

MODELES DE DIFFUSION

J. NEVEU & E. PARDOUX

Chapitre I

Fonctions aléatoires et Mouvement brownien

1.0. Introduction.

Les fonctions aléatoires, notamment les fonctions aléatoires gaussiennes et leurs développements orthogonaux sont introduits dans les deux premiers paragraphes en prélude à la définition du mouvement brownien donnée au paragraphe 3.

Ce mouvement brownien joue un rôle central dans la théorie des processus aléatoires, notamment parce que dans de nombreux problèmes théoriques ou appliqués, le mouvement brownien ou les diffusions que l'on en déduit, fournissent des modèles limites simples, sur lesquels de nombreux calculs peuvent être faits.

Les trajectoires d'un mouvement brownien bien que continues sont très irrégulières, comme il est expliqué au paragraphe 4. Le paragraphe suivant relie le mouvement brownien et le processus d'Ornstein Uhlenbeck qui s'en déduit par une intégrale stochastique, à l'équation de l'oscillateur harmonique soumis à une force aléatoire (équation de Langevin).

Le paragraphe 6 est consacré à un problème de changement de probabilités, tout à fait fondamental en Statistique des processus, qui conduit à la formule de Girsanov.

Un bref historique sur le mouvement brownien termine ce chapitre introductif.

1.1. Généralités sur les processus aléatoires.

Par définition une *fonction aléatoire réelle* (f.a.r. en abrégé) définie sur un intervalle réel T est une fonction de T dans R (ou R^d) qui dépend du hasard, c'est-à-dire qui dépend d'un ω (« petit oméga ») variant dans un espace de probabilité donné (Ω, \mathcal{A}, P) . C'est donc une fonction de *deux* arguments que nous écrirons

$$t, \omega \rightarrow X(t, \omega) \quad (t \in T \subset R, \omega \in \Omega)$$

Processus aléatoire, signal aléatoire sont des synonymes.

EXEMPLE. Si $A_1, \dots, A_n, \Phi_1, \dots, \Phi_n$ sont n v.a.r. positives et n angles aléatoires définis sur Ω , la formule

$$S(t, \omega) = \sum_1^n A_m(\omega) \sin [2\pi f_m t + \Phi_m(\omega)]$$

définit un signal aléatoire réel dépendant du temps réel $t \in R$, qui est une superposition de sinusoides de fréquences f_m fixes ($f_m \in R_+$), mais de phases et d'amplitudes aléatoires. Les signaux aléatoires de ce type (éventuellement un peu généralisés) servent à étudier par exemple les phénomènes de bruit dans la transmission de données, ou en Economie des phénomènes oscillatoires de consommation, etc.

L'observation d'une fonction aléatoire revient à choisir ω dans Ω , donc revient dans l'exemple précédent à se fixer les valeurs des A_m et Φ_m . On appelle alors réalisation ou plutôt *trajectoire* d'une fonction aléatoire X , toute fonction

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

obtenue en fixant ω et que l'on est donc susceptible d'observer. Il est évidemment naturel de supposer que les trajectoires d'une f.a.r. sont des fonctions « régulières »; nous ne considérerons ainsi dans ce cours que des f.a.r. à *trajectoires continues*. (Par contre la fonction aléatoire de comptage du processus de Poisson [Cours de probabilités, §2.5] possédait des « trajectoires en escaliers » discontinues).

La valeur à l'instant t ($t \in T$) de la fonction aléatoire X est l'application de Ω dans R ou R^d

$$\omega \rightarrow X(t, \omega) \quad (t \in T \text{ fixé})$$

que l'on notera simplement $X(t)$; c'est une variable aléatoire pourvu que cette application soit mesurable, ce que l'on supposera toujours. Les lois de probabilité des variables aléatoires $X(t)$ et plus généralement celles des vecteurs aléatoires

$$(X(t_1), \dots, X(t_k)) \quad (t_1 < \dots < t_k \quad \text{dans } T)$$

vont jouer un rôle primordial; leur donnée permettra en effet de calculer les espérances de la forme

$$\int_{\Omega} F[X(\cdot, \omega)] \, dP(\omega)$$

où F désigne une application de l'espace des trajectoires dans R ($F[x(\cdot)] = \sup_T x(t)$ ou $\int_T x(t)^2 dt$ par exemple) en approchant ces fonctionnelles F par des fonctionnelles ne dépendant que d'un nombre fini de $x(t)$ ($\sup_k x(t_k)$ ou $\sum_k (t_{k+1} - t_k) x(t_k)^2$ pour des suites (t_k) bien choisies, par exemple).

Résumons ce qui précède en la

Définition 1.1.1.

Une **fonction aléatoire réelle continue**, définie sur l'intervalle réel T et l'espace de probabilité (Ω, \mathcal{A}, P) est une application

$$(t, \omega) \rightarrow X(t, \omega) \quad \text{de } T \times \Omega \quad \text{dans } R \text{ (ou } R^d)$$

telle que :

- a) ses trajectoires $t \rightarrow X(t, \omega)$ soient continues sur T ($\omega \in \Omega$),
- b) ses valeurs $\omega \rightarrow X(t, \omega)$ soient des v.a. sur Ω ($t \in T$).

La loi temporelle de cette f.a.r. est définie pour la donnée des lois de probabilité de tous les vecteurs aléatoires

$$(X(t_1), \dots, X(t_k)) \quad (t_1 < \dots < t_k \quad \text{dans } T).$$

1.2. Espaces gaussiens.

Une v.a.r. X est dite gaussienne (= normale) si elle suit une densité de la forme

$$g_{\mu, \sigma}(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp[-(x - \mu)^2 / 2\sigma^2] \quad (x \in R) \quad (\mu \in R, \sigma > 0)$$

auquel cas $E(X) = \mu$, $\text{Var}(X) = \sigma^2 > 0$ ou si elle est p.s. constante soit $X = \mu$ p.s.,
 auquel cas $E(X) = \mu$ et $\text{Var}(X) = 0$. Pour une telle v.a., quel que soit $a \in \mathbb{C}$:

$$E[\exp(aX)] = \exp[aE(X) + \frac{1}{2}a^2 \text{Var}(X)] ;$$

inversement il suffit déjà que cette dernière formule soit exacte pour tout a réel, resp. pour tout a imaginaire pour que X soit gaussienne (Dans le 2^e cas, cela résulte de l'injectivité de la transformation de Fourier; dans le 1^{er} cas, on étend le domaine de validité de la formule à tout $a \in \mathbb{C}$ par prolongement analytique et on est alors ramené au 2^e cas).

Un vecteur aléatoire d -dimensionnel X est dit gaussien si toute combinaison linéaire $\sum_i^d c_i X_i$ de ses coordonnées X_i ($1 \leq i \leq d$) est une v.a.r. gaussienne au sens de l'alinéa précédent ou, ce qui est équivalent, si et seulement si pour tout $a \in \mathbb{R}^d$

$$(1.2.1) \quad E[\exp(a \cdot X)] = \exp[a \cdot E(X) + \frac{1}{2} a^t K(X) a]$$

où $a \cdot X = \sum_1^d a_i X_i$, $a \cdot E(X) = \sum_1^d a_i E(X_i)$ et où $a^t K(X) a$ désigne la forme quadratique $\sum_i \sum_j a_i a_j K_{ij}(X)$ associée à la matrice de covariance ($K_{ij}(X) = \text{Cov}(X_i, X_j)$; $1 \leq i, j \leq d$) du vecteur X . En particulier un vecteur aléatoire X dont les coordonnées X_i sont des v.a.r. gaussiennes réelles, *indépendantes* entre elles est un vecteur gaussien de covariance diagonale et réciproquement.

Exercice. Si X^1, \dots, X^k sont k ($k \geq 2$) vecteurs gaussiens indépendants entre eux, montrer que le vecteur aléatoire X formé par les coordonnées des X^i ($\dim(X) = \dim(X^1) + \dots + \dim(X^k)$) est encore gaussien et caractériser les vecteurs gaussiens ainsi obtenus par la forme particulière de leur covariance.

Une fonction aléatoire réelle ($X(t)$, $t \in T$) est dite gaussienne si tout vecteur fini-dimensionnel ($X(t_1), \dots, X(t_d)$) qui en est extrait est gaussien, donc si toute v.a.r.

$\sum_1^d a_i X(t_i)$ de l'espace vectoriel H_0 engendré par les $X(t)$ ($t \in T$) est gaussienne. Puisque les v.a.r. gaussiennes sont de carrés intégrables, il est naturel de se placer dans l'espace $L^2(\Omega)$ de toutes les v.a.r. de carrés intégrables définies sur l'espace de probabilité (Ω, \mathcal{A}, P) sous-jacent; cet espace est muni du produit scalaire

$$\langle X, Y \rangle = E(XY)$$

qui en fait un espace de Hilbert.

Ensuite pour utiliser les théorèmes de géométrie hilbertienne (bases orthonormales, projections orthogonales, etc.), il convient de fermer l'espace H_0 introduit ci-dessus. Comme toute limite L^2 de v.a.r. gaussiennes est encore gaussienne [démonstration facile par les fonctions caractéristiques], les v.a.r. de la fermeture de H_0 seront encore gaussiennes. Nous sommes ainsi conduits à poser la définition suivante, en ne considérant toutefois que des v.a.r. centrées.

Définition 1.2.1.

Un sous-espace (= sous-vectoriel fermé) H de l'espace $L^2(\Omega)$ construit sur un espace de probabilité (Ω, \mathcal{A}, P) donné, est appelé **espace gaussien** s'il est formé de v.a.r. gaussiennes centrées.

L'espace gaussien $H[X]$ associé à une f.a.r. gaussienne $(X(t), t \in T)$ est le sous-espace de $L^2(\Omega)$ engendré par les v.a.r. centrées $X(t) - E[X(t)]$ ($t \in T$), c'est-à-dire le sous-espace de $L^2(\Omega)$ formé par les combinaisons linéaires de ces v.a.r. centrées et leurs limites en moyenne quadratique.

Une f.a.r. gaussienne $(X(t), t \in T)$ définie sur un intervalle réel T est dite *continue en moyenne quadratique* si l'application $t \rightarrow X(t)$ de T dans $L^2(\Omega)$ est continue, c'est-à-dire si $E[(X(t) - X(s))^2] \rightarrow 0$ lorsque $t \rightarrow s$ et $s \in T$. On ne considère jamais d'autres fonctions gaussiennes sur T . [Remarque. Une f.a.r. à trajectoires continues est continue en moyenne quadratique lorsqu'elle est gaussienne; *la réciproque est fautive*.] \square

Formule de Loève-Karhunen. Etant donné une f.a.r. gaussienne $(X(t), t \in T)$ définie sur l'intervalle réel T et continue en moyenne quadratique, choisissons une base orthonormale $(\xi_n, n \in N)$ arbitrairement dans l'espace gaussien $H[X]$ associé et développons les v.a.r. $X(t) - E[X(t)]$ de $H(X)$ suivant cette base. La formule de Loève-Karhunen ainsi obtenue :

$$(1.2.2) \quad X(t, \omega) = E[X(t)] + \sum_N c_n(t) \xi_{n(\omega)}$$

avec

$$(1.2.3) \quad c_n(t) = \langle X(t) - E[X(t)], \xi_n \rangle_H = E[X(t) \xi_n] \quad (\text{car } E(\xi_n) = 0),$$

fournit une représentation de la f.a.r. X comme série de produits d'une fonction de t seulement et d'une v.a.r. (fonction de ω seulement).

En tant que suite orthonormale dans $H[X]$, la suite $(\xi_n, n \in N)$ est ici simplement une suite de v.a.r. gaussiennes réduites ($E(\xi_n) = 0, \text{Var}(\xi_n) = E(\xi_n^2) = 1$), orthogonales ($E(\xi_m \xi_n) = 0$ si $m \neq n$) et donc indépendantes. D'autre part la continuité en moyenne quadratique de la f.a.r. X entraîne la continuité des fonctions réelles c_n sur T (par la continuité du produit scalaire et la formule (1.2.3)).

Les convergences des séries de Loève-Karhunen doivent, par définition des développements orthonormaux, être prises de la manière suivante : pour tout $t \in T$ fixé, les approximations finies

$$X^q(t) = E[X(t)] + \sum_0^q c_n(t) \xi_n$$

convergent vers $X(t)$ dans L^2 lorsque $q \nearrow \infty$, soit

$$(1.2.4) \quad E[(X^q(t) - X(t))^2] \rightarrow 0 \quad (q \nearrow \infty, t \in T).$$

Mais ce second moment décroît avec q et est une fonction continue de t ; le lemme de Dini entraîne alors que la convergence précédente est uniforme en t sur tout compact de T (donc sur T si cet intervalle est lui-même compact). *En outre lorsque la f.a. est à trajectoires continues, un théorème de martingales vectorielles permet d'établir que $X^q(t)$ converge p.s. vers $X(t)$ uniformément sur tout compact de T .*

Les résultats précédents sont indépendants de la base $(\xi_n, n \in N)$ choisie dans $H[X]$; par contre la rapidité de la convergence vers 0 dans (1.2.4) dépend de ce choix. D'autre part en prenant une base orthonormale *infinie* $(\xi_n, n \in N)$ dans $H[X]$, nous avons supposé implicitement que l'espace $H[X]$ était de dimension infinie; dans le cas inverse, le développement de Loève-Karhunen se réduirait à une somme finie (ce cas ne se présente pratiquement jamais dans les applications!). On pourra noter aussi que $H[X]$ est nécessairement séparable car la continuité de X entraîne qu'il est engendré par la famille dénombrable de v.a.r. $(X(t) - E[X(t)], t \in Q \cap T)$.

1.3. Mouvement brownien et Intégrale de Wiener.

Définition 1.3.1.

On appelle mouvement brownien réel (sur R_+) une f.a.r. $(B(t), t \in R_+)$ à trajectoires continues telle que :

a) $B(0) = 0$; tout accroissement $B(t) - B(s)$ ($0 \leq s < t$) suit une loi gaussienne centrée de variance $t - s$ (loi notée $N(0, t - s)$);

b) quels que soient $d \geq 2$ et $0 = t_0 < t_1 < \dots < t_d$ dans R_+ , les accroissements $B(t_{i+1}) - B(t_i)$ ($0 \leq i \leq d$) soient indépendants.

Le mouvement brownien tire son importance fondamentale et son caractère universel du théorème de la limite centrale (Cours de probabilités : théorème 4.6.1). En effet, d'après ce théorème, si $(X_m, m \geq 1)$ désigne une suite de v.a.r. indépendantes et équidistribuées, de variance commune σ^2 finie et pour simplifier d'espérances nulles, les sommes

$S_n = \sum_1^n X_m$ renormalisées : $S_n / \sigma\sqrt{n}$ sont approximativement gaussiennes réduites

lorsque $n \rightarrow \infty$. En fait plus généralement, la promenade aléatoire toute entière $(S_n, n \geq 0)$ renormalisée en posant pour tout n (grand)

$$B^{(n)}(t) = S_{[nt]} / \sigma\sqrt{n} \quad (0 \leq t < \infty)$$

tend en loi vers un mouvement brownien au sens où la loi de $(B^{(n)}(t_1), \dots, B^{(n)}(t_d))$ tend vers celle de $(B(t_1), \dots, B(t_d))$ quels que soient $d \geq 2$ et $0 \leq t_1 < \dots < t_d$; on le voit en remarquant que des accroissements de la f.a.r. $B^{(n)}$

$$B^{(n)}(t_{i+1}) - B^{(n)}(t_i) = \sum_{[nt_i] < m \leq [nt_{i+1}]} X_m / \sigma\sqrt{n}$$

sont indépendants et asymptotiquement de lois gaussiennes centrées de variances $t_{i+1} - t_i$, lorsque $n \rightarrow \infty$, par le théorème de la limite centrale. Il est remarquable que ce résultat ne dépende pas de la loi initiale des X_m , pourvu que celle-ci ait un second moment fini et soit centrée.

REMARQUE. L'existence d'un mouvement brownien au sens de la définition précédente n'est pas évidente même si la convergence en loi des f.a.r. $B^{(n)}$ ci-dessus rend cette existence vraisemblable (à la propriété de continuité des trajectoires près). Nous reviendrons sur cette existence ultérieurement. \square

Proposition 1.3.2.

a) Un mouvement brownien est une f.a.r. gaussienne centrée de covariance

$$E [B(s) B(t)] = \min (s, t) \quad s, t \in R_+$$

dont les trajectoires sont continues, et réciproquement.

b) Si $(B(t), t \geq 0)$ est un mouvement brownien, il en est de même de

$$\left(\frac{1}{c} B(c^2 t), t \geq 0\right)$$

pour tout réel $c \neq 0$, ainsi que de

$$(t B(1/t), t \geq 0)$$

DEMONSTRATION. Toute combinaison linéaire de $B(t_i)$ ($0 \leq t_1 < t_2 < \dots < t_d$) est une combinaison linéaire de $B(t_i) - B(t_{i-1})$ (avec $t_0 = 0$) et est donc gaussienne centrée en tant que combinaison linéaire de v.a.r. gaussiennes centrées indépendantes. La covariance de la f.a. B s'obtient en écrivant que si $s \leq t$:

$$E [B(s) B(t)] = E [B(s)^2] + E[B(s) [B(t) - B(s)]] = E [B(s)^2] = s$$

On vérifie alors facilement la réciproque. La deuxième partie de la proposition se déduit également facilement de la première partie. (*Exercices !*) \square

Les accroissements $B(t) - B(s)$ ($0 \leq s < t$) d'un mouvement brownien ont des propriétés plus simples que les v.a. $B(t)$. D'une part en effet la définition du mouvement brownien entraîne que les accroissements $B(v_i) - B(u_i)$ ($u_i < v_i$) sont indépendants dès que les intervalles $]u_i, v_i]$ sont 2 à 2 disjoints. D'autre part la loi de $B(t + h) - B(t)$ ne dépend pas de t mais seulement de h . En outre si $s < t$, l'accroissement $B(t) - B(s)$ est indépendant de $(B(u), 0 \leq u \leq s)$ de sorte que la position $B(t)$ du mouvement brownien à l'instant t ne dépend de $(B(u), u \leq s)$ que par $B(s)$ et s'obtient en ajoutant à $B(s)$ une v.a. indépendante de loi $N(0, t - s)$.

Les accroissements infinitésimaux du mouvement brownien

$$dB(t) = B(t + dt) - B(t)$$

sont heuristiquement des variables gaussiennes centrées de variances dt , donc d'ordre de grandeur \sqrt{dt} ... Ceci laisse prévoir que les trajectoires $t \rightarrow B(t, \omega)$ ne peuvent pas être différentiables ; si les limites

$$\frac{dB(t)}{dt} = \lim_{h \rightarrow 0} \frac{B(t + h) - B(t)}{h}$$

existaient, elles devraient être gaussiennes centrées de variances $\lim_{h \rightarrow 0} (h/h^2) = \infty$... Nous démontrerons des résultats plus précis au paragraphe suivant ; dans le même ordre d'idées, le résultat que voici s'avèrera fondamental bien que relativement simple, dans l'étude de l'intégration stochastique.

Proposition 1.3.3.

Soit T un réel > 0 fixé. Alors les variations quadratiques du mouvement brownien $(B(t), t \geq 0)$:

$$V_T(\{t_0, \dots, t_n\}) = \sum_1^n [B(t_j) - B(t_{j-1})]^2$$

associées à toute suite finie $t_0 = < t_1 < \dots < t_n = T$ et à la partition correspondante de $]0, T]$ en intervalles $]t_{j-1}, t_j]$, convergent en moyenne quadratique vers T , soit

$$E([V_T(\{t_0, \dots, t_n\}) - T]^2) \rightarrow 0$$

lorsque les pas $\max_j (t_j - t_{j-1})$ de ces partitions tendent vers zéro.

Il convient de noter qu'à l'opposé de ce résultat, les variations quadratiques d'une fonction continûment différentiable $f: R_+ \rightarrow R$

$$v_T(\{t_0, \dots, t_n\}) = \sum_1^n [f(t_j) - f(t_{j-1})]^2 \text{ valent encore } \sum_1^n [f'(t_j^*) (t_j - t_{j-1})]^2$$

par le théorème des accroissements finis et tendent donc vers 0 lorsque $\max_j (t_j - t_{j-1}) \rightarrow 0$.

DEMONSTRATION. Le carré X^2 d'une v.a.r. gaussienne centrée X de variance σ^2 est tel que $E(X^2) = \sigma^2$ et que $\text{Var}(X^2) = 2\sigma^4$. Ceci appliqué aux v.a. $B(t_j) - B(t_{j-1})$ entraîne d'une part que

$$E[V_T(\{t_0, \dots, t_n\})] = \sum_1^n (t_j - t_{j-1}) = T$$

et d'autre part que

$$\text{Var}[V_T(\{t_0, \dots, t_n\})] = 2 \sum_1^n (t_j - t_{j-1})^2$$

compte tenu aussi de l'indépendance des accroissements $B(t_j) - B(t_{j-1})$ et donc de leurs carrés. Il s'en suit que

$$\begin{aligned} E([V_T(\{t_0, \dots, t_n\}) - T]^2) &= \text{Var}[V_T(\{t_0, \dots, t_n\})] \\ &\leq 2T \max (t_j - t_{j-1}) \\ &\rightarrow 0 \text{ lorsque } \max (t_j - t_{j-1}) \rightarrow 0. \quad \square \end{aligned}$$

Ainsi la fonction aléatoire $\dot{B}(t) = dB(t)/dt$, dérivée du mouvement brownien n'existe pas. Néanmoins pour les fonctions f de carrés sommables sur R_+ , nous réussirons à définir l'intégrale $\int_{R_+} f(t) \dot{B}(t) dt$ que nous noterons d'ailleurs plutôt $\int_{R_+} f(t) dB(t)$; cette « intégrale stochastique » ($f(t) dB(t)$ est une v.a. comme les $dB(t)$) jouera un rôle primordial dans la suite.

Dans la proposition suivante, nous notons comme d'habitude $L^2(R_+)$ l'espace des fonctions boréliennes $f: R_+ \rightarrow R$ telles que $\int_{R_+} f(t)^2 dt < \infty$, deux fonctions égales p.p. étant identifiées ; cet espace est hilbertien pour le produit scalaire

$$\langle f_1, f_2 \rangle = \int_{R_+} f_1(t) f_2(t) dt .$$

Il convient évidemment de ne pas confondre cet espace avec l'espace $L^2(\Omega)$ des v.a.r. de carrés intégrables !

Théorème 1.3.4. (Intégrale de Wiener).

Etant donné un mouvement brownien $(B(t), t \in R_+)$, on peut associer à toute fonction $f \in L^2(R_+)$ une v.a.r. de carré intégrable notée $\int_{R_+} f(t) dB(t)$, unique à une équivalence près, de telle manière que

a) *pour $f = 1_{[u,v]}$ où $0 \leq u < v < \infty$, on ait $\int_{R_+} f(t) dB(t) = B(v) - B(u)$*

b) *l'application $f \rightarrow \int_{R_+} f(t) dB(t)$ soit linéaire et isométrique de $L^2(R_+)$ dans $L^2(\Omega)$.*

(La propriété d'isométrie s'écrit :

$$E \left[\int_{R_+} f_1(t) dB(t) \int_{R_+} f_2(t) dB(t) \right] = \int_{R_+} f_1(t) f_2(t) dt \quad (f_1, f_2 \in L^2(R_+)) .$$

En outre le sous-espace de $L^2(\Omega)$ formé par les v.a.r. $\int_{R_+} f(t) dB(t)$ ($f \in L^2(R_+)$) coïncide avec l'espace gaussien $H[B]$ associé au mouvement brownien.

Enfin la v.a.r. $\int_{R_+} f(t) dB(t)$ est caractérisée par la double propriété d'appartenir à $H[B]$ et de vérifier pour tout $s \in R_+$ l'égalité

$$E \left[\int_{R_+} f(t) dB(t) \cdot B(s) \right] = \int_0^s f(t) dt .$$

DEMONSTRATION. Notons \mathcal{A} le sous-vectoriel de $L^2(R_+)$ formé des fonctions f s'écrivant sous la forme

$$f = \sum_1^d a_i 1_{[t_{i-1}, t_i]}$$

avec $d \geq 1$, $0 = t_0 < t_1 < \dots < t_d$ et $a_1, \dots, a_d \in R$. A chacune de ces fonctions f associons la v.a.r. notée $B(f)$ définie par

$$B(f) = \sum_1^d a_i [B(t_i) - B(t_{i-1})]$$

(bien que toute fonction $f \in \mathcal{A}$ ait plusieurs représentations de la forme ci-dessus résultant de l'identité $1_{[u,v]} = 1_{[u,w]} + 1_{[w,v]}$ si $u < w < v$, la v.a. $B(f)$ ne dépend manifestement que de f et non de la représentation de f choisie pour la définir, car $B_v - B_u = (B_w - B_u) + (B_v - B_w)$).

Il est immédiat de vérifier que $f \rightarrow B(f)$ est linéaire sur \mathcal{A} ; en outre

$$E[B(f)^2] = \sum_1^d \alpha_i^2 (t_i - t_{i-1}) = \int_{R_+} f(t)^2 dt \quad (f \in \mathcal{A})$$

Comme \mathcal{A} est dense dans $L^2(R_+)$, l'application $f \rightarrow B(f)$ peut ensuite s'étendre par continuité à tout $L^2(R_+)$. En effet tout $f \in L^2(R_+)$ est la limite d'au moins une suite $(f_n, n \in N)$ dans \mathcal{A} ; les v.a. $B(f_n)$ ($n \in N$) correspondantes forment alors une suite de Cauchy car

$$E([B(f_n) - B(f_m)]^2) = E[B(f_n - f_m)^2] = \int_{R_+} (f_n - f_m)^2 dt \rightarrow 0$$

lorsque $m, n \rightarrow \infty$; la limite $\lim_n B(f_n)$ qui existe dans l'espace complet $L^2(\Omega)$ peut alors servir de définition à $B(f)$ car toute autre suite $(f'_n, n \in N)$ dans \mathcal{A} convergent vers f donnerait la même limite $\lim_n B(f'_n)$ puisque $E([B(f_n) - B(f'_n)]^2) = \int_{R_+} (f_n - f'_n)^2 dt \rightarrow 0$ lorsque $n \rightarrow \infty$ si f_n et f'_n tendent asymptotiquement vers f .

L'application $f \rightarrow B(f)$ (que l'on a notée plutôt $\int f(t) dB(t)$ dans l'énoncé du théorème) étant ainsi définie, la vérification de ses propriétés peut être laissée au lecteur. \square

La proposition utile suivante montre que l'intégrale stochastique du théorème précédent vérifie la formule d'intégration par parties usuelles (celle qui serait immédiate si $\dot{B}(t)$ existait). Rappelons que $B(0) = 0$.

Proposition 1.3.5.

Pour toute fonction réelle continûment différentiable f sur R_+ et pour tout $T > 0$:

$$\int_0^T f(t) dB(t) + \int_0^T f'(t) B(t) dt = f(T) B(T)$$

En outre si $f(t) \rightarrow 0$ lorsque $t \rightarrow \infty$ et si $\int_0^\infty |f'(t)| \sqrt{t} dt < \infty$, la formule précédente est valable pour $T = \infty$ sous la forme

$$\int_{R_+} f(t) dB(t) = - \int_0^\infty f'(t) B(t) dt$$

DEMONSTRATION. Notons d'abord que pour tout T , la fonction $f(t) 1_{(t \leq T)}$ est bornée, à support compact et appartient donc à $L^2(R_+)$, ce qui permet de définir l'intégrale stochastique $\int_{R_+} f(t) 1_{(t \leq T)} dB(t)$ que l'on conviendra de noter $\int_0^T f(t) dB(t)$. Quant à l'intégrale

$\int_0^T f'(t) B(t) dt$ elle ne pose pas de problème de définition, puisqu'elle peut être définie ω par ω , les fonctions $t \rightarrow f'(t) B(t, \omega)$ étant continues.

D'après la caractérisation de l'intégrale stochastique de la fin du théorème précédent, tout revient à établir que pour tout $s \in R_+$:

$$E \left(\left[f(T) B(T) - \int_0^T f'(t) B(t) dt \right] B(s) \right) = \int_0^s f(t) 1_{(t \leq T)} dt$$

Or, en utilisant le théorème de Fubini et la formule $E[B(t) B(s)] = \min(t, s)$, cette formule équivaut à

$$f(T) s - \int_0^T f'(t) \min(t, s) dt = \int_0^s f(t) dt \quad (0 \leq s \leq T).$$

Puisque cette égalité se réduit à $0 = 0$ pour $s = 0$, elle équivaut encore par différentiation à

$$f(T) - \int_s^T f'(t) dt = f(s) \quad (0 \leq s \leq T)$$

et cette dernière formule est claire !

Sous les hypothèses de la seconde partie on a

$$f(t) = f(0) + \int_0^t f'(s) ds = - \int_t^\infty f'(s) ds \quad (t \in R_+)$$

et la fonction f est alors de carré intégrable sur tout R_+ car

$$\begin{aligned} \int_0^\infty f(t)^2 dt &= \int_0^\infty dt \left(\int_t^\infty f'(s) ds \right)^2 \leq \int_0^\infty \int_0^\infty \min(u, v) |f'(u)| |f'(v)| du dv \\ &\leq \int_0^\infty \int_0^\infty \sqrt{uv} |f'(u)| |f'(v)| du dv = \left[\int_0^\infty \sqrt{u} |f'(u)| du \right]^2. \end{aligned}$$

Il s'en suit que $\int_{R_+} f(t) dB(t)$ est définie. D'autre part comme $E(|B(t)|) = \sqrt{2t/\pi}$, car pour une v.a.r. gaussienne réduite $E(|\xi|) = \sqrt{2/\pi}$, on a

$$E \left[\int_0^\infty |f'(t)| |B(t)| dt \right] = \sqrt{2/\pi} \int_0^\infty |f'(t)| \sqrt{t} dt < \infty$$

et la v.a. $\int_0^\infty f'(t) B(t) dt$ est donc bien définie. L'égalité de la fin de la proposition se démontre alors comme dans le cas $T < \infty$. \square

1.4. Construction et régularité du mouvement brownien.

N.B. : Les résultats de ce paragraphe sont plus difficiles ; ils ne seront pas utilisés ultérieurement.

D'après le théorème 1.3.4, si $(\varphi_n, n \in N)$ désigne une base orthonormale de l'espace $L^2(R_+)$, les v.a.

$$\xi_n = \int_{R_+} \varphi_n(t) dB(t) \quad (n \in N)$$

construites dans ce théorème à partir d'un mouvement brownien B , forment une base orthonormale de l'espace gaussien $H[B]$ associé et sont donc des v.a.r. gaussiennes réduites indépendantes. En outre comme $B(s) = \int_{R_+} 1_{]0,s]}(t) dB(t)$ pour tout $s \in R_+$, aux développements orthonormaux des fonctions $1_{]0,s]}$ dans $L^2[R_+]$

$$1_{]0,s]} = \sum_N \tilde{\varphi}_n(s) \varphi_n \quad \text{où} \quad \tilde{\varphi}_n(s) = \langle 1_{]0,s]}, \varphi_n \rangle_{L^2(R_+)} = \int_0^s \varphi_n(t) dt$$

correspondent par isométrie les développements orthonormaux des v.a. $B(s)$ dans $H[B] \subset L^2(\Omega)$

$$B(s) = \sum_N \tilde{\varphi}_n(s) \xi_n \quad (s \in R_+) .$$

Cela suggère, pour construire un mouvement brownien, de se donner :

a) Une base orthonormale $(\xi_n, n \in N)$ de l'espace $L^2(R_+)$,

b) Une suite $(\xi_n, n \in N)$ de v.a.r. gaussiennes réduites indépendantes, sur un espace de probabilité convenable, puis de définir une fonction aléatoire $(B(s), s \in R_+)$ par les formules

$$B(s) = \sum_N \tilde{\varphi}_n(s) \xi_n \quad \text{où} \quad \tilde{\varphi}_n(s) = \int_0^s \varphi_n(t) dt \quad , \quad (s \in R_+) .$$

La convergence de ces séries est assurée car $\sum_N \tilde{\varphi}_n(s)^2 < \infty$ pour tout $s \in R_+$, puisque $\tilde{\varphi}_n(s)$ est le coefficient du développement de la fonction $1_{]0,s]}$ dans la base $(\varphi_n, n \in N)$ de $L^2(R_+)$; plus précisément

$$E[B(s)^2] = \sum_N \tilde{\varphi}_n(s)^2 = \int_{R_+} 1_{]0,s]}^2 dt = s \quad (s \in R_+)$$

et de même pour deux s et s' différents :

$$E[B(s) B(s')] = \sum_N \tilde{\varphi}_n(s) \tilde{\varphi}_n(s') = \int_{R_+} 1_{]0,s]} 1_{]0,s']} dt = \min(s, s') .$$

Ainsi $(B(s), s \in R_+)$ est une fonction aléatoire gaussienne centrée car toute combinaison linéaire des $B(s)$ est une série convergente de ξ_n , et sa covariance vaut $\min(s, t)$.

Pour établir que la f.a.r. est un mouvement brownien, il ne reste donc qu'à établir que ses trajectoires sont continues. Ceci s'établit facilement pour une base particulière de $L^2(R_+)$, la base de Haar, que nous supposons donc avoir choisie au début et que l'on définit comme suit. Si φ désigne la fonction réelle définie sur R par

$$\varphi(t) = \begin{cases} 1 & \text{sur }]0, 1/2] , \\ -1 & \text{sur }]1/2, 1] , \\ 0 & \text{ailleurs} \end{cases}$$

les fonctions de Haar $\varphi_{n,k}$ ($n, k \in N$) se déduisent de φ par les transformations affines :

$$\varphi_{n,k}(t) = 2^{n/2} \varphi(2^n t - k) \quad (t \in R_+) ;$$

notons que $\varphi_{n,k} = 0$ hors de $]k 2^{-n}, (k+1) 2^{-n}]$. Avec les fonctions

$$\psi_k = 1_{]k, k+1]} \quad (k \in N)$$

ces fonctions forment une base orthonormale de $L^2(R_+)$, le « système de Haar » (ψ_k ($k \in N$), $\varphi_{n,k}$ ($k, n \in N$)). Si $\tilde{\varphi}$ désigne la primitive de φ égale à

$$\tilde{\varphi}(t) = \begin{cases} t & \text{sur }]0, 1/2] , \\ 1 - t & \text{sur }]1/2, 1] , \\ 0 & \text{ailleurs} , \end{cases}$$

les primitives $\tilde{\varphi}_{n,k}$ des $\varphi_{n,k}$ valent

$$\tilde{\varphi}_{n,k}(t) = 2^{-n/2} \tilde{\varphi}(2^n t - k)$$

et sont nulles comme les $\varphi_{n,k}$ hors de $]k 2^{-n}, (k+1) 2^{-n}]$ respectivement.

En correspondance avec le système de Haar, nous nous sommes donnés des v.a.r. gaussiennes réduites indépendantes η_k ($k \in N$) et $\xi_{n,k}$ ($n, k \in N$). Posons alors sur R_+

$$\beta(s) = \sum_{k \in N} \tilde{\psi}_k(s) \eta_k \quad \text{et} \quad B^n(s) = \sum_{k \in N} \tilde{\varphi}_{n,k}(s) \xi_{n,k} \quad (n \in N)$$

puis

$$B(s) = \beta(s) + \sum_{n \in N} B^n(s) ;$$

nous savons d'après ce qui précède que B est une f.a. gaussienne centrée de covariance $\min(s, t)$ et nous voulons montrer que ses trajectoires sont continues. Or les trajectoires des f.a. β et B_n sont continues ; en effet les fonctions $\tilde{\psi}_k$ et $\tilde{\varphi}_{n,k}$ sont continues et comme pour un intervalle borné $[0, \nu]$ de R_+ ($\nu \in N$) on a $\tilde{\psi}_k \equiv 0$ sur $[0, \nu]$ si $k \geq \nu$, resp. si $k \geq \nu 2^n$, les restrictions de β et B^n à $[0, \nu]$ sont des combinaisons linéaires finies de fonctions $\tilde{\psi}_k$ et $\tilde{\varphi}_{n,k}$. Ainsi il nous suffira de montrer par une application du lemme de Borel-Cantelli que la série $\sum_n B^n(s, \omega)$ converge uniformément en s sur tout $[0, \nu]$, pour presque tout ω , pour qu'il s'en suive que presque toutes les trajectoires de B sont continues.

Or pour tout $n \in N$

$$B^n(s) = \tilde{\varphi}_{n,k}(s) \xi_{n,k} \quad \text{si} \quad k2^{-n} \leq s \leq (k+1)2^{-n} \quad (k \in N, s \in R_+)$$

et par conséquent, pour tout $v \in N$ fixé

$$\max (|B^n(s)|; 0 \leq s \leq v) = \frac{1}{2} 2^{-n/2} \max (|\xi_{n,k}|; 0 \leq k < v2^n)$$

puisque $1/2$ est la valeur maximum de la fonction positive $\tilde{\varphi}$. Il s'en suit que, pour tout $a \in R_+$:

$$\begin{aligned} & P[\max (|B^n(s)|; 0 \leq s \leq v) > a 2^{-n/2}] \\ &= P[\max (|\xi_{n,k}|; 0 \leq k < v2^n) > 2a] \leq v2^n P[|\xi| > 2a] \end{aligned}$$

par sous-additivité des probabilités. Mais pour la loi de Gauss réduite

$$P[|\xi| > 2a] \leq \exp(-2a^2) \quad \text{dès que} \quad a \geq 1$$

car par un calcul explicite sur cette loi

$$P(|\xi| > 2a) \leq E[|\xi| 1_{|\xi| > 2a}] / 2a = \exp(-2a^2) / a(2\pi)^{1/2}.$$

En prenant $a = \sqrt{n}$ dans l'inégalité précédente, nous trouvons finalement que

$$\begin{aligned} & \sum_{n \geq 1} P[\max (|B^n(s)|; 0 \leq s \leq v) > \sqrt{n} 2^{-n/2}] \\ & \leq \sum_{n \geq 1} v [2 \exp(-2)]^n < \infty ; \end{aligned}$$

le lemme de Borel-Cantelli montre alors que p.s. les termes de la série

$$\sum_{n \geq 1} \max (|B^n(s)|; 0 \leq s \leq v)$$

sont majorés à partir d'un certain rang aléatoire par ceux de la série convergente

$$\sum_{n \geq 1} \sqrt{n} 2^{-n/2} \quad . \square$$

Proposition 1.4.1.

Presque toute trajectoire du mouvement brownien n'est dérivable en aucun point de R_+ .

Pour tout $\alpha < 1/2$, presque toute trajectoire du mouvement brownien est hölderienne d'ordre α sur tout intervalle borné, ce qui signifie que pour tout $v \in N^$:*

$$\sup (|B(t) - B(s)| / |t - s|^\alpha; 0 \leq s, t \leq v, |t - s| < h) \rightarrow 0 \text{ p.s.}$$

lorsque $h \searrow 0$.

DEMONSTRATION. a) Si la fonction réelle $f: R_+ \rightarrow R$ est dérivable en un point de $[0, v]$ au moins, alors

$$|f((k+1)2^{-n}) - f(k2^{-n})| \leq a 2^{-n}$$

pour un entier $a > 0$ au moins, pour tout entier n assez grand et pour trois valeurs entières consécutives de k dans $[0, \nu 2^n[$ au moins... Or pour le mouvement brownien, l'événement

$$F_{a,n} = \bigcup_{0 < \ell < \nu 2^n} \bigcap_{k=\ell-1, \ell, \ell+1} \{ |B[(k+1)2^{-n}] - B[k2^{-n}]| < a 2^{-n} \}$$

a une probabilité

$$P(F_{a,n}) \leq \nu 2^n P(|\xi| < a 2^{-n/2})^3$$

puisque les v.a. $B[(k+1)2^{-n}] - B[k2^{-n}]$ ($k = \ell - 1, \ell, \ell + 1$) sont indépendantes et suivent la même loi que $\xi 2^{-n/2}$ si ξ est gaussienne réduite ; mais alors, lorsque $n \rightarrow \infty$

$$P(F_{a,n}) \leq c^{\text{te}} 2^{-n/2}$$

et grâce au lemme de Borel-Cantelli, un nombre fini de $F_{a,n}$ ($n \in N$) est réalisé, quel que soit l'entier $a > 0$. La première partie de la proposition en résulte immédiatement.

b) On reprend la démonstration de la continuité des trajectoires du mouvement brownien en commençant par remarquer que les fonctions $\tilde{\psi}_k$ et $\tilde{\varphi}_{n,k}$ ont des pentes maximum égales à 1 et $2^{n/2}$ respectivement de sorte que pour tous s, t dans $[0, \nu]$:

$$|\beta(t) - \beta(s)| \leq |t - s| \max (|\eta_k| ; 0 \leq k \leq \nu)$$

et

$$|B^n(t) - B^n(s)| \leq 2^{n/2} |t - s| \max (|\xi_{n,k}| ; 0 \leq k < \nu 2^n)$$

pour tout $n \in N$. Nous tiendrons compte aussi de ce que

$$\begin{aligned} |B^n(t) - B^n(s)| &\leq 2 \max (|B^n(u)| ; 0 \leq u \leq \nu) \\ &= 2^{-n/2} \max (|\xi_{n,k}| ; 0 \leq k < \nu 2^n) \end{aligned} ,$$

cette inégalité étant meilleure que la précédente lorsque $|t - s| > 2^{-n}$.

D'autre part nous avons vu précédemment que p.s. si n est assez grand, c'est-à-dire si $n \geq N(\omega)$ où N est un entier aléatoire p.s. fini,

$$\max (|\xi_{k,n}| ; 0 \leq k < \nu 2^n) \leq 2 \sqrt{n} .$$

Il s'en suit finalement que si $0 \leq s, t \leq \nu$:

$$\begin{aligned} |B(t) - B(s)| &\leq |\beta(t) - \beta(s)| + \sum |B^n(t) - B^n(s)| \\ &\leq |\beta(t) - \beta(s)| + \sum_{n < N} |B^n(t) - B^n(s)| + f(|t - s|) \end{aligned}$$

à condition de poser

$$f(u) = \sum_n 2 \sqrt{n} \min (2^{-n/2}, u 2^{n/2}) \quad (u \in R_+) .$$

Or il est facile de montrer que $f(u) = O(u^\alpha)$ pour tout $\alpha < 1/2$ [par contre $f(u) \sqrt{u} \rightarrow \infty$ lorsque $u \rightarrow 0$] et la seconde partie de la proposition est ainsi démontrée. \square

1.5. Equation de Langevin et Processus d'Ornstein-Uhlenbeck.

L'oscillateur harmonique classique sur R obéit à l'équation différentielle

$$(1.5.1) \quad \ddot{x}(t) - b \dot{x}(t) + c x(t) = f(t)$$

où $f(t)$ représente la force extérieure à laquelle il est soumis et $b, c \geq 0$ sont des constantes ≥ 0 . Si la force extérieure est due à des chocs aléatoires multiples se produisant aux instants k/n ($k \in Z$), la fonction $f(t)$ sera une fonction aléatoire telle que

$$\int_s^t f(u) du = \sum_{s \leq k/n \leq t} X_k$$

et donc si n est grand et le passage à la limite du §1.3 justifié, la fonction f sera donnée par

$$\int_s^t f(u) du = \sigma [B(t) - B(s)] \quad (s < t)$$

pour une constante $\sigma > 0$ et un mouvement brownien B ; autrement dit $f(t)$ sera égal à $\sigma \dot{B}(t)$...

De manière plus rigoureuse, l'équation stochastique de Langevin de l'oscillateur harmonique soumis à une force aléatoire brownienne s'écrit

$$\dot{X}(t) - \dot{X}(s) - b [X(t) - X(s)] + c \int_s^t X(u) du = \sigma (B(t) - B(s)) \quad (s < t)$$

ce que l'on note plus brièvement mais formellement

$$d \dot{X}(t) - b \dot{X}(t) dt + c X(t) dt = \sigma dB(t)$$

Pour résoudre cette équation, on procède en fait comme dans le cas d'une force $f(t)$ déterministe. Limitons nous pour simplifier au cas où $c = 0$ et rappelons qu'alors la solution de 1.5.1 pour des données initiales $x(0)$ et $\dot{x}(0)$ est donnée par

$$\dot{x}(t) = e^{-tb} \dot{x}(0) + \int_0^t e^{-(t-s)b} f(s) ds \quad ,$$

$$x(t) = x(0) + \int_0^t \dot{x}(s) ds \quad .$$

La première partie de la proposition suivante est donc naturelle.

Proposition 1.5.1.

La solution de l'équation de Langevin

$$(1.5.2) \quad dV(t) + b V(t) dt = \sigma^2 dB(t) \quad (t \in R_+)$$

pour une valeur initiale $V(0)$ donnée, vaut

$$(1.5.3) \quad V(t) = e^{-tb} V(0) + \int_0^t e^{-(t-s)b} \sigma dB(s) \quad (t \in R_+)$$

Si $V(0)$ est une v.a.r gaussienne (éventuellement dégénérée) indépendante du mouvement brownien B , la f.a.r $(V(t), t \in R_+)$ est gaussienne d'espérance et de covariance

$$E[V(t)] = e^{-tb} E[V(0)] \quad ,$$

$$\text{Cov}[V(s), V(t)] = e^{-sb} v e^{-tb} + \int_0^s e^{-(s-u)b} \sigma^2 e^{-(t-u)b} du$$

($s \leq t$) si v désigne la variance de $V(0)$.

En particulier si la v.a. $V(0)$ est centrée et si sa variance v vérifie l'égalité $2bv = \sigma^2$, la fonction aléatoire V est centrée et de covariance stationnaire (= ne dépendant que de $t - s$) égale à

$$(1.5.4) \quad \text{Cov}[V(s), V(t)] = v \exp(-(t-s)b) \quad (0 \leq s \leq t)$$

En général lorsque $T \rightarrow \infty$, l'espérance et la covariance de la f.a. $(V(T+t), t \in R_+)$ tendent vers celles du cas stationnaire.

La fonction aléatoire $(V(t), t \in R_+)$ est appelée processus d'Ornstein-Uhlenbeck.

DEMONSTRATION. La fonction aléatoire V définie par (1.5.2) peut encore s'écrire par la formule d'intégration par parties (proposition 1.3.5)

$$V(t) = e^{-tb} V(0) + \sigma B(t) - \int_0^t b e^{-(t-s)b} \sigma B(s) ds$$

En intégrant les deux membres de 0 à u , il vient alors après un calcul facile

$$b \int_0^u V(t) dt = (1 - e^{-ub}) V(0) + \int_0^u b e^{-(u-s)b} \sigma B(s) ds$$

de sorte que

$$\begin{aligned} V(u) - V(0) + b \int_0^u V(t) dt \\ &= V(u) - e^{-ub} V(0) + \int_0^u b e^{-(u-s)b} \sigma B(s) ds \\ &= \sigma B(u) \end{aligned}$$

pour tout $u \in R_+$. La formule ainsi obtenue

$$V(u) - V(0) + b \int_0^u V(t) dt = \sigma B(u) \quad (u \in R_+)$$

est l'équation de Langevin que l'on a écrit formellement dans la proposition sous forme différentielle

$$dV(u) + b V(u) du = \sigma dB(u) \quad .$$

La deuxième partie de la proposition se démontre par des calculs élémentaires et l'application du théorème 1.3.4. \square

Les processus d'Ornstein-Uhlenbeck multi-dimensionnels jouent un grand rôle dans les applications et sont les solutions de l'équation de Langevin multidimensionnelle

$$(1.5.5) \quad dV(t) + b V(t) dt = \sigma dB(t) \quad (t \in R_+)$$

où :

a) $(B(t), t \in R_+)$ désigne un mouvement brownien dans R^d , c'est-à-dire une fonction aléatoire à valeurs dans R^d ($B(t, \omega) \in R^d$) dont les coordonnées $(B_j(t), t \in R_+)$ sont des mouvements browniens scalaires indépendants entre eux ($j = 1, \dots, d$)

b) $b = (b_{ij}; 1 \leq i, j \leq d)$ et $\sigma = (\sigma_{ij}; 1 \leq i, j \leq d)$ sont deux matrices réelles $d \times d$

c) $(V(t), t \in R_+)$ désigne une fonction aléatoire à valeurs dans R^d . La solution de cette équation est encore donnée par (1.5.3), étant entendu que e^{-tb} désigne une matrice $d \times d$.

[Rappel. Si a désigne une matrice $d \times d$, par définition :

$e^a = \sum_{n \geq 0} \frac{1}{n!} a^n$ où a^n est la n^{e} puissance de la matrice a ($a^0 =$ matrice identité) ; donc

$$(e^a)_{ij} = \sum_{n \geq 0} \frac{1}{n!} (a^n)_{ij} \quad (1 \leq i, j \leq d) \quad .$$

Il est évidemment faux que $(e^a)_{ij} = \exp(a_{ij})$ sauf si a est une matrice diagonale. D'autre part

$$e^a \cdot e^{a'} = e^{a'} \cdot e^a = e^{a+a'} \quad \text{si } aa' = a'a$$

et donc en particulier $(e^{-tb}, t \geq 0)$ est un semi-groupe

$$e^{-tb} e^{-t'b} = e^{-(t+t')b} \quad (t, t' \in R_+) \quad] \quad .$$

Enfin la deuxième partie de la proposition 1.5.1 reste exacte dans le cas multidimensionnel, à condition de définir v comme la matrice de covariance de $V(0)$, de remplacer σ^2 par $\sigma \sigma^t$ (où σ^t désigne la matrice transposée de σ) et d'écrire l'équation scalaire $2bv = \sigma^2$ sous la forme matricielle

$$bv + v b^t = \sigma \sigma^t \quad ,$$

1.6. Formule de Cameron-Martin.

(Il s'agit du cas particulier de la formule de Girsanov correspondant au cas d'une intégrande de déterministe, le cas général étant traité dans le chapitre 2).

Si ξ est une v.a.r. gaussienne réduite, donc si ξ suit la densité

$$g(x) = (2\pi)^{-1/2} \exp(-x^2/2) \quad (x \in R) ,$$

la v.a. translatée $\xi + a$ suit pour tout réel a fixé, la densité $g(x - a)$ [noter le changement de signe !] qui s'écrit encore

$$g(x - a) = g(x) \exp(ax - a^2/2)$$

De manière plus probabiliste, nous pourrions alors écrire que

$$(1.6.1) \quad E[f(\xi + a)] = E[f(\xi) \exp(ax - a^2/2)]$$

au moins pour toute fonction f positive ou bornée, puisque les deux membres de cette formule valent respectivement

$$\int_R f(y) g(y - a) dy \quad \text{et} \quad \int_R f(x) \exp(ax - a^2/2) g(x) dx .$$

La formule (1.6.1) ramène le calcul de toute espérance relative à la v.a. $\xi + a$ à celui d'une espérance relative à la v.a. ξ ; il est remarquable que cette formule se généralise à une fonction aléatoire gaussienne quelconque (pour simplifier nous nous restreindrons au cas d'une f.a. centrée).

Théorème 1.6.1.

Soit $(X(t), t \in R_+)$ une fonction aléatoire gaussienne centrée et soit $m : R_+ \rightarrow R$ une fonction réelle qui puisse s'écrire

$$m(t) = E[X(t) Y]$$

pour une v.a.r. Y de l'espace gaussien associé à X . Alors

$$(1.6.2) \quad \begin{aligned} E[f(X(t_1) + m(t_1), \dots, X(t_n) + m(t_n))] \\ = E[f(X(t_1), \dots, X(t_n)) \exp(Y - E(Y^2)/2)] \end{aligned}$$

quels que soient $0 \leq t_1 < \dots < t_n$ dans R_+ et quelle que soit la fonction borélienne, positive ou bornée $f : R^n \rightarrow R$.

Plus généralement pour toute fonctionnelle borélienne F définie sur l'espace $C(R_+)$ des trajectoires de la f.a. X et à valeurs réelles positives ou bornées :

$$(1.6.3) \quad E[F(X(\cdot) + m(\cdot))] = E[F(X(\cdot)) \exp(Y - E(Y^2)/2)]$$

(pour fonctionnelle F on pourra considérer par exemple $\int_0^T x(t)^2 dt$ ou $\max_t x(t)$).

DEMONSTRATION. Pour une fonction f de forme exponentielle, soit

$$f(x_1, \dots, x_n) = \exp \left(\sum_1^n a_j x_j \right) \quad (a_j \in R)$$

nous avons en effet

$$\begin{aligned} & E[f(X(t_1) + m(t_1), \dots, X(t_n) + m(t_n))] \\ &= E[\exp(\sum_1^n a_j X(t_j))] \exp[\sum_1^n a_j m(t_j)] \\ &= \exp\left[\frac{1}{2} E\left(\left[\sum_1^n a_j X(t_j)\right]^2\right) + \sum_1^n a_j m(t_j)\right] \end{aligned}$$

puisque $\sum_1^n a_j X(t_j)$ est une v.a. gaussienne centrée, et de même

$$\begin{aligned} & E[f(X(t_1), \dots, X(t_n)) \exp(Y - E(Y^2)/2)] \\ &= E[\exp(\sum_1^n a_j X(t_j) + Y)] \exp\left[\frac{-1}{2} E(Y^2)\right] \\ &= \exp\left[\frac{1}{2} E\left[\left(\sum_1^n a_j X(t_j) + Y\right)^2\right] - \frac{1}{2} E(Y^2)\right] \\ &= \exp\left(\frac{1}{2} E\left(\left[\sum_1^n a_j X(t_j)\right]^2\right) + \sum_j^n a_j E[X(t_j) Y]\right) \end{aligned}$$

puisque $\sum_1^n a_j X(t_j) + Y$ est une v.a. gaussienne centrée. Ce calcul établit la formule de la proposition pour une fonction exponentielle, puisque $m(t) = E[X(t) Y]$.

La validité de la proposition pour une fonction f quelconque et plus généralement pour une fonctionnelle F s'obtient par une technique de prolongement que nous détaillerons pas. \square

La formule du théorème précédent, outre son importance dans les calculs, est fondamentale pour les applications statistiques. Pour l'expliquer, nous supposons que nous travaillons sur l'espace de probabilité $\Omega = C(R_+)$ formé par toutes les fonctions continues $\omega: R_+ \rightarrow R$ et que la fonction aléatoire $X(\cdot)$ sur Ω est la f.a. dite canonique $X(t, \omega) = \omega(t)$ (*Si ce n'était pas le cas, il conviendrait de projeter l'espace Ω sur $C(R_+)$ par l'application $\omega \rightarrow X(\cdot, \omega)$).

La formule (1.6.2) affirme qu'en remplaçant sur l'espace Ω , la probabilité P et l'espérance (ou intégrale)

$$E(Z) := \int_{\Omega} Z(\omega) P(d\omega) \quad (Z \text{ v.a.r.})$$

qui est lui est associée, pour la nouvelle probabilité Q définie sur Ω par

$$(1.6.4) \quad Q(A) = \int_A \exp[Y(\omega) - E(Y^2)/2] P(d\omega) \quad (A \in \mathfrak{A})$$

et par l'espérance E_Q correspondante, soit

$$\begin{aligned} E_Q(Z) &:= \int_{\Omega} Z(\omega) Q(d\omega) \\ &= \int_{\Omega} Z(\omega) \exp [Y(\omega) - E(Y^2)/2] P(d\omega) , \end{aligned}$$

on passe en fait de la loi de la fonction aléatoire $X(\cdot)$ à celle de la fonction aléatoire translatée $X(\cdot) + m(\cdot)$; en effet, (1.6.2) montre par exemple que

$$\begin{aligned} P[(X(t_1) + m(t_1), \dots, X(t_n) + m(t_n)) \in F] \\ = Q[(X(t_1), \dots, X(t_n)) \in F] \end{aligned}$$

pour tous $n \geq 1$, $t_1, \dots, t_n \in R_+$ et F dans R^n .

Ainsi au lieu de considérer les *deux fonctions aléatoires*

$$(X(t), t \in R_+) \quad \text{et} \quad (X(t) + m(t), t \in R_+)$$

sur Ω muni de la *seule probabilité* P , il est équivalent de considérer la *seule fonction aléatoire* $(X(t), t \in R_+)$ pour les *deux probabilités* P et Q sur Ω . Or ces deux points de vue correspondent à deux situations pratiques différentes.

Supposons que $m(\cdot)$ représente un signal temporel (déterministe) intéressant, susceptible d'être présent ou absent, tandis que $X(\cdot)$ représente un bruit additif. La modélisation de cette situation conduit évidemment à considérer les deux fonctions aléatoires $X(\cdot)$ correspondant à un bruit pur, et $m(\cdot) + X(\cdot)$ correspondant à un « signal utile » plongé dans un bruit, la probabilité P donnant la loi du bruit. Mais l'observateur de son côté n'a accès qu'à l'une de ces deux fonctions aléatoires dont il observe une trajectoire $\omega = (\omega(t), t \in R_+)$ et généralement il ignore laquelle des deux fonctions est en jeu, son problème étant précisément de le détecter. Or, d'après les résultats précédents la trajectoire observée ω possédait une probabilité a priori d'être observée égale à $P(d\omega)$ ou à $Q(d\omega)$ selon le cas; comme on a ici

$$\frac{Q(d\omega)}{P(d\omega)} = \exp [Y(\omega) - E(Y^2)/2] ,$$

on voit que si la valeur du second membre *pour le ω observé* est voisine de $+\infty$, resp. de 0 il est très vraisemblable que le signal utile $m(\cdot)$ était présent dans l'observation, resp. qu'il en était absent.

En extrapolant l'idée précédente, on aboutit à la règle ou stratégie du **maximum de vraisemblance**: *en possession de l'observation ω , on décide que le signal était présent si*

$$Y(\omega) > E(Y^2)/2 ,$$

c'est-à-dire si $Q(d\omega)/P(d\omega) > 1$. Cette règle n'est pas infaillible! Si le signal m est réellement absent, la probabilité de prendre une mauvaise décision vaut

$$p := P [Y > E(Y^2)/2]$$

tandis que si le signal m est réellement présent, la probabilité d'erreur vaut

$$q := Q [Y \leq E(Y^2)/2] ,$$

puisque suivant le cas, la probabilité P ou la probabilité Q régit l'expérience aléatoire.

Ces deux probabilités d'erreur p, q s'expriment ici simplement en fonction de $\sigma^2 := E(Y^2)$. En effet, Y suit la loi gaussienne centrée $N(0, \sigma^2)$ sous P et par conséquent, si ξ désigne une v.a.r. gaussienne réduite auxiliaire

$$p = \text{Proba.} (\sigma \xi > \sigma^2/2) = \text{Proba.} (\xi > \sigma/2) .$$

D'autre part, sous Q la v.a.r. Y suit la loi $N(\sigma^2, \sigma^2)$ car

$$\begin{aligned} \int \exp(u Y) dQ &= \int \exp((u+1) Y) dP \exp(-\sigma^2/2) \\ &= \exp([(u+1)^2 - 1] \sigma^2/2) \\ &= \exp(u\sigma^2 + u^2\sigma^2/2) \end{aligned}$$

pour tout $u \in R$. Par suite

$$q = \text{Proba.} (\sigma^2 + \sigma \xi \leq \sigma^2/2) = \text{Proba.} (\xi < -\sigma/2) .$$

Finalement, si g désigne la densité gaussienne réduite

$$p = q = \int_{\sigma/2}^{\infty} g(x) dx$$

et la règle de décision du maximum de vraisemblance peut être considérée comme satisfaisante si

$$\sigma > 4$$

(c'est-à-dire si $p = q$ est inférieure à 5%). On peut montrer d'autre part que la règle du maximum de vraisemblance est optimale au sens où elle minimise la somme $p + q$ des probabilités d'erreur parmi toutes les règles de décision possibles.

1.7. Bref historique du mouvement brownien.

R. Brown, botaniste anglais, décrit dès 1827 ce mouvement comme celui de fines particules organiques en suspension dans un gaz ou un fluide. Au 19^e siècle, après lui plusieurs physiciens reconnaissent que ce mouvement est très irrégulier et ne semble pas admettre de tangente ; on ne pourrait donc pas parler de sa vitesse ni a fortiori lui appliquer les lois de la Mécanique !

En 1905, A. Einstein construit un modèle probabiliste de diffusion pour décrire ce mouvement brownien ; il trouve que la densité de probabilité de $B(t)$ pour un état initial $B(0) = x$ donné vérifie l'équation de la chaleur et de ce fait est gaussienne. Sa théorie sera rapidement confortée par des mesures expérimentales de constantes de diffusion satisfaisantes.

La même année qu'Einstein, le physicien polonais Smoluchowski décrit le mouvement brownien comme limite en loi de promenades aléatoires.

En 1923, N. Wiener construit rigoureusement la fonction aléatoire du mouvement brownien ; il établit que ses trajectoires sont continues en introduisant la méthode du §1.4. pour la base trigonométrique de $L^2[0,1]$. P. Lévy découvre ensuite, avec d'autres mathématiciens, de nombreuses propriétés du mouvement brownien et introduit une première forme des équations différentielles stochastiques ; ses travaux sont rassemblés dans un traité paru en 1948 devenu célèbre.

Dès 1930, en suivant une idée de P. Langevin, Ornstein et Uhlenbeck étudient la fonction aléatoire qui porte leurs noms.

Pour de plus amples informations, on pourra consulter :

- A. Einstein, *Investigations on the theory of the brownian motion*, (Dover, réédition, 1956).
- P. Nelson, *Dynamical theories of brownian motion*.
- P. Lévy, *Le mouvement brownien*, Gauthier Villers, 1948.

Chapitre II

Le calcul stochastique d'Itô

2.0. Introduction.

Pour toute fonction réelle $t \rightarrow f(t)$ de carré intégrable définie sur R_+ , nous avons défini au chapitre 1 l'intégrale $\int_{R_+} f(t) dB(t)$ de f par rapport à un mouvement brownien B bien que les trajectoires de ce mouvement ne soient pas dérivables. Le but de ce chapitre est de généraliser cette intégrale au cas où f est remplacée par une fonction aléatoire $t, \omega \rightarrow \varphi(t, \omega)$ vérifiant une condition d'*adaptation* précisée ci-dessous et une condition d'intégrabilité appropriée.

L'intégrale stochastique d'Itô

$$\int_{R_+} \varphi(t) dB(t)$$

qui sera ainsi définie conduira à un calcul différentiel stochastique basé sur la célèbre formule d'Itô ; cela permettra de résoudre des équations différentielles stochastiques généralisant l'équation de Langevin du premier chapitre. Il conviendra d'ailleurs pour ces applications de définir l'intégrale stochastique pour des processus φ et B vectoriels et non seulement scalaires. La *formule de Girsanov* qui généralise aux intégrales aléatoires la formule de Cameron-Martin du 1^{er} chapitre, sera ensuite établie ; elle joue un rôle fondamental dans les applications, notamment dans les applications statistiques. En outre nous présenterons brièvement l'intégrale stochastique de Stratonovitch, qui est un peu différente de celle d'Itô.

2.1. Préliminaires.

Pour tout mouvement brownien B sur R_+ , par définition, la fonction aléatoire $(B(s), 0 \leq s \leq t)$ arrêtée à t ($t \in R_+$ fixé) est indépendante des accroissements ultérieurs $(B(u) - B(t), u \geq t)$. Pour définir l'intégrale stochastique $\int_{R_+} \varphi(t) dB(t)$ et en obtenir des propriétés intéressantes, nous verrons qu'il conviendra d'imposer à la fonction aléatoire φ une condition d'indépendance analogue : pour tout $t \in R_+$, la fonction aléatoire $(\varphi(s), 0 \leq s \leq t)$ arrêtée à t et le mouvement brownien arrêté à t sont conjointement indépendants de $(B(u) - B(t), u \geq t)$. Pour donner une formulation mathématique précise de cette condition, nous sommes obligés de revenir sur la notion de tribu et sur la définition de l'indépendance.

A.

Rappelons qu'un espace de probabilité est un triplet (Ω, \mathcal{A}, P) dans lequel \mathcal{A} désigne une tribu de parties de Ω et P une fonction d'ensembles σ -additive définie sur \mathcal{A} , à valeurs dans $[0, 1]$ telle que $P(\Omega) = 1$ (voir le § 1.5 du cours de Probabilités). Pour toute v.a. réelle vectorielle $X : \Omega \rightarrow R^d$, par définition (définition 3.2.4) les parties de Ω

$$\{ \omega : X(\omega) \in F \} = X^{-1}(F)$$

où F parcourt les boréliens de R^d , sont des évènements (= appartiennent à \mathcal{A}) ; ce sont les évènements qui ne dépendent que de X , ceux dont on sait qu'ils sont réalisés ou non lorsqu'on connaît $X(\omega)$, mais pas ω lui-même. De plus la classe

$$\tau(X) := (\{ X \in F \}, F \text{ borélien de } R^d)$$

est une sous-tribu de \mathcal{A} ; cette sous-tribu représente l'information apportée par X . On a d'ailleurs le résultat suivant.

Lemme 2.11.

Etant donné deux v.a.r. vectorielles $X: \Omega \rightarrow R^d$ et $Y: \Omega \rightarrow R^{d'}$ sur l'espace de probabilités (Ω, \mathcal{A}, P) , les conditions suivantes sont équivalentes :

- a) $Y = f(X)$ pour une fonction borélienne $f: R^d \rightarrow R^{d'}$ au moins,
- b) $\tau(Y) \subset \tau(X)$

Ces deux conditions expriment chacune à leur façon que Y est moins informative que X .

DEMONSTRATION. (a) \Rightarrow (b). En effet si \mathfrak{B} (resp. \mathfrak{B}') désigne la tribu des boréliens de R^d (resp. $R^{d'}$) :

$$\tau(Y) = Y^{-1}(\mathfrak{B}') = X^{-1} \circ f^{-1}(\mathfrak{B}') \subset X^{-1}(\mathfrak{B}) = \tau(X) \quad .$$

(b) \Rightarrow (a). On se ramène au cas où $d' = 1$. Alors l'hypothèse (b) entraîne que pour tout $a \in \mathbb{Q}$: $\{ Y < a \} = \{ X \in F_a \}$ pour un borélien F_a de R^d ; il suffit alors de définir la fonction $f: R^d \rightarrow \mathbb{R}$ par $f(x) = \inf (a : a \in \mathbb{Q}, x \in F_a)$ pour voir que

$$f(X(\omega)) = \inf (a : a \in \mathbb{Q}, X(\omega) \in F_a) = \inf (a : a \in \mathbb{Q}, Y(\omega) < a) = Y(\omega) \quad .$$

Enfin f est borélienne car $\{ f < b \} = \bigcup (F_a ; a \in \mathbb{Q}, a < b)$ est un borélien comme réunion dénombrable de boréliens. \square

La définition de l'indépendance de n v.a.r. X_1, \dots, X_n (définition 4.1.1. du cours) ne fait intervenir que les tribus $\tau(X_i)$ associées à ces v.a.r. et on peut donc plus généralement définir l'indépendance de n sous-tribus $\mathfrak{B}_1, \dots, \mathfrak{B}_n$ de \mathcal{A} dans l'espace (Ω, \mathcal{A}, P) par la propriété

$$(2.1.1) \quad P(B_1 \cap \dots \cap B_n) = P(B_1) \dots P(B_n) \quad \forall B_i \in \mathfrak{B}_i (1 \leq i \leq n)$$

On dira aussi qu'une v.a.r. X et une sous-tribu \mathfrak{B} sont indépendantes dans (Ω, \mathcal{A}, P) si $P(\{ X \in F \} \cap B) = P(X \in F) P(B)$ quels que soient le borélien F et l'évènement $B \in \mathfrak{B}$.

Enfin dans la théorie des processus aléatoires, il est utile d'associer à une famille quelconque (infinie notamment) de v.a., soit $(X_i, i \in I)$, la tribu $\tau(X_i, i \in I)$ dite engendrée par ces variables et qui est la plus petite sous-tribu de \mathcal{A} contenant tous les évènements $\{ X_i \in F \}$ ($i \in I, F$ borélien), c'est-à-dire toutes les sous-tribus $\tau(X_i)$ ($i \in I$). Par exemple pour un mouvement brownien $(B(t), t \in R_+)$, la tribu du passé à l'instant t ($t \geq 0$) que l'on notera \mathfrak{F}_t^B dans la suite, sera définie par

$$(2.1.2) \quad \mathfrak{F}_t^B = \tau(B(s), s \leq t) \quad .$$

Nous admettrons le résultat suivant

Lemme 2.1.2.

Pour qu'une v.a.r. Y soit indépendante de la tribu $\tau(X_i, i \in I)$ associée à une famille arbitraire de v.a.r. X_i , il suffit qu'elle soit indépendante de toutes les sous-familles finies $(X_{i_1}, \dots, X_{i_n})$ de v.a.r. X_i .

B.

Revenons au mouvement brownien $(B(t), t \in R_+)$. Les tribus du passé \mathfrak{F}_t^B ($t \in R_+$) introduites ci-dessus, croissent manifestement avec t : elles forment la *filtration propre* $(\mathfrak{F}_t^B, t \in R_+)$ du brownien B (une filtration se définit comme une famille croissante de sous-tribus de \mathfrak{A} dans un espace de probabilité $(\Omega, \mathfrak{A}, P)$).

Mais dans les applications faisant intervenir un mouvement brownien, il est fréquent que d'autres variables ou fonctions aléatoires interviennent qui ne dépendent pas du mouvement brownien. Alors pour de tels modèles, le passé à l'instant t est plus riche que celui donné par \mathfrak{F}_t^B ($t \in R_+$) et on est ainsi amené à définir les mouvements browniens de la manière suivante.

Définition 2.1.3.

Etant donné une filtration $(\mathfrak{F}_t, t \in R_+)$ dans l'espace de probabilité $(\Omega, \mathfrak{A}, P)$ représentant le passé à chaque instant, on appelle \mathfrak{F} -mouvement brownien toute fonction aléatoire $(B(t), t \in R_+)$ à trajectoires continues telle que

a) $\{B(t) \in F\} \in \mathfrak{F}_t$ pour tout borélien F et tout $t \in R_+$ (on dit alors que $B(t)$ est \mathfrak{F}_t -mesurable $\forall t \in R_+$ ou plus brièvement que $B(\cdot)$ est **adapté** à la filtration \mathfrak{F}).

b) pour tous $t < u$, l'accroissement $B(u) - B(t)$ est indépendant de \mathfrak{F}_t et suit la loi $N(0, u - t)$.

Cette définition est équivalente à celle du chapitre 1 lorsque \mathfrak{F} est la filtration propre \mathfrak{F}^B du mouvement brownien (pour le voir on utilisera le lemme ci-dessus ainsi que l'équivalence des deux définitions suivantes: a) X_1, X_2, \dots, X_n sont indépendants, b) pour tout m ($1 \leq m < n$), X_{m+1} est indépendant du vecteur (X_1, \dots, X_m)).

2.2. L'intégrale stochastique d'Itô.

Dans toute la suite, on supposera donnés une filtration $\mathfrak{F} = (\mathfrak{F}_t, t \in R_+)$ et un \mathfrak{F} -mouvement brownien $(B(t), t \in R_+)$ [Déf. 2.1.3]. Voici alors la classe des fonctions aléatoires que nous prendrons pour intégrandes dans l'intégrale d'Itô.

Définition 2.2.1.

Une fonction aléatoire $\varphi(t, \omega)$ définie sur $R_+ \times \Omega$, resp. sur $[0, T] \times \Omega$ ($T < \infty$) sera dite **progressivement mesurable** si pour tout $t \in R_+$, resp. $t \in [0, T]$, l'application

$$(2.2.1) \quad (s, \omega) \rightarrow \varphi(s, \omega) \quad \text{de} \quad [0, t] \times \Omega \quad \text{dans} \quad R$$

est $\mathfrak{B}([0,t]) \otimes \mathfrak{F}_t$ mesurable, où $\mathfrak{B}([0,t])$ désigne la tribu des parties boréliennes de $[0,t]$ et où $\mathfrak{B}([0,t]) \otimes \mathfrak{F}_t$ désigne la tribu produit de parties de $[0,t] \times \Omega$ qui par définition est engendrée par les rectangles $B \times F$ ($B \in \mathfrak{B}([0,t])$, $F \in \mathfrak{F}_t$).

On notera $M^2(R_+)$, resp. $M^2([0,t])$ l'espace des fonctions aléatoires progressivement mesurables $\varphi(t,\omega)$ sur R_+ , resp. $[0,T]$ telles que

$$(2.2.1) \quad E \left[\int_{R_+ \text{ resp. } [0,T]} \varphi^2(t,\cdot) dt \right] < \infty .$$

(on convient d'identifier deux f.a. φ et φ' qui ne diffèrent que sur un sous-ensemble de $R_+ \times \Omega$ de mesure $dt dP(\omega)$ égale à zéro). Cet espace est un espace de Hilbert pour le produit scalaire

$$\langle \varphi, \psi \rangle = E \left[\int_{R_+ \text{ resp. } [0,T]} \varphi(t,\cdot) \psi(t,\cdot) dt \right] .$$

Enfin on désignera encore par M^2 l'espace des f.a. progressivement mesurables φ sur $R_+ \times \Omega$ telles que pour tout $T < \infty$: $E \left(\int_0^T \varphi^2(t,\cdot) dt \right) < \infty$.

Notons qu'une f.a. progressivement mesurable est adaptée à la filtration \mathfrak{F} au sens où pour tout $t \in R_+$, resp. $[0,T]$, la v.a. $\varphi(t,\cdot)$ est \mathfrak{F}_t -mesurable.

Nous définirons l'intégrale d'Itô d'une fonction aléatoire φ progressivement mesurable par rapport au mouvement brownien B en considérant pour commencer le cas des « fonctions en escalier », c'est-à-dire des fonctions de la forme

$$(2.2.3) \quad \varphi(t,\omega) = \sum_{i=0}^{n-1} X_i(\omega) \mathbf{1}_{]t_i, t_{i+1}]}(t)$$

où $n \in \mathbb{N}$, $0 \leq t_0 < t_1 < \dots < t_n$ et où X_i est \mathfrak{F}_{t_i} -mesurable et de carré intégrable (ce que l'on note $X_i \in L^2(\mathfrak{F}_{t_i})$). L'intégrale stochastique de φ est égale par définition à la variable aléatoire

$$(2.2.4) \quad \int_{R_+} \varphi(t) dB(t) = \sum_{i=0}^{n-1} X_i [B(t_{i+1}) - B(t_i)] .$$

Lemme 2.2.2.

Sous les hypothèses précédentes, on a

$$(2.2.5) \quad E \left[\int_{R_+} \varphi(t) dB(t) \right] = 0 ,$$

$$(2.2.6) \quad E \left(\left[\int_{R_+} \varphi(t) dB(t) \right]^2 \right) = E \left[\int_{R_+} \varphi(t)^2 dt \right]$$

DEMONSTRATION. La première affirmation résulte de ce que $X_i(B(t_{i+1}) - B(t_i))$ est de carré

intégrable, comme produit de deux variables aléatoires de carré intégrable et indépendantes, et d'espérance nulle car un des deux termes du produit l'est. De nouveau par l'indépendance,

$$E \{ [X_i (B(t_{i+1}) - B(t_i))]^2 \} = E(X_i^2) E[(B(t_{i+1}) - B(t_i))^2] = E(X_i^2) (t_{i+1} - t_i)$$

Il reste à calculer, pour $i < j$:

$$E [X_i X_j (B(t_{i+1}) - B(t_i)) (B(t_{j+1}) - B(t_j))]]$$

Mais $B(t_{j+1}) - B(t_j)$ et $X_i X_j (B(t_{i+1}) - B(t_i))$ sont indépendantes car la seconde v.a.r. est \mathcal{F}_{t_i} mesurable, et $B(t_{j+1}) - B(t_j)$ est d'espérance nulle, donc la quantité ci-dessus est nulle. Le lemme est démontré. \square

Le lemme précédent va permettre de prolonger l'intégrale stochastique $\int_{R_+} \varphi(t) dB(t)$ à toutes les fonctions $\varphi \in M^2(R_+)$ grâce au résultat suivant.

Lemme 2.2.3.

Les fonctions φ en escaliers de la forme (2.2.3) forment un sous-ensemble dense de l'espace hilbertien $M^2(R_+)$.

DEMONSTRATION. a) Pour tout entier $n \geq 1$, considérons l'opérateur linéaire P_n défini sur l'espace $L^2(R_+)$ par

$$P_n f(t) = n \sum_{i=1}^{n^2} \int_{(i-1)/n}^{i/n} f(s) ds \cdot 1_{]i/n, (i+1)/n]}(t)$$

Sur chaque intervalle $]i/n, (i+1)/n]$ ($1 \leq i \leq n^2$), la fonction $P_n f$ est donc égale à sa moyenne sur l'intervalle **précédent** $]i/n, (i-1)/n]$ tandis qu'en dehors de $]1/n, (n+1)/n^2]$ elle vaut zéro. L'inégalité de Schwarz montre alors que

$$[P_n f(t)]^2 \leq n \int_{(i-1)/n}^{i/n} f(s)^2 ds \quad \text{si } t \in]i/n, (i+1)/n] \quad (1 \leq i \leq n^2)$$

et s'il s'en suit que

$$\int_{R_+} [P_n f(t)]^2 dt \leq \int_{R_+} f(t)^2 dt \quad (f \in L^2(R_+)) ;$$

l'opérateur P_n contracte donc la norme L^2 .

En outre il est facile de voir que lorsque $n \rightarrow \infty$

$$\|P_n f - f\|_{L^2(R_+)} \rightarrow 0$$

pour toute fonction continue à support compact et comme l'espace C_c de ces fonctions est dense dans $L^2(R_+)$, la convergence précédente est encore valable pour toute fonction f de $L^2(R_+)$.

b) Soit maintenant φ une fonction aléatoire dans $M^2(R_+)$. Alors pour tout entier $n \geq 1$, la fonction

$$P_n \varphi(t, \omega) := P_n [\varphi(\cdot, \omega)](t)$$

est du type (2.1.3), avec $t_i = i/n$, notamment parce que $\int_{(i-1)/n}^{i/n} \varphi(t, \cdot) dt$ est $\mathcal{F}_{i/n}$ mesurable de par la propriété (2.2.1) de progressive mesurabilité de φ . Mais $P_n \varphi$ converge vers φ dans $M^2(R_+)$ lorsque $n \rightarrow \infty$ car

$$\|P_n \varphi - \varphi\|_{M^2(R_+)}^2 = E \left(\int_{R_+} [P_n [\varphi(\cdot, \omega)](t) - \varphi(t, \omega)]^2 dt \right)$$

et comme pour presque tout ω , $\varphi(\cdot, \omega) \in L^2(R_+)$ et donc d'après (a)

$$\int [P_n [\varphi(\cdot, \omega)](t) - \varphi(t, \omega)]^2 dt \rightarrow 0 \quad \text{lorsque } n \rightarrow \infty ,$$

l'espérance ci-dessus tend aussi vers zéro lorsque $n \rightarrow \infty$ par convergence dominée puisque

$$\begin{aligned} \int [P_n [\varphi(\cdot, \omega)](t) - \varphi(t, \omega)]^2 dt & \\ & \leq (\|P_n [\varphi(\cdot, \omega)]\|_{L^2(R_+)} + \|\varphi(\cdot, \omega)\|_{L^2(R_+)})^2 \\ & \leq 4 \|\varphi(\cdot, \omega)\|_{L^2(R_+)}^2 \end{aligned}$$

et $E(\|\varphi(\cdot, \omega)\|_{L^2(R_+)}^2) < \infty$. \square

Revenons alors à la définition de l'intégrale stochastique. L'application

$$\varphi \rightarrow \int_{R_+} \varphi(t) dB(t)$$

que nous avons définie sur le sous-espace des fonctions en escaliers de la forme (2.1.3), sous-espace qui est dense dans $M^2(R_+)$, prend ses valeurs dans l'espace $L^2(\Omega)$ des v.a. de carrés intégrables et est isométrique d'après (2.2.6). Cette application se prolonge alors par continuité à tout l'espace $M^2(R_+)$ de manière unique, par exemple en posant pour toute fonction aléatoire $\varphi \in M^2(R_+)$

$$\begin{aligned} (2.2.7) \quad \int_{R_+} \varphi(t) dB(t) &= \lim_{n \rightarrow \infty} \int_{R_+} P_n \varphi(t) dB(t) \\ &= \lim_{n \rightarrow \infty} n \sum_{i=1}^{n^2} \int_{(i-1)/n}^{i/n} \varphi(s) ds [B((i+1)/n) - B(i/n)] , \end{aligned}$$

après avoir remarqué que le second membre est le n^{e} terme d'une suite de Cauchy dans $L^2(\Omega)$ puisque d'après (2.2.6)

$$\begin{aligned} & E \left[\int_{R_+} P_{n'} \varphi(t) dB(t) - \int_{R_+} P_n \varphi(t) dB(t) \right]^2 \\ &= E \left(\int_{R_+} [(P_{n'} \varphi - P_n \varphi)(t)]^2 dt \right) \rightarrow 0 \end{aligned}$$

lorsque $n, n' \rightarrow \infty$. La limite dans la formule (2.2.7) ci-dessus est donc une limite en moyenne quadratique sur Ω .

Résumons-nous.

Proposition 2.2.4. (Intégrale stochastique)

A toute fonction aléatoire $\varphi(t, \omega)$ dans $M^2(R_+)$ est associée de manière unique une v.a. de carré intégrale notée $\int_{R_+} \varphi(t) dB(t)$ telle que

$$E \left[\int_{R_+} \varphi(t) dB(t) \right] = 0 \quad , \quad E \left[\left(\int_{R_+} \varphi(t) dB(t) \right)^2 \right] = E \left[\int_{R_+} \varphi(t)^2 dt \right] \quad ,$$

de telle manière que si φ est de la forme (2.2.3), $\int_{R_+} \varphi(t) dB(t)$ soit donnée par (2.2.4).

2.3. L'intégrale stochastique de Itô comme Martingale.

Si $\varphi(\cdot, \omega)$ est progressivement mesurable, il en est de même de $\varphi 1_{[0,t]}$ pour chaque $t \in R_+$, où

$$\varphi 1_{[0,t]}(s, \omega) = \varphi(s, \omega) 1_{\{0 \leq s \leq t\}} \quad (s \in R_+, \omega \in \Omega) ;$$

il est alors naturel de poser

$$\int_0^t \varphi(s) dB(s) = \int_{R_+} \varphi 1_{[0,t]}(s) dB(s) \quad (t \in R_+)$$

pourvu que $\varphi 1_{[0,t]} \in M^2(R_+)$ pour tout $t \in R_+$, c'est-à-dire pourvu que $\varphi \in M^2$ (voir la définition 2.2.1). Quelles sont les propriétés de la fonction aléatoire

$$\left(\int_0^t \varphi(s) dB(s) \quad , \quad t \in R_+ \right)$$

ainsi construite ?

Il est d'abord facile de vérifier que l'application $t \rightarrow \int_0^t \varphi(s) dB(s)$ est continue en

moyenne quadratique sur R_+ ; nous admettrons que de plus les trajectoires de cette fonction aléatoire sont continues.

Désignons ensuite par

$$L^2(\mathfrak{F}_t) \quad (t \in R_+)$$

l'espace des v.a.r. de carrés intégrales, mesurables par rapport à \mathfrak{F}_t et par

$$E^{\mathfrak{F}_t}(\cdot)$$

le projecteur orthogonal de L^2 sur ce sous-espace. Alors à condition de poser

$$\int_t^u \varphi(s) dB(s) = \int_0^u \varphi(s) dB(s) - \int_0^t \varphi(s) dB(s) \quad (0 \leq t \leq u) ,$$

on a la

Proposition 2.3.1.

Soit $\varphi \in M^2$. Pour tout $t \in R_+$

$$\int_0^t \varphi(s) dB(s) \in L^2(\mathfrak{F}_t)$$

tandis que pour tous $t \leq u$ dans R_+

$$E^{\mathfrak{F}_t} \left[\int_t^u \varphi(s) dB(s) \right] = 0$$

ou ce qui est équivalent

$$E^{\mathfrak{F}_t} \left[\int_0^u \varphi(s) dB(s) \right] = \int_0^t \varphi(s) dB(s) \quad (t \leq u) .$$

On dit alors que $(\int_0^t \varphi(s) dB(s), t \in R_+)$ est une *martingale* par rapport à $(\mathfrak{F}_t, t \in R_+)$.

De plus si $\varphi, \psi \in M^2$ on a, quels que soient $t \leq u$:

$$E^{\mathfrak{F}_t} \left[\int_t^u \varphi(s) dB(s) \int_t^u \psi(s) dB(s) \right] = E^{\mathfrak{F}_t} \left[\int_t^u \varphi(s) \psi(s) ds \right]$$

DEMONSTRATION. Nous nous contenterons de démontrer cette proposition lorsque φ et ψ sont des fonctions en escaliers, le cas général s'obtenant ensuite par un passage à la limite.

D'abord, d'après (2.2.3) et (2.2.4), l'intégrale $\int_0^t \varphi(s) dB(s)$ est une somme de termes de la forme $X_i [B(t_{i+1}) - B(t_i)]$ avec $0 \leq t_i \leq t_{i+1} \leq t$ et $X_i \in L^2(\mathfrak{F}_{t_i})$, donc appartient à $L^2(\mathfrak{F}_t)$. De même $\int_t^u \varphi(s) dB(s)$ est une somme de termes de la forme précédente avec cette fois $t \leq t_i \leq t_{i+1} \leq u$ et $X_i \in L^2(\mathfrak{F}_{t_i})$; comme $B(t_{i+1}) - B(t_i)$ est une v.a.r. centrée indépendante de \mathfrak{F}_{t_i} , on a alors

$$E^{\mathfrak{F}_{t_i}} (X_i [B(t_{i+1}) - B(t_i)]) = X_i E^{\mathfrak{F}_{t_i}} [B(t_{i+1}) - B(t_i)] = 0$$

et par suite comme $L^2(\mathfrak{F}_t) \subset L^2(\mathfrak{F}_{t_i})$ puisque $t \leq t_i$, en projetant l'égalité précédente sur $L^2(\mathfrak{F}_t)$, on obtient que

$$E^{\mathfrak{F}_t}(X_i [B(t_{i+1}) - B(t_i)]) = 0 \quad .$$

La démonstration de la dernière égalité est analogue, mais utilise cette fois que

$$E^{\mathfrak{F}_{t_i}} ([B(t_{i+1}) - B(t_i)]^2) = t_{i+1} - t_i \quad . \quad \square$$

2.4. Le cas vectoriel.

Soit maintenant un mouvement brownien k -dimensionnel $\{B(t)\}$, et une f.a. à valeurs matrices $k \times d$ $\{\varphi(t)\}$. Supposons que pour tout $1 \leq i \leq d$ et $1 \leq j \leq k$, $\{\varphi_{ij}(t)\}$ soit un élément de $M^2(R_+)$ alors on peut définir

$$\int_{R_+} \varphi_{ij}(t) dB_j(t)$$

pour $1 \leq i \leq d$, $1 \leq j \leq k$ et on définit

$$\int_{R_+} \varphi(t) dB(t)$$

comme le vecteur aléatoire de dimension d dont la i -ème coordonnée est donnée par :

$$\sum_{j=1}^k \int_{R_+} \varphi_{ij}(t) dB_j(t)$$

Lorsque chaque $\{\varphi_{ij}(t)\}$ appartient à M^2 , on définit de même la f.a. de dimension d :

$$\int_0^t \varphi(s) dB(s) \quad ; \quad t \geq 0$$

qui possède les propriétés suivantes :

Proposition 2.4.1.

Soit $\varphi, \psi \in (M^2)^{d \times k}$. Alors pour tout $t \geq 0$,

- (i) $\int_0^t \varphi(s) dB(s)$ est un v.a. \mathcal{F}_t mesurable de dimension d
- (ii) $E \left(\int_0^t \varphi(s) dB(s) \right) = 0$
- (iii) $E \left[\left(\int_0^t \varphi(s) dB(s) \right) \left(\int_0^t \psi(s) dB(s) \right)^t \right] = E \int_0^t \varphi(s) \psi(s)^t ds$
- (iv) $E \left(\int_0^t \varphi(s) dB(s) \right) \int_0^t \psi(s) dB(s) = E \int_0^t \text{Tr} [\varphi(s) \psi(s)^t] ds$

Preuve. C'est une conséquence de la Proposition 2.3.1. (avec $t = 0$). \square

2.5. La formule d'Itô.

Considérons une fonction $x \in C^1(\mathbb{R}_+)$, et $\Phi \in C^1(\mathbb{R})$. Alors :

$$\Phi(x(t)) = \Phi(x(s)) + \int_0^t \Phi'(x(s)) x'(s) ds$$

cette formule élémentaire peut se réécrire (au moins de façon symbolique) :

$$\Phi(x(t)) = \Phi(x(s)) + \int_0^t \Phi'(x(s)) dx(s)$$

Puisque l'on sait définir des intégrales par rapport au mouvement brownien scalaire $\{B(t)\}$, on peut se demander si l'on a une formule analogue à celle que nous venons d'écrire pour une fonction de classe C^1 . Considérons tout d'abord le cas de la fonction $f(x) = x^2$. On a

$$\begin{aligned} B^2(t) &= \sum_{i=1}^n (B^2(t_i^n) - B^2(t_{i-1}^n)) \\ &= 2 \sum_{i=1}^n B(t_{i-1}^n) (B(t_i^n) - B(t_{i-1}^n)) + (B(t_i^n) - B(t_{i-1}^n))^2 \end{aligned}$$

En combinant la Remarque 3.4 et la Proposition 1.3.3, on obtient :

$$B^2(t) = 2 \int_0^t B(s) dB(s) + t$$

Plus généralement, on a ($\{B(t)\}$ désigne toujours un mouvement brownien scalaire) :

Proposition 2.5.1.

Si $\Phi \in C^2_b(R)$, pour tout $t > 0$,

$$\Phi(B(t)) = \Phi(B(s)) + \int_0^t \Phi'(B(s)) dB(s) + \frac{1}{2} \int_0^t \Phi''(B(s)) ds$$

DEMONSTRATION. On pose $t_i^n = ti/n$. Il résulte de la formule de Taylor que pour des $\theta_i^n \in]t_{i-1}^n, t_i^n]$

$$\begin{aligned} \Phi[B(t)] &= \Phi[B(0)] + \sum_{i=1}^n (\Phi[B(t_i^n)] - \Phi[B(t_{i-1}^n)]) \\ &= \Phi(B(s)) + \sum_{i=1}^n \Phi'(B(t_{i-1}^n)) (B(t_i^n) - B(t_{i-1}^n)) \\ &\quad + \frac{1}{2} \sum_{i=1}^n \Phi''(B(\theta_i^n)) (B(t_i^n) - B(t_{i-1}^n))^2 \\ &= \Phi(B(0)) + \int_0^t \Phi'(B_s) dB_s + \frac{1}{2} \int_0^t \Phi''(B_s) ds \end{aligned}$$

En effet, la convergence dans la dernière somme ci-dessus résulte de la continuité de $B(t)$ et des deux convergences :

$$\begin{aligned} & \left| \sum_{i=1}^n (\Phi''(B(\theta_i^n)) - \Phi''(B(t_{i-1}^n))) (B(t_i^n) - B(t_{i-1}^n))^2 \right| \\ & \leq \sup_i |\Phi''(B(\theta_i^n)) - \Phi''(B(t_{i-1}^n))| \sum_{i=1}^n (B(t_i^n) - B(t_{i-1}^n))^2 \\ & \rightarrow 0 \text{ en moyenne quadratique} \end{aligned}$$

$$\begin{aligned} & E \left[\left(\sum_{i=1}^n \Phi''(B(t_{i-1}^n)) [(B(t_i^n) - B(t_{i-1}^n))^2 - (t_i^n - t_{i-1}^n)] \right)^2 \right] \\ & = \sum_{i=1}^n E \{ (\Phi''(B(t_{i-1}^n)) [(B(t_i^n) - B(t_{i-1}^n))^2 - (t_i^n - t_{i-1}^n)])^2 \} \\ & \leq \sup_x |\Phi''(x)|^2 \sum_{i=1}^n E [(B(t_i^n) - B(t_{i-1}^n))^2 - (t_i^n - t_{i-1}^n)]^2 \\ & = \sup_x |\Phi''(x)|^2 \sum_{i=1}^n (t_i^n - t_{i-1}^n)^2 \rightarrow 0 \end{aligned}$$

et du fait que :

$$\sum_{i=1}^n \Phi''(B_{t_{i-1}^n}) (t_i^n - t_{i-1}^n) \rightarrow \int_0^t \Phi''(B_s) ds \quad \square$$

Remarquons que le calcul différentiel d'Itô diffère du calcul différentiel usuel par le terme complémentaire faisant intervenir la dérivée seconde. L'apparition de ce terme est due au fait que $\sum_{i=1}^n (B(t_i^n) - B(t_{i-1}^n))^2 \rightarrow t > 0$, contrairement à ce qui se passe avec une fonction de classe C^1 .

Nous allons maintenant énoncer une formule plus générale (toujours dans le cas scalaire !). Nous appellerons *fonction aléatoire d'Itô* une fonction aléatoire $\{X(t)\}$ qui se met sous la forme :

$$(*) \quad X(t) = X(0) + \int_0^t \psi(s) ds + \int_0^t \varphi(s) dB(s)$$

où $\varphi, \psi \in M^2$.

Théorème 2.5.2.

Si $\Phi \in C_b^2(\mathbb{R})$ et la f.a.r. de Itô $\{X(t)\}$ est donnée par (*), alors pour tout $t \geq 0$,

$$\begin{aligned} \Phi(X(t)) &= \Phi(X(0)) + \int_0^t \Phi'(X(s)) \psi(s) ds + \\ &+ \int_0^t \Phi'(X(s)) \varphi(s) dB(s) + \frac{1}{2} \int_0^t \Phi''(X(s)) \varphi^2(s) ds. \end{aligned}$$

DEMONSTRATION. Admettons que l'on s'est ramené à établir la formule dans le cas où les variables aléatoires $\int_0^t |\psi(s)| ds$ et $\sup_{0 \leq s \leq t} \varphi^2(s)$ sont bornées et $\{\varphi(s)\}$ est continu.

On fait le même calcul que dans la preuve précédente ($t_i^n = ti/n$) :

$$\begin{aligned} \Phi(X(t)) &= \Phi(X(0)) + \sum_{i=1}^n [\Phi(X(t_i^n)) - \Phi(X(t_{i-1}^n))] \\ &= \Phi(X(0)) + \sum_{i=1}^n \Phi'(X(t_{i-1}^n)) \int_{t_{i-1}^n}^{t_i^n} \psi(s) ds + \\ &+ \sum_{i=1}^n \Phi'(X(t_{i-1}^n)) \int_{t_{i-1}^n}^{t_i^n} \varphi(s) dB(s) + \frac{1}{2} \sum_{i=1}^n \Phi''(X(t_i^n)) (X(t_i^n) - X(t_{i-1}^n))^2 \end{aligned}$$

le passage à la limite dans les deux premiers termes est assez facile. Explicitons le second :

$$\begin{aligned} E \left(\left| \int_0^t \Phi'(X(s)) \varphi(s) dB(s) - \sum_{i=1}^n \Phi'(X(t_{i-1}^n)) \int_{t_{i-1}^n}^{t_i^n} \varphi(s) dB(s) \right|^2 \right) \\ = E \sum_{i=1}^n \int_{(i-1)/n}^{1/n} |\Phi'(X(s)) - \Phi'(X(t_{i-1}^n))|^2 \varphi^2(s) ds \end{aligned}$$

qui tend vers zéro par convergence dominée, et par la continuité de $\{X(t)\}$.

En ce qui concerne le dernier terme,

$$\begin{aligned} (X(t_{i+1}^n) - X(t_i^n))^2 &= \left(\int_{t(i-1)/n}^{ti/n} \psi(s) ds \right)^2 + 2 \int_{t(i-1)/n}^{ti/n} \varphi(s) ds \int_{t(i-1)/n}^{ti/n} \varphi(s) dB \\ &\quad + \left(\int_{t(i-1)/n}^{ti/n} \varphi(s) dB(s) \right)^2 \end{aligned}$$

Les sommes par rapport à i des deux premiers termes du membre de droite sont majorées par :

$$\sup_i \left| \int_{t(i-1)/n}^{ti/n} \psi(s) ds + 2 \int_{t(i-1)/n}^{ti/n} \varphi(s) dB(s) \right| \times \int_0^t |\psi(s)| ds$$

et le premier terme de ce produit tend vers zéro, par continuité uniforme d'une fonction continue sur le compact $[0, t]$.

Il est maintenant facile de voir que la preuve se termine comme à la Proposition 2.5, à condition d'établir la généralisation suivante de la Proposition 1.3.3 :

Lemme 2.5.3.

Si le processus $\{\varphi(t)\}$ est continu et borné

$$\sum_{i=1}^n \left[\left(\int_{t(i-1)/n}^{ti/n} \varphi(s) dB(s) \right)^2 - \int_{t(i-1)/n}^{ti/n} \varphi^2(s) ds \right]$$

tend vers 0 dans $L^1(\Omega)$ quand $n \rightarrow \infty$.

Preuve : Soit $\{\varphi_n(s)\}_{n \in \mathbb{N}}$ une suite de f.a.r. qui convergent vers $\{\varphi(s)\}$ dans $M^2(0, t)$.

$$\begin{aligned} &\sum_i \left(\int_{t(i-1)/n}^{ti/n} \varphi(s) dB(s) \right)^2 - \sum_i \int_{t(i-1)/n}^{ti/n} \varphi^2(s) ds \\ (**) \quad &= \sum_i \left[\left(\int_{t(i-1)/n}^{ti/n} \varphi(s) dB(s) \right)^2 - \left(\int_{t(i-1)/n}^{ti/n} \varphi_n(s) dB(s) \right)^2 \right] \\ &+ \sum_i \left[\left(\int_{t(i-1)/n}^{ti/n} \varphi_n(s) dB(s) \right)^2 - \int_{t(i-1)/n}^{ti/n} \varphi_n^2(s) ds \right] \\ &+ \int_0^t [\varphi_n^2(s) - \varphi^2(s)] ds \end{aligned}$$

Etudions successivement le dernier, le premier et le second terme dans le membre de droite de (**). On va utiliser de façon répétée l'inégalité de Cauchy-Schwarz.

$$\begin{aligned} E \left| \int_0^t [\varphi_n^2(s) - \varphi^2(s)] ds \right| &\leq (E \int_0^t (\varphi_n(s) - \varphi(s))^2 ds)^{1/2} \\ &\quad \times (E \int_0^t (\varphi_n(s) + \varphi(s))^2 ds) \end{aligned}$$

qui tend vers 0 quand $n \rightarrow \infty$.

$$\begin{aligned} E \left| \sum_i \left(\int_{t_{(i-1)/n}}^{ti/n} (\varphi(s) - \varphi_n(s)) dB(s) \right) \left(\int_{t_{(i-1)/n}}^{ti/n} (\varphi(s) + \varphi_n(s)) dB(s) \right) \right| \\ \leq [E \sum_i \left(\int_{t_{(i-1)/n}}^{ti/n} (\varphi(s) - \varphi_n(s)) dB(s) \right)^2]^{1/2} \\ \times [E \sum_i \left(\int_{t_{(i-1)/n}}^{ti/n} (\varphi(s) + \varphi_n(s)) dB(s) \right)^2]^{1/2} \\ = (E \int_0^t (\varphi(s) - \varphi_n(s))^2 ds)^{1/2} (E \int_0^t (\varphi(s) + \varphi_n(s))^2 ds)^{1/2} \end{aligned}$$

qui à nouveau tend vers 0 quand $n \rightarrow \infty$. On choisit maintenant

$$\varphi_n(s) = \sum_{i=1}^n \varphi(t_{i-1}^n) 1_{]t_{i-1}^n, t_i^n]}(s)$$

Il nous reste alors à étudier la différence :

$$\sum_i \varphi^2(t_{i-1}^n) [(B(t_i^n) - B(t_{i-2}^n))^2 - (t_i^n - t_{i-1}^n)]$$

dont on sait déjà (cf. Preuve de la Proposition 2.5.1) qu'elle tend vers 0 en moyenne quadratique. \square

Il nous reste à énoncer la formule d'Itô dans le cas vectoriel, qui se démontre exactement comme dans le cas scalaire. On appellera f.a. d'Itô de dimension d une f.a. $\{X(t)\}$ à valeurs dans R^d qui s'écrit :

$$(***) \quad X(t) = X(0) + \int_0^t \psi(s) ds + \int_0^t \varphi(s) dB(s), t \geq 0$$

où $\{B(t)\}$ est un mouvement brownien de dimension k (k entier quelconque), $\psi(t) = (\psi_1(t), \dots, \psi_d(t))'$ et chaque $\{\psi_i(t)\}$ appartient à M^2 , $\varphi(t) = (\varphi_{ij}(t))_{1 \leq i \leq d, 1 \leq j \leq k}$ et chaque $\{\varphi_{ij}(t)\}$ appartient à M^2 . L'égalité ci-dessus s'écrit donc composante par composante :

$$X_i(t) = X_i(0) + \int_0^t \psi_i(s) ds + \sum_{j=1}^k \int_0^t \varphi_{ij}(s) dB_j(s)$$

Etant donné $\Phi \in C^2(R^d)$, on notera $\Phi'(x)$ le vecteur ligne $\left(\frac{\partial \Phi}{\partial x_1}(x), \dots, \frac{\partial \Phi}{\partial x_d}(x) \right)$ et $\Phi''(x)$ la matrice $\left(\frac{\partial^2 \Phi}{\partial x_i \partial x_j}(x) \right)$.

Théorème 2.5.4.

Etant donné $\{X(t)\}$ une f.a. d'Itô de la forme (***) , et $\Phi \in C_b^2(\mathbb{R}^d)$, on a :

$$\begin{aligned} \Phi(X_t) &= \Phi(X_0) + \int_0^t \Phi'(X_s) \psi(s) ds + \\ &+ \int_0^t \Phi'(X_s) \varphi(s) dB(s) + \frac{1}{2} \int_0^t \text{Tr} [\Phi''(X_s) \varphi(s) \varphi^t(s)] ds \quad \square \end{aligned}$$

2.6. Remarque sur un autre type d'intégrale stochastique.

Dans cette section, toutes les f.a. seront scalaires. On vérifie aisément que dans le cas d'un intégrand continu $\{\varphi(t)\}$ l'intégrale d'Itô peut se définir comme :

$$\int_0^t \varphi(s) dB(s) = \lim_{n \rightarrow \infty} \sum_{i=1}^n \varphi(t_{i-1}^n) (B(t_i^n) - B(t_{i-1}^n))$$

avec à nouveau $t_i^n = \frac{i}{n} t$. En particulier,

$$\int_0^t B(s) dB(s) = \lim_{n \rightarrow \infty} \sum_{i=1}^n B(t_{i-1}^n) (B(t_i^n) - B(t_{i-1}^n))$$

Remarquons que l'on approche l'intégrale d'Itô par des « sommes de Darboux » un peu particulières, où l'on doit évaluer l'intégrand à l'extrémité gauche de l'intervalle $[t_{i-1}^n, t_i^n]$. Au vu du calcul fait au début de la section 2.5, il est assez clair qu'un autre choix conduit à une autre limite. En particulier, la limite suivante existe en moyenne quadratique :

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{B(t_{i-1}^n) + B(t_i^n)}{2} (B(t_{i+1}^n) - B(t_i^n))$$

Elle est appelée intégrale de Stratonovitch, et notée :

$$\int_0^t B(s) \circ dB(s)$$

Il est alors immédiat que

$$B^2(t) = 2 \int_0^t B(s) \circ dB(s)$$

Plus généralement, si $F \in C^1(\mathbb{R})$, on pose (la limite étant une limite en probabilité) :

$$\begin{aligned} \int_0^t F(B(s)) \circ dB(s) &= \lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{F(B(t_{i-1}^n)) + F(B(t_i^n))}{2} (B(t_i^n) - B(t_{i-1}^n)) \\ &= \int_0^t F(B(s)) dB(s) + \frac{1}{2} \int_0^t F'(B(s)) ds \end{aligned}$$

et en outre, si $\Phi \in C^2(\mathbb{R})$,

$$\Phi(B(t)) = \Phi(0) + \int_0^t \Phi'(B(s)) \circ dB(s)$$

La question qui se pose naturellement est de savoir quelles sont les f.a. dont on peut définir l'intégrale au sens de Stratonovich. Essentiellement, ce sont les f.a. qui se mettent sous la forme $\{F(X(t))\}$, avec $\{X(t)\}$ f.a. d'Itô et $F \in C^1(R)$.

Soit donc $\{X(t)\}$ un processus d'Itô de la forme

$$X(t) = X(0) + \int_0^t \psi(s) ds + \int_0^t \varphi(s) dB(s)$$

et $F \in C^1(R)$. Alors

$$\begin{aligned} \int_0^t F(X(s)) \circ dB(s) &= \sum_{i=1}^n \frac{F(X(t_{i-1}^n)) + F(X(t_i^n))}{2} (B(t_i^n) - B(t_{i-1}^n)) \\ &= \int_0^t F(X(s)) dB(s) + \frac{1}{2} \int_0^t F'(X(s)) \varphi(s) ds \end{aligned}$$

le dernier terme de la formule ci-dessus, qui est le terme de « correction » entre l'intégrale de Stratonovitch et celle d'Itô, s'obtient par un calcul analogue à celui fait dans la démonstration de la formule d'Itô.

En outre, si

$$X(t) = X(0) + \int_0^t \psi(s) ds + \int_0^t \varphi(s) \circ dB(s)$$

où $\varphi(s) = G(Y(s))$, $\{Y(t)\}$ f.a. d'Itô et $G \in C^1(R)$, et $\Phi \in C^2(R)$, alors on a la formule d'Itô-Stratonovitch :

$$\begin{aligned} F(X(t)) &= F(X(0)) + \int_0^t F'(X(s)) \psi(s) ds + \\ &+ \int_0^t F'(X(s)) \varphi(s) \circ dB(s) \end{aligned}$$

formule que l'on écrit souvent sous la forme condensée

$$F(X(t)) = F(X(0)) + \int_0^t F'(X(s)) \circ dX(s)$$

On voit donc que contrairement au calcul d'Itô, le calcul de Stratonovitch obéit aux règles usuelles du calcul différentiel. D'un autre côté, l'intégrale d'Itô possède d'autres propriétés très intéressantes, d'une part le fait que l'on peut calculer ses deux premiers moments et d'autre part le fait que la classe de processus que l'on peut intégrer au sens d'Itô est beaucoup plus vaste.

2.7. Généralisation de l'intégrale d'Itô.

Nous aurons besoin, en particulier à la section suivante, d'utiliser la formule d'Itô avec des fonctions Φ de classe C^2 , mais non nécessairement $C^2_{\mathbb{R}}$ (par exemple $\Phi(x) = x^2$, $\Phi(x) = e^x$). L'extension de la formule d'Itô à de telles fonctions est une conséquence immédiate d'une extension de l'intégrale d'Itô à la classe suivante d'intégrands.

Définition 2.7.1.

On notera $M_{loc}^2(0, T)$ l'espace des f.a.r. $\{\varphi(t) ; 0 \leq t \leq T\}$ qui vérifient :

(i) φ est progressivement mesurable

(ii) $\int_0^T \varphi(t)^2 dt < \infty$ p.s.

On notera en outre $M_{loc}^2 = \bigcap_{T>0} M_{loc}^2(0, T)$. \square

La généralisation correspondante de l'intégrale d'Itô s'appuie sur un outil technique qui est par ailleurs très utile :

Définition 2.7.2.

On appelle *temps d'arrêt* (par rapport à la filtration $\mathcal{F} = (\mathcal{F}_t)$) toute v.a. τ à valeurs dans $\mathbb{R}_+ \cup \{+\infty\}$ qui vérifie :

$$\forall t \geq 0, \{ \omega ; \tau(\omega) \leq t \} \in \mathcal{F}_t \quad \square$$

Les temps d'arrêt dont nous aurons besoin sont de la forme suivante :

Proposition 2.7.3.

Soit $(X(t), t \geq 0)$ une f.a. *progressivement mesurable et continue* à valeurs dans \mathbb{R}^d , et F un sous-ensemble *fermé* de \mathbb{R}^d . Alors la v.a. τ définie par :

$$\tau(\omega) = \begin{cases} \inf \{ t ; X(t, \omega) \in F \} & \text{si un tel } t \text{ existe} \\ +\infty & \text{sinon} \end{cases}$$

est un temps d'arrêt.

Preuve : Il suffit de remarquer que

$$\{ \omega ; \tau(\omega) \leq t \} = \bigcap_{p \in \mathbb{N}} \bigcup_{s \in \mathbb{Q} \cap [0, t]} \{ \omega ; X(s, \omega) \in F_{1/p} \}$$

où $F_{1/p}$ désigne l'ouvert $\{ x ; d(x, F) < 1/p \}$, et de se rappeler qu'une tribu est fermée pour les réunions et intersections dénombrables. \square

Etant donné $(\varphi(t)) \in M_{loc}^2$, et $n \in \mathbb{N}$, on pose :

$$\tau_n(\omega) = \begin{cases} \inf \{ t ; \int_0^t \varphi(s, \omega)^2 ds \geq n \} & \text{si un tel } t \text{ existe} \\ +\infty & \text{sinon} \end{cases}$$

Alors τ_n est un temps d'arrêt pour tout n , $\tau_n \uparrow +\infty$ quand $n \rightarrow +\infty$ et on vérifie aisément que

$$(1_{[0, \tau_n]}(t) \varphi(t)) \in M^2, \quad \forall n$$

On peut donc définir :

$$\int_0^t 1_{[0, \tau_n]}(s) \varphi(s) dB(s), t \geq 0$$

Il reste à montrer que quand $n \rightarrow \infty$,

$$\int_0^t 1_{[0, \tau_n]}(s) \varphi(s) dB(s) \text{ converge p.s.}$$

La limite de cette suite est par définition

$$\int_0^t \varphi(s) dB(s)$$

Nous ne détaillons pas le raisonnement. Contentons-nous d'énoncer le résultat technique qui est la clé de la construction (et qui est très intuitif !):

Lemme 2.7.4.

Si $(\varphi(t)) \in M^2$ et τ est un temps d'arrêt, alors $(1_{[0, \tau]}(t) \varphi(t)) \in M^2$, et

$$\int_0^t 1_{[0, \tau]}(s) \varphi(s) dB(s) = \int_0^{t \wedge \tau} \varphi(s) dB(s), t \geq 0 \quad \square$$

Remarquons enfin que si $\varphi \in M_{loc}^2$, $\{\int_0^t \varphi(s) dB(s), t \geq 0\}$ est une f.a.r. à trajectoires continues, mais la v.a.r.

$$\int_0^t \varphi(s) dB(s)$$

n'est pas nécessairement intégrable !

2.8. Le théorème de Girsanov.

Le but de cette section est de généraliser la formule de Cameron-Martin de la section 1.6. Nous allons tout d'abord réécrire la formule de Cameron-Martin sous la forme sous laquelle nous la généraliserons.

Soit (Ω, \mathcal{A}, P) un espace de probabilité muni d'une filtration $\mathcal{F} = (\mathcal{F}_t, t \in R_+)$ telle que $\mathcal{A} = \mathcal{F}_\infty$ et $(B(t), t \in R_+)$ un \mathcal{F} -mouvement brownien. Compte tenu de ce qui va suivre, il vaudrait mieux préciser que $(B(t), t \in R_+)$ est un P - \mathcal{F} -mouvement brownien. Soit enfin $f \in L^2(R_+)$. On pose :

$$\bar{B}(t) = B(t) - \int_0^t f(s) ds, t \in R_+$$

$$Z(t) = \exp \left(\int_0^t f(s) B(s) - \frac{1}{2} \int_0^t f^2(s) ds \right), t \in R_+$$

$$Z = Z(\infty)$$

En notant que $Z(t)$ est l'exponentielle d'une v.a.r. gaussienne, on constate par un calcul explicite que $E(Z(t)) = E(Z) = 1$. Définissons alors une nouvelle probabilité Q sur (Ω, \mathcal{A}) par :

$$\frac{dQ}{dP} = Z$$

$$\text{i.e.} \quad Q(A) = \int_A Z(\omega) dP(\omega), \quad A \in \mathcal{A}$$

Si F est une fonctionnelle positive ou bornée définie sur $C(R_+)$, on déduit de la définition de Q et du théorème 1.6.1. (avec $X = \bar{B}$, $Y = \int_0^\infty f(t) dB(t)$) :

$$\begin{aligned} E_Q(F(\bar{B})) &= E_P \left[F(B - \int_0^\infty f(s) ds) Z \right] \\ &= E_P [F(B)] \end{aligned}$$

C'est-à-dire que la loi de la f.a.r. $(\bar{B}(t), t \in R_+)$ sous Q coïncide avec la loi de la f.a.r. $(B(t), t \in R_+)$ sous P , i.e. $(\bar{B}(t), t \in R_+)$ est un Q - \mathcal{F} -mouvement brownien.

Remarquons que $\frac{Z}{Z_t}$ est indépendant de \mathcal{F}_t , donc

$$\begin{aligned} E^{\mathcal{F}_t}(Z/Z(t)) &= E(Z/Z(t)) \\ &= 1 \end{aligned}$$

ce qui entraîne immédiatement :

$$E^{\mathcal{F}_t}(Z) = Z_t$$

d'où l'on déduit que si $A \in \mathcal{F}_t$,

$$Q(A) = \int_A Z(t) dP$$

Autrement dit, si Q_t est définie par :

$$Q_t(A) = \int_A Z(t) dP, \quad A \in \mathcal{A}$$

Q et Q_t coïncident sur \mathcal{F}_t .

Supposons maintenant que $f \in L^2_{\text{loc}}(R_+)$, mais $f \notin L^2(R_+)$ (c'est le cas dans beaucoup d'applications). On ne peut plus définir Z , mais on peut définir $Z(t)$, $\forall t \in R_+$, et Q_t . On montre comme ci-dessus que pour tout $t \in R_+$, $(\bar{B}(s), 0 \leq s \leq t)$ est un Q_t - \mathcal{F} mouvement brownien. La question naturelle est alors de savoir s'il existe une probabilité Q sur (Ω, \mathcal{A}) dont la restriction à \mathcal{F}_t coïncide avec celle de Q_t pour tout $t \in R_+$, et donc telle que $(\bar{B}(t), t \in R_+)$ soit encore un Q - \mathcal{F} -mouvement brownien

(mais cette fois Q n'est plus nécessairement absolument continue par rapport à P sur \mathcal{A}). La réponse à cette question est oui et nous l'admettons ; cela résulte d'un théorème de base de théorie de la mesure.

Nous pouvons enfin passer au théorème de Girsanov. Nous remplaçons $f \in L^2(R_+)$ de la formule de Cameron-Martin par $\varphi \in M^2$, et posons :

$$Z(t) = \exp \left(\int_0^t \varphi(s) dB(s) - \frac{1}{2} \int_0^t \varphi^2(s) ds \right), t \geq 0$$

Il résulte de la formule d'Itô :

$$Z(t) = 1 + \int_0^t Z(s) \varphi(s) dB(s)$$

Si l'intégrale stochastique ci-dessus est d'intégrale nulle, alors $E(Z(t)) = 1$, et on peut définir Q_t par

$$Q_t(A) = \int_A Z_t dP, t \in R_+$$

Ceci est en particulier le cas sous l'hypothèse du :

Lemme 2.8.1.

Supposons qu'il existe une constante c telle que

$$\int_0^t \varphi(s)^2 ds \leq c \quad \text{p.s.}$$

Alors $E(Z(t)^2) < \infty$ et $E(Z(t)) = 1$.

Preuve : a) Supposons tout d'abord que le processus $(\varphi(s), 0 \leq s \leq t)$ est borné en valeur absolue par une constante k . Alors :

$$E(Z(t)^2) = 1 + E \int_0^t Z(s)^2 \varphi(s)^2 ds \leq 1 + k^2 \int_0^t E(Z(s)^2) ds,$$

d'où par le lemme de Gronwall (cf. Lemme 3.2.2. ci-dessous) :

$$E(Z(t)^2) \leq \exp(k^2 t)$$

et $Z(\cdot) \varphi(\cdot) \in M^2(0,t)$, ce qui entraîne que

$$E(Z(t)) = 1$$

b) Supposons maintenant que $\varphi(\cdot)$ satisfait l'hypothèse du lemme, et posons

$$\varphi_n(s) = \sup(\inf(\varphi(s), n), -n)$$

$$Z_n(t) = \exp\left(\int_0^t \varphi_n(s) dB(s) - \frac{1}{2} \int_0^t \varphi_n(s)^2 ds\right)$$

$$\begin{aligned} E(Z_n^2(t)) &= E\left(\exp\left[2 \int_0^t \varphi_n(s) dB(s) - \int_0^t \varphi_n(s)^2 ds\right]\right) \\ &\leq \left\{ E\left(\exp\left[4 \int_0^t \varphi_n(s) dB(s) - 8 \int_0^t \varphi_n(s)^2 ds\right]\right) \right\}^{1/2} \\ &\leq \left\{ E\left(\exp\left[6 \int_0^t \varphi_n(s)^2 ds\right]\right) \right\}^{1/2} \\ &\leq e^{3c} \end{aligned}$$

On a utilisé le résultat de la partie a pour conclure que :

$$E\left(\exp\left[4 \int_0^t \varphi_n(s) dB(s) - 8 \int_0^t \varphi_n(s)^2 ds\right]\right) = 1$$

Il est facile de vérifier que

$$Z_n(t) \rightarrow Z(t) \text{ p.s.}$$

la borne que nous venons d'établir sur $E(Z_n^2(t))$ entraîne d'après le lemme 2.8.2 ci-dessous que cette convergence a lieu dans $L^1(\Omega)$, donc :

$$\begin{aligned} E Z(t) &= \lim_{n \rightarrow \infty} E(Z_n(t)) \\ &= 1 \end{aligned}$$

et par le lemme de Fatou :

$$E(Z^2(t)) \leq e^{3c} \quad \square$$

Il nous reste à établir le :

Lemme 2.8.2.

Soit $(X_n, n \in \mathbb{N})$ et X des v.a.r., telles que $X_n \rightarrow X$ p.s., quand $n \rightarrow \infty$. Chacune des deux conditions suivantes entraîne que la convergence a lieu dans $L^1(\Omega)$:

(i) $\sup_{n \in \mathbb{N}} E(X_n^2) < \infty$

(ii) $X_n \geq 0$ p.s., $E(X_n) \rightarrow E(X)$ et $E(X) < \infty$.

Preuve : (i) Par Fatou, l'hypothèse entraîne que

$$E(X^2) < \infty$$

donc qu'il existe $c > 0$ t.q.

$$E(|X_n - X|^2) \leq c, \quad \forall n \in \mathbb{N}$$

Soit $M > 0$

$$\begin{aligned} E(|X_n - X|) &= E(|X_n - X| 1_{\{|X_n - X| \leq M\}}) + E(|X_n - X| 1_{\{|X_n - X| > M\}}) \\ &\leq E(|X_n - X| \wedge M) + \frac{1}{M} E(|X_n - X|^2) \end{aligned}$$

Par le théorème de convergence dominée,

$$\limsup_{n \rightarrow \infty} E(|X_n - X|) \leq \frac{c}{M}$$

Il reste à faire tendre M vers $+\infty$.

(ii) Par convergence dominée,

$$E[(X - X_n)^+] \rightarrow 0 \quad \text{quand } n \rightarrow \infty$$

et par l'hypothèse

$$E(X - X_n) \rightarrow 0, \quad n \rightarrow \infty$$

Il reste à remarquer que

$$|X - X_n| = 2(X - X_n)^+ - (X - X_n) \quad \square$$

Nous utiliserons aussi ci-dessous le fait que le produit d'une suite qui converge dans $L^1(\Omega)$ et d'une suite qui converge p.s. et est bornée, converge dans $L^1(\Omega)$; cela résulte du théorème de convergence dominée.

Nous pouvons maintenant établir le :

Théorème 2.8.3. (Girsanov).

Supposons que $E(Z_t) = 1$, pour tout $t \geq 1$. Alors il existe une unique probabilité Q sur (Ω, \mathcal{A}) telle que pour tout $t \in R_+$, $A \in \mathcal{F}_t$,

$$Q(A) = \int_A Z_t dP$$

et la f.a.r. $(\bar{B}(t), t \in R_+)$ définie par

$$\bar{B}(t) = B(t) - \int_0^t \varphi(s) ds$$

est un Q - \mathcal{F} -mouvement brownien.

Preuve. Nous allons nous contenter de montrer que t étant fixé tel que $E(Z_t) = 1$, si l'on définit Q sur (Ω, \mathcal{F}_t) par

$$Q(A) = \int_A Z_t dP, A \in \mathcal{F}_t$$

$(\bar{B}(s), 0 \leq s \leq t)$ est un Q - \mathcal{F} -mouvement brownien.

Notre démonstration sera proche de celle du théorème 1.6.1, mais nous ferons appel à la transformée de Fourier plutôt qu'à la transformée de Laplace. Nous allons démontrer que pour tout $p \in \mathbb{N}$; $0 < t_0 < t_1 < \dots < t_p \leq t$; $u_j \in \mathbb{R}, 1 \leq j \leq p$,

$$(*) \quad E(Z_t \exp [i \sum_1^p u_j \bar{B}(t_j)]) = E(\exp [i \sum_1^p u_j B(t_j)])$$

ce qui prouve que la loi de $(\bar{B}(t_1), \dots, \bar{B}(t_p))$ sous Q est celle de $(B(t_1), \dots, B(t_p))$ sous P .

Nous allons procéder en deux étapes.

a) *Preuve de (*) sous l'hypothèse supplémentaire* $\int_0^t \varphi^2(s) ds \leq c$ p.s.

Posons $\psi(s) = \varphi(s) + i \sum_{j=1}^p u_j 1_{[0, t_j]}(s)$. Alors $\int_0^t |\psi(s)|^2 ds \leq \bar{C}$ p.s. et il résulte d'une extension du Lemme 2.8.1 au cas d'un intégrand à valeurs complexes :

$$(**) \quad E(\exp[\int_0^t \psi(s) dB(s) - \frac{1}{2} \int_0^t \psi(s)^2 ds]) = 1$$

soit

$$E(Z_t \exp i \sum_1^p u_j \bar{B}(t_j)) = \exp \left(-\frac{1}{2} \sum_{k=1}^p u_k t_k \wedge t_k \right)$$

On conclut en calculant le second membre de (*), sachant que $(B(t_1), \dots, B(t_p))$ est un v.a. gaussien d'espérance nulle et de matrice de covariance Q t.q. $Q_{jk} = t_j \wedge t_k$.

b) *Le cas général.*

Il suffit de montrer que (***) est encore vraie. Soit $\{\varphi_n, n \in \mathbb{N}\} \subset M^2(0, t)$ t.q.

$$\int_0^t \varphi_n(s)^2 ds \leq c_n \quad \text{p.s.}$$

$$\varphi_n \rightarrow \varphi \quad \text{dans } M^2(0, t)$$

Posons

$$Z_n(t) = \exp \left[\int_0^t \varphi_n(s) dB(s) - \frac{1}{2} \int_0^t \varphi_n(s)^2 ds \right]$$

$Z_n \rightarrow Z(t)$ en probabilité, et quitte à extraire une sous-suite on peut supposer que la convergence a lieu p.s. En outre $Z_n(t) \geq P$ p.s. et

$$E(Z_n(t)) = E(Z(t))$$

donc d'après le lemme 2.8.2 (ii),

$$Z_n(t) \rightarrow Z(t) \quad \text{dans } L^1(\Omega)$$

Comme en outre

$$\exp \left[i \sum_{j=1}^p u_j (B(t_j) - \int_0^{t_j} \varphi_n(s) ds) + \frac{1}{2} \sum_{j,k=1}^p u_j u_k t_j \wedge t_k \right]$$

est une suite de v.a. complexes bornée en module, et qui converge en probabilité (donc p.s., après extraction d'une sous-suite) vers

$$\exp \left[i \sum_{j=1}^p u_j (B(t_j) - \int_0^{t_j} \varphi(s) ds) + \frac{1}{2} \sum_{j,k=1}^p u_j u_k t_j \wedge t_k \right]$$

On en déduit que la suite de v.a. complexes

$$\exp \left[\int_0^t \psi_n(s) dB(s) - \frac{1}{2} \int_0^t \psi_n(s)^2 ds \right]$$

(avec $\psi_n = \varphi_n + i \sum_{j=1}^p u_j 1_{[0,t_j]}$) converge dans $L^1(\Omega)$, donc les espérances convergent, ce qui établit (***) dans le cas général. \square

Il nous reste à donner des conditions plus faibles que celle du Lemme 9.1, qui entraînent que $E(Z(t)) = 1$.

En approchant $(\varphi(s); 0 \leq s \leq t)$ par des processus bornés, on déduit du lemme de Fatou que $E(Z(t)) \leq 1$, dès que $\varphi \in M_{loc}^2(0,t)$. En fait on a même, par le même raisonnement,

$$E^{s_s} \left(\frac{Z(t)}{Z(s)} \right) \leq 1 \text{ p.s.}, \quad 0 \leq s \leq t.$$

Donc la fonction $t \rightarrow E(Z(t))$ est décroissante et si $E(Z(t)) = 1$, $E(Z(s)) = 1$ pour tout $0 \leq s \leq t$.

Nous admettrons le premier résultat.

Lemme 2.8.4. (critère de Novikov). Si

$$E(\exp [\frac{1}{2} \int_0^t \varphi(s)^2 ds]) < \infty$$

alors

$$E(Z(t)) = 1 \quad \square$$

Le résultat suivant s'applique en particulier dans le cas où $(\varphi(s), 0 \leq s \leq t)$ est un processus gaussien.

Lemme 2.8.5.

Supposons qu'il existe c et $a > 0$ tels que

$$E(\exp [a\varphi(s)^2]) \leq c, \quad 0 \leq s \leq t$$

Alors $E(Z(t)) = 1$.

Preuve :

$$\text{Soit } \{\varphi_n, n \in \mathbb{N}\} \subset M^2(0,t) \text{ t. q. } \int_0^t \varphi_n(s)^2 ds \leq c_n \text{ p.s.}$$

$$\varphi_n \rightarrow \varphi \text{ dans } M^2(0,t)$$

et soit $0 \leq s \leq r \leq t$ t.q. $r-s \leq \frac{a}{6}$.

On associe Z_n à φ_n comme dans la preuve du Théorème 2.8.3.

$$\begin{aligned} \left(\frac{Z_n(r)}{Z_n(s)}\right)^2 &= \exp \left[2 \int_s^r \varphi_n(u) dB(u) - 4 \int_s^r \varphi_n(u)^2 du \right] \exp \left[3 \int_s^r \varphi_n^2(u) du \right] \\ E \left[\left(\frac{Z_n(r)}{Z_n(s)}\right)^2 \right] &\leq \left\{ E \left(\exp \left[4 \int_s^r \varphi_n(u) dB(u) - 8 \int_s^r \varphi_n(u)^2 du \right] \right) \right\}^{1/2} \\ &\quad \left\{ E \left(\exp \left[6 \int_s^r \varphi_n^2(u) du \right] \right) \right\}^{1/2} \\ &\leq \left\{ \int_s^r E \left(\exp \left[6(r-s) \varphi_n^2(u) \right] \right) \frac{du}{r-s} \right\}^{1/2} \\ &\leq c^{1/2} \end{aligned}$$

on a utilisé l'inégalité de Jensen avec la fonction convexe $x \rightarrow \exp x$ et la mesure de probabilité $\frac{du}{r-s}$ sur l'intervalle $[s,r]$.

On déduit de cette estimation, par le lemme 2.8.2 (i), que $\frac{Z_n(r)}{Z_n(s)} \rightarrow \frac{Z(r)}{Z(s)}$ dans $L^1(\Omega)$, et donc que

$$\begin{aligned} E^{s_0} \left(\frac{Z(r)}{Z(s)} \right) &= \lim_{n \rightarrow \infty} E^{s_0} \left(\frac{Z_n(r)}{Z_n(s)} \right) \\ &= 1 \end{aligned}$$

Soit n le plus petit entier plus grand ou égal à $\frac{6t}{a}$. On pose $t_k = \frac{kt}{n}$, $0 \leq k \leq n$.

$$Z(t) = \prod_{k=1}^n \frac{Z(t_k)}{Z(t_{k-1})}$$

Pour tout $1 \leq k \leq n$, $\frac{Z(t_k)}{Z(t_{k-1})}$ est \mathcal{F}_{t_k} mesurable, et d'après la première partie $E^{s_{t_{k-1}}} \left(\frac{Z(t_k)}{Z(t_{k-1})} \right) = 1$ p.s.

On en déduit aisément que $E(Z(t)) = 1$. \square

Chapitre III

Equations différentielles stochastiques

3.1. Introduction.

Au chapitre 1, on a déjà résolu l'équation différentielle stochastique (que nous abrègerons dorénavant par e.d.s.) de Langevin :

$$\begin{cases} dX(t) = aX(t) dt + dB(t) \\ X(0) \text{ donnée} \end{cases}$$

Maintenant que nous disposons de l'intégrale stochastique d'Itô, nous allons pouvoir résoudre beaucoup d'autres e.d.s. Par exemple, il résulte de la formule d'Itô que la f.a.r.

$$X(t) = \exp(B(t))$$

est solution de l'e.d.s. :

$$\begin{cases} dX(t) = \frac{1}{2} X(t) dt + X(t) dB(t) \\ X(0) = 1 \end{cases}$$

Cette équation admet-elle d'autres solutions que celle que nous avons trouvée ? La réponse à cette question est donnée par un théorème d'existence et d'unicité assez général, qui bien sûr concerne beaucoup d'équations pour lesquelles on n'a pas de solution explicite.

3.2. Un théorème d'existence et d'unicité.

Dans cette section, nous allons étudier l'E.D.S. :

$$(E) \begin{cases} dX(t) = f(X(t)) dt + g(X(t)) dB(t) \\ X(0) = X_0 \end{cases}$$

qui est une écriture formelle pour :

$$X(t) = X_0 + \int_0^t f(X(s)) ds + \int_0^t g(X(s)) dB(s), \quad t \geq 0$$

On supposera que $X(t)$ est d -dimensionnel et que $B(t)$ est k -dimensionnel. Donc f est une application de R^d dans R^d , g une application de R^d dans $R^{d \times k}$. Si $g_j(x)$ ($1 \leq j \leq k$) désigne la j -ième colonne de la matrice g , l'intégrale stochastique ci-dessus peut se réécrire :

$$\int_0^t g(X(s)) dB(s) = \sum_{j=1}^k \int_0^t g_j(X(s)) dB_j(s)$$

Le coefficient $f(X(t))$ de « dt » s'appelle la *dérivée*, et $g(X(t))$ - celui de « $dB(t)$ » - s'appelle le *coefficient de diffusion*.

On a le :

Théorème 3.2.1.

Etant donné un vecteur aléatoire d -dimensionnel X_0 \mathcal{F}_0 -mesurable (i.e. indépendant de $\{B(t), t \geq 0\}$) et de carré intégrable, si les applications f et g sont lipschitziennes, i.e. s'il existe une constante K telle que :

$$|f(x) - f(y)| + |g(x) - g(y)| \leq K|x-y|$$

pour tout $x, y \in R^d$, alors l'équation (E) admet une unique solution $\{X(t)\} \in M^2$.

REMARQUE. $g(x)$ étant une matrice $d \times k$, on prendra comme norme de $g(x)$:

$$|g(x)| = \sqrt{\text{Tr}(g(x)g^t(x))}$$

Avant de démontrer le théorème, établissons tout d'abord le :

Lemme 3.2.2. («Lemme de Gronwall»).

Soit $t \rightarrow x(t)$ une application de R_+ dans R , telle qu'il existe un réel $b \geq 0$ avec :

$$x(t) \leq a + b \int_0^t x(s)ds, \quad t \geq 0$$

Alors

$$x(t) \leq a \exp(bt), \quad t \geq 0$$

Preuve. Après multiplication par e^{-bt} , l'inégalité qui nous sert d'hypothèse peut se réécrire :

$$\frac{d}{dt} (e^{-bt} \int_0^t x(s)ds) \leq a e^{-bt}$$

d'où après intégration :

$$e^{-bt} \int_0^t x(s)ds \leq \frac{a}{b} (1 - e^{-bt})$$

$$a + b \int_0^t x(s)ds \leq a e^{bt}$$

Il suffit de réutiliser l'hypothèse pour conclure. \square

Preuve du théorème.

Unicité: Soit $\{X(t)\}, \{Y(t)\} \in M^2$ deux solutions de l'équation (E). Alors :

$$X(t) - Y(t) = \int_0^t [f(X(s)) - f(Y(s))] ds + \int_0^t (g(X(s)) - g(Y(s))) dB(s)$$

$$\begin{aligned} E(|X(t) - Y(t)|^2) &\leq 2E \left[\left(\int_0^t [f(X(s)) - f(Y(s))] ds \right)^2 \right] + \\ &\quad + 2E \left[\left(\int_0^t [f(X(s)) - g(X(s))] dB(s) \right)^2 \right] \\ &\leq 2tE \int_0^t |f(X(s)) - f(Y(s))|^2 ds \\ &\quad + 2E \int_0^t |g(X(s)) - g(Y(s))|^2 ds \\ &\leq 2(t+1)K \int_0^t E(|X(s) - Y(s)|^2) ds \end{aligned}$$

On a utilisé l'hypothèse de Lipschitz sur les coefficients à la dernière ligne. Il résulte alors du Lemme de Gronwall :

$$E(|X(t) - Y(t)|^2) = 0, \quad t \geq 0$$

Existence. On va utiliser la méthode itérative de Picard pour construire une suite qui converge vers la solution. Posons $X_0(t) \equiv 0$, et pour $n \geq 0$:

$$X_{n+1}(t) = X_0 + \int_0^t f(X_n(s)) ds + \int_0^t g(X_n(s)) dB(s)$$

Alors, pour $n \geq 1$,

$$\begin{aligned} X_{n+1}(t) - X_n(t) &= \int_0^t [f(X_n(s)) - f(X_{n-1}(s))] ds + \\ &\quad + \int_0^t [g(X_n(s)) - g(X_{n-1}(s))] dB(s) \end{aligned}$$

d'où l'on tire, par les mêmes calculs que pour l'unicité :

$$E(|X_{n+1}(t) - X_n(t)|^2) \leq 2(t+1) K \int_0^t E(|X_n(s) - X_{n-1}(s)|^2) ds$$

Par ailleurs,

$$E(|X_1(t)|^2) \leq 4E(|X_0|^2) + 4t^2 f^2(0) + 4t g^2(0)$$

Choisissons $T > 0$ arbitraire, et posons

$$C_T = \sup (2(T+1)K, 4E(|X_0|^2) + 4t^2 f^2(0) + 4t g^2(0))$$

Alors $E|X_1(t)|^2 \leq C_T$ et

$$E(|X_{n+1}(t) - X_n(t)|^2) \leq C_T \int_0^t E(|X_n(s) - X_{n-1}(s)|^2) ds, \quad 0 \leq t \leq T, n \geq 1$$

De ces deux inégalités, on déduit par récurrence sur n :

$$E|X_{n+1}(t) - X_n(t)|^2 \leq C_T^{n+1} \frac{t^n}{n!}, \quad 0 \leq t \leq T, n \geq 0$$

Donc

$$\|X_{n+1} - X_n\|_{M^2(0,T)} \leq \frac{(C_T T)^{n+1}}{(n+1)!}$$

et pour tout $n, p \in \mathbb{N}$,

$$\|X_{n+p} - X_n\|_{M^2(0,T)} \leq \sum_{\ell=1}^{\infty} \frac{(C_T T)^{n+\ell}}{(n+\ell)!},$$

donc $\{X_n\}$ est une suite de Cauchy dans $M^2(0,T)$. Les hypothèses faites sur les coefficients permettent alors de passer à la limite quand $n \rightarrow \infty$ dans l'égalité :

$$X_{n+1}(t) = X_0 + \int_0^t f(X_n(s)) ds + \int_0^t g(X_n(s)) d(s), \quad 0 \leq t \leq T,$$

d'où l'on déduit que $\{X(t)\}$ est solution de l'équation (E) sur $[0,T]$. Comme T est arbitraire, le théorème est démontré. \square

REMARQUE 3.2.3. EDS au sens de Stratonovich.

Remarquons que si g est de classe C^1 , et $(X(t))$ la solution de l'EDS (E), alors on peut définir l'intégrale de Stratonovich :

$$\int_0^t g_i(X(s)) \circ dB_i(s) = \int_0^t g_i(X(s)) dB_i(s) + \frac{1}{2} \int_0^t \frac{\partial g_i}{\partial x}(X(s)) g_i(X(s)) ds$$

où $\frac{\partial g_i}{\partial x}(x) g_i(x) = \sum_{j=1}^d \frac{\partial g_i}{\partial x_j}(x) g_{ji}(x)$, et on note

$$\int_0^t g(X(s)) \circ dB(s) = \int_0^t g(X(s)) dB(s) + \frac{1}{2} \sum_{i=1}^k \int_0^t \frac{\partial g_i}{\partial x}(X(s)) g_i(X(s)) ds$$

On dira alors que la f.a. $(X(t))$ est solution de l'EDS au sens de Stratonovich :

$$(ES) \quad X(t) = X(0) + \int_0^t f(X(s)) ds + \int_0^t g(X(s)) \circ dB(s)$$

si elle est solution de l'EDS au sens de Itô :

$$X(t) = X(0) + \int_0^t [f(X(s)) + \frac{1}{2} \sum_{i=1}^k \frac{\partial g_i}{\partial x}(X(s)) g_i(X(s))] ds + \int_0^t g(X(s)) dB(s)$$

Supposons dans la suite que les coefficients de cette dernière équation au sens de Itô sont globalement lipschitziens. Alors (ES) possède une solution unique.

Les EDS au sens de Stratonovich sont importantes pour au moins deux raisons.

La première est que si M est une variété différentiable de classe C^2 plongée dans \mathbb{R}^d , si $X(0) \in M$ p.s., f, g_1, \dots, g_k sont des champs de vecteurs sur M (i.e. $\forall x \in M, f(x), g_1(x), \dots, g_k(x)$ sont des vecteurs tangents en x à M), alors la solution de ES vérifie :

$$P(X(t) \in M, \forall t \geq 0) = 1.$$

Cette propriété serait fausse pour la solution de (E) .

La seconde est la propriété d'approximation suivante. Soit $t_k^* = \frac{k}{n}et$, ($t \in B_n$, $t \geq 0$) la f.a. dont les trajectoires ont pour courbes représentatives les lignes brisées joignant les points $\{(t_k^*, B(t_k^*))\}$, $k \in \mathbb{N}$. Alors les trajectoires de $B^n(t)$ sont p.p. de classe C^1 , et on peut considérer l'EDS « approchée » :

$$(En) \quad \begin{aligned} \frac{dX_n(t)}{dt} &= f(X_n(t)) + g(X_n(t)) \frac{dB_n}{dt}(t) \\ X_n(0) &= X_0 \end{aligned}$$

qui possède une solution unique. Le résultat important est que si $(X(t))$ désigne la solution de (ES) , pour tout $T > 0$,

$$\sup_{0 \leq t \leq T} |X_n(t) - X(t)| \rightarrow 0$$

en probabilité, quand $n \rightarrow \infty$. Il y a beaucoup d'autres façons d'approcher le mouvement brownien, et on peut démontrer beaucoup de théorèmes de convergence sur les e.d.s. approchées. Les approximations les plus « naturelles » du brownien conduisent en général à ce que la solution de (ES) soit la limite des solutions approchées.

3.3. Deux exemples.

EXEMPLE 3.3.1. Mouvement brownien géométrique.

Nous allons considérer le cas où la dérive et le coefficient de diffusion sont tous deux linéaires, en nous limitant au cas scalaire ($d = k = 1$), i.e. :

$$\begin{aligned} dX(t) &= a X(t) dt + b X(t) dB(t) \\ X(0) &= X_0 \end{aligned}$$

Il résulte aisément de la formule d'Itô que la solution est donnée explicitement par :

$$X(t) = X_0 \exp(\bar{a} t + bB(t))$$

où $\bar{a} = a - \frac{b^2}{2}$. Cette f.a.r. est appelée « mouvement brownien géométrique ». Rappelons que le mouvement brownien est un processus continu à accroissements indépendants et stationnaires, (i.e.) pour tout $n \in \mathbb{N}$, $0 \leq t_0 < t_1 < \dots < t_n$, $h \geq 0$, les variables aléatoires

$$B(t_1 + h) - B(t_0 + h), B(t_2 + h) - B(t_1 + h), \dots, B(t_n + h) - B(t_{n-1} + h)$$

sont indépendantes et de lois indépendantes de h . Il en est de même pour le « mouvement brownien avec dérive » :

$$Y(t) = \bar{a} t + bB(t)$$

Quand au processus

$$X(t) = \exp(Y(t))$$

ses « accroissements multiplicatifs » sont indépendants et stationnaires, i.e.

$$X(t_1 + h)/X(t_0 + h), X(t_2 + h)/X(t_1 + h), \dots, X(t_n + h)/X(t_{n-1} + h)$$

sont indépendantes et de lois indépendantes de h .

Notons en outre une propriété fondamentale du mouvement brownien géométrique :

$$P(X(t)/X_0 > 0, \forall t \geq 0) = 1$$

Donc pourvu que $X_0 > 0$, $X(t) > 0 \forall t \geq 0$. Il en résulte que le mouvement brownien géométrique est très utilisé pour modéliser des quantités qui ne changent pas de signe, par exemple les cours d'actions en bourse.

On pourrait plus généralement considérer l'équation :

$$\begin{aligned} dX(t) &= (a X(t) + c) dt + (b X(t) + d) dB(t) \\ X(0) &= X_0 \end{aligned}$$

dont la solution est une f.a.r. d'Ornstein-Uhlenbeck dans le cas particulier $b = 0$. La solution de cette dernière équation n'a plus la propriété de conserver un signe constant, sauf si $d = 0$ et $c X_0 > 0$.

EXEMPLE 3.3.2. Modèle « CIR » de fluctuation des taux d'intérêt.

Un des modèles les plus utilisés pour représenter les fluctuations des taux d'intérêt, appelé modèle de Cox-Ingersoll-Ross, est :

$$\begin{aligned} dX(t) &= c(\theta - X(t)) dt + \sigma \sqrt{X(t)} dB(t) \\ X(0) &= x \end{aligned}$$

avec $x > 0$, et à nouveau $(X(t), B(t))$ sont des processus scalaires ; c, θ, σ sont trois paramètres strictement positifs.

Bien que cette équation ne rentre pas dans le cadre du Théorème 2.1. (la fonction $x \rightarrow \sqrt{|x|}$ n'est pas lipschitzienne), on sait montrer que cette équation possède une unique solution positive.

Dans le cas $\sigma = 0$, la solution tend vers θ quand $t \rightarrow \infty$, et ceci d'autant plus vite que c est grand. Dans le cas $\sigma > 0$, le mouvement est perturbé par des fluctuations proportionnelles à $\sqrt{X(t)}$.

On n'a pas de formule explicite pour la solution de notre e.d.s. Cependant, on peut écrire des équations pour les deux premiers moments. En effet (la deuxième égalité résulte de la formule d'Itô) :

$$X(t) = x + c \int_0^t [\theta - X(s)] ds + \sigma \int_0^t \sqrt{X(s)} dB(s)$$

$$\begin{aligned} X^2(t) &= x^2 + 2c \int_0^t [\theta - X(s)] X(s) ds + 2\sigma \int_0^t (X(s))^{3/2} dB(s) + \sigma^2 \int_0^t X(s) ds \\ &= x^2 + (2c\theta + \sigma^2) \int_0^t X(s) ds - 2c \int_0^t X^2(s) ds + 2\sigma \int_0^t (X(s))^{3/2} dB(s) \end{aligned}$$

Nous admettrons que les deux intégrales stochastiques qui apparaissent dans les égalités ci-dessus sont d'espérances nulles. On obtient alors :

$$E X(t) = x + c [\theta t - \int_0^t E X(s) ds]$$

$$E X^2(t) = x^2 + (2c\theta + \sigma^2) \int_0^t E X(s) ds - 2c \int_0^t E X^2(s) ds$$

Posons

$$\mu(t) = \theta + (x - \theta) e^{-ct}, \quad t \geq 0$$

alors

$$E X(t) = \mu(t) \quad t \geq 0.$$

Soit maintenant $(Y(t), t \geq 0)$ la f.a.r. d'Ornstein-Uhlenbeck solution de :

$$dY(t) = c(\theta - Y(t)) dt + \sigma \sqrt{\mu(t)} dB(t)$$

$$Y(0) = x$$

On vérifie alors aisément que

$$E Y(t) = \mu(t)$$

$$= E X(t)$$

$$E Y^2(t) = E X^2(t), \quad t \geq 0.$$

$(Y(t), t \geq 0)$ est également utilisé pour modéliser les fluctuations des cours de taux d'intérêt, avec éventuellement $\mu(t)$ remplacé par sa valeur asymptotique θ . Certes, la f.a. $(Y(t))$ n'a pas la propriété de garder un signe constant. Cependant, si θ est suffisamment grand par rapport à σ , $Y(t)$ est rarement négatif, et les deux premiers moments de $Y(t)$ coïncident avec ceux de $X(t)$, pour tout $t \geq 0$. L'équation de $(Y(t))$ a l'avantage d'être plus simple que le modèle C.I.R.

3.4. Propriété de Markov de la solution d'une E.D.S.

Commençons par quelques généralités sur les processus de Markov. Soit (Ω, \mathcal{A}, P) un espace de probabilité muni d'une filtration $(\mathcal{F}_t, t \geq 0)$, et $(X(t), t \geq 0)$ une f.a. à valeurs dans \mathbb{R}^d ($X(t), t \geq 0$) adaptée à (\mathcal{F}_t) (i.e. t.q. $X(t)$ est \mathcal{F}_t mesurable pour tout $t \geq 0$).

Définition 3.4.1.

La f.a. $(X(t))$ est une \mathcal{F} -fonction aléatoire de Markov si pour tout $t, h > 0$, $A \in \mathcal{B}_d$,

$$P(X(t+h) \in A / \mathcal{F}_t) = P(X(t+h) \in A / X(t)).$$

Si en outre la quantité $P(X_{t+h} \in A / X_t = x)$ ne dépend que de h, x et A (et pas de t), la f.a. de Markov $(X(t))$ est dite *homogène*. \square

Cette définition généralise au cas du temps continu la notion de dépendance markovienne pour une suite aléatoire en temps discret $(X(t), t \in \mathbb{N})$, cf. cours de Probabilités de 1^{ère} année, section 1.9.

Une f.a. de Markov est l'analogie aléatoire de la solution d'une EDO du type :

$$\frac{dX(t)}{dt} = f(t, X(t))$$

par opposition à une EDO où l'évolution dépend du passé, par exemple :

$$\frac{dX(t)}{dt} = f(t, X(t), X(t-h))$$

Une f.a. de Markov homogène est l'analogie de la solution d'une EDO autonome, i.e. du type

$$\frac{dX(t)}{dt} = f(X(t))$$

où l'évolution ne dépend que de la position, et pas du temps.

Pour simplifier les notations, nous ne considérerons ci-dessous que des f.a. de Markov homogènes. Dans ce cas, le rôle joué en temps et espace discret par la matrice de transition et ses itérées est joué ici par la *probabilité de transition* :

$$Q(t, x, A) = P(X_{s+t} \in A / X_s = x)$$

où $Q : \mathbb{R}_+ \times \mathbb{R}^d \times \mathcal{B}_d \rightarrow [0, 1]$ vérifie les propriétés suivantes :

(i) $\forall A \in \mathcal{B}_d, (t, x) \rightarrow Q(t, x, A)$ est $\mathcal{B}_+ \otimes \mathcal{B}_d$ mesurable

(ii) $\forall (t,x) \in \mathbb{R}_+ \times \mathbb{R}^d, A \rightarrow Q(t,x,A)$ est une mesure de probabilité
et surtout Q satisfait l'équation de Chapman - Kolmogorov : $\forall s, t > 0, x \in \mathbb{R}^d, A \in \mathcal{B}_d,$

$$(iii) Q(s+t, x, A) = \int_{\mathbb{R}^d} Q(s, x, dy) Q(t, y, A)$$

Preuve de (iii).

$$Q(s+t, x, A) = P(X(s+t) \in A / X(0) = x) \\ E(E(1_A[X(s+t)]/\mathcal{F}_s) / X(0) = x)$$

Mais d'après la propriété de Markov

$$E(1_A[X(s+t)]/\mathcal{F}_s) = P(X(s+t) \in A/\mathcal{F}_s) \\ = P(X(s+t) \in A/X_s) \\ = Q(t, X_s, A)$$

Il reste à calculer $E(Q(t, X_s, A)/X(0) = x)$. Notons que pour tout $B \in \mathcal{B}_d,$

$$E(1_B(X_s)/X(0) = x) = Q(s, x, B) \\ = \int 1_B(y) Q(s, x, dy)$$

Il résulte des propriétés élémentaires de l'espérance conditionnelle et de l'intégrale par rapporte à la mesure $Q(s, x, \cdot)$ que cette dernière égalité est encore vraie avec $1_B(\cdot)$ remplacée par n'importe quelle fonction φ mesurable bornée de \mathbb{R}^d dans \mathbb{R} . On peut en particulier choisir

$$\varphi(y) = Q(t, y, A) \quad \square$$

Nous allons maintenant montrer que le mouvement brownien est une f.a. de Markov, ce qui résulte de la propriété du mouvement brownien d'être à accroissements indépendants. Nous aurons besoin du lemme suivant.

Lemme 3.4.2. Soit $\varphi : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ une application mesurable et bornée. Soit X un vecteur aléatoire de dimension n , Y un vecteur aléatoire de dimension m , définis sur un espace de probabilité (Ω, \mathcal{A}, P) , et \mathcal{G} une sous tribu de \mathcal{A} . Si X est \mathcal{G} mesurable et si Y et \mathcal{G} sont indépendants, alors

$$E[\varphi(X, Y)/\mathcal{G}] = E[\varphi(X, Y)/X]$$

Preuve (esquisse) : Il suffit de montrer que chacun des deux membres de l'égalité est égal à :

$$\int_{\mathbb{R}^m} \varphi(X, y) P_Y(dy)$$

où P_Y désigne la loi de probabilité de Y . Ce résultat est assez facile à établir lorsque $\varphi(x, y) = h(x)g(y)$. Le cas général s'en déduit par un passage à la limite. \square

Proposition 3.4.3. Un \mathfrak{F} -mouvement brownien $(B(t))$ est une \mathfrak{F} -f.a de Markov.

Preuve : Reprenant les notations de la Définition 3.4.1, on calcule :

$$P(B(t+h) \in A/\mathfrak{F}_t) = E(1_A(B(t) + B(t+h) - B(t))/\mathfrak{F}_t)$$

Il reste à appliquer le Lemme 3.4.2. avec $n = m = d$, $h(x,y) = 1_A(x+y)$, $X = B(t), Y = B(t+h) - B(t)$, $\mathfrak{G} = \mathfrak{F}_t$, pour conclure. \square

Notons que

$$\begin{aligned} P(B(t+h) \in A/B(t) = x) &= P(B(t+h) - B(t) + x \in A) \\ &= \frac{1}{(2\pi h)^{d/2}} \int_A \exp\left(-\frac{|y-x|^2}{2h}\right) dy \end{aligned}$$

i.e., le mouvement brownien est une f.a. de Markov homogène, dont la probabilité de transition est donnée par :

$$Q(t,x,dy) = q(t,x,y) dy$$

avec

$$q(t,x,y) = (2\pi t)^{-d/2} \exp\left(-\frac{|y-x|^2}{2t}\right)$$

Il n'est pas difficile de vérifier que la *densité de transition* $q(t,x,y)$ satisfait l'équation *progressive* (ou « en avant »).

$$\frac{\partial q}{\partial t}(t,x,y) = \frac{1}{2} \sum_{i=1}^d \frac{\partial^2 q}{\partial y_i^2}(t,x,y)$$

et l'équation *rétrograde* (ou « en arrière »)

$$\frac{\partial q}{\partial t}(t,x,y) = \frac{1}{2} \sum_{i=1}^d \frac{\partial^2 q}{\partial x_i^2}(t,x,y)$$

la variable y est la variable « de l'avant » : elle correspond à la position de la f.a. à l'instant courant t . La variable x est la variable « de l'arrière » (ou du passé) : elle correspond à la position de la f.a. à l'instant initial $t = 0$.

Considérons maintenant l'EDS :

$$\begin{cases} dX(t) = f(X(t)) dt + g(X(t)) dB(t) & , \quad t \geq 0 \\ X(0) = X_0 \end{cases}$$

Nous supposons satisfaites les hypothèses du Théorème 3.2.1. (\mathfrak{F}_t) désigne une filtration $t.q.$ X_0 est \mathfrak{F}_0 mesurable et $B(t)$ est un \mathfrak{F} -mouvement brownien. On a alors le :

Théorème 3.4.4. L'unique solution $(X(t))$ de l'EDS ci-dessus est une \mathfrak{F} -f.a. de Markov homogène.

Preuve (esquisse) : Nous allons nous contenter d'indiquer l'idée de la démonstration, sans rentrer dans les détails techniques. Fixons $t, h > 0$. Posons

$$\bar{B}(s) = B(t+s) - B(t) \quad , \quad s \geq 0$$

et considérons l'EDS :

$$\begin{cases} dY(s) = f(Y(s)) ds + g(Y(s)) d\bar{B}(s) \quad , \quad s \geq 0 \\ Y(0) = X(t) \end{cases}$$

Notons que $d\bar{B}(s) = dB(t+s)$, $s \geq 0$, et donc par le théorème d'unicité de la solution d'une EDS, on vérifie aisément que :

$$X(t+s) = Y(s) \quad , \quad s \geq 0$$

En particulier, $X(t+h) = Y(h)$.

On remarque d'autre part que, par la construction de la solution d'une EDS (preuve de la partie « Existence » du Théorème 3.2.1.), $Y(h)$ est une fonction mesurable de $X(t)$ et de $(\bar{B}(s), 0 \leq s \leq h)$. Mais $\bar{B}(\cdot)$ est indépendant de \mathfrak{F}_t . On en déduit par un raisonnement analogue à celui du Lemme 3.4.2 :

$$\begin{aligned} \text{c.e.} \quad & P(Y(h) \in A/\mathfrak{F}_t) = P(Y(h) \in A/X_t) \\ & P(X(t+h) \in A/\mathfrak{F}_t) = P(X(t+h) \in A/X_t) \end{aligned}$$

L'homogénéité résulte de ce que ni la loi de $\bar{B}(\cdot)$, ni les coefficients de l'équation ne dépendent de t . \square

Un processus de Markov solution d'une EDS est appelé *processus de diffusion*.

REMARQUE 3.4.5. On aurait pu faire dépendre les coefficients f et g de notre EDS du temps, i.e. considérer l'EDS :

$$dX(t) = f(t, X(t)) dt + g(t, X(t)) dB(t)$$

Pourvu que f et g soient des fonctions boréliennes de (t, x) , que la constante de Lipschitz K du Théorème 3.2.1. soit indépendante de t , et que $|f(t, 0)| + |g(t, 0)|$ soit borné, on aurait encore existence et unicité d'une solution dans M^2 . Cette solution serait encore une f.a. de Markov, mais évidemment non homogène. \square

La probabilité de transition $Q(t, x, \cdot)$ de la solution est caractérisée par le fait que $Q(t, x, \cdot)$ est la loi de probabilité du v. a. $X(t)$, si $X(\cdot)$ désigne la solution de l'EDS :

$$\begin{aligned} dX(t) &= f(X(t)) dt + g(X(t)) dB(t) \\ X(0) &= x \end{aligned}$$

L'équation progressive satisfaite par Q se réduit aisément de la formule d'Itô. Soit en effet $\varphi \in C_c^\infty(\mathbb{R}^d)$ (i.e. une fonction de classe C^∞ à support compact de \mathbb{R}^d dans \mathbb{R}). Par la formule d'Itô :

$$\varphi(X(t)) = \varphi(x) + \int_0^t L\varphi(X(s)) ds + \sum_{i=1}^h \int_0^t M_i \varphi(X(s)) dB_i(s)$$

$$\text{où } L = \frac{1}{2} \sum_{i,j=1}^d a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + f_i(x) \frac{\partial}{\partial x_i}$$

$$M_i = \sum_{i=1}^d g_{it}(x) \frac{\partial}{\partial x_i}$$

avec $a(x) = g(x) g^t(x)$. Les processus $(M_i \varphi(X(t)))_{t=1, \dots, h}$ étant bornés, les intégrales stochastiques ci-dessus sont d'espérance nulle, d'où :

$$E \varphi(X(t)) = \varphi(x) + \int_0^t E L \varphi(X(s)) ds, \quad t \geq 0$$

ce qui s'écrit encore :

$$Q(t, x, \varphi) = \varphi(x) + \int_0^t Q(s, x, L \varphi) ds$$

$$\text{où } Q(t, x, \psi) = \int_{\mathbb{R}^d} Q(t, x, dy) \psi(y)$$

ou encore, $(Q(t, x, \cdot), t \geq 0)$ satisfait l'équation aux dérivées partielles parabolique suivante au sens des distributions :

$$\frac{\partial Q}{\partial t}(t, x, \cdot) = L^*_y Q(t, x, \cdot)$$

$$Q(t, x, \cdot) \rightarrow \delta_x, \quad \text{quand } t \downarrow 0$$

où L^*_y désigne l'adjoint de l'opérateur L , agissant sur la variable y , i.e. :

$$L^*_y = \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2}{\partial y_i \partial y_j} (a_{ij}(y) \cdot) - \frac{\partial}{\partial y_i} (f_i(y) \cdot)$$

On peut alors déduire de l'équation de Chapman-Kolmogorov l'équation rétrograde satisfaite par $(Q(t, x, \varphi); t \geq 0, x \in \mathbb{R}^d)$ [supposant à nouveau $\varphi \in C_c^\infty(\mathbb{R}^d)$]:

$$\frac{\partial Q}{\partial t}(t, x, \varphi) = L_x Q(t, x, \varphi)$$

$$\varphi(0, x, \varphi) = \varphi(x)$$

REMARQUE 3.4.6. Supposons que l'on ait à notre disposition un théorème d'unicité pour cette dernière EDP. Alors le résultat que nous venons d'énoncer peut s'interpréter comme donnant une formule probabiliste pour la solution $(v(t, x))$ de l'EDP :

$$\frac{\partial v}{\partial t}(t, x) = L_x v(t, x)$$

$$v(0, x, \varphi) = \varphi(x)$$

à savoir :

$$v(t, x) = E \varphi(X^x(t))$$

où $(X^x(t))$ désigne la solution de notre EDS avec x comme condition initiale à l'instant 0. \square

Remarquons que le raisonnement qui a conduit à l'équation progressive démontre aussi le fait suivant : si pour tout $t \geq 0$, $\mu(t)$ désigne la loi de probabilité de $X(t)$, on a :

$$\frac{\partial \mu}{\partial t}(t) = L^* \mu(t)$$

$$\mu(0) = \text{loi de } X_0$$

Cette équation est appelée équation de Fokker-Planck. En particulier, si μ est une *mesure invariante* (i.e. une probabilité telle que loi de $X_0 = \mu \Rightarrow$ loi de $X(t) = \mu, \forall t \geq 0$), alors

$$L^* \mu = 0$$

Cette équation est appelée équation de Fokker-Planck stationnaire.

Réciproquement, sous des hypothèses très faibles sur les coefficients, si μ est une mesure sur \mathbb{R}^d t. q. $L^* \mu = 0$, i.e.

$$\int_{\mathbb{R}^d} L f(x) \mu(dx) = 0, \quad \forall f \in C_c^\infty(\mathbb{R}^d)$$

alors μ est une mesure invariante. On montre que si $a(x) = g(x) g^t(x) > 0, \forall x \in \mathbb{R}^d$, il existe au plus une mesure invariante, et celle-ci possède nécessairement une densité. Nous allons calculer cette densité dans deux cas particuliers.

EXEMPLE 3.4.7 : Mesure invariante d'une diffusion scalaire.

On suppose ici que $d=k=1$, et que $g(x) \neq 0, \forall x \in \mathbb{R}$. On a alors le résultat suivant :

Si la fonction

$$x \rightarrow \frac{1}{a(x)} \exp \left(2 \int_0^x \frac{f(y)}{a(y)} dy \right)$$

est intégrable sur \mathbb{R} , alors le processus de Markov scalaire solution de l'EDS

$$dX(t) = f(X(t)) dt + g(X(t)) dB(t)$$

possède l'unique mesure invariante μ de densité $(p(x), x \in \mathbb{R})$ par rapport à la mesure de Lebesgue sur \mathbb{R} , avec :

$$p(x) = \frac{c}{a(x)} \exp \left[2 \int_0^x \frac{f(y)}{a(y)} dy \right]$$

où la constante c est déterminée par la condition $\int_{-\infty}^{+\infty} p(x) dx = 1$. On vérifie en effet aisément que

$$\begin{aligned} L^* p(x) &= \frac{1}{2} (ap)''(x) - (fp)'(x) \\ &= 0 \quad \square \end{aligned}$$

EXEMPLE 3.4.8 : Mesure invariante d'une classe de diffusions sur \mathbb{R}^d .

Soit $V = \mathbb{R}^d \rightarrow \mathbb{R}$ une fonction de classe C^2 , à dérivées secondes bornées, telle que :

$$\int_{\mathbb{R}^d} \exp[V(x)] dx = 1$$

Alors la diffusion dans \mathbb{R}^d solution de l'EDS :

$$dX(t) = \frac{1}{2} \nabla V(X(t)) dt + dB(t),$$

$$\text{où } \nabla V(x) = \left(\frac{\partial V}{\partial x_1}(x), \dots, \frac{\partial V}{\partial x_d}(x) \right)^t, \text{ possède}$$

l'unique mesure invariante μ de densité $(p(x), x \in \mathbb{R}^d)$ par rapport à la mesure de Lebesgue sur \mathbb{R}^d , avec :

$$p(x) = \exp [V(x)]$$

on vérifie en effet aisément que

$$\begin{aligned} L^* p(x) &= \frac{1}{2} \sum_{i=1}^d \frac{\partial^2 p}{\partial x_i^2}(x) - \frac{1}{2} \frac{\partial}{\partial x_i} \left(\frac{\partial V}{\partial x_i} p \right)(x) \\ &= 0 \end{aligned}$$

Notons que si $V(x)$ est une forme quadratique homogène :

$$V(x) = - |Ax|^2$$

$$\text{alors } \frac{1}{2} \nabla V(x) = - A^* A x$$

et on retrouve le cas particulier d'une équation de Langevin, auquel cas la mesure invariante est gaussienne $N(0, (2 A^* A)^{-1})$ (notons que la condition d'intégrabilité de $e^{V(x)}$ impose $A^* A > 0$). \square

Bibliographie

- BOULEAU, N. *Processus stochastiques et applications*, Hermann 1988.
- DAUTREY, R. et all. *Méthodes probabilistes pour les équations de la physique*, Coll. CEA, Eyrolles 1989.
- FLEMING, W.H. & RISHEL, R.W. (1975), *Deterministic and stochastic optimal control*, Springer-Verlag.
- FRIEDMAN, A. (1975) *Stochastic differential equations and applications*, 2 vol. Acad. Press.
- GIHMAN, I.I. & SKOROHOD, A.V. (1972) *Stochastic differential equations*. Springer-Verlag.
- HARRISON, J.M. (1985), *Brownian motion and stochastic flow systems*, J. Wiley.
- IKEDA, N. & WATANABE, S. (1981) *Stochastic differential equations and diffusion processes*, North Holland/Kodurska.
- KALLIANPUR, G. (1980) *Stochastic filtering theory*. Springer-Verlag.
- KARATZAS, I. & SHREVE, S.E. (1988) *Brownian motion and stochastic calculus*. Springer-Verlag.
- KRYLOV, N.V. (1980) *Controlled diffusion processes*. Springer-Verlag.
- LIPTSER, R.S. & SHIRYAEV, A.N. (1977) *Statistics of random processes*. Springer-Verlag.
- McKEAN, H.P., Jr. (1969) *Stochastic integrals*, Acad. Press.
- METIVIER, M. (1982) *Semimartingales*, de Grunster.
- ROGERS, G.G. & WILLIAMS, D. (1979) *Diffusions, Markov processes and martingales*. (2 vol.) Wiley.
- WONG, E. & HAJEK, B. (1985) *Stochastic processes in engineering systems*. Springer-Verlag.