ANNALES SCIENTIFIQUES
DE L'UNIVERSITÉ DE CLERMONT-FERRAND 2
*Série Mathématiques*

P. BERNARD

# Une présentation du calcul des variations stochastique (calcul de Malliavin) pour les non-spécialistes

*Annales scientifiques de l'Université de Clermont-Ferrand 2*, tome 89, série *Mathématiques*, n° 23 (1986), p. 13-45

<http://www.numdam.org/item?id=ASCFM_1986__89_23_13_0>

*Ann. Sci. Univ. Clermont II, Sér. Math., Fasc. 23. 1986, pp. 13-45*

**UNE PRESENTATION du CALCUL des VARIATIONS STOCHASTIQUE**

**(CALCUL de MALLIAVIN)**

**POUR LES NON-SPECIALISTES**

P. BERNARD

Exposé «invité» au Séminaire d'Analyse du Département de Mathématiques de l'Université de Clermont II.

**Plan :**

**1. Considérations diverses.**
**Presentation naïve de quelques concepts fondamentaux du calcul stochastique.**

1.0. Dans le développement récent des méthodes mathématiques à la disposition des ingénieurs et physiciens, les méthodes probabilistes prennent une part croissante. Nous nous intéressons ici à la modélisation de phénomènes évoluant dans le temps, c'est-à-dire à ce qu'il est convenu d'appeler des systèmes dynamiques. Les aspects stochastiques se retrouvent alors à plusieurs niveaux. L'un de ces aspects, non considéré ici, réside dans les mouvements chaotiques observés dans certains systèmes physiques et associés au comportement asymptotique de systèmes dynamiques dissipatifs (les fameux «attracteurs étranges»). Il est à noter que ceci apparaît dans des systèmes régis par des équations purement déterministes. Un autre aspect réside dans la nécessité de prendre en compte dans la modélisation de systèmes dynamiques, outre les agents principaux du mouvement supposés bien connus, c'est-à-dire régis par des lois physiques parfaitement établies, des agents secondaires (des perturbations) connus avec beaucoup d'imprécision (loi physique connue approximativement, avec incertitude sur les paramètres) ou encore très mal connus (loi physique inconnue, la connaissance du phénomène se résumant à des statistiques d'observations), toutes les observations étant en outre entachées d'erreurs d'observation.

Ainsi en est-il de modèles mathématiques modernes de la physique quantique (mécanique quantique «à la Nelson», intégrale de Feynman) ; de modèles statistiques de la thermodynamique (avec en vedette le célébre et toujours fondamental mouvement brownien de Wiener-Levy) ; de modèles de développement des amas stellaires ; de modèles de la théorie du signal : dans le cadre de l'étude de la propagation des ondes électromagnétiques, «filtrage» d'un signal «bruité» transmis par un «système», opération visant à éliminer le bruit et récupérer le signal original (ceci sera le modèle fondamental justificatif des développements qui vont suivre).

Le modèle ci-dessus (filtrage) sera en fait impliqué dans la modélisation de nombreux problèmes concrets, autres que celui de la propagation des ondes électromagnétiques : en aéronautique par exemple, le problème du contrôle de la trajectoire d'un satellite et plus généralement de la navigation automatique (problème à l'origine d'un développement important, dans les années 60, des méthodes mathématiques et numériques du filtrage stochastique (filtrage de Kalman-Bucy)) ; en écologie, pour l'évolution d'un écosystème ; modélisation de la dynamique d'un système économique ; etc...

Le groupe de recherche «Méthodes mathématiques de la géophysique» au sein de l'unité associée n° 739 du C.N.R.S., à Clermont, a été amené à utiliser des modèles de ce type pour deux applications : modélisation de transport de polluants par diffusion turbulente atmosphérique, et analyse du comportement de structures soumises à sollicitation aléatoire, avec en projet une étude dans le cadre du génie parasismique . Dans un tel cadre (séismes), la nature de l'excitation amène tout naturellement à l'utilisation de méthodes stochastiques ; les caractéristiques de la sollicitation (forte ampleur pendant un intervalle de temps court) nous amenant à sortir du domaine de réponse élastique des structures, donc à utiliser des méthodes d'analyse stochastique non linéaire et à mener une étude en régime transitoire.

**1.1. Le mouvement brownien - L'équation de Langevin.**

**1.1.0. Le mouvement brownien de Wiener-Levy.**

Le mouvement brownien de Wiener-Levy d-dimensionnel issu de $x \in \mathbb{R}^n$ est le processus $(W_t\,, t > 0)$ à valeurs dans $\mathbb{R}^d$, à trajectoires continues, tel que

$$\forall\, n \in \mathbb{N} \qquad \forall 0 = t_o < t_1 < \ldots < t_n$$

$W_{t_o}, W_{t_1} - W_{t_o}, \ldots, W_{t_n} - W_{t_{n-1}}$ sont des variables aléatoires mutuellement indépendantes telles que

$W_{t_o} = x$ p - s - et la loi de $W_{t_j} - W_{t_{j-1}}$ soit la loi gaussienne centrée de densité

$$(2\pi\ (t_j - t_{j-1}))^{-d/2}\ \exp(-\ |x|^2 / 2\,(t_j - t_{j-1})).$$

Il est bien connu que presque toute trajectoire de ce processus est à variation non bornée sur tout intervalle.

(**Remarque.** Faisons un lien avec l'exposé de B. Brunet sur la dimension topologique. Une trajectoire du mouvement brownien est un objet curieux de ce point de vue. Considérons le mouvement brownien 1-dimensionnel, et une trajectoire $(t, W_\omega(t))$ de ce brownien comme sous-ensemble de $\mathbb{R}^2$. Sa dimension topologique est 1. Mais c'est une courbe «fractale» au sens de Mandelbrot. Sa dimension de Hausdorff n'est pas égale à sa dimension topologique (contrairement à ce qui se passe pour une courbe définie par une fonction $C^1$ par morceaux), et vaut 3/2.

L'ensemble des zéros de cette courbe est lui-même un fractal (un objet ressemblant à l'ensemble de Cantor), qui est un ensemble de mesure de Lebesgue 0 sur $\mathbb{R}$, de dimension de Hausdorff 1/2, alors que sa dimension topologique est 0).

**1.1.1. Quelques définitions.**

A cause de ce fait, on ne peut pas donner de sens, pour $\omega$ fixé, à

$\int_0^t y_\omega(s)\, dW_\omega(s)$ avec $t \to y_\omega(t)$ continue par exemple.

Mais on peut quand même, par l'intégrale stochastique de Ito, donner un sens à

$$Z(t, \omega) = \int_0^t Y(s, \omega)\, dW(s, \omega) \tag{1.1}$$

pour certains processus Y.

La classe des processus Y pour lesquels l'intégrale de Ito ci-dessus a un sens, contient en particulier, l'espace $\mathcal{L}_2$ de tous les processus $(Y(t, \omega), t \geq 0)$ sur $\Omega$, mesurables, $\mathcal{F}_t$-adaptés et tels que

$$\forall T > 0 \quad E\left[\int_0^T Y^2(s, \omega)\, ds\right] < \infty \tag{1.2}$$

Rappelons ici les définitions utiles :

La donnée de base est un espace de probabilité $(\Omega, \mathcal{F}, P)$ et une famille croissante $(\mathcal{F}_t, t \geq 0)$ de sous-tribus de $\mathcal{F}$, continue à droite.

$(X_t, t \geq 0)$ est $\mathcal{F}_t$-*adapté* ssi $X_t$ est $\mathcal{F}_t$-mesurable $\forall t \geq 0$.

$(X_t, t \geq 0)$ est *mesurable* ssi $X : [0\ \infty[ \times \Omega \ni (t,x) \rightsquigarrow X(t; \omega) \in \mathbb{R}^d$ est mesurable pour $\mathcal{B}([0, \infty)) \times \mathcal{F}/\mathcal{B}(\mathbb{R}^d)$ où $\mathcal{B}(T)$ désigne la tribu borélienne sur l'ensemble T.

*La tribu des ensembles $\mathcal{F}_t$-prévisibles* $\mathcal{S}$ est la plus petite tribu sur $[0, \infty) \times \Omega$ rendant mesurables tous les processus continus et $\mathcal{F}_t$-adaptés.

$(X_t, t \geq 0)$ est *prévisible* s'il est mesurable pour la tribu prévisible.

$(X_t, t \geq 0)$ est *progressivement mesurable* ssi pour tout $t > 0$, la restriction de X à $[o,t] \otimes \Omega$ est $\mathcal{B}([o,t]) \otimes \mathcal{F}_t$-mesurable.

Tout processus prévisible est progressivement mesurable, et tout processus de $\mathcal{L}_2$ est égal $\lambda \otimes P$ presque surement à un processus prévisible ( $\lambda$ : mesure de Lebesgue).

**1.1.2. Diffusions. Semi-groupes markoviens. Générateur infinitésimal.**

Considérons l'Equation Différentielle Stochastique (E.D.S.)

$$(1.3) \quad X_t = x + \int_0^t a(X_s)dB_s + \int_0^t b(X_s)ds$$

$$a : \mathbb{R}^d \longrightarrow \mathbb{R} \qquad b : \mathbb{R}^d \longrightarrow \mathbb{R}$$

où les coefficients sont supposés satisfaire des conditions de Lipschitz.

$\int_0^t a(X_s)dB_s$ est prise au sens d'Ito.

Cette équation admet alors une solution unique partant de x : $X_t^x$ , et si $f \in C_b^2 = \left\{ f : \mathbb{R}^f \longrightarrow \mathbb{R} \text{ de classe } C^2 \text{ et dont les deux premières dérivées sont bornées sur } \mathbb{R}^d \right\}$, on peut définir le semi-groupe

$$(1.4) \quad P_t(x,f) = E\left[f(X_t^x)\right] \qquad t \in \mathbb{R} \quad , \quad x \in \mathbb{R}^d.$$

Il existe un opérateur L, défini sur $C_b^2$ , appelé générateur infinitésimal du semi-groupe markovien $P_t$ , qui vérifie :

$$(1.5) \quad \frac{\partial}{\partial t} \quad P_t f = P_t Lf \qquad \forall f \in C_b^2 \qquad \forall t \in \mathbb{R}^+$$

L est un opérateur différentiel du second ordre semi-elliptique :

$$(1.6) \; Lf = \frac{1}{2} \sum_k a_k^i \, a_k^j \quad D_{ij}f + b^i D_i f \qquad f \in C_b^2$$

les indices croisés indiquant une sommation sous-entendue.

On veut étudier la régularité de la mesure $P_t(x, dy)$.

$$(P_t(x, f) = \int f(y)\ P_t(x, dy)).$$

Ainsi, pour le mouvement brownien de Wiener-Levy :

(1.7) $\quad P_t(x, dy) = p(t, x-y)dy$

et $L = \frac{1}{2}\ \Delta$

où $\Delta$ est le laplacien spatial et

$$p(t,x) = \frac{1}{(\sqrt{2\pi t})^d} \exp\left(-\frac{1}{2}\ \frac{\|x\|^2}{t}\right)$$

**1.2. L'équation de Langevin - L'opérateur d'Ornstein-Uhlenbeck.**

L'équation de Langevin régit le «mouvement brownien physique», tel qu'il a été décrit par Ornstein et Uhlenbeck dans un article datant de 1930. Il s'agit d'un modèle pour la dynamique d'une particule plongée dans un fluide, particule de petite taille, mais de dimension grande par rapport à celle des molécules du fluide. Langevin a imaginé de dissocier les forces génératrices du mouvement en deux parts : la force due au bombardement de la particule par les molécules du fluide mues par l'agitation thermique, et qui sera mathématiquement modélisée par un bruit blanc, (ce sera la partie aléatoire) ; la force de frottement due à la viscosité et qui tend à freiner le mouvement.

Formellement, la vitesse d'une telle particule sera solution de :

(1.8) $\quad \frac{dV}{dt} = -\alpha V + N_t$ (Equation de Langevin)

où $\alpha$ est une constante $> 0$ caractérisant la viscosité du fluide, $N_t$ un bruit blanc.

Ceci prend un sens mathématiquement correct sous la forme d'une équation différentielle stochastique d'Ito :

(1.9) $\quad dV_t = -\alpha V_t\, dt + dW_t$

où $W_t$ est un processus de Wiener-Levy 3-dimensionnel.

La solution $V_t^{v_o}$ de cette équation issue de $v_o$ est le *processus d'Ornstein-Uhlenbeck.*

C'est un processus de diffusion gaussien ; le noyau de probabilité de transition est solution de l'*équation de Kolmogorov en avant,* ou *équation de Fokker-Planck :*

(1.10) $$\frac{\partial p}{\partial t} - \frac{1}{2} L p = 0$$

où $\frac{1}{2}$ L est le générateur infinitésimal, et L s'écrit :

(1.11) $$L = \Delta_x - x.\nabla_x$$

où $\Delta_x$ est le laplacien spatial et $\nabla_x$ le gradient spatial.

L sera appelé *l'opérateur d'Ornstein-Uhlenbeck* et jouera un rôle important dans le calcul de Malliavin.

On montre que $V_t^x$ s'écrit : (1.12) $V_t^x = e^{-t} W^x_{e^{2t}-1}$

où $W_t^x$ est un processus de Wiener-Levy issu de x.

Le semi-groupe d'Ornstein-Uhlenbeck est alors lié à la loi $\mu$ du processus de Wiener-Levy par la formule de Mehler :

(1.13) $$P_t(u,f) = \int f(u\, e^{-t} + w \sqrt{1 - e^{-2t}}) \quad \mu(dw).$$

Il est à noter que, alors que le semi-groupe de la chaleur n'admet pas de mesure de probabilité invariante (i.e. une probabilité $\nu$ telle que

$\int P_t(x, dy)\, d\,\nu\,(x) = \nu\,(dy)$) (il admet une mesure invariante : la mesure de Lebesgue), le semi-groupe d'Ornstein-Uhlenbeck admet lui une mesure de probabilité invariante qui est gaussienne. (En physique statistique, cette mesure de probabilité est appelée *mesure de Gibbs,* et l'énoncé précédent est une version de la loi de Boltzmann pour les gaz parfaits).

**1.3. Martingales.**

Dans les problèmes physiques qui nous intéressent, la variable $X_t$ peut se décomposer en deux :

$$(14) \qquad X_t = \overline{X}_t + Y_t$$

où $\overline{X}_t$ représente le «mouvement moyen» et $Y_t$ les fluctuations du mouvement autour du mouvement moyen. Naturellement, on peut concevoir, si on considère maintenant $Y_t$, d'effectuer sur $Y_t$ la même décomposition. Et ainsi de suite. La question est alors : quand va-t-on s'arrêter ? La partie $Y_t$ ainsi obtenue à terme, celle représentant des fluctuations trop chaotiques pour qu'on puisse en extraire une quelconque tendance moyenne, sera représentée mathématiquement par le concept de *martingale*.

Une martingale sera donc un processus suffisamment fluctuant pour que la meilleure prédiction que l'on puisse faire de son comportement à l'instant t + h à partir de son passé jusqu'à l'instant t soit sa valeur à l'instant t.

**1.3.1. Définition.**

Etant donné $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t \geq 0})$ comme précédemment, le processus $(M_t, t \geq 0)$ à valeurs dans $\mathbb{R}$ sera une $\mathcal{F}_t$-martingale ssi :

(i) $M_t$ est P-intégrable $\forall t \geq 0$

(ii) $(M_t)$ est $\mathcal{F}_t$-adapté

(iii) $E(M_t / \mathcal{F}_s) = M_s$ p.s. $\quad \forall\, 0 \leq s < t$.

**1.3.2.**

Le processus $(M_t, t \geq 0)$ à valeurs dans $\mathbb{R}$ sera une $\mathcal{F}_t$-martingale locale ssi il est $\mathcal{F}_t$-adapté et s'il existe une suite $(\sigma_n, n \in \mathbb{N})$ de $\mathcal{F}_t$-temps d'arrêts telle que $\sigma_n < \infty$, $\sigma_n \uparrow \infty$ et $M(t \wedge \sigma_n)$ soit une $\mathcal{F}_t$-martingale pour chaque $n \in \mathbb{N}$.

**1.3.3. Notations.**

$\mathcal{M}_2^c$ désignera la famille des $\mathcal{F}_t$-martingales continues de carré intégrable telles que $M_0 = 0$ p.s.

$\mathcal{M} = \mathcal{M}_2^{c,loc}$ est la famille des $\mathcal{F}_t$-martingales locales continues de carré intégrable telles que $M_0 = 0$ p.s.

$\mathcal{A}_+$ est la famille de tous les processus $(A_t, t \geq 0)$ continus, $\mathcal{F}_t$-adaptés, tels que $A_0 = 0$ p.s. et $t \to A_t$ soit p.s. croissante.

$\mathcal{A}$ est la famille de tous les processus $(A_t, t \geq 0)$ continus, $\mathcal{F}_t$-adaptés, tels que $A_0 = 0$ p.s. et $t \to A_t$ soit p.s. à variations bornées sur tout intervalle fini.

$\mathcal{B}$ est la famille de tous les processus $(\Phi_t, t \geq 0)$ $\mathcal{F}_t$-prévisibles, tels que p.s. $t \to \Phi_t$ soit bornée sur tout intervalle borné.

Le processus $X = (X_t, t \geqslant 0)$ sera une *quasi-martingale* (ou semi-martingale continue) s'il s'écrit sous la forme :

(1.15) $\qquad X_t = X_0 + M_t + A_t \qquad \forall t \geq 0$

où $X_0$ est une variable aléatoire $\mathcal{F}_0$-mesurable, $M \in \mathcal{M}$, $A \in \mathcal{A}$.

L'écriture ci-dessus est la *décomposition canonique* (unique) de X.

La solution de l'E.D.S. (3) est une quasi-martingale.

### 1.3.4. Le crochet < , > de deux martingales.

Soit M, N deux éléments de $\mathcal{M}_2^c$. Le produit $M_t.N_t$ n'est pas en général une martingale. Mais c'est une quasi-martingale, et donc admet une décomposition canonique.

**Définition :**

Soient M et N $\in \mathcal{M}_2$. Il existe alors un processus à variations bornées unique : $A_t$ tel que $M_tN_t - A_t$ soit une martingale.

On note $A_t = \langle M,N \rangle_t$ et on l'appelle le «crochet» de M et N.

Si $M \in \mathcal{M}$, on peut considérer $\langle M, M \rangle_t$. On note alors $\langle M \rangle_t$, et l'on appelle *processus variation quadratique* de la martingale $M_t$. Dans le cas du mouvement brownien de

Wiener-Levy, $W_t$ est une martingale et le processus variation quadratique $<W>_t$ est déterministe : c'est $t \to t$.

**1.3.5. La formule d'Ito.**

C'est un résultat fondamental du calcul différentiel stochastique, qui donne la règle de dérivation des fonctions composées dans ce cadre.

Soit $X^i(t) = X^i(0) + M^i(t) + A^i(t) \qquad i = 1, \ldots, d \qquad d$

quasi-martingales, et $X(t) = (X^1(t), \ldots, X^d(t))$.

$X(t)$ sera une quasi-martingale d-dimensionnelle.

Soit F une fonction de classe $C^2$ sur $\mathbb{R}^d$.

Alors le processus $F[X(t)]$ est encore une quasi-martingale relativement à $(\mathcal{F}_t, t \geq 0)$ et l'on a :

$$
\begin{aligned}
(1.16) \quad F[X(t)] - F[X(0)] = & \sum_{i=1}^{d} \int_o^t F'_i[X(s)]\, d M^i(s) \\
& + \sum_{i=1}^{d} \int_o^t F'_i[X(s)]\, d A^i(s) \\
& + \frac{1}{2} \sum_{i,j=1}^{n} \int_o^t F''_{ij}[X(s)]\, d <M^i, M^j>(s)
\end{aligned}
$$

Dans le cas particulier où $M_t = W_t$, le Wiener-Levy, on a

$d <M>^i(s) = ds$ et $d<M^i, M^j>(s) = \delta_{ij}\, ds$.

Remarquons que, par rapport à la règle de dérivation des fonctions composées du calcul différentiel ordinaire, il apparaît un terme de plus, avec la variation quadratique, et une hypothèse plus fortes : c'est que F soit de classe $C^2$. C'est que pour une martingale, les variations «du second ordre», qui correspondent à la variation quadratique, ne sont plus négligeables devant les variations du 1er ordre de la partie à variations bornées.

**2. Introduction. L'exemple de l'aire de Levy.**

2.1. Soit $W_o^r = \mathscr{C}_0([0,1] \to \mathbb{R}^r)$ l'espace des fonctions continues de $[0,1]$ dans $\mathbb{R}^r$ nulles en 0. Muni de la norme uniforme, c'est un Banach. Notons P la mesure de Wiener r-dimensionnelle sur $W_o^r$, c'est-à-dire la loi du mouvement brownien r-dimensionnel, et $\mathscr{B}$ la tribu borélienne sur $W_o^r$.

L'espace de probabilité $(W_o^r, \mathscr{B}, \mathbb{P})$ est l'*espace de Wiener r-dimensionnel.*

Considérons maintenant l'équation différentielle stochastique

$$(2.1) \quad \begin{cases} dX^i(t) = \sum_{k=1}^{n} \sigma_k^i[X(t)]\, dW^k(t) + b^i[X(t)]\, dt & i = 1,\dots,d \\ X^i(0) = X^i & x \in \mathbb{R}^d \end{cases}$$

$$b^i : \mathbb{R}^d \to \mathbb{R} \qquad \sigma_k^i : \mathbb{R}^d \to \mathbb{R} \qquad i = 1,..,d \quad k = 1,..,n$$

On sait que, moyennant certaines hypothèses sur les coefficients, cette équation admet une solution unique.

Soit $X(t,x,w)$ cette solution, considérée sur l'espace de Wiener d-dimensionnel.

C'est le processus de diffusion issu de x de générateur infinitésimal

$$(2.2)\ L = \frac{1}{2} \sum_{i,j=1}^{d} a^{ij}(x) \frac{\partial^2}{\partial x^i\, \partial x^j} + \sum_{i=1}^{d} b^i(x) \frac{\partial}{\partial x^i}$$

où $[a^{ij}]$ est la matrice symétrique positive $\sigma \cdot \sigma^T$.

L'application $t \to X(t,x,w)$ pour x et w fixés est une *trajectoire* du processus.

L'application $x \to X(t,y,w)$ pour t et w fixés est une application de $\mathbb{R}^d \to \mathbb{R}^d$ qui est un difféomorphisme dès que les coefficients sont assez réguliers : c'est le flot stochastique associé à (2.1).

L'application $w \to X(t,x,w)$ pour t et x fixées est une application mesurable de $W_o^d \to \mathbb{R}^d$. Une telle application est appelée *fonctionnelle de Wiener.*

Or, l'application $w \longrightarrow X(t,xw)$ n'est en général pas continue de $W_o^d$ muni de la norme uniforme dans $\mathbb{R}^d$. Nous verrons cela dans l'exemple de l'aire de Levy.

Ceci montre qu'on ne peut, en général, espérer avoir la robustesse du filtrage non-linéaire par exemple pour la norme uniforme.

En effet, considérons le problème du filtrage non-linéaire :

$$(2.3)\quad \begin{cases} dX_t = X_o(X_t, Y_t)dt + X_i(X_t, Y_t)d\,W_t^i + \tilde{X}_j(X_t,Y_t)d\,\tilde{W}_t^j \\ dY_t = Z_o(Y_t)dt + Z_j(Y_t)\,d\tilde{W}_t^j \qquad i = 1, .., m \quad j = 1, .., d \end{cases}$$

où $W_t$ et $\tilde{W}_t$ sont des browniens indépendants et $X_o,\dots,X_m$ $\tilde{X}_1,\dots,\tilde{X}_d$ , $Z_o, \dots, Z_d$ sont des champs de vecteurs suffisamment réguliers.

$X_t$ représente un signal inconnu à estimer, $Y_t$ est l'observation :

c'est une fonction du signal, bruitée par un bruit d'observation.
Il s'agit d'estimer la loi de $X$ à partir de l'observation de $Y$ de $0$ à $t$.
(C'est donc une fonction que l'on observe, et le résultat de l'estimation est aussi une fonction).

On veut donc calculer, pour de bonnes fonctions f (continues bornées par exemple), $E\left[f(X_t) \,/\, \mathscr{B}_t^Y\right]$ , où $\mathscr{B}_t^Y$ est la tribu du passé de $Y_t$ : $\mathscr{B}_t^Y = \sigma\,(Y_s\,, 0 \leq s \leq t\,)$.

On suppose tous les coefficients de (2-1) $C^\infty$ bornés ainsi que leurs dérivées.

Le filtre sera dit *robuste* si, à une petite variation de l'observation, correspond une petite variation de l'estimation. C'est donc une propriété de régularité de l'opérateur $\Pi_t f$ qui, à une observation y , associe

$$\Pi_t f(y) = E\left[f(X_t)/\ \mathscr{B}_t^y\right] \in \mathscr{C}(\mathbb{R}).$$

Et bien, cette régularité ne pourra pas s'exprimer par une propriété de continuité sur $W_o^d$ muni de la norme uniforme.

**2.2 L'aire de Levy.**

Soit $(B^1, B^2)$ un couple de browniens 1-dimensionnels indépendants.

L'aire de Levy est : $\int_0^1 B_s^2\, dB_s^1 - B_s^1\, dB_s^2$ .

Si $\mu$ désigne la mesure de Wiener bidimensionnelle sur $W_o^2$, l' aire de Levy est la fonctionnelle de Wiener :

$$S(w) = \int_o^1 w_s^2\, d\, w_s^1 - w_s^1\, d\, w_s^2$$

$$S : W_o^2 \longrightarrow \mathbb{R}.$$

Notons $\hat{S}$ la restriction de S à $\mathscr{C}_o^{(2)}([0,1] \longrightarrow \mathbb{R}^2)$. Si S admettait une modification continue $\hat{S}$ pour la norme uniforme sur $W_o^2$ , alors sur $\mathscr{C}_o^{(2)}([0,1] \longrightarrow \mathbb{R}^2)$, $\hat{S}$ coïnciderait avec $\hat{\hat{S}}$ définie par

$$\hat{\hat{S}}(W) = \int_o^1 (w_s^2\, \dot{w}_s^1 - w_s^1\, \dot{w}_s^2)ds$$

Or, on sait que $\hat{\hat{S}}$ n'a pas de prolongement continu à $W_o^2$.

D'où une contradiction.

**3. Opérateurs hypo-elliptiques et régularité des noyaux de probabilité de transition de la diffusion associée.**

Considérons l'équation différentielle stochastique (E.D.S.) suivante, dans $\mathbb{R}^d$ :

$$d X_t = \sigma(X_t)dW(t) + b(X_t)dt \ , \quad X_o = x \qquad \text{en notation matricielle}$$

(C'est l'équation (2.1)).

La théorie classique des E.D.S. affirme que, si les coefficients satisfont une condition de Lipschitz, alors cette équation admet une solution unique : $X_t^x = X(t,x,w)$. On peut alors poser, pour f continue bornée sur $\mathbb{R}^d$ : $\quad P_t(x,f) = E[f(X_t^x)]$

et, si f a des dérivées aux deux premiers ordres continues bornées,

$$\frac{\partial}{\partial t} P_t(x,f) = P_t Lf$$

où L est le générateur infinitésimal déjà considéré au paragraphe 2.1.

Considérons maintenant le problème de la régularité du système $(P_t(x,dy))_{t \geq 0}$ de probabilités de transitions. On a alors deux méthodes et deux types de résultats :

- Soit on exige peu de régularité et beaucoup d'ellipticité (Stroock-Varadhan)
- Soit on admet une possible dégénérescence, mais on suppose une régularité $C^\infty$.

C'est ce cadre qui sera retenu ici, et l'on supposera donc désormais que $\sigma$ et b ont des dérivées de tous ordres continues et bornées.

**La condition d'hypo-ellipticité de Hörmander.**

On lit maintenant l'E.D.S. ci-dessus non plus au sens de Ito, mais au sens de Stratonovitch (ce qui revient simplement à modifier b).

Cette équation ainsi considérée a alors une interprétation géométrique : si nous interprétons $\sigma_k$ et b comme des champs de vecteurs sur $\mathbb{R}^d$, elle se transforme bien par les difféomorphismes de $\mathbb{R}^d$. Le générateur infinitésimal L s'écrit sous la forme des «sommes de carrés» :

$$L = \frac{1}{2} \sum_k (\sigma_k)^2 + b.$$

**Théorème.** (Condition de Hörmander)

Si, à chaque point x, le rang de l'algèbre de Lie engendrée par les champs de vecteurs

$$\sigma_k, [\sigma_k, \sigma_l], [b, \sigma_k]; [\sigma_k[\sigma_l, \sigma_m]], [b, [\sigma_k, \sigma_e]],$$

$$[b, [b, \sigma_k)]]; \ldots..$$

est égal à d, L est hypo-elliptique, ce qui implique que $P_t(x, dy)$ ait une densité $C^\infty$.

(Rappelons que l'opérateur différentiel L est dit hypo-elliptique si l'application définie dans $\mathscr{D}'(\mathbb{R}^d)$, espace des distributions $f \rightsquigarrow u$ où u est solution de $Lu = f$ n'augmente pas le support singulier des distributions).

La méthode probabiliste de démonstration de ce résultat, reposant sur le calcul de Malliavin, semble s'appliquer à des situations dans lesquelles les méthodes d'analyse ne fonctionnent pas.

**4. Espaces de Wiener abstraits.**

Soit W un Banach séparable, et $\mathscr{B}(W)$ sa tribu borélienne. Notons W' le dual topologique de W.

**4.1. Définition.**

Une mesure de probabilité $\mu$ sur $(W, \mathscr{B}(W))$ sera dite *mesure gaussienne centrée* si et seulement si :

$\forall n \in \mathbb{N}, \quad \forall l_1, \ldots, l_n \in W'$

$(l_1, \ldots, l_n)$ considéré comme vecteur aléatoire à valeurs dans $\mathbb{R}^n$ défini sur $(W, \mathscr{B}(W), \mu)$ a une loi gaussienne centrée, i.e. $\exists$ V matrice symétrique positive telle que

$\forall c \in \mathbb{R}^n$

$$\int_W \exp\left\{ i \sum_{j=1}^{d} c_j l_j(w) \right\} \mu(dw) = \exp\left\{ -\frac{1}{2} < Vc, c > \right\}$$

où $<\ ,\ >$ est le produit scalaire sur $\mathbb{R}^n$.

On peut toujours se ramener au cas où le support de $\mu$ est W.

**Théorème.**

Pour toute mesure gaussienne centrée $\mu$ sur $(W, \mathcal{B}(W))$, il existe un Hilbert séparable $H \subset W$ tel que l'inclusion

$i : H \hookrightarrow W$ soit continue à image dense, et

$$(4.1) \quad \int_W \exp\{i \; l(W)\} \, \mu(dw) = \exp\left(-\frac{1}{2} \|l\|_H^2\right)$$

$\forall\, l \in W'$.

**Remarque :**

$$H \hookrightarrow W \Longrightarrow W' \hookrightarrow H' \simeq H$$

et, pour $h \in H, l \in W$ $\qquad l(h) = \langle l,h \rangle_H$ .

**4.2. Définition.**

Le triplet $(W,H,\mu)$ est appelé *espace de Wiener abstrait.*

**Exemple.**

Revenons à l'espace de Wiener (concret ? ) $W = W_o^r$ ,

$\mu$ : mesure de Wiener r-dimensionnelle.

Alors :

$H = \Big\{ h = (h^i)_{i=1}^r \in W_o^r$ t.q. $h^i(t)$ soit absolument continue sur $[0, 1]$ et sa dérivée $\dot{h}^i \in L^2([0,1] \to \mathbb{R})$ , $i = 1,\ldots, r \Big\}$ muni du produit scalaire

$$\langle h, g \rangle_H = \sum_{i=1}^r \int_o^1 \dot{h}^i(t) \; \dot{g}^i(t)\, dt.$$

C'est l'espace autoreproduisant du mouvement brownien, et il est connu dans la littérature sous le nom d'*espace de Cameron-Martin.*

**5. Un calcul en dimension finie.**

Considérons dans ce paragraphe un espace vectoriel réel V de dimension finie n, et une mesure gaussienne centrée portée par V. Soit $V^*$ le dual de V.

$\mu$ est définie par :

$$\int_W \exp(i \;< l,v >)\;\; \mu(dv) \;=\; \exp(-\frac{1}{2}\;\|l\|_\mu^2\;) \qquad \forall\, l \in V^*$$

où $\|l\|_\mu$ est la norme associée au produit scalaire $< \Gamma l,\; l >_{V^*}$, $\Gamma$ étant la matrice des covariances de $\mu$ .

**5.1. Quelques notations.**

Soit $\vartheta$ un espace euclidien muni de la norme $\|\bullet\|_\vartheta$ et du produit scalaire $< \bullet, * >_\vartheta$ . Notons $\Lambda(V, \vartheta)$ l'espace euclidien des applications linéaires $l : V \to \vartheta$ muni du produit scalaire de Hilbert-Schmidt $< l, l' >_{H.S.} = \sum_{i=1}^{n} < l(e_i), l'(e_i) >_\vartheta$ pour

$(e_i)_{1 \leq i \leq n}$ B.O.N. de V (B.O.N. = Base Orthonormée).

Pour tout $k \in \mathbb{N}$, $\Lambda_k(V, \vartheta)$ désigne l'espace des applications k-linaires de V dans $\vartheta$ . On peut le définir par récurrence comme suit :

$$\Lambda_o(V, \vartheta) = \vartheta$$

$$\Lambda_k(V, \vartheta) = \Lambda(\,V, \Lambda_{k-1}(V, \vartheta)) \qquad k \geq 1.$$

Si $\vartheta = R$ , on note $\Lambda_k(V)$.

$\Lambda_1(V)$ est identifié à $V^*$ lui-même identifié à V et les deux normes coïncident.

$\Lambda(V, \vartheta)$ est aussi noté $V \otimes \vartheta$ et $\Lambda_k(V)$ : $\underbrace{V \otimes \ldots \otimes V}_{k}$.

Si $l \in \Lambda_2(V, \vartheta)$, la trace de l est l'élément Trl de $\vartheta$ défini par :

$$Tr\, l = \sum_{i=1}^{n} (l[e_i])\;[e_i]\;\text{pour } (e_i) \text{ B.O.N. de } V.$$

On notera $L_p(V \to \vartheta, \mu)$, $1 \leqslant p < \infty$, l'espace des fonctions de V dans $\vartheta$ de puissance $p^{\text{ième}}$ intégrables pour $\mu$, muni de la norme habituelle :

$$\| f \|_p = \left[ \int_V \|f(x)\|_\vartheta^p \quad \mu(dx) \right]^{1/p}.$$

**5.2. L'opérateur divergence relativement à $\mu$ .**

Si $f : V \to \vartheta$ est différentiable, on note Df sa dérivée.

$Df : V \to \Lambda \quad (V, \vartheta)$, définie par :

$$Df(x)[y] = \frac{d}{dt} f(x + ty)\big|_{t = 0} \qquad \forall y \in V.$$

Puis on définit la dérivée d'ordre k, $D^k f$, par récurrence :

$$D^k f : V \to \Lambda_k(V, \vartheta) \qquad\qquad D^k f = D(D^{k-1} f) \qquad k \geqslant 1.$$

Si f est seulement localement $\mu$-intégrable, $D^k f$ sera défini au sens des distributions.

Nous allons maintenant définir l'opérateur divergence $\delta$ :

Si $f : V \to V$ (resp. $f : V \to \Lambda \quad (V, \vartheta)$)

alors $\delta f : V \to \mathbb{R}$ (resp. $\delta f : V \to \vartheta$) est défini par

$$(5.1) \qquad \int_V \delta f(x) g(x) \quad \mu(dx) = - \int_V < f(x), Dg(x) >_V \quad \mu(dx)$$

$\forall g : V \to \mathbb{R} \quad C^\infty$ à support compact.

(Resp. : $\displaystyle\int_V < \delta f(x), g(x) >_V \quad \mu(dx) = - \int_V < f(x), Dg(x) >_{H.S.} \quad \mu(dx)$ )

$\forall g : V \to \vartheta \quad C^\infty$ à support compact.

Dans la norme sur V induite par dualité de la norme sur $V^*$ définie par la forme quadratique de matrice $\Gamma$ la matrice des covariances de $\mu$, $\mu$ est normale réduite.

On calcule alors, pour $(e_i)$ B.O.N. de V :

$$\delta f(x) = \sum_{i=1}^{n} \langle Df(x)\,[e_i], e_i \rangle_V - \langle f(x), x \rangle_V$$

$$(5.2) \qquad = \mathrm{Tr}\, Df(x) - f(x)\,[x] \qquad \text{pour } f, V \to V$$

Le dernier résultat étant encore valable pour $f : V \to \Lambda(V, \vartheta)$.

**5.3. L'opérateur d'Ornstein-Uhlenbeck.**

Pour f deux fois différentiable de $V \to \mathbb{R}$ (resp. $V \to \vartheta$), définissons

$Lf : V \to \mathbb{R}$ (resp. $V \to \vartheta$) par :

$$Lf(x) = \delta(Df)(x)$$

$$= \mathrm{Tr}\, D^2 f(x) - Df(x)\,[x]$$

On voit immédiatement que

$$(5.3) \qquad \int_V (Lf(x))g(x)\, \mu(dx) = \int_V f(x)\, Lg(x)\, \mu(dx) = -\int_V \langle Df(x), Dg(x) \rangle_V\, \mu(dx)$$

si $f, g : V \to \mathbb{R}$ sont $C^2$ et de carré $\mu$-intégrable ainsi que leurs deux premières dérivées.

Revenons à la mesure de Lebesgue $\lambda$ sur $\mathbb{R}^n$. Soit $A = \frac{1}{2}\Delta$ le générateur infinitésimal du mouvement brownien. Notons, pour f et g deux fonctions de classe $C^2$ sur $\mathbb{R}^n$ :

$$\Gamma(f,g) = \nabla f.\ \nabla g$$

On vérifie aussitôt que $\Gamma(f,g) = A(fg) - f\,A(g) - g\,A(f)$.

On a la propriété de symétrie

$$(5.4) \qquad \int_V g(x) A h(x)\, \lambda(dx) = \int_V A g(x)\, h(x)\, \lambda(dx)$$

et si $g^1, ..., g^k$ sont des fonctions de V dans $\mathbb{R}$ et F est une fonction de $\mathbb{R}^k$ dans $\mathbb{R}$ toutes assez régulières :

$$(5.5) \qquad \Gamma(F(g^1,...,g^k), h) = \sum_{i=1}^{k} \Gamma(g^i, h)\, D_i F(g^1,...,g^k).$$

Appliquant (5.5) aux fonctions composantes de X, $X : V \to \vartheta$ , suffisamment régulier on obtient :

$$(5.6) \qquad \Gamma(f(X), X^i) = \sum_j \Gamma(X^i, X^j)\, D_i f(X).$$

Considérons ici $\Gamma(X^i, X^j)$. En tout point $x \in V$ nous avons une application linéaire de V dans $\vartheta$ , définie par la matrice jacobienne

$$M_l^i = D_l X^i(x), \quad \text{et} \quad \Gamma(X^i, X^j) = \sum_{l=1}^{n} M_l^i M_l^j$$

$\Gamma$ peut donc être considérée comme une forme bilinéaire sur $\vartheta' \times \vartheta'$

$\Gamma(u,v) = {}^tMu . {}^tMv$. Cette forme est positive, et non dégénérée si et seulement si ${}^tM$ est injective, c'est-à-dire M est surjective.

On appelle généralement $(\Gamma(X^i, X^j))_{i,j}$ la *matrice de Malliavin,* et son déterminant $\sigma$ le *déterminant de Malliavin.*

La formule (5.6) permet d'écrire, pour toute fonction $\varphi\ C^\infty$ , si $b_{ij}$ dénote le facteur d'indices (i,j) dans la matrice de Malliavin :

$$\varphi\, \sigma\, D_i f(X) = \sum_j \varphi\, b_{ij}\, \Gamma(f(X), X^j)$$

En intégrant sur V et utilisant (5.4), on obtient la formule d'intégration par parties

$$\int_V \varphi(x)\, \sigma(x)\, D_i f(X(x))\, d\lambda(x) = -\int_V f(X(x))\, L_i\, \varphi(x)\, d\lambda(x)$$

où $L_i \varphi = \sum_j (A(\varphi b_{ij} X^j) - X^j A(\varphi b_{ij}) + \varphi b_{ij} AX^j)$

$$= \sum_j (\Gamma(X^j, b_{ij} \varphi) + 2 \varphi b_{ij} AX^j)$$

si X est deux fois différentiable et $L_i \varphi$ continue pour $\varphi \in C^\infty(V)$.

Prenons maintenant $\varphi(x) \equiv 1$. La formule ci-dessus dit alors que l'image de $\sigma \lambda$ par X est toujours absolument continue.

En effet, si $\eta = X_*(\eta)$, pour montrer que $\eta$ est absolument continue, il suffit de montrer que toutes ses dérivées au sens des distributions sont des mesures ([1]). C'est-à-dire :

$\forall \varphi \in \mathscr{D}(\mathbb{R}^d)$, à support compact contenu dans le compact K

(5.7) : $$|<D_i \eta, \varphi>| = |<\eta, D_i\varphi>|$$

$$= |\int_V D_i\varphi o X \, d\lambda| \leqslant C_k \|\varphi\|_\infty .$$

Ceci implique alors que $\eta$ est absolument continue (par rapport à la mesure de Lebesgue), et même que la densité est dans $L^{d/d-1}$.

En outre, $X_*(\sigma \lambda)$ absolument continue implique $X_*(\lambda)$ absolument continue si $\sigma \neq 0$ p.s.

Supposons $\sigma > 0$ partout. On peut alors remplacer $\varphi$ par $\varphi/\sigma$, et on obtient :

$$\int_V \varphi \, D_i f(X) \, d\lambda = -\int_V f(X)\overline{L}_i(\varphi) d\lambda \quad \text{où } \overline{L}_i(\varphi) = L_i(\varphi / \sigma)$$

Si $X \in C^\infty$, on peut itérer :

$$\int_V \varphi \, D^\alpha f(X) \, d\lambda = (-1)^{|\alpha|} \int_V f(X) \, \overline{L}^\alpha \varphi \, d\lambda .$$

Comme $\overline{L}^\alpha \varphi$ est une fonction continue, prenant $\varphi \equiv 1$, on en déduit que $X(\lambda)$ a une densité $C^\infty$ d'après le critère suivant (cf. [ 1 ]) :

$X(\lambda) = \eta$ a une densité $C^\infty$ si pour tout $\alpha$ multi-entier,et $\forall \varphi \in \mathscr{D}(\mathbb{R}^d)$,

$$| < \eta , D^\alpha \varphi > | = | \int_V D^\alpha \varphi \circ X \, d\lambda | \leqslant C_{k,\alpha} \| \varphi \|_\infty$$

La condition $\sigma \neq 0$ apparaissant ici a une interprétation simple : c'est que la différentielle de X soit surjective.

C'est le même type de calculs que nous allons mener, mais en remplaçant la mesure de Lebesgue $\lambda$ par la mesure gaussienne $\mu$ . Ceci afin de pouvoir généraliser en dimension infinie.

Posons $D^\infty = \left\{ f : V \to \mathbb{R} : f \in C^\infty, \ D^k f \in \bigcap_{p>1} L_p(V \to \Lambda_k(V), \mu) \ \forall k \in \mathbb{N} \right\}$

Si $F = (F^1, ..., F^d) : V \to \mathbb{R}^d$ avec $F^i \in D^\infty \ \forall i \in < 1 , d >$,

on pose $\sigma(x) = [\sigma^{ij}(x)]$

où $\sigma^{ij}(x) = < DF^i(x), DF^j(x) >_V$ . C'est la matrice de Malliavin.

Nous ferons les hypothèses suivantes :

A1. Pour $\mu$ -presque tout x, $\sigma(x)$ est strictement définie positive.

Alors, $\sigma^{-1}(x) = [\gamma_{ij}(x)]$ est définie $\mu$-p.s.

En outre, nous dirons que F satisfait l'hypothèse $(A.2)_p$ $1 < p < \infty$

si : F satisfait A1 et si $\det [\sigma^{-1}(x)] \in L_p(V, \mu)$.

On a : $$D\gamma_{ij} = - \sum_{k,l=1}^{d} \gamma_{ik} \gamma_{jl} D_\sigma^{kl} .$$

On en déduit que, si (A2).p est satisfaite pour $p > 2$, alors :

$$D\gamma_{ij} \in L_r(V \to V, \mu) \qquad \forall 1 \leq r < p/2$$

et, plus généralement, que si $(A2)_{.p}$ est satisfaite pour $p > k + 1$,

$$D^k \gamma_{ij} \in L_r(V \to \Lambda_k(V), \mu) \qquad \forall 1 \leqslant r < p/(k+1).$$

Soit $\phi : \mathbb{R}^d \to \mathbb{R}$, $C^\infty$ avec des dérivées de tous les ordres bornés.

Alors, $\phi \circ F \in D^\infty$, et :

$$D(\phi \circ F)(x) = \sum_{k=1}^{d} D_k(\phi) \circ F(x) \,.\, D F^k(x)$$

$$\Rightarrow \quad D_i(\phi) \circ F(x) = \sum_{j=1}^{d} \gamma_{ij}(x) \, < D(\phi \circ F)(x), \; DF^j(x) >_V.$$

Supposons dans ce qui suit que F satisfasse $(A2)_p$ pour un $p > 2$.

Si $g : V \to \mathbb{R}$ satisfait $g \in L_{q_o}(V, \mu)$ et $Dg \in L_{q_1}(V \to V, \mu)$ avec $1/q_o + 2/p < 1$, $1/q_1 + 1/p < 1$, alors, pour $i \in <1, d>$ :

$$\int_V D_i \phi \circ F(x) \, g(x) \, \mu(dx) =$$

$$= \sum_{j=1}^{d} \int_V < D(\phi \circ F)(x), \; \gamma_{ij}(x) g(x) \, DF^j(x) >_V \mu(dx)$$

$$= - \sum_{j=1}^{d} \int_V \phi \circ F(x) \;\; \delta \, [\gamma_{ij}(x) \, g(x) \, DF^j(x)] \;\; \mu(dx)$$

$$= \int_V \phi \circ F(x) \;\; \phi_i(x;g) \;\; \mu(dx)$$

$$\text{où} \quad \Phi_i(x\,;g) = - \sum_{j=1}^{d} \delta \left[\gamma_{ij}(x) \, g(x) \;\; DF^j(x)\right]$$

$$= - \sum_{j=1}^{d} \{ g(x) \; <D\gamma_{ij}(x), DF^j(x)>_V$$

$$+ \gamma_{ij}(x) \; < Dg(x), DF^i(x) >_V + \gamma_{ij}(x) \, g(x) \;\; LF(x).$$

Comme $DF^i \in \bigcap_{p>1} L_p(V \to V, \mu)$ et $LF^j \in \bigcap_{p>1} L_p(V, \mu)$, on en déduit que

$\Phi_i(\cdot\,;g) \in L_r(V, \mu)$ pour tout $r \geqslant 1$ tel que

$$1/r > \max(1/q_o + 2/p,\ 1/q_1 + 1/p).$$

Si l'on prend $g \equiv 1$, on en déduit, si $(A_2)_p$ est satisfaite pour un $p > 2$

$$\int_{R^\alpha} D_i\ \phi(\xi)\ \nu\ (d\ \xi) = \int_V \Phi \text{ o } F(x) \quad \Phi_i(x\ ;1) \quad \mu(dx)$$

où $\nu = F_*(\mu)$ est la mesure image de $\mu$ par $F : V \longrightarrow \mathbb{R}^d$.

Donc : $$\left|\int_{R^\alpha} D_i \Phi\ (\xi)\nu\ (d\ \xi)\right| \leqslant K \max_{\xi \in R^d} |\phi(\xi)|$$

avec $K = \| \Phi_i(.\ ;1) \|_1$.

Ceci montre l'absolue continuité de $\nu\ (d\ \xi)$ par rapport à $d\ \xi$ .

Puis, par un argument d'approximation, on déduit l'absolue continuité de $\nu$ sous la seule hypothèse (A1).

En poursuivant, on obtient le résultat suivant :

**Théorème 5.3. :**

Si $F : V \to \mathbb{R}^d$ vérifie $(A2)_p$ pour un $p > 2(d + k) + 2$, alors la mesure image $\nu$ de F par $\mu$ a une densité $C^k$ par rapport à la mesure de Lebesgue sur $\mathbb{R}^d$.

**6. Généralisation des opérateurs D, $\delta$ , L aux espaces de Wiener abstraits.**

Soit maintenant (W, H, $\mu$) espace de Wiener abstrait.

Nous nous proposons d'établir un calcul différentiel pour les fonctions $\mu$ -mesurables sur W (fonctionnelles de Wiener), et qui sera encore valable pour des fonctionnelles encore plus générales.

**6.1. Quelques notations.**

Si $E_1$ , $E_2$ sont des Hilbert séparables, $\Lambda(E_1, E_2)$ désigne l'espace des opérateurs de Hilbert-Schmidt de $E_1 \to E_2$ , i.e. l'espace des opérateurs linéaires continus $A : E_1 \to E_2$ tels que

$$\|A\|^2_{H.S.} = \sum_i |A(e_i)|^2_{E_2} < \infty \quad . \text{ pour } (e_i) \text{ B.O.N. de } E_1.$$

On le munit du produit scalaire d'Hilbert-Schmidt :

$$< A_1 , A_2 >_{H.S.} = \sum_i <A_1(e_i) , A_2(e_i)>_{E_2}$$

On définit alors $\Lambda_k(E_1, E_2)$ par récurrence :

$$\Lambda_o(E_1, E_2) = E_2 \qquad \Lambda_k(E_1, E_2) = \Lambda(E_1, \Lambda_{k-1}(E_1, E_2))$$

Si $E_2 = \mathbb{R}$, on note $\Lambda_k(E_1)$, qui coïncide avec $\underbrace{E_1 \otimes \dots \otimes E_1}_{k}$

Si E est un Hilbert, $E^o$ désignera un sous-espace vectoriel dense. (L'exemple fondamental sera : $E = H$ , $E^o = W'$).

On introduit alors le sous-espace $\overset{o}{\Lambda}(\overset{o}{E}_1, \overset{o}{E}_2)$ dense dans $\Lambda(E_1, E_2)$ par

$$\overset{o}{\Lambda}(\overset{o}{E}_1, \overset{o}{E}_2) = \left\{ A \in \Lambda(E_1, E_2) \ : \ A = \sum_{i=1}^{n} e_i \otimes f_i \text{ pour un n} \right.$$

$$\left. e_i \in E_1^o \quad f_i \in E_2^o \quad i = 1, \dots, n \right\}$$

Si $E_1 = H$ et $\mathring{E}_1 = W'$, alors $A \in \mathring{\Lambda}(W', \mathring{E}_2) \subset \Lambda(H, E_2)$ s'étend en une application linéaire continue $A : W \to \mathring{E}_2$ par

$$A(x) = \sum_{i=1}^{n} e_i(x).f_i$$

Ensuite, pour $k \in \mathbb{N}$, $\mathring{\Lambda}_k(\mathring{E}_1, \mathring{E}_2)$ est défini par récurrence.

Si $A \in \mathring{\Lambda}_2(\mathring{E}_1, \mathring{E}_2)$ alors A est de la forme

$$A = \sum_{\substack{i=1 \\ j}}^{n} e_i \otimes (e_j \otimes f_{ij}) \qquad \begin{array}{l} e_i \in \mathring{E}_1 \\ f_{ij} \in \mathring{E}_2 \end{array}$$

et la trace de A est un élément de $\mathring{E}_2$ définie par

$$\operatorname{Tr} A = \sum_{i,j,k=1}^{n} \langle e_i, \hat{e}_k \rangle_{E_1} \langle e_j, \hat{e}_n \rangle_{E_1} f_{ij}$$

où $(\hat{e}_k)$ est une B.O.N. de l'espace vectoriel engendré par $\{e_1, ..., e_n\}$ .

On va maintenant introduire les polynômes sur W.

Tous les opérateurs et les calculs seront faits pour ces polynômes, ce qui donne un cadre simple. Ensuite, on prendra la fermeture pour certaines semi-normes de ces espaces de polynômes.

$F : W \to \mathbb{R}$ est un polynôme ssi

$\exists n \quad \exists l_1, ..., l_n \in W \qquad F(w) = p(l_1(w), ..., l_n(w)) . \quad \forall w \in W$

$\exists p(x_1, ..., x_n)$ polynôme réel :

On notera $\mathcal{P}$ l'espace des polynômes réels sur W.

F est un polynôme sur W à valeurs dans $\mathring{E}$ ssi

$$F(w) = \sum_{i=1}^{n} F_i(w)e_i , \; n \in \mathbb{N}^* \quad F_i \in \mathcal{P} \quad (e_i) \text{ B.O.N. de } \mathring{E} .$$

On notera $\mathcal{P}(\mathring{E})$ l'espace de ces polynômes.

**Proposition 6.1.** $\mathcal{P}(E^\circ)$ est un sous-espace dense de $L_p(W \to E, \mu)$ pour tout $1 \leq p < \infty$ .

Notons $\mathcal{P}_n(E)$ l'espace des polynômes à valeurs dans E de d° $\leq$ n.

Soit $C_o = \{cstes\}$

$C_n$ = supplémentaire orthogonal de $C_o \oplus \ldots \oplus C_{n-1}$ dans $\overline{\mathcal{P}_n(E)}^2$

où $\overline{\mathcal{P}_n(E)}^2$ est la fermeture dans $L_2(W \to E, \mu)$.

On a alors : $$L_2(W \to E, \mu) = \bigoplus_{n=1}^{\infty} C_n$$

et ceci est la *décomposition en chaos de Wiener* de $L_2(W \to E, \mu)$.

**6.2. Opérateur dérivation D.**

Pour $F \in \mathcal{P}(E^\circ)$, on définit

$$DF \in \mathcal{P}\ (\Lambda^\circ(W^*, E^\circ)) \subset \mathcal{P}\ (\Lambda(H,E)) \text{ par}$$

$$DF(w)\ [u] = \frac{d}{dt} F(w + tu)\Big|_{t=0} \qquad \forall u \in W^*$$

**Attention ! On voit que l'on se limite ici à des dérivées dans les directions de H. Ceci est très important.**

**Dans le calcul des variations stochastiques, on ne peut considérer comme «déplacements admissibles» que les fonctions de l'espace autoreproduisant H.**

**L'exemple de l'aire de Levy montre que l'on aura couramment des fonctionnelles non continues pour la topologie uniforme sur $W_o^r$ , et donc non différentiables au sens classique sur $W_o^r$ , mais qui seront indéfiniment différentiables au sens introduit ici, c'est-à-dire, en se limitant à des dérivées dans les directions de H.**

On définit ensuite $D^kF$ par récurrence.

$$D^k F \in \mathcal{P}\ (\mathring{\Lambda}_k(W^*, E^\circ)) \subset \mathcal{P}\ (\Lambda_k(H,E)).$$

**6.3. Opérateur divergence : $\delta$**

$$\mathcal{P}(E^\circ) \underset{\delta}{\overset{D}{\rightleftarrows}} (\mathcal{P}(W^*, E^\circ))$$

$$\delta : \mathcal{P}(\mathring{\Lambda}(W^*, E^\circ)) \longrightarrow \mathcal{P}(E^\circ) \text{ est défini par}$$

$$\delta F(w) = \mathrm{Tr}\ DF(w) - F(w)[w]$$

suivant le calcul fait en dimension finie.

**6.4. Opérateur d'Ornstein-Uhlenbeck : L**

$$L : \mathcal{P}(E^\circ) \longrightarrow \mathcal{P}(E^\circ)$$

est défini par $\quad LF(w) = \delta(DF)(w)$

et donc $\quad LF(w) = \mathrm{Tr}\ D^2F(w) - DF(w)[w]$.

Remarquons que, si $F \in L^2(W \longrightarrow E, \mu) \cap \mathcal{P}(E)$, alors F se décompose selon les chaos de Wiener $\quad F = \sum_n F_n$

et on a pour $F_n \in C_n \qquad LF_n = (-n)F_n$

donc $\quad LF = \sum_n (-n) F_n$

On a alors une formule d'intégration par parties, comme dans le cas de la dimension finie :

**6.5. Formule d'intégration par parties :**

Pour tout $F \in \mathcal{P}_{(E)}$ et $G \in \mathcal{P}(\Lambda^\circ(W^*, E))$,

$$\int_W < DF(w), G(w) >_{H.S.} \mu(dw) = -\int_W < F(w), \delta G(w) >_E \mu(dw)$$

d' où l'on déduit :

$$\forall F \in \mathcal{P}_{(E)}, \quad \forall G \in \mathcal{P}_{(E)}$$

$$\int_W < LF(w), G(w) >_E \mu(dw) = \int_W < F(w), LG(w) >_E \mu(dw)$$

$$= -\int_W < DF(w), DG(w) >_{HS} \mu(dw)$$

**7. Espaces de Sobolev de fonctionnelles de Wiener - Espaces de différentiabilité**

On définit sur $\mathcal{P}(E)$ les normes de Sobolev $\| \ \|_{p,s}$ $\quad 1 \leq p < \infty \quad s \in \mathbb{R}$

par $\qquad \| F \|_{p,s} = \| (I - L)^{s/2} F \|_p$

où $(I - L)^{s/2}$ F est défini comme suit :

De la décomposition en chaos de Wiener $\qquad F = \sum_n F_n$

on a $\qquad LF = \sum_n (-n) F_n$

$$(I - L) F = \sum_n (1 + n) F_n$$

et on définit $\quad (I - L)^s F = \sum_n (1 + n)^s \quad F_n.$

L'espace de Sobolev $\mathbb{D}_p^s(E)$ sera le complété de $\mathcal{P}(E)$ pour $\| \ \|_{p,s}$.

Les normes de Sobolev $\| \ \|_{p,s}$ sont compatibles.

On a l'injection canonique $\mathbb{D}_{p'}^{s'}(E) \hookrightarrow \mathbb{D}_p^s(E)$ continue si $p \leq p'$
$\qquad 1 \leq s'$

et $(\mathbb{D}_p^s (E))^* = D_q^{-s}(E)$ $\qquad$ avec $1/p + 1/q = 1 \quad 1 < p < \infty$

sous l'identification canonique de $\mathbb{D}_2(E) = L_2(E, \mu)$ avec son dual.

Le résultat fondamental suivant a été établi par Meyer, et par P. Krée :

**Théorème 7.1.**

Pour tout $1 < p < \infty$ et $k \in \mathbb{N}$, il existe des constantes $0 < c_{p,k} < C_{p,k}$

telles que :

$$c_{p,k} \quad \| F \|_{p,k} \leq \| F \|_p + \| \ \| D^k F \|_{H.S.} \|_p \leq C_{p,k} \quad \| F \|_{p,k}$$

$\forall \ F \in \mathcal{P}(E).$

On déduit de ce théorème ce qui suit :

. L'opérateur $D^k : \mathcal{P}(E) \longrightarrow \mathcal{P}(\mathring{\wedge}_k(w^*, E))$ se prolonge en un opérateur linéaire borné

$$D^k : \mathbb{D}_p^{s+k}(E) \longrightarrow \mathbb{D}_p^s(\Lambda_k(H, E)) \qquad \forall\ 1 < p < \infty \quad \forall\ s \in \mathbb{R}$$

. Par dualité, $\delta : \mathscr{P}(\overset{\circ}{\Lambda}(W^*, E)) \longrightarrow \mathscr{P}(E)$

s'étend en un opérateur linéaire borné

$$\delta : \mathbb{D}_p^{s+1}(\Lambda(H, E)) \longrightarrow \mathbb{D}_p^s(E) \qquad \forall\ 1 < p < \infty \qquad \forall\ s \in \mathbb{R}$$

. L'opérateur $L : \mathscr{P}(E) \longrightarrow \mathscr{P}(E)$

se prolonge en un opérateur linéaire borné

$$L : \quad D_p^{s+2}(E) \leq D_p^s(E) \qquad \forall\ 1 < p < \infty \quad \forall\ s \in \mathbb{R}$$

On a le diagramme : si $0 < s < s'$ $\quad 1 \leq p < p'$

$$\begin{tikzcd}
\mathbb{D}_p^{s'}(E) \arrow[r, hook] & \mathbb{D}_p^{s}(E) \arrow[r, hook] & \mathbb{D}_p^{\circ}(E) = L_p(E) \arrow[r, hook] & \mathbb{D}_p^{-s}(E) \arrow[r, hook] & \mathbb{D}_p^{-s'}(E) \\
\mathbb{D}_{p'}^{s'}(E) \arrow[r, hook] \arrow[u, hook] & \mathbb{D}_{p'}^{s}(E) \arrow[r, hook] \arrow[u, hook] & \mathbb{D}_{p'}^{\bullet}(E) = L_{p'}(E) \arrow[r, hook] \arrow[u, hook] & \mathbb{D}_{p'}^{-s}(E) \arrow[r, hook] \arrow[u, hook] & \mathbb{D}_{p'}^{-s'}(E) \arrow[u, hook]
\end{tikzcd}$$

$\hookrightarrow$ désignant une injection continue à image dense.

Posons alors : $$\mathbb{D}^{\infty}(E) = \bigcap_{\substack{p>1 \\ s \in \mathbb{R}}} \mathbb{D}_p^s(E)$$

$$\mathbb{D}^{-\infty}(E) = \bigcup_{\substack{p>1 \\ s \in \mathbb{R}}} \mathbb{D}_p^s(E)$$

$\mathbb{D}^{\infty}(E)$ est un espace vectoriel dénombrablement normé complet et $\mathbb{D}^{-\infty}(E)$ son dual.

Pour $E = \mathbb{R}$, on note $\mathbb{D}^{\infty}$ et $\mathbb{D}^{-\infty}$ .

Les élements de $\mathbb{D}_p^s(E)$ pour $s < 0$ ne sont en général plus des fonctionnelles de Wiener :

on les appelle fonctionnelles de Wiener généralisées.

On peut aussi voir que :

$D^k : \mathbb{D}^{\infty}(E) \longrightarrow \mathbb{D}^{\infty}(\Lambda_k(H,E))$ (resp. $\mathbb{D}^{-\infty}(E) \longrightarrow \mathbb{D}^{-\infty}(\Lambda_k(H,E))$

$\delta \ : \mathbb{D}^{\infty}(\Lambda(H,E)) \longrightarrow \mathbb{D}^{\infty}(E)$ (resp. $\mathbb{D}^{-\infty}(\Lambda(H,E)) \longrightarrow \mathbb{D}^{-\infty}(E)$)

$L \ : \mathbb{D}^{\infty}(E) \longrightarrow \mathbb{D}^{\infty}(E)$ (resp. $\mathbb{D}^{-\infty}(E) \longrightarrow \mathbb{D}^{-\infty}(E)$)

sont des opérateurs continus.

On peut, par passage à la limite, obtenir des prolongements de formules d'intégration par parties.

**Théorème 7.2.**

(i) Soit $f \in C^{\infty}(\mathbb{R}^n)$ avec des dérivées à croissance polynomiale, et $F = (F^1, ..., F^n) \in D^{\infty}(\mathbb{R}^n)$. Alors, $f \circ F \in D^{\infty}$ et :

$$D(f \circ F) = \sum_{i=1}^{n} ((D_i f) \circ F) DF^i$$

$$L(f \circ F) = \sum_{i,j=1}^{n} ((D_i D_j f) \circ F) < DF^i, DF^j >_H + \sum_{i=1}^{n} ((D_i f) \circ F) \ LF^i$$

(ii) Pour $F, G \in D^{\infty}$ ,

$$< DF, DG >_H = \frac{1}{2} \{L(FG) - (LF)G - F(LG)\}$$

et $< DF, DG >_H \in D^{\infty}$

(iii) Si $F \in D^2_p$ , $G \in D^2_q$ et $J \in D^1_r$ , tels que $1/p + 1/q + 1/r \leq 1$, alors :

$$< D < DF, DG >_H , DJ > = < D^2F, DG \otimes DJ >_{HS} + < D^2G, DF \otimes DJ >_{HS}$$

(iv) Si $F \in D^1_p$ , $G \in D^1_q$ et $J \in D^1_q(H)$, avec $1/p + 1/q \leq 1$, alors :

$$D(FG) = FDG + G\,DF$$

$$\text{et} \quad \delta(FJ) = \langle DF, J \rangle_H + F\delta J$$

Si, en outre, $G \in D^2_q$, alors

$$\delta(FDG) = \langle DF, DG \rangle_H + FLG.$$

Les formules d'intégration par parties restent valables :

$$\int_W \langle DF(w), G(w) \rangle_H \ \mu(dw) = -\int_W F(w) \ \delta G(w) \ \mu(dw)$$

pour $F \in D^1_p$ et $G \in D^1_q$ tels que $1/p + 1/q \leq 1$, et :

$$\int_W (LF(w))G(w) \ \mu(dw) = \int_W F(w)(LG(w) \ \mu(dw) = -\int_W \langle DF(w), DG(w) \rangle_H \ \mu(dw)$$

pour $F \in D^2_p$ et $G \in D^2_q$ tels que $1/p + 1/q \leq 1$.

Nous disposons alors des outils pour prolonger en dimension infinie, sur les espaces de Wiener abstraits, les calculs du paragraphe 5.

Ceci permettra d'obtenir des théorèmes de régularité de l'image de la mesure $\mu$ par certaines fonctionnelles de Wiener, ou même par des fonctionnelles plus générales (fonctionnelles de Wiener généralisées) parmi lesquelles les distributions de Dirac $\delta_x$.

Mais, pour une présentation, nous en resterons là, et renvoyons le lecteur intéressé aux références suivantes.

## REFERENCES

[ 1 ] P. Malliavin. Stochastic calculus of Variations and hypoelliptic operators. Proc. of Intern. Symp. SDE. Kyoto - 1976 pp. 195-263.

[ 2] N. Ikeda - S. Watanabe. An introduction to Malliavin's Calculus. Taniguchi Symp. SA. Kata. 1983. pp. 1-52.

[ 3 ] P. Krée Cours de D.E.A. 1984-85. Paris VI.

[ 4 ] P.A. Meyer. The Malliavin Calculus and some pedagogy (a paraître)

[ 5 ] S. Watanabe. Stochastic Differential Equations and Malliavin Calculus. Tata Institute of Fondamental Research - Springer Verlag - 1984.

Reçu le 28 février 1986.

Université de Clermont II, U.E.R. Sciences, Mathématiques Pures, 63170 - Aubière, France.

Unité Associée au C.N.R.S. n° 739 : Méthodes Probabilistes et Analyse Numérique.

**Adresse personnelle : 3, rue Sounely, Pérignat les Sarliève, 63170 Aubière, France.**