ANNALES DE L'INSTITUT FOURIER

GÉRARD BEN AROUS

Développement asymptotique du noyau de la chaleur hypoelliptique sur la diagonale

Annales de l'institut Fourier, tome 39, nº 1 (1989), p. 73-99

<http://www.numdam.org/item?id=AIF_1989__39_1_73_0>

Ann. Inst. Fourier, Grenoble
39, 1 (1989), 73-99

DÉVELOPPEMENT ASYMPTOTIQUE DU NOYAU DE LA CHALEUR HYPOELLIPTIQUE SUR LA DIAGONALE

par Gérard BEN AROUS

Introduction.

Cet article continue l'étude entreprise dans [4] du comportement asymtotique, lorsque $t \to 0$, du noyau de la chaleur $p_t(x,y)$ associé à l'opérateur $L = 1/2 \sum_{1 \leq i \leq m} X_i^2$ sur $\mathbf{R}^d$ lorsque les champs de vecteur X_i sont de classe C^∞, bornés à dérivées bornées, et satisfont la condition d'Hörmander :

$$\forall x \in \mathbf{R}^d \quad \dim \mathrm{Lie}(X_1, \dots, X_m)(x) = d \tag{0.1}$$

où $\mathrm{Lie}(X_1, \dots, X_m)$ est l'algèbre de Lie engendrée par les champs X_i.

Léandre ([8] et [9]) a obtenu le comportement exponentiel de $p_t(x,y)$:

$$\lim_{t \to 0} 2t \log p_t(x,y) = -d^2(x,y) \tag{0.2}$$

où $d^2(x,y)$ est la métrique de Carnot-Carathéodory associée aux champs X_i. Jerison et Sanchez [5] ont encadré le noyau $p_t(x,y)$, pour tout $t > 0$:

$$A_1 \varphi(x,t)^{-1} e^{-C_1 d^2(x,y)/t} \leq p_t(x,y) \leq A_2 \varphi(x,t)^{-1} e^{-C_2 d^2(x,y)/t} \tag{0.3}$$

où A_1, A_2, C_1, C_2 sont des constantes positives et où $\varphi(x,t)$ est le volume de la boule (pour la distance de Carnot-Carathéodory) de centre x et de rayon $\sqrt{t}$.

Mots clés : Opérateurs hypoelliptiques – Noyau de la chaleur – Calcul de Malliavin – Processus de diffusion.

Classification A.M.S. : 35H05 – 35K05 – 58G32 – 60H07.

L'article [4] donnait le développement asymptotique (lorsque t tend vers zéro de $p_t(x,y)$:

$$(0.4) \qquad p_t(x,y) = t^{-d/2} e^{-d^2(x,y)/2t} \Big(\sum_{0 \le k \le N} c_k(x,y) t^k + O(t^{N+1}) \Big)$$

à tout ordre N, lorsque (x,y) n'appartient pas au cut-locus associé à L, c'est à dire lorsque l'unique géodésique minimisante joignant x à y est la projection d'une bicaractéristique (et que x et y sont non conjugués). On obtenait aussi la régularité des coefficients c_k sur le complémentaire du cut-locus ainsi que les équations de transport satisfaites par les c_k.

La diagonale rencontre le cut-locus exactement aux points où l'opérateur L n'est pas elliptique; ainsi le développement (0.4) ne renseigne pas sur le comportement asymptotique du noyau de la chaleur sur la diagonale, sauf dans le cas elliptique.

Nous abordons ici ce problème, avec les mêmes méthodes probabilistes qu'en [4], à savoir un développement de Taylor stochastique de la diffusion qui permet d'obtenir un développement asymptotique ("microlocal") de la transformée de Fourier de $p_t(x,y)$ puis l'inversion de la transformation de Fourier à l'aide du calcul de Malliavin. Il existe cependant deux différences de méthodes significatives : ici nous n'utilisons pas les théorèmes de grandes déviations par le biais de la méthode de Laplace de [3] comme en [4], ceci étant dû au fait que l'étude est limitée à la diagonale (le rôle des grandes déviations est tenu ici par le lemme de localisation du §2); par contre il est nécessaire ici de différencier les directions de la variable de Fourier et ceci est obtenu au moyen des relèvements de Rotschild et Stein sur un groupe nilpotent [11] (outil déjà utilisé par Jerison et Sanchez [5] et préssenti comme important pour ce problème dans l'article de L. Elie dans [1]).

Les coefficients de ce développement devraient, comme dans le cas riemannien, être des invariants importants de la géométrie associée à l'opérateur L. Par exemple, le premier coefficient devrait par analogie avec le cas riemannien, jouer le rôle de la densité d'une forme volume "graduée".

Nous obtenons, en fait, ici (au §7) en utilisant les résultats de [2], une description explicite du premier coefficient du développement asymptotique de $p_t(x,x)$, comme la valeur en 0 de la densité d'une variable aléatoire décrite, d'une part, par certaines constantes universelles (obtenues pour certaines à partir de la série de Campbell-Hausdorff et pour d'autres à partir des constantes de structures des groupes libres nilpotents) et, d'autre part, par les crochets des champs X_i et leurs relations de dépendance linéaire (ce qui fait intervenir de façon *universelle* la géométrie *spécifique* du problème).

Léandre obtient (cf [10]) le résultat principal de cet article (énoncé au §1) par une méthode qui semble éviter le relèvement de Rotschild et Stein et utiliser plutôt les résultats de [2] dans le cas non nilpotent.

Remarquons enfin que le résultat est ici énoncé sur $\mathbf{R}^d$, mais que le lemme de localisation du §2 permettrait de se placer sur une variété. Pour conclure il nous semble maintenant clair que la combinaison des méthodes de grandes déviations (au travers de la méthode de Laplace développée en [3]) combinées avec la méthode de relèvement de Rotschild et Stein utilisée comme ici, devrait permettre d'obtenir le développement du noyau de la chaleur sur le cut-locus et hors de la diagonale.

1. Énoncé du résultat.

Soient $X_1, \dots, X_m$ des champs de vecteurs C_b^∞ sur $\mathbf{R}^d$ tels que :

$$\dim \operatorname{Lie}(X_1, \dots, X_m)(x) = d \quad \forall x \in \mathbf{R}^d \tag{1.1}$$

où $\operatorname{Lie}(X_1, \dots, X_m)$ désigne l'algèbre de Lie engendrée par les champs X_i. Soit alors $p_t(x,y)$ le noyau de la chaleur associée à l'opérateur hypoelliptique $L = 1/2 \sum_{1 \le i \le m} X_i^2$, à savoir la densité de la loi de la diffusion x_t associée à L, i.e. la solution de l'équation de Stratonovitch :

$$dx_t = \sum_{1 \le i \le m} X_i(x_t) \circ dw_t^i \text{ avec } x_0 = x \tag{1.2}$$

où $(w^1, \dots, w^m)$ désigne un mouvement brownien m-dimensionnel. Nous allons étudier ici le développement asymptotique, lorsque $t \to 0$, du noyau de la chaleur sur la diagonale, c'est-à-dire de $p_t(x,x)$.

Pour cela introduisons les notations suivantes :

Si $J = (j_1, \dots, j_p)$ est un mot, avec $1 \le j_1 \le m$, on notera $|J| = p$, et X^J désignera le commutateur des champs X_i défini par :

$$X^J = [X_{j_1}, [X_{j_2}, \dots, [X_{j_{p-1}}, X_{j_p}] \dots]. \tag{1.3}$$

Définissons alors, pour $k \in \mathbb{N}$:

$$C_k(x) = \mathrm{Vect}\{X^J(x), |J| \leq k\} \tag{1.4}$$

et

$$r(x) = \inf\{k,\ \dim C_k(x) = d\}. \tag{1.5}$$

Par l'hypothèse (1.1) on sait que $r(x)$ est fini. On notera enfin :

$$Q(x) = \sum_{1 \leq k \leq r(x)} k(\dim C_k(x) - \dim C_{k-1}(x)). \tag{1.6}$$

On introduit alors une relation d'équivalence $\sim$ sur $\mathbf{R}^d$.

DÉFINITION (1.7). — *On dira que deux points x et y de $\mathbf{R}^d$ sont équivalents et on notera $x \sim y$ si pour tout entier $k \geq 0$ les dimensions de $C_k(x)$ et de $C_k(y)$ sont égales.*

Il est clair que la relation introduite dans la définition 1.7 est bien une relation d'équivalence. On notera $A(x)$ la classe d'équivalence de x pour cette relation. Sur une classe d'équivalence la "géométrie des crochets des champs X_i" est constante, ainsi l'entier $Q(x)$ y est constant. Ces classes d'équivalence ne sont pas en général de même dimension; dans le cas où les champs X_i sont analytiques ces classes d'équivalences sont des ensembles semi-analytiques.

On a alors le résultat central de cet article :

THÉORÈME. — 1) *Sous l'hypothèse (1.1) on a, pour $x_0 \in \mathbf{R}^d$ fixé, pour tout $N \in \mathbb{N}$, le développement asymptotique du noyau de la chaleur sur la diagonale :*

$$p_t(x_0, x_0) = t^{-Q(x_0)/2}\Big(\sum_{0 \leq k \leq N} c_k(x_0) t^k + t^{N+1} r_{N+1}(t, x_0)\Big) \tag{1.8}$$

avec :

$$\text{Il existe } t_0 > 0 \text{ tel que : } \sup_{t \leq t_0} |r_{N+1}(t, x_0)| < \infty \tag{1.9}$$

et

$$c_0(x_0) > 0. \tag{1.10}$$

2) *Le développement asymptotique (1.8) est uniforme sur tout compact K inclus dans une classe d'équivalence :*

$$\sup_{t \leq t_0, x \in K} |r_{N+1}(t,x)| < \infty. \tag{1.11}$$

De plus les coefficients c_k sont continus sur K.

3) Si la classe d'équivalence de x_0 est d'intérieur non vide alors les coefficients $c_k(x)$ du développement (1.8) ainsi que le reste $r_{N+1}(t,x)$ sont C^∞ en x au voisinage de x_0.

Remarque sur le plan de la preuve . — Nous allons prouver le théorème en commençant au §2 par montrer que l'étude asymptotique de $p_t(x_0,x_0)$ est un problème local; au §3 nous donnerons la preuve du théorème (à l'exception de (1.10)) en admettant deux lemmes essentiels (les lemmes 3.9 et 3.14) qui nécessitent l'usage des relèvements de Rotschild et Stein sur un groupe nilpotent. Ces deux lemmes seront prouvés au §6. Le §4 est consacré au rappel de ces résultats de relèvement, le §5 à l'étude de la diffusion relevée sur le groupe nilpotent et le §6 à la redescente du groupe sur $\mathbf{R}^d$ et à la preuve des lemmes admis. Enfin au §7 nous prouverons (1.10) et nous étudierons le premier coefficient c_0 de manière explicite.

Exemple. — Pour illustrer le théorème considérons un exemple très simple d'opérateur hypoelliptique, celui de l'opérateur de Grushin, où les calculs sont possibles directement sans avoir recours au théorème.

Soit, sur $\mathbf{R}^2$, l'opérateur $L = 1/2(X_1^2 + X_2^2)$, avec $X_1 = \partial_{x_1}$ et $X_2 = x_1\partial_{x_2}$.

Il y a alors deux classes d'équivalence, à savoir l'axe $x_1 = 0$ (où $Q(x) = 3$) et son complémentaire (où L est elliptique et donc où $Q(x) = 2$).

Le noyau de la chaleur $p_t(x,y)$ associé à L est la densité de la loi de la diffusion X_t^x donnée par :

$$X_t^x = (x_1 + w_t^1\,, x_2 + x_1 w_t^2 + \int_{0\leq s\leq t} w_s^1 dw_s^2) \tag{1.12}$$

où (w^1,w^2) est un mouvement brownien sur $\mathbf{R}^2$. Si l'on introduit l'aire stochastique $S_t = 1/2\left(\int_{0\leq s\leq t} w_s^1 dw_s^2 - \int_{0\leq s\leq t} w_s^2 dw_s^1\right)$ on a :

$$X_t^x = (x_1 + w_t^1, x_2 + x_1 w_t^2 + w_t^1 w_t^2/2 + S_t). \tag{1.13}$$

Par inversion de Fourier on a :

$$p_t(x,x) = (2\pi)^{-2} \int E[e^{i\xi(X_t^x - x)}]d\xi. \tag{1.14}$$

Or il est facile de calculer $E[e^{i\xi\cdot(X_t^x - x)}]$ en utilisant la formule de l'aire de Paul Lévy (voir par exemple Gaveau [6]) :

(1.15)

$$E[e^{i\alpha S_t}|w_t = z] = \alpha t(2sh(\alpha t/2))^{-1}\exp[(1-\alpha t(\coth(\alpha t/2))/2)|z|^2/2t].$$

On obtient ainsi, après le changement de variable $z = \sqrt{t}u$,

$$E[e^{i\xi\cdot(X_t^x - x)}] = (2\pi)^{-1}(\xi_2 t)(2sh(\xi_2 t/2))^{-1}\int e^{F(u,t,\xi,x)}du \tag{1.16}$$

avec

(1.17)

$$F(u,t,\xi,x) = i[\sqrt{t}(\xi_1 u_1 + x_1\xi_2 u_2) + t(\xi_2 u_1 u_2)/2] - |u|^2(\xi_2 t\coth(\xi_2 t/2))/4.$$

Si $x \neq 0$, c'est-à-dire si l'opérateur L est elliptique au voisinage de x (et donc $Q(x) = 2$), on effectue le changement de variable $\zeta = \sqrt{t}\xi$ dans (1.14) :

$$p_t(x,x) = (2\pi)^{-2}t^{-1}\int E[e^{i\zeta/\sqrt{t}(X_t^x - x)}]d\zeta, \tag{1.18}$$

et (1.16) montre aisément que :

$$p_t(x,x) \simeq (2\pi)^{-2}t^{-1}\int d\zeta e^{-(\zeta_1^2 + (x_1\zeta_2)^2)/2} = (2\pi t|x_1|)^{-1}. \tag{1.19}$$

Par contre, si $x_1 = 0$ (auquel cas $Q(x) = 3$) on a après le changement de variable $\zeta_1 = \sqrt{t}\xi_1$ et $\zeta_2 = t\xi_2$:

$$p_t(x,x) = (2\pi)^{-3}t^{-3/2}\int \varphi(\zeta)d\zeta \tag{1.20}$$

où $\varphi(\zeta) = \zeta_2(2sh(\zeta_2/2))^{-1}\int du e^{-|u|^2(\zeta_2\coth(\zeta_2/2))/4}e^{i(\zeta_1 u_1 + \zeta_2(u_1 u_2/2))}$.

Ainsi sur cet exemple très simple on vérifie que le comportement asymptotique de $p_t(x,x)$ est bien celui annoncé dans le théorème.

On obtient aussi la valeur du premier coefficient du développement asymptotique de $p_t(x,x)$.

Remarques. — 1) Si $x_1 \neq 0$, la valeur du premier coefficient est identique à celui que l'on obtiendrait en calculant le développement asymptotique de la densité de la diffusion gaussienne ayant mêmes moments d'ordre 1 et 2 que la diffusion X_t^x, c'est-à-dire ici de la diffusion associée à l'opérateur L où l'on aurait gelé les coefficents en x.

2) Dans cet exemple le premier coefficient du développement de $p_t(x,x)$ est indépendant de x sur l'axe $\{x_1 = 0\}$. Il y est donc évidemment non seulement continu, comme l'affirme le théorème mais aussi de classe C^∞. Le fait que cet axe (qui est ici $A(x)$) soit une sous variété de $\mathbf{R}^2$ est le point important. Il est clair que la preuve du théorème donnerait cette régularité du premier coefficient (et des suivants) dans le cas général dès que, comme ici, la classe d'équivalence $A(x)$ est une sous variété de $\mathbf{R}^d$.

3) Si, lorsque $x_1 = 0$, l'on considère de nouveau la variable gaussienne ayant les mêmes moments d'ordre 1 et 2 que X_t^x on obtient le même exposant de t dans le développement asymptotique de la densité, par contre le premier coefficient est différent.

4) On constate aussi sur cet exemple que la singularité du premier coefficient du développement asymtotique lorsque x change de "strate" (ou encore ici lorsque $x_1 \to 0$) est un pôle. Cette singularité est liée à la bifurcation de la géométrie au voisinage de l'axe $x_1 = 0$. Le problème de la nature de cette singularité n'est pas abordée dans cet article dans le cas général, mais nous espérons y consacrer un travail futur.

2. Localisation.

Nous allons commencer la preuve du théorème en montrant ici que le problème de l'estimation asymptotique de $p_t(x_0, x_0)$ est local.

Soient $\overline{X}_1, \ldots, \overline{X}_m$ des champs de vecteurs C_b^∞ sur $\mathbf{R}^d$ tels que :

(2.1) $\qquad \overline{X}_i(x) = X_i(x)$ pour x dans un voisinage V de x_0.

Soit alors $x_t^\varepsilon(x)$ (respectivement $\overline{x}_t^\varepsilon(x)$) la solution de l'équation de Stratonovich :

$$(2.2) \qquad dx_t^\varepsilon = \varepsilon \sum_{1 \le i \le m} X_i(x_t^\varepsilon) \circ dw_t^i \text{ avec } x_0^\varepsilon = x$$

(respectivement de :

$$(2.3) \qquad d\overline{x}_t^\varepsilon = \varepsilon \sum_{1 \le i \le m} \overline{X}_i(\overline{x}_t^\varepsilon) \circ dw_t^i \text{ avec } \overline{x}_0^\varepsilon = x).$$

On a alors le résultat de localisation :

LEMME (2.4). — *Pour tout compact K inclus dans V, pour tout $p \geq 1$, il existe des constantes K_p et $C > 0$ telles que, pour tout $\xi \in \mathbf{R}^d$:*

$$\sup_{x \in K} |E[e^{i\xi \cdot (x_1^\varepsilon(x) - x)} - e^{i\xi \cdot (\overline{x}_1^\varepsilon(x) - x)}]| \leq K_p |\xi|^{-2p} e^{-C/\varepsilon^2}.$$

Preuve. — Posons $f(u) = e^{i\xi \cdot (u-x)}$, l'intégration par parties du calcul de Malliavin (cf [12] et [13]) montre que pour $x \in K$:

$$E[((\Delta_u)^p f)(x_1^\varepsilon(x))] = E[f(x_1^\varepsilon(x)) l_m(\varepsilon, x)] \tag{2.5}$$

où l'on sait, par le théorème 8.43 de [12], qu'il existe un $\nu > 0$ tel que pour tout $q > 1$ on ait :

$$\sup_{x \in K} E[|l_m(\varepsilon, x)|^q]^{1/q} \leq K(m, q)\varepsilon^{-\nu}. \tag{2.6}$$

De même on a :

$$E[((\Delta_u)^p f)(\overline{x}_1^\varepsilon(x))] = E[f(\overline{x}_1^\varepsilon(x)) \overline{l}_m(\varepsilon, x)] \tag{2.7}$$

avec

$$\sup_{x \in K} E[|\overline{l}_m(\varepsilon, x)|^q]^{1/q} \leq \overline{K}(m, q)\varepsilon^{-\nu}. \tag{2.8}$$

Remarquons en effet pour cela que la condition 8.32 dans [12] est locale. D'où l'on déduit :

$$\begin{aligned} &E[e^{i\xi \cdot (x_1^\varepsilon - x)} - e^{i\xi \cdot (\overline{x}_1^\varepsilon - x)}] = \\ &\qquad |\xi|^{-2p} E[e^{i\xi \cdot (x_1^\varepsilon - x)} l_m(\varepsilon) - e^{i\xi \cdot (\overline{x}_1^\varepsilon - x)} \overline{l}_m(\varepsilon)]. \end{aligned} \tag{2.9}$$

De plus si $\tau^\varepsilon(x) = \inf\,(t, x_t^\varepsilon(x) \notin V)$, alors sur $\{\tau^\varepsilon(x) > 1\} x_t^\varepsilon(x)$ et $\overline{x}_t^\varepsilon(x)$ coïncident ainsi que $l_m(\varepsilon, x)$ et $\overline{l}_m(\varepsilon, x)$. On en déduit que :

$$\begin{aligned} &E[e^{i\xi \cdot (x_1^\varepsilon - x)} - e^{i\xi \cdot (\overline{x}_1^\varepsilon - x)}] = \\ &\qquad |\xi|^{-2p} E[(e^{i\xi \cdot (x_1^\varepsilon - x)} l_m(\varepsilon, x) - e^{i\xi \cdot (\overline{x}_1^\varepsilon - x)} \overline{l}_m(\varepsilon, x)) 1_{\tau^\varepsilon(x) > 1}], \end{aligned} \tag{2.10}$$

et donc par (2.6) et (2.8) que :

$$\sup_{x \in K} |E[e^{i\xi\cdot(x_1^\varepsilon - x)} - e^{i\xi\cdot(\overline{x}_1^\varepsilon - x)}]| \leq \tag{2.11}$$
$$K'(m,q)|\xi|^{-2p}\varepsilon^{-\nu}P(\tau^\varepsilon(x) > 1)^{1-1/q}.$$

De l'inégalité exponentielle usuelle :

$$\sup_{x \in K} P(\tau^\varepsilon(x) > 1) \leq K e^{-C/\varepsilon^2} \tag{2.12}$$

on déduit alors le lemme (2.4).

3. Construction d'une bonne carte et preuve du théorème.

Pour démontrer les théorèmes 1 et 2 nous allons introduire une carte adaptée au problème :

Avec les notations du §1 nous savons que : $\mathcal{C}_{r(x_0)} = \mathbf{R}^d$; considérons une base de $\mathbf{R}^d$: $\{X^J(x_0), J \in B\}$ où l'ensemble B de multi-indices est choisi de telle façon que la base soit triangulaire dans le sens suivant : Pour tout entier k inférieur ou égal à $r(x_0)$

$$\{X^J(x_0), J \in B, |J| \leq k\} \text{ engendre } \mathcal{C}_k(x_0). \tag{3.1}$$

La famille $\{X^J(x), J \in B\}$ reste une base pour x assez voisin de x_0, mais pas nécessairement une base triangulaire au sens de (3.1).

De fait cette base reste triangulaire au sens de (3.1) pour $x \in K$, si K est un compact inclus dans $A(x_0)$ et dans un voisinage assez petit de x_0.

Enfin il est clair qu'il existe un voisinage V_1 de x_0 et un voisinage W de 0 tels que, pour tout x de V_1, l'application φ_x :

$$(v_J)_{J \in B} \to \varphi_x(v_j) = \exp\left(\sum_{J \in B} v_J X^J\right)(x) \tag{3.2}$$

définisse un difféomorphisme de W sur son image $\varphi_x(W)$. De plus il existe des voisinages V_2 et U de x_0 tels que :

$$V_2 \subset \overline{V}_2 \subset V_1 \text{ et } \forall x \in V_2 \ U \subset \varphi_x(W). \tag{3.3}$$

Nous allons étudier la diffusion $x_t^\varepsilon(x)$ dans la carte donnée par φ_x.

Commençons par remplacer les champs X_i par les champs $\overline{X}_i = \theta X_i$, où θ est une fonction de troncature C^∞ telle que : $\theta = 1$ sur un voisinage V de x_0 tel que $V \subset \overline{V} \subset U$, et $\theta = 0$ hors de $\overline{V}$.

La solution $\overline{x}_t^\varepsilon(x)$ de l'équation définie par les champs $\overline{X}_i$:

$$d\overline{x}_t^\varepsilon = \varepsilon \sum_{1\le i\le m} \overline{X}_i(\overline{x}_t^\varepsilon) \circ dw_t^i \text{ avec } \overline{x}_0^\varepsilon = x \tag{3.4}$$

reste dans l'ouvert U et l'on peut considérer son expression dans la carte donnée par φ_x :

$$\overline{x}_t^\varepsilon(x) = \exp\Big(\sum_{J\in B} v_J(,\varepsilon,x)X^J\Big)(x) \tag{3.5}$$

ou encore

$$\varphi_x^{-1}(\overline{x}_t^\varepsilon(x)) = (v_J(t,\varepsilon,x))_{J\in B}. \tag{3.6}$$

On a alors le résultat classique de régularité :

Lemme (3.7). — *1) Pour tout $J \in B$ $v_J(\varepsilon,t,x)$ est une fonction C^∞ de (ε,x) pour*
2) Pour tout triplet $(j,k,n) \in \mathbb{N}^3$, pour tout multi-indice $\alpha \in \{1,\dots,d\}^j$, et pour tout $q \ge 2$, il existe $\varepsilon_0 > 0$ tel que :

$$\sup_{\varepsilon\le\varepsilon_0, x_0\in K} \sum_{i\le n} E[\sup_{0\le t\le 1} \|D^i(\partial_x)^\alpha(\partial_\varepsilon)^k(v_J(\varepsilon,tx))\|_{HS}^q] < \infty \tag{3.8}$$

où D désigne la dérivée au sens de Malliavin (ou de Shigékawa, voir [13]) et $\|\cdot\|_{HS}$ la norme Hilbert-Schmidt.

Preuve du lemme (3.7). — Cette preuve est identique à celle du lemme 3.10 de [4]. Notons $\Xi_t(\varepsilon,x_0)$ le processus constitué de toutes les dérivées :

$(\partial_{x_0})^\alpha(\partial_\varepsilon)^p(\varphi_x^{-1}(\overline{x}_t^\varepsilon(x)))$ avec $|\alpha| \le j, p \le k$. Ce processus Ξ_t est solution d'une équation des variations stochastiques. Cette équation est a priori écrite au sens de Stratonovitch mais on peut aisément l'écrire sous la forme d'une équation différentielle stochastique au sens d'Ito.

Du fait que les champs $\varphi_x^{-1} * \overline{X}_i$ qui définissent l'équation satisfaite par $\varphi_x^{-1}(\overline{x}_t^\varepsilon(x))$ sont bornés à dérivées (en la variable d'espace et en les paramètres ε et x) bornées, les coefficients de l'équation d'Ito des variations stochastiques satisfaite par Ξ_t vérifient la condition de croissance lente (2.10) de Kusuoka-Stroock [7]. Le lemme (3.7) est alors une conséquence immédiate de la majoration (2.21) du théorème (2.19) de [7].

On a alors le résultat essentiel dont la preuve est remise aux paragraphes suivants et requiert l'usage des relèvements de Rotschild et Stein sur un groupe nilpotent :

LEMME (3.9). — *Soient $x \in A(x_0) \cap U$ et $J \in B$, pour tout entier $k < |J|$ on a :*

$$\partial_\varepsilon^K (v_j(\varepsilon, t, x))_{|\varepsilon=0} = 0.$$

Considérons alors, pour $\lambda \in \mathbf{R}$, la dilatation T_λ définie sur $\mathbf{R}^d$ par :

$$T_\lambda((v_j)_{j\in B}) = (\lambda^{|J|} v_J)_{J\in B}. \tag{3.10}$$

Soit alors $U(t, \varepsilon, x)$ la diffusion définie (a priori pour $\varepsilon > 0$) par :

$$\begin{aligned} U(t,\varepsilon,x) &= T_{1/\varepsilon} \circ \varphi_x^{-1}(\overline{x}_t^\varepsilon(x)) \text{ ou encore} \\ U_J(t,\varepsilon,x) &= \varepsilon^{-|J|} v_J(t,\varepsilon,x) \text{ pour tout } J \in B. \end{aligned} \tag{3.11}$$

Le lemme (3.9) montre que $U(t, \varepsilon, x)$ se prolonge de façon C^∞ en $\varepsilon = 0$, pour $x \in A(x_0) \cap U$.

De plus le lemme (3.7) montre alors que (avec les mêmes notations) :

$$\sup_{\varepsilon \le \varepsilon_0, x\in K} \sum_{i\le n} E[\sup_{0\le t\le 1} \|D^i(\partial_\varepsilon)^k(U(t,\varepsilon,x_0))\|_{HS}^q] < \infty \tag{3.12}$$

où K est un compact inclus dans $A(x_0) \cap U$.

Si $A(x_0)$ est d'intérieur non vide, supposons pour simplifier les notations que U soit choisi assez petit pour être inclus dans l'intérieur de $A(x_0)$ alors on a, toujours par (3.7) et (3.9) :

$$\sup_{\varepsilon \le \varepsilon_0, x\in K} \sum_{i\le n} E[\sup_{0\le t\le 1} \|D^i(\partial_x)^\alpha(\partial_\varepsilon)^k(U(t,\varepsilon,x))\|_{HS}^q] < \infty \tag{3.13}$$

pour K compact inclus dans U.

On a alors le résultat de non dégénérescence, dont la preuve est aussi repoussée aux paragraphes suivants :

LEMME (3.14). — *Pour $x \in A(x_0) \cap U$, la matrice de covariance de Malliavin de $U(t,0,x) = (\partial_\varepsilon^{|J|} v_J(\varepsilon,t,x)_{|\varepsilon=0})_{J\in B}$ est non dégénérée pour tout $t > 0$; précisément, pour tout $p \geq 1$, pour K compact inclus dans $A(x_0) \cap U$:*

$$\sup_{x\in K} E[|\det\,(DU(t,0,x), DU(t,0,x))|^{-p}] < \infty.$$

Montrons ici comment, en admettant les lemmes (3.9) et (3.14) on achève la preuve des théorèmes 1 et 2.

Soit $x \in A(x_0) \cap U$, et $q_t(\varepsilon,x,y)$ la densité de la loi de $\varphi_x^{-1}(\overline{x}_t^\varepsilon(x))$; par inversion de la transformée de Fourier on a :

$$q_1(\varepsilon,x,0) = (2\pi)^{-d} \int E[e^{i\xi\cdot\varphi^{-1}(\overline{x}_1^\varepsilon(x))}]d\xi. \tag{3.15}$$

Le changement de variable linéaire $\zeta = T_\varepsilon \xi$ donne :

$$q_1(\varepsilon,x,0) = (2\pi)^{-d}\varepsilon^{-Q(x_0)} \int E[e^{iT_{1/\varepsilon}\zeta\cdot\varphi^{-1}(\overline{x}_1^\varepsilon(x))}]d\zeta \tag{3.16}$$

ou encore

$$q_1(\varepsilon,x,0) = (2\pi)^{-d}\varepsilon^{-Q(x_0)} \int E[e^{i\zeta\cdot U(1,\varepsilon,x)}]d\zeta. \tag{3.17}$$

Soit alors la fonction :

$$\chi(\varepsilon, x\zeta) = e^{i\zeta\cdot(U(1,\varepsilon,x)-U(1,0,x))}. \tag{3.18}$$

Nous venons de vérifier que :

(3.19) Pour $x \in A(x_0) \cap U$, par le lemme 3.9 et (3.12) $\chi(\cdot,x,\zeta)$ est C^∞ en ε au voisinage de 0.

(3.20) Si $A(x_0)$ est d'intérieur non vide alors, par le lemme 3.9 et (3.13), $\chi(\cdot,\cdot,\zeta)$ est C^∞ en (ε,x) au voisinage de $(0,x_0)$.

Considérons le développement de Taylor en $\varepsilon = 0$, à l'ordre N, de $\chi(\cdot,\zeta)$:

$$\chi(\varepsilon,x,\zeta) = \sum_{0\leq k\leq N} \varepsilon^k/k!(\partial_\varepsilon)^k\chi(0,x,\zeta) + \varepsilon^{N+1}S_{N+1}(\varepsilon,x,\zeta) \tag{3.21}$$

avec

$$S_{N+1}(\varepsilon,\xi,\zeta) = \int_{0\leq v\leq 1} (\partial_\varepsilon)^{N+1}\chi(\varepsilon v,x,\zeta)(1-v)^N/N!dv. \tag{3.22}$$

Par (3.12) il est clair que :

$$\sup_{\varepsilon\le\varepsilon_0, x\in K} E\left[\sum_{i\le n}\sup_{t\in[0,1]}\|D^i(\partial_\varepsilon)^m\chi(\varepsilon,x,\zeta)\|_{HS}^p\right] \le C|\zeta|^m. \tag{3.23}$$

On a alors :

$$E[e^{i\zeta\cdot U(1,\varepsilon,x)}] = \sum_{0\le k\le N}\alpha_k(\zeta,x)\varepsilon^k + \varepsilon^{N+1}T_{N+1}(\varepsilon,x,\zeta) \tag{3.24}$$

avec

$$\alpha_k(\zeta,x) = E[e^{i\zeta\cdot U(1,0,x)}\partial_\varepsilon^k\chi(0,x,\zeta)/k!] \tag{3.25}$$

et

$$T_{N+1}(\varepsilon,x,\zeta) = E[e^{i\zeta\cdot U(1,0,x)}S_{N+1}(\varepsilon,x,\zeta)]. \tag{3.26}$$

Par (3.22) et (3.23) on a donc :

$$\sup_{\varepsilon\le\varepsilon_0, x\in K}|T_{N+1}(\varepsilon,x,\zeta)| < \infty. \tag{3.27}$$

Ce qui montre que (3.24) est bien un développement limité de la transformée de Fourier $E[e^{i\zeta\cdot U(1,\varepsilon,x)}]$ uniforme sur K.

Il reste à intégrer ce développement en ζ, ce qui est, comme en [4], rendu possible par le calcul de Malliavin et, ici, le lemme 3.14.

En effet l'intégration par parties du calcul de Malliavin (cf [12], [13]) montre que, par (3.23) et le lemme(3.14) on a : $\forall p\ge 1$

$$\sup_{\varepsilon\le\varepsilon_0, x\in K}|E[e^{i\zeta\cdot U(1,0,x)}\partial_\varepsilon^k\chi(\rho,\zeta)]| \le C_p|\zeta|^{-2p}|\zeta|^k. \tag{3.28}$$

D'où l'on tire que, $\forall p\ge 1$:

$$\sup_{x\in K}|\alpha_k(\zeta,x)| \le C|\zeta|^{k-2p} \tag{3.29}$$

et que

$$\sup_{\varepsilon\le\varepsilon_0, x\in K}|T_{N+1}(\varepsilon,x,\zeta)| \le C|\zeta|^{N+1-2p}, \tag{3.30}$$

et donc que l'on peut intégrer en ζ le développement (3.24) et obtenir :

$$q_1(\varepsilon,x,0) = \varepsilon^{-Q(x_0)}\left[\sum_{0\le k\le N}\gamma_k(x)\varepsilon^k + \varepsilon^{N+1}s_{N+1}(\varepsilon,x)\right] \tag{3.31}$$

où :

(3.32)

$$\gamma_k(x) = (2\pi)^{-d} \int \alpha_k(\zeta, x) d\zeta \text{ et } s_{N+1}(\varepsilon, x) = (2\pi)^{-d} \int T_{N+1}(\varepsilon, x, \zeta) d\zeta.$$

On a ainsi, par (3.30) :

$$\sup_{\varepsilon \leq \varepsilon_0, x \in K} |s_{N+1}(\varepsilon, x)| < \infty. \tag{3.33}$$

Les coefficients $\alpha_k(\zeta, x)$ sont clairement continus en x sur K, il en est donc de même pour les $\gamma_k(x)$ par (3.29).

De plus il est facile de vérifier que :

$$\forall k \text{ impair } \gamma_k(x) = 0. \tag{3.34}$$

En effet la symétrie de la loi du mouvement brownien montre que :

(3.35) $E[e^{i\zeta \cdot U(1,\varepsilon,x)}] = E[e^{i(T_{-1}\zeta)\cdot U(1,-\varepsilon,x)}]$ où T_{-1} est donné en (3.10)

d'où l'on tire, par (3.24) que :

$$\alpha_k(\zeta, x) = (-1)^k \alpha_k(T_{-1}(\zeta), x) \tag{3.36}$$

et donc, par (3.32), que :

$$\gamma_k(x) = (-1)^k \gamma_k(x) \tag{3.37}$$

d'où (3.34).

Il est alors facile de terminer la preuve du 1) et du 2) du théorème : Soit $\overline{p}_t(\varepsilon, x, y)$ la densité de la loi de $\overline{x}_t^\varepsilon(x)$, on a de façon évidente :

$$\overline{p}_t(\varepsilon, x, y) = q_t(\varepsilon, x, \varphi_x^{-1}(y)) |\det(d\varphi_x^{-1}(y))| \tag{3.38}$$

et donc :

$$\overline{p}_1(\varepsilon, x, x) = q_1(\varepsilon, x, 0) |\det(d\varphi_x^{-1}(0))|. \tag{3.39}$$

Ainsi on a obtenu, par (3.31), un développement asymptotique de $\overline{p}_1(\varepsilon, x, x)$:

$$\overline{p}_1(\varepsilon, x, x) = \varepsilon^{-Q(x_0)} \left[\sum_{0 \leq k \leq N} c_k(x) \varepsilon^k + \varepsilon^{N+1} r_{N+1}(\varepsilon, x) \right] \tag{3.40}$$

avec :

(3.41)

$$c_k(x) = \gamma_k(x) |\det(d\varphi_x^{-1}(0))| \text{ et } r_{N+1}(\varepsilon, x) = s_{N+1}(\varepsilon, x) |\det(d\varphi_x^{-1}(0))|.$$

Enfin, grâce au lemme de localisation (2.4) on sait que :

$$\sup_{x\in K} |p_1(\varepsilon,x,x) - \overline{p}_1(\varepsilon,x,x)| \leq Ke^{-C/\varepsilon^2} \tag{3.42}$$

en effet

$$p_1(\varepsilon,x,x) - \overline{p}_1(\varepsilon,x,x) = (2\pi)^{-d}\int E[e^{i\xi\cdot(x_1^\varepsilon - x)} - e^{i\xi\cdot(\overline{x}_1^\varepsilon - x)}]d\xi,$$

et le lemme (2.4) donne (3.42).

Enfin, pour terminer la preuve du 1) et du 2) du théorème (à l'exception de (1.10) i.e. la positivité du premier coefficient qui sera démontrée au paragraphe 7), il suffit de remarquer que $p_{\varepsilon^2}(x,x)$ est la densité de la loi de la variable $x_{\varepsilon^2}(x)$ et que par scaling cette densité est égale à celle de $x_1^\varepsilon(x)$ c'est-à-dire à $p_1(\varepsilon,x,x)$.

Pour démontrer le 3) du théorème il suffit d'utiliser le fait, vu en (3.20), qui, si l'on ajoute l'hypothèse que $A(x_0)$ est d'intérieur non vide, on sait que $\chi(\varepsilon,x,\zeta)$ est C^∞ en (ε,x).

Le théorème de dérivation sous le signe somme permet simplement de montrer que c_k est C^∞ en x au voisinage de x_0. En effet il est facile de vérifier, par (3.25) et (3.13), que $\alpha_k(\zeta,x)$ est C^∞ en x au voisinage de x_0 et que, par le lemme (3.14) :

$$\sup_{x\in K} E[|(\partial_x)^\gamma \alpha_k(\zeta,x)|] \leq C_p|\zeta|^{k+|\gamma|-2p} \tag{3.43}$$

pour tout multi-indice γ et tout $p \geq 1$. Ainsi $\gamma_k(x) = (2\pi)^{-d}\int \alpha_k(\zeta,x)d\zeta$ est C^∞ en x au voisinage de x_0 et donc aussi c_k.

Ce qui achève la preuve du 3) du théorème.

4. Relèvement de Rotschild et Stein.

Nous allons ici rappeler les résultats de relèvement sur un groupe nilpotent de Rotschild et Stein [11], qui permettront aux §5 et 6 de démontrer le lemme 3.6 et donc d'achever la preuve du théorème. Dans la suite on notera r l'entier $r(x_0)$ défini en (1.5).

Soit F l'algèbre de Lie libre à m générateurs $y_1,\dots,y_m$, et soit $g(m,r) = F/F^{r+1}$. Alors, par définition $g(m,r)$ est l'algèbre de Lie libre nilpotente de pas r à m générateurs $Y_1,\dots,Y_m$, avec $Y_k = y_k + F^{r+1}$.

Nous identifierons $g(m,r)$ et le groupe nilpotent simplement connexe associé $N(m,r)$, qui n'est autre que $g(m,r)$ muni de la multiplication donnée par la série de Campbell-Hausdorff.

Introduisons la graduation naturelle de $g(m,r)$:

$$g(m,r) = V_1 \oplus \cdots \oplus V_r \tag{4.1}$$

avec :

$$V_k = \text{Vect}\ \{Y^J, |J| = k\}. \tag{4.2}$$

Cette graduation définit naturellement une famille de dilatations δ_t :

$$\delta_t\Big(\sum_{1\leq i\leq r} u_i\Big) = \sum_{1\leq i\leq r} t^i u_i \text{ avec } u_i \in V_i. \tag{4.3}$$

On dira qu'un opérateur différentiel P sur $g(m,r)$ (ou $N(m,r)$) est homogène de degré k si :

$$P(f\circ\delta_t) = t^k P(f)\circ\delta_t, \text{ pour tous } t>0 \text{ et } f\in C^\infty(g(m,r)).$$

Un opérateur différentiel P sera dit de poids $\leq k$ si chacun des termes du développement de Taylor en 0 de ses coefficients est homogène de degré $\leq k$. On sait alors (cf [11]) que si on note :

$$d(m,r) = \dim g(m,r) \text{ et } n = d(m,r) - d \tag{4.4}$$

on a :

Lemme (4.5). — *Il existe des champs de vecteurs $\widetilde{X}_i$ sur un voisinage de $(x_0,0)$ dans $\mathbf{R}^d\times\mathbf{R}^n = \mathbf{R}^{d(m,r)}$ tels que :*

$$\widetilde{X}_i(x,h_1,\dots,h_n) = X_i(x) + \sum_{1\leq k\leq n} a_i^k(x,h_1,\dots,h_{k-1})a_{h_k} \tag{4.6}$$

$$\{\widetilde{X}^J(x_0,0), |J|\leq r\} \text{ engendre } \mathbf{R}^d\times\mathbf{R}^n. \tag{4.7}$$

Étendons les champs $\widetilde{X}_i$ à l'espace $\mathbf{R}^d\times\mathbf{R}^n$ entier en prolongeant de façon C^∞ les fonctions a_i^k par 0 hors d'un voisinage de $(x_0,0)$ et en posant :

$$\widetilde{X}_i = \overline{X}_i + \sum a_i^k \partial_{h_k}.$$

Par (4.6) la famille $\{\widetilde{X}^J(x_0,0), J \in B\}$ est libre, puisque la famille $\{X^J(x_0), J \in B\}$ l'est. Complétons-la en une base de $\mathbf{R}^d \times \mathbf{R}^n$: $\{\widetilde{X}^K(x_0,0), K \in A\}$ où $B \subset A \subset \cup_{1\leq k\leq r}\{1,\dots,m\}^k$. Il est clair que la famille $\{\widetilde{X}^K(\xi), K \in A\}$ est toujours une base de $\mathbf{R}^d \times \mathbf{R}^n$ pour ξ assez voisin de $(x_0,0)$. Alors la famille $\{Y^K, K \in A\}$ est une base de $g(m,r)$ et l'application $\Theta_\xi(\eta) = \exp\left(\sum_{K\in A} u_K Y^K\right)$ où $\eta = \exp\left(\sum_{K\in A} u_K \widetilde{X}^K\right)(\xi)$ définit un difféomorphisme d'un voisinage de $(x_0,0)$ sur un voisinage de 0 dans $g(m,r)$. On a alors :

LEMME (4.8). — *$\Theta_\xi * \widetilde{X}_i = Y_i + R_i^\xi$ $\forall 1 \leq i \leq m$ où R_i^ξ est un champ de vecteurs de poids ≤ 0, variant de façon C^∞ avec ξ, (et où Y_i désigne, par un abus évident de notation, le champ de vecteurs invariant à gauche sur $g(m,r) \simeq N(m,r)$ défini par le générateur Y_i de $g(m,r)$).*

5. Étude de la diffusion relevée.

Considérons la solution $\widetilde{x}_t^\varepsilon(x)$ sur $\mathbf{R}^d \times \mathbf{R}^n$ de l'équation de Stratonovitch :

$$d\widetilde{x}_t^\varepsilon = \varepsilon \sum_{1\leq i\leq m} \widetilde{X}_i(\widetilde{x}_t^\varepsilon) \circ dw_t^i \text{ avec } \widetilde{x}_0^\varepsilon = (x,0) \text{ et } x \in K \tag{5.1}$$

(où K est toujours un compact inclus dans $A(x_0) \cap U$).

Nous allons étudier le développement de Taylor de $\widetilde{x}_t^\varepsilon(x)$ en $\varepsilon = 0$ dans la carte donnée par $\Theta_{(x,0)}$, c'est-à-dire le développement de Taylor de $z_t^\varepsilon(x) = \Theta_{(x,0)}(\widetilde{x}_t^\varepsilon(x))$.

Pour cela écrivons l'équation satisfaite par $z_t^\varepsilon(x)$:

$$dz_t^\varepsilon = \varepsilon \sum_{1\leq i\leq m} \Theta_{(x,0)*}\widetilde{X}_i(z_t^\varepsilon) \circ dw_t^i \text{ avec } z_0^\varepsilon = 0, \tag{5.2}$$

ou encore :

$$dz_t^\varepsilon = \varepsilon \sum_{1\leq i\leq m} Y_i(z_t^\varepsilon) \circ dw_t^i + R_i^{(x,0)}(z_t^\varepsilon) \circ dw_t^i. \tag{5.3}$$

Comme au lemme (3.7) on a :

LEMME (5.4). — *1) $z_t^\varepsilon(x)$ est une fonction C^∞ de (ε, x). 2) $\forall(j,k,n) \in \mathbb{N}^3$, pour tout multi-indice $\alpha \in \{1,\dots,d\}^j$, et pour tout $q \geq 2$, il existe $\varepsilon_0 > 0$ tel que :*

$$\text{(5.5)} \qquad \sup_{\varepsilon \leq \varepsilon_0, x \in K} \sum_{i \leq n} E[\sup_{0 \leq t \leq 1} \|D^i(\partial_x)^\alpha(\partial_\varepsilon)^k(z_t^\varepsilon(x))\|_{HS}^q] < \infty.$$

Enfin on peut préciser le développement de Taylor de z_t^ε en $\varepsilon = 0$:

LEMME (5.6). — *Pour tout $k \geq 0$, on a :*

$$(\partial_\varepsilon)^k z_t^\varepsilon(x)_{|\varepsilon=0} \in \oplus_{1 \leq j \leq k} V_j.$$

Preuve. — Il suffit de raisonner par récurrence sur le système stochastique triangulaire donnant les $(\partial_\varepsilon)^k z^\varepsilon_{t\,|\varepsilon=0}$ (obtenu naturellement en posant $\varepsilon = 0$ dans les équations aux variations stochastiques donnant les dérivées en ε de z_t^ε) en utilisant le fait que les champs Y_i soient homogènes de degré 1 et que les champs R_i soient de poids ≤ 0. On peut donner une preuve rapide en considérant la diffusion $g_t^\varepsilon = \delta_{1/\varepsilon}(z_t^\varepsilon)$. Écrivons l'équation satisfaite par g_t^ε :

$$\text{(5.7)} \qquad dg_t^\varepsilon = \sum_{1 \leq i \leq m} Y_i(z_t^\varepsilon) \circ dw_t^i + \varepsilon S_i(\varepsilon, z_t^\varepsilon) \circ dw_t^i,$$

avec

$$\text{(5.8)} \qquad S_i = \delta_{1/\varepsilon} * R_i.$$

A priori g_t^ε n'est défini que pour $\varepsilon > 0$, mais du fait que les champs R_i sont de poids ≤ 0, les fonctions $\varepsilon, x \to S_i(\varepsilon, x)$ admettent un prolongement C^∞ en $\varepsilon = 0$. Ceci montre que g_t^ε admet un prolongement C^∞ en $\varepsilon = 0$ ce qui montre le lemme (5.6).

De plus il est clair que $g_t^{(0)}$ est la solution de l'équation invariante :

$$\text{(5.9)} \qquad dg_t^{(0)} = \sum_{1 \leq i \leq m} Y_i(g_t^{(0)}) \circ dw_t^i \text{ avec } g_0^{(0)} = 0.$$

Du fait que les Y_i sont des générateurs de $g(m,r)$ on sait aussi que la matrice de covariance de Malliavin de $g_t^{(0)}$ (cf [12], [13]) est non dégénérée, i.e :

$$\text{(5.10)} \qquad E(\det\, (Dg_t^{(0)}, Dg_t^{(0)})^{-p}) < \infty.$$

Ceci nous permet d'obtenir le développement de Taylor stochastique de $\widetilde{x}_t^\varepsilon(x)$ dans la carte donnée par $\Theta_{(x,0)}$, i.e. par la base $\{X^K(x,0), K \in A\}$.

Notons :

(5.11)
$$\widetilde{x}_t^\varepsilon(x) = \exp\Big(\sum_{K\in A} u_K(t,x,\varepsilon)X^K\Big)(x,0).$$

On a par (5.6) le développement de Taylor des coordonnées de $\widetilde{x}_t^\varepsilon$:

(5.12)
$$u_K(t,x,\varepsilon) = \sum_{|K|\leq j\leq N} \varepsilon^j/j!(\partial_\varepsilon)^j u_K(t,x,\varepsilon)_{|\varepsilon=0} + \varepsilon^{N+1}R_{K,N+1}(t,x,\varepsilon))$$

avec

(5.13)
$$R_{K,N+1}(t,x,\varepsilon) = \int_{0\leq v\leq 1} (\partial_\varepsilon)^{N+1}u_K(t,x,\varepsilon v)(1-v)^N/N!dv.$$

De plus on connait explicitement le terme principal dans le développement (5.12), à savoir les dérivées "diagonales" des u_K : $((\partial_\varepsilon)^{|K|}u_K(t,x,\varepsilon)_{|\varepsilon=0})_{K\in A}$.

En effet on a :

(5.14)
$$g_t^{(0)} = \exp\sum_{K\in A}(\partial_\varepsilon)^{|K|}u_K(t,x,\varepsilon)_{|\varepsilon=0}/|K|!Y^K).$$

Ceci montre d'abord que ces dérivées sont indépendantes de x, ce qui est évident car le propre du relèvement de Rotschild et Stein est de construire une approximation, au premier ordre "gradué", qui soit universelle.

De plus ceci montre que, si on note M_t le vecteur constitué de ces dérivées "diagonales" des u_K en $\varepsilon = 0$, c'est-à-dire :

(5.15)
$$M_t = ((\partial_\varepsilon)^{|K|}u_K(t,x,\varepsilon)_{|\varepsilon=0})K \in A$$

alors, par (5.10), la matrice de covariance de Malliavin de M_t est inversible pour $t > 0$, i.e

(5.16)
$$E[|\det\,(DM_t, DM_t)|^{-p}] < \infty,\ \forall p \geq 1.$$

6. Preuve des lemmes 3.9 et 3.14.

Nous allons ici achever la preuve du théorème en démontrant les lemmes (3.9) et (3.14).

Pour cela il suffit maintenant d'utiliser les résultats du §5 sur la diffusion relevée $\widetilde{x}_t^\varepsilon$ et de les redescendre de $\mathbf{R}^d \times \mathbf{R}^n \simeq g(m,r)$ sur $\mathbf{R}^d$. Il est clair, par la forme (4.6) des champs $\widetilde{X}_i$, que si p désigne la projection de $\mathbf{R}^d \times \mathbf{R}^n$ sur $\mathbf{R}^d$, on a :

$$\overline{x}_t^\varepsilon(x) = p(\widetilde{x}_t^\varepsilon(x)) = p[\exp\Big(\sum_{K\in A} u_K(t,\varepsilon)\widetilde{X}^K\Big)(x,0)] \tag{6.1}$$

$$\overline{x}_t^\varepsilon = \exp\Big(\sum_{K\in A} u_K(t,x,\varepsilon)X^K\Big)(x). \tag{6.2}$$

Pour obtenir le développement de Taylor gradué de $\overline{x}_t^\varepsilon$ donné par le lemme (3.9) à partir de celui de $\widetilde{x}_t^\varepsilon$ obtenu au §5, il reste à supprimer (par dépendance linéaire) dans (6.2) les champs X^K avec $K \in A\backslash B$.

Considérons l'application ψ_x définie par :

$$\psi_x((u_K)_{K\in A}) = \exp\Big(\sum_{K\in A} u_K X^K\Big)(x). \tag{6.3}$$

L'application $\varphi_x^{-1} \circ \psi_x$, où φ_x est définie en (3.2), associe à $(u_K)_{K\in A}$ l'unique $(v_J)_{J\in B}$ tel que :

$$\exp\Big(\sum_{K\in A} u_K X^K\Big)(x) = \exp\Big(\sum_{J\in B} v_J X^J\Big)(x). \tag{6.4}$$

Ainsi on a :

$$v_J(t,x,\varepsilon) = (\varphi_x^{-1} \circ \psi_x)((u_K(t,x,\varepsilon))_{K\in A}) \tag{6.5}$$

où $v_J(t,x,\varepsilon)$ est introduit en (3.5) et $(u_K(t,x,\varepsilon))_{K\in A}$ l'est en (5.11). Puisque l'on sait, par le §5, que : $(\partial_\varepsilon)^j u_{K\,|\varepsilon=0} = 0$, si $j < |K|$, pour démontrer le lemme (3.9) il suffit de vérifier le lemme suivant.

LEMME (6.6). — *Pour tout $x \in K$, pour tout entier p, et toute famille de multi-indices $(K_1,\ldots,K_p)$ avec $K_i \in A$ et $\sum_{1\le i\le p} |K_i| < |J|$ où J est un multi-indice de B, on a :*

$$[\partial^p(\varphi_x^{-1} \circ \psi_x)/\partial u_{K1}\ldots\partial u_{K_p}]_{|u=0} = 0.$$

Puisque les $(X^J)_{J\in B}$ forment une base au voisinage de x_0, il existe des fonctions C^∞, définies au voisinage de $x_0, (d_J^K)_{J\in B, K\in A}$ telles que :

$$X^K = \sum_{J\in B} d_J^K X^J. \tag{6.7}$$

On vérifie alors simplement que :

$$[\partial(\varphi_x^{-1}\circ\psi_x)_J/\partial u_K]_{|u=0} = d_J^K(x) \tag{6.8}$$

$$[\partial^2(\varphi_x^{-1}\circ\psi_x)/\partial u_{K1}\partial u_{K2}]_{|u=0} = X^{K_1}(d_J^{K_2})(x) + X^{K_2}(d_J^{K_1})(x) \tag{6.9}$$

et, en itérant, que : $[\partial^p(\varphi_x^{-1}\circ\psi_x)_J/\partial u_{K_1}\dots\partial u_{K_p}]_{|u=0}$ est une combinaison linéaire des $X^{K_{\sigma(1)}}\dots X^{K_{\sigma(p-1)}}(d_J^{K_{\sigma(p)}})(x)$ avec $\sigma \in \sum_p$ (le groupe symétrique d'ordre p). Or la base $\{X^J(x), J\in B\}$ étant triangulaire au sens de (3.1) on a :

$$d_J^K(x) = 0 \text{ si } |K| < |J| \tag{6.10}$$

et, de même :

$$X^{K_1}\dots X^{K_{p-1}}(d_J^{K_p})(x) = 0 \text{ si } \sum_{1\le i\le p} |K_i| < |J|. \tag{6.11}$$

En effet, si l'on calcule le crochet $[X^{K_1}, X^{K_2}]$ on a, par (6.7) :

$$[X^{K_1}, X^{K_2}] = \sum_{J\in B} d_J^{K_1}[X^J, X^{K_2}] - X^{K_2}(d_J^{K_1})X^J \tag{6.12}$$

d'où, par (6.10) :

$$\Big(\sum_{J\in B} X^{K_2}(d_J^{K_1})X^J\Big)(x) \in C_{|K_1|+|K_2|}(x) \tag{6.13}$$

et donc

$$X^{K_2}(d_J^{K_1})(x) = 0 \text{ si } |K_1| + |K_2| < |J|, \tag{6.14}$$

toujours par le fait que la base est triangulaire. De même on a, par récurrence :

$$\begin{aligned}(6.15)\quad & [X^{K_1}, [X^{K_2}, \dots [X^{K_{p-1}}, X^{K_p}]\dots](x) \\ & - \sum_{J\in B} X^{K_1}\cdot X^{K_2}\dots X^{K_{p-1}}(d_J^{K_p})(x)X^J(x) \in C_{|K_1|+\dots+[k_p||}(x)\end{aligned}$$

et donc

(6.16) $$X^{K_1}X^{K_2}\dots X^{K_{p-1}}(d_J^{K_p})(x) = 0 \text{ si } \sum_{1\le i\le p} |K_i| < |J|.$$

Ce qui prouve 6.11.

Ceci achève la preuve du lemme (3.9).

Remarque. — Si $A(x_0)$ est d'intérieur vide (i.e. si la base $\{X^J(x), J \in B\}$ reste triangulaire au voisinage de x_0) on a : $d_J^K \equiv 0$ sur un voisinage de x_0, pour $|K| < |J|$. L'égalité (6.11) est alors évidente.

Passons enfin à la preuve du lemme (3.14) :

On a vu que :

(6.17) $$U_J(t,\varepsilon,x) = \varepsilon^{-|J|}(\varphi_x^{-1}\circ\psi_x)_J((u_K(t,\varepsilon,x))_{K\in A});$$

d'où l'on tire la valeur de la limite lorsque $\varepsilon \to 0$:

(6.18) $$U_J(t,0,x) = \sum_{K\in A,|K|=|J|} \partial(\varphi_x^{-1}\circ\psi_x)_J/\partial u_{K|u=0}u_K^{(0)}$$

(car on a vu que $\partial(\varphi_x^{-1}\circ\psi_x)_J/\partial u_{K|u=0} = 0$ si $|K| < |J|$ et que les termes correspondant aux K tels que $|K| > |J|$ n'interviennent évidemment pas puisqu'ils sont affectés d'un $\varepsilon^{|K|}$). Or on a aussi vu en (6.7) que :

(6.19) $$\partial(\varphi_x^{-1}\circ\psi_x)_J/\partial u_{K|u=0} = d_J^K(x);$$

d'où on obtient :

(6.20) $$U_J(t,\varepsilon,x) = \sum_{K\in A,|K|=|J|} d_J^K(x)(\partial_\varepsilon)^{|K|}u_K(t,x,\varepsilon)_{|\varepsilon=0}.$$

Or la matrice $\widetilde{D}(x) = (\widetilde{d}_J^K(x))_{J\in B,K\in A}$, à d lignes et $d+n(=d(m,r))$ colonnes, telle que :

(6.21) $\widetilde{d}_J^K(x) = d_J^K(x)$ si $|J| = |K|$ et $\widetilde{d}_J^K(x) = 0$ sinon

est de rang maximal, à savoir d (car $\{X^J(x), J \in B\}$ est une base triangulaire).

Ainsi :

(6.22) $$U(t,0,x) = \widetilde{D}(x)\cdot M_t$$

où M_t (introduit en (5.15)) a une matrice de covariance de Malliavin inversible pour $t > 0$ (par (5.16)) et $D(x)$ est une matrice de rang plein. On en déduit clairement que la matrice de covariance de Malliavin de $U(t,0,x)$ est inversible pour $t > 0$, ce qui prouve le lemme (3.14).

7. Étude du premier coefficient.

Nous allons commencer ici par démontrer (1.10), c'est-à-dire la stricte positivité du premier coefficient du développement (1.8), ce qui assure que (1.8) fournit bien un équivalent de $p_t(x,x)$. Pour cela il suffit bien sûr (par (3.41)) de démontrer la stricte positivité de $\gamma_0(x)$.

Pour cela commençons par remarquer que, par le lemme (3.14), on sait que, pour $x \in A(x_0) \cap U$, la loi de la variable aléatoire $U(1,0,x)$ admet une densité que nous noterons f_x, qui est C^∞ et même à décroissance rapide. Par inversion de Fourier on a donc :

$$(7.1) \qquad f_x(0) = (2\pi)^{-d} \int e^{i\zeta \cdot z} f_x(z) dz d\zeta = (2\pi)^{-d} \int E(e^{i\zeta \cdot U(1,0,x)}) d\zeta,$$

et donc, par (3.25) et (3.32), on obtient l'expression de γ_0 :

$$(7.2) \qquad \gamma_0(x) = f_X(0) \text{ et donc } c_0(x) = f_x(0) |\det\, (d\varphi_x^{-1}(0))|.$$

Ainsi montrer (1.9) revient à montrer que la densité de la loi de $U(1,0,x)$ est strictement positive en 0. Pour cela rappelons le fait, vu en (6.22) que :

$$(7.3) \qquad U(t,0,x) = \widetilde{D}(x) \cdot M_t.$$

Dans (7.3) $\widetilde{D}(x)$ est une matrice rectangulaire $d \times d+n$ de rang plein (à savoir d), ainsi pour vérifier que la densité en 0 de la loi de $U(t,0,x)$ est non nulle il suffit de vérifier que la densité de la loi de M_t est non identiquement nulle sur le noyau de l'application linéaire associée à $\widetilde{D}(x)$. Nous allons ici vérifier que la densité de la loi de M_t est non nulle en 0, ce qui prouvera donc (1.10). Pour cela rappelons que :

$$(7.4) \qquad M_t = ((\partial_\varepsilon)^{|K|} u_K(t,x,\varepsilon)_{|\varepsilon=0})_{K\in A}$$

et que

$$(7.5) \qquad g_t^{(0)} = \exp \sum_{K\in A} (\partial_\varepsilon)^{|K|} u_K(t,x,\varepsilon)_{|\varepsilon=0} / |K|! Y^K).$$

Ainsi montrer que la loi de M_t a une densité non nulle revient à démontrer que la loi de la diffusion invariante $g_t^{(0)}$ de générateur $1/2 \sum Y_i^2$ sur le groupe $N(m,r)$ (où l'algèbre $g(m,r)$) a une densité non nulle en 0. Or ce dernier point est évident car la densité $q_t(0,z)$ de la loi de $g_t^{(0)}$ est exactement le noyau de la chaleur associé à l'opérateur $1/2 \sum Y_i^2$.

Un raisonnement d'homogénéité évident montre en effet que :

$$q_t(0,0) = t^{-Q/2} q_1(0,0) \tag{7.6}$$

où Q est le grade de l'algèbre nilpotente graduée $g(m,r)$: $Q = \sum_{1 \leq k \leq r} k \dim V_k$. Ainsi si $q_t(0,0)$ s'annule pour un $t, q_t(0,0)$ est nul pour tout $t > 0$. Ce qui est impossible.

Ainsi la méthode précédente pour prouver le résultat de positivité du premier coefficient c_0 a consisté à se ramener au même résultat sur le groupe nilpotent.

Nous allons maintenant décrire plus "explicitement" le coefficient c_0. Pour cela nous allons rappeler les résultats de [2] qui donnent l'expression explicite de la diffusion invariante $g_t^{(0)}$ en fonction des crochets Y^K et des intégrales stochastiques de Stratonovitch itérées :

LEMME (7.7). — (théorème 13 de [2]) *Si J est le multi-indice $(j_1, \ldots, j_k)$, notons β_J le terme k-homogène de degré 1 en chaque variable dans la série de Campbell-Hausdorff à k variables $H(Y_{j_1}, \ldots, Y_{j_k})$, c'est-à-dire le coefficient de $s_1 s_2 \cdots s_k$ dans la série $H(x_1 Y_{j_1}, \ldots, s_k Y_{j_k})$. On a alors :*

$$g_t^{(0)} = \sum_{1 \leq k \leq r} \sum_{|J|=k} \beta_J \int \cdots \int_{0 \leq t_1 \leq \cdots \leq t_k \leq t} dw_{t_1}^{j_1} \circ \cdots \circ dw_{t_k}^{j_k}.$$

Ce lemme peut aussi être écrit autrement (théorème 16 de [2]) :

LEMME (7.8). — $g_t^{(0)} = \sum_{1 \leq k \leq r} \sum_{|L|=k} \alpha_L(t) Y^L$ avec

$$\alpha_L(t) = \sum_{\sigma \in \Sigma_k} c(\sigma) \int \cdots \int_{0 \leq t_1 \leq \cdots \leq t_k \leq t} dw_{t_1}^{1_{\sigma(1)}} \circ \cdots \circ dw_{t_k}^{1_{\sigma(k)}}$$

où Σ_k désigne le groupe symétrique d'ordre k et où les $c(\sigma)$ sont des constantes universelles définies à partir des la série de Campbell-Hausdorff ci-dessous.

Considérons la série de Campbell-Hausdorff à k termes $H(U_1, \ldots, U_k)$. Pour un multi-indice J donné le crochet U^J apparaît plusieurs fois dans cette série. Regroupons les termes où U^J apparaît et notons c_J son coefficient, de telle sorte que :

$$H(U_1,\ldots,U_k)=\sum_{m\geq 1}\sum_{|J|}c_J U^J. \tag{7.9}$$

Soit alors une permutation $\sigma\in\Sigma_k$, on vérifie (cf proposition 4 de [2]) que le coefficient c_J associé à un multi-indice J tel que :

$$j_{\sigma(1)}<j_{\sigma(2)}<\cdots j_{\sigma(k)} \tag{7.10}$$

est indépendant de J et ne dépend que de σ. On notera $c(\sigma)$ ce coefficient.

Notons (c_K^L) les constantes de structures du groupe $N(m,r)$ définies par :

$$Y^L=\sum_{K\in A,|K|=|L|}c_K^L Y^K. \tag{7.11}$$

On obtient ainsi l'expression de la diffusion invariante $g_t^{(0)}$ sur la base $\{Y^K, K\in A\}$:

$$g_t^{(0)}=\sum_{K\in A}\Big(\sum_{|L|=|K|}c_K^L\alpha_L(t)\Big)Y^K; \tag{7.12}$$

ce qui donne les termes principaux du développement de Taylor stochastique de $x_t^\varepsilon(x)$:

$$x_t^\varepsilon(x)=\exp\Big[\sum_{J\in B}\Big\{\sum_{K\in A,|K|=|J|}d_J^K(x)\Big(\sum_{|L|=|K|}c_K^L\alpha_L(t) +O(\varepsilon^{|J|+1})\Big)\Big\}X^J\Big](x), \tag{7.13}$$

c'est-à-dire la valeur explicite de $U(t,0,x)$:

$$U(t,0,x)=\Big(\sum_{K\in A,|K|=|J|}d_J^K(x)\sum_{|L|=|K|}c_K^L\alpha_L(t)\Big)_{J\in B}. \tag{7.14}$$

On obtient ainsi le résultat explicite :

THÉORÈME (7.15). — *Le premier coefficient $c_0(x)$ du développement asymptotique (1.7) est donné par :*
$c_0(x)=|\det \mathrm{d}\,\varphi_x^{-1}(O)|f_x(0)$ *où f_x est la densité de la loi de la variable*

aléatoire : $U(1,0,x) = \left(\sum_{K\in A, |K|=|J|} d_J^K(x) \sum_{|L|=|K|} c_K^L \alpha_L(1) \right)_{J\in B}$ *où les fonctions* d_J^K *sont définies par :* $X^K = \sum_{J\in B} d_J^K X^J, \forall K \in A$*, où les constantes* c_K^L *sont universelles, ce sont les constantes de structure du groupe* $N(m,r)$ *définies par :* $Y^L = \sum_{K\in A, |K|=|J|} c_K^L Y^K$*, pour* $|L| \leq r$*; et où, enfin les* $\alpha_L(1)$ *sont des variables aléatoires universelles définies par :*

$$\alpha_L(t) = \sum_{\sigma\in\Sigma_k} c(\sigma) \int \cdots \int_{0\leq t_1\leq\cdots\leq t_k\leq t} dw_{t_1}^{1_{\sigma(1)}} \circ \cdots \circ dw_{t_k}^{1_{\sigma(k)}}$$

où les constantes $c(\sigma)$ *sont universelles et définies à partir de la série de Campbell-Hausdorff (leur définition précise est donnée après (7.10)).*

Remarque. — 1) Le résultat précédent donne évidemment, dans le cas trivial où les champs X_i forment une base de $\mathbf{R}^d$ (et où nécessairement $m = d, r = 1$), c'est-à-dire dans le cas elliptique, le fait connu que :

$$c_0(x) = (2\pi)^{-d/2} |\det\, (d\varphi_x^{-1}(0))|.$$

En effet dans ce cas le groupe $N(m,1)$ est bien abélien et la variable aléatoire $U(1,0,x)$ se réduit au mouvement brownien w_1 au temps 1. La densité en 0 de sa loi vaut donc $(2\pi)^{-d/2}$.

2) Dans le cas où $r = 2$, le théorème (7.15) permet de calculer explicitement le coefficient $c_0(x)$ en utilisant la méthode employée par Gaveau et fondée sur la formule de l'aire de Paul Lévy.

3) Par contre dans les cas où $r \geq 3$, le calcul devient difficile, même dans une situation invariante sur un groupe libre nilpotent, car on ne dispose plus de cette formule de Paul Lévy pour le calcul de lois d'intégrales itérées d'ordre ≥ 3.

BIBLIOGRAPHIE

[1] R. AZENCOTT et Al., Géodésiques et diffusions en temps petit, Astérisque, 84-85 (1981).

[2] G. BEN AROUS, Flots et séries de Taylor stochastiques, Journal of Probability Theory and Related Fields, 81 (1989), 29-77.

[3] G. BEN AROUS, Méthode de Laplace et de la phase stationnaire sur l'espace de Wiener, Stochastics, vol 25 (1988), 125-153.

[4] G. BEN AROUS, Développement asymptotique du noyau de la chaleur hypoelliptique hors du cut-locus, Annales Scientifiques de l'École Normale Supérieure, 4e série, 21 (1988), 307-331.

[5] D.S. JERISON, A. SANCHEZ, Estimates for the heat kernel for the sum of squares of vector fields, Indiana University Mathematics Journal, Volume 35-4 (1986), 835-854.

[6] B. GAVEAU, Principe de moindre action, propagation de la chaleur, estimées sous-elliptiques sur certains groupes nilpotents, Acta Mathematica, 139 (1977), 96-153.

[7] S. KUSOAKA, D.W. STROOCK, Applications od the Malliavin calculus. Part 1, Proceedings of the conference at Katata (1982), Kinokuniya publishing Co. Tokyo and New York.

[8] R. LÉANDRE, Majoration en temps petit de la densité d'une diffusion dégénérée, à paraître.

[9] R. LÉANDRE, Minoration en temps petit de la densité d'une diffusion dégénérée, Journal of Functional Analysis, vol 74 (oct. 1987), 399-414.

[10] R. LÉANDRE, Développement asymptotique de la densité de diffusions dégénérées, Journal of Probability Theory and Related Fields, vol 76 (1987), 341-358.

[11] L.P. ROTSCHILD, E.M. STEIN, Hypoelliptic differential operators and nilpotent groups, Acta Mathematica, 137 (1976), 247-320 & Comm. in pure and applied Maths, 20 (1967), 659-685.

[12] D.W. STROOCK, Some applications of stochastic calculus to partial differential equations. École d'été de Saint-Flour XI-1981, Lectures notes in Mathematics, n° 976, Springer-Verlag.

[13] S. WATANABE, Stochastic differential equations and Malliavin calculus, Published for the Tata institute, Bombay (1984), Springer-Verlag.

Manuscrit reçu le 3 mars 1987,
révisé le 18 mars 1988.

Gérard BEN AROUS,
École Normale Supérieure
45 rue d'Ulm,
75230 Paris cedex 05.