PAUL-ANDRÉ MEYER

Variation des solutions d'une E.D.S., d'après J. M. Bismut

Séminaire de probabilités (Strasbourg), tome S16 (1982), p. 151-164
<http://www.numdam.org/item?id=SPS_1982__S16__151_0>

Université de Strasbourg
Séminaire de Probabilités 1980/81

VARIATION DES SOLUTIONS D'UNE E.D.S.

par P.A. Meyer, d'après J.M. Bismut

Nous présentons ici certains résultats de l'article de Bismut << Martingales, the Malliavin calculus and hypoellipticity under general Hörmander conditions >>, paru dans ZW, 56, 1981. L'idée du calcul des variations sur les solutions d'une e.d.s. est due à Malliavin, et consiste à montrer que la solution dépend différentiablement, en un certain sens, de la semimartingale directrice, et à calculer sa différentielle. Cela se retrouve aussi chez Bismut, mais les deux méthodes sont très différentes. Malliavin considère la semimartingale directrice comme la valeur à l'instant t d'un processus Y_t (un processus d'Ornstein-Uhlenbeck à valeurs dans l'espace des applications continues sur $\mathbb{R}_+$), et étudie la dépendance en t de la solution. Bismut n'utilise vraiment que la différentiabilité de la solution suivant des droites de l'espace des semimartingales, évitant ainsi tout véritable recours à la dimension infinie.

I. GENERALITES

Notre but est l'étude de la << variation des solutions >> d'une équation différentielle stochastique du type suivant

$$(1)\qquad Y_t(x) = x + \int_0^t a(Y_s(x))dZ_s$$

ou encore de sa variante au sens de Stratonovitch

$$(1')\qquad Y_t(x) = x + \int_0^t a(Y_s(x))*dZ_s \quad ,$$

lorsqu'on << varie >> la semimartingale directrice Z. L'équation de Stratonovitch est en fait la plus importante du point de vue géométrique (les résultats qui la concernent ont un numéro de la forme n').

Expliquons d'abord les notations. La semimartingale solution Y_t , la valeur initiale x, sont des éléments de $\mathbb{R}^d$, et leurs composantes sont notées Y^i_t , x^i (i=1,...,d). La semimartingale directrice Z - toujours supposée continue - est à valeurs dans $\mathbb{R}^{m+1}$, et ses composantes sont notées Z^λ_t (λ=0,1,...,m). Dans le cas particulier le plus important, appelé << cas brownien >>, on a $Z^0_t=t$, tandis que $(Z^\lambda_t)_{\lambda=1,\dots,m}$ est un mouvement brownien standard dans $\mathbb{R}^m$. Enfin, a(x) est une matrice $(a^i_\lambda(x))$ dont les éléments sont des fonctions de classe C^∞ , bornées ainsi que leurs dérivées partielles de tous les ordres.

On sait que l'équation (1') est alors un cas particulier de l'équation

(1), si l'on convient d'ajouter les crochets de Z parmi les semimartingales directrices. Rappelons aussi (cf. l'exposé sur les e.d.s. dans ce volume) que la forme (1') a un sens intrinsèque dans les variétés, en interprétant les (a^i_λ) comme des champs de vecteurs $a_\lambda = a^i_\lambda D_i$. Pour avoir une interprétation analogue de la forme (1), il est nécessaire de disposer d'une connexion Γ sur la variété.

Si nécessaire, on écrira $Y^Z_t(x)$ pour indiquer la dépendance de Y vis à vis de la semimartingale directrice Z.

LE FLOT DE L'E.D.S.. D'après les travaux de Malliavin, Baxendale, Elworthy, Eels... /Bismut, Kunita, Varadhan...[1] présentés dans le Sém. XV, on peut choisir une version $Y_t(\omega,x)$ de la solution de (1) ou (1'), possédant les propriétés suivantes, pour (presque) tout ω

i. $Y_t(\omega,x)$ est fonction continue de (t,x), de classe C^∞ en x.

ii. Pour tout multiindice α , la dérivée partielle $D^\alpha Y_t(\omega,x)$ par rapport à x est fonction continue de (t,x).

iii. L'application $x \mapsto Y_\cdot(\omega,x)$ est propre : pour tout intervalle compact $[0,T]$, on a $\lim_{x\to\infty} \inf_{t\in[0,T]} |Y_t(\omega,x)| = +\infty$

Il a été démontré aussi que, pour presque tout ω, $x \mapsto Y_t(\omega,x)$ est pour tout t un difféomorphisme de $\mathbb{R}^d$ sur lui même, mais nous n'utiliserons guère ce résultat.

Soit $U_t(\omega,x)$ la matrice de l'application linéaire tangente à $Y_t(\omega,x)$ au point x, c'est à dire la matrice carrée des dérivées partielles $D_j Y^i_t(\omega,x)$. C'est une fonction continue de (t,x), solution de l'équation linéaire

(2) $U_t(x) = I + \int_0^t dL_s(x)U_s(x)$ (2') $U_t(x) = I + \int_0^t * dL_s(x)U_s(x)$[2]

où (L_t) est la semimartingale matricielle

$$dL^i_{jt} = D_j a^i_\lambda(Y_s(x))dZ^\lambda_s \quad (\text{ resp. } *dZ^\lambda_s)$$

La convention de sommation est appliquée en λ comme d'habitude. La matrice $U_t(\omega,x)$ admet pour tout t une matrice inverse $V_t(\omega,x)$, dépendant continûment de (t,x), et donnée par l'équation suivante dans le cas d'Ito

$$(3) \qquad V_t(x) = I - \int_0^t V_s(x)(dL_s(x) - \langle dL_s(x), dL_s(x) \rangle)$$

(voir dans ce volume l'article de Karandikar ; cette formule figure aussi chez Bismut sous une forme moins générale). La formule analogue dans le cas de Stratonovitch est plus simple :

$$(3') \qquad V_t(x) = I - \int_0^t V_s(x) * dL_s(x) \ .$$

1. Kunita a signalé l'article de Blagoveščenskii-Freidlin (DAN 138, 1961), qui semble ne pas avoir attiré l'attention, et qui est remarquable.
2. $*d$ forme un tout (c'est le « d de Stratonovitch »).

LE CAS BROWNIEN. Comme nous l'avons dit au début, le << cas brownien >> est celui où $Z_t^0=t$, les Z_t^λ ($\lambda\geq 1$) étant les composantes d'un mouvement brownien standard.

Rappelons que (cf. Emery, ZW 41, 1978, p. 248, lemme 3), la solution de l'équation (1) avec valeur initiale Y_0 satisfait à une inégalité de la forme

(4) $$\|Y\|_{\underline{S}^p} \; (= \|Y^*\|_{L^p}) \leq C_p \|Y_0\|_{L^p}$$

sous les conditions suivantes : p est fixé dans $[1,\infty[$; la semimartingale directrice Z appartient à l'espace $\underline{\underline{H}}^\infty$ (même réf. p. 243), avec une norme suffisamment petite (i.e. majorée par une quantité qui dépend de p et de la constante de Lipschitz de la matrice a , ainsi que de la dimension de l'espace). Disons maintenant qu'une semimartingale Z est << de type brownien >> si l'application $u \mapsto Z^u$ (arrêt à u) est une application continue de $\mathbb{R}_+$ dans $\underline{\underline{H}}^\infty$, ce qui signifie encore que tout intervalle borné $[0,N]$ peut être partagé en un nombre fini d'intervalles $[t_i,t_{i+1}]$ où l'on peut appliquer l'inégalité (4). On a alors pour <u>tout</u> p fini

(5) $$\|Y_N^*\|_{L^p} \leq c\|Y_0\|_{L^p} \quad (c|x| \text{ dans le cas de (1)})^{(1)}$$

(ici c dépend dep. de N, et du module de continuité de l'application $u \mapsto Z^u$; le point important est l'appartenance à <u>tout</u> L^p). Les a_λ^i et leurs dérivées partielles étant bornés, les semimartingales directrices des équations (2) et (3) sont elles aussi << de type brownien >>, de sorte que l'on a par exemple

(6) $$\| \sup_{0\leq t\leq N} \|U_t(\cdot,x)\| \, \|_{L^p} \leq c|x|$$

de même pour V_t , ou pour $\det(V_t)$ qui est un polynôme en les coefficients de V_t. Tout cela a une grande importance pour les résultats de Malliavin.

II. VARIATION DES SEMIMARTINGALES DIRECTRICES

Dans ce paragraphe, nous allons étudier la << variation de la solution Y >> par la méthode élémentaire (bien connue en calcul des variations classique) consistant à dériver <u>le long d'une droite</u>. Nous cherchons donc à prouver la différentiabilité de $Y^{Z+h\zeta}$ ($h\in\mathbb{R}$), où ζ est une semimartingale, et en particulier à déterminer la différentielle

(7) $$\Phi(\zeta) = \frac{d}{dh} Y^{Z+h\zeta} \Big|_{h=0}$$

Dans une digression (intéressante, mais non indispensable pour les résultats ultérieurs) nous nous affranchirons partiellement de la dérivation le long des droites.

1. Cela vaut aussi pour les équations de Stratonovitch.

THEOREME 1. Il existe une version de $Y_t^{Z+h\zeta}(\omega,x)$ qui, pour presque tout ω, est de classe C^∞ en (h,x), avec des dérivées partielles de tous ordres continues en (t,h,x). De plus, la "différentielle" $\Phi(\zeta)=(\Phi_t^i(\omega))$ de $Y^\bullet$ au point Z est l'application linéaire qui à ζ associe la solution de l'e.d.s.

(8) $$\Phi_t \overset{(1)}{=} \int_0^t a(Y_s^Z(x))d\zeta_s + \int_0^t D_i a(Y_s^Z(x))\Phi_s^i\, dZ_s$$

DEMONSTRATION. Ce théorème est tout à fait simple : nous allons appliquer le grand théorème sur les flots d'e.d.s. à un système convenable. Nous doublons le nombre des semimartingales directrices, en associant à tout λ un indice λ' et en posant $Z^{\lambda'}=\zeta^\lambda$. Nous ajoutons à l'espace d'états une coordonnée supplémentaire x°. Puis nous considérons le système

$$\overline{Y}_t^i(x^\circ,x) = (x^\circ,x)^i + \sum\nolimits_{\mu=\lambda,\lambda'}\int_0^t a_\mu^i(\overline{Y}_s)dZ_s^\mu$$

avec pour $i=1,\dots,d$ $a_\lambda^i(x^\circ,x)=a_\lambda^i(x)$, $a_{\lambda'}^i(x^\circ,x)=f(x^\circ)a_\lambda^i(x)$

pour $i=0$, $a_\mu^i(x^\circ,x)=0$ (de sorte que $\overline{Y}_t^0(x^\circ,x)$ reste constant et égal à x°).

Pour x° fixé, $x^\circ=h$, la solution est simplement $\overline{Y}_t(h,x)=(h,Y_t^{Z+f(h)\zeta}(x))$. La fonction $f(h)$ ne sert pas à grand chose : on la prend C^∞ à support compact afin que le système ci-dessus soit globalement lipschitzien, avec $f(h)=h$ sur un intervalle entourant 0.

Le grand théorème sur les flots nous dit que $\overline{Y}$ admet une version C^∞ en (x°,x), avec une dérivée par rapport à x° donnée par (2). C'est exactement le résultat cherché.

Plus généralement, ce raisonnement se prête au calcul de la différentielle sur un espace de dimension finie : on remplace Z par $Z+h_1\zeta^1+\dots h_k\zeta^k$ et on a la différentiabilité par rapport à $(x,h_1,\dots h_k)$. C'est intéressant en principe pour les calculs de variation seconde.

L'équation (8) présente un grand intérêt. Elle a été étudiée de manière systématique par Yoeurp et Yor dans un important travail non publié, et développé à nouveau par Yoeurp (à paraître). Nous allons reprendre ici la méthode de résolution de l'équation (8) par Yoeurp et Yor (l'extension au cas matriciel est due à A. Uppman).

THEOREME 2. Soit (L_t) une matrice carrée d×d à coefficients semimartingales, continues et soit U_t la solution de l'équation linéaire

(9) $$U_t = I + \int_0^t dL_s U_s \quad (\text{ resp. } (9') : \ I + \int_0^t * dL_s U_s \ )^{(2)}.$$

1. Dans le cas de Stratonovitch : $*d\zeta$, $*dZ$. 2. Cf. (2)-(2') .

Soit (H_t) une semimartingale continue, nulle en 0 à valeurs dans $\mathbb{R}^d$, et soit (Φ_t) la solution de l'équation

$$(10) \qquad \Phi_t = H_t + \int_0^t dL_s\Phi_s \quad (\text{ resp. } \quad (10)' : *dL_s)$$

Alors on a[1]

$$(11) \qquad \Phi_t = U_t\int_0^t U_s^{-1}(dH_s - \langle dL,dH\rangle_s) = H_t - U_t\int_0^t H_s dU_s^{-1}$$

$$(11') \qquad \Phi_t = U_t\int_0^t U_s^{-1}*dH_s = H_t - U_t\int_0^t H_s*dU_s^{-1}$$

REMARQUES. Si H n'est pas nulle en 0, soit Φ' la solution correspondant à $H'=H-H_0$; alors $\Phi_t=H_0+\Phi'_t$.
Dans la formule (11), noter que le crochet est nul si H est à variation finie.

DEMONSTRATION. Nous cherchons Φ_t sous la forme U_tM_t, où (M_t) est une matrice colonne semimartingale. Nous écrivons la formule d'intégration par parties

$$d(UM) = (dU)M + UdM + \langle dU,dM \rangle$$

où $\langle dU,dM \rangle$ est la matrice colonne de coefficients $\langle dU,dM\rangle^i = \langle dU^i_k,dM^k\rangle$ (convention de sommation !). Comme $dU=(dL)U$

$$d\Phi = \underline{\underline{(dU)M}}+UdM+\langle dU,dM\rangle = dH + \underline{\underline{(dL)UM}}$$

Les termes soulignés s'éliminent, et nous avons

$$dM = U^{-1}(dH - \langle dU,dM\rangle)$$

Mais $dM = U^{-1}dH + d(\text{ var.finie })$ entraîne $\langle dU,dM\rangle = U^{-1}\langle dU,U^{-1}dH\rangle = U^{-1}\langle (dU)U^{-1},dH \rangle$. Or $UU^{-1}=I$, donc $(dU)U^{-1}= -UdU^{-1}+ d(\text{var.finie })$, et $U^{-1}\langle (dU)U^{-1},dH \rangle =-U^{-1}\langle U(dU^{-1}),dH \rangle = \langle dU^{-1},dH \rangle$. La formule obtenue mérite d'être explicitée :

$$(12) \qquad \Phi_t = U_t \int_0^t U_s^{-1}dH_s + \langle dU^{-1},dH \rangle_s$$

d'où l'on tire (11) en exprimant le dernier crochet au moyen de (3).
Dans (12), intégrons par parties :

$$U^{-1}dH + \langle dU^{-1},dH \rangle_s = d(U^{-1}H)-HdU^{-1}$$

et nous obtenons la seconde moitié de (11). Quant à (11'), elle n'est pas difficile, puisque le calcul de Stratonovitch obéit aux mêmes règles que le calcul ordinaire.

REMARQUE 1. Il est naturel que dH disparaisse de (11) : en effet, il est bien connu que l'équation (10) peut être résolue pour H seulement <u>continu adapté</u>. La seconde formule (11) , établie dans le cas des semimartingales continues, s'étendra par stabilité au cas des processus continus adaptés, mais une démonstration par un calcul direct mériterait d'être cherchée .

1. Il est souvent plus commode d'écrire $\langle dA,dB\rangle$ que $d\langle A,B\rangle$, et nous le faisons librement dans ces formules, et dans toute la suite.

REMARQUE 2. Oublions la structure linéaire de $\mathbb{R}^d$ pour ne garder que la structure différentiable. Il faut alors écrire (1') sous la forme

$$d^2Y_t(x) = a_\lambda(Y_s(x))*dZ_s^\lambda \qquad Y_0(x)=x$$

où les a_λ sont des champs de vecteurs sur $V=\mathbb{R}^d$ (cf. l'appendice de l'exposé de géométrie stochastique. On pourrait de même interpréter (1) en Ito, au moyen d'une connexion sur V). Dans ce cas, la variation de la solution

$$\zeta \mapsto \Phi\zeta = \frac{d}{dh}Y^{Z+h\zeta}\Big|_{h=0}$$

s'interprète comme une application linéaire de l'espace des semimartingales réelles ζ dans l'espace des semimartingales (Y_t^Z,Φ_t) à valeurs dans T(V) au dessus de Y_t^Z . Remarquons alors le caractère intrinsèque de la formule (11') : $*dH_t = a_\lambda(Y_t^Z)*d\zeta_t^\lambda$ "est" dans $T_{Y_t}(V)$; on le ramène dans $T_{Y_0}(V)$ par U_s^{-1} , et on intègre dans cet espace <u>fixe</u>, après quoi on revient dans $T_{Y_t}(V)$ par U_t. Les " " de la phrase précédente, dont la signification est plus que vague, sont inutiles lorsque ζ <u>est à variation finie</u>, le seul cas utilisé chez Bismut. En effet, on peut alors prendre des densités par rapport à un processus à variation finie scalaire, supprimer les * devant dH, et travailler avec de vrais vecteurs.

DIGRESSION. Nous allons maintenant exposer la méthode de Bismut, qui utilise le flot $Y_t^Z(x,\omega)$ correspondant à la semimartingale directrice Z comme une <<machine>> permettant de ramener l'intégration de (1) avec semimartingale directrice $Z+\zeta$ à celle d'équations différentielles auxiliaires, qui lorsque ζ est à variation finie sont <u>déterministes</u>, et non stochastiques.

L'idée, analogue à celle de la << méthode de variation des constantes >> pour la résolution des équations linéaires, consiste à rechercher la solution sous la forme $Y_t^Z(A_t(\omega),\omega)$, où (A_t) est une semimartingale convenable. Il faut montrer, bien sûr, que le résultat est une semimartingale dont on sait calculer explicitement la différentielle. C'est là l'objet des << formules d'Ito généralisées >> dues à Bismut, et retrouvées par Kunita (Sém. XV, p. 118-120) avec une démonstration plus simple [1]. Nous ne parlerons ici que du cas à variation finie, et nous admettrons le lemme suivant :

LEMME. Si (Y_t) est la solution de (1), et (A_t) est un processus à valeurs dans $\mathbb{R}^d$, continu, adapté et à variation finie, le processus $Y_t(A_t(.),.)$ est une semimartingale, et l'on a

1. D'importants progrès sur ces << formules d'Ito >> ont été réalisés récemment par Sznitman (article à paraître).

(13) $\qquad Y_t(A_t(.),.) = Y_0(A_0,.) + \int_0^t a(Y_s(A_s,.))dZ_s + \int_0^t U_s(A_s,.)dA_s$

Si (1) est remplacée par (1'), on obtient la formule (13') analogue en remplaçant dZ_s par $*dZ_s$ - et cette formule reste d'ailleurs vraie pour A non nécessairement à variation finie, en écrivant $*dA_s$ au lieu de dA_s.

THEOREME 3. 1) Supposons ζ à variation finie, et soit $Y_t^Z(x,.)$ la solution de l'équation (1). Soit $A_t(x,\omega)$ - on omettra la mention de x pour simplifier - la solution de l'équation différentielle

(14) $\qquad A_t(\omega)= x + \int_0^t U_s^{-1}(A_s(\omega),\omega)a(Y_s(A_s(\omega),\omega)d\zeta_s(\omega)$

Alors on peut prendre

(15) $\qquad Y_t^{Z+\zeta}(x,\omega) = Y_t^Z(A_t(x,\omega),\omega)$

2) Le même résultat vaut pour l'équation (1'), mais sans supposer ζ à variation finie : il suffit de remplacer $d\zeta_s$ par $*d\zeta_s$.

Ce théorème met en évidence les faits énoncés au début de la section. Si ζ est à variation finie, l'équation auxiliaire (14) est une équation différentielle non stochastique, qui pour ω fixé se résout de manière déterministe. On a donc un procédé mécanique, connaissant le flot $Y_t^Z(x,\omega)$, pour obtenir en même temps toutes les solutions $Y^{Z+\zeta}$ avec ζ à variation finie, trajectoire par trajectoire. Seule difficulté : l'équation (14) est seulement localement lipschitzienne, et peut donc exploser en temps fini. Bismut mentionne ce problème, et se borne à la remarque suivante : si la matrice a(x) est à support compact K, on a $Y_t^Z(\omega,x)=x$ pour $x\notin K$, donc $U_t(x,\omega)=U_t^{-1}(x,\omega)=I$, et on se trouve dans le cas lipschitzien, ce qui exclut les explosions.

La démonstration du théorème 3 se réduit à un calcul, et (comme nous n'en avons pas besoin) nous ne la donnons pas.

III. APPLICATION AU CAS BROWNIEN

Nous nous plaçons maintenant dans le cas brownien : $Z_t^0=t$, (Z_t^λ) est un mouvement brownien standard ($\lambda=1,\dots,n$). Nous désignons par ζ un processus à variation finie à valeurs dans $\mathbb{R}^{n+1}$ de la forme

$$\zeta_t^0 = 0 \quad , \quad \zeta_t^\lambda = \int_0^t u_s^\lambda ds \quad (u^\lambda \text{ prévisible borné ; } \lambda\geq 1).$$

et nous considérons la trajectoire $Y_{\cdot}^{Z+h\zeta}(x,\omega)$ comme un élément de $\underline{\underline{C}}(\mathbb{R}_+,\mathbb{R}^d)$. Il est bien connu que, si $\tau<\infty$ est fixé, le processus $Z+h\zeta$ est un mouvement brownien sur $[0,\tau]$, par rapport à la loi $M_\tau^h P$, où

$$M_\tau^h = \exp(-h\int_0^\tau u_s\cdot dZ_s - \tfrac{1}{2}h^2\int_0^\tau |u_s|^2 ds)$$

Donc $Y_{\cdot}^{Z+h\zeta}$ a même loi sous $M_\tau^h P$, sur l'intervalle $[0,\tau]$, que $Y_{\cdot}^Z$ sous P. Autrement dit, LEMME 4. Si φ est une fonction borélienne positive sur $\underline{C}(\mathbb{R}_+,\mathbb{R}^d)$, ne dépendant que des coordonnées d'indice $\leq\tau$, on a pour tout h

$$(16)\quad E[\varphi(Y^Z_\bullet)] = E[\varphi(Y^{Z+h\zeta}_\bullet)\exp(-h\int_0^\tau u_s\cdot dZ_s - \frac{h^2}{2}\int_0^\tau |u_s|^2 ds\ )]\ ^{(1)}$$

La quantité suivante est donc nulle pour tout h

$$E[\ \frac{1}{h}(\varphi(Y^{Z+h\zeta}_\bullet)-\varphi(Y^Z_\bullet))\exp(-h\int_0^\tau\ldots)]+E[\varphi(Y^Z_\bullet)\ \frac{\exp(-h\ldots)-1}{h}\ ]$$

Si nous supposons $\varphi\in L^p$ ($p>1$), il n'y a pas de difficulté à étudier le dernier terme : il tend vers $-E[\varphi(Y^Z_\bullet)\int_0^\tau u_s\cdot dZ_s]$. Si nous introduisons la martingale

$$m^\varphi_t = E[\varphi(Y^Z_\bullet)|\underline{F}_t]\quad (t\leq\tau)$$

qui est de la forme $E[\varphi(Y^Z_\bullet)] + \int_0^\tau j_{\lambda s}dZ^\lambda_s$ ($j_{0s}=0$), cette espérance vaut

$$(17)\qquad -E[\varphi(Y^Z_\bullet)\int_0^\tau u_s\cdot dZ_s] = -\ E[\ \int_0^\tau u^\lambda_s j_{\lambda s}\ ds\ ]$$

Du côté gauche, il est un peu difficile de faire une hypothèse tout à fait satisfaisante. Supposons que φ soit <u>fortement différentiable</u>. Comme une forme linéaire continue sur $\underline{C}([0,\tau],\mathbb{R}^d)$ s'écrit $w\mapsto \int_0^\tau \mu_i(dt)w^i_t$ (que nous écrirons $\int_0^\tau \mu(dt)w(t)$, μ étant une << ligne >> et w une << colonne >> . Alors la différentiabilité de φ s'écrit

$$\varphi(\pi+w)-\varphi(\pi) = \int_0^\tau \mu(\pi,dt)w(t)+ o(\|w\|)$$

et d'après (8) et (11), sous réserve de convergence dominée, on obtiendra pour le côté gauche

$$(18)\qquad E[\ \int_0^\tau \mu(Y^Z_\bullet,dt)\int_0^t U_tU_s^{-1}a(Y^Z_s)u_s ds\ ]$$
$$= E[\int_0^\tau(\ \int_s^\tau\mu\ (Y^Z_\bullet,dt)U_{t-s}a(Y^Z_s))u_s ds]$$

En écrivant que (17)+(18)=0, on obtient la valeur du processus j_s, considéré comme matrice ligne : c'est [1]

$$(19)\qquad j_{\lambda s} = E[\int_s^\tau \mu(Y^Z_\bullet,dt)U_{t-s}a_\lambda(Y^Z_s)\ |\ \underline{F}_s]$$ (il s'agit en réalité d'une proj. prévisible)

Cette belle formule est due à Haussmann. La première démonstration de Haussmann (SIAM J. Control and Opt. 16, 252-269, 1978) était fort compliquée. Haussmann en a publié dans Stochastics 3, 1979, 17-27 une démonstration plus simple, assez voisine de celle de Bismut que nous venons de reproduire (avec une simplification supplémentaire dans l'emploi du théorème de Girsanov, due à Williams).

Ici encore, soulignons le sens géométrique de cette formule. Si π est une trajectoire à valeurs dans $\mathbb{R}^d$=V, dans le calcul de la différentielle apparaît la fonction w qui doit être interprétée comme un transport

1. Ici et dans toute la page, le paramètre x est fixé, et on l'omet de la notation.

dans T(V) au dessus de π . Ainsi la matrice ligne $\mu(\pi,dt)$ " est " à valeurs dans $T^*_{\pi(t)}(V)$ - les " " peuvent être interprétés en un sens rigoureux, lorsque les $\mu(\pi,.)$ ont des densités par rapport à une mesure fixe sur $\mathbb{R}_+$. $a(Y^Z_s)$ est une application de $\mathbb{R}^{n+1}$ dans $T_{Y^Z_s}(V)$, $U_{t-s}a(Y^Z_s)$ nous amène dans $T_{Y^Z_t}(V)$, appliquer $\mu(Y^Z_\bullet,dt)$ nous laisse une forme linéaire sur $\mathbb{R}^{n+1}$, et finalement l'intégration a lieu dans le dual de $\mathbb{R}^{n+1}$, espace vectoriel fixe.

Donnons un énoncé explicite des résultats obtenus, sous une hypothèse simple (mais utilisant explicitement la norme euclidienne sur $\mathbb{R}^d$).

THEOREME 5. Supposons que φ soit une fonction fortement différentiable sur $\underline{\underline{C}}([0,\tau],\mathbb{R}^d)$, avec différentielle

$$\varphi(\pi+w) = \varphi(\pi) + \int_0^\tau \mu_i(\pi,dt)w^i(t) + o(\|w\|)$$

les masses totales des mesures μ_i étant uniformément bornées en π (ou plus généralement, à croissance au plus polynômiale en $\|\pi\|$). Alors on a, pour tout processus prévisible borné (u_s) à valeurs dans $\mathbb{R}^{n+1}$, tel que $u^0_s=0$

$$\text{(20)} \qquad E[\varphi(Y_\bullet)\int_0^\tau u_s\cdot dZ_s] = E[\int_0^\tau \mu(Y^\bullet,dt)U_t\int_0^t \bar{U}^{-1}_s a(Y_s)u_s ds\ ]^{(1)}$$

(x omis de la notation, ainsi que le Z de Y^Z_s).

DEMONSTRATION. Revenons à la formule suivant immédiatement (16), l'étude du terme de gauche. Il s'agit de montrer que l'expression sous le signe $\mathbb{E}$ est bornée dans un L^p ,p>1, lorsque h est petit. Pour ce qui est de

$$\exp(-h\int_0^\tau u_s.dZ_s - \frac{h^2}{2}\int|u_s|^2ds),$$

ce facteur a une norme bornée dans tout L^p (p fixé) pour h assez petit (inégalité de John-Nirenberg). Pour le premier facteur, nous écrivons

$$\frac{1}{h}(\varphi(Y^{Z+h\zeta})-\varphi(Y^Z)) = \frac{1}{h}\int_0^h ds \int_0^\tau \mu(Y^{Z+s\zeta},dt)\zeta(t)$$

et comme la norme de $\zeta(t)$ dans $\mathbb{R}^{n+1}$ est bornée par une constante, tout revient à majorer la norme de la mesure $\mu(Y^{Z+s\zeta})$ pour s petit. Mais l'hypothèse de croissance polynômiale nous ramène à montrer que $\|Y^{Z+s\zeta}\|_{\underline{C}([0,\tau])}$ appartient à **tout** L^p, avec une norme majorée en s . Nous avons signalé cette propriété à la fin de la section I.

L'<< INTEGRATION PAR PARTIES >> . Nous allons établir une dernière formule, un peu plus compliquée, avant de quitter Bismut au moment où il s'engage dans la technique du " calcul de Malliavin"

1. μ est une ligne (1,d), U un carré (d,d), a un rectangle (d,λ), u une colonne (λ,1), et le produit du tout un scalaire.

Nous allons prendre dans la formule (20)

$$\varphi(Y_{\bullet}) = g_k(Y_{\bullet})f(Y_\tau) \qquad k=1,\dots,d$$

où f est une fonction C^∞ à support compact sur $\mathbb{R}^d$, et g_k est une fonctionnelle sur $\underline{C}([0,\tau],\mathbb{R}^d)$, fortement différentiable - comme φ précédemment, avec les mêmes hypothèses. La raison pour laquelle nous écrivons g_k , c'est que nous pensons à g_k comme à la k-ième composante d'une fonctionnelle à valeurs dans l'espace cotangent initial T^*_x, notée g. Nous noterons $\nu_k(\pi,dt)$ la mesure associée à la différentielle de g_k , et si nous nous rappelons que $\nu_k(\pi,dt)$ est un élément de $T^*_{Y_t}$, nous avons en fait un tableau de mesures scalaires $\nu_{k\ell}(\pi,dt)$ avec deux indices en bas (une forme bilinéaire sur $T_x \times T_{Y_t}$).

Nous allons prendre d'autre part

$$u_{\lambda s} = u^k_{\lambda s} = (U_s^{-1}a_\lambda(Y_s))^k$$

Ce processus prévisible n'est pas borné, mais comme nous avons des bornes dans tous les L^p sur les matrices U_s^{-1}, l'extension ne pase pas de problème. Avec ces notations, la sommation sur k se fait toute seule, et nous obtenons du côté gauche

$$(21) \qquad E[f(Y_\tau) < g(Y_{\bullet}),\ \int_0^\tau U_s^{-1}a_\lambda(Y_s) * dZ_s^\lambda >^{(1)}] \qquad (\text{ sommation en } \lambda \geq 1 \text{ })$$

Du côté droit, il faut calculer la différentielle de $\varphi(Y_{\bullet})$, qui est un produit. Il apparaît donc deux termes. Le premier

$$(22) \qquad E[f(Y_\tau)\int_0^\tau \nu_{k\ell}(Y_{\bullet},dt)U^\ell_{jt}\int_0^t \sum_1^n (U_s^{-1}a_\lambda(Y_s))^j(U_s^{-1}a_\lambda(Y_s)^k)ds\ ]$$

et le second

$$(23) \qquad E[g_k(Y_{\bullet})D_\ell f(Y_\tau)U^\ell_{j\tau}\int_0^\tau \sum_1^n (U_s^{-1}a_\lambda(Y_s))^j(U_s^{-1}a_\lambda(Y_s))^k ds\ ]$$

On voit apparaître la forme bilinéaire positive sur $T^*_x \times T^*_x$, qui joue en fait un rôle <u>fondamental</u> dans l'étude de l'e.d.s.

$$(24) \qquad \gamma_t^{jk} = \int_0^t \Sigma(U_s^{-1}a_\lambda(Y_s))^j(U_s^{-1}a_\lambda(Y_s))^k ds \quad (\text{ somme sur } \lambda \geq 1 \text{ })$$

ou plus intrinsèquement $\gamma_t = \int_0^t \Sigma_\lambda\, U_s^{-1}a_\lambda(Y_s) \otimes U_s^{-1}a_\lambda(Y_s)\, ds$. Nous pouvons alors écrire

THEOREME 6. On a la << formule d'intégration par parties >>

$$(25) \quad E[g_k(Y_{\bullet})D_\ell f(Y_\tau)U^\ell_{j\tau}\gamma_\tau^{kj}] = E[f(Y_\tau) < g(Y_{\bullet}),\ \int_0^\tau U_s^{-1}a_\lambda(Y_s) * dZ_s^\lambda >] - E[f(Y_\tau)\int_0^\tau \nu_{k\ell}(Y_{\bullet},dt)U^\ell_{jt}\gamma_t^{kj}\ ] \ .$$

Si l'on veut, on peut interpréter la quantité du côté gauche, sous le signe $E[\ ]$, comme $\gamma(g,\ U^*_\tau(df|_{Y_\tau}))$, sans coordonnées. Du côté gauche, on a des dérivées en f, et du côté droit on n'en a plus.

1. On ne peut confondre ces <,> avec des crochets obliques de semimartingales !

APPENDICE : LIEN AVEC LE CALCUL DE MALLIAVIN

Nous nous proposons ici de faire un << pont >> entre le calcul des variations de Bismut, présenté dans les pages qui précèdent, et le calcul au moyen des opérateurs L et Γ, introduit par Malliavin, et présenté dans un autre exposé de ce volume (<< Note sur le processus d'Ornstein-Uhlenbeck >>, auquel nous renvoyons plus bas sous la référence [O-U]). Une bonne partie de cet appendice est purement formelle, consistant uniquement à expliquer le principe de la méthode de Malliavin. Puis nous montrons où se fait le lien avec la méthode de Bismut, et où interviennent les crochets de Lie de la condition de non-dégénérescence de Hörmander. Dans cette seconde partie, nous suivons à nouveau Bismut, et redonnons des démonstrations complètes.

I. Nous reprenons l'équation (1') dans le << cas brownien >>, que nous écrivons explicitement ci-dessous. Le mouvement brownien est réalisé sur son espace canonique, muni de la mesure de Wiener P

$$dY_t = a_0(Y_t)dt + \sum_{\lambda=1}^{n} a_\lambda(Y_t)*dB_t^\lambda \ , \ Y_0=x \ .$$

On fixe un instant $\tau>0$, et on se propose de montrer que la loi η de Y_τ a une densité C^∞ , sous certaines conditions de non-dégénérescence de la famille des champs a_λ $(\lambda\geqq 0)$.

Nous ne nous occuperons pas ici que de l'existence d'une densité pour la mesure η, laissant la régularité de côté. D'après un lemme d'analyse harmonique, pour montrer que η est absolument continue, il suffit de montrer que ses dérivées au sens des distributions sont des mesures bornées, autrement dit que pour $f\in \underline{\underline{C}}_c^\infty(\mathbb{R}^d)$

$$(26) \qquad |E[D_if(Y_\tau)]| = |<\eta, D_if>| \leqq C\sup_x|f(x)| \quad (i=1,\dots,d).$$

L'idée de Malliavin consiste à considérer Y_τ comme une fonctionnelle sur l'espace du mouvement brownien, et à montrer qu'elle est << de classe C^∞ >>, c.à d. que l'on peut lui appliquer autant de fois qu'on le veut les opérateurs L et Γ de [O-U]. Faisons alors un calcul formel, dont le principe est tout à fait simple.

Nous partons de la formule (26) de [O-U], en omettant la mention de τ pour alléger un peu

$$\Gamma((f(Y),Y^j) = \sum\nolimits_k D_kf(Y)\ \Gamma(Y^k,Y^j)$$

Posons $c^{kj}=\Gamma(Y^k,Y^j)$; identifiant cette forme quadratique positive à une matrice symétrique positive, appelons Δ son déterminant et b_{jk} ses cofacteurs (plus exactement Δ_τ , $b_{jk\tau}$, mais τ est omis pour l'instant). Nous avons

$$(27) \qquad \Delta D_kf(Y) = \sum\nolimits_j b_{kj}\Gamma(f(Y),Y^j)$$

Remplaçons Γ par sa définition : $L(Y^jf(Y)-Y^jLf(Y)-f(Y)LY^j$, multiplions encore par une fonction ϕ qui sera << bien régulière >> - dans beaucoup de cas, nous pourrons prendre $\phi=1/\Delta$, de manière à retomber sur (26) - et intégrons. Nous devons nous rappeler ici que L est symétrique : chaque fois que dans un produit L tombe sur f(Y), nous le faisons passer sur l'autre facteur. Nous obtenons

(28) $$E[\Delta\phi D_k f(Y)] = \textstyle\sum_j E[f(Y)(Y^j L(\phi b_{jk})-L(Y^j\phi b_{jk})-\phi b_{jk}LY^j]$$
$$=-\textstyle\sum_j E[f(Y)\{ 2\phi b_{jk}L(Y^j)+\Gamma(\phi b_{jk},Y^j)\}]$$

et nous voyons que le problème de la continuité absolue de η se ramène à la vérification de deux points (on prend $\phi=1$ pour commencer)

- L'intégrabilité des $\{\ \ \}_k$.
- La non dégénérescence de la forme (c^{jk}). En effet, si $\Delta>0$ p.s., (26) nous dira que l'image d'une mesure équivalente à la mesure de Wiener est absolument continue, d'où le même résultat pour l'image de la mesure de Wiener elle même.

Nous laisserons ici de côté le premier problème, pour nous attacher au second.

II. Choisissons une suite (ξ_n) de fonctions continues sur $\mathbb{R}_+$, à valeurs dans $\mathbb{R}^d$, nulles en 0, dont les dérivées $\dot\xi_n$ sont C^∞ à support compact et forment une base orthonormale de $L^2(\mathbb{R}_+,\mathbb{R}^d)$. Soient g et h deux fonctions C^∞ sur $\mathbb{R}^d$. Nous avons d'après le th. 9 de [O-U]

(29) $$\Gamma(g(Y_\tau),h(Y_\tau)) = \textstyle\sum_n \nabla_{\xi_n} g(Y_\tau)\, \nabla_{\xi_n} h(Y_\tau)$$

Mais nous avons calculé explicitement ces dérivées : ce sont les variations premières données par les théorèmes 1 et 2 :

$$\nabla_{\xi_n} g(Y_\tau) = \textstyle\sum_i D_i g(Y_\tau)\int_0^\tau (U_{\tau-s}a_\lambda(Y_s))^i \dot\xi^\lambda_{ns}ds \quad (\lambda=1,\dots n)$$

Calculons alors (29) : en utilisant le fait que les $\dot\xi_n$ forment une base, il vient

(30) $$\Gamma(g(Y_\tau),h(Y_\tau)) = D_i g(Y_\tau)D_j h(Y_\tau)c^{ij}_\tau$$
$$c^{ij}_\tau = \int_0^\tau \sum_{\lambda=1}^n (U_{\tau-s}a_\lambda(Y_s))^i(U_{\tau-s}a_\lambda(Y_s))^j ds$$

et nous avons explicitement les coefficients $c^{ij}_\tau=\Gamma(Y^i_\tau,Y^j_\tau)$ du paragraphe précédent.[1]

Les c^{ij}_τ sont les coefficients d'un tenseur symétrique (une forme bilinéaire positive sur $T^*_{Y_\tau}(V)$, image par U_τ du tenseur symétrique au point initial x

(31) $$S_\tau = \int_0^\tau \Sigma_1^n U_s^{-1}a_\lambda(Y_s)\otimes U_s^{-1}a_\lambda(Y_s)ds$$

Supposons que $S_\tau(\omega)$ définisse une forme quadratique dégénérée sur $T^*_x(V)$, et soit $v(\omega)$ un élément de longueur nulle. Par continuité, on a

1. Maintenant, nous suivons de nouveau Bismut mot pour mot.

$< v(\omega), U_s^{-1}a_\lambda(\omega) > = 0$ pour tout $\lambda=1,..,n$ et $t \leq \tau$. Soit $\mathcal{H}_t(\omega)$ le sous-espace de $T_x(V)$ engendré par les $U_s^{-1}a_\lambda(\omega)$, $s\leq t$, $\lambda=1,...,n$, et soit $\mathcal{H}_{0+}(\omega) = \bigcap_{t>0} \mathcal{H}_t$. En fait, $\mathcal{H}_t$ est $\underline{\underline{F}}_t$-mesurable, $\mathcal{H}_{0+}$ est $\underline{\underline{F}}_0$-mesurable, donc p.s. égal à un sous-espace fixe de $T_x(V)$ (cela exige que l'on définisse une bonne stucture mesurable sur l'ensemble des sous-espaces de $T_x(V)$, je n'insiste pas là dessus). Comme $v(\omega)$ est une forme orthogonale à $\mathcal{H}_{0+}(\omega)$, on voit que :

ou bien $S_\tau(\omega)$ est p.s. non dégénéré, ou bien il existe une forme $v\varepsilon T_x^*(V)$ non aléatoire orthogonale à $\mathcal{H}_{0+}$.

Mais soit $\sigma(\omega)=\inf\{t : \mathcal{H}_t(\omega)\neq\mathcal{H}_{0+}(\omega)\}$: σ est un temps d'arrêt >0, et l'on voit que v est en fait orthogonal à $\mathcal{H}_t(\omega)$ pour tout $t<\sigma(\omega)$.

III. Où les crochets de Lie font leur apparition. Nous utiliserons le lemme suivant[2], très proche de la formule $(52)_c$ de l'exposé << Géométrie Différentielle stochastique >>, la dérivée de Lie $\mathcal{L}$ remplaçant la dérivée covariante ∇ (rappelons que $[a,K]=\mathcal{L}_aK$ (Sém. XV, p. 68-69)).

LEMME 6. Soit K un champ de vecteurs sur $V=\mathbb{R}^d$. Alors on a

$$(32)\qquad U_t^{-1}K(Y_t) = K(x) + \int_0^t U_s^{-1}([a_\lambda,K])(Y_s)*dZ_s^\lambda \quad \text{(sommation sur tous les } \lambda \text{ y compris 0)}$$

DEMONSTRATION. Nous nous bornerons à esquisser cela : on a $d(U_t^{-1}K(Y_t))= U_t^{-1}*dK(Y_t) +(*dU_t^{-1})K(Y_t) = U_t^{-1}(*dK(Y_t)-*dL_tK(Y_t))$ (dL_t est définie après la formule (2), et dU_t^{-1} est donnée par la formule (3'))

D'autre part d'après l'expression de dL

$$*dL_tK(Y_t)^i = D_ja_\lambda^i(Y_t)K^j(Y_s)*dZ_t^\lambda$$
$$*dK(Y_t)^i = D_jK^i(Y_t)*dY_t^j = D_jK^i(Y_t)a_\lambda^j(Y_t)*dZ_t^\lambda$$

En formant $*dK-*dLK$ on voit donc apparaître les $a_\lambda^jD_jK^i-K^jD_ja_\lambda^i$, i.e. les composantes de $[a_\lambda,K]$.

REMARQUE. Si l'on écrit (32) avec $\mathcal{L}_{a_\lambda}K$ au lieu du crochet, on obtient une formule valable pour tous les champs de tenseurs. En particulier, pour une forme Θ, la formule (qui s'écrit avec $U_t^* : T_{Y_t}^* \to T_x^*$)

$$(33)\qquad U_t^*\Theta_{Y_t} = \Theta_x + \int_0^t U_s^*(\mathcal{L}_{a_\lambda}\Theta(Y_s))*dZ_s^\lambda$$

est la forme intrinsèque de la formule (2), qui était matricielle. On obtient en effet celle-ci en prenant $\Theta=dx^i$, $\mathcal{L}_{a_\lambda}dx^i=da_\lambda^i$. Cette formule a une propriété remarquable : elle s'étend aux formes d'ordre 2, permettant ainsi de calculer les extensions des U_t en opérateurs de l'espace tangent du second ordre $\tau_x(V)$ dans l'espace $\tau_{Y_t}(V)$.[3]

1. Il ne s'agit pas de "crochets obliques". 2. Bismut. Mécanique Aléatoire, p.158-163
3. Pour l'extension de $\mathcal{L}$ à l'ordre 2, voir Sém. XV, p. 68-69.

Développons la formule (32) en intégrales d'Ito . Soit $H(K)_t=U_t^{-1}K(Y_t)$, semimartingale dans $T_x(V)$. Nous avons (< , > est un crochet oblique !)

$$d\langle H(K), Z^\lambda\rangle_t = H(\mathcal{L}_{a_\mu} K)d\langle Z^\mu,Z^\lambda\rangle_t \quad (\text{ sommation sur } \mu=0,..,n)$$

et par conséquent, en remplaçant K par $\mathcal{L}_{a_\lambda} K$, pour développer l'intégrale de Stratonovitch dans (32)

$$(34) \quad H(K)_t = K(x) + \int_0^t H(\mathcal{L}_{a_\lambda} K)_s dZ_s^\lambda + \frac{1}{2}\int_0^t H(\mathcal{L}_{a_\mu}\mathcal{L}_{a_\lambda} K)_s d\langle Z^\mu,Z^\lambda\rangle_s$$

ou dans le cas brownien

$$(35) \quad U_t^{-1}K(Y_t) = K(x) + \sum_1^n \int_0^t U_s^{-1}\,[a_\lambda,K](Y_s)dZ_s^\lambda$$
$$+ \int_0^t U_s^{-1}[a_0,K](Y_s)ds + \frac{1}{2}\sum_1^n U_s^{-1}[a_\lambda,[a_\lambda,K]](Y_s)ds$$

IV. Revenons alors à la fin du II. Si v est non aléatoire, orthogonal à $\mathcal{H}_t(\omega)$ pour tout $t<\sigma(\omega)$, les semimartingales réelles $\langle v,U_t^{-1}K(Y_t)\rangle$ (crochets gras ! valeur d'une forme sur un vecteur) sont nulles, pour $K=a_1,\dots,a_n$. Donc les parties martingales sont nulles, et on voit que

l'orthogonalité de v aux $U_s^{-1}K(Y_s)$, $s<\sigma$ entraîne
l'orthogonalité aux $U_s^{-1}[a_\lambda,K](Y_s)$, $s<\sigma$, $\lambda=1,..,n$

En itérant, on voit que v est orthogonal à tous les crochets itérés $U_s^{-1}J(Y_s)$ ou $J=a_\lambda$, $[a_\lambda,a_\mu]$, $[a_\lambda,[a_\mu,a_\nu]]\dots$ correspondant à $\lambda,\mu,\nu..\geq 1$

Mais alors, en regardant les parties à variation finie, on voit que v est orthogonal aussi aux $U_s^{-1}[a_0,J](Y_s)$. Et finalement, nous sommes parvenus à une condition géométrique :

LEMME 7. Si au point x les champs $a_1,\dots,a_n$ et les crochets itérés formés à partir de $a_0,a_1,\dots,a_n$ engendrent $T_x(V)$, S_τ définit p.s. une forme non dégénérée pour $\tau>0$.

Noter le rôle particulier de a_0 !

L'ensemble des raisonnements faits plus haut montre que, si cette condition géométrique est satisfaite, la loi de $Y_t(x)$ est absolument continue. Bien sûr, cela peut se déduire de la formule de Bismut (25), mais il était peut être plus instructif de procéder autrement, et de renvoyer le lecteur à l'article de Bismut pour la comparaison...

Signalons enfin que la loi de $Y_t(x)$ admet en fait <u>une densité C^∞</u> sous la condition géométrique ci-dessus, et que cette conclusion reste exacte en supposant que les a_i (<u>y compris</u> a_o) et leurs crochets itérés engendrent $T_x(V)$ en tout point (alors que le lemme 7, lui, ne subsiste pas). Ces résultats figurent dans l'article de Bismut, et nous espérons avoir donné envie de les lire...