

ELEMENTS DE CALCUL STOCHASTIQUE

Monique Jeanblanc, Thomas Simon

IRBID, Septembre 2005

Table des matières

1	Notions générales de probabilités	5
1.1	Espace probabilisé	5
1.2	Indépendance, espérance conditionnelle	8
1.3	Théorème de Radon-Nikodym	12
1.4	Vecteurs Gaussiens	13
1.5	Convergence de variables aléatoires	14
2	Martingales et temps d'arrêt	17
2.1	Généralités sur les processus stochastiques	17
2.2	Temps d'arrêt	19
2.3	Martingales	20
3	Le mouvement brownien	23
3.1	Un peu d'histoire	23
3.2	Définitions, constructions, et premières propriétés	23
3.3	La propriété de Markov et ses conséquences	28
3.4	Equation de la chaleur	31
3.5	Mouvement brownien multidimensionnel	33
4	L'intégrale stochastique - Formule d'Itô	35
4.1	L'intégrale de Wiener	35
4.1.1	Le cas des fonctions en escalier	36
4.1.2	Le cas général	37
4.1.3	L'intégrale de Wiener vue comme processus gaussien	37
4.2	Application de l'intégrale de Wiener	39
4.2.1	Le processus d'Ornstein-Uhlenbeck	39
4.2.2	Le processus de Vařicek	42
4.3	L'intégrale stochastique générale	43
4.3.1	Cas des processus étagés	43
4.3.2	Cas général	44
4.4	Processus d'Itô	48
4.5	Formule d'Itô	49
4.6	Formule de Black & Scholes	51
5	Equations différentielles stochastiques	53
5.1	Définitions	53
5.2	Quelques propriétés	57
5.2.1	Propriété de Markov	57
5.2.2	Théorème de comparaison	57
5.2.3	Fonction d'échelle	58
5.2.4	Martingale exponentielle et condition de Novikov	59
5.3	Lien avec les équations aux dérivées partielles	59
5.3.1	Problème parabolique	59
5.3.2	Formule de Black & Scholes	60
5.3.3	Formule de Feynman-Kac	61
5.4	Processus de Bessel	63

5.4.1	Norme d'un mouvement Brownien n -dimensionnel	63
5.4.2	Définition générale	63
5.4.3	Probabilités de transition	64
5.5	Modèle de Cox-Ingersoll-Ross	64
6	Changement de probabilité - Théorème de Girsanov	67
6.1	Généralités	67
6.2	La formule de Cameron-Martin	67
6.3	Les deux théorèmes de Girsanov	69
6.4	Théorème de représentation prévisible	71
6.4.1	Représentation prévisible	71
6.5	Applications au mouvement brownien et aux EDS	71
6.5.1	Calcul d'espérances Exemple 1	71
6.5.2	Calcul d'espérances Exemple 2	72
6.5.3	Temps de passage du Brownien drifté	73
6.5.4	Solutions faibles d'EDS	73
6.6	Théorème fondamental de la finance : Probabilité risque neutre	74
6.6.1	Changement de numéraire	75
7	Introduction aux modèles poissonniens	79
7.1	Processus ponctuels sur \mathbb{R}^+	79
7.2	L'hypothèse de Poisson	80
7.3	La propriété de Markov forte et ses conséquences	83
7.4	Une formule d'Itô avec sauts	84
7.5	Des martingales	85
7.5.1	Exemple de changement de probabilité	86

Chapitre 1

Notions générales de probabilités

Dans ce chapitre introductif nous rassemblons les notions basiques de théorie des probabilités qui seront utilisées dans la suite du cours. Nous renvoyons aux ouvrages [7] [9] ou encore [24] pour les démonstrations non écrites ici, et aussi pour de plus amples informations. Ici comme dans les chapitres suivants, la plupart des exercices (dont les plus difficiles sont marqués d'un *) sont corrigés dans [8] ou dans [22].

1.1 Espace probabilisé

On considère un ensemble abstrait quelconque Ω , qui modélise le caractère aléatoire d'un certain phénomène et portera dans la suite le nom d'*ensemble fondamental*, ou d'*espace des chances*. On note $\mathcal{P}(\Omega)$ l'ensemble des parties de Ω . Rappelons que si $\text{Card } \Omega = n$, alors $\text{Card } \mathcal{P}(\Omega) = 2^n$. Par exemple si $\Omega = \{1, 2, 3\}$, alors

$$\mathcal{P}(\Omega) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \Omega\}.$$

Remarquons que $\mathcal{P}(\Omega)$ contient toujours Ω et \emptyset . Quand Ω est fini ou dénombrable, l'ensemble $\mathcal{P}(\Omega)$ décrit les *événements* aléatoires associés à l'expérience considérée. Quand Ω est infini non-dénombrable (ce sera le cas la plupart du temps), l'ensemble $\mathcal{P}(\Omega)$ est parfois trop gros pour pouvoir étudier mathématiquement la probabilité de ces événements. On a alors recours à la notion de tribu :

Définition 1.1 Une tribu ou σ -algèbre \mathcal{F} sur Ω est un sous-ensemble de $\mathcal{P}(\Omega)$ contenant Ω et vérifiant les propriétés suivantes :

- $A \in \mathcal{F} \implies A^c \in \mathcal{F}$ (stabilité par passage au complémentaire).
- $A_1, \dots, A_n, \dots \in \mathcal{F} \implies (A_1 \cap \dots \cap A_n \cap \dots) \in \mathcal{F}$ (stabilité par intersection dénombrable).

Il faut comprendre une tribu comme la quantité d'*information* associée à une certaine expérience aléatoire. Remarquons qu'une tribu contient forcément \emptyset et reste stable par intersection dénombrable, du fait de la première *loi de Morgan* :

$$\left(\bigcup_{i=1}^{+\infty} A_i \right)^c = \bigcap_{i=1}^{+\infty} A_i^c.$$

Les exemples les plus simples de tribus sur Ω sont $\{\emptyset, \Omega\}$ (tribu triviale), $\mathcal{P}(\Omega)$ (tribu maximale), ou $\{\emptyset, A, A^c, \Omega\}$ pour un certain $A \subset \Omega$ (tribu engendrée par A). Quand \mathcal{F} est une tribu sur Ω , on dit que (Ω, \mathcal{F}) est un *espace mesurable*. Une *sous-tribu* de \mathcal{F} est une tribu \mathcal{G} incluse dans \mathcal{F} , au sens où $A \in \mathcal{G} \implies A \in \mathcal{F}$. On montre très facilement la

Proposition 1.2 Toute intersection de tribus sur Ω est une tribu sur Ω .

Remarquons que cette propriété est fausse pour la réunion : une réunion de tribus n'est pas forcément une tribu. Par exemple si $A_1, A_2 \subset \Omega$, alors $\mathcal{F}_1 = \{\emptyset, A_1, A_1^c, \Omega\}$ et $\mathcal{F}_2 = \{\emptyset, A_2, A_2^c, \Omega\}$ sont des tribus, mais pas $\mathcal{F}_1 \cup \mathcal{F}_2$ en général, par exemple si $A_1 \cap A_2 \notin \mathcal{F}_1 \cup \mathcal{F}_2$. La définition suivante est fondamentale :

Définition 1.3 Soit \mathcal{A} une partie quelconque de $\mathcal{P}(\Omega)$. On appelle tribu engendrée par \mathcal{A} et on note $\sigma(\mathcal{A})$ l'intersection de toutes les tribus contenant \mathcal{A} .

Si \mathcal{F}_1 et \mathcal{F}_2 sont deux tribus, on note $\mathcal{F}_1 \vee \mathcal{F}_2$ la tribu engendrée par $\mathcal{F}_1 \cup \mathcal{F}_2$. C'est la plus petite tribu contenant les deux tribus $\mathcal{F}_1 \cup \mathcal{F}_2$. Remarquons que $\sigma(\mathcal{A})$ est toujours bien défini comme tribu : $\mathcal{P}(\Omega)$ est une tribu contenant \mathcal{A} et l'intersection d'une famille quelconque de tribus est une tribu. Cependant, il est en général difficile (voire impossible) de décrire $\sigma(\mathcal{A})$ plus précisément, sauf quand \mathcal{A} est un ensemble fini. Un exemple important est le suivant :

Définition 1.4 On appelle tribu des Boréliens de \mathbb{R} et on note $\mathcal{B}(\mathbb{R})$ (ou \mathcal{B} quand aucune ambiguïté n'est possible) la tribu engendrée par les intervalles ouverts $]a, b[$, $a > b \in \mathbb{R}$.

Dans la définition on aurait évidemment pu remplacer "ouvert" par "fermé" ou "ouvert à droite fermé à gauche" ...etc, vues les propriétés de stabilité des tribus. Soulignons qu'il est possible (mais difficile, voir [14]) de construire des sous ensembles de \mathbb{R} qui ne sont pas des Boréliens, autrement dit $\mathcal{B}(\mathbb{R}) \subsetneq \mathcal{P}(\mathbb{R})$.

Définition 1.5 Soit (Ω, \mathcal{F}) et (E, \mathcal{E}) deux espaces mesurables. Une application $f : \Omega \rightarrow E$ est dite $(\mathcal{F}, \mathcal{E})$ mesurable (ou mesurable lorsqu'il n'y a pas d'ambiguïté) si $f^{-1}(A) \in \mathcal{F}$ pour tout $A \in \mathcal{E}$, où l'on a noté

$$f^{-1}(A) = \{\omega \in \Omega \mid f(\omega) \in A\}.$$

Une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ est dite borélienne si elle est $(\mathcal{B}(\mathbb{R}), \mathcal{B}(\mathbb{R}))$ -mesurable, autrement dit si $f^{-1}(A) \in \mathcal{B}(\mathbb{R})$ pour tout $A \in \mathcal{B}(\mathbb{R})$. Par construction des Boréliens, il suffit que cette propriété soit vérifiée pour les intervalles ouverts A . En particulier, les fonctions continues sont boréliennes, puisque l'image réciproque d'un intervalle ouvert par une fonction continue est une réunion d'intervalles ouverts, donc un Borélien.

Définition 1.6 Soit (Ω, \mathcal{F}) un espace mesurable. Une variable aléatoire réelle (v.a.r.) X est une application mesurable de (Ω, \mathcal{F}) dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

L'exemple fondamental de variable aléatoire est la fonction indicatrice d'un ensemble $A \in \mathcal{F}$:

$$\mathbb{1}_A : \omega \mapsto \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{sinon.} \end{cases}$$

Remarquons que les variables aléatoires réelles sont stables par composition avec les fonctions boréliennes : si X est une v.a.r. \mathcal{G} -mesurable et f une fonction borélienne, alors $f(X)$ est une v.a.r. \mathcal{G} -mesurable. Les v.a.r. peuvent être approchées en limite croissante simple par des v.a.r. "en escalier" qui sont du type

$$\sum_{i=1}^n a_i \mathbb{1}_{A_i}$$

avec $A_i \in \mathcal{G}$. L'espace $(\mathbb{R}, \mathcal{B})$ est l'exemple plus simple d'espace d'état associé à une variable aléatoire. On peut cependant définir des v.a. sur des espaces plus gros, comme $\mathcal{C}(\mathbb{R}, \mathbb{R})$, l'espace des fonctions continues de \mathbb{R} dans \mathbb{R} , muni de sa tribu des Boréliens $\mathcal{B}(\mathcal{C}(\mathbb{R}, \mathbb{R}))$. Cette dernière est la tribu engendrée par les polynômes. Une v.a. à valeurs dans $\mathcal{C}(\mathbb{R}, \mathbb{R})$ portera le nom de processus.

Définition 1.7 La tribu engendrée par une variable aléatoire X définie sur (Ω, \mathcal{F}) est l'ensemble $\sigma(X) = \{X^{-1}(A), A \in \mathcal{B}(\mathbb{R})\}$.

On vérifie que $\sigma(X)$ est une tribu contenue dans \mathcal{F} : c'est la plus petite tribu sur Ω rendant X mesurable. Si \mathcal{G} est une sous-tribu de \mathcal{F} , alors une v.a.r. X est \mathcal{G} -mesurable si $\sigma(X) \subset \mathcal{G}$. On a la

Proposition 1.8 Soit X une v.a.r. et Y une v.a.r. $\sigma(X)$ -mesurable. Alors il existe une fonction borélienne $f : \mathbb{R} \rightarrow \mathbb{R}$ telle que $Y = f(X)$.

Nous introduisons maintenant la définition centrale de ce chapitre :

Définition 1.9 Soit (Ω, \mathcal{F}) un espace mesurable. Une probabilité sur (Ω, \mathcal{F}) est une application \mathbb{P} de \mathcal{F} dans $[0, 1]$ telle que

- $\mathbb{P}[\Omega] = 1$
- Si $A_1 \dots A_n \dots$ sont des parties de \mathcal{F} deux à deux disjointes, alors

$$\mathbb{P}\left[\bigcup_{n=0}^{\infty} A_n\right] = \sum_{n=0}^{\infty} \mathbb{P}[A_n].$$

La deuxième propriété porte le nom de σ -additivité. On utilise souvent l'écriture avec les fonctions indicatrices :

$$\mathbb{P}[A] = \int_A d\mathbb{P} = \int_{\Omega} \mathbb{1}_A d\mathbb{P}.$$

Un espace mesurable (Ω, \mathcal{F}) muni d'une probabilité \mathbb{P} s'écrit $(\Omega, \mathcal{F}, \mathbb{P})$ et on parle alors d'*espace de probabilité* ou d'*espace probabilisé*. On démontre aisément les propriétés suivantes :

- (a) $\mathbb{P}[A] + \mathbb{P}[A^c] = 1$ pour tout $A \in \mathcal{F}$.
- (b) $A \subset B \implies \mathbb{P}[A] \leq \mathbb{P}[B]$
- (c) $\mathbb{P}[A \cup B] + \mathbb{P}[A \cap B] = \mathbb{P}[A] + \mathbb{P}[B]$ pour tout $A, B \in \mathcal{F}$.
- (d) Si les A_n forment une suite croissante (resp. décroissante) d'éléments de \mathcal{F} , c'est-à-dire si $A_n \subset A_{n+1}$ (resp. $A_n \supset A_{n+1}$), et si $A = \bigcup_n A_n$ (resp. $A = \bigcap_n A_n$) alors $\mathbb{P}[A_n] \rightarrow \mathbb{P}[A]$ quand $n \uparrow +\infty$.

Exercice 1.10 (a) Montrer la généralisation suivante (connue sous le nom de Formule de Poincaré) de la propriété (c) : pour tout $A_1 \dots A_n \in \mathcal{F}$,

$$\mathbb{P}\left[\bigcup_{i=1}^n A_i\right] = \sum_{k=1}^n (-1)^{k+1} \sum_{1 \leq i_1 < \dots < i_k \leq n} \mathbb{P}[A_{i_1} \cap \dots \cap A_{i_k}].$$

Sur une table reposent n lettres distinctes correspondant à n enveloppes. Un postier est chargé de cacheter les lettres. Calculer la probabilité qu'aucune lettre ne soit dans la bonne enveloppe.¹ Quelle est la limite de cette probabilité quand $n \uparrow +\infty$?

(b)* Montrer le *principe d'inclusion-exclusion* : pour tout $A_1 \dots A_n \in \mathcal{F}$ et pour tout $p \in [1, \dots, n]$,

$$(-1)^p \left(\mathbb{P}\left[\bigcup_{i=1}^n A_i\right] - \left(\sum_{k=1}^p (-1)^{k+1} \sum_{1 \leq i_1 < \dots < i_k \leq n} \mathbb{P}[A_{i_1} \cap \dots \cap A_{i_k}] \right) \right) \geq 0.$$

(c)* Montrer l'*inégalité de Kounias* : pour tout $A_1 \dots A_n \in \mathcal{F}$

$$\mathbb{P}\left[\bigcup_{i=1}^n A_i\right] \geq \left(\sum_{i=1}^n \mathbb{P}[A_i] \right) - \max_{k=1 \dots n} \left(\sum_{j \neq k} \mathbb{P}[A_j \cap A_k] \right).$$

En quoi cette inégalité est-elle plus forte que le principe d'inclusion-exclusion pour $p = 2$?

Le théorème suivant est souvent utile pour identifier deux mesures de probabilité :

Théorème 1.11 [Théorème de classe monotone] Soient \mathbb{P} et \mathbb{Q} deux probabilités sur l'espace mesurable (Ω, \mathcal{F}) . Soit $\mathcal{C} \subset \mathcal{F}$ une famille stable par intersection finie et engendrant \mathcal{F} . Si $\mathbb{P}[A] = \mathbb{Q}[A]$ pour tout $A \in \mathcal{C}$, alors $\mathbb{P} = \mathbb{Q}$.

On prendra soin, pour appliquer ce théorème, de vérifier la stabilité par intersection de \mathcal{C} , c'est-à-dire de montrer que si $C_1 \in \mathcal{C}, C_2 \in \mathcal{C}$, alors $C_1 \cap C_2 \in \mathcal{C}$. Les intervalles ouverts ou fermés forment un exemple de famille \mathcal{C} sur la tribu des Boréliens. Le théorème de classe monotone permet de montrer l'unicité de la *mesure de Lebesgue* sur $[0, 1]$, qui est définie comme l'unique mesure de probabilité λ sur $\mathcal{B}([0, 1])$ telle que $\lambda([a, b]) = b - a$ pour tout $a < b$. L'existence de la mesure de Lebesgue est un problème délicat pour lequel nous renvoyons à [14].

Définition 1.12 Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité. Un ensemble $A \in \mathcal{F}$ est dit *négligeable* pour \mathbb{P} si $\mathbb{P}[A] = 0$.

Par exemple les singletons $\{x\}$ sont négligeables dans $[0, 1]$ pour la mesure de Lebesgue (on dit alors que cette dernière ne charge pas les points). Par σ -additivité, une réunion dénombrable d'ensembles négligeables est négligeable, mais pas forcément une réunion non dénombrable. On dit qu'une propriété est vraie *presque sûrement* (p.s.) si elle est vraie en dehors d'un ensemble négligeable. Un espace $(\Omega, \mathcal{F}, \mathbb{P})$ est dit *complet* si \mathcal{F} contient tous les ensembles G (dits négligeables au sens large) tels que $\inf\{\mathbb{P}[F] : F \in \mathcal{F}, G \subset F\} = 0$. Partant d'un espace de probabilité donné $(\Omega, \mathcal{F}, \mathbb{P})$, il est toujours possible de le *compléter* en considérant $\mathcal{F}' = \sigma(\mathcal{F}, \mathcal{N})$ où \mathcal{N} est l'ensemble des négligeables au sens large pour \mathbb{P} .

¹Ce problème est également connu sous le nom de *problème des échangistes*.

Définition 1.13 Soit X une v.a.r. définie sur $(\Omega, \mathcal{F}, \mathbb{P})$. La loi de X est la probabilité \mathbb{P}_X sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ définie comme mesure-image de \mathbb{P} par X : pour tout $A \in \mathcal{B}(\mathbb{R})$,

$$\mathbb{P}_X(A) = \mathbb{P}[X \in A] = \mathbb{P}[\{\omega \in \Omega / X(\omega) \in A\}].$$

En particulier, comme $\mathcal{B}(\mathbb{R}) = \sigma\{]-\infty, x] / x \in \mathbb{R}\}$, on voit par classe monotone que la loi de X est caractérisée par les quantités $F_X(x) = \mathbb{P}[X \leq x]$. On appelle F_X la *fonction de répartition* de X . On utilisera souvent la remarque élémentaire suivante : si X et Y sont deux v.a.r. telles que $\mathbb{P}[X \leq a] = \mathbb{P}[Y \leq a]$ pour tout $a \in \mathbb{R}$, alors X et Y ont même loi, ce que l'on notera

$$X \stackrel{d}{=} Y.$$

Attention : l'égalité en loi n'est pas l'égalité presque sûre. Par exemple, si X est une variable symétrique (par exemple une Gaussienne centrée), on a $X \stackrel{d}{=} -X$ mais ces deux variables ne sont pas égales presque sûrement. La fonction F_X est continue à droite et limitée à gauche (càdlàg), croissante, tend vers 0 en $-\infty$ et vers 1 en $+\infty$. On a $F_X(x) = \mathbb{P}[X < x]$ si $\mathbb{P}[X = x] = 0$, autrement dit si \mathbb{P}_X ne charge pas x . En particulier, F_X est une fonction continue sur \mathbb{R} si et seulement si \mathbb{P}_X ne charge aucun point (on dit alors que X n'est pas atomique).

Exemple 1.14 La fonction de répartition d'une *variable gaussienne centrée réduite* X (on notera alors $X \sim \mathcal{N}(0, 1)$) est donnée par

$$F_X(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-y^2/2} dy.$$

Ici on a $\mathbb{P}[X \leq x] = \mathbb{P}[X < x]$ par propriété de l'intégrale d'une fonction mesurable bornée.

Dans cet exemple, F_X est l'intégrale d'une fonction positive f_X qu'on appelle la *densité* de X . Si X a une densité, alors la fonction F_X est continue, mais on prendra garde que la réciproque est fautive : il existe des variables aléatoires non atomiques et sans densité. La densité f_X d'une variable aléatoire X est la dérivée de la fonction de répartition F_X quand cette dérivée existe. On peut alors écrire $\mathbb{P}[X \in A] = \int_A f_X(x) dx$. En particulier $\mathbb{P}[X \in [a, b]] = \int_a^b f_X(x) dx$. On utilise parfois la notation différentielle $\mathbb{P}[X \in dx]$ pour désigner $f(x)dx$.

1.2 Indépendance, espérance conditionnelle

Définition 1.15 Soit X une v.a.r. telle que

$$\int_{\Omega} |X(\omega)| d\mathbb{P} \stackrel{def}{=} \int_{\mathbb{R}} |x| d\mathbb{P}_X(x) < +\infty.$$

On définit alors l'espérance de X par

$$\mathbb{E}[X] = \int_{\Omega} X(\omega) d\mathbb{P} = \int_{\mathbb{R}} x d\mathbb{P}_X(x).$$

Il existe des variables aléatoires qui n'ont pas d'espérance, c'est-à-dire pour lesquelles l'intégrale $\int_{\Omega} X(\omega) d\mathbb{P}$ n'a pas de sens. Par exemple la variable ayant pour densité $f_X(x) = 0$ si $|x| < 1$ et $f_X(x) = 1/2x^2$ si $|x| \geq 1$ vérifie

$$\int_{\mathbb{R}} |x| d\mathbb{P}_X(x) = \int_{\mathbb{R} \cap [-1, 1]^c} |x| \frac{dx}{2x^2} = \int_1^{+\infty} \frac{dx}{x} = +\infty$$

et donc $\mathbb{E}[X]$ n'a pas de sens. On dit que X est *intégrable* quand on peut définir son espérance, autrement dit si $|X|$ a une espérance finie. On note $L^1(\Omega, \mathcal{F}, \mathbb{P})$ (ou $L^1(\Omega)$ quand il n'y a pas d'ambiguïté) l'ensemble des v.a. intégrables. De la même façon, on définit

$$L^p(\Omega, \mathcal{F}, \mathbb{P}) = \left\{ X \text{ v.a.r.} / \int_{\mathbb{R}} |x|^p d\mathbb{P}_X(x) < +\infty \right\}$$

pour tout $p \geq 1$ et $L^\infty(\Omega, \mathcal{F}, \mathbb{P}) = \{X \text{ v.a.r.} \exists K / X(\omega) \leq K < +\infty \text{ p.s.}\}$. Comme \mathbb{P} est une mesure finie, on montre aisément la chaîne d'inclusions : $L^\infty(\Omega) \subset \dots \subset L^p(\Omega) \subset \dots \subset L^1(\Omega)$. L'espace

$L^2(\Omega, \mathcal{F})$ des v.a. de carré intégrable joue un rôle essentiel : c'est un espace de Hilbert, muni du produit scalaire $\langle X, Y \rangle = E(XY)$.

Si $\Phi : \mathbb{R} \rightarrow \mathbb{R}$ est une fonction borélienne telle que $\Phi(X)$ soit intégrable (ce qui a lieu par exemple si Φ est bornée, puisque $L^\infty(\Omega) \subset L^1(\Omega)$), on note

$$\mathbb{E}[\Phi(X)] = \int_{\Omega} \Phi(X) d\mathbb{P} = \int_{\mathbb{R}} \Phi(x) d\mathbb{P}_X(x).$$

Dans cette formule, on a vu la loi de $\Phi(X)$ comme l'image par Φ de la loi de \mathbb{P}_X . On peut aussi utiliser la formule directe

$$\mathbb{E}[\Phi(X)] = \int_{\mathbb{R}} x d\mathbb{P}_{\Phi(X)}(x),$$

mais alors il faut calculer la mesure $\mathbb{P}_{\Phi(X)}$, ce qui est parfois difficile. La *fonction caractéristique* ou *transformée de Fourier* d'une v.a.r. X , est la fonction

$$\psi_X(t) = \mathbb{E}[e^{itX}] = \int_{\mathbb{R}} e^{itx} \mathbb{P}_X(dx).$$

Elle est toujours définie puisque $e^{itX} \in L^\infty$, et caractérise la loi de X au sens suivant :

Théorème 1.16 *Soient X et Y deux v.a.r. On a l'équivalence*

$$X \stackrel{d}{=} Y \iff \psi_X(t) = \psi_Y(t) \text{ pour tout } t \in \mathbb{R}.$$

Si la transformée de Fourier $\psi_X(t)$ appartient à $L^1(dt)$, c'est-à-dire si son module est intégrable par rapport à la mesure de Lebesgue dt , alors X a une densité f_X donnée par la *formule d'inversion de Fourier* :

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \psi_X(t) dt$$

Quand X est positive, on peut définir sa *transformée de Laplace* :

$$\varphi_X(\lambda) = \mathbb{E}[e^{-\lambda X}]$$

pour tout $\lambda > 0$, qui caractérise aussi la loi de X :

Théorème 1.17 *Soient X et Y deux v.a.r. positives. On a l'équivalence*

$$X \stackrel{d}{=} Y \iff \varphi_X(\lambda) = \varphi_Y(\lambda) \text{ pour tout } \lambda > 0.$$

Cependant, il n'y a pas de formule d'inversion simple pour les transformées de Laplace comme pour les transformées de Fourier. Dans certains cas, on peut définir les transformées de Laplace de variables non nécessairement positives :

Exemple 1.18 Transformée de Laplace d'une v.a. gaussienne : Soit X une variable gaussienne de loi $\mathcal{N}(m, \sigma^2)$, i.e. une v.a.r. de densité donnée par

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-m)^2/2\sigma^2}.$$

On a alors $\mathbb{E}[e^{\lambda X}] = e^{(\lambda m + \lambda^2 \sigma^2/2)}$ pour tout $\lambda \in \mathbb{R}$. Réciproquement, si X une v.a.r. ayant cette transformée de Laplace, alors $X \sim \mathcal{N}(m, \sigma^2)$.

Quand $\varphi_X(\lambda)$ est définie sur un voisinage ouvert contenant 0, le *principe du prolongement analytique* permet d'écrire $\psi_X(t) = \varphi_X(it)$. Ainsi, pour $X \sim \mathcal{N}(m, \sigma^2)$, on a $\psi_X(t) = e^{(itm - t^2 \sigma^2/2)}$.

Propriétés de l'espérance. (a) Linéarité : si X, Y sont des v.a.r. intégrables, alors pour tout $a, b \in \mathbb{R}$ $aX + bY$ est intégrable et $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$.

(b) Positivité : si $X \leq Y$ p.s. alors $\mathbb{E}[X] \leq \mathbb{E}[Y]$. En particulier, si $X \geq 0$ p.s. alors $\mathbb{E}[X] \geq 0$. De plus, si $X \geq 0$ p.s. et si $\mathbb{E}[X] = 0$, alors $X = 0$ p.s.

(c) Inégalité de Jensen : si Φ est une fonction convexe et X une v.a.r. telle que $\Phi(X)$ est intégrable, alors

$$\mathbb{E}[\Phi(X)] \geq \Phi(\mathbb{E}[X]).$$

En particulier, $x \mapsto x^p, p \geq 1$ et $x \mapsto e^x$ sont des fonctions convexes et on a $\mathbb{E}[X^p] \geq (\mathbb{E}[X])^p, p \geq 1$ et $\mathbb{E}[e^X] \geq e^{\mathbb{E}[X]}$. En revanche, $x \mapsto x^p, p \in [0, 1]$ et $x \mapsto \log x$ sont des fonctions concaves (soit l'opposé d'une fonction convexe), et on a les inégalités en sens inverse : $\mathbb{E}[X^p] \leq (\mathbb{E}[X])^p, p \in [0, 1]$ et $\mathbb{E}[\log X] \leq \log(\mathbb{E}[X])$.

Soit $X \in L^1$; considérons $X_n = X \mathbb{1}_{|X| > n} \rightarrow 0$ p.s. quand $n \uparrow +\infty$. Comme $|X_n| \leq |X| \in L^1$, le théorème de convergence dominée de Lebesgue assure que

$$\mathbb{E}[|X_n|] \rightarrow 0$$

quand $n \uparrow +\infty$. Ceci motive la définition suivante :

Définition 1.19 Une famille $\{X_i, i \in I\}$ de v.a.r. dans $L^1(\Omega)$ est dite uniformément intégrable (U.I.) si

$$\sup_{i \in I} \mathbb{E}[|X_i| \mathbb{1}_{|X_i| > n}] \rightarrow 0$$

quand $n \uparrow +\infty$.

On vient de voir que $\mathbb{E}[|X_i| \mathbb{1}_{|X_i| > n}] \rightarrow 0$ quand $n \uparrow +\infty$ pour chaque $i \in I$. La subtilité de l'uniforme intégrabilité vient de ce que l'on demande à cette convergence d'être *uniforme* sur I . Remarquons cependant que s'il existe $Y \in L^1(\Omega)$ telle que $|X_i| \leq Y$ pour tout $i \in I$, alors la famille $\{X_i, i \in I\}$ est U.I. La notion d'uniforme intégrabilité joue un rôle primordial en théorie des martingales. Venons-en maintenant à une définition bien connue :

Définition 1.20 Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité. On dit que deux sous-tribus \mathcal{F}_1 et \mathcal{F}_2 sont indépendantes si $\mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B]$ pour tout $A \in \mathcal{C}_1$ et $B \in \mathcal{F}_2$. On notera alors $\mathcal{F}_1 \perp \mathcal{F}_2$.

Par classe monotone, remarquons que pour que \mathcal{F}_1 et \mathcal{F}_2 soient indépendantes, il faut et il suffit que $\mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B]$ pour tout $A \in \mathcal{C}_1$ et $B \in \mathcal{C}_2$, où \mathcal{C}_i est une famille stable par intersection finie et telle que $\sigma(\mathcal{C}_i) = \mathcal{F}_i, i = 1, 2$.

Définition 1.21 Une variable aléatoire X est indépendante d'une sous-tribu \mathcal{G} si les tribus $\sigma(X)$ et \mathcal{G} sont indépendantes. En particulier, deux variables X et Y sont indépendantes si les tribus $\sigma(X)$ et $\sigma(Y)$ sont indépendantes.

Par théorème de classe monotone, on voit qu'une v.a. X est indépendante de la sous-tribu \mathcal{G} si et seulement si $\mathbb{P}[A \cap \{X \leq x\}] = \mathbb{P}[A]\mathbb{P}[X \leq x]$ pour tout $x \in \mathbb{R}$ et $A \in \mathcal{G}$. Ainsi, deux v.a.r. X et Y sont indépendantes si et seulement si $\mathbb{P}[X \leq x, Y \leq y] = \mathbb{P}[X \leq x]\mathbb{P}[Y \leq y]$ pour tout $x, y \in \mathbb{R}$. Un autre critère utile est le suivant :

$$X \perp Y \iff \mathbb{E}[e^{itX} e^{isY}] = \mathbb{E}[e^{itX}] \mathbb{E}[e^{isY}] \text{ pour tout } s, t \in \mathbb{R}.$$

Si X et Y sont à valeurs positives, on peut considérer les transformées de Laplace au lieu des transformées de Fourier. Enfin, si X et Y sont des variables indépendantes ayant une densité sur \mathbb{R} , alors le couple (X, Y) a une densité sur \mathbb{R}^2 donnée par la *formule du produit* :

$$f_{X,Y}(x, y) = f_X(x)f_Y(y).$$

Définition 1.22 Une famille de sous-tribus $\{\mathcal{F}_i, i \in I\}$ est indépendante si toutes les sous-familles finies le sont, c'est-à-dire si

$$\mathbb{P}[A_{i_1} \cap \dots \cap A_{i_n}] = \prod_{k=1}^n \mathbb{P}[A_{i_k}]$$

pour tout $n \geq 2$ et pour tout $A_{i_k} \in \mathcal{F}_{i_k}$ avec $i_{k_1} \neq i_{k_2}$ si $k_1 \neq k_2$.

La définition est identique pour une famille de variables aléatoires. L'exercice suivant (*lemme de Borel-Cantelli*) est classique :

Exercice 1.23 Soit $\{A_n, n \geq 1\}$ une famille d'événements de $(\Omega, \mathcal{F}, \mathbb{P})$. On pose

$$E = \bigcap_{n \geq 1} \left(\bigcup_{k \geq n} A_k \right).$$

- (a) Montrer que si la série de terme général $\mathbb{P}[A_n]$ est convergente, alors $\mathbb{P}[E] = 0$.
- (b) Montrer que si la série de terme général $\mathbb{P}[A_n]$ est divergente et que si les A_n sont mutuellement indépendants, alors $\mathbb{P}[E] = 1$.

La notion suivante jouera un rôle fondamental dans la suite :

Théorème 1.24 Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité, \mathcal{G} une sous-tribu de \mathcal{F} et X une v.a.r. intégrable. Il existe une unique v.a.r. Z \mathcal{G} -mesurable telle que

$$\mathbb{E}[X \mathbb{1}_A] = \mathbb{E}[Z \mathbb{1}_A]$$

pour tout $A \in \mathcal{G}$. On appelle Z l'espérance conditionnelle de X sachant \mathcal{G} et on note $Z = \mathbb{E}[X | \mathcal{G}]$. Elle est caractérisée par

$$\mathbb{E}(XG) = \mathbb{E}(ZG), \forall G, \text{ v.a. bornée } \mathcal{G}\text{-mesurable.}$$

Quand $X \in L^2(\Omega)$, il existe une importante interprétation hilbertienne de l'espérance conditionnelle : $\mathbb{E}[X | \mathcal{G}]$ est la projection de X sur l'espace des v.a. \mathcal{G} -mesurables de carré intégrable, c'est-à-dire l'unique v.a. \mathcal{G} mesurable qui minimise $\mathbb{E}[(X - Y)^2]$ parmi les v.a. Y qui sont \mathcal{G} -mesurables. Quand $\mathcal{G} = \sigma(Y)$ pour une certaine v.a. Y , on note parfois $\mathbb{E}[X | \mathcal{G}] = \mathbb{E}[X | Y]$ (espérance conditionnelle sachant Y). Le point important est qu'il s'agit d'une variable mesurable par rapport à $\sigma(Y)$ et donc une fonction déterministe de Y : il existe $\psi : \mathbb{R} \rightarrow \mathbb{R}$ borélienne telle que $\mathbb{E}[X | Y] = \psi(Y)$. En général il est cependant difficile de calculer ψ explicitement. La caractérisation de $\mathbb{E}[X | Y]$ est la suivante : c'est l'unique v.a. de la forme $\psi(Y)$ où ψ est une fonction borélienne (définie \mathbb{P}_Y p.s.) telle que

$$\mathbb{E}[X\varphi(Y)] = \mathbb{E}[\psi(Y)\varphi(Y)]$$

pour toute fonction φ borélienne bornée. En plus de la linéarité et de la positivité (analogues à celles de l'espérance classique), l'espérance conditionnelle possède les propriétés suivantes (les égalités entre v.a. sont p.s.) :

- (a) $\mathbb{E}[\mathbb{E}[X | \mathcal{G}]] = \mathbb{E}[X]$.
- (b) Si X est \mathcal{G} -mesurable, $\mathbb{E}[X | \mathcal{G}] = X$.
- (c) Si X est indépendante de \mathcal{G} , $\mathbb{E}[X | \mathcal{G}] = \mathbb{E}[X]$.
- (d) Si \mathcal{G} est la tribu triviale, $\mathbb{E}[X | \mathcal{G}] = \mathbb{E}[X]$.
- (e) Si \mathcal{G} et \mathcal{H} sont deux tribus telles que $\mathcal{H} \subset \mathcal{G}$, alors $\mathbb{E}[X | \mathcal{H}] = \mathbb{E}[\mathbb{E}[X | \mathcal{G}] | \mathcal{H}] = \mathbb{E}[\mathbb{E}[X | \mathcal{H}] | \mathcal{G}]$ (propriété d'emboîtement).

Si Y est \mathcal{G} -mesurable et XY intégrable, $\mathbb{E}[XY | \mathcal{G}] = Y\mathbb{E}[X | \mathcal{G}]$

- (g) Si $X \perp Y$ et si $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ est une fonction borélienne bornée, alors $\mathbb{E}[\varphi(X, Y) | Y] = \mathbb{E}[\varphi(X, y)]_{y=Y}$: pour calculer $\mathbb{E}[\varphi(X, Y) | Y]$, on explicite la fonction Ψ telle que $\Psi(y) = \mathbb{E}[\varphi(X, y)]$, puis on remplace y par Y pour obtenir la v.a. $\Psi(Y)$.

Exercice 1.25 Soit X une variable positive sur $(\Omega, \mathcal{F}, \mathbb{P})$ et \mathcal{G} une sous-tribu de \mathcal{F} . Montrer que p.s.

- (a) $\{\mathbb{E}[X | \mathcal{G}] = 0\} \subset \{X = 0\}$.
- (b) $\{X = +\infty\} \subset \{\mathbb{E}[X | \mathcal{G}] = +\infty\}$.

Exercice 1.26 * Soit $Z \in L^1(\Omega, \mathcal{F}, \mathbb{P})$.

- (a) Montrer que pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que $\mathbb{P}[A] < \delta \implies \mathbb{E}[|X| \mathbb{1}_A] < \varepsilon$.
- (b) Soit $G = \{\mathcal{G} \subset \mathcal{F}\}$ une famille de sous-tribus de \mathcal{F} . Dédurre du (a) que la famille $\{\mathbb{E}[X | \mathcal{G}], \mathcal{G} \in G\}$ est U.I.

Densité conditionnelle : soit (X, Y) un couple de v.a.r. ayant une densité $f(x, y)$. Les densités marginales de X et Y sont données par

$$f_X(x) = \int_{\mathbb{R}} f(x, y) dy \quad \text{et} \quad f_Y(y) = \int_{\mathbb{R}} f(x, y) dx.$$

Quand $X \perp Y$ on a la formule du produit : $f(x, y) = f_X(x)f_Y(y)$. Quand X et Y ne sont plus indépendantes, la formule du produit est remplacée par une formule de *désintégration* : on pose

$$f_{X/Y=y}(x) = \frac{f(x, y)}{f_Y(y)}$$

si $f_Y(y) \neq 0$ et $f_{X/Y=y}(x) = 0$ si $f_Y(y) = 0$, et on remarque que nécessairement $f(x, y) = f_{X/Y=y}(x)f_Y(y)$ p.s. en (x, y) . En effet, si $f_Y(y) = 0$, alors $f(x, y) = 0$ p.s. en x vu que $f(x, y) \geq 0$. La fonction $f_{X/Y=y}(x)$ peut être vue comme la *densité conditionnelle* de X sachant $Y = y$. En effet pour toute fonction ψ mesurable bornée on a

$$\begin{aligned} \int_{\mathbb{R}^2} \psi(x, y) f_Y(y) f_{X/Y=y}(x) dx dy &= \mathbb{E}[\psi(X, Y)] = \mathbb{E}[\mathbb{E}[\psi(X, Y) | Y]] \\ &= \int_{\mathbb{R}} f_Y(y) \mathbb{E}[\psi(X, y) | Y]_{Y=y} dy \end{aligned}$$

d'où l'on déduit

$$\int_{\mathbb{R}} \psi(x, y) f_{X/Y=y}(x) dy = \mathbb{E}[\psi(X, y) | Y]_{Y=y}$$

par identification, ce qui signifie bien que $f_{X/Y=y}(x)$ est la densité conditionnelle de X sachant $Y = y$. On peut définir $f_{Y/X=x}(y)$, densité conditionnelle de Y sachant $X = x$, de manière analogue.

1.3 Théorème de Radon-Nikodym

Définition 1.27 Soient \mathbb{P} et \mathbb{Q} deux probabilités définies sur le même espace (Ω, \mathcal{F}) .

- (a) On dit que \mathbb{P} est absolument continue par rapport à \mathbb{Q} et on écrit $\mathbb{P} \ll \mathbb{Q}$ si pour tout $A \in \mathcal{F}$, $\mathbb{P}[A] = 0 \Rightarrow \mathbb{Q}[A] = 0$ (les négligeables pour \mathbb{P} sont aussi négligeables pour \mathbb{Q}).
- (b) On dit que \mathbb{P} et \mathbb{Q} sont équivalentes et on écrit $\mathbb{P} \sim \mathbb{Q}$ si pour tout $A \in \mathcal{F}$, $\mathbb{P}[A] = 0 \Leftrightarrow \mathbb{Q}[A] = 0$ (\mathbb{P} et \mathbb{Q} ont les mêmes négligeables.)
- (c) On dit que \mathbb{P} et \mathbb{Q} sont étrangères et on écrit $\mathbb{P} \perp \mathbb{Q}$ s'il existe $A \in \mathcal{F}$ tel que $\mathbb{P}[A] = 0$ et $\mathbb{Q}[A] = 1$. (A est négligeable pour \mathbb{P} et certain pour \mathbb{Q} .)

Le théorème suivant, dû à Radon-Nikodym, sera très utile pour comprendre le théorème de Girsanov :

Théorème 1.28 Soient \mathbb{P} et \mathbb{Q} deux probabilités définies sur le même espace (Ω, \mathcal{F}) . Alors $\mathbb{P} \ll \mathbb{Q}$ si et seulement si il existe une variable $Z \geq 0$ \mathcal{F} -mesurable et d'espérance 1 sous \mathbb{Q} telle que pour tout $A \in \mathcal{F}$,

$$\mathbb{P}[A] = \mathbb{E}_{\mathbb{Q}}[Z \mathbb{1}_A].$$

On appelle Z la *densité de Radon-Nikodym* de \mathbb{P} par rapport à \mathbb{Q} et on écrit formellement $d\mathbb{P} = Z d\mathbb{Q}$ ou encore

$$Z = \frac{d\mathbb{P}}{d\mathbb{Q}}.$$

Ceci est motivé par l'écriture suivante, pour tout $A \in \mathcal{F}$,

$$\mathbb{P}[A] = \int_A d\mathbb{P}(\omega) = \int_A \frac{d\mathbb{P}}{d\mathbb{Q}}(\omega) d\mathbb{Q}(\omega) = \mathbb{E}_{\mathbb{Q}}[Z \mathbb{1}_A].$$

De plus, dans le cas où $\mathbb{P} \sim \mathbb{Q}$, on voit facilement que $Z > 0$ p.s. et que $d\mathbb{Q} = Z^{-1} d\mathbb{P}$.

Exercice 1.29 (a) Soit U une variable de Bernoulli sous \mathbb{P} définie par $\mathbb{P}[U = 0] = 1 - p$ et $\mathbb{P}[U = 1] = p$. Soit Z la variable positive définie par $Z = \lambda U + \mu(1 - U)$. Montrer que $\mathbb{E}[Z] = 1$. Si on pose $d\mathbb{Q} = Z d\mathbb{P}$, montrer que $\mathbb{Q}[U = 1] = \lambda p$: sous \mathbb{Q} , U est une variable de Bernoulli de paramètre λp .

(b) Soit X une v.a. de loi $\mathcal{N}(m, \sigma^2)$ sous \mathbb{P} et soit $Z = \exp[h(X - m) - h^2 \sigma^2 / 2]$. Si on pose $d\mathbb{Q} = Z d\mathbb{P}$, alors sous \mathbb{Q} , X est une v.a. de loi $\mathcal{N}(m + h\sigma^2, \sigma^2)$.

1.4 Vecteurs Gaussiens

Les variables gaussiennes $\mathcal{N}(m, \sigma^2)$ ont été introduites dans l'exemple 1.18. Leur généralisation sur \mathbb{R}^n conduit à la notion de *vecteur gaussien* :

Définition 1.30 Soit $X = (X_1, X_2, \dots, X_n)$ un vecteur aléatoire de \mathbb{R}^n . On dit que X est un vecteur gaussien si pour tout vecteur déterministe $u = (u_1, \dots, u_n)$, la variable aléatoire

$$(u, X) = \sum_{i=1}^n u_i X_i$$

est une variable gaussienne à valeurs réelles.

Par exemple, si X_1, \dots, X_n sont n Gaussiennes indépendantes, on sait que toute combinaison linéaire des X_i est une Gaussienne (par exemple si $X_1 \sim \mathcal{N}(m_1, \sigma_1^2)$, $i = 1, 2$, et si $X_1 \perp X_2$, alors $u_1 X_1 + u_2 X_2 \sim \mathcal{N}(u_1 m_1 + u_2 m_2, u_1^2 \sigma_1^2 + u_2^2 \sigma_2^2)$). Donc (X_1, \dots, X_n) est un vecteur gaussien.

Mais attention, cette propriété n'est plus vraie en général sans l'hypothèse d'indépendance. Par exemple si $X \sim \mathcal{N}(0, 1)$, si ε est une variable indépendante de X telle que $\mathbb{P}[\varepsilon = 1] = \mathbb{P}[\varepsilon = -1] = 1/2$ et si $Y = \varepsilon X$ alors on voit facilement que $Y \sim \mathcal{N}(0, 1)$. Cependant, (X, Y) n'est pas un vecteur gaussien puisque $X + Y$ n'est pas une Gaussienne : on voit immédiatement que $\mathbb{P}[X + Y = 0] = 1/2$ et ceci est impossible pour une loi gaussienne qui est soit concentrée en un seul point, soit absolument continue (et donc non-atomique).

La loi d'un vecteur gaussien est caractérisée par deux paramètres, comme dans le cas unidimensionnel :

Théorème 1.31 Soit $X = (X_1, \dots, X_n)$ un vecteur gaussien. Il existe un vecteur $\Gamma \in \mathbb{R}^n$ et une matrice $(n \times n)$, symétrique positive σ tels que la transformée de Fourier de X est donnée par

$$\mathbb{E} \left[e^{i(u, X)} \right] = \exp \left[i(u, \Gamma) - \frac{1}{2} {}^t u \sigma u \right].$$

Ce théorème généralise l'exemple 1.18 où l'on avait montré que $\mathbb{E} [e^{iuX}] = e^{(itm - t^2 \sigma^2 / 2)}$, soit $\Gamma_X = m$ et $\sigma_X = \sigma^2$ dans ce cas. On fera attention que m se généralise en un *vecteur* et σ en une *matrice*. L'interprétation de Γ et σ est la suivante : Γ est le vecteur espérance $(\mathbb{E}[X_1], \dots, \mathbb{E}[X_n])$ et σ la matrice de covariance :

$$\sigma_{i,j} = \text{cov}(X_i, X_j) = \mathbb{E}[X_i X_j] - \mathbb{E}[X_i] \mathbb{E}[X_j].$$

Comme pour le cas unidimensionnel, on note $X \sim \mathcal{N}(\Gamma, \sigma)$. Quand $\Gamma = 0$ et $\sigma = \text{Id}$, on dit que X est un vecteur gaussien *centré réduit*. La représentation suivante est très utile :

Proposition 1.32 Soit $X \sim \mathcal{N}(\Gamma, \sigma)$. Alors on peut écrire $X = \Gamma + \sigma Y$ où Y est un vecteur gaussien centré réduit.

On déduit de cette proposition que la loi de X admet une densité si et seulement si la matrice σ est inversible. La densité est alors donnée par

$$f(x_1, \dots, x_n) = \frac{1}{(2\pi)^{n/2} \sqrt{\det \sigma}} \exp \left[-\frac{1}{2} {}^t (x - \Gamma) \sigma^{-1} (x - \Gamma) \right]$$

où l'on a écrit $x = (x_1, \dots, x_n)$. On voit par cette formule que si σ est une matrice diagonale, alors la densité du vecteur gaussien est donnée par une formule produit, et donc les composantes de X sont indépendantes. En particulier, pour un couple gaussien (X, Y) , on a

$$X \perp Y \iff \text{cov}(X, Y) = 0,$$

ce qui se traduit par "orthogonalité et indépendance sont des notions équivalentes". Attention, cette propriété est *entièrement propre au cadre gaussien*.

Exercice 1.33 Soit (X, Y) est un vecteur gaussien bi-dimensionnel. Montrer qu'il existe α tel que $X - \alpha Y$ est indépendant de X .

Exercice 1.34 Soit (X, Y) un vecteur gaussien centré bi-dimensionnel avec $\mathbb{E}[X^2] = \mathbb{E}[Y^2] = 1$ et $\mathbb{E}[XY] = \rho$. On pose $S = X - Y$ et $T = X + Y$.

(a) Montrer que (S, T) est un vecteur gaussien et déterminer ses caractéristiques.

(b) On pose $W = \max(X, Y)$. Exprimer W en fonction de S et T et en déduire la densité de W , son espérance et sa variance. Même question pour $Z = \min(X, Y)$.

Exercice 1.35 Soit X un vecteur gaussien centré réduit sur \mathbb{R}^n . Déterminer la loi (appelée *loi du chi-deux à n degrés de liberté*) de $Z_n = X_1^2 + \dots + X_n^2$.

1.5 Convergence de variables aléatoires

On considère, sur un espace de probabilités fixé $(\Omega, \mathcal{F}, \mathbb{P})$, une suite de variables aléatoires $\{X_n, n \geq 0\}$ et l'on s'intéresse à la convergence de cette suite. Le caractère aléatoire de la suite met en évidence plusieurs types de convergence :

Convergence presque sûre. Une suite de variables aléatoires X_n converge p.s. vers X si

$$X_n(\omega) \rightarrow X(\omega) \quad \text{quand } n \rightarrow \infty,$$

pour \mathbb{P} presque tout ω . Cette notion de convergence dépend du choix de \mathbb{P} mais évidemment, si $\mathbb{Q} \sim \mathbb{P}$, alors $X_n \rightarrow X$ \mathbb{P} p.s. $\Leftrightarrow X_n \rightarrow X$ \mathbb{Q} p.s. Rappelons trois théorèmes fondamentaux :

Théorème 1.36 [Théorème de convergence monotone] Soit $\{X_n, n \geq 0\}$ une suite croissante (ou décroissante) de variables aléatoires et soit $X = \lim X_n$ p.s. Supposons que X soit intégrable. Alors $\mathbb{E}[X] = \lim \mathbb{E}[X_n]$.

Théorème 1.37 [Théorème de convergence dominée] Soit $\{X_n, n \geq 0\}$ une suite de variables aléatoires convergeant p.s. vers X . Supposons qu'il existe une variable aléatoire Y intégrable telle que $|X_n| \leq Y$, alors X est intégrable et $\mathbb{E}[|X_n - X|] \rightarrow 0$ quand $n \uparrow +\infty$.

Théorème 1.38 [Loi des grands nombres] Soit $\{X_n, n \geq 0\}$ une suite de variables aléatoires équidistribuées, indépendantes, et intégrables. Alors

$$\frac{1}{n} \sum_{i=1}^n X_i \rightarrow \mathbb{E}[X_1] \quad \text{p.s.}$$

Convergence en probabilité. Une suite $\{X_n, n \geq 0\}$ de variables aléatoires converge en probabilité vers X si

$$\mathbb{P}[|X_n - X| \geq \varepsilon] \rightarrow 0$$

quand $n \rightarrow +\infty$, pour tout $\varepsilon > 0$. Il est immédiat que la convergence presque sûre implique la convergence en probabilité. Réciproquement, on peut montrer avec le lemme de Borel-Cantelli que la convergence en probabilité de X_n vers X implique qu'une *sous-suite* des X_n converge vers X p.s.

Convergence dans les L^p . Une suite $\{X_n, n \geq 0\}$ de variables aléatoires dans L^p converge vers X dans L^p si

$$\|X_n - X\|_p = \mathbb{E}[|X_n - X|^p]^{1/p} \rightarrow 0$$

quand $n \rightarrow \infty$. Remarquons que si $X_n \rightarrow X$ dans L^p , alors $\mathbb{E}[X_n^p] \rightarrow \mathbb{E}[X^p]$, mais que la réciproque est fautive. Si une suite converge dans L^p , alors elle converge en probabilité et en particulier il existe une sous-suite qui converge p.s. Dans le cas $p = 1$, la réciproque est donnée par le théorème de convergence dominée. Remarquons enfin que la convergence dans L^p implique la convergence dans L^q pour tout $q \leq p$.

Convergence en loi. C'est la notion la plus faible, et aussi la plus délicate. Une suite de variables aléatoires X_n converge en loi vers X si

$$\mathbb{E}[\Phi(X_n)] \rightarrow \mathbb{E}[\Phi(X)]$$

quand $n \rightarrow \infty$, pour toute fonction Φ continue bornée. On note cette convergence $X_n \xrightarrow{d} X$. La convergence en loi est également définie par la convergence simple des fonctions caractéristiques, au sens où $\Psi_{X_n}(t) \rightarrow \Psi_X(t)$ pour tout t . Enfin, si X est une v.a. de fonction de répartition F continue, et si X_n est une suite de v.a. de fonctions de répartition F_n telles que $F_n(x)$ converge vers $F(x)$ pour tout x , alors X_n converge en loi vers X et réciproquement. La convergence en probabilité (et donc aussi la convergence p.s. et la convergence dans les L^p) implique la convergence en loi. Le théorème suivant, qui affine la loi des grands nombres, est l'un des résultats fondamentaux du calcul des probabilités, qui souligne le rôle central de la loi gaussienne :

Théorème 1.39 [Théorème Central Limite] *Soit $\{X_n, n \geq 0\}$ une suite de v.a. équidistribuées, indépendantes, et de variance finie σ^2 . Alors*

$$\frac{\sum_{i=1}^n X_i - n\mathbb{E}[X_1]}{\sigma\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, 1).$$

quand $n \uparrow +\infty$.

Chapitre 2

Martingales et temps d'arrêt

2.1 Généralités sur les processus stochastiques

On commence par introduire une notion mathématique qui joue un rôle central en théorie des marchés financiers :

Définition 2.1 Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité et $\{\mathcal{F}_t, t \geq 0\}$ une famille de sous-tribus de \mathcal{F} . On dit que $\{\mathcal{F}_t, t \geq 0\}$ est une filtration si c'est une famille croissante, au sens où $\mathcal{F}_s \subset \mathcal{F}_t$ si $s \leq t$.

Il faut comprendre \mathcal{F}_t comme "l'information au temps t " : plus le temps croît ($s \leq t$), plus on a d'informations ($\mathcal{F}_s \subset \mathcal{F}_t$). Une filtration $\{\mathcal{G}_t, t \geq 0\}$ est dite plus grosse que $\{\mathcal{F}_t, t \geq 0\}$ si $\mathcal{F}_t \subset \mathcal{G}_t$ pour tout $t \geq 0$. Une filtration est dite *normale* si elle vérifie deux propriétés supplémentaires :

- Les négligeables au sens large sont dans tous les \mathcal{F}_t au sens où $\mathbb{P}[A] = 0 \Rightarrow A \in \mathcal{F}_0$.
- La filtration est continue à droite au sens où

$$\mathcal{F}_t^X = \mathcal{F}_{t+}^X \stackrel{\text{def}}{=} \bigcap_{s>t} \mathcal{F}_s^X.$$

Si $\{\mathcal{F}_t, t \geq 0\}$ est une filtration sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, on dit que $(\Omega, \mathcal{F}, \{\mathcal{F}_t, t \geq 0\}, \mathbb{P})$ est un *espace de probabilité filtré*.

Définition 2.2 Soit $(\Omega, \mathcal{F}, \{\mathcal{F}_t, t \geq 0\}, \mathbb{P})$ est un espace de probabilité filtré. Une famille $X = \{X_t, t \geq 0\}$ de v.a. sur Ω est appelée un processus stochastique. On dit que le processus X est (\mathcal{F}_t) -adapté si X_t est \mathcal{F}_t -mesurable pour tout $t \geq 0$.

A un processus stochastique X on peut associer sa filtration naturelle complétée $\{\mathcal{F}_t^X, t \geq 0\}$, définie par $\mathcal{F}_t^X = \sigma(\sigma(X_s, s \leq t), \mathcal{N})$ où l'on rappelle que \mathcal{N} désigne la tribu des négligeables pour \mathbb{P} . La tribu \mathcal{F}_t^X représente l'information portée par $X_s, s \leq t$, et les négligeables. Evidemment, X est un processus (\mathcal{F}_t^X) -adapté. On peut montrer que si X a p.s. des trajectoires continues à droite, pourvues de limites à gauche (càdlàg), en particulier continues, alors la filtration $\{\mathcal{F}_t^X, t \geq 0\}$ est continue à droite.

Définition 2.3 On dit que deux processus X et Y sont égaux à modification près si $\mathbb{P}[X_t = Y_t] = 1$ pour tout t .

Remarquons que si deux processus continus à droite sont égaux à modification près, alors

$$\mathbb{P}[X_t = Y_t, \forall t \geq 0] = 1.$$

En effet, comme \mathbb{Q} est dénombrable et que l'intersection d'une famille dénombrable d'événements certains est un événement certain, on voit facilement que

$$\mathbb{P}[X_t = Y_t, \forall t \in \mathbb{Q}] = 1$$

si X et Y sont égaux à modification près. De plus, comme \mathbb{Q} est dense dans \mathbb{R} et X et Y sont continus à droite, les trajectoires de X et Y sont uniquement déterminées par leurs valeurs sur \mathbb{Q} puisque

$$X_t = \lim_{t_n \in \mathbb{Q} \downarrow t} X_{t_n}$$

pour tout $t \geq 0$. On a donc $\mathbb{P}[X_t = Y_t, \forall t \geq 0] = 1$.

Définition 2.4 On dit que deux processus X et Y sont égaux en loi et on écrit $X \stackrel{d}{=} Y$ si pour tout $n \in \mathbb{N}$ et pour tout (t_1, t_2, \dots, t_n) on a

$$\{X_{t_1}, X_{t_2}, \dots, X_{t_n}\} \stackrel{d}{=} \{Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}\}.$$

On dit aussi que X et Y sont des versions l'un de l'autre.

On remarque immédiatement que si X et Y sont continus à droite et modification l'un de l'autre, alors ils sont égaux en loi. Attention, la réciproque est fautive : en effet, si X est symétrique au sens où $X \stackrel{d}{=} -X$, alors X et $-X$ sont des versions l'un de l'autre mais évidemment $\mathbb{P}[X_t = -X_t, \forall t \geq 0] = 1$ si et seulement si $X \equiv 0$.

Propriétés fonctionnelles. (a) Un processus X est croissant si p.s. $t \rightarrow X_t$ est une fonction croissante, c'est-à-dire si $X_t(\omega) \leq X_s(\omega)$ pour tout $t \leq s$, p.s. De même, on dira qu'un processus est continu à droite, continu, dérivable, de classe C^2 ... etc si la fonction $t \mapsto X_t(\omega)$ vérifie p.s. la propriété considérée.

(b) Un processus X est dit à *variation finie* (V.F.) sur $[0, t]$ si

$$\sup_{0 \leq t_0 < \dots < t_n \leq t} \left(\sum_{k=0}^n |X_{t_{k+1}} - X_{t_k}| \right) < +\infty.$$

Cette dernière notion, assez délicate, joue un rôle fondamental en théorie de l'intégration. Elle signifie que la longueur de la courbe trajectoire de X est localement finie (quand cette longueur est infinie - ce qui sera le cas pour la courbe brownienne par exemple, on mesure la complexité de la courbe à l'aide des notions de dimension fractale et de mesure de Hausdorff). En particulier, X est V.F. s'il est dérivable. Dans le cas réel, on peut montrer que X est V.F. si et seulement si X est la différence de deux processus croissants.

Processus Gaussiens. Un processus X est appelé un *processus gaussien* si toute combinaison linéaire finie de X_t est une variable aléatoire gaussienne, autrement dit si pour tout $n \in \mathbb{N}$, $t_1 \dots t_n$, $a_1 \dots a_n$,

$$\sum_{i=1}^n a_i X_{t_i}$$

est une variable gaussienne. Un processus gaussien est caractérisé par deux fonctions : sa fonction espérance $t \mapsto m(t) = \mathbb{E}[X_t]$ et sa fonction de covariance $(s, t) \mapsto \Gamma(s, t) = \mathbb{E}[(X_t - m(t))(X_s - m(s))]$. La fonction $\Gamma(s, t)$ est de *type positif* au sens où pour tout $n \in \mathbb{N}$, $t_1 \dots t_n$, $a_1 \dots a_n$

$$\sum_{i,j=1}^n a_i a_j \Gamma(t_i, t_j) \geq 0.$$

Théorème 2.5 [Kolmogorov] Soit $m : \mathbb{R}^+ \rightarrow \mathbb{R}$ une fonction quelconque et $\Gamma : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}$ une fonction symétrique de type positif. Alors il existe un processus gaussien X de fonction espérance m et de fonction de covariance Γ .

L'espace *gaussien* engendré par un processus gaussien X est le sous espace de $L^2(\Omega)$ engendré par les v.a.r. centrées $\{X_t - \mathbb{E}[X_t], t \geq 0\}$, c'est-à-dire le sous-espace formé par les combinaisons linéaires de ces variables centrées et leurs limites en moyenne quadratique.

Définition 2.6 Un processus est dit *stationnaire* si $\{X_{t+s}, t \geq 0\} \stackrel{d}{=} \{X_t, t \geq 0\}$ pour tout $s \in \mathbb{R}^+$. En particulier, tous les X_t ont même loi.

On voit immédiatement qu'un processus gaussien stationnaire est tel que sa fonction espérance est *constante* et sa fonction de covariance $\Gamma(s, t)$ ne dépend que de $(s - t)$.

Théorème 2.7 [Bochner] Soit X un processus gaussien stationnaire. Il existe une unique mesure finie μ , appelée *mesure spectrale* associée à X , telle que

$$\Gamma(s, t) = \int_{\mathbb{R}} \cos((t - s)z) \mu(dz).$$

Exercice 2.8 Montrer que la fonction $(s, t) \mapsto \inf(s, t)$ est une fonction de type positif sur $\mathbb{R}^+ \times \mathbb{R}^+$, autrement dit que pour tout $t_1 \leq \dots \leq t_n$ et pour tous réels $a_1 \dots a_n$

$$\sum_{i,j=1}^n \inf(t_i, t_j) a_i a_j \geq 0.$$

Le processus gaussien centré ayant pour fonction de covariance $(s, t) \mapsto \inf(s, t)$ est appelé *mouvement brownien*.

2.2 Temps d'arrêt

Définition 2.9 Soit $(\Omega, \mathcal{F}, \{\mathcal{F}_t, t \geq 0\}, \mathbb{P})$ un espace de probabilité filtré. Un temps d'arrêt relativement à (\mathcal{F}_t) est une variable aléatoire τ à valeur dans $\mathbb{R}^+ \cup \{+\infty\}$ telle que $\{\tau \leq t\} \in \mathcal{F}_t$ pour tout $t \in \mathbb{R}^+$.

La définition est équivalente à $\{\tau > t\} \in \mathcal{F}_t$ pour tout $t \in \mathbb{R}^+$, par stabilité des tribus par passage au complémentaire. Un temps d'arrêt doit être vu comme un temps aléatoire qui suit l'évolution aléatoire sous-jacente. La raison pour laquelle l'événement $\{\tau \leq t\}$ et non l'événement $\{\tau = t\}$ intervient dans la définition du temps d'arrêt vient du temps discret : relativement à une filtration $\{\mathcal{F}_n, n \geq 0\}$, un temps d'arrêt est plus simplement défini par $\{\tau = n\} \in \mathcal{F}_n$ pour tout $n \in \mathbb{N}$ - voir [6] [24]. On voit facilement que ceci est équivalent à $\{\tau \leq n\} \in \mathcal{F}_n$ pour tout $n \in \mathbb{N}$. Typiquement, l'événement $\{\tau = t\}$ est de probabilité nulle en temps continu. C'est pourquoi on étend la définition en utilisant les événements $\{\tau \leq t\}$.

On voit facilement que si S et T sont des temps d'arrêt, alors $S \wedge T$ est un temps d'arrêt. L'exemple le plus simple de temps d'arrêt est un temps déterministe $t_0 \geq 0$: en particulier, si T est un temps d'arrêt, alors $T \wedge n$ est un temps d'arrêt pour tout $n \in \mathbb{N}$, propriété très utile dans la suite. Un exemple plus élaboré et extrêmement important de temps d'arrêt est le suivant :

Exemple 2.10 Soit $\{X_t, t \geq 0\}$ un processus (\mathcal{F}_t) -adapté à valeurs réelles, continu à droite et partant de 0 ($X_0 = 0$). Pour tout $a > 0$, le temps aléatoire suivant

$$T_a = \inf \{t > 0, X_t > a\},$$

appelé *premier temps de passage au seuil a* , est un (\mathcal{F}_t) -temps d'arrêt. En effet, pour tout $t > 0$, on a l'identification

$$\{T > t\} = \{X_s \leq a, \forall s \leq t\} = \cap_{s \in \mathbb{Q}^+, s \leq t} \{X_s \leq a\},$$

la deuxième égalité provenant de l'hypothèse cruciale de continuité à droite de X (qui est alors entièrement déterminé par ses évaluations sur les rationnels). Comme X est (\mathcal{F}_t) -adapté et par stabilité des tribus par intersection dénombrable, on voit que l'événement de droite est dans \mathcal{F}_t . Remarquons que si X est continu, on a de plus $T_a = \inf \{t > 0, X_t = a\}$ et $X_{T_a(\omega)}(\omega) = a$. On montre de même que $T_a = \inf \{t > 0, X_t < a\}$ est un temps d'arrêt pour tout $a < 0$. D'une manière générale, on peut montrer que si X est un processus càd et adapté à valeurs dans \mathbb{R}^d et si A est un Borélien de \mathbb{R}^d , alors $T_A = \inf \{t > 0, X_t \in A\}$ (premier temps d'atteinte de A) est un temps d'arrêt. Cette démonstration repose cependant sur des outils mathématiques de haute distinction - voir par exemple la bibliographie de [1]. On prendra garde qu'en général, le dernier temps de passage en A : $L_A = \sup \{t > 0, X_t \in A\}$ n'est *pas* un temps d'arrêt, même dans les situations les plus simples.

Définition 2.11 Soit $(\Omega, \mathcal{F}, \{\mathcal{F}_t, t \geq 0\}, \mathbb{P})$ un espace de probabilité filtré et τ un temps d'arrêt relativement à (\mathcal{F}_t) . On appelle filtration arrêtée en τ la σ -algèbre

$$\mathcal{F}_\tau = \{A \in \mathcal{F} \mid A \cap \{\tau \leq t\} \in \mathcal{F}_t \forall t \in \mathbb{R}^+\}.$$

Pour bien comprendre l'intérêt de cette notion, remarquons que τ est toujours \mathcal{F}_τ -mesurable : les Boréliens de \mathbb{R}^+ étant engendrés par les intervalles du type $]-\infty, u]$, il suffit de montrer que $\{\tau \leq u\} \in \mathcal{F}_\tau$. Mais c'est évident car $\{\tau \leq u\} \cap \{\tau \leq t\} = \cap \{\tau \leq t \wedge u\} \in \mathcal{F}_{t \wedge u} \subset \mathcal{F}_t$. De même, on montre la

Proposition 2.12 Si S et T sont des temps d'arrêt tels que p.s. $S \leq T$, alors $\mathcal{F}_S \subset \mathcal{F}_T$.

Preuve : Soit $A \in \mathcal{F}_S$. Montrons que $A \in \mathcal{F}_T$. Pour tout $t \geq 0$ on a $A \cap \{T \leq t\} = A \cap \{S \leq t\} \cap \{T \leq t\}$ puisque $S \leq T$. D'une part $\{T \leq t\} \in \mathcal{F}_t$ puisque T est un temps d'arrêt, d'autre part $A \cap \{S \leq t\} \in \mathcal{F}_t$ puisque $A \in \mathcal{F}_S$. Donc $A \cap \{S \leq t\} \cap \{T \leq t\} \in \mathcal{F}_t$ pour tout $t \geq 0$, ce qui est la propriété recherchée. \square

Une propriété importante, mais plus difficile à démontrer - voir [17] ou [24], est la suivante : soit $X : \mathbb{R}^+ \times \Omega \mapsto \mathbb{R}$ un processus et T un temps d'arrêt. Rappelons que la v.a. X_T est définie par $X_T(\omega) = X(T(\omega), \omega)$. On a la

Proposition 2.13 *Si le processus X est càd et adapté, alors la v.a. $X_T \mathbb{1}_{\{T < +\infty\}}$ est \mathcal{F}_T -mesurable.*

2.3 Martingales

Pour de plus amples informations sur ce sujet important, la lecture du livre de Williams [24] est vivement conseillée.

Définition 2.14 *Soit $(\Omega, \mathcal{F}, \{\mathcal{F}_t, t \geq 0\}, \mathbb{P})$ un espace de probabilité filtré et $X = \{X_t, t \geq 0\}$ un processus (\mathcal{F}_t) -adapté. On dit que X est une (\mathcal{F}_t) -martingale si*

- $\mathbb{E}[|X_t|] < +\infty$ (autrement dit $X_t \in L^1(\Omega)$) pour tout $t \geq 0$.
- $\mathbb{E}[X_t | \mathcal{F}_s] = X_s$ pour tout $s \leq t$.

On peut en fait définir X sur $(\Omega, \mathcal{F}, \mathbb{P})$ sans filtration préalable, en disant que c'est une martingale si $\mathbb{E}[|X_t|] < +\infty$ pour tout $t \geq 0$ et $\mathbb{E}[X_t | \mathcal{F}_s^X] = X_s$ pour tout $s \leq t$, où l'on rappelle que $\{\mathcal{F}_t^X, t \geq 0\}$ est la filtration naturelle de X . Si X est une (\mathcal{F}_t) -martingale, alors c'est nécessairement une (\mathcal{F}_t^X) -martingale. En effet, on a d'abord $\mathcal{F}_u^X = \sigma\{X_s^{-1}(A), A \in \mathcal{B}(\mathbb{R}), s \leq u\} \subset \mathcal{F}_u$, puisque X est (\mathcal{F}_t) -adapté et donc $X_s^{-1}(A) \in \mathcal{F}_s \subset \mathcal{F}_u$ pour tout $s \leq u$. D'où

$$\mathbb{E}[X_t | \mathcal{F}_s^X] = \mathbb{E}[\mathbb{E}[X_t | \mathcal{F}_s] | \mathcal{F}_s^X] = \mathbb{E}[X_s | \mathcal{F}_s^X] = X_s,$$

où l'on a successivement utilisé la propriété d'emboîtement de l'espérance conditionnelle, que X est une (\mathcal{F}_t) -martingale, et la mesurabilité de X_s par rapport à \mathcal{F}_s^X .

La propriété fondamentale (et constamment utilisée en finance) d'une martingale est que pour tout horizon $T > 0$, l'ensemble du processus $\{X_t, t \leq T\}$ est complètement déterminé par la *valeur terminale* X_T au sens où $X_t = \mathbb{E}[X_T | \mathcal{F}_t]$ pour tout $t \leq T$.

Définition 2.15 *Soit $\{X_t, t \geq 0\}$ une (\mathcal{F}_t) -martingale. On dit que c'est une martingale fermée si il existe $Z \in L^1(\Omega)$ tel que*

$$X_t = \mathbb{E}[Z | \mathcal{F}_t]$$

pour tout $t \geq 0$. Autrement dit, l'ensemble du processus est déterminé par une valeur terminale à l'horizon infini.

On verra plus tard que Z est en fait la limite de X_t quand $t \uparrow +\infty$. Le théorème suivant caractérise les martingales fermées par leur uniforme intégrabilité :

Théorème 2.16 *Soit $\{X_t, t \geq 0\}$ une \mathcal{F}_t -martingale. C'est une martingale fermée si et seulement si la famille de v.a. $\{X_t, t \geq 0\}$ est U.I.*

On remarque que dans ce théorème, l'inclusion $(\{X_t, t \geq 0\} \text{ fermée} \Rightarrow \{X_t, t \geq 0\} \text{ U.I.})$ est une conséquence immédiate de l'exercice 1.26 (b). L'inclusion réciproque est plus délicate, voir [24].

Définition 2.17 *Soit $(\Omega, \mathcal{F}, \{\mathcal{F}_t, t \geq 0\}, \mathbb{P})$ un espace de probabilité filtré et $X = \{X_t, t \geq 0\}$ un processus (\mathcal{F}_t) -adapté tel que $\mathbb{E}[|X_t|] < +\infty$ pour tout $t \geq 0$. On dit que X est une (\mathcal{F}_t) -sousmartingale (resp. (\mathcal{F}_t) -surmartingale) si $\mathbb{E}[X_t | \mathcal{F}_s] \geq X_s$ (resp. $\mathbb{E}[X_t | \mathcal{F}_s] \leq X_s$) pour tout $s \leq t$.*

Par analogie, on dira qu'une (\mathcal{F}_t) -sousmartingale (resp. une (\mathcal{F}_t) -surmartingale) est fermée s'il existe une v.a. Z \mathcal{F} -mesurable telle que $\mathbb{E}[Z | \mathcal{F}_t] \geq X_t$ (resp. $\mathbb{E}[Z | \mathcal{F}_t] \leq X_t$) pour tout $t \geq 0$. Il découle facilement des définitions qu'une martingale est un processus à espérance constante : $\mathbb{E}[X_t] = \mathbb{E}[X_s]$ ne dépend pas de t . Une martingale modélise ainsi un jeu *équitable*. Une sous-martingale est un processus à espérance croissante (jeu favorable), une surmartingale un processus à espérance décroissante (jeu défavorable). La manière la plus simple de construire une sous-martingale et d'ajouter à une martingale donnée une fonction croissante déterministe. Le théorème suivant, très célèbre, dit que la réciproque est presque vraie (la fonction croissante est remplacée par un processus aléatoire croissant) :

Théorème 2.18 [Décomposition de Doob-Meyer] Soit X une sous-martingale relativement à \mathcal{F}_t telle que la famille $\{X_T, T(\mathcal{F}_t)\text{-temps d'arrêt borné}\}$ est UI. Il existe une (\mathcal{F}_t) -martingale M et un processus croissant (\mathcal{F}_t) -adapté A tel que

$$X_t = M_t + A_t$$

pour tout $t \geq 0$. De plus, M et A sont uniques à constante additive près.

Par l'inégalité de Jensen, on voit immédiatement que si X est une (\mathcal{F}_t) -martingale, alors le processus $t \mapsto X_t^2$ est une (\mathcal{F}_t) -sous-martingale. Sous réserve d'UI, la décomposition de Doob-Meyer assure l'existence d'un unique processus croissant A tel que $t \mapsto X_t^2 - A_t$ soit une (\mathcal{F}_t) -martingale (on peut en fait se passer de la condition d'UI). On appelle A le *crochet* de la martingale X et on utilise la notation

$$A_t = \langle X \rangle_t$$

pour tout $t \geq 0$. Nous reviendrons sur le crochet des martingales au chapitre 4, où il jouera un rôle très important en calcul stochastique.

Si T est un temps d'arrêt et M une (\mathcal{F}_t) -martingale, le processus Z défini par $Z_t = M_{t \wedge T}$ est une (\mathcal{F}_t) martingale et $\mathbb{E}[M_{t \wedge T}] = \mathbb{E}[M_0]$.

On énonce maintenant, sans les démontrer, une série de résultats fondamentaux de la théorie des martingales, essentiellement dûs à J. L. Doob dans les années cinquante.

Théorème 2.19 [Inégalité maximale] Soit X une surmartingale réelle continue. Alors pour tout $t, \lambda > 0$,

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

L'intérêt de cette inégalité est de donner une vitesse de convergence (en $1/\lambda$) de la quantité $\mathbb{P}[\sup_{s \leq t} |X_s| \geq \lambda]$ vers 0 quand $\lambda \uparrow +\infty$. En fait, cette vitesse peut être améliorée quand X est une martingale ayant des moments d'ordre supérieur grâce au

Théorème 2.20 [Inégalités dans L^p] Soit $p \geq 1$ et X une martingale réelle continue telle que $X_t \in L^p$ pour tout $t \geq 0$. Alors pour tout $t \geq 0$

$$\mathbb{E}[\sup_{s \leq t} |X_s|^p] \leq q^p \mathbb{E}[|X_t|^p]$$

où q est le nombre conjugué de p : $1/p + 1/q = 1$.

Par l'inégalité de Markov, on déduit de ce théorème que si $X_t \in L^p$ pour tout $t \geq 0$, alors

$$\mathbb{P}[\sup_{s \leq t} |X_s| \geq \lambda] = \mathbb{P}[\sup_{s \leq t} |X_s|^p \geq \lambda^p] \leq \frac{q^p}{\lambda^p} \mathbb{E}[|X_t|^p].$$

D'où une meilleure vitesse de convergence, en $1/\lambda^p$.

Remarque 2.21 Pour appliquer ce théorème il ne suffit pas que $X_t \in L^p$ pour un seul $t \geq 0$, et la condition $X_t \in L^p$ pour tout $t \geq 0$ est nécessaire. En effet, il existe des martingales telles que $X_1 \in L^p$ et $X_2 \notin L^p$ pour un $p > 1$: prenons par exemple X une v.a. qui soit dans L^1 mais pas dans L^2 , et considérons une filtration $\{\mathcal{F}_t, t \geq 0\}$ qui soit triviale pour $0 \leq t \leq 1$ et égale à $\sigma(X)$ pour $t \geq 2$. La martingale $t \mapsto X_t = \mathbb{E}[X | \mathcal{F}_t]$ est telle que $X_1 = \text{cste} \in L^2$ et $X_2 \notin L^2$.

Théorème 2.22 [Théorème de convergence des martingales] Soit X une martingale continue. Les assertions suivantes sont équivalentes :

- (a) X est une martingale fermée par X_∞ .
- (b) X converge p.s. et dans L^1 vers X_∞ .
- (c) X est uniformément intégrable.

On prendra garde, dans ce théorème, que la limite de X_t est une *variable aléatoire* X_∞ . De plus, il existe des martingales qui ne sont pas uniformément intégrables, autrement dit qui ne convergent pas. Un exemple typique de martingale non convergente est le mouvement brownien, que nous verrons au chapitre suivant. Le théorème suivant est le plus important de cette série pour la finance :

Théorème 2.23 [Théorème d'arrêt] *Soit M une (\mathcal{F}_t) -martingale continue et S, T deux temps d'arrêt tels que $S \leq T$ p.s. Supposons que M soit uniformément intégrable ou que les temps d'arrêt soient bornés par une constante finie déterministe K . Alors M_T est une v. a. intégrable et*

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

Quand les deux temps d'arrêt ne sont pas bornés, l'uniforme intégrabilité de la martingale est une condition essentielle, comme nous le verrons au chapitre suivant avec le mouvement brownien. Quand la martingale n'est pas UI, une technique standard consiste à utiliser les temps d'arrêt $S \wedge n$ et $T \wedge n$ qui sont bornés par n , et faire tendre n vers $+\infty$. Quand la martingale est UI, nous avons vu qu'elle était fermée par M_∞ . Le théorème d'arrêt reste alors valable avec des valeurs infinies pour T , au sens où

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

La caractérisation des martingales donnée par la proposition suivante est quelquefois utile :

Proposition 2.24 *Soit X un processus (\mathcal{F}_t) -adapté et intégrable tel que $\mathbb{E}[X_\tau] = \mathbb{E}[X_0]$ pour tout temps d'arrêt borné τ . Alors le processus X est une martingale.*

On prendra garde que la condition $\mathbb{E}[X_t] = \mathbb{E}[X_0]$ pour tout $t \geq 0$ n'est pas suffisante pour que X soit une martingale. Un contre-exemple typique est le processus $X_t = \int_0^t M_u du$, où M est une martingale d'espérance nulle : X est un processus d'espérance constante (nulle), mais n'est pas une martingale. La proposition suivante montre enfin que les trajectoires d'une martingale continue non constante sont très irrégulières :

Proposition 2.25 *Si une martingale M continue est un processus à variations finies, alors elle est constante : $M_t = M_0$ p.s. pour tout $t \geq 0$.*

Chapitre 3

Le mouvement brownien

3.1 Un peu d'histoire

Avant d'être un objet mathématique rigoureux, le mouvement brownien a été étudié en Botanique, en Finance, et en Physique. Le botaniste R. Brown observe d'abord vers 1828 le mouvement irrégulier de particules de pollen en suspension dans l'eau. En 1877, Delsaux explique les changements incessants de direction de trajectoire par les chocs entre les particules de pollen et les molécules d'eau. Un mouvement de ce type est alors appelé *mouvement au hasard*. En 1900, L. Bachelier, en vue d'étudier les cours de la Bourse de Paris dans sa thèse, met en évidence le caractère markovien du mouvement brownien : la position d'une particule à l'instant $t+s$ dépend de sa position en t , et ne dépend pas de sa position avant t . Peu après, vers 1905, A. Einstein détermine la densité de transition du Brownien par l'intermédiaire de l'équation de la chaleur. La même année, Smoluchowski décrit le mouvement brownien comme une limite de promenades aléatoires. La première étude mathématique rigoureuse du Brownien est faite par

N. Wiener (1923), qui construit une mesure de probabilités sur l'espace des fonctions continues sous laquelle le processus canonique est un mouvement Brownien. Des recherches d'une influence considérable ont ensuite été menées par P. Lévy (1948), lequel s'est intéressé aux propriétés fines des trajectoires du Brownien. Ces objets ont été développés par les potentialistes américains à la suite de J. L. Doob, puis systématisés par les spécialistes de la "Théorie Générale des Processus" de l'école de Strasbourg, autour de P.-A. Meyer.

3.2 Définitions, constructions, et premières propriétés

Définition 3.1 Soit $X = \{X_t, t \geq 0\}$ un processus issu de 0 et à valeurs dans \mathbb{R} . On dit que X est un mouvement brownien s'il vérifie l'une des deux conditions suivantes (équivalentes) :

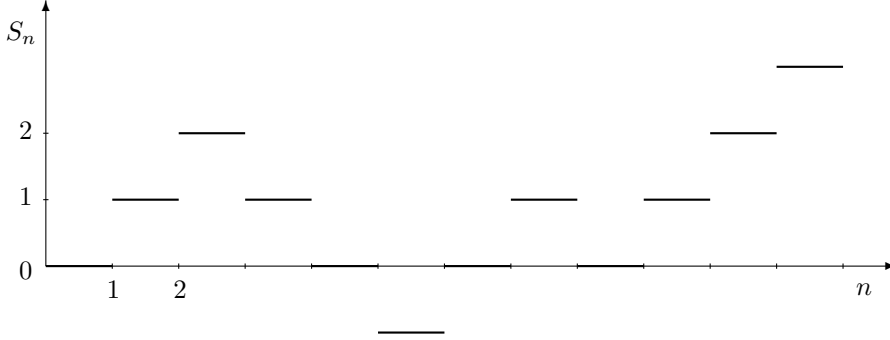
(a) X est un processus gaussien centré et de fonction de covariance $\mathbb{E}[X_s X_t] = s \wedge t$.

(b) X est un processus à accroissements indépendants et stationnaires tel que $X_t \sim \mathcal{N}(0, t)$ pour tout $t \geq 0$.

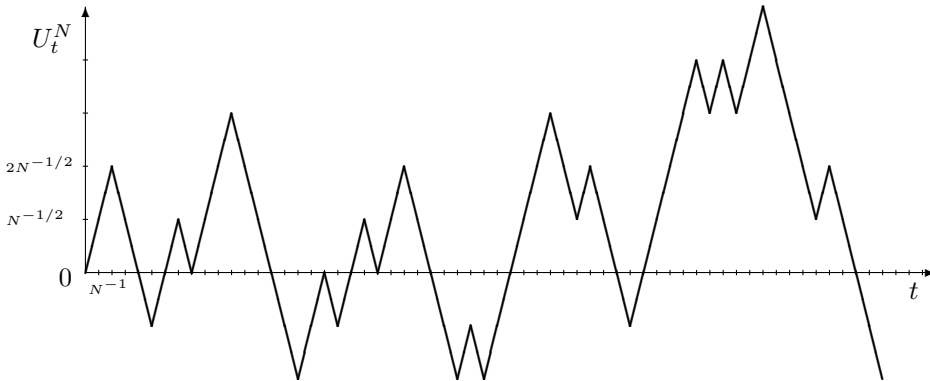
On montre que le MB est à trajectoires continues. Précisons que la propriété d'indépendance des accroissements signifie que pour tout $s \leq t$, la variable $X_t - X_s$ est indépendante de $\sigma\{X_u, u \leq s\}$, tribu du passé avant s . La propriété de stationnarité des accroissements signifie simplement que la loi de $X_t - X_s$ ne dépend que de $t - s$: pour $t \geq s$, $X_t - X_s \stackrel{d}{=} X_{t-s}$. Les processus à accroissements indépendants et stationnaires (P.A.I.S) portent aussi le nom de *processus de Lévy*. Leur structure a été caractérisée dans les années trente successivement par De Finetti, Khintchine et Lévy. Nous reviendrons plus loin sur ces processus très importants qu'utilisent de plus en plus les financiers mathématiciens. Contentons-nous pour l'instant de renvoyer aux deux monographies de référence [1], [21] ou encore [20] pour de plus amples informations.

Comme nous le disions plus haut, la preuve rigoureuse de l'existence du mouvement brownien repose sur la construction d'une mesure de probabilités sur l'espace des trajectoires, la *mesure de Wiener*. Plus précisément, on prend comme espace de probabilités sous-jacent l'espace $\Omega = \mathcal{C}(\mathbb{R}^+, \mathbb{R})$ des fonctions continues de \mathbb{R}^+ dans \mathbb{R} , on note $X_t(\omega) = \omega(t)$ pour tout $t \geq 0$ le processus canonique (ici, $\omega \in \Omega$ est une fonction), $\mathcal{F}_t^X = \sigma\{X_s, s \leq t\}$ la filtration canonique et $\mathcal{F} = \mathcal{F}_\infty$. La mesure de Wiener est

une probabilité sur $(\Omega, \mathcal{F}, \{\mathcal{F}_t^X\})$ sous laquelle X est un mouvement brownien au sens de la définition précédente. Nous renvoyons par exemple à [11] pour la construction de cette mesure. Signalons au passage que la construction de Wiener a donné naissance à une axiomatique plus générale, l'*axiomatique de Kolmogorov* (1933), laquelle permet de construire une classe beaucoup plus large de processus. Dans la suite, le mouvement brownien sera souvent noté B (comme Brownien) ou W (comme Wiener). Quand il sera nécessaire de considérer la mesure de probabilités sur l'espace des trajectoires, nous utiliserons plutôt la notation X du processus canonique. On peut aussi construire plus simplement le Brownien comme limite de promenades aléatoires renormalisées :



Cette propriété est notamment exploitée pour les simulations. Soit X une v.a. de Bernoulli, i. e. $\mathbb{P}[X = 1] = \mathbb{P}[X = -1] = 1/2$ et soit $\{X_n, n \geq 0\}$ une suite de v.a. indépendantes et équidistribuées de même loi que X . On considère $S_n = X_1 + \dots + X_n$ la *marche aléatoire symétrique simple*. On voit facilement que $\mathbb{E}[S_n] = 0$ et $\text{Var}[S_n] = n$. On procède ensuite à une double renormalisation en temps et en espace en fixant N et en considérant $U_{k/N} = S_k/\sqrt{N}$ pour $k = 0, \dots, N$. Remarquons que $U_{k/N}$ est indexé par $\{0, 1/N, \dots, 1\}$, que $\mathbb{E}[U_{k/N}] = 0$ et que $\text{Var}[U_{k/N}] = k/N$. On définit ensuite un processus à temps continu $\{U_t^N, t \in [0, 1]\}$ à partir de $U_{k/N}$ en imposant à la fonction $t \rightarrow U_t^N$ d'être affine entre k/N et $k + 1/N$:



Pour $t = 1$ on a

$$U_1^N = \frac{S_N - N\mathbb{E}[X_1]}{\sqrt{N\text{Var}[X_1]}},$$

de sorte que par le Théorème Central Limite, $U_1^N \rightarrow \mathcal{N}(0, 1)$ en loi. De même, on voit que $U_t^N \rightarrow \mathcal{N}(0, t)$ en loi pour tout $t \in [0, 1]$. On peut alors montrer alors que *toute* la trajectoire du processus U^N converge en loi vers celle du mouvement brownien B . Cette convergence en loi de processus porte aussi le nom de *principe d'invariance de Donsker*, et utilise de façon cruciale la notion de tension d'une famille de lois de probabilités. Nous renvoyons par exemple à [10] pour plus de détails sur ces deux notions.

Etablissons maintenant l'équivalence entre les points (a) et (b) dans la définition du Brownien : supposons d'abord que (a) soit vraie et fixons $r \leq s < t$. On a

$$\mathbb{E}[X_r(X_t - X_s)] = \mathbb{E}[X_r X_t] - \mathbb{E}[X_r X_s] = r \wedge t - r \wedge s = r - r = 0.$$

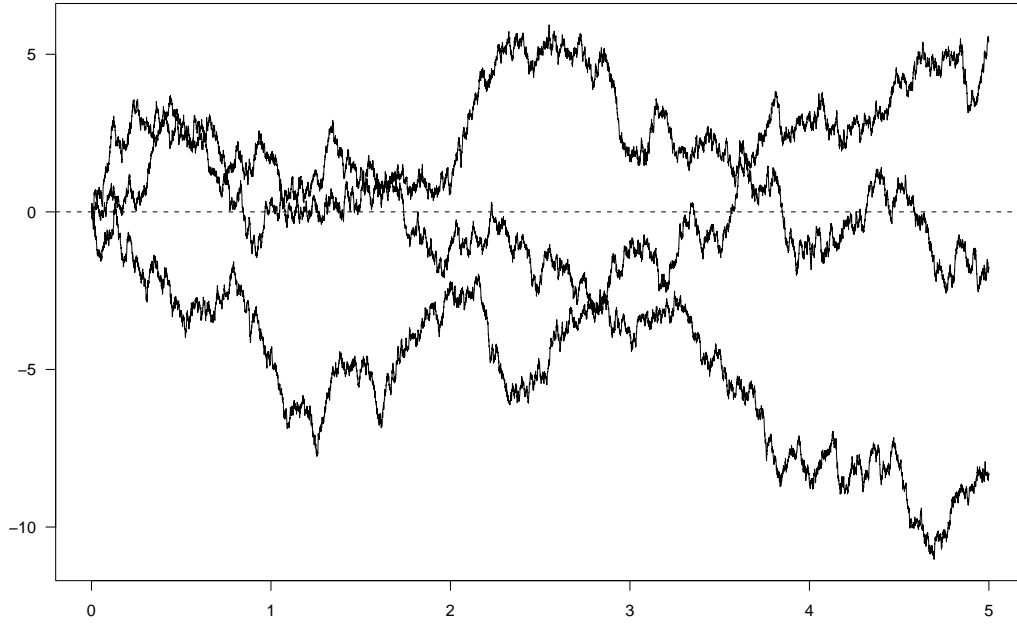


FIG. 3.1 – Trajectoires Browniennes

Donc $H = \text{Vect}\{X_r, r \leq s\}$ est orthogonal à $X_t - X_s$ dans L^2 . Mais comme X est un processus gaussien par hypothèse, ceci entraîne que $\sigma(H) = \sigma\{X_r, r \leq s\}$ est indépendante de $X_t - X_s$, et donc que X est à accroissements indépendants. Enfin, on sait que $X_t - X_s$ est une gaussienne centrée, et qu'il suffit donc de calculer sa variance pour déterminer sa loi. Or

$$\text{Var}(X_t - X_s) = \mathbb{E}[(X_t - X_s)^2] = \mathbb{E}[X_t^2] - 2\mathbb{E}[X_s X_t] + \mathbb{E}[X_s^2] = t - 2s + s = t - s,$$

d'où $X_t - X_s \sim \mathcal{N}(0, t - s)$, ce qui montre (b). Réciproquement, supposons que (b) ait lieu. Soient $a_1, \dots, a_n \in \mathbb{R}$ et $0 = t_0 \leq t_1 < \dots < t_n$. On peut redécomposer

$$\sum_{i=1}^n a_i X_{t_i} = \sum_{i=1}^n (a_i + \dots + a_n)(X_{t_i} - X_{t_{i-1}})$$

et on voit que le terme de droite est une somme de gaussiennes indépendantes (par indépendance des accroissements), donc une gaussienne. Ceci entraîne que X est un processus gaussien, et l'on voit immédiatement qu'il est centré. Pour calculer sa covariance, on écrit pour tout $s \leq t$

$$\mathbb{E}[X_s X_t] = \mathbb{E}[X_s(X_s + X_t - X_s)] = \mathbb{E}[X_s^2] + \mathbb{E}[X_s(X_t - X_s)] = s + 0 = s = s \wedge t.$$

□

Exercice 3.2 * Soit W un mouvement brownien et $0 \leq r < s < t$. Calculer les quantités $\mathbb{P}[W_r > 0]$, $\mathbb{P}[W_r > 0, W_s > 0]$ et $\mathbb{P}[W_r > 0, W_s > 0, W_t > 0]$.

Exercice 3.3 (a) Soit X un processus à accroissements indépendants. Montrer que si $\mathbb{E}[|X_t|] < +\infty$ (resp. $\mathbb{E}[|X_t|^2] < +\infty$), alors le processus $X_t^1 = X_t - \mathbb{E}[X_t]$ (resp. $X_t^2 = X_t^2 - \mathbb{E}[X_t^2]$) est une \mathcal{F}_t^X -martingale. Montrer aussi que si $\mathbb{E}[\exp \theta X_t] < +\infty$ pour un certain $\theta \in \mathbb{R}$, alors

$$X_t^3 = \frac{\exp \theta X_t}{\mathbb{E}[\exp \theta X_t]}$$

est une \mathcal{F}_t^X -martingale.

(b) Après avoir vérifié les hypothèses d'intégrabilité, identifier les processus X^1 , X^2 et X^3 dans le cas du mouvement brownien B : montrer que $X_t^1 = B_t$, $X_t^2 = (B_t)^2 - t$ et que $X_t^3 = \exp[\theta B_t - \theta^2 t/2]$.

Le point (b) de l'exercice précédent admet une célèbre (et très utile) réciproque, due à Paul Lévy :

Théorème 3.4 [Théorème de caractérisation de P. Lévy] *Soit $\{X_t, t \geq 0\}$ un processus continu issu de 0. Alors X est un mouvement brownien si l'une des deux conditions suivantes est vérifiée :*

- (a) *Les processus X et $t \mapsto (X_t)^2 - t$ sont des martingales.*
- (b) *Le processus $t \mapsto \exp[\theta X_t - \theta^2 t/2]$ est une martingale pour tout $\theta \in \mathbb{R}$.*

Remarque 3.5 Si B^1 et B^2 sont deux Browniens indépendants, alors le produit $B^1 B^2$ est une martingale. On peut le voir de deux façons :

- (a) En utilisant le fait que si \mathcal{F}, \mathcal{G} sont deux tribus et X, Y sont deux v.a. telles que $X \vee \mathcal{F}$ et \mathcal{G} sont indépendantes ainsi que $Y \vee \mathcal{G}$ et \mathcal{F} , alors $\mathbb{E}[XY | \mathcal{F} \vee \mathcal{G}] = \mathbb{E}[X | \mathcal{F}] \mathbb{E}[Y | \mathcal{G}]$.
- (b) En remarquant que $(B^1 + B^2)/\sqrt{2}$ est un processus gaussien de covariance $t \wedge s$, donc un Brownien. Donc $(B_t^1 + B_t^2)^2/2 - t$ est une martingale, de sorte que $B_t^1 B_t^2 = ((B_t^1 + B_t^2)^2 - t) - ((B_t^1)^2 - t)/2 - ((B_t^2)^2 - t)/2$, différence de trois martingales, est aussi une martingale.

Le théorème suivant, également célèbre, va maintenant permettre de préciser la régularité de la trajectoire brownienne :

Théorème 3.6 [Lemme de Kolmogorov] *Soit $\{X_t, t \geq 0\}$ un processus à valeurs dans \mathbb{R}^d tel qu'il existe c, ε, p avec*

$$\mathbb{E}[|X_t - X_s|^p] \leq c|t - s|^{1+\varepsilon}$$

pour tous $t, s \geq 0$. Alors p.s. les trajectoires de X sont α -Hölder pour tout $\alpha \in]0, \varepsilon/p[$.

Les démonstrations des théorèmes de Lévy et de Kolmogorov sont par exemple données dans [11] ou dans [17]. Rappelons qu'une fonction $f : \mathbb{R}^+ \rightarrow \mathbb{R}^d$ est α -Hölder si

$$\sup_{0 \leq s < t \leq T} \frac{|f(t) - f(s)|}{|t - s|^\alpha} < +\infty$$

pour tout $T \geq 0$. En particulier, on voit que pour tout $\beta \geq \alpha \geq 0$, f β -Hölder $\Rightarrow f$ α -Hölder $\Rightarrow f$ continue, et que f dérivable $\Rightarrow f$ 1-Hölder. En gros, la α -Hölderianité de f signifie que localement au voisinage de x_0 , la courbe de f est comprise entre celle de $x \mapsto f(x_0) + |x - x_0|^\alpha$ et celle de $x \mapsto f(x_0) - |x - x_0|^\alpha$. Pour le Brownien, on calcule

$$\begin{aligned} \mathbb{E}[|X_t - X_s|^p] &= \frac{1}{\sqrt{2\pi(t-s)}} \int_{\mathbb{R}} |x|^p e^{-|x|^2/2(t-s)} dx \\ &= \left(\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} |x|^p e^{-|x|^2/2} dx \right) |t-s|^{p/2} = c_p |t-s|^{1+(p-2)/2}, \end{aligned}$$

de sorte que par le lemme de Kolmogorov, le Brownien est α -Hölder pour tout $\alpha \in]0, 1/2 - 1/p[$. En faisant tendre $p \rightarrow \infty$, on voit que le Brownien est α -Hölder pour tout $\alpha \in]0, 1/2[$. On peut montrer que la valeur $1/2$ n'est pas atteinte :

$$\sup_{0 \leq s < t \leq T} \frac{|X_t - X_s|}{|t - s|^{1/2}} = +\infty$$

pour tout $T \geq 0$. En particulier, la courbe brownienne a la propriété remarquable d'être *partout continue et nulle part dérivable*. Signalons qu'il n'est pas évident de construire de telles fonctions sans faire appel aux probabilités, un exemple célèbre étant cependant la fonction de Weierstrass (1889) :

$$f(t) = \sum_{k=1}^{+\infty} \alpha^k \cos \beta^k t,$$

qui est partout continue et nulle part dérivable dès que $\alpha\beta > 1 + 3\pi/2$.

La filtration naturelle de B est $\mathcal{F}_t^B = \sigma\{B_s, s \leq t\} \vee \mathcal{N}$ où \mathcal{N} représente la tribu des négligeables pour \mathbb{P} . Comme B est à trajectoires continues, rappelons que $\{\mathcal{F}_t^B, t \geq 0\}$ est automatiquement càd :

$$\mathcal{F}_t^B = \bigcap_{s>t} \mathcal{F}_s^B.$$

Transformations du Brownien. Si $\{B_t, t \geq 0\}$ est un mouvement brownien, alors on peut montrer, en vérifiant leur caractère gaussien et en calculant leur espérance et leur covariance, que les processus suivants :

- (a) $t \mapsto -B_t$ (processus symétrique),
- (b) $t \mapsto c^{-1}B_{c^2t}$ (processus rééchelonné),
- (c) $t \mapsto tB_{1/t}$ si $t \neq 0$, $t \mapsto 0$ si $t = 0$ (processus inversé),

sont aussi des mouvements browniens. Remarquons que le fait que le processus (b) soit un Brownien est implicite dans la construction par discrétisation en $(N^{-1/2}, N^{-1})$ vue précédemment. L'invariance du processus (b) porte aussi le nom d'*invariance par Scaling* et signifie que le Brownien est un processus *autosimilaire*. Une étude exhaustive des processus autosimilaires, assez couramment utilisés en Finance, est le livre [20]. Venons-en maintenant à d'autres propriétés trajectorielles du Brownien :

Proposition 3.7 *Soit $\{B_t, t \geq 0\}$ un mouvement brownien. Alors p.s.*

- (a) $\limsup_{t \rightarrow +\infty} t^{-1/2}B_t = \limsup_{t \rightarrow 0} t^{-1/2}B_t = +\infty$.
- (b) $\liminf_{t \rightarrow +\infty} t^{-1/2}B_t = \liminf_{t \rightarrow 0} t^{-1/2}B_t = -\infty$.
- (c) $\lim_{t \rightarrow +\infty} t^{-1}B_t = 0$.

Preuve : Pour (a), on considère la v.a.

$$R = \limsup_{t \rightarrow +\infty} t^{-1/2}B_t = \limsup_{t \rightarrow +\infty} t^{-1/2}(B_t - B_s)$$

pour tout $s \geq 0$. Par indépendance des accroissements de B , R est indépendante de $\sigma\{B_u, u \leq s\}$ pour tout $s \geq 0$, et donc de $\sigma\{B_u, u \geq 0\}$ en faisant tendre $s \rightarrow \infty$. Or R est clairement mesurable par rapport à $\sigma\{B_u, u \geq 0\}$, de sorte que R est indépendante d'elle-même, et ceci n'est évidemment possible que si R est une constante déterministe (loi du tout ou rien). Supposons que R soit finie. Alors $\mathbb{P}[t^{-1/2}B_t \geq R+1] \rightarrow 0$ quand $t \rightarrow +\infty$. Mais par scaling,

$$\mathbb{P}[t^{-1/2}B_t \geq R+1] = \mathbb{P}[B_1 \geq R+1] \neq 0$$

pour tout $t \geq 0$, de sorte que nécessairement $R = +\infty$. De même, on montre que $\limsup_{t \rightarrow 0} t^{-1/2}B_t = +\infty$ p.s.

Le point (b) est une conséquence immédiate de la symétrie du Brownien, et le point (c) découle simplement de la loi des grands nombres. □

Cette proposition permet de mesurer grossièrement deux quantités reliées à la courbe brownienne, d'une part sa vitesse locale en chaque point t_0 - qui est plus grande que $|t - t_0|^{1/2}$, d'autre part le volume qu'elle occupe dans le demi-espace $\mathbb{R}^+ \times \mathbb{R}$ - qui est plus grand à l'infini que celui délimité par les courbes $t \mapsto \sqrt{t}$ et $t \mapsto -\sqrt{t}$, mais plus petit que celui délimité par les courbes $t \mapsto t$ et $t \mapsto -t$. Une question naturelle concerne alors la vitesse locale et le taux de croissance *exacts* du Brownien. La réponse est donné par le célèbre théorème suivant, dû à Khintchine :

Théorème 3.8 [Loi du logarithme itéré] *Soit $\{B_t, t \geq 0\}$ un mouvement brownien. Alors p.s.*

- (a) $\limsup_{t \rightarrow +0} B_t/\psi(t) = 1$ et $\liminf_{t \rightarrow 0} B_t/\psi(t) = -1$,
- (b) $\limsup_{t \rightarrow +\infty} B_t/\varphi(t) = 1$ et $\liminf_{t \rightarrow +\infty} B_t/\varphi(t) = -1$,

où l'on a posé $\psi(t) = \sqrt{2t \log \log(1/t)}$ pour tout $t < 1/e$, et $\varphi(t) = \sqrt{2t \log \log t}$ pour tout $t > e$.

Remarquons que les fonctions ψ et φ ressemblent beaucoup à $t \mapsto \sqrt{t}$ au sens où $\psi(t)/t^{1/2} \rightarrow +\infty$, $\psi(t)/t^\alpha \rightarrow 0$ pour tout $\alpha < 1/2$ quand $t \rightarrow 0$, et $\varphi(t)/t^{1/2} \rightarrow +\infty$, $\varphi(t)/t^\alpha \rightarrow 0$ pour tout $\alpha > 1/2$ quand $t \rightarrow +\infty$. Ainsi, une bonne image déterministe de la courbe brownienne, aussi bien localement que globalement, sont les fonctions $t \mapsto \sqrt{t}$ et $t \mapsto -\sqrt{t}$. Ceci était d'ailleurs déjà suggéré par la propriété de Scaling vue précédemment. Le théorème suivant sera utilisé dans la construction de l'intégrale stochastique. Il met une nouvelle fois en évidence le caractère "quadratique" de la courbe brownienne :

Théorème 3.9 Fixons $t > 0$ et posons $t_j = jt/2^n$ pour tout $n \in \mathbb{N}$ et $j = 0, \dots, 2^n$. Alors

$$Z_t^n = \sum_{j=1}^{2^n} |B_{t_j} - B_{t_{j-1}}|^2 \rightarrow t$$

quand $n \rightarrow \infty$, p.s. et dans L^2 .

Preuve : Remarquons déjà que $\mathbb{E}[Z_t^n] = t$. Pour la convergence dans L^2 il faut montrer que $\mathbb{E}[|Z_t^n - t|^2] \rightarrow 0$, soit $\text{Var}[Z_t^n] \rightarrow 0$ ce qui se déduit de

$$\text{Var}[Z_t^n] = \sum_{j=1}^{2^n} \text{Var}[B_{t_j} - B_{t_{j-1}}]^2 = \sum_{j=1}^{2^n} 2(t/2^n)^2 = 2^{1-n} t^2,$$

où la deuxième égalité vient de ce que si $X \sim \mathcal{N}(0, \sigma^2)$, alors $\text{Var} X^2 = 2\sigma^4$. On en déduit d'autre part que

$$\mathbb{E} \left[\sum_{n=1}^{\infty} |Z_t^n - t|^2 \right] = t \sum_{n=1}^{\infty} 2^{-n} < \infty,$$

d'où $\sum_{n=1}^{\infty} |Z_t^n - t|^2 < \infty$ p.s. (une v.a. positive d'espérance finie est finie), de sorte que le terme général de la série $|Z_t^n - t|^2$ converge nécessairement p.s. vers 0. □

Le théorème suivant, d'une portée surtout théorique mais fondamentale pour la théorie actuelle des équations différentielles stochastiques, montre cependant que le choix d'une subdivision non régulière peut rendre infinie la variation quadratique du Brownien :

Théorème 3.10 [Théorème de la p -variation forte] Pour tout $p \geq 1$, soit

$$V_t^p = \sup_{0=t_0 \leq t_1 \leq \dots \leq t_n=t} \left(\sum_{i=1}^n |B_{t_{i+1}}(\omega) - B_{t_i}(\omega)|^p \right)$$

la p -variation forte du Brownien sur $[0, t]$. Alors pour tout $t \geq 0$, $V_t^p = \infty$ p.s. $\Leftrightarrow p \leq 2$.

Il est facile de voir par l'inégalité de Hölder que si $p \geq q$, $V_t^p < \infty \Rightarrow V_t^q < \infty$ pour tout $t \geq 0$. En particulier, le théorème précédent entraîne que le Brownien n'est pas à un processus à variations finies, au sens du chapitre 1.

3.3 La propriété de Markov et ses conséquences

La propriété de Markov du Brownien W est ici une conséquence de l'indépendance de ses accroissements : on sait que, pour tout $s \geq 0$, le processus $\{W_u^s = W_{u+s} - W_s, u \geq 0\}$ est un mouvement brownien indépendant de \mathcal{F}_s^W et ceci entraîne le

Théorème 3.11 [Propriété de Markov simple] Soit W un mouvement brownien. Pour toute fonction f borélienne bornée et pour tout $s \leq t$,

$$\mathbb{E}[f(W_t) | \mathcal{F}_s^W] = \mathbb{E}[f(W_t) | W_s] = \frac{1}{\sqrt{2\pi(t-s)}} \int_{\mathbb{R}} f(y) e^{-(y-W_s)^2/2(t-s)} dy.$$

Preuve : On utilise les propriétés élémentaires de l'espérance conditionnelle :

$$\mathbb{E}[f(W_t) | \mathcal{F}_s^W] = \mathbb{E}[f(W_s + W_t - W_s) | \mathcal{F}_s^W] = \mathbb{E}[f(x + W_t - W_s)]_{x=W_s}$$

puisque W_s est \mathcal{F}_s^W -mesurable et $(W_t - W_s)$ est indépendant de \mathcal{F}_s^W . Comme $(W_t - W_s) \sim \mathcal{N}(0, t-s)$, on peut calculer

$$\begin{aligned} \mathbb{E}[f(x + W_t - W_s)] &= \frac{1}{\sqrt{2\pi(t-s)}} \int_{\mathbb{R}} f(y+y) e^{-(y)^2/2(t-s)} dy \\ &= \frac{1}{\sqrt{2\pi(t-s)}} \int_{\mathbb{R}} f(y) e^{-(y-x)^2/2(t-s)} dy. \end{aligned}$$

□

Une autre façon de décrire cette propriété est de dire que pour $t \geq s$, conditionnellement à \mathcal{F}_s^W , la v.a. W_t est de loi gaussienne d'espérance W_s et de variance $t - s$. Le théorème suivant étend la propriété de Markov à des temps aléatoires :

Théorème 3.12 [Propriété de Markov forte] *Soit W un mouvement brownien et T un (\mathcal{F}_t^W) -temps d'arrêt à valeurs finies. Pour toute fonction f borélienne bornée et pour tout $t \geq 0$,*

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

En particulier, pour tout temps d'arrêt fini T , le processus $\{W_t^T = W_{T+t} - W_T, t \geq 0\}$ est un Brownien indépendant de \mathcal{F}_T^W .

La propriété de Markov forte est sans conteste l'une des propriétés les plus utiles du Brownien. Une combinaison de cette propriété avec l'invariance par Scaling et le Théorème d'arrêt forme un cocktail d'une efficacité rarement démentie, et dont nous verrons quelques exemples par la suite.

Théorie des Fluctuations. Cette théorie concerne la loi jointe du processus bivarié $\{(W_t, S_t), t \geq 0\}$ où S est le processus (croissant) du supremum unilatère : $S_t = \sup_{s \leq t} W_s$. On a vu précédemment que $\limsup_{t \rightarrow +\infty} t^{-1/2} W_t = +\infty$, ce qui entraîne clairement que $S_t \rightarrow +\infty$ p.s. quand $t \rightarrow +\infty$. En particulier, si $T_a = \inf \{t > 0, W_t = a\}$ désigne le temps d'atteinte du Brownien au seuil $a > 0$, alors pour tout $M > 0$, $\{T_a < M\} = \{S_M > a\}$ et comme $\mathbb{P}[S_M > a] \uparrow 1$ quand $M \uparrow +\infty$, on voit que p.s. $T_a < +\infty$, autrement dit T_a est un temps d'arrêt fini (attention, T_a n'est pas un temps d'arrêt borné : en effet $\mathbb{E}(T_a) = \infty$). Le théorème suivant, utilisé en Finance pour évaluer les options barrières, précise la loi jointe du couple de v.a. (W_t, S_t) :

Théorème 3.13 [Théorème de Désiré André] *Pour tout $t, x \geq 0$ et pour tout $y \leq x$*

$$\mathbb{P}[S_t \geq x, W_t \leq y] = \mathbb{P}[W_t \geq 2x - y].$$

Preuve : On écrit d'abord

$$\mathbb{P}[S_t \geq x, W_t \leq y] = \mathbb{P}[T_x \leq t, W_t \leq y] = \mathbb{P}[T_x \leq t, W_{t-T_x+T_x} - W_{T_x} \leq y - W_{T_x}].$$

D'une part $W_{T_x} = x$ - par continuité de W , d'autre part $W_s^x = W_{s+T_x} - W_{T_x}$ est un Brownien indépendant de \mathcal{F}_{T_x} (donc de T_x). On en déduit que

$$\mathbb{P}[S_t \geq x, W_t \leq y] = \mathbb{P}[T_x \leq t, W_{t-T_x}^x \leq y - x] = \mathbb{P}[T_x \leq t, W_{t-T_x}^x \geq x - y]$$

où la deuxième égalité vient de la symétrie du processus W^x sachant T_x . Mais

$$\mathbb{P}[T_x \leq t, W_{t-T_x}^x \geq x - y] = \mathbb{P}[T_x \leq t, W_t - x \geq x - y] = \mathbb{P}[W_t \geq 2x - y]$$

où la deuxième égalité vient de l'inclusion $\{W_t \geq 2x - y\} \subset \{T_x \leq t\}$, laquelle découle immédiatement de ce que $2x - y \geq x$.

□

En dérivant par rapport à x puis y , on obtient la densité jointe du couple de v.a. (S_t, W_t) :

$$f_{(S_t, W_t)}(x, y) = \frac{2(2x - y)}{\sqrt{2\pi t^3}} e^{-(2x-y)^2/2t} \mathbb{1}_{\{x \geq y, x \geq 0\}}$$

pour tout $t \geq 0$. On peut aussi calculer la loi de S_t , soit en intégrant l'expression précédente par rapport à y , soit en écrivant pour tout $x \geq 0$

$$\begin{aligned} \mathbb{P}[S_t \geq x] &= \mathbb{P}[S_t \geq x, W_t \geq x] + \mathbb{P}[S_t \geq x, W_t \leq x] \\ &= \mathbb{P}[W_t \geq x] + \mathbb{P}[W_t \geq 2x - x] \\ &= 2\mathbb{P}[W_t \geq x] = \mathbb{P}[|W_t| \geq x] \end{aligned}$$

où la deuxième égalité est une conséquence du théorème de Désiré André et la dernière vient de la symétrie du Brownien. Par égalité des fonctions de répartition on en déduit alors que

$$S_t \stackrel{d}{=} |W_t|$$

pour tout $t \geq 0$. Ceci permet de calculer simplement la loi de T_a pour $a > 0$:

$$\mathbb{P}[T_a \leq t] = \mathbb{P}[S_t \geq a] = \mathbb{P}[|W_t| \geq a] = \mathbb{P}[W_t^2 \geq a^2] = \mathbb{P}[tW_1^2 \geq a^2] = \mathbb{P}[a^2W_1^{-2} \leq t]$$

de sorte que par égalité des fonctions de répartition,

$$T_a \stackrel{d}{=} \frac{a^2}{W_1^2}.$$

On calcule alors directement la densité

$$f_{T_a}(t) = \frac{a}{\sqrt{2\pi t^3}} e^{-a^2/2t} \mathbb{1}_{\{t \geq 0\}}.$$

On peut également montrer que

Théorème 3.14 *Les processus $\{S_t - W_t, t \geq 0\}$ et $\{|W_t|, t \geq 0\}$ ont même loi.*

La force de ce théorème vient de ce qu'il s'agit d'une égalité en loi *entre processus*, qui est bien plus forte que l'égalité en loi entre v.a. Par exemple on a vu que $S_t \stackrel{d}{=} |W_t|$ pour tout $t \geq 0$ mais les *processus* S et $|W|$ n'ont pas la même loi, puisque $|W|$ n'est pas croissant. En revanche, ce théorème entraîne le fait curieux que les v.a. S_t et $S_t - W_t$ ont même loi.

Nous souhaitons maintenant retrouver la densité de T_a par une autre méthode, très instructive, reposant sur le théorème d'arrêt. Soit $a > 0$. Comme $T_a \wedge n$ est un temps d'arrêt borné et que $t \mapsto e^{\lambda W_t - \lambda^2 t/2}$ est une martingale pour tout $\lambda \geq 0$, ce théorème entraîne que

$$\mathbb{E}[e^{\lambda W_{T_a \wedge n} - \lambda^2 (T_a \wedge n)/2}] = \mathbb{E}[e^{\lambda W_0 - \lambda^2 \cdot 0/2}] = 1.$$

On fait tendre $n \rightarrow +\infty$: $T_a \wedge n \rightarrow T_a$ et comme p.s. $T_a < +\infty$, $W_{T_a \wedge n} \rightarrow W_{T_a} = a$. De plus, comme $a, \lambda \geq 0$ et par définition de T_a , pour tout $n \geq 0$

$$e^{\lambda W_{T_a \wedge n} - \lambda^2 (T_a \wedge n)/2} \leq e^{\lambda W_{T_a \wedge n}} \leq e^{\lambda a}$$

et $e^{\lambda a}$ est constante finie, donc une v.a. intégrable. On peut appliquer le théorème de convergence dominée, qui donne

$$1 = \mathbb{E}[e^{\lambda W_{T_a \wedge n} - \lambda^2 (T_a \wedge n)/2}] \rightarrow \mathbb{E}[e^{\lambda a - \lambda^2 T_a}]$$

quand $n \rightarrow +\infty$, de sorte que

$$\mathbb{E}[\exp -\lambda T_a] = e^{-a\sqrt{2\lambda}}$$

pour tout $\lambda \geq 0$. Par inversion de la transformée de Laplace, on obtient alors la densité de T_a donnée précédemment.

Remarque 3.15 Pour $a < 0$, on peut remarquer que $T_a = \inf\{t \geq 0; B_t = a\} = \inf\{t \geq 0; W_t = -a\}$, avec $W = -B$. Il serait faux de conclure que $E(\exp -\lambda T_a) = \exp(-a\sqrt{2\lambda})$ pour tout a car, pour $a < 0$ et $\lambda > 0$, le membre de gauche est plus petit que 1 et celui de droite serait plus grand que 1. Il convient donc d'être vigilant en appliquant le théorème d'arrêt de Doob.

On peut également s'intéresser au premier temps de passage *bilatère* : $\hat{T}_a = \inf\{t \geq 0, |B_t| = a\}$ pour $a > 0$, et au processus supremum bilatère associé $S_t^* = \sup_{s \leq t} |W_s|$. Avec des méthodes analogues, on montre que

$$\mathbb{E}[e^{-\lambda \hat{T}_a}] = 1/\cosh a\sqrt{2\lambda}$$

pour tout $\lambda \geq 0$. L'inversion de cette transformation de Laplace donne cependant une densité plus compliquée obtenue sous forme de développement en série, voir [11].

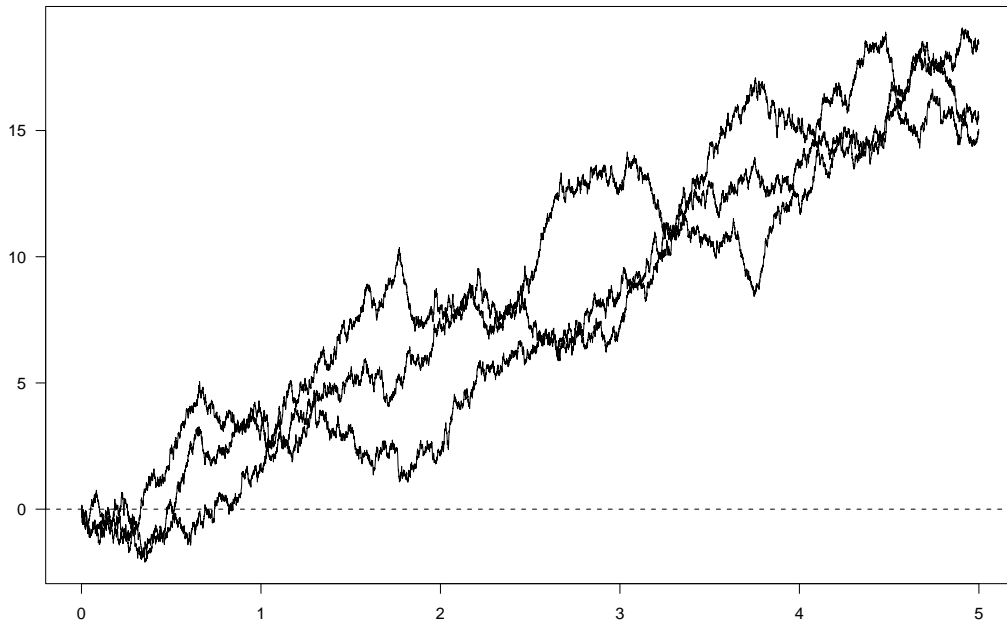


FIG. 3.2 – Mouvement Brownien avec drift

- Exercice 3.16** (a) Calculer la transformée de Laplace et la densité de T_a pour $a < 0$.
 (b) Montrer que $\mathbb{E}[T_a] = +\infty$ pour tout $a \neq 0$.
 (c)* Soit $a < 0 < b$. Montrer que $\mathbb{P}[T_a < T_b] = b/(b-a)$ et, si on pose $T = T_a \wedge T_b$, que

$$\mathbb{E}[e^{-\lambda T} \mathbb{1}_{\{T=T_a\}}] = \frac{\sinh b\sqrt{2\lambda}}{\sinh(b-a)\sqrt{2\lambda}} \quad \text{et} \quad \mathbb{E}[e^{-\lambda T}] = \frac{\cosh(a+b)\sqrt{\lambda/2}}{\cosh(b-a)\sqrt{\lambda/2}}$$

Terminons cette section en évoquant le *mouvement brownien drifté* :

$$W^\mu = \{W_t^\mu = W_t + \mu t, t \geq 0\}$$

pour tout $\mu \in \mathbb{R}$. Le processus W^μ est une sous-martingale si $\mu > 0$ et une surmartingale si $\mu < 0$. La variable W_t^μ est une Gaussienne d'espérance μt et de variance t . Un certain nombre de propriétés du Brownien se transmettent au Brownien drifté : propriété de Markov, théorèmes d'arrêt... mais attention pas toutes : par exemple, W^μ n'est invariant par scaling que pour $\mu = 0$. Nous reviendrons sur le Brownien drifté vers la fin de ce cours, dans le cadre du théorème de Girsanov.

Exercice 3.17 Soit W un mouvement brownien. On considère les processus $t \mapsto |W_t|$, $t \mapsto e^{W_t}$ et $t \mapsto \int_0^t W_s ds$.

- (a) Calculer leur espérance et leur variance.
 (b) Lesquels sont des processus gaussiens ? Des processus de Markov ?

3.4 Equation de la chaleur

Cette section peut être omise en première lecture. Soit $g(t, x)$ la densité gaussienne centrée de variance t . On note

$$q(t, x, y) = \frac{1}{\sqrt{2\pi t}} e^{-(y-x)^2/2t} = g(t, x-y)$$

la densité de transition du mouvement brownien. C'est de façon heuristique, la probabilité pour que le Brownien soit en y sachant que t instants auparavant, il se trouvait en x . Par stationnarité des accroissements du Brownien, c'est aussi la densité conditionnelle

$$\mathbb{P}[B_{t+s} \in dy \mid B_s = x] = q(t, x, y) dy.$$

La densité de transition q vérifie les équations *forward et backward*

$$\frac{\partial q}{\partial t}(t, x, y) = \frac{1}{2} \frac{\partial^2 q}{\partial y^2}(t, x, y) = \frac{1}{2} \frac{\partial^2 q}{\partial x^2}(t, x, y).$$

Si, pour une fonction mesurable bornée f , on considère la fonction

$$u_f(t, x) = \mathbb{E}[f(B_t + x)] = \int f(x + y)g(t, y)dy = \int f(y)q(t, x, y)dy,$$

on déduit de l'équation forward et du théorème de dérivation sous l'intégrale que u_f vérifie l'équation aux dérivées partielles

$$\begin{cases} u_f(0, x) &= f(x) \\ \frac{1}{2} \frac{\partial^2 u_f}{\partial x^2} &= \frac{\partial u_f}{\partial t} \end{cases}$$

Cette équation porte le nom d'*équation de la chaleur* (u_f représente l'évolution de la température dans un corps homogène, avec conditions initiales données par f). De plus, en dérivant sous l'intégrale, on a

$$\frac{\partial^2 u_f}{\partial x^2}(t, x) = \int_{\mathbb{R}} f''(x + y)g(t, y) dy = u_{f''}(t, x) = \mathbb{E}[f''(B_t + x)].$$

On peut ainsi écrire

$$\mathbb{E}[f(B_t + x)] - f(x) = u_f(t, x) - u_f(0, x) = \int_0^t \frac{\partial u_f}{\partial t}(s, x) ds = \frac{1}{2} \int_0^t \frac{\partial^2 u_f}{\partial x^2}(s, x) ds,$$

d'où

$$\mathbb{E}[f(B_t + x)] = f(x) + \frac{1}{2} \int_0^t \mathbb{E}[f''(B_s + x)] ds.$$

Une formule liant les processus $f(B_t + x)$ et $f''(B_t + x)$ sera donné dans le prochain chapitre (lemme d'Itô). On peut également considérer des fonctions temps-espace :

Proposition 3.18 Soit $f : \mathbb{R}^+ \times \mathbb{R} \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}_b^1 en temps et \mathcal{C}_b^2 en espace et B un Brownien. Alors

$$\mathbb{E}[f(t, x + B_t)] = f(0, x) + \int_0^t \mathbb{E} \left[f'_t(s, x + B_s) + \frac{1}{2} f''_{xx}(s, x + B_s) \right] ds.$$

Preuve : On pose $u_f(t, x) = \mathbb{E}[f(t, x + B_t)]$. En utilisant le théorème de dérivation des fonctions composées et l'équation de la chaleur, on en déduit que

$$\frac{du_f}{dt}(t, x) = \mathbb{E} \left[f'_t(t, x + B_t) + \frac{1}{2} f''_{xx}(t, x + B_t) \right].$$

En intégrant par rapport à t , on trouve

$$u_f(t, x) - u_f(0, x) = \int_0^t \mathbb{E} \left[f'_t(s, x + B_s) + \frac{1}{2} f''_{xx}(s, x + B_s) \right] ds.$$

□

On définit maintenant le *générateur* \mathcal{L} du processus espace-temps (t, B_t) comme l'opérateur agissant sur les fonctions de classe \mathcal{C}_b^1 en temps et \mathcal{C}_b^2 en espace :

$$\mathcal{L}(f)(t, x) = \partial_t f(t, x) + \frac{1}{2} \partial_{xx} f(t, x).$$

Le théorème suivant est très utile en Finance :

Théorème 3.19 Soit u de classe \mathcal{C}_b^1 en temps et \mathcal{C}_b^2 en espace telle que $\mathcal{L}u \equiv 0$. Alors le processus $t \mapsto u(t, B_t)$ est une martingale.

Preuve : On utilise à nouveau l'indépendance des accroissements du Brownien :

$$\mathbb{E}[u(t, B_t) | \mathcal{F}_s] = \mathbb{E}[u(s+t-s, x+B_{t-s})]_{x=B_s} = \mathbb{E}[u_{s,x}(t-s, B_{t-s})]_{x=B_s}$$

où l'on a noté $u_{s,x}(t, y) = u(s+t, x+y)$. Par la proposition précédente, on a donc

$$\mathbb{E}[u_{s,x}(t-s, B_{t-s})] = u_{s,x}(0, 0) + \int_0^t \mathbb{E}[\mathcal{L}u_{s,x}(w, B_w)] dw$$

Comme $\mathcal{L}u_{s,x} \equiv 0$ par hypothèse, il en résulte que

$$\mathbb{E}[u(t, B_t) | \mathcal{F}_s] = u_{s,x}(0, 0)_{x=B_s} = u(s, B_s).$$

□

3.5 Mouvement brownien multidimensionnel

Soit $\{B_t = (B_t^1, B_t^2, \dots, B_t^n), t \geq 0\}$ un processus n -dimensionnel. On dit que B est un Brownien multidimensionnel si les processus $B^i, i \leq n$ sont des Browniens réels indépendants. B est aussi un processus à accroissements indépendants et stationnaires et un processus gaussien de fonction de covariance

$$\mathbb{E}[(B_t, B_s)] = n(s \wedge t)$$

(où (\cdot, \cdot) désigne le produit scalaire dans \mathbb{R}^n). On a également une caractérisation de type Lévy : un processus n -dimensionnel B est un mouvement brownien si et seulement si tous les processus B^i et $(B_t^i B_t^j - \delta_{i,j} t, t \geq 0)$ sont des martingales (avec la notation de Kronecker $\delta_{i,j} = 0$ pour $i \neq j$ et $\delta_{i,i} = 1$). Pour tous a_1, \dots, a_n , il est facile de vérifier en calculant son espérance et sa covariance que le processus W défini par

$$W_t = \frac{1}{\sqrt{a_1^2 + \dots + a_n^2}} (a_1 B_t^1 + \dots + a_n B_t^n)$$

est un Brownien réel. On dira que deux mouvements Browniens réels B^1 et B^2 sont corrélés avec coefficient de corrélation ρ si le processus $t \mapsto B_t^1 B_t^2 - \rho t$ est une martingale. On "décorrèle" alors B^1 et B^2 en introduisant le processus

$$B_t^3 = \frac{1}{\sqrt{1-\rho^2}} (B_t^2 - \rho B_t^1).$$

Ce processus est une martingale et en écrivant

$$\begin{aligned} (B_t^3)^2 - t &= \frac{1}{1-\rho^2} ((B_t^2)^2 + \rho^2 (B_t^1)^2 - 2\rho B_t^2 B_t^1 - t(1-\rho^2)) \\ &= \frac{1}{1-\rho^2} [(B^2(t))^2 - t + \rho^2 [(B^1(t))^2 - t] - 2\rho [B^2(t)B^1(t) - \rho t], \end{aligned}$$

on montre que $(B_t^3)^2 - t$ est une martingale, de sorte que B^3 est un MB. On peut montrer que B^3 est indépendant de B^2 , de sorte que le produit $B^2 B^3$ est une martingale.

Chapitre 4

L'intégrale stochastique - Formule d'Itô

Dans ce chapitre on cherche à définir des variables aléatoires du type

$$\omega \mapsto Y_1(\omega) = \left(\int_0^1 X_s dB_s \right) (\omega) \quad (4.1)$$

où $\{X_t, t \geq 0\}$ est un certain processus et $\{B_t, t \geq 0\}$ un mouvement brownien. Le problème est bien sûr de donner un sens à l'élément différentiel dB_s puisque la fonction $s \mapsto B_s$ n'est pas dérivable. Un objet mathématique adéquat, introduit par K. Itô en 1942, est l'*intégrale stochastique*, laquelle permet de construire $Y_1(\omega)$ comme une limite de v.a. sous l'hypothèse - cruciale - d'adaptation du processus X à la filtration du Brownien porteur. La théorie plus récente du *calcul anticipatif* permet de lever cette hypothèse d'adaptation dans certains cas, ce qui est d'un intérêt évident en Finance, mais nous n'aborderons pas ce sujet ici. Dans la pratique on s'intéresse souvent à des quantités du type $F(Y_1(\omega))$ où F est une fonction réelle. Quand F est suffisamment régulière, le *Lemme d'Itô* permet alors d'exprimer $F(Y_1(\omega))$ au moyen d'intégrales stochastiques, ce qui est par exemple une étape importante dans le calcul des espérances. Cette méthodologie s'applique aux intégrales $\int_0^t X_s dB_s$ et l'on peut ainsi considérer le processus

$$\int_0^t X_s dB_s, t \geq 0$$

qui est, sous conditions d'intégrabilité de X une martingale. Le processus $F(Y_t)$ est alors décomposé en une martingale et un processus à variation finies.

4.1 L'intégrale de Wiener

Dans cette section, les fonctions et processus seront à valeurs réelles, les définitions et résultats se généralisant sans difficulté au cas \mathbb{R}^d . L'intégrale de Wiener est simplement une intégrale du type (4.1) avec X fonction *déterministe*, i.e. ne dépendant pas de ω . On fixe un horizon $T > 0$ déterministe (éventuellement $T = +\infty$) et on note

$$L^2([0, T], \mathbb{R}) = \left\{ f : [0, T] \rightarrow \mathbb{R} / \int_0^T |f(s)|^2 ds < \infty \right\}.$$

Remarquons que si $T < +\infty$, les fonctions continues et les fonctions bornées sont contenues dans $L^2([0, T], \mathbb{R})$. On peut montrer que muni du produit scalaire

$$\langle f, g \rangle = \int_0^T f(s)g(s) ds,$$

$L^2([0, T], \mathbb{R})$ est un espace de Hilbert, au sens où toute suite de $L^2([0, T], \mathbb{R})$ qui soit de Cauchy pour la norme

$$\|f\|_{2,T} = \sqrt{\langle f, f \rangle} = \left(\int_0^T f^2(s) ds \right)^{1/2}$$

converge vers un élément unique de $L^2([0, T], \mathbb{R})$. La propriété fondamentale des espaces de Hilbert est l'existence d'une base orthonormée dénombrable : il existe un système de fonctions $\{f_n, n \geq 0\}$ de $L^2([0, T], \mathbb{R})$ tel que $\langle f_n, f_p \rangle = 0$ si $n \neq p$ et $\langle f_n, f_p \rangle = 1$ si $n = p$, et tel que tout élément f de $L^2([0, T], \mathbb{R})$ s'écrive de manière unique sous la forme

$$f = \sum_{n \geq 1} a_n f_n$$

pour des coefficients a_n qu'on appelle les coordonnées de f dans la base $\{f_n, n \geq 0\}$. Dans le cas précis de $L^2([0, T], \mathbb{R})$, la base $\{f_n, n \geq 0\}$ peut être constituée de *fonctions en escalier* :

$$f_n(t) = \sum_{i=1}^{p_n} \alpha_i \mathbb{1}_{]t_i^{(n)}, t_{i+1}^{(n)}]}(t) \quad (4.2)$$

où $p_n \in \mathbb{N}$, les α_i sont réels et $\{t_i^{(n)}\}$ une suite croissante de $[0, T]$. On en déduit alors le lemme suivant :

Lemme 4.1 [Lemme hilbertien] *Soit $f \in L^2([0, T], \mathbb{R})$. Il existe une suite de fonctions en escalier $\{f_n\}$ telle que*

$$\|f - f_n\|_{2,T} \rightarrow 0$$

quand $n \rightarrow +\infty$.

4.1.1 Le cas des fonctions en escalier

Si f_n est la fonction donnée par (4.2), que l'on note simplement $f_n(t) = \sum_{i=1}^{p_n} \alpha_i \mathbb{1}_{]t_i, t_{i+1}]}(t)$ il est très facile de définir son intégrale de Wiener par :

$$I_T(f_n) = \int_0^T f_n(s) dB_s = \sum_{i=1}^{p_n} \alpha_i (B_{t_{i+1}} - B_{t_i})$$

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

$$\text{Var}[I_T(f_n)] = \sum_{i=1}^{p_n} \alpha_i^2 \text{Var}(B_{t_{i+1}} - B_{t_i}) = \sum_{i=1}^{p_n} \alpha_i^2 (t_{i+1} - t_i) = \int_0^T f_n^2(s) ds.$$

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

$$\begin{aligned} \mathbb{E}[I_T(f)I_T(g)] &= \frac{1}{2} (\text{Var}[I_T(f) + I_T(g)] - \text{Var}[I_T(f)] - \text{Var}[I_T(g)]) \\ &= \frac{1}{2} \left(\int_0^T (f+g)^2(s) ds - \int_0^T f^2(s) ds - \int_0^T g^2(s) ds \right) \\ &= \int_0^T f(s)g(s) ds. \end{aligned}$$

Cette dernière égalité est très importante et signifie que l'application

$$f \mapsto I_T(f)$$

est une *isométrie* de $L^2([0, T], \mathbb{R})$ dans $L^2(\Omega, \mathbb{P})$. On parle alors de la propriété d'isométrie de l'intégrale de Wiener, ce qui signifie

$$\langle I_T(f), I_T(g) \rangle_{L^2(\Omega)} = \langle f, g \rangle_{L^2(\mathbb{R})}.$$

4.1.2 Le cas général

Pour construire $I_T(f)$ quand f est un élément quelconque de $L^2([0, T], \mathbb{R})$, on utilise l'isométrie mise en place et le lemme suivant :

Lemme 4.2 [Lemme gaussien] *Soit $\{X_n, n \geq 0\}$ une suite de variables gaussiennes $\mathcal{N}(\mu_n, \sigma_n)$ convergeant vers une v.a. X dans L^2 (soit telle que*

$$\mathbb{E}[|X - X_n|^2] \rightarrow 0$$

quand $n \rightarrow +\infty$. Alors, $\mu_n \rightarrow \mu$ et $\sigma_n \rightarrow \sigma$ quand $n \rightarrow +\infty$ et $X \sim \mathcal{N}(\mu, \sigma)$.

Soit maintenant $f \in L^2([0, T], \mathbb{R})$ et soit, d'après le Lemme 4.1, $\{f_n, n \geq 0\}$ une suite de fonctions en escalier telle que $\|f - f_n\|_{2,T} \rightarrow 0$ quand $n \uparrow +\infty$. D'après le paragraphe précédent on peut construire les intégrales de Wiener $I_T(f_n)$ qui sont des gaussiennes centrées qui, par isométrie forment une suite de Cauchy. L'espace L^2 étant complet, cette suite converge vers une v.a. gaussienne notée $I_T(f)$. Par le Lemme 4.2, $I_T(f) \sim \mathcal{N}(0, \|f\|_{2,T}^2)$. Il reste à vérifier que la limite Y ne dépend que de f et non pas de la suite $\{f_n\}$ choisie. En particulier $I_T(f)$ n'est *jamaïs* une variable p.s. positive, même quand f elle-même est toujours positive. De plus, l'application $f \rightarrow I_T(f)$ est linéaire et isométrique de $L^2([0, T], \mathbb{R})$ dans $L^2(\Omega)$, au sens où $I_T(af + bg) = aI_T(f) + bI_T(g)$ et

$$\mathbb{E}[I_T(f)I_T(g)] = \int_0^T f(s)g(s)ds$$

pour tous $a, b \in \mathbb{R}$ et $f, g \in L^2([0, T], \mathbb{R})$. Enfin, $I_T(f)$ est une variable gaussienne mesurable par rapport à $\sigma\{B_t, 0 \leq t \leq T\}$ qui vérifie pour tout $t \in [0, T]$

$$\begin{aligned} \mathbb{E}[I_T(f)B_t] &= \mathbb{E}\left[\left(\int_0^T f(s)dB_s\right)\left(\int_0^T \mathbb{1}_{[0,t]}(s)dB_s\right)\right] \\ &= \int_0^T f(s)\mathbb{1}_{[0,t]}(s)ds = \int_0^t f(s)ds, \end{aligned}$$

où la deuxième égalité provient de la formule d'isométrie. Par propriété d'espace gaussien, cette formule caractérise l'intégrale stochastique : $I_T(f)$ est l'unique v.a. Z gaussienne mesurable par rapport à $\sigma\{B_t, 0 \leq t \leq T\}$ telle que

$$\mathbb{E}[ZB_t] = \int_0^t f(s)ds$$

pour tout $t \in [0, T]$.

4.1.3 L'intégrale de Wiener vue comme processus gaussien

On note $L_{loc}^2(\mathbb{R}^+, \mathbb{R})$ la réunion des $L^2([0, T], \mathbb{R})$ pour $T > 0$ et on considère $f \in L_{loc}^2(\mathbb{R}^+, \mathbb{R})$. Par le paragraphe précédent, le processus

$$M_t = \int_0^t f(s)dB_s$$

a donc bien un sens pour tout $t \geq 0$. On montre alors le

Théorème 4.3 *Le processus $\{M_t, t \geq 0\}$ est un processus gaussien (\mathcal{F}_t^B) -adapté, centré, de fonction de covariance*

$$\Gamma(s, t) = \int_0^{t \wedge s} f^2(u) du.$$

De plus, M est un P.A.I. au sens où

$$\{M_{t+s} - M_s, t \geq 0\} \perp \sigma\{B_u, u \leq s\}$$

pour tout $s \geq 0$.

Preuve : Si l'on note M_t^n la suite d'intégrales de Wiener associée à la suite $\{f_n\}$ approchant f dans L^2 , on voit que $t \mapsto M_t^n$ est un processus gaussien comme le Brownien. Par stabilité dans L^2 des espaces gaussiens, on en déduit que $t \mapsto M_t$ est un processus gaussien. L'expression de son espérance et sa fonction de covariance découle des calculs précédents, et il est évident par construction que M est (\mathcal{F}_t^B) -adapté. Pour montrer l'indépendance des accroissements, on écrit pour tout $s, t \geq 0$

$$\begin{aligned} M_{t+s} - M_s &= \int_s^{t+s} f(u) dB_u \\ &= \int_s^{t+s} f(u) d_u(B_u - B_s) \\ &\in \sigma\{B_u - B_s, u \in [s, t+s]\} \end{aligned}$$

et l'on utilise l'indépendance des accroissements du processus B , qui entraîne que

$$\sigma\{B_u - B_s, u \in [s, t+s]\} \perp \sigma\{B_u, u \leq s\}$$

□

Comme conséquence de l'exercice 3.3, on obtient alors immédiatement l'important

Corollaire 4.4 *Les processus $\{M_t, t \geq 0\}$ et $\{\tilde{M}_t, t \geq 0\}$, où*

$$\tilde{M}_t = M_t^2 - \int_0^t f^2(s) ds$$

sont des (\mathcal{F}_t^B) -martingales.

Remarquons que sauf lorsque f est constante, le processus M est un P.A.I. mais n'est pas un P.A.I.S. (processus à accroissements indépendants et stationnaires, ou encore *processus de Lévy*), au sens où l'égalité

$$M_{t+s} - M_s \stackrel{d}{=} M_t$$

n'est pas vérifiée pour tout t . Ceci se comprend intuitivement vue l'expression de M sous forme intégrale. C'est aussi une conséquence d'un résultat beaucoup plus profond, la *formule de Lévy-Khintchine* [1] [21], qui est une sorte de formule de structure sur les P.A.I.S. Elle entraîne que les seuls processus de Lévy qui soient des martingales continues sont du type cB_t pour un certain $c \in \mathbb{R}$. En particulier, M est un P.A.I.S. si et seulement si $f \equiv c$ pour un certain $c \in \mathbb{R}$.

Rappelons aussi, conséquence immédiate de la formule d'isométrie, que si $f, g \in L_{loc}^2(\mathbb{R}^+, \mathbb{R})$, alors

$$\mathbb{E} \left[\left(\int_0^t f(u) dB_u \right) \left(\int_0^s g(u) dB_u \right) \right] = \int_0^{t \wedge s} f(u) g(u) du.$$

La formule suivante, très utile en pratique, montre enfin qu'on aurait pu définir $I_T(f)$ à l'aide d'une intégrale déterministe (en travaillant ω par ω), pourvu que f soit dérivable :

Théorème 4.5 [Formule d'intégration par parties] *Soit $f \in \mathcal{C}^1(\mathbb{R}^+, \mathbb{R})$ et soit $\{B_t, t \geq 0\}$ un mouvement brownien réel. Alors*

$$I_t(f) = f(t)B_t - \int_0^t f'(s)B_s ds$$

pour tout $t \geq 0$.

Preuve : Fixons $t \geq 0$. Par propriété des espaces gaussiens, il suffit de vérifier que

$$\mathbb{E}[B_u I_t(f)] = \mathbb{E} \left[B_u \left(f(t)B_t - \int_0^t f'(s)B_s ds \right) \right]$$

pour tout $u \leq t$. En utilisant la formule d'intégration par parties classique, on calcule

$$\begin{aligned}
 \mathbb{E} \left[B_u \left(f(t)B_t - \int_0^t f'(s)B_s ds \right) \right] &= u f(t) - \int_0^t f'(s)(s \wedge u) ds \\
 &= u f(t) - \left(u \int_u^t f'(s) ds + \int_0^u f'(s)s ds \right) \\
 &= u f(u) - \int_0^u f'(s)s ds \\
 &= \int_0^u f(s) ds = \mathbb{E}[B_u I_t(f)].
 \end{aligned}$$

□

4.2 Application de l'intégrale de Wiener

4.2.1 Le processus d'Ornstein-Uhlenbeck

C'est le processus solution de l'équation suivante, dite *équation de Langevin* :

$$X_t = X_0 - a \int_0^t X_s ds + \sigma B_t, \quad (4.3)$$

où $a, \sigma \in \mathbb{R}$, X_0 est une certaine variable aléatoire (très souvent une constante en finance) et B un Brownien indépendant de X_0 . On écrit l'équation précédente sous la forme

$$dX_t = -aX_t dt + \sigma dB_t.$$

Le théorème suivant montre que la solution de cette équation est explicite en termes d'une intégrale de Wiener :

Théorème 4.6 *L'équation de Langevin a pour unique solution*

$$X_t = e^{-at} \left(X_0 + \sigma \int_0^t e^{as} dB_s \right). \quad (4.4)$$

Preuve : Notons Y_t le second membre de (4.4). En utilisant le théorème 4.5, on transforme l'intégrale

$$\int_0^t e^{as} dB_s = e^{at} B_t - a \int_0^t e^{as} B_s ds$$

et l'on en déduit que

$$Y_t = e^{-at} \left(X_0 - a\sigma \int_0^t e^{as} B_s ds \right) + \sigma B_t.$$

D'autre part, en appliquant le théorème de Fubini,

$$\begin{aligned}
 a \int_0^t Y_s ds &= aX_0 \int_0^t e^{-as} ds + a\sigma \int_0^t B_s ds - a^2\sigma \int_0^t e^{-as} \left(\int_0^s e^{au} B_u du \right) ds \\
 &= X_0(1 - e^{-at}) + a\sigma \int_0^t B_s ds - a^2\sigma \int_0^t e^{au} B_u \left(\int_u^t e^{-as} ds \right) du \\
 &= X_0(1 - e^{-at}) + a\sigma \int_0^t B_s ds - a^2\sigma \int_0^t B_u \left(\int_0^{t-u} e^{-as} ds \right) du \\
 &= X_0(1 - e^{-at}) + a\sigma \int_0^t B_s ds - a\sigma \int_0^t B_u du + a\sigma \int_0^t e^{a(u-t)} B_u du \\
 &= X_0(1 - e^{-at}) + a\sigma \int_0^t e^{a(u-t)} B_u du \\
 &= X_0 - e^{-at} \left(X_0 - a\sigma \int_0^t e^{au} B_u du \right) = X_0 - Y_t + \sigma B_t,
 \end{aligned}$$

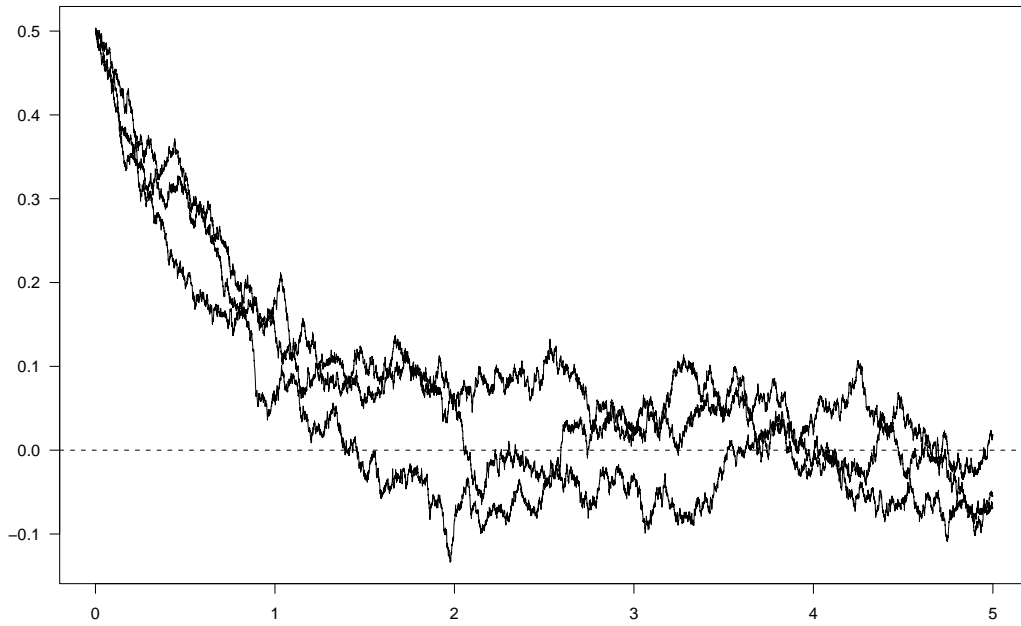


FIG. 4.1 – Processus O.U.

de sorte que

$$Y_t = X_0 - a \int_0^t Y_s ds + \sigma B_t$$

ce qui est bien l'équation (4.3). Remarquons maintenant que si X^1 et X^2 sont deux solutions de (4.3), alors $Z = X^1 - X^2$ vérifie l'EDO

$$Z_t = -a \int_0^t Z_s ds$$

dont l'unique solution est zéro (cette équation s'écrit $Z'(t) = -aZ(t)dt$, $Z(0) = 0$). Ainsi $X^1 \equiv X^2$ et (4.3) a bien une unique solution qui, d'après ce qui précède, est le processus Y .

□

Remarque 4.7 On aurait pu également poser $Y_t = e^{at}X_t$ et appliquer la formule d'intégration par parties (en admettant qu'elle est valable pour le processus X , ce que nous justifierons plus tard) pour trouver

$$dY_t = e^{at}dX_t + ae^{at}X_tdt = e^{at}\sigma dB_t,$$

équation dont la solution est

$$Y_t = X_0 + \sigma \int_0^t e^{as}dB_s,$$

de sorte que

$$X_t = e^{-at} \left(X_0 + \sigma \int_0^t e^{as}dB_s \right).$$

La propriété de Markov du processus d'Ornstein-Uhlenbeck est établie dans la proposition suivante. Elle est en fait valable dans un cadre beaucoup plus général que nous verrons au chapitre suivant.

Proposition 4.8 *L'unique solution de (4.3) est un processus de Markov simple : pour tout $s, t \geq 0$ et toute fonction réelle mesurable f ,*

$$\mathbb{E}[f(X_{t+s}) \mid \mathcal{F}_s^B] = \mathbb{E}[f(X_{t+s}) \mid X_s] = \Phi(t, X_s)$$

où l'on a noté $\Phi(t, x) = \mathbb{E}[f(X_t) \mid X_0 = x]$.

Preuve : On écrit

$$\begin{aligned} X_{t+s} &= e^{-a(t+s)} \left(X_0 + \sigma \int_0^{t+s} e^{au} dB_u \right) \\ &= e^{-at} X_s + \sigma e^{-a(t+s)} \int_s^{t+s} e^{au} dB_u \\ &= e^{-at} \left(X_s + \sigma \int_0^t e^{au} d\tilde{B}_u \right) \end{aligned}$$

où le processus \tilde{B} défini par $\tilde{B}_u = B_{s+u} - B_s$ est un Brownien indépendant de \mathcal{F}_s^B et donc aussi de $X_s \in \mathcal{F}_s^B$. En utilisant la propriété que si X et Y sont deux v.a. telles que $X \in \mathcal{F}_s^B$ et $Y \perp \mathcal{F}_s^B$ alors $\mathbb{E}[F(X, Y) \mid \mathcal{F}_s^B] = \mathbb{E}[F(x, Y)]_{x=X}$, on en déduit que

$$\mathbb{E}[f(X_{t+s}) \mid \mathcal{F}_s^B] = \mathbb{E}[f(X_{t+s}) \mid X_s] = \Phi(t, X_s)$$

pour toute fonction réelle mesurable f . □

La proposition suivante considère la situation où X_0 est une gaussienne indépendante du Brownien porteur.

Proposition 4.9 *Supposons $X_0 \sim \mathcal{N}(m, \sigma_0^2)$ indépendante de B . Alors la solution de (4.3) est un processus gaussien de fonction espérance*

$$\mathbb{E}[X_t] = me^{-at}$$

et de fonction de covariance

$$\text{cov}[X_s, X_t] = e^{-a(t+s)} \left(\sigma_0^2 + \frac{\sigma^2}{2a} (e^{2a(s \wedge t)} - 1) \right).$$

Preuve : On sait que les deux processus

$$t \mapsto e^{-at} X_0 \quad \text{et} \quad t \mapsto \sigma \int_0^t e^{-a(t-s)} dB_s$$

sont gaussiens et indépendants. Donc X , qui est leur somme, est un processus gaussien. Comme l'intégrale de Wiener est d'espérance nulle, on a

$$\mathbb{E}[X_t] = e^{-at} \mathbb{E}[X_0] = me^{-at}.$$

Enfin, par indépendance de X_0 et de B , et par isométrie de l'intégrale de Wiener,

$$\begin{aligned} \text{cov}[X_s, X_t] &= \text{cov}[X_0 e^{-as}, X_0 e^{-at}] + \text{cov} \left[\sigma \int_0^s e^{a(u-s)} dB_u, \sigma \int_0^t e^{a(u-t)} dB_u \right] \\ &= e^{-a(t+s)} \left(\text{Var}[X_0] + \sigma^2 \int_0^{s \wedge t} e^{2au} du \right) \\ &= e^{-a(t+s)} \left(\sigma_0^2 + \frac{\sigma^2}{2a} (e^{2a(s \wedge t)} - 1) \right). \end{aligned}$$

□

Remarque 4.10 Si X_0 est une constante (i.e. $\sigma_0^2 = 0$), on a

$$\text{cov}[X_s, X_t] = \frac{\sigma^2}{2a} (e^{-a|t-s|} - e^{-a(t+s)}) \quad \text{et} \quad \text{Var}[X_t] = \frac{\sigma^2}{2a} (1 - e^{-2at}).$$

4.2.2 Le processus de Vařicek

C'est une généralisation du modèle précédent, avec l'ajout d'une constante dans le drift :

$$dY_t = a(b - Y_t)dt + \sigma dB_t. \quad (4.5)$$

Cette équation est utilisée pour étudier l'évolution des taux d'intérêt dans le modèle financier dit de Vařicek. Un tel modèle est qualifié de mean-reverting (retour à la moyenne) car, pour a et b positifs, le processus Y_t tend vers b (en un sens à préciser), quand t tend vers l'infini. Ce comportement se justifie intuitivement : si $b > Y_t$, le drift $a(b - Y_t)$ est positif et le processus Y est, "en moyenne" croissant. Quand $Y_t = b$, le processus est en moyenne constant et si $b < Y_t$, le drift $a(b - Y_t)$ est négatif et le processus Y est, "en moyenne" décroissant. Si l'on pose $X_t = Y_t - b$, on voit que le processus X est solution de l'équation de Langevin, et l'on en déduit que (4.5) a une solution explicite donnée par

$$Y_t = (Y_0 - b)e^{-at} + b + \sigma \int_0^t e^{-a(t-u)} dB_u.$$

On a des propriétés analogues à celles du processus de Ornstein-Uhlenbeck : si $X_0 \sim \mathcal{N}(m, \sigma_0^2)$ est indépendante de B , Y est un processus gaussien de fonction espérance

$$\mathbb{E}[Y_t] = me^{-at} + b(1 - e^{-at})$$

et de fonction de covariance

$$\text{cov}[Y_s, Y_t] = \text{cov}[X_s, X_t] = e^{-a(t+s)} \left(\sigma_0^2 + \frac{\sigma^2}{2a} (e^{2a(s \wedge t)} - 1) \right).$$

De plus, l'égalité

$$Y_t = (Y_s - b)e^{-a(t-s)} + b + \sigma \int_s^t e^{-a(t-u)} dB_u$$

pour tout $s \leq t$ établit le caractère *gaussien* de Y . Cette expression explicite montre aussi que conditionnellement à $\mathcal{F}_s^B \vee \sigma(Y_0)$, Y_{t+s} est une gaussienne d'espérance

$$\mathbb{E}[Y_{t+s} | \mathcal{F}_s^B] = Y_s e^{-at} + b(1 - e^{-at})$$

et de variance

$$\text{Var}[Y_{t+s} | \mathcal{F}_s^B] = \frac{\sigma^2}{2a} (1 - e^{-2at}).$$

L'intégrale du processus de Vařicek intervient en Finance pour le calcul du prix $P(t, T)$ d'un zéro-coupon de maturité T , lorsque le taux court est supposé suivre un processus de Vařicek (hypothèse assez irréaliste, principalement car ce processus prend des valeurs négatives mais qui a constitué un des premiers modèles de taux). On montrera dans le cours de finance que ce prix $P(t, T)$ est obtenu par

$$P(t, T) = \mathbb{E} \left[\exp - \int_t^T Y_u du \middle| \mathcal{F}_t^B \right].$$

Par l'équation (4.5), on a

$$\begin{aligned} Z_t \stackrel{\text{def}}{=} \int_0^t Y_s ds &= a^{-1} (Y_0 - Y_t + abt + \sigma B_t) \\ &= a^{-1} \left(Y_0 + abt + \sigma B_t - (Y_0 - b)e^{-at} - b - \sigma \int_0^t e^{-a(t-u)} dB_u \right) \\ &= a^{-1} \left(b(at - 1 - e^{-at}) + Y_0(1 - e^{-at}) + \sigma B_t - \sigma e^{-at} \int_0^t e^{au} dB_u \right). \end{aligned}$$

Ainsi, quand $Y_0 \sim \mathcal{N}(m, \sigma_0^2)$ est indépendante de B , alors Z est un processus gaussien d'espérance

$$\mathbb{E}[Z_t] = bt + \frac{(m - b)(1 - e^{-at})}{a}$$

et de fonction de covariance

$$\begin{aligned} \text{cov}[Z_s, Z_t] &= \frac{\sigma_0^2}{a^2} (1 - e^{-as})(1 - e^{-at}) \\ &+ \frac{\sigma^2}{a^2} \left(t \wedge s + \frac{e^{-a(t+s)}(e^{2a(s \wedge t)} - 1)}{2a} - \frac{(1 + e^{-a|t-s|})(1 - e^{-a(s \wedge t)})}{a} \right). \end{aligned}$$

De même, on montre que pour $t > s$, conditionnellement à \mathcal{F}_s^B , la variable $Z_t - Z_s$ est une gaussienne d'espérance

$$M(s, t) = b(t - s) + a^{-1}(Y_s - b)(1 - e^{-a(t-s)})$$

et de variance

$$V(s, t) = \frac{\sigma_0^2}{a^2} (1 - e^{-a(t-s)})^2 + \frac{\sigma^2}{a^2} \left((t - s) + \frac{(1 - e^{-2a(t-s)})}{2a} - \frac{2(1 - e^{-a(t-s)})}{a} \right).$$

On connaît la transformée de Laplace d'une v.a. gaussienne, on en déduit alors la valeur du zéro-coupon :

$$\begin{aligned} P(t, T) &= \mathbb{E} \left[\exp - \int_t^T Y_u du \middle| \mathcal{F}_t^B \right] \\ &= \exp \left[M(t, T) + \frac{1}{2} V(t, T) \right]. \\ &= \exp \left[-b(t - s) - a^{-1}(Y_s - b)(1 - e^{-a(t-s)}) + \frac{1}{2} V(t, T) \right]. \end{aligned}$$

On calculera plus tard $dP(t, T)$ pour obtenir la dynamique du prix.

4.3 L'intégrale stochastique générale

On cherche maintenant à définir la v.a.

$$\int_0^t \theta_s dB_s$$

quand $\{\theta_s, s \geq 0\}$ est un processus stochastique. Le caractère aléatoire de θ va exiger des conditions supplémentaires par rapport au cas de l'intégrale de Wiener. On note $\{\mathcal{F}_t^B, t \geq 0\}$ la filtration naturelle du mouvement brownien B .

Définition 4.11 On dit que $\{\theta_t, t \geq 0\}$ est un bon processus ¹ s'il est (\mathcal{F}_t^B) -adapté, càglàd, et si

$$\mathbb{E} \left[\int_0^t \theta_s^2 ds \right] < +\infty$$

pour tout $t > 0$.

Comme dans le cas de l'intégrale de Wiener, la construction de $I_t(\theta)$ se fait par discrétisation :

4.3.1 Cas des processus étagés

Ce sont les processus du type

$$\theta_t^n = \sum_{i=0}^{p_n} \theta_i \mathbb{1}_{]t_i, t_{i+1}]}(t)$$

où $p_n \in \mathbb{N}$, $0 = t_0 \leq t_1 \leq \dots \leq t_{p_n}$ et $\theta_i \in L^2(\Omega, \mathcal{F}_{t_i}, \mathbb{P})$ pour tout $i = 0 \dots p_n$. On voit immédiatement que θ^n est un bon processus. On définit alors

$$I_t(\theta^n) = \int_0^t \theta_s^n dB_s = \sum_{i=0}^{p_n} \theta_i (B_{t_{i+1}} - B_{t_i})$$

¹Cette dénomination n'est pas standard

et on vérifie que pour $i \neq j$,

$$\mathbb{E}(\theta_i(B_{t_{i+1}} - B_{t_i})\theta_j(B_{t_{j+1}} - B_{t_j})) = 0$$

et que

$$\mathbb{E}[I_t(\theta^n)] = 0 \quad \text{et} \quad \text{Var}[I_t(\theta^n)] = \mathbb{E}\left[\int_0^t (\theta_s^n)^2 ds\right].$$

Cependant, on prendra garde que par le caractère aléatoire de θ^n , la variable $I_t(\theta^n)$ n'est pas une Gaussienne en général.

4.3.2 Cas général

Le principe est le même que pour l'intégrale de Wiener, mais les outils mathématiques sous-jacents plus compliqués que les lemmes hilbertien et gaussien du paragraphe précédent. Nous passerons les détails. Si θ est un bon processus, on montre d'abord qu'il existe $\{\theta^n, n \geq 0\}$ suite de processus étagés telle que

$$\mathbb{E}\left[\int_0^t (\theta_s - \theta_s^n)^2 ds\right] \rightarrow 0$$

quand $n \uparrow +\infty$, puis que pour tout $t > 0$ il existe une v.a. $I_t(\theta)$ de carré intégrable telle que

$$\mathbb{E}[|I_t(\theta) - I_t(\theta^n)|^2] \rightarrow 0$$

quand $n \uparrow +\infty$, avec $I_t(\theta^n)$ défini comme au paragraphe précédent. On pose alors naturellement

$$I_t(\theta) = \int_0^t \theta_s dB_s$$

pour tout $t \geq 0$. Par indépendance, on remarque d'abord que

$$\mathbb{E}[I_t(\theta^n)] = \sum_{i=0}^{p_n} \mathbb{E}[\theta_i] \mathbb{E}[B_{t_{i+1}} - B_{t_i}] = 0,$$

de sorte, en passant à la limite, que

$$\mathbb{E}[I_t(\theta)] = 0.$$

De même, on obtient

$$\begin{aligned} \text{Var}[I_t(\theta)] &= \lim_{n \uparrow +\infty} \text{Var}[I_t(\theta^n)] = \lim_{n \uparrow +\infty} \mathbb{E}[I_t(\theta^n)^2] \\ &= \lim_{n \uparrow +\infty} \mathbb{E}\left[\sum_{i=0}^{p_n} \theta_i^2 (t_{i+1} - t_i)\right] = \mathbb{E}\left[\int_0^t \theta_s^2 ds\right]. \end{aligned}$$

Insistons à nouveau sur le point que $I_t(\theta)$ n'est pas gaussienne en général, sauf lorsque θ est déterministe. En revanche, d'autres propriétés de l'intégrale de Wiener sont conservées :

Linéarité : Pour tous $t \geq 0$, $a_1, a_2 \in \mathbb{R}$ et θ^1, θ^2 bons processus, on a

$$I_t(a_1\theta^1 + a_2\theta^2) = a_1 I_t(\theta^1) + a_2 I_t(\theta^2).$$

Propriétés de martingale : Pour tout bon processus θ , les processus

$$t \mapsto I_t(\theta) \quad \text{et} \quad t \mapsto I_t(\theta)^2 - \int_0^t \theta_s^2 ds$$

sont des (\mathcal{F}_t^B) -martingales continues. On a donc, pour tout $s \leq t$

$$\mathbb{E}[I_t(\theta) \mid \mathcal{F}_s^B] = I_s(\theta)$$

(soit $\mathbb{E} [I_t(\theta) - I_s(\theta) \mid \mathcal{F}_s^B] = 0$), et on montre également que

$$\mathbb{E} [(I_t(\theta) - I_s(\theta))^2 \mid \mathcal{F}_s^B] = \mathbb{E} [(I_t(\theta))^2 - (I_s(\theta))^2 \mid \mathcal{F}_s^B] = E \left[\int_s^t \theta_u^2 du \mid \mathcal{F}_s^B \right].$$

En conséquence du théorème de Doob, on voit aussi que pour tout (\mathcal{F}_t^B) -temps d'arrêt τ et tout θ bon processus tel que

$$\mathbb{E} \left[\int_0^\tau \theta_s^2 ds \right] < \infty,$$

on a

$$\mathbb{E} [I_\tau(\theta)] = 0 \quad \text{et} \quad \mathbb{E} [I_\tau^2(\theta)] = \mathbb{E} \left[\int_0^\tau \theta_s^2 ds \right].$$

Enfin, on peut aussi appliquer les théorèmes 2.19 et 2.20. Par exemple, en prenant $p = q = 2$ dans le théorème 2.20, on obtient

$$\mathbb{E} \left[\left(\sup_{s \leq t} I_s(\theta) \right)^2 \right] \leq 4 \mathbb{E} [(I_t(\theta))^2] = 4 \int_0^t \mathbb{E} [\theta_u^2] du.$$

Propriété d'isométrie : Pour tous bons processus φ, θ et tout $s, t \geq 0$, on a

$$\mathbb{E} [I_s(\varphi) I_t(\theta)] = \mathbb{E} \left[\int_0^{s \wedge t} \theta_u \varphi_u du \right].$$

De plus, le processus

$$I_t(\theta) I_t(\varphi) - \int_0^t \theta_u \varphi_u du$$

est une (\mathcal{F}_t^B) -martingale.

La proposition suivante est un premier calcul sur les intégrales stochastiques, qui montre que la formule d'intégration par parties n'admet pas d'extension triviale au cas d'intégrale d'Itô. On y voit apparaître un terme déterministe à cause de la variation quadratique non nulle du Brownien. Ce terme "du deuxième ordre" se retrouvera dans la formule d'Itô.

Proposition 4.12 *Pour tout $t \geq 0$ on a*

$$\int_0^t B_s dB_s = \frac{1}{2} (B_t^2 - t).$$

Preuve : En prenant pour $\{t_i^n, i = 0 \dots 2^n\}$ la subdivision régulière de $[0, t]$, on écrit

$$\begin{aligned} \int_0^t B_s dB_s &= \lim_{n \rightarrow +\infty} \sum_{i=0}^{2^n-1} B_{t_i^n} (B_{t_{i+1}^n} - B_{t_i^n}) \\ &= \lim_{n \rightarrow +\infty} \left(\frac{1}{2} \left(\sum_{i=0}^{2^n-1} (B_{t_{i+1}^n}^2 - B_{t_i^n}^2 - (B_{t_{i+1}^n} - B_{t_i^n})^2) \right) \right) \\ &= \frac{1}{2} \left(B_t^2 - \lim_{n \rightarrow +\infty} \left(\sum_{i=0}^{2^n-1} (B_{t_{i+1}^n} - B_{t_i^n})^2 \right) \right) \\ &= \frac{1}{2} (B_t^2 - t). \end{aligned}$$

□

Dans la définition d'un bon processus, la condition d'intégrabilité

$$\mathbb{E} \left[\int_0^t \theta_s^2 ds \right] < +\infty$$

est parfois trop exigeante dans la pratique. Il est possible de définir $I_t(\theta)$ sous la seule condition

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

Cependant, $t \mapsto I_t(\theta)$ n'est plus nécessairement une martingale, et en particulier $\mathbb{E}[I_t(\theta)]$ peut être non nul. Pour définir $I_t(\theta)$, il faut introduire la notion importante de *martingale locale* :

Définition 4.13 Soit $\{\mathcal{F}_t, t \geq 0\}$ une filtration et $\{X_t, t \geq 0\}$ un processus (\mathcal{F}_t) -adapté. On dit que X est une (\mathcal{F}_t) -martingale locale s'il existe une suite $\{\tau_n, n \geq 0\}$ de (\mathcal{F}_t) -temps d'arrêt telle que

$$\mathbb{P}[\tau_n \rightarrow +\infty] = 1$$

et le processus $X^n : t \mapsto X_{t \wedge \tau_n}$ est une martingale pour tout $n \geq 0$.

Par le lemme de Fatou, on voit qu'une martingale locale positive est une surmartingale. Par le théorème de convergence dominée, une martingale locale uniformément intégrable est une vraie martingale. Fixons maintenant B un mouvement brownien.

Définition 4.14 On dit que $\{\theta_t, t \geq 0\}$ est un bon processus local s'il est càglàd, (\mathcal{F}_t^B) -adapté, et si

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

pour tout $t > 0$.

Soit θ un bon processus local. On pose

$$\tau_n = \inf \left\{ t > 0 / \int_0^t \theta_s^2 ds = n \right\}.$$

Comme

$$\{\tau_n > t\} = \left\{ \int_0^t \theta_s^2 ds < n \right\}$$

pour tout $n \in \mathbb{N}, t \geq 0$, on voit que τ_n est un (\mathcal{F}_t^B) -temps d'arrêt pour tout $n \in \mathbb{N}$ par adaptation de θ . De plus, l'hypothèse d'intégrabilité sur θ entraîne facilement que $\tau_n \rightarrow +\infty$ p.s. Enfin, par construction on a

$$\mathbb{E} \left[\int_0^{t \wedge \tau_n} \theta_s^2 ds \right] \leq n < +\infty.$$

Ainsi, par le paragraphe précédent, on peut définir $I_{t \wedge \tau_n}(\theta)$ qui est une martingale. Comme p.s. $\tau_n \rightarrow +\infty$, on peut définir $I_t(\theta)$ pour tout $t > 0$, qui est une martingale locale. De même, en prenant la même suite de temps d'arrêt, on montre que le processus

$$I_t(\theta)^2 - \int_0^t \theta_s^2 ds$$

est une martingale locale.

Exemple 4.15 Soit $\alpha \in \mathbb{R}$. On veut savoir quand l'intégrale stochastique

$$I_t(B^\alpha) = \int_0^t B_s^\alpha dB_s$$

a un sens, et quand le processus associé est une martingale. Remarquons déjà que

$$\begin{aligned} \mathbb{E} \left[\int_0^t (B_s^\alpha)^2 ds \right] &= \int_0^t \mathbb{E} [B_s^{2\alpha}] ds \\ &= \int_0^t s^\alpha \mathbb{E} [B_1^{2\alpha}] ds = \mathbb{E} [B_1^{2\alpha}] \int_0^t s^\alpha ds. \end{aligned}$$

Dans le terme de droite, l'intégrale est finie si et seulement si $\alpha > -1$. Donc, pour que $t \mapsto I_t(B^\alpha)$ soit bien défini et une martingale, il faut et il suffit que $\alpha > -1$ et $\mathbb{E}[B_1^{2\alpha}] < +\infty$. Or

$$\mathbb{E}[B_1^{2\alpha}] = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x^{2\alpha} e^{-x^2/2} dx$$

et cette intégrale est finie si et seulement si $\alpha > -1/2$ (puisque le seul problème d'intégrabilité est en 0). On en déduit finalement que

$$t \mapsto \int_0^t B_s^\alpha dB_s \text{ est une martingale } \iff \alpha > -1/2.$$

Quand $-1 < s \leq -1/2$, on cherche alors à savoir quand $I_t(B^\alpha)$ est malgré tout défini comme martingale locale. On utilise alors les théorèmes 3.7 et 3.8, qui permettent de comparer le comportement du Brownien au voisinage de 0 avec celui de $t \mapsto t^{1/2}$. On en déduit que

$$\int_0^t B_s^{2\alpha} ds < +\infty \text{ p.s. } \iff \alpha > -1.$$

D'où, d'après ce qui précède,

$$t \mapsto \int_0^t B_s^\alpha dB_s \text{ est défini et une martingale locale } \iff \alpha > -1.$$

Nous terminons sur ce paragraphe en revenant sur le *crochet* de deux martingales locales, reprenant les considérations du chapitre 2 après la décomposition de Doob-Meyer. On a vu que si Z est une (\mathcal{F}_t) -martingale continue de carré intégrable, alors $\langle Z \rangle$ est l'unique processus croissant continu (\mathcal{F}_t) -adapté tel que $t \mapsto Z_t^2 - \langle Z \rangle_t$ soit une (\mathcal{F}_t) -martingale. Quitte à utiliser les temps d'arrêt

$$\tau_n = \inf \{t \geq 0 / Z_t^2 = n\},$$

on peut maintenant étendre cette définition aux martingales locales : si Z est une martingale locale, $\langle Z \rangle$ est l'unique processus croissant continu (\mathcal{F}_t) -adapté tel que $t \mapsto Z_t^2 - \langle Z \rangle_t$ soit une (\mathcal{F}_t) -martingale locale. Par polarité, on peut définir le crochet de deux (\mathcal{F}_t) -martingales locales M et N en écrivant

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

Le crochet $\langle M, N \rangle$ est aussi l'unique processus à variation finie tel que le processus $MN - \langle M, N \rangle$ soit une martingale locale. On a alors la

Proposition 4.16 *Soit M une martingale locale continue. Alors M est une martingale L^2 si et seulement si $\mathbb{E}[\langle M \rangle_t] < +\infty$ pour tout $t \geq 0$.*

Enfin, la proposition suivante donne enfin de $\langle M, N \rangle$ une importante construction trajectorielle :

Proposition 4.17 *Soient M et N deux martingales locales continues. Alors p.s. pour tout $t \geq 0$,*

$$\langle M, N \rangle_t = \lim_{n \rightarrow +\infty} \sum_{i=1}^{2^n} (M_{t_i^n} - M_{t_{i-1}^n})(N_{t_i^n} - N_{t_{i-1}^n})$$

où $\{t_i^n, i = 0 \dots 2^n\}$ désigne la subdivision régulière sur $[0, t]$.

Remarque 4.18 Par l'inégalité de Cauchy-Schwarz, le terme de droite est bien défini pour deux processus à variation quadratique finie. Cependant, on voit aussi que ce crochet sera nul dès qu'une variation quadratique est nulle, en particulier *dès qu'un des deux processus est à variation finie.*

Il résulte aussi de cette proposition le fait crucial que le crochet $\langle M, N \rangle$ reste inchangé si l'on effectue un changement de probabilité équivalente. On dit enfin que deux martingales continues sont *orthogonales* si leur crochet est nul, c'est-à-dire si leur produit est une martingale. Par exemple, deux Browniens indépendants sont des martingales orthogonales. La proposition 4.12 établit que le crochet du Brownien B est $\langle B \rangle_t = t$. On peut aussi calculer le crochet de deux Browniens corrélés avec coefficient ρ : par définition, $\langle B_1, B_2 \rangle_t = \rho t$. On peut aussi bien entendu calculer le crochet d'intégrales stochastiques générales, et les propriétés de martingale entraînent immédiatement que

$$\langle I(\theta) \rangle_t = \int_0^t \theta_s^2 ds \quad \text{et} \quad \langle I(\theta), I(\varphi) \rangle_t = \int_0^t \theta_s \varphi_s ds.$$

4.4 Processus d'Itô

Ce sont des processus écrits sous la forme

$$X_t = x + \int_0^t b_s ds + \int_0^t \sigma_s dB_s \quad (4.6)$$

où b est un processus \mathcal{F}_t^B -adapté tel que

$$\int_0^t |b_s| ds < +\infty \quad \text{p.s.}$$

pour tout $t \geq 0$, et σ un bon processus local. On utilise la notation formelle

$$\begin{aligned} dX_t &= b_t dt + \sigma_t dB_t \\ X_0 &= x. \end{aligned}$$

Le coefficient b s'appelle la *dérivée* (ou le *drift*) du processus, et σ son coefficient de diffusion. On appelle aussi le processus

$$t \mapsto x + \int_0^t b_s ds$$

la *partie à variation finie* de X , et le processus

$$t \mapsto \int_0^t \sigma_s dB_s$$

la *partie martingale* de X (c'est a priori une martingale locale). Comme, d'après la proposition 2.25, une martingale locale à variation finie est un processus constant, on en déduit que la décomposition (4.6) du processus X est unique, au sens où si X admet une autre décomposition

$$X_t = x + \int_0^t \tilde{b}_s ds + \int_0^t \tilde{\sigma}_s dB_s,$$

alors $b \equiv \tilde{b}$ et $\sigma \equiv \tilde{\sigma}$. En particulier, X sous la forme (4.6) est une martingale locale si et seulement si $b \equiv 0$. En fait, cette représentation des martingales locales dans une filtration Brownienne est caractéristique, indépendamment de ce que le processus soit *a priori* un processus d'Itô :

Théorème 4.19 [Théorème de représentation des martingales locales] *Soit B un mouvement brownien et M une (\mathcal{F}_t^B) -martingale locale continue. Alors il existe $x \in \mathbb{R}$ et θ bon processus local tel que*

$$M_t = x + \int_0^t \theta_s dB_s.$$

Ce théorème est extrêmement important en Finance et est lié à la complétion d'un marché financier. L'hypothèse fondamentale est que M est une martingale *par rapport à la filtration naturelle du Brownien*. Signalons enfin que dans certains cas, la *formule de Clark-Ocone* - voir [13] - donne une représentation explicite du processus θ . Si X^1 et X^2 sont deux processus d'Itô de décomposition

$$X_t^i = x + \int_0^t b_s^i ds + \int_0^t \sigma_s^i dB_s$$

pour $i = 1, 2$, leur crochet est par définition le crochet de leurs parties martingales. Autrement dit

$$\langle X^1, X^2 \rangle = \langle I(\sigma^1), I(\sigma^2) \rangle.$$

Attention, à cause de la partie à variation finie, le processus $t \mapsto X_t^1 X_t^2 - \langle X^1, X^2 \rangle_t$ n'est pas une martingale locale en général. En revanche, comme $X^i - I(\sigma^i)$ est un processus à variation finie, on a toujours

$$\langle X^1, X^2 \rangle_t = \lim_{n \rightarrow +\infty} \sum_{i=1}^{2^n} (X_{t_i^n}^1 - X_{t_{i-1}^n}^1)(X_{t_i^n}^2 - X_{t_{i-1}^n}^2)$$

du fait de la remarque 4.18.

4.5 Formule d'Itô

Dans ce paragraphe, on se donne un processus d'Itô réel X de décomposition (4.6) et une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ suffisamment régulière. La formule d'Itô vise à donner une formule de changement de variable pour le processus $f(X_t)$ qui sera un processus d'Itô. Imaginons un instant que $\sigma \equiv 0$ et que b soit \mathcal{C}^0 . Alors X est un processus \mathcal{C}^1 et $(f \circ X)' = (f' \circ X)X'$. D'où, par la formule de changement de variables classique,

$$\begin{aligned} f(X_t) &= f(X_0) + \int_0^t (f \circ X)'(s) ds \\ &= f(x) + \int_0^t f'(X_s)X'_s ds = f(x) + \int_0^t f'(X_s) dX_s. \end{aligned}$$

Cette formule garderait encore un sens quand $\sigma \neq 0$, en posant

$$dX_s = b_s ds + \sigma_s dB_s.$$

Mais en fait, cette formule du 1er ordre n'est plus vraie quand $\sigma \neq 0$, à cause du caractère quadratique de la partie martingale de X . On a une formule du 2ème ordre :

Théorème 4.20 [Première formule d'Itô] *Supposons f de classe C^2 . Alors*

$$f(X_t) = f(x) + \int_0^t f'(X_s) dX_s + \frac{1}{2} \int_0^t f''(X_s) \sigma_s^2 ds.$$

Si f est à dérivées bornées, le processus $f(X_t) - \int_0^t f'(X_s) b_s ds - \frac{1}{2} \int_0^t f''(X_s) \sigma_s^2 ds$ est une martingale.

Cette formule s'écrit sous forme condensée

$$\begin{aligned} df(X_t) &= f'(X_t) dX_t + \frac{1}{2} f''(X_t) \sigma_t^2 dt \\ &= f'(X_t) b_t dt + \frac{1}{2} f''(X_t) \sigma_t^2 dt + f'(X_t) \sigma_t dB_t \\ &= f'(X_t) b_t dt + \frac{1}{2} f''(X_t) d\langle X \rangle_t + f'(X_t) \sigma_t dB_t. \end{aligned}$$

On utilise souvent (dans les articles de finance écrits par des non-mathématiciens) la notation

$$df(X_t) = f'(X_t) dX_t + \frac{1}{2} f''(X_t) dX_t \cdot dX_t$$

avec la table de multiplication

	dt	dB_t
dt	0	0
dB_t	0	dt

En particulier, $t \mapsto f(X_t)$ est un processus d'Itô de dérive

$$\int_0^t \left(f'(X_s) b_s + \frac{1}{2} f''(X_s) \sigma_s^2 \right) ds$$

et de partie martingale

$$\int_0^t f'(X_s) \sigma_s dB_s.$$

Quand les dérivées sont bornées, l'intégrale stochastique apparaissant dans la formule est une vraie martingale, et on en déduit :

$$\begin{aligned} \mathbb{E}[f(X_t)] &= \mathbb{E}[f(X_0)] + \mathbb{E} \left[\int_0^t \left(f'(X_s) b_s + \frac{1}{2} f''(X_s) \sigma_s^2 \right) ds \right] \\ &= \mathbb{E}[f(X_0)] + \int_0^t \mathbb{E} \left[f'(X_s) b_s + \frac{1}{2} f''(X_s) \sigma_s^2 \right] ds. \end{aligned}$$

En faisant $X_t = B_t$, on retrouve ainsi le premier résultat du paragraphe 3.4. On peut également calculer de la même façon des espérances conditionnelles :

$$\begin{aligned}\mathbb{E}[f(X_t) | \mathcal{F}_s^B] &= f(X_s) + \mathbb{E}\left[\int_s^t \left(f'(X_u)b_u + \frac{1}{2}f''(X_u)\sigma_u^2\right) du \middle| \mathcal{F}_s^B\right] \\ &= f(X_s) + \int_s^t \mathbb{E}\left[f'(X_u)b_u + \frac{1}{2}f''(X_u)\sigma_u^2 \middle| \mathcal{F}_s^B\right] du.\end{aligned}$$

La deuxième formule d'Itô fait intervenir le temps en première variable. Elle sera très utile pour étudier des EDS inhomogènes.

Théorème 4.21 [Deuxième formule d'Itô] *Soit f une fonction définie sur $\mathbb{R}_+ \times \mathbb{R}$ de classe \mathcal{C}^1 par rapport à t , de classe \mathcal{C}^2 par rapport à x . On a*

$$f(t, X_t) = f(0, X_0) + \int_0^t f'_t(s, X_s) ds + \int_0^t f'_x(s, X_s) dX_s + \frac{1}{2} \int_0^t f''_{xx}(s, X_s) \sigma_s^2 ds.$$

Comme précédemment, on peut écrire cette formule sous forme différentielle :

$$\begin{aligned}df(t, X_t) &= \left(f'_t(t, X_t) + \frac{1}{2}f''_{xx}(t, X_t)\sigma_t^2\right) dt + f'_x(t, X_t) dX_t \\ &= f'_t(t, X_t) dt + f'_x(t, X_t) dX_t + \frac{1}{2}f''_{xx}(t, X_t) d\langle X \rangle_t.\end{aligned}$$

Exemple 4.22 Le mouvement brownien géométrique, ou processus log-normal est défini par l'équation

$$X_t = x + \mu \int_0^t X_s ds + \sigma \int_0^t X_s dB_s$$

avec $\mu, \sigma \in \mathbb{R}$. Si on pose $Y_t = e^{-\mu t} X_t$ pour tout $t \geq 0$, la deuxième formule d'Itô donne

$$dY_t = \sigma e^{-\mu t} X_t dB_t = \sigma Y_t dB_t$$

Nous verrons au chapitre suivant dans un cadre général (c'est aussi une conséquence de la première formule d'Itô) que Y s'écrit

$$Y_t = x \exp[\sigma B_t - \sigma^2 t/2].$$

On en déduit que

$$X_t = x \exp[\mu t + \sigma B_t - \sigma^2 t/2].$$

On peut aussi considérer le cas où μ et σ sont des fonctions déterministes :

$$X_t = x + \int_0^t \mu(s) X_s ds + \int_0^t \sigma(s) X_s dB_s$$

On dit alors que X est un Brownien géométrique à coefficients déterministes. Toujours par la deuxième formule d'Itô, on montre que le processus

$$t \mapsto X_t \exp\left[-\int_0^t \mu(s) ds\right]$$

est une martingale locale. C'est en fait une vraie martingale et

$$X_t = X_0 \exp\left[\int_0^t \mu(s) ds + \int_0^t \sigma(s) ds - \frac{1}{2} \int_0^t \sigma^2(s) ds\right]$$

Enfin, la troisième formule d'Itô permet de traiter des processus bivariés :

Théorème 4.23 [Troisième formule d'Itô] Soient X^1 et X^2 deux processus d'Itô issus de x_1 (resp. de x_2) de coefficient de dérive b^1 (resp. b^2), de coefficient de diffusion σ^1 (resp. σ^2) et portés respectivement par deux Browniens B^1 et B^2 corrélés avec coefficient ρ . On suppose que b^i, σ^i sont $(\mathcal{F}_t^{B^i})$ -adaptés. Soit f une fonction de \mathbb{R}^2 dans \mathbb{R} de classe \mathcal{C}^2 à dérivées bornées. On a

$$\begin{aligned} f(X_t^1, X_t^2) &= f(x_1, x_2) + \int_0^t f'_1(X_s^1, X_s^2) dX_s^1 + \int_0^t f'_2(X_s^1, X_s^2) dX_s^2 \\ &\quad + \frac{1}{2} \int_0^t \left(f''_{11}(X_s^1, X_s^2) (\sigma_s^1)^2 + 2\rho f''_{12}(X_s^1, X_s^2) \sigma_s^1 \sigma_s^2 + f''_{22}(X_s^1, X_s^2) (\sigma_s^2)^2 \right) ds \end{aligned}$$

où f'_i désigne la dérivée par rapport à x_i et f''_{ij} la dérivée seconde par rapport à x_j puis x_i , $i, j = 1, 2$.

En différentiel, la troisième formule d'Itô peut prendre une forme sommatoire :

$$df(X_t^1, X_t^2) = \sum_{i=1,2} f'_i(X_t^1, X_t^2) dX_t^i + \frac{1}{2} \left(\sum_{i,j=1,2} c_{ij} f''_{ij}(X_t^1, X_t^2) \sigma_t^i \sigma_t^j \right) dt$$

avec $c_{ij} = 1$ si $i = j$ et $c_{ij} = \rho$ sinon. Un exemple important d'application de la troisième formule d'Itô est la *formule d'intégration par parties*, où l'on choisit la fonction $f(x, y) = xy$:

Proposition 4.24 [Formule d'intégration par parties] Avec les mêmes notations que dans le théorème 4.23, on a

$$X_t^1 X_t^2 = x_1 x_2 + \int_0^t X_s^1 dX_s^2 + \int_0^t X_s^2 dX_s^1 + \rho \int_0^t \sigma_s^1 \sigma_s^2 ds.$$

On retrouve l'expression du crochet de X_1 et X_2 :

$$\langle X_1, X_2 \rangle_t = \rho \int_0^t \sigma_s^1 \sigma_s^2 ds,$$

et la formule d'intégration par parties s'écrit

$$d(X^1 X^2)_t = X_t^1 dX_t^2 + X_t^2 dX_t^1 + d\langle X^1, X^2 \rangle_t.$$

4.6 Formule de Black & Scholes

On considère un marché financier comportant un actif dit sans risque de taux constant r et de prix $S_t^0 = e^{rt}$ (soit $dS_t^0 = rS_t^0 dt$) et un actif risqué dont le prix S vérifie

$$dS_t = b S_t dt + \sigma S_t dB_t$$

soit

$$S_t = S_0 \exp [\sigma B_t + (b - \sigma^2/2)t]$$

avec B mouvement brownien et $b, \sigma \in \mathbb{R}$. On fixe un horizon $T > 0$ et on souhaite donner le prix d'un actif financier qui versera $h(S_T)$ à la date T . Le cas d'un call Européen de maturité T et de strike K correspond au cas $h(x) = (x - K)^+$. On procède par duplication (hedging) : on forme un portefeuille et d' α_t parts de l'actif sans risque (le montant de la richesse investie dans cet actif est αe^{rt}) et de β_t parts de l'actif risqué. On va trouver un portefeuille auto-finançant (ne nécessitant pas de mise de fonds autre qu'à la date 0) de valeur terminale $h(S_T)$. La valeur de ce portefeuille à la date t est

$$V_t = \alpha_t S_t^0 + \beta_t S_t.$$

La condition d'*auto-financement* se formalise par

$$dV_t = \alpha_t dS_t^0 + \beta_t dS_t;$$

soit

$$\begin{aligned} dV_t &= \alpha_t r S_t^0 dt + \beta_t dS_t \\ &= r V_t dt + \beta_t S_t ((b - r)dt + \sigma dB_t) \end{aligned}$$

(La valeur initiale du portefeuille sera la valeur de l'actif financier. On suppose que la valeur V_t du portefeuille à la date t est une fonction déterministe du temps et de la valeur de l'actif risqué, soit $V_t = V(t, S_t)$. En utilisant la deuxième formule d'Itô, on calcule

$$\begin{aligned} dV_t &= \left(\frac{\partial V}{\partial t}(t, S_t) + b S_t \frac{\partial V}{\partial x}(t, S_t) + \frac{\sigma^2 S_t^2}{2} \frac{\partial^2 V}{\partial x^2}(t, S_t) \right) dt \\ &+ \left(\sigma S_t \frac{\partial V}{\partial x}(t, S_t) \right) dB_t. \end{aligned}$$

En utilisant et en identifiant avec la condition d'auto-financement les parties martingales on obtient

$$\sigma \beta_t S_t + \sigma S_t \frac{\partial V}{\partial x}(t, S_t) = 0 \quad \text{soit} \quad \beta_t = - \frac{\partial V}{\partial x}(t, S_t),$$

ce qui entraîne alors en identifiant les parties à variation finie

$$r S_t \frac{\partial V}{\partial x}(t, S_t) + \frac{\partial V}{\partial t}(t, S_t) + \frac{\sigma^2 S_t^2}{2} \frac{\partial^2 V}{\partial x^2}(t, S_t) - r V(t, S_t) = 0$$

avec pour condition terminale $V(T, S_T) = h(S_T)$. Comme S_t est une v.a. qui peut prendre toutes les valeurs de \mathbb{R}^+ , on en déduit que V satisfait l'EDP

$$\boxed{r x \frac{\partial V}{\partial x}(t, x) + \frac{\partial V}{\partial t}(t, x) + \frac{\sigma^2 x^2}{2} \frac{\partial^2 V}{\partial x^2}(t, x) - r V(t, x) = 0} \quad (4.7)$$

avec pour condition terminale $V(T, x) = h(x)$. On notera que le coefficient b a disparu ! Dans le cas d'un call européen $h(x) = (x - K)^+$, et pour $\sigma > 0$, cette équation se résout alors en :

$$V(t, x) = x \mathcal{N}(d_1) - K e^{-r(T-t)} \mathcal{N}(d_2)$$

où \mathcal{N} est la fonction de répartition d'une v.a. gaussienne standard :

$$\mathcal{N}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du,$$

et avec les notations

$$d_1 = \frac{1}{2\sigma\sqrt{T-t}} \left(\ln \left(x e^{r(T-t)} / K \right) + \frac{1}{2} \sigma^2 (T-t) \right) \quad \text{et} \quad d_2 = d_1 - \sigma\sqrt{T-t}.$$

La quantité

$$\frac{\partial C}{\partial x}(t, S_t) = \mathcal{N}(d_1)$$

qui représente le nombre de parts de l'actif sous jacent utilisées pour répliquer l'option s'appelle le Delta de l'option et représente aussi la sensibilité du prix de l'option par rapport au prix du sous jacent. Le couple $(V(t, S_t) - \beta_t S_t, \beta_t)$ représente le portefeuille de couverture.

Remarque 4.25 Comme conséquence de la formule d'Itô appliquée aux EDS, on verra plus tard une formule probabiliste pour le prix du call :

$$C(t, S_t) = e^{r(t-T)} \mathbb{E}[(S_T - K)^+ | \mathcal{F}_t]$$

où dans le calcul de l'espérance on considère que S a pour dynamique

$$dS_t = r S_t dt + \sigma S_t dW_t.$$

Cette formule est fondamentale en Finance, et fait intervenir un changement de probabilité.

Chapitre 5

Equations différentielles stochastiques

Les équations différentielles stochastiques (EDS) sont les équations qui régissent l'évolution de la plupart des prix des actifs financiers, et ce chapitre en donne une courte introduction. Pour une étude plus approfondie, nous suggérons la lecture de [15] [16] [17], dans un ordre croissant de difficulté.

5.1 Définitions

Rappelons qu'une équation différentielle ordinaire (EDO) sur $\mathbb{R}^+ \times \mathbb{R}$ est un système du type

$$\begin{cases} y_0 &= y \\ y'_t &= f(t, y_t) \end{cases} \quad (5.1)$$

où $y : \mathbb{R}^+ \rightarrow \mathbb{R}$ est la fonction inconnue et $f : \mathbb{R}^+ \times \mathbb{R} \rightarrow \mathbb{R}$ est une fonction donnée. L'étude mathématique des équations différentielles ordinaires, d'une importance considérable pour les applications notamment en physique, s'est développée au début du 19-ème siècle avec la théorie des fonctions spéciales. Ces dernières peuvent être vues comme la généralisation de la fonction exponentielle, laquelle est solution de

$$\begin{cases} y_0 &= 1 \\ y'_t &= y_t. \end{cases}$$

La théorie a pris un nouvel essor à la fin du 19-ème siècle avec les travaux de Lie et de Poincaré, et continue aujourd'hui de mobiliser une communauté très importante de mathématiciens. En général, il est impossible de donner une solution explicite à une EDO. On peut cependant chercher à savoir s'il existe une solution, et si elle est unique. Un critère est le :

Théorème 5.1 [Cauchy-Lipschitz] *Supposons qu'il existe une constante $K > 0$ telle que pour tous $t \in \mathbb{R}^+$, $x, y \in \mathbb{R}$*

$$\begin{cases} |f(t, x) - f(t, y)| &\leq K|x - y| & \text{(condition de Lipschitz globale)} \\ |f(t, x)| &\leq K(1 + |x|) & \text{(condition de croissance linéaire)} \end{cases}$$

Alors l'EDO (5.1) a une solution unique définie sur \mathbb{R}^+ .

Remarquons que la condition de Lipschitz globale est assez naturelle pour que la solution soit bien définie sur tout \mathbb{R}^+ . Si on considère par exemple $y_0 = 1$ et $f(t, y) = y^2$, alors on voit facilement que l'unique solution du système (5.1) correspondant est $y_t = 1/(1 - t)$, et que cette solution "explose" en $t = 1$ avec valeur $+\infty$.

Une équation différentielle stochastique (EDS) est une perturbation de (5.1) avec un terme aléatoire modélisant un "bruit" autour du phénomène déterministe décrit par (5.1). La perturbation la plus simple est l'ajout d'un Brownien, où l'on considère

$$\begin{cases} Y_0 &= y \\ dY_t &= f(t, Y_t)dt + \sigma dB_t \end{cases} \quad (5.2)$$

soit, sous forme intégrale (la seule qui ait un sens mathématique, puisque le Brownien n'est pas dérivable) :

$$Y_t = y + \int_0^t f(s, X_s) ds + \sigma B_t$$

pour tout $t \geq 0$. On a coutume d'utiliser des majuscules pour les solutions d'EDS et des minuscules pour les solutions d'EDO. Le caractère martingalien du Brownien entraîne que pour σ petit, la trajectoire non dérivable de la solution de (5.2) va suivre en gros celle régulière et déterministe de (5.1), en oscillant aléatoirement autour. Mais quand σ est grand, la trajectoire de (5.2) n'a plus rien à voir avec celle de (5.1).

Le cadre général des EDS concerne la situation où le coefficient σ dépend aussi du temps et de la solution Y_t (en finance, on parle alors de modèle à volatilité locale). On peut encore gagner en généralité en considérant un modèle multidimensionnel. Ceci donne la

Définition 5.2 Soient $d, m \in \mathbb{N}$, $b : \mathbb{R}^+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ et $\sigma : \mathbb{R}^+ \times \mathbb{R}^d \rightarrow \mathcal{M}_{d,m}(\mathbb{R})$ deux fonctions mesurables bornées ($b(x) = \{b_i(x), 1 \leq i \leq d\}$ est un champ de vecteurs et $\sigma(x) = \{\sigma_{ij}(x), 1 \leq i \leq d, 1 \leq j \leq m\}$ un champ de matrices). Soit $x \in \mathbb{R}^d$ une condition initiale. Une solution de l'EDS

$$E_x(b, \sigma) : \begin{cases} X_0 &= x \\ dX_t &= b(t, X_t) dt + \sigma(t, X_t) dB_t \end{cases}$$

est constituée par :

- (a) Un espace de probabilité filtré $(\Omega, \mathcal{F}, \{\mathcal{F}_t, t \geq 0\}, \mathbb{P})$
- (b) Un (\mathcal{F}_t) -mouvement brownien $B = (B_1, \dots, B_m)$ à valeurs dans \mathbb{R}^m .
- (c) Un processus $X = \{X_t, t \geq 0\}$ continu \mathcal{F}_t -adapté tel que les intégrales

$$\int_0^t b(s, X_s) ds \quad \text{et} \quad \int_0^t \sigma(s, X_s) dB_s$$

aient un sens et tel que l'égalité

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

soit satisfaite pour tout $t \in \mathbb{P}$ p.s. Autrement dit, pour tout $i = 1, \dots, d$ on a

$$X_t^i = x + \int_0^t b^i(s, X_s) ds + \sum_{j=1}^m \int_0^t \sigma_{ij}(s, X_s) dB_s^j.$$

Le caractère aléatoire des EDS impose plusieurs notions d'existence et d'unicité. On dit qu'il y a

- (1) Existence d'une solution faible si $E_x(b, \sigma)$ admet une solution X .
- (2) Existence d'une solution forte si $E_x(b, \sigma)$ admet une solution X qui soit adaptée à la filtration du Brownien porteur.
- (3) Unicité faible si tous les processus X solutions de $E_x(b, \sigma)$ ont même loi.
- (4) Unicité trajectorielle si, l'espace de probabilité et le Brownien porteur étant fixés, deux solutions quelconques X et X' de $E_x(b, \sigma)$ sont indistinguables au sens où

$$\mathbb{P}[\exists t \in \mathbb{R} / X_t \neq X'_t] = 0.$$

L'exemple suivant montre que l'on peut avoir (1) et (3) (existence et unicité faible), mais ni (2) ni (4) :

Exemple 5.3 Soit $\{W_t, t \geq 0\}$ un Brownien standard. On considère le processus

$$B_t = \int_0^t \text{sgn}(W_s) dW_s$$

où sgn est la fonction définie par

$$\text{sgn}(x) = \begin{cases} 1 & \text{si } x \geq 0 \\ -1 & \text{si } x < 0. \end{cases}$$

Comme $\text{sgn}^2(x) = 1$, on remarque facilement que le processus B est bien défini et que B est une martingale continue de crochet t . Par le théorème de caractérisation de Paul Lévy, B est donc un mouvement brownien. On considère maintenant l'EDS

$$\begin{cases} X_0 &= 0 \\ dX_t &= \text{sgn}(X_t) dB_t \end{cases}$$

Par construction, on voit que W est solution de cette équation. De plus, par le même argument que précédemment, on voit que toutes les solutions de cette équation sont des mouvements browniens et sont donc égaux en loi : on a existence et unicité faibles. En revanche le point (4) n'est pas vérifié : le processus $-W$ est aussi solution de l'équation avec pour tout $t > 0$,

$$\mathbb{P}[W_t \neq -W_t] = \mathbb{P}[2W_t \neq 0] = 1,$$

de sorte que W et $-W$ sont deux processus solution qui ne sont pas indistinguables. De plus, par construction on voit intuitivement (et on peut démontrer rigoureusement) que la filtration naturelle du processus B est celle de $|X|$, laquelle est bien sûr plus petite que celle de X , puisqu'elle ne prend pas en compte le signe de X . Donc toute solution X n'est pas (\mathcal{F}_t^B) -adaptée et il n'y a pas existence de solution forte.

Le théorème suivant nous dit en revanche que (1) + (4) \implies (2) + (3) :

Théorème 5.4 [Yamada-Watanabe] *Supposons que $E_x(b, \sigma)$ admette une solution faible et que toutes ses solutions soient indistinguables. Alors $E_x(b, \sigma)$ admet une unique solution forte.*

Le théorème suivant est l'analogue du théorème de Cauchy-Lipschitz pour les EDS. Il fournit les conditions standard d'existence et d'unicité de solution forte :

Théorème 5.5 [Théorème d'existence sous conditions lipschitziennes] *Supposons que pour tout compact K de \mathbb{R}^d , il existe une constante $M_K > 0$ telle que*

(a) $|b_i(t, x) - b_i(t, y)| + |\sigma_{ij}(t, x) - \sigma_{ij}(t, y)| \leq M_K |x - y|$ pour tout $i = 1 \dots d$, $j = 1 \dots m$, $t \geq 0$, $x, y \in K$ (condition de Lipschitz locale)

et qu'il existe une constante $M > 0$ telle que

(b) $|b_i(t, x)| + |\sigma_{ij}(t, x)| \leq M(1 + |x|)$ pour tout $i = 1 \dots d$, $j = 1 \dots m$, $t \geq 0$, $x, y \in \mathbb{R}^d$ (condition de croissance linéaire),

alors il existe une unique solution forte à $E_x(b, \sigma)$, de durée de vie infinie.

La démonstration de l'existence faible repose sur une méthode de point fixe, un peu trop longue à détailler ici. L'idée est de construire une suite X^n par

$$X_t^n = x + \int_0^t b^i(s, X_s^{n-1}) ds + \sum_{j=1}^m \int_0^t \sigma_{ij}(s, X_s^{n-1}) dB_s^j,$$

et de montrer la convergence de cette suite vers une solution.

Nous allons en revanche démontrer l'unicité trajectorielle, qui suffit pour avoir existence et unicité d'une solution forte, d'après le théorème de Yamada-Watanabe. On suppose pour simplifier que b et σ sont globalement lipschitziennes en espace, avec constante M . L'argument repose sur le lemme suivant qui est extrêmement utile :

Lemme 5.6 [Gronwall] *Soit $T > 0$ et g une fonction positive mesurable bornée telle que*

$$g(t) \leq a + b \int_0^t g(s) ds$$

pour tout $t \leq T$, où a et b sont des constantes positives. Alors

$$g(t) \leq ae^{bt}$$

pour tout $t \leq T$.

La preuve est facile par itération de la condition sur g sous l'intégrale : on écrit

$$\begin{aligned}
g(t) &\leq a + b \int_0^t \left(a + b \int_0^s g(u) du \right) ds \\
&\leq a + abt + b^2 \int_{0 \leq u \leq s \leq t} g(u) du ds \\
&\leq a + abt + b^2 \int_{0 \leq u \leq s \leq t} \left(a + b \int_0^u g(v) dv \right) du ds \\
&\leq a + abt + ab^2 \int_{0 \leq u \leq s \leq t} du ds + b^3 \int_0^t \int_0^s \int_0^u g(v) dv du ds \\
&\leq a + abt + a \frac{b^2 t^2}{2} + b^3 \int_0^t \int_0^s \int_0^u g(v) dv du ds \\
&\vdots \\
&\leq a + abt + a \frac{b^2 t^2}{2} + a \frac{b^3 t^3}{3!} + \dots = ae^{bt}.
\end{aligned}$$

□

On fixe alors $t > 0$ et on considère deux solutions distinctes X et Y de $E_x(b, \sigma)$. On a

$$\begin{aligned}
\mathbb{E}(|X_t - Y_t|^2) &= \mathbb{E} \left| \int_0^t b(s, X_s) - b(s, Y_s) ds + \int_0^t \sigma(s, X_s) - \sigma(s, Y_s) dB_s \right|^2 \\
&\leq 2 \left(\mathbb{E} \left(\int_0^t |b(s, X_s) - b(s, Y_s)| ds \right)^2 + \mathbb{E} \left(\int_0^t |\sigma(s, X_s) - \sigma(s, Y_s)| dB_s \right)^2 \right) \\
&\leq 2M^2 \left(t^2 \int_0^t \mathbb{E} |X_s - Y_s|^2 ds + \int_0^t \mathbb{E} |X_s - Y_s|^2 ds \right) \\
&\leq K \int_0^t \mathbb{E} |X_s - Y_s|^2 ds
\end{aligned}$$

pour une certaine constante $K > 0$, où dans la première inégalité on a utilisé la formule $(a + b)^2 \leq 2(a^2 + b^2)$ et dans la deuxième inégalité on a utilisé la lipschitzianité de b et σ , l'inégalité de Hölder pour la première intégrale et l'isométrie de l'intégrale stochastique pour la deuxième. On en déduit le résultat en utilisant le lemme de Gronwall.

□

Dans la pratique, le théorème d'existence lipschitzien est parfois insuffisant pour ce qui concerne le coefficient de diffusion. On peut cependant notablement améliorer ce théorème dans le cas réel :

Théorème 5.7 [Yamada-Watanabe II] Soit $d = m = 1$. Supposons que b et σ soient à croissance linéaire, que b vérifie la condition de Lipschitz locale et que $|\sigma(t, x) - \sigma(t, y)|^2 \leq \rho(|x - y|)$ pour tout $t \geq 0$, où ρ est une fonction borélienne de $]0, \infty[$ dans lui-même telle que

$$\int_{|z| \leq \varepsilon} \frac{dz}{\rho^2(z)} = +\infty.$$

pour tout $\varepsilon > 0$. Alors $E_x(b, \sigma)$ admet une unique solution forte.

Une classe très importante d'EDS sont les équations de Bessel :

$$X_t = x + 2 \int_0^t \sqrt{X_s} dW_s + \delta t \quad (5.3)$$

sur \mathbb{R} , où W est un Brownien réel et $x, \delta > 0$. Par le deuxième théorème de Yamada-Watanabe appliqué avec la fonction $\rho(x) = \sqrt{x}$, on voit que (5.3) admet une unique solution forte. De plus, on peut montrer que cette solution est toujours positive et de durée de vie infinie. Cependant, le théorème d'existence lipschitzien n'aurait pas permis d'obtenir ce résultat, car $\sigma(x) = \sqrt{x}$ n'est pas lipschitzienne en zéro. Le processus X est un processus de Bessel carré de dimension δ . Nous reviendrons plus tard sur les processus de Bessel et sur leurs liens avec le mouvement brownien.

5.2 Quelques propriétés

5.2.1 Propriété de Markov

Supposons que $E_x(b, \sigma)$ ait une unique solution forte $\{X_t(x), t \geq 0\}$. Par linéarité de l'intégrale, on voit que pour tout $s, t \geq 0$,

$$X_{s+t}(x) = X_s(x) + \int_s^{s+t} b(u, X_u(x)) du + \int_s^{s+t} \sigma(u, X_u(x)) dB_u.$$

Comme la variable $X_s(x)$ est \mathcal{F}_s^B -mesurable, elle est indépendante du processus des accroissements $\{B_u - B_s, u \geq s\}$, qui est lui-même un Brownien B' partant de 0. On peut donc écrire

$$X_{s+t}(x) = X_s(x) + \int_0^t b(s+u, X_{s+u}(x)) du + \int_0^t \sigma(s+u, X_{s+u}(x)) dB'_u$$

et par unicité, ceci entraîne que $X_{s+t}(x) = X_t^s(X_s(x))$ pour tout $t \geq 0$, où $X_t^s(x)$ est l'unique solution de $E_x(b(s+.,.), \sigma(s+.,.))$ portée par le Brownien B' . Pour tout $s, t \geq 0$ et toute fonction borélienne f , on déduit alors de l'indépendance de B' et \mathcal{F}_s^B que

$$\mathbb{E}[f(X_{t+s}) | \mathcal{F}_s^B] = \mathbb{E}[f(X_{t+s}) | X_t] = \Phi_t(s, X_s)$$

où $\Phi_t(s, x) = \mathbb{E}[f(X_t^s(x))]$. Ceci signifie que X vérifie la *propriété de Markov inhomogène*. C'est un résultat important qui permet de calculer facilement certaines espérances conditionnelles. Dans le cas où les coefficients b et σ ne dépendent pas du temps, X vérifie de plus la *propriété de Markov homogène*, au sens où

$$\mathbb{E}[f(X_{t+s}) | \mathcal{F}_s^B] = \mathbb{E}[f(X_{t+s}) | X_s] = \Phi_t(X_s)$$

avec $\Phi_t(x) = \mathbb{E}[f(X_t(x))]$. La différence entre homogène et inhomogène provient de la dépendance en temps. Dans le cas homogène, on peut étendre la propriété aux temps d'arrêt finis :

Théorème 5.8 [Propriété de Markov forte] *Soit X l'unique solution forte de $E_x(b, \sigma)$ avec b et σ ne dépendant pas du temps. Soit T un temps d'arrêt fini p.s. pour la filtration naturelle du Brownien porteur. Alors pour tout $t \geq 0$ et toute fonction f mesurable bornée,*

$$\mathbb{E}[f(X_{T+t}) | \mathcal{F}_T^B] = \mathbb{E}[f(X_{T+t}) | X_T] = \Phi_t(X_T)$$

où l'on a noté $\Phi_t(x) = \mathbb{E}[f(X_t(x))]$.

Dans le cas inhomogène, on montre que le processus espace-temps $t \mapsto (t, X_t)$ vérifie la propriété de Markov forte, au sens où pour tout temps d'arrêt fini T et toute fonction f de deux variables,

$$\mathbb{E}[f(T+t, X_{T+t}) | \mathcal{F}_T^B] = \mathbb{E}[f(T+t, X_{T+t}) | T, X_T] = \Phi_t(T, X_T)$$

avec la notation $\Phi_t(s, x) = \mathbb{E}[f(s+t, X_t^s(x))]$. On voit donc que en général, un couple de processus (X, Y) peut être fortement markovien sans que ses composantes le soient.

5.2.2 Théorème de comparaison

Ce théorème permet de comparer presque sûrement deux EDS uni-dimensionnelles, et s'avère souvent extrêmement utile en pratique. La preuve, qui utilise des arguments semblables à ceux de l'unicité trajectoirelle vus précédemment, peut être trouvée dans [11].

Théorème 5.9 *Soit $\{W_t, t \geq 0\}$ un mouvement brownien réel, b_1, b_2 et σ trois fonctions globalement lipschitziennes, $x_1 \geq x_2$ deux réels. On considère les deux EDS*

$$X_t^i = x_i + \int_0^t b_i(X_s^i) ds + \int_0^t \sigma(X_s^i) dW_s$$

pour $i = 1, 2$. Supposons que $b_1(x) \geq b_2(x)$ pour tout $x \in \mathbb{R}$. Alors $X_t^1 \geq X_t^2$ p.s. pour tout $t \geq 0$.

5.2.3 Fonction d'échelle

On considère l'EDS homogène

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

où b et σ sont des fonctions de \mathbb{R} dans \mathbb{R} globalement lipschitziennes. D'après la formule d'Itô, on voit que si f est une fonction de \mathbb{R} dans \mathbb{R} de classe \mathcal{C}^2 à dérivées bornées et telle que

$$b(x)f'(x) + \frac{1}{2}\sigma^2(x)f''(x) = 0$$

pour tout $x \in \mathbb{R}$, alors le processus $t \mapsto f(X_t)$ est une martingale (martingale locale si f n'est pas à dérivées bornées). La fonction f est appelée *fonction d'échelle* du processus X . Elle est déterminée à deux constantes près c_1 et c_2 par la formule

$$f(x) = \int_{c_1}^x \exp\left(-2 \int_{c_2}^u b(v)/\sigma^2(v) dv\right) du.$$

L'opérateur \mathcal{L} qui à $f \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$ fait correspondre

$$\mathcal{L}f : x \mapsto b(x)f'(x) + \frac{1}{2}\sigma^2(x)f''(x)$$

s'appelle le *générateur infinitésimal* du processus X ou encore son *Dynkin*. On peut montrer qu'il vérifie

$$\begin{aligned} \mathcal{L}f(x) &= \lim_{t \rightarrow 0} \frac{\mathbb{E}_x[f(X_t)] - f(x)}{t} \\ &= \lim_{t \rightarrow 0} \frac{\mathbb{E}[f(X_t) \mid X_0 = x] - f(x)}{t} \end{aligned}$$

pour tout $x \in \mathbb{R}^d$. Le générateur infinitésimal a aussi un sens dans le cas inhomogène : si l'on considère l'EDS

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

alors on peut définir son générateur infinitésimal comme l'opérateur \mathcal{L} agissant sur les fonctions de $\mathbb{R}^+ \times \mathbb{R}$ une fois dérivables en temps et deux fois en espace par

$$\mathcal{L}(f)(t, x) = b(t, x)f'_x(t, x) + \frac{\sigma^2(t, x)}{2}f''_{xx}(t, x).$$

D'après la formule d'Itô, on voit que si f est une fonction de $\mathbb{R}_+ \times \mathbb{R}$ dans \mathbb{R} telle que

$$\mathcal{L}(f)(t, x) + f'_t(t, x) = 0,$$

alors le processus $\{f(t, X_t), t \geq 0\}$ est une martingale locale. C'est de plus une vraie martingale sous des conditions supplémentaires d'intégrabilité, par exemple quand la fonction $\sigma f'_x$ est bornée.

On peut enfin introduire un coefficient exponentiel (qui joue le rôle d'un coefficient d'actualisation en finance) pour retrouver le prix d'un call européen vu au chapitre précédent. En effet, supposons que X est un brownien géométrique

$$dX_t = rX_t dt + \sigma X_t dB_t$$

et soit f une fonction de $\mathbb{R}_+ \times \mathbb{R}$ dans \mathbb{R} telle que $\sigma f'_x$ est bornée et

$$f'_t(t, x) + \mathcal{L}(f)(t, x) = rf(t, x), \quad (5.4)$$

alors le processus $t \mapsto e^{-rt}f(t, X_t)$ est une martingale. En particulier, si f vérifie $\forall x, f(T, x) = h(x)$ comme condition au bord, alors

$$f(t, X_t) = e^{r(t-T)} \mathbb{E}[h(X_T) \mid \mathcal{F}_t].$$

On donnera des applications en finance dans la section 5.3.2.

5.2.4 Martingale exponentielle et condition de Novikov

Soit θ un bon processus local et Z_0 une constante. Par la formule d'Itô, on démontre que l'unique solution de l'EDS

$$Z_t = Z_0 + \int_0^t \theta_s Z_s dB_s \quad (5.5)$$

est

$$Z_t = Z_0 \exp \left[\int_0^t \theta_s dB_s - \frac{1}{2} \int_0^t \theta_s^2 ds \right].$$

Le processus Z , noté $\mathcal{E}_t(\theta \star B)$ est appelé *l'exponentielle de Doléans-Dade* de $\theta \star B$. Par (5.5), c'est une martingale locale positive dès que $Z_0 > 0$. Le critère suivant, de preuve difficile, permet de savoir quand l'exponentielle de Doléans-Dade est une martingale :

Théorème 5.10 [Condition de Novikov] *Supposons que*

$$\mathbb{E} \left[\exp \left(\frac{1}{2} \int_0^t \theta_s^2 ds \right) \right] < \infty$$

pour tout $t > 0$. Alors $t \mapsto \mathcal{E}_t(\theta \star B)$ est une vraie martingale.

Quand la condition de Novikov n'est pas satisfaite, $\mathcal{E}(\theta \star B)$ est une martingale locale positive, donc une surmartingale, et $\mathbb{E}[Z_t] \leq \mathbb{E}[Z_s] \leq Z_0$ pour tout $t \geq s \geq 0$. On ne connaît pas de conditions plus faciles à vérifier que la condition de Novikov, sauf dans le cas particulier suivant :

Proposition 5.11 *Supposons $\theta_t = f(t, B_t)$ où f est une fonction globalement lipschitzienne. Alors $t \mapsto \mathcal{E}_t(\theta \star B)$ est une vraie martingale.*

5.3 Lien avec les équations aux dérivées partielles

5.3.1 Problème parabolique

On considère l'opérateur $\mathcal{A}f(t, x) = f'_t(t, x) + \mathcal{L}f(t, x)$ où \mathcal{L} est le générateur infinitésimal du processus X solution de l'EDS

$$X_t = X_0 + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dB_s. \quad (5.6)$$

On cherche les solutions du problème parabolique suivant

$$\begin{cases} \mathcal{A}f(t, x) = 0 & \text{pour tout } x \in \mathbb{R} \text{ et } t \in [0, T] \\ f(T, x) = g(x) & \text{pour tout } x \in \mathbb{R}. \end{cases} \quad (5.7)$$

où g est une fonction de \mathbb{R} dans \mathbb{R} . Si f est une solution de (5.7) et si pour tout $u \geq t \geq 0$ on note $X_u^{x,t}$ une solution de

$$X_u^{x,t} = X_t^{x,t} + \int_t^u b(s, X_s^{x,t}) ds + \int_t^u \sigma(s, X_s^{x,t}) dB_s$$

avec pour condition initiale $X_t^{x,t} = x$, la formule d'Itô conduit alors à

$$f(u, X_u^{x,t}) = f(t, x) + \int_t^u f'_x(s, X_s^{x,t}) \sigma(s, X_s^{x,t}) dB_s.$$

En faisant $u = T$ en particulier, on en déduit que

$$g(X_T^{x,t}) = f(t, x) + \int_t^T f'_x(s, X_s^{x,t}) \sigma(s, X_s^{x,t}) dB_s.$$

Si f'_x et σ vérifient des conditions d'intégrabilité suffisante, alors l'intégrale stochastique est une martingale. On en déduit une importante représentation probabiliste de la solution de (5.7) :

$$f(t, x) = \mathbb{E} [g(X_T^{x,t})].$$

On écrit parfois ce résultat sous la forme $f(t, x) = \mathbb{E}_{x,t} [g(X_T)]$, où X est la solution de (5.6) prise sous la probabilité $\mathbb{P}_{x,t}$ qui est telle que processus X prend la valeur x à l'instant t . On peut s'intéresser à un problème un peu plus général que (5.7) :

$$\begin{cases} \mathcal{A}f(t, x) = \alpha f(t, x) & \text{pour tout } x \in \mathbb{R} \text{ et } t \in [0, T] \\ f(T, x) = g(x) & \text{pour tout } x \in \mathbb{R}. \end{cases} \quad (5.8)$$

avec $\alpha > 0$. Avec les notations précédentes, si f est solution de (5.8), la formule d'Itô entre t et T conduit à

$$e^{-\alpha T} f(T, X_T^{x,t}) = e^{-\alpha t} f(t, x) + \int_t^T e^{-\alpha s} f'_x(s, X_s^{x,t}) \sigma(s, X_s^{x,t}) dB_s.$$

A nouveau sous des conditions d'intégrabilité suffisantes sur f'_x et σ , l'intégrale stochastique est une martingale et on en déduit la représentation probabiliste

$$f(t, x) = \mathbb{E}_{x,t} \left[e^{\alpha(t-T)} g(X_T) \right].$$

5.3.2 Formule de Black & Scholes

On considère à nouveau un sous jacent de dynamique

$$dS_t = S_t(bdt + \sigma dB_t).$$

Les résultats précédents entraînent que la solution de l'équation de Black & Scholes

$$xr \frac{\partial C}{\partial x}(t, x) + \frac{\partial C}{\partial t}(t, x) + \frac{\sigma^2 x^2}{2} \frac{\partial^2 C}{\partial x^2}(t, x) = rC(t, x)$$

avec $C(T, x) = (x - K)^+$, est donnée par

$$C(t, x) = \mathbb{E} \left[e^{r(t-T)} (\tilde{S}_T^{x,t} - K)^+ \right] \quad (5.9)$$

avec

$$d\tilde{S}_t = \tilde{S}_t(rdt + \sigma dW_t)$$

où W est un mouvement Brownien. La solution de cette équation vérifie

$$\tilde{S}_T = \tilde{S}_t \exp(\sigma(W_T - W_t)(r - \sigma^2/2)(T - t))$$

soit

$$\tilde{S}_T^{x,t} = x e^{\sigma\sqrt{T-t}G + (r - \sigma^2/2)(T-t)}$$

avec $G \sim \mathcal{N}(0, 1)$. Pour $t = 0$, le calcul donne

$$\begin{aligned} \mathbb{E} \left[e^{-rT} (\tilde{S}_T^x - K)^+ \right] &= e^{-rT} \left(\mathbb{E} \left[\tilde{S}_T^x \mathbb{1}_{\{\tilde{S}_T^x \geq K\}} \right] - K \mathbb{P} \left[\tilde{S}_T^x \geq K \right] \right) \\ &= x e^{-rT} \mathbb{E} \left[e^{\sigma\sqrt{T}G + (r - \sigma^2/2)(T)} \mathbb{1}_{\{\sigma\sqrt{T}G + (r - \sigma^2/2)(T) \geq \ln(K/x)\}} \right] \\ &\quad - K e^{-rT} \mathbb{P} \left[\sigma\sqrt{T}G + (r - \sigma^2/2)(T) \geq \ln(K/x) \right] \end{aligned}$$

L'espérance et la probabilité se calculent alors en explicitant les intégrales qui font intervenir la densité gaussienne. Nous verrons plus loin que le calcul du premier terme peut se déduire du second. L'expression (5.9) permet également de calculer le 'Delta' de l'option : dans le cas $t = 0$, on a

$$C(0, x) = \mathbb{E} \left[e^{-rT} (x M_T e^{rT} - K)^+ \right]$$

avec la notation $\tilde{S}_t = x M_t e^{rt}$ et où M est une martingale. En dérivant par rapport à x sous l'espérance, on retrouve

$$\frac{\partial C}{\partial x}(0, x) = \mathbb{E} \left[M_T \mathbb{1}_{\{x M_T e^{rT} \geq K\}} \right] = \mathcal{N}(d_1)$$

avec la formule pour d_1 donnée au chapitre précédent.

5.3.3 Formule de Feynman-Kac

Nous venons de voir comment les processus permettent de résoudre explicitement certaines EDP. Les mêmes techniques peuvent être aussi employées pour l'étude de certaines EDO du deuxième ordre. On s'intéresse par exemple au problème suivant, dit de *Sturm-Liouville* :

$$(\alpha + k)f = \frac{f''}{2} + g \quad (5.10)$$

où $\alpha > 0$, $k : \mathbb{R} \rightarrow \mathbb{R}^+$ est une fonction continue et $g : \mathbb{R} \rightarrow \mathbb{R}$ une fonction continue telle que :

$$\int_{\mathbb{R}} |g(x+y)| e^{-|y|\sqrt{2\alpha}} dy < +\infty$$

pour tout $x \in \mathbb{R}$. On a le

Théorème 5.12 [Formule de Feynman-Kac] *La fonction f définie par :*

$$f(x) = \mathbb{E} \left[\int_0^\infty g(x + B_t) \exp \left(-\alpha t - \int_0^t k(x + B_s) ds \right) dt \right],$$

où $\{B_t, t \geq 0\}$ est un mouvement brownien standard issu de 0, est l'unique solution bornée et \mathcal{C}^2 de (5.10).

Donnons une idée de la démonstration. Tout d'abord, les conditions de positivité sur α et k et de bornitude sur g garantissent l'existence d'une solution continue et bornée à l'équation (5.10). Soit $\{Z_t, t \geq 0\}$ le processus à variation finie défini par :

$$Z_t = \alpha t + \int_0^t k(x + B_s) ds.$$

On applique la première formule d'Itô au processus

$$U_t(x) = f(x + B_t) e^{-Z_t} + \int_0^t g(x + B_s) e^{-Z_s} ds$$

où f est une fonction \mathcal{C}^2 , et on obtient

$$\begin{aligned} dU_t(x) &= f'(x + B_t) e^{-Z_t} dB_t \\ &+ \left(\frac{f''(x + B_t)}{2} + g(x + B_t) - (\alpha + k(x + B_s)) f(x + B_t) \right) e^{-Z_t} dt. \end{aligned}$$

Si f est solution de (5.10), alors

$$\frac{f''(x)}{2} + g(x) - (\alpha + k(x)) f(x) = 0,$$

de sorte que

$$U_t(x) = f(x) + \int_0^t f'(x + B_s) e^{-Z_s} dB_s$$

est une martingale locale. Comme Z est positive et f' bornée, U est en fait une vraie martingale et on a

$$f(x) = \mathbb{E}[U_t(x)]$$

pour tout $t \geq 0$. Pour obtenir la formule du théorème, il suffit alors de démontrer que

$$\mathbb{E}[f(B_t) e^{-Z_t}] \longrightarrow 0$$

quand $t \rightarrow +\infty$.

□

Ce résultat nous donne en particulier la transformée de Laplace en temps de la variable

$$g(B_t) \exp - \int_0^t k(B_s) ds$$

pour toute fonction g , et donc par inversion la loi du couple

$$\left(B_t, \int_0^t k(B_s) ds \right).$$

Ce couple est très important en Finance, puisqu'il permet de pricer certaines options exotiques avec temps d'occupation. La formule de Feynman-Kac permet aussi de calculer la densité de la variable aléatoire

$$A_t^+ = \int_0^t \mathbb{1}_{[0, \infty[}(B_s) ds$$

qui représente le temps d'occupation de \mathbb{R}^+ par le Brownien. En effet, si on pose $k(x) = \beta \mathbb{1}_{x \geq 0}$ et $g(x) = 1$, on en déduit que pour $\alpha, \beta > 0$ la fonction

$$f(x) = \mathbb{E} \left[\int_0^\infty \exp \left(-\alpha t - \beta \int_0^t \mathbb{1}_{[0, \infty)}(x + B_s) ds \right) dt \right]$$

est solution de l'EDO

$$\begin{cases} \alpha f(x) = 1 - \beta f(x) + \frac{f''(x)}{2} & \text{si } x \geq 0, \\ \alpha f(x) = 1 + \frac{f''(x)}{2} & \text{si } x \leq 0. \end{cases}$$

L'unique solution bornée et continue de cette EDO est donnée par :

$$f(x) = \begin{cases} Ae^{-x\sqrt{2(\alpha+\beta)}} + \frac{1}{\alpha+\beta} & \text{si } x \geq 0, \\ Be^{x\sqrt{2\alpha}} + \frac{1}{\alpha} & \text{si } x \leq 0. \end{cases}$$

En imposant la continuité de f et f' en zéro, on trouve

$$A = \frac{1}{\sqrt{\alpha(\alpha+\beta)}} - \frac{1}{(\alpha+\beta)} \quad \text{et} \quad B = \frac{1}{\sqrt{\alpha(\alpha+\beta)}} - \frac{1}{\alpha}.$$

On en déduit que

$$\int_0^\infty e^{-\alpha t} \mathbb{E} \left[e^{-\beta A_t^+} \right] dt = f(0) = \frac{1}{\sqrt{\alpha(\alpha+\beta)}}$$

et, en utilisant l'égalité

$$\int_0^\infty e^{-\alpha t} \left(\int_0^t du \frac{e^{-\beta u}}{\pi \sqrt{u(t-u)}} \right) dt = \frac{1}{\sqrt{\alpha(\alpha+\beta)}},$$

que la densité de A_t^+ est donnée par :

$$\mathbb{P} [A_t^+ \in du] = \frac{\mathbb{1}_{\{u < t\}} du}{\pi \sqrt{u(t-u)}}.$$

La fonction de répartition de cette loi est

$$\mathbb{P} [A_t^+ \leq \theta] = \int_0^\theta \frac{ds}{\pi \sqrt{s(t-s)}} = \int_0^{\theta/t} \frac{du}{\pi \sqrt{u(1-u)}} = \frac{2}{\pi} \arg \sin \sqrt{\frac{\theta}{t}}$$

et l'on donne alors à la loi de A_t^+ le nom de *loi de l'arcsinus*. On remarque enfin que

$$\mathbb{P} [A_t^+ \leq \theta] = \mathbb{P} [tA_1^+ \leq \theta],$$

ce qui montre que les variables A_t^+ et tA_1^+ ont même loi. On aurait pu aussi obtenir ce résultat directement par scaling du Brownien.

5.4 Processus de Bessel

5.4.1 Norme d'un mouvement Brownien n -dimensionnel

Soit $n > 1$ et $B = (B_1, B_2, \dots, B_n)$ un mouvement Brownien n -dimensionnel. Soit X défini par $X_t = \|B_t\|$, soit $X_t^2 = \sum_{i=1}^n (B_i)^2(t)$. En appliquant la formule d'Itô $dX_t^2 = \sum_{i=1}^n 2B_i(t)dB_i(t) + n dt$. Le processus β défini par

$$d\beta_t = \frac{1}{X_t} B_t \cdot dB_t = \frac{1}{\|B_t\|} \sum_{i=1}^n B_i(t)dB_i(t), \quad \beta_0 = 0,$$

est une martingale continue et son crochet est t (en effet, $(\beta_t^2 - t, t \geq 0)$ est une martingale). Il en résulte que β est un mouvement Brownien. L'égalité $d(X_t^2) = 2B_t \cdot dB_t + n dt$ s'écrit

$$d(X_t^2) = 2X_t d\beta_t + n dt.$$

Une nouvelle application de la formule d'Itô conduit à

$$dX_t = d\beta_t + \frac{n-1}{2} \frac{dt}{X_t}$$

où β est un mouvement Brownien, ainsi, en posant $V_t = X_t^2$

$$dV_t = 2\sqrt{V_t}d\beta_t + n dt.$$

On dit que X est un processus de Bessel (BES) de dimension n , et que V est un processus de Bessel carré (BESQ) de dimension n .

5.4.2 Définition générale

Soit W un mouvement Brownien. En utilisant l'inégalité $|\sqrt{x} - \sqrt{y}| \leq \sqrt{|x - y|}$, le second théorème d'existence montre que pour $\delta \geq 0$ et $\alpha \geq 0$, l'équation

$$dZ_t = \delta dt + 2\sqrt{|Z_t|} dW_t, \quad Z_0 = \alpha$$

a une unique solution forte. Cette solution est un processus de Bessel carré de dimension δ , que l'on désigne par BESQ^δ . En particulier, si $\alpha = 0$ et $\delta = 0$, la solution $Z \equiv 0$ est l'unique solution. En utilisant le théorème de comparaison, si $0 \leq \delta \leq \delta'$ et si ρ et ρ' sont des processus de Bessel carré de dimension δ et δ' partant du même point, alors $0 \leq \rho_t \leq \rho'_t$ a.s.

Dans le cas $\delta > 2$, le processus de Bessel carré BESQ^δ partant de α n'atteint jamais 0. Si $0 < \delta < 2$, le processus ρ atteint 0 en un temps fini. Si $\delta = 0$ le processus reste en 0 dès qu'il atteint ce point.

Définition 5.13 (BESQ^δ) Pour tout $\delta \geq 0$ et $\alpha \geq 0$, l'unique solution forte de

$$\rho_t = \alpha + \delta t + 2 \int_0^t \sqrt{\rho_s} dW_s$$

processus de Bessel carré de dimension δ , partant de α et est noté BESQ^δ .

Définition 5.14 (BES^δ) Soit ρ un BESQ^δ partant de α . Le processus $R = \sqrt{\rho}$ est un processus de Bessel de dimension δ , partant de $a = \sqrt{\alpha}$ et est noté BES^δ .

Définition 5.15 Le nombre $\nu = (\delta/2) - 1$ (soit $\delta = 2(\nu + 1)$) est l'indice du processus de Bessel, et un processus de Bessel d'indice ν est noté $\text{BES}^{(\nu)}$.

Pour $\delta > 1$, un BES^δ est solution de

$$R_t = \alpha + W_t + \frac{\delta-1}{2} \int_0^t \frac{1}{R_s} ds. \quad (5.11)$$

5.4.3 Probabilités de transition

Les fonctions de Bessel modifiées I_ν et K_ν solutions de

$$x^2 u''(x) + xu'(x) - (x^2 + \nu^2)u(x) = 0$$

sont données par :

$$I_\nu(z) = \left(\frac{z}{2}\right)^\nu \sum_{n=0}^{\infty} \frac{z^{2n}}{2^{2n} n! \Gamma(\nu + n + 1)}$$

$$K_\nu(z) = \frac{\pi(I_{-\nu}(z) - I_\nu(z))}{2 \sin \pi z}$$

Les probabilités de transition $q_t^{(\nu)}$ d'un BESQ $^{(\nu)}$ sont

$$q_t^{(\nu)}(x, y) = \frac{1}{2t} \left(\frac{y}{x}\right)^{\nu/2} \exp\left(-\frac{x+y}{2t}\right) I_\nu\left(\frac{\sqrt{xy}}{t}\right) \quad (5.12)$$

et le processus de Bessel d'indice ν a une probabilité de transition $p_t^{(\nu)}$ donnée par

$$p_t^{(\nu)}(x, y) = \frac{y}{t} \left(\frac{y}{x}\right)^\nu \exp\left(-\frac{x^2 + y^2}{2t}\right) I_\nu\left(\frac{xy}{t}\right), \quad (5.13)$$

5.5 Modèle de Cox-Ingersoll-Ross

Pour modéliser des taux, Cox-Ingersoll-Ross étudient l'équation suivante

$$dr_t = k(\theta - r_t)dt + \sigma\sqrt{r_t}dB_t \quad (5.14)$$

L'unique solution est un processus positif pour $k\theta \geq 0$ (utiliser le second théorème d'existence). Il n'est pas possible d'obtenir une formule explicite. Soit r^x le processus solution de (5.14) avec $r_0^x = x$.

Le changement de temps $A(t) = \sigma^2 t/4$ réduit l'étude de (5.14) au cas $\sigma = 2$: En effet, si $Z_t = r_{\sigma^2 t/4}$, alors

$$dZ_t = k'(\theta - Z_t)dt + 2\sqrt{Z_t}dB_t$$

avec $k' = k\sigma^2/4$ et où B est un mouvement Brownien.

Le processus de CIR process (5.14) est un BESQ changé de temps : en effet,

$$r_t = e^{-kt} \rho\left(\frac{\sigma^2}{4k}(e^{kt} - 1)\right)$$

où $(\rho(s), s \geq 0)$ est un BESQ $^\delta(\alpha)$, avec $\delta = \frac{4k\theta}{\sigma^2}$.

On peut montrer que, si $T_0^x \stackrel{\text{def}}{=} \inf\{t \geq 0 : r_t^x = 0\}$ et $2k\theta \geq \sigma^2$ alors $P(T_0^x = \infty) = 1$. Si $0 \leq 2k\theta < \sigma^2$ et $k > 0$ alors $P(T_0^x < \infty) = 1$ et si $k < 0$ on a $P(T_0^x < \infty) \in]0, 1[$. (Cela se fait au moyen du théorème de comparaison)

Cependant, on peut calculer l'espérance de la v.a. r_t au moyen de l'égalité $E(r_t) = r_0 + k(\theta t - \int_0^t E(r_s)ds)$, en admettant que l'intégrale stochastique est une martingale, ce qui est le cas. On calcule sans difficultés supplémentaires l'espérance conditionnelle, en utilisant le caractère Markovien :

Théorème 5.16 Soit r le processus vérifiant

$$dr_t = k(\theta - r_t)dt + \sigma\sqrt{r_t}dB_t.$$

L'espérance conditionnelle et la variance conditionnelle sont données par

$$E(r_t | \mathcal{F}_s) = r_s e^{-k(t-s)} + \theta(1 - e^{-k(t-s)}),$$

$$\text{Var}(r_t | \mathcal{F}_s) = r_s \frac{\sigma^2(e^{-k(t-s)} - e^{-2k(t-s)})}{k} + \frac{\theta\sigma^2(1 - e^{-k(t-s)})^2}{2k}.$$

Preuve : Par définition, on a pour $s \leq t$

$$r_t = r_s + k \int_s^t (\theta - r_u) du + \sigma \int_s^t \sqrt{r_u} dB_u,$$

et en appliquant la formule d'Itô

$$\begin{aligned} r_t^2 &= r_s^2 + 2k \int_s^t (\theta - r_u) r_u du + 2\sigma \int_s^t (r_u)^{3/2} dB_u + \sigma^2 \int_s^t r_u du \\ &= r_s^2 + (2k\theta + \sigma^2) \int_s^t r_u du - 2k \int_s^t r_u^2 du + 2\sigma \int_s^t (r_u)^{3/2} dB_u. \end{aligned}$$

En admettant que les intégrales stochastiques qui interviennent dans les égalités ci-dessus sont d'espérance nulle, on obtient, pour $s = 0$

$$E(r_t) = r_0 + k \left(\theta t - \int_0^t E(r_u) du \right),$$

et

$$E(r_t^2) = r_0^2 + (2k\theta + \sigma^2) \int_0^t E(r_u) du - 2k \int_0^t E(r_u^2) du.$$

Soit $\Phi(t) = E(r_t)$. En résolvant l'équation $\Phi(t) = r_0 + k(\theta t - \int_0^t \Phi(u) du)$ qui se transforme en l'équation différentielle $\Phi'(t) = k(\theta - \Phi(t))$ et $\Phi(0) = r_0$, on obtient

$$E(r(t)) = \theta + (r_0 - \theta)e^{-kt}.$$

De la même façon, on introduit $\psi(t) = E(r_t^2)$ et en résolvant $\Psi'(t) = (2k\theta + \sigma^2)\Phi(t) - 2k\Psi(t)$, on calcule

$$\text{Var}[r_t] = \frac{\sigma^2}{k} (1 - e^{-kt}) [r_0 e^{-kt} + \frac{\theta}{2} (1 - e^{-kt})].$$

L'espérance et la variance conditionnelle de r s'obtiennent en appliquant la propriété de Markov :

$$E(r_t | \mathcal{F}_s) = \theta + (r_s - \theta)e^{-k(t-s)} = r_s e^{-k(t-s)} + \theta(1 - e^{-k(t-s)}),$$

$$\text{Var}(r_t | \mathcal{F}_s) = r_s \frac{\sigma^2(e^{-k(t-s)} - e^{-2k(t-s)})}{k} + \frac{\theta \rho^2 (1 - e^{-k(t-s)})^2}{2k}.$$

□.

On va utiliser les méthodes du chapitre précédent pour calculer $E\left(\exp - \int_t^T r_u du | \mathcal{F}_t\right)$.

Calcul du prix d'un zéro-coupon

Proposition 5.17 *Soit*

$$dr_t = a(b - r_t)dt + \sigma \sqrt{r_t} dB_t.$$

Alors

$$E\left(\exp - \int_t^T r_u du | \mathcal{F}_t\right) = G(t, r_t)$$

avec

$$G(t, x) = \Phi(T - t) \exp[-x\Psi(T - t)]$$

$$\Psi(s) = \frac{2(e^{\gamma s} - 1)}{(\gamma + a)(e^{\gamma s} - 1) + 2\gamma}, \quad \Phi(s) = \left(\frac{2\gamma e^{(\gamma+a)\frac{s}{2}}}{(\gamma + a)(e^{\gamma s} - 1) + 2\gamma} \right)^{\frac{2ab}{\rho^2}}, \quad \gamma^2 = a^2 + 2\rho^2.$$

Preuve : Soit $r^{x,t}$ la solution de

$$dr_s^{x,t} = a(b - r_s^{x,t})ds + \rho \sqrt{r_s^{x,t}} dB_s, \quad r_t^{x,t} = x$$

et $R_s^t = \exp\left(-\int_t^s r_u^{x,t} du\right)$. La propriété de Markov implique qu'il existe G telle que

$$\exp\left(-\int_t^s r_u^{x,t} du | \mathcal{F}_t\right) = G(t, r_t)$$

On admet que G est de classe $C^{1,2}$. On applique la formule d'Itô à $G(s, r_s^{x,t})R_s^t$ qui est une martingale. Il vient

$$G(T, r_T^{x,t})R_T^t = G(t, x) + \int_t^T R_s^t (-r_s^{x,t}G + \frac{\partial G}{\partial t} + a(b - r_s^{x,t})\frac{\partial G}{\partial x} + \frac{1}{2}\sigma^2 r_s^{x,t}\frac{\partial^2 G}{\partial x^2})(s, r_s)ds + M_T - M_t$$

où M_t est une intégrale stochastique. Si l'on choisit G telle que

$$-xG + \frac{\partial G}{\partial t} + a(b - x)\frac{\partial G}{\partial x} + \frac{1}{2}\sigma^2 x\frac{\partial^2 G}{\partial x^2} = 0 \quad (5.15)$$

et $G(T, x) = 1, \forall x$, il vient

$$R_T^t = R_t G(t, r_t) + M_T - M_t,$$

où M est une martingale. En particulier, $E\left(\exp\left(-\int_0^T r_s ds\right)\right) = E(R_T) = R_0 G(0, x)$. En se plaçant entre t et T , on obtient

$$E\left(\exp\left(-\int_t^T r_u^{x,t} du\right)\right) = G(t, x)$$

Il reste à calculer la solution de l'équation aux dérivées partielles (5.15). Un calcul assez long montre que

$$G(t, x) = \Phi(T - t) \exp[-x\Psi(T - t)]$$

avec

$$\begin{aligned} \Psi(s) &= \frac{2(e^{\gamma s} - 1)}{(\gamma + a)(e^{\gamma s} - 1) + 2\gamma} & \Phi(s) &= \left(\frac{2\gamma e^{(\gamma+a)\frac{s}{2}}}{(\gamma + a)(e^{\gamma s} - 1) + 2\gamma} \right)^{\frac{2ab}{\rho^2}} \\ \gamma^2 &= a^2 + 2\sigma^2. \end{aligned}$$

□

Si l'on note $P(t, T)$ le prix du zero-coupon associé,

$$P(t, T) = \Phi(T - t) \exp[-r_t \Psi(T - t)]$$

on montre que

$$dP(t, T) = B(t, T) (r_t dt + \sigma(T - t, r_t) dB_t)$$

avec $\sigma(u, r) = \sigma\Psi(u)\sqrt{r}$.

Chapitre 6

Changement de probabilité - Théorème de Girsanov

Avec la formule d'Itô, le théorème de Girsanov est l'outil fondamental du calcul stochastique. Il décrit comment changer de façon absolument continue la loi de certains processus. Plus précisément, on considère un processus $\{X_t, t \geq 0\}$ défini sur un espace de probabilité $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ et on perturbe la mesure \mathbb{P} à l'aide d'une (\mathcal{F}_t) -martingale exponentielle. On obtient alors une nouvelle mesure de probabilité \mathbb{Q} , sous laquelle le processus X suit une autre loi, qu'il aurait pu être délicat d'étudier directement. On étudie alors cette loi en revenant sous la mesure \mathbb{P} avec la martingale exponentielle. Les applications de cette méthode sont multiples, aussi bien en Finance que pour le mouvement brownien proprement dit.

6.1 Généralités

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et Z une v.a. \mathcal{F} -mesurable positive d'espérance 1. On définit une nouvelle probabilité \mathbb{Q} sur \mathcal{F} par $\mathbb{Q}(A) = \mathbb{E}(Z \mathbb{1}_A)$. On a, pour toute v.a. \mathbb{Q} intégrable $\mathbb{E}_{\mathbb{Q}}(X) = \mathbb{E}_{\mathbb{P}}(ZX)$. Si l'espace de probabilité est muni d'une filtration, et si Z_T est une v.a. \mathcal{F}_T -mesurable positive d'espérance 1, on définit \mathbb{Q} sur \mathcal{F}_T par $\mathbb{Q}(A) = \mathbb{E}(Z_T \mathbb{1}_A)$. La probabilité \mathbb{Q} est équivalente à \mathbb{P} sur \mathcal{F}_T si Z_T est stricte positivité. Enfin, pour tout $t < T$ et tout $A \in \mathcal{F}_t$ on a

$$\mathbb{Q}_T[A] = \mathbb{E}_{\mathbb{P}}[\mathbb{E}_{\mathbb{P}}[Z_T | \mathcal{F}_t] \mathbb{1}_A] = \mathbb{E}_{\mathbb{P}}[Z_t \mathbb{1}_A],$$

en posant $Z_t = \mathbb{E}_{\mathbb{P}}[Z_T | \mathcal{F}_t]$ et

$$d\mathbb{Q}|_{\mathcal{F}_t} = Z_t d\mathbb{P}|_{\mathcal{F}_t}.$$

Lemme 6.1 *Soit $\{M_t, t \geq 0\}$ un processus. C'est une (\mathcal{F}_t) -martingale sous \mathbb{Q} si et seulement si le processus $t \mapsto Z_t M_t$ est une (\mathcal{F}_t) -martingale sous \mathbb{P} .*

Preuve : Supposons que M soit une \mathbb{Q} -martingale. Pour tout $s \leq t \leq T$ et tout $A \in \mathcal{F}_s$ on a

$$\mathbb{E}_{\mathbb{P}}[Z_t M_t \mathbb{1}_A] = \mathbb{E}_{\mathbb{Q}}[M_t \mathbb{1}_A] = \mathbb{E}_{\mathbb{Q}}[M_s \mathbb{1}_A] = \mathbb{E}_{\mathbb{P}}[Z_s M_s \mathbb{1}_A]$$

et l'on en déduit que ZM est une \mathbb{P} -martingale. La réciproque se démontre de la même façon. \square

6.2 La formule de Cameron-Martin

Avant de passer à cette formule qui est de type fonctionnel, donnons un exemple en dimension finie. Soient (X_1, \dots, X_n) des variables gaussiennes centrées réduites indépendantes, construites sur un même espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$. Pour tout $(\mu_1, \dots, \mu_n) \in \mathbb{R}^n$ on calcule

$$\mathbb{E} \left[\exp \sum_{i=1}^n \mu_i X_i \right] = \prod_{i=1}^n \mathbb{E} [\exp [\mu_i X_i]] = \prod_{i=1}^n \exp \left[\frac{\mu_i^2}{2} \right] = \exp \frac{1}{2} \left[\sum_{i=1}^n \mu_i^2 \right],$$

ce qui entraîne que

$$\mathbb{E} \left[\exp \sum_{i=1}^n \left(\mu_i X_i - \frac{\mu_i^2}{2} \right) \right] = 1.$$

Ainsi, on peut définir une nouvelle probabilité \mathbb{Q} sur (Ω, \mathcal{F}) en posant

$$Z(\omega) = \exp \sum_{i=1}^n \left(\mu_i X_i(\omega) - \frac{\mu_i^2}{2} \right)$$

et $\mathbb{Q}(d\omega) = Z(\omega)\mathbb{P}(d\omega)$, autrement dit

$$\mathbb{Q}[A] = \mathbb{E}_{\mathbb{P}}[Z\mathbb{1}_A] = \int_A Z(\omega)\mathbb{P}(d\omega)$$

pour tout $A \in \mathcal{F}$. La mesure \mathbb{Q} est bien une probabilité sur (Ω, \mathcal{F}) puisque $\mathbb{E}_{\mathbb{P}}[Z] = 1$ et que Z est strictement positive. De plus, comme p.s. $Z > 0$, les probabilités \mathbb{P} et \mathbb{Q} sont *équivalentes*, au sens où

$$\mathbb{P}[A] = 0 \iff \mathbb{Q}[A] = 0$$

pour tout $A \in \mathcal{F}$. La variable Z désigne la *densité* de \mathbb{P} par rapport à \mathbb{Q} , telle qu'elle est avait été donnée au chapitre 1 par le théorème de Radon-Nikodym. On cherche maintenant à savoir quelle est la loi du n -uple (X_1, \dots, X_n) sous cette nouvelle probabilité \mathbb{Q} . Pour cela on écrit

$$\begin{aligned} \mathbb{Q}(X_1 \in dx_1, \dots, X_n \in dx_n) &= e^{\sum_{i=1}^n (\mu_i x_i - \mu_i^2/2)} \mathbb{P}(X_1 \in dx_1, \dots, X_n \in dx_n) \\ &= (2\pi)^{-n/2} e^{\sum_{i=1}^n (\mu_i x_i - \mu_i^2/2)} e^{-\sum_{i=1}^n \frac{x_i^2}{2}} dx_1 \cdots dx_n \\ &= (2\pi)^{-n/2} e^{-\frac{1}{2} \sum_{i=1}^n (x_i - \mu_i)^2} dx_1 \cdots dx_n \end{aligned}$$

et l'on en déduit que sous \mathbb{Q} ,

$$(X_1, \dots, X_n) \sim \mathcal{N}(\mu, \text{Id}),$$

où $\mu = (\mu_1, \dots, \mu_n)$.

La formule de Cameron-Martin obéit au même principe de changement de probabilité, sauf que celui-ci a lieu sur l'espace des fonctions continues et donc en *dimension infinie*. On se donne $\{W_t, t \geq 0\}$ un mouvement brownien sur $(\Omega, \mathcal{F}, \mathbb{P})$, et on note $\{\mathcal{F}_t^W, t \geq 0\}$ sa filtration naturelle complétée. Pour tout $m \in \mathbb{R}$, on sait que le processus

$$t \mapsto Z_t^m = \exp \left[mW_t - \frac{m^2 t}{2} \right]$$

est une (\mathcal{F}_t^W) -martingale positive. Remarquons déjà l'analogie entre cette martingale Z^m et la variable Z définie plus haut. On fixe alors un horizon $T > 0$ et on construit une nouvelle mesure de probabilité \mathbb{Q}_T^m sur (Ω, \mathcal{F}_T) en posant

$$\mathbb{Q}_T^m[A] = \mathbb{E}_{\mathbb{P}}[Z_T^m \mathbb{1}_A].$$

La formule de Cameron-Martin spécifie alors la loi de W sous \mathbb{Q}_T^m :

Théorème 6.2 [Formule de Cameron-Martin] *Sous la mesure \mathbb{Q}_T^m , le processus*

$$\tilde{W} : t \mapsto W_t - mt, \quad t \leq T$$

est un mouvement brownien.

Preuve : Remarquons d'abord que les tribus \mathcal{F}_t^W et $\mathcal{F}_t^{\tilde{W}}$ sont égales pour tout $t \leq T$. On fixe $\lambda \in \mathbb{R}$ et on considère le processus

$$L_t^\lambda = \exp \left[\lambda \tilde{W}_t - \lambda^2 t/2 \right].$$

Pour tout $s \leq t \leq T$ et tout $A \in \mathcal{F}_s^W$, on remarque que

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}} [L_t^\lambda \mathbb{1}_A] &= \mathbb{E}_{\mathbb{P}} [L_t^\lambda Z_t \mathbb{1}_A] = \mathbb{E}_{\mathbb{P}} \left[\exp \left[\lambda \tilde{W}_t - \lambda^2 t/2 + m W_t - m^2 t/2 \right] \mathbb{1}_A \right] \\ &= \mathbb{E}_{\mathbb{P}} \left[\exp \left[\lambda W_t - \lambda m t - \lambda^2 t/2 + m W_t - m^2 t/2 \right] \mathbb{1}_A \right] \\ &= \mathbb{E}_{\mathbb{P}} \left[\exp \left[(\lambda + m) W_t - (\lambda + m)^2 t/2 \right] \mathbb{1}_A \right] \\ &= \mathbb{E}_{\mathbb{P}} \left[\exp \left[(\lambda + m) W_s - (\lambda + m)^2 s/2 \right] \mathbb{1}_A \right] \\ &= \mathbb{E}_{\mathbb{P}} [L_s^\lambda Z_s \mathbb{1}_A] = \mathbb{E}_{\mathbb{Q}} [L_s^\lambda \mathbb{1}_A] \end{aligned}$$

où dans la quatrième ligne on a utilisé la propriété de (\mathcal{F}_t^W) -martingale sous \mathbb{P} du processus

$$t \mapsto \exp \left[(\lambda + m) W_t - (\lambda + m)^2 t/2 \right],$$

puisque W est un Brownien sous \mathbb{P} . On en déduit que

$$\mathbb{E}_{\mathbb{Q}} [L_t^\lambda \mathbb{1}_A] = \mathbb{E}_{\mathbb{Q}} [L_s^\lambda \mathbb{1}_A]$$

et ceci signifie que

$$\mathbb{E}_{\mathbb{Q}} [L_t^\lambda | \mathcal{F}_s^W] = L_s^\lambda$$

pour tout $s \leq t \leq T$, donc que L^λ est une (\mathcal{F}_t^W) -martingale pour tout $\lambda \in \mathbb{R}$. Par le théorème de caractérisation de Paul Lévy (théorème 3.4), ceci signifie que \tilde{W} est un mouvement brownien. \square

6.3 Les deux théorèmes de Girsanov

On se donne à nouveau $\{W_t, t \geq 0\}$ un mouvement brownien sur $(\Omega, \mathcal{F}, \mathbb{P})$, et on note $\{\mathcal{F}_t^W, t \geq 0\}$ sa filtration naturelle complétée. Soit $\{\theta_t, t \geq 0\}$ un bon processus local vérifiant la condition de Novikov. Par les résultats du chapitre précédent, on sait que l'unique solution de l'EDS

$$Z_t^\theta = 1 + \int_0^t \theta_s Z_s^\theta dW_s$$

qui s'écrit

$$Z_t^\theta = \exp \left[\int_0^t \theta_s dW_s - \frac{1}{2} \int_0^t \theta_s^2 ds \right],$$

est une (\mathcal{F}_t^W) -martingale. Comme précédemment, on fixe un horizon $T > 0$ et on définit la mesure

$$\mathbb{Q}_T^\theta(d\omega) = Z_T(\omega) \mathbb{P}(d\omega)$$

sur $(\Omega, \mathcal{F}_T^W)$, qui est une probabilité équivalente à \mathbb{P} sur \mathcal{F}_T^W . La loi de W sous \mathbb{Q}_T^θ est alors donnée par le théorème suivant, lequel peut être vu comme une généralisation de la formule de Cameron-Martin au cas $\theta \neq m$:

Théorème 6.3 [Théorème de Girsanov] *Sous la mesure \mathbb{Q}_T^θ , le processus*

$$\tilde{W} : t \mapsto W_t - \int_0^t \theta_s ds, \quad t \leq T$$

est un mouvement brownien.

Ce théorème s'avère extrêmement utile en pratique. La preuve s'établit de deux façons. L'une repose sur le lemme 6.1. On démontre d'abord directement que $Z\tilde{W}$ est une martingale sous \mathbb{P} , de sorte que par le lemme 6.1, \tilde{W} est une martingale sous \mathbb{Q}_T^θ . De plus, le crochet sous \mathbb{P} de \tilde{W} vu comme processus d'Itô est celui de sa partie martingale, donc $\langle \tilde{W} \rangle_t = t$. Mais on a vu que ce crochet était invariant par changement de probabilité équivalent, et l'on en déduit que sous \mathbb{Q}_T^θ , \tilde{W} est une martingale continue de crochet $\langle \tilde{W} \rangle_t = t$, autrement dit un mouvement brownien. \square

La deuxième méthode repose sur le théorème de Girsanov abstrait, d'intérêt moins évident a priori, mais qui reste valable dans un cadre très général.

Théorème 6.4 [Théorème de Girsanov abstrait] *Soit \mathbb{P} et \mathbb{Q} deux mesures de probabilité équivalentes sur un espace filtré $(\Omega, \{\mathcal{F}_t, t \leq T\})$. On suppose que toutes les (\mathcal{F}_t) -martingales sont continues. Alors sous \mathbb{P} il existe L une (\mathcal{F}_t) -martingale continue telle que pour tout $t \leq T$ et tout $A \in \mathcal{F}_t$,*

$$\mathbb{Q}[A] = \mathbb{E}_{\mathbb{P}}[\exp[L_t - \langle L \rangle_t/2] \mathbb{1}_A].$$

De plus, si M une martingale locale continue sous \mathbb{P} , alors le processus

$$\tilde{M} : t \mapsto M_t - \langle M, L \rangle_t$$

est une martingale locale continue sous \mathbb{Q} .

Expliquons d'abord la première partie de ce théorème concernant l'existence de la martingale L . Elle repose sur le théorème de Radon-Nikodym, qui assure l'existence d'un processus densité $\{D_t, t \geq 0\}$ de \mathbb{P} par rapport à \mathbb{Q} , au sens où pour tout $t \leq T$ et tout $A \in \mathcal{F}_t$,

$$\mathbb{Q}[A] = \mathbb{E}_{\mathbb{P}}[D_t \mathbb{1}_A].$$

Comme $A \in \mathcal{F}_T$, on a aussi

$$\mathbb{Q}[A] = \mathbb{E}_{\mathbb{P}}[D_T \mathbb{1}_A] = \mathbb{E}_{\mathbb{P}}[\mathbb{E}_{\mathbb{P}}[D_T | \mathcal{F}_t] \mathbb{1}_A]$$

de sorte, par identification, que

$$D_t = \mathbb{E}_{\mathbb{P}}[D_T | \mathcal{F}_t]$$

pour tout $t \leq T$. On en déduit que D est une martingale UI continue sous \mathbb{P} . De plus elle est strictement positive, par équivalence entre \mathbb{P} et \mathbb{Q} . On peut alors (on peut effectivement définir une intégrale stochastique par rapport à une martingale continue) définir le processus

$$L_t = \log D_0 + \int_0^t \frac{dD_s}{D_s}$$

et on applique la formule d'Itô à

$$\begin{aligned} R_t = \exp[L_t - \langle L \rangle_t/2] &= D_0 + \int_0^t R_s dL_s - \frac{1}{2} \int_0^t R_s d\langle L \rangle_s + \frac{1}{2} \int_0^t R_s d\langle L \rangle_s \\ &= D_0 + \int_0^t R_s dL_s. \end{aligned}$$

Mais par définition de L , on a aussi

$$D_t = D_0 + \int_0^t D_s dL_s,$$

de sorte que D et R sont solutions fortes de la même EDS. Par unicité, on en déduit que $R \equiv D$, d'où

$$D_t = \exp[L_t - \langle L \rangle_t/2].$$

Ceci entraîne finalement que pour tout $t \leq T$ et tout $A \in \mathcal{F}_t$,

$$\mathbb{Q}[A] = \mathbb{E}_{\mathbb{P}}[\exp[L_t - \langle L \rangle_t/2] \mathbb{1}_A].$$

□

On voit enfin facilement comment le théorème 6.4 entraîne le théorème 6.3. La martingale L du théorème 6.4 est ici

$$L_t = \int_0^t \theta_s dB_s$$

et on a donc

$$\langle B, L \rangle_t = \int_0^t \theta_s ds.$$

Comme B est une martingale locale continue sous \mathbb{P} , \tilde{B} est une martingale locale continue sous \mathbb{Q}_T^θ . On démontre comme ci-dessus que son crochet est $\langle \tilde{B} \rangle_t = t$, et l'on en déduit que \tilde{B} est un mouvement brownien sous \mathbb{Q}_T^θ .

□

6.4 Théorème de représentation prévisible

Soit B un mouvement brownien et \mathbf{F} sa filtration naturelle, soit $\mathcal{F}_t^B = \sigma(B_s, s \leq t)$. (il est important que (\mathcal{F}_t^B) soit la filtration naturelle).

6.4.1 Représentation prévisible

Théorème 6.5 Soit B un mouvement Brownien et (\mathcal{F}_t^B) sa filtration naturelle. Soit M une (\mathcal{F}_t^B) -martingale, telle que $\sup_{t \leq T} E[M_t^2] < \infty$. Il existe un unique processus prévisible H vérifiant $E(\int_0^T H_s^2 ds) < \infty$, tel que

$$\forall t \in [0, T], \quad M_t = M_0 + \int_0^t H_s dB_s.$$

Si M est une (\mathcal{F}_t) -martingale locale, il existe un unique processus prévisible H tel que $\int_0^T H_s^2 ds < \infty$ et

$$\forall t \quad M_t = M_0 + \int_0^t H_s dB_s$$

Preuve : On montre tout d'abord que si F est une v.a. \mathcal{F}_∞ mesurable, de carré intégrable, elle admet la représentation

$$F = E(F) + \int_0^\infty H_s dB_s$$

Pour cela, notant \mathcal{H} l'ensemble de v.a. possédant la propriété de représentation, on montre que \mathcal{H} est fermé dans L^2 et contient les v.a. de la forme $F = \mathcal{E}(f \star B)_\infty$ avec $f = \sum_i \lambda_i \mathbb{1}_{]t_{i-1}, t_i]}$, qui sont totales dans L^2 . On en déduit le résultat pour des martingales bornées, puis le cas général par densité et localisation. \square

Remarque 6.6 Ce résultat est important en finance pour exhiber un portefeuille de couverture. On remarque que $d\langle M, B \rangle_t = H_t dt$. Si M est une martingale de la forme $f(t, B_t)$, on trouve $H_t = f'_x(t, B_t)$ en appliquant la formule d'Itô. Ce théorème se généralise aux Browniens d -dimensionnels. Les conditions d'intégrabilité exigées sur H sont importantes. Dudley a montré que pour toute v.a. $\zeta \in \mathcal{F}_T$ on peut trouver un processus prévisible H tel que $\zeta = \int_0^T H_s dB_s$. (le cas où ζ est positif non nul est important en finance, puisqu'il génère un arbitrage).

Corollaire 6.7 Toutes les (\mathcal{F}_t^B) -martingales locales sont continues.

6.5 Applications au mouvement brownien et aux EDS

6.5.1 Calcul d'espérances Exemple 1

Le théorème de Girsanov permet de calculer diverses espérances de fonctionnelles du mouvement brownien. Par exemple on peut s'intéresser à

$$\mathbb{E} \left[B_t \exp \left[\int_0^T \theta_s dB_s - \frac{1}{2} \int_0^T \theta_s^2 ds \right] \right]$$

pour $t < T$ et θ fonction déterministe. On effectue un changement de probabilité en posant

$$L_t = \exp \left[\int_0^t \theta_s dB_s - \frac{1}{2} \int_0^t \theta_s^2 ds \right]$$

et on calcule

$$\begin{aligned} \mathbb{E}_{\mathbb{P}} [B_t L_T] &= \mathbb{E}_{\mathbb{P}} [B_t L_t] = \mathbb{E}_{\mathbb{Q}} [B_t] = \mathbb{E}_{\mathbb{Q}} \left[\tilde{B}_t + \int_0^t \theta_s ds \right] \\ &= \mathbb{E}_{\mathbb{Q}} \left[\int_0^t \theta_s ds \right] = \int_0^t \theta_s ds. \end{aligned}$$

On en déduit que

$$\mathbb{E} \left[B_t \exp \left[\int_0^T \theta_s dB_s - \frac{1}{2} \int_0^T \theta_s^2 ds \right] \right] = \int_0^t \theta_s ds$$

et en faisant en particulier $\theta \equiv 1$, que

$$\mathbb{E} [B_t \exp B_t] = te^{t/2},$$

ce qu'il aurait été plus difficile de calculer directement.

6.5.2 Calcul d'espérances Exemple 2

Intéressons-nous maintenant à un autre exemple plus élaboré, le calcul de

$$I = \mathbb{E} \left[\exp - \left[\alpha B_t^2 + \frac{\beta^2}{2} \int_0^t B_s^2 ds \right] \right]$$

où B est un brownien issu de a . On pose $x = a^2$ et on effectue un changement de probabilité $\mathbb{P} \rightarrow \mathbb{P}^\beta$ avec densité sur \mathcal{F}_t^W

$$\frac{d\mathbb{P}^\beta}{d\mathbb{P}} = L_t^\beta = \exp - \left[\frac{\beta}{2} (B_t^2 - x - t) + \frac{\beta^2}{2} \int_0^t B_s^2 ds \right].$$

En utilisant la formule d'intégration par parties, on calcule

$$L_t^\beta = \exp - \left[\beta \int_0^t B_s dB_s + \frac{\beta^2}{2} \int_0^t B_s^2 ds \right]$$

et on vérifie avec la proposition 5.11 que L^β est une martingale sous \mathbb{P} . Sous \mathbb{P}^β , le théorème 6.3 entraîne que

$$B_t = a + W_t - \beta \int_0^t B_s ds$$

avec W Brownien. Donc B est un *processus d'Ornstein-Uhlenbeck* sous \mathbb{P}^β , et l'on sait alors par le paragraphe 4.2.1 que B_t une v.a. gaussienne d'espérance $ae^{-\beta t}$ et de variance $\frac{1}{2\beta}(1 - e^{-2\beta t})$. On en déduit que

$$I = \mathbb{E}^\beta \left[L_t^{-1} \exp - \left[\alpha B_t^2 + \frac{\beta^2}{2} \int_0^t B_s^2 ds \right] \right] = \mathbb{E}^\beta \left[\exp \left[-\alpha B_t^2 + \frac{\beta}{2} (B_t^2 - x - t) \right] \right].$$

Après quelques calculs simples et longs, on obtient

$$I = (\cosh \beta t + 2\alpha \sinh \beta t / \beta)^{-1/2} \exp \left[-\frac{x\beta(1 + 2\alpha \coth \beta t / \beta)}{2(\cosh \beta t + 2\alpha / \beta)} \right].$$

En faisant $a = \alpha = 0$ et $\beta = \sqrt{2\lambda}$, cette formule donne la transformée de Laplace de la norme L^2 du Brownien :

$$\mathbb{E} \left[\exp - \lambda \int_0^t B_s^2 ds \right] = \left(\cosh \sqrt{2\lambda} t \right)^{-1/2}.$$

Signalons pour terminer que la simplicité de cette formule provient en fait du caractère hilbertien de L^2 . Les choses deviennent nettement plus compliquées quand on s'intéresse aux quantités

$$\mathbb{E} \left[\exp - \lambda \int_0^t B_s^p ds \right]$$

avec $p \neq 2$.

6.5.3 Temps de passage du Brownien drifté

On considère W un mouvement Brownien standard issu de 0 et T_b son premier temps de passage au seuil b :

$$T_b = \inf \{t > 0, W_t = b\}.$$

Au chapitre 3, nous avons vu que la densité de T_b est donnée par

$$\mathbb{P}[T_b \in dt] = \frac{|b|}{\sqrt{2\pi t^3}} \exp - \left[\frac{b^2}{2t} \right] dt.$$

En particulier $\mathbb{P}[T_b < +\infty] = 1$, ce qui signifie que p.s. le Brownien franchit tout seuil réel en temps fini. On considère maintenant

$$T_b^\mu = \inf \{t > 0, W_t + \mu t = b\},$$

premier temps de passage au seuil b du Brownien drifté $W^\mu : t \mapsto W_t + \mu t$, pour $\mu \in \mathbb{R}$. On définit

$$\mathbb{P}^\mu(d\omega) = \exp[\mu W_t - \mu^2 t/2] \mathbb{P}(d\omega)$$

sur \mathcal{F}_t^W . La formule de Cameron-Martin entraîne alors que le processus $W^{(-\mu)}$ avec $W_t^{(-\mu)} = W_t - \mu t$ est un Brownien sous \mathbb{P}^μ , donc que la loi de W sous \mathbb{P}^μ est celle de W^μ sous \mathbb{P} : utiliser que $W_t = W_t^{(-\mu)} + \mu t$ est, sous \mathbb{P}^μ un MB de drift μ . Ceci permet de calculer la densité de T_b^μ en écrivant

$$\begin{aligned} \mathbb{P}[T_b^\mu \leq t] &= \mathbb{P}^\mu[T_b \leq t] = \mathbb{E}[\exp[\mu W_t - \mu^2 t/2] \mathbb{1}_{\{T_b \leq t\}}] \\ &= \mathbb{E}[\mathbb{1}_{\{T_b \leq t\}} \mathbb{E}(\exp[\mu W_t - \mu^2 t/2] | \mathcal{F}_{T_b})] \\ &= \mathbb{E}[\exp[\mu W_{T_b} - \mu^2 T_b/2] \mathbb{1}_{\{T_b \leq t\}}] \\ &= e^{\mu b} \mathbb{E}[\exp - [\mu^2 T_b/2] \mathbb{1}_{\{T_b \leq t\}}] \\ &= e^{\mu b} \int_0^t e^{-\mu^2 s/2} \left(\frac{|b| e^{-b^2/2s}}{\sqrt{2\pi s^3}} \right) ds \\ &= \int_0^t \frac{|b|}{\sqrt{2\pi s^3}} \exp - \left[\frac{(b - \mu s)^2}{2s} \right] ds, \end{aligned}$$

où dans la deuxième formule on a utilisé le théorème d'arrêt de Doob. Ceci entraîne par dérivation que

$$\mathbb{P}[T_b^\mu \in dt] = \frac{|b|}{\sqrt{2\pi t^3}} \exp - \left[\frac{(b - \mu t)^2}{2t} \right] dt.$$

On peut aussi calculer

$$\mathbb{P}[T_b^\mu < +\infty] = e^{\mu b} \mathbb{E}[\exp - [\mu^2 T_b/2]] = e^{\mu b - |\mu b|},$$

de sorte que $\mathbb{P}[T_b^\mu < +\infty] = 1$ si $\mu b \geq 0$ et $\mathbb{P}[T_b^\mu < +\infty] < 1$ si $\mu b < 0$. Autrement dit, pour tout $b < 0$, il existe des trajectoires browniennes driftées vers le haut qui restent toujours au dessus du seuil b . Enfin, soit par calcul direct avec la densité, soit en utilisant une nouvelle fois la formule de Cameron-Martin, on peut calculer la transformée de Laplace de T_b^μ :

$$\mathbb{E}[\exp - \lambda T_b^\mu] = \exp \left[\mu b - |b| \sqrt{\mu^2 + 2\lambda} \right].$$

Le temps de passage au seuil b du Brownien drifté W^μ peut aussi être interprété comme le premier temps où la courbe brownienne touche la *frontière* donnée par la courbe $t \mapsto b - \mu t$.

6.5.4 Solutions faibles d'EDS

Le théorème de Girsanov permet de démontrer l'existence d'une solution faible à des EDS n'admettant pas forcément de solution forte. Considérons par exemple

$$X_t = \int_0^t a(X_s) ds + B_t \tag{6.1}$$

où B est un Brownien réel et $a : \mathbb{R} \rightarrow \mathbb{R}$ une fonction bornée. La condition sur a est trop faible pour pouvoir appliquer le théorème d'existence lipschitzien, et on peut construire une fonction a telle que (6.1) n'admette pas de solution forte. En revanche, si on part de l'espace de probabilité $(\Omega, \{\mathcal{F}_t^B, t \geq 0\}, \mathbb{P})$ et si on pose

$$Z_t = \exp \left[\int_0^t a(B_s) dB_s - \frac{1}{2} \int_0^t a^2(B_s) ds \right],$$

alors Z est une \mathcal{F}_t^B -martingale sous \mathbb{P} , puisque le caractère borné de a entraîne

$$\exp \left[\frac{1}{2} \int_0^t a^2(B_s) ds \right] \leq e^{t\|a\|_\infty^2/2} < +\infty$$

pour tout $t \geq 0$, de sorte que la condition de Novikov est bien remplie. Le théorème de Girsanov entraîne que sous \mathbb{Q}^a définie par

$$\mathbb{Q}^a(d\omega) = Z_t(\omega)\mathbb{P}(d\omega)$$

sur \mathcal{F}_t^B , le processus

$$W_t = B_t - \int_0^t a(B_s) ds$$

est un mouvement brownien. Ainsi, sous \mathbb{Q}^a , le processus B est solution de

$$X_t = W_t + \int_0^t a(X_s) ds$$

qui est précisément l'équation (6.1). On a donc construit une probabilité \mathbb{Q}^a et un processus (B, W) tels que W soit un mouvement brownien et B une solution de (6.1) sous \mathbb{Q}^a : on a donc construit une solution faible de (6.1). Ce fait met en évidence l'effet *régularisant* du Brownien W dans (6.1), si l'on songe que l'EDO

$$x_t = \int_0^t a(x_s) ds$$

n'admet pas de solution en général quand a est seulement supposée bornée. Une extension de ce résultat a été donnée par Beneš dans le cas d'un coefficient de diffusion non constant :

Théorème 6.8 [Théorème de Beneš] *Soit $a : \mathbb{R} \rightarrow \mathbb{R}$ une fonction bornée et $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ une fonction globalement lipschitzienne et ne s'annulant pas. Soit B un mouvement brownien réel. L'équation*

$$X_t = \int_0^t a(X_s) ds + \int_0^t \sigma(X_s) dB_s$$

admet une solution faible qui est unique en loi.

Dans ce théorème, l'hypothèse cruciale est que le coefficient de diffusion ne s'annule pas. Elle signifie que le processus "diffuse tout le temps", ce qui permet d'appliquer le théorème de Girsanov. D'une manière générale, les conditions de non-annulation du coefficient de diffusion jouent un rôle fondamental dans l'étude de la régularité des solutions d'EDS, via la théorie du *calcul de Malliavin* [13].

6.6 Théorème fondamental de la finance : Probabilité risque neutre

Une opportunité d'arbitrage est une stratégie d'investissement qui permet, à partir d'une mise de fonds initiale nulle, d'obtenir une richesse terminale positive, non nulle. Dans le cas d'un marché où sont négociés un actif sans risque, c'est-à-dire d'un actif dont le prix suit la dynamique

$$dS_t^0 = S_t^0 r dt$$

et un actif risqué de dynamique

$$dS_t = S_t(\mu_t dt + \sigma_t dB_t)$$

une stratégie d'investissement est un couple (π^0, π) de processus adaptés. La quantité $\pi_t^0 S_t^0$ est le montant monétaire (positif ou négatif) placé sur l'actif sans risque, π_t (positif ou négatif) est le nombre de parts d'actif financier dont l'agent dispose à la date t . La richesse associée est $X_t = \pi_t^0 S_t^0 + \pi_t S_t$. La stratégie est dite auto-finançante si $dX_t = \pi_t^0 dS_t^0 + \pi_t dS_t$ (ne pas confondre avec la formule d'Itô). Il en résulte que

$$dX_t = rX_t dt + \pi_t(dS_t - rS_t dt).$$

On généralise immédiatement au cas où il y a plusieurs actifs risqués.

Définition 6.9 Une probabilité risque neutre est une probabilité \mathbb{Q} équivalente à la probabilité \mathbb{P} , telle que les prix des actifs financiers actualisés par l'actif sans risque, soit $S_t^{(i)}/S_t^0$ soient des \mathbb{Q} -martingales.

Dans le cas où il y a un seul actif risqué, ceci signifie que

$$S_t \exp\left(-\int_0^t r_s ds\right) = S_t R_t$$

est une martingale. On voit immédiatement que si \mathbb{Q} existe

$$d\mathbb{Q}|_{\mathcal{F}_t} = \exp\left(-\int_0^t \theta_s dB_s - \frac{1}{2} \int_0^t \theta_s^2 ds\right)$$

avec $\theta_s = \frac{\mu_s - r_s}{\sigma_s}$ est la prime de risque.

Si \mathbb{Q} existe, la valeur de tout portefeuille auto-finançant (soit X), actualisée est une martingale.

Dans un modèle de prix d'actifs financiers, on doit veiller à ce que le modèle ne présente pas d'opportunités d'arbitrage. Dans un marché comportant un actif sans risque, c'est-à-dire d'un actif dont le prix suit la dynamique

$$dS_t^0 = S_t^0 r_t dt$$

le théorème fondamental s'énonce

Théorème 6.10 Le marché est sans arbitrage si et seulement si il existe une probabilité risque-neutre.

Si la probabilité risque neutre est unique, la valeur d'un actif financier H , payé à la date T , est calculée comme l'espérance sous la probabilité risque neutre du payoff actualisé, soit $V_t = \mathbb{E}_{\mathbb{Q}}(H R_T)$, où $R_T = \int_0^T r_s ds$. On parle alors de marché complet (cette notion est liée au théorème de représentation). Si la probabilité risque neutre n'est pas unique, on parle de marché incomplet. On peut se référer à l'ouvrage de Bjork ou à celui de Dana.

6.6.1 Changement de numéraire

Il est parfois très utile d'exprimer les prix en valeur relative par rapport à un autre processus de prix. **numéraire**.

Définition 6.11 Un numéraire est un actif financier de prix strictement positif.

Si M est un numéraire (par exemple le prix d'un zéro-coupon) on peut évaluer la valeur V_t d'une stratégie en terme de ce numéraire, soit V_t/M_t . Il est important de vérifier que les propriétés fondamentales ne sont pas modifiées :

Proposition 6.12 Supposons qu'il y a d actifs risqués dans le marché, dont les prix $(S_t^{(i)}; i = 1, \dots, d, t \geq 0)$ sont des processus d'Itô, et tels que $S^{(1)}$ est strictement positif. Soit $V_t = \sum_{i=1}^d \pi_t^i S_t^{(i)}$ le valeur du portefeuille $\pi_t = (\pi_t^i, i = 1, \dots, d)$. Si $(\pi_t, t \geq 0)$ est auto-financiant, i.e. si $dV_t = \sum_{i=1}^d \pi_t^i dS_t^{(i)}$, et si l'on choisit $S_t^{(1)}$ comme numéraire, alors

$$dV_t^1 = \sum_{i=2}^d \pi_t^i dS_t^{(i,1)}$$

où $V_t^1 = V_t/S_t^{(1)}$, $S_t^{(i,1)} = S_t^{(i)}/S_t^{(1)}$.

Preuve : Nous donnons la preuve pour deux actifs. Soit V la valeur d'un portefeuille auto-finançant construit sur deux actifs de dynamique $S^{(i)}, i = 1, 2$. Alors

$$\begin{aligned} dV_t &= \pi_t^1 dS_t^{(1)} + \pi_t^2 dS_t^{(2)} = \pi_t^2 dS_t^{(2)} + (V_t - \pi_t^2 S_t^{(2)}) dS_t^{(1)} / S_t^{(1)} \\ &= \pi_t^2 dS_t^{(2)} + (V_t^1 - \pi_t^2 S_t^{(2,1)}) dS_t^{(1)}. \end{aligned}$$

On exprime les prix en terme du numéraire $S^{(1)}$. A partir de $V_t^1 S_t^{(1)} = V_t$ on obtient

$$dV_t = V_t^1 dS_t^{(1)} + S_t^{(1)} dV_t^1 + d\langle S^{(1)}, V^1 \rangle_t. \quad (6.2)$$

L'égalité $S_t^{(2,1)} S_t^{(1)} = S_t^{(2)}$ implique

$$dS_t^{(2)} - S_t^{(2,1)} dS_t^{(1)} = S_t^{(1)} dS_t^{(2,1)} + d\langle S^{(1)}, S^{(2,1)} \rangle_t$$

d'où en utilisant (6.2)

$$\begin{aligned} dV_t^1 &= \frac{1}{S_t^{(1)}} \left(dV_t - V_t^1 dS_t^{(1)} - d\langle V^1, S^{(1)} \rangle_t \right) \\ &= \frac{1}{S_t^{(1)}} \left(\pi_t^2 dS_t^{(2)} - \pi_t^2 S_t^{(2,1)} dS_t^{(1)} - d\langle V^1, S^{(1)} \rangle_t \right) \\ &= \pi_t^2 dS_t^{(2,1)} + \frac{\pi_t^2}{S_t^{(1)}} d\langle S^{(2,1)}, S^{(1)} \rangle_t - \frac{1}{S_t^{(1)}} d\langle V^1, S^{(1)} \rangle_t \end{aligned}$$

Cette dernière égalité implique $d\langle V^1, S^{(1)} \rangle_t = \pi_t^2 d\langle S^{(2,1)}, S^{(1)} \rangle_t$, d'où $dV_t^1 = \pi_t^2 dS_t^{(2,1)}$. □

On associe à un numéraire M le changement de probabilité suivant : Soit Q la probabilité risque neutre, telle que le processus $(M_t R_t, t \geq 0)$ est une Q -martingale. Soit Q^M défini par $Q^M|_{\mathcal{F}_t} = (M_t R_t)Q|_{\mathcal{F}_t}$.

Proposition 6.13 *Soit $(X_t, t \geq 0)$ le prix d'un actif financier et M un numéraire. Le prix de X , dans le numéraire M , soit $(X_t/M_t, 0 \leq t \leq T)$ est une Q^M -martingale.*

Preuve : Si X est un processus de prix, le processus actualisé $\tilde{X}_t \stackrel{\text{def}}{=} X_t R_t$ est une Q -martingale. Il en résulte que X_t/M_t est une Q^M -martingale si et seulement si $(X_t/M_t)M_t R_t = R_t X_t$ est une Q -martingale. □

Le calcul du terme $E_Q(S_T e^{-rT} \mathbb{1}_{S_T \geq a})$ qui apparait dans la formule de Black et Scholes est immédiat : il suffit d'utiliser que, sous Q le processus $M_t = S_t e^{-rt}/S_0$ est une martingale strictement positive d'espérance 1, et de poser $d\hat{Q} = M_t dQ$

$$\begin{aligned} E_Q(S_T e^{-rT} \mathbb{1}_{S_T \geq a}) &= E_{\hat{Q}}(S_0 \mathbb{1}_{S_T \geq a}) \\ &= S_0 \hat{Q}(S_T \geq a). \end{aligned}$$

Il reste à exprimer la dynamique de S sous \hat{Q} .

Un changement de numéraire est également très efficace pour calculer le prix d'une option d'échange, qui est $E((S_T^1 - S_T^2)^+)$.

Un exemple important : Probabilité forward-neutre

La valeur à la date t d'un flux déterministe F reçu à la date T est

$$FP(t, T) = F E_Q[\exp - \int_t^T r(u) du | \mathcal{F}_t].$$

Si ce flux est aléatoire, la valeur à la date t de ce flux est

$$E_Q[F \exp - \int_t^T r(u) du | \mathcal{F}_t].$$

Par hypothèse A.O.A, le processus $R(t)P(t, T)$ est une Q -martingale, son espérance est constante, égale à $P(0, T)$.

Pour tout T , le processus $\zeta_t^T := \frac{R(t)P(t, T)}{P(0, T)}$ est une Q -martingale positive d'espérance 1. On peut donc utiliser ζ_t^T comme densité de changement de probabilité. Soit Q_T la mesure de probabilité définie sur (Ω, \mathcal{F}_T) par $Q_T(A) = E_Q(\zeta_t^T 1_A)$ pour tout $A \in \mathcal{F}_t$. Lorsque T est fixé, on notera $\zeta_t = \zeta_t^T$.

Définition 6.14 La probabilité Q_T définie sur \mathcal{F}_T , par $\frac{dQ_T}{dQ} = \zeta_t^T$ est appelée probabilité forward-neutre de maturité T .

Avec cette notation

$$E_{Q_T}(F | \mathcal{F}_t) = E_Q(F \frac{\zeta_T}{\zeta_t} | \mathcal{F}_t).$$

Lorsque r est déterministe, $Q_T = Q$.

La mesure Q_T est la martingale mesure associée au choix du zéro-coupon de maturité T comme numéraire.

Chapitre 7

Introduction aux modèles poissonniens

Avec le mouvement Brownien, le processus de Poisson est l'autre processus fondamental du calcul stochastique. Une des raisons principales de cette importance est la formule de Lévy-Khintchine, que nous n'aborderons pas ici, selon laquelle tout P.A.I.S s'écrit comme le mélange d'un Brownien et d'une certaine intégrale de comptage le long d'un processus de Poisson. Nous renvoyons par exemple au chapitre introductif de [1] pour plus de détails sur cette formule, d'intérêt constant en Finance. Nous ne traiterons dans ce chapitre que le cas simple des processus de Poisson ponctuels d'intensité *finie*. Les *distributions exponentielles* de paramètre $\lambda > 0$, dont nous rappelons que la densité s'écrit

$$x \mapsto \lambda e^{-\lambda x} \mathbb{1}_{\{x \geq 0\}},$$

il y joueront un rôle central, analogue à celui des distributions gaussiennes pour le Brownien. A la fin du chapitre, nous donnerons un analogue de la formule d'Itô pour le processus de Poisson. En fait, il s'agit d'une formule de changement de variable entièrement déterministe, et qui ne nécessite aucune hypothèse particulière sur la fonction f .

7.1 Processus ponctuels sur \mathbb{R}^+

Un processus ponctuel est la donnée d'une suite ordonnée de variables aléatoires positives

$$0 = T_0 < T_1 < T_2 < \dots < T_n < \dots$$

qui modélisent les instants d'arrivée (ou les instants de saut) d'un certain phénomène. On fait l'hypothèse que deux instants d'arrivée ne peuvent pas coïncider, soit $T_{n+1} \neq T_n$ p.s. On suppose aussi que la suite ne s'accumule pas en un point, autrement dit $T_n \uparrow +\infty$ p.s. Pour tout $i \geq 1$ on note $\tau_i = T_i - T_{i-1}$ le i -ème *délai d'arrivée*. On remarque la formule

$$T_n = \sum_{i=1}^n \tau_i. \tag{7.1}$$

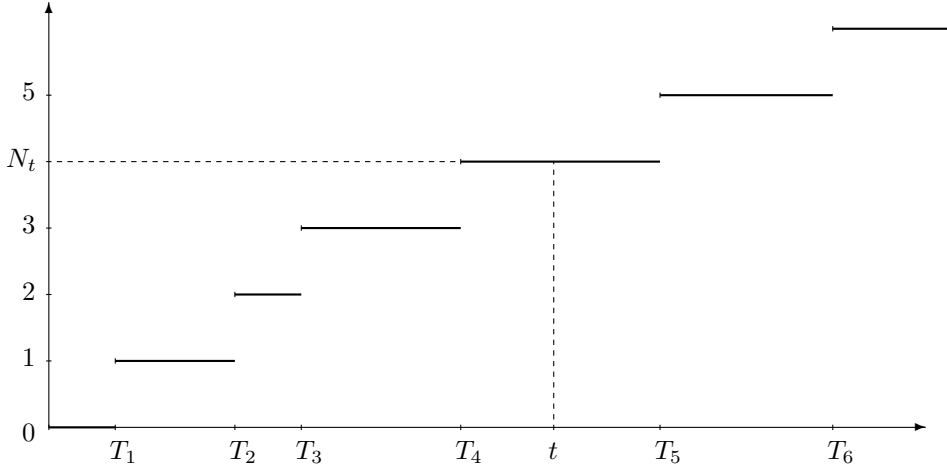
Pour tout $t \geq 0$ on introduit

$$N_t = \sum_{n \geq 1} \mathbb{1}_{\{T_n \leq t\}}$$

et on dit que $\{N_t, t \geq 0\}$ est le *processus de comptage* associé à la suite $\{T_n, n \geq 0\}$. C'est un processus à valeurs dans \mathbb{N} , qui compte le nombre de sauts ayant eu lieu avant t . En connaissant ce processus, on peut retrouver la suite $\{T_n, n \geq 0\}$ grâce à l'identification entre événements :

$$\{N_t = n\} = \{T_n \leq t < T_{n+1}\}.$$

Cette identification est immédiate en faisant un dessin :



On a aussi trois autres identifications entre événements, toutes également utiles :

$$\{N_t < n\} = \{T_n > t\}, \{N_t \geq n\} = \{T_n \leq t\} \text{ et } \{N_s < n \leq N_t\} = \{s < T_n \leq t\}.$$

On fera attention que $\{N_t, t \geq 0\}$ est un processus *continu à droite* seulement, et non pas continu. Dans les événements ci-dessus, le fait que les inégalités soient strictes ou larges n'est donc pas anodin. Remarquons enfin que $N_t < +\infty$ p.s. pour tout $t > 0$. En effet, on a

$$\{N_t < +\infty\} = \bigcup_{n \geq 1} \{N_t < n\} = \bigcup_{n \geq 1} \{T_n > t\} = \lim_{n \rightarrow +\infty} \{T_n > t\}$$

et le dernier événement est de probabilité 1 puisque $T_n \uparrow +\infty$ p.s. par hypothèse.

7.2 L'hypothèse de Poisson

Cette hypothèse formule une hypothèse d'indépendance et de stationnarité sur le processus $\{N_t, t \geq 0\}$ identique à celle du Brownien :

Définition 7.1 Soit $\{T_n, n \geq 0\}$ un processus ponctuel sur \mathbb{R}^+ . On dit que c'est un processus ponctuel de Poisson (P.P.P) si le processus de comptage associé $\{N_t, t \geq 0\}$ est un P.A.I.S. Autrement dit, si $\mathcal{F}_t^N = \sigma\{N_s, s \leq t\} \vee \mathcal{N}$ désigne la filtration naturelle complétée du processus N , alors pour tout $t, s \geq 0$ la v.a. $N_{t+s} - N_t$ est indépendante de \mathcal{F}_t^N et a même loi que N_s .

Le théorème suivant précise la structure des P.P.P. sur \mathbb{R}^+ . C'est un cas particulier de la formule de Lévy-Khintchine évoquée dans l'introduction.

Théorème 7.2 [Théorème de structure poissonnien] Soit $\{T_n, n \geq 0\}$ un processus ponctuel de Poisson sur \mathbb{R}^+ . Alors il existe $\theta > 0$ tel que pour tout $t > 0$, la v.a. N_t suit la loi de Poisson d'intensité θt . Autrement dit, pour tout $t > 0, k \in \mathbb{N}$,

$$\mathbb{P}[N_t = k] = e^{-\theta t} \frac{(\theta t)^k}{k!}.$$

On dit que θ est l'intensité du processus $\{T_n, n \geq 0\}$.

Preuve : Rappelons que si $X \sim \text{Poisson}(\lambda)$, alors la fonction caractéristique de X est

$$\mathbb{E}[u^X] = \sum_{k=0}^{+\infty} u^k \mathbb{P}[X = k] = e^{-\lambda} \sum_{k=0}^{+\infty} u^k \frac{\lambda^k}{k!} = e^{\lambda(u-1)}$$

pour tout $u \in]0, 1]$. Par caractérisation, il suffit donc de montrer qu'il existe $\theta > 0$ tel que

$$\mathbb{E}[u^{N_t}] = e^{\theta t(u-1)}$$

pour tout $u \in]0, 1]$ et $t \geq 0$. On pose $f_t(u) = \mathbb{E}[u^{N_t}]$ pour tout $u \in]0, 1]$. C'est une fonction strictement positive et, par le théorème de convergence dominée de Lebesgue, continue à droite. Fixons $u \in]0, 1]$. En utilisant l'hypothèse d'indépendance et de stationnarité des accroissements de N_t , on remarque que pour tout $t, s \geq 0$,

$$\begin{aligned} f_t(u)f_s(u) &= \mathbb{E}[u^{N_t}] \mathbb{E}[u^{N_s}] = \mathbb{E}[u^{N_t}] \mathbb{E}[u^{N_{t+s}-N_t}] \\ &= \mathbb{E}[u^{N_t} u^{N_{t+s}-N_t}] = \mathbb{E}[u^{N_{t+s}}] = f_{t+s}(u). \end{aligned}$$

Ceci entraîne que à u fixé, la fonction $t \mapsto g_t(u) = \log(f_t(u))$ est additive au sens où

$$g_{t+s}(u) = g_t(u) + g_s(u)$$

pour tout $t, s \geq 0$. Comme elle est continue à droite, un théorème d'analyse nous dit qu'il existe un paramètre $\theta(u) \in \mathbb{R}$ tel que $g_t(u) = t\theta(u)$, autrement dit

$$f_t(u) = e^{t\theta(u)}$$

et il suffit de montrer que $\theta(u) = (u-1)\theta$ pour un certain $\theta > 0$ indépendant de u . Pour cela, on remarque que

$$\begin{aligned} \theta(u) &= \lim_{t \downarrow 0} \frac{e^{t\theta(u)} - 1}{t} = \lim_{t \downarrow 0} \frac{\mathbb{E}[u^{N_t}] - 1}{t} \\ &= \lim_{t \downarrow 0} \frac{1}{t} \left(\sum_{k=1}^{+\infty} (u^k - 1) \mathbb{P}[N_t = k] \right) \\ &= (u-1) \left(\lim_{t \downarrow 0} \frac{\mathbb{P}[N_t = 1]}{t} \right) + \lim_{t \downarrow 0} \left(\sum_{k=2}^{+\infty} (u^k - 1) \frac{\mathbb{P}[N_t = k]}{t} \right). \end{aligned}$$

Comme $u \in]0, 1]$, on a

$$\left| \sum_{k=2}^{+\infty} (u^k - 1) \frac{\mathbb{P}[N_t = k]}{t} \right| \leq \sum_{k=2}^{+\infty} \left(\frac{\mathbb{P}[N_t = k]}{t} \right) \leq \frac{\mathbb{P}[N_t \geq 2]}{t}$$

et nous allons montrer que le terme de droite tend vers 0 quand $t \rightarrow 0$. En effet, pour tout $t > 0$, on a l'inclusion entre événements :

$$\bigcup_{n \geq 0} \{N_{nt} = 0, N_{(n+1)t} \geq 2\} \subset \{T_2 < T_1 + t\}$$

la réunion à gauche étant constituée d'événements *disjoints* par croissance de N_t . D'où

$$\begin{aligned} \mathbb{P} \left[\bigcup_{n \geq 0} \{N_{nt} = 0, N_{(n+1)t} \geq 2\} \right] &= \sum_{n \geq 0} \mathbb{P}[N_{nt} = 0, N_{(n+1)t} \geq 2] \\ &= \sum_{n \geq 0} \mathbb{P}[N_{nt} = 0, N_{(n+1)t} - N_{nt} \geq 2] \\ &= \sum_{n \geq 0} \mathbb{P}[N_{nt} = 0] \mathbb{P}[N_t \geq 2] \\ &= \mathbb{P}[N_t \geq 2] \sum_{n \geq 0} f_{nt}(0) \\ &= \mathbb{P}[N_t \geq 2] \sum_{n \geq 0} e^{nt\theta(0)}. \end{aligned}$$

Comme la série est nécessairement convergente, ceci entraîne que $\theta(0) < 0$. On pose alors $\theta = -\theta(0) > 0$ et on trouve

$$\mathbb{P} \left[\bigcup_{n \geq 0} \{N_{nt} = 0, N_{(n+1)t} \geq 2\} \right] = \mathbb{P}[N_t \geq 2] (1 - e^{-t\theta}),$$

d'où

$$\begin{aligned} \lim_{t \downarrow 0} \frac{\mathbb{P}[N_t \geq 2]}{t} &= \lim_{t \downarrow 0} \left(\frac{1 - e^{-t\theta}}{t} \right) \mathbb{P} \left[\bigcup_{n \geq 0} \{N_{nt} = 0, N_{(n+1)t} \geq 2\} \right] \\ &\leq \theta \left(\lim_{t \downarrow 0} \mathbb{P}[T_2 < T_1 + t] \right) = 0 \end{aligned}$$

par croissance de la suite $\{T_n\}$. Ceci entraîne finalement que

$$\begin{aligned} \theta(u) &= (u-1) \left(\lim_{t \downarrow 0} \frac{\mathbb{P}[N_t = 1]}{t} \right) = (u-1) \left(\lim_{t \downarrow 0} \frac{\mathbb{P}[N_t \geq 1]}{t} \right) \\ &= (u-1) \left(\lim_{t \downarrow 0} \frac{(1 - \mathbb{P}[N_t = 0])}{t} \right) \\ &= (u-1) \left(\lim_{t \downarrow 0} \frac{(1 - e^{-\theta t})}{t} \right) = \theta(u-1) \end{aligned}$$

pour un certain $\theta > 0$, ce qui termine la preuve du théorème. □

Comme conséquence de l'indépendance et de la stationnarité des accroissements de N , on obtient immédiatement que la loi conditionnelle de la v.a. $(N_{t+s} - N_t)$ sachant \mathcal{F}_t^N est la loi de Poisson de paramètre θs , autrement dit que

$$\mathbb{P}[N_{t+s} - N_t = k \mid \mathcal{F}_t^N] = \mathbb{P}[N_s = k] = e^{-\theta s} \frac{(\theta s)^k}{k!}$$

pour tout $k \geq 0$. On peut aussi calculer la loi de T_1 en écrivant

$$\mathbb{P}[T_1 > t] = \mathbb{P}[N_t < 1] = \mathbb{P}[N_t = 0] = e^{-\theta t},$$

ce qui entraîne par comparaison des fonctions de répartition que

$$T_1 \sim \text{Exp}(\theta).$$

Enfin, on peut aussi déterminer la loi du délai avant le premier instant de saut de la suite $\{T_n\}$ ayant lieu après un instant fixé t . En observant la figure précédente, on constate que ce délai est donné par la variable

$$T_{N_t+1} - t.$$

Or, pour tout $s \geq 0$,

$$\mathbb{P}[T_{N_t+1} - t > s] = \mathbb{P}[T_{N_t+1} > t + s] = \mathbb{P}[N_{t+s} = N_t] = \mathbb{P}[N_s = 0] = e^{-\theta s}$$

ce qui entraîne, de même, que

$$T_{N_t+1} - t \stackrel{d}{=} T_1 \sim \text{Exp}(\theta).$$

Au paragraphe suivant, nous généraliserons ces résultats en calculant la loi de *toute* la suite $\{T_n, n \geq 0\}$.

Exercice 7.3 Soit $\{T_n, n \geq 1\}$ un processus ponctuel de Poisson et soit N le processus de comptage associé. Montrer que pour tout $t > 0$ la loi conditionnelle de T_1 sachant $\{N_t = 1\}$ est la loi uniforme sur $[0, t]$.

Exercice 7.4 (a) Soient R_1, \dots, R_n des variables uniformes sur $[0, t]$ et indépendantes. Soit S_1, \dots, S_n leur *réarrangement croissant*, au sens où $S_i = R_{\sigma(i)}$ pour tout $i = 1 \dots n$, σ étant une certaine permutation aléatoire de $[1, \dots, n]$ choisie de telle sorte que

$$S_1 \leq \dots \leq S_n.$$

Montrer que la densité du n -uplet (S_1, \dots, S_n) s'écrit

$$f(s_1, \dots, s_n) = \frac{n!}{t^n} \mathbb{1}_{\{s_1 < \dots < s_n\}}.$$

(b) Soit $\{T_n, n \geq 1\}$ un processus ponctuel de Poisson et soit N le processus de comptage associé. Montrer que pour tout $t > 0$ la loi conditionnelle du n -uplet (T_1, \dots, T_n) sachant $\{N_t = n\}$ est la même que celle du n -uplet (S_1, \dots, S_n) défini précédemment.

7.3 La propriété de Markov forte et ses conséquences

Soit $\{N_t, t \geq 0\}$ un processus de comptage de Poisson d'intensité $\theta > 0$. Par hypothèse, c'est un P.A.I.S. au sens où pour tous $s, t \geq 0$, la v.a. $N_{t+s} - N_t$ est indépendante de \mathcal{F}_t^N et a même loi que N_s , $\{\mathcal{F}_t^N, t \geq 0\}$ désignant la filtration naturelle complétée du processus N . Rappelons qu'un *temps d'arrêt* T pour cette filtration est une v.a. positive telle que

$$\{T \leq t\} \in \mathcal{F}_t^N$$

pour tout $t \geq 0$. Si T est un temps d'arrêt, on peut alors définir la σ -algèbre

$$\mathcal{F}_T^N = \{A \in \mathcal{F}_\infty^N / A \cap \{T \leq t\} \in \mathcal{F}_t^N \text{ pour tout } t \geq 0\}.$$

La propriété de Markov forte étend la propriété de P.A.I.S. aux (\mathcal{F}_t^N) -temps d'arrêt. Remarquons que dans ce cas précis, cette propriété est plus forte que la définition habituelle, puisqu'on a homogénéité aussi bien dans le temps que *dans l'espace*.

Théorème 7.5 *Pour tout temps d'arrêt T et tout temps déterministe $t \geq 0$, la v.a. $N_{T+t} - N_T$ est indépendante de \mathcal{F}_T^N et a même loi que N_t .*

Preuve : Nous montrons d'abord la propriété quand T prend une quantité dénombrable de valeurs $t_1 \dots t_n \dots$. Soit $A \in \mathcal{F}_T^N$, $p \geq 1$, $s_1, \dots, s_p \geq 0$ et $n_1, \dots, n_p \in \mathbb{N}$. On introduit la notation $N_s^t = N_{t+s} - N_t$ pour tous $t, s \geq 0$ et on écrit

$$\begin{aligned} \mathbb{P}[A, N_{s_i}^T = n_i, 1 \leq i \leq p] &= \sum_{j \geq 1} \mathbb{P}[A, T = t_j, N_{s_i}^{t_j} = n_i, 1 \leq i \leq p] \\ &= \sum_{j \geq 1} \mathbb{P}[A, T = t_j, N_{s_i}^{t_j} = n_i, 1 \leq i \leq p] \\ &= \sum_{j \geq 1} \mathbb{P}[A, T = t_j] \mathbb{P}[N_{s_i}^{t_j} = n_i, 1 \leq i \leq p] \\ &= \sum_{j \geq 1} \mathbb{P}[A, T = t_j] \mathbb{P}[N_{s_i} = n_i, 1 \leq i \leq p] \\ &= \mathbb{P}[A] \mathbb{P}[N_{s_i} = n_i, 1 \leq i \leq p] \end{aligned}$$

où dans la troisième ligne on a utilisé que $A \cap \{T = t_j\} \in \mathcal{F}_{t_j}^N$ et l'indépendance des accroissements de N , et dans la quatrième la stationnarité de ces accroissements. Ceci montre que le processus N^T est indépendant de A et a même loi que N pour tout $A \in \mathcal{F}_T^N$, et termine la preuve quand T prend des valeurs dénombrables. Dans le cas général, on approche T par des nombres dyadiques et on passe à la limite (nous laissons les détails ou renvoyons au chapitre 1 de [1] pour le lecteur intéressé).

□

Rappelons que la suite $\{\tau_n, n \geq 1\}$ des délais entre deux instants de saut d'un P.P.P. est donnée par la formule $\tau_n = T_n - T_{n-1}$. La propriété de Markov forte va nous permettre de calculer la loi de la suite de ces délais :

Corollaire 7.6 *La suite $\{\tau_n, n \geq 1\}$ est une suite i.i.d. avec $\tau_1 \sim \exp(\theta)$.*

Preuve : Nous avons déjà vu que $\tau_n \sim \tau_1 \sim \exp(\theta)$. Pour montrer que la suite est i.i.d. on écrit, pour tout $n \in \mathbb{N}$ et tout $t_1 \dots t_n \geq 0$,

$$\begin{aligned} \mathbb{P}[\tau_i > t_i, 1 \leq i \leq n] &= \mathbb{P}[N_{t_i}^{T_i} = 0, 1 \leq i \leq n] \\ &= \prod_{i=1}^n \mathbb{P}[N_{t_i} = 0] \\ &= \prod_{i=1}^n \mathbb{P}[\tau_i > t_i], \end{aligned}$$

ce qui termine la preuve. □

Comme conséquence de la formule (7.1), on en déduit que si $\{T_n, n \geq 1\}$ est un processus ponctuel de Poisson d'intensité $\theta > 0$, alors pour tout $n \geq 1$ T_n est la somme de n variables exponentielles indépendantes de paramètre $\theta > 0$.

7.4 Une formule d'Itô avec sauts

Nous terminons ce chapitre par une courte évocation du calcul stochastique discontinu, via la formule d'Itô avec sauts. Soit N un processus de comptage de Poisson. Rappelons que la fonction $t \mapsto N_t$ est une fonction càdlàg, constante par morceaux avec des sauts de taille 1. Soit $t \mapsto \theta_t$ un processus *quelconque*. On peut définir l'intégrale stochastique

$$I_t^\theta = \int_0^t \theta_s dN_s = \sum_{n \geq 1} \theta_{T_n} \mathbb{1}_{\{T_n \leq t\}},$$

ce que l'on va également écrire

$$I_t^\theta = \sum_{0 < s \leq t} \theta_s \Delta N_s.$$

Plus précisément, $I_t^\theta = 0$ sur $\{t < T_1\}$ (car N n'a pas varié), $I_t^\theta = \theta_{T_1}$ sur $\{T_1 \leq t < T_2\}$ (car N a sauté une fois en T_1 et n'a pas varié ensuite), $I_t^\theta = \theta_{T_1} + \theta_{T_2}$ sur $\{T_2 \leq t < T_3\}$ (car N a sauté une fois en T_1 et n'a pas varié, puis une fois en T_2 et n'a pas varié ensuite)... etc. Remarquons de plus que si $\theta \equiv 1$, on a

$$I_t^\theta = \sum_{n \geq 1} \mathbb{1}_{T_n \leq t} = N_t$$

de sorte que

$$N_t = \int_0^t dN_s$$

comme il est souhaitable. Soit maintenant $f : \mathbb{R} \rightarrow \mathbb{R}$ une fonction mesurable *quelconque*. On a $f(N_t) = f(0)$ si $t < T_1$, $f(N_t) = f(N_{T_1})$ si $T_1 \leq t < T_2$... ce qu'on peut récrire

$$\begin{aligned} f(N_t) &= f(0) + \sum_{i=1}^n (f(N_{T_i}) - f(N_{T_{i-1}})) \\ &= f(0) + \sum_{i=1}^n (f(N_{T_i}) - f(N_{T_i-})) \\ &= f(0) + \sum_{s \leq t} (f(N_s) - f(N_{s-})) \end{aligned}$$

sur l'événement $\{T_n \leq t < T_{n+1}\}$. Comme le dernier terme ne dépend pas de n on en peut en déduire le

Théorème 7.7 [Formule d'intégration pour le processus de Poisson] *Pour toute fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ mesurable on a*

$$f(N_t) = f(0) + \sum_{s \leq t} (f(N_s) - f(N_{s-}))$$

On remarque que la preuve de cette formule est beaucoup plus élémentaire que pour le Brownien. De plus aucune hypothèse particulière n'est faite sur f . Supposons maintenant que f soit de classe \mathcal{C}^1 . Avec la définition précédente pour l'intégrale, on a

$$\begin{aligned} \int_0^t f'(N_{s-}) dN_s &= \sum_{i=1}^n f'(N_{T_i-}) \\ &= \sum_{s \leq t} f'(N_{s-}) \Delta N_s \end{aligned}$$

sur l'événement $\{T_{n-1} \leq t < T_n\}$. Remarquons que le dernier terme ne dépend pas de n . En retranchant et en ajoutant on en déduit alors le

Théorème 7.8 [Deuxième formule d'intégration pour le processus de Poisson] *Pour toute fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 on a*

$$f(N_t) = f(0) + \int_0^t f'(N_{s-}) dN_s + \sum_{s \leq t} (f(N_s) - f(N_{s-}) - f'(N_{s-}) \Delta N_s).$$

Cette formule est plus compliquée mais son intérêt est qu'elle peut se généraliser à une classe plus grande de processus à sauts, les *processus de Lévy*, lesquels comportent en général un nombre infini de sauts. Nous renvoyons à [5] et à [16] pour plus de détails sur ce sujet. Dans le cas plus difficile où le processus de Lévy est à variation forte infinie, une vraie théorie de l'intégrale stochastique est nécessaire, et la clé en est la procédure de retranchement et d'ajout du terme comportant la dérivée vue ci-dessus, laquelle porte le nom de *compensation*. Ce mécanisme de compensation joue d'ailleurs un rôle fondamental dans toute la théorie générale des processus à sauts, voir [4].

7.5 Des martingales

Proposition 7.9 *Pour tout $\alpha \in \mathbb{R}, \beta > -1$, pour toute fonction bornée h , les processus suivants sont des (\mathcal{F}_t) -martingales :*

- (i) $M_t = N_t - \lambda t$,
- (ii) $M_t^2 - \lambda t = (N_t - \lambda t)^2 - \lambda t$,
- (iii) (a) $\exp(\alpha N_t + \lambda t(1 - e^\alpha))$,
- (b) $\exp[\ln(1 + \beta)N_t - \lambda \beta t] = (1 + \beta)^{N_t} e^{-\lambda \beta t}$,
- (c) $\exp\left[\int_0^t h(s) dN_s - \lambda \int_0^t (1 - e^{h(s)}) ds\right]$

Preuve : Soit $s < t$. En utilisant l'indépendance des accroissements du PP, on obtient

- (i) $E(M_t - M_s | \mathcal{F}_s) = E(N_t - N_s) - \lambda(t - s) = 0$
- (ii) $E(M_t^2 - M_s^2 | \mathcal{F}_s) = E[(M_t - M_s)^2 | \mathcal{F}_s] = E[(N_t - N_s + \lambda(t - s))^2 | \mathcal{F}_s]$
 $E[(N_t - N_s)^2] - \lambda^2(t - s)^2 = \lambda(t - s)$
- (iii) $E[\exp(\alpha(N_t - N_s) + \lambda(t - s)(1 - e^\alpha)) | \mathcal{F}_s] = 1$. Les autres formules sont faciles.

On peut généraliser : si H est un processus borné (\mathcal{F}_t) -prévisible, les processus suivants sont des martingales

$$M_t^H \stackrel{def}{=} \int_0^t H_s dN_s - \lambda \int_0^t H_s ds = \int_0^t H_s dM_s \quad (7.2)$$

$$Y_t^H \stackrel{def}{=} (M_t^H)^2 - \lambda \int_0^t H_s^2 ds \quad (7.3)$$

$$Z_t^H \stackrel{def}{=} \exp \left(\int_0^t H_s dN_s + \lambda \int_0^t (1 - e^{H_s}) ds \right) \quad (7.4)$$

7.5.1 Exemple de changement de probabilité

Si N est un PP d'intensité λ , alors, d'après la proposition 7.9 (iii,b) avec $\beta = 1$, le processus

$$L_t = 2^{N_t} e^{-\lambda t} = e^{N_t - \lambda t} \prod_{0 < s \leq t} (1 + \Delta N_s) e^{-\Delta N_s}$$

est une martingale et est solution de $dL_t = L_{t-} dM_t$. Soit Q la probabilité définie par $\frac{dQ}{dP}|_{\mathcal{F}_t} = L_t$. Il est facile de vérifier que

$$E^Q(x^{N_t}) = \exp(2\lambda t(x - 1))$$

et que le processus N a des A.I. sous Q . En effet,

$$\begin{aligned} E_Q \left(\prod_{i=1}^n x_i^{N_{t_{i+1}} - N_{t_i}} \right) &= E_P \left(\prod_{i=1}^n (2x_i)^{N_{t_{i+1}} - N_{t_i}} e^{-\lambda t} \right) \\ &= \left(\prod_{i=1}^n e^{-\lambda(t_{i+1} - t_i)} e^{\lambda(t_{i+1} - t_i) 2x_i} e^{-\lambda t} \right) \end{aligned}$$

Il en résulte que N est un Q -processus de Poisson process d'intensité 2λ .

Bibliographie

- [1] J. BERTOIN. *Lévy processes*. Cambridge University Press, Cambridge, 1996.
- [2] T. BJÖRK. *Arbitrage theory in continuous time*. Oxford University Press, Oxford, 1998.
- [3] R.A. DANA ET M. JEANBLANC. *Marchés financiers en temps continu* Economica, Paris, 1998. Traduction Springer 2002.
- [4] C. DELLACHERIE ET P.-A. MEYER. *Probabilités et Potentiel*. Volumes 1 à 5. Hermann, Paris, 1975-1992.
- [5] R. J. ELLIOTT. *Stochastic Calculus and Applications*. Springer, Berlin, 1982.
- [6] H. FÖLLMER ET A. SCHIED *Stochastic Finance : an introduction in discrete time*. Walter de Gruyter, Berlin, 2003.
- [7] G. GRIMMETT ET D. STIRZAKER. *Probability and Random Processes*. Oxford University Press, New-York, 1982.
- [8] G. GRIMMETT ET D. STIRZAKER. *One Thousand Exercises in Probability*. Oxford University Press, New-York, 2001.
- [9] J. JACOD ET P. PROTTER. *Probability Essentials*. Springer, Berlin, 1997.
- [10] J. JACOD ET A. N. SHIRYAEV. *Limit Theorems for Stochastic Processes*. Springer, Berlin, 1987.
- [11] I. KARATZAS ET S. E. SHREVE. *Brownian Motion and Stochastic Calculus*. Springer, Berlin, 1991.
- [12] D. LAMBERTON ET B. LAPEYRE. *Introduction au calcul stochastique appliqué à la finance*. Ellipses, Paris, 1991.
- [13] P. MALLIAVIN. *Stochastic Analysis*. Springer, Berlin, 1997.
- [14] J. NEVEU. *Bases mathématiques du calcul des probabilités*. Masson, Paris, 1964.
- [15] B. OKSENDAL. *Stochastic Differential Equations*. Springer, Berlin, 6th edition, 2002.
- [16] P. PROTTER. *Stochastic Integration and Differential Equations*. Springer, Berlin, 2nd edition, version 2.1. 2005.
- [17] D. REVUZ ET M. YOR. *Continuous Martingales and Brownian Motion*. Springer, Berlin, 3th edition 1999.
- [18] L. C. G. ROGERS ET D. WILLIAMS. *Diffusions, Markov Processes, and Martingales. Volume 1 : Foundations*. Wiley, Chichester, 1994.
- [19] L. C. G. ROGERS ET D. WILLIAMS. *Diffusions, Markov Processes, and Martingales. Volume 2 : Itô Calculus*. Wiley, Chichester, 1987.
- [20] G. SAMORODNITSKY ET M. S. TAQQU. *Stable Non-Gaussian Random Processes*. Chapman & Hall, New York, 1994.
- [21] K.-I. SATO. *Lévy Processes and Their Infinitely Divisible Distributions*. Cambridge University Press, Cambridge, 1999.
- [22] D. STIRZAKER. *Elementary Probability*. Cambridge University Press, Cambridge, 1994.
- [23] A.N. SHIRYAEV . *Essentials of stochastic Finance*. World Scientific, Singapore, 1999.
- [24] D. WILLIAMS. *Probability with Martingales*. Cambridge University Press, Cambridge, 1991.

Index

- Application mesurable, 6
- Bon processus, 43
- Bon processus local, 46
- Brownien géométrique, 58
- Changement de numéraire, 75
- Changement de temps, 64
- Condition
 - de Lipschitz, 53
 - de Novikov, 59
- Crochet, 47
 - de processus d'Itô, 48
- Décomposition de Doob-Meyer, 20
- Drift, 48
- Ensemble négligeable, 7
 - au sens large, 7
- Espérance
 - conditionnelle, 11
- Espace de probabilité complet, 7
- Espace Gaussien, 18
- Espace mesurable, 5
- Exponentielle de Doléans-Dade, 59
- Filtration, 17
- Fonction
 - caractéristique, 9
 - d'échelle, 58
 - de Bessel, 64
 - de répartition, 8
 - en escalier, 36
- Formule
 - d'intégration par parties, 38, 51
- Générateur infinitésimal, 58, 59
- Indice, 63
- Intégrable
 - v.a. -, 8
- Intégrale
 - de Wiener, 35
- Lemme
 - de Gronwall, 56
- Martingale, 20
 - locale, 46
- Martingales
 - orthogonales, 47
- Modification, 17
- Mouvement Brownien, 23
 - géométrique, 50
- Mouvements Browniens corrélés, 33
- Premier temps de passage, 19
 - bilatère d'un MB, 30
 - d'un MB drifté, 73
 - d'un MB, 30
- Presque sûrement, 7
- Probabilité, 6
- Probabilités
 - équivalentes, 12
 - absolument continues, 12
- Processus
 - à accroissements indépendants et stationnaires, 23, 38
 - à variation finie, 18
 - adapté, 17
 - càdlàg, 17
 - croissant, 18
 - d'Ornstein-Uhlenbeck, 39, 72
 - de Vařicek, 42
 - de Bessel, 56, 63
 - de Bessel carré, 63
 - de CIR, 65
 - de Cox-Ingersoll-Ross, 64
 - de Lévy, 38
 - de Markov, 40
 - gaussien, 18, 23, 37
 - stationnaire, 18
 - stochastique, 17
- Propriété de Markov, 28, 57
- Radon-Nikodym, 12
- Représentation prévisible, 71
- Scaling, 27
- Solution
 - faible, 54
 - forte, 54
- Sousmartingale, 20
- Surmartingale, 20
- Temps d'arrêt, 19
- Théorème
 - de Bochner, 18
 - de classe monotone, 7
 - de convergence dominée, 14
 - de convergence monotone, 14
 - de représentation des martingales locales, 48
 - de Yamada-Watanabe, 55
 - de Yamada-Watanabe II, 56
- Transformée
 - de Fourier, 9
 - de Laplace, 9
- Tribu, 5
 - des Boréliens, 6
 - engendrée, 6
- Unicité
 - faible, 54
 - trajectorielle, 54
- Uniforme intégrabilité, 10
- Variable gaussienne centrée réduite, 8
- Version, 18