
MODÉLISATION STOCHASTIQUE

Aspects numériques

Notes de cours

MARIE BILLAUD-FRIESS

↻ ATTENTION. Ce polycopié est régulièrement mis à jour et corrigé. Si vous découvrez des erreurs, n'hésitez pas à me les communiquer. Toutes les remarques ou questions permettant d'en améliorer le contenu peuvent être envoyées par mail à l'adresse ci-dessous.

Dernière mise à jour : 26 février 2025

Marie Billaud-Friess
marie.billaud_friess@centrale-med.fr

Centrale Méditerranée
Institut de Mathématiques de Marseille

TABLE DES MATIÈRES

1	Mouvement brownien	1
1.1	Généralités sur les processus stochastiques	1
1.1.1	Définitions	1
1.1.2	Mesurabilité	2
1.1.3	Loi de probabilité	2
1.2	Les processus gaussiens	2
1.2.1	Processus du second ordre	3
1.2.2	Loi normale et vecteurs gaussiens	3
1.2.3	Processus gaussien	4
1.3	Bref historique	4
1.3.1	Définitions	5
1.3.2	Variante du mouvement brownien	7
1.4	Simulation du mouvement brownien	8
1.4.1	Discrétisation en temps	8
1.4.2	Méthode de Cholesky	8
1.4.3	Procédure progressive	9
1.4.4	Décomposition de Karhunen-Loève	10
2	Éléments du calcul stochastique	13
2.1	Objectifs de ce chapitre	13
2.2	Filtration et mesurabilité	13
2.2.1	Intégrale stochastique et propriétés	15
2.3	Processus d'Itô, formule de Itô	17
2.3.1	Formule de Itô pour le mouvement brownien	17
2.3.2	Processus de Itô et formule de Itô pour un processus de Itô	18
3	Intégration numérique des EDS	21
3.1	Notations	21
3.1.1	Quelques exemples	21
3.2	Existence et unicité de la solution	22
3.3	Intégration numérique des EDS	24
3.3.1	Schémas d'approximation	25
3.3.2	Erreur d'approximation	26
3.3.3	Convergence forte	26
3.3.4	Convergence faible	27
3.3.5	Aspects pratiques	27
4	Représentation probabiliste des EDP (lien EDP-EDS)	29
4.1	Equation de la chaleur	29
4.1.1	Un résultat formel	29
4.1.2	Une première formule de Feynman-Kac	29
4.2	Cas général : formule de Feynman-Kac	31
4.2.1	Générateur infinitésimal	31
4.2.2	Equation rétrograde	31
4.3	Stratégie numérique	32
5	Notations et conventions	35
5.1	Abbréviations	35
5.2	Notations	35
5.3	Tribus et variables aléatoires	35

Chapitre 1

MOUVEMENT BROWNIEN

1.1 Généralités sur les processus stochastiques

Dans ce chapitre, on donne les définitions principales concernant les processus stochastiques nécessaires à la compréhension du mouvement brownien, pour plus de détails voir e.g. [1].

1.1.1 Définitions

Dans la suite on notera $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilités.

Définition 1.1.1 (Processus stochastique). *Soient $I \subset \mathbb{R}$ et (E, \mathcal{E}) un espace mesurable. Un **processus stochastique** à valeurs dans E est défini sur $I \times \Omega$ par*

$$\begin{aligned} X : I \times \Omega &\rightarrow E \\ (t, \omega) &\mapsto X(t, \omega) = X_t(\omega). \end{aligned}$$

Par la suite, comme pour les variables aléatoires (v.a.), la dépendance en ω sera omise.

On notera

- Un processus stochastique comme $\{X(t), t \in I\}$ ou encore $\{X_t\}_{t \in I}$ voire plus simplement X .
- Le paramètre t est appelé **temps**.
- L'espace E est appelé **l'espace d'états**. En particulier, si $(E, \mathcal{E}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$ on parle de processus à valeurs entières ou si $(E, \mathcal{E}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ on parlera de processus à valeurs réelles.
- I est un ensemble d'indices (intervalle, ou ensemble fini ou dénombrable). Si I est dénombrable (e.g. $I = \mathbb{N}$) on parle de processus à **temps discret** ou **suite aléatoire**. Si I est un intervalle de \mathbb{R} (i.e. $I = [0, T], T > 0$ ou $I = \mathbb{R}_+$) on parle de processus à **temps continu** ou **fonction aléatoire**.
- Etant donnés $s < t$ de I , la différence $X_t - X_s$ est appelée **accroissement** du processus. Le processus X est dit à **accroissements indépendents** si pour tout $n \in \mathbb{N}^*$ et tout $(t_1, \dots, t_n) \in I^n$ les variables aléatoires $X_{t_1} - X_0, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$ sont indépendantes.

Définition 1.1.2 (Trajectoire d'un processus stochastique). *Pour $\omega \in \Omega$ donné, la **trajectoire** d'un processus stochastique $\{X_t\}_{t \in I}$ est la fonction déterministe $X(\cdot, \omega)$ de I dans E donnée par $t \mapsto X(t, \omega)$ pour tout $t \in I$.*

Exemple 1.1.3.

1. (nombre naissances) *Le nombre de naissances sur une période $[0, T]$ est un processus stochastique N à temps continu défini sur $I = [0, T]$ et à valeurs entières $H = \mathbb{N}$. Le processus de Poisson permet de modéliser une telle situation en partant de $N_0 = 0$ alors $N_t \sim \mathcal{P}(\lambda t), \lambda > 0$.*
2. (marche en temps) *Une marche en temps X définie comme la somme des v.a. réelles $\{\xi_i\}_{i \geq 0}$ aux instants t_i de $\mathbb{T} = \{t_0, t_1, \dots, t_n, \dots\}$ est un processus à temps discret et à valeurs réelles $X_{t_i} = \sum_{j \leq i} \xi_j$, où $t_i \in \mathbb{T}$.*

Définition 1.1.4 (Processus stochastique à trajectoires continues). **Soit $I \subset \mathbb{R}$ un intervalle.** *Un processus stochastique $\{X_t\}_{t \in I}$ à temps continu est dit à **trajectoires continues** si la fonction $t \mapsto X(t, \omega)$ est continue pour tout $\omega \in \Omega$.*

REMARQUE. D'autres définitions de la continuité sont possibles, par exemple on ayant une continuité presque sûrement (i.e. p.p. $\omega \in \Omega$) ou encore continuité en moyenne quadratique (voir e.g. [4, §5.5]).

Exemple 1.1.5.

1. *Le processus de Poisson n'est pas à trajectoires continues (fonction en escalier).*

1.1.2 Mesurabilité

On pourra utiliser les deux définitions suivantes du processus stochastique.

Définition 1. Une variable aléatoire à valeurs dans E , pour $t \in I$ donné : $X_t : \omega \mapsto X(t, \omega)$.

Définition 2. Une variable aléatoire à valeurs dans E^I (applications de I dans E) : $X : \omega \mapsto (X(t, \omega))_{t \in I}$.
pour peu que les notions de mesurabilité soient aussi bien définies, en particulier pour E^I .

Définition 1.1.6 (Tribu produit). Pour $k \geq 1$, $t_1, \dots, t_k \in I$ et $B_1, \dots, B_k \in \mathcal{E}$, l'ensemble

$$C(t_1, \dots, t_k; B_1, \dots, B_k) = \{(x_t)_{t \in I} \in E^I, x_{t_i} \in B_i, i = 1, \dots, k\},$$

est appelé **cylindre**. La **tribu produit** définie sur E^I est notée $\mathcal{E}^{\otimes I}$, et correspond à la plus petite tribu sur E^I contenant toutes les cylindres.

Lemme 1.1.7 (Définitions équivalentes). Soit $\{X_t\}_{t \in I}$, une famille de fonctions (v.a.) $X_t : \Omega \rightarrow E$. La fonction de Ω dans E $X_t : \omega \mapsto X(t, \omega)$, pour tout $t \in I$ est mesurable pour la tribu \mathcal{E} pour \mathcal{E} **si et seulement si** la fonction de Ω dans E^I $X : \omega \mapsto (X(t, \omega))_{t \in I}$ est mesurable pour la tribu produit $\mathcal{E}^{\otimes I} := \bigotimes_{t \in I} \mathcal{E}$.

Le lemme précédent, nous assure donc que les **Définitions 1&2** du processus X sont équivalentes.

1.1.3 Loi de probabilité

A la lumière des définitions précédentes, on peut désormais (au même titre que pour les v.a.) se poser la question de la définition de la loi d'un processus aléatoire.

Définition 1.1.8 (Loi d'un processus aléatoire). On appelle **loi du processus aléatoire** X la mesure de probabilité \mathbb{P}_X définie sur $\mathcal{E}^{\otimes I}$ par

$$\mathbb{P}_X(A) := \mathbb{P}(X \in A), \quad A \in \mathcal{E}^{\otimes I}.$$

La loi de X est déterminée¹ par ses marginales fini-dimensionnelles i.e. la lois des vecteurs $(X_{t_1}, \dots, X_{t_k})$ avec $k \geq 1$ et $t_1, \dots, t_k \in I$.

Nous donnons dans la suite diverses définitions permettant de comparer deux processus en terme de loi et indépendance.

Définition 1.1.9. Soient $X = \{X_t\}_{t \in I}$ et $Y = \{Y_t\}_{t \in J}$ deux processus stochastiques définis sur un même espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$.

i) Supposons $I = J$, les processus X et Y sont **égaux en loi** (au sens de l'égalité en loi des lois marginales fini-dimensionnelles [1]), si pour tout entier $k \in \mathbb{N}^*$ et tout $(t_1, \dots, t_k) \in I^k$, les vecteurs aléatoires $(X_{t_1}, \dots, X_{t_k})$ et $(Y_{t_1}, \dots, Y_{t_k})$ ont la même loi dans E^k . Plus précisément, pour tout $B \in \underbrace{\mathcal{E} \otimes \dots \otimes \mathcal{E}}_{k \text{ times}}$

on a

$$\mathbb{P}((X_{t_1}, \dots, X_{t_k}) \in B) = \mathbb{P}((Y_{t_1}, \dots, Y_{t_k}) \in B).$$

ii) Les processus X et Y sont **indépendants** si et seulement si $(X_{t_1}, \dots, X_{t_n})$ et $(Y_{s_1}, \dots, Y_{s_m})$ sont indépendants pour tout n, t_1, \dots, t_n et m, s_1, \dots, s_m . i.e. tous les vecteurs de dimension finis extraits de X et Y sont indépendants².

1.2 Les processus gaussiens

Avant de donner la définition des processus gaussiens, nous donnons quelques définitions générales sur les processus du second ordre.

1. En effet, cela est intimement lié à la définition de la tribu produit. Voir [1, §1,2] pour plus de détails.

2. Cela signifie en d'autres termes que les tribus générées par les processus $\sigma(X_t, t \in I)$ et $\sigma(Y_t, t \in J)$ sont indépendantes [1, §1].

1.2.1 Processus du second ordre

Définition 1.2.1 (Processus du second ordre). *Un processus stochastique est dit d'ordre 2 ou du second ordre si pour tout $t \in I$ on a X_t dans $L^2(\Omega)$ ³. On désigne par*

- $m : I \rightarrow \mathbb{R}$ la fonction **moyenne** donnée par $m(t) = \mathbb{E}(X_t)$,
- $C : I \times I \rightarrow \mathbb{R}$ la fonction **covariance** donnée par $C(s, t) = \text{cov}(X_t, X_s) = \mathbb{E}(X_t X_s) - \mathbb{E}(X_t)\mathbb{E}(X_s)$.

Définition 1.2.2 (Fonction de type positif). *Une fonction $C : I \times I \rightarrow \mathbb{R}$ est dite de type positif (ou positive) si pour tout $n \in \mathbb{N}^*$*

$$\forall (x_1, \dots, x_n)^T \in \mathbb{R}^n \text{ et } \forall (t_1, \dots, t_n)^T \in I^n$$

on a

$$\sum_{i=1}^n \sum_{j=1}^n x_i x_j C(t_i, t_j) \geq 0.$$

Exemple 1.2.3. *Soit X un processus de second ordre, alors*

- C est symétrique car $C(s, t) = \text{cov}(X_t, X_s) = \text{cov}(X_s, X_t) = C(t, s)$,
- C est de type positif en effet pour $(x_1, \dots, x_n)^T \in \mathbb{R}^n$ et $(t_1, \dots, t_n)^T \in I^n$ on a

$$\begin{aligned} \sum_{i=1}^n \sum_{j=1}^n x_i x_j C(t_i, t_j) &= \sum_{i=1}^n \sum_{j=1}^n x_i x_j \text{cov}(X_{t_i}, X_{t_j}) = \mathbb{E} \left(\sum_{i=1}^n \sum_{j=1}^n (x_i X_{t_i})(x_j X_{t_j}) \right) \\ &- \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}(x_i X_{t_i}) \mathbb{E}(x_j X_{t_j}) = \mathbb{E} \left(\left(\sum_{i=1}^n x_i X_{t_i} \right)^2 \right) - \mathbb{E} \left(\sum_{i=1}^n x_i X_{t_i} \right)^2 = \mathbb{V} \left(\sum_{i=1}^n x_i X_{t_i} \right) \geq 0. \end{aligned}$$

1.2.2 Loi normale et vecteurs gaussiens

Dans cette partie nous discutons un exemple qui aura une importance capitale par la suite : les *processus gaussiens*. On rappelle pour commencer, quelques propriétés sur les variables et vecteurs gaussiens.

Une variable aléatoire $X \in \mathbb{R}$ est dite **gaussienne** ou suit **une loi normale**, si il existe $\mu \in \mathbb{R}$ et $\sigma \neq 0$ tels que sa densité de probabilité $p : \mathbb{R} \rightarrow \mathbb{R}$ est donnée par

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left(-\frac{(x - \mu)^2}{2\sigma^2} \right) \quad x \in \mathbb{R},$$

et sa fonction caractéristique $\varphi : \mathbb{R} \rightarrow \mathbb{C}$ est définie par

$$\varphi(t) = \mathbb{E}(\exp(itX)) = \exp(i\mu t) \exp \left(-\frac{\sigma^2 t^2}{2} \right) \quad t \in \mathbb{R}.$$

On notera $X \sim \mathcal{N}(\mu, \sigma^2)$. De plus l'espérance est $\mathbb{E}(X) = \mu$ et la variance $\mathbb{V}(X) = \sigma^2$. Plus généralement, $X \in L^p(\Omega)$, $p \geq 1$. Les moments centrés sont nuls pour $q = 2k + 1$ et pour $q = 2k$ on a $\mathbb{E}(|X - \mu|^q) = \frac{(2k)!}{2^k k!} \sigma^{2k}$, $k \geq 1$.

Proposition 1.2.4. *La v.a. X est la translation et la dilation d'une gaussienne standard $G \sim \mathcal{N}(0, 1)$ t.q. $X = \mu + \sigma G$ pour tout $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}^*$.*

Démonstration. Voir ci-dessous pour le cas général. ■

On considère désormais, des variables aléatoires (vecteurs aléatoires) X à valeurs dans \mathbb{R}^n , $n \in \mathbb{N}^*$, équipé du produit scalaire canonique noté $\langle \cdot, \cdot \rangle$ de norme associée $\| \cdot \|$.

1. Un vecteur X est dit **gaussien** ou suit **une loi normale multivariée** si toute combinaison linéaire de ses composantes $\langle X, \lambda \rangle = \sum_{i=1}^n \lambda_i X_i$ (avec $\lambda = (\lambda_1, \dots, \lambda_n)^T \in \mathbb{R}^n$) est une variable aléatoire gaussienne.

³ L'espace de Lebesgue $L^p(\Omega)$ des variables aléatoires réelles de moment d'ordre $p \geq 1$ fini pour la mesure de probabilité \mathbb{P} est noté $L^p(\Omega, \mathcal{A}, \mathbb{P}) = L^2(\Omega) := \{X; \int_{\Omega} \|X(\omega)\|^p d\mathbb{P}(\omega) < \infty\}$

2. Un vecteur gaussien est caractérisé par $\mu = (\mu_1, \dots, \mu_n)^T \in \mathbb{R}^n$ et la matrice $\Sigma \in \mathbb{R}^{n \times n}$ symétrique et définie positive. On notera $X \sim \mathcal{N}(\mu, \Sigma)$ et sa densité de probabilité $p : \mathbb{R}^n \rightarrow \mathbb{R}$ est donnée par

$$p(x) = \frac{1}{(2\pi)^{n/2} \sqrt{\det \Sigma}} \exp\left(-\frac{\langle x - \mu, \Sigma^{-1}(x - \mu) \rangle}{2}\right), \quad x \in \mathbb{R}^n.$$

Sa fonction caractéristique $\varphi : \mathbb{R}^n \rightarrow \mathbb{C}$ est définie par

$$\varphi(t) = \mathbb{E}(\exp(i\langle t, X \rangle)) = \exp(i\langle t, \mu \rangle) \exp\left(-\frac{1}{2}\langle t, \Sigma t \rangle\right), \quad t \in \mathbb{R}^n,$$

avec $\mu \in \mathbb{R}^n$ et $\Sigma \in \mathbb{R}^{n \times n}$. On notera $X \sim \mathcal{N}(\mu, \Sigma)$. La matrice Σ est la **matrice de covariance** du vecteur gaussien t.q. $\Sigma = (\text{cov}(X_i, X_j))_{1 \leq i, j \leq n}$ et $\mu = \mathbb{E}(X) = (\mathbb{E}(X_1), \dots, \mathbb{E}(X_n))^T$ est l'**espérance** de X . Si $\mathbb{E}(X) = 0$, le vecteur X est dit centré.

De façon générale un vecteur gaussien s'obtient par le biais d'une transformation affine d'un vecteur gaussien dont les composantes sont des variables aléatoires gaussiennes centrées réduites i.e. de loi $\mathcal{N}(0, I_n)$.

Proposition 1.2.5. Soient $k, n \in \mathbb{N}^*$ et X un vecteur gaussien de loi $\mathcal{N}(0, I_k)$. Etant donnés $\mu \in \mathbb{R}^n$ et $L \in \mathbb{R}^{n \times k}$. Alors $\mu + LX \sim \mathcal{N}(\mu, LL^T)$.

Démonstration. On calcule la fonction caractéristique de $Y = m + LX$. Pour tout $t \in \mathbb{R}^n$ on a

$$\begin{aligned} \varphi_Y(t) &= \mathbb{E}(\exp(i\langle t, Y \rangle)) = \mathbb{E}(\exp(i\langle t, m + LX \rangle)) = \exp(i\langle t, m \rangle) \mathbb{E}(\exp(i\langle t, LX \rangle)), \\ &= \exp(i\langle t, m \rangle) \underbrace{\mathbb{E}(\exp(i\langle L^T t, X \rangle))}_{\varphi_X(L^T t)} = \exp(i\langle t, m \rangle) \exp\left(-\frac{1}{2}\langle L^T t, L^T t \rangle\right), \\ &= \exp(i\langle t, m \rangle) \exp\left(-\frac{1}{2}\langle t, LL^T t \rangle\right). \end{aligned}$$

donc $Y \sim \mathcal{N}(m, LL^T)$. ■

1.2.3 Processus gaussien

Définition 1.2.6. (Processus gaussien) Un processus stochastique $X = \{X_t\}_{t \in I}$ à valeurs réelles de second ordre est appelé **processus gaussien** si $(X_{t_1}, \dots, X_{t_n})$ est gaussien, pour tout $(t_1, \dots, t_n) \in I^n$, $n \in \mathbb{N}^*$.

La loi d'un processus gaussien $\{X_t\}_{t \in I}$ est complètement caractérisée par sa fonction moyenne $m : I \rightarrow \mathbb{R}$ et sa fonction covariance $C : I \times I \rightarrow \mathbb{R}$. En particulier, on a que la matrice de covariance vérifie $\Sigma = (C(t_i, t_j))_{1 \leq i, j \leq n}$ et elle définie positive car c'est un processus du second ordre!⁴

Exemple 1.2.7. Exemple de processus Gaussiens du second ordre centrés $m(t) = 0$.

- Processus d'Orstein-Uhlenbeck : $C(s, t) = \exp(-\frac{|t-s|}{2})$.
- Bruit blanc : $C(s, t) = \delta_s(t)$.
- Mouvement brownien ou processus de Wiener : $C(s, t) = \min(s, t)$.

Le mouvement brownien possède d'autres propriétés : comme celle de Markov qui assure qu'il est aveugle de ce qui se passe avant t_0 (et est même indépendant) ou d'inversion en temps (cf. [2]), mais surtout des propriétés de Martingale utiles pour la construction de l'intégrale stochastique (voir cours A. Falcó).

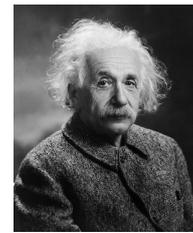
1.3 Bref historique

Dans ce paragraphe nous donnons un bref historique de l'apparition du **mouvement brownien** (aussi appelé **processus de Wiener**) et de la théorie du calcul stochastique qui en a découlé.

- **En 1827** : le naturaliste botaniste écossais *Robert Brown* (voir figure 1.3.1) décrit le comportement erratique de particules en suspension dans un gaz ou un fluide par un **processus stochastique** appelé le **mouvement brownien**.

4. C'est une conséquence de l'égalité en loi.

- **En 1901** : le mathématicien français *Louis Bachelier* (voir figure 1.3.1) introduit dans sa thèse intitulée : « Théorie de la spéculation » le premier modèle mathématique du mouvement brownien appliqué à des modèles financiers.
- **En 1905** : *Albert Einstein* (voir figure 1.3.1) utilise le mouvement brownien pour décrire le mouvement d'une particule physique. Il montre que la densité de probabilité de la particule d'être à une position donnée à un temps donné est une densité gaussienne (i.e. la densité du m.b.s. au temps t), solution fondamentale de l'équation de la chaleur.
- **En 1923** : l'américain *Norbert Wiener* (voir figure 1.3.1) donne une première définition mathématique rigoureuse du mouvement brownien en construisant une mesure de probabilité sur l'espace des fonctions (continues) réelles. En d'autres termes, il définit le mouvement brownien comme un processus aléatoire (à trajectoires continues), c'est à dire une famille $\{W_t; t \in \mathbb{R}_+\}$ de variables aléatoires sur un espace de probabilités $(\Omega, \mathcal{A}, \mathbb{P})$ (telle que l'application $t \mapsto W_t(\omega)$ est continue pour tout $\omega \in \Omega$). C'est pour cela qu'on appelle aussi le mouvement brownien : *processus de Wiener*. Wiener s'est aussi intéressé à la régularité des trajectoires $t \mapsto W_t$ de ce processus. Les trajectoires du Brownien sont continues mais non dérivables ! Wiener définit également l'intégrale par rapport au mouvement brownien qu'on appellera plus tard l'intégrale de Wiener. C'est le point de vue de Wiener que nous allons adopter dans la suite de ce cours.
- **Dans les années 50** : les travaux du mathématicien japonais *Kiyoshi Itô* (voir figure 1.3.1) permettent d'étendre les outils du calcul différentiel au mouvement brownien, c'est l'apparition du calcul stochastique (intégrale d'Itô, formule d'Itô...). En parallèle le mathématicien français *Paul Lévy* (voir figure 1.3.1) met aussi en évidence de nombreuses propriétés sur la trajectoire du mouvement Brownien.



Robert Brown (1773 - 1858) Louis Bachelier (1870 - 1946) Albert Einstein (1879 - 1955)



Norbert Wiener (1879 - 1964) Paul Lévy (1886 - 1971) Kiyoshi Itô (1915 - 2008)

FIGURE 1.3.1 – Chercheurs ayant contribué à la construction du mouvement brownien. (Photos from Wikimedia Commons, the free media repository.)

1.3.1 Définitions

Dans cette partie on considère mouvement brownien en dimension 1 (i.e. $d = 1$).

Définition 1.3.1 (Mouvement brownien standard en dimension 1). *Un mouvement brownien standard (m.b.s.) ou processus de Wiener est un processus stochastique, à temps continu et à trajectoires continues, noté W ou $\{W_t\}_{t \in \mathbb{R}_+}$ à valeurs réelles TEL QUE*

1. $W_0 = 0$,
2. tout accroissement $W_t - W_s$, $0 \leq s < t$, suit une loi gaussienne centrée de variance $(t - s)$,
3. pour tout $0 = t_0 < t_1 < t_2 < \dots < t_n$ les accroissements $\{W_{t_{i+1}} - W_{t_i}, 0 \leq i \leq n-1\}$ sont indépendants.

A l'instant t , W_t est une variable aléatoire gaussienne de moyenne $\mathbb{E}(W_t) = 0$ et de variance $\mathbb{E}(W_t^2) = t$ qui croît avec le temps. Ainsi on peut vérifier que W_t est une v.a. comprise dans l'intervalle $[-1.96\sqrt{t}, 1.96\sqrt{t}]$ avec une probabilité de 95% voir Figure 1.3.2. En effet comme $\frac{W_t}{\sqrt{t}} \sim \mathcal{N}(0, 1)$ alors $\mathbb{P}\left(\left|\frac{W_t}{\sqrt{t}}\right| \leq 1.96\right) = 0.95 \Leftrightarrow \mathbb{P}(-1.96\sqrt{t} \leq W_t \leq 1.96\sqrt{t}) = 0.95$.

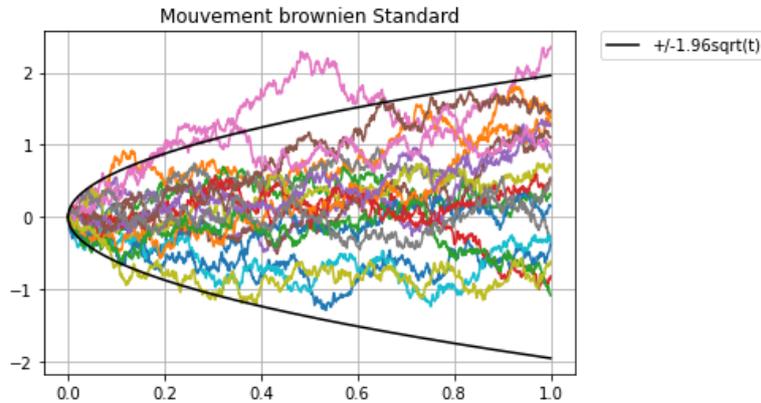


FIGURE 1.3.2 – Trajectoires du mouvement brownien.

REMARQUE.

1. Les trajectoires de W ne sont pas dérivables p.s. (voir TD).
2. Les trajectoires de W ne sont pas monotones p.s..

Théorème 1.3.2. *Le mouvement brownien existe.*

On peut considérer une preuve constructive de l'existence du mouvement brownien à la Wiener. Cette approche que l'on utilise aussi numériquement s'appuyant sur des séries de Fourier ou décomposition de Karhunen-Loève. Notons qu'il existe d'autres constructions : due à Kolmogorov, ou encore Donsker en interprétant le m.b.s. comme la limite d'une marche aléatoire (cf. TP).

On donne désormais une définition équivalente du m.b.s. .

Proposition 1.3.3 (Définition alternative du mouvement brownien). *Un mouvement brownien standard (m.b.s.) ou processus de Wiener noté W est un processus gaussien centré d'espérance $m(t) = \mathbb{E}(W_t) = 0$ et de fonction de covariance $C(s, t) = \text{cov}(W_t, W_s) = \min(s, t)$, $s, t \in \mathbb{R}_+$.*

Démonstration.

1. Comme $W_t \sim \mathcal{N}(0, t)$ alors $\mathbb{E}(W_t) = 0$ et $C(t, s) = \text{cov}(W_t, W_s) = \mathbb{E}(W_t W_s)$ ainsi
 - si $t = s$, $C(t, t) = \mathbb{E}(W_t^2) = t = \min(s, t)$, en remarquant que $\mathbb{E}(W_t^2) = \mathbb{V}(W_t) + \mathbb{E}(W_t)^2 = t$
 - si $t \neq s$ avec $s < t$ on a

$$\begin{aligned} \mathbb{E}(W_t W_s) &= \mathbb{E}((W_t - W_s)W_s + W_s^2) = \mathbb{E}((W_t - W_s)(W_s - W_0)) + s, \\ &= \underbrace{\mathbb{E}(W_t - W_s)\mathbb{E}(W_s - W_0)}_{=0} + s = s = \min(s, t). \end{aligned}$$

car $W_t - W_s, W_s - W_0$ indépendants et centrés. On remarque d'ailleurs que par définition la fonction covariance est bien symétrique. On a un processus de second ordre.

2. Introduisons la séquence des accroissements $\{W_{t_i} - W_{t_{i-1}}, 1 \leq i \leq n\}$. On va montrer que $(W_{t_1}, \dots, W_{t_n})^T$ est gaussien.

Par définition du mouvement brownien on a $(W_{t_1} - W_{t_0}, \dots, W_{t_n} - W_{t_{n-1}})^T$ est un vecteur gaussien, dont les composantes sont indépendantes ainsi $W_{t_i} - W_{t_{i-1}} \sim \mathcal{N}(0, t_i - t_{i-1})$ et $(W_{t_1} - W_{t_0}, \dots, W_{t_n} - W_{t_{n-1}})^T \sim \mathcal{N}(0, \Sigma)$ avec $\Sigma = \text{diag}(t_1 - t_0, \dots, t_n - t_{n-1}) \in \mathbb{R}^{n \times n}$.

Notons T la matrice de $\mathbb{R}^{n \times n}$ triangulaire définie par

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ \dots & & & \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

On a $(W_{t_1}, \dots, W_{t_n})^T = T(W_{t_1} - W_{t_0}, \dots, W_{t_n} - W_{t_{n-1}})^T$.

Ainsi par lemme du cours $(W_{t_1}, \dots, W_{t_n})^T \sim \mathcal{N}(0, TDT^T)$. Ce qui prouve que W est gaussien et il est caractérisé par ses fonctions moyenne et covariance. ■

Nous énonçons désormais quelques propriétés des trajectoires du mouvement brownien.

Proposition 1.3.4. *Soit W un m.b.s., le processus X défini comme il suit est aussi un m.b.s. .*

- i) (Symétrie) $X = -W$,
- ii) (Propriété d'échelle) $X_t = \frac{1}{c}W_{c^2t}$, pour $c \neq 0, t \geq 0$,
- iii) (Retournement du temps) $X_t = W_T - W_{T-t}$ pour $t \in [0, T], T > 0$.

Démonstration.

Par construction, tous les processus considérés sont des processus gaussiens (on effectue uniquement des transformation du temps et des combinaisons linéaires).

Il reste donc à vérifier que leur fonction de covariance vérifie $C(s, t) = \min(s, t)$ pour tout $s, t \in I$.

- i) On a directement $C(s, t) = \text{Cov}(X_t, X_s) = \text{Cov}(-W_t, -W_s) = \min(s, t)$.
- ii) Comme $C(s, t) = \mathbb{E}(W_t^c W_s^c) = \frac{1}{c^2} \mathbb{E}(W_{c^2t} W_{c^2s}) = \frac{1}{c^2} \min(c^2t, c^2s) = \min(t, s)$.
- iii) On a

$$\begin{aligned} C(s, t) &= \mathbb{E}(X_t X_s) = \mathbb{E}((W_T - W_{T-t})(W_T - W_{T-s})) = \mathbb{E}(W_T^2) - \mathbb{E}(W_{T-t} W_T) \\ &\quad - \mathbb{E}(W_T W_{T-s}) + \mathbb{E}(W_{T-s} W_{T-t}) = T - (T-s) - (T-t) + \min(T-s, T-t) = t + s - \max(s, t) = \min(s, t), \end{aligned}$$

en utilisant que $C(s, t) = \text{cov}(W_s, W_t) = \mathbb{E}(W_s W_t)$ (car W_s de moyenne nulle). ■

Le mouvement brownien possède d'autres propriétés : comme celle de Markov qui assure qu'il est aveugle de ce qui se passe avant t_0 (et est même indépendant) ou d'inversion en temps, mais surtout des propriétés de Martingale utiles pour la construction de l'intégrale stochastique (voir e.g. [1, 2]).

1.3.2 Variantes du mouvement brownien

Définition 1.3.5 (Mouvement brownien d-dimensionnel standard). *Un mouvement brownien d-dimensionnel standard est un processus noté $\{W_t\}_{t \in \mathbb{R}_+}$ à valeurs dans \mathbb{R}^d en temps continu et à trajectoires continues t.q. $W_t = ((W_1)_t, \dots, (W_d)_t)^T$ où les $\{(W_i)_t\}_{t \in \mathbb{R}_+}$ sont des m.b.s. indépendants pour tout $1 \leq i \leq d$.*

Définition 1.3.6 (Mouvement brownien ayant une valeur initiale non nulle). *Un mouvement brownien $\{X_t\}_{t \in \mathbb{R}_+}$ issu de X_0 est un processus de la forme $X_t = X_0 + W_t$ avec X_0 indépendant de W_t .*

On peut aussi considérer $X_0 = x$ une C.I. déterministe.

Définition 1.3.7 (Mouvement brownien avec dérive ou arithmétique). *Un mouvement brownien $\{X_t\}_{t \in \mathbb{R}_+}$ issu de X_0 , de dérive b et de coefficient de diffusion σ est un processus de la forme $X_t = X_0 + bt + \sigma W_t$ avec X_0 indépendant de W_t .*

REMARQUE. Dans la dernière définition, on appelle aussi $\{X_t\}_{t \in \mathbb{R}_+}$ **mouvement brownien arithmétique**. C'est un processus à accroissements indépendant, stationnaires et gaussiens mais non centrées en effet $\mathbb{E}(X_t) = \mathbb{E}(X_0) + bt$.

1.4 Simulation du mouvement brownien

L'objectif de cette partie est de simuler une réalisation d'un mouvement brownien standard.

1.4.1 Discrétisation en temps

Considérons un m.b.s. $W = \{W_t\}_{t \geq 0}$, on veut simuler W sur l'intervalle de temps $I = [0, T]$. Pour cela, on se donne un maillage uniforme contenant $N + 1$ points $0 = t_0 < t_1 < t_2 < \dots < t_N$ où $t_i = i\Delta t$ et $\Delta t = \frac{T}{N}$. On simule une réalisation de W aux instants t_i de façon à obtenir le vecteur $W^N = (W_{t_1}, \dots, W_{t_N})^T$ (rappel $W_{t_0} = 0$). On verra comment calculer W^N par la suite par différentes stratégies.

Afin de construire une approximation "globale" $W^{\Delta t}$ de W pour tout $t \in [0, T]$ on peut par exemple construire un interpolant aux points $\{W_{t_1}, \dots, W_{t_N}\}$.

APPROXIMATION \mathbb{P}_0 PAR MORCEAUX

Le processus approché $\{W_t^{\Delta t}\}_{t \in [0, T]}$ est **constant par morceaux** et défini par

$$W_s^{\Delta t} = W_{t_i}, \quad s \in [t_i, t_{i+1}), i \in \{0, \dots, N-1\},$$

de cette façon on a $W_{t_i}^{\Delta t} = W_{t_i}$ et $W^{\Delta t}$ est seulement **continu** à droite.

APPROXIMATION \mathbb{P}_1 PAR MORCEAUX

Le processus approché $\{W_t^{\Delta t}\}_{t \in [0, T]}$ est **continu et affine par morceaux** et défini par

$$W_s^{\Delta t} = \frac{s - t_i}{\Delta t} W_{t_{i+1}} - \frac{s - t_{i+1}}{\Delta t} W_{t_i}, \quad s \in [t_i, t_{i+1}], i \in \{0, \dots, N-1\}.$$

On s'intéresse désormais au calcul de la séquence W^N .

1.4.2 Méthode de Cholesky

Cette méthode permet de simuler les trajectoires de tout processus gaussien centré et en particulier le m.b.s.. Avant de rentrer dans les détails, rappelons le résultat sur la décomposition de Cholesky d'une matrice symétrique définie positive (voir [6, p.80-81]).

Théorème 1.4.1 (Décomposition de Cholesky). *Soit A une matrice réelle de taille $n \times n$, symétrique et définie positive. Alors il existe une unique matrice triangulaire inférieure L de taille $n \times n$ dont les coefficients diagonaux sont tous positifs telle que $A = LL^T$*

On procède de la façon suivante.

1. Pour simuler un processus gaussien centré noté X aux instants t_1, \dots, t_N on introduit sa matrice de covariance $\Sigma \in \mathbb{R}^{N \times N}$ définie par $\Sigma_{ij} = \text{cov}(X_{t_i}, X_{t_j}), 1 \leq i, j \leq N$ (définie positive) de coefficients donnés par

$$\Sigma_{ij} = \text{cov}(W_{t_i}, W_{t_j}) = \min(t_i, t_j) = \frac{T}{N} \min(i, j), 1 \leq i, j \leq N.$$

2. Puis on calcule L la matrice de factorisation de Cholesky donnée par $\Sigma = LL^T$.
3. Etant donnée un vecteur gaussien centré réduit Z de taille n t.q. $Z \sim \mathcal{N}(0, Id)$ alors LZ a la même loi que $(X_{t_1}, \dots, X_{t_N}) \sim \mathcal{N}(0, LL^T)$ par la Proposition 1.2.5. La méthode est résumée dans l' Algorithme 1 et une trajectoire du brownien obtenue par ce biais est tracée sur la Figure 1.4.1.

Algorithm 1 Méthode de Cholesky pour le calcul du mouvement brownien.

```

1 # modules
2 import numpy as np
3 import matplotlib.pyplot as plt
4 from random import random
5
6 # grille temporelle
7 N = 1000
8 tps = np.linspace(0,1,N+1)
9 dt = 1/N
10
11 # matrice de covariance
12 Sigma = np.zeros([N+1,N+1])
13 for i in range(0,N+1):
14     for j in range(0,N+1):
15         Sigma[i,j] = dt*min(i,j)
16 Sigma = Sigma[1:N+1, 1:N+1]
17
18 # factorization de Cholesky (Sigma sym def pos)
19 L = np.linalg.cholesky(Sigma)
20
21 #vecteur de v.a. normales centrees reduites
22 Z = np.random.normal(0,1,N)
23
24 # brownien
25 W = L.dot(Z)
26 W = np.concatenate(([0],W))
27
28 # trace
29 plt.figure(1)
30 plt.grid()
31 plt.plot(tps,W,label="W")
32 plt.title('Mouvement brownien Standard - Methode Cholesky')
33 plt.legend(bbox_to_anchor=(1.05, 1), loc='upper left', borderaxespad=0.)
34 plt.show()

```

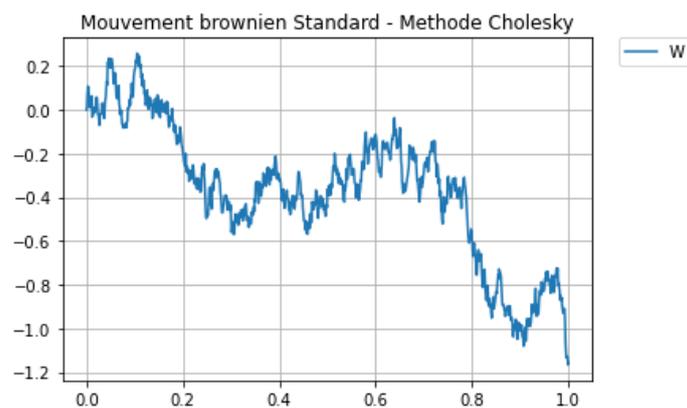


FIGURE 1.4.1 – Trajectoire du mouvement brownien sur $[0, 1]$ obtenu avec l’Algorithme 1 pour $N = 1000$.

1.4.3 Procédure progressive

L’idée est de calculer itérativement la séquence W^N à l’aide de v.a. gaussiennes centrées réduites indépendantes.

1. Générer une séquence (ξ_1, \dots, ξ_N) de n v.a. de loi $\mathcal{N}(0, 1)$.
2. En posant $W_{t_0} = W_0 = 0$ calculer pour $i = 1, \dots, N$

$$W_{t_i} = W_{t_{i-1}} + \sqrt{\Delta t} \xi_i,$$

car $W_{t_i} - W_{t_{i-1}} \sim \mathcal{N}(0, \Delta t)$. La méthode est résumée dans l' Algorithme 2 et une trajectoire du brownien obtenue par ce biais est tracée sur la Figure 1.4.2.

Algorithm 2 Méthode progressive pour le calcul du mouvement brownien.

```

1 #modules
2 import numpy as np
3 import matplotlib.pyplot as plt
4 from random import random
5
6 # grille temporelle
7 N = 1000
8 tps = np.linspace(0,1,N+1)
9 dt = 1/N
10
11 # matrice de covariance
12 W = np.zeros(np.size(tps,0))
13 for i in range(1,N+1):
14     W[i] = W[i-1]+np.random.normal(0,1)*np.sqrt(dt)
15
16 # trace
17 plt.figure(1)
18 plt.grid()
19 plt.plot(tps,W,label="W")
20 plt.title('Mouvement brownien Standard - Methode incrementale')
21 plt.legend(bbox_to_anchor=(1.05, 1), loc='upper left', borderaxespad=0.)
22 plt.show()

```

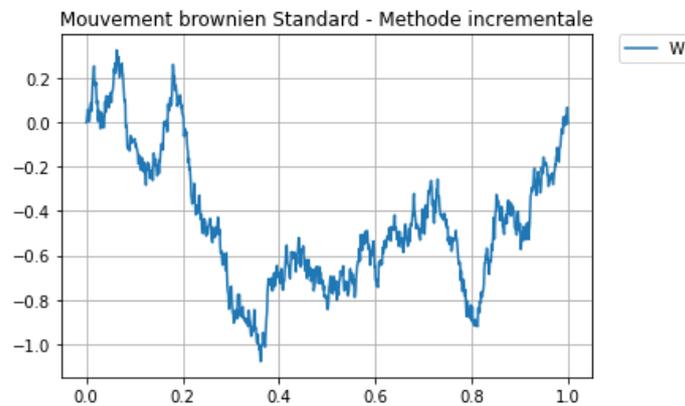


FIGURE 1.4.2 – Trajectoire du mouvement brownien sur $[0, 1]$ obtenu avec l'Algorithme 2 pour $N = 1000$.

C'est cette méthode que nous utiliserons en pratique pour la résolution numérique d'EDS.

1.4.4 Décomposition de Karhunen-Loève

Une autre méthode consiste à construire générer les W_{t_n} à partir de la décomposition de Karhunen-Loève (1.4.1) du m.b.s.

$$W_t = \sum_{m=0}^{\infty} \xi_m \phi_m(t), \quad t \in [0, T] \quad (1.4.1)$$

avec $\phi_m(t) = \frac{2\sqrt{2T}}{(2m+1)\pi} \sin\left(\frac{2m+1}{2T}\pi t\right)$ des fonctions orthogonales pour le produit scalaire L^2 sur $[0, T]$.

Pour cela, on évalue aux noeuds du maillage temporel $\{t_1, \dots, t_n\}$, le processus tronqué W_t^K de W_t aux K premiers termes de la série donné par

$$W_t^K = \sum_{m=0}^K \xi_m \phi_m(t), \quad t \in [0, T].$$

La méthode est résumée dans l' Algorithme 2 et une trajectoire du brownien obtenue par ce biais est tracée sur la Figure 1.4.2.

Algorithm 3 Méthode de Karhunen-Loève pour le calcul du mouvement brownien.

```

1 #modules
2 import numpy as np
3 import matplotlib.pyplot as plt
4 from random import random
5
6 # grille temporelle
7 N = 100
8 T = 1
9 tps = np.linspace(0,T,N+1)
10 dt = 1/N
11
12 # fonctions de base
13 def phi(t,m):
14     return 2*np.sqrt(2*T)/((2*m+1)*np.pi)*np.sin((2*m+1)*np.pi*t/(2*T));
15
16 # decomposition de KL
17 KMAX = 200
18 sample = np.random.normal(0,1,KMAX+1)
19 TK = [20, 40, 50, 100,200]
20 ind = np.linspace(0,KMAX,KMAX)
21 for l in range(0,np.size(TK)):
22     print(l)
23     K = TK[l]
24     W = np.zeros(np.size(tps,0))
25     xi = sample[0:K]
26     print(l)
27     for i in range(0,N+1):
28         for k in range(0,K):
29             W[i] = W[i]+xi[k]*phi(tps[i],k)
30
31 # trace
32 plt.figure(1)
33 plt.grid()
34 plt.plot(tps,W,label= 'K= %i' %K)
35 plt.title('Mouvement brownien Standard - Methode de Karhunen-Loeve')
36 plt.legend(bbox_to_anchor=(1.05, 1), loc='upper left', borderaxespad=0.)
37 plt.show()

```

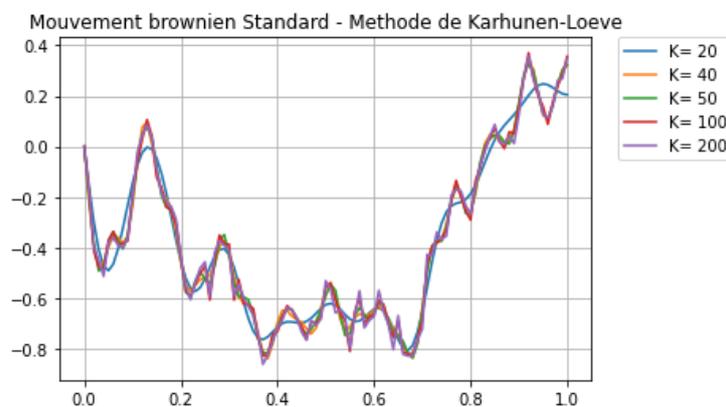


FIGURE 1.4.3 – Trajectoire du mouvement brownien sur $[0, 1]$ obtenu avec l'Algorithme 3 pour $N = 100$ et $K \in \{20, 40, 50, 100, 200\}$.

♠ On donne pour terminer un résultat sur l'erreur de troncature commise lorsqu'on approche W_t par W_t^K (développement de KL tronqué à $K + 1$ termes) qui nous dit que **l'erreur quadratique** se comporte en $O(K^{-1/2})$.

Proposition 1.4.2. *Pour $K \in \mathbb{N}$ et $t \in I = [0, T]$ l'erreur quadratique est donnée par*

$$\mathbb{V}(W_t) - \mathbb{V}(W_t^K) \leq \frac{2T}{\pi^2 K}$$

de même l'erreur en moment d'ordre 2 i.e. pour la norme de $L^2(\Omega, L^2(0, T))$ est donnée par

$$\mathbb{E} \left(\|W - W^K\|_{L^2(0, T)}^2 \right) \leq \frac{T}{\pi^2 K}.$$

Démonstration. Faite en TD. ■

Chapitre 2

ELÉMENTS DU CALCUL STOCHASTIQUE

Ce qu'il faut avoir retenu du cours d'A. Falcó : Filtration, Martingale, Intégrale de Wiener, de Itô, Processus de Itô, Formule de Itô, IPP

2.1 Objectifs de ce chapitre

Considérons $g : [0, T] \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 et $h : [0, T] \rightarrow \mathbb{R}$ intégrable.

MOTIVATION ① On peut écrire l'intégrale (au sens de Riemann)

$$\int_0^T h(t)dg(t) := \int_0^T h(t)g'(t)dt := \lim_{n \rightarrow \infty} \int_0^T h_n(t)g'(t)dt,$$

où $h_n(t) := \sum_{i=0}^{n-1} \xi_i \mathbf{1}_{[t_i, t_{i+1})}(t)$ est une fonction définie par morceaux t.q. $\int_0^T |h_n(t) - h(t)|dt \rightarrow 0$ quand $n \rightarrow \infty$. Par analogie on aimerait définir de même la notion d'intégrale pour

$$\int_0^T H_t(\omega)dW_t(\omega)$$

où W est un m.b.s. (de trajectoire non \mathcal{C}^1 , c.f. TD) et où H dénote un processus stochastique. Dans ce cas quelle notion de mesurabilité doit on donner à H et comment définir l'intégrale ?

MOTIVATION ② Désormais, considérons

$$\int_0^T g(t)dg(t) := \int_0^T g(t)g'(t)dt = \frac{1}{2} \int_0^T G'(t)dt = \frac{1}{2}(g^2(T) - g^2(0)).$$

où l'avant dernière égalité est obtenue en posant $G(t) = g^2(t)$. De la même façon que précédente, on souhaiterait pouvoir calculer

$$\int_0^T W_t dW_t$$

en posant $G_t = W_t^2$.

2.2 Filtration et mesurabilité

On considère par la suite $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilités. On appelle **sous-tribu** \mathcal{G} de \mathcal{A} , une tribu sur Ω incluse dans \mathcal{A} .

Définition 2.2.1 (Variables aléatoires et sous-tribus). Soit X une v.a. définie sur (Ω, \mathcal{A}) à valeurs dans (E, \mathcal{E}) .

- La variable aléatoire est mesurable par rapport à \mathcal{G} sous-tribu de \mathcal{A} si $\{X \in B\} \in \mathcal{G}$ pour tout $B \in \mathcal{E}$. On dit que X est \mathcal{G} -mesurable.
- La sous-tribu engendrée par X est notée $\sigma(X) := \{\{X \in B\}, B \in \mathcal{E}\}$. C'est une sous-tribu de \mathcal{A} .

Définition 2.2.2 (Filtration). Soit I un sous-ensemble de \mathbb{R} . Une **filtration** est une famille de sous-tribus, notée $\{\mathcal{F}_t\}_{t \in I}$ de \mathcal{A} t.q. pour tout $s, t \in I$ avec $s \leq t$ on a $\mathcal{F}_s \subset \mathcal{F}_t$. On la notera plus succinctement $\mathcal{F} := \{\mathcal{F}_t\}_{t \in I}$

Une filtration \mathcal{F} est donc une suite croissante de sous-tribus de \mathcal{A} et \mathcal{F}_t représente l'information disponible à l'instant t . La filtration va nous permettre de définir une notion de mesurabilité pour un processus X .

Définition 2.2.3 (Processus adapté). Soit $X = \{X_t\}_{t \in I}$ un processus stochastique. Ce processus est dit **adapté** à la filtration $\{\mathcal{F}_t\}_{t \in I}$, si pour tout $t \in I$ alors X_t est \mathcal{F}_t -mesurable.

On peut donc désormais définir une notion de mesurabilité de $X : (t, \omega) \rightarrow \mathbb{R}$ sur $[0, T] \times \Omega$ uniquement vis à vis de cette information disponible à chaque instant t , et non plus de la "grosse" tribu produit $\mathcal{B}([0, T]) \times \mathcal{A}$. Attention, ne pas confondre avec cette nouvelle notion de mesurabilité avec celle que nous avons introduite dans le Chapitre 1 pour les processus stochastiques. En effet, cette dernière donne une notion de mesurabilité pour la famille de v.a. indexée par t i.e. $\{X_t\}_{t \in I}$ à valeurs dans E !

Définition 2.2.4. Un processus stochastique $X = \{X_t\}_{t \in I}$ est dit **progressivement mesurable** si pour tout $t \in I$, l'application de $[0, t] \times \Omega$ dans \mathbb{R} définie par $(s, \omega) \rightarrow X_s(\omega)$ est mesurable pour la tribu produit $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$.

Un processus stochastique progressivement mesurable est adapté, le contraire n'est pas vrai en général.

On retiendra les propriétés pratiques suivantes en lien avec l'espérance conditionnelle (voir 5).

1. Soit X un processus stochastique. Si X_t est \mathcal{F}_t -mesurable : $\mathbb{E}(X_t | \mathcal{F}_t) = X_t$.
2. Soit X un processus stochastique. Si X_t est indépendant de \mathcal{F}_t : $\mathbb{E}(X_t | \mathcal{F}_t) = \mathbb{E}(X_t)$.
3. Soit X un processus stochastique et ξ une v.a. \mathcal{F}_t -mesurable : $\mathbb{E}(\xi X_t | \mathcal{F}_t) = \xi \mathbb{E}(X_t | \mathcal{F}_t)$.

Exemple 2.2.5 (Filtration engendrée par un processus stochastique). Soit $X = \{X_t\}_{t \in I}$ un processus stochastique. La filtration engendrée par X est définie par

$$\mathcal{F}_t^X = \sigma(\{X_s\}_{s \in I, s \leq t}),$$

pour tout $t \in I$. La filtration propre d'un processus stochastique X représente l'information contenue dans l'observation de X entre les instants 0 et t . On remarquera que si $\mathcal{F}_t^X \subset \mathcal{F}_t$ alors X est \mathcal{F} -adapté.

Nous terminons par énoncer la définition d'un m.b.s. adapté à une filtration $\mathcal{F} := \{\mathcal{F}_t\}_{t \geq 0}$.

Définition 2.2.6. On appellera un \mathcal{F} -mouvement brownien, un m.b.s. W t.q.

- W est adapté à la filtration \mathcal{F} ,
- pour tout $0 \leq s \leq t$ on a, $W_t - W_s$ indépendant de \mathcal{F}_s .

Cela permettra par la suite d'assurer que W_t est \mathcal{F}_t -mesurable. Une conséquence est que par indépendance dans l'espérance conditionnelle (voir rappels dans la section 5) $\mathbb{E}(W_t - W_s) = \mathbb{E}(W_t - W_s | \mathcal{F}_s) = \mathbb{E}(W_t | \mathcal{F}_s) - W_s = 0 \Leftrightarrow \mathbb{E}(W_t | \mathcal{F}_s) = W_s$. C'est ce qu'on appelle une propriété de **martingale**.

Définition 2.2.7 (Martingale). Une famille de v.a. $\{M_t, t \in I\}$ adaptée à la filtration $\{\mathcal{F}_t\}_{t \in I}$ est une **martingale** si

$$\mathbb{E}(M_t | \mathcal{F}_s) = M_s, \text{ p.s. } s < t,$$

en particulier on a $\mathbb{E}(M_t) = \mathbb{E}(M_0)$.

Proposition 2.2.8. Soit W un m.b.s. On définit les processus définis par $W_t^2 - t, \exp(\sigma W_t - \frac{\sigma^2}{2} t)$ sont des martingales pour $\{\mathcal{F}^W\}_{t \geq 0}$ la filtration propre du brownien W .

Démonstration. Soient $s, t \in [0, T], s < t$.

- Comme $\mathbb{E}(W_t | \mathcal{F}_s^W) - W_s = \mathbb{E}(W_t - W_s | \mathcal{F}_s^W) = 0$ (car $W_t - W_s$ indépendant de tout se qui se passe avant s) on a le résultat.
- On a $\mathbb{E}(W_t^2 - t | \mathcal{F}_s^W) - (W_s^2 - s) = \mathbb{E}(W_t^2 - t | \mathcal{F}_s^W) - \mathbb{E}(W_s^2 - s | \mathcal{F}_s^W) = \mathbb{E}((W_t - W_s)^2 + 2W_s(W_t - W_s) | \mathcal{F}_s^W) - (t - s) = (t - s) \left(\mathbb{E}\left(\frac{(W_t - W_s)^2}{t - s} | \mathcal{F}_s^W\right) - 1 \right)$. Or $W_t - W_s \sim \mathcal{N}(0, t - s)$ alors $\frac{(W_t - W_s)^2}{t - s} \sim \chi^2(1)$. Ainsi $\mathbb{E}\left(\frac{(W_t - W_s)^2}{t - s} | \mathcal{F}_s^W\right) = \mathbb{E}\left(\frac{(W_t - W_s)^2}{t - s}\right) = 1$ donc $\mathbb{E}(W_t^2 - t | \mathcal{F}_s^W) - (W_s^2 - s) = 0$. Ce qui prouve le résultat.
- On a $\mathbb{E}(\exp(\sigma W_t - \frac{\sigma^2}{2} t) | \mathcal{F}_s^W) - \exp(\sigma W_s - \frac{\sigma^2}{2} s) = \mathbb{E}(\exp(\sigma W_t - \frac{\sigma^2}{2} t) - \exp(\sigma W_s - \frac{\sigma^2}{2} s) | \mathcal{F}_s^W) = \exp(\sigma W_s - \frac{\sigma^2}{2} s) \exp(-\frac{\sigma^2}{2} (t - s)) [\mathbb{E}(\exp(\sigma(W_t - W_s)) | \mathcal{F}_s^W) - \exp(\frac{\sigma^2}{2} (t - s))]$. Cependant $\exp(\sigma(W_t - W_s)) \sim \log \mathcal{N}(0, \sigma^2(t - s))$ ainsi $\mathbb{E}(\exp(\sigma(W_t - W_s)) | \mathcal{F}_s^W) = \exp(\frac{\sigma^2}{2} (t - s))$ donc $\mathbb{E}(\exp(\sigma W_t - \frac{\sigma^2}{2} t) | \mathcal{F}_s^W) - \exp(\sigma W_s - \frac{\sigma^2}{2} s) = 0$.

■

2.2.1 Intégrale stochastique et propriétés

On présente dans ce qui suit des éléments de compréhension (pour les preuves cf. cours A. Falcó ou encore [1, 7]) de l'intégrale stochastique.

On considère par la suite $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilités. On travaille avec la filtration $\mathcal{F} = \{\mathcal{F}_t\}_{t \geq 0}$ et on note W un m.b.s. adapté à cette filtration. On supposera que \mathcal{F}_0 est complète (i.e. soit $A \in \mathcal{F}_0$ et $B \subseteq A$ et $\mathbb{P}(A) = 0$ alors $B \in \mathcal{F}_0$). Enfin, on notera $I = [0, T] \subset \mathbb{R}_+$.

L'objectif de ce paragraphe est de construire l'intégrale suivante

$$\int_0^T H_t dW_t,$$

pour des processus stochastiques H (dont les propriétés sont à définir pour que l'intégrale ait du sens, et soit bien définie).

PROCESSUS PROGRESSIVEMENT MESURABLES.

Soit $t = 0 < t_1 < \dots < t_n = T$ avec $t_i = iT/n$. Par analogie avec la définition de l'intégrale (au sens de Riemann) pour les fonctions régulières, une idée serait de construire le processus stochastique défini par morceaux H^n par

$$H_t^n = \sum_{i=0}^{n-1} \xi_i \mathbf{1}_{[t_i, t_{i+1})}(t)$$

et de définir

$$\int_0^T H_t dW_t = \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} \xi_i (W_{t_{i+1}} - W_{t_i}).$$

Dans la suite nous noterons

- $\Lambda^2(I)$ l'ensemble des processus progressivement mesurables t.q. $\int_0^T \mathbb{E}(H_t^2) dt < \infty$.
- $\Lambda_0^2(I) \subset \Lambda^2(I)$ contenant tous les processus stochastiques par morceaux $H^n = \{H_t^n\}_{t \in I}$ t.q. $t = 0 < t_1 < \dots < t_n = T$ et des v.a. ξ_0, \dots, ξ_{n-1} t.q. $H_t = \sum_{i=0}^{n-1} \xi_i \mathbf{1}_{[t_i, t_{i+1})}(t)$.^a

a. Cette convention est importante et correspond à l'intégrale de Itô. En effet il est possible de définir des processus stochastiques par morceaux différemment, ce qui amène à une autre définition de l'intégrale.

Lemme 2.2.9. *Un processus stochastique H^n appartient à $\Lambda_0^2(I)$ si et seulement si les v.a. $\xi_i, i = 0, \dots, n-1$ sont dans $L^2(\Omega)$ et \mathcal{F}_{t_i} -mesurables.*

Démonstration. On peut se convaincre assez facilement que $\int_0^T \mathbb{E}(H_t^2) dt = \sum_{i=0}^{n-1} \mathbb{E}(\xi_i^2)(t_{i+1} - t_i) < \infty$ si et seulement si $\mathbb{E}(\xi_i^2) < \infty$. Par ailleurs, si et seulement si les ξ_i sont \mathcal{F}_{t_i} -mesurables, cela assure que H est progressivement mesurable (voir détails [7, Lemma 10.1.3, Chapitre 10.1]). ■

Ce lemme indique que les v.a. ξ_i qui définissent les processus H^n de $\Lambda_0^2(I)$ ne dépendent que de l'information disponible à l'instant t_i !

Définition 2.2.10 (Intégrale stochastique pour les processus de $\Lambda_0^2(I)$). *Pour tout $H^n \in \Lambda_0^2(I)$, on définit l'intégrale stochastique comme*

$$\int_0^T H_t^n dW_t = \sum_{i=0}^{n-1} \xi_i (W_{t_{i+1}} - W_{t_i}). \quad (2.2.1)$$

Proposition 2.2.11 (Intégrale stochastique dans $\Lambda_0^2(I)$). *L'intégrale stochastique définie par (2.2.1) est un processus progressivement mesurable qui satisfait les propriétés suivantes.*

— (Chasles) Pour tout $T' \in [0, T]$ alors

$$\int_0^T H_t^n dW_t = \int_0^{T'} H_t^n dW_t + \int_{T'}^T H_t^n dW_t.$$

— (Linéarité) Pour tout $\lambda, \mu \in \mathbb{R}$ et $H^n, G^n \in \Lambda_0^2(I)$ on a

$$\int_0^T (\lambda H_t^n + \mu G_t^n) dW_t = \lambda \int_0^T H_t^n dW_t + \mu \int_0^T G_t^n dW_t.$$

— (Isométrie) On a

$$\mathbb{E} \left(\int_0^T H_t^n dW_t \right) = 0,$$

et

$$\mathbb{E} \left(\left(\int_0^T H_t^n dW_t \right)^2 \right) = \mathbb{E} \left(\int_0^T (H_t^n)^2 dt \right).$$

La dernière égalité est appelée **isométrie de Itô**.

Ici nous avons présenté ce résultat dans le cas scalaire pour des processus à valeurs réelles. Ces résultats peuvent être étendus dans le cas vectoriel, attention toutefois à travailler avec les normes et produits scalaires adéquats (voir e.g. [4, Theorem 8.14]).

Démonstration. La progressive mesurabilité et les deux premières propriétés sont directes. Pour la propriété d'espérance nulle, on écrit

$$\mathbb{E} \left(\int_0^t H_s^n dW_s \right) = \mathbb{E} \left(\int_0^t H_s^n dW_s \right) = \mathbb{E} \left(\sum_{i=0}^{n-1} \xi_i (W_{t_{i+1}} - W_{t_i}) \right) = \sum_{i=0}^{n-1} \mathbb{E} \left(\mathbb{E} (\xi_i (W_{t_{i+1}} - W_{t_i}) | \mathcal{F}_{t_i}) \right)$$

Or ξ_i est \mathcal{F}_{t_i} -mesurable, par propriété de l'espérance conditionnelle, il vient

$$\mathbb{E} \left(\int_0^t H_s^n dW_s \right) = \sum_{i=0}^{n-1} \mathbb{E} (\xi_i \mathbb{E} ((W_{t_{i+1}} - W_{t_i}) | \mathcal{F}_{t_i})) = \sum_{i=0}^{n-1} \mathbb{E} (\xi_i \mathbb{E} (W_{t_{i+1}} - W_{t_i})) = 0.$$

Les accroissements brownien étant indépendant de \mathcal{F}_{t_i} et centrés, on en déduit la dernière égalité.

En fait on pourrait même dire directement $\mathbb{E} ((W_{t_{i+1}} - W_{t_i}) | \mathcal{F}_{t_i}) = 0$. C'est la propriété de martingale.

Pour l'isométrie, on écrit

$$\mathbb{E} \left(\int_0^t (H_s^n)^2 dW_s \right) = \underbrace{\mathbb{E} \left(\sum_{i=0}^{n-1} \xi_i^2 (W_{t_{i+1}} - W_{t_i})^2 \right)}_{=A} + \underbrace{\mathbb{E} \left(\sum_{j \neq i} \xi_i \xi_j (W_{t_{i+1}} - W_{t_i})(W_{t_{j+1}} - W_{t_j}) \right)}_{=B}$$

En procédant de même que précédemment pour le premier terme on a

$$A = \sum_{i=0}^{n-1} \mathbb{E} \left(\mathbb{E} (\xi_i^2 (W_{t_{i+1}} - W_{t_i})^2 | \mathcal{F}_{t_i}) \right) = \sum_{i=0}^{n-1} \mathbb{E} \left(\xi_i^2 \underbrace{\mathbb{E} ((W_{t_{i+1}} - W_{t_i})^2)}_{=h} \right) = \sum_{i=0}^{n-1} \mathbb{E} (\xi_i^2) h = \int_0^T \mathbb{E} ((H_s^n)^2) ds.$$

Pour le second terme on utilise encore l'indépendance des accroissements par rapport à la filtration

$$B = 2 \sum_{j < i} \mathbb{E} \left(\mathbb{E} (\xi_i \xi_j (W_{t_{i+1}} - W_{t_i})(W_{t_{j+1}} - W_{t_j})) \right) = 2 \sum_{j < i} \mathbb{E} \left(\mathbb{E} (\xi_i \xi_j (W_{t_{i+1}} - W_{t_i})(W_{t_{j+1}} - W_{t_j}) | \mathcal{F}_{t_i}) \right)$$

Or par définition de la filtration, on a $\mathcal{F}_{t_j} \subset \mathcal{F}_{t_i}$ ainsi les ξ_i, ξ_j et $W_{t_{j+1}} - W_{t_j}$ sont \mathcal{F}_{t_i} -mesurables ainsi

$$B = 2 \mathbb{E} \left(\sum_{j < i} \xi_i \xi_j (W_{t_{j+1}} - W_{t_j}) \mathbb{E} (W_{t_{i+1}} - W_{t_i} | \mathcal{F}_{t_i}) \right) = 0,$$

et par indépendance des accroissement par rapport à \mathcal{F}_{t_i} alors $\mathbb{E} (W_{t_{i+1}} - W_{t_i} | \mathcal{F}_{t_i}) = \mathbb{E} (W_{t_{i+1}} - W_{t_i}) = 0$. ■

Nous admettons les résultats qui suivent sans les démontrer (pour plus de détails voir e.g. [1, 7]).

CAS DES PROCESSUS $\Lambda^2(I)$ (VOIR [1]).

Désormais, on souhaite étendre cette définition aux processus plus généraux, i.e. dans $\Lambda^2(I)$. Par la suite on équipe $\Lambda^2(I)$ de la norme $\|\cdot\|_{\Lambda^2(I)}$ définie par

$$\|H\|_{\Lambda^2(I)}^2 = \int_0^T \mathbb{E}(H_t^2) dt.$$

Nous avons les propriétés suivantes.

1. L'espace $\Lambda^2(I)$ muni de la norme $\|\cdot\|_{\Lambda^2(I)}$ est un espace de Banach.
2. L'espace $\Lambda_0^2(I)$ est dense dans $\Lambda^2(I)$.
3. L'application $\mathcal{J}_T : \Lambda_0^2(I) \rightarrow L^2(\Omega)$ définie par $\mathcal{J}_T : H^n \mapsto \int_0^T H_t^n dW_t$ est une isométrie c.à.d.

$$\|H^n\|_{\Lambda^2(I)}^2 = \mathbb{E} \left(\int_0^T H_t^n dW_t \right).$$

On retiendra que : pour $H \in \Lambda^2(I)$ et pour peut qu'il existe $H^n \in \Lambda_0^2(I)$ t.q. $\|H - H^n\|_{\Lambda^2(I)}^2 \rightarrow 0$. On définit l'intégrale stochastique de H comme

$$\int_0^T H_t dW_t = \lim_{n \rightarrow \infty} \int_0^T H_t^n dW_t, \text{ dans } L^2(\Omega).$$

Ainsi, les résultats de la proposition 2.2.11 sont toujours vrais, pour $H \in \Lambda^2(I)$.

CAS DES PROCESSUS $\Lambda_{loc}(I)$.

Il est possible de relaxer l'hypothèse $H \in \Lambda^2(I)$ en considérant uniquement des processus stochastiques dans $\Lambda_{loc}(I)$ et t.q.

$$\forall T > 0 \quad \int_0^T H_t^2 dt < \infty, \quad \text{presque sûrement.}$$

Bien sur si $H \in \Lambda^2(I)$ alors $H \in \Lambda_{loc}(I)$. L'idée est ensuite de construire l'intégrale stochastique pour tout $H \in \Lambda_{loc}(I)$ un processus (localisé) $H^M \in \Lambda^2(I)$ t.q.

$$\int_0^T H_t dW_t = \lim_{M \rightarrow \infty} \mathcal{J}(H^M) \text{ presque sûrement.}$$

Les résultats de la Proposition 2.2.11 restent vrais pour $\Lambda_{loc}(I)$. Attention toutefois au point 3, qui est satisfait si de plus $\mathbb{E}(\int_0^T H_t^2 dt) < \infty$). Les preuves nécessitent la définition de temps d'arrêt que nous n'avons pas introduit dans ce cours.

2.3 Processus d'Itô, formule de Itô

On termine cette partie avec un outil fondamental du calcul stochastique la formule de Itô. On note W un m.b.s. adapté à la filtration $\mathcal{F} = \{\mathcal{F}_t\}_{t \in I}$.

2.3.1 Formule de Itô pour le mouvement brownien

Théorème 2.3.1 (Formule de Itô pour le mouvement brownien $\phi(W_t)$). *Soit $\phi \in \mathcal{C}^2(\mathbb{R})$ alors pour tout $t > 0$ on a*

$$\phi(W_t) = \phi(0) + \int_0^t \phi'(W_s) dW_s + \frac{1}{2} \int_0^t \phi''(W_s) ds, \text{ p.s.,} \quad (2.3.1)$$

ou de façon équivalente

$$d\phi(W_t) = \phi'(W_t) dW_t + \frac{1}{2} \phi''(W_t) dt. \quad (2.3.2)$$

REMARQUES. En fait, comme $\phi \in \mathcal{C}^2$ alors on a $t \mapsto \phi'(W_t)$ et $t \mapsto \phi''(W_t)$ continues p.s. donc bornées p.s. sur $[0, T]$, en particulier cela assure qu'on a bien un processus $\{\phi'(W_t)\}_{t \geq 0}$ dans $\Lambda_{loc}([0, T])$. Les intégrales dans le membre de droite de l'équation sont donc bien définies.

La formule relève une analogie avec un développement de Taylor d'ordre deux de la fonction ϕ , qui est exploité en pratique pour la preuve (cf. e.g. cours A. Falco).

Exemple 2.3.2. Soit W un m.b.s. adapté à la filtration \mathcal{F} . Montrer que pour tout $t \geq 0$

$$\int_0^t W_s dW_s = \frac{1}{2}(W_t^2 - t), p.s..$$

Considérons, $\phi(x) = x^2$ on a $\phi \in \mathcal{C}^2(\mathbb{R})$ avec $\phi'(x) = x, \phi''(x) = 1$. La formule de Itô donne

$$\frac{1}{2}W_t^2 = \frac{1}{2}W_0^2 + \int_0^t W_s dW_s + \frac{1}{2} \int_0^t ds \Rightarrow \int_0^t W_s dW_s = \frac{1}{2}(W_t^2 - t).$$

2.3.2 Processus de Itô et formule de Itô pour un processus de Itô

Définition 2.3.3 (Processus de Itô). Soit $\{b_t\}_{t \in I}$ et $\{\sigma_t\}_{t \in I}$ deux processus progressivement mesurables (donc \mathcal{F} -adaptés), à valeurs dans \mathbb{R} et t.q. $\int_0^T |b_t| dt < \infty$ et $\int_0^T \sigma_t^2 dt < \infty$ presque sûrement. Enfin soit X_0 une .v.a. \mathcal{F}_0 -mesurable. On appelle **processus de Itô** $\{X_t\}_{t \in I}$, le processus à valeurs réelles partant de X_0 défini par

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s, \quad t \in I.$$

ou bien sous forme différentielle

$$dX_t = b_t dt + \sigma_t dW_t.$$

Définition 2.3.4 (Variation quadratique d'un processus de Itô). Soit X un processus de Itô. On appelle **variation quadratique** de X , le processus positif et croissant $\langle X \rangle = \{\langle X \rangle_t\}_{t \geq 0}$ défini par

$$\langle X \rangle_t = \int_0^t \sigma_s^2 ds.$$

On utilise aussi la notation $d\langle X \rangle_t = \sigma_t^2 dt$.

REMARQUE. La variation quadratique vérifie pour tout $\lambda \in \mathbb{R}, \langle \lambda X \rangle_t = \lambda^2 \langle X \rangle_t$.

Théorème 2.3.5 (Formule de Itô pour un processus de Itô). Soit X un processus de Itô et $\phi \in \mathcal{C}^2(\mathbb{R})$ alors $\phi(X)$ est un processus de Itô, et pour tout $t \in I$ on a

$$\phi(X_t) = \phi(X_0) + \int_0^t \phi'(X_s) b_s + \frac{1}{2} \phi''(X_s) \sigma_s^2 ds + \int_0^t \sigma_s \phi'(X_s) dW_s, \quad (2.3.3)$$

ou encore sous forme différentielle

$$d\phi(W_t) = (\phi'(X_t) b_t + \frac{1}{2} \phi''(X_t) \sigma_t^2) dt + \sigma_t \phi'(X_t) dW_t = \phi'(X_t) dX_t + \frac{1}{2} \phi''(X_t) d\langle X \rangle_t. \quad (2.3.4)$$

Exemple 2.3.6 (Cas du mouvement brownien géométrique). Montrons que $X_t = \exp\left((b - \frac{\sigma^2}{2}t) + \sigma W_t\right) x_0$ est solution de $dX_t = bX_t dt + \sigma X_t dW_t$, avec la condition initiale $X_0 = x_0$. Si $t = 0$, alors $X_0 = x_0$. Soit $\phi : \mathbb{R}_+^* \rightarrow \mathbb{R}$ avec $\phi(x) = \ln(x)$, on a bien $f \in \mathcal{C}^2(\mathbb{R}_+^*)$. En posant $Y_t = \ln X_t$ par la formule d'Itô Y satisfait $Y_0 = \ln x_0$ et

$$dY_t = d \ln(X_t) = 0 dt + \frac{1}{X_t} (bX_t dt + \sigma X_t dW_t) - \frac{1}{2} \frac{\sigma^2 X_t^2}{X_t^2} dt = (b - \frac{1}{2} \sigma^2) dt + \sigma dW_t.$$

Ainsi Y est mouvement brownien algébrique donné par $Y_t = \ln x_0 + (b - \frac{1}{2} \sigma^2)t + \sigma W_t$ on en déduit donc X_t en passant à l'exponentielle.

Généralisations

Soit $W = (W_1, \dots, W_d)$ un mouvement brownien d -dimensionnel, i.e. t.q. pour tout $k = 1, \dots, d$ les processus $W_k, k = 1, \dots, d$ sont des m.b.s. adaptés à la filtration $\{\mathcal{F}_t\}_{t \geq 0}$.

Définition 2.3.7 (Processus de Itô pour un mouvement brownien d -dimensionnel). *Soit deux processus progressivement mesurables (adaptés) b à valeurs dans \mathbb{R} et $\sigma = \{(\sigma_{1,t}, \dots, \sigma_{d,t})^T\}_{t \in I}$ à valeurs dans \mathbb{R}^d t.q. $\int_0^t |b_s| ds < \infty$ p.s. et $\sigma_k \in \Lambda_{loc}$ pour $k = 1, \dots, d$. Enfin soit X_0 une v.a. \mathcal{F}_0 -mesurable. Le **processus de Itô** $\{X_t\}_{t \in I}$ partant de X_0 est le processus à valeurs réelles défini par*

$$X_t = X_0 + \int_0^t b_s ds + \sum_{k=1}^d \int_0^t \sigma_{k,s} dW_{k,s}, \quad t \in I.$$

De façon équivalente,

$$dX_t = b_t dt + \sigma_t \cdot dW_t,$$

(ici \cdot désigne le produit scalaire euclidien dans \mathbb{R}^d).

Définition 2.3.8 (Covariation quadratique d'un processus de Itô). *Soit X un processus de Itô. On appelle la **covariation quadratique** de X, Y deux processus de Itô à valeurs vectorielles de diffusions σ^X, σ^Y respectivement. Le processus stochastique $\langle X, Y \rangle = \{\langle X, Y \rangle_t\}_{t \geq 0}$ défini par*

$$\langle X, Y \rangle_t = \int_0^t \sum_{k=1}^d \sigma_{k,s}^X \sigma_{s,k}^Y ds,$$

ainsi $d\langle X, Y \rangle_t = \sum_{k=1}^d \sigma_{k,s}^X \sigma_{s,k}^Y dt$.

REMARQUE. En particulier pour W^1, W^2 deux m.b.s. indépendants, on a [1, §3.2.5] $d\langle W^i, W^j \rangle_t = \delta_{ij} dt$.

Soit W un brownien standard, et b, σ, β, γ des processus adaptés au sens de la définition précédente. Considérons deux processus d'Itô définis par

$$dX_t = b_t dt + \sigma_t dW_t \quad \text{et} \quad dY_t = \beta_t dt + \gamma_t dW_t,$$

alors on a la *formule d'intégration par parties* suivante¹ :

$$d(X_t Y_t) = X_t dY_t + Y_t dX_t + d\langle X, Y \rangle_t,$$

ou de façon équivalente

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \int_0^t \sigma_s \gamma_s ds.$$

REMARQUE. En comparaison avec la formule d'IPP déterministe, on a un terme additionnel lié à la covariation quadratique des deux processus. Ce dernier terme est cependant nul si un des deux processus est une intégrale en temps uniquement.

Théorème 2.3.9 (Formule de Itô q -dimensionnelle et dépendant du temps). *Soit $\phi \in \mathcal{C}^{1,2}(I \times \mathbb{R}^q, \mathbb{R})$ et $X = (X_1, \dots, X_q)^T$ un processus d'Itô q -dimensionnel i.e., pour tout $i = 1, \dots, q$ les X_i sont des processus de d'Itô pour un m.b.s. d -dimensionnel de processus progressivement mesurables (voir Définition 2.3.7). Alors $Y_t = \phi(t, X_t)$ est le processus de dimension 1 donné par*

$$d\phi(t, X_t) = \frac{\partial}{\partial t} \phi(t, X_t) dt + \sum_{i=1}^q \frac{\partial}{\partial x_i} \phi(t, X_t) dX_{i,t} + \frac{1}{2} \sum_{i,j=1}^q \frac{\partial^2}{\partial x_i \partial x_j} \phi(t, X_t) d\langle X_i, X_j \rangle_t.$$

où $dX_{i,t} = b_{i,t} dt + \sigma_{i,t} \cdot dW_t, i = 1, \dots, q$ ou encore

$$d\phi(t, X_t) = \underbrace{\partial_t \phi(t, X_t)}_{\text{gradient}} dt + \underbrace{\nabla \phi(t, X_t)}_{\text{hessienne}} \cdot dX_t + \frac{1}{2} \text{tr}(\nabla^2 \phi(t, X_t) \sigma_t \sigma_t^T) dt.$$

1. Voir le cours Antonio ou [1, 7] pour la notion de variation quadratique $\langle X \rangle_t$ ou de covariation quadratique $\langle dX, dY \rangle_t$.

REMARQUE : Ici on a exprimé le dernier terme en utilisant les règles de calcul sur les variations quadratiques. Comme σ est supposé à valeurs matricielle dans $\mathbb{R}^{q \times d}$ il vient

$$\sum_{i,j=1}^q \frac{\partial^2}{\partial x_i \partial x_j} \phi(t, X_t) \sum_{k=1}^d \sigma_{ik,t} \sigma_{jk,t} dt = \sum_{i,j=1}^q \frac{\partial^2}{\partial x_i \partial x_j} \phi(t, X_t) d\langle X_i, X_j \rangle_t,$$

avec X_i, X_j les i, j -ième composantes de X .

Chapitre 3

INTÉGRATION NUMÉRIQUE DES EDS

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilités équipé d'une filtration $\mathcal{F} = \{\mathcal{F}_t\}_{t \in I}$ t.q. \mathcal{F}_0 est complète. On notera $I = \mathbb{R}_+$ ou encore $I = [0, T]$ avec $T > 0$. Dans la suite, on utilise $W = (W_1, \dots, W_d)^T$ un m.b.s. d -dimensionnel, t.q. les $W_k, k = 1, \dots, d$ sont adaptés à la filtration \mathcal{F} .

On munira \mathbb{R}^q du produit scalaire $\langle \cdot, \cdot \rangle$ et de la norme $\|\cdot\|$ euclidiens et $\mathbb{R}^{q \times d}$ du produit scalaire $\langle \cdot, \cdot \rangle_F$ et de la norme $\|\cdot\|_F$ de Frobenius.

3.1 Notations

On considère $b : I \times \mathbb{R}^q \rightarrow \mathbb{R}^q$ et $\sigma : I \times \mathbb{R}^q \rightarrow \mathbb{R}^{q \times d}$ des fonctions mesurables. On cherche le processus de **diffusion** X à valeurs dans \mathbb{R}^q solution de l'équation différentielle stochastique (EDS) au sens d'Itô définie par

$$\begin{cases} dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, & t \in (0, T], \\ X_0 = \xi, \end{cases} \quad (3.1.1)$$

avec $\xi = (\xi_1, \dots, \xi_q)^T$ une v.a. \mathcal{F}_0 -mesurable. Les termes $b(t, X_t)$ et $\sigma(t, X_t)$ sont appelés termes de **dérive** et **diffusion** respectivement. On peut écrire de façon équivalente (par composante) cette équation sous la forme

$$\begin{cases} dX_{i,t} = b_i(t, X_t)dt + \sum_{j=1}^d \sigma_{ij}(t, X_t)dW_{j,t}, & t \in (0, T], \\ X_{0,i} = \xi_i. \end{cases}$$

Nous allons donner quelques conditions suffisantes (mais non nécessaires), pour assurer l'existence d'une **solution (forte)** au problème (3.1.1).

3.1.1 Quelques exemples

Exemple 3.1.1 (Mouvement brownien arithmétique $d = q = 1$). *Etant donné W un mouvement brownien et en choisissant pour diffusion et dérive des constantes i.e. $\sigma(s, x) = \sigma$ et $b(s, x) = b$ respectivement, le système*

$$\begin{cases} dX_t = bdt + \sigma dW_t, \\ X_0 = x_0, \end{cases}$$

*régit l'évolution d'un processus stochastique X appelé **mouvement brownien arithmétique** t.q.*

$$X_t = x_0 + bt + \sigma W_t.$$

Exemple 3.1.2 (Mouvement brownien géométrique $d = q = 1$). *Etant donné W un mouvement brownien et en choisissant pour diffusion et dérive $\sigma(s, x) = \sigma x$ et $b(s, x) = bx$ respectivement, le système*

$$\begin{cases} dX_t = bX_t dt + \sigma X_t dW_t, \\ X_0 = x_0, \end{cases}$$

*régit l'évolution d'un processus stochastique X appelé **mouvement brownien géométrique** t.q.*

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

Le mouvement brownien géométrique n'est rien d'autre que l'exponentielle d'un mouvement brownien arithmétique (obtenu à l'aide de la formule de Itô).

Exemple 3.1.3 (Processus d'Orstein-Uhlenbeck $d = q = 1$). *En physique, économie ou finance il est courant d'utiliser le processus d'Orstein-Uhlenbeck X donné par*

$$X_t = \theta + (x_0 - \theta) \exp(-at) + \sigma \int_0^t \exp(-a(t-s)) dW_s$$

solution de

$$\begin{cases} dX_t &= -a(X_t - \theta)dt + \sigma dW_t, \\ X_0 &= x_0, \end{cases}$$

correspondant au coefficients de diffusion et dérive donnés par $\sigma(s, x) = \sigma$ et $b(s, x) = -a(x - \theta)$.

Exemple 3.1.4 (Un modèle pour décrire le déplacement d'un poisson $q = 4, d = 1$). *En modélisant l'environnement du poisson (mer, lac...) comme une surface, le mouvement de celui-ci peut être décrit par le biais de trois processus $X \in \mathbb{R}^2$ décrivant sa position, $\Theta \in \mathbb{R}$ sa vitesse angulaire et $K \in \mathbb{R}$ sa courbure vérifiant le système d'équations différentielles suivant*

$$\begin{cases} dX_t &= \tau(\Theta_t, K_t)dt \\ d\Theta_t &= K_t dt, \\ dK_t &= -K_t dt + \sqrt{2\alpha} dW_t, \\ X_0 &= x_0, \quad K_t = k_0, \quad \Theta_0 = \theta_0, \end{cases}$$

avec $\alpha > 0$ et $\tau(\theta, K) = \frac{1}{1+2|K|}(\cos(\theta), \sin(\theta))^T$.

3.2 Existence et unicité de la solution

Dans le cas déterministe c'est à dire $\sigma := 0$, alors on retrouve un EDO. Dans ce cas sous les hypothèses du théorème de Cauchy-Lipschitz assurent l'existence et l'unicité de la solution. Voyons comment cela se traduit pour une EDS.

Théorème 3.2.1 (Existence et unicité de la solution (forte) [1, 7]). *Supposons qu'il existe une constante $0 < K < \infty$ telle que*

i) *(globalement lipschitz en x) telle que tout $x, y \in \mathbb{R}^q$ et $t \in I$*

$$\|b(t, x) - b(t, y)\| + \|\sigma(t, x) - \sigma(t, y)\|_F \leq K\|x - y\|,$$

ii) *(fonctions à croissance linéaire¹) pour tout $x \in \mathbb{R}^q$ et $t \in I$*

$$\|b(t, x)\| + \|\sigma(t, x)\|_F \leq K(1 + \|x\|).$$

Etant donnée $\xi \in L^2(\mathcal{F}_0)$, le problème (3.1.1) admet pour unique solution (forte), le processus de Itô X , t.q. $X_k \in \Lambda^2(I), k = 1, \dots, q$, donné par

$$X_t = \xi + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s, \quad t \in I.$$

On appelle X **processus de diffusion**.

Comme pour les EDO, la preuve de l'existence utilise un point fixe défini par le biais d'un processus itératif de Picard. L'hypothèse lipschitz assure quant à lui l'unicité. Pour une preuve voir [4] pour les EDS autonomes ou encore [1], enfin [3, 5] dans le cas général. Ici on fait une preuve dans la lignée de [1, 7].

Démonstration. Comme $\xi \in L^2(\mathcal{F}_0)$ et b, σ à croissance linéaire, on peut montrer que les composantes de X sont des processus progressivement mesurables (adaptés) pour la filtration \mathcal{F} t.q. $\int_0^T \mathbb{E}(X_t)^2 dt < \infty$ donc dans $\Lambda^2(I)$. On admettra ce résultat, pour une preuve détaillée voir [7, Lemma 11.1.2]. Il reste donc à prouver l'existence et l'unicité.

1. Cette condition est aussi parfois écrite comme $\|b(t, x)\|^2 + \|\sigma(t, x)\|_F^2 \leq K(1 + \|x\|^2)$ dans certains ouvrages. Elle découle tout simplement de l'inégalité $(a + b)^2 \leq 2(a^2 + b^2)$.

Ici nous allons présenter dans le cas $q = d = 1$ la preuve, il en est de même dans le cas vectoriel. Soient $X, Y \in \Lambda^2(I)$ deux solutions de (3.1.1). Alors pour tout $t \in I$ on a

$$X_t - Y_t = \int_0^t (b(s, X_s) - b(s, Y_s)) ds + \int_0^t (\sigma(s, X_s) - \sigma(s, Y_s)) dW_s$$

En utilisant que $(a + b)^2 \leq 2(a^2 + b^2)$ et en passant à l'espérance, il vient

$$\mathbb{E}[(X_t - Y_t)^2] \leq 2\mathbb{E} \left[\left(\int_0^t (b(s, X_s) - b(s, Y_s)) ds \right)^2 + \left(\int_0^t (\sigma(s, X_s) - \sigma(s, Y_s)) dW_s \right)^2 \right]$$

Puis en appliquant Cauchy-Schwarz à la première équation, l'isométrie de Itô à la seconde on a

$$\mathbb{E}[(X_t - Y_t)^2] \leq 2t\mathbb{E} \left(\int_0^t (b(s, X_s) - b(s, Y_s))^2 ds \right) + 2\mathbb{E} \int_0^t (\sigma(s, X_s) - \sigma(s, Y_s))^2 ds$$

Cependant les dérivées et diffusion étant lipschitziennes on a

$$\mathbb{E}[(X_t - Y_t)^2] \leq \underbrace{2K^2(T+1)}_C \mathbb{E} \left(\int_0^t (X_s - Y_s)^2 ds \right) = C \int_0^t \mathbb{E}[(X_s - Y_s)^2] ds$$

par Fubini/Tonelli (en effet $X, Y \in \Lambda^2(I)$ donc $\int_0^t \mathbb{E}[(X_s - Y_s)^2] ds < \infty$).

On remarque qu'en appliquant le lemme de Gronwall pour $x(t) = \mathbb{E}[(X_t - Y_t)^2]$ et $x(0) = 0$ on a

$$x(t) \leq x(0) + C \int_0^t x(s) ds \Rightarrow x(t) \leq x(0) \exp(Ct) = 0.$$

D'où l'unicité.

On construit l'application $G : \mathbb{R} \rightarrow \mathbb{R}$ t.q.

$$(GX)_t = \xi + \int_0^t b(s, X(s)) ds + \int_0^t \sigma(s, X(s)) dW_s$$

Par construction, $GX \in \Lambda^2(I)$ (on a X solution de (3.1.1) $\Leftrightarrow GX = X$). On a donc pour tout $X, Y \in \Lambda^2(I)$

$$\mathbb{E}[(GX)_t - (GY)_t]^2 \leq C \int_0^t \mathbb{E}[(X_s - Y_s)^2] ds$$

Ainsi par récurrence, pour $n \geq 1$

$$\mathbb{E}[(G^n X)_t - (G^n Y)_t]^2 \leq C \int_0^t \mathbb{E}[(G^{n-1} X)_s - (G^{n-1} Y)_s]^2 ds \leq C^n \int_{s_1=0}^t \dots \int_{s_{n-1}=0}^{s_{n-1}} \mathbb{E}[(X_{s_n} - Y_{s_n})^2] ds_n \dots ds_1$$

En utilisant Fubini (attention aux bornes) on a

$$\mathbb{E}[(GX)_t - (GY)_t]^2 \leq C^n \int_{s_n=0}^t \mathbb{E}[(X_{s_n} - Y_{s_n})^2] \underbrace{\int_{s_{n-1}=s_n}^t \dots \int_{s_1=s_2}^t ds_1 \dots ds_{n-1} ds_n}_{(t-s_n)^{n-1}/(n-1)!}.$$

On conclut donc en repassant à l'intégrale à gauche sur t .

$$\int_0^T \mathbb{E}[(G^n X)_t - (G^n Y)_t]^2 dt \leq C^n \int_0^T \int_0^t \mathbb{E}[(X_s - Y_s)^2] \frac{(t-s)^{n-1}}{(n-1)!} ds dt$$

à nouveau par Fubini il vient

$$\int_0^T \mathbb{E}[(G^n X)_t - (G^n Y)_t]^2 dt \leq C^n \frac{T^n}{n!} \int_0^T \mathbb{E}[(X_s - Y_s)^2] ds.$$

Dès lors que $n \geq 1$ est assez grand i.e. t.q. $C^n \frac{T^n}{n!} < 1$ alors G^n est une contraction. Par ailleurs, $\Lambda^2(I)$ étant un Banach (voir Chapitre 2), le théorème du point fixe de Banach (cf. cours analyse fonctionnelle) assure que G^n admet un unique point fixe $X \in \Lambda^2(I)$ t.q. $G^n X = X$. Cependant on observe aussi que $G^{n+1} X = G(G^n X) = GX = G^n(GX)$. Donc GX est aussi un point fixe de G^n . Par unicité du point fixe de G^n il vient nécessairement $X = G(X)$ qui est solution de (3.1.1). ■

Exemple 3.2.2. *Considérons, l'équation de Langevin avec $a > 0$*

$$dX_t = -aX_t dt + \sigma dW_t, \quad X_0 = x_0 \in \mathbb{R}.$$

On est dans le cas homogène en temps. Il suffit donc de vérifier que la dérive et la diffusion sont Lipschitz, ce qui est le cas car $x \mapsto -ax$ est linéaire, et $x \mapsto \sigma$ constante. Donc l'équation de Langevin admet une unique solution que l'on peut calculer à l'aide de la formule de Itô. En posant $Y_t = X_t e^{at} = f(t, X_t)$ alors $dY_t = aX_t e^{at} dt + e^{at} dX_t = e^{at} \sigma dW_t$ d'où $Y_t = Y_0 + \int_0^t e^{as} \sigma dW_s \Rightarrow X_t = X_0 e^{-at} + \int_0^t e^{a(s-t)} \sigma dW_s$.

3.3 Intégration numérique des EDS

Nous nous intéressons à l'approximation numérique d'un processus de diffusion X solution de (3.1.1). Comme pour les EDO, le processus X peut être explicitement calculé seulement pour des cas particuliers des coefficients de dérive b et de diffusion σ e.g. constants ou linéaires pour lesquels X est brownien arithmétique ou géométrique (cf cours A. Falcó). Généralement, il n'est pas possible d'utiliser les outils du calcul stochastique comme la formule d'Itô pour trouver une expression explicite de X . On utilise donc des méthodes numériques.

Soit un maillage uniforme $0 = t^0 < t^1 < \dots < t^n < \dots < t^N$ de $[0, T]$, $t^n = n\Delta t$, $n = 0, 1, \dots, N$ où $\Delta t = \frac{T}{N}$. Le processus X en tout instant t^n i.e. X_{t^n} est approché par la v.a. X^n calculé par un schéma d'approximation. Dans ce cours nous nous focaliserons sur les schémas d'**Euler-Maruyama et Milstein**. Si on s'intéresse à la trajectoire complète on peut reconstruire par interpolation un processus approché $X^{\Delta t}$ (cf. plus loin pour l'analyse d'erreur).

♠ Nous introduisons de **façon formelle** le schéma d'Euler-Maruyama. Considérons

$$\begin{aligned} X_{t^{n+1}} &= \int_0^{t^{n+1}} b(s, X_s) ds + \int_0^{t^{n+1}} \sigma(s, X_s) dW_s, \\ X_{t^n} &= \int_0^{t^n} b(s, X_s) ds + \int_0^{t^n} \sigma(s, X_s) dW_s, \end{aligned}$$

après soustraction membre à membre nous obtenons

$$X_{t^{n+1}} = X_{t^n} + \int_{t^n}^{t^{n+1}} b(s, X_s) ds + \int_{t^n}^{t^{n+1}} \sigma(s, X_s) dW_s.$$

On propose d'approcher sur $[t^n, t^{n+1})$ les fonctions de dérive et de diffusion par des fonctions constantes par morceaux. Ainsi on obtient une approximation X^n de X_{t^n} pour tout t^n donnée par

$$X^{n+1} = X^n + b(t^n, X^n) \Delta t + \sigma(t^n, X^n) \Delta W^n.$$

avec un accroissement $\Delta W^n = W_{t^{n+1}} - W_{t^n}$. Cette première approximation définit le schéma d'Euler-Maruyama.

Nous verrons, qu'il est possible de construire un tel schéma de la même manière que des schémas d'ordre supérieur en considérant des développements de Taylor des termes de dérive et diffusion (c.f. TD), pour l'instant nous présentons directement ces schémas ainsi que leurs propriétés.

3.3.1 Schémas d'approximation

Définition 3.3.1 (Schéma d'Euler-Maruyama). *Pour un pas de temps $\Delta t > 0$ et une condition initiale $X^0 = X_0$, l'approximation X^n de X_{t^n} , au temps $t^n = n\Delta t$, par le schéma d'Euler-Maruyama est défini par*

$$X^{n+1} = X^n + b(t^n, X^n)\Delta t + \sigma(t^n, X^n)\Delta W^n, \quad (3.3.1)$$

pour $\Delta W^n = W_{t^{n+1}} - W_{t^n}$.

En pratique, les incréments du m.b.s. ΔW^n sont des vecteurs gaussiens i.e. $\Delta W^n \sim \mathcal{N}(0, \Delta t I_q)$. On peut donc simplement les simuler en faisant des tirages aléatoires indépendants (voir Algorithme 2).

Le schéma d'Euler-Maruyama peut être implémenté de la façon qui suit, voir Algorithme 4.

Algorithm 4 Fonction pour le schéma d'Euler-Maruyama.

```

1 # EulerMaruyama
2 def EulerMaruyama(u0,t,N,d,m):
3
4     print("Euler Maruyama")
5     # initialisation
6     u = np.zeros((m,np.size(t,0)));
7     u[:,0] = u0;
8
9     # boucle
10    for n in range(1,N):
11        dt = t[n+1]-t[n];
12        dW = np.sqrt(dt)*np.random.normal(0,1,m);
13        u[:,n+1] = u[:,n]+dt*b(u[:,n])+np.multiply(dW,sigma(u[:,n]));
14    return u;
```

On s'intéresse désormais à un second schéma d'approximation **plus précis** (en un sens à préciser). Pour simplifier la présentation et pour les applications considérées dans ce cours, on présente uniquement le schéma de Milstein pour une fonction de diffusion diagonale $\sigma \in \mathbb{R}^{d \times d}$ i.e. t.q. $\sigma_{ij}(x) = 0$ pour tout $i \neq j$. Pour le cas général voir [4].

Définition 3.3.2 (Schéma de Milstein). *Pour un pas de temps $\Delta t > 0$ et une condition initiale $X^0 = X_0$, l'approximation X^n de X_{t^n} , au temps $t^n = n\Delta t$, par le schéma de Milstein est définie pour $k = 1, \dots, d$ par*

$$X_k^{n+1} = X_k^n + b_k(t^n, X^n)\Delta t + \sigma_{kk}(t^n, X^n)\Delta W_k^n + \frac{1}{2} \frac{\partial \sigma_{kk}}{\partial x_k}(t^n, X^n) \sigma_{kk}(t^n, X^n) ((\Delta W_k^n)^2 - \Delta t) \quad (3.3.2)$$

pour $\Delta W^n = W_{t^{n+1}} - W_{t^n}$.

Le schéma de Milstein peut être implémenté simplement de la façon qui suit, voir Algorithme 5.

Algorithm 5 Fonction pour le schéma de Milstein avec σ diagonale.

```

1 # EMilstein diag
2 def Milstein(u0,t,N,d,m):
3
4     print("Milstein Diag")
5     # initialisation
6     u = np.zeros((m,np.size(t,0)));
7     u[:,0] = u0;
8
9     # boucle
10    for n in range(1,N):
11        dt = t[n+1]-t[n];
12        dW = np.sqrt(dt)*np.random.normal(0,1,m);
13        dW2 = (np.multiply(dW,dW)-dt)/2;
14        bn = b(u[:,n]);
15        sn = sigma(u[:,n]); dsn = dsigma(u[:,n]);
16        u[:,n+1] = u[:,n]+dt*bn+np.multiply(dW,sn) + np.multiply(np.multiply(dsn,sn),dW2);
17    return u;
```

3.3.2 Erreur d'approximation

Comme pour les schémas d'approximation pour les EDO [6], on se posera la question de la stabilité numérique et de l'erreur d'approximation des méthodes considérées. Nous discuterons de la stabilité (en un sens à préciser en TD.) Dans ce chapitre nous focalisons sur l'erreur d'approximation des schémas présentés dans le paragraphe précédent.

Pour cela nous allons distinguer deux types de convergence reposant sur deux erreurs entre le processus exact et son approximation au temps $T > 0$ t.q. $T = N\Delta t$.

— **La convergence forte** : on considère l'espérance de l'erreur absolue :

$$\mathcal{E}_T := \mathbb{E}(\|X_T - X^N\|).$$

On parle ici de convergence forte car elle est définie par le biais de la norme dans les espaces L^p qui engendre une topologie forte.

— **La convergence faible** : on considère l'erreur absolue entre l'espérance d'une fonction mesurable $\phi : \mathbb{R}^q \rightarrow \mathbb{R}$ du processus exact et de son approximation :

$$e_T := |\mathbb{E}(\phi(X_T)) - \mathbb{E}(\phi(X^N))|.$$

REMARQUE : Lorsque ϕ est **lipschitz**, on peut montrer que la convergence forte implique la convergence faible. En effet par propriété de l'espérance

$$e_T = |\mathbb{E}(\phi(X_T) - \phi(X^N))| \leq K\mathbb{E}(\|X_T - X^N\|) = K\mathcal{E}_T.$$

Maintenant, on donne les ordres de convergence forte et faible associées aux schémas d'Euler-Maruyama et de Milstein. Pour cela, nous supposons que X est solution de (3.1.1) avec les hypothèses suivantes.

(Hyp.1) Les fonctions de dérive et de diffusion sont lipschitziennes, et à croissance linéaire, i.e. pour tout $t \in I$ et $x, y \in \mathbb{R}^q$

$$\|b(t, x) - b(t, y)\| + \|\sigma(t, x) - \sigma(t, y)\|_F \leq K\|x - y\|, \quad \|b(t, x)\| + \|\sigma(t, x)\|_F \leq K(1 + \|x\|).$$

(Hyp.2) Les fonctions b et σ sont Hölder continues d'exposant $\frac{1}{2}$ en temps i.e. pour tout $t, s \in I$ et $x \in \mathbb{R}^q$ on a

$$\|b(s, x) - b(t, x)\| + \|\sigma(t, x) - \sigma(s, x)\|_F \leq K|t - s|^{1/2}.$$

Ces hypothèses seront nécessaires pour effectuer les majorations intervenant dans l'étude d'erreur des schémas. En outre, on rappelle que la première hypothèse assure l'existence et l'unicité de la solution de (3.1.1).

Nous introduisons également $X^{\Delta t}$ le processus continu défini par interpolation (temporelle) \mathbb{P}_1 entre les instants $\{t^n, n \geq 0\}$, par

$$X_s^{\Delta t} = X^n + b^n(s - t^n) + \sigma^n(W_s - W_{t^n}), \quad s \in [t^n, t^{n+1}],$$

où $b^n = b(t^n, X^n)$ et $\sigma^n = \sigma(t^n, X^n)$. Par construction on a $X_{t^n}^{\Delta t} = X^n$ pour $n \geq 0$.

3.3.3 Convergence forte

Définition 3.3.3 (Convergence forte). *Un schéma numérique X^n de pas de temps $\Delta t > 0$ converge fortement à l'ordre $\gamma > 0$ si pour tout $T > 0$, il existe une constante réelle strictement positive $C := C(T) < \infty$ pouvant dépendre de T telle que*

$$\mathcal{E}_T := \mathbb{E}(\|X_T - X^N\|) \leq C\Delta t^\gamma, \quad \Delta t \in (0, 1].$$

Théorème 3.3.4. *Le schéma d'Euler-Maruyama est fortement convergent d'ordre $\boxed{\gamma = \frac{1}{2}}$. Plus précisément sous les hypothèses **(Hyp.1)**-**(Hyp.2)**, il existe une constante réelle strictement positive $\bar{C} < \infty$ t.q.*

$$\sup_{t \in I} \mathbb{E}(\|X_t - X_t^{\Delta t}\|^2) \leq C\Delta t.$$

REMARQUE. Attention, il y a un carré sur le résultat de convergence donnant l'erreur forte en effet celle-ci est donnée en norme L^2 au carré! Le théorème précédent nous donne

$$C\Delta t \geq \sup_{t \in I} \mathbb{E}(\|X_t - X_t^{\Delta t}\|^2) \geq \mathbb{E}(\|X_T - X_T^{\Delta t}\|^2) \geq \mathbb{E}(\|X_T - X^N\|)^2 = \mathcal{E}_T^2$$

par l'inégalité de Lyapunov². Donc $\mathcal{E}_T \leq \sqrt{C\Delta t}$ qui caractérise la convergence forte.

Théorème 3.3.5. *Le schéma de Milstein est fortement convergent d'ordre $\boxed{\gamma = 1}$. Plus précisément sous les hypothèses (Hyp.1)-(Hyp.2) et $\sigma \in \mathcal{C}^{0,1}$ il existe une constante réelle $0 < C < \infty$ t.q.*

$$\sup_{t \in I} \mathbb{E}(\|X_t - X_t^{\Delta t}\|^2) \leq C\Delta t^2.$$

3.3.4 Convergence faible

Définition 3.3.6 (Convergence faible). *Un schéma numérique X^n de pas de temps $\Delta t > 0$ converge faiblement à l'ordre $\beta > 0$ si pour tout $T > 0$ et $\phi : \mathbb{R}^q \rightarrow \mathbb{R}$ dans \mathcal{C}_b^∞ (\mathcal{C}_b^∞ ensemble des fonctions \mathcal{C}^∞ de dérivée bornée), il existe une constante strictement positive $C := C(T) < \infty$ telle que*

$$e_T := |\mathbb{E}(\phi(X_T)) - \mathbb{E}(\phi(X^N))| \leq C\Delta t^\beta, \quad \Delta t \in (0, 1].$$

REMARQUE. La convergence forte \implies la convergence faible avec le même ordre car $\phi \in \mathcal{C}_b^\infty$ implique qu'elle est lipschitzienne. Cela implique que Euler-Maruyama et Milstein ont des ordres de convergence faible respectivement de 1/2 et 1. En réalité, la convergence forte peut impliquer un ordre de convergence faible plus grossier qu'il ne l'est en réalité pour la convergence faible. Cela est vrai en particulier pour le schéma d'Euler-Maruyama.

Théorème 3.3.7. *Supposons que b, σ soient \mathcal{C}_b^∞ ³ et vérifient les hypothèses (Hyp.1)-(Hyp.2) alors le schéma d'Euler-Maruyama est **faiblement** convergent d'ordre $\beta = 1$.*

Théorème 3.3.8. *Supposons que b, σ soient \mathcal{C}_b^∞ par rapport à leur première variable et $\sigma \in \mathcal{C}^{0,1}$ et vérifient les hypothèses (Hyp.1)-(Hyp.2) alors le schéma de Milstein est **faiblement** convergent d'ordre $\beta = 1$.*

3.3.5 Aspects pratiques

Le calcul de l'erreur de convergence forte requiert le calcul de $\mathbb{E}(\|X_t - X_t^{\Delta t}\|^2)$ pour différents pas de temps de Δt . Premièrement, il s'agit d'être capable d'évaluer numériquement l'espérance de l'erreur à la solution $\|X_t - X_t^{\Delta t}\|^2$ pour Δt donné. Le second point délicat est de savoir calculer cette solution quand bien même la solution exacte X ne soit pas accessible.

① **Nous supposons pour commencer que la solution exacte est accessible.** Pour estimer $\sup_t \mathbb{E}(\|X_t - X_t^{\Delta t}\|^2)$ on utilise une approximation de **Monte-Carlo** et on prend le sup pour les instant du maillage temporel ainsi on a

$$\sup_t \mathbb{E}(\|X_t - X_t^{\Delta t}\|^2) \approx \sup_{0 \leq n \leq N} \frac{1}{M} \sum_{i=1}^M \|X_{t^n}(\omega_i) - X^n(\omega_i)\|^2 =: \mathcal{E}_M$$

où les $X_{t^n}(\omega_i), X^n(\omega_i)$ sont des réalisations indépendantes de $X_t, X_t^{\Delta t}$ **simulées pour les mêmes trajectoires indépendantes du mouvement brownien** $W(\omega_i)$. Il faut garder à l'esprit qu'une approximation de Monte-Carlo converge très lentement avec un taux en $O(\frac{1}{\sqrt{M}})$. Il existe un certain nombre de stratégies pour améliorer ce taux : échantillonnage antithétique, méthode de Monte-Carlo Multi-niveau (voir [2, 4]) que nous ne détaillerons pas ici. On supposera que M est choisi suffisamment grand pour négliger cette erreur.

Nous proposons de calculer \mathcal{E}_M^n pour la résolution d'une équation différentielle régissant un mouvement brownien géométrique, de sorte à disposer d'une solution exacte. Nous résolvons le problème de l'exemple

2. Soit X une v.a. dans $L^s(\Omega)$ et les entiers $0 < r < s < \infty$ alors $\mathbb{E}(|X|^r)^{1/r} \leq \mathbb{E}(|X|^s)^{1/s}$. En particulier si $r = 1, s = 2$ on a $\mathbb{E}(|X|)^2 \leq \mathbb{E}(|X|)$.

3. On peut relaxer cette hypothèse et demander moins de régularité i.e. $\mathcal{C}^{1,4}$ de dérivée bornée (voir [2]).

3.1.2. Pour les paramètres $b = -1, \sigma = 2$ et $N \in \{50, 100, 200, 400, 500, 800, 1000\}$ de sorte que la méthode d'Euler-Maruyama reste stable (voir TD et [4, §8]). Nous avons fixé le nombre de réalisations à $M = 1000$. Les résultats numériques obtenus pour les schémas d'Euler-Maruyama et Milstein sont représentés sur la figure 3.3.1. A gauche on obtient une réalisation de les solutions numériques pour $N = 100$ sur l'intervalle $I = [0, 1]$ comparée à la solution exacte, à droite nous avons tracé l'erreur en fonction de Δt . Primo les solutions se superposent, deuzio nous retrouvons les ordres théoriques attendus.

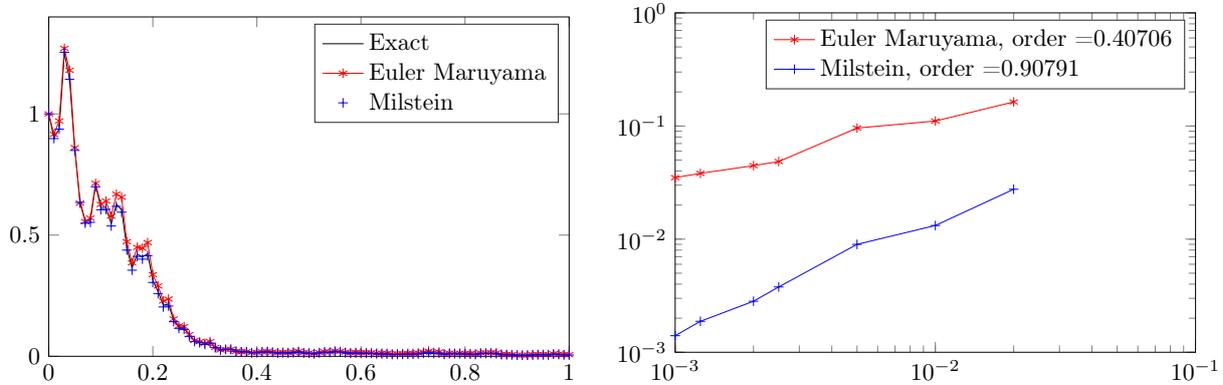


FIGURE 3.3.1 – Exemple du mouvement brownien géométrique : (gauche) trajectoire de X pour $I = [0, 10]$ obtenue pour $N = 100$, (droite) erreurs de convergence en fonction du pas de temps Δt pour $M = 1000$.

② **Désormais nous supposons que la solution exacte n'est pas accessible.** Dans ce cas on utilise une solution de référence notée X^{ref} obtenue pour un maillage très fin de pas $\Delta t^{ref} = \frac{T}{N_{ref}}$ pour une réalisation du brownien donnée ainsi on remplace les $X_{t^n}(\omega_i)$ par $X^{n,ref}(\omega_i)$ dans les formules précédentes, i.e.

$$\sup_t \mathbb{E} (\|X_t - X_t^{\Delta t}\|^2) \approx \sup_{0 \leq n \leq N} \frac{1}{M} \sum_{i=1}^M \|X^{n,ref}(\omega_i) - X^n(\omega_i)\|^2 =: \mathcal{E}_{ref,M}$$

L'idée est alors de calculer l'erreur pour différents maillages grossiers déduits du maillage fin en choisissant $\Delta t = \kappa \Delta t^{ref}$ avec $\kappa \in \mathbb{N}$, de sorte que la solution de référence en tout instant t^n du maillage grossier coïncide avec la solution de référence déjà calculée pour le maillage fin. Il en est de même pour le brownien discret, pour une réalisation du brownien discret on déduira le brownien pour n'importe quel instant $t^n = n\Delta t$ du maillage grossier en choisissant

$$W_{t^{n+1}} = W_{t^n} + (W_{(n+1)\kappa\Delta t^{ref}} - W_{n\kappa\Delta t^{ref}})$$

Cette fois-ci on considère de nouveau le mouvement brownien géométrique, pour lequel on va utiliser une solution de référence obtenue pour un maillage avec $N_{ref} = 2000$. Ici on choisit $\kappa \in \{40, 20, 10, 5, 4, 2\}$. De nouveau on retrouve les résultats théoriques attendus voir figure 3.3.2

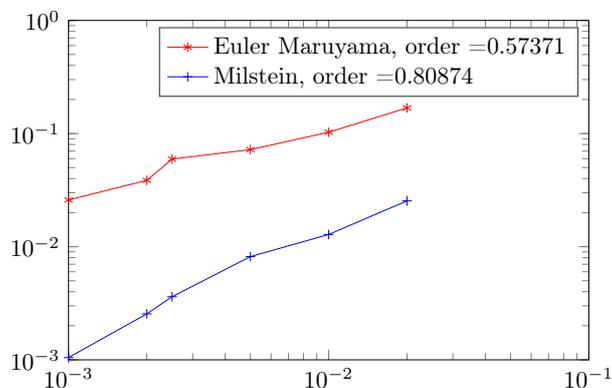


FIGURE 3.3.2 – Exemple du mouvement brownien géométrique : (gauche) trajectoire de X pour $I = [0, 10]$ obtenue pour $N = 100$, (droite) erreurs de convergence en fonction du pas de temps Δt pour $M = 1000$.

Chapitre 4

REPRÉSENTATION PROBABILISTE DES EDP (LIEN EDP-EDS)

4.1 Equation de la chaleur

4.1.1 Un résultat formel

Dans le paragraphe qui suit, nous allons faire le lien (de façon assez formelle) entre l'équation de la chaleur 1D (en domaine non borné) et un mouvement brownien. Nous considérons dans le cas mono-dimensionnel. Soit W un m.b.s. et $x \in \mathbb{R}$ donné. A l'instant $t > 0$, la variable aléatoire $x + W_t$ est de loi gaussienne de moyenne x et de variance t . Ainsi sa densité de probabilité est donnée par

$$g(y; t, x) := \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(y-x)^2}{2t}\right), \quad y \in \mathbb{R}.$$

Par ailleurs un calcul direct donne

$$\begin{aligned} \partial_t g(y; t, x) &= -\frac{1}{2} \frac{1}{\sqrt{2\pi t^3}} \exp\left(-\frac{(y-x)^2}{2t}\right) + \frac{1}{\sqrt{2\pi t}} \frac{(y-x)^2}{2} \exp\left(-\frac{(y-x)^2}{2t}\right), \\ &= -\frac{1}{2\sqrt{2\pi t^3}} \left(1 - \frac{(y-x)^2}{t}\right) \exp\left(-\frac{(y-x)^2}{2t}\right). \end{aligned}$$

et

$$\partial_x g(y; t, x) = \frac{(y-x)}{\sqrt{2\pi t^3}} \exp\left(-\frac{(y-x)^2}{2t}\right),$$

ainsi

$$\partial_{xx}^2 g(y; t, x) = -\frac{1}{\sqrt{2\pi t^3}} \exp\left(-\frac{(y-x)^2}{2t}\right) + \frac{(y-x)^2}{\sqrt{2\pi t^5}} \exp\left(-\frac{(y-x)^2}{2t}\right) = 2\partial_t g(y; t, x).$$

Pour toute fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ mesurable et bornée il vient

$$\int_{\mathbb{R}} \partial_t g(y; t, x) f(y) dy = \frac{1}{2} \int_{\mathbb{R}} \partial_{xx}^2 g(y; t, x) f(y) dy,$$

en notant $u(t, x) = \mathbb{E}(f(x + W_t)) = \int_{\mathbb{R}} g(y; t, x) f(y) dy$, on obtient en appliquant le théorème de dérivation des intégrales à paramètres de Lebesgue la relation

$$\partial_t u(t, x) = \frac{1}{2} \partial_{xx}^2 u(t, x), \quad u^0(x) = f(x). \quad (4.1.1)$$

Notons que pour $t \rightarrow 0^+$, on a la loi normale impropre où la densité de probabilité coïncide avec la mesure de Dirac au point x i.e.

$$\lim_{t \rightarrow 0^+} g(\cdot; t, x) = \delta_x.$$

4.1.2 Une première formule de Feynman-Kac

Etant donnée $f : \mathbb{R} \rightarrow \mathbb{R}$, considérons le problème de Cauchy suivant, associé à l'équation de la chaleur :

$$\begin{cases} \partial_t u(t, x) &= \frac{1}{2} \partial_{xx}^2 u(t, x), t > 0, \quad x \in \mathbb{R}, \\ u(0, x) &= f(x). \end{cases} \quad (4.1.2)$$

On se place pour la suite dans $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilités équipé d'une filtration $\mathcal{F} = \{\mathcal{F}_t\}_{t \in I}$ t.q. \mathcal{F}_0 est complète. Dans la suite, on utilise W un m.b.s. t.q. adapté à la filtration \mathcal{F} .

Proposition 4.1.1 (Feynman-Kac pour la chaleur). *Soit $f \in \mathcal{C}^0(\mathbb{R})$. Supposons qu'il existe une solution $u \in \mathcal{C}^{1,2}([0, T] \times \mathbb{R}, \mathbb{R})$ du problème de Cauchy défini par (4.1.2) telle que, pour tout $t \in [0, T]$ et $x \in \mathbb{R}$ on ait*

$$(\partial_x u(s, W_s^x))_{s \in [0, t]} \in \Lambda^2([0, t]),$$

alors pour tout $(t, x) \in [0, T] \times \mathbb{R}$, la solution s'écrit

$$u(t, x) = \mathbb{E}(f(W_t^x)).$$

Ici W^x , t.q. $W_t^x = W_t + x$, est le processus brownien de C.I. non nulle, i.e. partant x à l'instant $t = 0$.

Démonstration. Pour $t > 0$ fixé et $s \in [0, t]$. On peut appliquer la formule de Itô au processus

$$Y_s = \phi(s, W_s) = u(t - s, W_s^x)$$

car ϕ suffisamment régulière par propriétés de u et $W_s^x = W_s + x$. Alors on a

$$\begin{aligned} dY_s &= \partial_s \phi(s, W_s) ds + \partial_x \phi(s, W_s) \cdot dW_s + \frac{1}{2} \partial_{xx}^2 \phi(s, W_s) ds, \\ &= \left(\cancel{-\partial_t u(t-s, W_s^x) ds} + \partial_x u(t-s, W_s^x) \cdot dW_s + \frac{1}{2} \cancel{\partial_{xx}^2 u(t-s, W_s^x) ds} \right) \end{aligned}$$

Cependant par l'EDP, on peut annuler un certain nombre de termes, ce qui donne

$$dY_s = \partial_x u(t-s, W_s^x) \cdot dW_s.$$

En intégrant par rapport à s de 0 à t on finit par obtenir

$$Y_t - Y_0 = \int_0^t \partial_x u(t-s, W_s^x) \cdot dW_s.$$

Or $Y_0 = u(t, W_0^x) = u(t, x)$ et $Y_t = u(t-t, W_t^x) = u(0, W_t^x) = f(x)$ par définition de la condition initiale. Ainsi

$$\Rightarrow f(W_t^x) - u(t, x) = \int_0^t \partial_x u(t-s, W_s^x) \cdot dW_s \Rightarrow u(t, x) = \mathbb{E}(f(W_t^x)).$$

on note que l'intégrale stochastique à droite a du sens par les hypothèses sur la dérivée spatiale. On déduit le résultat en prenant l'espérance et en utilisant la propriété d'espérance nulle de l'intégrale stochastique d'un processus $\Lambda^2([0, T])$. Ici $Y_0 = u(t-t, W_t^x) = u_0(W_t^x) = f(W_t^x)$ où $W_0^x = W_0 + x = x$. ■

REMARQUES.

- Il faut remarquer qu'on vient conditionner l'espérance en imposant que le processus démarre de x autrement dit $W_t^x = W_t + x$ ou encore $u(t, x) = \mathbb{E}(f(W_t^x) | W_0^x = x)$.
- Attention, ce n'est pas un théorème d'existence, mais simplement une propriété de la solution problème de Cauchy qui admet une représentation probabiliste pour peu que celle-ci soit suffisamment régulière !
- On peut généraliser (voir TD) ce résultat en introduisant un terme de dissipation dans l'équation i.e. $\partial_t u = \frac{1}{2} \partial_{xx}^2 u - ku$ avec k est bornée. Dans ce cas on obtient

$$u(t, x) = \mathbb{E} \left(f(W_t^x) \exp \left(- \int_0^t k(W_s^x) ds \right) \right).$$

- Pour faire le lien entre l'équation de la chaleur et le mouvement brownien on a utilisé la densité de $x + W_t$ qui est explicitement donnée par une gaussienne. Dans le cas général, il est possible de faire le lien EDP-EDS par le biais de processus de diffusion régit par une EDS voir e.g. [1, 2, 7].

4.2 Cas général : formule de Feynman-Kac

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilités équipé d'une filtration $\mathcal{F} = \{\mathcal{F}_t\}_{t \in I}$ t.q. \mathcal{F}_0 est complète. On notera $I = [0, T]$ avec $T > 0$. Dans la suite, on utilise $W = (W_1, \dots, W_d)^T$ un m.b.s. d -dimensionnel, t.q. les $W_k, k = 1, \dots, d$ sont adaptés à la filtration \mathcal{F} .

4.2.1 Générateur infinitésimal

Soit $T > 0$, considérons le processus de diffusion X solution

$$\begin{cases} dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, t \in (0, T] \\ X_0 \text{ donnée,} \end{cases} \quad (4.2.1)$$

où $b : [0, T] \times \mathbb{R}^q \rightarrow \mathbb{R}^q$ et $\sigma : [0, T] \times \mathbb{R}^q \rightarrow \mathbb{R}^{q \times d}$.

Définition 4.2.1 (Générateur infinitésimal). Le **générateur infinitésimal** associé à l'EDS (4.2.1) noté \mathcal{L}_t est l'opérateur différentiel donné par

$$\mathcal{L}_t(x) := \frac{1}{2} \sum_{i,j=1}^q (\sigma(t, x)\sigma(t, x)^T)_{ij} \partial_{x_i x_j}^2 + \sum_{i=1}^q (b(t, x))_i \partial_{x_i}.$$

Exemple 4.2.2 (Equation d'advection diffusion). Si X est le processus stochastique tel que

$$dX_t = b(X_t)dt + dW_t, \quad X_0 \text{ donné,}$$

avec $b : \mathbb{R}^q \rightarrow \mathbb{R}^q$ et W un m.b.s. q -dimensionnel, alors le générateur infinitésimal associé est l'opérateur de convection-diffusion

$$\mathcal{L}(x) := \frac{1}{2} \Delta + b(x) \cdot \nabla.$$

4.2.2 Equation rétrograde

Par la suite nous donnons succinctement un cadre général pour faire le lien entre des EDS et EDP paraboliques sur l'intervalle $I = [0, T], T > 0$ en domaine infini \mathbb{R}^q . Dans le cas parabolique on se donnera x une condition de Cauchy au temps T (équation "rétrograde"). Considérons l'équation aux dérivées partielles avec la donnée finale

$$\begin{cases} \partial_t v(t, x) + \mathcal{L}_t(x)v(t, x) - k(x)v(t, x) + g(x) = 0, & t < T, x \in \mathbb{R}^q, \\ v(T, x) = f(x), & x \in \mathbb{R}^q. \end{cases} \quad (4.2.2)$$

Théorème 4.2.3 (Equation parabolique (rétrograde)). Supposons que

1. $f : \mathbb{R}^q \rightarrow \mathbb{R}$ et $k, g : I \times \mathbb{R}^q \rightarrow \mathbb{R}$ soient continues et f continues et k bornée inférieurement,
2. si de plus pour tout $t \in [0, T]$ et $x \in \mathbb{R}$ on ait

$$(\sigma^T(s, X_s^{t,x}) \nabla v(s, X_s^{t,x}))_{s \in [t, T]} \in \Lambda^2([t, T]),$$

3. où $v : I \times \mathbb{R}^q \rightarrow \mathbb{R}$ de classe $\mathcal{C}^{1,2}$ est solution de (4.2.2).

Alors, pour tout $(t, x) \in I \times \mathbb{R}^d$, la solution v admet la représentation probabiliste suivante

$$v(t, x) = \mathbb{E} \left(f(X_T^{t,x}) e^{-\int_t^T k(X_s^{t,x}) ds} + \int_t^T g(X_s^{t,x}) e^{-\int_t^s k(X_r^{t,x}) dr} ds \right).$$

où $X^{t,x}$ est le processus de diffusion partant de x à l'instant t .

Démonstration. Procéder comme pour l'équation de la chaleur en partant de $v(s, X_s^{t,x}) \exp(-\int_{u=t}^s k(u, X_u^{t,x}) du)$. ■

De même que précédemment, on a énoncé ici un théorème de représentation, voir [7, Proposition 11.2.6] ou encore [1, Théorème 6.6] pour les hypothèses requises pour l'existence.

Pour les EDP à coefficients ne dépendant pas du temps, il est possible de considérer le changement de variable $t \mapsto T - t$ pour retrouver une équation "progressive" i.e. avec une donnée connue au temps initial [1, Chapitre 6.2.2.]. Notons $u(t, x) = v(T - t, x)$ alors formellement on a

$$\begin{cases} \partial_t u(t, x) &= \mathcal{L}(x)u(t, x) - k(x)u(t, x) + g(x), t \in (0, T], \quad x \in \mathbb{R}^q, \\ u(0, x) &= f(x), x \in \mathbb{R}^q. \end{cases} \quad (4.2.3)$$

avec

$$u(t, x) = \mathbb{E} \left(f(X_t^x) e^{-\int_0^t k(X_r^x) dr} + \int_0^t g(X_s^x) e^{-\int_0^s k(X_r^x) dr} ds \right),$$

où X^x est le processus de diffusion translaté issu de x au temps $t = 0$.

4.3 Stratégie numérique

Imaginons que l'on veuille calculer une approximation de u (solution du problème en domaine non borné 4.2.3) admettant la représentation probabiliste suivante

$$u(t, x) = \mathbb{E}(\phi(f, g, k, X)) = \mathbb{E} \left(f(X_t^x) e^{-\int_0^t k(X_r^x) dr} + \int_0^t g(X_s^x) e^{-\int_0^s k(X_r^x) dr} ds \right),$$

pour $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^q$. On peut combiner une méthode de MC avec un schéma de discrétisation adéquat pour approcher les trajectoires du processus X (cf. équation de la chaleur).

Soit un maillage uniforme $0 = t_0 < t_1 < \dots < t_N = T$ de $[0, T]$, $t_n = n\Delta t$, $n = 0, 1, \dots$ et $\Delta t = \frac{T}{N}$, le processus X_t est approché par $X_t^{\Delta t}$ t.q. la v.a. $X_{t_n}^{\Delta t} = X^n$ en tout instant t^n où X^n est l'approximation calculée par le schéma d'Euler-Maruyama. (On peut interpréter la trajectoire du processus discret comme une façon de discrétiser en espace.) On peut calculer une approximation de $\phi(f, g, k, X)$ via

$$\phi(f, g, k, X) \approx \phi_{\Delta t}(f, g, k, X^{\Delta t}) = f(X_T^{\Delta t}) e^{-\Delta t \sum_{j=0}^{N-1} k(X^j)} + \Delta t \sum_{i=0}^{N-1} g(X^i) e^{-\Delta t \sum_{j=0}^{i-1} k(X^j)}.$$

Ici on approche les intégrales en supposant que les processus sont constants sur $[t^n, t^{n+1}]$. Puis on calcule l'estimateur $u_M^{\Delta t}(t, x)$ de $u(t, x)$ en estimant l'espérance via un MC

$$\mathbb{E}(\phi(f, g, k, X)) \approx \frac{1}{M} \sum_{m=1}^M \phi_{\Delta t}(f, g, k, X^{\Delta t}(\omega_m))$$

où $\phi_{\Delta t}(f, g, k, X^{\Delta t}(\omega_m))$ est la m -ième réalisation $\phi_{\Delta t}(f, g, k, X^{\Delta t})$ (i.e. associée à $X^{\Delta t}(\omega_m)$ la m -ième réalisation de $X^{\Delta t}$). L'erreur que l'on commet en utilisant cette méthode est la suivante

$$\begin{aligned} \mathbb{E}(\phi(f, g, k, X)) - \frac{1}{M} \sum_{m=1}^M \phi_{\Delta t}(f, g, k, X^{\Delta t}(\omega_m)) &= \underbrace{\mathbb{E}(\phi(f, g, k, X)) - \mathbb{E}(\phi_{\Delta t}(f, g, k, X^{\Delta t}))}_{\text{erreur d'intégration}} \\ &+ \underbrace{\mathbb{E}(\phi_{\Delta t}(f, g, k, X^{\Delta t})) - \frac{1}{M} \sum_{m=1}^M \phi_{\Delta t}(f, g, k, X^{\Delta t}(\omega_m))}_{\text{erreur statistique}}. \end{aligned}$$

Il y a deux niveaux d'approximation introduisant deux types d'erreurs :

- l'une qui introduit un **erreur statistique** du au nombre de réalisations effectuée, mais qui converge par la loi des grands nombres quand $M \rightarrow \infty$.
- l'autre qui est liée à l' **erreur de discrétisation** pour le processus X , qui est sensée décroître lors que $\Delta t \rightarrow 0$.

En pratique, l'idée est de choisir de façon optimale $\Delta t, M$ pour que les deux types d'erreur que l'on commet soient du même ordre !

Exemple 4.3.1 (Application à l'équation de la chaleur 1D). On cherche à estimer u solution de (4.1.2) (pour $d = 1$) pour un nuage de point $\{x_i\}_{i=1}^N$ dans l'intervalle $[a, b]$. En tout point (x_i, t^n) , on a

$$u(x_i, t^n) \approx \frac{1}{M} \sum_{m=1}^M f(x_i + W^n(\omega_m)).$$

Nous illustrons les cas où $u^0(x) = \cos(x)$ pour lequel $u(t, x) = \exp(-\frac{1}{2}t) \cos(x)$ (courbes du haut de la Figure 4.3.1) et $u^0(x) = x^2$ pour lequel $u(t, x) = x^2 + t$ (courbes du bas de la Figure 4.3.1) à $t = 1$ pour le premier cas et $t = 0.1$ pour le second. On a utilisé 5000 réalisations du mouvement brownien et un maillage temporel contenant 501 noeuds.

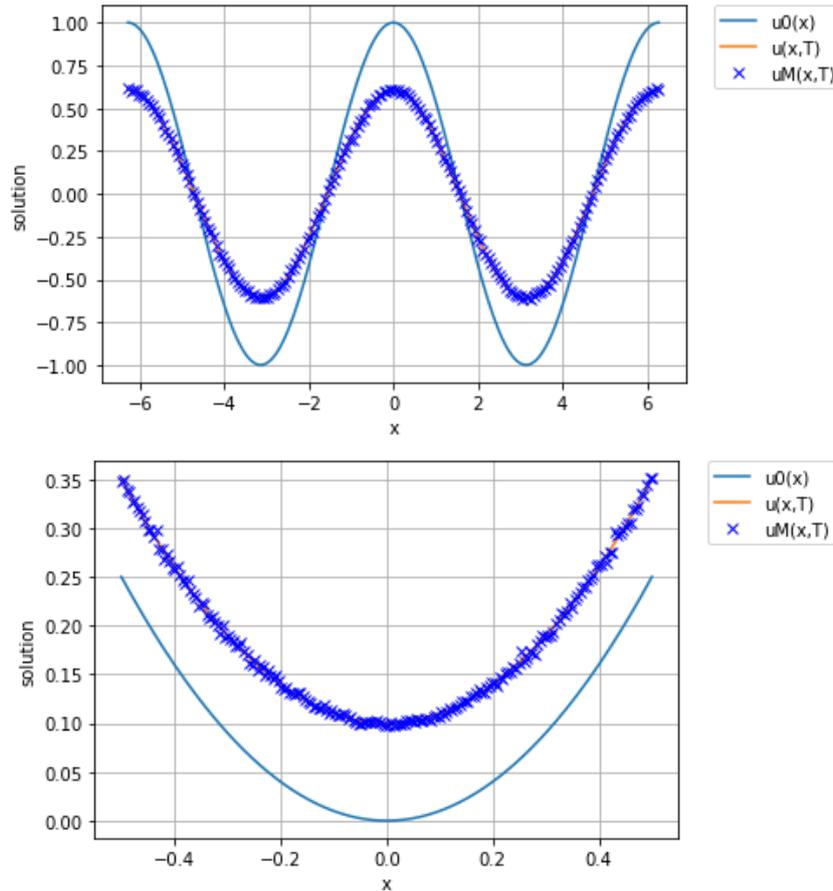


FIGURE 4.3.1 – Comparaison solution exacte u et son estimation $u_M^{\Delta t}$.

VERS DES MÉTHODES DE SIMULATION "PROBABILISTES"

On peut alors s'interroger sur l'intérêt de cette approche par rapport à une méthode de résolution déterministe par schéma de discrétisation de type EF, VF, DF...

Méthodes déterministes	Méthodes de type Monte Carlo
<ul style="list-style-type: none"> ✓ Evaluation globale de la solution u ✓ Convergence liée à $\Delta t, \Delta x$ ✓ Maillage fin ✗ Grande dimension (cond., CFL, mal. dim.) 	<ul style="list-style-type: none"> ✗ Evaluation ponctuelle (regression, interpolation) ✗ Convergence lente M.C. ✓ Convergence inconditionnelle ✓ Complexité \sim pour le linéaire en grande dimension

Une méthode probabiliste ne permet que d'accéder à des valeurs ponctuelles de la solution et non dans sa globalité (il faut alors envisager une reconstruction globale via une méthode d'interpolation ou de moindres

carrés). Par ailleurs, la méthode de MC peut converger très lentement en fonction de M en effet dans ce cas l'erreur au carré se comporte comme $\mathbb{V}(u_M^{\Delta t}(x, t))/M$ et elle dépend de surcroît de la variance de l'estimateur. On peut introduire des méthodes de réduction de variance : méthode d'échantillonnage antithétique, méthodes adaptatives (voir MNP ou [2]) ou encore des méthodes de MC multi-niveaux.

Chapitre 5

NOTATIONS ET CONVENTIONS

5.1 Abréviations

- *v.a.* : variable aléatoire
- *i.i.d.* : indépendantes et identiquement distribuées
- *m.b.s.* : mouvement brownien standard
- *p.s.* presque sûrement
- *p.p.* presque partout

5.2 Notations

- $(\Omega, \mathcal{A}, \mathbb{P})$ espace de probabilité avec Ω univers, \mathcal{M} tribu et \mathbb{P} mesure de probabilité.
- Etant donnés x, y dans \mathbb{R}^d , $\|x\| = \left(\sum_{i=1}^d x_i^2\right)^{1/2}$ est la norme canonique de x induite par le produit scalaire $\langle x, y \rangle = \sum_{i=1}^d x_i y_i$.
- Etant donnés A, B dans $\mathbb{R}^{n \times m}$, $\|A\|_F = \left(\sum_{i=1}^n \sum_{j=1}^m A_{ij}^2\right)^{1/2}$ est la norme de Frobenius de A induite par le produit scalaire $\langle A, B \rangle_F = \sum_{i=1}^n \sum_{j=1}^m A_{ij} B_{ij}$.
- X v.a. de loi $\mathcal{L}(\theta)$ est noté $X \sim \mathcal{L}(\theta)$.
- Une suite de v.a. $\{X_n\}_{n \geq 0}$ convergente en loi vers $\mathcal{L}(\theta)$ est notée $X_n \underset{n \rightarrow \infty}{\rightsquigarrow} \mathcal{L}(\theta)$. Soit $X \sim \mathcal{L}(\theta)$ on note aussi la convergence en loi $X_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} X$.
- Une suite de v.a. $\{X_n\}_{n \geq 0}$ convergente en probabilité vers X est notée $X_n \xrightarrow[n \rightarrow \infty]{P} X$.
- Une suite de v.a. $\{X_n\}_{n \geq 0}$ convergente presque sûrement vers X est notée $X_n \xrightarrow[n \rightarrow \infty]{p.s.} X$.
- Une suite de v.a. $\{X_n\}_{n \geq 0}$ convergente presque en moment d'ordre 2 vers X est notée $X_n \xrightarrow[n \rightarrow \infty]{L^2} X$ ou de façon équivalente $\mathbb{E}(\|X_n - X\|^2) \xrightarrow[n \rightarrow \infty]{} 0$.

5.3 Tribus et variables aléatoires

Soit Ω un ensemble et \mathcal{F} un ensemble de sous ensembles de Ω est une **tribu** sur Ω si :

- $\emptyset \in \mathcal{F}$,
- si $A \in \mathcal{F}$ alors $\Omega \setminus A \in \mathcal{F}$,
- si $\{A_n\}_{n \geq 1} \in \mathcal{F}$ alors $\cup_{n \geq 1} A_n \in \mathcal{F}$.

Une **sous-tribu** \mathcal{E} de \mathcal{F} est une tribu telle que $\mathcal{E} \subset \mathcal{F}$.

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilité où Ω est un ensemble, \mathcal{M} une tribu définie sur Ω et \mathbb{P} une mesure de probabilités. Une **variable aléatoire (v.a.)** à valeurs réelles $X : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, $d \geq 1$, est une **fonction mesurable** c-à-d que pour tout borélien $B \in \mathcal{B}(\mathbb{R}^d)$ l'ensemble

$$X^{-1}(B) = \{\omega \in \Omega, X(\omega) \in B\}$$

aussi noté $\{X \in B\}$ appartient à \mathcal{M} . On note $\mathbb{E}(X)$ l'**espérance de X** définie par

$$\mathbb{E}(X) = \int_{\Omega} X d\mathbb{P}.$$

La **tribu engendrée par X** est la tribu notée $\sigma(X)$ définie par

$$\sigma(X) = \{X \in B, B \in \mathcal{B}(\mathbb{R}^d)\} \subset \mathcal{M}$$

c'est la plus petite sous tribu de \mathcal{M} rendant X mesurable.

Espérance conditionnelle

Soit X une variable aléatoire à valeurs réelles définie sur l'espace de probabilités $(\Omega, \mathcal{A}, \mathbb{P})$ et $\mathcal{F} \subset \mathcal{M}$ une sous tribu. On suppose que X est intégrable, i.e. $\mathbb{E}(|X|) < \infty$.

Alors **l'espérance conditionnelle de X par rapport à la tribu \mathcal{F}** est la variable aléatoire Y \mathcal{F} -mesurable t.q.

$$\mathbb{E}(XU) = \mathbb{E}(YU) \text{ pour toute v.a. } U \text{ } \mathcal{F}\text{-mesurable.}$$

On note la variable aléatoire $Y = \mathbb{E}(X|\mathcal{F})$. On rappelle les propriétés de l'espérance conditionnelle par rapport à la sous tribu $\mathcal{F} \subset \mathcal{M}$. et pour tout variable aléatoire X .

1. $\mathbb{E}(\mathbb{E}(X|\mathcal{F})) = \mathbb{E}(X)$.
2. Pour tout a, b réels on a $\mathbb{E}(aX + bY|\mathcal{F}) = a\mathbb{E}(X|\mathcal{F}) + b\mathbb{E}(Y|\mathcal{F})$.
3. Si $\mathcal{E} \subset \mathcal{F}$ alors $\mathbb{E}(\mathbb{E}(X|\mathcal{F})|\mathcal{E}) = \mathbb{E}(X|\mathcal{E})$ et $\mathbb{E}(\mathbb{E}(X|\mathcal{E})|\mathcal{F}) = \mathbb{E}(X|\mathcal{E})$.
4. Si $X \leq Y$ p.s. alors $\mathbb{E}(X|\mathcal{F}) \leq \mathbb{E}(Y|\mathcal{F})$.
5. Si X est une variable aléatoire intégrable et indépendante de \mathcal{F} alors $\mathbb{E}(X|\mathcal{F}) = \mathbb{E}(X)$.

Par ailleurs, étant données deux variables aléatoires X, Y définies sur $(\Omega, \mathcal{A}, \mathbb{P})$. On définit **l'espérance conditionnelle de Y par rapport à X** comme l'espérance conditionnelle de Y par rapport à la tribu $\sigma(X)$ i.e.

$$\mathbb{E}(Y|X) = \mathbb{E}(Y|\sigma(X)).$$

Les propriétés précédentes sont toujours vérifiées et on a pour tout Z $\sigma(X)$ -mesurable,

1. $\mathbb{E}(Z|X) = \mathbb{E}(Z|\sigma(X)) = Z$.
2. $\mathbb{E}(ZY|X) = \mathbb{E}(ZY|\sigma(X)) = Z\mathbb{E}(Y|\sigma(X))$.

BIBLIOGRAPHIE

- [1] Francis Comets and Thierry Meyre. Calcul stochastique et modèles de diffusions. May 2020.
- [2] E. Gobet. Méthodes de Monte-Carlo et processus stochastiques : du linéaire au non-linéaire. Editions de l'Ecole Polytechnique, 2013.
- [3] Peter E. Kloeden and Eckhard Platen. Numerical Solution of Stochastic Differential Equations. Springer, Berlin, Germany, 1992.
- [4] Gabriel J. Lord, Catherine E. Powell, and Tony Shardlow. An Introduction to Computational Stochastic PDEs. Cambridge University Press, Cambridge, England, UK, July 2014.
- [5] Bernt Øksendal. Stochastic Differential Equations. Springer, Berlin, Germany, 1995.
- [6] Alfio Quarteroni, Riccardo Sacco, and Fausto Saleri. Numerical Mathematics. Springer, Berlin, Germany, 2007.
- [7] Julien Reygner. Méthodes numériques probabilistes, Notes de cours. 2024.