

Processus continu approfondis

Année 2023 - 2024

DRAFT

François Simenhaus
Bureau B 640
simenhaus@ceremade.dauphine.fr

Table des matières

1 Révisions	5
1.1 Lemme de classe monotone	5
1.2 Autres rappels de théorie de la mesure	8
1.3 Variables et vecteurs gaussiens	9
2 Généralités sur les processus	13
2.1 Définitions. Tribu et loi.	13
2.1.1 Un processus comme fonction aléatoire	13
2.1.2 Loi d'un processus	16
2.1.3 Indépendance de deux processus	16
2.1.4 Processus canonique	16
2.1.5 Le théorème de Daniell - Kolmogorov	17
2.2 Processus continus	20
2.2.1 Tribu(s)	20
2.2.2 Modification continue et critère de Kolmogorov	22
3 Mouvement brownien	24
3.1 Définitions	24
3.2 Existence	25
3.3 Mesure de Wiener et processus canonique	26
3.4 Premières propriétés	26
3.5 Propriété de Markov fort	30
3.6 Théorème de Donsker	32
4 Martingales continues	34
4.1 Définition	34
4.2 Convergence et régularisation	35
4.3 Théorème d'arrêt de Doob	36
4.4 Inégalités maximales de Doob	37
5 Intégrale stochastique - Itô (1950)	38
5.1 Intégrale de Stieljies	38
5.2 Intégrale d'Itô	41
5.3 Formule d'Itô	54
6 Équations différentielles stochastiques	59
6.1 Existence et unicité d'une solution forte	59
6.2 Équation de Langevin et processus d'Ornstein-Uhlenbeck	63

Un amphi d'1h30 par semaine et une séance de TD de la même durée. La note est déterminée par

$$\text{Note} = \max(\text{Examen} ; 0,4 \text{ Partiel} + 0,6 \text{ Examen}).$$

Ces notes de cours sont largement inspirées des notes et livres de Fabrice Baudoin [1], Francis Comets [5], Jean-François Legall [7], Marc Yor et Daniel Revuz [8]. Je vous invite aussi à consulter les photocopiés de Djalil Chafaï [4], Philippe Bougerol [3] et Nadine Guillotin [6], que vous trouverez facilement en ligne.

Dans tout ce cours on considère un espace de probabilités (Ω, \mathcal{F}, P) . Tous les objets aléatoires considérés sont construits sur cet espace sauf quand on le mentionnera explicitement. On notera E l'espérance associée. Si X est une variable aléatoire et $A \in \mathcal{F}$ un événement, on utilisera parfois la notation (classique)

$$E(X, A) := E(X1_A).$$

Définition 1. *Un processus indexé par un ensemble d'indices T et à valeurs dans un espace mesurable (E, \mathcal{E}) est une famille $(X_t)_{t \in T}$ de variables aléatoires définies sur (Ω, \mathcal{F}, P) et à valeurs dans (E, \mathcal{E}) .*

Des exemples d'ensemble d'indices courants :

1. $T = \mathbb{N}$ et on parle alors de *processus discrets*. C'est le cadre du cours que vous avez suivi au premier semestre. Les exemples les plus importants de tels processus sont bien sûr les martingales en temps discret et les chaînes de Markov.
2. $T = \mathbb{R}^2$ et on parle alors de *champs aléatoires*.¹

Dans ce cours, on s'intéressera principalement aux **processus en temps continu** et à valeurs dans \mathbb{R} :

$$T = \mathbb{R}^+ \text{ et } (E, \mathcal{E}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

1. Voir par exemple la page wikipedia consacrée au champs libre gaussien pour un exemple célèbre

(parfois $\mathbb{T} = \mathbb{R}$ et $(E, \mathcal{E}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ avec $d \geq 1$ entier).

On s'intéressera également principalement aux **processus dont les trajectoires sont continues**. On va notamment définir, construire et étudier le **mouvement brownien** $(B_t)_{t \geq 0}$. Un de nos objectifs est de donner un sens à l'**équation différentielle stochastique** suivante :

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

Cette équation peut être comprise de la façon suivante : le déplacement infinitésimal de la particule X au temps t est la somme de deux termes,

1. un terme correspondant au champs de vitesse $b : b(X_t)dt$. Si on n'avait que ce terme notre equation serait une EDO $X'_t = b(X_t)$ similaire à celles que vous avez étudiées l'année dernière en L3 ;
2. un terme rendant compte des chocs microscopiques dûs à l'agitation thermique $\sigma : \sigma(X_t)dB_t$.

Pour donner du sens à cette équation nous aurons besoin d'une part de construire le mouvement brownien mais également de construire l'**intégrale stochastique** par rapport à ce mouvement brownien.

1 Révisions

Dans ce chapitre nous proposons quelques points de révisions qui nous seront très utiles dans la suite du cours. Deux sections principales : une première permet de faire le point sur l'« argument de classe monotone » qui est souvent invoqué lors de l'étude des processus ; une seconde comprend quelques révisions sur les variables et vecteurs gaussiens qui nous seront très utiles dans la mesure où le mouvement brownien est un processus gaussien ; quelques résultats de théorie de la mesure que nous utiliserons également pendant ce semestre complètent cette partie.

1.1 Lemme de classe monotone

Ce lemme est particulièrement utile dans le cadre de la théorie des processus. De manière générale, dans quel contexte l'utilise-t-on ?

Pour montrer qu'une propriété \mathcal{P} est vraie pour tout événement A d'une tribu \mathcal{F} on procède souvent de la façon suivante :

1. on trouve une classe (c'est-à-dire une collection de parties de Ω) I engendrant \mathcal{F} (i.e. $\sigma(I) = \mathcal{F}$) telle que $\mathcal{P}(A)$ soit vraie pour tout $A \in I$;
2. on montre que la classe $\{A \text{ tel que } \mathcal{P}(A) \text{ est vrai}\}$ forme une tribu ;

et on conclut que \mathcal{P} est vraie (au moins) sur $\sigma(I) = \mathcal{F}$. Le problème est que dans certains cas le second point est faux car la classe que l'on considère ne vérifie pas la stabilité par union dénombrable mais seulement la stabilité par union dénombrable *croissante*. C'est le cas par exemple, si on considère μ et ν deux probabilités sur (Ω, \mathcal{F}) et que l'on s'intéresse à la classe $\{A \text{ tel que } \mu(A) = \nu(A)\}$. Et c'est dans ces cas là que le lemme de classe monotone est utile.

Définition 2. Une classe $\mathcal{M} \subset \mathcal{P}(\Omega)$ est une **classe monotone** si

1. $\Omega \in \mathcal{M}$
2. Pour tout $A, B \in \mathcal{M}$ et $A \subset B$, $B \setminus A \in \mathcal{M}$
3. Pour toute suite **croissante** $(A_n)_{n \geq 0}$ à valeurs dans \mathcal{M} ,

$$\bigcup_{n \geq 0} \uparrow A_n \in \mathcal{M}.$$

Le terme λ -système est parfois utilisé dans la littérature plutôt que celui de classe monotone. On notera aussi qu'il existe d'autres définitions

équivalentes d'une classe monotone dans la littérature². Pour ces deux points on se contentera dans ce cours du terme "classe monotone" et de la définition donnée ci-dessus.

On définit $\lambda(\mathcal{C})$ la classe monotone engendrée par une classe I en utilisant la même idée que pour la tribu engendrée par I qui est elle notée $\sigma(I)$. Rappelons que :

$$\sigma(I) = \bigcap_{\mathcal{T} \text{ tribu t.q. } I \subset \mathcal{T}} \mathcal{T}.$$

On définit donc

$$\lambda(I) = \bigcap_{\mathcal{M} \text{ classe monotone t.q. } I \subset \mathcal{M}} \mathcal{M}.$$

Pour que cette définition ait du sens il faut bien sûr vérifier qu'une intersection quelconque de classe monotone est encore une classe monotone et que l'intersection est non vide puisque $\mathcal{P}(\Omega)$ est bien une classe monotone et contient I . On note qu'une tribu est aussi une classe monotone et on a donc toujours

$$\lambda(I) \subset \sigma(I). \quad (1)$$

Revenons à notre problème initial : on se trouve dans un cas où la seconde partie de la démonstration ne fonctionne pas car la classe

$$\{A \text{ tel que } \mathcal{P}(A) \text{ est vrai}\}$$

n'est pas une tribu mais seulement une classe monotone. On obtient donc que la propriété A est vraie sur $\lambda(I)$. Le but du lemme de classe monotone est de montrer, au prix d'une hypothèse additionnelle, la réciproque de (1) et donc $\lambda(I) = \sigma(I)$.

Lemme 1 (Lemme de classe monotone). *Soit $I \subset \mathcal{P}(\Omega)$ une classe **stable par intersection finie**. Alors*

$$\sigma(I) = \lambda(I).$$

Preuve. L'idée de la preuve est de montrer qu'avec l'hypothèse additionnelle de stabilité par intersection finie $\lambda(I)$ est en fait une tribu. C'est l'objet de l'Exercice 1 du TD 1. □

Voici quelques applications importantes :

2. Voir par exemple la page wikipedia anglaise du lemme de classe monotone

1. Soient μ et ν deux probabilités qui coïncident sur une classe $I \subset \mathcal{P}(\Omega)$ stable par intersection finie et telle que $\sigma(I) = \mathcal{F}$ (cela signifie donc $\mu(A) = \nu(A)$ pour tout $A \in I$). Alors $\mu = \nu$, c'est-à-dire

$$\mu(A) = \nu(A), \quad \text{pour tout } A \in \mathcal{F}.$$

En effet, la classe $\mathcal{M} = \{A \text{ tel que } \mu(A) = \nu(A)\}$ est une classe monotone (le vérifier!) et contient I . Donc $\mathcal{M} \supset \lambda(I)$. De plus, I est stable par intersection finie donc, d'après le lemme de classe monotone, $\lambda(I) = \sigma(I)$. Donc $\mu = \nu$.

On notera bien que μ et ν ne coïncident pas forcément si on enlève l'hypothèse "*I est stable par intersection finie*" comme le prouve l'exemple suivant : on considère $\Omega = \{1, 2, 3, 4, 5\}$, la tribu $\mathcal{F} = \mathcal{P}(\Omega)$ et la classe $I = \{\{1, 2, 3\}, \{2, 4\}, \{3, 4, 5\}\}$ dont on notera qu'elle n'est pas stable par intersection finie. On vérifie aisément que $\sigma(I) = \mathcal{F}$. Notons μ la probabilité uniforme et ν la probabilité telle que $\nu(\{1\}) = 3/10$, $\nu(\{2\}) = 1/10$, $\nu(\{3\}) = 1/5$, $\nu(\{4\}) = 3/10$, $\nu(\{5\}) = 1/10$. Il est facile de vérifier que μ et ν coïncident sur I mais pas sur $\mathcal{P}(\Omega)$.

Voici deux exemples importants où nous utilisons cette propriété :

- (a) On se place sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Si deux mesures μ et ν coïncident sur les intervalles ouverts bornés (i.e. $\mu(]a, b[) = \nu(]a, b[)$ pour tout $a < b$) alors elles sont égales. C'est ainsi que l'on prouve, finalement sans grands efforts, l'unicité de la mesure de Lebesgue. L'existence en revanche est une toute autre histoire!
- (b) La fonction de répartition caractérise la loi car la classe $\{] - \infty, a], a \in \mathbb{R}\}$ est stable par intersection finie et engendre également la tribu borélienne.
2. On considère une famille $(X_t)_{t \in T}$ de variables aléatoires. On rappelle que par définition $\sigma(X_t, t \in T)$ est la plus petite tribu rendant tous les $X_t, t \in T$, mesurables [vérifier que cette définition a bien du sens et que, quand $|T| < +\infty$, $\sigma(X_t, t \in T) = \{(X_t)_{t \in T} \in A, A \in \mathcal{B}(\mathbb{R})^{|T|}\}$. Voir l'Exercice 5 du TD 1 sur ce point] On considère également \mathcal{G} une sous tribu de \mathcal{F} . Alors les deux propositions suivantes sont équivalentes :

- (a) \mathcal{G} et $\sigma(X_t, t \in T)$ sont indépendantes
- (b) Pour tout $S \subset T$ fini, \mathcal{G} et $\sigma(X_t, t \in S)$ sont indépendantes.

Preuve. On fixe $A \in \mathcal{G}$ et on introduit

$$\mathcal{M} = \{B \in \mathcal{F}, P(A \cap B) = P(A)P(B)\}.$$

On veut montrer que $\mathcal{M} = \sigma(X_t, t \in T)$. On montre pour cela que \mathcal{M} est une classe monotone. D'après l'hypothèse, cette classe contient

$$\mathcal{C} = \bigcup_{|J| < +\infty} \sigma(X_j, j \in J).$$

Cette classe est stable par intersection finie car si $A \in \sigma(X_j, j \in J_1)$ et $B \in \sigma(X_j, j \in J_2)$ où J_1 et J_2 sont deux sous-ensembles finis de J alors $A \cap B \in \sigma(X_j, j \in J_1 \cup J_2)$.

donc $\lambda(\mathcal{C}) = \sigma(\mathcal{C})$ et on en déduit que \mathcal{M} contient $\sigma(\mathcal{C})$. Comme $\sigma(\mathcal{C})$ rend mesurable tous les $X_t, t \in T$ on en déduit que \mathcal{M} contient $\sigma(X_t, t \in T)$. \square

En corollaire de cette propriété, on peut également montrer (Exercice 3 TD 1) que deux familles de variables aléatoires $(X_t)_{t \in T}$ et $(Y_s)_{s \in S}$ sont indépendantes si et seulement si pour toutes familles finies $t_1, \dots, t_n \in T$ et $s_1, \dots, s_m \in S$, les vecteurs $(X_{t_1}, \dots, X_{t_n})$ et $(Y_{s_1}, \dots, Y_{s_m})$ sont indépendants.

Nous verrons dans ce cours de nombreuses autres applications de ce lemme de classe monotone. Je vous conseille donc de bien le maîtriser.

1.2 Autres rappels de théorie de la mesure

Voici quelques résultats qu'il faut savoir montrer et utiliser.

Lemme 2. Soit X une fonction à valeurs dans un ensemble E et $\mathcal{C} \subset \mathcal{P}(E)$. Alors X est une variable aléatoire de (Ω, \mathcal{F}) dans $(E, \sigma(\mathcal{C}))$ si et seulement si pour tout $B \in \mathcal{C}$, $\{X \in B\} \in \mathcal{F}$.

Preuve. Exercice! \square

Pour toute variable aléatoire X on note ϕ_X la fonction caractéristique de X définie pour tout $\xi \in \mathbb{R}$ par

$$\phi_X(\xi) = \mathbb{E}(e^{i\xi X}).$$

Nous rappelons maintenant le théorème fondamental dû à Paul Lévy :

Théorème 1 (Théorème de Lévy). 1. Une suite $(X_n)_{n \geq 1}$ de variables aléatoires converge en loi vers la variable aléatoire X si et seulement si pour tout $\xi \in \mathbb{R}$, $(\phi_{X_n}(\xi))_{n \geq 1}$ converge vers $\phi_X(\xi)$.

2. On suppose qu'il existe une fonction ϕ telle que pour tout $\xi \in \mathbb{R}$, $(\phi_{X_n}(\xi))_{n \geq 1}$ converge vers $\phi(\xi)$. Alors les trois points suivants sont équivalents :

- (a) $(X_n)_{n \geq 1}$ converge en loi;
- (b) ϕ est une fonction caractéristique;
- (c) ϕ est continue;
- (d) ϕ est continue en 0.

C'est un théorème difficile (ce qui ne veut pas dire qu'il faut renoncer à en étudier la preuve).

1.3 Variables et vecteurs gaussiens

Il faut bien réviser toutes les notions relatives aux variables et vecteurs gaussiens parce qu'elles nous seront très utiles lors de l'étude du mouvement brownien.

Une **variable aléatoire gaussienne** de paramètre $m \in \mathbb{R}$ et $\sigma^2 > 0$, et de loi notée $\mathcal{N}(m, \sigma^2)$ (dite aussi loi normale) est soit (dans le cas où $\sigma^2 = 0$) une variable constante égale à m , soit (dans le cas où $\sigma^2 > 0$) une variable réelle admettant pour densité

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad x \in \mathbb{R},$$

par rapport à la mesure de Lebesgue. On dit qu'elle est **centrée** si $m = 0$ et **réduite** si $\sigma^2 = 1$ (elle suit donc une $\mathcal{N}(0, 1)$) et on a alors

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}.$$

On peut vérifier que si X suit une $\mathcal{N}(m, \sigma^2)$ alors $E(X) = m$ et $\text{Var}(X) = \sigma^2$. Si U suit une $\mathcal{N}(0, 1)$ alors pour tout $m, \sigma \in \mathbb{R}$ la variable $m + \sigma U$ suit une $\mathcal{N}(m, \sigma^2)$.

La **fonction caractéristique** d'une variable U centrée réduite est donnée par

$$\phi_U(\xi) = e^{-\xi^2/2}, \quad \xi \in \mathbb{R},$$

(vérifier que vous savez le prouver : il y a plusieurs méthodes mais une des plus simples est de montrer que ϕ_U satisfait une EDO facile à résoudre). Plus généralement pour tout $z \in \mathbb{C}$,

$$E(e^{zU}) = e^{z^2/2}.$$

On en déduit facilement de l'expression de ϕ_X lorsque X suit une $\mathcal{N}(m, \sigma^2)$:

$$\phi_X(\xi) = e^{i\xi m} e^{-\frac{\xi^2 \sigma^2}{2}}, \quad \xi \in \mathbb{R}.$$

En conséquence si X_1 et X_2 sont deux gaussiennes indépendantes et de lois respectives $\mathcal{N}(m_1, \sigma_1^2)$ et $\mathcal{N}(m_2, \sigma_2^2)$, d'après la caractérisation de Lévy, $X_1 + X_2$ suit une $\mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$.

Les **moments de la gaussienne centrée réduite** s'obtiennent facilement à partir du développement de la fonction caractéristique ϕ_U . Pour tout $p \geq 1$, $E(U^{2p}) = \frac{(2p)!}{2^p p!}$ et bien sûr $E(U^{2p-1}) = 0$.

La proposition suivante est un classique à connaître et à savoir prouver (voir Exercice 6 TD 1).

Proposition 1 (Une limite (même en loi !) d'une suite de gaussiennes est une gaussienne). Soit $(X_n)_{n \geq 0}$ une suite de variables gaussiennes telles que pour tout $n \geq 0$, X_n suit une $\mathcal{N}(m_n, \sigma_n^2)$. On suppose également que $(X_n)_{n \geq 0}$ converge **en loi** vers une variable aléatoire X . Alors :

1. X suit une $\mathcal{N}(m, \sigma^2)$ où $m = \lim_{n \rightarrow +\infty} m_n$ et $\sigma^2 = \lim_{n \rightarrow +\infty} \sigma_n^2$.
2. Si de plus $(X_n)_{n \geq 0}$ converge **en probabilité** vers X alors $(X_n)_{n \geq 0}$ converge aussi dans L^p pour tout $1 \leq p < +\infty$.

Preuve. C'est l'exercice 5 de la feuille 1. C'est un classique mais ça n'est pas un exercice facile pour autant. \square

On passe maintenant aux **vecteurs gaussiens**.

Définition 3. Un vecteur aléatoire (X_1, \dots, X_n) est un vecteur gaussien si pour tout $(t_1, \dots, t_n) \in \mathbb{R}^n$ la variable

$$t \cdot X = \sum_{i=1}^n t_i X_i$$

est gaussienne.

Si le vecteur $(X_i)_{i=1, \dots, n}$ est gaussien cela implique bien sûr que pour tout $i = 1, \dots, n$, la variable X_i est gaussienne mais la réciproque est fautive.

À partir du cas unidimensionnel on peut facilement retrouver la **fonction caractéristique d'un vecteur gaussien** : pour tout vecteur ξ dans \mathbb{R}^n ,

$$\phi_X(\xi) = E(e^{i^t \xi X}) = \phi_{t\xi X}(1).$$

Or $t\xi X$ suit une $\mathcal{N}(t\xi m, t\xi \Gamma \xi)$ avec

$$m = E(X) \quad \text{et} \quad \Gamma = E((X - m)^t (X - m)),$$

la moyenne et matrice de variance de X . On obtient donc finalement

$$\phi_X(\xi) = e^{i^t \xi m - \frac{t \xi \Gamma \xi}{2}},$$

et on note que la loi d'un vecteur gaussien est caractérisée par son vecteur moyenne et sa matrice de variance.

Une conséquence importante est la **caractérisation de l'indépendance** entre les coordonnées d'un vecteur gaussien :

Proposition 2. *Si X est un vecteur gaussien alors les composantes sont indépendantes si et seulement si Γ est diagonale, c'est-à-dire, si les covariances sont nulles.*

Plus généralement, si $(X_i)_{i \in I}$ est un vecteur gaussien et $I = \cup_{j=1, \dots, m}^{(disj.)} I_j$, alors les sous-vecteurs $(X_i)_{i \in I_j}$, $j = 1, \dots, m$ sont indépendants si et seulement si pour tout $1 \leq k < \ell \leq m$ et tout $a, b \in I_k \times I_\ell$, $Cov(X_a, X_b) = 0$ (c'est-à-dire que Γ est diagonale par blocs avec les blocs correspondant à la partition de I en I_j , $j = 1, \dots, m$).

Preuve. Le sens direct est toujours vrai : il n'y a rien de spécifique dans le cas gaussien. Pour la réciproque il faut montrer que la fonction caractéristique se factorise. \square

Comme dans le cas unidimensionnel on peut écrire tout vecteur gaussien centré comme **transformée linéaire d'une normale centrée réduite** (c'est-à-dire $m = 0$ et $\Gamma = I_n$). En effet pour toute matrice Γ symétrique positive il existe P orthogonale (i.e. ${}^t P P = P {}^t P = I_n$) et $\lambda_1, \dots, \lambda_d > 0$ ($d \leq n$) telles que

$${}^t P \Gamma P = \begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_d & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}.$$

On en déduit que $\Gamma = A {}^t A$ avec

$$A = P \begin{pmatrix} \sqrt{\lambda_1} & & & & & \\ & \ddots & & & & \\ & & \sqrt{\lambda_d} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}.$$

Si U est un vecteur gaussien centré réduit (i.e. U suit une $\mathcal{N}(0, I_n)$). On vérifie que

$$X \stackrel{(loi)}{=} AU.$$

C'est aussi une façon simple de voir que pour tout Γ symétrique positive il existe une $\mathcal{N}(0, \Gamma)$ puisqu'on peut la construire explicitement à partir de U centrée réduite (et qui est seulement un vecteur de gaussiennes centrées réduites indépendantes donc facile à construire).

Densité. Pour tout vecteur aléatoire X à valeurs dans \mathbb{R}^n les deux points suivants sont équivalents :

1. X suit une $\mathcal{N}(m, \Gamma)$ et $\text{rg}(\Gamma) = n$
2. X admet pour densité par rapport à la mesure de Lebesgue sur \mathbb{R}^n :

$$f_X(x) = \frac{1}{\sqrt{2\pi}^n \sqrt{\det\Gamma}} e^{-\frac{1}{2}(x-m)\Gamma^{-1}(x-m)}.$$

On le prouve par exemple en écrivant « habilement » le vecteur X comme somme de vecteurs gaussiens indépendants.

2 Généralités sur les processus

2.1 Définitions. Tribu et loi.

Définition 4. *Un processus indexé par un ensemble d'indices \mathbb{T} et à valeurs dans un espace mesurable (E, \mathcal{E}) est une famille $(X_t)_{t \in \mathbb{T}}$ de variables aléatoires définies sur $(\Omega, \mathcal{F}, \mathbb{P})$ et à valeurs dans (E, \mathcal{E}) .*

Des exemples d'ensemble d'indices courants :

1. $\mathbb{T} = \mathbb{N}$ et on parle alors de *processus discrets*. C'est le cadre du cours que vous avez suivi au premier semestre. Les exemples les plus importants de tels processus sont bien sûr les martingales en temps discret et les chaînes de Markov.
2. $\mathbb{T} = \mathbb{R}^2$ et on parle alors de *champs aléatoires*.

Dans ce cours, on s'intéressera au **processus en temps continu** et à valeurs dans \mathbb{R} :

$$\mathbb{T} = \mathbb{R}^+ \text{ et } (E, \mathcal{E}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

(parfois $\mathbb{T} = \mathbb{R}$ ou $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ avec $d \geq 1$ entier).

2.1.1 Un processus comme fonction aléatoire

On remarque que à $\omega \in \Omega$ fixé, un processus $(X_t)_{t \in \mathbb{T}}$ (ici $\mathbb{T} = \mathbb{R}^+$) définit une fonction :

$$\begin{aligned} X(\omega) : \mathbb{R}^+ &\rightarrow \mathbb{R} \\ t &\mapsto X_t(\omega). \end{aligned}$$

Une première façon de considérer le processus X est de le voir comme une **fonction aléatoire**. Est-ce possible? Quelle tribu \mathcal{T} considérer sur l'ensemble $\mathcal{A}(\mathbb{R}^+, \mathbb{R})$ des fonctions de \mathbb{R}^+ dans \mathbb{R} pour que

$$X : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T})$$

soit mesurable?

Définition 5 (Tribu cylindrique). *On appelle cylindre un sous ensemble de $\mathcal{A}(\mathbb{R}^+, \mathbb{R})$ de la forme*

$$\{f \in \mathcal{A}(\mathbb{R}^+, \mathbb{R}) \text{ tel que } f(t_1) \in B_1, \dots, f(t_n) \in B_n\}$$

où $n \geq 1$ est un entier, t_1, \dots, t_n sont des réels positifs et B_1, \dots, B_n sont dans $\mathcal{B}(\mathbb{R})$. On appelle tribu cylindrique, et on note \mathcal{T} , la tribu sur $\mathcal{A}(\mathbb{R}^+, \mathbb{R})$ engendrée par les cylindres :

$$\mathcal{T} = \sigma(C, \text{Cylindre de } \mathcal{A}(\mathbb{R}^+, \mathbb{R})),$$

c'est-à-dire la plus petite tribu qui contient tous les cylindres.

On définit en fait la même tribu en remplaçant dans la définition des cylindres les boréliens B_i , $i = 1, \dots, n$ par des intervalles I_i , $i = 1, \dots, n$ ouverts ou encore ouverts seulement d'un côté (voir Exercice 3 du TD 2). On notera surtout l'autre définition courante et utile de cette tribu :

Proposition 3. *La tribu \mathcal{T} est la plus petite tribu qui rendent toutes les applications coordonnées mesurables. Autrement dit*

$$\mathcal{T} = \sigma(\pi_t, t \geq 0)$$

où pour tout $t \geq 0$, π_t est l'application coordonnée en t défini par :

$$\begin{aligned} \pi_t : \mathcal{A}(\mathbb{R}^+, \mathbb{R}) &\rightarrow \mathbb{R} \\ f &\mapsto f(t). \end{aligned}$$

On rappelle que les applications coordonnées sont les fonctions π_t , $t \geq 0$, définies par :

$$\begin{aligned} \pi_t : \mathcal{A}(\mathbb{R}^+, \mathbb{R}) &\rightarrow \mathbb{R} \\ f &\mapsto f(t). \end{aligned}$$

Preuve. Pour tout $t \geq 0$ on note que π_t est $\mathcal{T} - \mathcal{B}(\mathbb{R})$ -mesurable puisque pour tout $B \in \mathcal{B}(\mathbb{R})$,

$$\{\pi_t \in B\} = \{f \in \mathcal{A}(\mathbb{R}^+, \mathbb{R}), f(t) \in B\}$$

est un cylindre donc dans \mathcal{T} .

Réciproquement, notons $\tilde{\mathcal{T}}$ une tribu qui rend toutes les applications coordonnées mesurables et montrons qu'elle contient tout les cylindres. On considère donc un cylindre

$$C = \{f \text{ t.q. } f(t_1) \in A_1, \dots, f(t_n) \in A_n\}.$$

On peut le réécrire

$$C = \bigcap_{i=1}^n \{\pi_{t_i} \in A_i\}$$

et on en déduit que $C \in \tilde{\mathcal{T}}$. On a donc bien montré que $\mathcal{T} \subset \tilde{\mathcal{T}}$. \square

Nous pouvons maintenant considérer un processus comme **fonction aléatoire** :

Proposition 4. *Si $(X_t)_{t \in \mathbb{R}^+}$ est un processus alors*

$$\begin{aligned} X : (\Omega, \mathcal{F}, \mathbb{P}) &\rightarrow (\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T}) \\ \omega &\rightarrow (t \mapsto X_t(\omega)) \end{aligned}$$

est mesurable de $(\Omega, \mathcal{F}, \mathbb{P})$ dans $(\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T})$ et définit donc une fonction aléatoire.

Preuve. Comme $\mathcal{T} = \sigma(C, C \text{ cylindre de } \mathcal{A}(\mathbb{R}^+, \mathbb{R}))$ il suffit de vérifier que pour tout cylindre C , $\{X \in C\} \in \mathcal{F}$. On se donne donc un cylindre

$$C = \bigcap_{i=1}^n \{\pi_{t_i} \in B_i\}$$

et on vérifie que

$$\{X \in C\} = \{\omega \text{ t.q. } (t \mapsto X_t(\omega)) \in C\} = \bigcap_{i=1}^n \{X_{t_i} \in B_i\}$$

appartient bien à \mathcal{F} . □

On vient en fait de montrer que

$$\sigma(X) = X^{-1}(\mathcal{T}) \subset \sigma(X_t, t \in \mathbb{T}).$$

L'inclusion réciproque est également vraie (c'est plus facile ! voir TD 2) et on a donc que pour tout processus $(X_t)_{t \in \mathbb{T}}$

$$\sigma(X_t, t \in \mathbb{T}) = X^{-1}(\mathcal{T}). \quad (2)$$

On rappelle (voir encore le TD 2) que pour toute famille finie (X_1, \dots, X_n) de variables aléatoires réelles, $\sigma(X_1, \dots, X_n)$, la plus petite tribu rendant tous les X_i ($i = 1, \dots, n$) mesurables coïncide avec $\sigma((X_1, \dots, X_n))$ la tribu engendrée par (X_1, \dots, X_n) vu comme vecteur aléatoire à valeurs dans $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. En munissant $\mathcal{A}(\mathbb{R}^+, \mathbb{R})$ de la tribu cylindrique \mathcal{T} on peut donc étendre cette propriété des familles finies de variables aléatoires aux processus.

Par ailleurs, \mathcal{T} permet de considérer X comme une fonction aléatoire mais notons aussi que (2) montre aussi que \mathcal{T} est suffisamment riche pour que son image réciproque englobe tout $\sigma(X_t, t \in \mathbb{T})$.

2.1.2 Loi d'un processus

Nous pouvons donc définir la **loi du processus** $(X_t)_{t \in \mathbb{T}}$ comme la loi image sur $(\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T})$ de P par la fonction aléatoire X : pour tout $A \in \mathcal{T}$,

$$P_X(A) = P((X_t)_{t \in \mathbb{T}} \in A).$$

Proposition 5. *Les mesures de probabilités sur $(\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T})$ sont caractérisées par leurs valeurs sur les cylindres. Autrement dit, si μ et ν sont deux probabilités sur $(\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T})$ telles que pour tout cylindre C de $\mathcal{A}(\mathbb{R}^+, \mathbb{R})$, $\mu(C) = \nu(C)$ alors $\mu = \nu$.*

Preuve. C'est encore une utilisation du lemme de classe monotone. En effet, puisque $\mathcal{T} = \sigma(C, C \text{ cylindre de } \mathcal{A}(\mathbb{R}^+, \mathbb{R}))$ il nous suffit de vérifier que l'ensemble des cylindres forme une classe stable par intersection finie (ce qui n'est pas très difficile). \square

Pour un processus $(X_t)_{t \in \mathbb{T}}$, on appelle **marginale fini dimensionnelle** toute famille finie de marginales (c'est-à-dire toute sous-famille finie de $(X_t)_{t \in \mathbb{T}}$) :

$$(X_{t_i})_{i=1, \dots, n} \quad n \geq 1, t_i \in \mathbb{T} \text{ pour tout } i = 1, \dots, n.$$

et **lois fini dimensionnelles** les lois des marginales fini dimensionnelles. On déduit donc de la Proposition 5 que deux processus $(X_t)_{t \in \mathbb{T}}$ et $(Y_t)_{t \in \mathbb{T}}$ qui ont même lois fini dimensionnelles ont même loi :

si pour tout $n \geq 1$, tous $t_1, \dots, t_n \in \mathbb{T}$ et tout $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R})$,

$$P(X_{t_1} \in B_1, \dots, X_{t_n} \in B_n) = P(Y_{t_1} \in B_1, \dots, Y_{t_n} \in B_n),$$

alors $P_X = P_Y$ (i.e. $X \stackrel{(loi)}{=} Y$).

2.1.3 Indépendance de deux processus

D'après le lemme de classe monotone (encore! voir Exercice 3 TD 1), **deux processus $(X_t)_{t \in \mathbb{T}}$ et $(Y_t)_{t \in \mathbb{T}}$ sont indépendants** si et seulement si pour tout $n, m \geq 1$, tous $t_1, \dots, t_n \in \mathbb{T}$, et tout $s_1, \dots, s_m \in \mathbb{T}$

$$(X_{t_1}, \dots, X_{t_n}) \perp\!\!\!\perp (Y_{s_1}, \dots, Y_{s_m}).$$

2.1.4 Processus canonique

On s'intéresse maintenant à la question de **l'existence de loi** sur $\mathcal{A}(\mathbb{R}^+, \mathbb{R})$. Première question : je considère une mesure de probabilité μ sur $(\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T})$,

existe-t-il un processus qui ait pour loi μ ? On pourra, pour s'échauffer, essayer de répondre à la même question pour $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (Exercice 4 TD 2). Pour répondre à cette question on introduit le **processus canonique** : il s'agit des applications coordonnées $(\pi_t)_{t \in \mathbb{R}^+}$ vues comme variables aléatoires sur l'espace de probabilité $(\Omega, \mathcal{F}, P) = (\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T}, \mu)$: pour tout $t \geq 0$,

$$\begin{aligned} \pi_t : (\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T}, \mu) &\rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R})) \\ \omega &\rightarrow \omega(t) \end{aligned}$$

On note que par définition même de \mathcal{T} (voir Proposition 3), π_t est mesurable pour tout $t \geq 0$. On notera bien que, vu comme fonction aléatoire, le processus π , est en fait l'identité :

$$\begin{aligned} \pi : (\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T}) &\rightarrow (\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T}) \\ \omega &\rightarrow (t \rightarrow \pi_t(\omega)) \end{aligned}$$

On vérifie donc bien que pour tout $A \in \mathcal{T}$, $\mu_\pi(A) = \mu(\pi \in A) = \mu(A)$ ce qui signifie que π a pour loi μ ! Donc pour toute probabilité sur $(\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T})$ je sais construire un processus qui a cette probabilité pour loi en considérant le processus canonique.

2.1.5 Le théorème de Daniell - Kolmogorov

On a déjà vu que les lois fini dimensionnelles caractérisaient la loi d'un processus. Est ce que, pour toute famille de lois fini dimensionnelles, il existe un processus dont les marginales correspondent à ces lois? Il est clair qu'on ne peut pas espérer que cela soit vrai en toute généralité : si on considère les projections fini dimensionnelles d'une loi sur $\mathcal{A}(\mathbb{R}^+, \mathbb{R})$ elles ont clairement la propriété de **compatibilité** suivante (qui est donc une condition nécessaire) :

on considère une probabilité μ sur $(\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T})$. Pour rendre les choses un peu plus concrètes on peut imaginer que μ est la loi d'un processus $(X_t)_{t \geq 0}$ (donc $\mu = P_X$ mais on utilisera aussi parfois pour plus de lisibilité $\mu = X(P)$). Pour tout $I \subset \mathbb{R}^+$ fini, on note μ_I la loi image de μ par l'application de restriction à I

$$\begin{aligned} \pi_I : (\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T}) &\rightarrow (\mathbb{R}^I, \mathcal{B}(\mathbb{R}^I)) \\ \omega &\rightarrow (\omega_t)_{t \in I} \end{aligned}$$

On a donc $\mu_I = \mu_{\pi_I}$. De façon équivalente on aurait pu définir μ_I comme la loi de $(X_t)_{t \in I}$ puisque pour tout $A \in \mathcal{B}(\mathbb{R}^I)$,

$$\mu_I(A) = \mu(\pi_I \in A) = P((X_t)_{t \geq 0} \in \{\pi_I \in A\}) = P(\pi_I \circ (X_t)_{t \geq 0} \in A) = P((X_t)_{t \in I} \in A).$$

La loi μ_I est donc la projection de la loi μ sur \mathbb{R}^I et l'ensemble des lois μ_I , $I \subset \mathbb{R}^+$ finie, est l'ensemble des lois finies dimensionnelles de $(X_t)_{t \geq 0}$. On définit également pour toutes parties $J \subset I \subset \mathbb{R}^+$ finies

$$\begin{aligned} \pi_J^I : (\mathbb{R}^I, \mathcal{B}(\mathbb{R}^I)) &\rightarrow (\mathbb{R}^J, \mathcal{B}(\mathbb{R}^J)) \\ (\omega_t)_{t \in I} &\rightarrow (\omega_t)_{t \in J}. \end{aligned}$$

On a $\pi_J^I \circ \pi_I = \pi_J$ et donc pour tout $A \in \mathcal{B}(\mathbb{R}^J)$,

$$\pi_J^I(\mu_I)(A) = \mu_I(\pi_J^I \in A) = \mu(\pi_I \in \{\pi_J^I \in A\}) = \mu(\pi_J \in A) = \mu_J(A).$$

On vient de vérifier que $\pi_J^I(\mu_I)$ est la loi de $(X_t)_{t \in J}$. Ça semble compliqué à écrire mais on n'a rien dit de très spectaculaire : je peux obtenir la loi de $(X_t)_{t \in J}$ de deux manières : soit en extrayant directement le vecteur $(X_t)_{t \in J}$ du processus $(X_t)_{t \geq 0}$ soit en extrayant d'abord le vecteur $(X_t)_{t \in I}$ puis de ce vecteur le sous-vecteur $(X_t)_{t \in J}$. La famille des lois finies dimensionnelles μ_I vérifie donc la **condition de compatibilité** : pour toutes parties $J \subset I \subset \mathbb{R}^+$ finies

$$\pi_J^I(\mu_I) = \mu_J. \quad (3)$$

On a donc déjà dégagé une condition nécessaire finalement assez triviale pour qu'une famille de lois finies dimensionnelles soit « issues » d'une loi commune sur $\mathcal{A}(\mathbb{R}^+, \mathbb{R})$ (i.e. famille de lois image d'une mesure μ par les applications de restrictions associées aux parties finies de \mathbb{R}^+). Ce qui est spectaculaire, c'est que la réciproque est vraie :

Théorème 2 (Théorème de Daniell (1918) - Kolmogorov (1933)). *Pour tout $I \subset \mathbb{R}^+$ fini on suppose donnée une probabilité μ_I sur $(\mathbb{R}^I, \mathcal{B}(\mathbb{R}^I))$. On suppose de plus que ces probabilités vérifient la condition de compatibilité (3). Alors, il existe une unique probabilité μ sur $(\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T})$ telle que pour tout $I \subset \mathbb{R}^+$ finie*

$$\mu_I = \mu_{\pi_I}.$$

On admet la preuve de ce théorème pour cette année (et probablement l'année prochaine!). On en déduit que pour toute famille compatible de probabilités fini dimensionnelles il existe un processus qui a ces lois finis dimensionnelles. En effet le théorème de Daniell Kolmogorov nous fournit une probabilité sur $\mathcal{A}(\mathbb{R}^+, \mathbb{R})$ et on construit ensuite le processus recherché grâce au processus canonique.

Un exemple très important d'utilisation de ce théorème est l'existence de processus gaussiens.

Définition 6. Un processus $(X_t)_{t \in \mathbb{T}}$ est dit *gaussien* si toutes ses lois finies dimensionnelles le sont : pour tout $n \geq 0$, pour tout $t_1, \dots, t_n \in \mathbb{T}$ et tout $\lambda_1, \dots, \lambda_n \in \mathbb{R}$,

$$\lambda \cdot (X_{t_1}, \dots, X_{t_n}) = \sum_{i=1}^n \lambda_i X_{t_i}$$

est une variable gaussienne.

Proposition 6. La loi d'un processus gaussien est caractérisée par :

1. sa **moyenne** :

$$\begin{aligned} m : \mathbb{T} &\rightarrow \mathbb{R} \\ t &\rightarrow \mathbb{E}(X_t). \end{aligned}$$

2. sa **covariance** :

$$\begin{aligned} R : \mathbb{T} \times \mathbb{T} &\rightarrow \mathbb{R} \\ (s, t) &\rightarrow \text{Cov}(X_s, X_t) = \mathbb{E}([X_s - \mathbb{E}(X_s)][X_t - \mathbb{E}(X_t)]). \end{aligned}$$

En effet, d'après la Proposition 5, la loi d'un processus est caractérisée par ses lois finies dimensionnelles et, comme ici les lois fini dimensionnelles sont gaussiennes, elles sont caractérisées par leurs moyennes et covariances (qui déterminent la fonction caractéristique). On note que la fonction de covariance est

1. **symétrique** : pour tout $s, t \in \mathbb{T}$, $R(s, t) = R(t, s)$,
2. **positive** (vous lirez parfois de type positive) : pour tout $n \geq 1$ et tout $t_1, \dots, t_n \in \mathbb{T}$ la matrice $(R(t_i, t_j))_{1 \leq i, j \leq n}$ est positive puisque pour tout $\lambda \in \mathbb{R}^n$,

$${}^t \lambda R \lambda = \mathbb{E} \left(\sum_{i=1}^n \lambda_i (X_{t_i} - m(t_i))^2 \right) \geq 0$$

Proposition 7 (Existence de processus gaussien). Soit $m : \mathbb{R}^+ \rightarrow \mathbb{R}$ et $R : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}$ de type positif et symétrique. Alors il existe un processus gaussien $(X_t)_{t \geq 0}$ de moyenne m et variance R . Il est unique en loi.

Preuve. Utilisez le Théorème 2, voir TD2. □

Voici un exemple important de processus gaussien qui va nous occuper une bonne partie du semestre. On cherche à construire un **processus gaussien centré à accroissements indépendants et stationnaires et de variance t au temps $t \geq 0$** .

On dit d'un processus $(X_t)_{t \geq 0}$ qu'il est à

1. **accroissements indépendants** si pour tout $n \geq 1$ et tous réels $0 \leq t_1 \leq \dots \leq t_n$, les variables aléatoires $X_{t_1} - X_0, \dots, X_{t_n} - X_{t_{n-1}}$ sont indépendantes.
2. **accroissements stationnaires** si pour tout $0 \leq s < t$ les variables aléatoires $X_t - X_s$ et $X_{t-s} - X_0$ ont même loi.

Revenons à notre problème : si un tel processus existe il vérifie pour tout $0 \leq s < t$,

$$R(s, t) = E(X_s X_t) = E((X_t - X_s)X_s) + E(X_s^2) \stackrel{(PAIS)}{=} E(X_t - X_s)E(X_s) + E(X_s^2) = s.$$

La fonction $(s, t) \rightarrow s \wedge t$ est bien de type positif et on peut donc définir le processus gaussien associé à cette fonction de covariance R et à la fonction moyenne $m = 0$. On a presque construit le mouvement brownien : il nous manque la continuité qui va nous occuper dans la prochaine section.

2.2 Processus continus

On a travaillé jusqu'à présent dans le cadre très général $(\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T})$. Comme cet espace est énorme (ou que la tribu \mathcal{T} est trop petite), on note que certains ensembles que l'on voudrait étudier ne sont pas des événements. On peut par exemple montrer (TD2) que les ensembles

$$\begin{aligned} & \{f \in \mathcal{A}(\mathbb{R}^+, \mathbb{R}), \sup_{t \in [0,1]} f(t) < 1\} \text{ et} \\ & \{f \in \mathcal{A}(\mathbb{R}^+, \mathbb{R}), \exists t \in [0,1] f(t) = 0\}, \end{aligned}$$

ne sont pas dans \mathcal{T} (et l'ensemble des fonctions continues non plus d'ailleurs...exercice encore!).

2.2.1 Tribu(s)

Dans ce cours on s'intéresse principalement à des **processus continus**. On va donc travailler dans l'espace des fonctions continues $\mathcal{C}(\mathbb{R}^+, \mathbb{R})$. On définit sur cet espace la tribu cylindrique \mathcal{C} , plus petite tribu rendant les applications coordonnées mesurables. Dans ce cadre les applications coordonnées sont définies pour tout $t \geq 0$ par

$$\begin{aligned} \pi_t : \mathcal{C}(\mathbb{R}^+, \mathbb{R}) & \rightarrow \mathbb{R} \\ \omega & \rightarrow \omega(t) \end{aligned}$$

De façon équivalente, on peut définir la tribu \mathcal{C} comme la plus petite tribu contenant les cylindres de $\mathcal{C}(\mathbb{R}^+, \mathbb{R})$ c'est-à-dire un ensemble de la forme

$$C : \{f \in \mathcal{C}(\mathbb{R}^+, \mathbb{R}) \text{ tel que } f(t_1) \in B_1, \dots, f(t_n) \in B_n\},$$

où t_1, \dots, t_n sont des réels positifs et B_1, \dots, B_n des boréliens de \mathbb{R} .

Définition 7. On dit d'un processus $(X_t)_{t \geq 0}$ qu'il est **continu** si pour tout $\omega \in \Omega$, la fonction

$$t \in \mathbb{R} \rightarrow X_t(\omega)$$

est continue.

Là encore on peut voir X comme une fonction aléatoire

$$X : (\Omega, \mathcal{F}, P) \rightarrow (\mathcal{C}(\mathbb{R}^+, \mathbb{R}), \mathcal{C}).$$

Il y a cependant sur $\mathcal{C}(\mathbb{R}^+, \mathbb{R})$ une autre tribu « naturelle » : celle de la convergence uniforme sur les compacts. On va se contenter d'étudier ces différentes tribus quand $T = [0, 1]$ et non \mathbb{R}^+ pour se ramener à la topologie de la norme uniforme. On ne perd pas grand chose en compréhension à se limiter à ce cadre et on pourra étudier le cadre général dans l'exercice 10 du TD2. On rappelle donc que la norme uniforme sur $\mathcal{C}([0, 1], \mathbb{R})$ est définie pour tout fonction f de cet espace par

$$\|f\|_\infty = \sup\{|f(t)|; t \in [0, 1]\}.$$

Cette norme définit une distance qui elle-même définit une topologie (=un ensemble de parties ouvertes). On peut donc considérer la tribu borélienne \mathcal{B} sur $\mathcal{C}([0, 1], \mathbb{R})$, la plus petite tribu contenant tous les ouverts pour la norme uniforme.

Proposition 8. Les tribus cylindrique et borélienne sur $\mathcal{C}([0, 1], \mathbb{R})$ coïncident :

$$\mathcal{B} = \mathcal{C}.$$

Preuve. $\underline{\mathcal{B} \supset \mathcal{C}}$. On va montrer que \mathcal{B} rend mesurable toutes les applications coordonnées. En effet pour tout $t \geq 0$,

$$\pi_t : (\mathcal{C}(\mathbb{R}^+, \mathbb{R}), \|\cdot\|_\infty) \rightarrow (\mathbb{R}, |\cdot|)$$

est continue donc mesurable.

$\underline{\mathcal{B} \subset \mathcal{C}}$. Comme $\mathcal{C}([0, 1], \mathbb{R})$ muni de la norme uniforme est séparable, la tribu borélienne est engendrée par les boules ouvertes ou encore, par les boules fermées (Exercice 13 du TD2). Il nous suffit donc de montrer que toutes les boules fermées sont dans \mathcal{C} . On considère une fonction $f \in \mathcal{C}([0, 1], \mathbb{R})$ et $\varepsilon > 0$. En utilisant la continuité, on obtient

$$\begin{aligned} B_f(f, \varepsilon) &= \{g \in \mathcal{C}([0, 1], \mathbb{R}) \text{ tel que } \|g - f\|_\infty \leq \varepsilon\} \\ &= \bigcap_{t \in [0, 1]} \{g \in \mathcal{C}([0, 1], \mathbb{R}) \text{ tel que } g(t) \in [f(t) - \varepsilon, f(t) + \varepsilon]\} \\ &= \bigcap_{t \in [0, 1] \cap \mathbb{Q}} \{\pi_t \in [f(t) - \varepsilon, f(t) + \varepsilon]\}. \end{aligned}$$

On en déduit que $B_f(f, \varepsilon) \in \mathcal{C}$ et cela conclut la preuve. \square

On notera que $(\mathcal{C}([0, 1], \mathbb{R}), \mathcal{B})$ constitue un espace plus confortable pour travailler que $(\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T})$: par exemple la fonction « sup » sur cet espace est cette fois mesurable (Exercice 6 TD 2).

2.2.2 Modification continue et critère de Kolmogorov

On considère deux processus $(X_t)_{t \geq 0}$ et $(Y_t)_{t \geq 0}$ (non nécessairement continus). On dit que

1. $(Y_t)_{t \geq 0}$ est une **modification** de $(X_t)_{t \geq 0}$ si pour tout $t \geq 0$

$$P(X_t = Y_t) = 1.$$

2. $(X_t)_{t \geq 0}$ et $(Y_t)_{t \geq 0}$ sont **indistinguables** si

$$P(\forall t \geq 0, X_t = Y_t) = 1.$$

Cela signifie que les processus sont égaux *en tant que fonction aléatoire*.

L'ensemble que l'on considère dans la définition de « indistinguishable » est il bien un événement ? On peut vérifier (exercice !) que c'est bien un élément de \mathcal{C} mais que ça n'est par contre pas un événement de \mathcal{T} . La façon dont est énoncé cette définition est donc un petit abus : il faut comprendre que le complémentaire de cette partie est négligeable c'est-à-dire incluse dans un événement de probabilité nulle ou de façon équivalente : il existe $\Omega' \in \mathcal{T}$ tel que $P(\Omega') = 1$ et pour tout $\omega \in \Omega'$ et tout $t \geq 0$, $X_t(\omega) = Y_t(\omega)$.

Proposition 9. *On considère deux processus $(X_t)_{t \geq 0}$ et $(Y_t)_{t \geq 0}$*

1. *Si X est une modification de Y alors X et Y ont même loi.*
2. *Si X et Y sont indistinguables alors X est une modification de Y .*
3. *Si X est une modification de Y et que X et Y sont continus (où seulement càd ou seulement làg) alors X et Y sont indistinguables.*

Preuve. 1. Si X est une modification de Y alors les deux processus ont même lois fini dimensionnelles et donc même loi.

2. C'est facile : $\cap_{t \geq 0} \{X_t = Y_t\} \subset \{X_t = Y_t\}$ pour tout $t \geq 0$.

3. En utilisant la continuité on obtient que $\cap_{t \in \mathbb{R}^+} \{X_t = Y_t\} = \cap_{t \in \mathbb{Q}^+} \{X_t = Y_t\}$ et donc

$$P(\exists t \geq 0, X_t \neq Y_t) = P(\exists t \in \mathbb{Q}^+, X_t \neq Y_t) \leq \sum_{t \in \mathbb{Q}^+} P(X_t \neq Y_t) = 0.$$

□

On considère un processus $(X_t)_{t \geq 0}$. À quelle condition (suffisante) existe-t-il une modification continue de ce processus ? Autrement dit, peut-on rendre X continue en ne changeant pour tout $t \geq 0$ les valeurs de X_t que pour un événement de probabilité nulle.

Théorème 3 (Critère de Kolmogorov). *Soit $(X_t)_{0 \leq t \leq 1}$ un processus. On suppose qu'il existe $q, \varepsilon, C > 0$ tels que pour tous $s, t \in [0, 1]$,*

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

Alors il existe une modification Y de X dont les trajectoires sont hölderiennes d'exposant $\alpha \in]0, \varepsilon/q[$: pour tout $\alpha \in]0, \varepsilon/q[$ et tout ω , il existe une constante $C_\alpha(\omega)$ tel que pour tout $s, t \in [0, 1]$,

$$|Y_s(\omega) - Y_t(\omega)| \leq C_\alpha(\omega)|t - s|^\alpha.$$

En particulier Y est une modification continue de X (unique à indistaguabilité près).

Remarque 1. *Si on travaille avec des processus sur \mathbb{R}^+ et non $[0, 1]$, on obtient seulement que la modification est « localement hölderienne » (c'est-à-dire hölderienne sur tout compact) car les constantes pourraient diverger.*

Preuve. Au tableau ! On suivra la preuve de [7]. □

3 Mouvement brownien

[Intro]

3.1 Définitions

Dans cette partie on propose plusieurs définitions du mouvement brownien (et on montre qu'elles sont bien équivalentes!). Relire la définition d'un processus à accroissement stationnaires avant de vous lancer dans ces définitions.

Définition 8 (B1). On appelle **mouvement brownien** tout processus **continu** $(B_t)_{t \geq 0}$ (c'est-à-dire tel que pour tout $\omega \in \Omega$ la fonction $t \in \mathbb{R}^+ \rightarrow B_t(\omega)$ est continue) à **accroissements indépendants gaussiens** et tel que $B_0 = 0$ p.s. et pour tout $0 \leq s \leq t$

$$B_t - B_s \rightsquigarrow \mathcal{N}(0, t - s).$$

On note que la variance en t du mouvement brownien signe le comportement diffusif de ce processus.

Définition 9 (B2). On appelle **mouvement brownien** tout processus **continu** $(B_t)_{t \geq 0}$ **gaussien centré** de fonction de variance définie pour tout $0 \leq s \leq t$ par

$$R(s, t) = s \wedge t.$$

Définition 10 (B3). On appelle **mouvement brownien** tout processus **continu** $(B_t)_{t \geq 0}$ tel que $B_0 = 0$ p.s. et pour tout $0 \leq s \leq t$

$$\begin{aligned} B_t - B_s &\perp \sigma(B_r; 0 \leq r \leq s), \\ B_t - B_s &\rightsquigarrow \mathcal{N}(0, t - s). \end{aligned}$$

Preuve de l'équivalence des trois définitions. B1 \implies B2. Pour tout $t \geq 0$, $E(B_t - B_0) = 0$ et comme $B_0 = 0$ p.s. on obtient que $(B_t)_{t \geq 0}$ est centré. Montrons que ce processus est gaussien. Soit $n \geq 1$ et $0 \leq t_1 \leq \dots \leq t_n$, on a

$$\sum_{i=1}^n \lambda_i B_{t_i} = \sum_{i=1}^n \mu_i (B_{t_i} - B_{t_{i-1}})$$

avec $t_0 = 0$ et

$$\mu_i = \sum_{j=i}^n \lambda_j, \quad i = 1, \dots, n.$$

Comme les accroissements sont indépendants et gaussiens, on en déduit que la combinaison linéaire est gaussienne et donc que $(B_t)_{t \geq 0}$ est gaussien. Enfin pour tout $0 \leq s \leq t$,

$$R(s, t) = E(B_s B_t) = E((B_t - B_s)B_s) + E(B_s^2) = s.$$

B2 \implies B3. Comme $(B_t)_{t \geq 0}$ est centré et $E(B_0^2) = R(0, 0) = 0$, on en déduit que $B_0 = 0$ p.s. Soit $0 \leq s \leq t$. Par un argument de classe monotone il suffit de montrer que pour tout $n \geq 1$ et tout $0 \leq r_1 \leq \dots \leq r_n \leq s$, $B_t - B_s$ est indépendant de $(B_{r_1}, \dots, B_{r_n})$. Comme $(B_t)_{t \geq 0}$ est un processus gaussien il nous suffit de vérifier que les covariances sont nulles. Or pour tout $1 \leq i \leq n$

$$E((B_t - B_s)B_{r_i}) = r_i - r_i = 0.$$

Enfin l'accroissement $B_t - B_s$ est bien gaussien centré et on vérifie que sa variance vaut

$$E((B_t - B_s)^2) = t + s - 2(s \wedge t) = t - s.$$

B3 \implies B1. Il nous faut montrer uniquement l'indépendance des accroissements. On se donne donc une famille $0 = t_0 \leq t_1 \leq \dots \leq t_n$ et des fonctions continues bornées ϕ_1, \dots, ϕ_n , et comme $B_{t_n} - B_{t_{n-1}} \perp \sigma(B_r; 0 \leq r \leq t_{n-1})$,

$$E(\phi_n(B_{t_n} - B_{t_{n-1}}) \dots \phi_1(B_{t_1} - B_{t_0})) = E(\phi_n(B_{t_n} - B_{t_{n-1}}))E(\phi_{n-1}(B_{t_{n-1}} - B_{t_{n-2}}) \dots \phi_1(B_{t_1} - B_{t_0}))$$

et on conclut facilement en itérant. \square

3.2 Existence

On s'intéresse maintenant à la question de l'existence d'un tel processus. Heureusement on a déjà fait le travail! On utilise la Définition B2. On a déjà vu comment le théorème de Daniell Kolmogorov (Théorème 2) permet de construire un processus $(B_t)_{t \geq 0}$ gaussien de moyenne $m = 0$ et de fonction de variance $R(s, t) = s \wedge t$ ($s, t \in \mathbb{R}^+$). Il nous reste à nous assurer de l'existence d'une modification continue d'un tel processus (qui aura donc même loi!). Pour cela on va utiliser le critère de Kolmogorov (Théorème 3) : pour tout $0 \leq s < t$, $B_t - B_s$ suit une loi gaussienne de variance $t - s$ donc a même loi que $\sqrt{t - s}N$ où N est une gaussienne centrée réduite. Donc pour tout $q > 0$,

$$E(|B_t - B_s|^q) = |t - s|^{q/2} E(|N|^q) = c_q |t - s|^{q/2},$$

où c_q désigne le q -ième moment de la gaussienne centrée réduite. On peut donc appliquer le critère de Kolmogorov en prenant pour tout $q > 2$, $\varepsilon =$

$q/2 - 1$. On en déduit que B a une modification \tilde{B} qui est localement hölderienne d'exposant α pour tout

$$\alpha < \frac{q/2 - 1}{q} = \frac{1}{2} - \frac{1}{q}.$$

On obtient donc une modification qui est localement hölderienne d'exposant arbitrairement proche de $1/2$ en faisant tendre q vers l'infini. En particulier \tilde{B} est continu et cela conclut la preuve de l'existence du mouvement brownien. On a, au passage, également prouvé la

Proposition 10. *Soit $(B_t)_{t \geq 0}$ un mouvement brownien. Alors p.s. les trajectoires de $(B_t)_{t \geq 0}$ sont localement hölderiennes pour tout $\gamma \in [0, \frac{1}{2}[$.*

Preuve. On vient de voir que B a une modification X qui satisfait cette propriété. Or B et X sont continus tous deux donc ces deux processus sont indistinguables. \square

3.3 Mesure de Wiener et processus canonique

On appelle **mesure de Wiener** et on note W la loi du mouvement brownien en tant que fonction aléatoire sur l'espace $(\mathcal{C}(\mathbb{R}^+, \mathbb{R}), \mathcal{B})$. On appelle construction canonique du mouvement brownien ou **mouvement brownien canonique**, le processus canonique sur cet espace :

$$\begin{aligned} B_t : (\mathcal{C}(\mathbb{R}^+, \mathbb{R}), \mathcal{B}, W) &\rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R})) \\ \omega &\mapsto B_t(\omega) = \omega(t) = \pi_t(\omega) \end{aligned}$$

3.4 Premières propriétés

Proposition 11. *Soit $(B_t)_{t \geq 0}$ un mouvement brownien. Alors*

1. [**Symétrie**] $(-B_t)_{t \geq 0}$ est aussi un mouvement brownien ;
2. [**Invariance par changement d'échelle diffusif**] pour tout $\lambda > 0$, $(B_t^\lambda)_{t \geq 0}$ défini pour tout $t \geq 0$ par

$$B_t^\lambda = \frac{1}{\lambda} B_{\lambda^2 t}$$

est aussi un mouvement brownien ;

3. [**Markov simple**] pour tout $s \geq 0$, le processus $(B_t^{(s)})_{t \geq 0}$ défini pour $t \geq 0$ par

$$B_t^{(s)} = B_{s+t} - B_s$$

est un mouvement brownien indépendant de $\sigma(B_r, r \leq s)$.

Preuve. Pour les trois points on note que la continuité des trajectoires ne pose pas problème. L'item 1. est très facile avec la Définitions B1 ou B2. Pour le second item on peut utiliser la deuxième définition et calculer la fonction de covariance K^λ de $(B_t^\lambda)_{t \geq 0}$: pour tout $s, t \geq 0$

$$K^\lambda(s, t) = \frac{1}{\lambda^2} \mathbb{E}(B_{\lambda^2 s} B_{\lambda^2 t}) = \frac{1}{\lambda^2} (\lambda^2 s \wedge \lambda^2 t) = s \wedge t.$$

Pour le dernier point on utilise à nouveau la Définition B2 : le processus est bien gaussien centré et il ne nous reste que à calculer la covariance $K^{(s)}$: pour tout $t, u \geq 0$,

$$K^{(s)}(t, u) = \mathbb{E}((B_{s+t} - B_s)(B_{s+u} - B_s)) = (s+t) \wedge (s+u) - s = u \wedge t.$$

□

Pour poursuivre notre étude du brownien, nous allons nous appuyer sur la propriété suivante (voir [7] pour cette séquence du cours)

Théorème 4 (Loi du 0 – 1 de Blumenthal). *Soit $(B_t)_{t \geq 0}$ un mouvement brownien. Pour tout $t \geq 0$, on note*

$$\mathcal{F}_t = \sigma(B_s; s \leq t) \quad \text{et} \\ \mathcal{F}_{0+} = \bigcap_{s > 0} \mathcal{F}_s.$$

Alors la tribu \mathcal{F}_{0+} est grossière c'est-à-dire que pour tout $A \in \mathcal{F}_{0+}$, $\mathbb{P}(A) \in \{0, 1\}$.

Intuitivement \mathcal{F}_{0+} représente une information : c'est ce que l'on peut dire en observant un morceau arbitrairement petit du brownien. Par exemple

$$\{\exists \varepsilon > 0 \text{ tel que pour tout } 0 \leq t \leq \varepsilon, B_t \geq 0\} \in \mathcal{F}_{0+}.$$

[preuve?]

Preuve. On va montrer que \mathcal{F}_{0+} est indépendante d'elle-même ce qui implique le résultat. Soit $A \in \mathcal{F}_{0+}$, $0 < t_1 < \dots < t_k$ et $g : \mathbb{R}^k \rightarrow \mathbb{R}$ continue bornée. Comme g est continue,

$$\lim_{\varepsilon \rightarrow 0}^{p.s.} 1_A g(B_{t_1} - B_\varepsilon, \dots, B_{t_k} - B_\varepsilon) = 1_A g(B_{t_1}, \dots, B_{t_k})$$

et qu'on a de plus une domination de la suite par $\|g\|_\infty$, on obtient par le théorème de convergence dominée,

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E}(1_A g(B_{t_1} - B_\varepsilon, \dots, B_{t_k} - B_\varepsilon)) = \mathbb{E}(1_A g(B_{t_1}, \dots, B_{t_k})).$$

Pour tout $\varepsilon > 0$, $A \in \mathcal{F}_\varepsilon$ et, par la propriété de Markov simple, pour tout $t > \varepsilon$, $B_t - B_\varepsilon$ est indépendant de \mathcal{F}_ε . Donc pour $\varepsilon > 0$ assez petit

$$E(1_A g(B_{t_1} - B_\varepsilon, \dots, B_{t_k} - B_\varepsilon)) = P(A)E(g(B_{t_1} - B_\varepsilon, \dots, B_{t_k} - B_\varepsilon)).$$

En utilisant à nouveau le théorème de convergence dominée on obtient finalement

$$E(1_A g(B_{t_1}, \dots, B_{t_k})) = P(A)E(g(B_{t_1}, \dots, B_{t_k})).$$

Comme cela vaut pour tout k et tout k -uplet on a en fait prouvé que

$$\mathcal{F}_{0+} \perp\!\!\!\perp \sigma(B_t, t > 0).$$

Or $\sigma(B_t, t > 0) = \sigma(B_t, t \geq 0)$ car $B_0 = \lim_{t \rightarrow 0, t > 0} B_t$ et, d'autre part, pour tout $\varepsilon > 0$,

$$\mathcal{F}_{0+} \subset \mathcal{F}_\varepsilon \subset \sigma(B_t, t \geq 0).$$

On a donc bien montré que \mathcal{F}_{0+} est indépendante d'elle-même. \square

Ce résultat important (et intéressant par lui-même) permet de dérouler de nouvelles propriétés pour le brownien.

Corollaire 1. *Soit $(B_t)_{t \geq 0}$ un mouvement brownien. On a alors les propriétés suivantes :*

1. *p.s. pour tout $\varepsilon > 0$, $\sup_{0 \leq s \leq \varepsilon} B_s > 0$ et $\inf_{0 \leq s \leq \varepsilon} B_s < 0$.*
2. *p.s. pour tout $\varepsilon > 0$ $(B_t)_{t \geq 0}$ a un zéro sur $]0, \varepsilon[$.*
3. *p.s. $(B_t)_{t \geq 0}$ n'est monotone sur aucun intervalle.*

Preuve. 1. Une remarque avant de démarrer la preuve : le brownien est un processus continu donc $\{\sup_{0 \leq s \leq \varepsilon} B_s > 0\}$ est bien un événement. On considère une suite $(\varepsilon_n)_{n \geq 1}$ décroissant vers 0 et on définit

$$A = \bigcap_{n \geq 1} \downarrow \left\{ \sup_{0 \leq s \leq \varepsilon_n} B_s > 0 \right\},$$

et on doit donc montrer que $P(A) = 1$. Il est clair que pour tout $t > 0$, $A \in \mathcal{F}_t$ puisque $A \in \mathcal{F}_{\varepsilon_n}$ pour tout n et en particulier pour n assez grand tel que $\varepsilon_n < t$. On en déduit que $A \in \mathcal{F}_{0+}$ et donc, en utilisant la loi du 0 – 1 de Blumenthal que $P(A) \in \{0, 1\}$. Or

$$P(A) = \lim_{n \rightarrow +\infty} \downarrow P\left(\sup_{0 \leq s \leq \varepsilon_n} B_s > 0\right)$$

et comme pour tout $n \geq 1$,

$$P\left(\sup_{0 \leq s \leq \varepsilon_n} B_s > 0\right) \geq P(B_{\varepsilon_n} > 0) \geq \frac{1}{2},$$

on conclut que $P(A) = 1$.

2. On déduit du point précédent que p.s. pour tout $\varepsilon > 0$, il y a un zéro de $(B_t)_{t \geq 0}$ sur $]0, \varepsilon[$. Cela implique bien sûr que p.s. $(B_t)_{t \geq 0}$ a une infinité de zéros au voisinage de 0.
3. On déduit du point 1. que p.s. pour tout $\varepsilon > 0$, $(B_t)_{t \geq 0}$ n'est pas monotone sur $[0, \varepsilon]$. On considère maintenant un intervalle aux bords rationnels $[s, t]$. D'après la propriété de Markov simple le processus $(B_t^{(s)})_{t \geq 0}$ est un mouvement brownien et donc n'est pas monotone sur $[0, t - s]$. On a donc montré

$$\forall 0 < s < t \in \mathbb{Q} \quad p.s. \quad (B_t)_{t \geq 0} \text{ n'est pas monotone sur } [s, t],$$

et comme \mathbb{Q} est dénombrable,

$$p.s. \quad \forall 0 < s < t \in \mathbb{Q} \quad (B_t)_{t \geq 0} \text{ n'est pas monotone sur } [s, t],$$

enfin, par densité de \mathbb{Q} dans \mathbb{R} et continuité des trajectoires de $(B_t)_{t \geq 0}$,

$$p.s. \quad \forall 0 < s < t \in \mathbb{R} \quad (B_t)_{t \geq 0} \text{ n'est pas monotone sur } [s, t].$$

□

On note, pour tout $a \in \mathbb{R}$, T_a le temps d'atteinte de a :

$$T_a = \inf\{s \geq 0 \text{ tel que } B_s = a\}.$$

Corollaire 2 (du corollaire). *Soit $(B_t)_{t \geq 0}$ un mouvement brownien. Alors p.s. pour tout $a \in \mathbb{R}$, $T_a < +\infty$.*

Preuve. Commençons par montrer que $T_1 < +\infty$ p.s. Comme $\{\sup_{0 \leq s \leq 1} B_s > 0\} = \{\exists n \geq 1; \sup_{0 \leq s \leq 1} B_s > 1/n\}$, on déduit du Corollaire 1

$$\lim_{n \rightarrow +\infty} \uparrow \mathbb{P}(\sup_{0 \leq s \leq 1} B_s > 1/n) = 1.$$

Or pour tout $n \geq 1$, en utilisant la propriété de scaling,

$$\mathbb{P}(\sup_{0 \leq s \leq 1} B_s > 1/n) = \mathbb{P}(\sup_{0 \leq s \leq n^2} nB_{s/n^2} > 1) = \mathbb{P}(\sup_{0 \leq s \leq n^2} B_s > 1).$$

Or $\lim_{n \rightarrow +\infty} \uparrow \mathbb{P}(\sup_{0 \leq s \leq n^2} B_s > 1) = \mathbb{P}(\sup_{s \geq 0} B_s > 1)$ et on en déduit que p.s. $\sup_{s \geq 0} B_s > 1$. En utilisant à nouveau le scaling, on obtient que pour tout $a > 0$,

$$\mathbb{P}(\sup_{s \geq 0} B_s > a) = \mathbb{P}(\sup_{s \geq 0} \frac{1}{a} B_{a^2 s} > 1) = \mathbb{P}(\sup_{s \geq 0} B_s > 1) = 1.$$

On obtient bien sûr un résultat similaire pour l'inf en travaillant avec $(-B_t)_{t \geq 0}$.

□

On a donc montré que p.s. $\limsup_{t \rightarrow +\infty} B_t = -\liminf_{t \rightarrow +\infty} B_t = +\infty$. En effet, comme les trajectoires sont continues, $\{\limsup_{t \rightarrow +\infty} B_t < +\infty\} \subset \{\sup_{t \rightarrow +\infty} B_t < +\infty\}$ qui est de probabilité nulle. On déduit facilement (toujours en utilisant la continuité des trajectoires) :

Corollaire 3 (du corollaire du corollaire). *Soit $(B_t)_{t \geq 0}$ un mouvement brownien. Alors p.s. l'ensemble des zéros du brownien est non borné.*

3.5 Propriété de Markov fort

Dans cette section on cherche à remplacer, dans la propriété de Markov, le temps déterministe par un temps aléatoire. On n'a bien sûr aucune chance que cela soit vrai en toute généralité (pourquoi ? Donnez un exemple !) et nous introduisons d'abord la notion de temps d'arrêt (en temps continu ici...mais la définition est similaire à ce que vous avez vu au premier semestre en temps discret). On définit donc pour tout $t \geq 0$,

$$\begin{aligned}\mathcal{F}_t &= \sigma(B_s, s \leq t) \\ \mathcal{F}_\infty &= \sigma(B_s, s \geq 0).\end{aligned}$$

La famille de tribu $(\mathcal{F}_t)_{t \geq 0}$ est une **filtration** c'est-à-dire que pour tout $0 \leq s \leq t$,

$$\mathcal{F}_s \subset \mathcal{F}_t.$$

Définition 11. *Une variable aléatoire T à valeurs dans $[0, +\infty]$ est un **temps d'arrêt** si pour tout $t \geq 0$,*

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

Si T est un temps d'arrêt, on définit la tribu

$$\mathcal{F}_T = \{A \in \mathcal{F}_\infty \text{ tel que pour tout } t \geq 0, A \cap \{T \leq t\} \in \mathcal{F}_t\}.$$

Il faut bien sûr montrer que la définition ci-dessus est pertinente, c'est-à-dire que \mathcal{F}_T est bien une tribu (Exercice !). On peut maintenant énoncer la propriété de Markov forte :

Théorème 5 (Propriété de Markov forte). *Soit $(B_t)_{t \geq 0}$ un mouvement brownien et T un temps d'arrêt tel que $P(T < +\infty) > 0$. On définit le processus*

$$B_t^{(T)} = 1_{\{T < +\infty\}}(B_{T+t} - B_T), \quad t \geq 0.$$

Alors, sous $P(\cdot | T < +\infty)$, $(B_t^{(T)})_{t \geq 0}$ est un mouvement brownien indépendant de \mathcal{F}_T .

Remarque 2. On note que sur l'événement $\{T = +\infty\}$, on a $B_t^{(T)} = 0$ pour tout $t \geq 0$.

Preuve. (voir [7]) On se contente du cas $T < +\infty$ p.s. Notre but : montrer que pour tout $A \in \mathcal{F}_T$, tout $0 \leq t_1 \leq \dots \leq t_p$ ($p \geq 1$) et toute fonction continue bornée $F : \mathbb{R}^p \rightarrow \mathbb{R}$,

$$\mathbb{E}(1_A F(B_{t_1}^{(T)}, \dots, B_{t_p}^{(T)})) = \mathbb{P}(A) \mathbb{E}(F(B_{t_1}, \dots, B_{t_p})). \quad (\star)$$

Cela permet en effet de montrer tout ce dont on a besoin :

1. par un argument de classe monotone, que le processus $B^{(T)}$ est indépendant de la tribu \mathcal{F}_T ,
2. en prenant $A = \Omega$, que le processus $(B^{(T)})_{t \geq 0}$ a même lois fini dimensionnelles que $(B_t)_{t \geq 0}$ et donc, à nouveau par le lemme de classe monotone que $(B^{(T)})_{t \geq 0}$ a même loi que $(B_t)_{t \geq 0}$. Cela prouve bien que $(B^{(T)})_{t \geq 0}$ est un mouvement brownien puisque la continuité des trajectoires ne pose pas de problème.

Il reste donc à montrer (\star) et pour cela on discrétise T afin de se ramener à Markov simple. On note T_n le plus petit rationnel de la forme $k/2^n$ supérieur ou égal à T (donc $T_n = \lceil T 2^n \rceil / 2^n$ et on pourra montrer qu'il s'agit aussi d'un temps d'arrêt). Pour tout $t \geq 0$, on a donc en utilisant la continuité du brownien, $B_t^{(T_n)} \rightarrow B_t^{(T)}$ et comme F est continue on obtient

$$\lim_{n \rightarrow +\infty} F(B_{t_1}^{(T_n)}, \dots, B_{t_p}^{(T_n)}) = F(B_{t_1}^{(T)}, \dots, B_{t_p}^{(T)}),$$

et par le théorème de convergence dominée

$$\lim_{n \rightarrow +\infty} \mathbb{E}(1_A F(B_{t_1}^{(T_n)}, \dots, B_{t_p}^{(T_n)})) = \mathbb{E}(1_A F(B_{t_1}^{(T)}, \dots, B_{t_p}^{(T)})).$$

Or pour tout $n \geq 1$

$$\mathbb{E}(1_A F(B_{t_1}^{(T_n)}, \dots, B_{t_p}^{(T_n)})) = \sum_{k=0}^{+\infty} \mathbb{E}(1_A 1_{\frac{k-1}{2^n} < T \leq \frac{k}{2^n}} F(B_{\frac{k}{2^n} + t_1} - B_{\frac{k}{2^n}}, \dots, B_{\frac{k}{2^n} + t_p} - B_{\frac{k}{2^n}})).$$

Comme $A \in \mathcal{F}_T$, $A \cap \{\frac{k-1}{2^n} < T \leq \frac{k}{2^n}\} \in \mathcal{F}_{k/2^n}$ et d'après Markov simple, on obtient pour tout $k \geq 0$,

$$\begin{aligned} & \mathbb{E}(1_A 1_{\frac{k-1}{2^n} < T \leq \frac{k}{2^n}} F(B_{\frac{k}{2^n} + t_1} - B_{\frac{k}{2^n}}, \dots, B_{\frac{k}{2^n} + t_p} - B_{\frac{k}{2^n}})) \\ &= \mathbb{P}\left(A \cap \left\{\frac{k-1}{2^n} < T \leq \frac{k}{2^n}\right\}\right) \mathbb{E}(F(B_{t_1}, \dots, B_{t_p})) \end{aligned}$$

En sommant sur k on obtient (\star) . □

3.6 Théorème de Donsker

Une propriété importante du Brownien, et même en fait une autre façon de l'introduire et de le construire, est de le voir comme limite d'une marche aléatoire sur \mathbb{Z} après un changement d'échelle pertinent. C'est l'objet du théorème de Donsker.

Commençons par rappeler la définition de la marche aléatoire simple sur \mathbb{Z} . On considère donc une famille $(X_i)_{i \geq 1}$ de variables aléatoires i.i.d. de loi commune la loi uniforme sur $\{-1, 1\}$. On définit alors la marche par

$$S_0 = 0 \quad \text{et}$$

$$S_n = \sum_{i=1}^n X_i, \quad n \geq 1.$$

La variance de S_n est n et l'ordre de grandeur de la distance à l'origine de la marche au temps n est donc \sqrt{n} . Si on veut observer quelque chose de non dégénéré à grande échelle, nous devons donc renormaliser la marche en considérant un « scaling diffusif » : contracter pour n grand le temps par n (c'est-à-dire ramener l'intervalle $[0, n]$ à l'intervalle $[0, 1]$) et l'espace par \sqrt{n} (c'est-à-dire ramener l'intervalle $[-\sqrt{n}, \sqrt{n}]$ à l'intervalle $[-1, 1]$). Cela nous conduit à définir pour tout $n \geq 1$ le processus

$$S_t^{(n)} = \frac{1}{\sqrt{n}} \left\{ \sum_{i=1}^{\lfloor nt \rfloor} X_i + (nt - \lfloor nt \rfloor) X_{\lfloor nt \rfloor + 1} \right\}, \quad t \geq 0.$$

On a aussi transformé, par cette opération, la marche discrète en un processus continu en reliant simplement les points par des droites ! Le brownien apparait alors comme limite de cette suite (en n) de processus :

Théorème 6 (Théorème de Donsker). *La suite de processus $(S_t^{(n)})_{0 \leq t \leq 1}$, $n \geq 1$ converge en loi pour la topologie de la convergence uniforme vers le processus $(B_t)_{0 \leq t \leq 1}$ c'est-à-dire que pour toute fonction $F : \mathcal{C}([0, 1], \mathbb{R}) \rightarrow \mathbb{R}$, continue (pour la topologie uniforme) et bornée*

$$\mathbb{E}(F(S^{(n)})) \xrightarrow{n \rightarrow +\infty} \mathbb{E}(F(B)).$$

Preuve. Il est facile de voir la convergence des marginales finies dimensionnelles puisque ça n'est rien d'autre que le théorème centrale limite multidimensionnelle. Il manque un second ingrédient pour obtenir la convergence en loi de la suite de processus qui est la **tension** de cette suite de loi sur $\mathcal{C}([0, 1], \mathbb{R})$. C'est hors programme mais vous verrez peut-être cela l'année prochaine ! Une référence classique sur ces question de convergence de processus est [2]. □

[Ajouter un dessin]

DRAFT

4 Martingales continues

Ce chapitre est volontairement court. Il s'agit d'apprendre sur les martingales continues uniquement ce dont nous avons besoin pour étudier l'intégrale stochastique et les équations différentielles stochastique dans les deux prochains chapitres. Le théorème d'arrêt est également souvent un outil utile pour étudier le brownien (nous le verrons dans de nombreux exercices). De nombreux résultats sont admis car la preuve consiste souvent à se ramener au cas discret et à utiliser les résultats que vous avez prouvés au premier semestre dans le cours *Processus discrets*. Pour celles et ceux qui souhaitent en savoir plus je vous conseille [7].

4.1 Définition

Définition 12 (filtration). Une filtration $(\mathcal{F}_t)_{t \geq 0}$ est une suite croissante de sous-tribus de \mathcal{F} .

Un exemple important est la **filtration engendrée par le processus** $(X_t)_{t \geq 0}$: pour tout $t \geq 0$

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

La tribu \mathcal{F}_t contient l'information relative à la trajectoire de X jusqu'au temps t . Un processus $(X_t)_{t \geq 0}$ est dit **adapté** à la filtration $(\mathcal{F}_t)_{t \geq 0}$ si pour tout $t \geq 0$, X_t est \mathcal{F}_t -mesurable.

Nous avons besoin maintenant d'énoncer une condition technique qui nous sera utile dans de nombreux théorèmes :

Définition 13 (Conditions usuelles ou conditions habituelles). Une filtration $(\mathcal{F}_t)_{t \geq 0}$ satisfait aux **conditions usuelles** si

1. la tribu \mathcal{F}_0 contient les négligeables (ce qui implique que pour tout $t \geq 0$, \mathcal{F}_t contient les négligeables) ;
2. pour tout $t \geq 0$, $\mathcal{F}_{t+} := \bigcap_{\varepsilon > 0} \mathcal{F}_{t+\varepsilon}$ coïncide avec \mathcal{F}_t (on dit que la filtration est continue à droite).

On rappelle qu'un sous-ensemble de Ω est dit négligeable s'il est inclus dans un événement (un élément de \mathcal{F} donc) de probabilité nulle. En pratique on peut toujours transformer une filtration pour qu'elle satisfasse aux conditions usuelles en considérant (et complétant) \mathcal{F}_{t+} au lieu de \mathcal{F}_t . Cette condition nous permet notamment de nous assurer qu'une modification de processus adapté est encore adapté. Cela va nous servir en particulier pour les régularisations de martingales.

Définition 14 (sur/sous/ \emptyset martingales). Un processus $(M_t)_{t \geq 0}$ est une (sur/sous/ \emptyset) (\mathcal{F}_t) -martingale (où (\mathcal{F}_t) est une filtration) s'il est

1. **adapté** à $(\mathcal{F}_t)_{t \geq 0}$,
2. **intégrable** i.e. pour tout $t \geq 0$, $E(|M_t|) < +\infty$,
3. pour tout $0 \leq s \leq t$, $E(M_t | \mathcal{F}_s) = M_s$ (resp \leq, \geq)

4.2 Convergence et régularisation

Dans cette section on étudie le comportement asymptotique et la régularisation.

Proposition 12 (Convergence p.s.). Soit $(M_t)_{t \geq 0}$ une martingale à trajectoires **càdlàg** et borné dans L^1 . Alors il existe $M_\infty \in L^1$ tel que

$$M_t \xrightarrow{p.s.} M_\infty.$$

Preuve. [Quelques idées seulement!] On regarde en fait les limsup et liminf pour t dans \mathcal{D} où \mathcal{D} est un ensemble dénombrable dense dans \mathbb{R} . On peut alors appliquer le théorème de convergence pour les martingales discrètes. On revient au cas continu en utilisant la continuité à droite. Pour montrer que M_∞ est dans L^1 on utilise le lemme de Fatou (exercice!). De manière générale, pensez à réviser les résultats du premier semestre sur les martingales discrètes. \square

On rappelle les deux définitions possibles (et bien sûr équivalentes) d'une famille de variables aléatoires *uniformément intégrable* :

Définition 15. On dit qu'une famille $(X_i)_{i \in I}$ (où I est un ensemble quelconque) de variables aléatoires est *uniformément intégrable* si

$$\lim_{a \rightarrow +\infty} \sup_{i \in I} E(|X_i| 1_{|X_i| > a}) = 0.$$

Définition 16. On dit qu'une famille $(X_i)_{i \in I}$ (où I est un ensemble quelconque) de variables aléatoires est *uniformément intégrable* si

1. la famille $(X_i)_{i \in I}$ est bornée dans L^1 **et**
2. pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que pour tout $A \in \mathcal{F}$ tel que $P(A) < \delta$ et tout $i \in I$,

$$E(|X_i| 1_A) < \varepsilon.$$

Proposition 13 (Convergence L^1). Soit $(M_t)_{t \geq 0}$ une martingale à trajectoires **càdlàg**. Les trois conditions sont équivalentes :

1. $(M_t)_{t \geq 0}$ converge dans L^1 et p.s. vers M_∞ ,
2. $(M_t)_{t \geq 0}$ est fermée,
3. $(M_t)_{t \geq 0}$ est uniformément intégrable.

Preuve. On rappelle qu'une martingale $(M_t)_{t \geq 0}$ est dite fermée s'il existe une variable $M_\infty \in L^1$ telle que pour tout $t \geq 0$, $M_t = E(M_\infty | \mathcal{F}_t)$. La preuve fonctionne comme dans le cas discret que vous avez vu au premier semestre. Exercice! \square

Théorème 7 (Théorème de régularisation de Doob). *Soit $(M_t)_{t \geq 0}$ une martingale satisfaisant aux conditions habituelles. Alors il existe une modification $(\tilde{M}_t)_{t \geq 0}$ de $(M_t)_{t \geq 0}$ telle que*

1. $(\tilde{M}_t)_{t \geq 0}$ est une martingale
2. $(\tilde{M}_t)_{t \geq 0}$ est à trajectoires localement bornées et càdlàg.

Preuve. La preuve est admise. Elle s'appuie sur le théorème de convergence p.s. On pourra à titre d'exercice montrer que lorsque la filtration satisfait aux conditions usuelles une modification de martingale est encore une martingale. \square

4.3 Théorème d'arrêt de Doob

On considère dans toute cette partie une filtration $(\mathcal{F}_t)_{t \geq 0}$.

Théorème 8 (d'arrêt de Doob). *Soit $(M_t)_{t \geq 0}$ une martingale à trajectoires càdlàg et uniformément intégrable. Soit S et T deux temps d'arrêt tels que $S \leq T$. Alors*

1. M_S et M_T sont intégrables,
2. $M_S = E(M_T | \mathcal{F}_S)$,
3. $M_S = E(M_\infty | \mathcal{F}_S)$.

En particulier, dans ce cas, $E(M_S) = E(M_\infty) = E(M_0)$.

On note que dans ce théorème la variable M_T est bien définie même si T n'est pas supposé fini p.s. puisque la variable M_∞ est bien définie. Sur l'événement $\{T = +\infty\}$, on a donc $M_T = M_\infty$. À nouveau on ne détaillera pas la preuve de ce théorème qui repose sur le théorème similaire dans le cas discret. Pour ce ramener au cas discret on introduira la suite de temps d'arrêt $T_n = \lceil 2^n T \rceil / 2^n$ qui converge p.s. en décroissant vers T . On notera les autres formes utiles du théorème d'arrêt :

Théorème 9. *Soit $(M_t)_{t \geq 0}$ une martingale à trajectoires càdlàg et S et T deux temps d'arrêt **bornés** tels que $S \leq T$. Alors*

1. M_S et M_T sont intégrables,
2. $M_S = E(M_T | \mathcal{F}_S)$,

Preuve. Ce théorème se déduit aisément du précédent puisque, en notant C une borne supérieure de T , on a que $(M_{t \wedge C})_{t \geq 0}$ est une martingale fermée par M_C donc uniformément intégrable. \square

Théorème 10 (d'arrêt de Doob). *Soit $(M_t)_{t \geq 0}$ une martingale à trajectoires càdlàg et T un temps d'arrêt. Alors*

1. le processus $(M_{t \wedge T})_{t \geq 0}$ est une martingale,
2. si $(M_t)_{t \geq 0}$ est uniformément intégrable alors $(M_{t \wedge T})_{t \geq 0}$ est aussi uniformément intégrable et pour tout $t \geq 0$

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

4.4 Inégalités maximales de Doob

Ces inégalité sont très utiles et à connaître. Elles nous serviront en particulier pour l'étude de l'intégrale stochastique.

Théorème 11. *Soit $(M_t)_{t \geq 0}$ une martingale **càd**. On note $M_t^* = \sup_{0 \leq s \leq t} |M_s|$. Alors*

1. Pour tout $p \geq 1$ et tout $t \geq 0$, pour tout $\lambda > 0$,

$$P(M_t^* \geq \lambda) \leq \frac{E(|M_t|^p)}{\lambda^p}.$$

2. Si $p > 1$ et $t > 0$

$$\|M_t^*\|_p \leq \frac{p}{p-1} \|M_t\|_p.$$

La preuve est un exercice classique que vous avez sans doute vu au premier semestre dans le cas discret. Pour se rappeler de la première inégalité (qui permet de montrer la seconde) on peut remarquer qu'il s'agit en fait d'une « super » inégalité de Markov : au lieu de contrôler la probabilité que la seule variable finale M_t soit grande, on peut en fait (et c'est bien parce que $(M_t)_{t \geq 0}$ est une martingale) contrôler le supremum de toute la trajectoire entre 0 et t . On notera aussi que le contrôle dans le second point n'est valable que pour $p > 1$ (ou plutôt qu'il ne donne rien pour $p = 1$ puisque le membre de droite est infini).

5 Intégrale stochastique - Itô (1950)

On souhaite intégrer par rapport au mouvement brownien c'est-à-dire donner du sens, lorsque $(u_s)_{s \geq 0}$ est un processus et $(B_t)_{t \geq 0}$ un mouvement brownien, à l'expression

$$\int u_s dB_s.$$

La première idée (qui ne va pas marcher !) est de définir cette intégrale ω par ω c'est-à-dire de donner du sens pour tout ω à

$$\int u_s(\omega) dB_s(\omega).$$

Cela revient à se demander dans quel cadre on peut donner une définition satisfaisante d'une fonction par rapport à une autre.

5.1 Intégrale de Stieltjes

La théorie de Stieltjes-Riemman permet de construire l'intégrale d'une fonction continue f par rapport à une autre fonction g à la condition que celle-ci soit à variations bornées. On ne présente ici qu'un résumé de la construction de

$$\int_0^t f(s) dg(s),$$

et nous renvoyons à [7, 4, 6] pour des exposés différents et beaucoup plus complets (que nous avons utilisés pour ce chapitre).

1. **Le cas où g est positive, continue à droite et croissante sur $[0, t]$.** La fonction g définit alors une mesure μ sur $[0, t]$ par $\mu(\{0\}) = 0$ et $\mu([a, b]) = g(b) - g(a)$ pour tout $0 \leq a \leq b \leq t$. Autrement dit g est la fonction de répartition de μ (même si le terme est plutôt utilisé pour une mesure de probabilité en général). On peut alors définir

$$\int_0^t f(s) dg(s) = \int_0^t f(s) \mu(ds),$$

pour toute fonction f positive ou dans $L^1(\mu)$. On notera que si f est continue sur $[0, t]$ alors pour toute suite de subdivision $0 = t_0^n \leq t_1^n \leq \dots \leq t_{p_n}^n = t$ de pas tendant vers 0

$$\int_0^t f(s) dg(s) = \lim_{n \rightarrow +\infty} \sum_{i=1}^{p_n} f(t_{i-1}^n) (g(t_i^n) - g(t_{i-1}^n)). \quad (4)$$

En effet pour tout $n \geq 1$, $\sum_{i=1}^{p_n} f(t_{i-1}^n)(g(t_i^n) - g(t_{i-1}^n)) = \int_0^t f_n(s) dg(s)$ où f_n est la fonction constante par morceaux définie par $f(t) = f(t_{i-1}^n)$ si $t \in]t_{i-1}^n, t_i^n]$ et on conclut avec le théorème de convergence dominée.

2. **Le cas où g est à variations bornées.** On appelle variations d'une fonction continue $g : [0, t] \rightarrow \mathbb{R}$ la quantité :

$$V(t) = \sup_{\substack{n \geq 1 \\ t_0 \leq \dots \leq t_n}} \left\{ \sum_{i=0}^{n-1} |g(t_{i+1}) - g(t_i)| \right\}$$

où le supremum porte sur l'ensemble des partitions $0 = t_0 \leq \dots \leq t_n = t$ de $[0, t]$. Quand $V(t)$ est fini on dit que g est à *variations finies* ou *bornées*. On notera $V(g, t)$ quand on voudra spécifier la fonction que l'on considère. On dira que $g : \mathbb{R} \rightarrow \mathbb{R}$ est à variations bornées quand $t \rightarrow V(g, t)$ est bornée. On peut montrer que lorsqu'une fonction est à variations finies, le sup qui intervient dans la définition de la variation est également une limite quand on considère n'importe quelle suite de subdivisions dont le pas tend vers 0.

Si les fonctions à variations finies nous intéressent ici c'est en raison du résultat suivant :

Proposition 14. *Soit g une fonction continue sur $[0, t]$. Les deux propositions suivantes sont équivalentes :*

- (a) g est à variations finies ;
- (b) g est la différence de deux fonctions positives croissantes continues.

Preuve. Le sens réciproque (b) \implies (a) est le plus facile. En effet, on peut déjà vérifier qu'une fonction croissante est bien à variations finies. On peut de plus montrer, en utilisant l'inégalité triangulaire, que la somme de deux fonctions à variations finies et encore à variations finies.

Montrons la réciproque, d'abord dans le cas où $g(0) = 0$. On vérifie facilement que $V : s \rightarrow V(g, s)$ est croissante sur $[0, t]$ et positive. Montrons que c'est également le cas pour $W = V - g$. On considère donc $0 \leq s_1 \leq s_2 \leq t$ et on veut montrer que $V(s_1) - g(s_1) \leq V(s_2) - g(s_2)$. C'est bien le cas puisque

$$V(s_1) + g(s_2) - g(s_1) \leq V(s_1) + |g(s_2) - g(s_1)| \leq V(s_2).$$

De plus pour tout $0 \leq s \leq t$ et toute subdivision $(s_i)_{i=1, \dots, n}$ de $[0, s]$,

$$g(s) \leq \sum_{i=1}^n |g(s_i) - g(s_{i-1})| \leq V(s),$$

donc $V - g$ est positive. Quand $g(0)$ est non nulle on l'ajoute selon son signe à V ou à W . Il reste à montrer que V est continue (ce qui implique que W l'est également) : exercice (pas si facile)! \square

Avec cette proposition on peut construire l'intégrale par rapport à une fonction g à variations finies. On la décompose en $g = g^+ - g^-$. On associe à g^+ et g^- deux mesures μ^+ et μ^- et on définit la mesure signée $\mu = \mu^+ - \mu^-$ (une mesure signée est la différence de deux mesures positives). La décomposition de g en g^+ et g^- n'est pas unique et donc le couple de mesures associée (μ^+, μ^-) non plus mais on peut montrer que la mesure μ ne dépend pas de ce choix (on l'admet pour cette fois!). On définit alors la mesure positive $|\mu| = \mu^+ + \mu^-$ et pour $f \in L^1(|\mu|)$ on pose

$$\int_0^t f(s)dg(s) = \int_0^t f(s)\mu(ds) = \int_0^t f(s)\mu^+(ds) - \int_0^t f(s)\mu^-(ds).$$

Quand f est continue (4) est encore vraie.

Le problème est que cette théorie de l'intégration ne marche pas avec le brownien dont les trajectoires, p.s., ne sont pas à variations finies :

Proposition 15. *Soit $(B_t)_{t \geq 0}$ un mouvement brownien et $t \in \mathbb{R}^+$. Pour toute suite de subdivisions $0 = t_0^n \leq \dots \leq t_{p_n}^n = t$ de pas tendant vers 0*

$$\sum_{k=1}^{p_n} (B_{t_k^n} - B_{t_{k-1}^n})^2 \xrightarrow{L^2} t.$$

On en déduit que p.s. les trajectoires du mouvement brownien sont à variations infinies.

Preuve (voir [5]). On note $W_n = \sum_{k=1}^{p_n} (B_{t_k^n} - B_{t_{k-1}^n})^2$ et Δ_n le pas de la n -ième subdivision (pour alléger les notations on n'indique plus d'exposant n). Comme les accroissements sont stationnaires, pour tout $n \geq 1$,

$$E(W_n) = \sum_{k=1}^{p_n} (t_k - t_{k-1}) = t.$$

On obtient donc, comme les accroissements sont indépendants,

$$\|W_n - t\|_2^2 = \text{Var}(W_n) = \sum_{k=1}^{p_n} \text{Var}[(B_{t_k} - B_{t_{k-1}})^2].$$

Or, pour tout $1 \leq k \leq p_n$

$$\text{Var}[(B_{t_k} - B_{t_{k-1}})^2] = (t_k - t_{k-1})^2 \text{Var}(B_1^2) = 2(t_k - t_{k-1})^2,$$

et on en déduit

$$\|W_n - t\|_2^2 = 2 \sum_{k=1}^{p_n} (t_k - t_{k-1})^2 \leq 2\Delta_n t,$$

ce qui conclut la preuve du premier point puisque $\Delta_n \rightarrow 0$ quand n tend vers l'infini.

On en déduit qu'il existe une sous-suite de la suite des sous-divisions telle que $(W_{\phi(n)})_{n \geq 1}$ tendent vers t p.s. On remarque ensuite que

$$W_{\phi(n)} \leq \sup\{|B_{t_k} - B_{t_{k-1}}|, 1 \leq k \leq p_n\} V(B, t).$$

Comme $\sup\{|B_{t_k} - B_{t_{k-1}}|, 1 \leq k \leq p_n\} \rightarrow 0$ p.s. car le brownien est continue donc uniformément continue sur $[0, t]$, on déduit du premier point que p.s. $V(B, t) = +\infty$. \square

Remarque 3. *On peut montrer que la variation quadratique (qui consiste à prendre le carré des incréments plutôt que leurs valeurs absolus) d'une fonction \mathcal{C}^1 est nulle (voir Exercice ?). La proposition précédente nous dit donc que les variations du brownien sur petits temps sont bien plus grandes que pour des fonctions « lisses ». Intuitivement pour $\delta > 0$ petit $B_{t+\delta} - B_t$ est d'ordre $\sqrt{\delta}$ qui est bien beaucoup plus grand que δ au voisinage de 0.*

Il n'est donc pas possible de définir trajectoire par trajectoire l'intégrale par rapport au mouvement brownien et il va nous falloir changer de stratégie !

5.2 Intégrale d'Itô

L'idée générale est de définir l'intégrale comme le prolongement d'une isométrie entre deux espaces de Hilbert. Si on considère un processus « simple » de la forme

$$u_t = \sum_{i=1}^n F_i 1_{]t_{i-1}, t_i]}(t) \quad t \geq 0,$$

avec les F_i des variables aléatoires, il est naturel de vouloir que

$$\int u_t dB_t = \sum_{i=1}^n F_i (B_{t_i^n} - B_{t_{i-1}^n}).$$

On pourrait ensuite approximer un processus général par une suite de processus simples et passer à la limite L^2 (la partie précédente nous dit qu'il n'est pas possible d'avoir une convergence p.s.). Attention cependant : la façon dont on approxime le processus peut changer la limite comme le montre l'exemple suivant : que vaut $\int_0^T B_t dB_t$? On peut approximer $(B_t)_{0 \leq t \leq T}$ de façon anticipative par ϕ^n :

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

ou non anticipative par ψ_n :

$$\psi_t^n = \sum_{i=0}^{n-1} B_{t_i^n} 1_{]t_i^n, t_{i+1}^n]}.$$

Si on ne prend pas garde dans la définition de l'intégrale, on obtient alors

$$\begin{aligned} \int_0^T \phi_t^n dB_t - \int_0^T \psi_t^n dB_t &= \sum_{i=1}^{n-1} B_{t_{i+1}^n} (B_{t_{i+1}^n} - B_{t_i^n}) - \sum_{i=1}^{n-1} B_{t_i^n} (B_{t_{i+1}^n} - B_{t_i^n}) \\ &= \sum_{i=1}^{n-1} (B_{t_{i+1}^n} - B_{t_i^n})^2. \end{aligned}$$

Cette dernière quantité converge dans L^2 vers T d'après la Proposition 15 et on voit donc que le choix de l'approximation n'est pas anodin ! On va faire le choix d'une **approximation non anticipative** qui permettra notamment que l'intégrale stochastique soit une martingale.

On considère dans cette section une filtration $(\mathcal{F}_t)_{t \geq 0}$ qui satisfait aux conditions usuelles (voir la Définition 13) et $(B_t)_{t \geq 0}$ un \mathcal{F} -mouvement brownien c'est-à-dire un mouvement brownien adapté à \mathcal{F} et tel que pour tout $0 \leq s \leq t$:

$$B_t - B_s \perp \mathcal{F}_s.$$

Définition 17. Un processus $(X_t)_{t \geq 0}$ est dit :

1. **mesurable** si

$$X : (\mathbb{R}^+ \times \Omega, \mathcal{B}(\mathbb{R}^+) \otimes \mathcal{F}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

est mesurable.

2. **progressivement mesurable** si pour tout $t \geq 0$

$$X : ([0, t] \times \Omega, \mathcal{B}([0, t]) \otimes \mathcal{F}_t) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

est mesurable.

On remarque que si $(X_t)_{t \geq 0}$ est progressivement mesurable alors il est adapté et mesurable :

1. *Adapté.* Soit $t \geq 0$ et $A \in \mathcal{B}(\mathbb{R})$. Comme $(X_s)_{s \geq 0}$ est progressivement mesurable, $\{X|_{[0,t]} \in A\} = \{(s, \omega) \in [0, t] \times \Omega \text{ tel que } X_s(\omega) \in A\} \in \mathcal{B}([0, t]) \otimes \mathcal{F}_t$. On rappelle que si \mathcal{E} et \mathcal{F} sont des tribus sur E et F alors pour tout $A \in \mathcal{E} \otimes \mathcal{F}$ et tout $x \in E$, la section de A en x , $A_x = \{y \in F \text{ tel que } (x, y) \in A\}$ appartient à \mathcal{F} . Ici on en déduit que $\{X_t \in A\} = \{X|_{[0,t]} \in A\}_t \in \mathcal{F}_t$.
2. *Mesurable.* On vérifie que pour tout $A \in \mathcal{B}(\mathbb{R})$,

$$\begin{aligned} \{X \in A\} &= \{(s, \omega) \in \mathbb{R} \times \Omega \text{ tel que } X_s(\omega) \in A\} \\ &= \cup_{n \geq 1} \{(s, \omega) \in [0, n] \times \Omega \text{ tel que } X_s(\omega) \in A\}. \end{aligned}$$

Or, comme le processus est progressivement mesurable, pour tout $n \geq 1$, $\{(s, \omega) \in [0, n] \times \Omega \text{ tel que } X_s(\omega) \in A\} \in \mathcal{B}([0, n]) \otimes \mathcal{F}_n \subset \mathcal{B}(\mathbb{R}^+) \otimes \mathcal{F}$.

Proposition 16 (pour dédramatiser). *Soit $(X_t)_{t \geq 0}$ un processus adapté et càd. Alors $(X_t)_{t \geq 0}$ est progressivement mesurable.*

Preuve. Exercice : voir TD 5. □

On peut maintenant définir l'ensemble des processus que nous allons pouvoir intégrer par rapport au brownien :

Définition 18. *On note $L^2(Prog)$ l'ensemble des processus $(u_s)_{s \geq 0}$ progressivement mesurables tels que*

$$\mathbb{E} \int u_s^2 ds < +\infty.$$

Proposition 17. *L'espace $L^2(Prog)$ est un Hilbert.*

Preuve. La ligne de la preuve est de montrer que $L^2(Prog)$ est en fait un L^2 « normal » et c'est une conséquence du lemme suivant :

Lemme 3. *On définit*

$$\mathcal{P} = \{A \in \mathcal{B}(\mathbb{R}^+) \otimes \mathcal{F} \text{ tel que } X : (t, \omega) \rightarrow 1_A(t, \omega) \text{ est progressivement mesurable}\}.$$

*Alors \mathcal{P} est une tribu (appelée **tribu progressive**) et un processus $(X_t)_{t \geq 0}$ est progressivement mesurable si et seulement si la fonction*

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

est \mathcal{P} -mesurable.

Preuve. Vérifions déjà que \mathcal{P} est une tribu. Pour $A \in \mathcal{B}(\mathbb{R}^+) \otimes \mathcal{F}$, on note $I(A)|_t$ la fonction définie sur $([0, t] \times \Omega, \mathcal{B}([0, t] \otimes \mathcal{F}_t))$ par $I(A)|_t(s, \omega) = 1_A(s, \omega)$. Avec ces notations, $A \in \mathcal{P}$ si et seulement si pour tout $t \geq 0$, $I(A)|_t$ est mesurable.

1. On en déduit facilement que $\mathbb{R}^+ \times \Omega \in \mathcal{P}$.
2. Soit $A \in \mathcal{P}$ et $t \geq 0$. Comme $I(A^c)|_t = 1 - I(A)|_t$ c'est une fonction mesurable également. D'où $A^c \in \mathcal{P}$.
3. Soit $(A_n)_{n \geq 1}$ une suite de \mathcal{P} et $t \geq 0$. On a $I(\cup_{n \geq 1} A_n)|_t = 1 - \lim_{N \rightarrow +\infty} \prod_{n=1}^N I(A_n^c)|_t$. On en déduit que $\cup_{n \geq 1} A_n \in \mathcal{P}$.

On considère maintenant un processus $(X_t)_{t \geq 0}$ et $B \in \mathcal{B}(\mathbb{R})$. On suppose X progressivement mesurable. On doit montrer que $\{X \in B\} \in \mathcal{P}$ donc que pour tout $t \geq 0$, $I(X \in B)|_t$ est mesurable. Or $\{I(X \in B)|_t = 1\} = \{(s, \omega) \in [0, t] \times \Omega \text{ tel que } X(s, \omega) \in B\} \in \mathcal{B}([0, t] \otimes \mathcal{F}_t)$ car X est progressivement mesurable. Réciproquement supposons que X est \mathcal{P} -mesurable et fixons $t \geq 0$. Avec l'égalité que l'on vient d'établir montre que X est progressivement mesurable. \square

Il permet en tout cas de voir les processus progressivement mesurables comme des variables aléatoires de carré intégrable sur $(\mathbb{R}^+ \times \Omega, \mathcal{P}) : L^2(Prog)$ est donc un Hilbert comme n'importe quel L^2 ,

$$L^2(Prog) = L^2(\mathbb{R}^+ \times \Omega, \mathcal{P}, \text{Leb} \otimes P).$$

Le produit scalaire associé est défini pour tous processus $u, v \in L^2(Prog)$ par

$$\langle u, v \rangle = E \int u_t v_t dt,$$

et la norme par

$$\|u\|_{L^2(Prog)}^2 = E \int u_t^2 dt.$$

\square

On s'intéresse maintenant à un sous-ensemble de processus de $L^2(Prog)$ pour lesquels l'intégrale stochastique sera définie « facilement ».

Définition 19. Un processus $(u_t)_{t \geq 0}$ est dit **en escalier ou processus simple prévisible** s'il est de la forme

$$u_t : \sum_{i=0}^{n-1} F_i 1_{[t_i, t_{i+1}[}(t) \quad t \geq 0,$$

où $n \geq 0$ est un entier, $0 = t_0 \leq \dots \leq t_n$, et pour tout $i = 1, \dots, n$ la variable F_i est \mathcal{F}_{t_i} -mesurable et de carré intégrable. On note \mathcal{E} l'ensemble des processus en escalier.

On notera qu'un processus en escalier est non anticipatif et càdlàg. On va pouvoir définir facilement l'intégrale stochastique sur \mathcal{E} mais pour pouvoir la prolonger, il faut nous assurer que \mathcal{E} est assez gros :

Proposition 18. *L'ensemble \mathcal{E} est un s.e.v. dense de $L^2(Prog)$.*

Preuve. Il est facile de vérifier que \mathcal{E} a une structure d'espace vectoriel. On vérifie ensuite $\mathcal{E} \subset L^2(Prog)$. On considère $0 \leq t_1 \leq t_2$ et F une variable aléatoire \mathcal{F}_{t_1} -mesurable. On doit montrer que le processus $u : F1_{[t_1, t_2[}$ est progressivement mesurable. Pour tout $T \geq 0$,

$$\begin{aligned} \{u|_{[0, T]} \in A\} = & \{(t, \omega) \text{ tel que } t < t_1, t \leq T, 0 \in A\} \\ & \cup \{(t, \omega) \text{ tel que } t_1 \leq t < t_2, t \leq T, F \in A\} \\ & \cup \{(t, \omega) \text{ tel que } t_2 < t, t \leq T, 0 \in A\} \end{aligned}$$

Seul l'ensemble du milieu est un peu plus délicat : il est égal à $([t_1, t_2[\cap [0, T]) \times \{F \in A\}$ qui appartient bien à $\mathcal{B}([0, T]) \otimes \mathcal{F}_T$ et on notera que ça n'est pas le cas si F est seulement \mathcal{F}_{t_2} -mesurable.

Il nous reste à prouver la densité. Deux méthodes (au moins!) sont possibles. La première consiste à approximer $u \in L^2(Prog)$ par une suite de processus dans \mathcal{E} (en passant par les processus bornés puis les bornés à support compact, puis on ajoute continu et enfin on arrive à une suite approximante dans \mathcal{E}) ; on la trouve notamment dans [1, 5, 4] et je vous recommande de vous y reporter car c'est instructif. Nous allons ici suivre une autre ligne que l'on trouve développée dans [7] qui consiste à montrer que $\mathcal{E}^\perp = \{0\}$. Nous avons besoin pour cela de deux lemmes (qui sont par ailleurs intéressants pour eux-mêmes).

Lemme 4. *Soit $u \in L^2(prog)$ tel que pour tout $t \geq 0$ et tout $\omega \in \Omega$ $\int_0^t |u_s(\omega)| ds < +\infty$. Alors $t \rightarrow \int_0^t u_s ds$ est à variations finies et adapté.*

Preuve. (voir aussi [7] pour cette preuve) *Adapté.* On fixe $t \geq 0$. On rappelle que $u|_{[0, t]}$ est $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$ -mesurable et on doit montrer que $\int_0^t u_s ds$ est \mathcal{F}_t -mesurable. Si u est de la forme $1_A 1_{[\alpha, \beta]}$ avec $0 \leq \alpha \leq \beta \leq t$ et $A \in \mathcal{F}_t$, on vérifie bien que $\int_0^t u_s ds = 1_A(\beta - \alpha)$ est \mathcal{F}_t mesurable. On considère maintenant la classe $\mathcal{M} = \{\Gamma \in \mathcal{B}([0, t]) \otimes \mathcal{F}_t \text{ tel que } \int_0^t 1_\Gamma ds \text{ soit } \mathcal{F}_t\text{-mesurable}\}$. Il s'agit d'une classe monotone qui contient tout les Γ de la forme $[\alpha, \beta] \times A$ et par lemme de classe monotone on en déduit que \mathcal{M} contient tout $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$. On vient donc de traiter le cas de $u = 1_\Gamma$ où $\Gamma \in \mathcal{B}([0, t]) \otimes \mathcal{F}_t$. On obtient ensuite facilement le cas d'un processus u étagé puis d'un processus quelconque $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$ -mesurable par approximation par des étagées et le théorème de convergence dominée par rapport à la mesure de Lebesgue sur $[0, t]$.

Variations finies. De plus $\int_0^t u_s ds = \int_0^t u_s^+ ds - \int_0^t u_s^- ds$ peut s'écrire comme différence de deux fonctions croissantes donc est bien à variations finies. \square

Lemme 5. Soit $(M_t)_{t \geq 0}$ une martingale continue, à variations finies et telle que $M_0 = 0$ p.s. Alors $(M_t)_{t \geq 0}$ est indistinguable de 0.

Preuve. (voir aussi [5] pour cette preuve) On suppose dans un premier temps que $(M_t)_{t \geq 0}$ est à variations bornées par K p.s. i.e., p.s., pour tout $t \geq 0$, $V(M(\omega), t) < K$. Pour toute subdivision de $[0, t]$

$$\begin{aligned} \mathbb{E}(M_t^2) &= \mathbb{E}\left(\sum_{i=0}^{n-1} M_{t_{i+1}}^2 - M_{t_i}^2\right) \\ &\stackrel{Mart.}{=} \sum_{i=0}^{n-1} \mathbb{E}[(M_{t_{i+1}} - M_{t_i})^2] \\ &\leq K \mathbb{E}(\max |M_{t_{i+1}} - M_{t_i}|). \end{aligned}$$

On considère une suite de subdivisions de pas δ_n tendant vers 0. On a alors, par uniforme continuité, convergence p.s. de $|M_{t_{i+1}} - M_{t_i}|$ vers 0. De plus $|M|$ est bornée par K donc on peut utiliser le théorème de convergence dominée et on en déduit que $\mathbb{E}(M_t^2) = 0$ puis que p.s. $M_t = 0$. En utilisant que M est une version du processus nul et la continuité, on obtient que M est indistinguable de 0.

Si M est seulement à variations finies, on introduit la suite de temps d'arrêt

$$\tau_k = \inf\{s \text{ tel que } V(s) \geq k\}, \quad k \geq 1.$$

Les martingales arrêtées $(M_{\tau_k \wedge t})_{t \geq 0}$ sont à variations bornées et en utilisant ce que l'on vient de prouver on en déduit que pour tout $k \geq 1$ p.s. $M_{\tau_k \wedge t} = 0$ pour tout $t \geq 0$. On échange le quantificateur qui porte sur un ensemble dénombrable et le p.s. pour conclure. \square

Avec ces deux lemmes en mains, revenons à notre preuve de la densité de \mathcal{E} dans $L^2(Prog)$. On considère donc $u \in \mathcal{E}^\perp$ et on veut montrer que $u = 0$. On définit pour cela le processus

$$M_t = \int_0^t u_s ds \quad t \geq 0.$$

Il s'agit ici d'une intégrale bien définie ω par ω p.s. puisque

$$\mathbb{E}|M_t| \leq \mathbb{E} \int_0^{+\infty} |u_s| 1_{[0,t]}(s) ds \stackrel{C.S.}{\leq} \sqrt{t} \mathbb{E} \left(\int_0^{+\infty} |u_s|^2 ds \right)^{1/2} < +\infty.$$

On note même que $M_t \in L^1$. Montrons que le processus $(M_t)_{t \geq 0}$ est une martingale. On vient de voir qu'il est intégrable et le Lemme 4 nous assure qu'il est adapté. Pour montrer la propriété de martingale, on considère $0 \leq s \leq t$ et F une variable \mathcal{F}_s -mesurable. On pose $v = F1_{[s,t]}$. Comme $v \in \mathcal{E}$, on a $\langle v \cdot u \rangle_{L^2(Prog)} = 0$. Or

$$\langle v \cdot u \rangle_{L^2(Prog)} = \mathbb{E} \int_0^{+\infty} v_r u_r dr = \mathbb{E} \left(F \int_s^t u_r dr \right) = \mathbb{E} (F(M_t - M_s)).$$

On en déduit que $(M_t)_{t \geq 0}$ est une martingale.

De plus $(M_t)_{t \geq 0}$ est aussi un processus à variation finie d'après le Lemme 4. On en déduit, d'après le Lemme 5 que p.s. pour tout $t \geq 0$, $M_t = 0$. On a donc p.s. pour tout $t \geq 0$, $\int_0^t u_r dr = 0$ et on en déduit que $u_r = 0$ p.p. et p.s. Pour montrer ce dernier point on peut par exemple montrer que p.s. u est dans l'orthogonal de l'ensemble des fonctions en escalier qui est dense dans $L^2(\mathbb{R})$ (mais peut-être peut on trouver un argument plus court). \square

Nous avons maintenant tout ce qu'il nous faut pour définir l'intégrale stochastique.

Théorème 12 (Intégrale stochastique d'Itô). *Il existe une unique application linéaire*

$$I : L^2(Prog) \rightarrow L^2(\Omega, \mathcal{F}, \mathbb{P})$$

tel que

1. Pour tout processus $u = \sum_{i=0}^{n-1} F_i 1_{[t_i, t_{i+1}[}$ dans \mathcal{E} ,

$$I(u) = \sum_{i=0}^{n-1} F_i (B_{t_{i+1}} - B_{t_i}). \quad (5)$$

2. Pour tout $u \in L^2(Prog)$,

$$\|I(u)\|_{L^2(\mathcal{F})}^2 = \|u\|_{L^2(Prog)}^2$$

autrement dit

$$\mathbb{E}(I(u)^2) = \mathbb{E} \int u_s^2 ds$$

encore autrement dit, I est une **isométrie**.

Preuve. L'idée est d'utiliser le théorème de prolongement des isométries entre deux Hilbert. Comme on a déjà prouvé la densité de \mathcal{E} dans $L^2(Prog)$ à la

Proposition 18, il ne nous reste qu'à montrer que la relation (5) définit une isométrie de \mathcal{E} dans $L^2(\mathcal{F})$. On vérifie que

$$\begin{aligned} \|I(u)\|_{L^2(\mathcal{F})}^2 &= \mathbb{E}\left[\left(\sum_{i=0}^{n-1} F_i (B_{t_{i+1}} - B_{t_i})\right)^2\right] \\ &= \mathbb{E}\left[\sum_{i=0}^{n-1} F_i^2 (B_{t_{i+1}} - B_{t_i})^2 + 2 \sum_{0 \leq i < j \leq n-1} \mathbb{E}(F_i F_j (B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j}))\right] \\ &= A + B. \end{aligned}$$

Or

$$\begin{aligned} A &= \sum_{i=0}^{n-1} \mathbb{E}\left(\underbrace{F_i^2}_{\mathcal{F}_{t_i}\text{-mes.}} \underbrace{(B_{t_{i+1}} - B_{t_i})^2}_{\text{indep. de } \mathcal{F}_{t_i}}\right) \\ &= \sum_{i=0}^{n-1} \mathbb{E}(F_i^2)(t_{i+1} - t_i) \\ &= \mathbb{E} \int_0^t u_s^2 ds \\ &= \|u\|_{L^2(Prog)}^2. \end{aligned}$$

De plus B est nul car pour tout $i < j$,

$$\mathbb{E}\left(\underbrace{F_i F_j (B_{t_{i+1}} - B_{t_i})}_{\mathcal{F}_{t_j}\text{-mes.}} \underbrace{(B_{t_{j+1}} - B_{t_j})}_{\text{indep. de } \mathcal{F}_{t_j}}\right) = 0.$$

□

On note au passage des premières propriétés de l'intégrale stochastique. Comme I est une isométrie, pour tout $u, v \in L^2(Prog)$

$$\langle I(u), I(v) \rangle_{L^2(\mathcal{F})} = \langle u, v \rangle_{L^2(Prog)},$$

ce qui se réécrit

$$\mathbb{E}\left(\int u_s dB_s \int v_s dB_s\right) = \mathbb{E} \int u_s v_s ds.$$

Par ailleurs pour tout $u \in \mathcal{E}$,

$$\mathbb{E}(I(u)) = \sum_{i=0}^{n-1} \mathbb{E}(F_i (B_{t_{i+1}} - B_{t_i})).$$

Or pour tout i , F_i est \mathcal{F}_{t_i} -mesurable et comme $(B_t)_{t \geq 0}$ est un \mathcal{F} -mouvement brownien, on en déduit que $F_i \perp B_{t_{i+1}} - B_{t_i}$ et donc $E(I(u)) = 0$. On étend cette propriété à tout $u \in L^2(Prog)$ en approximant u par une suite $(u_n)_{n \geq 1}$ de processus prévisibles simples. On obtient alors que $(I(u_n))_{n \geq 1}$ converge dans $L^2(\mathcal{F})$ et donc dans $L^1(\mathcal{F})$ vers $I(u)$ ce qui implique la convergence des espérances. On a donc bien montré que pour tout $u \in L^2(Prog)$,

$$E \int u_t dB_t = 0.$$

On considère un processus $u \in L^2(Prog)$. On note que pour tout $t \geq 0$ le processus $u1_{[0,t]}$ est également dans $L^2(Prog)$ (il est \mathcal{P} -mesurable comme produit de variables mesurables et le fait qu'il soit de carré intégrable ne pose pas problème) et on peut donc définir son intégrale

$$X_t = \int_0^t u_s dB_s = \int u_s 1_{[0,t]}(s) dB_s.$$

Le choix d'une définition non anticipative des processus en escalier nous permet d'obtenir la propriété suivante

Proposition 19. *Soit $u \in L^2(Prog)$. Alors le processus*

$$X_t = \int_0^t u_s dB_s \quad t \geq 0,$$

est une martingale bornée dans L^2 continue (i.e. qui admet une modification continue). Pour tout $t \geq 0$,

$$\|X_t\|_2^2 = E \int_0^t u_s^2 ds.$$

On notera que cette proposition serait fautive si on n'avait pas fait le choix d'une définition non anticipative de l'intégrale stochastique.

Preuve. La ligne de la preuve pour montrer, et la continuité, et le caractère martingale, est de montrer d'abord ces propriétés quand $u \in \mathcal{E}$ puis de montrer qu'elles sont conservées par approximation.

Notons pour commencer que pour tout $u \in L^2(Prog)$ et tout $t \geq 0$,

$$\|X_t\|_2^2 = \|I(u1_{[0,t]})\|_2^2 = E \int |u_s 1_{[0,t]}(s)|^2 ds = E \int_0^t u_s^2 ds.$$

On en déduit que pour tout $t \geq 0$, X_t est dans L^2 (donc dans L^1) et même que le processus $(X_t)_{t \geq 0}$ est borné dans L^2 par $E \int u_s^2 ds$ qui est bien fini puisque $u \in L^2(Prog)$.

Montrons maintenant que $(X_t)_{t \geq 0}$ est une martingale. On suppose donc dans un premier temps que $u = \sum_{i=0}^{n-1} F_i 1_{[t_i, t_{i+1}[}$ est un processus de \mathcal{E} . On vérifie que pour tout $t \geq 0$,

$$X_t = \sum_{i=0}^{n-1} F_i (B_{t_{i+1} \wedge t} - B_{t_i \wedge t}) \quad (6)$$

et on en déduit que X_t est \mathcal{F}_t -mesurable et donc que $(X_t)_{t \geq 0}$ est adapté. Montrons maintenant la propriété de martingale et fixons pour cela $0 \leq s \leq t$. Si s et t sont dans le même intervalle de la subdivision, c'est-à-dire que pour un certain k , s et t sont dans $]t_k, t_{k+1}]$ alors

$$\int_0^t u_r dB_r - \int_0^s u_r dB_r = F_k(B_t - B_s).$$

On conclut facilement car F_k est \mathcal{F}_{t_k} -mesurable donc \mathcal{F}_s -mesurable donc $E(F_k(B_t - B_s) | \mathcal{F}_s) = F_k E(B_t - B_s | \mathcal{F}_s) = 0$. On passe maintenant au cas où s et t ne sont pas dans le même intervalle : $t_\ell < s \leq t_{\ell+1} \leq t_k < t \leq t_{k+1}$. On a alors

$$\int_0^t u_r dB_r - \int_0^s u_r dB_r = F_\ell(B_{t_{\ell+1}} - B_s) + \sum_{i=\ell+1}^{k-1} F_i(B_{t_{i+1}} - B_{t_i}) + F_k(B_t - B_{t_k}).$$

Lorsqu'on considère l'espérance conditionnelle par rapport à \mathcal{F}_s on peut traiter le premier terme comme dans le cas précédent, puis pour $i \in \{\ell + 1, \dots, k - 1\}$,

$$\begin{aligned} E(F_i(B_{t_{i+1}} - B_{t_i}) | \mathcal{F}_s) &= E[E[F_i(B_{t_{i+1}} - B_{t_i}) | \mathcal{F}_{t_i}] | \mathcal{F}_s] \\ &= E[F_i E[(B_{t_{i+1}} - B_{t_i}) | \mathcal{F}_{t_i}] | \mathcal{F}_s] \\ &= 0, \end{aligned}$$

où on a utilisé que F_i est \mathcal{F}_{t_i} -mesurable. On notera bien que ce calcul n'est plus possible si le processus que l'on intègre est anticipatif. On traite le dernier terme de façon similaire.

On ne suppose plus désormais que $u \in \mathcal{E}$ mais seulement qu'il est dans $L^2(Prog)$ et on considère une suite approximante $(u^n)_{n \geq 0}$ dans \mathcal{E} . Pour tout $t \geq 0$

$$\|u^n 1_{[0,t]} - u 1_{[0,t]}\|_{L^2(Prog)} \leq \|u^n - u\|_{L^2(Prog)}$$

donc $(u^n 1_{[0,t]})$ converge dans $L^2(Prog)$ vers $u 1_{[0,t]}$ et on en déduit que $(\int_0^t u_s^n dB_s)$ converge dans $L^2(\mathcal{F})$ vers $\int_0^t u_s dB_s$. Comme pour tout $n \geq 1$, $\int_0^t u_s^n dB_s$ est

\mathcal{F}_t -mesurable (car $u^n \in \mathcal{E}$) on en déduit que $\int_0^t u_s dB_s$ est également \mathcal{F}_t -mesurable. On a bien prouvé que $(X_t)_{t \geq 0}$ est adapté.

On passe à la propriété de martingale. On fixe $0 \leq s \leq t$. Pour tout $n \geq 1$, comme $u^n \in \mathcal{E}$,

$$\mathbb{E} \left(\int_0^t u_r^n dB_r | \mathcal{F}_s \right) = \int_0^s u_r^n dB_r.$$

On a déjà vu en utilisant que l'intégrale stochastique est une isométrie que $(\int_0^s u_r^n dB_r)$ converge dans $L^2(\mathcal{F})$ vers $\int_0^s u_r dB_r$. De plus comme l'espérance conditionnelle est un opérateur continu sur $L^2(\mathcal{F})$, on obtient que $(\mathbb{E}(\int_0^t u_r^n dB_r | \mathcal{F}_s))$ converge, toujours dans $L^2(\mathcal{F})$, vers $\mathbb{E}(\int_0^t u_r dB_r | \mathcal{F}_s)$. On a bien montré la propriété de martingale.

Il nous reste à montrer la continuité de $(X_t)_{t \geq 0}$. Là encore, lorsque $u \in \mathcal{E}$, il est facile de conclure en utilisant (6) car les trajectoires du brownien sont continues. Pour le cas général, on fixe $T > 0$ et on considère une suite approximante (u^n) . D'après l'inégalité de Doob (que l'on peut utiliser car on a bien la continuité de la martingale pour les processus de \mathcal{E}), pour tout $n, m \geq 0$,

$$\begin{aligned} \mathbb{P} \left(\sup_{0 \leq t \leq T} \left| \int_0^t u_s^n - u_s^m dB_s \right| > \varepsilon \right) &\leq \frac{\mathbb{E} \left(\left| \int_0^T u_s^n - u_s^m dB_s \right|^2 \right)}{\varepsilon^2} \\ &\stackrel{Iso.}{\leq} \frac{\mathbb{E} \left(\int_0^T (u_s^n - u_s^m)^2 ds \right)}{\varepsilon^2} \\ &= \frac{\|u^n - u^m\|_{L^2(Prog)}^2}{\varepsilon^2}. \end{aligned}$$

On peut donc extraire une sous-suite $(n_k)_{k \geq 1}$ telle que pour tout $k \geq 1$

$$\mathbb{P} \left(\sup_{0 \leq t \leq T} \left| \int_0^t u_s^{n_k} - u_s^{n_{k+1}} dB_s \right| > \frac{1}{2^k} \right) \leq \frac{1}{2^k}$$

et par le lemme de Borel Cantelli on obtient que p.s. pour k assez grand

$$\sup_{0 \leq t \leq T} \left| \int_0^t u_s^{n_k} - u_s^{n_{k+1}} dB_s \right| \leq \frac{1}{2^k}.$$

On en déduit que p.s. la suite de fonction $(t \rightarrow \int_0^t u_s^{n_k} dB_s)$ converge uniformément vers une limite qui est donc continue. Cette limite est bien $t \rightarrow \int_0^t u_s dB_s$ car pour tout $0 \leq t \leq T$, on a convergence L^2 de $(\int_0^t u_s^{n_k} dB_s)$ vers $\int_0^t u_s dB_s$. \square

On a bien avancé... mais on ne peut intégrer, pour l'instant, que des processus qui sont dans $L^2(Prog)$ donc qui satisfont, en plus de la condition de progressive mesurabilité, la condition $E \int u_t^2 dt < +\infty$. C'est en fait assez restrictif et on peut faire mieux à moindre coût.

Définition 20. On note L_{loc}^2 l'ensemble des processus progressivement mesurables tels que pour tout $t \geq 0$,

$$p.s. \int_0^t u_s^2 ds < +\infty.$$

On remarque que dans cette définition on pourrait échanger « p.s. » et « \forall » puisque $\int_0^t u_s^2 ds$ est croissant en t et qu'il suffit donc de vérifier que ces quantités sont finies pour t entier. Il est clair que $L_{loc}^2 \not\subseteq L^2(Prog)$ mais l'intérêt de considérer cet ensemble réside dans la **propriété de localisation** suivante :

Lemme 6. Soit $u \in L_{loc}^2$. Alors il existe une suite croissante de temps d'arrêt $(T_n)_{n \geq 1}$ telle que

1. p.s. $T_n \nearrow +\infty$ ($n \rightarrow +\infty$),
2. $E \int_0^{T_n} u_t^2 dt < +\infty$ pour tout $n \geq 1$.

Preuve. On considère pour tout $n \geq 1$,

$$T_n = \inf\{t \geq 0 \text{ tel que } \int_0^t u_s^2 ds \geq n\}. \quad (7)$$

On vérifie qu'il s'agit bien de temps d'arrêt car pour tout $r \geq 0$

$$\{T_n \leq r\} = \left\{ \int_0^r u_s^2 ds \geq n \right\},$$

et, en raisonnant comme au Lemme 4, on vérifie que $\int_0^r u_s^2 ds$ est \mathcal{F}_r -mesurable. Il est clair par ailleurs que T_n est croissant et que sa limite est $+\infty$ par définition même de L_{loc}^2 . \square

On en déduit que si $u \in L_{loc}^2$ alors pour tout $n \geq 1$, $u1_{[0, T_n]}$ est dans $L^2(Prog)$ (exercice : vérifier le caractère progressivement mesurable en montrant que pour tout temps d'arrêt T le processus $t \rightarrow 1_{[0, T]}(t)$ est progressivement mesurable) et on peut donc définir

$$\int_0^{T_n} u_s dB_s = \int u_s 1_{[0, T_n]}(s) dB_s$$

et pour tout $t \geq 0$,

$$M_t^{(n)} := \int_0^t u_s 1_{[0, T_n]}(s) dB_s = \int u_s 1_{[0, T_n]}(s) 1_{[0, t]}(s) dB_s.$$

Pour tous $m \geq n$ et $t \geq 0$, on vérifie que $M_{t \wedge T_n}^{(m)} = M_t^{(n)}$ ce qui implique que p.s. pour n assez grand (tel que $T_n \geq t$) $M_t^{(n)}$ est constante et cette limite est par définition

$$\int_0^t u_s dB_s = \lim_{n \rightarrow +\infty} \int_0^t u_s 1_{[0, T_n]}(s) dB_s \quad \text{p.s.}$$

Voici les premières propriétés importantes de cette intégrale stochastique définie sur L_{loc}^2 :

1. **Quand $u \in L^2(Prog)$, les deux définitions coïncident.** En effet, dans ce cas la suite de processus $(s \rightarrow u_s 1_{[0, t]}(s) 1_{[0, T_n]}(s))_{n \geq 0}$ converge dans $L^2(Prog)$ vers $s \rightarrow u_s 1_{[0, t]}(s)$ par théorème de convergence dominée et on en déduit que $(\int_0^t u_s 1_{[0, T_n]}(s) dB_s)_{n \geq 0}$ converge dans $L^2(\mathcal{F})$ vers $\int_0^t u_s dB_s$
2. **Le processus $(\int_0^t u_s dB_s)_{t \geq 0}$ est adapté et continu.** En effet pour tout $n \geq 1$, $\int_0^t u_s 1_{[0, T_n]}(s) dB_s$ est bien \mathcal{F}_t -mesurable comme intégrale d'un processus dans $L^2(Prog)$ et on conclut comme au point précédent. De plus, pour tout $0 \leq v \leq t$, p.s. dès que $T_n \geq t$, $\int_0^v u_s dB_s$ coïncide avec $\int_0^v u_s 1_{[0, T_n]}(s) dB_s$ ce qui donne la continuité de notre processus sur $[0, t]$.
3. Le processus $(\int_0^t u_s dB_s)_{t \geq 0}$ n'est pas forcément une martingale bornée dans L^2 ni même forcément une martingale. Par contre c'est une **martingale locale** :

Définition 21. *Un processus adapté et continu $(M_t)_{t \geq 0}$ est appelé martingale locale s'il existe une suite croissante de temps d'arrêt $(T_n)_{n \geq 1}$ telle que*

- (a) p.s. $T_n \nearrow +\infty$ ($n \rightarrow +\infty$),
- (b) pour tout $n \geq 1$, $(M_{t \wedge T_n})_{t \geq 0}$ est une martingale uniformément intégrable.

On dit d'une telle suite de temps d'arrêt qu'elle réduit $(M_t)_{t \geq 0}$.

On notera que cette définition n'implique pas que $(M_t)_{t \geq 0}$ soit intégrable.

Ainsi si $u \in L_{loc}^2$, en utilisant les $(T_n)_{n \geq 1}$ définis en (7), pour tout $n \geq 1$ et $t \geq 0$,

$$\int_0^{t \wedge T_n} u_s dB_s = \int_0^t 1_{[0, T_n]} u_s dB_s$$

est une martingale bornée dans L^2 donc est uniformément intégrable. On obtient bien que $(\int_0^t u_s dB_s)_{t \geq 0}$ est une martingale locale.

4. L'intégrale définie sur L_{loc}^2 n'est plus une isométrie et on a seulement l'inégalité

$$\mathbb{E} \left[\left(\int_0^t u_s dB_s \right)^2 \right] \leq \mathbb{E} \left(\int_0^t u_s^2 ds \right). \quad (8)$$

En effet si $\mathbb{E} \left(\int_0^t u_s^2 ds \right) < +\infty$, alors $u1_{[0,t]} \in L^2(Prog)$ et on a donc l'égalité. Et si $\mathbb{E} \left(\int_0^t u_s^2 ds \right) = +\infty$, il n'y a rien à prouver !

Maintenant que nous avons défini l'intégrale stochastique nous pouvons définir une classe importante de processus :

Définition 22 (Processus d'Itô). *Un processus $(X_t)_{t \geq 0}$ est un processus d'Itô s'il s'écrit*

$$p.s. \text{ pour tout } t \geq 0 \quad X_t = X_0 + \int_0^t v_s ds + \int_0^t u_s dB_s, \quad (9)$$

où

1. X_0 est \mathcal{F}_0 -mesurable ;
2. le processus v est progressivement mesurable et vérifie $\int_0^t |v_s| ds < +\infty$ pour tout $t \geq 0$,
3. le processus u appartient à L_{loc}^2 .

Pour un processus d'Itô u , on utilisera souvent la notation

$$dX_t = v_t dt + u_t dB_t,$$

qui ne signifie rien d'autre que (9).

5.3 Formule d'Itô

Commençons par un calcul peu rigoureux pour comprendre la nécessité de la formule d'Itô. On considère une fonction ϕ de classe \mathcal{C}^2 sur \mathbb{R} et x de

classe \mathcal{C}^1 sur $[0, t]$. On obtient alors le développement suivant

$$\begin{aligned}\phi(x(t)) &= \phi(x(0)) + \sum_{i=1}^n \left(\phi\left(x\left(\frac{it}{n}\right)\right) - \phi\left(x\left(\frac{(i-1)t}{n}\right)\right) \right) \\ &= \phi(x(0)) + \sum_{i=1}^n \phi' \left(x\left(\frac{(i-1)t}{n}\right) \right) \left(x\left(\frac{it}{n}\right) - x\left(\frac{(i-1)t}{n}\right) \right) \\ &\quad + \sum_{i=1}^n \frac{1}{2} \phi'' \left(x\left(\frac{(i-1)t}{n}\right) \right) \left(x\left(\frac{it}{n}\right) - x\left(\frac{(i-1)t}{n}\right) \right)^2 \\ &\quad + \text{Reste.}\end{aligned}$$

Quand on fait tendre n vers l'infini, la première somme converge vers $\int_0^t \phi'(x(s))x'(s)ds = \int_0^t \phi'(x(s))dx(s)$. La seconde converge vers 0 car la dérivée seconde est bornée et la variation quadratique de x est nulle. On obtient donc finalement la formule classique :

$$\phi(x(t)) = \phi(x(0)) + \int_0^t \phi'(x(s))dx(s).$$

Ce calcul devient faux si on prend pour x le mouvement brownien car la variation quadratique n'est plus nulle et le second terme ne peut plus être négligé. La formule d'Itô propose donc une version corrigée qui tient compte de ce terme :

Théorème 13 (Formule d'Itô). *Soit $\phi \in \mathcal{C}^2(\mathbb{R})$ telle que ϕ, ϕ' et ϕ'' sont bornées. Alors p.s., pour tout $t \geq 0$,*

$$\phi(B_t) = \phi(B_0) + \int_0^t \phi'(B_s)dB_s + \frac{1}{2} \int_0^t \phi''(B_s)ds.$$

Preuve. (cette présentation vient de [5]) On pose $t_i = \frac{ti}{n}$ pour $i = 0, \dots, n$ et on écrit le développement de Taylor à l'ordre 2 sur chaque pas de temps :

$$\phi(B_t) = \phi(B_0) + \sum_{i=1}^n \phi'(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}}) + \frac{1}{2} \sum_{i=1}^n \phi''(B_{\theta_i})(B_{t_i} - B_{t_{i-1}})^2,$$

où pour tout i , θ_i est un point adapté de $[t_{i-1}, t_i]$ (on a utilisé le théorème des valeurs intermédiaires).

Première étape. On interprète $\sum_{i=1}^n \phi'(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}})$ comme $\int_0^t X_s^n dB_s$ où $X_s^n = \sum_{i=1}^n \phi'(B_{t_{i-1}})1_{[t_{i-1}, t_i]}(s)$ pour $s \in [0, t]$. Montrons que $(X_s^n)_{s \geq 0}$

converge dans $L^2(Prog)$ vers $(\phi'(B_s))_{s \geq 0}$. On a en effet, en utilisant la continuité de ϕ' , convergence P-p.s. et s-p.p. de X_s^n vers $\phi'(B_s)$. Comme de plus ϕ' est bornée on en déduit par théorème de convergence dominée

$$\mathbb{E} \int [X_s^n - \phi'(B_s)]^2 ds \rightarrow 0.$$

En utilisant que l'intégrale d'Itô est une isométrie, on obtient

$$\sum_{i=1}^n \phi'(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}}) \xrightarrow{L^2} \int \phi'(B_s) dB_s.$$

Deuxième étape. Pour étudier la seconde somme

$$U_n = \sum_{i=1}^n \phi''(B_{\theta_i})(B_{t_i} - B_{t_{i-1}})^2,$$

on fait deux déplacements : on remplace θ_i par t_{i-1} pour définir

$$V_n = \sum_{i=1}^n \phi''(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}})^2,$$

puis $(B_{t_i} - B_{t_{i-1}})^2$ par $t_i - t_{i-1}$ et on définit

$$W_n = \sum_{i=1}^n \phi''(B_{t_{i-1}})(t_i - t_{i-1}).$$

Contrôlons dans un premier temps la distance L^1 entre U_n et V_n :

$$\begin{aligned} \mathbb{E}|U_n - V_n| &\leq \mathbb{E} \left(\sup_i |\phi''(B_{\theta_i}) - \phi''(B_{t_{i-1}})| \sum_{i=1}^n (B_{t_i} - B_{t_{i-1}})^2 \right) \\ &\stackrel{C.S.}{\leq} \mathbb{E}[\sup_i |\phi''(B_{\theta_i}) - \phi''(B_{t_{i-1}})|^2]^{1/2} \mathbb{E} \left(\left(\sum_{i=1}^n (B_{t_i} - B_{t_{i-1}})^2 \right)^2 \right)^{1/2} \end{aligned}$$

Pour le premier terme dans le produit : ce qui est dans l'espérance tend vers 0 p.s. par uniforme continuité et est dominé par 2 fois le supremum de ϕ'' . On en déduit que ce premier terme tend vers 0 par le théorème de convergence dominée. Pour le second terme du produit on se rappelle que $\sum_{i=1}^n (B_{t_i} - B_{t_{i-1}})^2$ tend vers t dans L^2 et on en déduit que ce second terme converge vers t . Finalement on a prouvé que

$$U_n - V_n \xrightarrow{L^1} 0.$$

Troisième étape. On contrôle maintenant la distance L^2 entre V_n et W_n :

$$\begin{aligned} \mathbb{E}(|V_n - W_n|^2) &= \mathbb{E} \left(\left| \sum_{i=1}^n \phi''(B_{t_{i-1}}) [(B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})] \right|^2 \right) \\ &= \sum_{i=1}^n \mathbb{E} \left(|\phi''(B_{t_{i-1}})|^2 [(B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})]^2 \right) \\ &\leq \|\phi''\|_\infty^2 \sum_{i=1}^n \mathbb{E} \left([(B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})]^2 \right) \end{aligned}$$

où la seconde égalité vient de ce que les espérances des termes croisés s'annulent. Pour $i = 1, \dots, n$, en notant X une variable aléatoire de loi $\mathcal{N}(0, t_i - t_{i-1})$,

$$\mathbb{E} \left([(B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})]^2 \right) = \mathbb{E}((X^2 - \mathbb{E}(X^2))^2) = \text{Var}(X^2) = (t_i - t_{i-1})^2 \text{Var}(U^2),$$

où U est une gaussienne centrée réduite. On en déduit finalement que $V_n - W_n$ tend vers 0 dans L^2 .

Quatrième étape. Étudions enfin la convergence de W_n : c'est une somme de Riemann donc p.s.

$$W_n \rightarrow \int_0^t \phi''(B_s) ds.$$

Comme de plus la suite (W_n) est dominée par $\|\phi''\|_\infty t$ on a également par le théorème de convergence dominée convergence dans L^p pour tout $p \geq 1$. Finalement on a bien montré que (U_n) converge dans L^1 vers $\int_0^t \phi''(B_s) ds$ et cela conclut la preuve pour t fixé. Pour échanger le p.s. et le $t \geq 0$ il suffit de se restreindre aux rationnels et d'utiliser la continuité. \square

Pour conclure cette partie nous donnons d'autres versions plus générales de la formule d'Itô. Nous n'allons pas les prouver cette année mais il est important de les connaître car elles sont très utiles en exercices.

Théorème 14 (Formule d'Itô pour un processus d'Itô). *Soit*

$$X_t = X_0 + \int_0^t u_s dB_s + \int_0^t v_s ds$$

un processus d'Itô avec u dans L_{loc}^2 et v progressivement mesurable tel que pour tout $t \geq 0$ p.s. $\int_0^t |v_s| ds < +\infty$. Soit ϕ une fonction \mathcal{C}^2 . Alors p.s. pour tout $t \geq 0$

$$\begin{aligned} \phi(X_t) &= \phi(X_0) + \int_0^t \phi'(X_s) dX_s + \frac{1}{2} \int_0^t \phi''(X_s) d\langle X \rangle_s \\ &= \phi(X_0) + \int_0^t \phi'(X_s) u_s dB_s + \int_0^t \phi'(X_s) v_s ds + \frac{1}{2} \int_0^t \phi''(X_s) u_s^2 ds. \end{aligned}$$

Théorème 15 (Fonction du temps et d'un processus d'Itô). *On considère un processus d'Itô $(X_t)_{t \geq 0}$ comme dans le théorème précédent et $\phi : \mathbb{R}^+ \times \mathbb{R} \rightarrow \mathbb{R}$ une fonction \mathcal{C}^2 . Alors p.s. pour tout $t \geq 0$*

$$\begin{aligned} \phi(t, X_t) &= \phi(0, X_0) + \int_0^t \partial_s \phi(X_s) ds + \int_0^t \partial_x \phi(X_s) dX_s + \frac{1}{2} \int_0^t \partial_{xx} \phi(X_s) d \langle X \rangle_s \\ &= \phi(0, X_0) + \int_0^t \partial_x \phi(X_s) u_s dB_s + \int_0^t \partial_s \phi(X_s) + \partial_x \phi(X_s) v_s + \frac{1}{2} \partial_{xx} \phi(X_s) u_s^2 ds. \end{aligned}$$

DRAFT

6 Équations différentielles stochastiques

Voir [7] pour une référence utilisée dans cette section. Le but de cette section est de donner un sens et résoudre l'équation différentielle stochastique (EDS) suivante

$$\begin{aligned}dX_t &= b(X_t)dt + \sigma(X_t)dB_t \\ X_0 &= Z\end{aligned}$$

où

- Z est une variable aléatoire appelée *condition initiale*,
- b est une fonction mesurable localement bornée de \mathbb{R} dans \mathbb{R} appelée *drift ou dérive*,
- σ est une fonction mesurable localement bornée de \mathbb{R} dans \mathbb{R} appelée *coefficient de diffusion*,
- $(B_t)_{t \geq 0}$ est un mouvement brownien.

On note (\mathcal{F}_t) la filtration engendrée par $(B_t)_{t \geq 0}$ (que l'on suppose complétée).

6.1 Existence et unicité d'une solution forte

Définition 23 (Solution forte). On appelle solution forte de notre équation différentielle stochastique un processus $(X_t)_{t \geq 0}$ à trajectoires continues tel que

1. $(X_t)_{t \geq 0}$ est adapté à $(\mathcal{F}_t)_{t \geq 0}$
2. p.s. pour tout $t \geq 0$,

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

La solution est dite *forte* au sens où elle est fonction de la condition initiale et du brownien qui sont des données du problème.

Théorème 16. On suppose que $Z \in L^2$ et que b et σ sont *lipschitziennes*. Alors l'équation différentielle stochastique

$$\begin{aligned}dX_t &= b(X_t)dt + \sigma(X_t)dB_t \\ X_0 &= Z\end{aligned}$$

admet une *unique solution forte*.

Preuve. Unicité. Soient $(X_t)_{t \geq 0}$ et $(Y_t)_{t \geq 0}$ deux solutions fortes. On fixe $M > 0$ et on note τ le temps d'atteinte de la distance M par $(X_t)_{t \geq 0}$ ou $(Y_t)_{t \geq 0}$:

$$\tau = \inf\{t \geq 0 \text{ tel que } |X_t| \geq M \text{ ou } |Y_t| \geq M\}.$$

On fixe $T > 0$ et on considère $0 \leq t \leq T$. On obtient

$$\begin{aligned}
\mathbb{E}(|X_{t \wedge \tau} - Y_{t \wedge \tau}|^2) &= \mathbb{E}(|\int_0^{t \wedge \tau} b(X_s) - b(Y_s) ds + \int_0^{t \wedge \tau} \sigma(X_s) - \sigma(Y_s) dB_s|^2) \\
&\leq 2 \left\{ \mathbb{E}(|\int_0^{t \wedge \tau} b(X_s) - b(Y_s) ds|^2) + \mathbb{E}(|\int_0^{t \wedge \tau} \sigma(X_s) - \sigma(Y_s) dB_s|^2) \right\} \\
&\stackrel{C.S.+^{(8)}}{\leq} 2 \left\{ \mathbb{E}(t \int_0^{t \wedge \tau} (b(X_s) - b(Y_s))^2 ds) + \mathbb{E}(\int_0^{t \wedge \tau} (\sigma(X_s) - \sigma(Y_s))^2 ds) \right\} \\
&\leq 2K^2(T+1) \int_0^t \mathbb{E}(|X_{s \wedge \tau} - Y_{s \wedge \tau}|^2) ds.
\end{aligned}$$

En posant $u(s) = \mathbb{E}(|X_{s \wedge \tau} - Y_{s \wedge \tau}|^2)$ pour $0 \leq s \leq T$ on obtient donc $u(0) = 0$ et pour tout $0 \leq t \leq T$,

$$u(t) \leq 2K^2(T+1) \int_0^t u(s) ds.$$

Nous allons pouvoir utiliser le

Lemme 7 (Lemme de Gronwall). *Soit u une fonction positive localement bornée sur \mathbb{R}^+ telle que pour tout $t \geq 0$*

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

où a et b sont des constantes positives. Alors pour tout $t \geq 0$

$$u(t) \leq ae^{bt}.$$

Preuve du Lemme de Gronwall. Par récurrence pour tout $n \geq 1$

$$u(t) \leq a + ab + \dots + a \frac{(bt)^n}{n!} + b^{n+1} \int_0^t \int_0^{t_1} \dots \int_0^{t_n} u(t_{n+1}) dt_{n+1} \dots dt_1$$

Comme u est bornée sur $[0, t]$ (disons par une constante C) le dernier terme est majoré par $C \frac{(bt)^{n+1}}{(n+1)!}$ et tend donc vers 0. D'où le résultat. \square

Dans notre cas $a = 0$ et u est bornée par $4M^2$ par définition de τ . On en déduit que $u(t)$ est nulle pour tout $t \geq 0$ et donc que p.s. $X_{t \wedge \tau} = Y_{t \wedge \tau}$. On fait tendre M vers l'infini pour conclure que X et Y sont modifications l'une de l'autre. Comme ces deux processus sont continus, on obtient également qu'ils sont indistinguables.

Existence. On montre l'existence par un procédé de Picard (i.e. un argument de point fixe). On définit donc le processus

$$X_t^{(0)} = Z \quad t \geq 0,$$

et pour tout $n \geq 0$,

$$X_t^{(n+1)} = Z + \int_0^t b(X_s^{(n)}) ds + \int_0^t \sigma(X_s^{(n)}) dB_s \quad t \geq 0. \quad (10)$$

On fixe $T \geq 0$. On va utiliser l'inégalité de Doob pour $(\int_0^t \sigma(X_s^{(n)}) dB_s)_{0 \leq t \leq T}$ et on doit donc vérifier qu'il s'agit d'une vraie martingale. Pour cela, on montre par récurrence (voir [7]) par un calcul similaire à celui que l'on vient de réaliser pour montrer l'unicité que pour tout $n \geq 0$,

$$\sup_{0 \leq t \leq T} \mathbb{E}((X_t^{(n)})^2) < +\infty.$$

On en déduit que $(X_t^{(n)})_{0 \leq t \leq T}$ est dans $L^2(Prog, [0, T])$ et, en utilisant que σ est lipschitzien, que $(\sigma(X_t^{(n)}))_{0 \leq t \leq T}$ également. Pour tout $n \geq 1$, le processus $(\int_0^t \sigma(X_s^{(n)}) dB_s)_{0 \leq t \leq T}$ est donc une vraie martingale bornée dans L^2 et on va pouvoir utiliser Doob. On obtient donc pour tout $t \leq T$,

$$\begin{aligned} \mathbb{E}(\sup_{0 \leq s \leq t} |X_s^{(n+1)} - X_s^{(n)}|^2) &\leq 2\mathbb{E} \left(\sup_{0 \leq s \leq t} \left| \int_0^s b(X_u^{(n)}) - b(X_u^{(n-1)}) du \right|^2 \right) \\ &\quad + 2\mathbb{E} \left(\sup_{0 \leq s \leq t} \left| \int_0^s \sigma(X_u^{(n)}) - \sigma(X_u^{(n-1)}) dB_u \right|^2 \right) \\ &\stackrel{Doob}{\leq} 2\mathbb{E} \left(\left(\int_0^t |b(X_u^{(n)}) - b(X_u^{(n-1)})| du \right)^2 \right) \\ &\quad + 8\mathbb{E} \left(\left| \int_0^t \sigma(X_u^{(n)}) - \sigma(X_u^{(n-1)}) dB_u \right|^2 \right) \\ &\stackrel{C.S.+(8)}{\leq} 2(4+T)K^2 \mathbb{E} \left(\int_0^t |X_u^{(n)} - X_u^{(n-1)}|^2 du \right) \\ &\leq C \int_0^t \mathbb{E}(\sup_{0 \leq v \leq u} |X_v^{(n)} - X_v^{(n-1)}|^2) du, \end{aligned} \quad (11)$$

en posant $C = 2(4+T)K^2$. Pour $n = 1$, on obtient donc

$$\mathbb{E}(\sup_{0 \leq s \leq t} |X_s^{(2)} - X_s^{(1)}|^2) \leq C \int_0^t \mathbb{E}(\sup_{0 \leq v \leq u} |X_v^{(1)} - X_v^{(0)}|^2) du \leq C a t,$$

en posant $a = \mathbb{E}(\sup_{t \leq T} |X_t^{(1)} - X_t^{(0)}|^2)$ (qui est bien fini car $Z \in L^2$). Par itération on en déduit que pour tout $n \geq 1$, et tout $0 \leq t \leq T$,

$$\left\| \sup_{0 \leq t \leq T} |X_t^{(n+1)} - X_t^{(n)}| \right\|_2^2 = \mathbb{E} \left(\sup_{0 \leq s \leq t} |X_s^{(n+1)} - X_s^{(n)}|^2 \right) \leq a C^n \frac{t^n}{n!},$$

ce qui implique

$$\sum_{n \geq 0} \left\| \sup_{0 \leq t \leq T} |X_t^{(n+1)} - X_t^{(n)}| \right\|_2 < +\infty.$$

On en déduit que

$$\sum_{n \geq 0} \sup_{0 \leq t \leq T} |X_t^{(n+1)} - X_t^{(n)}|$$

qui est la limite p.s. d'une suite croissante de variables positives et aussi limite au sens de L^2 . En particulier la limite est finie p.s. :

$$\sum_{n \geq 0} \sup_{0 \leq t \leq T} |X_t^{(n+1)} - X_t^{(n)}| < +\infty \quad p.s.$$

Cela implique que p.s. la suite des $(X_t^{(n)})_{0 \leq t \leq T}$ est de Cauchy dans l'espace des fonctions continues sur $[0, T]$ muni de la norme uniforme. Presque sûrement, cette suite converge donc uniformément vers une limite que l'on note $(X_t)_{0 \leq t \leq T}$. Les trajectoires de $(X_t)_{0 \leq t \leq T}$ sont continues comme limites uniformes de fonctions continues. De plus $(X_t)_{0 \leq t \leq T}$ est adapté à la filtration du brownien comme limite des $(X_t^{(n)})_{0 \leq t \leq T}$ qui sont adaptés à cette même filtration (on a même mieux que la convergence p.s. de $X_t^{(n)}$ vers X_t).

Il reste à montrer que ce processus est solution de notre EDS. Revenons pour cela à l'équation de récurrence (10). On remarque tout d'abord par un calcul similaire à (11) que

$$\mathbb{E} \left(\sup_{0 \leq t \leq T} |X_t^{(n)} - X_t|^2 \right) \xrightarrow{n \rightarrow +\infty} 0 \quad (12)$$

ce qui implique que pour tout $0 \leq t \leq T$, $X_t^{(n)}$ converge dans L^2 vers X_t . En utilisant que σ est lipschitzienne, on obtient

$$\mathbb{E} \int_0^T (\sigma(X_s^{(n)}) - \sigma(X_s))^2 ds \leq K^2 \mathbb{E} \int_0^T (X_s^{(n)} - X_s)^2 ds,$$

on en déduit, en utilisant (12), que pour tout $0 \leq t \leq T$, $(\int_0^t \sigma(X_s^{(n)}) dB_s)_{n \geq 0}$ converge dans L^2 vers $\int_0^t \sigma(X_s) dB_s$. De même

$$\mathbb{E} \left| \int_0^T b(X_s^{(n)}) ds - \int_0^T b(X_s) ds \right|^2 \leq K^2 T \mathbb{E} \left(\int_0^T |X_s^{(n)} - X_s|^2 ds \right)$$

qui tend donc également vers 0. On a donc convergence L^2 de $(\int_0^T b(X_s^{(n)})ds)_{n \geq 0}$ vers $\int_0^T b(X_s)ds$. On peut donc passer à la limite L^2 dans (10) et on obtient bien que $(X_t)_{0 \leq t \leq T}$ est solution forte de notre EDS. Comme T est arbitraire et qu'il y a unicité cela conclut la preuve. \square

On peut en fait prouver avec les mêmes outils le théorème suivant, plus fort, qui traite également le cas inhomogène en temps.

Théorème 17. *Soit $Z \in L^2$. On suppose que σ et b sont continues sur $\mathbb{R}^+ \times \mathbb{R}$ et lipschitziennes en la variable d'espace (i.e. pour tout $t \geq 0$ fixé, $b(t, \cdot)$ et $\sigma(t, \cdot)$ sont lipschitziennes). Alors il existe une unique solution forte à l'équation*

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

6.2 Équation de Langevin et processus d'Ornstein-Uhlenbeck

[Se reporter à [5] pour un exposé repris dans cette partie]. Paul Langevin en 1908 propose une équation pour rendre compte du déplacement d'une particule en suspension dans un fluide en partant des lois de la mécanique newtonienne. On suppose que la particule est soumise à une force de friction dans le fluide (proportionnelle à sa vitesse) et à des chocs dus à l'agitation thermique du liquide. Si on note $(x(t))_{t \geq 0}$ le déplacement d'une particule de masse 1 alors pour tout $t \geq 0$

$$x''(t) = -bx'(t) + F(t)$$

où b est le coefficient de friction et F la force extérieure qui s'exerce sur la particule. La vitesse $v = x'$ de la particule satisfait donc

$$v'(t) = -bv(t) + F(t).$$

En considérant que la force F représente des chocs indépendants on peut raisonnablement supposer que $\int_a^b F(t)dt = \sigma(B(b) - B(a))$ où $(B_t)_{t \geq 0}$ est un mouvement brownien et σ est un réel mesurant l'agitation du milieu. On obtient ainsi l'EDS suivante pour la vitesse

$$dX_t = -bX_t dt + \sigma dB_t$$

avec condition initiale $Z \in L^2$ indépendante de $(B_t)_{t \geq 0}$. Pour déterminer la solution de cette équation on applique la formule d'Itô à $(e^{bt} X_t)_{t \geq 0}$:

$$\begin{aligned} e^{bt} X_t &= Z + \int_0^t b e^{bs} X_s ds + \int_0^t e^{bs} dX_s \\ &= Z + \sigma \int_0^t e^{bs} dB_s. \end{aligned}$$

Le processus

$$X_t = Ze^{-bt} + \sigma \int_0^t e^{-b(t-s)} dB_s, \quad t \geq 0$$

est donc solution de l'équation de Langevin. Ce processus s'appelle processus d'Ornstein-Uhlenbeck. La partie intégrale est un processus de Wiener (voir TD 5) puisque le processus intégré est déterministe et il s'agit donc d'un processus gaussien. Pour tout $0 \leq r \leq t$

$$E(X_t) = E(Z)e^{-bt}$$

et

$$\begin{aligned} \text{Cov}(X_r, X_t) &= e^{-b(t+r)} \text{Var}(Z) + \sigma^2 \int_0^r e^{-b(t-s)} e^{-b(r-s)} ds \\ &= e^{-b(t+r)} \text{Var}(Z) + \frac{\sigma^2}{2b} (e^{-b(t-r)} - e^{-b(t+r)}). \end{aligned}$$

On obtient donc les asymptotiques suivantes : $E(X_t) \rightarrow 0$ quand t tend vers l'infini et pour tout $h \geq 0$

$$\lim_{r \rightarrow +\infty} \text{Cov}(X_r, X_{r+h}) = \frac{\sigma^2}{2b} e^{-bh}.$$

et en particulier $E(X_t^2)$ tend vers $\frac{\sigma^2}{2b}$ ce qui signifie que l'énergie cinétique moyenne converge. Le processus $(X_t)_{t \geq 0}$ converge en loi vers une gaussienne $\mathcal{N}(0, \frac{\sigma^2}{2b})$ puisque une limite de gaussiennes est une gaussienne. Cette probabilité est invariante au sens où si Z a pour loi $\mathcal{N}(0, \frac{\sigma^2}{2b})$ alors pour tout $t \geq 0$, X_t suit la même loi : en effet X_t est gaussien et il est facile de calculer explicitement sa moyenne (nulle) et sa variance ($\sigma^2/2b$). Plus généralement dans ce cas, le processus $(X_t)_{t \geq 0}$ est un processus gaussien centré de covariance

$$K_X(r, t) = \frac{\sigma^2}{2b} e^{-b(t-r)}, \quad 0 \leq r \leq t.$$

Ce processus est stationnaire : pour tout $t_0 \geq 0$ le processus $(X_{t_0+t})_{t \geq 0}$ a même loi que $(X_t)_{t \geq 0}$.

On revient pour finir à la position $(Y_t)_{t \geq 0}$ de la particule elle-même et non à sa vitesse, dans le cas où la vitesse est donnée par le processus gaussien ci-dessus, et où la particule part d'une position $x \in \mathbb{R}$ déterministe. On obtient alors pour tout $t \geq 0$

$$Y_t = x + \int_0^t X_s ds.$$

Le processus $(Y_t)_{t \geq 0}$ est gaussien (car X est gaussien et on peut voir l'intégrale comme une suite de sommes de Riemann) centré en x et de covariance définie pour tout $0 \leq r \leq t$ par

$$\begin{aligned} K_Y(r, t) &= \int_0^r \int_0^t \text{Cov}(X_{s_1}, X_{s_2}) ds_1 ds_2 \\ &= \frac{\sigma^2}{b^2} r - \frac{\sigma^2}{2b^3} (2 - 2e^{-bt} - 2e^{-br} + e^{-b(t-r)} + e^{-b(t+r)}). \end{aligned}$$

Si on fait tendre les coefficients de friction et d'agitation b et σ vers l'infini en même temps de telle sorte que le rapport σ/b converge vers une constante $\kappa \in]0, +\infty[$ alors

$$K_Y(r, t) \rightarrow \kappa^2 r.$$

Dans cette limite notre processus converge en loi vers $(x + \kappa B_t)_{t \geq 0}$ où B est un brownien puisque dans le cas gaussien il suffit de vérifier la convergence des moyennes et des fonctions de variance. On retrouve ainsi dans cette limite le processus proposé par Einstein pour décrire la particule en suspension dans un liquide.

Références

- [1] F. Baudoin. *Diffusion processes and stochastic calculus*. EMS Textbooks in Mathematics. European Mathematical Society (EMS), Zürich, 2014.
- [2] P. Billingsley. *Wiley Series in Probability and Statistics : Probability and Statistics*. John Wiley & Sons, Inc., New York, second edition, 1999.
- [3] P. Bougerol. <https://www.lpsm.paris/pageperso/bougerol/CalcSto15-16.pdf>.
- [4] D. Chafaï. <https://djalil.chafai.net/docs/M2/m2-stochastic-calculus-course-2020-2021.pdf>.
- [5] F. Comets and T. Meyre. *Calcul stochastique et modèles de diffusions : Cours et exercices corrigés*. Dunod, 2006. Coll. Sciences sup.
- [6] N. Guillin. <http://math.univ-lyon1.fr/guillin/livre.pdf>.
- [7] J.-F. Le Gall. *Mouvement brownien, martingales et calcul stochastique*, volume 71 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer, Heidelberg, 2013.
- [8] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.