GRADIENT ESTIMATES FOR SOME DIFFUSION SEMIGROUPS

Jean Picard

Laboratoire de Mathmatiques Appliques (CNRS-UMR 6620)

Universit Blaise Pascal

63177 Aubire Cedex, France

E-mail: `Jean.Picard@math.univ-bpclermont.fr`

Abstract. Consider the semigroup P_t of an elliptic diffusion; we describe a simple stochastic method providing gradient estimates on $P_t f$. If N is a manifold endowed with a connection, the method can also be applied to the associated nonlinear semigroup Q_t acting on N-valued maps. With a localization technique, we deduce gradient estimates for real harmonic functions or N-valued harmonic maps. Moreover, the results are extended to a class of hypoelliptic diffusions.

Keywords. Gradient estimates, Harmonic functions, Harmonic maps, Elliptic diffusions, Hypoelliptic diffusions.

Mathematics Subject Classification (2000). 60J60 60H07 58E20 58J65

1 Introduction

Consider the semigroup $P_t = e^{tL}$ of a diffusion with generator L on a manifold M. If f is a real-valued function defined on M, the derivative $d(P_t f)$ of the function $x \mapsto P_t f(x)$ takes its values in the cotangent bundle $T^\star M$, and we are interested in estimations of this derivative with stochastic methods. More precisely, we look for estimates of the form

$$\big|d(P_t f)\big| \le C_{q,t}\, P_t(|f|^q)^{1/q} \tag{1.1}$$

for $t > 0$ and $q > 1$. This problem can be generalised in several ways. First, it can be localized; the function $h(t,x) = P_t f(x)$ is solution of the heat equation

$$\frac{\partial}{\partial t} h = Lh, \qquad h(0,.) = f \tag{1.2}$$

on $\mathbb{R}_+ \times M$, and if D is an open subset of M, one can consider a function h which is only solution on $\mathbb{R}_+ \times D$, and estimate its derivative $dh(t,x)$ with respect to x. For instance, by considering functions $h(x)$ which do not depend on t, we want to obtain estimates for the derivative of functions which are harmonic on D. Notice that such functions have a well known stochastic interpretation; they map the diffusion to a real martingale. A second generalisation is to replace the target space $\mathbb{R}$ of the function by a manifold N endowed with a connection; the connection provides the exponential map $\exp_y : T_y N \to N$ and its reverse map $\exp_y^{-1}$ which is defined on a neighbourhood of y; if $f : M \to N$ is a C^2 map, one can consider the tension field of f, namely

$$L_N f : M \to TN \quad x \mapsto L_N f(x) = L(\exp_{f(x)}^{-1} f)(x) \in T_{f(x)} N.$$

Then the semigroup P_t is replaced by a nonlinear semigroup Q_t which acts on functions $f : M \to N$ and is given by the equation

$$\frac{d}{dt} Q_t f = L_N(Q_t f), \qquad Q_0 f = f. \tag{1.3}$$

This semigroup has a stochastic interpretation similar to the real-valued case, but real martingales have to be replaced by N-valued martingales (we refer to [14] for basic results about manifold-valued martingales). One can prove that $Q_t f$ is defined for any $t \ge 0$ if N satisfies some convexity assumptions, for instance if N is a regular geodesic ball of a Riemannian manifold (see the definition in Example 3.2); the stochastic method for the construction of $Q_t f$ is to prove the existence of N-valued martingales with prescribed terminal values, see [17, 22, 23, 1]. The derivative $d(Q_t f)$ takes its values in the bundle of linear maps $L(TM, TN) = \bigcup_{(x,y)} L(T_x M, T_y N)$, and we would like to obtain an estimate similar to the real-valued case (1.1).

First notice that if L satisfies Hrmander's conditions, then it is well known that its semigroup has a smooth density and can be written as

$$P_t f(x) = \int_M f(z) p(t, x, z) dz. \tag{1.4}$$

Then estimates on the derivative of $p(t, x, z)$ with respect to x such as [20] imply that $P_t f$ is smooth, and moreover, one can deduce estimates on $d(P_t f)$. However, this method cannot be applied for $d(Q_t f)$ because the nonlinear semigroup has no representation of type (1.4); a calculation using local coordinates enables to study the smoothness of $Q_t f$ (see [24], see also [16] for symmetric diffusions with an analytical method), but this technique does not provide good gradient estimates.

Thus we are looking for a direct method which does not use the density of the diffusion. A first possibility is to use a coupling method; this method has been worked out in the elliptic case for $P_t f$ in [7]; it has also been extended in [19] to the study of the N-valued semigroup $Q_t f$, but it is not simple (one needs to study the Hlder continuity before the Lipschitz continuity), and it does not cover the case where N is a general regular geodesic ball. Another possibility is to apply the Bismut formula of [6] which gives an expression for $d(P_t f)$ involving f but not its derivatives; this type of formula has been widely studied and enables to obtain some estimates in the elliptic case, see [12, 26, 27, 29, 28]. Some formulae can also be given for the manifold-valued case, see [4, 2], but gradient estimates for general regular geodesic balls are again not given. The method has also been extended in [3] to the hypoelliptic case, but it does not seem easy to deduce gradient estimates for the manifold-valued case.

Here, we consider a method which relates these gradient estimates to the estimation of the quadratic variation of a martingale (real or on N); then, by applying Burkholder and Doob inequalities, one notices that this quadratic variation is dominated by the final value of the martingale (or of some function of the martingale in the manifold case, see [8] for the Burkholder inequalities in this case). This method is worked out in the elliptic case, in Section 2 for the real-valued case, and in Section 3 for the manifold-valued case; in particular it provides estimates of $d(Q_t f)$ when N is a regular Riemannian geodesic ball, or when N is a small enough subset of a manifold endowed with a connection. Moreover, we show in Section 4 that the method also works for a class of hypoelliptic diffusions; more precisely, if L is written in Hrmander's form

$$L = \Xi_0 + \sum_{i=1}^{n} \Xi_i^2/2, \tag{1.5}$$

we obtain the estimate when the tangent space at a point x is linearly generated by Ξ_i and the commutators $[\Xi_i, \Xi_j]$ taken at x ; the simplest example is the hypoelliptic Brownian motion on the Heisenberg group; this result is obtained by using the invariance by rotation of the Wiener process. The more general hypoelliptic case (where one considers Lie brackets of arbitrary length) apparently cannot be dealt with by our method, so the problem of obtaining gradient estimates for $Q_t f$ in this framework is still open.

2 The elliptic case

The results of this section are not new and can be deduced from Malliavin's calculus or Bismut's formula as it was described in the introduction; however the method of this section will again be applied subsequently (for the manifold-valued and hypoelliptic cases), and we prefer to introduce it in a simple framework. We suppose that M is a finite dimensional C^∞ manifold and that X_t is a diffusion on M with generator L written in Hrmander's form (1.5) with C^∞ vector fields Ξ_i. If W_t^i, $1 \le i \le n$, are standard independent real Wiener processes, this diffusion can be written as the solution $X_t = X_t^x$ of the Stratonovich equation

$$\delta X_t = \Xi_0(X_t)\delta t + \sum_{i=1}^{n} \Xi_i(X_t)\delta W_t^i, \quad X_0 = x. \tag{2.1}$$

One can choose a C^∞ modification of the stochastic flow $x \mapsto X_t^x$. The superscript x will often be omitted.

2.1 Gradient estimates on the semigroup

Suppose that the diffusion is elliptic, so that the vector space generated by $\Xi_i(x)$, $1 \le i \le n$, is the whole tangent space $T_x M$, and let us estimate $d(P_t f)$ for a bounded Borel function f and a time $t > 0$. The function $d(P_t f)$ takes its values in the cotangent bundle $T^\star M$, and we need a Riemannian metric on M in order to estimate it; this provides a norm on $T_x M \sim T_x^\star M$; however, the order of the estimate will not depend on this metric if M is compact.

Theorem 1. *Assume that M is compact and that the diffusion is elliptic. Let $q > 1$. Then there exists a constant C such that*

$$\big|d(P_t f)(x)\big| \le \frac{C}{\sqrt{t} \wedge 1} P_t(|f|^q)(x)^{1/q} \tag{2.2}$$

for any x in M, any $t > 0$ and any bounded Borel function f.

Remark 2.1. In all this work, the letter C will denote a positive constant which may change from a formula to the other.

Proof. It is sufficient to prove the result for smooth functions f because the general case can be dealt with by approximating f by $P_\varepsilon f$ as $\varepsilon \downarrow 0$. The proof can also be reduced to the time interval $0 \le t \le 1$; then, for $t > 1$, one can write $P_t f = P_1 P_{t-1} f$, and therefore

$$\big|d(P_t f)\big| \le C\, P_1\big(|P_{t-1} f|^q\big)^{1/q} \le C\, P_t(|f|^q)^{1/q}.$$

So let $t \le 1$. The process

$$Y_s^x = P_{t-s} f(X_s^x), \qquad 0 \le s \le t$$

is a martingale which is given by

$$Y_s = P_t f(x) + \sum_i \int_0^s \Xi_i\big(P_{t-u} f\big)(X_u) dW_u^i. \tag{2.3}$$

By differentiating Y_s^x with respect to x, if $J_s : T_x M \to T_{X_s^x} M$ is the Jacobian of the map $x \mapsto X_s^x$, we obtain a process

$$Y_s' = d(P_{t-s} f)(X_s) J_s,$$

with values in $T_x^\star M$. We deduce from the compactness of M that $\mathbb{E}|Y_s'|^2$ is bounded with respect to x; in particular, we can exchange the conditional expectation and the differentiation in order to prove that Y_s' inherits the martingale property of Y_s. Thus

$$Y_0' = \frac{1}{t} \mathbb{E} \int_0^t Y_s' ds, \tag{2.4}$$

so

$$\begin{aligned} \big|d(P_t f)(x)\big| &= \frac{1}{t} \Big|\mathbb{E} \int_0^t d(P_{t-s} f)(X_s) J_s ds\Big| \\ &\le \frac{1}{\sqrt{t}} \Big\|\sup_{s \le 1} |J_s|\Big\|_{q'} \Big\|\Big(\int_0^t \big|d(P_{t-s} f)(X_s)\big|^2 ds\Big)^{1/2}\Big\|_q \end{aligned} \tag{2.5}$$

with $1/q+1/q'=1$, and where $|J_s|$ is the operator norm of J_s. The process J_t is obtained by differentiating the equation (2.1) of X_t, and since M is compact, standard estimates show that the $L^{q'}$ norm is bounded (independently of x). Moreover, it follows from (2.3) and the ellipticity assumption that the integral in the right hand side of (2.5) is dominated by the quadratic variation

$$\langle Y,Y\rangle_t = \int_0^t \sum_i \Big|\Xi_i(P_{t-s}f)(X_s)\Big|^2 ds \tag{2.6}$$

of the martingale Y_s, so

$$\big|d(P_t f)(x)\big| \le \frac{C}{\sqrt{t}} \big\|\langle Y,Y\rangle_t^{1/2}\big\|_q.$$

Finally, we deduce from the Burkholder and Doob inequalities that

$$\big|d(P_t f)(x)\big| \le \frac{C}{\sqrt{t}}\|Y_t\|_q = \frac{C}{\sqrt{t}}\,P_t(|f|^q)(x)^{1/q}.$$

□

Remark 2.2. In (2.4), one can replace Y'_s, and therefore J_s, by its conditional expectation $\overline{J}_s$ given $(X_u; u\le s)$, so that

$$d(P_t f)(x) = \frac{1}{t}\mathbb{E}\int_0^t d(P_{t-s}f)(X_s)\overline{J}_s ds. \tag{2.7}$$

In particular, let M be a compact Riemannian manifold and suppose that $L=\Delta/2$ for the Laplace-Beltrami operator Δ; then X_t is the Brownian motion on M. One can consider M as a Riemannian submanifold of an Euclidean space $\mathbb{R}^n$ with canonical basis $(e_i; 1\le i\le n)$; then X_t can be viewed as the solution of an equation (2.1) where $\Xi_0(x)=0$ and $\Xi_i(x)$, $1\le i\le n$, are the orthogonal projections of the vectors e_i on T_xM; with this representation, the process X_t becomes a gradient Brownian system, see [10] for details; it is uniformly elliptic. In this case, $\overline{J}_s$ is computed in [13, 11] (this is called filtering out redundant noise). It appears that it is the solution of the covariant differential equation

$$\frac{D}{ds}\overline{J}_s = -\frac{1}{2}\,\mathrm{Ric}^\star(X_s)\overline{J}_s$$

for the Ricci curvature considered as a linear operator on the tangent space; thus a lower bound on the Ricci curvature implies an upper bound on $|\overline{J}_s|$ and can be used in (2.7); then one can go on as in the previous proof. The advantage of this approach is that it provides bounds with geometrical meaning; in functional analytic terms, it is related to the Bochner formula (see [25]). For more general diffusions, these bounds can also be obtained by applying the iterated "carr du champ" (Γ_2) technique, and an analogue of the Ricci curvature is again involved, see [5] or Lemma 1.3 of [21]. However, the method of Theorem 1 can be found simpler if the diffusion is given as a solution of a stochastic differential equation.

Remark 2.3. Another advantage of $\overline{J}_s$ with respect to J_s is that $\overline{J}_s$ is in L^∞. Thus (2.5) can be written with $q = 1$. However, the Doob inequality is not valid for $q = 1$, so we cannot obtain (2.2); we only can use a modification of Doob's inequality (see (25.2) in [9]) to get

$$\big|d(P_t f)(x)\big| \leq \frac{C}{\sqrt{t \wedge 1}}\Big(1 + P_t(|f| \log^+ |f|)(x)\Big).$$

Remark 2.4. One can add a constant to the function f in Theorem 1, so for any real y_0 one has

$$\big|d(P_t f)(x)\big| \leq \frac{C}{\sqrt{t \wedge 1}} P_t(|f - y_0|^q)(x)^{1/q}. \tag{2.8}$$

In particular, for $q = 2$ and $y_0 = P_t f(x)$, we obtain the standard deviation of $f(X_t^x)$ in the right-hand side.

Remark 2.5. One can also estimate higher order derivatives of $P_t f$. For the second order derivative, one studies the martingale Y_s'' which is the second order derivative of Y_s^x. The value Y_0'' is given by an equation similar to (2.4), and the right hand side is estimated by means of the quadratic variations of Y_s and Y_s'.

2.2 Localization and estimates for harmonic functions

The above procedure can be localized in order to estimate derivatives of functions $h(t, x)$ which are solutions of the heat equation (1.2) on a part of $\mathbb{R}_+ \times M$; the localization can also be used to study $P_t f$ when M is not compact (see also subsection 2.3 below). If D is an open subset of M, we consider the space-time process $Z_s^{t,x} = (t - s, X_s^x)$, $0 \leq s \leq t$, and we let τ be the first exit time of $\mathbb{R}_+ \times D$ for this process. Then a smooth function $h(t, x)$ is solution of the heat equation $\partial h/\partial t = Lh$ on $\mathbb{R}_+ \times D$ if $h(Z_{s\wedge\tau}^{t,x})$ is a martingale for any (t, x). Like previously, we fix a Riemannian metric on M.

Theorem 2. *Suppose that the diffusion is elliptic and consider a smooth bounded function h which is solution of the heat equation on $\mathbb{R}_+ \times D$. Let ρ be the distance function to the complement of D, and let $q > 1$. Let $dh(t,x)$ be the derivative of $h(t,x)$ with respect to x. For any compact subset K of M, there exists a $C_K > 0$ which does not depend on h such that*

$$|dh(t,x)| \leq \frac{C_K}{1 \wedge \sqrt{t} \wedge \rho(x)} \Big(\mathbb{E}|h(Z_\tau^{t,x})|^q \Big)^{1/q} \tag{2.9}$$

for $x \in K \cap D$ and $t > 0$. In particular, if $h(t,x) = h(x)$ is harmonic on D, then

$$|dh(x)| \leq \frac{C_K}{\rho(x) \wedge 1} \Big(\mathbb{E}|h(X_\tau^x)|^q \Big)^{1/q}$$

where $h(X_\tau^x)$ is defined as the limit of $h(X_t^x)$ on $\{\tau = \infty\}$.

Reduction to the compact case. We first verify that the proof of Theorem 2 can be reduced to the case of a compact manifold M. We have to prove that any point x_0 of M has a neighbourhood, for instance a ball B_0, on which the estimate (2.9) holds. Let B_1 be a ball which is slightly larger than B_0, let τ_1 be the exit time of $\mathbb{R}_+ \times (D \cap B_1)$, and let ρ_1 is the distance function to the complement of $D \cap B_1$. We have

$$\mathbb{E}|h(Z_\tau^{t,x})|^q \geq \mathbb{E}|h(Z_{\tau_1}^{t,x})|^q \quad \text{and} \quad \rho(x) \wedge 1 \leq C\, \rho_1(x)$$

for $x \in B_0$, where the first inequality holds because $Y_s = h(Z_s^{t,x})$ is a martingale up to time τ. We deduce that it is sufficient to prove the estimate (2.9) for $D \cap B_1$ instead of D. The ball B_1 can be embedded isometrically in a compact manifold, so we can suppose that M is compact and take $K = M$.

Proof of Theorem 2 in the compact case. Like previously, it is sufficient to consider the case $t \leq 1$. Consider on $\mathbb{R}_+ \times M$ the distance function

$$\delta_\star((t,x),(t',x')) = \max\Big(\frac{\delta_M(x,x')}{\alpha}, \sqrt{|t'-t|} \Big)$$

for the Riemannian distance δ_M on M, and where $\alpha > 1$ is a constant which will be chosen later. We also consider the function $\rho_\star(t,x)$ which is the distance to the complement of $\mathbb{R}_+ \times D$, so that

$$\rho_\star(t,x) = \frac{\rho(x)}{\alpha} \wedge \sqrt{t}.$$

Fix (t,x). For $0 < r < \rho_\star^2(t,x) \le 1$, let

$$\begin{aligned}\tau(r) &= \inf\Big\{ s \ge 0; \delta_\star((t,x), Z_s) \ge \sqrt{r} \Big\} \\ &= \inf\Big\{ s \ge 0; \delta_M(x, X_s) \ge \alpha\sqrt{r} \Big\} \wedge r\end{aligned}$$

By differentiating the martingale property of $Y_s = h(Z_s)$, we verify that $Y'_s = dh(Z_s)J_s$ is a local martingale up to time τ, where J_s is the Jacobian of X_s^x as in the proof of Theorem 1 (there is a small difficulty due to the time τ, but the result can be proved with a time change on the diffusion outside a compact subset of D, so that it does not quit D; this implies that Y'_s is a local martingale up to the exit time of the arbitrary compact subset). The condition on r implies that $\tau(r) < \tau$, so

$$\begin{aligned}|Y'_0| &\le \frac{1}{r}\Big\| \int_0^r |Y'_{s\wedge\tau(r)}| ds \Big\|_1 \\ &\le \frac{1}{r}\Big\| \int_0^{\tau(r)} |Y'_s| ds \Big\|_1 + \frac{1}{r}\Big\| (r - \tau(r)) |Y'_{\tau(r)}| \Big\|_1 \\ &\le \frac{1}{\sqrt{r}} \Big\| \Big(\int_0^{\tau(r)} |dh(Z_s)|^2 ds \Big)^{1/2} \Big\|_q \Big\| \sup_{0\le s\le r} |J_s| \Big\|_{q'} \\ &\quad + \mathbb{P}[\tau(r) < r]^{1/q'} \Big\| Y'_{\tau(r)} \Big\|_q\end{aligned}$$

if $1/q + 1/q' = 1$. The moments of $\sup_s |J_s|$ are finite, and the first L^q norm is estimated as in Theorem 1 by means of the quadratic variation of $Y_s = h(Z_s)$, and therefore by $\|Y_{\tau(r)}\|_q \le \|Y_\tau\|_q$. In the second term, standard estimates show that the probability of

$$\{\tau(r) < r\} = \Big\{ \sup_{s<r} \delta_M(x, X_s) \ge \alpha\sqrt{r} \Big\}$$

can be made arbitrarily small if α is chosen large enough. Thus we choose α so that

$$|Y'_0| \le \frac{C}{\sqrt{r}} \|Y_\tau\|_q + \frac{1}{4}\|Y'_{\tau(r)}\|_q. \tag{2.10}$$

On the other hand

$$\rho_\star(Z_{\tau(r)}) \ge \rho_\star(Z_0) - \sqrt{r},$$

so

$$\begin{aligned}\rho_\star(Z_0)|Y'_0| \le &\frac{C}{\sqrt{r}} \rho_\star(Z_0) \|Y_\tau\|_q \\ &+ \frac{1}{4} \frac{\rho_\star(Z_0)}{\rho_\star(Z_0) - \sqrt{r}} \Big\| \rho_\star(Z_{\tau(r)}) Y'_{\tau(r)} \Big\|_q.\end{aligned}$$

If we choose $r = \rho_\star^2(Z_0)/4$, then we obtain

$$\rho_\star(Z_0)|Y'_0| \leq C\big\|Y_\tau\big\|_q + \frac{1}{2}\Big\|\rho_\star(Z_{\tau(r)})Y'_{\tau(r)}\Big\|_q.$$

More generally, if now τ' is any optional time such that $\tau' < \tau$, the same method enables to estimate $\rho_\star(Z_{\tau'})|Y'_{\tau'}|$, and by taking the L^q norm, we show that there exists an optional time $\tau' < \tau'' < \tau$ such that

$$\Big\|\rho_\star(Z_{\tau'})Y'_{\tau'}\Big\|_q \leq C\big\|Y_\tau\big\|_q + \frac{1}{2}\Big\|\rho_\star(Z_{\tau''})Y'_{\tau''}\Big\|_q.$$

By taking the supremum over all optional times which are less than τ, we obtain

$$\sup_{\tau'}\Big\|\rho_\star(Z_{\tau'})Y'_{\tau'}\Big\|_q \leq 2C\big\|Y_\tau\big\|_q.$$

We deduce (2.9) by taking $\tau' = 0$. □

2.3 Uniform estimates in the non compact case

If M is not compact, we now wonder whether Theorem 1 holds, or whether Theorem 2 holds for a constant $C_K = C$ which does not depend on K. We first have to choose on M a Riemannian metric so that the diffusion is uniformly elliptic, that is

$$\sum_i \big|\Xi_i f(x)\big|^2 \geq c\big|df(x)\big|^2$$

for any smooth function f. Then, in order to work out the estimates of Theorem 1, we first have to justify the exchange of the conditional expectation and differentiation in (2.4); then we need the uniform boundedness of the moments of $\sup_{s\leq 1}|J_s|$; to this end, we have to choose a convenient representation (1.5) of the generator. For Theorem 2, we also have to verify that the probability of $\{\tau(r) < r\}$ is uniformly small (for $r \leq 1$) if α is chosen large enough.

Example 2.1. If $M = \mathbb{R}^d$, one can write the equation (2.1) as an It equation

$$dX_t = b(X_t)dt + \sum_i \xi_i(X_t)dW_t^i, \qquad X_0 = x.$$

If the matrix $\xi\xi^\star$ is bounded and uniformly elliptic and if moreover b and the Jacobian matrices of b and ξ_i are bounded, then one can verify with the above procedure that Theorems 1 and 2 hold uniformly.

Example 2.2. Let M be a closed submanifold of an Euclidean space; suppose that M is endowed with its induced Riemannian metric and $L = \Delta/2$; then the diffusion is uniformly elliptic. If we apply the method of Remark 2.2, then $\overline{J}_s$ is conveniently estimated as soon as the Ricci curvature of M is bounded below; however, obtaining an intrinsic condition ensuring the justification of (2.4) is not so easy, see chapters 6 and 10 of [25]. On the other hand, the estimation of the exit time of small balls (more precisely of $\mathbb{P}[\tau(r) < r]$) can also be worked out when the Ricci curvature is bounded below with the technique of [15], and we obtain (2.9) uniformly.

Example 2.3. Suppose that there exists a group G which acts transitively on M, and that the diffusion is invariant under this action; this means that $(X_t^{g.x})$ and $(g.X_t^x)$ have the same law; an example is the Brownian motion on an homogeneous space. Then it is clear that the estimation of $dh(t,x)$ for solutions h of (1.2) can be reduced to the estimation at a fixed $x = x_0$, so Theorem 2 holds uniformly, and Theorem 1 also holds.

3 Estimation for harmonic maps

Let us now consider the nonlinear semigroup Q_t defined for N-valued maps by (1.3) for a given connection on the manifold N. We suppose that $Q_t f$ is well defined, and that $(t,x) \mapsto Q_t f(x)$ is smooth for $t > 0$; in the elliptic case, this holds under some convexity conditions on N, see [19, 2, 24] for probabilistic proofs. We want to estimate $d(Q_t f)$ and prove an analogue of (2.8). The stochastic interpretation of Q_t is similar to P_t; the connection enables to consider a notion of N-valued continuous martingale, and the process $Y_s^x = Q_{t-s} f(X_s^x)$ is for any x a N-valued martingale, so $Q_t f(x)$ is the initial value of the martingale with final value $f(X_t^x)$ (this martingale is unique under the convexity assumptions). The aim of this section is to prove the following result (see also the extension to the nonlinear heat equation at the end of the section).

Theorem 3. *Suppose that M is compact, that the diffusion is elliptic, and that N is a relatively compact open subset of a manifold $\widetilde{N}$ endowed with an extension of the connection of N. We suppose that $\widetilde{N}$ satisfies the following convexity conditions; there exists a $p \geq 1$, a Riemannian distance δ on $\widetilde{N}$ and nonnegative functions ϕ and ψ on $\widetilde{N} \times \widetilde{N}$ such that*

$$\phi(y_0, y) = 0 \iff \psi(y_0, y) = 0 \iff y_0 = y$$

and

1. *The function* ϕ *is convex and* $\phi(y_0, y) \sim \delta^p(y_0, y)$ *as* $y \to y_0$.

2. *For any* y_0*, the function* $y \mapsto \psi(y_0, y)$ *is* C^2 *and strictly convex (its Hessian is positive definite).*

Then, for any $q > p$*, there exists a* C *such that*

$$|d(Q_t f)(x)| \leq \frac{C}{\sqrt{t} \wedge 1} \mathbb{E}\big[\delta^q(y_0, f(X_t))\big]^{1/q} \tag{3.1}$$

for any y_0 *in* N*, any* x *in* M*, any* $t > 0$ *and any function* $f : M \to N$ *such that the solution* $Q_t f$ *of* (1.3) *is well defined,* N*-valued, and smooth for* $t > 0$.

Remark 3.1. The manifold $\widetilde{N} \times \widetilde{N}$ is endowed with the product connection; saying that ϕ is convex means that it is convex along the geodesic curves; if U_t^1 and U_t^2 are N-valued martingales, this implies that $\phi(U_t^1, U_t^2)$ is a submartingale. Notice also that the result (3.1) is stated for the distance δ, but it also holds for other Riemannian distances on $\widetilde{N}$, since all these distances are equivalent on N.

The convexity conditions of the theorem look stronger than the "p-convexity" conditions used in [23, 1]. However, one can give basically the same examples of manifolds N satisfying them.

Example 3.1. On any manifold endowed with a connection and for any $p > 1$, any point has neighbourhoods $N \subset \widetilde{N}$ satisfying the above conditions. The function constructed in Proposition 2.5 of [2] can be proved to be a convenient function ϕ for some distance δ, and one can take $\psi = \delta^2$.

Example 3.2. One can choose for $N \subset \widetilde{N}$ regular geodesic balls in a Riemannian manifold with distance δ_0; this means that the sectional curvatures of the manifold are bounded above by some $\kappa \geq 0$, and that $N \subset \widetilde{N}$ are balls (for the distance δ_0) with empty cut loci and with radii less than $\pi/(2\kappa)$ (there is no condition on the radii if $\kappa = 0$). In this situation, the existence of ϕ has been obtained in [18] by generalizing the case of the sphere (notice that δ is generally different from δ_0); moreover, this function ϕ is strictly convex outside the diagonal (one can also use the method of [23] to construct ϕ, but one needs some extra work in order to obtain the strict convexity); the value of p depends on the radius of $\widetilde{N}$. On the other hand, there exists a $c > 0$ such that $y \mapsto \delta^2(y_0, y)$ is strictly convex for $\phi(y_0, y) \leq c$; then the function

$$\psi = \delta^2 + \beta\big((\phi - c/2)^+\big)^3$$

satisfies the convexity condition of Theorem 3 if β is chosen large enough.

Example 3.3. If $\widetilde{N}$ is a Cartan-Hadamard manifold with distance δ_0 (a simply connected Riemannian manifold with nonpositive sectional curvatures), then any ball $N \subset \widetilde{N}$ satisfies the conditions with $p = 1$, $\phi = \delta = \delta_0$ and $\psi = \delta^2$. Moreover, the constants involved in the proof below do not depend on the size of the ball, so the estimate holds for the whole (non compact) manifold and the result of Theorem 3 becomes quite similar to the result (2.8) for the real-valued case.

Proof of Theorem 3. As in Theorem 1, we suppose that $Q_t f$ is smooth for $t \geq 0$ (otherwise approximate f by $Q_\varepsilon f$) and we let $t \leq 1$. Let us consider the N-valued martingale $Y_s^x = Q_{t-s} f(X_s^x)$ and its derivative $Y_s' = d(Q_{t-s} f)(X_s)$ with respect to x; this is a process with values in the bundle $L(TM, TN)$. Let $|Y_s'|$ be its operator norm (associated to the Riemannian metrics δ_M and δ on M and N). For any x and x', the convexity of ϕ implies that $\phi(Y_s^x, Y_s^{x'})$ is a submartingale, so by dividing by $\delta_M(x, x')$ and taking the limit as $x' \to x$, we deduce that $|Y_s'|^p$ is a submartingale. Thus

$$|Y_0'|^p \leq \frac{1}{t} \mathbb{E} \int_0^t |Y_s'|^p ds.$$

Case $p \leq 2$. In this case,

$$\begin{aligned} |Y_0'| &\leq \frac{1}{\sqrt{t}} \Big\| \Big(\int_0^t |Y_s'|^2 ds \Big)^{1/2} \Big\|_p \\ &\leq \frac{1}{\sqrt{t}} \Big\| \sup_{s \leq 1} |J_s| \Big\|_{q'} \Big\| \Big(\int_0^t \big| d(Q_{t-s} f)(X_s) \big|^2 ds \Big)^{1/2} \Big\|_q \end{aligned}$$

for $q > p$ and $1/q + 1/q' = 1/p$. The last term can again be dominated by means of the quadratic variation $\langle\!\langle Y \rangle\!\rangle$ of Y_s (computed for instance for the distance δ) so that

$$|Y_0'| \leq \frac{C}{\sqrt{t}} \Big\| \langle\!\langle Y \rangle\!\rangle_t^{1/2} \Big\|_q.$$

On the other hand, the derivatives of ψ are dominated by $\psi^{1/2}$, and ψ is strictly convex, so we can apply the Burkholder inequalities of [8] for manifold-valued martingales to obtain

$$|Y_0'| \leq \frac{C}{\sqrt{t}} \Big\| \sup_{s \leq t} \psi(Y_0, Y_s)^{1/2} \Big\|_q.$$

Since $\psi \asymp \delta^2 \asymp \phi^{2/p}$, we get

$$|Y_0'| \leq \frac{C}{\sqrt{t}} \Big\| \sup_{s \leq t} \phi(Y_0, Y_s) \Big\|_{q/p}^{1/p}.$$

Finally we can apply the Doob inequality to the submartingale $\phi(Y_0, Y_s)$ and obtain

$$|Y_0'| \le \frac{C}{\sqrt{t}} \big\|\phi(Y_0, Y_t)\big\|_{q/p}^{1/p} \le \frac{C'}{\sqrt{t}} \big\|\delta(Y_0, Y_t)\big\|_q.$$

Moreover

$$\begin{aligned} \big\|\delta(Y_0, Y_t)\big\|_q &\le \delta(y_0, Y_0) + \big\|\delta(y_0, Y_t)\big\|_q \\ &\le C\big|\phi(y_0, Y_0)\big|^{1/p} + \big\|\delta(y_0, Y_t)\big\|_q \\ &\le C\big\|\phi(y_0, Y_t)\big\|_{q/p}^{1/p} + \big\|\delta(y_0, Y_t)\big\|_q \le C'\big\|\delta(y_0, Y_t)\big\|_q, \end{aligned} \tag{3.2}$$

so the proof of the theorem is complete in the case $p \le 2$.

Case $p > 2$. In this case, we have

$$\begin{aligned} |Y_0'|^p &\le \frac{1}{r}\mathbb{E}\int_0^r |Y_s'|^p ds \\ &\le \frac{1}{r}\mathbb{E}\Big[\sup_{s\le r} |Y_s'|^{p-2} \int_0^r |Y_s'|^2 ds\Big] \\ &\le \frac{1}{r}\mathbb{E}\Big[\sup_{s\le r} |Y_s'|^p\Big]^{1-2/p} \mathbb{E}\Big[\Big(\int_0^r |Y_s'|^2 ds\Big)^{p/2}\Big]^{2/p} \\ &\le \frac{C}{r}\mathbb{E}\Big[\sup_{s\le r} |Y_s'|^p\Big]^{1-2/p} \mathbb{E}\Big[\langle\!\langle Y\rangle\!\rangle_r^{q/2}\Big]^{2/q}. \end{aligned}$$

By applying the Doob inequality to the first term, and the Burkholder and Doob inequalities to the second term (as in the case $p \le 2$), we obtain

$$|Y_0'|^p \le \frac{C}{r}\mathbb{E}\Big[|Y_r'|^q\Big]^{(p-2)/q} \mathbb{E}\Big[\delta(Y_0, Y_r)^q\Big]^{2/q}.$$

On the other hand, for any $c > 0$, there exists a $C > 0$ such that

$$xy \le c\, x^{p/(p-2)} + C\, y^{p/2}.$$

for positive x and y. By applying this property for $c = 1/4$, we get

$$\begin{aligned} |Y_0'| &\le \frac{C}{r^{1/p}} \big\|Y_r'\big\|_q^{(p-2)/p} \big\|\delta(Y_0, Y_r)\big\|_q^{2/p} \\ &\le \frac{1}{4}\|Y_r'\|_q + \frac{C'}{\sqrt{r}} \big\|\delta(Y_0, Y_r)\big\|_q. \end{aligned} \tag{3.3}$$

We have obtained an estimate for Y_0'; for $u < u + r \le t$, we deduce similarly an estimate for Y_u', and by taking the L^q norm, we obtain

$$\|Y_u'\|_q \le \frac{1}{4}\|Y_{u+r}'\|_q + \frac{C}{\sqrt{r}} \big\|\delta(Y_u, Y_{u+r})\big\|_q.$$

Moreover, a technique similar to (3.2), based on the convexity of ϕ enables to prove that

$$\big\|\delta(Y_u, Y_{u+r})\big\|_q \le C\big\|\delta(y_0, Y_t)\big\|_q,$$

so

$$\|Y'_u\|_q \le \frac{1}{4}\|Y'_{u+r}\|_q + \frac{C}{\sqrt{r}}\big\|\delta(y_0, Y_t)\big\|_q.$$

By choosing $r = \frac{3}{4}(t-u)$, we deduce that for any $0 \le u < t$, there exists $u < s < t$ given by $s = u + r = u + \frac{3}{4}(t-u)$ such that

$$\sqrt{t-u}\|Y'_u\|_q \le \frac{1}{2}\sqrt{t-s}\|Y'_s\|_q + C\big\|\delta(y_0, Y_t)\big\|_q.$$

If we take the supremum with respect to u, we obtain

$$\sup_{0\le u<t}\Big(\sqrt{t-u}\|Y'_u\|_q\Big) \le 2C\big\|\delta(y_0, Y_t)\big\|_q$$

and deduce the estimate (3.1) for $d(Q_t f)(x) = Y'_0$. □

If now $h(t,x)$ is solution of the nonlinear heat equation $\partial h/\partial t = L_N h$ on $\mathbb{R}_+ \times D$, one can apply jointly the technique of Theorem 3 and the localization procedure of Theorem 2. If for instance $p > 2$, one considers $Y_s = h(Z_s)$, $Y'_s = dh(Z_s)J_s$, and $|Y'_s|^p$ is a submartingale up to time τ; by stopping the processes at $\tau(r)$ one obtains

$$|Y'_0|^p \le \frac{1}{r}\mathbb{E}\int_0^{\tau(r)} |Y'_s|^p ds + \frac{1}{r}\mathbb{E}\Big[(r-\tau(r))|Y'_{\tau(r)}|^p\Big].$$

The first term can be estimated as in (3.3), and since the probability of $\{\tau(r) < r\}$ can be made small, the second term can be estimated by the second term of (2.10). Thus we obtain

$$|Y'_0| \le \frac{1}{4}\big\|Y'_{\tau(r)}\big\|_q + \frac{1}{4}\big\|Y'_{\tau(r)}\big\|_q + \frac{C}{\sqrt{r}}\big\|\delta(y_0, Y_{\tau(r)})\big\|_q$$

Then we multiply by $\rho_\star(Z_0)$, and the study can be completed as in Section 2. The result is

$$|dh(t,x)| \le \frac{C_K}{1\wedge\sqrt{t}\wedge\rho(x)}\mathbb{E}\big[\delta^q(y_0, h(Z_\tau^{t,x}))\big]^{1/q}.$$

4 The hypoelliptic case

Let us now consider the non elliptic case; as it has been explained in the introduction (see also Remark 4.2 below), we cannot deal with the general case, but have to restrict to the case where the tangent space is generated by the C^∞ vector fields Ξ_i and their commutators; a particular (non compact) case is the classical Brownian motion on the Heisenberg group; it is a three-dimensional process consisting of a two-dimensional standard Brownian motion and of its Lvy area. Like previously, the following result is not new in the real-valued case (it can be obtained with Malliavin's calculus), but the method can be extended to the manifold-valued case (this is sketched at the end of the section). We do not try to prove precise estimates with the L^q norm of $f(X_t)$ as in previous sections, but only with the supremum $\|f\|_\infty$ of $|f|$.

Theorem 4. *Suppose that M is compact and that T_xM is for any $x \in M$ generated by the vector fields Ξ_i and $\Xi_{ij} = [\Xi_i, \Xi_j]$ taken at point x. Then there exists a $C > 0$ such that*

$$\big|d(P_t f)(x)\big| \le \frac{C}{t \wedge 1} \|f\|_\infty$$

for any $t > 0$ and any bounded Borel function f.

Proof. We are going to prove the result for f smooth and $t \le 1$. We will denote the vector fields Ξ_i and Ξ_{ij} taken at x by $\xi_i(x)$ and $\xi_{ij}(x) \in T_xM$ (we keep the upper case letters to denote the vector fields considered as acting on functions). We consider the relation

$$d(P_t f)(x) = \frac{2}{t} \mathbb{E} \int_0^{t/2} d(P_{t-s} f)(X_s) J_s ds.$$

Our hypoellipticity condition implies that we can write the Jacobian J_s taken on some vector $\overrightarrow{e} \in T_xM$ as

$$J_s \overrightarrow{e} = \sum_i \psi_s^i \xi_i(X_s) + \sum_{ij} \psi_s^{ij} \xi_{ij}(X_s). \tag{4.1}$$

with $|\psi_s^i|, |\psi_s^{ij}| \le C\,|J_s|$; in particular the moments of ψ_s^i and ψ_s^{ij} are bounded. Thus

$$\begin{aligned} d(P_t f)(x)\overrightarrow{e} &= \frac{2}{t} \sum_i \mathbb{E} \int_0^{t/2} \Xi_i(P_{t-s} f)(X_s) \psi_s^i ds \\ &\quad + \frac{2}{t} \sum_{ij} \mathbb{E} \int_0^{t/2} \Xi_{ij}(P_{t-s} f)(X_s) \psi_s^{ij} ds. \end{aligned} \tag{4.2}$$

The first type of terms can be estimated by means of the quadratic variation of the martingale Y_s like previously, and is of order $1/\sqrt{t}$; thus it is sufficient to study the second type, and therefore to estimate $\Xi_{ij}(P_t f)$. To this end, embed M into an Euclidean space, extend the vector fields Ξ_i to C^∞ vector fields with compact support, and write the equation for X_s in It's form

$$X_s = x + \int_0^s b(X_u)du + \sum_k \int_0^s \xi_k(X_u)dW_u^k. \tag{4.3}$$

Then

$$\xi_i(X_s) = \xi_i(x) + \sum_k \Xi_k \xi_i(x) W_s^k + O(s) \tag{4.4}$$

in the spaces L^q for s small, and

$$\begin{aligned} \xi_i(X_s)W_s^j &= \int_0^s \xi_i(X_u)dW_u^j + \int_0^s W_u^j d\xi_i(X_u) + \langle \xi_i(X), W^j \rangle_s \\ &= \xi_i(x)W_s^j + \sum_k \Xi_k \xi_i(x) \Big(\int_0^s W_u^k dW_u^j \\ &\quad + \int_0^s W_u^j dW_u^k \Big) + \Xi_j \xi_i(x) s + O(s^{3/2}). \end{aligned} \tag{4.5}$$

Now, for (i, j) fixed, let us apply an infinitesimal rotation on the Wiener process (W_s^j, W_s^i), and an infinitesimal modification on the initial condition in the direction $t\xi_{ij}(x)$ (t is now a small positive parameter; it will be later interpreted as a time parameter); this means that (W_s^j, W_s^i) is replaced by

$$(W_s^j \cos\varepsilon + W_s^i \sin\varepsilon, -W_s^j \sin\varepsilon + W_s^i \cos\varepsilon)$$

and that x is replaced by $x + \varepsilon t \xi_{ij}(x)$; we denote the perturbed process by X_s^ε. Then the differentiation of (4.3) shows that the derivative V_s of $\varepsilon \mapsto X_s^\varepsilon$ at $\varepsilon = 0$ is solution of

$$\begin{aligned} V_s = t\xi_{ij}(x) &+ \int_0^s (db)(X_u)V_u du + \sum_k \int_0^s (d\xi_k)(X_u)V_u dW_u^k \\ &+ \int_0^s \xi_j(X_u)dW_u^i - \int_0^s \xi_i(X_u)dW_u^j. \end{aligned} \tag{4.6}$$

This process is of order $t + \sqrt{s}$ for s and t small, and

$$V_s = t\xi_{ij}(x) + \xi_j(x)W_s^i - \xi_i(x)W_s^j + O(t\sqrt{s} + s).$$

By applying this estimate to the terms V_u of (4.6), and by using also (4.4) , we obtain the more precise expansion

$$\begin{aligned} V_s = & t\xi_{ij}(x) + \xi_j(x)W_s^i - \xi_i(x)W_s^j \\ & + \Xi_j\xi_k(x)\int_0^s W_u^i dW_u^k - \Xi_i\xi_k(x)\int_0^s W_u^j dW_u^k \\ & + \Xi_k\xi_j(x)\int_0^s W_u^k dW_u^i - \Xi_k\xi_i(x)\int_0^s W_u^k dW_u^j + O(t\sqrt{s} + s^{3/2}) \end{aligned}$$

where the expressions involving k have to be summed. In this equation, the vector fields are taken at x; we use (4.5) in order to give an approximate expression of V_s as a linear combination of vector fields taken at X_s. By using $\Xi_k\xi_j - \Xi_j\xi_k = \xi_{kj}$, we check after a calculation that

$$\begin{aligned} V_s = & (t-s)\xi_{ij}(X_s) + \xi_j(X_s)W_s^i - \xi_i(X_s)W_s^j \\ & + \xi_{ki}(X_s)\int_0^s W_u^j dW_u^k - \xi_{kj}(X_s)\int_0^s W_u^i dW_u^k + O(t\sqrt{s} + s^{3/2}). \end{aligned}$$

The law of X_s is sensible to the perturbation of the initial condition, but not to the rotation of the Wiener process; this means that X_s^ε is still a diffusion with the same semigroup, so $(P_{s-u}f)(X_u^\varepsilon)$ is a martingale for any ε; after differentiation, we deduce that $d(P_{s-u}f)(X_u)V_u$, $0 \le u \le s$ is a martingale with initial value $\Xi_{ij}(P_s f)(x)t$. Thus

$$\begin{aligned} \Xi_{ij}(P_s f)(x) &= \mathbb{E}\big[df(X_s)V_s\big] \,/\, t \\ &= \mathbb{E}\Big[(1-s/t)\Xi_{ij}f(X_s) + \Xi_j f(X_s)W_s^i/t \\ &\quad - \Xi_i f(X_s)W_s^j/t + \Xi_{ki}f(X_s)\int_0^s W_u^j dW_u^k/t \\ &\quad - \Xi_{kj}f(X_s)\int_0^s W_u^i dW_u^k/t\Big] + \mathbb{E}\Big[O(\sqrt{s} + s^{3/2}/t)|df(X_s)|\Big]. \end{aligned} \tag{4.7}$$

We let now t be the time parameter, and we write

$$\Xi_{ij}(P_t f) = \frac{n}{t}\int_0^{t/n} \Xi_{ij}(P_s P_{t-s} f)ds$$

for some integer $n \ge 2$ which will be chosen later; we express the right hand side by applying (4.7) with $P_{t-s}f$ instead of f. Some of the terms can be estimated; the term

$$\begin{aligned} & \frac{n}{t^2}\int_0^{t/n} \Xi_j(P_{t-s}f)(X_s)W_s^i ds \\ & \quad \le \frac{n}{t^2}\Big(\int_0^{t/n} \big|\Xi_j(P_{t-s}f)(X_s)\big|^2 ds\Big)^{1/2}\Big(\int_0^{t/n} (W_s^i)^2 ds\Big)^{1/2} \end{aligned}$$

can be estimated as in Theorem 1 by means of the quadratic variation of the martingale $P_{t-s}f(X_s)$ and is $O(1/t)\|f\|_\infty$. From our hypoellipticity assumption, the derivative $d(P_{t-s}f)$ can be expressed as a combination of $\Xi_k(P_{t-s}f)$ and of $\Xi_{kl}(P_{t-s}f)$, and

$$\frac{n}{t}\int_0^{t/n}\big|\Xi_k(P_{t-s}f)(X_s)\big|O(\sqrt{s}+s^{3/2}/t)ds = O(1)\,\|f\|_\infty$$

by using again the quadratic variation. We obtain

$$\begin{aligned}\Xi_{ij}(P_tf)(x) = \mathbb{E}\Big[&\frac{n}{t}\int_0^{t/n}(1-\frac{s}{t})\Xi_{ij}(P_{t-s}f)(X_s)ds\\ &+\frac{n}{t}\int_0^{t/n}\Xi_K(P_{t-s}f)(X_s)\Big(\frac{G_{ij}^K(R_s)}{t}+O(\sqrt{s})\Big)ds\Big]\\ &+O(1/t)\|f\|_\infty,\end{aligned}$$

where the expression is summed over indices $K=(k_1,k_2)$, where R_s consists of the double integrals of the process W_s, and where G_{ij}^K are linear forms. Thus the vector $\Xi(P_tf)=(\Xi_K(P_tf))$ is solution of

$$\Xi(P_tf)(x) = \frac{n}{t}\mathbb{E}\int_0^{t/n}(I-A_s)\Xi(P_{t-s}f)(X_s)ds + O(1/t)\|f\|_\infty$$

for a matrix-valued process A_s satisfying

$$A_s = \frac{s}{t}I - \frac{G(R_s)}{t} + O(\sqrt{s})$$

for a linear map G. The procedure can be iterated in order to express $\Xi(P_{t-s}f)(X_s)$, and we obtain

$$\begin{aligned}\Xi(P_tf)(x) =&\frac{n^2}{t^2}\mathbb{E}\int_0^{t/n}\int_{t/n}^{2t/n}(I-A_{s_1})(I-A_{s_1s_2})\Xi(P_{t-s_2}f)(X_{s_2})ds_2ds_1\\ &+O(1/t)\|f\|_\infty\end{aligned}$$

with

$$A_{s_1s_2} = \frac{s_2-s_1}{t}I - \frac{G(R_{s_1s_2})}{t} + O(\sqrt{s_2-s_1})$$

and where $R_{s_1s_2}$ consists of the double integrals of the increments of W on $[s_1,s_2]$. After n iterations, we get

$$\begin{aligned}\Xi(P_tf)(x) =&\frac{n^n}{t^n}\mathbb{E}\int_0^{t/n}\dots\int_{(n-1)t/n}^{t}(I-A_{s_1})\dots(I-A_{s_{n-1}s_n})\\ &\Xi(P_{t-s_n}f)(X_{s_n})ds_n\dots ds_1 + O(1/t)\|f\|_\infty.\end{aligned} \tag{4.8}$$

Consider now the process

$$\overline{A}_{s_1s_2} = \frac{s_2 - s_1}{t} I - \frac{G(R_{s_1s_2})}{t}.$$

The variable $G(R_{s_{i-1}s_i})$ is of order $s_i - s_{i-1}$, and its conditional expectation given the process W up to time s_{i-1} is 0. By applying classical techniques (such as those which are used for the time discretization of stochastic differential equations), we can deduce that

$$(I - \overline{A}_{s_1}) \dots (I - \overline{A}_{s_{n-1}s_n})$$

converges in the spaces L^q to $e^{-1}I$ as $n \to \infty$, uniformly for $(s_1, \dots, s_n)$ in the integration domain of (4.8). Thus the L^1 norm of the operator norm of this variable is close to e^{-1} if n is chosen large enough; let us fix such an n. Then, if t is small enough, we deduce that

$$\Big\| \big| (I - A_{s_1}) \dots (I - A_{s_{n-1}s_n}) \big| \Big\|_1 \leq 1/2 \tag{4.9}$$

and therefore

$$|\Xi(P_t f)(x)| \leq \frac{1}{2} \sup_{s \leq t} \|\Xi(P_s f)\|_\infty + \frac{C}{t} \|f\|_\infty.$$

By using $P_u = P_t P_{u-t}$, this equation implies

$$\|\Xi(P_u f)\|_\infty \leq \frac{1}{2} \sup_{u-t \leq s \leq u} \|\Xi(P_s f)\|_\infty + \frac{C}{t} \|f\|_\infty$$

for $u \geq t$. Thus, if

$$F(t) = \sup_{t \leq u \leq 1} \|\Xi(P_u f)\|_\infty,$$

then

$$F(4t) \leq \frac{1}{2} F(3t) + \frac{C}{t} \|f\|_\infty$$

for $t \leq 1/4$, and therefore

$$\sup_{0 \leq t \leq 1/4} \big(4tF(4t)\big) \leq \frac{2}{3} \sup_{0 \leq t \leq 1/3} \big(3tF(3t)\big) + 4C\|f\|_\infty.$$

Thus $t\,F(t)$ is dominated by $\|f\|_\infty$, and we conclude the proof from (4.2). □

Remark 4.1. Consider (4.2) with $\overrightarrow{e} = \xi_k(x)$; then the left hand side is $\Xi_k(P_t f)(x)$; moreover, an analysis of (4.1) shows that ψ_s^{ij} is in this case of order $\sqrt{s}$ (notice that $\psi_0^{ij} = 0$). We can deduce that

$$\left|\Xi_k(P_t f)(x)\right| \leq \frac{C}{\sqrt{t} \wedge 1}\|f\|_\infty.$$

Thus, as $t \downarrow 0$, the Lipschitz coefficient of $P_t f$ is of order $1/\sqrt{t}$ for the intrinsic subriemannian distance of the diffusion on M, and it is of order $1/t$ for Riemannian distances.

Remark 4.2. If now we consider the general hypoelliptic case, then we have higher order brackets of (Ξ_i) in (4.2). As it has been said in the proof, the vector fields Ξ_i can be estimated on $P_t f$ by means of a quadratic variation; roughly speaking, they can be interpreted as Cameron-Martin perturbations on the driving Brownian motion W_t. Then the vector fields Ξ_{ij} were estimated on $P_t f$ by means of rotations on W_t. However, we have no other absolutely continuous perturbation to estimate higher order brackets, so we think that the method cannot be extended to more general hypoelliptic situations.

The localization (see the framework of Theorem 2) does not cause much problem; we stop the processes at $\tau(r)$, and $\Xi h(t,x)$ is now expressed by means of variables $A_{s_{i-1}\wedge\tau(r), s_i\wedge\tau(r)}$; then we consider separately (like previously) the events $\{\tau(r) = r\}$ (on which we use the above estimations) and $\{\tau(r) < r\}$ (which has small probability). We obtain

$$\left|Y_0'\right| \leq \frac{C}{r}\|h\|_\infty + \frac{1}{4}\left\|Y_{\tau(r)}'\right\|_q$$

with $Y_s' = \Xi h(Z_s)$. We multiply by $\rho_\star^2(Z_0)$ instead of $\rho_\star(Z_0)$, and proceed as in Theorem 2. The result is

$$\left|dh(t,x)\right| \leq \frac{C_K}{t \wedge \rho^2(x) \wedge 1}\|h\|_\infty.$$

The extension to the manifold-valued case (framework of Theorem 3) can also be worked out; the main difference is that the equality (4.7) should be replaced by an inequality with an L^p norm; then the estimation (4.9) for the L^1 norm can also be done with the L^p norm, and we can prove

$$\left|dh(t,x)\right| \leq \frac{C_K}{t \wedge \rho^2(x) \wedge 1} \sup_{(z,z')} \delta(h(z), h(z')).$$

Acknowledgement. The author wishes to thank the referees for valuable comments on a first version of this work.

References

[1] M. Arnaudon, *Differentiable and analytic families of continuous martingales in manifolds with connection*, Probab. Theory Relat. Fields **108** (1997), 219–257.

[2] M. Arnaudon, X.M. Li, and A. Thalmaier, *Manifold-valued martingales, changes of probabilities, and smoothness of finely harmonic maps*, Ann. Inst. Henri Poincar, Probab. Stat. **35** (1999), no. 6, 765–791.

[3] M. Arnaudon and A. Thalmaier, *The differentiation of hypoelliptic diffusion semigroups*, preprint.

[4] ———, *Complete lifts of connections and stochastic Jacobi fields*, J. Math. Pures Appl. **77** (1998), 283–315.

[5] D. Bakry, *Transformations de Riesz pour les semigroupes symtriques*, Sminaire de Probabilits XIX (J. Azma and M. Yor, eds.), Lectures Notes in Math., vol. 1123, Springer, 1985.

[6] J.M. Bismut, *Large deviations and the Malliavin calculus*, Progress in Mathematics, vol. 45, Birkhuser, 1984.

[7] M. Cranston, *Gradient estimates on manifolds using coupling*, J. Functional Analysis **99** (1991), no. 1, 110–124.

[8] R.W.R. Darling, *Martingales on noncompact manifolds: Maximal inequalities and prescribed limits*, Ann. Inst. Henri Poincar, Probab. Stat. **32** (1996), no. 4, 431–454.

[9] C. Dellacherie and P. A. Meyer, *Probabilits et potentiel, Chapitres V VIII*, Hermann, 1980.

[10] D. Elworthy, *Geometric aspects of diffusions on manifolds*, Ecole d'Et de Probabilits de Saint-Flour XV–XVII, 1985–87 (P.L. Hennequin, ed.), Lect. Notes in Mathematics, vol. 1362, Springer, 1988.

[11] K.D. Elworthy, Y. Le Jan, and X.M. Li, *On the geometry of diffusion operators and stochastic flows*, Lect. Notes in Mathematics, vol. 1720, Springer, 1999.

[12] K.D. Elworthy and X.M. Li, *Formulae for the derivatives of heat semigroups*, J. Functional Analysis **125** (1994), no. 1, 252–286.

[13] K.D. Elworthy and M. Yor, *Conditional expectations for derivatives of certain stochastic flows*, Sminaire de Probabilits XXVII (M. Yor J. Azma, P.A. Meyer, ed.), Lect. Notes in Mathematics, vol. 1557, Springer, 1993, pp. 159–172.

[14] M. Emery, *Stochastic calculus in manifolds*, Universitext, Springer, 1989.

[15] P. Hsu, *Heat semigroup on a complete Riemannian manifold*, The Annals of Probability **17** (1989), no. 3, 1248–1254.

[16] J. Jost and C.J. Xu, *Subelliptic harmonic maps*, Trans. American Math. Soc. **350** (1998), no. 11, 4633–4649.

[17] W.S. Kendall, *Probability, convexity, and harmonic maps with small image I: Uniqueness and fine existence*, Proc. London Math. Soc. **61** (1990), no. 3, 371–406.

[18] ______, *Convexity and the hemisphere*, J. London Math. Soc. **43** (1991), no. 2, 567–576.

[19] ______, *Probability, convexity and harmonic maps II: Smoothness via probabilistic gradient inequalities*, J. Functional Analysis **126** (1994), no. 1, 226–257.

[20] S. Kusuoka and D. Stroock, *Applications of the Malliavin calculus, Part III*, J. Fac. Sci. Univ. Tokyo **34** (1987), 391–442.

[21] M. Ledoux, *The geometry of Markov diffusion generators*, Annales Fac. Sciences Toulouse **IX** (2000), no. 2, 305–366.

[22] J. Picard, *Martingales on Riemannian manifolds with prescribed limit*, J. Functional Analysis **99** (1991), no. 2, 223–261.

[23] ______, *Barycentres et martingales sur une varit*, Ann. Inst. Henri Poincar, Probab. Stat. **30** (1994), no. 4, 647–702.

[24] ______, *Smoothness of harmonic maps for hypoelliptic diffusions*, The Annals of Probability **28** (2000), no. 2, 643–666.

[25] D.W. Stroock, *An introduction to the analysis of paths on a Riemannian manifold*, Mathematical Surveys and Monographs, vol. 74, American Mathematical Society, 2000.

[26] D.W. Stroock and O. Zeitouni, *Variations on a theme by Bismut*, Astrisque, vol. 236, pp. 291–301, Soc. Math. de France, 1996.

[27] A. Thalmaier, *On the differentiation of heat semigroups and Poisson integrals*, Stochastics Stoch. Reports **61** (1997), 297–321.

[28] A. Thalmaier and F.Y. Wang, *Gradient estimates for harmonic functions on regular domains in Riemannian domains*, J. Functional Analysis **155** (1998), no. 1, 109–124.

[29] F.Y. Wang, *On estimation of the logarithmic Sobolev constant and gradient estimates of heat semigroups*, Probab. Theory Relat. Fields **108** (1997), 87–101.