Séminaire de Probabilités, XXXII, 188–214.
Lecture Notes in Mathematics **1686**,
Springer: Berlin, 1998.

Stability of stochastic differential equations in manifolds

Marc Arnaudon
Institut de Recherche Mathématique Avancée
Université Louis Pasteur et CNRS
7, rue René Descartes
F–67084 Strasbourg Cedex
France
arnaudon@math.u-strasbg.fr

Anton Thalmaier
Naturwissenschaftliche Fakultät I – Mathematik
Universität Regensburg
D–93040 Regensburg
Germany
anton.thalmaier@mathematik.uni-regensburg.de

Abstract. — We extend the so-called topology of semimartingales to continuous semimartingales with values in a manifold and with lifetime, and prove that if the manifold is endowed with a connection ∇ then this topology and the topology of compact convergence in probability coincide on the set of continuous ∇-martingales. For the topology of manifold-valued semimartingales, we give results on differentiation with respect to a parameter for second order, Stratonovich and Itô stochastic differential equations and identify the equation solved by the derivative processes. In particular, we prove that both Stratonovich and Itô equations differentiate like equations involving smooth paths (for the Itô equation the tangent bundles must be endowed with the complete lifts of the connections on the manifolds). As applications, we prove that differentiation and antidevelopment of C^1 families of semimartingales commute, and that a semimartingale with values in a tangent bundle is a martingale for the complete lift of a connection if and only if it is the derivative of a family of martingales in the manifold.

1. Introduction

Let $(\Omega, (\mathscr{F}_t)_{0\le t<\infty}, \mathbb{P})$ denote a filtered probability space, M a smooth connected manifold endowed with a connection ∇. Then the tangent bundle TM inherits a connection ∇' (usually denoted by ∇^c), the complete lift of ∇ (see [Y-I] for details). Let X be a continuous semimartingale with values in M. The antidevelopment of X

in $T_{X_0}M$ is the semimartingale Z solving the Stratonovich equation

$$p(\delta Z) = U_0 U^{-1} \delta X, \quad Z_0 = 0, \tag{1.1}$$

where U is a horizontal lift of X taking values in the frame bundle on M and p is the canonical projection in TM of a vertical vector of TTM. The map $\mathcal{A}$ will denote the antidevelopment with respect to ∇ and $\mathcal{A}'$ the antidevelopment with respect to ∇'.

The initial motivation of this paper was to answer the following question: For some open interval I in $\mathbb{R}$, consider a family $(X_t(a))_{a\in I,\, t\in[0,\xi(a)[}$ of continuous martingales $X(a)$ in M, each with lifetime $\xi(a)$, differentiable in a for the topology of compact convergence in probability. Is then also $\big(X(a), \mathcal{A}(X(a))\big)$ differentiable in a, and if the answer is positive, do we have the relation $s\big(\partial_a \mathcal{A}(X(a))\big) = \mathcal{A}'\left(\partial_a X(a)\right)$ (where ∂_a denotes differentiation with respect to a and s is the map $TTM \to TTM$ defined by $s(\partial_a\partial_t x(t,a)) = \partial_t\partial_a x(t,a)$, if $(t,a) \mapsto x(t,a)$ is smooth and takes its values in M)?

A positive answer will be given to this question, and this result will be obtained as a particular case of general theorems on stability of stochastic differential equations.

In this paper equations of the general type

$$\mathcal{D}Z(a) = f\big(X(a), Z(a)\big)\, \mathcal{D}X(a) \tag{1.2}$$

between two manifolds M and N are studied, where $\mathcal{D}X(a)$ denotes the (formal) differential of order 2 of $X(a)$, and f is a Schwartz morphism between the second order bundles τM and τN. The topology of semimartingales, defined in [E1] for $\mathbb{R}$-valued processes, will be adapted to manifold-valued semimartingales with lifetime. In particular, it will be shown that the map $(X, f, Z_0) \mapsto (X, Z)$ is continuous, where Z is the maximal solution starting from Z_0 to $\mathcal{D}Z = f(X,Z)\mathcal{D}X$, with appropriate topologies on both sides.

When applied to a certain family of semimartingales and an appropriate Schwartz morphism, this result will tell us that if $a \mapsto X(a)$ is C^1 in the topology of semimartingales, and further if f is C^1 with locally Lipschitz derivative, $Z(a)$ the maximal solution to (1.2) with $(Z_0(a))_{a\in I}$ C^1 in probability, then $a \mapsto (X(a), Z(a))$ is C^1 in the topology of semimartingales and the derivative $\partial_a Z(a)$ is the maximal solution to

$$\mathcal{D}\partial_a Z(a) = f'\big(\partial_a X(a), \partial_a Z(a)\big)\, \mathcal{D}\partial_a X(a) \tag{1.3}$$

where f' is a Schwartz morphism between the second order bundles τTM and τTN.

As a corollary, we obtain results on differentiability of solutions to Stratonovich and Itô equations. It will be shown that they can be differentiated in the same way as solutions to ordinary differential equations (for the Itô case, the Itô differentials of the derivative process have to be defined with the complete lifts of the connections).

If M is endowed with a connection ∇, then it will be shown that, as in the flat case, the topology of semimartingales and the topology of uniform convergence in probability on compact sets coincide on the set of martingales. Using these results it will be possible to prove commutativity of antidevelopment and differentiation.

Acknowledgement. — We would like to thank Michel Émery for his comments and suggestions to improve this paper.

2. Topologies of semimartingales and of uniform convergence in probability on compact sets

2.1. $\mathbb{R}^d$*-valued processes*

In this section we define topologies of uniform convergence in probability and of semimartingales for processes with lifetime. We investigate their main properties.

Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions. If ξ is a predictable stopping time, we denote by $D_{\mathrm{c}}(\mathbb{R}^d;\xi)$ the set of continuous adapted $\mathbb{R}^d$-valued processes with lifetime ξ, and by $\mathscr{S}(\mathbb{R}^d;\xi)$ the set of $\mathbb{R}^d$-valued continuous semimartingales with lifetime ξ. These sets are described as follows: an element of $D_{\mathrm{c}}(\mathbb{R}^d;\xi)$ (resp. $\mathscr{S}(\mathbb{R}^d;\xi)$) is the image under an isomorphic time change $A\colon [0,\xi[\, \to \{\xi>0\}\times[0,\infty)$ of an $\mathbb{R}^d$-valued continuous adapted process (resp. semimartingale) defined on the probability space $\Big(\{\xi>0\}, \big(\mathscr{F}_{A_s^{-1}}\big)_{s\geq 0}, \mathbb{P}(\,\cdot\,|\xi>0)\Big)$. They can be endowed with a complete metric space structure, as in the case $\xi=\infty$, which gives respectively the topology of compact convergence in probability and the topology of semimartingales (see [E1]). Let $\mathscr{T}$ denote the set of predictable stopping times and let

$$\hat D_{\mathrm{c}}(\mathbb{R}^d) = \bigcup_{\xi\in\mathscr{T}} D_{\mathrm{c}}(\mathbb{R}^d;\xi)\,, \quad \hat{\mathscr{S}}(\mathbb{R}^d) = \bigcup_{\xi\in\mathscr{T}} \mathscr{S}(\mathbb{R}^d;\xi)\,.$$

The sum $X+Y$, difference $X-Y$, product (X,Y) of two processes with lifetime is a process with lifetime the infimum of the lifetimes of the two processes. The lifetime of a process X will be denoted by ξ_X.

If T is a predictable stopping time, we can define the operations of stopping at T and killing at T on the sets $\hat D_{\mathrm{c}}(\mathbb{R}^d)$ and $\hat{\mathscr{S}}(\mathbb{R}^d)$: let X be an element of $\hat D_{\mathrm{c}}(\mathbb{R}^d)$ or $\hat{\mathscr{S}}(\mathbb{R}^d)$. Then the process X^T stopped at time T is the continuous process with lifetime $+\infty\, 1_{\{T<\xi_X\}} + \xi_X\, 1_{\{T\geq\xi_X\}}$ which coincides with X on $[0, T\wedge\xi_X[$ and is constant on $[T,\infty[\,\cap\,\{T<\xi_X\}$; the process $X^{T\text{-}}$ killed at time T is the continuous process which has lifetime $T\wedge\xi_X$ and coincides with X on $[0,T\wedge\xi_X[$. If ξ is any predictable stopping time, then by $T<\xi$ we will mean $T<\xi$ on $\{\xi>0\}$ and $T=0$ on $\{\xi=0\}$.

Let us define a topology on the sets $\hat D_{\mathrm{c}}(\mathbb{R}^d)$ and $\hat{\mathscr{S}}(\mathbb{R}^d)$. If $X\in\hat D_{\mathrm{c}}(\mathbb{R}^d)$ with lifetime ξ_X, T a predictable stopping time such that $T<\xi_X$ and $\varepsilon>0$, one defines neighbourhoods of X with the sets

$$V_{\mathrm{cp}}(X,T,\varepsilon) = \Big\{ Y\in\hat D_{\mathrm{c}}(\mathbb{R}^d),\ \mathbb{E}\Big[1\wedge \sup_{0<t\leq T} \|Y_t - X_t\|\Big] < \varepsilon \Big\}$$

(with the convention that $\sup\emptyset = 0$ and $\|Z_t\| = +\infty$ if $t\geq\xi_Z$) and

$$W_{\mathrm{cp}}(X,\varepsilon) = \Big\{ Y\in\hat D_{\mathrm{c}}(\mathbb{R}^d),\ \mathbb{P}\Big(\{\xi_Y > \xi_X+\varepsilon\}\cap\big\{\lim_{t\to\xi_X} X_t \text{ exists}\big\}\Big) < \varepsilon \Big\}$$

(the second condition will insure that the topology is separated).

Analogously, one defines neighbourhoods of $X\in\hat{\mathscr{S}}(\mathbb{R}^d)$ by setting

$$V(X,T,\varepsilon) = \Big\{ Y\in\hat{\mathscr{S}}(\mathbb{R}^d),\ \mathbb{E}\big[1\wedge v(Y-X)_T\big] < \varepsilon \Big\}$$

where if $Z = Z_0 + M + A$ is the canonical decomposition of $Z \in \hat{\mathscr{S}}(\mathbb{R}^d)$,

$$v(Z)_t = \sum_{i=1}^{d} \left(|Z_0^i| + <M^i, M^i>_t^{1/2} + \int_0^t |dA^i| \right)$$

(with the convention that $v(Z)_t = +\infty$ if $t \geq \xi_Z$) and

$$W(X, \varepsilon) = \left\{ Y \in \hat{\mathscr{S}}(\mathbb{R}^d),\ \mathbb{P}\Big(\{\xi_Y > \xi_X + \varepsilon\} \cap \{ \lim_{t \to \xi_X} X_t \text{ exists}\} \Big) < \varepsilon \right\}.$$

Proposition and definition 2.1. — *The basis of neighbourhoods*

$$\left.\begin{array}{ll} V_{\mathrm{cp}}(X, T, \varepsilon) \cap W_{\mathrm{cp}}(X, \varepsilon'), & X \in \hat{D}_{\mathrm{c}}(\mathbb{R}^d) \\ \big(\textit{resp. } V(X, T, \varepsilon) \cap W(X, \varepsilon'), & X \in \hat{\mathscr{S}}(\mathbb{R}^d) \big) \end{array}\right\} \begin{array}{l} \varepsilon, \varepsilon' > 0,\ T \textit{ predictable} \\ \textit{stopping time such that } T < \xi_X, \end{array}$$

defines a separated topology on $\hat{D}_{\mathrm{c}}(\mathbb{R}^d)$ (resp. $\hat{\mathscr{S}}(\mathbb{R}^d)$) such that every point has a countable basis of neighbourhoods. This topology will be called the topology of compact convergence in probability (resp. the topology of semimartingales).

Remarks. — 1) If for the topology in $\hat{D}_{\mathrm{c}}(\mathbb{R}^d)$ (resp. $\hat{\mathscr{S}}(\mathbb{R}^d)$) $(X^n)_{n\in\mathbb{N}}$ converges to X, then $\xi_{X^n} \wedge \xi_X$ converges in probability to ξ_X and ξ_{X^n} converges to ξ_X in probability on the set $\{\lim_{t\to\xi_X} X_t \text{ exists}\}$.

2) Let $\xi \in \mathscr{T}$. The topology of the complete metric space $(D_{\mathrm{c}}(\mathbb{R}^d, \xi), d_{\mathrm{cp}})$ (resp. $(\mathscr{S}(\mathbb{R}^d, \xi), d_{\mathrm{sm}})$) defined in [E1] is exactly the topology induced by $\hat{D}_{\mathrm{c}}(\mathbb{R}^d)$ (resp. $\hat{\mathscr{S}}(\mathbb{R}^d)$) on $D_{\mathrm{c}}(\mathbb{R}^d, \xi)$ (resp. $\mathscr{S}(\mathbb{R}^d, \xi)$).

Proof of Proposition 2.1. — We are going to prove this for $\hat{D}_{\mathrm{c}}(\mathbb{R}^d)$. To see that every point has a countable basis of neighbourhoods, one shows that it is sufficient to consider an increasing sequence of predictable stopping times $(T_m)_{m\in\mathbb{N}}$ converging to ξ_X and such that $T_m < \xi_X$ for all m.

Let us show that the topology is separated. If $X \neq Y$, then two situations can occur. Either there exists $\varepsilon > 0$ and a predictable stopping time T with $T < \xi_X \wedge \xi_Y$ and $\mathbb{E}\Big[1 \wedge \sup\limits_{0<t\leq T} \|Y_t - X_t\|\Big] > 2\varepsilon$ in which case $V_{\mathrm{cp}}(X, T, \varepsilon) \cap V_{\mathrm{cp}}(Y, T, \varepsilon) = \emptyset$, or $Y^{\xi_X -} = X^{\xi_Y -}$ with $\mathbb{P}(\xi_Y < \xi_X) > 0$ and there exists $\varepsilon > 0$ and a predictable stopping time T satisfying $T < \xi_X$ such that $\mathbb{P}\Big(\Big\{\xi_Y + 2\varepsilon < T,\ \lim\limits_{t\to\xi_Y} Y_t \text{ exists}\Big\}\Big) > 2\varepsilon$; in this case, one verifies that $V_{\mathrm{cp}}(X, T, \varepsilon) \cap W_{\mathrm{cp}}(Y, \varepsilon) = \emptyset$. ∎

Remark. — Convergence for the topology of semimartingales implies compact convergence in probability.

For $1 \leq p \leq \infty$ and $\xi \in \mathscr{T}$, let $S^p(\mathbb{R}^d, \xi)$ denote the Banach space of processes $X \in D_{\mathrm{c}}(\mathbb{R}^d, \xi)$ such that $\|X\|_{S^p(\mathbb{R}^d, \xi)} = \|X_\xi^*\|_{L^p} < \infty$, where $X_t^* = \sup_{s<t} \|X_s\|$ on $0 \leq t \leq \xi$ and $\sup \emptyset = 0$. Let $\hat{S}^p(\mathbb{R}^d) = \cup_{\xi\in\mathscr{T}} S^p(\mathbb{R}^d, \xi)$.

Definition 2.2. — *We say that a sequence $(X^n)_{n\in\mathbb{N}}$ in $\hat D_c(\mathbb{R}^d)$ converges to $X \in \hat D_c(\mathbb{R}^d)$ locally in $\hat D_c(\mathbb{R}^d)$ (resp. $\hat S^p(\mathbb{R}^d)$) if the following two conditions are satisfied*:

(i) *There exists an increasing sequence of stopping times $(T_m)_{m\in\mathbb{N}}$ converging to ξ_X such that for any m, $T_m < \xi_X$, $(X^n)^{T_m-}$ belongs to $D_c(\mathbb{R}^d, T_m)$ (resp. $S^p(\mathbb{R}^d, T_m)$) for n sufficiently large and converges in $D_c(\mathbb{R}^d, T_m)$ (resp. $S^p(\mathbb{R}^d, T_m)$) to X^{T_m-}.*

(ii) *The lifetimes ξ_{X^n} converge in probability to the lifetime ξ_X on the set $\{\lim_{t\to\xi_X} X_t$ exists$\}$.*

For $1 \le p \le \infty$ and $\xi \in \mathscr{T}$, let $H^p(\mathbb{R}^d,\xi)$ be the space of processes $X \in \mathscr{S}(\mathbb{R}^d,\xi)$ such that $\|X\|_{H^p(\mathbb{R}^d,\xi)} = \|v(X)_\xi\|_{L^p} < \infty$. Let $\hat H^p(\mathbb{R}^d) = \cup_{\xi\in\mathscr{T}} H^p(\mathbb{R}^d,\xi)$.

Definition 2.3. — *We say that a sequence $(X^n)_{n\in\mathbb{N}}$ in $\hat{\mathscr{S}}(\mathbb{R}^d)$ converges to $X \in \hat{\mathscr{S}}(\mathbb{R}^d)$ locally in $\hat{\mathscr{S}}(\mathbb{R}^d)$ (resp. $\hat H^p(\mathbb{R}^d)$) if the following two conditions are satisfied*:

(i) *There exists an increasing sequence of stopping times $(T_m)_{m\in\mathbb{N}}$ converging to ξ_X such that for any m, $T_m < \xi_X$, $(X^n)^{T_m-}$ belongs to $\mathscr{S}(\mathbb{R}^d, T_m)$ (resp. $H^p(\mathbb{R}^d, T_m)$) for n sufficiently large and converges in $\mathscr{S}(\mathbb{R}^d, T_m)$ (resp. $H^p(\mathbb{R}^d, T_m)$) to X^{T_m-}.*

(ii) *The lifetimes ξ_{X^n} converge in probability to the lifetime ξ_X on the set $\{\lim_{t\to\xi_X} X_t$ exists$\}$.*

Note that local convergences are not derived from topologies. Their relation to topologies is described in the following proposition which is the analogue for processes with lifetime of [E1] Proposition 1 and Theorem 2.

Proposition 2.4. — *Let $p \in [1,\infty]$ and let $E \subset \hat D_c(\mathbb{R}^d)$ (resp. $E \subset \hat{\mathscr{S}}(\mathbb{R}^d)$). Let F be the sequential closure of E for local convergence in $\hat D_c(\mathbb{R}^d)$ (resp. $\hat{\mathscr{S}}(\mathbb{R}^d)$), let G be the closure of E for the topology of compact convergence in probability (resp. for the topology of semimartingales), and let K_p be the sequential closure of E for local convergence in $\hat S^p(\mathbb{R}^d)$ (resp. $\hat H^p(\mathbb{R}^d)$).*

Then $F = G = K_p$.

Remark. — Proposition 2.4 can be rewritten as follows: let $(X^n)_{n\in\mathbb{N}}$ be a sequence of elements of $\hat D_c(\mathbb{R}^d)$ (resp. $\hat{\mathscr{S}}(\mathbb{R}^d)$). Then the following three conditions are equivalent:

(i) for every subsequence $(Y^n)_{n\in\mathbb{N}}$, there exists a subsubsequence $(Z^n)_{n\in\mathbb{N}}$ which converges to X^0 locally in $\hat D_c(\mathbb{R}^d)$ (resp. $\hat{\mathscr{S}}(\mathbb{R}^d)$),

(ii) $(X^n)_{n\in\mathbb{N}}$ converges to X^0 in the topology of compact convergence in probability (resp. in the topology of semimartingales),

(iii) for every subsequence $(Y^n)_{n\in\mathbb{N}}$, there exists a subsubsequence $(Z^n)_{n\in\mathbb{N}}$ which converges to X^0 locally in $\hat S^p(\mathbb{R}^d)$ (resp. $\hat H^p(\mathbb{R}^d)$).

Proof of Proposition 2.4. — 1) Second equality: We will give the proof for compact convergence in probability. The proof for semimartingale convergence is similar.

To prove $K_1 \subset G$, it is sufficient to verify that if X^n converges to X locally in $\hat S^1(\mathbb{R}^d)$, then X^n converges to X for the topology of compact convergence in probability, and this is almost evident.

We are left to prove that $G \subset K_\infty$, i.e. that for every sequence X^n converging to X for the topology of compact convergence in probability, there exists a subsequence which converges to X locally in $\hat{S}^\infty(\mathbb{R}^d)$. One easily shows that condition (ii) of local convergence is satisfied, without extracting a subsequence. By extracting a subsequence to obtain an a.s. convergence of $\xi_{X^n} \wedge \xi_X$ and by stopping at a time smaller than ξ_X but close to ξ_X in probability, one may assume that all the terms of the sequence belong to $D_{\mathrm{c}}(\mathbb{R}^d, \infty)$. One can also assume that $X = 0$. It is then sufficient to show that we can find a stopping time T as big as we want for the topology of convergence in probability and a subsequence $(X^{n_k})_{k\in\mathbb{N}}$ such that $(X^{n_k})^T$ converges to 0 in $S^\infty(\mathbb{R}^d, \infty)$ (a sequence of stopping time increasing to ∞ and a diagonal subsequence give then the result). But for every $M \in \mathbb{N}^*$, $(X^n)^*_M$ converges in probability to 0. By extracting a subsequence one can assume that the convergence is almost sure. The end of the proof is similar to the proof of Egoroff's theorem: let $\varepsilon > 0$, $T^n_m = M \wedge \inf\{t > 0,\ \|(X^n)^*_t\| \geq 1/m\}$, $S^n_m = \inf_{k\geq n} T^k_m$, $n(m)$ such that $\mathbb{P}\big(S^{n(m)}_m < M - 1\big) < \dfrac{\varepsilon}{2^m}$, and $R = \inf_{m\in\mathbb{N}^*} S^{n(m)}_m$. Then R is as close to ∞ as we want and $(X^n)^R$ converges a.s. uniformly to 0.

2) The proof of the first equality is identical as the one for infinite times. ∎

As a corollary, using the demonstration of Theorem 2 in [E1], one can show that a sequence $(X^n)_{n\in\mathbb{N}}$ of elements of $\hat{\mathscr{S}}(\mathbb{R}^d)$ converges to $X \in \hat{\mathscr{S}}(\mathbb{R}^d)$ if and only if it converges in $\hat{D}_{\mathrm{c}}(\mathbb{R}^d)$ and for all bounded predictable process H with values in $\mathbb{R}^d$, $\left(\int_0^\cdot H dX^n\right)^{\xi_X -}$ converges in $\hat{D}_{\mathrm{c}}(\mathbb{R})$ to $\int_0^\cdot H dX$ (compare with the definition of the topology of semimartingales in [E1]).

Definition 2.5. — *Let $E, F = \hat{D}_{\mathrm{c}}(\mathbb{R}^d)$ or $\hat{\mathscr{S}}(\mathbb{R}^d)$, and let ϕ: $E \to F$ be a map. We will say that ϕ is lower semicontinuous if for every sequence $(X^n)_{n\in\mathbb{N}}$ of elements in E converging to $X \in E$, the sequence $\big((\phi(X^n))^{\xi_{\phi(X)}-}\big)_{n\in\mathbb{N}}$ converges to $\phi(X)$.*

An important example of a lower semicontinuous map is $X \mapsto p(\phi(X)) \in \hat{\mathscr{S}}(\mathbb{R}^d)$ if $X \mapsto \phi(X) \in \hat{\mathscr{S}}(\mathbb{R}^{d+d'})$ is continuous and p: $\mathbb{R}^{d+d'} \to \mathbb{R}^d$ the canonical projection.

Note also that if $X \mapsto \phi(X)$ is lower semicontinuous, and if both X and $\phi(X)$ are in $\hat{D}_{\mathrm{c}}$ (or $\hat{\mathscr{S}}$) and the lifetime of $\phi(X)$ is greater or equal to the lifetime of X, then $X \mapsto (X, \phi(X))$ is continuous.

With Proposition 2.4, one can investigate continuity properties for operations on the sets of continuous adapted processes and of semimartingales. For $m \in \mathbb{N}$, let $C^m(\mathbb{R}^d)$ denote the set of real-valued C^m functions on $\mathbb{R}^d$, endowed with the topology of uniform convergence on compact sets of the derivatives up to order m.

Proposition 2.6. — 1) *The map*

$$\begin{aligned} C^0(\mathbb{R}^d) \times \hat{D}_{\mathrm{c}}(\mathbb{R}^d) &\longrightarrow \hat{D}_{\mathrm{c}}(\mathbb{R}) \\ (h, X) &\longmapsto h(X) \end{aligned}$$

is lower semicontinuous.

2) *The maps*

$$\begin{aligned} C^2(\mathbb{R}^d)\times \hat{\mathscr{S}}(\mathbb{R}^d) &\longrightarrow \hat{\mathscr{S}}(\mathbb{R}) \\ (h,X) &\longmapsto h(X) \end{aligned}$$

and

$$\begin{aligned} \hat{\mathscr{S}}(\mathbb{R}^d) &\longrightarrow \hat{\mathscr{S}}(\mathbb{R}) \\ X &\longmapsto M^i,\ A^i,\ <M^i,M^j> \end{aligned}$$

are lower semicontinuous, where $X=X_0+M+A$ *is the decomposition of* X *into the value at* 0, *a local martingale and a process with finite variation.*

3) *Let* T *be a predictable stopping time. The map*

$$\begin{aligned} \hat{D}_{\mathrm{c}}(\mathbb{R}^d)\ (\textit{resp. } \hat{\mathscr{S}}(\mathbb{R}^d)) &\longrightarrow \hat{D}_{\mathrm{c}}(\mathbb{R}^d)\ (\textit{resp. } \hat{\mathscr{S}}(\mathbb{R}^d)) \\ X &\longmapsto X^{T-} \end{aligned}$$

is continuous, and

$$\begin{aligned} \hat{D}_{\mathrm{c}}(\mathbb{R}^d)\ (\textit{resp. } \hat{\mathscr{S}}(\mathbb{R}^d)) &\longrightarrow \hat{D}_{\mathrm{c}}(\mathbb{R}^d)\ (\textit{resp. } \hat{\mathscr{S}}(\mathbb{R}^d)) \\ X &\longmapsto X^{T} \end{aligned}$$

is lower semicontinuous and continuous at the points X *with lifetime* ξ_X *such that* $\mathbb{P}(\xi_X=T)=0$.

4) *Let* U *be an open subset of* $\mathbb{R}^d$. *If* X *belongs to* $\hat{D}_{\mathrm{c}}(\mathbb{R}^d)$, *let* $T_U(X)$ *denote the exit time of* X *from* U, *i.e.,* $T_U(X)=\inf\{t>0,\ X_t\notin U\}$ (*with* $\inf\emptyset=+\infty$). *Then*

$$\begin{aligned} \hat{D}_{\mathrm{c}}(\mathbb{R}^d)\ (\hat{\mathscr{S}}(\mathbb{R}^d)) &\longrightarrow \hat{D}_{\mathrm{c}}(\mathbb{R}^d)\ (\hat{\mathscr{S}}(\mathbb{R}^d)) \\ X &\longmapsto X^{T_U(X)-} \end{aligned}$$

is lower semicontinuous, and

$$\begin{aligned} \hat{D}_{\mathrm{c}}(\mathbb{R}^d)\ (\hat{\mathscr{S}}(\mathbb{R}^d)) &\longrightarrow \hat{D}_{\mathrm{c}}(U)\ (\hat{\mathscr{S}}(U)) \\ X &\longmapsto X^{T_U(X)-} \end{aligned}$$

is continuous.

In part 4), $\hat{D}_{\mathrm{c}}(U)$ $(\hat{\mathscr{S}}(U))$ is the set of elements of $\hat{D}_{\mathrm{c}}(\mathbb{R}^d)$ $(\hat{\mathscr{S}}(\mathbb{R}^d))$ which take their values in U, endowed with a topology defined in the same manner.

Proof. — 1) By Proposition 2.4, is sufficient to show that for every sequence (h^n,X^n) converging to (h,X), there exists a subsequence (h^{n_k},X^{n_k}) such that $h^{n_k}(X^{n_k})$ satisfies condition (i) of local convergence to $h(X)$ in $\hat{S}^\infty(\mathbb{R}^d)$. But using again Proposition 2.4, by extracting a subsequence, we can assume that the X^n are locally bounded and converge locally a.s. uniformly to X. We conclude using the fact that h is uniformly continuous on compact sets and h^n converges to h uniformly on compact sets.

2) The proof is analogous to 1) using the equality

$$v(h(X)) = \left|h(X_0)\right| + \left(\int_0 D_i h(X) D_j h(X)\, d{<}M^i, M^j{>}\right)^{1/2} + \int_0 \left| \frac{1}{2} D_{ij} h(X)\, d{<}M^i, M^j{>} + D_i h(X)\, dA^i \right|$$

and condition (i) of local convergence in $\hat{H}^\infty(\mathbb{R}^d)$.

3) The proof is left to the reader.

4) We only give a sketch of the proof for the second assertion. It is sufficient to prove that for every T satisfying $T < T_U(X) \wedge \xi_X$, $T_U(X^n) \wedge T$ converges in probability to T, and that $T_U(X^n)$ converges in probability to $T_U(X)$ on the event $\left\{\lim_{t \to \xi_X \wedge T_U(X)} X_t \text{ exists in } U\right\}$. But this is a consequence of the existence for every subsequence $(X^{n_k})_{k \in \mathbb{N}}$ of a subsubsequence which converges locally a.s. uniformly. $\square$

A consequence of 1) is that if F is a closed subspace of $\mathbb{R}^d$, then taking $h(x) = \operatorname{dist}(x, F)$ shows that the subset of $\hat{D}_c(\mathbb{R}^d)$ ($\hat{\mathscr{S}}(\mathbb{R}^d)$) consisting of F-valued processes is closed. This topological subspace will be denoted by $\hat{D}_c(F)$ ($\hat{\mathscr{S}}(F)$).

Property 4) is very useful for the study of manifold-valued processes and stochastic differential equations. It removes problems in connection with the exit time from domains of definition. It allows localization in time.

We are now interested in differentiability properties.

Definition 2.7. — *Let $a \mapsto X(a) \in \hat{\mathscr{S}}(\mathbb{R}^d)$ be defined on some interval I in $\mathbb{R}$.*

1) *The map $a \mapsto X(a)$ is differentiable in $\hat{\mathscr{S}}(\mathbb{R}^d)$ at $a_0 \in I$ if it is continuous at a_0 and if there exists $Y \in \hat{\mathscr{S}}(\mathbb{R}^d)$ such that $\dfrac{X(a) - X(a_0)}{a - a_0}$ converges in $\hat{\mathscr{S}}(\mathbb{R}^d)$ to Y as $a \to a_0$. Then $(X(a_0), Y)$ is called the derivative of X at a_0.*

2) *The map $a \mapsto X(a)$ is C^1 in $\hat{\mathscr{S}}(\mathbb{R}^d)$ if for all $a_0 \in I$, $a \mapsto X(a)$ is differentiable in $\hat{\mathscr{S}}(\mathbb{R}^d)$ at a_0, and if the derivative $a \mapsto Y(a)$ is continuous in $\hat{\mathscr{S}}(\mathbb{R}^{2d})$. The semimartingale $Y(a)$ is denoted by $\partial_a X(a)$.*

3) *For $k \geq 1$, the map $a \mapsto X(a)$ is C^{k+1} in $\hat{\mathscr{S}}(\mathbb{R}^d)$ if $a \mapsto X(a)$ is C^1 in $\hat{\mathscr{S}}(\mathbb{R}^d)$ and $\partial_a X(a)$ is C^k in $\hat{\mathscr{S}}(\mathbb{R}^{2d})$.*

Remarks. — 1) In the first part of the definition, one asks $a \mapsto X(a)$ to be continuous at a_0 only to guarantee that $\xi_{X(a)}$ converges in probability to $\xi_{X(a_0)}$ on the set $\left\{\lim_{t \to \xi(a_0)} X_t(a_0) \text{ exists}\right\}$.

2) In the same manner, replacing $\hat{\mathscr{S}}(\mathbb{R}^d)$ by $\hat{D}_c(\mathbb{R}^d)$ in Definition 2.7, the notion of a map $a \mapsto X(a) \in \hat{D}_c(\mathbb{R}^d)$ being C^k in $\hat{D}_c(\mathbb{R}^d)$ can be defined.

The following proposition says that regularity of paths implies regularity in $\hat{D}_c(\mathbb{R}^d)$.

PROPOSITION 2.8. — *Let $k \geq 0$. Suppose $a \mapsto X(a) \in \hat{D}_{\mathrm{c}}(\mathbb{R}^d)$, with lifetime $\xi(a)$, is defined on an open interval I in $\mathbb{R}$. Assume that ω-almost surely, $a \mapsto \xi(a)(\omega)$ is lower semicontinuous and continuous at a_0 if $\lim_{t\to\xi(a_0)(\omega)} X_t(a_0)$ exists, $a \mapsto X_t(a)(\omega)$ is of class C^k on its domain for all t, and that the map $(t,a) \mapsto \partial_a^k\big(X_t(a)(\omega)\big)$, defined on $\big\{(t,a) \in \mathbb{R}_+ \times I,\ 0 \leq t < \xi(a)(\omega)\big\}$, is continuous.*

Then $a \mapsto X(a)$ is C^k in $\hat{D}_{\mathrm{c}}(\mathbb{R}^d)$.

Proof. — Let us first consider the case $k=0$. Let $(a_\ell)_{\ell\in\mathbb{N}^*}$ be a sequence of elements of I converging to $a_0 \in I$. Then $\xi(a_\ell)$ converges almost surely to $\xi(a_0)$ on the set $\Big\{\lim_{t\to\xi(a_0)(\omega)} X_t(a_0) \text{ exists}\Big\}$, hence for $\varepsilon > 0$, $X(a_\ell) \in W_{\mathrm{cp}}\big(X(a_0),\varepsilon\big)$ for ℓ sufficiently large. Since $\xi(a_0) \wedge \xi(a_\ell)$ converges almost surely to $\xi(a_0)$, the stopping times $T'_m = \inf_{\ell\geq m} \xi(a_0) \wedge \xi(a_\ell)$ are predictable, increasing in m, and converge still almost surely to $\xi(a_0)$. Thus there exists a sequence of predictable stopping times $(T_m)_{m\in\mathbb{N}^*}$ increasing almost surely to $\xi(a_0)$, such that almost surely, for all m, $T_m < T'_m$ on $\{T'_m > 0\}$.

By the second part of Proposition 2.4, it is sufficient to show that $X(a_\ell)^{T_m-}$ converges in $\hat{D}_{\mathrm{c}}(\mathbb{R}^d)$ to $X(a_0)^{T_m-}$ as ℓ tends to ∞. But on $\{T_m > 0\}$, almost surely, there exists $\varepsilon(\omega) > 0$ such that the map

$$\begin{aligned} [0,T_m(\omega)] \times [a_0 - \varepsilon(\omega), a_0 + \varepsilon(\omega)] &\longrightarrow \mathbb{R}^d \\ (t,a) &\longmapsto X_t(a)(\omega) \end{aligned}$$

is well-defined and uniformly continuous. Thus $\lim_{\ell\to\infty} \sup_{0\leq t\leq T_m} \|X_t(a_\ell) - X_t(a_0)\| = 0$ almost surely on $\{T_m > 0\}$, and this gives the convergence of $X(a_\ell)^{T_m-}$ to $X(a_0)^{T_m-}$ in $\hat{D}_{\mathrm{c}}(\mathbb{R}^d)$. Hence we have the result.

If $k=1$, let a_0, $(a_\ell)_{\ell\in\mathbb{N}^*}$, $(T_m)_{m\in\mathbb{N}^*}$ be as above. It is sufficient to prove that for every m,

$$\frac{X^{T_m-}(a_\ell) - X^{T_m-}(a_0)}{a_\ell - a_0}$$

converges to $\partial_a X^{T_m-}(a_0)$ in $\hat{D}_{\mathrm{c}}(\mathbb{R}^d)$, as ℓ tends to ∞. Almost surely on $\{T_m > 0\}$, there exists $\varepsilon(\omega) > 0$ such that the map

$$\begin{aligned} [0,T_m(\omega)] \times [a_0 - \varepsilon(\omega), a_0 + \varepsilon(\omega)] &\longrightarrow \mathbb{R}^d \\ (t,a) &\longmapsto \partial_a X_t(a)(\omega) \end{aligned}$$

is defined and uniformly continuous. But, for such ω, t, a, we have

$$\left\| \frac{X_t(a) - X_t(a_0)}{a - a_0} - \partial_a X_t(a_0) \right\| \leq \sup_{\|b-a_0\|\leq\|a-a_0\|} \big\|\partial_a X_t(b) - \partial_a X_t(a_0)\big\|,$$

hence

$$\begin{aligned} \sup_{0\leq t\leq T_m} &\left\| \frac{X_t(a_\ell) - X_t(a_0)}{a_\ell - a_0} - \partial_a X_t(a_0) \right\| \\ &\leq \sup_{\|b-a_0\|\leq\|a_\ell-a_0\|} \ \sup_{0\leq t\leq T_m} \big\|\partial_a X_t(b) - \partial_a X_t(a_0)\big\| \end{aligned}$$

and the left-hand side converges almost surely to 0 as ℓ tends to ∞. It implies that $\dfrac{X^{T_m\text{-}}(a_\ell) - X^{T_m\text{-}}(a_0)}{a_\ell - a_0}$ converges to $\partial_a X^{T_m\text{-}}(a_0)$ in $\hat{D}_{\mathrm{c}}(\mathbb{R}^d)$, as ℓ tends to ∞.

If $k \geq 2$, one can prove in the same way by induction that for $\ell \leq k$, $a \mapsto X(a)$ is C^ℓ in $\hat{D}_{\mathrm{c}}(\mathbb{R}^d)$, and almost surely, for all t, $\left(\partial_a^\ell X\right)_t = \partial_a^\ell (X_t)$. □

REMARK. — Proposition 2.8 is false with $\hat{\mathscr{S}}(\mathbb{R}^d)$.

2.2. *Manifold-valued processes*

Let M be a connected smooth manifold endowed with a connection ∇. With respect to some fixed filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$, for every predictable stopping time ξ, let let $D_{\mathrm{c}}(M,\xi)$ denote the set of M-valued adapted continuous processes with lifetime ξ, and $\mathscr{S}(M,\xi)$ the set of M-valued continuous semimartingales with lifetime ξ. The spaces $D_{\mathrm{c}}(F;\xi)$, $\mathscr{S}(F;\xi)$, $\hat{D}_{\mathrm{c}}(F)$, $\hat{\mathscr{S}}(F)$, where F is a closed subset of M are defined by analogy with the previous definitions.

Let ϕ: $M \to \mathbb{R}^d$ be a smooth proper embedding. Then $\phi(M)$ is a closed subset of $\mathbb{R}^d$. As a consequence, $(\hat{D}_{\mathrm{c}}(\phi(M)), d_{\mathrm{cp}})$, resp. $(\hat{\mathscr{S}}(\phi(M)), d_{\mathrm{sm}})$, is a topological subspace of $\hat{D}_{\mathrm{c}}(\mathbb{R}^d)$, resp. $\hat{\mathscr{S}}(\mathbb{R}^d)$. By means of the diffeomorphism ϕ: $M \to \phi(M)$, we obtain complete topological space structures on $\hat{D}_{\mathrm{c}}(M)$ and $\hat{\mathscr{S}}(M)$.

DEFINITION 2.9. — *Let ϕ: $M \to \mathbb{R}^d$ be a smooth proper embedding.*

1) *The topology of compact convergence in probability on $\hat{D}_{\mathrm{c}}(M)$ is the topology induced by the diffeomorphism ϕ: $M \to \phi(M)$ and the topological space $\hat{D}_{\mathrm{c}}(\phi(M))$.*

2) *The topology of semimartingales on $\hat{\mathscr{S}}(M)$ is the topology induced by the diffeomorphism ϕ: $M \to \phi(M)$ and the topological space $\hat{\mathscr{S}}(\phi(M))$.*

Since every smooth function on M is of the form $g \circ \phi$ for some smooth g: $\mathbb{R}^d \to \mathbb{R}$, it is easy to see that the induced structures are independent of the choice of the proper embedding ϕ.

Independent of the proper embedding ϕ are also the notions of local convergence in $\hat{S}^\infty(\phi(M))$ and of local convergence both in $\hat{S}^\infty(\phi(M))$ and in $\hat{H}^\infty(\phi(M))$. This is of great importance in the sequel.

With a proper embedding ϕ, we can also define differentiability for families of processes in $\hat{D}_{\mathrm{c}}(M)$ (resp. $\hat{\mathscr{S}}(M)$). In this case, if $a \mapsto \phi(X(a))$ is differentiable at a_0 and Z is the derivative of $\phi(X(a))$ at a_0, then it is easy to verify that Z takes its values in $T\phi(TM)$ and the derivative of $X(a)$ at a_0 is the process $\partial_a X(a_0) = (T\phi)^{-1}(Z)$ with values in $\hat{D}_{\mathrm{c}}(TM)$ (resp. $\hat{\mathscr{S}}(TM)$).

Let $\hat{\mathscr{M}}_\nabla(M)$ be the set of continuous martingales with lifetime in $\hat{D}_{\mathrm{c}}(M)$. By [E4 4.43], $\hat{\mathscr{M}}_\nabla(M)$ is closed in $\hat{D}_{\mathrm{c}}(M)$. This implies that it is also closed in $\hat{\mathscr{S}}(M)$.

PROPOSITION 2.10. — *On $\hat{\mathscr{M}}_\nabla(M)$, the topology of compact convergence in probability and the topology of semimartingales coincide.*

To establish this result, we need some lemmas.

LEMMA 2.11. — *Every point x of M has a compact neighbourhood V, contained in the domain of a chart h, together with a smooth convex function ψ: $V\times V\to\mathbb{R}_+$ which satisfies the following conditions*:

1) *For all $x,y\in V$, $\psi(x,y)=0$ if and only if $x=y$,*

2) *There exists a constant $c>0$ such that for all $(X,Y)\in T_xM\times T_yM$, $x,y\in V$, with coordinates $\bar X=dh(X)$, $\bar Y=dh(Y)\in\mathbb{R}^d$,*

$$(\nabla\otimes\nabla)d\psi(x,y)\big((X,Y),(X,Y)\big)\ge c\,\|\bar Y-\bar X\|^2,$$

3) *For every Riemannian metric δ on V there exists a constant $A>0$ such that $\psi\le A\,\delta^2$.*

It is proven in [K] that convex geometry (the existence of a convex function ψ satisfying 1)) implies that every V-valued martingale has almost surely a limit at infinity.

Proof. — We show that the function ψ defined in [E4 4.59] has the desired properties. For $x_0\in M$, take an exponential chart (h,V) centered at x_0, and define

$$\psi(x,y)=\frac{1}{2}\left(\varepsilon^2+\|h(x)+h(y)\|^2\right)\|h(x)-h(y)\|^2.$$

Note that ψ satisfies 1) and 3). It is proven in [E4 4.59] that, if V is sufficiently small, one can choose $\varepsilon>0$ and $0<\beta<1$ such that if $U=(U_1,U_2)\in TV\oplus TV$ is a tangent vector with coordinates $(\bar X,\bar Y)\in\mathbb{R}^d\oplus\mathbb{R}^d$ where $\bar X=dh(U_1)$, $\bar Y=dh(U_2)$, then

$$\begin{aligned}(\nabla\otimes\nabla)d\psi(U,U)&\ge(1-\beta)\left(\varepsilon^2\|\bar X-\bar Y\|^2+\|h(x)-h(y)\|^2\|\bar X+\bar Y\|^2\right)\\&\ge(1-\beta)\,\varepsilon^2\,\|\bar X-\bar Y\|^2.\end{aligned}$$

This gives 2). ∎

LEMMA 2.12. — *Let V, δ be as in Lemma 2.11. There exists a constant $C>0$ such that if Y and Z are V-valued martingales, $h(Y)=(Y^1,\dots,Y^d)$ and $h(Z)=(Z^1,\dots,Z^d)$ in coordinates, then*

$$\mathbb{E}\left[\sum_{i=1}^d <Y^i-Z^i,Y^i-Z^i>_\infty\right]\le C\,\mathbb{E}\left[\delta^2(Y_\infty,Z_\infty)\right].$$

REMARK. — In particular, applying this result with a constant Z, we deduce that the expectation of the quadratic Riemannian variation of Y is bounded by a constant independent of Y.

Proof of Lemma 2.12. — Let ψ be as in Lemma 2.11. The Itô formula applied to ψ and (Y,Z) gives

$$\begin{aligned}\psi(Y_\infty,Z_\infty)=\psi(Y_0,Z_0)&+\int_0^\infty\left\langle d\psi,d^{\nabla\otimes\nabla}(Y,Z)\right\rangle\\&+\frac{1}{2}\int_0^\infty(\nabla\otimes\nabla)d\psi(Y,Z)\big(d(Y,Z)\otimes d(Y,Z)\big)\end{aligned}$$

where $d^{\nabla\otimes\nabla}$ denotes the Itô differential with respect to the product connection in $M\times M$. Using the fact that (Y,Z) is a martingale, we obtain

$$\mathbb{E}\left[\psi(Y_\infty,Z_\infty)\right]=\mathbb{E}\left[\psi(Y_0,Z_0)\right]+\frac{1}{2}\mathbb{E}\left[\int_0^\infty(\nabla\otimes\nabla)d\psi(Y,Z)\big(d(Y,Z)\otimes d(Y,Z)\big)\right],$$

hence by 2) and 3) of Lemma 2.11, we have

$$A\,\mathbb{E}\left[\delta^2(Y_\infty,Z_\infty)\right]\ge\mathbb{E}\left[\psi(Y_\infty,Z_\infty)\right]\ge\frac{c}{2}\,\mathbb{E}\left[\sum_{i=1}^d <Y^i-Z^i,Y^i-Z^i>_\infty\right].$$

This gives the result, with $C=2A/c$. ∎

Proof of Proposition 2.10. — We may assume that M is a closed subset of $\mathbb{R}^d$, and have to show that every sequence $(X^n)_{n\in\mathbb{N}}$ of ∇-martingales converging in $\hat{D}_{\mathrm{c}}(M)$ to a ∇-martingale X converges in $\hat{\mathscr{S}}(M)$ to the same limit. By means of the second equality of Proposition 2.4 with $p=1$, it is sufficient to prove the existence of a subsequence which converges to X locally in $H^1(\mathbb{R}^d,\infty)$. Since we are allowed to extract subsequences and since we have to prove only local convergence, by using the second equality of Proposition 2.4 with $p=\infty$, we may assume that $(X^n)_{n\in\mathbb{N}}$ converges to X in $S^\infty(\mathbb{R}^d,\infty)$. Still using the fact that it is sufficient to prove local convergence, we may further assume the existence of a finite increasing sequence of stopping times such that if S and T are two consecutive times in this sequence, then on $[S,T[$ all the $(X^n)_{n\in\mathbb{N}}$ and X take values in a compact set V as considered in Lemma 2.11. Finally, since the sequence of stopping times is finite, it is sufficient to prove convergence on one of the intervals $[S,T[$. Hence we assume that $(X^n)_{n\in\mathbb{N}}$ is a sequence of V-valued ∇-martingales converging to X in $S^\infty(\mathbb{R}^d,\infty)$, and it is sufficient to prove its convergence to X locally in $H^1(\mathbb{R}^d,\infty)$.

Since we are dealing with martingales, the finite variation parts of the coordinates satisfy

$$d\widetilde{(X^n)}^i=-\frac{1}{2}\sum_{j,k=1}^d\Gamma^i_{jk}(X^n)\,d<(X^n)^j,(X^n)^k>,$$

$$d\widetilde{X}^i=-\frac{1}{2}\sum_{j,k=1}^d\Gamma^i_{jk}(X)\,d<X^j,X^k>$$

where Γ^i_{jk} are the Christoffel symbols of the connection. This gives the bound

$$\begin{aligned}\|X^n-X\|_{H^1(\mathbb{R}^d,\infty)}\le\mathbb{E}\Bigg[\sum_{i=1}^d\big|(X^n)^i_0-X^i_0\big|&\sum_{i=1}^d<(X^n)^i-X^i,(X^n)^i-X^i>^{1/2}_\infty\\&+\sum_{i,j,k=1}^d\int_0^\infty\Bigg(\big|\Gamma^i_{jk}(X^n)-\Gamma^i_{jk}(X)\big|\,\big|d<X^j,X^k>\big|\\&+\big|\Gamma^i_{jk}(X^n)\big|\,\Big(\big|d<(X^n)^j-X^j,(X^n)^k>\big|+\big|d<X^j,(X^n)^k-X^k>\big|\Big)\Bigg)\Bigg].\end{aligned}$$

The Christoffel symbols are Lipschitz on V, hence by dominated convergence, $\sum_{i,j,k=1}^{d}\int_0^\infty \left|\Gamma_{jk}^i(X^n)-\Gamma_{jk}^i(X)\right|\left|d{<}X^j,X^k{>}\right|$ converges to 0 almost surely, and still by dominated convergence and the remark after Lemma 2.12, its expectation converges to 0. Since the Christoffel symbols are bounded on V, the last terms can be bounded by

$$C\,\mathbb{E}\left[\sum_{i,j=1}^{d}{<}(X^n)^i-X^i,(X^n)^i-X^i{>}_\infty^{1/2}\left({<}(X^n)^j,(X^n)^j{>}_\infty^{1/2}+{<}X^j,X^j{>}_\infty^{1/2}\right)\right]$$

with a constant $C>0$. Using Hölder's inequality and uniform boundedness of the expectations of the quadratic variations of V-valued martingales, we are led to show that $\mathbb{E}\left[\sum_{i=1}^{d}{<}(X^n)^i-X^i,(X^n)^i-X^i{>}_\infty\right]$ converges to 0. But, by means of Lemma 2.12,

$$\mathbb{E}\left[\sum_{i=1}^{d}{<}(X^n)^i-X^i,(X^n)^i-X^i{>}_\infty\right]\le C\,\mathbb{E}\left[\delta^2(X_\infty^n,X_\infty)\right]$$

with a constant $C>0$, and this gives the result. ∎

3. Regularity of solutions of stochastic differential equations

Let M and N be connected smooth manifolds. In this section, we will study stability of second order stochastic differential equations of the type

$$\mathcal{D}Z=f(X,Z)\,\mathcal{D}X \tag{3.1}$$

where $f\in\Gamma(\tau(M)^*\otimes\tau(N))$ is a Schwartz morphism, X belongs to $\hat{\mathscr{S}}(M)$ and Z to $\hat{\mathscr{S}}(N)$.

Remark. — If P is a submanifold of $M\times N$ such that the canonical projection $P\to M$ is a surjective submersion, and if f is only defined on P and constrained to P (see [E3]), then one can extend f in a smooth way to $M\times N$, and one knows that a solution (X,Z) of (3.1) with $(X_0,Z_0)\in P$ will stay on P.

Proposition 3.1. — *Let $(X^n)_{n\in\mathbb{N}}$ be a sequence of elements in $\hat{\mathscr{S}}(M)$ converging to X in $\hat{\mathscr{S}}(M)$, let $(Z_0^n)_{n\in\mathbb{N}}$ be a sequence of N-valued random variables converging to Z_0 in probability, and let $(f^n)_{n\in\mathbb{N}}$ be a sequence of locally Lipschitz Schwartz morphisms in $\Gamma(\tau(M)^*\otimes\tau(N))$ with uniform Lipschitz constant on compact sets, converging to a Schwartz morphism $f\in\Gamma(\tau(M)^*\otimes\tau(N))$. If Z^n is the maximal solution starting from Z_0^n to $\mathcal{D}Z^n=f^n(X^n,Z^n)\,\mathcal{D}X^n$, then (X^n,Z^n) converges in $\hat{\mathscr{S}}(M\times N)$ to (X,Z) where Z is the solution to $\mathcal{D}Z=f(X,Z)\,\mathcal{D}X$ starting from Z_0. Moreover, if ξ_{X^n} converges in probability to ξ_X then Z^n converges to Z in $\hat{\mathscr{S}}(N)$.*

Proof. — Let ξ_Z be the lifetime of Z. We will show that $(Z^n)^{\xi_Z-}$ converges to Z and that $\lim_{t\to\xi_Z} Z_t$ does not exist on $\{\xi_Z < \xi_X\}$, which is stronger than the results of Proposition 3.1. The second point is known, let us prove the first one. We have to show that there exists a stopping time T as close to ξ_Z as we want and a subsequence Z^{n_k} converging to Z. Hence we can assume that X^n, X take their values in a compact subset K_M, Z_0^n in a compact subset K_N and that X^n converge in $H^\infty(K_M,\infty)$ and in $S^\infty(K_M,\infty)$ to X. We can also assume that Z takes its values in K_N and has lifetime ∞. Consider Schwartz morphisms f_K^n, f_K satisfying the same convergence assumptions as f^n and f, with compact support K containing a neighbourhood of the product $K_M \times K_N$. Using the continuity results of Proposition 2.6 and [E2] theorem 0, we obtain that the solution Z_K^n of $\mathcal{D}Z_K^n = f_K^n(X^n, Z_K^n)\,\mathcal{D}X^n$ with $(Z_K^n)_0 = Z_0^n$ converge in $\mathscr{S}(N,\infty)$ to the solution Z_K of $\mathcal{D}Z_K = f_K(X, Z_K)\,\mathcal{D}X$ with $(Z_K)_0 = Z_0$. This implies that a subsequence converges locally in $\hat{H}^\infty(N)$ and in $\hat{S}^\infty(N)$, but then locally, for indices sufficiently large, the solutions to the truncated equation coincide with the solutions to the original equation. This gives the claim. $\square$

Immediate consequences of Proposition 3.1 are the following results.

Corollary 3.2. — 1) *Let $\Gamma^1(\tau(M)^*\otimes\tau(N))$ be the set of C^1 Schwartz morphisms endowed with the topology of uniform convergence on compact sets of the maps and their derivatives, and let $L^0(N)$ be the set of N-valued random variables endowed with the topology of convergence in probability. Then the map*

$$\hat{\mathscr{S}}(M) \times \Gamma^1(\tau(M)^* \otimes \tau(N)) \times L^0(N) \longrightarrow \hat{\mathscr{S}}(M \times N),$$

defined by $(X, f, Z_0) \mapsto (X, Z)$ with Z the maximal solution of $\mathcal{D}Z = f(X,Z)\,\mathcal{D}X$, is continuous.

2) *Let $\Gamma^1(\tau(M)^*)$ be the set of C^1 forms of order 2 endowed with the topology of uniform convergence on compact sets of the maps and their derivatives. Then the map*

$$\begin{aligned} \Gamma^1(\tau(M)^*) \times \hat{\mathscr{S}}(M) &\longrightarrow \hat{\mathscr{S}}(\mathbb{R}) \\ (\theta, X) &\longmapsto \int_0 \langle \theta(X), \mathcal{D}X \rangle \end{aligned}$$

is lower semicontinuous.

Example. — Here we give an example of a sequence of deterministic paths converging uniformly to a constant path, but such that parallel transports above the elements of this sequence do not converge. This shows in particular that in 1) we cannot replace the topology of semimartingales in $\hat{\mathscr{S}}(M)$ by the topology of compact convergence in probability, unless we restrict for instance to the sets of martingales with respect to a given connection.

Let M be a simply connected surface endowed with a rotationally invariant metric about $o \in M$, represented in polar coordinates as $ds^2 = dr^2 + g^2(r)\,d\vartheta^2$ for some smooth function g. Let $t \mapsto x(t) \in M$ be a path in M, defined on the unit interval $[0,1]$, with polar coordinates $r(t) \equiv \varepsilon$ and $\vartheta(t) = \alpha t$ for some $\alpha > 0$. A straightforward calculation shows that the rotational speed of a parallel transport above x in polar coordinates is $-\alpha g'(\varepsilon)$. Hence the rotational speed in an exponential chart with centre o which realizes an isometry at o is $\alpha(1 - g'(\varepsilon))$ (note that this gives 0 if the metric is flat).

In the following, M is taken to be an open subset of the sphere S^2. Thus we have $g(r) = \sin r$, and $\alpha(1-g'(\varepsilon)) = \alpha(1-\cos\varepsilon)$. Consider the sequence of paths $(x^n)_{n\in\mathbb{N}}$ defined in polar coordinates by $\vartheta^n(t) = 2\pi nt$ and $r^n(t) \equiv \varepsilon_n = \arccos(1-\frac{1}{2n})$ (hence $2\pi n(1-\cos\varepsilon_n) = \pi$). Since $\varepsilon_n \to 0$, we get uniform convergence of $(x^n)_{n\in\mathbb{N}}$ to the constant path o. But for all n, the rotation at time 1 of a parallel transport above x^n is π. Hence parallel transports above x^n do not converge to a parallel transport above o.

In the sequel we are seeking differentiability results. This requires some geometric preliminaries. We will use the maps ϕ: $M \times N \times M \to N$ defined by Cohen [C1] and [C2] to describe stochastic differential equations in manifolds with càdlàg semimartingales.

Definition 3.3. — *Let $k \in \mathbb{N}$. A Schwartz morphism $f \in \Gamma(\tau(M)^* \otimes \tau(N))$ (resp. a section $e \in \Gamma(TM^* \otimes TN)$) is said to be of class C^k_{Lip} if f (resp. e) is C^k with locally Lipschitz derivatives of order k.*

We say that a measurable map ϕ: $M \times N \times M \to N$ is of class $C^{k,\infty}_{\mathrm{Lip}}$ if there exists a neighbourhood of the submanifold $\{(x,z,x),\ (x,z) \in M \times N\}$ on which ϕ is C^∞ with respect to the third variable and all the derivatives with respect to this variable are C^k with locally Lipschitz derivatives of order k (with respect to the three variables).

Lemma (and Definition) 3.4. — *Let $k \in \mathbb{N}$. For every Schwartz morphism $f \in \Gamma(\tau(M)^* \otimes \tau(N))$ of class C^k_{Lip}, there exists a map ϕ: $M \times N \times M \to N$ of class $C^{k,\infty}_{\mathrm{Lip}}$ such that for all $(x,z) \in M \times N$*

$$f(x,z) = \tau_3\,\phi(x,z,x)$$

where $\tau_3\phi$ denotes the second order derivative of ϕ with respect to the third variable. Such a map ϕ will be called a Cohen map associated to f.

In particular, a Cohen map satisfies $\phi(x,z,x) = z$ for all $(x,z) \in M \times N$.

Proof. — First, we remark that it is sufficient to construct ϕ in a neighbourhood of the submanifold $\{(x,z,x),\ (x,z) \in M \times N\}$ and to extend it then in a measurable way to $M \times N \times M$.

Let ∇^M (resp. ∇^N) be a connection on M (resp. N). There exists a neighbourhood of the diagonal of $M \times M$ on which the maps $(x,z) \mapsto v(x,z) = \dot\gamma(0)$ and $(x,z) \mapsto u(x,z) = \ddot\gamma(0)$ are smooth, where γ is the geodesic such that $\gamma(0) = x$ and $\gamma(1) = z$. There exists a neighbourhood of the null section in TN on which the exponential map, denoted by $\exp^N$, is smooth. If $u \in \tau N$ is a second order vector, denote by $F(u) \in TN$ its first order part with respect to the connection ∇^N (see [E4] for the definition).

Thus there exists a neighbourhood V of $\{(x,y,x),\ (x,y) \in M \times N\}$ such that the map

$$\begin{aligned}\phi\colon V &\to N\\ (x,y,z) &\mapsto \exp^N\Big(f(x,y)\,v(x,z) + \frac{1}{2}F\big(f(x,y)\,u(x,z)\big)\Big)\end{aligned}$$

is defined and satisfies the regularity assumptions. We have to verify the equation $\tau_3\phi(x,y,x) = f(x,y)$. For this, it is sufficient to check that these maps coincide on elements of $\tau_x M$ of the form $\dot\gamma(0)$ and $\ddot\gamma(0)$ where γ is a geodesic with $\gamma(0) = x$ and $\gamma(1) = z$. A change of time gives

$$\phi(x,y,\gamma(t)) = \exp^N\Big(t\,f(x,y)\,v(x,z) + \frac{t^2}{2}F\big(f(x,y)\,u(x,z)\big)\Big)$$

Taking successively first and second order derivatives with respect to t at time 0 gives the result. ∎

THEOREM 3.5. — *Let $a \mapsto X(a)$ be C^1 from I to $\hat{\mathscr{S}}(M)$, let $f \in \Gamma(\tau(M)^* \otimes \tau(N))$ be a Schwartz morphism of class C^1_{Lip}, and $a \mapsto Z(a)$ the maximal solution of*

$$\mathcal{D}Z(a) = f\big(X(a), Z(a)\big)\,\mathcal{D}X(a) \tag{3.2}$$

where $a \mapsto Z_0(a)$ is C^1 in probability. Then the map $a \mapsto (X(a), Z(a))$ defined on I and with values in $\hat{\mathscr{S}}(M \times N)$ is C^1, and the process $\partial_a Z(a)$ is the maximal solution of

$$\mathcal{D}\partial_a Z(a) = f'\big(\partial_a X(a), \partial_a Z(a)\big)\,\mathcal{D}\partial_a X(a) \tag{3.3}$$

with initial condition $\partial_a Z_0(a)$ where f' is the Schwartz morphism of class C^0_{Lip} defined as follows: if $f(x,z) = \tau_3\,\phi(x,z,x)$ with a $C^{1,\infty}_{\mathrm{Lip}}$ Cohen map ϕ associated to f, then $f'(u,v) = \tau_3\,T\phi(u,v,u)$ for $(u,v) \in TM \times TN$, i.e., $T\phi$ is a $C^{0,\infty}_{\mathrm{Lip}}$ Cohen map associated to f'. If moreover $a \mapsto \xi_{X(a)}$ is continuous in probability, then $a \mapsto Z(a)$ is C^1 in $\hat{\mathscr{S}}(N)$.

REMARK. — If P is a submanifold of $M \times N$ such that the canonical projection $P \to M$ is a surjective submersion, and if f is only defined on P and is constrained to P, then one can show that f' is constrained to TP. As a consequence, by the remark at the beginning of this section, if $\big(\partial_a X_0(a), \partial_a Z_0(a)\big)$ belongs to TP, then $\big(\partial_a X(a), \partial_a Z(a)\big)$ takes its values in TP.

LEMMA 3.6. — *Let P, Q, R, S be manifolds, $\varphi\colon Q \to P$ and $\psi\colon R \to S$ maps, and let $\phi\colon Q \times R \times Q \to R$ and $\phi'\colon P \times S \times P \to S$ be Cohen maps such that $\phi' \circ (\varphi, \psi, \varphi) = \psi \circ \phi$. Then, for all $(x,y) \in Q \times R$, we have*

$$\tau_3\phi'\big(\varphi(x), \psi(y), \varphi(x)\big) \circ \tau\varphi(x) = \tau\psi(y) \circ \tau_3\phi(x,y,x).$$

If semimartingales X, Z take values in Q, resp. R, and satisfy the equation $\mathcal{D}Z = \tau_3\,\phi(X,Z,X)\,\mathcal{D}X$, then $U = \varphi(X)$ and $V = \psi(Z)$ satisfy

$$\mathcal{D}V = \tau_3\,\phi'(U,V,U)\,\mathcal{D}U.$$

Proof. — It is sufficient and easy to prove the first equality with second order derivatives of curves. The second equality is a consequence of the first one. ∎

Proof of Theorem 3.5. — Assume that $0 \in I$. Using Proposition 3.1, it is sufficient to prove that $a \mapsto (X(a), Z(a))$ is differentiable at $a = 0$ and that the derivative of $a \mapsto Z(a)$ is the maximal solution of (3.3).

Let ∇^M be a connection on M. There exists an open neighbourhood Δ_1^M of the diagonal Δ^M in $M \times M$ such that for $a \neq 0$ the function

$$\begin{aligned} \varphi_a^M \colon \Delta_1^M &\longrightarrow U_a^M := \varphi_a^M(\Delta_1) \\ (x,y) &\longmapsto \frac{1}{a}(\exp_x^N)^{-1} y \end{aligned}$$

is well-defined and a diffeomorphism. The same objects on N are denoted with the superscript N. Let ϕ be a $C^{1,\infty}_{\mathrm{Lip}}$ Cohen map associated to f. It is easy to see that

$$\mathcal{D}(Z(0), Z(a)) = \tau_3\,(\phi,\phi)\big((X(0),X(a)),(Z(0),Z(a)),(X(0),X(a))\big)\,\mathcal{D}\big(X(0),X(a)\big).$$

Let $T^M(a)$ be the exit time of $(X(0), X(a))$ of Δ_1^M, $\widetilde{X}(a) = (X(0), X(a))^{T^M(a)-}$ and $\widetilde{Z}(a)$ be the maximal solution to

$$\mathcal{D}\widetilde{Z}(a) = \tau_3\,(\phi,\phi)\big(\widetilde{X}(a),\widetilde{Z}(a),\widetilde{X}(a)\big)\big)\,\mathcal{D}\widetilde{X}(a)$$

with initial condition $(Z_0(0), Z_0(a))$. Let then $T^N(a)$ be the exit time of $\widetilde{Z}(a)$ of Δ_1^N. Using Proposition 2.6, it is easy to see that $T^M(a) \wedge \xi_{X(0)}$ converges in probability to $\xi_{X(0)}$ as a tends to 0, and then that $T^N(a) \wedge \xi_{Z(0)}$ converges in probability to $\xi_{Z(0)}$.

By Lemma 3.6, defining $\widetilde{V}(a) = \varphi_a^N\big(\widetilde{Z}(a)^{T^N(a)-}\big)$ and $\widetilde{Y}(a) = \varphi_a^M\big(\widetilde{X}(a)^{T^N(a)-}\big)$ for $a \neq 0$, we have that $\widetilde{V}(a)$ is the maximal solution in $\mathscr{S}(TM)$ of

$$\mathcal{D}\widetilde{V}(a) = \tau_3\big(\varphi_a^N \circ (\phi,\phi) \circ \big((\varphi_a^M)^{-1},(\varphi_a^N)^{-1},(\varphi_a^M)^{-1}\big)\big)\big(\widetilde{Y}(a),\widetilde{V}(a),\widetilde{Y}(a)\big)\,\mathcal{D}\widetilde{Y}(a)$$

with initial condition $\widetilde{V}_0(a) = \varphi_a^N\big(Z_0(0), Z_0(a)\big)$ on $\{(Z_0(0), Z_0(a)) \in \Delta_1^N\}$. For $u \in U_a^M$ (resp. $u \in U_a^N$), denote by $\ell_a^M(u)$ (resp. $\ell_a^N(u)$) the second coordinate of $(\varphi_a^M)^{-1}(u)$ (resp. $(\varphi_a^N)^{-1}(u)$). Then the mapping

$$(a,u,v,w) \mapsto \begin{cases} \varphi_a^N\big(\phi(\pi(u),\pi(v),\pi(w)), \phi\big(\ell_a^M(u),\ell_a^N(v),\ell_a^M(w)\big)\big) & \text{if } a \neq 0, \\ T\phi(u,v,w) & \text{if } a = 0, \end{cases}$$

defined on an open subset of $(-1,1) \times TM \times TN \times TM$ containing the elements of the form $(0,u,v,u)$ with $(u,v) \in TM \times TN$, depends C^∞ on the last variable and its derivatives with respect to this variable are locally Lipschitz (as functions of all four variables). This implies the convergence of $\tau_3\big(\varphi_a^N \circ (\phi,\phi) \circ \big((\varphi_a^M)^{-1},(\varphi_a^N)^{-1},(\varphi_a^M)^{-1}\big)\big)$ to $\tau_3 T\phi$ as $a \to 0$, and the existence of uniform Lipschitz constants on compact sets. Since $a \mapsto X(a)$ is differentiable at $a = 0$, $T^M(a) \wedge \xi_{X(0)}$ converges in probability to $\xi_{X(0)}$ and $T^N(a) \wedge \xi_{Z(0)}$ converges in probability to $\xi_{Z(0)}$, we have that $\widetilde{Y}(a)$ converges to $Y(0)^{\xi_{Z(0)}}$ with $Y(0) := \partial_a X(0)$; on the other side, $\widetilde{V}_0(a)$ converges in probability to $\partial_a Z_0(0) = V_0(0)$ on $\{\xi_Z(0) > 0\}$; hence we get by Proposition 3.1 that $(\widetilde{Y}(a), \widetilde{V}(a))$ converges to $(Y(0), V(0))$ where $V(0)$ is the maximal solution of

$$\mathcal{D}V(0) = \tau_3\,T\phi\big(Y(0),V(0),Y(0)\big)\,\mathcal{D}Y(0)$$

with initial condition $V_0(0) = \partial_a Z_0(0)$. This implies that $a \mapsto (X(a), Z(a))$ is differentiable at $a = 0$ and that its derivative is $(Y(0), V(0))$. $\square$

We now want to investigate Stratonovich and Itô equations. In the following, if $(t,a) \mapsto x(t,a)$ is a map defined on an open subset of $\mathbb{R}^2$ and with values in a manifold M, $\dot{x}(t,a)$ will denote its derivative with respect to t, and $\ddot{x}(t,a)$ will denote the second order tangent vector such that for all smooth function g on M, $\ddot{x}(t,a)(g) = \partial_t^2(g \circ x)(t,a)$. For a smooth function g on M, $d^2 g$ will denote the second order form defined by $\langle d^2 g, \ddot{x}(t,a)\rangle = \ddot{x}(t,a)(g)$ (see [E4]).

LEMMA 3.7. — *Let J, I be two intervals in $\mathbb{R}$. Suppose that $(t,a) \mapsto x(t,a) \in M$ and $(t,a) \mapsto z(t,a) \in N$ are $C^{2,1}$ maps defined on $J \times I$, and satisfy for each a*

$$\ddot{z}(0,a) = \tau_3\, \phi\big(x(0,a), z(0,a), x(0,a)\big)\, \ddot{x}(0,a) \tag{3.4}$$

where ϕ: $M \times N \times M \to N$ is a $C^{1,\infty}_{\mathrm{Lip}}$ Cohen map. Then

$$(\partial_a z)^{\cdot\cdot}(0,a) = \tau_3\, T\phi\big(\partial_a x(0,a), \partial_a z(0,a), \partial_a x(0,a)\big)(\partial_a x)^{\cdot\cdot}(0,a).$$

Proof. — It is sufficient to prove

$$\big\langle d^2\ell, (\partial_a z)^{\cdot\cdot}(0,a)\big\rangle = \big\langle d^2\ell, \tau_3\, T\phi\big(\partial_a x(0,a), \partial_a z(0,a), \partial_a x(0,a)\big)(\partial_a x)^{\cdot\cdot}(0,a)\big\rangle \tag{3.5}$$

and

$$\big\langle d\ell, (\partial_a z)^{\cdot}(0,a)\big\rangle^2 = \big\langle d\ell, T_3\, T\phi\big(\partial_a x(0,a), \partial_a z(0,a), \partial_a x(0,a)\big)(\partial_a x)^{\cdot}(0,a)\big\rangle^2 \tag{3.6}$$

for $\ell = g \circ \pi$: $TN \to \mathbb{R}$ and $\ell = dg$: $TN \to \mathbb{R}$ where g: $N \to \mathbb{R}$ is smooth. Equations (3.5) and (3.6) for $\ell = g \circ \pi$: $TN \to \mathbb{R}$, $g \in C^\infty(N, \mathbb{R})$ are direct consequences of assumption (3.4). To establish (3.5) for $\ell = dg$: $TN \to \mathbb{R}$, we define $z'(t,a) = \phi\big(x(0,a), z(0,a), x(t,a)\big)$. Then

$$\begin{aligned}
\langle d^2\ell, (\partial_a z)^{\cdot\cdot}(0,a)\rangle &= \partial_t^2 \partial_a (g \circ z)(0,a) = \partial_a \partial_t^2 (g \circ z)(0,a) = \partial_a \big\langle d^2 g, \ddot{z}(0,a)\big\rangle \\
&= \partial_a \big\langle d^2 g, \tau_3\, \phi\big(x(0,a), z(0,a), x(0,a)\big)\ddot{x}(0,a)\big\rangle \\
&= \partial_a \big\langle d^2 g, (z')^{\cdot\cdot}(0,a)\big\rangle = \partial_a \partial_t^2 (g \circ z')(0,a) = \partial_t^2 \partial_a (g \circ z')(0,a) \\
&= \partial_t^2 dg \circ T\phi\big(\partial_a x(0,a), \partial_a z(0,a), \partial_a x(t,a)\big) \\
&= \big\langle d^2\ell, \tau_3\, T\phi\big(\partial_a x(0,a), \partial_a z(0,a), \partial_a x(0,a)\big)\, (\partial_a x)^{\cdot\cdot}(0,a)\big\rangle.
\end{aligned}$$

Finally, to verify (3.6) for $\ell = dg$: $TN \to \mathbb{R}$, we have to prove that

$$\big(\partial_t \partial_a (g \circ z)(0,a)\big)^2 = \big(\partial_t \partial_a (g \circ z')(0,a)\big)^2.$$

We first derive from (3.4) that

$$\big(\partial_t (g \circ z)(0,a)\big)^2 = \big(\partial_t (g \circ z')(0,a)\big)^2,$$

and by taking the square of the derivative with respect to a,

$$\big(\partial_t (g \circ z)(0,a)\big)^2 \big(\partial_t \partial_a (g \circ z)(0,a)\big)^2 = \big(\partial_t (g \circ z')(0,a)\big)^2 \big(\partial_t \partial_a (g \circ z')(0,a)\big)^2.$$

Let $a_0 \in I$. If $\big(\partial_t (g \circ z)(0,a_0)\big)^2 \neq 0$, equality (3.6) is satisfied for $a = a_0$. Now consider the case $\big(\partial_t (g \circ z)(0,a_0)\big)^2 = 0$. If $\big(\partial_t \partial_a (g \circ z)(0,a_0)\big)^2 \neq 0$ or $\big(\partial_t \partial_a (g \circ z')(0,a_0)\big)^2 \neq 0$, then we have $\big(\partial_t (g \circ z)(0,a)\big)^2 \neq 0$ in a neighbourhood of a_0 (a_0 excepted) and (3.6) is satisfied for $a = a_0$ by continuity. ∎

Definition 3.8. — *A Cohen map $\phi\colon M\times N\times M\to N$ is said to be a Cohen map of Stratonovich type if in addition it has the following property: if a C^2 curve (γ,α) in $M\times N$ satisfies $\dot\alpha(t)=T_3\phi\big(\gamma(t),\alpha(t),\gamma(t)\big)\,\dot\gamma(t)$ then $\ddot\alpha(t)=\tau_3\,\phi\big(\gamma(t),\alpha(t),\gamma(t)\big)\,\ddot\gamma(t)$.*

Proposition 3.9. — *Let $k\geq 1$ and e be a $C^k_{\rm Lip}$ section of the vector bundle $T^*M\times TN$ over $M\times N$. Then there exists a $C^{k-1,\infty}_{\rm Lip}$ Cohen map ϕ of Stratonovich type such that $e(x,z)=T_3\phi(x,z,x)$ for all $(x,z)\in M\times N$. If ϕ is a $C^{k,\infty}_{\rm Lip}$ Cohen map of Stratonovich type, then $T\phi\colon TM\times TN\times TM\to TN$ is a $C^{k-1,\infty}_{\rm Lip}$ Cohen map of Stratonovich type.*

Proof. — The existence of ϕ of class $C^{k-1,\infty}_{\rm Lip}$ is a consequence of [E3 Theorem 8], which gives the existence of a unique Schwartz morphism of Stratonovich type f of class $C^{k-1}_{\rm Lip}$ associated to e, together with Lemma 3.4.

Let ϕ be a $C^{k,\infty}_{\rm Lip}$ Cohen map of Stratonovich type; we want to show that $T\phi$ is also a Cohen map of Stratonovich type. Let $t\mapsto\beta(t)$ be a smooth curve with values in TN and $t\mapsto\delta(t)$ a smooth curve with values in TM such that

$$\dot\beta(t)=T_3T\phi\big(\delta(t),\beta(t),\delta(t)\big)\,\dot\delta(t). \tag{3.7}$$

We have to prove that

$$\ddot\beta(t)=\tau_3\,T\phi\big(\delta(t),\beta(t),\delta(t)\big)\,\ddot\delta(t).$$

This will be done by means of Lemma 3.7. More precisely, let $(t,a)\mapsto x(t,a)$ satisfy $\partial_a x(t,0)=\delta(t)$, and let $(t,a)\mapsto z(t,a)\in M$ be a solution of

$$\dot z(t,a)=T_3\phi\big(x(t,a),z(t,a),x(t,a)\big)\,\dot x(t,a) \tag{3.8}$$

with the property $\partial_a z(0,0)=\beta(0)$. It is easy to verify that $\beta(t)=\partial_a z(t,0)$ then already for all t, by exploiting uniqueness of solutions to (3.7) with given initial conditions and by calculating $\langle dh,(\partial_a z)^{\boldsymbol\cdot}(t,0)\rangle$ for $h=dg$ and $h=g\circ\pi$ where $g\colon N\to\mathbb{R}$ is smooth. Since ϕ is a Cohen map of Stratonovich type, together with equation (3.8), we get from Lemma 3.7

$$(\partial_a z)^{\boldsymbol{\cdot\cdot}}(t,a)=\tau_3\,T\phi\big(\partial_a x(t,a),\partial_a z(t,a),\partial_a x(t,a)\big)\,(\partial_a x)^{\boldsymbol{\cdot\cdot}}(t,a)$$

which can be rewritten for $a=0$ as

$$\ddot\beta(t)=\tau_3\,T\phi\big(\delta(t),\beta(t),\delta(t)\big)\,\ddot\delta(t).$$

This proves that $T\phi$ is indeed a Cohen map of Stratonovich type. $\square$

Rephrased in terms of Cohen maps of Stratonovich type, the following result is a consequence of [E3 Theorem 8].

Proposition 3.10. — *Let $k \geq 1$ and let e be a C^k_{Lip} section of $T^*M \times TN$ over $M \times N$. Let ϕ be a $C^{k-1,\infty}_{\mathrm{Lip}}$ Cohen map satisfying $e(x,z) = T_3\phi(x,z,x)$. The equations $\delta Z = T_3\phi(X,Z,X)\,\delta X$ and $\mathcal{D}Z = \tau_3\,\phi(X,Z,X)\,\mathcal{D}X$ are equivalent if and only if ϕ is a Cohen map of Stratonovich type.*

For the rest of this section we assume that both M and N are endowed with connections ∇^M and ∇^N. On the tangent bundles TM and TN we consider the corresponding complete lifts $(\nabla^M)'$ and $(\nabla^N)'$ of these connections (see [Y-I] for a definition).

We will say that a Schwartz morphism $f \in \Gamma(\tau(M)^* \otimes \tau(N))$ is semi-affine if for every ∇^M-geodesic γ with values in M and defined at time 0, for every $y \in N$, $f(\gamma(0), y)\,\ddot\gamma(0)$ is the second derivative of a ∇^N-geodesic in N (see [E3] for details). In fact $f(\gamma(0), y)\,\ddot\gamma(0)$ is the second order derivative $\ddot\alpha(0)$ of the geodesic α which satisfies $\alpha(0) = y$ and $\dot\alpha(0) = f(\gamma(0), \alpha(0))\,\dot\gamma(0)$.

Definition 3.11. — *We say that a Cohen map ϕ is a Cohen map of Itô type (with respect to the connections ∇^M and ∇^N) if $\tau_3\,\phi(x,z,x)\colon \tau_x M \to \tau_z N$ is a semi-affine morphism.*

Proposition 3.12. — *Let $k \geq 0$ and let e be a C^k_{Lip} section of $T^*M \times TN$ over $M \times N$. There exists a $C^{k,\infty}_{\mathrm{Lip}}$ Cohen map ϕ of Itô type such that $e(x,z) = T_3\phi(x,z,x)$ for all $(x,z) \in M \times N$. If $k \geq 1$ and ϕ is a $C^{k,\infty}_{\mathrm{Lip}}$ Cohen map of Itô type, then $T\phi$ is a $C^{k-1,\infty}_{\mathrm{Lip}}$ Cohen map of Itô type (with respect to the connections $(\nabla^M)'$ and $(\nabla^N)'$).*

Proof. — The existence of ϕ is a consequence of [E3 Lemma 11] which gives the existence of a unique Schwartz morphism of Itô type associated to e, together with Lemma 3.4.

Let ϕ be a Cohen map of Itô type; we want to show that $T\phi$ is also a Cohen map of Itô type. We have to prove that for all $(y_0, v_0) \in TM \times TN$, $\tau_3\,T\phi(y_0, v_0, y_0)$ is semi-affine, i.e., if $t \mapsto y(t)$ is a $(\nabla^M)'$-geodesic in TM with $y(0) = y_0$, then the $(\nabla^N)'$-geodesic $t \mapsto v(t)$ in TN with $\dot v(0) = T_3T\phi(y_0, v_0, y_0)\,\dot y(0)$ satisfies $\ddot v(0) = \tau_3\,T\phi(y_0, v_0, y_0)\,\ddot y(0)$.

Let $(t,a) \mapsto x(t,a) \in M$ satisfy $\partial_a|_{a=0}\,x(t,a) = y(t)$ and such that $t \mapsto x(t,a)$ is a ∇^M-geodesic for all a. Note that this is possible because y is a Jacobi field. Let $(t,a) \mapsto z(t,a) \in N$ be such that for all a, $t \mapsto z(t,a)$ is a ∇^N-geodesic with

$$\dot z(0,a) = T_3\phi\big(x(0,a), z(0,a), x(0,a)\big)\,\dot x(0,a)$$

and $\partial_a|_{a=0}\,z(0,a) = v(0)$. Since $t \mapsto x(t,a)$ and $t \mapsto z(t,a)$ are geodesics and $T_3\phi(x,z,x)$ is semi-affine, we deduce that

$$\ddot z(0,a) = \tau_3\,\phi\big(x(0,a), z(0,a), x(0,a)\big)\,\ddot x(0,a).$$

Now we can apply Lemma 3.7 to obtain

$$(\partial_a z)^{\cdot\cdot}(0,a) = \tau_3\,T\phi(\partial_a x(0,a), \partial_a z(0,a), \partial_a x(0,a))\,(\partial_a x)^{\cdot\cdot}(0,a). \tag{3.9}$$

It remains to prove that $(\partial_a z)^{\cdot\cdot}(0,0)=\ddot v(0)$ and $(\partial_a x)^{\cdot\cdot}(0,0)=\ddot y(0)$. But $t\mapsto \partial_a z(t,a)$ and $t\mapsto \partial_a x(t,a)$ are geodesics for $(\nabla^N)'$ and $(\nabla^M)'$, hence it is sufficient to know that $(\partial_a z)^{\cdot}(0,0)=\dot v(0)$ and $(\partial_a x)^{\cdot}(0,0)=\dot y(0)$ (the last equality is already known). For this, we want to calculate $\langle dh,(\partial_a z)^{\cdot}(0,a)\rangle$ for $h=dg$ and $h=g\circ\pi$ with $g\colon N\to\mathbb{R}$ smooth (we will do the verification only for $h=dg$). Let $h=dg$, then

$$\begin{aligned}\langle dh,(\partial_a z)^{\cdot}(0,a)\rangle &= \partial_t|_{t=0}\big\langle dg,\partial_a z(t,a)\big\rangle\\ &= \partial_t|_{t=0}\partial_a(g\circ z)(t,a)=\partial_a\partial_t|_{t=0}(g\circ z)(t,a)\\ &= \partial_a\partial_t|_{t=0}(g\circ\phi)\big(x(0,a),z(0,a),x(t,a)\big)\\ &= \partial_t|_{t=0}\partial_a(g\circ\phi)\big(x(0,a),z(0,a),x(t,a)\big)\\ &= \partial_t|_{t=0}\big\langle dg,T\phi\big(\partial_a x(0,a),\partial_a z(0,a),\partial_a x(t,a)\big)\big\rangle\\ &= \big\langle dh,T_3T\phi\big(\partial_a x(0,a),\partial_a z(0,a),\partial_a x(0,a)\big)(\partial_a x)^{\cdot}(0,a)\big\rangle.\end{aligned}$$

In particular, for $a=0$, this gives

$$\big\langle dh,(\partial_a z)^{\cdot}(0,0)\big\rangle=\big\langle dh,T_3T\phi\,(y_0,v_0,y_0)\,\dot y(0)\big\rangle.$$

Since $\dot v(0)=T_3T\phi(y_0,v_0,y_0)\,\dot y(0)$ we obtain $\dot v(0)=(\partial_a z)^{\cdot}(0,0)$ which finally gives with (3.9)

$$\ddot v(0)=(\partial_a z)^{\cdot\cdot}(0,0)=\tau_3\,T\phi(y_0,v_0,y_0)\,\ddot y(0).$$

This proves that $T\phi$ is a Cohen map of Itô type. ∎

Rewritten with Cohen maps of Itô type, we get the following result as a consequence of [E3 Theorem 12].

Proposition 3.13. — *Let $k\ge 0$ and e be a C^k_{Lip} section of $T^*M\times TN$ over $M\times N$. Let ϕ be a $C^{k,\infty}_{\mathrm{Lip}}$ Cohen map satisfying $e(x,z)=T_3\phi(x,z,x)$ for all $(x,z)\in M\times N$. Then the equations $d^{\nabla^N}Z=T_3\phi(X,Z,X)\,d^{\nabla^M}X$ and $\mathcal{D}Z=\tau_3\,\phi(X,Z,X)\,\mathcal{D}X$ are equivalent if and only if ϕ is a Cohen map of Itô type.*

The main motivation in our study of Cohen maps of Stratonovich and Itô type is the following result.

Corollary 3.14. — 1) *Let $k\ge 0$ and e be a C^{k+1}_{Lip} section of the vector bundle $T^*M\times TN$ over $M\times N$. Assume that $a\mapsto X(a)$ is C^k in $\hat{\mathscr{S}}(M)$, and $a\mapsto Z(a)$ is the maximal solution of*

$$\delta Z(a)=e\big(X(a),Z(a)\big)\,\delta X(a)\tag{3.10}$$

where $a\mapsto Z_0(a)$ is C^k in probability. Then $a\mapsto (X(a),Z(a))$ is C^k in $\hat{\mathscr{S}}(M\times N)$, and if $k\ge 1$, the derivative $\partial_a Z(a)$ is the maximal solution of

$$\delta\partial_a Z(a)=e'\big(\partial_a X(a),\partial_a Z(a)\big)\,\delta\partial_a X(a)\tag{3.11}$$

with initial condition $\partial_a Z_0(0)$ where e' is the C^k_{Lip} section of $T^*TM\times TTN$ over $TM\times TN$ defined as follows: if $e(x,z)=T_3\phi(x,z,x)$ with a $C^{k+1,\infty}_{\mathrm{Lip}}$ Cohen map ϕ then $e'(u,v)=T_3T\phi(u,v,u)$ for $(u,v)\in TM\times TN$. If moreover $a\mapsto \xi_{X(a)}$ is continuous in probability, then $a\mapsto Z(a)$ is C^k in $\hat{\mathscr{S}}(N)$.

2) Let $k \geq 0$ and e be a C^k_{Lip} section of the vector bundle $T^*M \times TN$ over $M \times N$. Assume that M (resp. N) is endowed with a connection ∇^M (resp. ∇^N), and denote by $(\nabla^M)'$ (resp. $(\nabla^N)'$) the complete lift of ∇^M (resp. ∇^N) in TM (resp. TN). Assume that $a \mapsto X(a)$ is C^k in $\hat{\mathscr{S}}(M)$, $a \mapsto Z(a)$ is the maximal solution of

$$d^{\nabla^N} Z(a) = e\big(X(a), Z(a)\big)\, d^{\nabla^M} X(a) \tag{3.12}$$

where $a \mapsto Z_0(a)$ is C^k in probability. Then $a \mapsto (X(a), Z(a))$ is C^k in $\hat{\mathscr{S}}(M \times N)$, and the derivative $\partial_a Z(a)$ is the maximal solution of

$$d^{(\nabla^N)'} \partial_a Z(a) = e'\big(\partial_a X(a), \partial_a Z(a)\big)\, d^{(\nabla^M)'} \partial_a X(a) \tag{3.13}$$

with initial condition $\partial_a Z_0(0)$ where e' is the C^{k-1}_{Lip} section of $T^*TM \times TTN$ over $TM \times TN$ defined in 1). If moreover $a \mapsto \xi_{X(a)}$ is continuous in probability, then $a \mapsto Z(a)$ is C^k in $\hat{\mathscr{S}}(N)$.

Remark. — We like to stress the pleasant point that both Stratonovich and Itô equations differentiate like equations involving smooth paths.

Proof of Corollary 3.14. — 1) We only have to consider the case $k \geq 1$. Let ϕ be a $C^{k,\infty}_{\mathrm{Lip}}$ Cohen map of Stratonovich type such that $T_3\phi(x, z, x) = e(x, z)$ for all $(x, z) \in M \times N$. By Proposition 3.10, equation (3.10) is equivalent to

$$\mathcal{D} Z(a) = \tau_3\, \phi\big(X(a), Z(a), X(a)\big)\, \mathcal{D} X(a).$$

Applying Theorem 3.5, we can differentiate with respect to a and we get

$$\mathcal{D} \partial_a Z(a) = \tau_3\, T\phi\big(\partial_a X(a), \partial_a Z(a), \partial_a X(a)\big)\, \mathcal{D} \partial_a X(a). \tag{3.14}$$

But by Proposition 3.9, $T\phi$ is a $C^{k-1,\infty}_{\mathrm{Lip}}$ Cohen map of Stratonovich type, and again by Proposition 3.10, equation (3.14) is equivalent to

$$\delta \partial_a Z(a) = T_3 T\phi\big(\partial_a X(a), \partial_a Z(a), \partial_a X(a)\big)\, \delta \partial_a X(a)$$

which is precisely equation (3.11).

2) The proof of 1) carries over verbatim, replacing Stratonovich by Itô, Proposition 3.10 by Proposition 3.13, and Proposition 3.9 by Proposition 3.12. ∎

We want to rephrase equation (3.13) in terms of covariant derivatives. For this we need some definitions and lemmas. Let R^M denote the curvature tensor of the connection ∇^M on M, which is assumed here to be torsion-free. If J is a semimartingale with values in TM endowed with the horizontal lift $(\nabla^M)^h$ of ∇^M (see [Y-I] for a definition), let DJ denote its covariant derivative, i.e. the projection of the vertical part of $d^{(\nabla^M)^h} J$, thus $DJ = v_J^{-1}(d^{(\nabla^M)^h} J)^v$ with $v_j\colon T_xM \to T_jTM$ denoting the vertical lift for $j \in T_xM$. We observe that also $DJ = /\!/_{0,\cdot}\, d(/\!/_{0,\cdot}^{-1} J)$ where $/\!/_{0,t}$ means parallel translation along $\pi(J)$. Indeed, this equality is verified if J is a smooth curve, and since by [Y-I] (9.2) p. 114, J is a geodesic if and only if $(\pi(J), /\!/_{0,\cdot}^{-1} J)$ is a geodesic in $M \times T_{\pi(J_0)}M$ for the product connection, using [E3] corollary 16, it extends to semimartingales as an Itô equation.

LEMMA 3.15. — *Let J be a semimartingale with values in TM, and $X = \pi(J)$ its projection to M. Then*

$$d^{(\nabla^M)'} J = d^{(\nabla^M)^h} J + \frac{1}{2} v_J \left(R^M(J, dX)dX\right) \tag{3.15}$$

where $v_j(u)$ is the vertical lift above $j \in T_xM$ of a vector $u \in T_xM$.

Proof. — Following [E3], if $\widetilde{\nabla}$ is a connection on TM, the Itô differential $d^{\widetilde{\nabla}} J$ may be written as $F^{\widetilde{\nabla}}(\mathcal{D}J)$ where $F^{\widetilde{\nabla}} \colon \tau TM \to TTM$ is the projection defined as follows: if A and B are vector fields on TM, then $F(A) = A$ and $F(AB) = \frac{1}{2}\big(\widetilde{\nabla}_A B + \widetilde{\nabla}_B A + [A, B]\big)$. The result is a direct consequence of the following Lemma. ∎

For $\ell \in \tau M$, let $b(\ell) \in TM \odot TM$ denote its symmetric bilinear part, i.e., $\langle df \otimes dg, b(\ell)\rangle = \frac{1}{2}\big[\ell(fg) - f\,\ell(g) - g\,\ell(f)\big]$ for f, g smooth functions on M.

LEMMA 3.16. — *Let L be an element of $\tau_u TM$ with $u \in T_xM$. Then*

$$\Big(F^{(\nabla^M)'} - F^{(\nabla^M)^h}\Big)(L) = v_u\Big(\big(R^M(u, \cdot\,)\,\cdot\,\big) b(\pi_* L)\Big)$$

where $\pi_* \colon \tau TM \to \tau M$ is induced by $\pi \colon TM \to M$.

Proof. — It is sufficient to prove this for $L_u = (AB)_u$ with A and B horizontal or vertical vector fields. But since among these possibilities $(\nabla^M)'_A B$ and $(\nabla^M)^h_A B$ coincide except if both A and B are horizontal, we can restrict to this case. Let A (resp. B) be the horizontal lift of $\bar{A}$ (resp. $\bar{B}$). Then by [Y-I],

$$(\nabla^M)'_{A_u} B - (\nabla^M)^h_{A_u} B = v_u\left(R^M(u, \bar{A}_x)\bar{B}_x\right)$$

where $x = \pi(u)$, and this gives the result, since $b(\pi_* L_u) = \frac{1}{2}\left(\bar{A}_x \otimes \bar{B}_x + \bar{B}_x \otimes \bar{A}_x\right)$. ∎

COROLLARY 3.17. — *Let $k \geq 0$ and e be a C^k_{Lip} section of the vector bundle $T^*M \times TN$ over $M \times N$. Assume that M (resp. N) is endowed with a torsion-free connection ∇^M (resp. ∇^N). Assume that $a \mapsto X(a)$ is C^k in $\hat{\mathscr{S}}(M)$, $a \mapsto Z(a)$ the maximal solution of*

$$d^{\nabla^N} Z(a) = e\big(X(a), Z(a)\big)\, d^{\nabla^M} X(a) \tag{3.16}$$

where $a \mapsto Z_0(a)$ is C^k in probability. Then $a \mapsto \big(X(a), Z(a)\big)$ is C^k in $\hat{\mathscr{S}}(M \times N)$, and the derivative $\partial_a Z(a)$ is the maximal solution of the covariant stochastic differential equation

$$\begin{aligned} D\partial_a Z &= e(X, Z)\, D\partial_a X + \nabla e(\partial_a X, \partial_a Z)\, d^{\nabla^M} X \\ &\quad + \frac{1}{2}\Big(e(X, Z) R^M\big(\partial_a X, dX\big)dX - R^N\big(\partial_a Z, e(X, Z)dX\big)e(X, Z)dX\Big) \end{aligned} \tag{3.17}$$

with initial condition $\partial_a Z_0(0)$. If moreover $a \mapsto \xi_{X(a)}$ is continuous in probability, then $a \mapsto Z(a)$ is C^k in $\hat{\mathscr{S}}(N)$.

REMARKS. — 1) If ∇^M and ∇^N are allowed to have torsion, one can write first a covariant equation of the form (3.17) with respect to the symmetrized connections $\bar\nabla^M$ and $\bar\nabla^N$. With the obvious notations, expressing $\bar D\partial_a X$, $\bar D\partial_a X$, $\bar R^M$ and $\bar R^N$ in terms of $D\partial_a X$, $D\partial_a X$, R^M, R^N and the torsion tensors, one obtains then a covariant equation with respect to ∇^M and ∇^N.

2) Starting with (3.11) in Corollary 3.14, one can also easily determine a Stratonovich covariant equation, identical to the equation for smooth processes.

Proof of Corollary 3.17. — Applying Lemma 3.15 to part 2) of Corollary 3.14 gives the following equation for $\partial_a Z$:

$$d^{(\nabla^N)^h}\partial_a Z = e'\big(\partial_a X, \partial_a Z\big)\left(d^{(\nabla^M)^h}\partial_a X + \frac{1}{2}v_{\partial_a X}\left(R^M(\partial_a X, dX)dX\right)\right) - \frac{1}{2}v_{\partial_a Z}\big(R^N(\partial_a Z, dZ)dZ\big).$$

But, if $u, w \in T_xM$, $z \in TN$, we have $e'(u,z)v_u(w) = v_z\left(e\big(\pi(u),\pi(z)\big)w\right)$, and by definition, if $h_u^{\nabla^M}(w) \in T_uTM$ is the horizontal lift of w, then $v_z\big(\nabla e(u,z)w\big)$ is the vertical part of $e'(u,z)h_u^{\nabla^M}(w)$. These equalities applied to $u = \partial_a X$, $z = \partial_a Z$, and successively to $w = D\partial_a X$, $w = d^{\nabla^M}X$ and $w = \frac{1}{2}R^M(\partial_a X, dX)dX$, give the desired equation. ∎

As an application of Corollary 3.14, we get differentiability results for stochastic integrals, considered as particular instances of stochastic differential equations:

COROLLARY 3.18. — 1) *Let $k \ge 0$ and α be a C^{k+1}_{Lip} section of the vector bundle T^*M. Assume that $a \mapsto X(a)$ is C^k in $\hat{\mathscr{S}}(M)$.*

Then $a \mapsto \left(X(a), \int_0^\cdot \langle\alpha(X(a)), \delta X(a)\rangle\right)$ *is C^k in $\hat{\mathscr{S}}(M\times\mathbb{R})$.*

2) *Let $k \ge 0$ and α be a C^k_{Lip} section of the vector bundle $T^*M \times TN$ over $M \times N$. Assume that M (resp. N) is endowed with a connection ∇^M. Assume that $a \mapsto X(a)$ is C^k in $\hat{\mathscr{S}}(M)$.*

Then $a \mapsto \left(X(a), \int_0^\cdot \langle\alpha(X(a)), d^{\nabla^M}X(a)\rangle\right)$ *is C^k in $\hat{\mathscr{S}}(M\times\mathbb{R})$.*

4. Application to antidevelopment

If M is a manifold, we will denote by $s\colon TTM \to TTM$ the following canonical isomorphism: if $(t,a) \mapsto x(t,a)$ is a smooth M-valued map defined on some open subset of $\mathbb{R}^2$, then $\partial_t\partial_a x(t,a) = s\left(\partial_a\partial_t x(t,a)\right)$.

THEOREM 4.1. — *Let M be a manifold endowed with a connection ∇. Denote by ∇' the complete lift of ∇ on TM. Let $\mathcal{A}'$ denote the antidevelopment with respect to ∇'. Let $a \mapsto X(a) \in \hat{\mathscr{S}}(M)$ be a map of class C^1 defined on some interval I of $\mathbb{R}$. Then $a \mapsto \big(X(a), \mathcal{A}(X(a))\big) \in \hat{\mathscr{S}}(TM\times TM)$ is of class C^1 and*

$$s\big(\partial_a\mathcal{A}(X(a))\big) = \mathcal{A}'\big(\partial_a X(a)\big).$$

Moreover, if $a \mapsto \xi_{X(a)}$ is continuous in probability, then $a \mapsto \mathcal{A}(X(a))$ is C^1 in $\hat{\mathscr{S}}(TN)$.

Before proving this result we introduce some definitions and lemmas. Let M be a manifold of dimension m. If ∇ is a connection on M, we consider the complete lift ∇' of ∇ on TM, which is characterized by the relation $\nabla'_{X^c}Y^c = (\nabla_X Y)^c$ valid for all vector fields $X, Y \in \Gamma(TM)$. Here X^c denotes the complete lift of X, i.e. the vector field in $\Gamma(TTM)$ defined by $X^c_u = s\big(TX(u)\big)$ (see [Y-I] for details). Recall that the geodesics for ∇' are the Jacobi fields for ∇.

Let $L(M)$ be the principal bundle of linear frames on TM: thus $L_x(M)$ is the set of linear isomorphisms $\mathbb{R}^m \to T_xM$ for each $x \in M$. There is a canonical embedding $\jmath\colon TL(M) \to L(TM)$ defined as follows: if $W \in TL(M)$ is equal to $(\partial_a U)(0)$ where $a \mapsto U(a)$ is a smooth path in $L(M)$, and if $v \in T\mathbb{R}^m = \mathbb{R}^{2m}$ is equal to $(\partial_a z)(0)$ where $a \mapsto z(a)$ is a smooth path in $\mathbb{R}^m$, then one has $\jmath(W)v = s\left((\partial_a(Uz))(0)\right)$. Let $(e_1, \ldots, e_m, e_{\bar 1}, \ldots, e_{\bar n})$ be the standard basis of $T\mathbb{R}^n$. Then $\jmath((\partial_a U)(0))e_\alpha = s\left((\partial_a(Ue_\alpha))(0)\right)$ and $\jmath((\partial_a U)(0))e_{\bar\alpha}$ is the vertical lift $(Ue_\alpha)^v(0)$ of $(Ue_\alpha)(0)$ above $\partial_a(\pi \circ U)(0)$ where $\pi\colon L(M) \to M$ is the canonical projection.

LEMMA 4.2. — *If $a \mapsto X(a) \in \hat{\mathscr{S}}(M)$ is C^1 and $U(a) \in \hat{\mathscr{S}}(L(M))$ is a horizontal lift of $X(a)$ such that $a \mapsto U_0(a)$ is C^1 in probability, then $a \mapsto U(a)$ is C^1 in $\hat{\mathscr{S}}(L(M))$ and $\jmath(\partial_a U(a))$ is a horizontal lift of $\partial_a X(a)$ with respect to the connection ∇'.*

Proof. — The fact that $a \mapsto U(a)$ is C^1 is a direct consequence of Corollary 3.14 and the fact that $\big\{\lim_{t\to\xi_{U(a)}} U_t(a) \text{ exists}\big\}$ is included in $\big\{\lim_{t\to\xi_{X(a)}} X_t(a) \text{ exists}\big\}$. Another consequence of Corollary 3.14 is that it suffices to prove the assertion with both $X_t(a) = x(t,a)$ and $a \mapsto U_0(a) = u(0,a)$ deterministic and smooth. Write $u(t,a) = U_t(a)$.

It is sufficient to verify that for all $i \in \{1, \ldots, m\}$, $t \mapsto \jmath(\partial_a u(t,a))\, e_i$ and $t \mapsto \jmath(\partial_a u(t,a))\, e_{\bar\imath}$ are parallel transports. But by [Y-I chapt. I, prop. 6.3], we have

$$\nabla'_{(\partial_a x)^{\cdot}(t,a)} s\big(\partial_a(u(t,a)e_i)\big) = s\left(\partial_a\big(\nabla_{\dot x(t,a)} u(t,a)e_i\big)\right) = 0$$

and

$$\nabla'_{(\partial_a x)^{\cdot}(t,a)} \big(u(t,a)e_i\big)^v = \big(\nabla_{\dot x(t,a)} u(t,a)e_i\big)^v = 0.$$

This proves Lemma 4.2. ∎

Proof of Theorem 4.1. — The fact that $a \mapsto \big(X(a), \mathcal{A}(X(a))\big)$ is C^1 is a consequence of Corollary 3.14. We can calculate as if dealing with smooth deterministic paths. Let $a \mapsto U_0(a) \in L_{X_0(a)}(M)$ be C^1 in probability and denote by $U(a)$ the parallel transport of $U_0(a)$ along $X(a)$. Write $Z(a) = \mathcal{A}(X(a))$. Then we have the equation

$$U(a)U_0^{-1}(a)\, p\big(\delta Z(a)\big) = \delta X(a)$$

where if $z \in T_xM$ and $v \in T_zTM$ is a vertical vector, $p(v)$ denotes its canonical projection onto T_xM.

Denoting by m the dimension of M, we define a family of $\mathbb{R}^m$-valued processes $R(a)$ by $R(a) = U_0^{-1}(a)Z(a)$, hence

$$\delta R(a) = U_0^{-1}(a)\, p\big(\delta Z(a)\big). \tag{4.1}$$

We have $U(a)\delta R(a) = \delta X(a)$ and by differentiation with respect to a, using the definition of $\jmath$,

$$\jmath\big(\partial_a U(a)\big)\, s\big(\delta\partial_a R(a)\big) = \delta\partial_a X(a). \tag{4.2}$$

On the other hand, differentiation of (4.1) gives

$$\jmath\big(\partial_a U_0(a)\big)\, s\big(\delta\partial_a R(a)\big) = p'\left(\delta s\big(\partial_a Z(a)\big)\right) \tag{4.3}$$

where p' on the vertical vectors to TTM is defined like p. Putting together (4.2) and (4.3) gives

$$\jmath\big(\partial_a U(a)\big)\left(\jmath\big(\partial_a U_0(a)\big)\right)^{-1} p'\left(\delta s\big(\partial_a Z(a)\big)\right) = \delta\partial_a X(a).$$

But $\jmath\big(\partial_a U(a)\big)$ is the parallel transport of $\jmath(U_0(a))$ above $\partial_a X(a)$ by Lemma 4.2, hence $s\big(\partial_a Z(a)\big) = \mathcal{A}'\left(\partial_a X(a)\right)$. ☐

COROLLARY 4.3. — *Let J be a TM-valued semimartingale with lifetime $\xi = \xi(0)$. There exists a C^1 family $(X(a))_{a\in\mathbb{R}}$ of elements in $\hat{\mathscr{S}}(M)$ such that the equality $J = \partial_a X(0)$ is satisfied. In particular, if $\xi(a)$ is the lifetime of $X(a)$, then $\xi(a)\wedge\xi(0)$ converges in probability to $\xi(0)$ as a tends to 0. The semimartingale J is a ∇'-martingale if and only if one can choose $(X(a))_{a\in\mathbb{R}}$ such that $X(a)$ is a ∇-martingale for each $a \in \mathbb{R}$.*

Proof. — With the notations of Theorem 4.1 define $V = \mathcal{A}'(J)$, and for $a \in \mathbb{R}$,

$$Z(a) = T\exp\Big(s\big(as(V)\big)\Big).$$

Note that the lifetime of $Z(a)$ can be 0 if $\exp aJ_0$ is not defined. A straightforward calculation shows that $s\big(\partial_a Z(0)\big) = V$. Define now $X(a)$ as the stochastic development of $Z(a)$. We have the relation $Z(a) = \mathcal{A}\big(X(a)\big)$; by Corollary 3.14, the map $a \mapsto (Z(a), X(a))$ is C^∞ in $\hat{\mathscr{S}}(M)$ and in particular, $\xi(a)\wedge\xi(0)$ converges in probability to $\xi(0)$ as a tends to 0. By Theorem 4.1, the antidevelopment of the derivative of $a \mapsto X(a)$ at $a = 0$ is $s\big(\partial_a Z(0)\big) = V$. This implies that $\partial_a X(0) = J$.

If J is a martingale, then V is a local martingale (with possibly finite lifetime). It is easy to see that for each $a \in \mathbb{R}$, $Z(a)$ is also a local martingale, hence its development $X(a)$ is a martingale. ☐

References

[C1] Cohen (S.) — *Géométrie différentielle stochastique avec sauts 1*, Stochastics and Stochastics Reports, t. **56**, 1996, p. 179–203.

[C2] Cohen (S.) — *Géométrie différentielle stochastique avec sauts 2: discrétisation et applications des EDS avec sauts*, Stochastics and Stochastics Reports, t. **56**, 1996, p. 205–225.

[E1] Emery (M.) — *Une topologie sur l'espace des semimartingales*, Séminaire de Probabilités XIII, Lecture Notes in Mathematics, Vol 721, Springer, 1979, p. 260–280.

[E2] Emery (M.) — *Equations différentielles stochastiques lipschitziennes: étude de la stabilité*, Séminaire de Probabilités XIII, Lecture Notes in Mathematics, Vol 721, Springer, 1979, p. 281–293.

[E3] Emery (M.) — *On two transfer principles in stochastic differential geometry*, Séminaire de Probabilités XXIV, Lecture Notes in Mathematics, Vol 1426, Springer, 1989, p. 407–441.

[E4] Emery (M.) — *Stochastic calculus in manifolds.* — Springer, 1989.

[K] Kendall (W.S.) — *Convex geometry and nonconfluent* Γ*-martingales II: Well-posedness and* Γ*-martingale convergence*, Stochastics, t. **38**, 1992, p. 135–147.

[Pi] Picard (J.) — *Convergence in probability for perturbed stochastic integral equations*, Probability Theory and Related Fields, t. **81**, 1989, p. 383–451.

[Pr] Protter (P.) — *Stochastic Integration and Differential Equations. A New Approach.* — Springer, 1990.

[Y-I] Yano (K.), Ishihara (S.) — *Tangent and Cotangent Bundles.* — Marcel Dekker, Inc., New York, 1973.