text{loc}}(M)$, il existe une unique martingale locale issue de* 0*, notée $H \cdot M$, telle que pour toute martingale locale N,*
$$\langle H \cdot M, N\rangle = H \cdot \langle M, N\rangle.$$
La propriété (5.2) *reste vérifiée. De plus, si $M \in H^2$ et $H \in L^2(M)$, cette définition étend celle du Théorème 5.3.*

Remarque. La propriété de la proposition 5.4 reste aussi vraie sous des hypothèses convenables.

Démonstration. On note
$$T_n = \inf\{t \geq 0, \int_0^t (1 + H_s^2) \, d\langle M, M\rangle_s \geq n\},$$
de sorte que (T_n) est une suite de temps d'arrêt croissant vers $+\infty$. Puisque
$$\langle M^{T_n}, M^{T_n}\rangle_t = \langle M, M\rangle_{t \wedge T_n} \leq n,$$
la martingale arrêtée M^{T_n} est dans H^2 (Théorème 4.7). De plus, il est aussi clair que
$$\int_0^\infty H_s^2 \, d\langle M^{T_n}, M^{T_n}\rangle_s \leq n.$$

Donc, $H \in L^2(M^{T_n})$, et on peut définir pour chaque n l'intégrale stochastique $H \cdot M^{T_n}$. En utilisant la propriété caractéristique (5.1), on vérifie très facilement que si $m > n$ on a
$$H \cdot M^{T_n} = (H \cdot M^{T_m})^{T_n}.$$
Cela montre qu'il existe un (unique) processus noté $H \cdot M$ tel que, pour tout n,
$$(H \cdot M)^{T_n} = H \cdot M^{T_n}.$$
Puisque les processus $(H \cdot M)^{T_n}$ sont des martingales de H^2, $H \cdot M$ est une martingale locale.

Soit N une martingale locale, qu'on peut supposer issue de 0 et soient $T'_n = \inf\{t \geq 0, |N_t| \geq n\}$, $S_n = T_n \wedge T'_n$. Alors,
$$\begin{aligned}
\langle H \cdot M, N\rangle^{S_n} &= \langle (H \cdot M)^{S_n}, N^{S_n}\rangle \\
&= \langle (H \cdot M^{T_n})^{S_n}, N^{S_n}\rangle \\
&= \langle H \cdot M^{T_n}, N^{S_n}\rangle \\
&= H \cdot \langle M^{T_n}, N^{S_n}\rangle \\
&= H \cdot \langle M, N\rangle^{S_n} \\
&= (H \cdot \langle M, N\rangle)^{S_n}
\end{aligned}$$
d'où l'égalité $\langle H \cdot M, N\rangle = H \cdot \langle M, N\rangle$. Le fait que cette égalité écrite pour toute martingale locale N caractérise $H \cdot M$ se démontre exactement comme dans le Théorème 5.3.

La propriété (5.2) est obtenue dans ce cadre par les mêmes arguments que dans la preuve du Théorème 5.3 (ces arguments utilisent seulement la propriété caractéristique (5.1) qu'on vient d'étendre). Enfin, si $M \in H^2$ et $H \in L^2(M)$, l'égalité $\langle H \cdot M, H \cdot M\rangle = H^2 \cdot \langle M, M\rangle$ entraîne d'abord que $H \cdot M \in H^2$, et ensuite la propriété caractéristique (5.1) montre que les définitions des Théorèmes 5.3 et 5.5 coïncident. $\square$

Remarque. Discutons maintenant l'extension des formules de moments énoncées avant le théorème 5.5. Soient M une martingale locale, $H \in L^2_{\mathrm{loc}}(M)$ et $t \in [0, \infty]$. Alors, **sous la condition**
$$E[\langle H \cdot M, H \cdot M\rangle_t] = E\Big[\int_0^t H_s^2 \, d\langle M, M\rangle_s\Big] < \infty,$$
on peut appliquer à $(H \cdot M)^t$ le Théorème 4.7, et on a
$$E\Big[\int_0^t H_s \, dM_s\Big] = 0, \qquad E\Big[\Big(\int_0^t H_s \, dM_s\Big)^2\Big] = E\Big[\int_0^t H_s^2 \, d\langle M, M\rangle_s\Big].$$

Nous allons maintenant étendre l'intégrale stochastique aux semimartingales continues. On dit qu'un processus progressif H est localement borné si

$$\text{p.s.}\quad \forall t \geq 0, \qquad \sup_{s\leq t} |H_s| < \infty.$$

En particulier, tout processus continu adapté est localement borné. De plus, si H est localement borné, pour tout processus à variation finie V on a

$$\text{p.s.}\quad \forall t \geq 0, \qquad \int_0^t |H_s|\,|dV_s| < \infty.$$

De même, pour toute martingale locale M, $H \in L^2_{\mathrm{loc}}(M)$.

Définition 5.4 *Soit $X = X_0 + M + V$ une semimartingale continue, et soit H un processus (progressif) localement borné. L'intégrale stochastique $H \cdot X$ est alors définie par*

$$H \cdot X = H \cdot M + H \cdot V,$$

et on note

$$(H \cdot X)_t = \int_0^t H_s\,dX_s.$$

Propriétés.

(i) L'application $(H, X) \mapsto H \cdot X$ est bilinéaire.

(ii) $H \cdot (K \cdot X) = (HK) \cdot X$, si H et K sont localement bornés.

(iii) Pour tout temps d'arrêt T, $(H \cdot X)^T = H\mathbf{1}_{[0,T]} \cdot X = H \cdot X^T$.

(iv) Si X est une martingale locale, resp. si X est un processus à variation finie, alors il en va de même pour $H \cdot X$.

(v) Si H est un processus progressif de la forme $H_s(\omega) = \sum_{i=0}^{p-1} H_{(i)}(\omega)\,\mathbf{1}_{]t_i,t_{i+1}]}(s)$, où, pour chaque i, $H_{(i)}$ est $\mathcal{F}_{t_i}$-mesurable, alors

$$(H \cdot X)_t = \sum_{i=0}^{p-1} H_{(i)}\,(X_{t_{i+1}\wedge t} - X_{t_i \wedge t}).$$

Ces propriétés découlent facilement des résultats obtenus quand X est une martingale, resp. un processus à variation finie. Remarquer que dans la propriété (v) on ne suppose pas que les variables $H_{(i)}$ soient bornées.

Nous terminons ce paragraphe par un résultat technique d'approximation qui sera utile dans la suite.

Proposition 5.6 *Soit X une semimartingale continue et soit H un processus continu adapté. Alors, pour tout $t > 0$, pour toute suite $0 = t_0^n < \cdots < t_{p_n}^n = t$ de subdivisions de $[0, t]$ de pas tendant vers 0, on a*

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

au sens de la convergence en probabilité.

Démonstration. On peut traiter séparément les parties martingale et à variation finie de X. La partie à variation finie est traitée par le Lemme 4.2. On peut donc supposer que $X = M$ est une martingale locale issue de 0. Pour chaque n, définissons un processus H^n par

$$H_s^n = \begin{cases} H_{t_i^n} & \text{si } t_i^n < s \leq t_{i+1}^n \\ 0 & \text{si } s > t. \end{cases}$$

Posons enfin pour tout $p \geq 1$

$$T_p = \inf\{s \geq 0 : |H_s| + \langle M, M\rangle_s \geq p\},$$

et remarquons que H, H^n et $\langle M, M\rangle$ sont bornés sur l'intervalle $]0, T_p]$. D'après la théorie L^2 de l'intégrale stochastique, pour tout p fixé,

$$E\left[\left((H^n \mathbf{1}_{[0,T_p]} \cdot M^{T_p})_t - (H\mathbf{1}_{[0,T_p]} \cdot M^{T_p})_t\right)^2\right] = E\Big[\int_0^{t\wedge T_p} (H_s^n - H_s)^2 d\langle M, M\rangle_s\Big]$$

converge vers 0 quand $n \to \infty$ par convergence dominée. En utilisant (5.2), on en déduit que

$$\lim_{n\to\infty} (H^n \cdot M)_{t\wedge T_p} = (H \cdot M)_{t\wedge T_p},$$

dans L^2. Il ne reste alors plus qu'à remarquer que $P[T_p > t] \uparrow 1$ quand $p \uparrow \infty$. □

5.2 La formule d'Itô

La formule d'Itô est l'outil de base du calcul stochastique. Elle montre qu'une fonction de classe C^2 de p semimartingales continues est encore une semimartingale continue, et exprime explicitement la décomposition de cette semimartingale.

Théorème 5.7 (Formule d'Itô) *Soient $X^1, \ldots, X^p$ p semimartingales continues, et soit F une fonction de classe C^2 de $\mathbb{R}^p$ dans $\mathbb{R}$. Alors,*

$$\begin{aligned} F(X_t^1, \ldots, X_t^p) &= F(X_0^1, \ldots, X_0^p) + \sum_{i=1}^p \int_0^t \frac{\partial F}{\partial x^i}(X_s^1, \ldots, X_s^p)\, dX_s^i \\ &\quad + \frac{1}{2} \sum_{i,j=1}^p \int_0^t \frac{\partial^2 F}{\partial x^i \partial x^j}(X_s^1, \ldots, X_s^p)\, d\langle X^i, X^j \rangle_s. \end{aligned}$$

Démonstration. On traite d'abord le cas $p = 1$ et on note $X = X^1$. Considérons une suite $0 = t_0^n < \cdots < t_{p_n}^n = t$ de subdivisions emboitées de $[0, t]$ de pas tendant vers 0. Alors,

$$F(X_t) = F(X_0) + \sum_{i=0}^{p_n - 1} (F(X_{t_{i+1}^n}) - F(X_{t_i^n})),$$

et d'après la formule de Taylor, on a

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

où

$$\inf_{\theta \in [0,1]} F''(X_{t_i^n} + \theta(X_{t_{i+1}^n} - X_{t_i^n})) \leq f_{n,i} \leq \sup_{\theta \in [0,1]} F''(X_{t_i^n} + \theta(X_{t_{i+1}^n} - X_{t_i^n})).$$

D'après la Proposition 5.6 avec $H_s = F'(X_s)$, on a

$$\lim_{n \to \infty} \sum_{i=0}^{p_n - 1} F'(X_{t_i^n})(X_{t_{i+1}^n} - X_{t_i^n}) = \int_0^t F'(X_s)\, dX_s,$$

en probabilité. Pour compléter la preuve, il suffit donc de montrer que

$$\lim_{n \to \infty} \sum_{i=0}^{p_n - 1} f_{n,i}(X_{t_{i+1}^n} - X_{t_i^n})^2 = \int_0^t F''(X_s)\, d\langle X, X \rangle_s, \tag{5.3}$$

en probabilité.

Commençons par observer que pour $m < n$,

$$\begin{aligned} &\left| \sum_{i=0}^{p_n - 1} f_{n,i}(X_{t_{i+1}^n} - X_{t_i^n})^2 - \sum_{j=0}^{p_m - 1} f_{m,j} \sum_{\{i, t_j^m \leq t_i^n < t_{j+1}^m\}} (X_{t_{i+1}^n} - X_{t_i^n})^2 \right| \\ &\qquad \leq Z_{m,n} \left(\sum_{i=0}^{p_n - 1} (X_{t_{i+1}^n} - X_{t_i^n})^2 \right), \end{aligned}$$

avec

$$Z_{m,n} = \sup_{0\le j\le p_m-1}\left(\sup_{\{i,t_j^m\le t_i^n<t_{j+1}^m\}}|f_{n,i}-f_{m,j}|\right).$$

Grâce à la continuité de F'', on vérifie facilement que $Z_{m,n} \longrightarrow 0$ p.s. quand $m,n\to\infty$. Il découle alors de la Proposition 4.11 que pour $\varepsilon>0$ donné, on peut choisir m assez grand de façon que, pour tout $n>m$,

$$P\left[Z_{m,n}\left(\sum_{i=0}^{p_n-1}(X_{t_{i+1}^n}-X_{t_i^n})^2\right)\ge\varepsilon\right]\le\varepsilon.$$

Ensuite, pour cette valeur fixée de m, la Proposition 4.11 montre aussi que

$$\begin{aligned}\lim_{n\to\infty}\sum_{j=0}^{p_m-1}f_{m,j}\sum_{\{i,t_j^m\le t_i^n<t_{j+1}^m\}}(X_{t_{i+1}^n}-X_{t_i^n})^2 &= \sum_{j=0}^{p_m-1}f_{m,j}\Big(\langle X,X\rangle_{t_{j+1}^m}-\langle X,X\rangle_{t_j^m}\Big)\\ &= \int_0^t h_m(s)\,d\langle X,X\rangle_s,\end{aligned}$$

où $h_m(s)=f_{m,j}$ si $t_j^m\le s<t_{j+1}^m$. Il est clair que $h_m(s)\longrightarrow F''(X_s)$ p.s. quand $m\to\infty$. Donc, quitte à prendre m encore plus grand, on peut supposer que

$$P\Big[\Big|\int_0^t h_m(s)\,d\langle X,X\rangle_s-\int_0^t F''(X_s)\,d\langle X,X\rangle_s\Big|\ge\varepsilon\Big]\le\varepsilon.$$

Finalement, en combinant ce qui précède on voit que pour $n>m$ assez grand,

$$P\Big[\Big|\sum_{i=0}^{p_n-1}f_{n,i}(X_{t_{i+1}^n}-X_{t_i^n})^2-\int_0^t F''(X_s)\,d\langle X,X\rangle_s\Big|\ge 3\varepsilon\Big]\le 3\varepsilon,$$

ce qui termine la preuve de (5.3).

Dans le cas où p est quelconque, la formule de Taylor donne

$$\begin{aligned}F(X^1_{t_{i+1}^n},\dots,X^p_{t_{i+1}^n})-F(X^1_{t_i^n},\dots,X^p_{t_i^n}) &= \sum_{k=1}^p\frac{\partial F}{\partial x^k}(X^1_{t_i^n},\dots,X^p_{t_i^n})\,(X^k_{t_{i+1}^n}-X^k_{t_i^n})\\ &\quad+\sum_{k,l=1}^p\frac{f_{n,i}^{k,l}}{2}(X^k_{t_{i+1}^n}-X^k_{t_i^n})(X^l_{t_{i+1}^n}-X^l_{t_i^n})\end{aligned}$$

avec

$$\inf_{\theta\in[0,1]}\frac{\partial^2F}{\partial x_k\partial x_l}(X^1_{t_i^n}+\theta(X^1_{t_{i+1}^n}-X^1_{t_i^n}),\dots)\le f_{n,i}^{k,l}\le\sup_{\theta\in[0,1]}\frac{\partial^2F}{\partial x_k\partial x_l}(X^1_{t_i^n}+\theta(X^1_{t_{i+1}^n}-X^1_{t_i^n}),\dots)$$

La Proposition 5.6 donne à nouveau le résultat recherché pour les termes faisant intervenir les dérivées premières. De plus une légère modification des arguments ci-dessus montre que pour tous $k, l \in \{1, \ldots, p\}$,

$$\lim_{n\to\infty} \sum_{i=0}^{p_n-1} f_{n,i}^{k,l}(X^k_{t^n_{i+1}} - X^k_{t^n_i})(X^l_{t^n_{i+1}} - X^l_{t^n_i}) = \int_0^t \frac{\partial^2 F}{\partial x_k \partial x_l}(X^1_s, \ldots, X^p_s) d\langle X^k, X^l \rangle_s,$$

ce qui termine la preuve du théorème. $\square$

Un cas particulier important de la formule d'Itô est la formule d'intégration par parties. Si X et Y sont deux semimartingales continues, on a

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s \, dY_s + \int_0^t Y_s \, dX_s + \langle X, Y \rangle_t.$$

En particulier, si $Y = X$,

$$X_t^2 = X_0^2 + 2 \int_0^t X_s \, dX_s + \langle X, X \rangle_t.$$

Lorsque $X = M$ est une martingale locale, on sait que $M^2 - \langle M, M \rangle$ est une martingale locale. La formule précédente montre que cette martingale locale est

$$M_0^2 + 2 \int_0^t M_s \, dM_s,$$

ce qu'on aurait pu voir directement sur la démonstration faite dans le Chapitre 4 (notre construction de $\langle M, M \rangle$ faisait intervenir des approximations de l'intégrale stochastique $\int_0^t M_s dM_s$).

Un $(\mathcal{F}_t)$-mouvement brownien B (issu de 0) est un mouvement brownien qui est $(\mathcal{F}_t)$-adapté et à accroissements indépendants par rapport à $(\mathcal{F}_t)$. Un $(\mathcal{F}_t)$-mouvement brownien est bien sûr une martingale, et on a déjà remarqué que sa variation quadratique est $\langle B, B \rangle_t = t$. Plus généralement, si U est une variable aléatoire $\mathcal{F}_0$-mesurable, un $(\mathcal{F}_t)$-mouvement brownien issu de U est un processus qui s'écrit $B_t = U + B'_t$, où B' est un $(\mathcal{F}_t)$-mouvement brownien issu de 0.

Pour un $(\mathcal{F}_t)$-mouvement brownien B, la formule d'Itô s'écrit

$$F(B_t) = F(B_0) + \int_0^t F'(B_s) \, dB_s + \frac{1}{2} \int_0^t F''(B_s) ds.$$

En prenant $X_t^1 = t$, $X_t^2 = B_t$, on a aussi pour toute fonction F de classe C^2 sur $\mathbb{R}_+ \times \mathbb{R}$,

$$F(t, B_t) = F(0, B_0) + \int_0^t \frac{\partial F}{\partial x}(s, B_s) \, dB_s + \int_0^t (\frac{\partial F}{\partial t} + \frac{1}{2} \frac{\partial^2 F}{\partial x^2})(s, B_s) \, ds.$$

On dit qu'un processus $(\mathcal{F}_t)$-adapté $B_t = (B_t^1, \ldots, B_t^p)$ à valeurs dans $\mathbb{R}^d$ et issu de 0 est un $(\mathcal{F}_t)$-mouvement brownien si $B^1, \ldots, B^p$ sont des mouvement browniens réels indépendants et si B est à accroissements indépendants par rapport à la filtration $(\mathcal{F}_t)$. Plus généralement, un processus B à valeurs dans $\mathbb{R}^d$ est un $(\mathcal{F}_t)$-mouvement brownien si on peut écrire $B_t = B_0 + B'_t$ où B_0 est $\mathcal{F}_0$-mesurable et B' est un $(\mathcal{F}_t)$-mouvement brownien issu de 0.

Soit B un $(\mathcal{F}_t)$-mouvement brownien en dimension d. Il est facile d'adapter la preuve du Théorème 2.11 pour montrer que si T est un temps d'arrêt de la filtration $(\mathcal{F}_t)_{t\geq 0}$ qui est fini p.s., alors le processus $(B_t^{(T)} = B_{T+t} - B_T, t \geq 0)$ est aussi un mouvement brownien et est indépendant de $\mathcal{F}_T$.

Si $B_t = (B_t^1, \ldots, B_t^p)$ est un $(\mathcal{F}_t)$-mouvement brownien en dimension d, on a vu au Chapitre précédent que $\langle B^i, B^j \rangle = 0$ lorsque $i \neq j$. La formule d'Itô montre alors que, pour toute fonction F de classe C^2 sur $\mathbb{R}^p$,

$$\begin{aligned}&F(B_t^1, \ldots, B_t^p)\\&= F(B_0^1, \ldots, B_0^p) + \sum_{i=1}^{p} \int_0^t \frac{\partial F}{\partial x_i}(B_s^1, \ldots, B_s^p)\, dB_s^i + \frac{1}{2}\int_0^t \Delta F(B_s^1, \ldots, B_s^p)\, ds,\end{aligned}$$

et on a une formule analogue pour $F(t, B_t^1, \ldots, B_t^p)$.

Nous utilisons maintenant la formule d'Itô pour dégager une classe importante de martingales, qui généralise les martingales exponentielles rencontrées pour les processus à accroissements indépendants. On dit qu'un processus à valeurs dans $\mathbb{C}$ est une martingale locale si sa partie réelle et sa partie imaginaire sont des martingales locales.

Proposition 5.8 *Soit M une martingale locale et pour tout $\lambda \in \mathbb{C}$, soit*

$$\mathcal{E}(\lambda M)_t = \exp\Big(\lambda M_t - \frac{\lambda^2}{2}\langle M, M\rangle_t\Big).$$

Le processus $\mathcal{E}(\lambda M)$ est une martingale locale.

Démonstration. Si $F(x, r)$ est une fonction de classe C^2 sur $\mathbb{R} \times \mathbb{R}_+$, la formule d'Itô entraîne que

$$\begin{aligned}F(M_t, \langle M, M\rangle_t) &= F(M_0, 0) + \int_0^t \frac{\partial F}{\partial x}(M_s, \langle M, M\rangle_s)\, dM_s\\&\quad + \int_0^t \Big(\frac{\partial F}{\partial r} + \frac{1}{2}\frac{\partial^2 F}{\partial x^2}\Big)(M_s, \langle M, M\rangle_s)\, d\langle M, M\rangle_s.\end{aligned}$$

Donc, $F(M_t, \langle M, M\rangle_t)$ est une martingale locale dès que F vérifie l'équation

$$\frac{\partial F}{\partial r} + \frac{1}{2}\frac{\partial^2 F}{\partial x^2} = 0.$$

Or cette équation est clairement vérifiée par la fonction $F(x, r) = \exp(\lambda x - \frac{\lambda^2}{2} r)$ (plus précisément par les parties réelle et imaginaire de cette fonction). □

5.3 Quelques applications de la formule d'Itô

Nous commençons par un important théorème de caractérisation.

Théorème 5.9 (Lévy) *Soit $X = (X^1, \ldots, X^d)$ un processus continu $(\mathcal{F}_t)$-adapté. Il y a équivalence entre*

(i) *X est un $(\mathcal{F}_t)$-mouvement brownien en dimension d.*

(ii) *Les processus $X^1, \ldots, X^d$ sont des $(\mathcal{F}_t)$-martingales locales continues et de plus $\langle X^i, X^j \rangle_t = \delta_{ij} t$ (δ_{ij} est le symbole de Kronecker, $\delta_{ij} = \mathbf{1}_{\{i=j\}}$).*

En particulier, une $(\mathcal{F}_t)$-martingale locale continue M est un $(\mathcal{F}_t)$-mouvement brownien si et seulement si $\langle M, M \rangle_t = t$.

Démonstration. Nous savons déjà que (i) $\Rightarrow$ (ii). Montrons la réciproque. Soit $\xi \in \mathbb{R}^d$. Alors, $\xi \cdot X_t = \sum_{j=1}^d \xi_j X_t^j$ est une martingale locale de processus croissant

$$\sum_{j=1}^d \sum_{k=1}^d \xi_j \xi_k \langle X^j, X^k \rangle_t = |\xi|^2 t.$$

D'après la Proposition 5.8, $\exp(i\xi \cdot X_t + \frac{1}{2}|\xi|^2 t)$ est une martingale locale. Cette martingale locale étant bornée sur les intervalles $[0, T]$, $T > 0$, est une vraie martingale. Donc, pour $s < t$,

$$E[\exp(i\xi \cdot X_t + \frac{1}{2}|\xi|^2 t) \mid \mathcal{F}_s] = \exp(i\xi \cdot X_s + \frac{1}{2}|\xi|^2 s).$$

Il en découle que, pour $A \in \mathcal{F}_s$,

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

En prenant $A = \Omega$, on voit que $X_t - X_s$ est un vecteur gaussien centré de covariance $(t - s)\mathrm{Id}$. De plus, fixons $A \in \mathcal{F}_s$ avec $P[A] > 0$, et notons P_A la probabilité $P_A[\cdot] = P[A]^{-1} P[\cdot \cap A]$. On voit que

$$P_A[\exp(i\xi \cdot (X_t - X_s))] = \exp(-\frac{1}{2}|\xi|^2 (t - s))$$

ce qui veut dire que la loi de $X_t - X_s$ sous P_A est aussi celle du vecteur gaussien de covariance $(t - s)\mathrm{Id}$. Donc, pour toute fonction mesurable positive f sur $\mathbb{R}^d$, on a

$$E_A[f(X_t - X_s)] = E[f(X_t - X_s)],$$

soit encore

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

Comme cela est vrai pour tout $A \in \mathcal{F}_s$, $X_t - X_s$ est indépendant de $\mathcal{F}_s$.

Finalement, si $t_0 = 0 < t_1 < \ldots < t_p$, le vecteur $(X^i_{t_j} - X^i_{t_{j-1}}; 1 \le i \le d, 1 \le j \le p)$ est un vecteur gaussien car obtenu en regroupant p vecteurs gaussiens indépendants. De plus, ses composantes sont orthogonales donc indépendantes. Compte-tenu de la loi obtenue précédemment pour $X_t - X_s$, cela suffit pour dire que X est un $(\mathcal{F}_t)$-mouvement brownien. □

Théorème 5.10 (Dubins-Schwarz) *Soit M une martingale locale continue issue de 0 et telle que $\langle M, M\rangle_\infty = \infty$ p.s. Il existe alors un mouvement brownien β tel que*

$$p.s.\ \forall t \ge 0, \qquad M_t = \beta_{<M,M>_t}.$$

Remarques. (i) On peut s'affranchir de l'hypothèse $\langle M, M\rangle_\infty = \infty$, mais il faut alors éventuellement "grossir" l'espace de probabilité.
(ii) Le mouvement brownien β n'est pas adapté par rapport à la filtration $(\mathcal{F}_t)$, mais par rapport à une filtration "changée de temps".

Démonstration. Pour tout $r \ge 0$, on définit un temps d'arrêt τ_r en posant

$$\tau_r = \inf\{t \ge 0,\ \langle M, M\rangle_t > r\}.$$

L'hypothèse assure que $\tau_r < \infty$ p.s. De plus, la fonction $r \mapsto \tau_r$ est croissante càdlàg et, si $r > 0$,

$$\lim_{s\uparrow\uparrow r} \tau_s = \tau_{r-} = \inf\{t \ge 0,\ \langle M, M\rangle_t = r\}.$$

On pose $\beta_r = M_{\tau_r}$. Le processus β est adapté par rapport à la filtration $\mathcal{G}_r = \mathcal{F}_{\tau_r}$. Remarquons aussi que cette filtration satisfait les conditions habituelles (rappelons que si T_n est une suite de temps d'arrêt décroissant vers T on a $\mathcal{F}_T = \cap_n \mathcal{F}_{T_n}$).

Lemme 5.11 *Les intervalles de constance de M et $\langle M, M\rangle$ sont p.s. les mêmes. En d'autres termes, on a p.s. pour tous $a < b$,*

$$M_t = M_a,\ \forall t \in [a,b] \iff \langle M, M\rangle_t = \langle M, M\rangle_a,\ \forall t \in [a,b].$$

Démonstration. Exercice (de la gauche vers la droite, utiliser l'approximation habituelle de $\langle M, M\rangle_t$; de la droite vers la gauche, appliquer le Corollaire 4.8 au processus $M_{(a+t)\wedge T} - M_a$, pour un temps d'arrêt T convenable).

Revenons à la preuve du Théorème 5.10. On a p.s. pour tout $r > 0$,

$$\lim_{s\uparrow\uparrow r} \beta_s = \lim_{s\uparrow\uparrow r} M_{\tau_s} = M_{\tau_{s-}} = M_{\tau_s},$$

où la dernière égalité provient du Lemme 5.11 et du fait que pour $t \in [\tau_{s-}, \tau_s]$ on a $\langle M, M\rangle_t = s$. Comme les trajectoires de β sont clairement continues à droite, on conclut que le processus β est continu.

Nous vérifions ensuite que β_s et $\beta_s^2 - s$ sont des martingales relativement à la filtration $(\mathcal{G}_s)$. Pour tout $n \geq 1$, les martingales locales arrêtées M^{τ_n} et $(M^{\tau_n})^2 - \langle M, M\rangle^{\tau_n}$ sont des vraies martingales uniformément intégrables (d'après le Théorème 4.7, en observant que $\langle M^{\tau_n}, M^{\tau_n}\rangle_\infty = \langle M, M\rangle_{\tau_n} = n$). Le théorème d'arrêt entraîne alors que pour $r \leq s \leq n$,

$$E[\beta_s \mid \mathcal{G}_r] = E[M^{\tau_n}_{\tau_s} \mid \mathcal{F}_{\tau_r}] = M^{\tau_n}_{\tau_r} = \beta_r$$

et

$$\begin{aligned} E[\beta_s^2 - s \mid \mathcal{G}_r] = E[(M^{\tau_n}_{\tau_s})^2 - \langle M^{\tau_n}, M^{\tau_n}\rangle_{\tau_s} \mid \mathcal{F}_{\tau_r}] &= (M^{\tau_n}_{\tau_r})^2 - \langle M^{\tau_n}, M^{\tau_n}\rangle_{\tau_r} \\ &= \beta_r^2 - r. \end{aligned}$$

Ensuite, le cas $d = 1$ du Théorème 5.9 montre que β est un $(\mathcal{G}_r)$-mouvement brownien. Finalement, par définition de β, on a p.s. pour tout $t \geq 0$,

$$\beta_{<M,M>_t} = M_{\tau_{<M,M>_t}}.$$

Mais à nouveau le Lemme 5.11 montre que la fonction $r \mapsto \tau_r$ est constante sur les intervalles $[t, \tau_{<M,M>_t}]$. On conclut que p.s. pour tout $t \geq 0$ on a $M_t = \beta_{<M,M>_t}$. $\square$

Nous énonçons maintenant des inégalités importantes reliant une martingale et sa variation quadratique. Si M est une martingale locale continue, on note $M_t^* = \sup_{s \leq t} |M_s|$.

Théorème 5.12 (Inégalités de Burkholder-Davis-Gundy) *Pour tout réel $p > 0$, il existe des constantes c_p, $C_p > 0$ telles que, pour toute martingale locale M issue de 0,*

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

Remarque. Si T est un temps d'arrêt quelconque, en remplaçant M par la martingale locale arrêtée M^T, on obtient les mêmes inégalités avec T à la place de ∞.

Démonstration. Observons d'abord que l'on peut se restreindre au cas où M est bornée, quitte à remplacer M par M^{T_n}, si $T_n = \inf\{t \geq 0, |M_t| = n\}$, et à faire ensuite tendre n vers ∞. Nous démontrons seulement l'inégalité de droite dans le cas $p \geq 2$ (c'est ce cas particulier que nous utiliserons dans le chapitre 6). On applique la formule d'Itô à la fonction $|x|^p$:

$$|M_t|^p = \int_0^t p|M_s|^{p-1}\mathrm{sgn}(M_s)\, dM_s + \frac{1}{2}\int_0^t p(p-1)|M_s|^{p-2}\, d\langle M, M\rangle_s.$$

Puisque M est bornée, donc en particulier $M \in H^2$, le processus

$$\int_0^t p|M_s|^{p-1}\mathrm{sgn}(M_s)\, dM_s$$

est une vraie martingale dans H^2. On trouve alors

$$\begin{aligned} E[|M_t|^p] &= \frac{p(p-1)}{2} E\Big[\int_0^t |M_s|^{p-2}\, d\langle M, M\rangle_s\Big] \\ &\leq \frac{p(p-1)}{2} E[(M_t^*)^{p-2}\langle M, M\rangle_t] \\ &\leq \frac{p(p-1)}{2} E[(M_t^*)^p]^{(p-2)/p} E[\langle M, M\rangle_t^{p/2}]^{2/p}, \end{aligned}$$

d'après l'inégalité de Hölder. D'autre part, d'après l'inégalité de Doob,

$$E[(M_t^*)^p] \leq \Big(\frac{p}{p-1}\Big)^p E[|M_t|^p]$$

et en combinant cette égalité avec la précédente, on arrive à

$$E[(M_t^*)^p] \leq \Big(\Big(\frac{p}{p-1}\Big)^p \frac{p(p-1)}{2}\Big)^{p/2} E[\langle M, M\rangle_t^{p/2}].$$

Il ne reste plus qu'à faire tendre t vers ∞. □

Représentation des martingales. Nous allons établir que, dans le cas où la filtration sur Ω est engendrée par un mouvement brownien, toutes les martingales peuvent être représentées comme intégrales stochastiques par rapport à ce mouvement brownien.

Théorème 5.13 *Supposons que la filtration $(\mathcal{F}_t)$ sur Ω est (l'augmentation habituelle de) la filtration canonique d'un mouvement brownien B issu de 0. Alors, pour toute variable aléatoire $Z \in L^2(\Omega, \mathcal{F}_\infty)$, il existe un (unique) processus $h \in L^2(B)$ tel que*

$$Z = E[Z] + \int_0^\infty h(s,\omega)\, dB_s.$$

En conséquence, pour toute martingale M bornée dans L^2 (respectivement pour toute martingale locale M), il existe un (unique) processus $h \in L^2(B)$ (resp. $h \in L^2_{\mathrm{loc}}(B)$) et une constante $C \in \mathbb{R}$ tels que

$$M_t = C + \int_0^t h(s,\omega)\, dB_s.$$

Lemme 5.14 *Sous les hypothèses du théorème précédent, l'espace vectoriel engendré par les variables aléatoires*

$$\exp\Big(i\sum_{j=1}^n \lambda_j(B_{t_j}-B_{t_{j-1}})\Big)$$

pour $0=t_0<t_1<\cdots<t_n$ *et* $\lambda_1,\ldots,\lambda_n\in\mathbb{R}$, *est dense dans* $L^2_{\mathbb{C}}(\Omega,\mathcal{F}_\infty)$.

Démonstration. Il suffit de montrer que si $Z\in L^2_{\mathbb{C}}(\Omega,\mathcal{F}_\infty)$ vérifie

$$E\Big[Z\,\exp\Big(i\sum_{j=1}^n \lambda_j(B_{t_j}-B_{t_{j-1}})\Big)\Big]=0 \tag{5.4}$$

pour tout choix de $0=t_0<t_1<\cdots<t_n$ et $\lambda_1,\ldots,\lambda_n\in\mathbb{R}$, alors $Z=0$. Fixons $0=t_0<t_1<\cdots<t_n$ et, pour tous $\sigma>0$ et $m\in\mathbb{R}$, notons $g_{m,\sigma}(x),\ x\in\mathbb{R}$, la densité de la loi $\mathcal{N}(m,\sigma^2)$. On déduit de la condition (5.4) que, pour tous $\sigma_1,\ldots,\sigma_n>0$ et $m_1,\ldots,m_n\in\mathbb{R}$, on a

$$\begin{aligned} 0 &= \int_{\mathbb{R}^n}\prod_{j=1}^n g_{m_j,\sigma_j}(\lambda_j)\,E\Big[Z\,\exp\Big(i\sum_{j=1}^n \lambda_j(B_{t_j}-B_{t_{j-1}})\Big)\Big]\,d\lambda_1\ldots d\lambda_n\\ &= E\Big[Z\prod_{j=1}^n \exp\Big(im_j(B_{t_j}-B_{t_{j-1}})-\frac{\sigma_j^2(B_{t_j}-B_{t_{j-1}})^2}{2}\Big)\Big]\end{aligned}$$

où la deuxième égalité résulte d'une application (facile à justifier) du théorème de Fubini et de la formule pour la transformée de Fourier de la loi gaussienne. On a donc pour tous $\alpha_1,\ldots,\alpha_n>0$ et $m_1,\ldots,m_n\in\mathbb{R}$,

$$E\Big[Z\,\exp\Big(\sum_{j=1}^n(im_j(B_{t_j}-B_{t_{j-1}})-\alpha_j(B_{t_j}-B_{t_{j-1}})^2)\Big)\Big]=0.$$

Une application du théorème de Stone-Weierstrass montre que les combinaisons linéaires complexes de la fonction constante égale à 1 et des fonctions de la forme

$$(y_1,\ldots,y_n)\longrightarrow \exp\Big(\sum_{j=1}^n(im_jy_j-\alpha_j\,y_j^2)\Big)$$

avec $\alpha_1,\ldots,\alpha_n>0$ et $m_1,\ldots,m_n\in\mathbb{R}$, sont denses dans l'espace $C_\ell(\mathbb{R}^n,\mathbb{C})$ des fonctions continues de $\mathbb{R}^n$ dans $\mathbb{C}$ qui ont une limite à l'infini, muni de la norme de la convergence uniforme. Par un passage à la limite, on obtient donc, pour toute fonction $\varphi\in C_\ell(\mathbb{R}^n,\mathbb{C})$,

$$E[Z\,\varphi(B_{t_1},B_{t_2}-B_{t_1},\ldots,B_{t_n}-B_{t_{n-1}})]=0.$$

Il en découle que l'égalité

$$E[Z\,\mathbf{1}_U(B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}})] = 0$$

est vraie d'abord pour tout ouvert borné U de $\mathbb{R}^n$, puis pour tout borélien U de $\mathbb{R}^n$ par un argument de classe monotone.

On a donc obtenu l'égalité $E[Z\,\mathbf{1}_A] = 0$ pour tout $A \in \sigma(B_{t_1}, \ldots, B_{t_n})$. A nouveau un argument de classe monotone montre que cette égalité reste vraie pour tout $A \in \sigma(B_t, t \geq 0)$ puis par complétion pour tout $A \in \mathcal{F}_\infty$. On conclut alors que $Z = 0$. $\square$

Démonstration du Théorème 5.13. On montre d'abord la première assertion. Pour cela, on note $\mathcal{H}$ l'espace vectoriel des variables aléatoires $Z \in L^2(\Omega, \mathcal{F}_\infty)$ qui ont la propriété de l'énoncé. Remarquons que l'unicité de h est facile puisque si h et h' correspondent à la même variable Z, on a

$$E\Big[\int_0^\infty (h(s,\omega) - h'(s,\omega))^2 ds\Big] = E\Big[\Big(\int_0^\infty h(s,\omega)\,dB_s - \int_0^\infty h'(s,\omega)\,dB_s\Big)^2\Big] = 0,$$

d'où $h = h'$ dans $L^2(B)$. Si $Z \in \mathcal{H}$ correspond à h, on a

$$E[Z^2] = E[Z]^2 + E\Big[\int_0^\infty h(s,\omega)^2\,ds\Big].$$

Il en découle facilement que si (Z_n) est une suite dans $\mathcal{H}$ qui converge dans $L^2(\Omega, \mathcal{F}_\infty)$ vers Z, les processus h_n associés à Z_n forment une suite de Cauchy dans $L^2(B)$ donc convergent vers $h \in L^2(B)$. D'après les propriétés de l'intégrale stochastique, on a aussitôt $Z = E[Z] + \int_0^\infty h(s,\omega)\,dB_s$. Donc $\mathcal{H}$ est fermé.

Ensuite, pour $0 = t_0 < t_1 < \cdots < t_n$ et $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, notons $f(s) = \sum_{j=1}^n \lambda_j \mathbf{1}_{]t_{j-1},t_j]}(s)$, et $\mathcal{E}_t^f$ la martingale exponentielle $\mathcal{E}(i\int_0^\cdot f(s)\,dB_s)$ (cf Proposition 5.8). La formule d'Itô montre que

$$\exp\Big(i\sum_{j=1}^n \lambda_j(B_{t_j} - B_{t_{j-1}}) + \frac{1}{2}\sum_{j=1}^n \lambda_j^2(t_j - t_{j-1})\Big) = \mathcal{E}_\infty^f = 1 + i\int_0^\infty \mathcal{E}_s^f f(s)\,dB_s$$

et il en découle que les variables de la forme $\mathrm{Re}\Big(\exp\Big(i\sum_{j=1}^n \lambda_j(B_{t_j} - B_{t_{j-1}})\Big)\Big)$ sont dans $\mathcal{H}$. D'après le Lemme 5.14, les variables de ce type sont denses dans $L^2(\Omega, \mathcal{F}_\infty)$. On conclut que $\mathcal{H} = L^2(\Omega, \mathcal{F}_\infty)$.

Si M est une martingale bornée dans L^2, alors $M_\infty \in L^2(\Omega, \mathcal{F}_\infty)$, s'écrit sous la forme

$$M_\infty = E[M_\infty] + \int_0^\infty h(s,\omega)\,dB_s,$$

avec $h \in L^2(B)$. Il en découle aussitôt que

$$M_t = E[M_\infty \mid \mathcal{F}_t] = E[M_\infty] + \int_0^t h(s,\omega)\,dB_s$$

et l'unicité de h s'obtient comme ci-dessus.

Enfin, si M est une martingale locale (continue), on a d'abord $M_0 = C \in \mathbb{R}$ parce que $\mathcal{F}_0$ est P-triviale (ce qu'on peut déduire soit de la première partie de la preuve soit du chapitre 2). Si $T_n = \inf\{t \geq 0, |M_t| \geq n\}$ on peut appliquer ce qui précède à M^{T_n} et trouver un processus $h_n \in L^2(B)$ tel que

$$M_t^{T_n} = C + \int_0^t h_n(s,\omega)\,dB_s.$$

Par unicité on a, si $m < n$, $h_n(s,\omega) = h_m(s,\omega)$, ds p.p. sur $[0, T_m]$, p.s. Il est alors facile de construire $h \in L^2_{\mathrm{loc}}(B)$ tel que, pour tout m, $h(s,\omega) = h_m(s,\omega)$, ds p.p. sur $[0, T_m]$, p.s. La formule de l'énoncé découle ensuite de la construction de l'intégrale stochastique $\int_0^t h(s,\omega) dB_s$, et l'unicité de h est aussi facile par un argument de localisation. $\square$

Remarque. Sous les hypothèses du théorème 5.13, notons $\mathcal{N}$ la classe des P-négligeables de $\sigma(B_t, t \geq 0)$ et pour tout $t \geq 0$, $\mathcal{G}_t = \sigma(B_s, 0 \leq s \leq t) \vee \mathcal{N}$. A priori on a $\mathcal{F}_t = \mathcal{G}_{t+}$. En fait, le théorème 5.13 entraîne que $\mathcal{G}_t = \mathcal{G}_{t+} = \mathcal{F}_t$ (le cas $t = 0$ est le Théorème 2.8). En effet, si Z est une variable $\mathcal{F}_t$-mesurable bornée, on a

$$Z = \int_0^t h(s,\omega)\,dB_s = \lim_{\varepsilon \downarrow 0} (L^2) \int_0^{t-\varepsilon} h(s,\omega)\,dB_s,$$

et quitte à prendre une sous-suite on voit que Z est limite p.s. de variables $\mathcal{G}_t$-mesurables (car $\mathcal{F}_{t-\varepsilon} \subset \mathcal{G}_t$ si $\varepsilon > 0$).

5.4 Le théorème de Girsanov

Nous considérons comme précédemment dans ce chapitre, un espace de probabilité filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ satisfaisant les conditions habituelles. Notre objectif est d'étudier comment se transforment les notions de semimartingales et de martingales lorsqu'on remplace la probabilité P par une probabilité Q absolument continue par rapport à P.

Proposition 5.15 *Supposons que $Q \ll P$ sur $\mathcal{F}_\infty$. Pour tout $t \in [0, \infty]$, soit*

$$D_t = \frac{dQ}{dP}_{|\mathcal{F}_t}$$

la dérivée de Radon-Nikodym de Q par rapport à P sur la tribu $\mathcal{F}_t$. Le processus D est une $\mathcal{F}_t$-martingale uniformément intégrable. On peut donc remplacer D par une modification càdlàg. Après ce remplacement, on a aussi pour tout temps d'arrêt T,

$$D_T = \frac{dQ}{dP}_{|\mathcal{F}_T}.$$

Enfin, si on suppose que Q est équivalente à P sur $\mathcal{F}_\infty$, on a p.s. pour tout $t \geq 0$, $D_t > 0$.

Démonstration. Pour $A \in \mathcal{F}_t$, on a

$$Q(A) = E_Q[\mathbf{1}_A] = E_P[\mathbf{1}_A D_\infty] = E_P[\mathbf{1}_A E_P[D_\infty \mid \mathcal{F}_t]]$$

et par unicité de la dérivée de Radon-Nikodym sur $\mathcal{F}_t$, il en découle que

$$D_t = E_P[D_\infty \mid \mathcal{F}_t], \qquad \text{p.s.}$$

Donc D est une martingale u.i. (fermée par D_∞), et quitte à remplacer D par une modification on peut supposer que ses trajectoires sont càdlàg.

Ensuite, si T est un temps d'arrêt, on a pour $A \in \mathcal{F}_T$, d'après le théorème d'arrêt,

$$Q(A) = E_Q[\mathbf{1}_A] = E_P[\mathbf{1}_A D_\infty] = E_P[\mathbf{1}_A D_T],$$

d'où, puisque D_T est $\mathcal{F}_T$-mesurable,

$$D_T = \frac{dQ}{dP}_{|\mathcal{F}_T}.$$

Enfin, considérons le temps d'arrêt $S = \inf\{t \geq 0,\, D_t = 0\}$. Par continuité à droite, $D_S = 0$ p.s. sur $\{S < \infty\}$. En prenant $A = \{S < \infty\} \in \mathcal{F}_S$, l'égalité $Q[A] = E_P[\mathbf{1}_A D_S]$ entraîne que $Q[A] = 0$ donc aussi $P[A] = 0$. $\square$

Proposition 5.16 *Soit D une martingale locale (continue) strictement positive. Il existe alors une unique martingale locale continue L telle que*

$$D_t = \exp(L_t - \frac{1}{2}\langle L, L\rangle_t) = \mathcal{E}(L)_t.$$

De plus L est donnée par la formule

$$L_t = \log D_0 + \int_0^t D_s^{-1}\, dD_s.$$

Démonstration. L'unicité est une conséquence immédiate du Théorème 4.5. Ensuite, puisque D est strictement positive, on peut appliquer la formule d'Itô à $\log D_t$, et il vient

$$\log D_t = \log D_0 + \int_0^t \frac{dD_s}{D_s} - \frac{1}{2}\int_0^t \frac{d\langle D,D\rangle_s}{D_s^2} = L_t - \frac{1}{2}\langle L,L\rangle_t.$$

□

Théorème 5.17 (Girsanov) *Soit Q une mesure de probabilité équivalente à P sur la tribu $\mathcal{F}_\infty$. Soit D la martingale associée à Q par la Proposition 5.15. On suppose que D est continue. Soit L la martingale locale associée à D par la Proposition 5.16. Alors, si M est une $(\mathcal{F}_t, P)$-martingale locale continue, le processus*

$$\widetilde{M} = M - \langle M, L\rangle$$

est une $(\mathcal{F}_t, Q)$-martingale locale continue.

Démonstration. Montrons d'abord que, si T est un temps d'arrêt et si X est un processus continu adapté tel que $(XD)^T$ est une P-martingale, alors X^T est une Q-martingale. Puisque, d'après la Proposition 5.15, $E_Q[|X_{T\wedge t}|] = E_P[|X_{T\wedge t}D_{T\wedge t}|] < \infty$, on a d'abord $X_t^T \in L^1(Q)$. Ensuite, soient $A \in \mathcal{F}_s$ et $s < t$. Puisque $A \cap \{T > s\} \in \mathcal{F}_s$, on a, en utilisant le fait que $(XD)^T$ est une P-martingale,

$$E_P[\mathbf{1}_{A\cap\{T>s\}} X_{T\wedge t} D_{T\wedge t}] = E_P[\mathbf{1}_{A\cap\{T>s\}} X_{T\wedge s} D_{T\wedge s}].$$

D'après la Proposition 5.15,

$$D_{T\wedge t} = \frac{dQ}{dP}_{|\mathcal{F}_{T\wedge t}}, \quad D_{T\wedge s} = \frac{dQ}{dP}_{|\mathcal{F}_{T\wedge s}},$$

et donc, puisque $A \cap \{T > s\} \in \mathcal{F}_{T\wedge s} \subset \mathcal{F}_{T\wedge t}$, il vient

$$E_Q[\mathbf{1}_{A\cap\{T>s\}} X_{T\wedge t}] = E_Q[\mathbf{1}_{A\cap\{T>s\}} X_{T\wedge s}].$$

D'autre part, il est immédiat que

$$E_Q[\mathbf{1}_{A\cap\{T\le s\}} X_{T\wedge t}] = E_Q[\mathbf{1}_{A\cap\{T\le s\}} X_{T\wedge s}].$$

En combinant avec ce qui précède, on obtient $E_Q[\mathbf{1}_A X_{T\wedge t}] = E_Q[\mathbf{1}_A X_{T\wedge s}]$, d'où le résultat annoncé.

Comme conséquence immédiate, on voit que si XD est une P-martingale locale, alors X est une Q-martingale locale.

Soit maintenant M une P-martingale locale continue. On applique ce qui précède à $X = \widetilde{M}$, en remarquant que d'après la formule d'Itô,

$$\begin{aligned}\widetilde{M}_t D_t &= M_0 D_0 + \int_0^t \widetilde{M}_s\, dD_s + \int_0^t D_s\, dM_s - \int_0^t D_s\, d\langle M, L\rangle_s + \langle M, D\rangle_t \\ &= M_0 D_0 + \int_0^t \widetilde{M}_s\, dD_s + \int_0^t D_s\, dM_s\end{aligned}$$

puisque $d\langle M, L\rangle_s = D_s^{-1} d\langle M, D\rangle_s$ d'après la Proposition 5.16. On voit ainsi que $\widetilde{M}D$ est une P-martingale locale, donc $\widetilde{M}$ est une Q-martingale locale. $\square$

Conséquences.

(a) Une $(\mathcal{F}_t, P)$ martingale locale continue M reste une $(\mathcal{F}_t, Q)$-semimartingale continue, dont la décomposition est $M = \widetilde{M} + \langle M, L\rangle$. On voit ainsi que la classe des $(\mathcal{F}_t, P)$-semimartingales continues est contenue dans la classe des $(\mathcal{F}_t, Q)$ semimartingales continues.

En fait ces deux classes coïncident. En effet, sous les hypothèses du Théorème 5.17, P et Q jouent des rôles symétriques. Pour le voir, appliquons le Théorème 5.17 à $M = -L$. On voit que $-\widetilde{L} = -L + \langle L, L\rangle$ est une martingale locale continue, et $\langle \widetilde{L}, \widetilde{L}\rangle = \langle L, L\rangle$. Donc,

$$\mathcal{E}(-\widetilde{L})_t = \exp(-L_t + \langle L, L\rangle_t - \frac{1}{2}\langle L, L\rangle_t) = \Big(\mathcal{E}(L)_t\Big)^{-1} = D_t^{-1}.$$

Cela montre que sous les hypothèses du Théorème 5.17, on peut échanger les rôles de P et Q quitte à remplacer D par D^{-1} et L par $-\widetilde{L}$.

(b) Soient X et Y deux semimartingales continues (relativement à P ou à Q). La valeur du crochet $\langle X, Y\rangle$ est la même sous les deux probabilités P et Q. En effet, ce crochet est toujours donné par l'approximation de la Proposition 4.11 (cette observation a été utilisée implicitement dans (a) ci-dessus).

De même, si H est un processus localement borné, l'intégrale stochastique $H \cdot X$ est la même sous P et sous Q (pour le voir, utiliser l'approximation par des processus élémentaires).

Notons, toujours sous les hypothèses du Théorème 5.17, $\widetilde{M} = \mathcal{G}_Q^P(M)$. L'application $\mathcal{G}_Q^P$ envoie l'ensemble des P-martingales locales continues dans l'ensemble des Q-martingales locales continues. On vérifie facilement que $\mathcal{G}_P^Q \circ \mathcal{G}_Q^P = \mathrm{Id}$. De plus, la transformation $\mathcal{G}_Q^P$ commute avec l'intégrale stochastique: si H est un processus localement borné, $H \cdot \mathcal{G}_Q^P(M) = \mathcal{G}_Q^P(H \cdot M)$.

(c) Si $M = B$ est un $(\mathcal{F}_t, P)$-mouvement brownien, alors $\widetilde{B} = B - \langle B, L\rangle$ est une $(\mathcal{F}_t, Q)$-martingale locale de variation quadratique $\langle \widetilde{B}, \widetilde{B}\rangle_t = \langle B, B\rangle_t = t$. Donc, d'après le Théorème 5.9, $\widetilde{B}$ est un $(\mathcal{F}_t, Q)$-mouvement brownien.

(d) On utilise souvent le Théorème de Girsanov “à horizon fini”. Pour $T > 0$ fixé, on se donne une filtration $(\mathcal{F}_t)_{t\in[0,T]}$ indexée par $t \in [0,T]$ au lieu de $t \in [0,\infty]$. On suppose que cette filtration satisfait les conditions habituelles (l'hypothèse de complétion signifiant que chaque tribu $\mathcal{F}_t$ contient les P-négligeables de $\mathcal{F}_T$). Si Q est une autre probabilité équivalente à P sur $\mathcal{F}_T$, on définit comme ci-dessus la martingale $(D_t, t \in [0,T])$ et, si D a une modification continue, la martingale $(L_t, t \in [0,T])$. L'analogue du Théorème 5.17 reste alors bien sûr vrai.

Dans les applications du Théorème de Girsanov, on construit la probabilité Q de la manière suivante. On part d'une martingale locale continue L telle que $L_0 = 0$. Alors $\mathcal{E}(L)_t$ est une martingale locale continue à valeurs strictement positives, donc une surmartingale (Proposition 4.4) ce qui assure l'existence de la limite $\mathcal{E}(L)_\infty$ avec, d'après le lemme de Fatou, $E[\mathcal{E}(L)_\infty] \leq 1$. Si on a

$$E[\mathcal{E}(L)_\infty] = 1 \tag{5.5}$$

alors il est facile de voir (exercice) que $\mathcal{E}(L)$ est en fait une vraie martingale uniformément intégrable. En posant $Q = \mathcal{E}(L)_\infty \cdot P$, on est dans le cadre du Théorème 5.17. Il est donc très important de pouvoir donner des conditions qui assurent l'égalité (5.5).

Théorème 5.18 *Soit L une martingale locale continue telle que $L_0 = 0$. Considérons les propriétés suivantes*

(i) $E[\exp \frac{1}{2}\langle L, L\rangle_\infty] < \infty$*;*

(ii) *L est une martingale uniformément intégrable, et* $E[\exp \frac{1}{2}L_\infty] < \infty$*;*

(iii) *$\mathcal{E}(L)$ est une martingale uniformément intégrable.*

Alors, (i)$\Rightarrow$(ii)$\Rightarrow$(iii).

Démonstration. (i)$\Rightarrow$(ii) La propriété (i) entraîne que $E[\langle L, L\rangle_\infty] < \infty$ donc aussi que L est une vraie martingale bornée dans L^2 (Théorème 4.7). Ensuite,

$$\exp \frac{1}{2} L_\infty = \mathcal{E}(L)_\infty^{1/2} \exp(\frac{1}{2}\langle L, L\rangle_\infty)^{1/2}$$

d'où grâce à l'inégalité de Cauchy-Schwarz,

$$E[\exp \frac{1}{2}L_\infty] \leq E[\mathcal{E}(L)_\infty]^{1/2} E[\exp(\frac{1}{2}\langle L, L\rangle_\infty)]^{1/2} \leq E[\exp(\frac{1}{2}\langle L, L\rangle_\infty)]^{1/2} < \infty.$$

(ii)$\Rightarrow$(iii) Puisque L est une martingale uniformément intégrable, on a $L_t = E[L_\infty \mid \mathcal{F}_t]$ d'où

$$\exp \frac{1}{2} L_t \leq E[\exp \frac{1}{2} L_\infty \mid \mathcal{F}_t].$$

Ceci montre d'abord que $\exp \frac{1}{2} L_t \in L^1$. Ensuite, par convexité, $\exp \frac{1}{2} L_t$ est une sous-martingale, et l'inégalité précédente montre que cette sous-martingale est fermée par sa limite $\exp \frac{1}{2} L_\infty$. D'après le théorème d'arrêt pour les sur(sous)martingales fermées (Théorème 3.12) on a pour tout temps d'arrêt T, $\exp \frac{1}{2} L_T \leq E[\exp \frac{1}{2} L_\infty \mid \mathcal{F}_T]$, ce qui montre que la famille $\{\exp \frac{1}{2} L_T,\ T$ temps d'arrêt$\}$ est uniformément intégrable.

Pour $0 < a < 1$, posons $Z_t^{(a)} = \exp(\frac{aL_t}{1+a})$. Alors, on vérifie facilement que

$$\mathcal{E}(aL)_t = (\mathcal{E}(L)_t)^{a^2} \, (Z_t^{(a)})^{1-a^2}.$$

Si $\Gamma \in \mathcal{F}_\infty$ et T est un temps d'arrêt, l'inégalité de Hölder donne

$$E[\mathbf{1}_\Gamma \mathcal{E}(aL)_T] \leq E[\mathcal{E}(L)_T]^{a^2} E[\mathbf{1}_\Gamma Z_T^{(a)}]^{1-a^2} \leq E[\mathbf{1}_\Gamma Z_T^{(a)}]^{1-a^2} \leq E[\mathbf{1}_\Gamma \exp \frac{1}{2} L_T]^{2a(1-a)},$$

en utilisant le fait que $\mathcal{E}(L)$ est une surmartingale positive pour la deuxième inégalité, puis l'inégalité de Jensen. Comme la famille des $\{\exp \frac{1}{2} L_T,\ T$ temps d'arrêt$\}$ est uniformément intégrable, l'inégalité précédente montre que la famille des $\{\mathcal{E}(aL)_T,\ T$ temps d'arrêt$\}$ l'est aussi. Cela entraîne facilement que $\mathcal{E}(aL)$ est une (vraie) martingale uniformément intégrable. Il en découle que

$$1 = E[\mathcal{E}(aL)_\infty] \leq E[\mathcal{E}(L)_\infty]^{a^2} E[Z_\infty^{(a)}]^{1-a^2} \leq E[\mathcal{E}(L)_\infty]^{a^2} E[\exp \frac{1}{2} L_\infty]^{2a(1-a)}.$$

Lorsque $a \to 1$, cette dernière inégalité entraîne $E[\mathcal{E}(L)_\infty] \geq 1$ d'où $E[\mathcal{E}(L)_\infty] = 1$. $\square$

Chapitre 6

Equations Différentielles Stochastiques

6.1 Motivation et définitions générales

Le but des équations différentielles stochastiques est de fournir un modèle mathématique pour une équation différentielle perturbée par un bruit aléatoire. Partons d'une équation différentielle ordinaire de la forme

$$y'(t) = b(y(t)),$$

soit encore sous forme différentielle,

$$dy_t = b(y_t)\, dt.$$

Une telle équation est utilisée pour décrire l'évolution d'un système physique. Si l'on prend en compte les perturbations aléatoires, on ajoute un terme de bruit, qui sera de la forme $\sigma\, dB_t$, où B désigne un mouvement brownien et σ est pour l'instant une constante qui correspond à l'intensité du bruit. On arrive à une équation différentielle "stochastique" de la forme

$$dy_t = b(y_t)\, dt + \sigma\, dB_t,$$

ou encore sous forme intégrale, la seule qui ait un sens mathématique,

$$y_t = y_0 + \int_0^t b(y_s)\, ds + \sigma\, B_t.$$

On généralise cette équation en autorisant σ à dépendre de l'état à l'instant t:

$$dy_t = b(y_t)\, dt + \sigma(y_t)\, dB_t,$$

soit sous forme intégrale,

$$y_t = y_0 + \int_0^t b(y_s)\,ds + \int_0^t \sigma(y_s)\,dB_s.$$

Remarquons que le sens donné à cette équation dépend de la théorie de l'intégrale stochastique développée dans le chapitre précédent. On généralise encore un peu en autorisant σ et b à dépendre du temps t, et en se plaçant dans un cadre vectoriel. Cela conduit à la définition suivante.

Définition 6.1 *Soient d et m des entiers positifs, et soient σ et b des fonctions mesurables localement bornées définies sur $\mathbb{R}_+ \times \mathbb{R}^d$ et à valeurs respectivement dans $M_{d\times m}(\mathbb{R})$ et $\mathbb{R}^d$, où $M_{d\times m}(\mathbb{R})$ désigne l'ensemble des matrices $d \times m$ à coefficients réels. On note $\sigma = (\sigma_{ij})_{1\le i\le d, 1\le j\le m}$ et $b = (b_i)_{1\le i\le d}$.*

Une solution de l'équation

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

est la donnée de

- *un espace de probabilité filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ satisfaisant les conditions habituelles;*
- *sur cet espace un $(\mathcal{F}_t)$-mouvement brownien dans $\mathbb{R}^m$, $B = (B^1, \dots, B^m)$;*
- *un processus $(\mathcal{F}_t)$-adapté continu $X = (X^1, \dots, X^d)$ à valeurs dans $\mathbb{R}^d$, tel que*

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

soit encore, coordonnée par coordonnée, pour tout $i \in \{1, \dots, d\}$,

$$X_t^i = X_0^i + \sum_{j=1}^m \int_0^t \sigma_{ij}(s, X_s)\,dB_s^j + \int_0^t b_i(s, X_s)\,ds.$$

Lorsque de plus $X_0 = x \in \mathbb{R}^d$, on dira que le processus X est solution de $E_x(\sigma, b)$.

Il existe plusieurs notions d'existence et d'unicité pour les équations différentielles stochastiques.

Définition 6.2 *On dit pour l'équation $E(\sigma, b)$ qu'il y a*

- *existence faible si pour tout $x \in \mathbb{R}^d$ il existe une solution de $E_x(\sigma, b)$;*
- *existence et unicité faibles si de plus toutes les solutions de $E_x(\sigma, b)$ ont même loi;*

· *unicité trajectorielle si, l'espace de probabilité filtré* $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ *et le mouvement brownien* B *étant fixés, deux solutions* X *et* X' *telles que* $X_0 = X'_0$ *p.s. sont indistinguables.*

On dit de plus qu'une solution X *de* $E_x(\sigma, b)$ *est une solution forte si* X *est adapté par rapport à la filtration canonique de* B*.*

Remarque. Il peut y avoir existence et unicité faibles sans qu'il y ait unicité trajectorielle. L'exemple le plus simple est obtenu en partant d'un mouvement brownien réel β issu de $\beta_0 = y$, et en posant

$$B_t = \int_0^t \mathrm{sgn}\,(\beta_s)\, d\beta_s,$$

avec $\mathrm{sgn}\,(x) = 1$ si $x \geq 0$, -1 si $x < 0$. Alors on a aussi

$$\beta_t = y + \int_0^t \mathrm{sgn}\,(\beta_s)\, dB_s.$$

De plus le théorème de Lévy (Théorème 5.9) montre que B est aussi un mouvement brownien (issu de 0). On voit ainsi que β est solution de l'équation différentielle stochastique

$$dX_t = \mathrm{sgn}\,(X_s)\, dB_s, \qquad X_0 = y,$$

pour laquelle il y a donc existence faible. A nouveau le théorème de Lévy montre que n'importe quelle solution de cette équation doit être un mouvement brownien, ce qui donne l'unicité faible.

En revanche, il n'y a pas unicité trajectorielle pour cette équation. En effet, on voit facilement, dans le cas $\beta_0 = 0$, que β et $-\beta$ sont deux solutions issues de 0 correspondant au même mouvement brownien B (remarquer que $\int_0^t \mathbf{1}_{\{\beta_s = 0\}}\, ds = 0$, ce qui entraîne $\int_0^t \mathbf{1}_{\{\beta_s = 0\}}\, dB_s = 0$). On peut aussi voir que β n'est pas solution forte de l'équation: on montre que la filtration canonique de B coïncide avec la filtration canonique de $|\beta|$, qui est strictement plus petite que celle de β.

Le théorème suivant, que nous n'utiliserons pas dans la suite, relie les différentes notions d'existence et d'unicité.

Théorème [Yamada-Watanabe] *S'il y a existence faible et unicité trajectorielle, alors il y a aussi unicité faible. De plus, pour tout espace de probabilité filtré* $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ *et tout* $(\mathcal{F}_t)$*-mouvement brownien* B*, il existe pour chaque* $x \in \mathbb{R}^d$ *une (unique) solution forte de* $E_x(\sigma, b)$*.*

6.2 Le cas lipschitzien

Dans ce paragraphe nous nous plaçons sous les hypothèses suivantes.
Hypothèses. Les fonctions σ et b sont continues sur $\mathbb{R}_+ \times \mathbb{R}^d$ et lipschitziennes en la variable x: il existe une constante K telle que, pour tous $t \geq 0$, $x, y \in \mathbb{R}^d$,

$$|\sigma(t,x) - \sigma(t,y)| \leq K\,|x-y|,$$
$$|b(t,x) - b(t,y)| \leq K\,|x-y|.$$

Théorème 6.1 *Sous les hypothèses précédentes, il y a unicité trajectorielle pour $E(\sigma, b)$. De plus, pour tout espace de probabilité filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ et tout $(\mathcal{F}_t)$-mouvement brownien B, il existe pour chaque $x \in \mathbb{R}^d$ une (unique) solution forte de $E_x(\sigma, b)$.*

Le théorème entraîne en particulier qu'il y a existence faible pour $E(\sigma, b)$. L'unicité faible découlera du théorème suivant (elle est aussi une conséquence de l'unicité trajectorielle si on utilise le théorème de Yamada-Watanabe).

Remarque. On peut affaiblir l'hypothèse de continuité en la variable t, qui n'intervient essentiellement que pour majorer $\sup_{0 \leq t \leq T} |\sigma(t,x)|$ et $\sup_{0 \leq t \leq T} |b(t,x)|$ pour x fixé. On peut aussi "localiser" l'hypothèse sur le caractère lipschitzien de σ et b (la constante K dépendra du compact sur lequel on considère t et x, y), à condition de conserver une condition de croissance linéaire

$$|\sigma(t,x)| \leq K(1+|x|), \quad |b(t,x)| \leq K(1+|x|).$$

Ce type de condition, qui sert à éviter l'explosion de la solution, intervient déjà dans les équations différentielles ordinaires.

Démonstration. Pour alléger les notations, on traite seulement le cas $d = m = 1$. Commençons par établir l'unicité trajectorielle. On se donne (sur le même espace, avec le même mouvement brownien B) deux solutions X et X' telles que $X_0 = X'_0$. Pour $M > 0$ fixé, posons

$$\tau = \inf\{t \geq 0,\ |X_t| \geq M \text{ ou } |X'_t| \geq M\}.$$

On a alors, pour tout $t \geq 0$,

$$X_{t\wedge\tau} = X_0 + \int_0^{t\wedge\tau} \sigma(s, X_s)\, dB_s + \int_0^{t\wedge\tau} b(s, X_s)\, ds$$

et on a une équation analogue pour $X'_{t\wedge\tau}$. En faisant la différence entre ces deux équations, et en remarquant que X et X' sont bornées par M sur l'intervalle $]0, \tau]$, on trouve si $t \in [0,T]$, $T > 0$,

$$
\begin{aligned}
&E[(X_{t\wedge\tau} - X'_{t\wedge\tau})^2] \\
&\leq 2\left(E\Big[\Big(\int_0^{t\wedge\tau}(\sigma(s,X_s)-\sigma(s,X'_s))dB_s\Big)^2\Big] + E\Big[\Big(\int_0^{t\wedge\tau}(b(s,X_s)-b(s,X'_s))ds\Big)^2\Big]\right) \\
&\leq 2\left(E\Big[\int_0^{t\wedge\tau}(\sigma(s,X_s)-\sigma(s,X'_s))^2 ds\Big] + T\,E\Big[\int_0^{t\wedge\tau}(b(s,X_s)-b(s,X'_s))^2 ds\Big]\right) \\
&\leq 2K^2(1+T)\,E\Big[\int_0^{t\wedge\tau}(X_s - X'_s)^2 ds\Big] \\
&\leq 2K^2(1+T)\,E\Big[\int_0^{t}(X_{s\wedge\tau} - X'_{s\wedge\tau})^2 ds\Big]
\end{aligned}
$$

Donc la fonction $h(t) = E[(X_{t\wedge\tau} - X'_{t\wedge\tau})^2]$ vérifie pour $t \in [0,T]$

$$h(t) \leq C \int_0^t h(s)\,ds$$

avec $C = 2K^2(1+T)$.

Lemme 6.2 *Soit $T > 0$ et soit g une fonction positive mesurable bornée sur l'intervalle $[0,T]$. Supposons qu'il existe deux constantes $a \geq 0$, $b \geq 0$ telles que pour tout $t \in [0,T]$,*

$$g(t) \leq a + b\int_0^t g(s)\,ds.$$

Alors on a pour tout $t \in [0,T]$,

$$g(t) \leq a\,\exp(bt).$$

Démonstration du lemme. En itérant la condition sur g, on trouve que pour tout $n \geq 1$,

$$g(t) \leq a + a(bt) + a\frac{(bt)^2}{2} + \cdots + a\frac{(bt)^n}{n!} + b^{n+1}\int_0^t ds_1 \int_0^{s_1} ds_2 \cdots \int_0^{s_n} ds_{n+1} g(s_{n+1}).$$

Si g est majorée par A, le dernier terme ci-dessus est majoré par $A(bt)^{n+1}/(n+1)!$, donc tend vers 0 quand $n \to \infty$. Le résultat recherché en découle. □

Revenons à la preuve du théorème. La fonction h est bornée par $4M^2$ et vérifie l'hypothèse du lemme avec $a = 0$, $b = C$. On obtient donc $h = 0$, soit $X_{t\wedge\tau} = X'_{t\wedge\tau}$. En faisant tendre M vers ∞, on a $X_t = X'_t$ ce qui achève la preuve de l'unicité trajectorielle.

Pour la deuxième assertion, nous construisons la solution par la méthode d'approximation de Picard. On pose

$$
\begin{aligned}
X_t^0 &= x, \\
X_t^1 &= x + \int_0^t \sigma(s,x)\,dB_s + \int_0^t b(s,x)\,ds, \\
X_t^n &= x + \int_0^t \sigma(s,X_s^{n-1})\,dB_s + \int_0^t b(s,X_s^{n-1})\,ds.
\end{aligned}
$$

Les intégrales stochastiques ci-dessus sont bien définies puisqu'il est clair par récurrence que pour chaque n, X_t^n est continu et adapté, donc le processus $\sigma(t,X_t^n)$ l'est aussi.

Fixons un réel $T>0$, et raisonnons sur l'intervalle $[0,T]$. Vérifions d'abord par récurrence sur n qu'il existe une constante C_n telle que pour tout $t\in[0,T]$,

$$
E[(X_t^n)^2] \leq C_n. \tag{6.1}
$$

Cette majoration est triviale si $n=0$. Ensuite, si elle est vraie à l'ordre $n-1$, on utilise les majorations

$$
|\sigma(s,y)| \leq K' + K|y|, \quad |b(s,y)| \leq K' + K|y|, \qquad \forall s\in[0,T],\ y\in\mathbb{R},
$$

pour écrire

$$
\begin{aligned}
E[(X_t^n)^2] &\leq 3\left(|x|^2 + E\Big[\Big(\int_0^t \sigma(s,X_s^{n-1})\,dB_s\Big)^2\Big] + E\Big[\Big(\int_0^t b(s,X_s^{n-1})\,ds\Big)^2\Big]\right) \\
&\leq 3\left(|x|^2 + E\Big[\int_0^t \sigma(s,X_s^{n-1})^2\,ds\Big] + t\,E\Big[\int_0^t b(s,X_s^{n-1})^2\,ds\Big]\right) \\
&\leq 3\left(|x|^2 + 4(1+T)E\Big[\int_0^t (K'^2 + K^2(X_s^{n-1})^2)ds\Big]\right) \\
&\leq 3(|x|^2 + 4T(1+T)(K'^2 + K^2 C_{n-1})) =: C_n
\end{aligned}
$$

Pour justifier le calcul du moment d'ordre deux de l'intégrale stochastique, on a utilisé le fait que $E[\int_0^t \sigma(s,X_s^{n-1})^2\,ds] < \infty$, ce qui découle de la majoration ci-dessus pour σ et de l'hypothèse de récurrence.

La majoration (6.1), et l'hypothèse sur σ entraînent que la martingale locale $\int_0^t \sigma(s,X_s^n)\,dB_s$ est pour chaque n une vraie martingale bornée dans L^2 sur l'intervalle $[0,T]$. Nous utilisons cette remarque pour majorer par récurrence

$$
E\Big[\sup_{0\leq t\leq T} |X_t^{n+1} - X_t^n|^2\Big].
$$

On a

$$X_t^{n+1} - X_t^n = \int_0^t (\sigma(s, X_s^n) - \sigma(s, X_s^{n-1}))\, dB_s + \int_0^t (b(s, X_s^n) - b(s, X_s^{n-1}))\, ds,$$

d'où, en utilisant l'inégalité de Doob (Proposition 3.6) à la troisième ligne,

$$\begin{aligned}
&E\Big[\sup_{0\le s\le t} |X_s^{n+1} - X_s^n|^2\Big] \\
&\le 2\, E\Big[\sup_{0\le s\le t} \Big|\int_0^s (\sigma(u, X_u^n) - \sigma(u, X_u^{n-1}))\, dB_u\Big|^2 + \sup_{0\le s\le t} \Big|\int_0^s (b(u, X_u^n) - b(u, X_u^{n-1}))\, du\Big|^2\Big] \\
&\le 2\Big(4\, E\Big[\Big(\int_0^t (\sigma(u, X_u^n) - \sigma(u, X_u^{n-1}))\, dB_u\Big)^2\Big] + E\Big[\Big(\int_0^t |b(u, X_u^n) - b(u, X_u^{n-1}|\, du\Big)^2\Big]\Big) \\
&\le 2\Big(4\, E\Big[\int_0^t (\sigma(u, X_u^n) - \sigma(u, X_u^{n-1}))^2\, du\Big] + T\, E\Big[\int_0^t (b(u, X_u^n) - b(u, X_u^{n-1})^2\, du\Big]\Big) \\
&\le 2(4+T)K^2\, E\Big[\int_0^t |X_u^n - X_u^{n-1}|^2\, du\Big] \\
&\le C_T\, E\Big[\int_0^t \sup_{0\le r\le u} |X_r^n - X_r^{n-1}|^2\, du\Big]
\end{aligned}$$

en notant $C_T = 2(4+T)K^2$. Si $g_n(u) = E[\sup_{0\le r\le u} |X_r^n - X_r^{n-1}|^2]$, on voit donc que

$$g_{n+1}(t) \le C_T \int_0^t g_n(u)\, du. \tag{6.2}$$

D'autre part, (6.1) et les inégalités précédentes montrent que chacune des fonctions g_n est bornée sur $[0,T]$. En particulier, il existe une constante C'_T telle que $g_0(t) \le C'_T$ pour $t \in [0,T]$. Une récurrence simple utilisant (6.2) montre alors que pour tous $n \ge 0$, $t \in [0,T]$,

$$g_n(t) \le C'_T (C_T)^n \frac{t^n}{n!}.$$

En particulier, $\sum_{n=0}^\infty g_n(T)^{1/2} < \infty$, ce qui entraîne que p.s.

$$\sum_{n=0}^\infty \sup_{0\le t\le T} |X_t^{n+1} - X_t^n| < \infty,$$

et donc p.s. la suite $(X_t^n, 0 \le t \le T)$ converge uniformément sur $[0,T]$ vers un processus limite $(X_t, 0 \le t \le T)$ qui est adapté et continu. On vérifie par récurrence que chaque processus X^n est adapté par rapport à la filtration canonique de B, et donc X l'est aussi.

Enfin, les estimations précédentes montrent aussi que

$$E\Big[\sup_{0\leq s\leq T}|X_s^n - X_s|^2\Big] \leq \Big(\sum_{k=n}^{\infty} g_k(T)^{1/2}\Big)^2 \longrightarrow 0$$

et on en déduit aussitôt que

$$\int_0^t \sigma(s,X_s)\,dB_s = \lim_{n\to\infty}\int_0^t \sigma(s,X_s^n)\,dB_s$$
$$\int_0^t b(s,X_s)\,ds = \lim_{n\to\infty}\int_0^t b(s,X_s^n)\,ds,$$

dans L^2. En passant à la limite dans l'équation de récurrence pour X^n, on trouve que X est solution (forte) de $E_x(\sigma,b)$ sur $[0,T]$. $\square$

Dans l'énoncé suivant, $W(d\mathrm{w})$ désigne la mesure de Wiener sur l'espace canonique $C(\mathbb{R}_+,\mathbb{R}^m)$ des fonctions continues de $\mathbb{R}_+$ dans $\mathbb{R}^m$ ($W(d\mathrm{w})$ est la loi de $(B_t, t\geq 0)$ si B est un mouvement brownien en dimension m issu de 0).

Théorème 6.3 *Sous les hypothèses du théorème précédent, il existe pour chaque $x\in\mathbb{R}$ une application $F_x : C(\mathbb{R}_+,\mathbb{R}^m) \longrightarrow C(\mathbb{R}_+,\mathbb{R}^d)$ mesurable lorsque $C(\mathbb{R}_+,\mathbb{R}^m)$ est muni de la tribu borélienne complétée par les ensembles W-négligeables, telle que les propriétés suivantes soient vérifiées:*

(i) *pour tout $t\geq 0$, $F_x(\mathrm{w})_t$ coïncide $W(d\mathrm{w})$ p.s. avec une fonction mesurable de $(\mathrm{w}(r), 0\leq r\leq t)$;*

(ii) *pour tout $\mathrm{w}\in C(\mathbb{R}_+,\mathbb{R}^m)$, l'application $x\longrightarrow F_x(\mathrm{w})$ est continue;*

(iii) *pour tout $x\in\mathbb{R}^d$, pour tout choix de l'espace de probabilité filtré $(\Omega,\mathcal{F},(\mathcal{F}_t),P)$ et du $(\mathcal{F}_t)$-mouvement brownien B en dimension m, le processus X_t défini par $X_t = F_x(B)_t$ est la solution unique de $E(\sigma,b)$ avec valeur initiale x; de plus si Z est une variable aléatoire $\mathcal{F}_0$-mesurable, le processus $F_Z(B)_t$ est la solution unique avec valeur initiale Z.*

Remarque. L'assertion (iii) montre en particulier qu'il y a unicité en loi pour $E(\sigma,b)$: les solutions de $E_x(\sigma,b)$ sont toutes de la forme $F_x(B)$ et ont donc la même loi qui est la mesure image de $W(d\mathrm{w})$ par F_x.

Démonstration. A nouveau on traite le cas $d=m=1$. Notons $\mathcal{N}$ la classe des sous-ensembles W-négligeables de $C(\mathbb{R}_+,\mathbb{R})$, et pour tout $t\in[0,\infty]$,

$$\mathcal{G}_t = \sigma(\mathrm{w}(s), 0\leq s\leq t)\vee\mathcal{N}.$$

D'après la remarque de la fin du paragraphe 5.3 (appliquée au mouvement brownien canonique $B_t(\mathrm{w}) = \mathrm{w}(t)$ sous W), la filtration $(\mathcal{G}_t)$ est continue à droite, donc satisfait les conditions habituelles. Pour chaque $x \in \mathbb{R}$, notons X^x la solution de $E_x(\sigma, b)$ associée à l'espace $(C(\mathbb{R}_+, \mathbb{R}), \mathcal{G}_\infty, (\mathcal{G}_t), W)$ et au mouvement brownien $B_t(\mathrm{w}) = \mathrm{w}(t)$. Cette solution existe et est unique à indistinguabilité près d'après le Théorème 6.1.

Soient $x, y \in \mathbb{R}$ et T_n le temps d'arrêt défini par

$$T_n = \inf\{t \geq 0,\ |X_t^x| \geq n \text{ ou } |X_t^y| \geq n\}.$$

Soit $p \geq 2$ et $T \geq 1$. En utilisant les inégalités de Burkholder-Davis-Gundy (Théorème 5.12) puis l'inégalité de Hölder on obtient pour $t \in [0, T]$

$$\begin{aligned}
&E\Big[\sup_{s\leq t} |X^x_{s\wedge T_n} - X^y_{s\wedge T_n}|^p\Big] \\
&\leq C_p\Big(|x-y|^p + E\Big[\sup_{s\leq t}\Big|\int_0^{s\wedge T_n} (\sigma(r, X_r^x) - \sigma(r, X_r^y))dB_r\Big|^p\Big] \\
&\qquad + E\Big[\sup_{s\leq t}\Big|\int_0^{s\wedge T_n} (b(r, X_r^x) - b(r, X_r^y))dr\Big|^p\Big]\Big) \\
&\leq C_p\Big(|x-y|^p + C'_p E\Big[\Big(\int_0^{t\wedge T_n} (\sigma(r, X_r^x) - \sigma(r, X_r^y))^2 dr\Big)^{p/2}\Big] \\
&\qquad + E\Big[\Big(\int_0^{t\wedge T_n} |b(r, X_r^x) - b(r, X_r^y)| dr\Big)^p\Big]\Big) \\
&\leq C_p\Big(|x-y|^p + C'_p t^{\frac{p}{2}-1} E\Big[\int_0^t |\sigma(r\wedge T_n, X^x_{r\wedge T_n}) - \sigma(r\wedge T_n, X^y_{r\wedge T_n})|^p dr\Big] \\
&\qquad + t^{p-1}\, E\Big[\int_0^t |b(r\wedge T_n, X^x_{r\wedge T_n}) - b(r\wedge T_n, X^y_{r\wedge T_n})|^p dr\Big]\Big) \\
&\leq C''_p\Big(|x-y|^p + T^p \int_0^t E[|X^x_{r\wedge T_n} - X^y_{r\wedge T_n}|^p]\, dr\Big),
\end{aligned}$$

où la constante $C''_p < \infty$ dépend de p (et de la constante K intervenant dans l'hypothèse sur σ et b) mais pas de n ni de x, y et T.

Puisque la fonction $t \longrightarrow E\Big[\sup_{s\leq t} |X^x_{s\wedge T_n} - X^y_{s\wedge T_n}|^p\Big]$ est évidemment bornée, le Lemme 6.2 entraîne alors pour $t \in [0, T]$

$$E\Big[\sup_{s\leq t} |X^x_{s\wedge T_n} - X^y_{s\wedge T_n}|^p\Big] \leq C''_p |x-y|^p \exp(C''_p T^p t),$$

d'où en faisant tendre n vers ∞,

$$E\Big[\sup_{s\leq t} |X^x_s - X^y_s|^p\Big] \leq C''_p |x-y|^p \exp(C''_p T^p t).$$

La topologie sur l'espace $C(\mathbb{R}_+,\mathbb{R})$ est définie par la distance

$$d(\mathrm{w},\mathrm{w}') = \sum_{k=1}^{\infty} \alpha_k \Big(\sup_{s\le k} |\mathrm{w}(s) - \mathrm{w}'(s)| \wedge 1 \Big),$$

pour un choix arbitraire de la suite de réels $\alpha_k > 0$, tels que la série $\sum \alpha_k$ soit convergente. On peut choisir les coefficients α_k tels que

$$\sum_{k=1}^{\infty} \alpha_k \, \exp(C''_p k^{p+1}) < \infty.$$

Alors les estimations précédentes et l'inégalité de Jensen montrent que

$$E[d(X^x,X^y)^p] \le (\sum_{k=1}^{\infty} \alpha_k)^{p-1} \sum_{k=1}^{\infty} \alpha_k \, E\Big[\sup_{s\le k} |X^x_s - X^y_s|^p \Big] \le \bar{C}_p \, |x-y|^p.$$

D'après le lemme de Kolmogorov (Théorème 2.5), appliqué au processus $(X^x, x\in\mathbb{R})$ à valeurs dans $E = C(\mathbb{R}_+,\mathbb{R})$ muni de la distance d, il existe une modification à trajectoires continues, notée $(\tilde{X}^x, x\in\mathbb{R})$, du processus $(X^x, x\in\mathbb{R})$. On note $F_x(\mathrm{w}) = \tilde{X}^x(\mathrm{w}) = (\tilde{X}^x_t(\mathrm{w}))_{t\ge 0}$. La propriété (ii) découle alors de ce qui précède.

L'application $\mathrm{w} \longrightarrow F_x(\mathrm{w})$ est mesurable de $C(\mathbb{R}_+,\mathbb{R})$ muni de la tribu $\mathcal{G}_\infty$ dans $C(\mathbb{R}_+,\mathbb{R})$ muni de la tribu borélienne $\mathcal{C} = \sigma(\mathrm{w}(s), s\ge 0)$. De même, pour chaque $t\ge 0$, $F_x(\mathrm{w})_t = \tilde{X}^x_t(\mathrm{w})$ est $\mathcal{G}_t$-mesurable donc coïncide $W(d\mathrm{w})$ p.s. avec une fonction mesurable de $(\mathrm{w}(s), 0\le s\le t)$. On a ainsi obtenu la propriété (i).

Montrons maintenant la première partie de l'assertion (iii). Pour cela, fixons l'espace de probabilité filtré $(\Omega,\mathcal{F},(\mathcal{F})_t,P)$ et le $(\mathcal{F}_t)$-mouvement brownien B. Il faut voir que $F_x(B)$ est alors solution de $E_x(\sigma,b)$. Remarquons déjà que ce processus est (trivialement) continu et aussi adapté puisque $F_x(B)_t$ coïncide p.s. avec une fonction mesurable de $(B_r, 0\le r\le t)$, d'après (i). D'autre part, par construction de F_x on a $W(d\mathrm{w})$ p.s.

$$F_x(\mathrm{w})_t = x + \int_0^t \sigma(s, F_x(\mathrm{w})_s)\, d\mathrm{w}(s) + \int_0^t b(s, F_x(\mathrm{w})_s) ds,$$

et l'intégrale stochastique $\int_0^t \sigma(s,F_x(\mathrm{w})_s) d\mathrm{w}(s)$ peut être définie par

$$\int_0^t \sigma(s,F_x(\mathrm{w})_s)\, d\mathrm{w}(s) = \lim_{k\to\infty} \sum_{i=0}^{2^{n_k}-1} \sigma(\frac{it}{2^{n_k}}, F_x(\mathrm{w})_{it/2^{n_k}})\,(\mathrm{w}(\frac{(i+1)t}{2^{n_k}}) - \mathrm{w}(\frac{it}{2^{n_k}})),$$

$W(d\mathrm{w})$ p.s. le long d'une sous-suite (n_k) bien choisie (d'après la Proposition 5.6). On peut maintenant remplacer w par B (qui a pour loi W !) et trouver p.s.,

$$\begin{aligned} F_x(B)_t &= x + \lim_{k\to\infty} \sum_{i=0}^{2^{n_k}-1} \sigma(\frac{it}{2^{n_k}}, F_x(B)_{it/2^{n_k}})\,(B_{(i+1)t/2^{n_k}} - B_{it/2^{n_k}}) + \int_0^t b(s,F_x(B)_s) ds \\ &= x + \int_0^t \sigma(s,F_x(B)_s) dB_s + \int_0^t b(s,F_x(B)_s) ds, \end{aligned}$$

à nouveau grâce à la Proposition 5.6. On obtient ainsi que $F_x(B)$ est la solution recherchée.

Il reste à établir la seconde partie de l'assertion (iii). On fixe à nouveau l'espace de probabilité filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ et le $(\mathcal{F}_t)$-mouvement brownien B. Soit Z une variable aléatoire $\mathcal{F}_0$-mesurable. Si dans l'équation intégrale stochastique vérifiée par $F_x(B)$ on remplace formellement x par Z, on obtient que $F_Z(B)$ est solution de $E(\sigma, b)$ avec valeur initiale Z. Cependant, ce remplacement formel n'est pas si facile à justifier, et nous allons donc l'expliquer avec soin.

Remarquons d'abord que l'application $(x, \omega) \longrightarrow F_x(B)_t$ est continue par rapport à la variable x et $\mathcal{F}_t$-mesurable par rapport à ω. On en déduit aisément que cette application est mesurable pour la tribu produit $\mathcal{B}(\mathbb{R}) \otimes \mathcal{F}_t$. Comme Z est $\mathcal{F}_0$-mesurable, il en découle facilement que $F_Z(B)_t$ est $\mathcal{F}_t$-mesurable. Donc le processus $F_Z(B)$ est (continu et) adapté. Définissons $G(x, \mathrm{w}) \in C(\mathbb{R}_+, \mathbb{R})$, pour $x \in \mathbb{R}$ et $\mathrm{w} \in C(\mathbb{R}_+, \mathbb{R})$, par l'égalité

$$G(x, \mathrm{w})_t = \int_0^t b(s, F_x(\mathrm{w})_s)\, ds.$$

Soit aussi $H(x, \mathrm{w}) = F_x(\mathrm{w}) - x - G(x, \mathrm{w})$. Par construction on a $W(d\mathrm{w})$ p.s.

$$H(x, \mathrm{w})_t = \int_0^t \sigma(s, F_x(\mathrm{w})_s)\, d\mathrm{w}(s),$$

Donc, si

$$H_n(x, \mathrm{w})_t = \sum_{i=0}^{2^n-1} \sigma(\frac{it}{2^n}, F_x(\mathrm{w})_{it/2^n})\, (\mathrm{w}(\frac{(i+1)t}{2^n}) - \mathrm{w}(\frac{it}{2^n})),$$

la Proposition 5.6 montre que

$$H(x, \mathrm{w})_t = \lim_{n\to\infty} H_n(x, \mathrm{w})_t,$$

en probabilité sous $W(d\mathrm{w})$, pour chaque $x \in \mathbb{R}$. En utilisant le fait que Z et B sont indépendants (parce que Z est $\mathcal{F}_0$-mesurable), on déduit de cette dernière convergence que

$$H(Z, B)_t = \lim_{n\to\infty} H_n(Z, B)_t$$

en probabilité. Toujours grâce à la Proposition 5.6, cette dernière limite est l'intégrale stochastique

$$\int_0^t \sigma(s, F_Z(B)_s)\, dB_s.$$

On a ainsi montré que

$$\int_0^t \sigma(s, F_Z(B)_s)\, dB_s = H(Z, B)_t = F_Z(B)_t - Z - \int_0^t b(s, F_Z(B)_s)\, ds$$

ce qui prouve que $F_Z(B)$ est solution de $E(\sigma, b)$ avec valeur initiale Z. $\square$

6.3 La propriété de Markov forte des solutions d'équations différentielles stochastiques

Dans ce paragraphe, nous conservons les hypothèses du paragraphe 2, et nous supposons de plus que $\sigma(t,y) = \sigma(y)$, $b(t,y) = b(y)$. Pour chaque $x \in \mathbb{R}^d$, on note P_x la loi sur $C(\mathbb{R}_+, \mathbb{R}^d)$ des solutions de $E_x(\sigma, b)$ (la mesure image de la mesure de Wiener par l'application F_x du Théorème 6.3). Le Théorème 6.3 (ii) montre que l'application $x \longrightarrow P_x$ est continue pour la topologie de la convergence étroite. Il en découle facilement par un argument de classe monotone que pour toute fonction Φ borélienne de $C(\mathbb{R}_+, \mathbb{R})$ dans $\mathbb{R}_+$, l'application $x \longrightarrow E_x[\Phi]$ est mesurable.

Théorème 6.4 *Soit X une solution de $E(\sigma, b)$ sur un espace de probabilité filtré $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$. Soit aussi T un temps d'arrêt fini p.s. Alors, si Φ est une fonction borélienne de $C(\mathbb{R}_+, \mathbb{R})$ dans $\mathbb{R}_+$,*

$$E[\Phi(X_{T+t}, t \geq 0) \mid \mathcal{F}_T] = E_{X_T}[\Phi],$$

ou de manière équivalente, pour toute variable positive U $\mathcal{F}_T$-mesurable,

$$E[U\,\Phi(X_{T+t}, t \geq 0)] = E[U\,E_{X_T}[\Phi]].$$

On traduit cet énoncé en disant que le processus X possède la propriété de Markov forte par rapport à la filtration $(\mathcal{F}_t)$: pour tout temps d'arrêt fini T, la loi conditionnelle du "futur" $(X_{T+t}, t \geq 0)$ connaissant le passé $\mathcal{F}_T$ est la loi de X partant de X_T, qui ne dépend que du présent à l'instant T. Dans le cas particulier $\sigma = \mathrm{Id}$, $b = 0$, on retrouve la propriété de Markov forte du mouvement brownien, formulée de manière légèrement différente.

Démonstration. On suppose comme ci-dessus que $m = d = 1$. Notons $B^{(T)}_u = B_{T+u} - B_T$ qui est un mouvement brownien indépendant de $\mathcal{F}_T$ (voir les remarques précédant la Proposition 5.8). Posons aussi $X'_t = X_{T+t}$ et remarquons que le processus X' est adapté par rapport à la filtration $\mathcal{F}'_t = \mathcal{F}_{T+t}$, qui satisfait les conditions habituelles. De plus,

$$X'_t = X_T + \int_T^{T+t} \sigma(X_s)\,dB_s + \int_T^{T+t} b(X_s)\,ds,$$

d'après l'équation satisfaite par X. On a bien sûr:

$$\int_T^{T+t} b(X_s)\,ds = \int_0^t b(X'_s)\,ds.$$

On voudrait faire le même changement de variables dans l'intégrale stochastique. Cela est justifié par le lemme suivant.

Lemme 6.5 *Si h est un processus continu adapté, on a*

$$\int_T^{T+t} h(s,\omega)\,dB_s = \int_0^t h(T+u,\omega)\,dB_u^{(T)}.$$

La preuve du lemme est facile en approchant h par des combinaisons linéaires finies de processus de la forme $h(s,\omega) = \varphi(\omega)\mathbf{1}_{]T+r,T+r']}(s)$, où φ est $\mathcal{F}_{T+r}$-mesurable.

On déduit du Lemme 6.5 que

$$\int_T^{T+t} \sigma(X_s)\,dB_s = \int_0^t \sigma(X'_u)\,dB_u^{(T)},$$

et on a donc

$$X'_t = X_T + \int_0^t \sigma(X'_s)\,dB_s^{(T)} + \int_0^t b(X'_s)\,ds.$$

Remarquons alors que X' est adapté par rapport à la filtration $(\mathcal{F}'_t)$, $B^{(T)}$ est un $(\mathcal{F}'_t)$-mouvement brownien, et X_T est $\mathcal{F}'_0$-mesurable. D'après la dernière assertion du Théorème 6.3, on doit avoir p.s. $X' = F_{X_T}(B^{(T)})$. Le résultat du théorème en découle aisément: puisque X_T est $\mathcal{F}_T$-mesurable et $B^{(T)}$ est indépendant de $\mathcal{F}_T$, on a

$$E[\Phi(X'_t, t \geq 0) \mid \mathcal{F}_T] = E[\Phi(F_{X_T}(B^{(T)}) \mid \mathcal{F}_T] = \int W(d\mathrm{w})\,\Phi(F_{X_T}(\mathrm{w})) = E_{X_T}[\Phi].$$

□

Dans le cadre de la théorie qui sera développée dans le chapitre suivant, le théorème précédent montre que toute solution de $E_x(\sigma, b)$ est un processus de Markov (même fortement markovien), dont le semigroupe est donné par $P_t f(y) = E_y[f(X_t)]$, pour toute fonction f borélienne bornée. On peut vérifier que ce semigroupe possède la propriété de Feller, ce qui donne une autre approche de la propriété de Markov forte. Les solutions d'équations différentielles stochastiques (à coefficients lipschitziens ici) forment une classe particulièrement importante de processus de Markov à trajectoires continues dans $\mathbb{R}^d$.

On appelle procesus de diffusion un processus fortement markovien et à trajectoires continues du type considéré dans le Théorème 6.4. Il est important de pouvoir attacher aux processus de diffusion des martingales (locales) continues. Dans le théorème suivant, on note σ^* la transposée de la matrice σ.

Théorème 6.6 *Pour toute fonction F de classe C^2 sur $\mathbb{R}^d$ posons*

$$LF(x) = \frac{1}{2}\sum_{i=1}^d \sum_{j=1}^d (\sigma\sigma^*)_{ij}(x)\,\frac{\partial^2 F}{\partial x_i \partial x_j}(x) + \sum_{i=1}^d b_i(x)\,\frac{\partial F}{\partial x_i}(x).$$

Alors, si X est une solution de $E(\sigma, b)$,

$$F(X_t) - \int_0^t LF(X_s)\, ds$$

est une martingale locale.

Démonstration. Il suffit d'appliquer la formule d'Itô à $F(X_t)$. $\square$

On traduit le théorème précédent en disant que L est le générateur infinitésimal (étendu) du processus de Markov X. Dans le formalisme du Chapitre 7 ci-dessous, le domaine du générateur de X contient toutes les fonctions de classe C^2 sur $\mathbb{R}^d$ qui tendent vers 0 à l'infini ainsi que leurs dérivées, et la valeur du générateur sur une telle fonction f est Lf (comparer avec le Théorème 7.6). Remarquons que L est un opérateur différentiel du second ordre. Le processus X permet de donner une approche ou une interprétation probabiliste de nombreux résultats analytiques concernant l'opérateur L. Ces liens entre probabilités et analyse ont été une motivation importante pour l'étude des équations différentielles stochastiques.

Chapitre 7

Théorie générale des processus de Markov

7.1 Définitions générales et problème d'existence

Soit $(E, \mathcal{E})$ un espace mesurable. Un noyau de transition de E dans E est une application $Q : E \times \mathcal{E} \longrightarrow [0,1]$ qui possède les deux propriétés :

(i) Pour tout $x \in E$, l'application $A \to Q(x, A)$ est une mesure de probabilité sur E.

(ii) Pour tout $A \in \mathcal{E}$, l'application $x \to Q(x, A)$ est mesurable.

Remarque. Si $f : E \longrightarrow \mathbb{R}$ est mesurable bornée (resp. positive), l'application Qf définie par

$$Qf(x) = \int Q(x, dy)\, f(y)$$

est aussi mesurable bornée (resp. positive) sur E.

Définition 7.1 *Une famille $(Q_t)_{t \geq 0}$ de noyaux de transition sur E est un semi-groupe de transition si elle satisfait les trois propriétés suivantes :*

(i) *Pour tout $x \in E$, $Q_0(x, dy) = \delta_x(dy)$.*

(ii) *Pour tous $s, t \geq 0$ et $A \in \mathcal{E}$,*

$$Q_{t+s}(x, A) = \int_E Q_t(x, dy)\, Q_s(y, A)$$

(relation de Chapman-Kolmogorov).

(iii) *Pour tout* $A \in E$*, l'application* $(t, x) \to Q_t(x, A)$ *est mesurable pour la tribu* $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E}$*.*

Remarque. Dans le cas où E est dénombrable ou fini, Q_t est caractérisé par la donnée de la "matrice" $(Q_t(x, \{y\}))_{x,y \in E}$.

Soit $B(E)$ l'espace vectoriel des fonctions mesurables bornées sur E, qui est muni de la norme $\|f\| = \sup\{|f(x)|, x \in E\}$. Alors l'application $B(E) \ni f \to Q_t f$ est une contraction de $B(E)$. De plus la relation de Chapman-Kolmogorov équivaut à l'identité d'opérateurs

$$Q_{t+s} = Q_t Q_s$$

ce qui permet de voir $(Q_t)_{t \geq 0}$ comme un semigroupe de contractions de $B(E)$.

On se donne maintenant un espace de probabilité filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0,\infty]}, P)$.

Définition 7.2 *Soit* $(Q_t)_{t \geq 0}$ *un semigroupe de transition sur* E*. Un processus de Markov de semigroupe* $(Q_t)_{t \geq 0}$ *(relativement à la filtration* $(\mathcal{F}_t)$*) est un processus adapté* $(X_t)_{t \geq 0}$ *à valeurs dans* E *tel que, pour tous* $s, t \geq 0$ *et* $f \in B(E)$*,*

$$E[f(X_{s+t}) \mid \mathcal{F}_s] = Q_t f(X_s).$$

Remarque. Si la filtration n'est pas spécifiée, on prend $\mathcal{F}_t = \sigma(X_r, 0 \leq r \leq t)$.

On peut interpréter la définition comme suit. En prenant $f = \mathbf{1}_A$, on a

$$P[X_{s+t} \in A \mid \mathcal{F}_s] = Q_t(X_s, A)$$

et en particulier

$$P[X_{s+t} \in A \mid X_r, 0 \leq r \leq s] = Q_t(X_s, A).$$

Donc la loi conditionnelle de X_{s+t} connaissant le "passé" $(X_r, 0 \leq r \leq s)$ à l'instant s est donnée par $Q_t(X_s, \cdot)$, et cette loi conditionnelle ne dépend que du "présent" X_s.

Conséquence de la définition. Soit μ la loi de X_0. Alors, si $0 < t_1 < t_2 < \cdots < t_p$ et $A_0, A_1, \ldots, A_p \in \mathcal{E}$,

$$\begin{aligned}
&P(X_0 \in A_0, X_{t_1} \in A_1, X_{t_2} \in A_2, \ldots, X_{t_p} \in A_p) \\
&= \int_{A_0} \mu(dx_0) \int_{A_1} Q_{t_1}(x_0, dx_1) \int_{A_2} Q_{t_2 - t_1}(x_1, dx_2) \cdots \int_{A_p} Q_{t_p - t_{p-1}}(x_{p-1}, dx_p).
\end{aligned}$$

Plus généralement, si $f_0, f_1, \ldots, f_p \in B(E)$,

$$\begin{aligned}
&E[f_0(X_0) f_1(X_{t_1}) \cdots f_p(X_{t_p})] \\
&= \int \mu(dx_0) f_0(x_0) \int Q_{t_1}(x_0, dx_1) f_1(x_1) \int Q_{t_2 - t_1}(x_1, dx_2) f_2(x_2) \cdots \int Q_{t_p - t_{p-1}}(x_{p-1}, dx_p) f_p(x_p).
\end{aligned}$$

Cette dernière formule se démontre par récurrence sur p à partir de la définition. Remarquons qu'inversement si cette formule est vraie pour tout choix de $0 < t_1 < t_2 < \cdots < t_p$ et $f_0, f_1, \ldots, f_p \in B(E)$, $(X_t)_{t\geq 0}$ est un processus de Markov de semigroupe $(Q_t)_{t\geq 0}$ relativement à la filtration canonique $\mathcal{F}_t^0 = \sigma(X_r, 0 \leq r \leq t)$.

Exemple. Si $E = \mathbb{R}$, on peut prendre pour $t > 0$,

$$Q_t(x, dy) = p_t(y - x)\, dy$$

avec

$$p_t(y - x) = \frac{1}{\sqrt{2\pi t}} \exp -\frac{|y - x|^2}{2t}.$$

Le processus de Markov associée est le mouvement brownien (en fait le pré-mouvement brownien).

Nous abordons maintenant la question de l'existence d'un processus de Markov associé à un semigroupe de transition donné. Pour cela, nous aurons besoin d'un théorème général de construction de processus aléatoires, le théorème de Kolmogorov, que nous admettrons sans démonstration.

Soit $\Omega^* = E^{\mathbb{R}_+}$ l'espace de toutes les applications $\omega : \mathbb{R}_+ \longrightarrow E$. On munit Ω^* de la tribu $\mathcal{F}^*$ qui est la plus petite tribu rendant mesurables les applications coordonnées $\omega \longrightarrow \omega(t)$ pour $t \in \mathbb{R}_+$. Soit $F(\mathbb{R}_+)$ l'ensemble des parties finies de $\mathbb{R}_+$, et pour tout $U \in F(\mathbb{R}_+)$, soit $\pi_U : \Omega^* \longrightarrow E^U$ l'application qui à une application $\omega : \mathbb{R}_+ \longrightarrow E$ associe sa restriction à U. Si $U, V \in F(\mathbb{R}_+)$ et $U \subset V$, on note de même $\pi_U^V : E^V \longrightarrow E^U$ l'application de restriction.

On rappelle qu'un espace topologique est dit polonais si sa topologie est séparable (il existe une suite dense) et peut être définie par une distance pour laquelle l'espace est complet.

Théorème 7.1 *On suppose que E est un espace polonais muni de sa tribu borélienne $\mathcal{E}$. On se donne pour tout $U \in F(\mathbb{R}_+)$ une mesure de probabilité μ_U sur E^U, et on suppose que la famille $(\mu_U, U \in F(\mathbb{R}_+))$ est compatible au sens suivant : si $U \subset V$, μ_U est l'image de μ_V par π_U^V. Il existe alors une (unique) mesure de probabilité μ sur $(\Omega^*, \mathcal{F}^*)$ telle que $\pi_U(\mu) = \mu_U$ pour tout $U \in F(\mathbb{R}_+)$.*

Remarque. L'unicité de μ est une conséquence immédiate du lemme de classe monotone.

Ce théorème permet de construire des processus aléatoires ayant des lois marginales de dimension finie prescrites. En effet, notons $(X_t)_{t\geq 0}$ le processus canonique sur Ω^* :

$$X_t(\omega) = \omega(t), \qquad t \geq 0.$$

Si μ est une mesure de probabilité sur Ω^* et $U = \{t_1, \ldots, t_p\} \in F(\mathbb{R}_+)$, la loi du vecteur $(X_{t_1}, \ldots, X_{t_p})$ sous μ est $\pi_U(\mu)$. Le théorème de Kolmogorov se traduit donc en disant qu'étant donné une famille de lois marginales $(\mu_U, U \in F(\mathbb{R}_+))$ satisfaisant la condition de compatibilité (qui est manifestement nécessaire pour la conclusion recherchée), on peut construire une probabilité μ sur l'espace Ω^* sous laquelle les lois marginales de dimension finie du processus canonique X sont les $\mu_U, U \in F(\mathbb{R}_+)$.

Corollaire 7.2 *On suppose que E satisfait l'hypothèse du théorème précédent et que $(Q_t)_{t\geq 0}$ est un semigroupe de transition sur E. Soit γ une mesure de probabilité sur E. Il existe alors une (unique) mesure de probabilité P sur Ω^* sous laquelle le processus canonique $(X_t)_{t\geq 0}$ est un processus de Markov de semigroupe $(Q_t)_{t\geq 0}$ et la loi de X_0 est γ.*

Démonstration. Soit $U = \{t_1, \ldots, t_p\} \in F(\mathbb{R}_+)$, avec $0 \leq t_1 < \cdots < t_p$. On définit alors une mesure de probabilité P^U sur E^U en posant

$$\begin{aligned}&\int P^U(dx_1 \ldots dx_p)\, 1_A(x_1, \ldots, x_p)\\ &= \int \gamma(dx_0) \int Q_{t_1}(x_0, dx_1) \int Q_{t_2 - t_1}(x_1, dx_2) \cdots \int Q_{t_p - t_{p-1}}(x_{p-1}, dx_p) 1_A(x_1, \ldots, x_p)\end{aligned}$$

pour tout partie mesurable A de E^U (de manière évidente on a identifié E^U à E^p en identifiant $\omega \in E^U$ au vecteur $(\omega(t_1), \ldots, \omega(t_p))$).

En utilisant la relation de Chapman-Kolmogorov, on vérifie aisément que les mesures P^U satisfont la condition de compatibilité. Le théorème de Kolmogorov fournit alors l'existence (et l'unicité) de P. Le fait que les lois marginales de $(X_t)_{t\geq 0}$ sous P soient les P^U suffit pour dire que $(X_t)_{t\geq 0}$ est sous P un processus de Markov de semigroupe $(Q_t)_{t\geq 0}$, relativement à sa filtration canonique. □

Pour $x \in E$, notons P_x la mesure obtenue dans le Corollaire lorsque $\gamma = \delta_x$. Alors, l'application $x \to P_x$ est mesurable au sens où pour tout $A \in \mathcal{F}^*$ l'application $x \to P_x(A)$ est mesurable. En effet cette dernière propriété est vraie lorsque A dépend d'un nombre fini de coordonnées (dans ce cas on a une formule explicite pour $P_x(A)$) et il suffit ensuite d'utiliser un argument de classe monotone. De plus, pour toute mesure de probabilité γ sur E, la mesure définie par

$$P_{(\gamma)}(A) = \int \gamma(dx)\, P_x(A)$$

est l'unique mesure de probabilité sur Ω^* sous laquelle le processus canonique $(X_t)_{t\geq 0}$ est un processus de Markov de semigroupe $(Q_t)_{t\geq 0}$ et la loi de X_0 est γ.

En résumé, le corollaire ci-dessus permet de construire (sous une hypothèse topologique sur E) un processus de Markov $(X_t)_{t\geq 0}$ de semigroupe $(Q_t)_{t\geq 0}$ partant avec une loi initiale donnée. Plus précisément, on obtient même une famille de probabilités $(P_x, x \in E)$ telles que sous P_x le processus de Markov X part de x.

Cependant, un inconvénient de la méthode utilisée est qu'elle ne donne aucune information sur les trajectoires de X. Nous remédierons à cet inconvénient plus tard, mais cela nécessitera des hypothèses supplémentaires sur le semigroupe $(Q_t)_{t\geq 0}$. Pour terminer ce paragraphe nous introduisons un outil important, la notion de résolvante d'un semigroupe.

Définition 7.3 *Soit $\lambda > 0$. La λ-résolvante du semigroupe $(Q_t)_{t\geq 0}$ est l'opérateur $R_\lambda : B(E) \longrightarrow B(E)$ défini par*

$$R_\lambda f(x) = \int_0^\infty e^{-\lambda t} Q_t f(x)\, dt$$

pour $f \in B(E)$ et $x \in E$.

Remarque. La propriété (iii) de la définition d'un semigroupe de transition est utilisée ici pour obtenir la mesurabilité de l'application $t \to Q_t f(x)$.

Propriétés.

(i) $\|R_\lambda f\| \leq \frac{1}{\lambda}\|f\|$.

(ii) Si $0 \leq f \leq 1$, alors $0 \leq \lambda R_\lambda f \leq 1$.

(iii) Si $\lambda, \mu > 0$,

$$R_\lambda - R_\mu + (\lambda - \mu) R_\lambda R_\mu = 0$$

(équation résolvante).

Démonstration. Les propriétés (i) et (ii) sont faciles. Démontrons seulement (iii). On peut supposer $\lambda \neq \mu$. Alors,

$$\begin{aligned}
R_\lambda(R_\mu f)(x) &= \int_0^\infty e^{-\lambda s} Q_s\Big(\int_0^\infty e^{-\mu t} Q_t f\, dt\Big)(x)\, ds \\
&= \int_0^\infty e^{-\lambda s}\Big(\int Q_s(x,dy)\int_0^\infty e^{-\mu t} Q_t f(y)\, dt\Big) ds \\
&= \int_0^\infty e^{-\lambda s}\Big(\int e^{-\mu t} Q_{s+t} f(y)\, dt\Big) ds \\
&= \int_0^\infty e^{-(\lambda-\mu)s}\Big(\int e^{-\mu(s+t)} Q_{s+t} f(y)\, dt\Big) ds \\
&= \int_0^\infty e^{-(\lambda-\mu)s}\Big(\int_s^\infty e^{-\mu r} Q_r f(y)\, dr\Big) ds \\
&= \int_0^\infty Q_r f(y)\, e^{-\mu r}\Big(\int_0^r e^{-(\lambda-\mu)s} ds\Big) dr \\
&= \int_0^\infty Q_r f(y)\Big(\frac{e^{-\mu r} - e^{-\lambda r}}{\lambda - \mu}\Big) dr
\end{aligned}$$

d'où le résultat recherché. □

Exercice. Dans le cas du mouvement brownien, vérifier que

$$R_\lambda f(x) = \int r_\lambda(y-x) f(y)\, dy$$

avec

$$r_\lambda(y-x) = \frac{1}{\sqrt{2\lambda}} \exp(-|y-x|\sqrt{2\lambda}).$$

7.2 Semigroupes de Feller

A partir de maintenant, nous supposons que E est un espace topologique métrisable localement compact et dénombrable à l'infini (E est réunion dénombrable de compacts) muni de sa tribu borélienne. Ces propriétés entraînent que E est polonais.

On note $C_0(E)$ l'espace des fonctions continues de E dans $\mathbb{R}$ qui tendent vers 0 à l'infini. L'espace $C_0(E)$ est un espace de Banach pour la norme

$$\|f\| = \sup_{x \in E} |f(x)|.$$

Définition 7.4 *Soit $(Q_t)_{t\geq 0}$ un semigroupe de transition sur E. On dit que $(Q_t)_{t\geq 0}$ est un semigroupe de Feller si :*

(i) $\forall f \in C_0(E)$, $Q_t f \in C_0(E)$.

(ii) $\forall f \in C_0(E)$, $\|Q_t f - f\| \longrightarrow 0$ *quand* $t \to 0$.

Remarque. On montre qu'on peut remplacer (ii) par la condition apparemment plus faible

$$\forall f \in C_0(E),\ \forall x \in E,\ Q_t f(x) \underset{t\to 0}{\longrightarrow} f(x).$$

Nous admettrons cela pour traiter certains exemples qui suivent.

La condition (ii) entraîne que, pour tout $s \geq 0$,

$$\lim_{t\downarrow 0} \|Q_{s+t} f - Q_s f\| = \lim_{t\downarrow 0} \|Q_s(Q_t f - f)\| = 0$$

puisque Q_s est une contraction de $C_0(E)$. La convergence est même uniforme quand s varie dans $\mathbb{R}_+$, ce qui assure que la fonction $t \to Q_t f$ est uniformément continue de $\mathbb{R}_+$ dans $C_0(E)$, dès que $f \in C_0(E)$.

Dans la suite de ce paragraphe, on se donne un semigroupe de Feller $(Q_t)_{t\geq 0}$ sur E. On vérifie aisément que, pour tout $\lambda > 0$, $R_\lambda f \in C_0(E)$ si $f \in C_0(E)$.

Proposition 7.3 *Soit $\mathcal{R} = \{R_\lambda f : f \in C_0(E)\}$. Alors $\mathcal{R}$ ne dépend pas du choix de $\lambda > 0$. De plus $\mathcal{R}$ est un sous-espace dense de $C_0(E)$.*

Démonstration. Si $\lambda \neq \mu$, l'équation résolvante donne

$$R_\lambda f = R_\mu(f + (\mu - \lambda)R_\lambda f).$$

Donc toute fonction de la forme $R_\lambda f$ avec $f \in C_0(E)$ s'écrit aussi sous la forme $R_\mu g$ avec $g \in C_0(E)$. Cela donne la première assertion.

La deuxième assertion découle de ce que, pour toute $f \in C_0(E)$,

$$\lambda R_\lambda f = \lambda \int_0^\infty e^{-\lambda t} Q_t f \, dt \underset{\lambda \to \infty}{\longrightarrow} f$$

d'après la propriété (ii) de la définition d'un semigroupe de Feller. $\square$

Définition 7.5 *On pose*

$$D(L) = \{f \in C_0(E) : \frac{Q_t f - f}{t} \text{ converge dans } C_0(E) \text{ quand } t \downarrow 0\}$$

et pour toute $f \in D(L)$,

$$Lf = \lim_{t \downarrow 0} \frac{Q_t f - f}{t}.$$

Alors $D(L)$ *est un sous-espace vectoriel de* $C_0(E)$ *et* $L : D(L) \longrightarrow C_0(E)$ *est un opérateur linéaire appelé le générateur du semigroupe* $(Q_t)_{t \geq 0}$. *L'ensemble* $D(L)$ *est appelé domaine de* L.

Proposition 7.4 *Soit* $\lambda > 0$. *On a :*

(i) *Pour tous* $\lambda > 0$ *et* $g \in C_0(E)$, $R_\lambda g \in D(L)$ *et* $(\lambda - L)R_\lambda g = g$.

(ii) *Si* $g \in D(L)$, $R_\lambda(\lambda - L)g = g$.

En conséquence, $D(L) = \mathcal{R}$ *et les opérateurs* $R_\lambda : C_0(E) \to \mathcal{R}$ *et* $\lambda - L : D(L) \to C_0(E)$ *sont inverses l'un de l'autre.*

Démonstration. (i) Si $g \in C_0(E)$, on a pour tout $\varepsilon > 0$,

$$\begin{aligned} \varepsilon^{-1}(Q_\varepsilon R_\lambda g - R_\lambda g) &= \varepsilon^{-1}\Big(\int_0^\infty e^{-\lambda t} Q_{\varepsilon + t} g \, dt - \int_0^\infty e^{-\lambda t} Q_t g \, dt\Big) \\ &= \varepsilon^{-1}\Big((1 - e^{-\lambda \varepsilon}) \int_0^\infty e^{-\lambda t} Q_{\varepsilon + t} g \, dt - \int_0^\varepsilon e^{-\lambda t} Q_t g \, dt\Big) \\ &\underset{\varepsilon \to 0}{\longrightarrow} \lambda R_\lambda g - g \end{aligned}$$

en utilisant la propriété (ii) de la définition d'un semigroupe de Feller (et le fait que cette propriété entraîne la continuité de l'application $t \to Q_t g$ de $\mathbb{R}_+$ dans $C_0(E)$). Le calcul précédent montre que $R_\lambda g \in D(L)$ et $L(R_\lambda g) = \lambda R_\lambda g - g$.

(ii) Soit $f \in D(L)$. Pour tout $t \geq 0$,

$$\varepsilon^{-1}(Q_{t+\varepsilon} f - Q_t f) = Q_t(\varepsilon^{-1}(Q_\varepsilon f - f)) \underset{\varepsilon \downarrow 0}{\longrightarrow} Q_t(Lf).$$

De plus la convergence précédente est uniforme en $t \in \mathbb{R}_+$. Cela suffit pour dire que pour tout $x \in E$, la fonction $t \to Q_t f(x)$ est dérivable sur $\mathbb{R}_+$ et sa dérivée est $Q_t(Lf)(x)$, qui est une fonction continue de t. En conséquence, pour tout $t \geq 0$,

$$Q_t f(x) = f(x) + \int_0^t Q_s(Lf)(x)\, ds.$$

Il en découle que

$$\begin{aligned}\int_0^\infty e^{-\lambda t} Q_t f(x)\, dt &= \frac{f(x)}{\lambda} + \int_0^\infty e^{-\lambda t}\Big(\int_0^t Q_s(Lf)(x)\, ds\Big) dt \\ &= \frac{f(x)}{\lambda} + \int_0^\infty \frac{e^{-\lambda s}}{\lambda} Q_s(Lf)(x)\, ds.\end{aligned}$$

On a ainsi obtenu l'égalité

$$\lambda R_\lambda f = f + R_\lambda L f$$

d'où le résultat annoncé. □

Corollaire 7.5 *Le semigroupe $(Q_t)_{t \geq 0}$ est déterminé par la donnée du générateur L (et bien entendu de son domaine $D(L)$).*

Démonstration. Soit f une fonction positive dans $C_0(E)$. Alors $R_\lambda f$ est l'unique élément de $D(L)$ tel que $(\lambda - L)R_\lambda f = f$. Par ailleurs la donnée de $R_\lambda f(x) = \int_0^\infty e^{-\lambda t} Q_t f(x) dt$ pour tout $\lambda > 0$ suffit à déterminer la fonction continue $t \to Q_t f(x)$. Or Q_t est déterminé par la donnée de $Q_t f$ pour toute fonction positive f dans $C_0(E)$. □

Exemple. Le mouvement brownien. Il est facile de vérifier que le semigroupe du mouvement brownien est de Feller. De plus, nous avons vu que, pour $f \in C_0(\mathbb{R})$,

$$R_\lambda f(x) = \int \frac{1}{\sqrt{2\lambda}} \exp(-\sqrt{2\lambda}|y - x|)\, f(y)\, dy.$$

Nous allons utiliser cette formule pour calculer le générateur du mouvement brownien. Si $h \in D(L)$ on sait qu'il existe $f \in C_0(\mathbb{R})$ telle que $h = R_\lambda f$. En prenant $\lambda = \frac{1}{2}$, on a

$$h(x) = \int \exp(-|x - y|)\, f(y)\, dy.$$

On justifie facilement l'application du théorème de dérivation sous le signe intégrale, pour obtenir que h est dérivable sur $\mathbb{R}$, et

$$h'x) = -\int \mathrm{sgn}(x-y)\,\exp(-|x-y|)\,f(y)\,dy$$

avec la notation $\mathrm{sgn}(z) = \mathbf{1}_{\{z>0\}} - \mathbf{1}_{\{z<0\}}$. Nous allons montrer aussi que h' est dérivable sur $\mathbb{R}$. Soit $x_0 \in \mathbb{R}$. Alors pour $x > x_0$,

$$\begin{aligned} h'(x) - h'(x_0) &= \int \Big(\mathrm{sgn}(y-x)\exp(-|x-y|) - \mathrm{sgn}(y-x_0)\exp(-|x_0-y|)\Big) f(y)dy \\ &= \int_{x_0}^{x} \Big(-\exp(-|x-y|) - \exp(-|x_0-y|)\Big) f(y)\,dy \\ &\quad + \int_{\mathbb{R}\setminus[x_0,x]} \mathrm{sgn}(y-x_0)\Big(\exp(-|x-y|) - \exp(-|x_0-y|)\Big) f(y)\,dy. \end{aligned}$$

On en déduit aisément que

$$\frac{h'(x)-h'(x_0)}{x-x_0} \underset{x\downarrow x_0}{\longrightarrow} -2f(x_0) + h(x_0).$$

On obtient la même limite quand $x \uparrow x_0$, et on voit ainsi que h est deux fois dérivable, et $h'' = -2f + h$.

Par ailleurs, puisque $h = R_{1/2}f$, la Proposition 7.4 montre que

$$(\frac{1}{2} - L)h = f$$

d'où $Lh = -f + \frac{1}{2}h = \frac{1}{2}h''$.

En conclusion, on a montré que

$$D(L) \subset \{h \in C^2(\mathbb{R}) : h \text{ et } h'' \in C_0(\mathbb{R})\}$$

et que pour $h \in D(L)$, on a $Lh = \frac{1}{2}h''$.

En fait, l'inclusion précédente est une égalité. Pour le voir on peut raisonner comme suit. Si g est une fonction de classe C^2 telle que g et g'' sont dans $C_0(\mathbb{R})$, on peut poser $f = \frac{1}{2}(g - g'') \in C_0(\mathbb{R})$, puis $h = R_{1/2}f \in D(L)$. Le raisonnement ci-dessus montre qu'alors $h'' = -2f + h$, d'où $(h-g)'' = h - g$. Puisque la fonction $h - g$ est dans $C_0(\mathbb{R})$ elle doit être identiquement nulle, et on a $g = h \in D(L)$.

Remarque. En général il est très difficile de calculer le domaine exact du générateur. La proposition suivante permet souvent d'identifier des éléments de ce domaine au moyen de martingales associées au processus de Markov de semigroupe $(Q_t)_{t\geq 0}$.

Nous considérons à nouveau un semigroupe de Feller général $(Q_t)_{t\geq 0}$. Nous supposons donnés un processus $(X_t)_{t\geq 0}$ et une famille $(P_x)_{x\in E}$ de mesure de probabilités sur E, telle que, sous P_x, $(X_t)_{t\geq 0}$ est un processus de Markov de semigroupe $(Q_t)_{t\geq 0}$, relativement à une filtration $(\mathcal{F}_t)_{t\geq 0}$, et $P_x(X_0 = x) = 1$. Pour donner un sens aux intégrales qui apparaissent ci-dessous, nous supposons aussi que les trajectoires de $(X_t)_{t\geq 0}$ sont càdlàg.

Théorème 7.6 *Soient $h, g \in C_0(E)$. Les deux conditions suivantes sont équivalentes :*

(i) $h \in D(L)$ *et* $Lh = g$.

(ii) *Pour tout $x \in E$, le processus*

$$h(X_t) - \int_0^t g(X_s)\,ds$$

est une $(\mathcal{F}_t)$-martingale sous P_x.

Démonstration. On établit d'abord (i)⇒(ii). Soit donc $h \in D(L)$ et $g = Lh$. Comme cela a été observé dans la preuve de la Proposition 7.4, on a alors pour tout $t \geq 0$,

$$Q_t h = h + \int_0^t Q_r g\,dr.$$

Il en découle que, pour $t \geq 0$ et $s \geq 0$,

$$E_x[h(X_{t+s}) \mid \mathcal{F}_t] = Q_s h(X_t) = h(X_t) + \int_0^s Q_r g(X_t)\,dr.$$

D'autre part,

$$\begin{aligned} E_x\Big[\int_t^{t+s} g(X_r)\,dr \,\Big|\, \mathcal{F}_t\Big] &= \int_t^{t+s} E_x[g(X_r) \mid \mathcal{F}_t]\,dr \\ &= \int_t^{t+s} Q_{r-t}\,g(X_t)\,dr \\ &= \int_0^s Q_r g(X_t)\,dr. \end{aligned}$$

L'interversion de l'intégrale et de l'espérance conditionnelle dans la première égalité est facile à justifier en utilisant la propriété caractéristique de l'espérance conditionnelle. Il découle de ce qui précède que

$$E_x\Big[h(X_{t+s}) - \int_0^{t+s} g(X_r)\,dr \,\Big|\, \mathcal{F}_t\Big] = h(X_t) - \int_0^t g(X_r)\,dr$$

d'où la propriété (ii).

Inversement, supposons que (ii) est réalisée. Alors pour tout $t \geq 0$,

$$E_x\Big[h(X_t) - \int_0^t g(X_r)\,dr\Big] = h(x)$$

et par ailleurs

$$E_x\Big[h(X_t) - \int_0^t g(X_r)\,dr\Big] = Q_t h(x) - \int_0^t Q_r g(x)\,dr.$$

En conséquence,

$$\frac{Q_t h - h}{t} = \frac{1}{t}\int_0^t Q_r g\,dr \underset{t\downarrow 0}{\longrightarrow} g$$

dans $C_0(E)$, d'après la propriété (ii) de la définition d'un semigroupe de Feller. On conclut que $h \in D(L)$ et $Lh = g$. □

Exemple. Dans le cas du mouvement brownien, la formule d'Itô montre que si $h \in C^2(\mathbb{R})$,

$$h(X_t) - \frac{1}{2}\int_0^t h''(X_s)\,ds$$

est une martingale locale. Cette martingale locale devient une vraie martingale lorsqu'on suppose aussi que h et h'' sont dans $C_0(\mathbb{R})$ (donc bornées). On retrouve ainsi le fait que $Lh = \frac{1}{2}h''$ pour de telles fonctions h.

7.3 La régularité des trajectoires

Notre objectif dans cette partie est de montrer que l'on peut construire le processus de Markov associé à un semigroupe de Feller de manière à ce que ses trajectoires soient càdlàg (c'est-à dire continues à droite avec des limites à gauche en tout point). Nous commençons par un résultat valable pour un semigroupe général.

Lemme 7.7 *Soit X un processus de Markov de semigroupe $(Q_t)_{t\geq 0}$ à valeurs dans un espace mesurable $(E, \mathcal{E})$. Soit $h \in B(E)$ une fonction à valeurs positives, et soit $\lambda > 0$. Le processus*

$$e^{-\lambda t} R_\lambda h(X_t)$$

est une surmartingale.

Démonstration. Pour tout $s \geq 0$,

$$Q_s R_\lambda h = \int_0^\infty e^{-\lambda t} Q_{s+t} h\,dt$$

et donc

$$e^{-\lambda s}Q_sR_\lambda h = \int_0^\infty e^{-\lambda(s+t)}Q_{s+t}h\,dt = \int_s^\infty e^{-\lambda t}Q_th\,dt \leq R_\lambda h.$$

Il suffit ensuite d'écrire, pour tous $s, t \geq 0$,

$$E[e^{-\lambda(t+s)}R_\lambda h(X_{t+s}) \mid \mathcal{F}_t] = e^{-\lambda(t+s)}Q_sR_\lambda h(X_t) \leq e^{-\lambda t}R_\lambda h(X_t),$$

ce qui donne la propriété de surmartingale recherchée. $\square$

Nous revenons maintenant au cas d'un semigroupe de Feller $(Q_t)_{t\geq 0}$, l'espace E étant supposé métrique localement compact et dénombrable à l'infini comme dans la partie précédente. Nous supposons donnés un processus $(X_t)_{t\geq 0}$ et une famille de mesures de probabilité $(P_x)_{x\in E}$ telle que, sous P_x, $(X_t)_{t\geq 0}$ est un processus de Markov de semigroupe $(Q_t)_{t\geq 0}$ (relativement à une filtration $(\mathcal{F}_t)_{t\in[0,\infty]}$) et $P_x(X_0 = x) = 1$. Nous avons vu dans le début de ce chapitre que ces conditions sont réalisées en prenant pour $(X_t)_{t\geq 0}$ le processus canonique sur l'espace $\Omega^* = E^{\mathbb{R}_+}$ et en construisant les mesures P_x à l'aide du théorème de Kolmogorov.

Nous notons $\mathcal{N}$ la classe des ensembles $\mathcal{F}_\infty$-mesurables qui sont de P_x-probabilité nulle pour tout $x \in E$. On définit ensuite une nouvelle filtration $(\widetilde{\mathcal{F}}_t)_{t\in[0,\infty]}$ en posant $\widetilde{\mathcal{F}}_\infty = \mathcal{F}_\infty$ et pour tout $t \geq 0$,

$$\widetilde{\mathcal{F}}_t = \mathcal{F}_{t+} \vee \sigma(\mathcal{N}).$$

On vérifie aisément que la filtration $(\widetilde{\mathcal{F}}_t)$ est continue à droite.

Théorème 7.8 *On peut construire un processus $(\widetilde{X}_t)_{t\geq 0}$ adapté à la filtration $(\widetilde{\mathcal{F}}_t)$, dont les trajectoires sont des fonctions càdlàg à valeurs dans E, et tel que*

$$\widetilde{X}_t = X_t\ , \qquad P_x\ p.s. \quad \forall x \in E.$$

De plus sous chaque probabilité P_x, $(\widetilde{X}_t)_{t\geq 0}$ est un processus de Markov de semigroupe $(Q_t)_{t\geq 0}$, relativement à la filtration $(\widetilde{\mathcal{F}}_t)_{t\in[0,\infty]}$, et $P_x(\widetilde{X}_0 = x) = 1$.

Démonstration. Soit $E_\Delta = E \cup \{\Delta\}$ le compactifié d'Alexandrov de E obtenu en ajoutant à E le point à l'infini Δ. Toute fonction $f \in C_0(E)$ se prolonge en une fonction continue sur E_Δ en posant $f(\Delta) = 0$.

Notons $C_0^+(E)$ l'ensemble des fonctions positives dans $C_0(E)$. On peut alors trouver une suite $(f_n)_{n\in\mathbb{N}}$ de fonctions de $C_0^+(E)$ qui sépare les points de E_Δ, au sens où, pour tous $x, y \in E_\Delta$ avec $x \neq y$, il existe un entier n tel que $f_n(x) \neq f_n(y)$. Alors

$$\mathcal{H} = \{R_pf_n : p \geq 1, n \in \mathbb{N}\}$$

est aussi un sous-ensemble dénombrable de $C_0^+(E)$ qui sépare les points de E_Δ (utiliser le fait que $\|pR_pf - f\| \longrightarrow 0$ quand $p \to \infty$).

Si $h \in \mathcal{H}$, le lemme précédent montre qu'il existe un entier $p \geq 1$ tel que $e^{-pt}h(X_t)$ est une surmartingale sous P_x, pour tout $x \in E$. Soit D un sous-ensemble dénombrable dense de $\mathbb{R}_+$. Alors les limites

$$\lim_{D \ni s \downarrow\downarrow t} h(X_s) \ , \qquad \lim_{D \ni s \uparrow\uparrow t} h(X_s)$$

existent simultanément pour tout $t \in \mathbb{R}_+$ (la deuxième seulement si $t > 0$) en dehors d'un ensemble $\mathcal{F}_\infty$-mesurable N_h tel que $P_x(N_h) = 0$ pour tout $x \in E$. En effet, il suffit de définir le complémentaire de N_h comme l'ensemble des $\omega \in \Omega$ pour lesquels la fonction $D \ni s \longrightarrow e^{-ps}h(X_s)$ fait un nombre fini de montées le long de tout intervalle $[a, b]$, $a, b \in \mathbb{Q}$, et d'utiliser les résultats classiques sur les nombres de montées des surmartingales (cf Chapitre 3). On pose

$$N = \bigcup_{h \in \mathcal{H}} N_h$$

de sorte que $N \in \mathcal{N}$. Alors, si $\omega \notin N$, les limites

$$\lim_{D \ni s \downarrow\downarrow t} X_s(\omega) \ , \qquad \lim_{D \ni s \uparrow\uparrow t} X_s(\omega)$$

existent pour tout $t \geq 0$ dans E_Δ (en effet, si par exemple $X_s(\omega)$ a deux valeurs d'adhérence différentes dans E_Δ quand $D \ni s \downarrow\downarrow t$, on obtient une contradiction en choisissant une fonction $h \in \mathcal{H}$ qui sépare ces deux valeurs). Cela permet de poser, pour $\omega \in \Omega \backslash N$ et pour tout $t \geq 0$,

$$\widetilde{X}_t(\omega) = \lim_{D \ni s \downarrow\downarrow t} X_s(\omega).$$

Si $\omega \in N$, on pose $\widetilde{X}_t(\omega) = x_0$ pour tout $t \geq 0$, où x_0 est un point fixé de E. Alors, pour tout $t \geq 0$, $\widetilde{X}_t$ est une v.a. $\widetilde{\mathcal{F}}_t$-mesurable à valeurs dans E_Δ. De plus, pour tout $\omega \in \Omega$, l'application $t \to \widetilde{X}_t(\omega)$, vue comme application à valeurs dans E_Δ, est càdlàg par construction.

Montrons maintenant que, pour tout $t \geq 0$,

$$P_x(X_t = \widetilde{X}_t) = 1, \qquad \forall x \in E.$$

Soient $f, g \in C_0(E)$ et soit une suite (s_n) dans D telle que $s_n \downarrow\downarrow t$. Alors, pour tout $x \in E$,

$$\begin{aligned} E_x[f(X_t)g(\widetilde{X}_t)] &= \lim_{n \to \infty} E_x[f(X_t)g(X_{s_n})] \\ &= \lim_{n \to \infty} E_x[f(X_t)Q_{s_n - t}g(X_t)] \\ &= E_x[f(X_t)g(X_t)] \end{aligned}$$

puisque $Q_{s_n-t}g \longrightarrow g$ d'après la définition d'un semigroupe de Feller. L'égalité obtenue suffit pour dire que les deux couples $(X_t, \widetilde{X}_t)$ et (X_t, X_t) ont même loi sous P_x, et donc $P_x(X_t = \widetilde{X}_t) = 1$.

Montrons ensuite que $(\widetilde{X}_t)_{t\geq 0}$ vérifie la propriété de définition d'un processus de Markov de semigroupe $(Q_t)_{t\geq 0}$ relativement à la filtration $(\widetilde{\mathcal{F}}_t)$. Il suffit de voir que, pour tous $s \geq 0$, $t > 0$ et $A \in \widetilde{\mathcal{F}}_s$, $f \in C_0(E)$, on a

$$E_x[1_A\, f(\widetilde{X}_{s+t})] = E_x[1_A\, Q_t f(\widetilde{X}_s)].$$

Puisque $\widetilde{X}_s = X_s$ p.s. et $\widetilde{X}_{s+t} = X_{s+t}$ p.s., il revient au même de montrer

$$E_x[1_A\, f(X_{s+t})] = E_x[1_A\, Q_t f(X_s)].$$

Puisque A coïncide p.s. avec un élément de $\mathcal{F}_{s+}$ on peut supposer $A \in \mathcal{F}_{s+}$. Soit (s_n) une suite décroissant strictement vers s, de sorte que $A \in \mathcal{F}_{s_n}$ pour tout n. Alors, dès que $s_n \leq s + t$,

$$E_x[1_A\, f(X_{s+t})] = E_x[1_A\, E_x[f(X_{s+t} \mid \mathcal{F}_{s_n}]] = E_x[1_A\, Q_{s+t-s_n} f(X_{s_n})].$$

Mais $Q_{s+t-s_n} f$ converge (uniformément) vers $Q_t f$ par les propriétés des semigroupes de Feller, et puisque $X_{s_n} = \widetilde{X}_{s_n}$ p.s. on sait aussi que X_{s_n} converge p.s. vers $X_s = \widetilde{X}_s$ p.s. On obtient donc l'égalité recherchée en faisant tendre n vers ∞.

Il reste finalement à montrer que les fonctions $t \longrightarrow \widetilde{X}_t(\omega)$ sont càdlàg à valeurs dans E, et pas seulement dans E_Δ (on sait déjà que pour chaque $t \geq 0$, $\widetilde{X}_t(\omega) = X_t(\omega)$ p.s. est p.s. dans E, mais cela ne suffit pas pour montrer que les trajectoires, et leurs limites à gauche, restent dans E). Fixons une fonction $g \in C_0^+(E)$ telle que $g(x) > 0$ pour tout $x \in E$. La fonction $h = R_1 g$ vérifie alors la même propriété. Posons pour tout $t \geq 0$,

$$Y_t = e^{-t} h(\widetilde{X}_t).$$

Alors le Lemme 7.7 montre que $(Y_t)_{t\geq 0}$ est une surmartingale positive relativement à la filtration $(\widetilde{\mathcal{F}}_t)$. De plus, les trajectoires de $(Y_t)_{t\geq 0}$ sont càdlàg.

Pour tout entier $n \geq 1$, posons

$$T_{(n)} = \inf\{t \geq 0 : Y_t < \frac{1}{n}\}.$$

Alors $T_{(n)}$ est un temps d'arrêt de la filtration $(\widetilde{\mathcal{F}}_t)$, comme temps d'entrée dans un ouvert pour un processus adapté à trajectoires càdlàg (rappelons que la filtration $(\widetilde{\mathcal{F}}_t)$ est continue à droite). De même,

$$T = \lim_{n\to\infty} \uparrow T_{(n)}$$

est un temps d'arrêt. Le résultat voulu découlera de ce que $P_x(T < \infty) = 0$ pour tout $x \in E$. En effet il est clair que pour tout $t \in [0, T_{(n)}[$, $\widetilde{X}_t \in E$ et $\widetilde{X}_{t-} \in E$, et il suffira de redéfinir $\widetilde{X}_t(\omega) = x_0$ pour tout $t \geq 0$ lorsque ω appartient à l'ensemble $\{T = \infty\} \in \mathcal{N}$. On utilise le lemme suivant.

Lemme 7.9 (Théorème d'arrêt pour les surmartingales positives) *Si* $(Z_t)_{t\geq 0}$ *est une surmartingale positive à trajectoires continues à droite, et si* U *et* V *sont deux temps d'arrêt avec* $U \leq V$*, alors*

$$E[Z_V \mathbf{1}_{\{V<\infty\}}] \leq E[Z_U \mathbf{1}_{\{U<\infty\}}].$$

On applique le lemme à $Z = Y$ et $U = T_{(n)}$, $V = T + q$, où q est un rationnel positif. On trouve

$$E_x[Y_{T+q} \mathbf{1}_{\{T<\infty\}}] \leq E_x[Y_{T_{(n)}} \mathbf{1}_{\{T_{(n)}<\infty\}}] \leq \frac{1}{n}.$$

En faisant tendre n vers ∞, on a donc

$$E_x[Y_{T+q} \mathbf{1}_{\{T<\infty\}}] = 0,$$

d'où $Y_{T+q} = 0$ p.s. sur $\{T < \infty\}$. Par continuité à droite, on conclut que $Y_t = 0$, $\forall t \in [T, \infty[$, p.s. sur $\{T < \infty\}$. Mais on sait que p.s. $\forall k \in \mathbb{N}$, $Y_k = e^{-k} h(\widetilde{X}_k) > 0$ puisque $\widetilde{X}_k \in E$ p.s. Cela suffit pour conclure que $P_x(T < \infty) = 0$. $\square$

7.4 La propriété de Markov forte

Dans tout ce paragraphe, nous supposons donnés un processus $(X_t)_{t\geq 0}$ à trajectoires càdlàg à valeurs dans un espace métrique E et une famille de mesures de probabilité $(P_x)_{x\in E}$ telle que, sous P_x, $(X_t)_{t\geq 0}$ est un processus de Markov de semigroupe $(Q_t)_{t\geq 0}$ (relativement à une filtration $(\mathcal{F}_t)_{t\in[0,\infty]}$) et $P_x(X_0 = x) = 1$.

Nous commençons par un énoncé de la propriété de Markov simple qui est une généralisation facile de la définition d'un processus de Markov. On note $\mathbb{D}(E)$ l'espace des fonctions càdlàg $f : \mathbb{R}_+ \longrightarrow E$. On munit $\mathbb{D}(E)$ de la tribu $\mathcal{D}$ engendrée par les applications coordonnées $f \longrightarrow f(t)$. Remarquons que l'application de Ω dans $\mathbb{D}(E)$ qui à ω associe la trajectoire $(X_t(\omega))_{t\geq 0}$ est mesurable.

Théorème 7.10 (Propriété de Markov simple) *Soit* $s \geq 0$ *et soit* $\Phi : \mathbb{D}(E) \longrightarrow \mathbb{R}_+$ *une application mesurable. Alors, pour tout* $x \in E$,

$$E_x[\Phi((X_{s+t})_{t\geq 0}) \mid \mathcal{F}_s] = E_{X_s}[\Phi((X_t)_{t\geq 0})].$$

Remarque. Le terme de droite est la composée de X_s et de l'application $y \to E_y[\Phi((X_t)_{t\geq 0})]$. Pour voir que cette application est mesurable, il suffit de traiter le cas où $\Phi = 1_A$, $A \in \mathcal{D}$. Lorsque A ne dépend que d'un nombre fini de coordonnées, on a une formule explicite, et un argument de classe monotone complète le raisonnement.

Démonstration. Comme dans la remarque ci-dessus, on se ramène facilement au cas où $\Phi = 1_A$ et

$$A = \{f \in \mathbb{D}(E) : f(t_1) \in B_1, \dots, f(t_p) \in B_p\}$$

où $0 \leq t_1 < t_2 < \cdots < t_p$ et $B_1, \dots, B_p$ sont des parties mesurables de E. Dans ce cas on doit montrer

$$\begin{aligned}&P_x(X_{s+t_1} \in B_1, \dots, X_{s+t_p} \in B_p \mid \mathcal{F}_s)\\&= \int_{B_1} Q_{t_1}(X_s, dx_1) \int_{B_2} Q_{t_2-t_1}(x_1, dx_2) \cdots \int_{B_p} Q_{t_p-t_{p-1}}(x_{p-1}, dx_p).\end{aligned}$$

En fait on montre plus généralement que si $\varphi_1, \dots, \varphi_p \in B(E)$,

$$\begin{aligned}&E_x[\varphi_1(X_{s+t_1}) \cdots \varphi_p(X_{s+t_p}) \mid \mathcal{F}_s]\\&= \int Q_{t_1}(X_s, dx_1)\varphi_1(x_1) \int Q_{t_2-t_1}(x_1, dx_2)\varphi_2(x_2) \cdots \int Q_{t_p-t_{p-1}}(x_{p-1}, dx_p)\varphi_p(x_p).\end{aligned}$$

Si $p = 1$ c'est la définition d'un processus de Markov. On raisonne ensuite par récurrence en écrivant :

$$\begin{aligned}&E_x[\varphi_1(X_{s+t_1}) \cdots \varphi_p(X_{s+t_p}) \mid \mathcal{F}_s]\\&= E_x[\varphi_1(X_{s+t_1}) \cdots \varphi_{p-1}(X_{s+t_{p-1}})\, E_x[\varphi_p(X_{s+t_p}) \mid \mathcal{F}_{s+t_{p-1}}] \mid \mathcal{F}_s]\\&= E_x[\varphi_1(X_{s+t_1}) \cdots \varphi_{p-1}(X_{s+t_{p-1}})\, Q_{t_p-t_{p-1}}\varphi_p(X_{s+t_{p-1}}) \mid \mathcal{F}_s]\end{aligned}$$

d'où facilement le résultat voulu. $\square$

Dans le théorème suivant, on suppose de plus que $(Q_t)_{t\geq 0}$ est un semigroupe de Feller (et donc E est localement compact et dénombrable à l'infini).

Théorème 7.11 (Propriété de Markov forte) *Soit T un temps d'arrêt de la filtration $(\mathcal{F}_{t+})$, et soit $\Phi : \mathbb{D}(E) \longrightarrow \mathbb{R}_+$ une application mesurable. Alors, pour tout $x \in E$,*

$$E_x[\mathbf{1}_{\{T<\infty\}}\Phi((X_{T+t})_{t\geq 0}) \mid \mathcal{F}_T] = \mathbf{1}_{\{T<\infty\}}E_{X_T}[\Phi((X_t)_{t\geq 0})].$$

Démonstration. Puisque le terme de droite est $\mathcal{F}_T$-mesurable, il suffit de montrer que, pour $A \in \mathcal{F}_T$ fixé,

$$E_x[\mathbf{1}_{A\cap\{T<\infty\}}\Phi((X_{T+t})_{t\geq 0})] = E_x[\mathbf{1}_{A\cap\{T<\infty\}}E_{X_T}[\Phi((X_t)_{t\geq 0})]].$$

Comme ci-dessus, on peut se limiter au cas où

$$\Phi(f) = \varphi_1(f(t_1)) \cdots \varphi_p(f(t_p))$$

où $0 \leq t_1 < t_2 < \cdots < t_p$ et $\varphi_1, \ldots, \varphi_p \in B(E)$. Il suffit en fait de traiter le cas $p = 1$: si ce cas est établi, on raisonne par récurrence en écrivant

$$\begin{aligned}
&E_x[\mathbf{1}_{A\cap\{T<\infty\}}\varphi_1(X_{T+t_1}) \cdots \varphi_p(X_{T+t_p})] \\
&= E_x[\mathbf{1}_{A\cap\{T<\infty\}}\varphi_1(X_{T+t_1}) \cdots \varphi_{p-1}(X_{T+t_{p-1}})\, E_x[\varphi_p(X_{T+t_p}) \mid \mathcal{F}_{T+t_{p-1}}]] \\
&= E_x[\mathbf{1}_{A\cap\{T<\infty\}}\varphi_1(X_{T+t_1}) \cdots \varphi_{p-1}(X_{T+t_{p-1}}) \int Q_{t_p-t_{p-1}}(X_{T+t_{p-1}}, dx_p)\varphi_p(x_p)].
\end{aligned}$$

On fixe donc $t \geq 0$ et $\varphi \in B(E)$ et on veut montrer

$$E_x[\mathbf{1}_{A\cap\{T<\infty\}}\varphi(X_{T+t})] = E_x[\mathbf{1}_{A\cap\{T<\infty\}}E_{X_T}[\varphi(X_t)]].$$

On peut supposer que $\varphi \in C_0(E)$, par un raisonnement standard de classe monotone.

Notons $[T]_n$ le plus petit nombre réel de la forme $i2^{-n}$ strictement supérieur à T. Alors,

$$\begin{aligned}
E_x[\mathbf{1}_{A\cap\{T<\infty\}}\varphi(X_{T+t})] &= \lim_{n\to\infty} E_x[\mathbf{1}_{A\cap\{T<\infty\}}\varphi(X_{[T]_n+t})] \\
&= \lim_{n\to\infty} \sum_{i=1}^{\infty} E_x[\mathbf{1}_{A\cap\{(i-1)2^{-n}\leq T<i2^{-n}\}}\varphi(X_{i2^{-n}+t})] \\
&= \lim_{n\to\infty} \sum_{i=1}^{\infty} E_x[\mathbf{1}_{A\cap\{(i-1)2^{-n}\leq T<i2^{-n}\}}Q_t\varphi(X_{i2^{-n}})] \\
&= \lim_{n\to\infty} E_x[\mathbf{1}_{A\cap\{T<\infty\}}Q_t\varphi(X_{[T]_n})] \\
&= E_x[\mathbf{1}_{A\cap\{T<\infty\}}Q_t\varphi(X_T)]
\end{aligned}$$

d'où le résultat voulu. Dans la première (et la dernière) égalité, on utilise la continuité à droite des trajectoires. Dans la troisième égalité, on se sert du fait que $A \cap \{(i-1)2^{-n} \leq T < i2^{-n}\} \in \mathcal{F}_{i2^{-n}}$ parce que $A \in \mathcal{F}_T$ et T est un temps d'arrêt de la filtration $(\mathcal{F}_{t+})$. Enfin, dans la dernière égalité, on utilise le fait que $Q_t\varphi$ est continue parce que $\varphi \in C_0(E)$ et le semigroupe est de Feller. □

La propriété de Feller du semigroupe joue un rôle crucial dans la preuve qui précède. En l'absence de cette propriété, le résultat devient faux. Pour le voir on peut considérer l'exemple très simple suivant. En prenant $E = \mathbb{R}_+$, on construit très facilement un processus $(X_t)_{t\geq 0}$ et une famille de probabilités $(P_x)_{x\in\mathbb{R}_+}$ tels que :

- si $x > 0$, P_x p.s., $\forall t \geq 0,\ X_t = x + t$;
- si $x = 0$, P_0 p.s., $\forall t \geq 0,\ X_t = (t - T)_+$,

où la v.a. T suit sous P_0 une loi exponentielle de paramètre 1. Alors, sous P_x, $(X_t)_{t\geq 0}$ est un processus de Markov de semigroupe

$$Q_t(x,dy) = \delta_{x+t}(dy) \qquad \text{si } x > 0\,,$$
$$Q_t(0,dy) = e^{-t}\delta_0(dy) + \int_0^t ds\, e^{-s}\,\delta_{t-s}(dy)\,.$$

De plus, les trajectoires de X sont continues. Toutes les hypothèses du théorème précédent sont vérifiées, à l'exception de la propriété de Feller du semigroupe. Cependant, la propriété de Markov forte est en défaut. En effet, $T = \inf\{t \geq 0 : X_t > 0\}$ est un temps d'arrêt de la filtration $\mathcal{F}_{t+}$, tel que $X_T = 0$, P_0 p.s., mais la loi de $(X_{T+t})_{t\geq 0}$ sous P_0 est très différente de la loi de $(X_t)_{t\geq 0}$ sous P_0.

7.5 Deux classes importantes de processus de Feller

7.5.1 Processus de Lévy

Considérons un processus $(Y_t)_{t\geq 0}$ à valeurs dans $\mathbb{R}$ (ou dans $\mathbb{R}^d$) qui vérifie les trois propriétés suivantes :

(i) $Y_0 = 0$ p.s.

(ii) Pour tous $0 \leq s < t$, la v.a. $Y_t - Y_s$ est indépendante de $(Y_r, 0 \leq r \leq s)$ et a même loi que Y_{t-s}.

(iii) Y_t converge en probabilité vers 0 quand $t \downarrow 0$.

Remarquons qu'on ne suppose pas que les trajectoires de Y sont càdlàg, mais on remplace cette hypothèse par la condition beaucoup plus faible (iii). La théorie qui précède va nous montrer qu'on peut cependant trouver une modification de Y dont les trajectoires sont càdlàg.

Pour tout $t \geq 0$, notons $Q_t(0,dy)$ la loi de Y_t, et pour tout $x \in \mathbb{R}$, soit $Q_t(x,dy)$ la mesure-image de $Q_t(0,dy)$ par l'application $y \to x+y$.

Proposition 7.12 *La famille $(Q_t)_{t\geq 0}$ forme un semigroupe de Feller sur $\mathbb{R}$. De plus $(Y_t)_{t\geq 0}$ est un processus de Markov de semigroupe $(Q_t)_{t\geq 0}$.*

Démonstration. Montrons d'abord que $(Q_t)_{t\geq 0}$ est un semigroupe de transition. Si $\varphi \in B(\mathbb{R})$, $s,t \geq 0$ et $x \in \mathbb{R}$,

$$\begin{aligned}
\int Q_t(x,dy)\int Q_s(y,dz)\varphi(z) &= \int Q_t(0,dy)\int Q_s(0,dz)\varphi(x+y+z)\\
&= E[\varphi(x+Y_t+(Y_{t+s}-Y_s))]\\
&= E[\varphi(x+Y_{t+s})]\\
&= \int Q_{t+s}(x,dz)\varphi(z)
\end{aligned}$$

d'où la relation de Chapman-Kolmogorov. Il faudrait aussi vérifier la mesurabilité de l'application $(t, x) \longrightarrow Q_t(x, A)$, mais cela découle en fait des propriétés de continuité plus fortes que nous allons établir pour montrer la propriété de Feller.

Commençons par la première propriété d'un semigroupe de Feller. Si $\varphi \in C_0(\mathbb{R})$, l'application

$$x \longrightarrow Q_t\varphi(x) = E[\varphi(x + Y_t)]$$

est continue par convergence dominée, et, toujours par convergence dominée on a

$$E[\varphi(x + Y_t)] \underset{x\to\infty}{\longrightarrow} 0$$

ce qui montre que $Q_t\varphi \in C_0(\mathbb{R})$. Ensuite,

$$Q_t\varphi(x) = E[\varphi(x + Y_t)] \underset{t\to 0}{\longrightarrow} \varphi(x)$$

grâce à la propriété (iii). La continuité uniforme de φ montre même que cette convergence est uniforme en x. Cela termine la preuve de la première assertion. La preuve de la seconde est facile en utilisant la propriété (ii). □

On déduit du Théorème 7.8 qu'on peut trouver une modification de $(Y_t)_{t\geq 0}$ dont les trajectoires sont càdlàg (en fait dans le Théorème 7.8 on supposait qu'on avait une famille de probabilités $(P_x)_{x\in E}$ correspondant aux différents points de départ possibles, mais la même preuve s'applique sans changement au cas où on considère le processus sous une seule mesure de probabilité).

7.5.2 Processus de branchement continu

Un processus de Markov $(X_t)_{t\geq 0}$ à valeurs dans $E = \mathbb{R}_+$ est appelé processus de branchement continu si son semigroupe $(Q_t)_{t\geq 0}$ vérifie pour tous $x, y \in \mathbb{R}_+$ et tout $t \geq 0$,

$$Q_t(x, \cdot) * Q_t(y, \cdot) = Q_t(x + y, \cdot).$$

Exercice. Vérifier alors que si X et X' sont deux processus de Markov indépendants (relativement à la même filtration) de semigroupe $(Q_t)_{t\geq 0}$, issus respectivement de x et de x', alors $(X_t + X'_t)_{t\geq 0}$ est aussi un processus de Markov de semigroupe $(Q_t)_{t\geq 0}$.

On fait les deux hypothèses de régularité suivantes :

(i) $Q_t(x, \{0\}) < 1$ pour tous $x > 0$ et $t > 0$;

(ii) $Q_t(x, \cdot) \longrightarrow \delta_x(\cdot)$ quand $t \to 0$, au sens de la convergence étroite des mesures de probabilité.

Proposition 7.13 *Sous les hypothèses précédentes, le semigroupe* $(Q_t)_{t\geq 0}$ *est de Feller. De plus, pour tout* $\lambda > 0$,

$$\int Q_t(x, dy)\, e^{-\lambda y} = e^{-x\psi_t(\lambda)}$$

où les fonctions $\psi_t :]0, \infty[\longrightarrow]0, \infty[$ *vérifient* $\psi_t \circ \psi_s = \psi_{t+s}$ *pour tous* $s, t \geq 0$.

Démonstration. Commençons par la deuxième assertion. Si $x, y > 0$, l'égalité $Q_t(x, \cdot) * Q_t(y, \cdot) = Q_t(x + y, \cdot)$ entraîne que

$$\Big(\int Q_t(x, dz)\, e^{-\lambda z}\Big)\Big(\int Q_t(y, dz)\, e^{-\lambda z}\Big) = \int Q_t(x + y, dz)\, e^{-\lambda z}.$$

Ainsi la fonction

$$x \longrightarrow -\log\Big(\int Q_t(x, dz)\, e^{-\lambda z}\Big)$$

est linéaire et croissante sur $\mathbb{R}_+$ donc de la forme $x\psi_t(\lambda)$ pour une constante $\psi_t(\lambda) > 0$ (le cas $\psi_t(\lambda) = 0$ est écarté à cause de (i)). Pour obtenir l'égalité $\psi_t \circ \psi_s = \psi_{t+s}$, on écrit

$$\begin{aligned}
\int Q_{t+s}(x, dz)\, e^{-\lambda z} &= \int Q_t(x, dy) \int Q_s(y, dz)\, e^{-\lambda z} \\
&= \int Q_t(x, dy)\, e^{-y\psi_s(\lambda)} \\
&= e^{-x\psi_t(\psi_s(\lambda))}.
\end{aligned}$$

Il reste à vérifier le caractère fellérien du semigroupe. Posons, pour tout $\lambda > 0$, $\varphi_\lambda(x) = e^{-\lambda x}$. Alors,

$$Q_t\varphi_\lambda = \varphi_{\psi_t(\lambda)} \in C_0(\mathbb{R}_+).$$

De plus, une application du théorème de Stone-Weierstrass montre que l'espace vectoriel engendré par les fonctions φ_λ, $\lambda > 0$ est dense dans $C_0(\mathbb{R}_+)$. Il en découle aisément que $Q_t\varphi \in C_0(\mathbb{R}_+)$ pour toute fonction $\varphi \in C_0(\mathbb{R}_+)$.

Enfin, si $\varphi \in C_0(\mathbb{R}_+)$, pour tout $x \geq 0$,

$$Q_t\varphi(x) = \int Q_t(x, dy)\, \varphi(y) \underset{t\to 0}{\longrightarrow} \varphi(x)$$

d'après la propriété (ii). En utilisant une remarque suivant la définition des semigroupes de Feller, cela suffit pour montrer que $\|Q_t\varphi - \varphi\| \longrightarrow 0$ quand $t \to 0$, ce qui termine la preuve. $\square$

Quelques références

K.L. Chung, R.J. Williams. Introduction to stochastic integration. Birkhäuser, 1990.

R. Durrett. Brownian motion and martingales in analysis. Wadsworth, 1984.

N. Ikeda, S. Watanabe. Stochastic differential equations and diffusion processes. North-Holland, 1988.

I. Karatzas, S. Shreve. Brownian motion and stochastic calculus. Springer, 1987.

D. Revuz, M. Yor. Continuous martingales and Brownian motion. Springer, 1991.

L.C.G. Rogers, D. Williams. Diffusions, Markov processes and martingales, vol.2. Wiley, 1987.