

# 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

$$(1.1) \quad \delta X = A(X) \delta Z + A_0(X) dt$$

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  $\mathbb{R}^r \otimes TM$ .

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

$$(1.2) \quad L = A_0 + \frac{1}{2} \sum_{i=1}^r A_i^2$$

where  $A_i = A(\cdot)e_i \in \Gamma(TM)$  and  $e_i$  the  $i$ th 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  $X_t(x)$ ,  $0 \leq t < \zeta(x)$ , is the maximal strong solution to (1.1) with

---

*Date:* March 26, 2011.

*1991 Mathematics Subject Classification.* Primary 58G32, 60H30; Secondary 60H10.

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  $t \geq 0$ , i.e.  $\zeta(\cdot, \omega)$  is lower semicontinuous on  $M$ .
- (ii)  $X_t(\cdot, \omega): M_t(\omega) \rightarrow M$  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  $C^\infty$ -topology.

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

Let

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

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

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

$$(H1) \quad \text{Lie}(A_i, [A_0, A_i]: i = 1, \dots, r)(x) = T_x M \quad \text{for all } x \in M.$$

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

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

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

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

see [8] 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  $u: D \rightarrow \mathbb{R}$ , given on some domain  $D$  in  $M$  by its boundary values  $u|_{\partial D}$  via

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

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 [5] 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 Section 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 [22, 24]. 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 [6] 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,

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

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,

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

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:*

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

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

*Proof.* See Kunita [16], section 5. □

We have the famous “invertibility of the Malliavin covariance matrix” under the Hörmander condition (H1), e.g., see Bismut [8], Prop. 4.1.

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

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

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

Almost surely, for each  $t > 0$ ,

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

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, e.g., [20, 4, 17, 21].

**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  $1 \leq p < \infty$ . In the same way,

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

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  $X_{\cdot}(x)$  from some relatively compact open neighbourhood  $D \neq M$  of  $x$ . Also note that  $\tau_D(x) \in L^p$  for all  $1 \leq p < \infty$ , e.g. [7], Lemma (1.21).

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

$$(2.7) \quad (X_{s*}^{-1}A)_x: \mathbb{R}^r \rightarrow T_x M, \quad z \mapsto \sum_{i=1}^r (X_{s*}^{-1}A_i)_x z^i.$$

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

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

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 [8].

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  $t > 0$ ,

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

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  $G_t^\lambda$  defined by

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

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

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

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

*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

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

In particular, we have

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

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

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

This gives the claim by definition of the bracket.  $\square$

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

$$(3.4) \quad d(P_t f)_x v = \mathbb{E} [(f \circ X_t(x)) \Phi_t v]$$

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

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

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

is independent of  $\lambda$  for any  $C^1$ -function  $f$  on  $M$ . Thus

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

which gives

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

Note that

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

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

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

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

$$a_s \equiv a_s^n := (X_{s^*}^{-1}A)_x 1_{\{s \leq \tau_n\}} \in T_x M \otimes (\mathbb{R}^r)^*$$

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

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

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

$$(3.6) \quad d(P_t f)_x v = \mathbb{E}[(df)_{X_t(x)} X_{t^*} v] = \mathbb{E}[(f \circ X_t(x)) \cdot \Phi_t v]$$

where

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

which can be further evaluated by means of (3.3). Eq. (3.3) also allows to conclude that  $\Phi_t \in \cap_{p \geq 1} L^p$ .  $\square$

#### 4. INTEGRATION BY PARTS AT THE LEVEL OF LOCAL MARTINGALES

Let  $F(\cdot, X.(x))$ ,  $x \in M$  be a family of local martingales where  $F$  is differentiable in the second variable with a derivative jointly continuous in both variables. Here  $X.(x) \equiv (X_t(x))_{t \geq 0}$  denotes the solution to the SDE (1.1) starting from  $x$ . We are mainly interested in the following two cases:

- (1)  $F(\cdot, X.(x)) = u \circ X.(x)$  for some  $L$ -harmonic function  $u$  on  $M$ , and
- (2)  $F(\cdot, X.(x)) = (P_{t-\cdot} f)(X.(x))$  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  $F(\cdot, X.(x))$ ,  $x \in M$  be a family of local martingales as described above. Then, for any predictable  $\mathbb{R}^r$ -valued process  $k$  in  $L_{\text{loc}}^2(Z)$ ,*

$$(4.1) \quad dF(\cdot, X.(x))(T_x X.) \int_0^\cdot (X_{s^*}^{-1}A)_x k_s ds - F(\cdot, X.(x)) \int_0^\cdot \langle k, dZ \rangle, \quad x \in M,$$

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

*Proof (by means of Girsanov).* For  $\varepsilon$  varying locally about 0, consider the SDE

$$(4.2) \quad \delta X^\varepsilon = A(X^\varepsilon) \delta Z^\varepsilon + A_0(X^\varepsilon) dt$$

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

$$(4.3) \quad F(\cdot, X^\varepsilon(x)) G^\varepsilon$$

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

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

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

$$\frac{\partial}{\partial \varepsilon} \Big|_{\varepsilon=0} F(\cdot, X^\varepsilon(x)) G^\varepsilon = \frac{\partial}{\partial \varepsilon} \Big|_{\varepsilon=0} F(\cdot, X^\varepsilon(x)) + F(\cdot, X(x)) \frac{\partial}{\partial \varepsilon} \Big|_{\varepsilon=0} G^\varepsilon$$

is also a local martingale. Taking into account that

$$\frac{\partial}{\partial \varepsilon} \Big|_{\varepsilon=0} X_r^\varepsilon(x) = X_{r*} \int_0^r X_{s*}^{-1} A(X_s(x)) k_s ds$$

and

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

we get the claim.  $\square$

*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  $T_x^*M$ , see [2]. Thus also

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

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

$$h = \int_0^\cdot (X_{s*}^{-1} A)_x k_s ds$$

and taking into account that

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

the claim follows by noting that

$$\int_0^\cdot dF(s, \cdot)_{X_s(x)} X_{s*} dh_s \stackrel{\text{m}}{=} F(\cdot, X(x)) \int_0^\cdot \langle k, dZ \rangle$$

where  $\stackrel{\text{m}}{=}$  denotes equality modulo local martingales.  $\square$

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

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

is a local martingale in  $T_x M$  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

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

For the last term it is trivial to observe that

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

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

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

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  $f: M \rightarrow \mathbb{R}$  be bounded measurable. Fix  $x \in M$  and  $v \in T_x M$ . Then, for any predictable  $\mathbb{R}^r$ -valued process  $k$  in  $L_{\text{loc}}^2(Z)$ ,*

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

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

Note that  $(dP_{t-} f)_{X_\bullet(x)} (T_x X_\bullet) v$  is a local martingale as the derivative of the local martingale  $(P_{t-} f)(X_\bullet(x))$  at  $x$  in the direction  $v$ , see [2].

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

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

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

**Problem 5.3** (Control Problem). Let  $x \in M$  and  $v \in T_x M$ . Consider the random dynamical system

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

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  $X_\cdot(x)$  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, i.e.,  $A(x): \mathbb{R}^r \rightarrow T_x M$  surjective for each  $x \in M$ . Then

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

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  $A(\cdot)$ , i.e.  $A(x)A^*(x) = \text{id}_{T_x M}$  for  $x \in M$ , the process  $h$  may be any adapted process with values in  $T_x M$  and with absolutely continuous sample paths (e.g., paths in the Cameron-Martin space  $\mathbb{H}(\mathbb{R}_+, T_x M)$ ) 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  $\sigma$ , 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

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

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  $f: M \rightarrow \mathbb{R}$  be bounded measurable,  $x \in M$ ,  $v \in T_x M$ ,  $t > 0$ . 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  $X_\cdot(x)$  from  $D$ . Suppose there exists an  $\mathbb{R}^r$ -valued predictable process  $k$  such that*

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

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

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

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  $(P_{t-\sigma} f)(X_\sigma(x)) = \mathbb{E}^{\mathcal{F}_\sigma} [f(X_t(x)) 1_{\{t < \zeta(x)\}}]$ . See Theorem 2.4 in [22] for technical details.  $\square$

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

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

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

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

In the elliptic case, formulas of type (5.2) and (5.3) have been used in [24] to establish gradient estimates for  $P_t f$  and for harmonic functions  $u$ , see also [10] for extensions from functions to sections. Nonlinear generalizations of the elliptic case, e.g., to harmonic maps and solutions of the nonlinear heat equations, are treated in [3].

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  $h, \Phi$  with  $h_0 = v$  and  $\Phi_0 = 0$  where  $h$  is  $T_x M$ -valued and  $\Phi$  real-valued such that

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

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 [18, 19], this means that  $\xi := (T_x X) h$  is a  $F$ -quasiderivative for  $X$  along  $\xi$  at  $x$  and  $\Phi$  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  $\Phi$  itself should already be a local martingale, say  $\Phi_s = \int_0^s \langle k_r, dZ_r \rangle$ . Thus

$$n \stackrel{\text{m}}{=} \int_0^\cdot (dF_r)_{X_r(x)} X_{r*} dh_r + \int_0^\cdot (dF_r)_{X_r(x)} A(X_r(x)) k_r dr$$

which implies

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

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

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

or

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

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

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

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

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

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

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

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

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

Consequently

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

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

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

and  $\left( \int_0^\sigma |k_s|^2 ds \right)^{1/2} \in L^{1+\varepsilon}$ . 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

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

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

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

(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

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

where  $X_s = X_s(0)$  and  $\varphi$  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  $Z_\sigma^1 \neq 0$  with probability 1, we may conclude that also

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

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),

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

which shows in particular that

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

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

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

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

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

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

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

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

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

In particular,

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

Thus (5.1) is given by

$$\dot{h}_s = (k_s^1, k_s^2, Z_s^1 k_s^2 - Z_s^2 k_s^1)$$

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  $\int_0^\sigma Z_s^1 k_s^2 ds - \int_0^\sigma Z_s^2 k_s^1 ds = -v^3$ . On the other hand, an integration by parts yields

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

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  $\sigma$  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

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

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:

$$\langle c_s(0)u, u \rangle = (u^1 - u^3 Z_s^2)^2 + (u^2 + u^3 Z_s^1)^2, \quad u \in T_0 M.$$

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  $\sigma$  (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  $\sigma$  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  $\alpha$  such that for any continuous (non necessarily adapted) process  $u_t$ , taking values in  $\{w \in T_x M, \|w\| = 1\}$  and converging to  $u$  almost surely,

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

**Condition (C2).** There exists a positive constant  $\alpha$  such that for any  $u_0 \in \{w \in T_x M, \|w\| = 1\}$ , there exists a neighbourhood  $V_{u_0}$  of  $u_0$  in  $\{w \in T_x M, \|w\| = 1\}$ , such that

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

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*

$$(6.5) \quad \begin{cases} \dot{h}_s = (X_{s^*}^{-1}A)_x k_s \\ h_0 = v. \end{cases}$$

*there exists a (non necessarily finite) stopping time  $\sigma$  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

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

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  $\varphi_s$  takes its values in  $\{0, 1\}$ .

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

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

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

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  $s < \sigma$ ,

$$(6.7) \quad \begin{aligned} \varphi_s &= \mathbf{1}_{\cup_n [T_{2n+1}, T_{2n+2}]}(s), \\ k_s &= -\varphi_s \frac{1}{\|h_s\|} \sum_{i=1}^r \langle (X_{s^*}^{-1} A_i)_x, h_s \rangle e_i, \end{aligned}$$

where  $(e_1, \dots, e_r)$  denotes the canonical basis of  $\mathbb{R}^r$ . Then  $h_t$  solves Eq. (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

$$(6.8) \quad \int_0^\sigma \|k_s\|^2 ds \leq \|h_0\|.$$

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

First we remark that  $h_t$  converges almost surely as  $t$  tends to  $\sigma$ . This is due to the fact that

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

(recall  $d\|h\| \leq 0$ ); 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  $\sigma$ , 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  $\sigma$ , but which is not adapted. On the set where  $h_t$  does not converge to 0, we have

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

which implies, by Condition (C1), that this set has probability 0.  $\square$

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

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

For  $u \in T_0M$ ,  $\|u\| = 1$ , 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

$$\begin{aligned} \cos(c_s(0)u, u) &= \frac{(u^1 - u^3 Z_s^2)^2 + (u^2 + u^3 Z_s^1)^2}{\left( (u^1 - u^3 Z_s^2)^2 + (u^2 + u^3 Z_s^1)^2 + (-Z_s^2 u^1 + Z_s^1 u^2 + \|Z_s\|^2 u^3)^2 \right)^{1/2}}. \end{aligned}$$

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  $\sigma$  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  $F(\cdot, X(x))$ ,  $x \in M$  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  $\sigma$  be a stopping time which is dominated by the first exit time of  $X(x)$  from some relatively compact neighbourhood of  $x$ . We first note that

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

where  $X_{\sigma*}$  is the derivative process at the random time  $\sigma$ . Eq. (7.1) follows from the fact that the local martingale  $F(\cdot, X(x))$ , differentiated in the direction  $v$  at  $x$ , is again a local martingale, and under the given assumptions a uniformly integrable martingale when stopped at  $\sigma$ . 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

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

and

$$(7.3) \quad \sigma(r) = \inf\{s \geq 0 : T(s) \geq r\} \leq \tau_D(x).$$

Note that  $T(r) \rightarrow \infty$  as  $r \nearrow \tau_D(x)$ , almost surely, see [24]. Fix  $t_0 > 0$  and consider

$$(7.4) \quad \ell_s = \frac{1}{t_0} \rho \left( \int_0^s \varphi^{-2}(X_r(x)) dr \right) v$$

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  $s \geq t_0$ . Then  $\ell_0 = 0$  and  $\ell_s = v$  for  $s \geq \sigma(t_0)$ .

Now for perturbations  $X^\lambda$  of  $X$ , as in Section 3, let

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

and  $\sigma^\lambda(r) = \inf\{s \geq 0 : T^\lambda(s) \geq r\}$ . We introduce the abbreviation  $\partial_\lambda = (\frac{\partial}{\partial \lambda_1}, \dots, \frac{\partial}{\partial \lambda_n})$ . Then  $\partial_\lambda|_{\lambda=0} \ell_s^\lambda$  exists and lies in  $\cap_{p>1} L^p$ , see [24], Section 4

(the arguments there before Theorem 4.1 extend easily to general exponents  $p$ ). In a similar way, using  $T^\lambda \circ \sigma^\lambda = \text{id}$ , we see that

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

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  $\ell_s$  in an obvious way.

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

$$(7.5) \quad F(\cdot, X_\cdot^\lambda(x)) G^\lambda$$

of local martingales where

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

Then

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

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

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

then

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

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  $\ell$  is given by (7.4). Taking expectations gives

(7.8)

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

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

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

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

$$d_s^k = (X_{s*}^{-1}A)_x^* 1_{\{s \leq \tau_k\}},$$

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

We return to the term (7.9). Observe that

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

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

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

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

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

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

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

$$(7.11) \quad d(P_t f)_x v = \mathbb{E} \left[ f(X_t(x)) 1_{\{t < \zeta(x)\}} \Phi_t v \right]$$

for the minimal semigroup  $P_t f$  defined by (1.3) where  $\Phi_t$  is a  $T_x^*M$ -valued random variable which is  $L^p$ -integrable for any  $1 \leq p < \infty$  and local in the following sense: For any relatively compact neighbourhood  $D$  of  $x$  in  $M$  there is a choice for  $\Phi_t$  which is  $\mathcal{F}_\sigma$ -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 Eq. (7.10) gives

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

Again by taking into account that

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

we get the claimed formula.  $\square$

**Theorem 7.2.** *Let  $M$  be compact with smooth boundary  $\partial M \neq \emptyset$  and  $u \in C(M)$  be  $L$ -harmonic on  $M \setminus \partial M$ . Assume that (H1) holds. Let  $x \in M \setminus \partial M$  and  $v \in T_x M$ . Denote  $\tau(x)$  the first hitting time of  $X_\cdot(x)$  at  $\partial M$ . Then the following formula holds:*

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

where  $\Phi_{\tau(x)}$  is a  $T_x^* M$ -valued random variable which is in  $L^p$  for any  $1 \leq p < \infty$  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  $\mathcal{F}_\sigma$ -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.  $\square$

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

$$(7.13) \quad dS_t = \sigma(S_t) dW_t + \mu(S_t) dt,$$

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

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

where  $S_t$  is given as solution to (7.13) and

$$(7.14) \quad A_t = \int_0^t S_r dr.$$

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

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

with the vector fields

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

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  $\mu = 0$ , i.e.,

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

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

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

**Remark 7.5.** In the more general situation of Example 7.3, i.e.,

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

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

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

where the weight  $\pi_T$  is explicitly given and may be implemented numerically in Monte-Carlo simulations; compare with [14]. See [9] for extensions to jump diffusions, and [15] for weights  $\pi_T$  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 torsionfree connection  $\nabla$ . In stochastic terms, a smooth map  $u: M \rightarrow N$  is harmonic (with respect to  $L$ ) if it takes  $L$ -diffusions on  $M$  to  $\nabla$ -martingales on  $N$ . Likewise, a smooth map  $u: [0, t] \times M \rightarrow N$  is said to solve the nonlinear heat equation, if  $u(t - \cdot, X_\cdot(x))$  is a  $\nabla$ -martingale on  $N$  for any  $L$ -diffusion  $X_\cdot(x)$  on  $M$ .

Henceforth, we fix a family  $F(\cdot, X.(x))$ ,  $x \in M$  of  $\nabla$ -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  $\nabla$ -martingales on  $N$  and include the following two cases:

- (1)  $F(\cdot, X.(x)) = u \circ X.(x)$  for some harmonic map  $u: M \rightarrow N$ , and
- (2)  $F(\cdot, X.(x)) = u(t - \cdot, X.(x))$  where  $u$  solves the heat equation for maps  $M \rightarrow N$ .

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  $\nabla$ , then the geodesic (damped or deformed) transport  $\Theta_{0,t}: T_{Y_0}N \rightarrow T_{Y_t}N$  on  $N$  along  $Y$  is defined by the following covariant equation along  $Y$ :

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

where  $\//_{0,t}: T_{Y_0}N \rightarrow T_{Y_t}N$  is parallel translation on  $N$  along  $Y$  and  $R$  the curvature tensor to  $\nabla$ , see [3]. Finally, recall the notion of anti-development of  $Y$ , resp. “deformed anti-development” of  $Y$ ,

$$(8.2) \quad \mathcal{A}(Y) = \int_0^\cdot \//_{0,s}^{-1} \delta Y_s, \quad \mathcal{A}_{\text{def}}(Y) = \int_0^\cdot \Theta_{0,s}^{-1} \delta Y_s$$

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

**Theorem 8.1.** *Let  $F(\cdot, X.(x))$ ,  $x \in M$  be a family of  $\nabla$ -martingales on  $N$ , as described above. Then, for any predictable  $\mathbb{R}^r$ -valued process  $k$  in  $L_{\text{loc}}^2(Z)$ ,*

$$(8.3) \quad \Theta_{0,\cdot}^{-1} dF(\cdot, X.(x)) (T_x X.) \int_0^\cdot (X_{s*}^{-1} A)_x k_s ds - \mathcal{A}_{\text{def}}(F(\cdot, X.(x))) \int_0^\cdot \langle k, dZ \rangle$$

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

*Proof.* Observe that by [3],

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

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

$$\mathcal{A}_{\text{def}}(F(\cdot, X.(x))) = \int_0^\cdot \Theta_{0,s}^{-1} dF(s, \cdot)_{X_s(x)} A(X_s(x)) dZ_s.$$

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 non-linear setting by means of the local martingale (8.3).

**Theorem 8.2.** *Let  $u: [0, t] \times M \rightarrow N$  be a solution of the nonlinear heat equation,  $x \in M$ ,  $v \in T_x M$ . 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*

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

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

$$(8.4) \quad du(t, \cdot)_x v = \mathbb{E} \left[ \mathcal{A}_{\text{def}}(u(t - \cdot, X_\bullet(a)))_\sigma \int_0^\sigma \langle k, dZ \rangle \right].$$

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

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

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

$$(8.5) \quad (du)_x v = \mathbb{E} \left[ \mathcal{A}_{\text{def}}(u(X_\bullet(x)))_{\tau(x)} \int_0^{\tau(x)} \langle k, dZ \rangle \right].$$

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

$$(8.6) \quad \Theta_{0,\cdot}^{-1} dF(\cdot, X_\bullet(x)) (T_x X_\bullet) \int_0^\cdot (X_{s*}^{-1} A)_x a_s ds - \mathcal{A}_{\text{def}}(F(\cdot, X_\bullet(x))) \int_0^\cdot a_r^* dZ_r$$

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

$$(8.7) \quad a_s = (X_{s*}^{-1} A)_x^* 1_{\{s \leq \tau\}},$$

where  $\tau$  may be any predictable stopping time, we see that

$$(8.8) \quad n_s = \Theta_{0,s}^{-1} dF(s, \cdot)_{X_s(x)} X_{s*} C_{s \wedge \tau}(x) - \mathcal{A}_{\text{def}}(F(\cdot, X_\bullet(x)))_s \int_0^{s \wedge \tau} (X_{r*}^{-1} A)_x dZ_r$$

is a local martingale. Let

$$(8.9) \quad Y = \mathcal{A}_{\text{def}}(F(\cdot, X_\bullet(x))) \quad \text{and} \quad Y^\lambda = \mathcal{A}_{\text{def}}(F(\cdot, X_\bullet^\lambda(x))).$$

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

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

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

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

$$(8.11) \quad \text{vert} [\partial_\lambda \Big|_{\lambda=0} Y^\lambda] = \Theta_{0,\cdot}^{-1} J - J_0 + \int_0^\cdot \Theta_{0,s}^{-1} (\nabla \Theta_{0,s}) dY_s$$

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

$$(8.12) \quad (\nabla \Theta_{0,\cdot}) u = v_J^{-1} ((\Theta_{0,\cdot}^c, h_{J_0}(u))^{\text{vert}}).$$

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

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:

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

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

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

for some stopping time  $\sigma$ . 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, Li [12, 13] 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  $F(\cdot, X.(x))$  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, Xue-Mei Li and A. Thalmaier, *Manifold-valued martingales, changes of probabilities, and smoothness of finely harmonic maps*, Ann. Inst. H. Poincaré Probab. Statist. **35** (1999), 765–792.
- [2] M. Arnaudon and A. Thalmaier, *Stability of stochastic differential equations in manifolds*, Séminaire de Probabilités, XXXII, 188–214. Lecture Notes in Math. **1686**. Berlin: Springer, 1998.
- [3] M. Arnaudon and A. Thalmaier, *Complete lifts of connections and stochastic Jacobi fields*, J. Math. Pures Appl. **77** (1998), 283–315.
- [4] D.R. Bell, *The Malliavin calculus*, Pitman Monographs and Surveys in Pure and Applied Mathematics, vol. 34, Longman Scientific & Technical, Harlow, 1987.
- [5] D.R. Bell, *Divergence theorems in path space. III. Hypoelliptic diffusions and beyond*, J. Funct. Anal. **251** (2007), no. 1, 232–253.
- [6] 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.
- [7] G. Ben Arous, S. Kusuoka and D. Stroock, *The Poisson kernel for certain degenerate elliptic operators*, J. Funct. Anal. **56** (1984), 171–209.
- [8] J.-M. Bismut, *Martingales, the Malliavin calculus and hypoellipticity under general Hörmander's conditions*, Z. Wahrsch. Verw. Gebiete **56** (1981), 469–505.
- [9] T. Cass and P. Friz, *The Bismut-Elworthy-Li formula for jump-diffusions and applications to Monte Carlo pricing in finance*, arXiv math/0604311.
- [10] B.K. Driver and A. Thalmaier, *Heat equation derivative formulas for vector bundles*, J. Funct. Anal. **183** (2001), no. 1, 42–108.
- [11] K.D. Elworthy, Y. Le Jan and Xue-Mei Li, *Integration by parts formulae for degenerate diffusion measures on path spaces and diffeomorphism groups*, C. R. Acad. Sci. Paris, Sér. I. **323** (1996), 921–926.
- [12] K.D. Elworthy, Y. Le Jan, and Xue-Mei 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.
- [13] K.D. Elworthy, Y. Le Jan, and Xue-Mei Li, *On the geometry of diffusion operators and stochastic flows*, Lecture Notes in Mathematics, vol. 1720, Springer-Verlag, Berlin, 1999.
- [14] E. Fournié, J.-M. Lasry, J. Lebuchoux, P.-L. Lions, N. Touzi, *Applications of Malliavin calculus to Monte Carlo methods in finance*, Finance Stoch. **3** (1999), no. 4, 391–412.
- [15] 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), no. 5, 2132–2153.

- [16] H. Kunita, *On the decomposition of solutions of stochastic differential equations*, In: D. Williams (Ed.) *Stochastic integrals*. Proc. Sympos., Durham, 1980. Lecture Notes in Math. **851** (1981), 213–255.
- [17] S. Kusuoka, *The Malliavin calculus and its applications*, Sugaku Expositions **3** (1990), 127–144.
- [18] N.V. Krylov, *Quasiderivatives for solutions of Itô's stochastic equations and their applications*, Stochastic analysis and related topics (Oslo, 1992), 1–44, Stochastics Monographs **8**, Gordon and Breach, Montreux, 1993.
- [19] 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.
- [20] J. Norris, *Simplified Malliavin calculus*, Séminaire de Probabilités, XX, 101–130. Lecture Notes in Math. **1204**. Berlin: Springer, 1986.
- [21] D. Nualart, *The Malliavin calculus and related topics*, New York: Springer, 1995.
- [22] A. Thalmaier, *On the differentiation of heat semigroups and Poisson integrals*, Stochastics and Stochastics Reports **61** (1997), 297–321.
- [23] A. Thalmaier, *Some remarks on the heat flow for functions and forms*, Electron. Comm. Probab. **3** (1998), 43–49.
- [24] A. Thalmaier and F.-Y. Wang, *Gradient estimates for harmonic functions on regular domains in Riemannian manifolds*, J. Funct. Anal. **155** (1998), 109–124.

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

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