Illinois Journal of Mathematics
Volume 54, Number 4, Winter 2010, Pages 1285–1311
S 0019-2082

THE DIFFERENTIATION OF HYPOELLIPTIC DIFFUSION SEMIGROUPS

MARC ARNAUDON AND ANTON THALMAIER

Dedicated to the memory of Donald Burkholder

Abstract. Basic derivative formulas are presented for hypoelliptic heat semigroups and harmonic functions extending earlier work in the elliptic case. According to our approach, special emphasis is placed on integration by parts formulas at the level of local martingales. Combined with the optional sampling theorem, this turns out to be an efficient way of dealing with boundary conditions, as well as with difficulties related to finite lifetime of the underlying diffusion. Our formulas require hypoellipticity of the diffusion in the sense of Malliavin calculus (integrability of the inverse Malliavin covariance) and are formulated in terms of the derivative flow, the Malliavin covariance and its inverse. Finally, some extensions to the nonlinear setting of harmonic mappings are discussed.

1. Introduction

Let M be a smooth n-dimensional manifold. On M consider a globally defined Stratonovich SDE of the type

$$\delta X = A(X)\delta Z + A_0(X)\,dt \tag{1.1}$$

with $A_0 \in \Gamma(TM)$, $A \in \Gamma(\mathbb{R}^r \otimes TM)$ for some r, and Z an $\mathbb{R}^r$-valued Brownian motion on some filtered probability space satisfying the usual completeness conditions. Here $\Gamma(TM)$, resp. $\Gamma(\mathbb{R}^r \otimes TM)$, denote the smooth sections over M of the tangent bundle TM, resp. the vector bundle $\mathbb{R}^r \otimes TM$.

Received April 2, 2010; received in final form March 28, 2011.
2010 *Mathematics Subject Classification.* Primary 58G32, 60H30. Secondary 60H10.

Solutions to (1.1) are diffusions with generator given in Hörmander form as

$$L = A_0 + \frac{1}{2}\sum_{i=1}^{r} A_i^2, \tag{1.2}$$

where $A_i = A(\cdot)e_i \in \Gamma(TM)$ and e_i the ith standard unit vector in $\mathbb{R}^r$.

There is a partial flow $X_t(\cdot)$, $\zeta(\cdot)$ to (1.1) such that for each $x \in M$ the process $X_t(x)$, $0 \le t < \zeta(x)$, is the maximal strong solution to (1.1) with starting point $X_0(x) = x$ and explosion time $\zeta(x)$. Adopting the notation $X_t(x,\omega) = X_t(x)(\omega)$, resp. $\zeta(x,\omega) = \zeta(x)(\omega)$ and

$$M_t(\omega) = \{x \in M : t < \zeta(x,\omega)\},$$

it further means that there exists a set $\Omega_0 \subset \Omega$ of full measure such that for all $\omega \in \Omega_0$ the following conditions hold:

(i) $M_t(\omega)$ is open in M for $t \ge 0$, that is, $\zeta(\cdot,\omega)$ is lower semicontinuous on M.
(ii) $X_t(\cdot,\omega) : M_t(\omega) \to M$ is a diffeomorphism onto an open subset $R_t(\omega)$ of M.
(iii) The map $s \mapsto X_s(\cdot,\omega)$ is continuous from $[0,t]$ to $C^\infty(M_t(\omega), M)$, for $t > 0$, when the latter is equipped with the C^∞-topology.

Thus, the differential $T_xX_t : T_xM \to T_{X_t}M$ of the map $X_t : M_t \to M$ is well-defined at each point $x \in M_t$, for all $\omega \in \Omega_0$. We also write X_{t*} for TX_t.

Let

$$(P_tf)(x) = \mathbb{E}\big[\big(f \circ X_t(x)\big)\mathbb{1}_{\{t<\zeta(x)\}}\big] \tag{1.3}$$

be the minimal semigroup associated to (1.1), acting on bounded measurable functions $f : M \to \mathbb{R}$.

Let $\mathrm{Lie}(A_0, A_1, \ldots, A_r)$ denote the Lie algebra generated by $A_0, \ldots, A_r$, that is, the smallest $\mathbb{R}$-vector space of vector fields on M containing $A_0, \ldots, A_r$ and being closed under Lie brackets. We suppose that (1.2) is non-degenerate in the sense that the ideal generated by $(A_1, \ldots, A_r)$ in $\mathrm{Lie}(A_0, A_1, \ldots, A_r)$ is the full tangent space at each point $x \in M$:

$$\mathrm{Lie}\big(A_i, [A_0, A_i] : i = 1, \ldots, r\big)(x) = T_xM \quad \text{for all } x \in M. \tag{H1}$$

Note that (H1) is equivalent to the following Hörmander condition for $\frac{\partial}{\partial t} + L$ on $\mathbb{R} \times M$:

$$\dim \mathrm{Lie}\left(\frac{\partial}{\partial t} + A_0, A_1, \ldots, A_r\right)(t,x) = n + 1 \quad \text{for all } (t,x) \in \mathbb{R} \times M.$$

By Hörmander's theorem, under (H1) the semigroup (1.3) is strongly Feller (mapping bounded measurable functions on M to bounded continuous functions on M) and has a smooth density $p \in C^\infty(]0,\infty[\times M \times M)$ such that

$$P\{X_t(x) \in dy, t < \zeta(x)\} = p(t,x,y)\,\mathrm{vol}(dy), \quad t > 0, x \in M,$$

see [7] for a probabilistic discussion.

In this paper, we are concerned with the problem of finding stochastic representations, under hypothesis (H1), for the derivative $d(P_t f)$ of (1.3) which do not involve derivatives of f. Analogously, in the situation of L-harmonic functions $u: D \to \mathbb{R}$, given on some domain D in M by its boundary values $u|\partial D$ via

$$u(x) = \mathbb{E}\big[u \circ X_{\tau(x)}(x)\big], \tag{1.4}$$

formulas are developed for du not involving derivatives of the boundary function; here $\tau(x)$ is the first exit time of $X(x)$ from D. See [4] for related integration by parts formulas on path space.

The paper is organized as follows. In Section 2, we collect some background on Malliavin calculus related to hypoelliptic diffusions. In Section 3, we explain our approach to integration by parts in the hypoelliptic case which leads to differentiation formulas for hypoelliptic semigroups. Section 4 is devoted to integration by parts formulas at the level of local martingales. In Section 5, control theoretic aspects related to differentiation formulas are discussed. It is shown that the solvability of a certain control problem leads to simple formulas in particular cases, however the method turns out not to cover the full hypoelliptic situation. We deal with the general situation in Section 7 where we refine the arguments of Sections 4 and 5 to give probabilistic representations for the derivative of semigroups and L-harmonic functions in the hypoelliptic case. A crucial step in this approach is the use of the optional sampling theorem to obtain local formulas by appropriate stopping times, as in the elliptic case [20], [21], [22]. Our formulas are in terms of the derivative flow and Malliavin's covariance; hence, they are neither unique nor intrinsic: the appearing terms depend on the specific SDE and not just on the generator.

Finally, in Section 8, we deal with possible extensions to nonlinear situations, like the case of harmonic maps and nonlinear heat equations for maps taking values in curved targets.

All presented formulas do not require full Hörmander's Lie algebra condition (H1) but rather invertibility and integrability of the inverse Malliavin covariance which is known to be slightly weaker, but still sufficient to imply hypoellipticity of $\frac{\partial}{\partial t} + L$. In particular, (H1) is allowed to fail on a collection of hypersurfaces. The reader is referred to [5] for precise statements in this direction.

2. Hypoellipticity and the Malliavin covariance

Let $B \in \Gamma(TM)$ be a vector field on M. We consider the push-forward $X_{t*}B$ (resp. pull-back $X_{t*}^{-1}B$) of B under the partial flow $X_t(\cdot)$ to the system (1.1), more precisely,

$$\begin{aligned} (X_{t*}B)_x &= (T_{X_t^{-1}(x)}X_t)B_{X_t^{-1}(x)}, \quad x \in R_t, \\ \big(X_{t*}^{-1}B\big)_x &= (T_{X_t(x)}X_t)^{-1}B_{X_t(x)}, \quad x \in M_t. \end{aligned} \tag{2.1}$$

Note that $X_{t*}B$, resp. $X_{t*}^{-1}B$, are smooth vector fields on R_t, resp. M_t, well-defined for all $\omega\in\Omega_0$. By definition,

$$(2.2)\qquad \begin{aligned}(X_{t*}B)_x f &= B_{X_t^{-1}(x)}(f\circ X_t),\quad x\in R_t,\\ \bigl(X_{t*}^{-1}B\bigr)_x f &= B_{X_t(x)}\bigl(f\circ X_t^{-1}\bigr),\quad x\in M_t,\end{aligned}$$

for germs f of smooth functions at x.

THEOREM 2.1. *The pushed vector fields $X_{t*}B$ and $X_{t*}^{-1}B$ as defined by* (2.1) *satisfy the following SDEs*:

$$(2.3)\qquad \delta(X_{t*}B)=\sum_{i=1}^r [X_{t*}B,A_i]\delta Z_t^i+[X_{t*}B,A_0]\,dt,$$

$$(2.4)\qquad \delta\bigl(X_{t*}^{-1}B\bigr)=\sum_{i=1}^r \bigl(X_{t*}^{-1}[A_i,B]\bigr)\delta Z_t^i+\bigl(X_{t*}^{-1}[A_0,B]\bigr)\,dt.$$

Proof. See Kunita [14], Section 5. □

We have the famous "invertibility of the Malliavin covariance matrix" under the Hörmander condition (H1), for example, see Bismut [7], Prop. 4.1.

THEOREM 2.2. *Suppose* (H1) *holds. Let σ be a predictable stopping time, $x\in M$. Then, a.s., for any predictable stopping time $\tau<\zeta(x)$, on $\{\sigma<\tau\}$*

$$\sum_{i=1}^r\int_\sigma^\tau \bigl(X_{s*}^{-1}A_i\bigr)_x\otimes\bigl(X_{s*}^{-1}A_i\bigr)_x\,ds\in T_xM\otimes T_xM$$

is a positive definite quadratic form on T_x^*M.

Almost surely, for each $t>0$,

$$(2.5)\qquad C_t(x):=\sum_{i=1}^r\int_0^t \bigl(X_{s*}^{-1}A_i\bigr)_x\otimes\bigl(X_{s*}^{-1}A_i\bigr)_x\,ds$$

defines a smooth section C_t of the bundle $TM\otimes TM$ over M_t with the property that under condition (H1) the symmetric "random matrices" $C_t(x)$ are invertible for $x\in M_t$ and $t>0$.

The following property is a key result in the Stochastic Calculus of Variations, for example, [3], [15], [18], [19].

REMARK 2.3. Under hypothesis (H1) and certain boundedness conditions on the vector fields $A_0,A_1,\dots,A_r$ (which are satisfied, for instance, if M is compact) we have $(\det C_t(x))^{-1}\in L^p$ for all $1\le p<\infty$. In the same way,

$$(2.6)\qquad \bigl(\det C_\sigma(x)\bigr)^{-1}\in L^p\quad \text{for } 1\le p<\infty$$

if $\sigma=\tau_D(x)$ or $\sigma=\tau_D(x)\wedge t$ for some $t>0$ where $\tau_D(x)$ is the first exit time of $X_\bullet(x)$ from some relatively compact open neighbourhood $D\ne M$ of x. Also note that $\tau_D(x)\in L^p$ for all $1\le p<\infty$, e.g. [6], Lemma 1.21.

We adopt the following notation. By definition, $C_t(x) \in T_xM \otimes T_xM$ which we may read as $C_t(x) : T_x^*M \to T_xM$, consequently $C_t(x)^{-1} : T_xM \to T_x^*M$. On the other hand,

$$(X_{s*}^{-1}A)_x : \mathbb{R}^r \to T_xM, \qquad z \mapsto \sum_{i=1}^r (X_{s*}^{-1}A_i)_x z^i. \tag{2.7}$$

Denoting by $(X_{s*}^{-1}A)_x^* : T_x^*M \to (\mathbb{R}^r)^* \equiv \mathbb{R}^r$ the adjoint (dual) map to (2.7), the Malliavin covariance writes as

$$C_t(x) = \int_0^t (X_{s*}^{-1}A)_x (X_{s*}^{-1}A)_x^* \, ds. \tag{2.8}$$

We identify $(\mathbb{R}^r)^*$ and $\mathbb{R}^r$.

If a non-degenerate inner product $\langle \cdot, \cdot \rangle$ on T_xM is given, we may think of $C_t(x) \in T_xM \otimes T_xM$ in equal terms as a positive definite bilinear form on T_xM:

$$\langle C_t(x)u, v\rangle = \sum_{i=1}^r \int_0^t \langle (X_{s*}^{-1}A_i)_x, u\rangle \langle (X_{s*}^{-1}A_i)_x, v\rangle \, ds, \quad u, v \in T_xM.$$

3. A basic integration by parts argument

In this section, we explain an elementary strategy for integration by parts formulas which will serve us as a guideline in the sequel. The argument is inspired by Bismut's original approach to Malliavin calculus [7].

Consider again the SDE (1.1) and assume (H1) to be satisfied. For simplicity, we suppose that M is compact. Let a be a predictable process taking values in $T_xM \otimes (\mathbb{R}^r)^* \equiv T_xM \otimes \mathbb{R}^r$ and $\lambda \in T_x^*M$ such that for each $t > 0$,

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^t |a_s\lambda|^2 \, ds\right)\right] < \infty, \quad \lambda \text{ in a neighbourhood about } 0. \tag{3.1}$$

We use the $\mathbb{R}^r$-valued process $a\lambda$ to perturb the Brownian motion Z,

$$dZ^\lambda = dZ + a\lambda \, dt,$$

and consider the Girsanov exponential $G_\bullet^\lambda$ defined by

$$G_t^\lambda = \exp\left(-\int_0^t \langle a_s\lambda, dZ_s\rangle - \frac{1}{2}\int_0^t |a_s\lambda|^2 \, ds\right). \tag{3.2}$$

Write X^λ for the flow to our SDE driven by the perturbed Brownian motion Z^λ, analogously $C_\bullet^\lambda(x)$ etc. By definition, $C_\bullet^\lambda(x) \in T_xM \otimes T_xM$ is a linear map from T_x^*M to T_xM and $C_\bullet^\lambda(x)^{-1} : T_xM \to T_x^*M$.

LEMMA 3.1. *For any vector field $B \in \Gamma(TM)$, we have*

$$\frac{\partial}{\partial \lambda_k}\bigg|_{\lambda=0} (X_{t*}^\lambda)^{-1}(B) = \sum_{i=1}^r \left[\int_0^t X_{s*}^{-1}(A_i) a_s^{ik} \, ds, X_{t*}^{-1}(B)\right] \tag{3.3}$$

in terms of the Lie bracket $[\cdot,\cdot]$. The index k refers to the coordinates in T_x^*M and the index i to the components in $\mathbb{R}^r$.

Proof. Note that $X_t^\lambda(x) = X_t \circ \varrho_t^\lambda(x)$ where $\varrho^\lambda(x)$ solves

$$\begin{cases} d\varrho_t^\lambda = X_{t*}^{-1}(A)\big(\varrho_t^\lambda\big)a_t\lambda\,dt, \\ \varrho_0^\lambda = x. \end{cases}$$

In particular, we have

$$\frac{\partial}{\partial\lambda_k}\bigg|_{\lambda=0}\varrho_t^\lambda = \sum_{i=1}^r \int_0^t X_{s*}^{-1}(A_i)a_s^{ik}\,ds.$$

Moreover, from $X_{t*}^\lambda(x) = (T_{\varrho_t^\lambda(x)}X_t)(T_x\varrho_t^\lambda)$ we conclude that

$$\begin{aligned} \big((X_{t*}^\lambda)^{-1}B\big)_x &= \big(T_x\varrho_t^\lambda\big)^{-1}(T_{\varrho_t^\lambda(x)}X_t)^{-1}B\big(X_t\circ\varrho_t^\lambda(x)\big) \\ &\equiv \big(T_x\varrho_t^\lambda\big)^{-1}\big(X_{t*}^{-1}B\big)_{\varrho_t^\lambda(x)}. \end{aligned}$$

This gives the claim by definition of the bracket. □

THEOREM 3.2. *Let M be compact and $f \in C^1(M)$. Assume that (H1) is satisfied. Then, for each $v \in T_xM$,*

$$d(P_tf)_xv = \mathbb{E}\big[\big(f\circ X_t(x)\big)\Phi_tv\big], \tag{3.4}$$

where Φ is an adapted process with values in T_x^*M such that each Φ_t is L^p for any $1 \le p < \infty$.

Proof. We fix x and identify T_xM with $\mathbb{R}^n$. By Girsanov's theorem, for $v \in T_xM$, the expression

$$H_k(\lambda) = \sum_\ell \mathbb{E}\big[\big(f\circ X_t^\lambda(x)\big)\cdot G_t^\lambda \cdot \big(C_t^\lambda(x)^{-1}\big)_{k\ell}v^\ell\big]$$

is independent of λ for any C^1-function f on M. Thus,

$$\sum_k \frac{\partial}{\partial\lambda_k}\bigg|_{\lambda=0} H_k(\lambda) = 0$$

which gives

$$\begin{aligned} &\sum_{i,k,\ell}\mathbb{E}\left[(D_if)\big(X_t(x)\big)\left(X_{t*}\int_0^t \big(X_{s*}^{-1}A\big)_x a_s\,ds\right)_{ik}\big(C_t(x)^{-1}\big)_{k\ell}v^\ell\right] \\ &\quad = -\sum_{k,\ell}\mathbb{E}\left[f\big(X_t(x)\big)\frac{\partial}{\partial\lambda_k}\bigg|_{\lambda=0}\big(G_t^\lambda\big(C_t^\lambda(x)^{-1}\big)_{k\ell}v^\ell\big)\right] \\ &\quad = -\sum_{k,\ell}\mathbb{E}\left[f\big(X_t(x)\big)\left(\left(\frac{\partial}{\partial\lambda_k}\bigg|_{\lambda=0}G_t^\lambda\right)\big(C_t(x)^{-1}\big)_{k\ell}\right.\right. \\ &\qquad \left.\left. + \frac{\partial}{\partial\lambda_k}\bigg|_{\lambda=0}\big(C_t^\lambda(x)^{-1}\big)_{k\ell}\right)v^\ell\right]. \end{aligned}$$

Note that

$$\frac{\partial}{\partial\lambda_k}\bigg|_{\lambda=0} G_t^\lambda = -\left(\int_0^t a_s^*\,dZ_s\right)_k,$$

where a^* taking values in $T_xM\otimes(\mathbb{R}^r)^*$ is defined as the adjoint to a. Furthermore,

$$\frac{\partial}{\partial\lambda_k}\bigg|_{\lambda=0} C_t^\lambda(x)^{-1} = -C_t(x)^{-1}\left(\frac{\partial}{\partial\lambda_k}\bigg|_{\lambda=0} C_t^\lambda(x)\right)C_t(x)^{-1}.$$

Recall that $(X_{s*}^{-1}A)_x\in(\mathbb{R}^r)^*\otimes T_xM$. We set

$$a_s\equiv a_s^n := \big(X_{s*}^{-1}A\big)_x^*\mathbb{1}_{\{s\le\tau_n\}}\in T_xM\otimes\big(\mathbb{R}^r\big)^*,$$

where (τ_n) is an increasing sequence of stopping times such that $\tau_n\nearrow t$ and such that each $a_\bullet^n$ satisfies condition (3.1). This gives a formula of the type

$$\mathbb{E}\big[(df)_{X_t(x)}X_{t*}C_{\tau_n}(x)C_t(x)^{-1}v\big]=\mathbb{E}\big[\big(f\circ X_t(x)\big)\cdot\Phi_t^n v\big]. \tag{3.5}$$

Finally, taking the limit as $n\to\infty$, we get

$$d(P_tf)_xv=\mathbb{E}\big[(df)_{X_t(x)}X_{t*}v\big]=\mathbb{E}\big[\big(f\circ X_t(x)\big)\cdot\Phi_tv\big], \tag{3.6}$$

where

$$\begin{aligned}\Phi_tv = {}&\left(\int_0^t\big(X_{s*}^{-1}A\big)_x\,dZ_s\right)C_t^{-1}(x)v\\ &+\sum_{k,\ell}\left(C_t(x)^{-1}\left(\frac{\partial}{\partial\lambda_k}\bigg|_{\lambda=0}C_t^\lambda(x)\right)C_t(x)^{-1}\right)_{k\ell}v^\ell\end{aligned}$$

which can be further evaluated by means of (3.3). Equation (3.3) also allows to conclude that $\Phi_t\in\bigcap_{p\ge1}L^p$. □

4. Integration by parts at the level of local martingales

Let $F(\cdot,X_\bullet(x))$, $x\in M$ be a family of local martingales where F is differentiable in the second variable with a derivative jointly continuous in both variables. Here, $X_\bullet(x)\equiv(X_t(x))_{t\ge0}$ denotes the solution to the SDE (1.1) starting from x. We are mainly interested in the following two cases:

(1) $F(\cdot,X_\bullet(x))=u\circ X_\bullet(x)$ for some L-harmonic function u on M, and
(2) $F(\cdot,X_\bullet(x))=(P_{t-\bullet}f)(X_\bullet(x))$ for some bounded measurable f on M and $t>0$.

Let dF denote the differential of F with respect to the second variable.

THEOREM 4.1. *Let $F(\cdot,X_\bullet(x))$, $x\in M$ be a family of local martingales as described above. Then, for any predictable $\mathbb{R}^r$-valued process k in $L^2_{\mathrm{loc}}(Z)$,*

$$\begin{aligned}&dF\big(\cdot,X_\bullet(x)\big)(T_xX_\bullet)\int_0^\bullet\big(X_{s*}^{-1}A\big)_xk_s\,ds\\ &\quad-F\big(\cdot,X_\bullet(x)\big)\int_0^\bullet\langle k,dZ\rangle,\quad x\in M,\end{aligned} \tag{4.1}$$

is a family of local martingales. Here Z denotes again the driving Brownian motion on $\mathbb{R}^r$.

Proof. (By means of Girsanov). For ε varying locally about 0, consider the SDE

$$\delta X^\varepsilon = A\big(X^\varepsilon\big)\delta Z^\varepsilon + A_0\big(X^\varepsilon\big)\,dt \tag{4.2}$$

with the perturbed driving process $dZ^\varepsilon := dZ + \varepsilon k\,dt$. Then, for each ε,

$$F\big(\cdot, X^\varepsilon_\bullet(x)\big)G^\varepsilon_\bullet \tag{4.3}$$

is again a local martingale when the Girsanov exponential $G^\varepsilon_\bullet$ is defined by

$$G^\varepsilon_r = \exp\bigg(-\int_0^r \varepsilon\langle k, dZ\rangle - \frac{1}{2}\varepsilon^2\int_0^r |k|^2\,ds\bigg).$$

Moreover, the local martingale (4.3) depends C^1 on the parameter ε (in the topology of compact convergence in probability), thus

$$\frac{\partial}{\partial\varepsilon}\bigg|_{\varepsilon=0} F\big(\cdot, X^\varepsilon_\bullet(x)\big)G^\varepsilon_\bullet = \frac{\partial}{\partial\varepsilon}\bigg|_{\varepsilon=0} F\big(\cdot, X^\varepsilon_\bullet(x)\big) + F\big(\cdot, X_\bullet(x)\big)\frac{\partial}{\partial\varepsilon}\bigg|_{\varepsilon=0} G^\varepsilon_\bullet$$

is also a local martingale. Taking into account that

$$\frac{\partial}{\partial\varepsilon}\bigg|_{\varepsilon=0} X^\varepsilon_r(x) = X_{r*}\int_0^r X^{-1}_{s*}A\big(X_s(x)\big)k_s\,ds$$

and

$$\frac{\partial}{\partial\varepsilon}\bigg|_{\varepsilon=0} G^\varepsilon_r = -\int_0^r \langle k, dZ\rangle,$$

we get the claim. □

Alternative proof of Theorem 4.1. First, note that $m_s := dF(s,\cdot)_{X_s(x)}X_{s*}$, as the derivative of a family of local martingales, is a local martingale in T^*_xM, see [1]. Thus, also

$$n_s := m_s h_s - \int_0^s m_r\,dh_r$$

is a local martingale for any T_xM-valued adapted process h locally of bounded variation. Choosing

$$h = \int_0^\bullet \big(X^{-1}_{s*}A\big)_x k_s\,ds$$

and taking into account that

$$F\big(\cdot, X_\bullet(x)\big) = \int_0^\bullet dF(s,\cdot)_{X_s(x)}A\big(X_s(x)\big)\,dZ,$$

the claim follows by noting that

$$\int_0^\bullet dF(s,\cdot)_{X_s(x)}X_{s*}\,dh_s \overset{\mathrm{m}}{=} F\big(\cdot, X_\bullet(x)\big)\int_0^\bullet \langle k, dZ\rangle,$$

where $\overset{\mathrm{m}}{=}$ denotes equality modulo local martingales. □

Let a be a predictable process taking values in $T_xM \otimes (\mathbb{R}^r)^*$ as in the last section. The calculation above shows that

$$n_s := dF(s,\cdot)_{X_s(x)} X_{s*}\left(\int_0^s \left(X_{r*}^{-1}A\right)_x a_r\,dr\right) - F\big(s, X_s(x)\big)\int_0^s a_r^*\,dZ_r$$

is a local martingale in T_xM which implies that

$$N_s := n_s h_s - \int_0^s n_r\,dh_r$$

is also a local martingale for any T_x^*M-valued adapted process h locally of bounded variation. In particular, choosing again $a_s = (X_{s*}^{-1}A)_x^*$, we get

$$\begin{aligned} N_s = {} & dF(s,\cdot)_{X_s(x)} X_{s*} C_s(x) h_s - F\big(s, X_s(x)\big)\left(\int_0^s \left(X_{r*}^{-1}A\right)_x dZ_r\right) h_s \\ & - \int_0^s dF(r,\cdot)_{X_r(x)} X_{r*} C_r(x)\,dh_r \\ & + \int_0^s F\big(r, X_r(x)\big)\left(\int_0^r \left(X_{\rho*}^{-1}A\right)_x dZ_\rho\right) dh_r. \end{aligned}$$

For the last term, it is trivial to observe that

$$\begin{aligned} & \int_0^s F\big(r, X_r(x)\big)\left(\int_0^r \left(X_{\rho*}^{-1}A\right)_x dZ_\rho\right) dh_r \\ & \quad \overset{\mathrm{m}}{=} F\big(s, X_s(x)\big)\int_0^s \left(\int_0^r \left(X_{\rho*}^{-1}A\right)_x dZ_\rho\right) dh_r. \end{aligned}$$

Now the idea is to take h of the special form $h_s = C_s(x)^{-1}k_s$ for some adapted T_xM-valued process k locally pathwise of bounded variation such that in addition $k_\tau = v$ and $k_s = 0$ for s close to 0. Then the remaining problem is to replace

$$\int_0^s dF(r,\cdot)_{X_r(x)} X_{r*} C_r(x)\,dh_r \tag{4.4}$$

modulo local martingales by expressions not involving derivatives of F. This however seems to be difficult in general, but in Section 7 we show that, more easily, the expectation of (4.4) can be rewritten in terms not involving derivatives of F.

5. Hypoelliptic diffusions and control theory

The following two corollaries are immediate consequences of Theorem 4.1.

COROLLARY 5.1. *Let $f : M \to \mathbb{R}$ be bounded measurable. Fix $x \in M$ and $v \in T_xM$. Then, for any predictable $\mathbb{R}^r$-valued process k in $L^2_{\mathrm{loc}}(Z)$,*

$$(dP_{t-\bullet}f)_{X_\bullet(x)}(T_xX_\bullet)\left[v + \int_0^\bullet \left(X_{s*}^{-1}A\right)_x k_s\,ds\right] - (P_{t-\bullet}f)\big(X_\bullet(x)\big)\int_0^\bullet \langle k, dZ\rangle$$

is a local martingale on the interval $[0, t\wedge\zeta(x)[$.

Note that $(dP_{t-\bullet}f)_{X_\bullet(x)}(T_xX_\bullet)v$ is a local martingale as the derivative of the local martingale $(P_{t-\bullet}f)(X_\bullet(x))$ at x in the direction v, see [1].

COROLLARY 5.2. *Assume that M is compact with nonempty smooth boundary ∂M. Let $u \in C(M)$ be L-harmonic on $M \backslash \partial M$. Fix $x \in M \backslash \partial M$ and $v \in T_xM$. Then, for any predictable $\mathbb{R}^r$-valued process k in $L^2_{\mathrm{loc}}(Z)$,*

$$(du)_{X_\bullet(x)}(T_xX_\bullet)\left[v + \int_0^\bullet \left(X_{s*}^{-1}A\right)_x k_s\,ds\right] - u\big(X_\bullet(x)\big)\int_0^\bullet \langle k, dZ\rangle$$

is a local martingale on the interval $[0, \tau(x)[$ where $\tau(x)$ is the first hitting time of $X_\bullet(x)$ at ∂M.

PROBLEM 5.3 (Control problem). Let $x \in M$ and $v \in T_xM$. Consider the random dynamical system

$$\begin{cases} \dot{h}_s = \left(X_{s*}^{-1}A\right)_x k_s, \\ h_0 = v. \end{cases} \tag{5.1}$$

Let $\sigma = \tau_D(x)$, resp., $\sigma = \tau_D(x) \wedge t$ for some $t > 0$, where $\tau_D(x)$ is the first exit time of $X_\bullet(x)$ from some relatively compact open neighbourhood D of x. We are concerned with the problem of finding predictable processes k taking values in $\mathbb{R}^r$ such that $h_\sigma = 0$, a.s.

EXAMPLE 5.4. Assume L to be elliptic, that is, $A(x) : \mathbb{R}^r \to T_xM$ surjective for each $x \in M$. Then

$$k_s = A^*\big(X_s(x)\big)T_xX_s\dot{h}_s$$

solves Problem 5.3 if the terms are defined as follows: $A^*(\cdot) \in \Gamma(T^*M \otimes \mathbb{R}^r)$ is a smooth section and (pointwise) right-inverse to $A(\cdot)$, that is, $A(x)A^*(x) = \mathrm{id}_{T_xM}$ for $x \in M$, the process h may be any adapted process with values in T_xM and with absolutely continuous sample paths (e.g., paths in the Cameron–Martin space $\mathbb{H}(\mathbb{R}_+, T_xM)$) such that $h_0 = v$ and $h_\sigma = 0$, a.s. Thus, for elliptic L, there are "controls" k transferring system (5.1) from v to 0 in time σ, moreover it is even possible to follow prescribed trajectories $s \mapsto h_s$ from v to 0. In the hypoelliptic case, this cannot be achieved in general, since the right-hand side in

$$(T_xX_s)\dot{h}_s = A\big(X_s(x)\big)k_s$$

is allowed to be degenerate.

Under the assumption that Problem 5.3 has an affirmative solution, we get differentiation formulas in a straightforward way.

THEOREM 5.5. *Let $f : M \to \mathbb{R}$ be bounded measurable, $x \in M$, $v \in T_xM$, $t > 0$. Let D be a relatively compact open neighbourhood of x and $\sigma = \tau_D(x) \wedge t$ where $\tau_D(x)$ is the first exit time of $X_\bullet(x)$ from D. Suppose there exists an $\mathbb{R}^r$-valued predictable process k such that*

$$\int_0^\sigma \left(X_{s*}^{-1}A\right)_x k_s\,ds \equiv v, \quad a.s.,$$

and $(\int_0^\sigma |k_s|^2\,ds)^{1/2} \in L^{1+\varepsilon}$ *for some* $\varepsilon > 0$. *Then*

$$d(P_t f)_x v = \mathbb{E}\left[f\big(X_t(x)\big) \mathbb{1}_{\{t<\zeta(x)\}} \int_0^\sigma \langle k, dZ\rangle \right], \tag{5.2}$$

where $P_t f$ *is the minimal semigroup defined by* (1.3).

Proof. It is enough to check that the local martingale defined in Theorem 4.1 is actually a uniformly integrable martingale on the interval $[0,\sigma]$. The claim then follows by taking expectations, noting that $(P_{t-\sigma} f)(X_\sigma(x)) = \mathbb{E}^{\mathscr{F}_\sigma}[f(X_t(x))\mathbb{1}_{\{t<\zeta(x)\}}]$. See Theorem 2.4 in [20] for technical details. □

Along the same lines, now exploiting Corollary 5.2, the following result can be derived.

THEOREM 5.6. *Let* M *be compact with smooth boundary* $\partial M \neq \varnothing$ *and let* $u \in C(M)$ *be* L-*harmonic on* $M\backslash\partial M$. *Let* $x \in M\backslash\partial M$ *and* $v \in T_x M$. *Denote* $\tau(x)$ *the first hitting time of* $X_{\boldsymbol{\cdot}}(x)$ *at* ∂M. *Suppose there exists an* $\mathbb{R}^r$-*valued predictable process* k *such that*

$$\int_0^{\tau(x)} \big(X_{s*}^{-1} A\big)_x k_s\,ds \equiv v, \quad a.s.,$$

and $(\int_0^{\tau(x)} |k_s|^2\,ds)^{1/2} \in L^{1+\varepsilon}$ *for some* $\varepsilon > 0$. *Then the following formula holds*:

$$(du)_x v = \mathbb{E}\left[u\big(X_{\tau(x)}(x)\big) \int_0^{\tau(x)} \langle k, dZ\rangle \right]. \tag{5.3}$$

In the elliptic case, formulas of type (5.2) and (5.3) have been used in [22] to establish gradient estimates for $P_t f$ and for harmonic functions u, see also [9] for extensions from functions to to sections. Nonlinear generalizations of the elliptic case, for example, to harmonic maps and solutions of the nonlinear heat equations, are treated in [2].

As explained, differentiation formulas may be obtained from the local martingales (4.1) by taking expectations if there is a "control" (k_s) transferring the system (5.1) from $h_0 = v$ to $h_\sigma = 0$. Solvability of the "control problem" is more or less necessary for this approach, as is explained in the following remark.

REMARK 5.7. Consider the general problem of finding semimartingales h, Φ with $h_0 = v$ and $\Phi_0 = 0$ where h is $T_x M$-valued and Φ real-valued such that

$$n_s = (dF_s)_{X_s(x)} X_{s*} h_s + F_s\big(X_s(x)\big)\Phi_s, \quad s \ge 0 \tag{5.4}$$

is a local martingale for any space–time transformation F of the diffusion $X(x)$ such that $F_s(X_s(x)) \equiv F(s, X_s(x))$ is a local martingale. In the notion of quasiderivatives, as used by Krylov [16], [17], this means that $\xi := (T_x X)h$ is a F-quasiderivative for X along ξ at x and Φ its F-accompanying process. Suppose that h takes paths in the Cameron–Martin space $\mathbb{H}(\mathbb{R}_+, T_x M)$. Then, by choosing $F \equiv 1$, we see that Φ itself should already be a local martingale, say $\Phi_s = \int_0^s \langle k_r, dZ_r \rangle$. Thus,

$$n \overset{\mathrm{m}}{=} \int_0^{\bullet} (dF_r)_{X_r(x)} X_{r*}\, dh_r + \int_0^{\bullet} (dF_r)_{X_r(x)} A\big(X_r(x)\big) k_r\, dr$$

which implies

$$\int_0^{\bullet} (dF_r)_{X_r(x)} X_{r*}\, dh_r + \int_0^{\bullet} (dF_r)_{X_r(x)} A\big(X_r(x)\big) k_r\, dr \equiv 0,$$

that is, $(dF_s)_{X_s(x)} X_{s*} \dot{h}_s + (dF_s)_{X_s(x)} A(X_s(x)) k_s \equiv 0$ for all F of the above type. Hence, assuming local richness of transformations F of this type, we get for $s \geq 0$,

$$X_{s*} \dot{h}_s + A\big(X_s(x)\big) k_s \equiv 0$$

or

$$\dot{h}_s + \big(X_{s*}^{-1} A\big)_x k_s = 0,$$

which means that k solves the "control problem."

Coming back to Problem 5.3, we note that since the problem is unaffected by changing M outside of D, we may assume that M is already compact. It is also enough to deal with the case $\sigma = \tau_D(x)$ where D has smooth boundary.

PROBLEM 5.8 (Modified control problem). Let

$$c_s(x) = \frac{d}{ds} C_s(x) = \sum_{i=1}^{r} \big(X_{s*}^{-1} A_i\big)_x \otimes \big(X_{s*}^{-1} A_i\big)_x.$$

Confining the consideration to $\mathbb{R}^r$-valued processes k of the special form

$$k_s = \sum_{i=1}^{r} \big\langle \big(X_{s*}^{-1} A_i\big)_x, u_s \big\rangle e_i \tag{5.5}$$

for some adapted $T_x M$-valued process u, we observe that Problem 5.3 reduces to finding predictable $T_x M$-valued processes u such that

$$\begin{cases} \dot{h}_s = c_s(x) u_s, \\ h_0 = v \quad \text{and} \quad h_\sigma = 0. \end{cases} \tag{5.6}$$

This Problem 5.8, as well as Problem 5.3, have an affirmative solution in many cases. However, in the general situation, both problems are not solvable under hypothesis (H1), as will be shown in the next section.

6. Solvability of the control problem: Examples and counterexamples

We start discussing an example with solvability of the control conditions in a non-elliptic situation.

EXAMPLE 6.1. Let $M=\mathbb{R}^2$ and $A_0\equiv 0$, $A_1(x)=(1,0)$, $A_2(x)=(0,x_1)$. Then $[A_1,A_2](x)=(0,1)$. The solution to

$$\delta X = A(X)\delta Z$$

starting from $x=(x^1,x^2)$ is given by

$$X_t(x)=\left(x^1+Z_t^1, x^2+x^1Z_t^2+\int_0^t Z_s^1\,dZ_s^2\right).$$

Consequently

$$\left(X_{s*}^{-1}A\right)(x)=\begin{pmatrix}1 & 0\\ -Z_s^2 & X_s^1\end{pmatrix},$$

and the control problem at $x=0$ comes down to finding k such that

$$\dot{h}_s=\begin{pmatrix}1 & 0\\ -Z_s^2 & Z_s^1\end{pmatrix}k_s,\qquad h_0=v,\qquad h_\sigma=0,$$

and $(\int_0^\sigma |k_s|^2\,ds)^{1/2}\in L^{1+\varepsilon}$. We may assume that $|v|=1$, and will further assume that $\sigma=\tau_D$ or $\sigma=\tau_D\wedge t$ where D is some relatively compact neighbourhood of the origin in $\mathbb{R}^2$. (After possibly shrinking D, we may also assume that D is open with smooth boundary.) Note that

$$c_s(0)=\left(X_{s*}^{-1}A\right)_0\left(X_{s*}^{-1}A\right)_0^*=\begin{pmatrix}1 & -Z_s^2\\ -Z_s^2 & |Z_s|^2\end{pmatrix}.$$

Thus if $\lambda_{\min}(s)$ denotes the smallest eigenvalue of $c_s(0)$, then

$$\lambda_{\min}(s)\geq\frac{(Z_s^1)^2}{1+|Z_s|^2}. \tag{6.1}$$

(Indeed, let $a:=Z_s^1$, $b:=Z_s^2$, and $x:=1+|Z_s|^2=1+a^2+b^2$; then

$$\lambda_{\min}(s)=\frac{x-\sqrt{x^2-4a^2}}{2}=\frac{x}{2}\left[1-\sqrt{1-\frac{4a^2}{x^2}}\right]\geq\frac{a^2}{x},$$

where we used $1-\sqrt{1-x}\geq x/2$.)

We construct h by solving the equation

$$\dot{h}_s=-\varphi^{-2}(X_s,Z_s)c_s(0)\frac{h_s}{|h_s|},\qquad h_0=v, \tag{6.2}$$

where $X_s=X_s(0)$ and φ is chosen in such a way that

$$\sigma':=\inf\{s\geq 0: h_s=0\}\leq\sigma.$$

More precisely, take $\varphi_1\in C^2(\bar{D})$ with $\varphi_1|\partial D=0$ and $\varphi_1>0$ in D. Similarly, for some large ball B in $\mathbb{R}^2$ about 0 (containing D), let $\varphi_2\in C^2(\bar{B})$ with

$\varphi_2|\partial B = 0$ and $\varphi_2 > 0$ in B. Let $\varphi(x,z) := \varphi_1(x)\varphi_2(z)$. We only deal with the case $\sigma = \tau_D$, the case $\sigma = \tau_D \wedge t$ is dealt with an obvious modification of (6.2). Now, arguing as in the elliptic case, one shows

$$\int_0^\sigma \varphi^{-2}(X_s, Z_s)\, ds = \infty, \quad \text{a.s.}$$

Consequently, since $Z^1_\sigma \neq 0$ with probability 1, we may conclude that also

$$\int_0^\sigma \varphi^{-2}(X_s, Z_s)\frac{(Z^1_s)^2}{1+|Z_s|^2}\, ds = \infty, \quad \text{a.s.}$$

Note that

$$\frac{d}{ds}|h_s| = \frac{\langle \dot h_s, h_s\rangle}{|h_s|} = \frac{-\varphi^{-2}(X_s,Z_s)\langle c_s(0)h_s, h_s\rangle}{|h_s|^2},$$

and hence by means of (6.1),

$$1 - |h_t| \geq \int_0^t \varphi^{-2}(X_s, Z_s)\lambda_{\min}(s)\, ds \geq \int_0^t \varphi^{-2}(X_s, Z_s)\frac{(Z^1_s)^2}{1+|Z_s|^2}\, ds$$

which shows in particular that

$$\sigma' \leq \inf\left\{t \geq 0 : \int_0^t \varphi^{-2}(X_s, Z_s)\frac{(Z^1_s)^2}{1+|Z_s|^2}\, ds = 1\right\}.$$

It remains to verify the integrability condition, that is, $(\int_0^{\sigma'} |k_s|^2\, ds)^{1/2} \in L^{1+\varepsilon}$ where

$$k_s = -\varphi^{-2}(X_s, Z_s)\big(X_{s*}^{-1}A\big)_0^*\frac{h_s}{|h_s|}.$$

But, since on the interval $[0,\sigma]$ the Brownian motion Z stays in a compact ball B, and thus

$$\left|\big(X_{s*}^{-1}A\big)_0^*\frac{h_s}{|h_s|}\right| \leq C$$

for some constant C, we are left to check

$$\left(\int_0^{\sigma'} \varphi^{-4}(X_s, Z_s)\, ds\right)^{1/2} \in L^{1+\varepsilon}$$

which is done as in the elliptic case.

Contrary to Example 6.1, the next example gives a negative result showing that in general Problem 5.3 is not always solvable.

EXAMPLE 6.2 (J. Picard). Let $M = \mathbb{R}^3$ and take

$$A_0(x) = (0,0,0), \qquad A_1(x) = (1,0,0), \qquad A_2(x) = \big(0,1,x^1\big)$$

which obviously satisfy (H1). Then SDE (1.1) reads as

$$X_t(x) = x + \left(Z^1_t, Z^2_t, x^1 Z^2_t + \int_0^t Z^1_s\, dZ^2_s\right).$$

In particular,

$$\big(X_{t*}^{-1}A_1\big)(0)=\big(1,0,-Z_t^2\big),\qquad \big(X_{t*}^{-1}A_2\big)(0)=\big(0,1,Z_t^1\big).$$

Thus, (5.1) is given by

$$\dot h_s=\big(k_s^1,k_s^2,Z_s^1k_s^2-Z_s^2k_s^1\big),$$

where the problem is to find h such that $h_0=v=(v^1,v^2,v^3)$ and $h_\sigma=0$. By extracting the third coordinate, we get $\int_0^\sigma Z_s^1k_s^2\,ds-\int_0^\sigma Z_s^2k_s^1\,ds=-v^3$. On the other hand, an integration by parts yields

$$\int_0^\sigma Z_s^2k_s^1\,ds-\int_0^\sigma Z_s^1k_s^2\,ds=-\int_0^\sigma h_s^1\,dZ_s^2+\int_0^\sigma h_s^2\,dZ_s^1,$$

where the condition on the integrability of k implies that $-\int_0^\sigma h_s^1\,dZ_s^2+\int_0^\sigma h_s^2\,dZ_s^1$ is L^1 with expectation equal to 0. Combining both facts, we conclude that there is no solution satisfying the integrability condition if $v^3\neq 0$.

Note that if σ is not in L^1, then the condition on the integrability of k does not imply any more that $\int_0^\sigma h_s^1\,dZ_s^2+\int_0^\sigma h_s^2\,dZ_s^1$ is in L^1.

REMARK 6.3. In Example 6.2, Malliavin's covariance is explicitly given by

$$\begin{aligned}\big\langle C_t(0)u,u\big\rangle&=\sum_{i=1}^2\int_0^t\big\langle\big(X_{r*}^{-1}A_i\big)(0),u\big\rangle^2\,dr\\&=\int_0^t\big[\big(u^1-u^3Z_r^2\big)^2+\big(u^2+u^3Z_r^1\big)^2\big]\,dr.\end{aligned}$$

Of course, $C_t(0)-C_s(0)=\int_s^t c_r(0)\,dr$ is non-degenerate for all $s<t$, nevertheless $\lambda_{\min}c_s(0)=0$ for each fixed s, indeed:

$$\big\langle c_s(0)u,u\big\rangle=\big(u^1-u^3Z_s^2\big)^2+\big(u^2+u^3Z_s^1\big)^2,\quad u\in T_0M.$$

The negative result of Example 6.2 depends very much on the fact that $\sigma=\sigma_D$ is the first exit time of the diffusion from a relatively compact neighbourhood of its starting point. The situation changes completely if we allow arbitrarily large stopping times σ (not necessarily exit times from compact sets).

In the remainder of this section, we give sufficient conditions for solvability of the control problem. We assume that diffusions with generator L have infinite lifetime, but do no longer assume that the stopping time σ is of a given type. The question whether in this situation, given solvability of the control problem, the local martingales defined in Theorem 4.1 are still uniformly integrable martingales, needs to be checked from case to case.

We consider the following two conditions:

CONDITION (C1). There exists a positive constant α such that for any continuous (non-necessarily adapted) process u_t, taking values in $\{w \in T_xM, \|w\| = 1\}$ and converging to u almost surely,

$$\int_0^\infty \langle c_s(x)u_s, u_s\rangle \mathbb{1}_{\{\cos(c_s(x)u_s,u_s)>\alpha\}}\, ds = \infty \quad \text{a.s.} \tag{6.3}$$

CONDITION (C2). There exists a positive constant α such that for any $u_0 \in \{w \in T_xM, \|w\| = 1\}$, there exists a neighbourhood V_{u_0} of u_0 in $\{w \in T_xM, \|w\| = 1\}$, such that

$$\int_0^\infty \inf_{u\in V_{u_0}} \left(\langle c_s(x)u, u\rangle \mathbb{1}_{\{\cos(c_s(x)u,u)>\alpha\}}\right) ds = \infty \quad \text{a.s.} \tag{6.4}$$

The following result is immediate:

PROPOSITION 6.4. *Condition (C2) implies Condition (C1).*

Now we prove that the control problem is solvable under Condition (C1).

PROPOSITION 6.5. *Under Condition (C1), the control problem is solvable. More precisely, considering the random dynamical system*

$$\begin{cases} \dot h_s = \left(X_{s*}^{-1}A\right)_x k_s, \\ h_0 = v, \end{cases} \tag{6.5}$$

there exists a (non-necessarily finite) stopping time σ and a predictable $\mathbb{R}^r$-valued process $k \in L^2(Z)$ such that the process h given by (6.5) satisfies $h_\sigma = 0$, a.s.

Proof. We look for a solution of the control problem satisfying an equation of the type

$$\dot h_s = -\varphi_s \frac{1}{\|h_s\|} c_s(x)h_s \tag{6.6}$$

with $c_s(x)u = \sum_{i=1}^r (X_{s*}^{-1}A_i)_x\langle (X_{s*}^{-1}A_i)_x, u\rangle$, and where φ_s takes its values in $\{0, 1\}$.

Assuming that (C1) is satisfied, we construct a sequence of stopping times $(T_n)_{n\ge 0}$ and a continuous process h inductively as follows:

(i) $T_0 = 0$;
(ii) for $n \ge 0$, if $h_{T_{2n}} = 0$, then $T_{2n+2} = T_{2n+1} = T_{2n}$;
(iii) for $n \ge 0$, if $h_{T_{2n}} \ne 0$, h_t is constant on $[T_{2n}, T_{2n+1}]$ where

$$T_{2n+1} = \inf\left\{t > T_{2n}, \cos\left(c_t(x)h_{T_{2n}}, h_{T_{2n}}\right) > \alpha\right\},$$

and h_t solves

$$\dot h_s = -\frac{1}{\|h_s\|} c_s(x)h_s \quad \text{on } [T_{2n+1}, T_{2n+2}],$$

where $T_{2n+2} = \inf\{t > T_{2n+1}, \cos(c_t(x)h_t, h_t) < \alpha/2 \text{ or } h_t = 0\}$.

Let

$$\sigma = \inf\{t > 0, h_t = 0\} \quad (= \infty \text{ if this set is empty}),$$

and for $s < \sigma$,

$$\varphi_s = \mathbb{1}_{\bigcup_n [T_{2n+1}, T_{2n+2}[}(s),$$
$$k_s = -\varphi_s \frac{1}{\|h_s\|} \sum_{i=1}^{r} \langle (X_{s*}^{-1} A_i)_x, h_s \rangle e_i, \tag{6.7}$$

where $(e_1, \ldots, e_r)$ denotes the canonical basis of $\mathbb{R}^r$. Then h_t solves equation (6.6), $\dot{h}_s = (X_{s*}^{-1} A)_x k_s$, and since

$$\|k_s\|^2 = -\varphi_s \left\langle \dot{h}_s, \frac{h_s}{\|h_s\|} \right\rangle = -\frac{d}{ds}\|h_s\|,$$

we have

$$\int_0^\sigma \|k_s\|^2 \, ds \le \|h_0\|. \tag{6.8}$$

To conclude it is sufficient to prove that solutions h_t satisfy $\lim_{s\to\sigma} h_s = 0$.

First, we remark that h_t converges almost surely as t tends to σ. This is due to the fact that

$$\|dh\| = \frac{d\|h\|}{\cos(h, dh)} = -\frac{d\|h\|}{\cos(h, c_s(x)h)} \le -\frac{2}{\alpha} d\|h\|$$

(recall $d\|h\| \le 0$); hence h has a total variation bounded by $2\|h_0\|/\alpha$.

We define $u_t = h_0/\|h_0\|$ on the set where h_t converges to 0 as t tends to σ, and $u_t = h_t/\|h_t\|$ on the set where h_t does not converge to 0. This provides a process which converges as t tends to σ, but which is not adapted. On the set where h_t does not converge to 0, we have

$$\|h_0\| \ge -\int_0^\sigma d\|h\| \ge \int_0^\infty \langle c_s(x)u_s, u_s \rangle \mathbb{1}_{\{\cos(c_s(x)u_s, u_s) > \alpha\}} \, ds,$$

which implies, by Condition (C1), that this set has probability 0. □

EXAMPLE 6.6. Consider again Example 6.2, with $M = \mathbb{R}^3$,

$$A_0(x) = (0,0,0), \qquad A_1(x) = (1,0,0), \qquad A_2(x) = (0, 1, x^1).$$

For $u \in T_0 M$, $\|u\| = 1$, we have

$$\langle c_s(0)u, u \rangle = (u^1 - u^3 Z_s^2)^2 + (u^2 + u^3 Z_s^1)^2$$

and

$$\cos(c_s(0)u, u) = \frac{(u^1 - u^3 Z_s^2)^2 + (u^2 + u^3 Z_s^1)^2}{((u^1 - u^3 Z_s^2)^2 + (u^2 + u^3 Z_s^1)^2 + (-Z_s^2 u^1 + Z_s^1 u^2 + \|Z_s\|^2 u^3)^2)^{1/2}}.$$

From there it is straightforward to verify that Condition (C2) is realized in this case. With Proposition 6.5 we obtain Condition (C1), and with

Proposition 6.4 we get solvability of the control problem. We stress again that now we allow σ to be arbitrarily large. Then, contrary to the negative result of Example 6.2, we are able to find h such that $h_0 = v$, $h_\sigma = 0$, $\dot{h}_s = (k_s^1, k_s^2, Z_s^1 k_s^2 - Z_s^2 k_s^1)$, and $\int_0^\sigma |k_s|^2\,ds \in L^1$.

7. Derivative formulas in the hypoelliptic case

In this section, the results of the Sections 3 and 4 are extended to derive general differentiation formulas for heat semigroups and L-harmonic functions in the hypoelliptic case.

Let again $F(\cdot, X_\bullet(x))$, $x \in M$, be a family of local martingales where the transformation F is differentiable in the second variable with a derivative jointly continuous in both variables. We fix $x \in M$ and $v \in T_x M$. Let σ be a stopping time which is dominated by the first exit time of $X_\bullet(x)$ from some relatively compact neighbourhood of x. We first note that

$$dF(0,\cdot)_x v \equiv \mathbb{E}\big[dF(\sigma,\cdot)_{X_\sigma(x)} X_{\sigma *} v\big], \tag{7.1}$$

where $X_{\sigma *}$ is the derivative process at the random time σ. Equation (7.1) follows from the fact that the local martingale $F(\cdot, X_\bullet(x))$, differentiated in the direction v at x, is again a local martingale, and under the given assumptions a uniformly integrable martingale when stopped at σ. Our aim is to replace the right-hand side of (7.1) by expressions not involving derivatives of F. To this end the local martingales of Section 4 are exploited.

We start with an elementary construction. Let $D \subset M$ be a nonempty relatively compact domain and $\varphi \in C^2(\bar{D})$ such that $\varphi|\partial D = 0$ and $\varphi > 0$ on D. For $x \in D$ let

$$T(s) = \int_0^s \varphi^{-2}\big(X_r(x)\big)\,dr, \quad s \le \tau_D(x), \tag{7.2}$$

and

$$\sigma(r) = \inf\big\{s \ge 0 : T(s) \ge r\big\} \le \tau_D(x). \tag{7.3}$$

Note that $T(r) \to \infty$ as $r \nearrow \tau_D(x)$, almost surely, see [22]. Fix $t_0 > 0$ and consider

$$\ell_s = \frac{1}{t_0}\rho\bigg(\int_0^s \varphi^{-2}\big(X_r(x)\big)\,dr\bigg) v \tag{7.4}$$

for some $\rho \in C^1(\mathbb{R}_+, \mathbb{R})$ such that $\rho(s) = 0$ for s close to 0 and $\rho(s) = t_0$ for $s \ge t_0$. Then $\ell_0 = 0$ and $\ell_s = v$ for $s \ge \sigma(t_0)$.

Now for perturbations X^λ of X, as in Section 3, let

$$\ell_s^\lambda = \frac{1}{t_0}\rho\bigg(\int_0^s \varphi^{-2}\big(X_r^\lambda(x)\big)\,dr\bigg) v$$

and $\sigma^\lambda(r) = \inf\{s \ge 0 : T^\lambda(s) \ge r\}$. We introduce the abbreviation $\partial_\lambda = (\frac{\partial}{\partial\lambda_1}, \ldots, \frac{\partial}{\partial\lambda_n})$. Then $\partial_\lambda|_{\lambda=0}\ell_s^\lambda$ exists and lies in $\bigcap_{p>1} L^p$, see [22], Section 4 (the arguments there before Theorem 4.1 extend easily to general exponents p). In a similar way, using $T^\lambda \circ \sigma^\lambda = \mathrm{id}$, we see that

$$\partial_\lambda|_{\lambda=0}\sigma^\lambda = -\frac{1}{T' \circ \sigma}\big(\partial_\lambda|_{\lambda=0}T^\lambda\big) \circ \sigma.$$

For our applications, it is occasionally useful to modify the above construction such that already $\ell_s = v$ for $s \geq \sigma(t_0) \wedge t$ where $t > 0$ is fixed. This can easily be achieved by adding a term of the type $\tan(\pi r/2t)$ to the right-hand side of (7.2) and by changing the definition of ℓ_s in an obvious way.

Now let again $F(\cdot, X_\bullet(x))$ be a local martingale, as in Section 4, and consider the variation

$$F\big(\cdot, X_\bullet^\lambda(x)\big)G_\bullet^\lambda \tag{7.5}$$

of local martingales where

$$G_t^\lambda = \exp\biggl(-\int_0^t \langle a_s\lambda, dZ_s\rangle - \frac{1}{2}\int_0^t |a_s\lambda|^2\, ds\biggr). \tag{7.6}$$

Then

$$n_s = dF(s,\cdot)_{X_s(x)}X_{s*}\biggl(\int_0^s X_{r*}^{-1}A\big(X_r(x)\big)a_r\, dr\biggr) - F\big(s, X_s(x)\big)\int_0^s a_r^*\, dZ_r$$

is a local martingale in T_xM. Observe that n is the derivative of (7.5) at 0 with respect to λ, that is, $n_s = \partial_\lambda|_{\lambda=0}F(s, X_s^\lambda(x))G_s^\lambda$. In particular, taking

$$a_s = \big(X_{s*}^{-1}A\big)_x^*, \tag{7.7}$$

then

$$n_s = dF(s,\cdot)_{X_s(x)}X_{s*}C_s(x) - F\big(s, X_s(x)\big)\int_0^s \big(X_{r*}^{-1}A\big)_x\, dZ_r.$$

This implies that also

$$N_s := n_s h_s - \int_0^s n_r\, dh_r$$

is a local martingale for any T_x^*M-valued adapted process h locally of bounded variation. We choose $h_s = C_s(x)^{-1}\ell_s$ where ℓ is given by (7.4). Taking expectations gives

$$\begin{aligned} dF(0,\cdot)_x v &= \mathbb{E}\big[dF(\sigma,\cdot)_{X_\sigma(x)}X_{\sigma*}v\big] \\ &= \mathbb{E}\biggl[F\big(\sigma, X_\sigma(x)\big)\biggl(\int_0^\sigma \big(X_{s*}^{-1}A\big)_x\, dZ_s\biggr)C_\sigma^{-1}(x)v \\ &\qquad + \int_0^\sigma n_s\, dh_s\biggr], \end{aligned} \tag{7.8}$$

where $\sigma := \sigma(t_0)$. We deal separately with the term

$$\mathbb{E}\biggl[\int_0^\sigma n_s\, dh_s\biggr] = \mathbb{E}\biggl[\int_0^\sigma \partial_\lambda|_{\lambda=0}\big[F\big(s, X_s^\lambda(x)\big)G_s^\lambda\big]\, d\big(C_s(x)^{-1}\ell_s\big)\biggr]. \tag{7.9}$$

To avoid integrability problems, it may be necessary, as in proof of Theorem 3.2, to go through the calculation first with (7.7) replaced by

$$a_s^k = \big(X_{s*}^{-1}A\big)_x^* \mathbb{1}_{\{s\le\tau_k\}},$$

where (τ_k) is an appropriate increasing sequence of stopping times such that $\tau_k \nearrow \sigma$, and to take the limit as $k\to\infty$ in the final formula. Note that, without loss of generality, σ may be assumed to be bounded. We shall omit this technical modification here.

We return to the term (7.9). Observe that

$$\begin{aligned}&\mathbb{E}\bigg[\int_0^{\sigma^\lambda} F\big(s,X_s^\lambda(x)\big)G_s^\lambda\, d\big(C_s^\lambda(x)^{-1}\ell_s^\lambda\big)\bigg]\\ &\quad\equiv \int_0^\infty \mathbb{E}\bigg[\mathbb{1}_{\{s\le\sigma^\lambda\}}F\big(s,X_s^\lambda(x)\big)G_s^\lambda \frac{d}{ds}\big(C_s^\lambda(x)^{-1}\ell_s^\lambda\big)\bigg]\,ds\end{aligned}$$

is independent of λ. Thus, differentiating with respect to λ at $\lambda=0$ gives

$$\begin{aligned}&\mathbb{E}\bigg[\int_0^\sigma n_s\,dh_s\bigg]\\ &\quad= -\mathbb{E}\bigg[\int_0^\sigma F_s\, d\big[\partial_\lambda|_{\lambda=0}\big(C_s^\lambda(x)^{-1}\ell_s^\lambda\big)\big] + \partial_\lambda|_{\lambda=0}\int_0^{\sigma^\lambda} F_s\, d\big(C_s(x)^{-1}\ell_s\big)\bigg]\\ &\quad= -\mathbb{E}\bigg[F_\sigma\big[\partial_\lambda|_{\lambda=0}\big(C_s^\lambda(x)^{-1}\ell_s^\lambda\big)\big]_{s=\sigma} + F_\sigma\bigg(\frac{d}{ds}\bigg|_{s=\sigma} C_s(x)^{-1}\ell_s\bigg)\big(\partial_\lambda|_{\lambda=0}\sigma^\lambda\big)\bigg],\end{aligned}$$

where $F_s\equiv F(s,X_s(x))$. Note that all terms in the last line are nicely integrable. Substituting this back into equation (7.8), we find a formula of the wanted type:

$$dF(0,\cdot)_x v = \mathbb{E}\big[F\big(\sigma,X_\sigma(x)\big)\Phi_\sigma v\big], \tag{7.10}$$

where Φ_σ takes values in T_x^*M and is L^p-integrable for any $1\le p<\infty$. Summarizing the above discussion, we conclude with the following two theorems.

THEOREM 7.1. *Let M be a smooth manifold and $f: M\to\mathbb{R}$ a bounded measurable function. Assume that* (H1) *holds. Let $x\in M$, $v\in T_xM$, $t>0$. Then*

$$d(P_tf)_x v = \mathbb{E}\big[f\big(X_t(x)\big)\mathbb{1}_{\{t<\zeta(x)\}}\Phi_t v\big] \tag{7.11}$$

for the minimal semigroup P_tf defined by (1.3) *where Φ_t is a T_x^*M-valued random variable which is L^p-integrable for any $1\le p<\infty$ and local in the following sense: For any relatively compact neighbourhood D of x in M there is a choice for Φ_t which is $\mathscr{F}_\sigma$-measurable where $\sigma=t\wedge\tau_D(x)$ and $\tau_D(x)$ is the first exit time of X from D when starting at x.*

Proof. Let $F(\cdot,X_\cdot(x)) = (P_{t-\cdot}f)(X_\cdot(x))$. Then Equation (7.10) gives

$$d(P_tf)_x v = \mathbb{E}\big[F\big(\sigma,X_\sigma(x)\big)\Phi_\sigma\big]$$

Again by taking into account that

$$(P_{t-\sigma}f)\big(X_\sigma(x)\big) = \mathbb{E}^{\mathscr{F}_\sigma}\big[f\big(X_t(x)\big)\mathbb{1}_{\{t<\zeta(x)\}}\big],$$

we get the claimed formula. □

THEOREM 7.2. *Let M be compact with smooth boundary $\partial M \neq \varnothing$ and $u \in C(M)$ be L-harmonic on $M\backslash\partial M$. Assume that* (H1) *holds. Let $x \in M\backslash\partial M$ and $v \in T_xM$. Denote $\tau(x)$ the first hitting time of $X_\bullet(x)$ at ∂M. Then the following formula holds*:

$$(du)_x v = \mathbb{E}\big[u\big(X_{\tau(x)}(x)\big)\Phi_{\tau(x)}v\big], \tag{7.12}$$

where $\Phi_{\tau(x)}$ is a T_x^*M-valued random variable which is in L^p for any $1 \le p < \infty$ and local in the following sense: *For any relatively compact neighbourhood D of x in M, there is a choice for $\Phi_{\tau(x)}$ which is already $\mathscr{F}_\sigma$-measurable where $\sigma = \tau_D(x)$ is the first exit time of X from D when starting at x.*

Proof. The proof is completely analogous to the proof of Theorem 7.1. □

EXAMPLE 7.3 (Greek Deltas for Asian Options). Consider the following SDE on the real line:

$$dS_t = \sigma(S_t)\,dW_t + \mu(S_t)\,dt, \tag{7.13}$$

where W_t is a real Brownian motion. In Mathematical Finance, one likes to calculate so-called *Greek Deltas for Asian Options* which are expressions of the form

$$\Delta_0 = \frac{\partial}{\partial S_0}\mathbb{E}\big[f(S_T, A_T)\big], \quad T > 0,$$

where S_t is given as solution to (7.13) and

$$A_t = \int_0^t S_r\,dr. \tag{7.14}$$

We may convert (7.13) to Stratonovich form

$$dS_t = \sigma(S_t)\delta W_t + m(S_t)\,dt$$

and consider $X_t := (S_t, A_t)$ as a diffusion on $\mathbb{R}^2$. Then

$$d\begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} = \begin{pmatrix} \sigma\big(X_t^1\big) \\ 0 \end{pmatrix} \circ dW_t + \begin{pmatrix} m\big(X_t^1\big) \\ X_t^1 \end{pmatrix} dt$$

with the vector fields

$$A_0 = \begin{pmatrix} m(x_1) \\ x_1 \end{pmatrix}, \qquad A_1 = \begin{pmatrix} \sigma(x_1) \\ 0 \end{pmatrix}.$$

Observe that

$$[A_1, A_0] = \begin{pmatrix} \sigma(x_1)m'(x_1) - \sigma'(x_1)m(x_1) \\ \sigma(x_1) \end{pmatrix}.$$

Thus if $\sigma > 0$, then $X_t = (S_t, A_t)$ defines a hypoelliptic diffusion on $\mathbb{R}^2$.

EXAMPLE 7.4 (Trivial example). In the special case $\sigma > 0$ constant and $\mu = 0$, that is,

$$\begin{cases} dS_t = \sigma\, dW_t, \\ dA_t = S_t\, dt, \end{cases}$$

one easily checks

$$X_{t*} = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix} \quad \text{and} \quad X_{t*}^{-1}(A_1) \otimes X_{t*}^{-1}(A_1) = \sigma^2 \begin{pmatrix} 1 & -t \\ -t & t^2 \end{pmatrix},$$

and hence

$$C_T(x) = \sigma^2 \begin{pmatrix} T & -T^2/2 \\ -T^2/2 & T^3/3 \end{pmatrix}.$$

Consequently, the integration by parts argument of Section 3 immediately gives

$$\frac{\partial}{\partial S_0}\mathbb{E}\big[f(S_T, A_T)\big] = \frac{6}{\sigma T}\mathbb{E}\left[f(S_T, A_T)\left(\frac{1}{T}\int_0^T W_t\, dt - \frac{1}{3}W_T\right)\right].$$

REMARK 7.5. In the more general situation of Example 7.3, that is,

$$dS_t = \sigma(S_t)\, dW_t + \mu(S_t)\, dt \quad \text{and} \quad A_t = \int_0^t S_r\, dr,$$

Theorem 7.1 may be applied to give a formula of the type

$$\Delta_0 = \frac{\partial}{\partial S_0}\mathbb{E}\big[f(S_T, A_T)\big] = \mathbb{E}\big[f(S_T, A_T)\pi_T\big],$$

where the weight π_T is explicitly given and may be implemented numerically in Monte-Carlo simulations; compare with [12]. See [8] for extensions to jump diffusions, and [13] for weights π_T in terms of anticipating integrals.

8. The case of non-Euclidean targets

The aim of this section is to adapt our method, to some extent, to the nonlinear case of harmonic maps between manifolds. In addition to the manifold M, carrying a hypoelliptic L-diffusion, we fix another manifold N, endowed with a torsion free connection ∇. In stochastic terms, a smooth map $u: M \to N$ is harmonic (with respect to L) if it takes L-diffusions on M to ∇-martingales on N. Likewise, a smooth map $u: [0,t] \times M \to N$ is said to solve the nonlinear heat equation, if $u(t - \boldsymbol{\cdot}, X_{\boldsymbol{\cdot}}(x))$ is a ∇-martingale on N for any L-diffusion $X_{\boldsymbol{\cdot}}(x)$ on M.

Henceforth, we fix a family $F(\boldsymbol{\cdot}, X_{\boldsymbol{\cdot}}(x))$, $x \in M$ of ∇-martingales on N where F is differentiable in the second variable with a derivative jointly continuous in both variables. In particular, such transformations F map hypoelliptic L-diffusions on M into ∇-martingales on N and include the following two cases:

(1) $F(\boldsymbol{\cdot}, X_{\boldsymbol{\cdot}}(x)) = u \circ X_{\boldsymbol{\cdot}}(x)$ for some harmonic map $u: M \to N$, and

(2) $F(\cdot, X_\cdot(x)) = u(t - \cdot, X_\cdot(x))$ where u solves the heat equation for maps $M \to N$.

Theorem 4.1 is easily extended to this situation. Recall that, if Y is a continuous semimartingale taking values in a manifold N endowed with a torsionfree connection ∇, then the geodesic (damped or deformed) transport $\Theta_{0,t} : T_{Y_0}N \to T_{Y_t}N$ on N along Y is defined by the following covariant equation along Y:

$$\begin{cases} d\big(//_{0,\cdot}^{-1}\Theta_{0,\cdot}\big) = -\frac{1}{2}//_{0,\cdot}^{-1}R\big(\Theta_{0,\cdot}, dY\big)\, dY, \\ \Theta_{0,0} = \mathrm{id}, \end{cases} \tag{8.1}$$

where $//_{0,t} : T_{Y_0}N \to T_{Y_t}N$ is parallel translation on N along Y and R the curvature tensor to ∇, see [2]. Finally, recall the notion of anti-development of Y, resp. "deformed anti-development" of Y,

$$\mathscr{A}(Y) = \int_0^\cdot //_{0,s}^{-1}\delta Y_s, \qquad \mathscr{A}_{\mathrm{def}}(Y) = \int_0^\cdot \Theta_{0,s}^{-1}\delta Y_s \tag{8.2}$$

which by definition both take values in $T_{Y_0}N$. Note that an N-valued semimartingale is a ∇-martingale if and only if $\mathscr{A}(Y)$, or equivalently $\mathscr{A}_{\mathrm{def}}(Y)$, is a local martingale.

THEOREM 8.1. *Let $F(\cdot, X_\cdot(x))$, $x \in M$, be a family of ∇-martingales on N, as described above. Then, for any predictable $\mathbb{R}^r$-valued process k in $L^2_{\mathrm{loc}}(Z)$,*

$$\begin{aligned} &\Theta_{0,\cdot}^{-1}\, dF\big(\cdot, X_\cdot(x)\big)(T_xX_\cdot)\int_0^\cdot \big(X_{s*}^{-1}A\big)_x k_s\, ds \\ &\qquad - \mathscr{A}_{\mathrm{def}}\big(F\big(\cdot, X_\cdot(x)\big)\big)\int_0^\cdot \langle k, dZ\rangle \end{aligned} \tag{8.3}$$

is a local martingale in $T_{F(0,x)}N$. Here $\Theta_{0,\cdot}$ denotes the geodesic transport on N along the martingale $F(\cdot, X_\cdot(x))$.

Proof. Observe that by [2],

$$m_s := \Theta_{0,s}^{-1}\, dF(s,\cdot)_{X_s(x)}X_{s*}$$

is local martingale taking values in $T_xM \otimes T_{F(0,x)}N$, and that by definition,

$$\mathscr{A}_{\mathrm{def}}\big(F\big(\cdot, X_\cdot(x)\big)\big) = \int_0^\cdot \Theta_{0,s}^{-1}\, dF(s,\cdot)_{X_s(x)}A\big(X_s(x)\big)dZ_s.$$

The rest of the (alternative) proof to Theorem 4.1 carries over with straightforward modifications. $\square$

It is straightforward to extend Theorem 5.5 and Theorem 5.6 to the nonlinear setting by means of the local martingale (8.3).

THEOREM 8.2. *Let $u:[0,t]\times M\to N$ be a solution of the nonlinear heat equation, $x\in M$, $v\in T_xM$. Let D be a relatively compact open neighbourhood of x and $\sigma=\tau_D(x)\wedge t$ where $\tau_D(x)$ is the first exit time of $X_\bullet(x)$ from D. Suppose there exists an $\mathbb{R}^r$-valued predictable process k such that*

$$\int_0^\sigma \big(X_{s*}^{-1}A\big)_x k_s\,ds\equiv v,\qquad a.s.$$

and $(\int_0^\sigma |k_s|^2\,ds)^{1/2}\in L^{1+\varepsilon}$ for some $\varepsilon>0$. Then the following formula holds:

$$du(t,\cdot)_x v=\mathbb{E}\bigg[\mathscr{A}_{\rm def}\big(u\big(t-\cdot,X_\bullet(a)\big)\big)_\sigma\int_0^\sigma\langle k,dZ\rangle\bigg].\tag{8.4}$$

THEOREM 8.3. *Let M be compact with smooth boundary $\partial M\neq\varnothing$. For $x\in M\backslash\partial M$, let $\tau(x)$ be the first hitting time of ∂M with respect to the process $X_\bullet(x)$. Given $v\in T_xM$, we suppose that there exists an $\mathbb{R}^r$-valued predictable process k such that*

$$\int_0^{\tau(x)} \big(X_{s*}^{-1}A\big)_x k_s\,ds\equiv v,\qquad a.s.$$

and $(\int_0^{\tau(x)} |k_s|^2\,ds)^{1/2}\in L^{1+\varepsilon}$ for some $\varepsilon>0$. Then, for any $u\in C^\infty(M,N)$ which is harmonic on $M\backslash\partial M$, the following formula holds:

$$(du)_x v=\mathbb{E}\bigg[\mathscr{A}_{\rm def}\big(u\big(X_\bullet(x)\big)\big)_{\tau(x)}\int_0^{\tau(x)}\langle k,dZ\rangle\bigg].\tag{8.5}$$

Note that if a is a predictable process taking values in $T_xM\otimes(\mathbb{R}^r)^*$, as in Section 4, then

$$\begin{aligned}\Theta_{0,\bullet}^{-1}\,dF\big(\cdot,X_\bullet(x)\big)(T_xX_\bullet)&\int_0^\bullet \big(X_{s*}^{-1}A\big)_x a_s\,ds\\ &-\mathscr{A}_{\rm def}\big(F\big(\cdot,X_\bullet(x)\big)\big)\int_0^\bullet a_r^*\,dZ_r\end{aligned}\tag{8.6}$$

gives a local martingale in $T_xM\otimes T_{F(0,x)}N$. In particular, setting

$$a_s=\big(X_{s*}^{-1}A\big)_x^*\mathbb{1}_{\{s\le\tau\}},\tag{8.7}$$

where τ may be any predictable stopping time, we see that

$$\begin{aligned}n_s=\Theta_{0,s}^{-1}\,dF(s,\cdot)_{X_s(x)}X_{s*}C_{s\wedge\tau}(x)&\\ -\mathscr{A}_{\rm def}\big(F\big(\cdot,X_\bullet(x)\big)\big)_s&\int_0^{s\wedge\tau}\big(X_{r*}^{-1}A\big)_x\,dZ_r\end{aligned}\tag{8.8}$$

is a local martingale. Let

$$Y=\mathscr{A}_{\rm def}\big(F\big(\cdot,X_\bullet(x)\big)\big)\quad\text{and}\quad Y^\lambda=\mathscr{A}_{\rm def}\big(F\big(\cdot,X_\bullet^\lambda(x)\big)\big)\tag{8.9}$$

for variations $X^\lambda(x)$ of $X(x)$, as in Section 3, and recall that, again with the choice (8.7),

$$J_s=\partial_\lambda|_{\lambda=0}F\big(s,X_s^\lambda(x)\big)=dF(s,\cdot)_{X_s(x)}X_{s*}C_{s\wedge\tau}(x).\tag{8.10}$$

By definition, $J_\bullet w$ is a vector field on N along the martingale $F(\cdot, X_\bullet(x))$ for each $w \in T_x^* M$. Imitating the strategy of Section 7, the idea is to differentiate $Y_\bullet^\lambda G_\bullet^\lambda$ with respect to λ.

LEMMA 8.4. *Keeping the notations as above, we have*

$$\operatorname{vert}\bigl[\partial_\lambda|_{\lambda=0} Y^\lambda\bigr] = \Theta_{0,\bullet}^{-1} J - J_0 + \int_0^\bullet \Theta_{0,s}^{-1}(\nabla\Theta_{0,s})\, dY_s, \tag{8.11}$$

where $\nabla\Theta_{0,\bullet} : T_{F(0,x)}N \to T_{F(\bullet, X_\bullet(x))}N$ *is defined by*

$$(\nabla\Theta_{0,\cdot})u = v_J^{-1}\bigl(\bigl(\Theta_{0,\bullet}^c h_{J_0}(u)\bigr)^{\mathrm{vert}}\bigr). \tag{8.12}$$

In particular, $\operatorname{vert}[\partial_\lambda|_{\lambda=0} Y^\lambda]$ *and* $\Theta_{0,\bullet}^{-1} J - J_0$ *differ only by a local martingale. Here* $\Theta_{0,\bullet}^c$ *denotes the geodesic transport on* TN *along* J *with respect to the complete lift* ∇^c *of the connection* ∇.

We are not going to prove Lemma 8.4 here. We just remark that, again with the choice (8.7) for the process a, we end up with the following local martingale:

$$\begin{aligned} m &:= \operatorname{vert}\bigl[\partial_\lambda|_{\lambda=0}\bigl(Y^\lambda G^\lambda\bigr)\bigr] \\ &= \Theta_{0,\bullet}^{-1} J - J_0 + \int_0^\bullet \Theta_{0,s}^{-1} \nabla\Theta_{0,s}\, dY_s - Y \int_0^{\bullet\wedge\tau} \bigl(X_{s*}^{-1} A\bigr)_x\, dZ_s. \end{aligned} \tag{8.13}$$

Then a procedure along the lines of Section 7 leads to a formula for $dF(0,\cdot)_x v$ which is analogous to the linear case, but with an additional term of the type

$$\mathbb{E}\biggl[\biggl(\int_0^\sigma \Theta_{0,s}^{-1} \nabla_{J_s}\Theta_{0,s}\, dY_s\biggr) C_\sigma^{-1}(x) v\biggr] \tag{8.14}$$

for some stopping time σ. At the moment, it seems unclear whether it is possible to avoid this extra term.

9. Concluding remarks

1. The presented differentiation formulas are not intrinsic: they involve the derivative flow which depends on the particular SDE and not just on the generator. It is possible to make the formulas more intrinsic by using the framework of Elworthy, Le Jan and Li [10], [11] on geometry of SDEs (e.g., filtering out redundant noise and working with connections induced by the SDE).

2. In this paper, we exploited perturbations of the driving Brownian motion and a change of measure as method for constituting variational formulas. There are of course other ways of performing perturbations leading to local martingales which are related to integration by parts formulas. For instance, one observes that the local martingale property of $F(\cdot, X_\bullet(x))$ is preserved under

(i) a change of measure via Girsanov's theorem,

(ii) a change of time,
(iii) rotations of the Brownian motion Z.

In particular, (iii) seems to be promising in the hypoelliptic context since it leads to contributions in the direction of the bracket $[A_i, A_j]$. So far however, it is unclear to us how to relate such variations to regularity results under hypoellipticity conditions.

References

[1] M. Arnaudon and A. Thalmaier, *Stability of stochastic differential equations in manifolds*, Séminaire de Probabilités, XXXII, Lecture Notes in Math, vol. 1686, Springer, Berlin, 1998, pp. 188–214. MR 1655151

[2] M. Arnaudon and A. Thalmaier, *Complete lifts of connections and stochastic Jacobi fields*, J. Math. Pures Appl. **77** (1998), 283–315. MR 1618537

[3] D. R. Bell, *The Malliavin calculus*, Pitman Monographs and Surveys in Pure and Applied Mathematics, vol. 34, Longman Scientific & Technical, Harlow, 1987. MR 0902583

[4] D. R. Bell, *Divergence theorems in path space. III. Hypoelliptic diffusions and beyond*, J. Funct. Anal. **251** (2007), 232–253. MR 2353706

[5] D. R. Bell and S.-E. A. Mohammed, *An extension of Hörmander's theorem for infinitely degenerate second-order operators*, Duke Math. J. **78** (1995), 453–475. MR 1334203

[6] G. Ben Arous, S. Kusuoka and D. Stroock, *The Poisson kernel for certain degenerate elliptic operators*, J. Funct. Anal. **56** (1984), 171–209. MR 0738578

[7] J.-M. Bismut, *Martingales, the Malliavin calculus and hypoellipticity under general Hörmander's conditions*, Z. Wahrsch. Verw. Gebiete **56** (1981), 469–505. MR 0621660

[8] T. Cass and P. Friz, *The Bismut–Elworthy–Li formula for jump-diffusions and applications to Monte Carlo pricing in finance*, available at: arXiv:math/0604311.

[9] B. K. Driver and A. Thalmaier, *Heat equation derivative formulas for vector bundles*, J. Funct. Anal. **183** (2001), 42–108. MR 1837533

[10] K. D. Elworthy, Y. Le Jan and X.-M. Li, *Concerning the geometry of stochastic differential equations and stochastic flows*, New trends in stochastic analysis (Charingworth, 1994), World Sci. Publ., River Edge, NJ, 1997, pp. 107–131. MR 1654511

[11] K. D. Elworthy, Y. Le Jan and X.-M. Li, *On the geometry of diffusion operators and stochastic flows*, Lecture Notes in Mathematics, vol. 1720, Springer-Verlag, Berlin, 1999. MR 1735806

[12] E. Fournié, J.-M. Lasry, J. Lebuchoux, P.-L. Lions and N. Touzi, *Applications of Malliavin calculus to Monte Carlo methods in finance*, Finance Stoch. **3** (1999), 391–412. MR 1842285

[13] B. Forster, E. Lütkebohmert and J. Teichmann, *Absolutely continuous laws of jump-diffusions in finite and infinite dimensions with applications to mathematical finance*, SIAM J. Math. Anal. **40** (2008/09), 2132–2153. MR 2471915

[14] H. Kunita, *On the decomposition of solutions of stochastic differential equations*, Stochastic Integrals, Proc. Sympos., Durham, 1980. Lecture Notes in Math, vol. 851 (D. Williams, ed.), 1981, pp. 213–255. MR 0620992

[15] S. Kusuoka, *The Malliavin calculus and its applications*, Sugaku Expositions **3** (1990), 127–144. MR 0780358

[16] N. V. Krylov, *Quasiderivatives for solutions of Itô's stochastic equations and their applications*, Stochastic analysis and related topics (Oslo, 1992), Stochastics Monographs, vol. 8, Gordon and Breach, Montreux, 1993, pp. 1–44. MR 1268004

[17] N. V. Krylov, *Probabilistic methods of investigating interior smoothness of harmonic functions associated with degenerate elliptic operators*, Publications of the Ennio de Giorgi Mathematical Research Center, Scuola Normale Superiore, Pisa, 2004. MR 2144644

[18] J. Norris, *Simplified Malliavin calculus*, Séminaire de Probabilités, XX, Lecture Notes in Math., vol. 1204, Springer, Berlin, 1986, pp. 101–130. MR 0942019

[19] D. Nualart, *The Malliavin calculus and related topics*, Springer, New York, 1995. MR 1344217

[20] A. Thalmaier, *On the differentiation of heat semigroups and Poisson integrals*, Stochastics and Stochastics Reports **61** (1997), 297–321. MR 1488139

[21] A. Thalmaier, *Some remarks on the heat flow for functions and forms*, Electron. Comm. Probab. **3** (1998), 43–49. MR 1637977

[22] A. Thalmaier and F.-Y. Wang, *Gradient estimates for harmonic functions on regular domains in Riemannian manifolds*, J. Funct. Anal. **155** (1998), 109–124. MR 1622800

Marc Arnaudon, Laboratoire de Mathématiques et Applications, CNRS: UMR6086, Université de Poitiers, Téléport 2–BP 30179, F-86962 Futuroscope Chasseneuil, France

E-mail address: marc.arnaudon@math.univ-poitiers.fr

Anton Thalmaier, Unité de Recherche en Mathématiques, FSTC, Université du Luxembourg, 6, rue Richard Coudenhove-Kalergi, L-1359 Luxembourg, Grand-Duchy of Luxembourg

E-mail address: anton.thalmaier@uni.lu