SÉMINAIRE DE PROBABILITÉS (STRASBOURG)

DOMINIQUE BAKRY
MICHEL ÉMERY

Diffusions hypercontractives

Séminaire de probabilités (Strasbourg), tome 19 (1985), p. 177-206
<http://www.numdam.org/item?id=SPS_1985__19__177_0>

DIFFUSIONS HYPERCONTRACTIVES

par D. Bakry et M. Émery

La notion d'hypercontractivité d'un semi-groupe d'opérateurs est née il y a une quinzaine d'années, dans des articles de théorie quantique des champs (Glimm, Gross, Nelson, Simon). Nous n'en aborderons que l'aspect probabiliste, qui décrit le comportement de semi-groupes de noyaux markoviens ; notre outil essentiel sera l'opérateur " carré du champ itéré " (pour les intimes : le Γ_2), déjà utilisé par le premier auteur dans un travail sur les transformations de Riesz (dans ce volume). Notre résultat principal est une condition suffisante (une forme forte de positivité de Γ_2) pour qu'un semi-groupe soit hypercontractif. Ceci fournit de nouveaux résultats d'hypercontractivité, mais aussi de nouvelles démonstrations simplifiant ou améliorant des résultats connus. Soulignons que la condition suffisante que nous obtenons est très loin d'être nécessaire : non seulement nous ne traitons que de diffusions, laissant ainsi échapper des cas aussi importants que le processus de Poisson sur l'espace à deux points, mais nous sommes loin de retrouver certains cas connus de diffusions hypercontractives, telles que les mouvements browniens sur toutes les variétés riemanniennes compactes.

Dans un premier paragraphe, nous expliciterons une hypothèse en vigueur dans toute la suite : le processus considéré est une diffusion. La seconde partie donnera la définition de Γ_2 et son calcul sur quelques exemples. Dans la troisième, nous ferons connaissance avec l'hypercontractivité ; dans la dernière, enfin, nous établirons le théorème que nous avons en vue.

DIFFUSIONS MARKOVIENNES

Tous les processus que nous rencontrerons dans la suite seront des diffusions. Nous allons d'abord préciser ce que nous entendons par là et expliciter l'équivalence entre cette hypothèse et une formule de changement de variable pour le générateur infinitésimal L, ou pour le carré du champ associé

$$\Gamma(f,g) = \tfrac{1}{2}[L(fg) - fLg - gLf] \quad .$$

Cette section est donc consacrée à l'énoncé rigoureux et à la démonstration du fait suivant : Pour qu'un processus de Markov admettant un carré du champ Γ soit continu, il faut et il suffit que Γ soit une dérivation.

Tout ceci nous a été expliqué par Meyer, et semble faire plus ou moins partie du folklore de la théorie des processus markoviens et des espaces de Dirichlet. Si la proposition 2 présente quelque nouveauté, la paternité en revient à Meyer.

On suppose données une algèbre $\underline{\underline{A}}$ de fonctions sur l'espace d'états E, et une application linéaire L de $\underline{\underline{A}}$ dans $\underline{\underline{A}}$. Pour que les résultats de ce paragraphe soient faciles à étendre au cas sous-markovien, nous ne supposerons pas que $\underline{\underline{A}}$ contient les constantes (c'est un peu du luxe : l'étude de l'hypercontractivité, plus bas, restera toujours dans le cadre markovien). Nous ne supposons aucune propriété de densité de $\underline{\underline{A}}$, qui peut être réduite à $\{0\}$; mais bien sûr les résultats seront d'autant plus intéressants et exploitables que $\underline{\underline{A}}$ sera plus riche. On appelle opérateur carré du champ associé à L l'application bilinéaire symétrique de $\underline{\underline{A}} \times \underline{\underline{A}}$ dans $\underline{\underline{A}}$ donnée par $\Gamma(f,g) = \frac{1}{2}[L(fg) - fLg - gLf]$. On notera le coefficient $\frac{1}{2}$ (sa présence varie d'un article à l'autre), qui justifie son nom : si L est le laplacien dans $\mathbb{R}^n$, $\Gamma(f,g)$ est le produit scalaire $\operatorname{grad} f \operatorname{grad} g$. On remarquera aussi que tout opérateur bilinéaire symétrique sur $\underline{\underline{A}}$ n'est pas nécessairement le carré du champ associé à un L : des conditions de compatibilité de nature algébrique sont nécessaires, par exemple la symétrie en f, g et h de $\Gamma(f,gh) + f\Gamma(g,h)$.

Nous dirons que Γ est une dérivation si l'on a identiquement pour toutes f, g et h, $\Gamma(fg,h) = f\Gamma(g,h) + g\Gamma(f,h)$. Ceci équivaut, de manière purement algébrique, à une formule de changement de variable pour L.

LEMME 1. Les conditions qui suivent sont équivalentes :

(1a) Pour toute f dans $\underline{\underline{A}}$, $\Gamma(f^2,f) = 2f\Gamma(f,f)$.

(2a) Pour toute f dans $\underline{\underline{A}}$, $L(f^3) = 3fL(f^2) - 3f^2Lf$.

(1b) Pour toutes $f^1, \dots, f^n$, g dans $\underline{\underline{A}}$ et tout polynôme u à n variables sans terme constant,

$$\Gamma(u(f^1, \dots, f^n), g) = \sum_i D_i u(f^1, \dots, f^n)\, \Gamma(f^i, g) \quad .$$

(2b) Pour toutes $f^1, \dots, f^n$ dans $\underline{\underline{A}}$ et tout polynôme u à n variables sans terme constant,

$$L(u(f^1, \dots, f^n)) = \sum_i D_i u(f^1, \dots, f^n)\, L(f^i) + \sum_{i,j} D_{ij} u(f^1, \dots, f^n)\, \Gamma(f^i, f^j).$$

Lorsque ces conditions sont réalisées, l'opérateur L est local au sens suivant : Si f et g sont deux fonctions de $\underline{\underline{A}}$, et si $f = 0$ sur $\{g \neq 0\}$, alors $Lf = 0$ sur $\{g \neq 0\}$.

Pour $u(x,y) = xy$, (1b) exprime que Γ est une dérivation ; les formules $\Gamma(u(f),f) = u'(f)\Gamma(f,f)$ et $L(u(f)) = u'(f)Lf + u''(f)\Gamma(f,f)$ que l'on obtient pour $n = 1$ sont très utiles dans la pratique. La formule $L(f^2) = 2fLf + 2\Gamma(f,f)$ n'est autre que la définition de Γ .

Démonstration du lemme. La formule (2a) peut être mise sous la forme

$$L(f^3) = 3f^2Lf + 6f\Gamma(f,f) \quad ;$$

c'est donc un cas particulier de (2b) ; de même, (1a) est un cas particulier de (1b). Puisque (1a) entraîne (2a) (remplacer Γ par sa définition), il reste à montrer que (2a) implique (1b) et (2b).

En remplaçant dans (2a) f par $af + bg + ch$ (a,b,c réels et f, g, h dans $\underline{\underline{A}}$), on obtient une identité polynômiale en a, b et c ; en identifiant le coefficient de $6abc$, on trouve

$$L(fgh) = fL(gh) + gL(hf) + hL(fg) - fgLh - ghLf - hfLg \quad ,$$

qui peut encore s'écrire $\Gamma(fg,h) = f\Gamma(g,h) + g\Gamma(f,h)$. La relation (1b) se vérifie inductivement sur le polynôme u : en écrivant

$$u(f^1, \dots, f^n) = af^1 + f^1 v(f^1, \dots, f^n) + w(f^2, \dots, f^n)$$

où v et w sont des polynômes sans terme constant, on abaisse le degré de u en f^1, et on se ramène à établir (1b) pour v et w, puis on recommence... La même méthode permet ensuite de vérifier (2b) ; l'équivalence est ainsi complètement établie.

Il reste maintenant à montrer que L est local : si $fg=0$, alors $gLf=0$. En utilisant le fait que Γ est une dérivation, on peut écrire

$$\begin{aligned} g^2L(fg) - 2g\Gamma(g,fg) &= g^3Lf + g^2fLg + 2g^2\Gamma(f,g) - 2fg\Gamma(g,g) - 2g^2\Gamma(f,g) \\ &= g^3Lf + g^2fLg - 2fg\Gamma(g,g) \quad . \end{aligned}$$

Il est clair sur cette identité que $fg=0$ impose $g^3Lf=0$, d'où $gLf=0$. ∎

Nous allons maintenant établir l'équivalence annoncée entre ces formules de changement de variable pour le générateur d'un processus et le fait que celui-ci est une diffusion. En réalité, il ne s'agit pas tout-à-fait d'une équivalence : nous nous plaçons dans des cadres légèrement différents pour les deux sens d'implication (mais voir les remarques qui suivent les propositions 1 et 2). Le lecteur observera que l'une des difficultés rencontrées lorsqu'on établit qu'un processus est une diffusion, l'existence même du carré du champ, est ici complètement passée sous silence, puisque nous supposons d'emblée que L opère sur une algèbre.

a) <u>Si</u> L <u>est un générateur de diffusion</u>, Γ <u>est une dérivation</u>.

Nous supposons donc ici que L est le générateur infinitésimal d'une diffusion sur E, au sens suivant : Pour tout x de E, il existe un processus $(X_t)_{t\geq 0}$ défini sur un espace filtré $(\Omega,\underline{\underline{F}},P,(\underline{\underline{F}}_t)_{t\geq 0})$ pouvant dépendre de x, à valeurs dans E, tel que $X_0=x$ et que, pour chaque $f\in\underline{\underline{A}}$, le processus réel $f\circ X$ soit une semimartingale continue, de décomposition canonique

$$f\circ X_t = f(x) + M_t^f + \int_0^t Lf(X_s)\,ds$$

(où la martingale locale continue M^f dépend de x). La continuité de chaque $f\circ X$ exprime la continuité de X pour la topologie engendrée par $\underline{\underline{A}}$.

PROPOSITION 1. Sous ces hypothèses, Γ est une dérivation. Plus précisément, les formules de changement de variable (1b) et (2b) sont valides pour tous les $u, f^1, \dots f^n, g$ tels que u soit une fonction C^2 et que $f^1, \dots, f^n, g$ et $u(f^1, \dots, f^n)$ soient dans l'algèbre. En outre, pour f et g dans $\underline{\underline{A}}$, le processus à variation finie $\langle M^f, M^g \rangle_t$ est dérivable en t, de dérivée $2\Gamma(f,g)(X_t)$.

Comme la fonction $\Gamma(f,g)$ est dans $\underline{\underline{A}}$, le processus $\Gamma(f,g)(X_t)$ est continu et la dernière phrase dit simplement que $\langle M^f, M^g \rangle_t = 2\int_0^t \Gamma(f,g)(X_s)\,ds$.

Démonstration. On écrit, pour u fonction C^2 telle que $u\circ f$ soit aussi dans $\underline{\underline{A}}$,

$$\begin{aligned} M_t^{u\circ f} + \int_0^t L(u\circ f)(X_s)\,ds &= u\circ f(X_t) - u\circ f(x) \\ &= \int_0^t u'\circ f(X_s)\, d(f\circ X)_s + \tfrac{1}{2}\int_0^t u''\circ f(X_s)\, d[f\circ X, f\circ X]_s \\ &= \int_0^t u'\circ f(X_s)\, dM_s^f + \int_0^t u'\circ f(X_s)\, Lf(X_s)\,ds \\ &\quad + \tfrac{1}{2}\int_0^t u''\circ f(X_s)\, d\langle M^f, M^f\rangle_s \quad . \end{aligned}$$

En identifiant les parties à variation finie, cela donne

$$\int_0^t (L(u\circ f) - u'\circ f\, Lf)(X_s)\,ds = \tfrac{1}{2}\int_0^t u''\circ f(X_s)\, d\langle M^f, M^f\rangle_s \quad .$$

Pour $u(x) = x^2$, ceci s'écrit $\langle M^f, M^f\rangle_t = 2\int_0^t \Gamma(f,f)(X_s)\,ds$; d'où le crochet $\langle M^f, M^g\rangle$ par polarisation. Pour u quelconque, on a alors

$$\int_0^t (L(u\circ f) - u'\circ f\, Lf - u''\circ f\, \Gamma(f,f))(X_s)\,ds = 0 \ .$$

L'expression sous le signe somme, de la forme $g(X_s)$ avec g dans $\underline{\underline{A}}$, est continue par rapport à s . En dérivant, on a donc $g(X_t) = 0$, d'où, pour $t = 0$, $g(x) = 0$. Comme x est arbitraire, la formule du changement de variable est établie pour $L(u\circ f)$. La formule $\Gamma(u\circ f, f) = u'\circ f\,\Gamma(f,f)$ s'en déduit sans peine, et le cas d'une fonction u de plusieurs variables se traite de même. ∎

REMARQUE. Si l'on s'était donné un seul processus, de loi initiale fixée, et non toute une famille, seul l'argument final serait en défaut ; on obtiendrait alors les mêmes formules de changement de variable, mais elles ne seraient valables que sur un complémentaire d'ensemble polaire.

b) Réciproquement, si Γ est une dérivation, le processus est continu.

Cette réciproque se place dans un cadre un peu moins strict : comme dans la remarque qui précède, nous n'aurons besoin que d'un processus, et non de toute une famille indexée par E .

PROPOSITION 2. Soit X un processus à valeurs dans E , défini sur un espace filtré $(\Omega,\underline{\underline{F}},P,(\underline{\underline{F}}_t)_{t\geq 0})$, tel que pour toute f dans $\underline{\underline{A}}$, le processus

$$M_t^f = f\circ X_t - f\circ X_0 - \int_0^t Lf\circ X_s\, ds$$

soit une martingale locale. Si Γ est une dérivation, alors pour f et g dans $\underline{\underline{A}}$ les semimartingales $f\circ X$ et M^f sont continues, et $\frac{d}{dt}\langle M^f,M^g\rangle_t = 2\Gamma(f,g)(X_t)$.

Démonstration. Elle repose sur un lemme de théorie générale des processus, directement inspiré d'un résultat très semblable de Meyer, Stricker et Zheng (voir Meyer-Zheng [5]). Nous noterons $H_n(x,a)$ les polynômes d'Hermite à deux variables définis par la série génératrice $\sum_{n\geq 0} H_n(x,a)\, z^n = \exp\left(xz - \frac{1}{2}az^2\right)$; ils diffèrent de la définition usuelle par un facteur $n!$ sans importance, sont de degré n en x , et vérifient $\frac{\partial}{\partial x}H_n = H_{n-1}$; $\frac{\partial}{\partial a}H_n = -\frac{1}{2}H_{n-2}$ (avec la convention $H_n = 0$ pour $n<0$). Il est bien connu — et nous le redémontrerons au passage — que, si M est une martingale locale continue et $A = \langle M,M\rangle$ sa variation quadratique, alors, pour tout n , $H_n(M,A)$ est une martingale locale. Le lemme en est une réciproque.

LEMME 2. Soient M une martingale locale, A un processus à variation finie continu et nul en 0 . On suppose que, pour $n\leq 4$, les semimartingales $H_n(M,A)$ sont des martingales locales. Alors M est continue et $A = \langle M-M_0\, ,\, M-M_0\rangle$.

Ce lemme sera établi plus loin ; terminons d'abord de démontrer la proposition. Il suffit, par le lemme, de vérifier que, pour $f\in\underline{\underline{A}}$ et $n\leq 4$, la semimartingale $N = H_n(M^f,A^f)$ est une martingale locale, où l'on a posé $A_t^f = 2\int_0^t \Gamma(f,f)(X_s)\, ds$. Nous utiliserons pour cela l'identité suivante, qui découle de la formule de changement de variable $L(f^k) = kf^{k-1}Lf + k(k-1)f^{k-2}\Gamma(f,f)$ (où, par convention, que les constantes soient ou non dans l'algèbre, on a posé $L(f^0) = 0$) :

$$(3)\qquad \sum_{k=0}^{n} \frac{1}{k!}\left[\, b_{n-k}\, L(f^k) - f^k\left(b_{n-k-1}\, Lf + b_{n-k-2}\,\Gamma(f,f)\right)\right] = 0 \quad ,$$

pour toute suite $(b_n)_{n\in\mathbb{Z}}$ de réels nuls pour $n<0$.

En appliquant la formule de Taylor

$$H_n(x+y\,,a) = \sum_{k\leq n} \frac{y^k}{k!}\, H_{n-k}(x,a)$$

on obtient pour N l'expression

$$N = \sum_k \frac{1}{k!}\, f^k\circ X_t\; H_{n-k}(-f\circ X_0 - \int_0^t Lf(X_s)\,ds\,,\, A_t^f) \quad .$$

Par la formule d'intégration par parties $d(YB)_t = B_t dY_t + Y_t dB_t$ (Y semimartingale, B processus à variation finie continu), on en déduit

$$\begin{aligned} dN_t &= \sum_k \frac{1}{k!}\,[\, H_{n-k}(\dots)\,(dM_t^{f^k} + L(f^k)(X_t)\,dt\,)^{(*)} \\ &\quad + f^k\circ X_t\,(H_{n-k-1}(\dots)(-Lf(X_t))dt - \tfrac{1}{2}\, H_{n-k-2}(\dots)\,2\,\Gamma(f,f)(X_t)dt)\,] \\ &= \sum_k \frac{1}{k!}\, H_{n-k}(\dots)\, dM_t^{f^k} \quad \text{en vertu de l'identité (3)} \quad . \end{aligned}$$

Ainsi N est une martingale locale, le lemme s'applique, et on obtient donc la continuité de M^f et $f\circ X$, ainsi que la formule $\langle M^f,M^f\rangle_t = A_t^f = 2\int_0^t \Gamma(f,f)(X_s)\,ds$. ∎

REMARQUE. Puisque Γ est une dérivation, les formules de changement de variable ont lieu pour les polynômes sans terme constant. La remarque qui suit la proposition 1 montre qu'elles s'étendent aux autres fonctions, mais seulement hors d'un ensemble polaire. On pourrait lever cette difficulté en supposant comme dans le a) que nous disposons de toute une famille de processus, indexés par leur valeur initiale ; dans ce cadre, on obtiendrait une équivalence complète entre continuité du processus et formules de changement de variable.

Il reste maintenant à établir le lemme 2. Puisque $H_0(x,a) = 1$, $H_1(x,a) = x$ et $H_2(x,a) = \frac{1}{2}(x^2 - a)$, pour $n=0$ l'hypothèse ne dit rien ; pour $n=1$ elle répète que M est une martingale locale et pour $n=2$ elle dit que M est localement de carré intégrable, de variation quadratique prévisible $\langle M,M\rangle = A$. Le véritable contenu du lemme, c'est que, si de plus $H_3(M,A)$ et $H_4(M,A)$ sont des martingales locales, alors M est continue. En réalité, la démonstration utilisera seulement l'hypothèse un peu plus faible : $H_4(M_t,A_t) - \int_0^t M_{s-}\,d(H_3(M_s,A_s))$ est une martingale locale.

(*) Que les constantes soient ou non dans l'algèbre, $dM_t^{f^k} + L(f^k)(X_t)\,dt = d(f^k\circ X)_t$ doit être pris par convention égal à zéro pour $k=0$.

Démonstration du lemme 2. Pour $n \leq 4$, nous savons que $H_n(M,A)$ est une martingale locale. Or

$$\begin{aligned}
H_n(M_t,A_t) &- H_n(M_0,A_0) \\
&= \int_0^t H_{n-1}(M_{s-},A_s)\,dM_s + \tfrac{1}{2}\int_0^t H_{n-2}(M_{s-},A_s)\,d(\langle M^c,M^c\rangle - A)_s \\
&\qquad + \sum_{s\leq t}[H_n(M_s,A_s) - H_n(M_{s-},A_s) - H_{n-1}(M_{s-},A_s)\Delta M_s] \\
&= \int_0^t H_{n-1}(M_{s-},A_s)\,dM_s + \tfrac{1}{2}\int_0^t H_{n-2}(M_{s-},A_s)\,d([M,M]-A)_s \\
&\qquad + \sum_{s\leq t}[H_n(M_s,A_s)-H_n(M_{s-},A_s)-H_{n-1}(M_{s-},A_s)\Delta M_s-\tfrac{1}{2}H_{n-2}(M_{s-},A_s)\Delta M_s^2].
\end{aligned}$$

Par hypothèse, $M^2 - A = 2H_2(M,A)$ est une martingale locale. Donc $[M,M]-A$ aussi, et ceci établit que

$$\sum_{s\leq t}[H_n(M_s,A_s) - H_n(M_{s-},A_s) - H_{n-1}(M_{s-},A_s)\Delta M_s - \tfrac{1}{2}H_{n-2}(M_{s-},A_s)\Delta M_s^2]$$

en est également une. Mais, par la formule de Taylor, cette dernière peut s'écrire

$$\sum_{s\leq t}\Big[\sum_{k=3}^{n}\frac{1}{k!}H_{n-k}(M_{s-},A_s)\,\Delta M_s^k\Big] \quad ;$$

pour $n = 3$ et 4, on obtient que $K_t = \sum_{s\leq t}\Delta M_s^3$ et $L_t = \sum_{s\leq t}(4M_{s-}\Delta M_s^3 + \Delta M_s^4)$ sont aussi des martingales locales, donc $\sum_{0<s\leq t}\Delta M_s^4 = L_t - 4\int_0^t M_{s-}\,dK_s$ est elle aussi une martingale locale. Ceci entraîne que M est continue. ∎

L'énoncé du lemme semble curieux : pourquoi $n \leq 4$? La seule chose que nous sachions est que les hypothèses $H_1(M,A)$, $H_2(M,A)$ et $H_3(M,A)$ martingales locales ne suffisent pas à entraîner la continuité de M. Voici un contre-exemple : Soient $N'_t - t$ et $N''_t - t$ deux processus de Poisson compensés indépendants, et M la martingale non continue $N' - N''$. Puisque $[M,M] = N' + N''$ est compensé par le processus croissant $A = 2t$, $H_1(M,A) = M$ et $H_2(M,A) = \frac{1}{2}(M^2 - A)$ sont des martingales. Quant à $H_3(M,A)$, il résulte de la preuve du lemme que c'est une martingale si et seulement si $K_t = \sum_{s\leq t}\Delta M_s^3$ en est une ; or ceci est vrai puisque, M variant uniquement par sauts d'amplitude ± 1, on a $K = M$.

LE CARRÉ DU CHAMP ITÉRÉ

Cette section met en scène l'acteur principal : le carré du champ itéré Γ_2 . C'est l'opérateur bilinéaire symétrique sur $\underline{\underline{A}}$ défini à partir de L et Γ par

$$\Gamma_2(f,g) = \tfrac{1}{2}[\,L\Gamma(f,g) - \Gamma(Lf,g) - \Gamma(f,Lg)\,] .$$

(On pourrait itérer cette définition et introduire des opérateurs Γ_n pour tout n . Nous ne les utiliserons pas ; remarquons seulement que ceci conduirait à poser $\Gamma_1 = \Gamma$, $\Gamma_0(f,g) = fg$: Γ_2 est à Γ ce que celui-ci est au produit.)

Exemples. Voici maintenant l'expression de Γ_2 pour quelques générateurs L de processus markoviens.

a) Si $E = \mathbb{R}$, $Lf = af'' + bf'$ pour deux fonctions a et b , alors $\Gamma(f,f) = af'^2$ et $\Gamma_2(f,f) = a^2f''^2 + aa'f'f'' + \frac{1}{2}(\,aa'' + a'b - 2ab'\,)\,f'^2$. Cette expression se simplifie lorsque L est non dégénéré : il se met alors sous la forme $L = H^2 + bH$, où $H(f) = \alpha f'$, et on a alors $\Gamma(f,f) = (Hf)^2$ et $\Gamma_2(f,f) = (H^2f)^2 - H(b)(Hf)^2$. On remarque tout de suite que, contrairement à Γ qui est toujours positif dans le cas probabiliste $a \geqq 0$, $\Gamma_2(f,f)$ ne l'est pas nécessairement.

b) Si $E = \mathbb{R}^n$ et si L est le laplacien, alors $\Gamma(f,g)$ est le produit scalaire $\mathrm{grad}\,f\ \mathrm{grad}\,g$ et $\Gamma_2(f,g) = \sum_{i,j} D_{ij}f\,D_{ij}g$. C'est de là que lui vient le nom de "gradient itéré" parfois employé, mais trompeur, car $\Gamma_2(f,f)$ est quadratique en l'argument f .

c) Plus généralement, sur un espace E quelconque, soit L un opérateur de la forme $\sum_i D_i^2 + X$, où X et chaque D_i sont des dérivations ($D_i(fg) = fD_ig + gD_if$). Alors, $\Gamma(f,g) = \Sigma\, D_if\,D_ig$ (car $D_i^2(fg) = fD_i^2g + gD_i^2f + 2\,D_if\,D_ig$) ; réciproquement, cette expression de Γ entraîne que $L - \Sigma\, D_i^2$ est une dérivation. Si, en outre, $[L,D_i]$ ($= LD_i - D_iL$, commutateur de L et D_i) est égal à aD_i pour tout i , avec une fonction a quelconque, mais indépendante de i , alors

$$\Gamma_2(f,g) = \sum_{i,j} D_iD_jf\,D_iD_jg + a\,\Gamma(f,g) ,$$

car $L\Gamma(f,g) = \Sigma(\,D_if\,LD_ig + D_ig\,LD_if + 2\Gamma(D_if,D_ig)\,)$ et $\Gamma(f,Lg) + \Gamma(Lf,g) = \Sigma(\,D_if\,D_iLg + D_ig\,D_iLf\,)$.

Ceci permet de calculer Γ et Γ_2 pour le laplacien sur la sphère n-dimensionnelle, c'est-à-dire la partie orthoradiale du laplacien dans $\mathbb{R}^{n+1}$ écrite pour les fonctions définies sur la sphère unité $E = S^n$ de $\mathbb{R}^{n+1}$ (nous en verrons un autre calcul dans l'exemple d). L'opérateur différentiel défini dans $\mathbb{R}^{n+1}$ par

$$R_{ij} = x_i D_j - x_j D_i \qquad (1 \leq i < j \leq n+1)$$

est un vecteur tangent à S^n (quand on le calcule en un point de S^n), c'est donc une dérivation sur les fonctions de S^n , et le laplacien sphérique n'est autre que $L = \sum_{i,j} (R_{ij})^2$. Comme L commute avec les rotations de S^n , il commute aussi avec les R_{ij} , qui sont des générateurs infinitésimaux de groupes de rotations. Donc ici $a = 0$, et

$$\Gamma(f,f) = \sum_{i,j} (R_{ij}f)^2 \qquad ;$$

$$\Gamma_2(f,f) = \sum_{i,j,k,\ell} (R_{ij} R_{k\ell} f)^2 \qquad .$$

La même méthode s'applique aux générateurs du type Ornstein-Uhlenbeck. Dans ce cas, l'algèbre $\underline{\underline{A}}$ est graduée : $\underline{\underline{A}} = \bigoplus_{n \geq 0} \underline{\underline{A}}_n$ et L opère sur $\underline{\underline{A}}_n$ par multiplication par $-n$; on a en outre $L = \Sigma D_i^2 + X$ où chaque D_i envoie $\underline{\underline{A}}_{n+1}$ dans $\underline{\underline{A}}_n$. Il est clair que l'on a alors $[L, D_i] = D_i$, d'où

$$\Gamma_2(f,g) = \sum_{i,j} D_i D_j f \, D_i D_j g + \Gamma(f,g) \qquad .$$

Ceci s'applique, en particulier, au générateur $L = (\frac{d}{dx})^2 - x\frac{d}{dx}$ sur $\mathbb{R}$: prendre $D = \frac{d}{dx}$, $X = -xD$, $\underline{\underline{A}}$ = algèbre des polynômes, $\underline{\underline{A}}_n$ = sous-espace de dimension 1 engendré par le $n^{\text{ième}}$ polynôme d'Hermite. La même formule vaut aussi, pour la même raison au fond, pour le processus d'Ornstein-Uhlenbeck introduit par Malliavin sur l'espace de Wiener $C([0,\infty[)$: voir Meyer [3], formule (83).

d) Étendons l'exemple a) au cas où $E = \mathbb{R}^n$, ou, plus généralement, E variété C^∞ à n dimensions. On suppose que L est un opérateur différentiel d'ordre 2 , sans terme constant, à coefficients réguliers : dans un système de coordonnées locales,

$$Lf(x) = \alpha^{ij}(x) \, D_{ij} f(x) + \beta^i(x) \, D_i f(x) \qquad .$$

Nous ferons le calcul sous l'hypothèse d'<u>ellipticité</u> de L : pour chaque x , la matrice symétrique α^{ij} (tenseur deux fois contravariant) est définie positive. Sous cette condition, il existe une unique structure riemannienne sur E telle que, si

Δ désigne le laplacien sur E associé à cette structure riemannienne (opérateur de Beltrami), $L-\Delta$ soit du premier ordre : puisque dans une carte locale on a toujours $\Delta = g^{ij}(D_{ij} - \Gamma^k_{ij}D_k)$, l'unique métrique répondant au problème est donnée par $g^{ij} = \alpha^{ij}$, c'est-à-dire $g_{ij} = (\alpha^{ij})^{-1}$. (Tout ceci est bien familier aux probabilistes : c'est simplement le truc, classique depuis Itô, consistant à écrire la diffusion de générateur L à l'aide d'équations différentielles stochastiques browniennes.) On est ainsi ramené au cas où $L = \Delta + b$, avec b champ de vecteurs (opérateur différentiel d'ordre 1). Dans les cas qui nous intéresseront, b sera un champ de gradients ; ceci revient à dire que L est symétrique pour une mesure invariante μ , qui est alors liée à b par la relation de Kolmogorov $b = \operatorname{grad} \operatorname{Log} \frac{d\mu}{dr}$, où r est la mesure riemannienne.

PROPOSITION 3. <u>Sur une variété riemannienne (ou pseudo-riemannienne), soient b un champ de vecteurs, Δ le laplacien, et L l'opérateur $\Delta + b$. Le carré du champ et le carré du champ itéré associés à L sont donnés par</u>

$$\Gamma(f,g) = (\operatorname{grad} f \,|\, \operatorname{grad} g)$$

$$\Gamma_2(f,g) = (\operatorname{Hess} f \,|\, \operatorname{Hess} g) + (\operatorname{Ric} - \nabla^{\mathrm{sym}} b)(\operatorname{grad} f, \operatorname{grad} g) \quad ; \tag{4a}$$

<u>lorsque</u> b <u>est un champ de gradients,</u> $b = \operatorname{grad} h$, <u>ceci s'écrit</u>

$$\Gamma_2(f,g) = (\operatorname{Hess} f \,|\, \operatorname{Hess} g) + (\operatorname{Ric} - \operatorname{Hess} h)(\operatorname{grad} f, \operatorname{grad} g) \quad . \tag{4b}$$

Pour $b = 0$, c'est la formule de Bochner-Lichnerowicz-Weitzenböck (voir Berger-Gauduchon-Mazet [1] p. 131 — qui comporte une coquille : $|\Delta f|^2$ doit y être remplacé par $(d(\Delta f), df)$). La notation $\operatorname{Hess} f$ désigne la forme hessienne $\nabla \operatorname{grad} f$, c'est-à-dire le tenseur symétrique deux fois contravariant dont l'action sur les vecteurs est donnée par $\operatorname{Hess} f\,(U,V) = (\nabla_U \operatorname{grad} f \,|\, V)$; en coordonnées locales, $(\operatorname{Hess} f)_{ij} = (D_{ij} - \Gamma^k_{ij} D_k) f$, où Γ^k_{ij} sont les symboles de Christoffel. Les parenthèses $(\ |\ )$ forment le produit scalaire local des vecteurs ou des tenseurs ; ainsi, $(\operatorname{Hess} f \,|\, \operatorname{Hess} g) = g^{ik} g^{j\ell} (\operatorname{Hess} f)_{ij} (\operatorname{Hess} g)_{k\ell}$. Enfin, Ric désigne le tenseur de courbure de Ricci de la variété riemannienne (avec la convention de signe qui le rend positif pour les variétés à courbure positives !) et $\nabla^{\mathrm{sym}} b$ n'est autre que la dérivée covariante symétrique du champ de vecteurs b :

$$\nabla^{\mathrm{sym}} b(U,V) = \tfrac{1}{2}\,[(\nabla_U b \,|\, V) + (\nabla_V b \,|\, U)] \quad .$$

Démonstration de la proposition 3. Il est clair que $\nabla^{sym}\operatorname{grad} h = \operatorname{Hess} h$, donc (4b) découle de (4a). La formule donnant Γ n'est autre que l'identité

$$\Delta(fg) = f\Delta g + g\Delta f + 2(\operatorname{grad} f \mid \operatorname{grad} g) ,$$

jointe au fait que, b étant une dérivation, Γ ne dépend pas de b . Il nous reste à établir (4a).

Nous utiliserons pour cela la formule suivante : si F et G sont deux champs de gradients sur E ,

$$\text{(5)} \qquad \operatorname{grad}(F|G) = \nabla_F G + \nabla_G F \quad .$$

(Pour une formule plus générale, voir Meyer [4] formule $(55)_c$.) En effet, si $F = \operatorname{grad} f$ et $G = \operatorname{grad} g$, alors pour tout champ de vecteurs Z ,

$$\begin{aligned}(Z|\nabla_F G + \nabla_G F) &= F(G|Z) + G(F|Z) - Z(F|G) + ([Z,F]|G) + ([Z,G]|F)\\ &\qquad \text{(voir [1] page 25)}\\ &= FZg + GZf - Z(F|G) + (ZF - FZ)g + (ZG - GZ)f\\ &= Z[Fg + Gf - (F|G)] = Z(F|G) = (Z|\operatorname{grad}(F|g)) \quad .\end{aligned}$$

La relation (5) est établie, nous allons maintenant démontrer (4a). Nous nous plaçons d'abord dans le cas où b est nul : $L = \Delta$.

Posons $F = \operatorname{grad} f$, $G = \operatorname{grad} g$; nous allons établir (4a) en un point x fixé dans la suite. Choisissons une base orthonormée (e_i) de l'espace tangent T_xE (dans le cas pseudo-riemannien, $(e_i|e_i) = \pm 1$), et définissons un champ de repères orthonormés au voisinage de x en décidant que $e_i(y)$ est obtenu à partir de $e_i(x) = e_i$ par transport parallèle le long de la géodésique qui joint x et y . Ceci entraîne $\nabla_U e_i = 0$ en x pour tout $U \in T_x$. Le calcul de Γ_2 est alors le suivant :

$$\begin{aligned}\Delta\Gamma(f,g) &= \operatorname{div}\operatorname{grad}(F|G) = \operatorname{div}(\nabla_F G + \nabla_G F) \quad \text{grâce à (5)}\\ &= \sum_i (e_i|\nabla_{e_i}\nabla_F G) + \sum_i (e_i|\nabla_{e_i}\nabla_G F)\\ \Gamma(\Delta f,g) &= (\operatorname{grad}\operatorname{div} F \mid G) = G(\operatorname{div} F) = \sum_i G(e_i|\nabla_{e_i} F)\\ &= \sum_i (\nabla_G e_i|\nabla_{e_i} F) + \sum_i (e_i|\nabla_G \nabla_{e_i} F) \underset{\boxed{\text{en } x}}{=} \sum_i (e_i|\nabla_G\nabla_{e_i} F) \quad .\\ \Gamma_2(f,g) &= \tfrac{1}{2}\sum_i (e_i|\nabla_{e_i}\nabla_F G + \nabla_{e_i}\nabla_G F - \nabla_G\nabla_{e_i} F - \nabla_F\nabla_{e_i} G)\\ &= \tfrac{1}{2}\sum_i (e_i|\nabla_{[e_i,F]} G + \nabla_{[e_i,G]} F + R(F,e_i)G + R(G,e_i)F) ,\end{aligned}$$

formule valable au point x , où R représente le tenseur de courbure riemannienne.
Les deux derniers termes, sommés en i , donnent

$$\tfrac{1}{2}\,[\,\mathrm{Ric}(F,G)+\mathrm{Ric}(G,F)\,] = \mathrm{Ric}(F,G) \quad .$$

Restent les deux premiers termes. Il suffit de calculer le premier, le second s'en déduira par échange de f et g . Il vient

$$\begin{aligned}
\sum_i (e_i|\nabla_{[e_i,F]}G) &= \sum_i \mathrm{Hess}\,g\,(e_i,[e_i,F]) = \sum_i (\nabla_{e_i}G|[e_i,F]) \\
&\qquad\qquad\text{(symétrie de la hessienne)} \\
&= \sum_i (\nabla_{e_i}G|\nabla_{e_i}F) - \sum_i (\nabla_{e_i}G|\nabla_F e_i) \\
&= \sum_i (\nabla_{e_i}G|\nabla_{e_i}F) \qquad \text{au point } x \\
&= \sum_{i,j} (\nabla_{e_i}F|e_j)\,(\nabla_{e_j}G|e_j) = \sum_{i,j} \mathrm{Hess}\,f\,(e_i,e_j)\ \mathrm{Hess}\,g \\
&= \sum_{i,j} \mathrm{Hess}\,f\,(e_i,e_j)\ \mathrm{Hess}\,g\,(e_i,e_j) = (\mathrm{Hess}\,f|\mathrm{Hess}\,g)\ .
\end{aligned}$$

En définitive, pour $L = \Delta$,

$$\Gamma_2(f,g) = (\mathrm{Hess}\,f|\mathrm{Hess}\,g) + \mathrm{Ric}(\mathrm{grad}\,f,\mathrm{grad}\,g) \quad ,$$

et (4a) est établie dans ce cas.

Nous n'avons plus, pour le cas général, qu'à <u>évaluer le terme correctif dû à la présence de</u> b . Il vaut

$$\begin{aligned}
&\tfrac{1}{2}\,[b(\mathrm{grad}\,f|\mathrm{grad}\,g) - (\mathrm{grad}\,b(f)|\mathrm{grad}\,g) - (\mathrm{grad}\,f|\mathrm{grad}\,b(g))] \\
&\quad = \tfrac{1}{2}\,[b(F|G) - G(b|F) - F(b|G)] \\
&\quad = \tfrac{1}{2}\,[(\nabla_b F|G) + (F|\nabla_b G) - (\nabla_G b|F) - (b|\nabla_G F) \\
&\qquad\qquad - (\nabla_F b|G) - (b|\nabla_F G)] \\
&\quad = \tfrac{1}{2}\,[\mathrm{Hess}\,f\,(b,G) + \mathrm{Hess}\,g\,(b,F) - (\nabla_G b|F) - \mathrm{Hess}\,f\,(G,b) \\
&\qquad\qquad -(\nabla_F b|G) - \mathrm{Hess}\,g\,(F,b)\,] \\
&\quad = -\,\nabla^{\mathrm{sym}}b(F,G) \qquad . \qquad \blacksquare
\end{aligned}$$

Dans le cas où $b = \mathrm{grad}\,h$, ce dernier calcul fournit une formule liant la hessienne, c'est-à-dire au fond la connection riemannienne, au carré du champ Γ : sur une variété riemannienne, ou pseudo-riemannienne, on a toujours

(6) $$\mathrm{Hess}\,h\,(\mathrm{grad}\,f,\mathrm{grad}\,g) = \tfrac{1}{2}\,[\Gamma(\Gamma(h,f),g) + \Gamma(f,\Gamma(h,g)) - \Gamma(h,\Gamma(f,g))] \quad .$$

Ceci n'est autre que l'expression classique des symboles de Christoffel Γ^i_{jk} en fonction des dérivées du tenseur métrique :

$$\Gamma^{jik} = g^{j\ell}g^{km}\Gamma^{i}_{lm} = \frac{1}{2}\,[g^{i\ell}D_\ell g^{jk} - g^{j\ell}D_\ell g^{ik} - g^{k\ell}D_\ell g^{ij}] \;.$$

Nous l'avons écrite à l'envers, avec tous les indices en haut ; elle est bien sûr équivalente à la formule traditionnelle. Sous cette forme, elle a l'avantage de garder un sens dans le cas dégénéré où on se donne un L non nécessairement elliptique : la formule (6) permet encore de définir $\mathrm{Hess}\,h\,(\mathrm{grad}\,f,\mathrm{grad}\,g)$ bien qu'on dispose alors seulement de la "co-métrique" g^{ij} (forme quadratique sur l'espace cotangent), mais non de g_{ij} pour abaisser les indices. Dans ce cas, on peut toujours écrire $(\mathrm{Hess}\,h)^{ij} = g^{i\ell}g^{jm}D_{\ell m}h - \Gamma^{ikj}D_k h$, mais $(\mathrm{Hess}\,h)_{ij}$ n'a aucun sens.

Notons aussi, au passage, une formule équivalente à (6) quand Γ est une dérivation (ce qui est le cas ici) :

$$\mathrm{Hess}\,h\,(\mathrm{grad}\,f,\mathrm{grad}\,g) = \frac{1}{2}\,[\Gamma_2(fg,h) - f\Gamma_2(g,h) - g\Gamma_2(f,h)] \quad ;$$

elle exprime précisément ce qui manque à Γ_2 pour en être aussi une.

e) Ni cet exemple, ni les suivants, ne sont des diffusions : Γ n'est pas une dérivation. Sur $E = \{-1,1\}$, le générateur $Lf(x) = f(-x) - f(x)$ (correspondant au processus qui change de site à des instants poissonniens) donne lieu à

$$\Gamma_2(f,f) = 2\,\Gamma(f,f) = (Lf)^2 \;.$$

f) Si μ est une mesure de probabilité sur un espace E quelconque, et si L est donné par $Lf(x) = \mu(f) - f(x)$, on a $L^2 = -L$,

$$\Gamma(f,f) = \frac{1}{2}\,[\mu(f^2) - \mu(f)^2 + (Lf)^2] = \frac{1}{2}\,[\mathrm{var}\,f + (Lf)^2] \quad ,$$

$$\Gamma_2(f,f) = \frac{1}{2}\,[\mathrm{var}\,f + \Gamma(Lf,Lf)] = \frac{1}{2}\,[\,\mathrm{var}\,f + \Gamma(f,f)\,] \quad .$$

g) Ce dernier exemple est un cas particulier d'une situation étudiée par Surgailis [12] (voir aussi Ruiz de Chavez [11]). On se donne un ensemble S muni d'une mesure m positive, finie et diffuse ; E est l'ensemble de toutes les parties finies de S , et, pour $x \in E$, $Lf(x)$ est défini par

$$Lf(x) = \int_{u\in S} [f(x\cup\{u\}) - f(x)]\,m(du) + \sum_{u\in x} [f(x-\{u\}) - f(x)] \;.$$

Ce générateur correspond au processus à valeurs dans E décrivant l'existence de particules dans S qui naissent à des instants poissonniens (les naissances forment un processus de Poisson d'intensité $m \times dt$ dans l'espace-temps), restent durant toute leur existence au point où elles sont nées, et disparaissent après une durée de vie exponentielle ; il admet comme mesure invariante réversible la probabilité sur E qui est la loi du processus ponctuel de Poisson d'intensité m sur S . En posant $D_u^+f(x) = f(x \cup \{u\}) - f(x)$ et, pour $u \in x$, $D_u^-f(x) = f(x-\{u\}) - f(x)$ (ce ne sont pas des dérivations), le calcul du carré du champ et du carré du champ itéré donne

$$2\,\Gamma(f,f)\,(x) = \int [D_u^+f(x)]^2\, m(du) + \sum_{u \in x} [D_u^-f(x)]^2 \quad ;$$

$$\begin{aligned} 4\,\Gamma_2(f,f)\,(x) = & \iint [D_u^+D_v^+f(x)]^2\, m(du)\, m(dv) \\ & + 2 \int \sum_{v \in x} [D_u^+D_v^-f(x)]^2\, m(du) \\ & + \sum_{\substack{v \in x \\ u \in x-\{v\}}} [D_u^-D_v^-f(x)]^2 \\ & + 3 \int [D_u^+f(x)]^2\, m(du) + \sum_{u \in x} [D_u^-f(x)]^2 \quad . \end{aligned}$$

DÉFINITION DE L'HYPERCONTRACTIVITÉ

De quoi s'agit-il ? Si X et Y sont deux variables aléatoires à valeurs dans un espace quelconque E, on a toujours

$$\|E[f(Y)|X]\|_{L^p} \leq \|f(Y)\|_{L^p} \leq \infty \quad :$$

c'est la propriété de <u>contractivité</u> de l'espérance conditionnelle (ici et dans toute la suite, les exposants p, q, ... sont dans $[1,\infty[$). Mais il peut se faire que, pour certains couples $p<q$, on ait

(7) $$\text{pour toute } f, \quad \|E[f(Y)|X]\|_{L^q} \leq \|f(Y)\|_{L^p} \quad ;$$

il est naturel d'appeler cela <u>hypercontractivité</u> du vecteur (X,Y) (bien que n'interviennent ici en fait que les tribus $\sigma(X)$ et $\sigma(Y)$). En suivant Neveu [ ], on peut mettre ceci sous une forme plus symétrique, en introduisant l'exposant q' conjugué de q :

(7') $$\forall f, g \quad E[g(X)f(Y)] \leq \|g(X)\|_{L^{q'}} \|f(Y)\|_{L^p} \quad ,$$

la condition $q>p$ devenant maintenant $(q'-1)(p-1)<1$.

Par exemple, l'hypercontractivité du semi-groupe d'Ornstein-Uhlenbeck (due à Nelson [7] ; nous verrons cela plus loin) signifie que si (X,Y) est un vecteur gaussien de coefficient de corrélation ρ, alors (7') a lieu pour tout couple p, q' tel que $(q'-1)(p-1) \geq \rho^2$ (Neveu [8] a donné de cette propriété une éblouissante démonstration à l'aide d'intégrales stochastiques).

Quand $\rho^2 = 1$, (7') se réduit à l'inégalité de Hölder usuelle, toujours vraie (contractivité) ; à l'opposé, quand $\rho = 0$, (7') écrite écrite pour $p = q' = 1$ exprime l'indépendance de X et Y. Ainsi, la relation $(q'-1)(p-1) \geq \rho^2 \Rightarrow$ (7'), ou, de façon équivalente, $q \leq 1+(p-1)\rho^2 \Rightarrow$ (7), est une sorte de mesure de dépendance, les variables X et Y étant d'autant plus indépendantes que ρ est petit.

Nous dirons qu'un processus aléatoire $(X_t)_{t\geq 0}$ à valeurs dans un espace E est hypercontractif si la dépendance des variables X_0 et X_t au sens ci-dessus décroît exponentiellement avec le temps t ("exponentiellement hypercontractif" serait préférable, d'autant plus qu'on rencontre d'autres formes d'hypercontractivité, plus faibles que celle-ci ; par exemple en n'exigeant pas une décroissance exponentielle, ou en autorisant une constante dans l'équation (7)).

DÉFINITION. Le processus X est hypercontractif s'il existe une constante $\lambda > 0$ (dite constante d'hypercontractivité) telle que, pour tous $p \geq 1$, $q \geq 1$, $t \geq 0$,

(8) $$q-1 \leq (p-1)e^{\lambda t} \Rightarrow \forall f \quad \|E[f(X_t)|X_0]\|_{L^q} \leq \|f(X_t)\|_{L^p} .$$

Lorsque X est un processus de Markov stationnaire, de semi-groupe de transition $(P_t)_{t\geq 0}$ et de loi invariante μ , ceci s'écrit simplement

(9) $$q-1 \leq (p-1)e^{\lambda t} \Rightarrow \forall f \quad \|P_t f\|_{L^q(\mu)} \leq \|f\|_{L^p(\mu)} .$$

Plus la constante λ est grande, plus le semi-groupe est hypercontractif ; si λ est une constante d'hypercontractivité, il en va de même de toute constante plus petite.

La décroissance exponentielle de la dépendance, c'est-à-dire le facteur $e^{\lambda t}$ dans (8) et (9), est alors justifié par sa compatibilité parfaite avec la propriété de semi-groupe : si P_s contracte L^p dans L^q pour $q-1 = (p-1)e^{\lambda s}$, et si P_t contracte L^q dans L^r pour $r-1 = (q-1)e^{\lambda t}$, alors P_{s+t} contracte L^p dans L^r pour $r-1 = (p-1)e^{\lambda(s+t)}$.

Nous allons maintenant indiquer diverses formulations équivalentes à l'hypercontractivité lorsque le processus est une diffusion. La plus importante d'entre elles est l'inégalité (ou plutôt les inégalités) de Sobolev logarithmique, due à Gross [2] (Gross se place dans un cadre bien plus général que celui des diffusions markoviennes).

Nous nous donnons une algèbre $\underline{\underline{A}}$ de fonctions bornées sur E , sur laquelle opèrent les fonctions de classe $\mathcal{C}^\infty$ (en particulier, $\underline{\underline{A}}$ contient les constantes). Sur $\underline{\underline{A}}$ agissent un opérateur L et un semigroupe $(P_t)_{t\geq 0}$ d'opérateurs markoviens, engendré par L : pour $f \in \underline{\underline{A}}$ et $x \in E$, $P_t f(x)$ est dérivable en t , de dérivée $\frac{d}{dt}P_t f(x) = LP_t f(x) = P_t Lf(x)$. Le caractère markovien des P_t entraîne la positivité de $\Gamma(f,f) = \frac{d}{dt}\Big|_{t=0} [\frac{1}{2}(P_t(f^2) - (P_t f)^2)]$. [Ces hypothèses techniques (fonctions bornées, stabilité par les fonctions C^∞ , stabilité par P_t , dérivation de P_t

identiquement en tout point) sont déraisonnables ; elles nous serviront à justifier tous les calculs formels (commutation d'intégrales et de passages à la limite, dérivation sous le signe somme, ...). Dans la pratique, le plus souvent, on dispose de plusieurs espaces de fonctions sur E, possédant chacun quelques unes de ces hypothèses, et il faut, dans les démonstrations, passer constamment d'un espace à l'autre. D'ailleurs, la plupart des démonstrations qui suivent n'emploient chacune qu'une partie des hypothèses ; nous ne cherchons pas à trier ce qui est utilisé ici ou là.]

Nous ferons également des hypothèses de nature probabiliste sur le comportement du processus :

__Diffusion__. Le carré du champ Γ est une dérivation, les formules de changement de variable du lemme 1 sont en vigueur, comme dans la proposition 1, pour toute fonction u de classe C^∞.

__Stationnarité__, __réversibilité__. Il existe une loi de probabilité μ sur E (la loi de X_t pour tout t) telle que $\underline{\underline{A}}$ soit incluse dans $L^2(E,\mu)$ (donc aussi dans tous les L^p) et que, pour f et g dans $\underline{\underline{A}}$, on ait $\langle Lf,g\rangle = \langle f,Lg\rangle$ (nous notons $\langle f,g\rangle$ l'intégrale $\int fg\,d\mu$; nous emploierons aussi $\langle f\rangle = \int f\,d\mu$; $\|f\|_p$ désignera la norme de f dans $L^p(\mu)$). Ceci implique que, pour $s\in[0,t]$, on a
$\frac{d}{ds}\langle P_{t-s}f,P_sg\rangle = \langle P_{t-s}f,LP_sg\rangle - \langle LP_{t-s}f,P_sg\rangle = 0$, donc $\langle P_tf,g\rangle = \langle f,P_tg\rangle$.
Les formules d'intégration par parties $\langle\Gamma(f,g)\rangle = -\langle f,Lg\rangle$ et $\langle\Gamma_2(f,g)\rangle = \langle Lf,Lg\rangle$ (qui résultent de $L1=0$) seront abondamment utilisées par la suite. Nous supposerons aussi que $\underline{\underline{A}}$ est dense dans tous les espaces L^p, et nous poserons

$$\|P_t\|_{p,q} = \sup_{f\in L^p} \|P_tf\|_q/\|f\|_p = \sup_{f\in \underline{\underline{A}}^+} \|P_tf\|_q/\|f\|_p ,$$

$\underline{\underline{A}}^+$ désignant les fonctions de $\underline{\underline{A}}$ telles que $\inf f>0$ (ou, ce qui revient au même, les fonction positives dont le logarithme est dans $\underline{\underline{A}}$).

PROPOSITION 4. __Soit__ $\lambda>0$. __Les six conditions suivantes sont équivalentes__ (elles traduisent toutes l'hypercontractivité) :

(9) __Pour tous__ $p>1$, $t\geq 0$, $1\leq q\leq 1+(p-1)e^{\lambda t}$, __on a__ $\|P_t\|_{p,q} \leq 1$.

(10) __Pour un__ $p>1$, __et pour tous__ $t\geq 0$, $1\leq q\leq 1+(p-1)e^{\lambda t}$, __on a__ $\|P_t\|_{p,q} \leq 1$.

(11) __Pour tous__ $t>0$, $1\leq q\leq e^{\lambda t}$, $f\in\underline{\underline{A}}^+$, __on a__ $\|\exp P_t \operatorname{Log} f\|_q \leq \|f\|_1$.

(12) En posant $U(x) = x \operatorname{Log} x$, on a, pour tous $t \geq 0$ et $f \in \underline{\underline{A}}^+$

$$<U \circ P_t f> \ \leq \ e^{-\lambda t} <U \circ f> + (1 - e^{-\lambda t}) U(<f>) \quad .$$

(13) Pour un $p \geq 1$ et toute $f \in \underline{\underline{A}}^+$

$$<f^p, \operatorname{Log} f> \ \leq \ <f^p> \operatorname{Log} \|f\|_p \ + \ \frac{p}{\lambda} <f^{p-2}, \Gamma(f,f)> \quad .$$

(14) Même condition, pour tout $p \geq 1$.

Les inégalités (13) et (14) sont les " inégalités de Sobolev logarithmiques " de Gross ; pour $p = 2$, elles montrent que f est dans $L^2 \operatorname{Log} L$ dès que f et $(\Gamma(f,f))^{\frac{1}{2}}$ sont dans L^2 , fournissant aux inégalités de Sobolev usuelles un substitut qui ne dépend pas de la dimension. Le point important dans cette proposition est l'équivalence entre ces inégalités de Sobolev logarithmiques et l'hypercontractivité exprimée par (9) ou (10) (théorème de Gross [2]). Le terme $\frac{p}{\lambda} <f^{p-2}, \Gamma(f,f)>$ de l'inégalité (13) peut, par changement de variable et intégration par parties, être réécrit sous la forme $- \frac{p}{\lambda(p-1)} <f^{p-1}, Lf>$ (pour $p = 1$: $- \frac{1}{\lambda} < \operatorname{Log} f, Lf>$) .

Démonstration.

(13) $\Leftrightarrow$ (14) : En remplaçant dans (13) f par $f^{r/p}$, on obtient

$$\frac{r}{p} <f^r, \operatorname{Log} f> \ \leq \ \frac{r}{p} <f^r> \operatorname{Log} \|f\|_r \ + \ \frac{p}{\lambda} <f^{r-2\frac{r}{p}}, (\frac{r}{p} f^{\frac{r}{p}-1})^2 \Gamma(f,f)>$$

qui n'est autre que la même inégalité écrite pour r ; ces inégalités, énoncées pour les différentes valeurs de p , sont donc toutes équivalentes entre elles (ceci reste vrai pour p plus petit que 1).

(9) $\Rightarrow$ (10) $\Rightarrow$ (13), (14) $\Rightarrow$ (9) : Pour $q(t) = 1 + (p-1)e^{\lambda t}$, en calculant la dérivée $\frac{d}{dt} \operatorname{Log} \|P_t f\|_{q(t)}$, on trouve, en posant pour abréger $h = P_t f$,

$$- \frac{q'}{q^2} \log <h^q> \ + \ \frac{1}{q <h^q>} \left[<qh^{q-1}, Lh> + <q' h^q, \operatorname{Log} h> \right]$$

où $q' = q'(t) = \lambda(q-1)$. [Les dérivations sous le signe somme sont justifiées par nos lourdes hypothèses, qui assurent que tout est borné. On pourrait s'en tirer à moindres frais, par exemple en contrôlant les quantités de la forme $\sup_t P_t g$ à l'aide du lemme de Rota.] Ceci peut se mettre sous une forme faisant apparaître l'inégalité de Sobolev logarithmique :

$$\frac{\lambda(q-1)}{q} \quad \frac{1}{\langle h^q\rangle} \left[\langle h^q, \mathrm{Log}\, h\rangle - \langle h^q\rangle \mathrm{Log}\, \|h\|_q + \frac{q}{\lambda(q-1)} \langle h^{q-1}, Lh\rangle \right] \quad .$$

Si l'on a l'hypercontractivité sous la forme (10), cette dérivée pour $t=0$ ne peut être que négative, d'où (13). Réciproquement, si l'on a (14), cette dérivée est négative pour tout t, la quantité $\|P_t f\|_{q(t)}$ est fonction décroissante de t pour f dans $\underline{\underline{A}}^+$, d'où l'inégalité $\|P_t f\|_{q(t)} \leqq \|f\|_{q(0)} = \|f\|_p$ vraie pour f dans $\underline{\underline{A}}^+$, donc aussi pour toute f dans L^p, et (9) est établie.

<u>(11) ⇒ (13)</u>, <u>(14) ⇒ (11)</u> : L'argument est très semblable. En posant $q(t) = \exp(\lambda t)$ et $\exp(q(t) P_t \mathrm{Log}\, f) = h$, on écrit

$$\begin{aligned}
\frac{d}{dt} \mathrm{Log}\, \| \exp P_t \mathrm{Log}\, f \|_{q(t)} &= \frac{d}{dt}\left(\frac{1}{q} \mathrm{Log}\, \langle h\rangle\right) \\
&= -\frac{q'}{q^2} \mathrm{Log} \langle h\rangle + \frac{1}{q\langle h\rangle} \langle h, q' P_t \mathrm{Log}\, f + qLP_t \mathrm{Log}\, f\rangle \\
&= \frac{\lambda}{q\langle h\rangle} \left[-\langle h\rangle \mathrm{Log} \langle h\rangle + \langle h, \mathrm{Log}\, h\rangle + \frac{1}{\lambda} \langle h, L\, \mathrm{Log}\, h\rangle \right] \quad ;
\end{aligned}$$

le terme $\langle h, L\,\mathrm{Log}\, h\rangle$ s'intègre par parties et donne $\langle Lh, \mathrm{Log}\, h\rangle$, ce qui fait apparaître l'inégalité de Sobolev logarithmique correspondant à $p=1$, et permet, comme ci-dessus, de conclure à l'équivalence.

<u>(12) ⇒ (13)</u>, <u>(14) ⇒ (12)</u> : La méthode est encore la même, en utilisant cette fois

$$\begin{aligned}
&\frac{d}{dt} \left[e^{\lambda t} \left(\langle U \circ P_t f\rangle - U(\langle f\rangle) \right) \right] \\
&= e^{\lambda t} \left[\lambda(\langle U \circ P_t f\rangle - U(\langle f\rangle)) + \langle U' \circ P_t f, LP_t f\rangle \right] \\
&= \lambda\, e^{\lambda t} \left[\langle P_t f, \mathrm{Log}\, P_t f\rangle - \langle f\rangle \mathrm{Log} \langle f\rangle + \frac{1}{\lambda} \langle \mathrm{Log}\, P_t f, LP_t f\rangle \right] \quad .
\end{aligned}$$

Puisque $\langle f\rangle = \langle P_t f\rangle$, on est encore ramené à l'inégalité de Sobolev logarithmique pour $p=1$. La proposition est ainsi entièrement démontrée. ∎

CONDITION SUFFISANTE D'HYPERCONTRACTIVITÉ

Nous restons dans le cadre qui nous a permis d'établir la proposition 4 (une grande algèbre fourre-tout) avec en particulier les hypothèses de diffusion et de stationnarité et réversibilité du processus, auxquelles nous ajoutons

Ergodicité : Pour f dans $\underline{\underline{A}}$, $Lf = 0 \Rightarrow f = \text{constante}$. Ceci entraîne que, pour toute f, $P_t f$ tend vers $\langle f \rangle$ μ-presque partout (et donc dans tous les L^p) quand t tend vers l'infini.

Avant de continuer, nous groupons sous forme de lemmes des bribes de calculs qui resserviront plusieurs fois.

LEMME 3. La formule de changement de variable pour Γ_2 (en vigueur, bien sûr, dès que Γ est une dérivation) s'énonce

$$\Gamma_2(u\circ f, u\circ f) = (u'\circ f)^2 \Gamma_2(f,f) + (u'\circ f)(u''\circ f)\,\Gamma(f,\Gamma(f,f)) + (u''\circ f)^2\, \Gamma^2(f,f) \quad . \tag{15}$$

Démonstration. Posons, pour simplifier la typographie, $U' = u'\circ f$; $U'' = u''\circ f$; $U''' = u'''\circ f$. Il vient

$$\begin{aligned}
\Gamma_2(u\circ f, u\circ f) &= \tfrac{1}{2} L[U'^2\Gamma(f,f)] - \Gamma(u\circ f\,, U'Lf + U''\Gamma(f,f)) \\
&= \tfrac{1}{2} U'^2 L\Gamma(f,f) + \tfrac{1}{2}\Gamma(f,f)\,L(U'^2) + \Gamma(U'^2, \Gamma(f,f)) \\
&\qquad - U'\,\Gamma(f\,, U'Lf + U''\Gamma(f,f)) \\
&= \tfrac{1}{2} U'^2 L\Gamma(f,f) + U'U''Lf\,\Gamma(f,f) + (U'U''' + U''^2)\Gamma^2(f,f) \\
&\qquad + 2\,U'U''\,\Gamma(f,\Gamma(f,f)) \\
&\qquad - U'^2\Gamma(f,Lf) - U'U''\,Lf\,\Gamma(f,f) - U'U''\,\Gamma(f,\Gamma(f,f)) \\
&\qquad - U'U'''\,\Gamma^2(f,f) \quad ,
\end{aligned}$$

d'où le résultat. ∎

LEMME 4. Soient u une fonction de classe C^∞ sur un intervalle ouvert I de $\mathbb{R}$, f une fonction de $\underline{\underline{A}}$ prenant ses valeurs dans un compact de I . Alors

$$\langle u\circ f, \Gamma(f,Lf)\rangle = -\langle u\circ f, \Gamma_2(f,f)\rangle - \tfrac{1}{2}\langle u'\circ f, \Gamma(f,\Gamma(f,f))\rangle \quad ; \tag{16}$$

$$\langle u'\circ f, Lf\,\Gamma(f,f)\rangle = -\langle u'\circ f, \Gamma(f,\Gamma(f,f))\rangle - \langle u''\circ f, \Gamma^2(f,f)\rangle \quad ; \tag{17}$$

$$\langle u\circ f, (Lf)^2\rangle = \langle u\circ f, \Gamma_2(f,f)\rangle + \tfrac{3}{2}\langle u'\circ f, \Gamma(f,\Gamma(f,f))\rangle + \langle u''\circ f, \Gamma^2(f,f)\rangle \; . \tag{18}$$

Démonstration. Elle se fait par changements de variable (qu'il est facile de justifier en remplaçant u par une fonction C^∞ sur toute la droite, et qui coïncide avec u sur l'image de f) et par intégrations par parties.

$$\begin{aligned} <u\circ f,\Gamma(f,Lf)> &= -<u\circ f,\Gamma_2(f,f)> + \tfrac{1}{2}<u\circ f,L\Gamma(f,f)> \\ &= -<u\circ f,\Gamma_2(f,f)> - \tfrac{1}{2}<\Gamma(u\circ f,\Gamma(f,f))> \quad , \text{ d'où (16).} \end{aligned}$$

Nous conservons les notations U' , U'' de la démonstration précédente.

$$\begin{aligned} <U',Lf\,\Gamma(f,f)> &= <Lf,U'\Gamma(f,f)> = -<\Gamma(f,U'\Gamma(f,f))> \\ &= -<U',\Gamma(f,\Gamma(f,f))> - <\Gamma(f,f),\Gamma(f,U')> \quad , \text{ d'où (17).} \end{aligned}$$

$$\begin{aligned} <u\circ f,(Lf)^2> &= <Lf\,,\,u\circ f\,Lf> = -<\Gamma(f\,,\,u\circ f\,Lf)> \\ &= -<u\circ f,\Gamma(f,Lf)> - <U',Lf\,\Gamma(f,f)> \quad , \end{aligned}$$

et (18) résulte de (16) et (17). ∎

Nous sommes maintenant en mesure de donner une condition suffisante d'hypercontractivité, condition technique apparemment invérifiable. La suite sera consacrée à la recherche d'hypothèses maniables assurant cette condition.

PROPOSITION 5. Soient $\lambda>0$, U une fonction convexe C^∞ définie sur un intervalle ouvert I , et $u=U''\geq 0$ sa dérivée seconde. On suppose que, pour toute fonction f de $\underline{\underline{A}}$ à valeurs dans un compact de I , on ait

(19) $$<u\circ f,\Gamma_2(f,f)-\tfrac{\lambda}{2}\Gamma(f,f)> + <u'\circ f,\Gamma(f,\Gamma(f,f))> + \tfrac{1}{2}<u''\circ f,\Gamma^2(f,f)> \geq 0 \quad .$$

Alors on a aussi

(20) $$<U\circ f> - U(<f>) \leq \frac{1}{\lambda}<u\circ f,\Gamma(f,f)>$$

et, pour tout $t\geq 0$,

(21) $$<U\circ P_t f> \leq e^{-\lambda t}<U\circ f> + (1-e^{-\lambda t})\,U(<f>) \quad .$$

Démonstration. Les conclusions (20) et (21) — qui expriment l'hypercontractivité lorsque $U=x\operatorname{Log}x$ — sont équivalentes : cela se vérifie exactement comme l'équivalence (12) ⇔ inégalité de Sobolev logarithmique dans la proposition 4.

Les formules (16) et (17) permettent de remplacer l'hypothèse par

$$<u'\circ f,\,Lf\,\Gamma(f,f)> + 2<u\circ f,\Gamma(f,Lf)> + \lambda<u\circ f,\Gamma(f,f)> \leq 0 \quad ;$$

en y substituant P_tf à f , on obtient

$$\frac{d}{dt}<u\circ P_tf,\Gamma(P_tf,P_tf)> + \lambda<u\circ P_tf,\Gamma(P_tf,P_tf)> \leq 0 \quad ,$$

ou encore $\frac{d}{dt}[\, e^{\lambda t} \langle u\circ P_t f, \Gamma(P_t f, P_t f)\rangle \,] \leqq 0$. On en déduit

$$\langle u\circ P_t f, \Gamma(P_t f, P_t f)\rangle \leqq e^{-\lambda t} \langle u\circ f, \Gamma(f,f)\rangle \quad ,$$

ce qui permet d'écrire

$$\begin{aligned} \langle U\circ f\rangle - U(\langle f\rangle) &= \langle U\circ P_0 f\rangle - \langle U\circ P_\infty f\rangle = -\int_0^\infty \left(\frac{d}{dt}\langle U\circ P_t f\rangle\right) dt \\ &= -\int_0^\infty \langle U'\circ P_t f, LP_t f\rangle\, dt = \int_0^\infty \langle \Gamma(U'\circ P_t f, P_t f)\rangle\, dt \\ &= \int_0^\infty \langle u\circ P_t f, \Gamma(P_t f, P_t f)\rangle\, dt \\ &\leqq \langle u\circ f, \Gamma(f,f)\rangle \int_0^\infty e^{-\lambda t}\, dt = \frac{1}{\lambda}\langle u\circ f, \Gamma(f,f)\rangle \quad , \end{aligned}$$

et la proposition est démontrée. ∎

Les deux critères que nous allons en tirer sont probablement bien plus importants que les raffinements qui vont suivre.

COROLLAIRE 1. <u>Si l'on a, pour toute</u> f <u>dans</u> $\underline{\underline{A}}$,

(22) $\qquad \langle e^f, \Gamma_2(f,f) - \frac{\lambda}{2}\Gamma(f,f)\rangle \geqq 0 \quad ,$

<u>alors a lieu l'hypercontractivité avec constante</u> λ .

<u>Démonstration</u>. La formule de changement de variable (15) pour $f = \operatorname{Log} g$ donne

$$\begin{aligned} & e^f\, [\Gamma_2(f,f) - \tfrac{\lambda}{2}\Gamma(f,f)] \\ &\quad = g\, [g^{-2}(\Gamma_2(g,g) - \tfrac{\lambda}{2}\Gamma(g,g)) - g^{-3}\Gamma(g,\Gamma(g,g)) + g^{-4}\Gamma^2(g,g)\,] \;, \end{aligned}$$

donc l'hypothèse (19) de la proposition est vérifiée avec $I =]0,\infty[$, $u = \frac{1}{x}$ et $U = x\operatorname{Log} x$. Dans ce cas, la conclusion (20) n'est autre que l'inégalité de Sobolev logarithmique avec $p = 1$. ∎

COROLLAIRE 2. <u>Si l'on a, pour toute</u> f <u>dans</u> $\underline{\underline{A}}$,

(23) $\qquad \Gamma_2(f,f) \geqq \frac{\lambda}{2}\Gamma(f,f) \quad ,$

<u>alors l'hypercontractivité a lieu, avec constante</u> λ .

C'est une conséquence immédiate du précédent.

Exemples. a) <u>Processus d'Ornstein-Uhlenbeck</u>. Nous avons vu plus haut que les générateurs du type d'Ornstein-Uhlenbeck vérifient

$$\Gamma_2(f,f) - \Gamma(f,f) = \sum_{i,j} (D_i D_j f)^2 \geqq 0 \quad ;$$

ils admettent donc, lorsque nos calculs sont justifiés, 2 pour constante d'hypercontractivité. Par exemple, le semi-groupe sur $\mathbb{R}$, de générateur $L = D^2 - xD$, est symétrique par rapport à la loi gaussienne standard μ ; il opère sur les polynômes d'Hermite H_n (ce sont ici les polynômes d'Hermite usuels, à une variable) par $LH_n = -n H_n$. Le semi-groupe est défini dans $L^2(\mu)$ par $P_t H_n = e^{-nt} H_n$; il est explicitement donné par le noyau $p_t(x,y) = [2\pi(1-e^{-t})]^{-\frac{1}{2}} \exp - \frac{(y-e^{-t}x)^2}{2(1-e^{-t})}$.

Ici, l'algèbre naturelle serait celle des polynômes, mais elle ne vérifie pas nos hypothèses ; il est plus commode de travailler sur l'algèbre des fonctions de la forme $a+s$, où a est une constante et s une fonction de Schwartz. Le corollaire 2 s'applique dans ce cadre, et l'hypercontractivité a lieu.

b) Mouvements browniens sur les sphères n-dimensionnelles. (Nous les prendrons de générateur Δ et non $\frac{1}{2}\Delta$; pour retrouver le cas probabiliste usuel, le lecteur devra donc diviser par 2 nos constantes d'hypercontractivité.) Dans ce cas, l'algèbre est constituée de toutes les fonctions C^∞, et, puisque le tenseur de courbure de Ricci vaut $Ric_{ij} = (n-1)r^{-2} g_{ij}$ (où r est le rayon de la sphère), on a, par la proposition 3, $\Gamma_2(f,f) \geq \frac{n-1}{r^2}\Gamma(f,f)$, ce qui fournit la constante d'hypercontractivité $\lambda = 2(n-1)r^{-2}$. Ceci n'est pas optimal : la meilleure constante est connue (Mueller-Weissler [6]) et vaut $2nr^{-2}$. Nous allons voir bientôt comment on peut retrouver ce résultat à l'aide de la proposition 5 ; remarquons pour l'instant que le cas du cercle ($n=1$) nous échappe, puisque nous trouvons $\lambda = 0$. C'est dû au fait que l'hypothèse utilisée (23) est locale, alors que l'hypercontractivité du brownien circulaire est une propriété globale (la constante d'hypercontractivité dépend du rayon r, et tend vers zéro avec $\frac{1}{r}$; donc aucune méthode purement locale ne suffit). Ceci suggère d'essayer plutôt d'employer le corollaire 1 ; dans cet ordre d'idées, nous avons seulement réussi à établir $\langle e^f, f''^2 - \frac{1}{2} f'^2\rangle \geq 0$ (ici, $E = S^1 = \mathbb{R}/2\pi\mathbb{Z}$), qui fournit l'hypercontractivité avec $\lambda = 1$; alors que l'on sait (Rothaus [9], Weissler [13]) que la valeur optimale est 2. S'il existait des fonctions f sur S^1 ne vérifiant pas l'inégalité $\langle e^f, f''^2 - f'^2\rangle \geq 0$, cela montrerait que la condition suffisante (22) du corollaire 1 n'est pas nécessaire.

Voici un résultat un peu plus précis que le corollaire 2 : il fournit, comme la proposition 5, une gamme d'inégalités incluant l'hypercontractivité, et améliore un peu les constantes. La démonstration consistera à vérifier l'hypothèse technique (19) de la proposition 5.

THEOREME. 1) On suppose que, pour deux constantes $a>0$ et $b\in[0,1[$, on ait, pour toute f de $\underline{\underline{A}}$,

(24) $$\Gamma_2(f,f) \geq a\,\Gamma(f,f) + b\,(Lf)^2 \quad .$$

Alors, si U est une fonction C^∞ et convexe définie sur un intervalle ouvert I , dont la dérivée seconde $u = U''$ est strictement positive et d'inverse $\frac{1}{u}$ concave, on a, pour toute f de $\underline{\underline{A}}$ à valeurs dans un compact de I ,

(20) $$\langle U\circ f\rangle - U(\langle f\rangle) \leq \frac{1}{\lambda}\langle u\circ f, \Gamma(f,f)\rangle \quad ,$$

avec $\lambda = \frac{2a}{1-b}$. En particulier, pour $U(x) = x\,\mathrm{Log}\,x$, l'hypercontractivité a lieu.

2) Le même résultat subsiste si l'hypothèse (24) est remplacée par

(25) $$\Gamma_2(f,f) \leq -a\,\Gamma(f,f) + b\,(Lf)^2$$

avec des constantes $a>0$, $b\in\,]1,4]$. La constante obtenue est dans ce cas $\lambda = \frac{2a}{b-1}$.

REMARQUE. Les hypothèses $b<1$ dans le premier cas et $b>1$ dans le second sont automatiquement satisfaites : elles découlent de $\Gamma(f,f)\geq 0$ et de $\langle\Gamma_2(f,f)\rangle = \langle(Lf)^2\rangle$ (intégrer sur E (24) et (25)). Par contre la limitation $b\leq 4$ nous semble artificielle, et plus probablement due à notre méthode de calcul qu'à la nature des choses.

Démonstration. Premier cas (hypothèse (24)). Cette hypothèse appliquée à $v\circ f$ (où v est un polynôme du second degré) donne, par la formule (15) de changement de variable pour Γ_2 ,

$$\begin{aligned}(v'\circ f)^2\Gamma_2(f,f) + (v'\circ f)(v''\circ f)\Gamma(f,\Gamma(f,f)) + (v''\circ f)^2\Gamma^2(f,f)\\ \geq a(v'\circ f)^2\Gamma(f,f) + b\,[(v'\circ f)Lf + (v''\circ f)\Gamma(f,f)]^2 \quad .\end{aligned}$$

Fixons $x\in E$ et $f\in\underline{\underline{A}}$. Les deux nombres $v'\circ f(x)$ et $v''\circ f(x)$ peuvent, par un choix approprié de v , être pris égaux à deux réels α et β donnés arbitrairement a priori. Donc la forme quadratique en α et β

$$A\alpha^2 + B\alpha\beta + C\beta^2 = \alpha^2\Gamma_2(f,f)(x) + \alpha\beta\Gamma(f,\Gamma(f,f))(x) + \beta^2\Gamma^2(f,f)(x) - \alpha^2 a\Gamma(f,f)(x) - b[\alpha(Lf(x))^2 + \beta\Gamma(f,f)(x)]^2$$

est positive, ce qui revient à dire que la matrice $\begin{pmatrix} A & \frac{1}{2}B \\ \frac{1}{2}B & C \end{pmatrix}$ est de type positif.

Mais, par concavité de $\frac{1}{u}$, $uu'' \geq 2u'^2$, donc, pour k et ξ réels, la matrice $\begin{pmatrix} u(\xi) & ku'(\xi) \\ ku'(\xi) & \frac{1}{2}k^2u''(\xi) \end{pmatrix}$ est aussi de type positif. Il en résulte

$$Au(\xi) + kBu'(\xi) + \tfrac{1}{2}k^2Cu''(\xi) \geq 0 \quad ,$$

comme trace du produit de deux matrices de type positif. Prenant $\xi = f(x)$, ceci démontre que la fonction

$$\begin{aligned} g = {} & (u\circ f)\Gamma_2(f,f) + k(u'\circ f)\Gamma(f,\Gamma(f,f)) + \tfrac{1}{2}k^2(u''\circ f)\Gamma^2(f,f) \\ & - a(u\circ f)\Gamma(f,f) - b\,[\,(u\circ f)(Lf)^2 + 2k(u'\circ f)\,Lf\,\Gamma(f,f) \\ & + \tfrac{1}{2}k^2(u''\circ f)\Gamma^2(f,f)\,] \end{aligned}$$

est positive au point fixé x. Elle est donc positive partout, et ceci entraîne $\langle g\rangle \geq 0$. L'utilisation des formules (17) et (18) permet d'écrire

$$\frac{1}{1-b}\langle g\rangle = \langle u\circ f,\Gamma_2(f,f)\rangle - \frac{a}{1-b}\langle u\circ f,\Gamma(f,f)\rangle + r(b,k)\,\langle u'\circ f,\Gamma(f,\Gamma(f,f))\rangle + s(b,k)\,\langle u''\circ f,\Gamma^2(f,f)\rangle \geq 0$$

où les coefficients valent respectivement $r = \frac{1}{1-b}(k - \frac{3}{2}b + 2kb)$ et $s = \frac{1}{2}k^2 + \frac{b}{1-b}(2k-1)$. Cette inégalité a lieu pour tout k ; en lui fixant la valeur $k = (1+\frac{b}{2})/(1+2b)$, on obtient $r = 1$, d'où

$$\langle u\circ f,\Gamma_2(f,f) - \tfrac{\lambda}{2}\Gamma(f,f)\rangle + \langle u'\circ f,\Gamma(f,\Gamma(f,f))\rangle + s\langle u''\circ f,\Gamma^2(f,f)\rangle \geq 0 \quad .$$

Ceci est presque l'hypothèse (19) de la proposition 5 : nous avons le coefficient s au lieu de $\frac{1}{2}$ devant le dernier terme. Mais ce terme $\langle u''\circ f,\Gamma^2(f,f)\rangle$ est toujours positif, car la concavité de $\frac{1}{u}$ implique que u est convexe. D'autre part, la valeur choisie pour k donne, par un calcul aisé, $s \leq \frac{1}{2}$, ce qui entraîne a fortiori la condition (19). Il ne reste plus qu'à appliquer la proposition 5.

Deuxième cas (hypothèse (25)). On procède de façon tout-à-fait semblable. La dernière étape ramène aussi à vérifier que, pour $k = (1+\frac{b}{2})/(1+2b)$, le coefficient $s = \frac{1}{2}k^2 + \frac{b}{b-1}(1-2k)$ est majoré par $\frac{1}{2}$; ceci n'est vrai que pour $b \leq 4$, d'où la restriction. ∎

Puisque le théorème donne des résultats pour d'autres fonctions convexes que $U = x \operatorname{Log} x$, il est tentant de l'appliquer aux fonctions puissance. De fait, on vérifie que $U(x) = x^p$ satisfait les conditions du théorème pour $1 < p \leq 2$, sur $I =]0,\infty[$ ($I = \mathbb{R}$ pour $p=2$). Sous la forme intégrée (21), la conclusion peut s'énoncer

$$<(P_t f)^p> - <f>^p \leq e^{-\lambda t}(<f^p> - <f>^p) \qquad (f \geq 0 \text{ ou } p=2) .$$

Pour $p=2$, c'est presque une trivialité : cette inégalité exprime que, pour f d'intégrale nulle (i.e. orthogonale aux constantes, qui forment le noyau de L), $\|P_t f\|_2$ décroît exponentiellement vers zéro, en $e^{-\frac{1}{2}\lambda t}$; ceci traduit simplement un trou entre 0 et $\frac{1}{2}\lambda$ dans le spectre de $-L$. Or ceci s'obtient directement en intégrant sur E l'hypothèse (24) ou (25), car l'inégalité $<\Gamma_2(f,f)> \geq \frac{\lambda}{2} <\Gamma(f,f)>$ équivaut elle aussi à cette lacune spectrale.

Comme nous l'avions annoncé, ce théorème permet de retrouver la constante d'hypercontractivité optimale pour le semi-groupe brownien sur une sphère n-dimensionnelle ($n \geq 2$) : on a dans ce cas, par la proposition 3,

$$\Gamma_2(f,f) = \|\operatorname{Hess} f\|^2 + \operatorname{Ric}(\operatorname{grad} f, \operatorname{grad} f) \quad ;$$

nous avons vu que le terme de courbure est égal à $(n-1)r^{-2}\Gamma(f,f)$; puisque Δf est la trace de $\operatorname{Hess} f$, on a $\|\operatorname{Hess} f\|^2 \geq \frac{1}{n}(\Delta f)^2$ (c'est simplement l'identité $\Sigma_{i,j}(H_{ij})^2 \geq \frac{1}{n}(\Sigma_i H_{ii})^2$), et le théorème s'applique avec $a = (n-1)r^{-2}$ et $b = \frac{1}{n}$, d'où $\lambda = \frac{2a}{1-b} = 2nr^{-2}$ (après simplification par $n-1$: ceci ne donne toujours pas l'hypercontractivité du cercle — d'ailleurs, nous avons vu qu'elle n'est pas du ressort de nos méthodes locales).

Plus généralement, les mêmes considérations montrent que <u>sur une variété n-dimensionnelle compacte dont la courbure de Ricci a toutes ses valeurs propres minorées par</u> $\varepsilon > 0$, <u>le mouvement brownien est hypercontractif</u>, <u>avec</u> $\lambda = \frac{2\varepsilon}{1-1/n}$.

Nos méthodes ne permettent pas de sortir du cas où la courbure est positive, contrairement à Rothaus [10] qui prouve l'hypercontractivité sur toutes les variétés riemanniennes compactes. En revanche, elles donnent une estimation géométrique simple de λ et sont relativement robustes : si l'on ajoute à Δ un champ de gradients suffisamment lipschitzien, le résultat subsiste.

Un autre exemple est l'hypercontractivité des semi-groupes ultrasphériques, étudiée par Mueller et Weissler [6] : le laplacien de la sphère n-dimensionnelle projeté sur l'intervalle $[-1,1]$ (considéré comme diamètre de cette sphère) devient l'opérateur

$$Lf(x) = (1-x^2)f''(x) - nxf'(x) \quad .$$

défini sur $C^\infty([-1,1])$. Cet opérateur peut être écrit pour n non entier (il perd alors son interprétation géométrique) ; nous supposerons $n>0$ pour que L soit symétrique par rapport à la mesure de probabilité $\mu(dx) = C(n)(1-x^2)^{\frac{n}{2}-1}dx$. Nous avons ici

$$\Gamma(f,f) = (1-x^2)f'^2(x) \quad ;$$
$$\Gamma_2(f,f) = [(1-x^2)f''(x) - xf'(x)]^2 + (n-1)f'^2(x) \quad ,$$

d'où

$$\Gamma_2(f,f) \geq (n-1)\Gamma(f,f) + \frac{1}{n}(Lf)^2 \qquad \text{si} \quad n>1 \quad ,$$
$$\Gamma_2(f,f) \leq -(1-n)\Gamma(f,f) + \frac{1}{n}(Lf)^2 \qquad \text{si} \quad 0<n<1 \quad .$$

Si $(J_k)_{k\geq 0}$ désigne la suite des polynômes de Jacobi (polynômes orthogonaux pour μ , normalisés dans $L^2(\mu)$), alors $LJ_k = -\lambda_k J_k$ pour $\lambda_k = k(k+n-1)$. Le semi-groupe est défini sur $L^2(\mu)$ par $P_t J_k = e^{-\lambda_k t} J_k$; on a évidemment $P_t 1 = 1$, et le problème est de vérifier la positivité des P_t à partir de celle de Γ . Rappelons brièvement la méthode de Mueller et Weissler. On montre tout d'abord que, pour tout polynôme f , $\langle|P_t f|\rangle \leq \langle|f|\rangle$ de la façon suivante : si $\varphi_k(x)$ est une suite de fonctions positives, C^∞, convexes, qui croît vers $|x|$,

$$\frac{d}{dt}\langle\varphi_k \circ P_t f\rangle = \langle\varphi_k' \circ P_t f, LP_t f\rangle = -\langle\Gamma(\varphi_k' \circ P_t f, P_t f)\rangle$$
$$= -\langle\varphi_k'' \circ P_t f, \Gamma(P_t f, P_t f)\rangle \leq 0 \quad ,$$

d'où $\langle\varphi_k \circ P_t f\rangle \leq \langle\varphi_k \circ f\rangle$, et, à la limite, $\langle|P_t f|\rangle \leq \langle|f|\rangle$. La positivité de $P_t f$ pour $f\geq 0$ découle alors de $\langle|P_t f|\rangle \leq \langle|f|\rangle = \langle f\rangle = \langle P_t f\rangle$.

Pour vérifier l'hypercontractivité, on a le choix quant à l'algèbre $\underline{\underline{A}}$: on peut prendre $C^\infty([-1,1])$, mais il faut alors établir qu'elle est stable par P_t (en fait, Mueller et Weissler démontrent, dans leur lemme 1.16, que pour $t>0$, P_t envoie $L^2(\mu)$ dans $C^\infty([-1,1])$). On peut aussi prendre l'algèbre plus petite des polynômes, mais elle n'est pas stable par composition avec les fonctions C^∞ , ce qui oblige à passer constamment d'une algèbre à l'autre dans les démonstrations, mais présente l'avantage d'éviter le recours au lemme analytique de Mueller et Weissler.

En tout état de cause, à l'aide des estimations précédentes sur Γ_2, le théorème donne l'hypercontractivité avec $\lambda = 2n$, les deux cas $n > 1$ et $n < 1$ étant couverts séparément par les deux parties du théorème (le cas $n = 1$ s'obtient par passage à la limite dans l'équation (19), ou plus facilement, dans (22) qui lui est équivalente ; remarquons qu'il s'agit là de l'hypercontractivité du mouvement brownien sur l'intervalle $[-\frac{\pi}{2}, \frac{\pi}{2}]$ réfléchi aux deux extrémités). Mais, en raison de la restriction $b \leqq 4$ dans le théorème, ceci ne marche que pour $n \geqq \frac{1}{4}$, et, pour les petites "dimensions" $0 < n < \frac{1}{4}$, nous ne retrouvons pas le résultat de Mueller et Weissler.

C'est aussi l'occasion de remarquer que l'hypothèse (25) que nous venons d'utiliser ($\Gamma_2(f,f) \leqq -a\Gamma(f,f) + b(Lf)^2$, $b > 1$) est spécifique de la dimension un ; plus précisément, si (25) est satisfaite pour un générateur du second ordre non dégénéré sur une variété, alors, en appliquant la proposition 3 à une fonction de gradient nul en x, mais de hessienne arbitraire, on trouve une majoration du carré de la hessienne $\|\mathrm{Hess}\, f\|^2$ par le carré de sa trace $(\Delta f)^2$ à une constante près, ce qui n'est possible qu'en dimension 1.

Nous avons systématiquement utilisé les formules de changement de variable, qui expriment la continuité du processus. Sans cette hypothèse, la situation est bien moins claire. Le cas de l'espace à deux points (exemple e) du deuxième paragraphe) pourrait faire croire que le théorème se laisse généraliser aux processus à sauts, car $\Gamma_2 = 2\Gamma$ et on vérifie, par un calcul direct (Gross [2]), l'inégalité de Sobolev logarithmique et l'hypercontractivité. Mais l'exemple g) qui suit exhibe le phénomène inverse : bien que Γ_2 soit minoré par $\frac{1}{2}\Gamma$, Surgailis [12] a établi que le processus n'est pas hypercontractif. Ceci se voit facilement sur les fonctions "exponentielles" sur E, de la forme

$$f_a(x) = \exp[-\int a\, dm] \prod_{u \in x} (1 + a(u))$$

où a est une fonction sur S. Sur ces fonctions, le semi-groupe est donné par $P_t f_a = f_{ae^{-t}}$; en prenant pour a une fonction constante, que l'on fait tendre vers l'infini, il apparaît que, pour tous $t>0$ et $1\leq p<q$, on a $\|P_t\|_{p,q} = \infty$ (où les normes sont calculées pour la probabilité μ sur E qui est la loi du processus ponctuel de Poisson d'intensité m) ; on observe aussi que, sous la forme faisant intervenir L (ainsi que, pour $p\geq 2$, sous la forme faisant intervenir Γ), les inégalités de Sobolev logarithmiques sont invalides.

RÉFÉRENCES

[1] M. Berger, P. Gauduchon et E. Mazet. Le spectre d'une variété riemannienne. Lecture Notes in Math. 194, Springer.

[2] L. Gross. Logarithmic Sobolev Inequalities. Amer J. Math. 97, 1975.

[3] P. A. Meyer. Note sur les processus d'Ornstein-Uhlenbeck. Sém. Prob. XVI, Lecture Notes in Math. 920, Springer.

[4] P. A. Meyer. Géométrie différentielle stochastique (bis). Sém. Prob. XVI B, Lecture Notes in Math. 921, Springer.

[5] P. A. Meyer et W. A. Zheng. Tightness criteria for laws of semimartingales. Ann. I.H.P. (à paraître).

[6] C. Mueller et F. Weissler. Hypercontractivity for the Heat Semigroup for Ultraspheric Polynomials and on the n-Sphere. J. Func. Anal. 48, 1982.

[7] E. Nelson. The free Markov Field. J. Funct. Anal. 12, 1973.

[8] J. Neveu. Sur l'espérance conditionnelle par rapport à un mouvement brownien. Ann. I.H.P. 2, 1976.

[9] O. Rothaus. Logarithmic Sobolev inequalities and the spectrum of Sturm-Liouville operators. J. Func. Anal. 39, 1980.

[10] O. Rothaus. Diffusion on compact Riemannian manifolds and logarithmic Sobolev inequalities. J. Func. Anal 42, 1981.

[11] J. Ruiz de Chavez. Thèse de troisième cycle. Strasbourg (à paraître).

[12] D. Surgailis. On Poisson multiple stochastic integrals and associated equilibrium Markov processes. Proc. IFIP-ISI international conf. on random fields, Bangalore 1982. Lect. Notes in Inf. Control 49, Springer.

[13] F. Weissler. Logarithmic Sobolev inequalities and hypercontractive estimates on the circle. J. Func. Anal. 37, 1980.