ANNALES DE L'I. H. P., SECTION B

SERGE COHEN
ANNE ESTRADE

# Non-symmetric approximations for manifold-valued semimartingales

*Annales de l'I. H. P., section B*, tome 36, nº 1 (2000), p. 45-70
<http://www.numdam.org/item?id=AIHPB_2000__36_1_45_0>

*Ann. Inst. Henri Poincaré, Probabilités et Statistiques* 36, 1 (2000) 45–70


# Non-symmetric approximations for manifold-valued semimartingales

by

**Serge COHEN** [a,1], **Anne ESTRADE** [b,2]

[a] Université de Versailles – St Quentin en Yvelines, 45, avenue des Etats-Unis, 78035 Versailles, France or CERMICS-ENPC

[b] MAPMO – UMR 6628 Département de mathématiques, Université d'Orléans, B.P.6759, 45067 Orléans Cedex 2, France

Article received on 23 June 1998, revised 23 March 1999

ABSTRACT. – We study general approximations of continuous semimartingales in a manifold. Classically the limits of integrals with respect to the approximated semimartingales yield Stratonovich integrals. Nevertheless several authors have remarked that a skew-symmetric extra-term may appear for specific approximations when the manifold is a vector space. We give the geometric meaning of the skew-symmetric term and an interpretation in term of a "second order non-symmetric intrinsic calculus". This stochastic non-symmetric calculus is further extended to stochastic differential equations between manifolds. A particular emphasis is pointed on the role of interpolators in approximations. © 2000 Éditions scientifiques et médicales Elsevier SAS

*Key words:* Approximation of semimartingales, regularization of stochastic differential equations, stochastic calculus on manifolds

*AMS classification*: Primary 58F30, secondary 58C35, 60H05, 60H10

RÉSUMÉ. – Notre but est l'étude de l'approximation de semimartingales continues sur des variétés. Quand on passe à la limite pour un

[1] E-mail: cohen@math.uvsq.fr.

[2] E-mail: estrade@labomath.univ-orleans.fr.

procédé approximant, il est classique d'obtenir l'intégrale de Stratonovich d'une forme différentielle. Dans le cas où la variété est un espace vectoriel plusieurs auteurs ont remarqué l'apparition d'un terme antisymétrique pour certaines approximations particulières. D'un point de vue géométrique ce terme correspond à un crochet antisymétrique de semimartingales. C'est le pendant antisymétrique du crochet oblique usuel. Nous proposons ainsi une généralisation du calcul stochastique d'ordre 2, en oubliant toutes les conditions de symétries qui existaient dans ce calcul. De plus nous appliquons ces nouveaux outils à l'approximation d'équations différentielles stochastiques entre variétés. Un intérêt particulier est porté aux approximations de semimartingales issues de règles d'interpolation. 

## INTRODUCTION

The approximation of stochastic processes and its relationship with integral calculus is an old and fascinating subject. Semimartingales that are often used to define stochastic integrals, are somehow sensitive to approximation: The limit of $\int X^\delta \, dY^\delta$ when $\delta$ goes to 0 depends on the approximation of both processes $X$ and $Y$. One of the most famous stochastic calculus, the Stratonovich calculus, can be viewed as the limit of such approximated integrals when a linear interpolation of a discrete sample of the driving process $Y$ is used to construct $Y^\delta$. If we take the Stratonovich calculus as the starting point of our study we can make two remarks.

- The Stratonovich calculus can be also presented by adding to the Ito calculus a correction term which is related to the quadratic variation of semimartingales. In this paper we consider the approximation associated to the linear interpolation as symmetric because the correction term is symmetric in $(X, Y)$.
- Although Stratonovich calculus is very popular for applications since the seminal work of Wong and Zakai [17], it yields also very important consequences for stochastic differential geometry since it obeys the ordinary change of variable formula. For instance we shall recall the *transfer principle*: most geometrical constructions involving smooth curves formally extend to semimartingales via Stratonovich integrals.

Hence our first aim has been to find the geometrical nature of different methods of approximations proposed for vector valued semimartingales, where the limit cannot be expressed as a Stratonovich integral ([4,6, 10,11]). One can also find a very general study of approximations of stochastic differential equation in [15]. Actually the common feature of theses examples is that the limit of $\int X^\delta \, dY^\delta$ involves a skew-symmetric term which has to be added to the usual Stratonovich correction.

In Section 1 we show that an intrinsic skew-symmetric bracket can be constructed for manifold-valued approximated semimartingale. In many respects the skew-symmetric bracket behaves like its symmetric counterpart, the $b$-quadratic variation introduced by [3]. But we shall always remember that the skew-symmetric bracket depends on the approximation rule used to transform semimartingales into finite variation processes whereas the $b$-quadratic variation does not need additional stucture to be defined. To understand better the skew-symmetric bracket we study its relationship with interpolation rules. Actually we propose a very natural definition of a general interpolation which is a smooth family of paths $\{I(x, y, t), t \in [0, 1]\}$ indexed by every pair of points $(x, y)$ of the manifold and we show how to compute the skew-symmetric bracket associated to the general interpolation rule. At this point it explains why various extra assumptions are given in [1,3] which are only useful to ensure convergence to Stratonovich integrals.

In Section 2 we look at the influence of non-symmetric approximation of the driving semimartingale on the integral of a 1-form. We then introduce a second order calculus that takes into account the non-symmetric part of the approximation rules and which extends the second order calculus of [3,12,16]. Moreover this general second order calculus yields a probabilistic interpretation of the "Leibnitz" differentiation operator $d_2$ which has been recently reintroduced by [13].

The last section of this article is devoted to the approximation of stochastic differential equation between manifolds. Once again we prove that a correction term has to be added to the usual Stratonovich limit equation, when the driving semimartingale is non-symmetrically approximated. From a technical point of view the approximation theorems of [10] or of [4] are used when the manifolds are embedded in some vector spaces. In particular we consider the stochastic exponential of a Lie group which can be defined by a stochastic differential equation between the Lie algebra and the corresponding Lie group. As an example, we take the case of the Heisenberg group and we compute the solution of the corrected limit equation.

In this article we have tried to keep the presentation as elementary as possible and hence we have used embeddings of manifolds in vector spaces to write most of the equations, but all the notions we introduce are of course independent of the choice of embeddings.

Let us now introduce precisely the geometrical setting and its relation with stochastic calculus.

By a manifold $M$ we shall always mean a finite dimensional $C^\infty$-manifold which can be properly embedded into $\mathbf{R}^m$ for some positive $m$ (by Whitney's embedding theorem, this will be done as soon as $M$ admits a countable atlas). Then, even when not specified, $M$ will be endowed with an embedding, denoted by a family $(x^i)_{1\leqslant i\leqslant m}$ of maps from $M$ to $\mathbf{R}$. In particular any point $p$ in $M$, when considered as embedded in $\mathbf{R}^m$, is written $(p^i)_{1\leqslant i\leqslant m}$ and any smooth function $f$ on $M$ has the form $f(p)=\tilde{f}(p^1,\ldots,p^m)$ for a smooth $\tilde{f}$ on $\mathbf{R}^m$ which we still denote by $f$. When we are concerned with two manifolds $M$ and $N$, in order to avoid confusion, we use $(x^i)_{1\leqslant i\leqslant m}$ as embedding of $M$ into $\mathbf{R}^m$ and $(y^\alpha)_{1\leqslant \alpha\leqslant n}$ as embedding of $N$ into $\mathbf{R}^n$.

Moreover each embedding of $M$ induces a distance $d_M$ on $M$. Although two distances induced by two embeddings are not equal, they are equivalent on compact sets and so we can make use of anyone of those distances to produce estimates.

Manifold-valued continuous processes will play the crucial role. In this article we fix a filtered probability space $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geqslant 0},\mathbf{P})$. An $M$-valued process defined on $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geqslant 0},\mathbf{P})$ is called a semimartingale if for any smooth $f$ from $M$ to $\mathbf{R}$, $f(X)$ is a real-valued semimartingale. An $M$-valued process $X$ is said to have finite variation on bounded intervals if for any positive $T$ and any sequence $(\tau^n)_{n\in\mathbf{N}}$ of partitions of $[0,T]$ whose mesh $|\tau^n|$ goes to 0,

$$\sup_n\Big(\sum_{t_k^n\in\tau^n} d_M(X_{t_k^n},X_{t_{k+1}^n})\Big)<+\infty. \tag{1}$$

In this case, the left hand side of (1) defines the $d_M$-total variation of $X$ on $[0,T]$ denoted by $\int_0^T|dX_s|$. Since the uniform convergence in probability is used througout, we remind the reader of its definition. A family of $M$-valued continuous processes $(X^\delta)_{\delta>0}$ is said to converge in probability uniformly in $t$ on bounded intervals if $\forall\varepsilon,t\geqslant 0$

$$\lim_{\delta\to 0}\mathbf{P}\Big(\sup_{s\leqslant t} d_M(X_s^\delta,X_s)>\varepsilon\Big)=0.$$

For more details about stochastic calculus in manifolds we refer to Emery's book [3], in particular for the definition of Ito and Stratonovich integrals along $M$-valued continuous semimartingales. Most of our notations are similar to this book. The right bracket $[\,,\,]$ will always be used with its geometric meaning: Lie bracket for vector fields, and never with its stochastic meaning. When refering to the quadratic variation of two real-valued continuous semimartingales, we use the angle bracket $\langle\,,\,\rangle$.

At last, Einstein summation convention will be used throughout.

## 1. THE SKEW-SYMMETRIC BRACKETS

### 1.1. Approximation of semimartingales in a manifold

Consider an $M$-valued continuous semimartingale $X$ on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbf{P})$. Even when $M$ is a vector space, every reasonable definition for an approximation family $(X^\delta)_{\delta>0}$ of $X$ includes some conditions (see [4,6,10]): a pointwise convergence in probability and a convergence of paths strong enough to allow convergence of stochastic integrals. Those conditions appear in the following definition.

DEFINITION 1.1. – *By an approximation of $X$ we mean a family $(X^\delta)_{\delta>0}$ of $M$-valued continuous processes with finite variation on bounded intervals such that*:

(A0) *for all $\delta > 0$, the process $X^\delta = (X^\delta_t, t \geqslant 0)$ is adapted to the filtration $(\mathcal{F}_t)_{t\geqslant 0}$*;

(A1) *as $\delta$ goes to 0, $X^\delta_t$ goes to $X_t$ in probability uniformly in $t$ on bounded intervals*;

(A2) *for all $t \geqslant 0$, $\int_0^t d_M(X_s, X^\delta_s)|dX^\delta_s|$ is bounded in probability uniformly in $\delta$*;

(A3) *there exist an embedding $(x^i)_{1\leqslant i\leqslant m}$ of $M$ in $\mathbf{R}^m$ and a $m \times m$-dimensional continuous process $(S^{ij})_{1\leqslant i,j\leqslant m}$ with finite variation on bounded intervals such that, as $\delta$ goes to* 0,

$$S^{\delta,ij}_t := \int_0^t \left(X^i_s - X^{\delta,i}_s\right) dX^{\delta,j}_s - \frac{1}{2}\langle X^i, X^j\rangle_t \tag{2}$$

*goes to $S^{ij}_t$, in probability uniformly in $t$ on bounded intervals.*

*Remarks*. – 1. Condition (A2) on $(X^\delta)_{\delta>0}$ means precisely that the family $(S^\delta)_{\delta>0}$ satisfies the "U.T."-criterion of [7], or the "C2-2(ii)"-

condition of [9]. It is exactely the condition which ensures, with (A3), the convergence of any integral along $S^\delta$ toward the same integral along $S$: for all $b = (b_{ij})_{1 \leqslant i,j \leqslant m} \in \mathcal{C}(M)$,

$$\lim_{\delta \to 0} \int_0^{\cdot} b_{ij}(X)\, dS^{\delta,ij} = \int_0^{\cdot} b_{ij}(X)\, dS^{ij} \tag{3}$$

in probability uniformly on bounded intervals.

2. By Ito's formula, for all $t \geqslant 0$ and $\delta > 0$,

$$S_t^{\delta,ij} + S_t^{\delta,ji} = (X_0^i - X_0^{\delta,i})(X_0^j - X_0^{\delta,j}) - (X_t^i - X_t^{\delta,i})(X_t^j - X_t^{\delta,j}) + \int_0^t (X_s^i - X_s^{\delta,i})\, dX_s^j + \int_0^t (X_s^j - X_s^{\delta,j})\, dX_s^i$$

which goes to 0 as $\delta$ tends to 0. Therefore the matrices $S_t = (S_t^{ij})_{1 \leqslant i,j \leqslant m}$ of (A3) above are skew-symmetric for all $t \geqslant 0$. This remark explains why the term "$-\frac{1}{2}\langle X^i, X^j \rangle$" appears in Eq. (2).

Condition (A3) involves a specific embedding of $M$ into $\mathbf{R}^m$. The next proposition proves that if this condition is realized for one embedding then it is realized for all.

PROPOSITION 1.2. – *Suppose $(X^\delta)_{\delta>0}$ is an approximation of $X$ which satisfies condition* (A3) *with embedding $(x^i)_{1 \leqslant i \leqslant m}$ of $M$ into $\mathbf{R}^m$ and limit process $(S^{ij})_{1 \leqslant i,j \leqslant m}$. Let $(\tilde{x}^k)_{1 \leqslant k \leqslant m}$ be any other embedding of $M$ into $\mathbf{R}^m$.*

*Then $(X^\delta)_{\delta>0}$ satisfies condition* (A3) *with embedding $(\tilde{x}^k)_{1 \leqslant k \leqslant m}$ and the limit process $(\widetilde{S}^{kl})_{1 \leqslant k,l \leqslant m}$ is given by*

$$\widetilde{S}_t^{kl} = \int_0^t \frac{\partial \phi^k}{\partial x^i}(X_s) \frac{\partial \phi^l}{\partial x^j}(X_s)\, dS_s^{ij}, \tag{4}$$

*where $\phi$ is a smooth diffeomorphism of $\mathbf{R}^m$ such that*

$$(\tilde{x}^k)_{1 \leqslant k \leqslant m} = \phi\big((x^i)_{1 \leqslant i \leqslant m}\big).$$

*Proof.* – The diffeomorphism $\phi$ has bounded first and second derivatives on compact subsets of $\mathbf{R}^m$.

Formula (2) applied with the new embedding $(\tilde{x}^k)_{1 \leqslant k \leqslant m}$ gives

$$\widetilde{S}_t^{kl,\delta} = \int_0^t \left(\phi^k(X_s) - \phi^k(X_s^\delta)\right) \frac{\partial \phi^l}{\partial x^j}(X_s^\delta)\, dX_s^{\delta,j}$$
$$- \frac{1}{2} \int_0^t \frac{\partial \phi^k}{\partial x^i}(X_s) \frac{\partial \phi^l}{\partial x^j}(X_s)\, d\langle X^i, X^j\rangle_s.$$

On the other hand, by formulas (2) and (3) with the original embedding,

$$\int_0^t \frac{\partial \phi^k}{\partial x^i}(X_s) \frac{\partial \phi^l}{\partial x^j}(X_s)\, dS_s^{ij}$$

is the limit as $\delta$ goes to 0 of

$$\int_0^t \frac{\partial \phi^k}{\partial x^i}(X_s) \frac{\partial \phi^l}{\partial x^j}(X_s)\left( (X_s^i - X_s^{\delta,i})\, dX_s^{\delta,j} - \frac{1}{2} d\langle X^i, X^j\rangle_s \right).$$

Then, to prove the lemma and formula (4), we have to prove that

$$\Delta_t^\delta := \int_0^t \Bigg[ \left(\phi^k(X_s) - \phi^k(X_s^\delta)\right) \frac{\partial \phi^l}{\partial x^j}(X_s^\delta)$$
$$- \frac{\partial \phi^k}{\partial x^i}(X_s) \frac{\partial \phi^l}{\partial x^j}(X_s)\left(X_s^i - X_s^{\delta,i}\right) \Bigg]\, dX_s^{\delta,j} \tag{5}$$

goes to 0 in probability uniformly on bounded intervals. This will be done using Taylor expansion which enables us to write

$$\left| \phi^k(X_s) - \phi^k(X_s^\delta) - \frac{\partial \phi^k}{\partial x^i}(X_s)\left(X_s^i - X_s^{\delta,i}\right) \right| \leqslant \frac{1}{2} \|D^2\phi\|_\infty d_M(X_s, X_s^\delta)^2$$

and

$$\left| \frac{\partial \phi^l}{\partial x^j}(X_s^\delta) - \frac{\partial \phi^l}{\partial x^j}(X_s) \right| \leqslant \|D^2\phi\|_\infty d_M(X_s, X_s^\delta).$$

Putting those inequalities in (5) gives

$$|\Delta_t^\delta| \leqslant \frac{3}{2} \|D^2\phi\|_\infty \|D\phi\|_\infty \int_0^t d_M(X_s, X_s^\delta)^2 |dX_s^\delta|.$$

We conclude that $\Delta^\delta$ converges to 0 using conditions (A1) and (A2) satisfied by the approximation $(X^\delta)_{\delta>0}$ of $X$. □

In [3], Emery associates to $X$ a symmetric bracket $(dX, dX)$ named the $b$-quadratic variation ([3] Theorem 3.8). In a similar way, we now associate to the approximation $(X^\delta)_{\delta>0}$ a "skew-symmetric bracket", denoted by $d\mathcal{A}(X,X)$ in the following proposition.

PROPOSITION 1.3. – *Suppose $(X^\delta)_{\delta>0}$ is an approximation of $X$. There exists a unique linear mapping $b \mapsto \int_0^\cdot b\, d\mathcal{A}(X,X)$ from the space of all bilinear forms on $M$ to the space of real valued continuous processes with finite variation, such that for all $f, g \in C^2(M)$*

$$\text{(i)}\quad \int_0^\cdot f b\, d\mathcal{A}(X,X) = \int_0^\cdot f(X)\, d\Big(\int b\, d\mathcal{A}(X,X)\Big)$$

$$\text{(ii)}\quad \int_0^\cdot df \otimes dg\, d\mathcal{A}(X,X)$$

$$= \lim_{\delta\to 0}\Big(\int_0^\cdot \big(f(X) - f(X^\delta)\big)\, dg(X^\delta) - \frac{1}{2}\langle f(X), g(X)\rangle\Big)$$

*in probability uniformly on bounded intervals.*

*Moreover, if $(x^i)_{1\leqslant i\leqslant m}$ is an embedding of $M$ into $\mathbf{R}^m$ and $(S^{ij})_{1\leqslant i,j\leqslant m}$ is the limit process involved in condition* (A3) *for this embedding, then for any bilinear form $b$ on $M$ with $b = b_{ij}\, dx^i \otimes dx^j$,*

$$\int b\, d\mathcal{A}(X,X) = \int b_{ij}(X)\, dS^{ij}. \tag{6}$$

*Proof. – Uniqueness.* Let $b$ be a bilinear form on $M$ and write $b = b_{ij}\, dx^i \otimes dx^j$ for an embedding of $M$ into $\mathbf{R}^m$. Using (ii), (2) and condition (A3),

$$\begin{aligned}\int dx^i \otimes dx^j\, d\mathcal{A}(X,X) &= \lim_{\delta\to 0}\Big(\int (X^i - X^{\delta,i})\, dX^{\delta,j} - \frac{1}{2}\langle X^i, X^j\rangle\Big)\\ &= \lim_{\delta\to 0} S^{\delta,ij}\\ &= S^{ij}\end{aligned}$$

and then with (i), $\int b\,d\mathcal{A}(X,X)$ is uniquely defined by

$$\int b\,d\mathcal{A}(X,X) = \int b_{ij}(X)\,dS^{ij}.$$

*Existence.* Let $b$ be a bilinear form on $M$. We define $\int b\,d\mathcal{A}(X,X)$ by the expression $\int b_{ij}(X)\,dS^{ij}$ for some embedding of $M$ into $\mathbf{R}^m$.

Now, for $f, g \in \mathcal{C}^2$, writting $fb = fb_{ij}\,dx^i \otimes dx^j$ yields immediately (i). For (ii), write

$$df \otimes dg = \frac{\partial f}{\partial x^i}\frac{\partial g}{\partial x^j}\,dx^i \otimes dx^j$$

and use (3) to obtain

$$\int df \otimes dg\,d\mathcal{A}(X,X) = \lim_{\delta\to 0}\int \frac{\partial f}{\partial x^i}(X)\frac{\partial g}{\partial x^j}(X)\,dS^{\delta,ij}.$$

The same calculation as in the proof of Proposition 1.2, with $f$ and $g$ instead of $\phi^k$ and $\phi^l$, concludes the proof. □

Unlike the $b$-quadratic variation $(dX, dX)$ of $X$, the skew-symmetric bracket $d\mathcal{A}(X,X)$ depends on the approximation $(X^\delta)_{\delta>0}$ of $X$. As in [10], we introduce the following definition.

DEFINITION 1.4. – *We say that the approximation $(X^\delta)_{\delta>0}$ of $X$ is symmetric if the associated skew-symmetric bracket $d\mathcal{A}(X,X)$ vanishes.*

### 1.2. Examples of approximations

Intrinsic approximations of semimartingales in manifolds have been proposed by various authors. Basically two techniques are used to approximate semimartingales: regularization and interpolation. The first one consists in smoothing the semimartingale with some "intrinsic" average whereas in the second method a discrete sample of the paths is first considered and an interpolation rule is used.

All the geometric approximations that we have found in the literature are symmetric in the sense of Definition 1.4. Actually there is nothing surprising in it since a general rule seems to be: "If you want to get some intrinsic Stratonovich Calculus at the limit, you have to use symmetric approximation". This idea will be developed later on. Let us first consider regularization techniques.

*Example* 1. – The main point when one wants to regularize semimartingale on a manifold is to have some intrinsic expectation. One way to have expectation with a geometric meaning is to use "Barycentre" (introduced by Picard in [14]). Such a regularization is obtained in [2]: for all $\delta > 0$ and $t \geqslant 0$, $X_t^\delta$ is defined as the barycentre of all $X_s$ where $s$ runs uniformly in the interval $[t-\delta, t]$. Note that when the manifold is (an open subset of) $\mathbf{R}^m$ with its euclidian structure, $X_t^\delta$ is nothing but the usual regularization $X_t^\delta = \frac{1}{\delta}\int_{t-\delta}^t X_s\, ds$ (see [8]).

It is proved in Lemma 2.14 of [2] that for continuous $X$ such approximations are symmetric. We stress the fact that barycentres (see Remark 2.5 in [2]) are also related to interpolation rules.

*Example* 2. – The first definition of an interpolation rule has been introduced by Emery (Definition 7.9 in [3]):

DEFINITION 1.5. – *An interpolation rule is a measurable mapping $I$ from $M \times M \times [0,1]$ to $M$ such that*:

(i) $I(x,x,t) = x$ , $I(x,y,0) = x$ *and* $I(x,y,1) = y$;

(ii) $t \mapsto I(x,y,t)$ *is smooth*;

(iii) *for* $k = 1, 2$ *and* 3, $\frac{\partial^k}{\partial t^k} I(x,y,t) \in O(d_M(x,y)^k)$ *uniformly on bounded intervals.*

Although we can have different kinds of interpolation rules, the method to derive an approximation of a semimartingale from an interpolation rule is standard. Let us recall the general method which is also used in Example 3. For $\delta > 0$, let $(t_k^\delta)_{k\in\mathbf{N}}$ be a subdivision of $\mathbf{R}^+$ such that $\sup_{k\in\mathbf{N}} |t_{k+1}^\delta - t_k^\delta| \leqslant \delta$, and define the process $X^\delta$ by

$$X_t^\delta = I\left(X(t_k^\delta), X(t_{k+1}^\delta), \frac{t - t_k^\delta}{t_{k+1}^\delta - t_k^\delta}\right) \tag{7}$$

for $t \in [t_k^\delta, t_{k+1}^\delta]$, where $I$ is an interpolation rule.

Then $(X^\delta)_\delta$ satisfies Definition 1.1 with the slight difference that $X^\delta$ is $(\mathcal{F}_{t+\delta})_{t\geqslant 0}$ adapted in (A0): since the other assumptions are clearly fulfilled, the only point to prove is the convergence in assumption (A3). Actually our claim is that, for an interpolation rule as in Definition 1.5, this convergence holds toward 0 (straightforward application of Theorem 7.14 in [3]). Consequently we get a symmetric approximation and its relationship with Stratonovich Calculus is explained in [3].

*Example* 3. – Another interpolation rule has been proposed in [1] (Definition 11) which is recalled here.

DEFINITION 1.6. – *An interpolation rule is a $\mathcal{C}^3$ mapping from $M \times M \times [0,1]$ to $M$ such that:*

(i) $I(x,x,t) = x$ , $I(x,y,0) = x$ *and* $I(x,y,1) = y$;

(ii) *there exists a $\mathcal{C}^2$ function $\lambda : [0,1] \to \mathbf{R}$ such that $d_x I_{t,x} = \lambda(t)\, id_{T_x M}$ where $d_x$ is the tangent map at point $x$ of the partial map $I_{t,x}(y) = I(x,y,t)$ and where $id_{T_x M}$ is the identity map of the tangent space $T_x M$.*

Then an approximation of any semimartingale $X$ is defined with the general formula (7). Let us remark that the approximation based on this interpolation rule is also a symmetric approximation. It is a consequence of Proposition 3 in [1].

It is interesting to compare Definition 1.5 with Definition 1.6: the main difference is assumption (iii) in Definition 1.5 and assumption (ii) in Definition 1.6. They seem technical but necessary assumptions to get a Stratonovich Calculus at the limit. Hence the main questions are: Can we explain why we need such assumptions to get Stratonovich Calculus? What happens if we omit such assumptions? These questions are answered in the next section.

*Example* 4. – Let us now mention a non-symmetric approximation of a 2-dimensional Brownian motion introduced by Mac Shane [11] (see also [6] p. 484). Consider $M = \mathbf{R}^2$ with its euclidian structure and $X = (B^1, B^2)$ a 2-dimensional Brownian motion. Let $\phi^1, \phi^2 \in \mathcal{C}^1([0,1])$ such that $\phi^i(0) = 0$ and $\phi^i(1) = 1$ for $i = 1, 2$.

Let $\delta > 0$ be fixed. The approximating process $B^\delta$ is built via the following interpolation scheme:
for $i = 1, 2$ and for $t \in [k\delta, (k+1)\delta]$,

$$B_t^{\delta,i} = \begin{cases} B^i(k\delta) + \phi^i\left(\dfrac{t-k\delta}{\delta}\right)\Delta_k B^i & \text{if } \Delta_k B^1.\Delta_k B^2 > 0 \\ B^i(k\delta) + \phi^{3-i}\left(\dfrac{t-k\delta}{\delta}\right)\Delta_k B^i & \text{if } \Delta_k B^1.\Delta_k B^2 \leqslant 0 \end{cases} \tag{8}$$

where it is convenient to write for $k \in \mathbf{N}$ and $i = 1, 2$,

$$\Delta_k B^i := B^i\big((k+1)\delta\big) - B^i(k\delta).$$

To show that Mac Shane approximation is non-symmetric you have to compute the limit of $\int_0^t (B^1 - B^{\delta,1})\, dB^{\delta,2}$ when $\delta$ goes to 0. With the help of the Law of Large Numbers (make use that the Brownian motion has stationary and independent increments) you get

$$S_t^{12} = \frac{t}{\pi}\left(\int_0^1 (u-\phi^1(u))\dot{\phi}^2(u)\,du - \int_0^1 (u-\phi^2(u))\dot{\phi}^1(u)\,du\right).$$

Hence Mac Shane approximation may be non-symmetric for some functions $\phi^1$, $\phi^2$. Moreover it has two special features: first, it is strongly dependent on the linear structure of $\mathbf{R}^2$ because the approximation scheme (8) depends explicitely of the increments of the process. Secondly, although the approximation scheme (8) can be applied to any semimartingale $X$, it is clear that $\int_0^t (X^1 - X^{\delta,1})\,dX^{\delta,2}$ will not converge in general.

At this point we adress the problem to find a non-symmetric approximation which is intrinsic.

### 1.3. Approximations by general interpolation

A careful look at the proofs of the convergence of Theorem 7.14 in [3] and Proposition 3 in [1] shows that some assumption is needed, in the definition of interpolation rules, to have the increment "$I(x,y,t) - x$" (considered in $\mathbf{R}^m$ where the manifold is embedded) that depends linearly on the increment "$y - x$". Moreover one can guess that an approximation is non-symmetric if this linear operator is not homothetic. Let us state this precisely in the following definition.

DEFINITION 1.7. – *A general interpolation rule is a measurable mapping $I$ from $M \times M \times [0,1]$ to $M$ such that:*

(i) $I(x,x,t) = x$, $I(x,y,0) = x$, $I(x,y,1) = y$;

(ii) $t \mapsto I(x,y,t)$ *is* $\mathcal{C}^1$ *on* $[0,1]$ *and* $(x,y) \mapsto I(x,y,t)$ *is* $\mathcal{C}^2$ *on a neighbourhood of* $\{(x,x); x \in M\}$ *uniformly in* $(x,y)$.

Example 4 was the starting point of our interest in this problem. But the interpolation (8) is not a general interpolation in the sense of Definition 1.7 because it is not given by a deterministic interpolation rule. Moreover this random interpolation rule is not smooth ($\mathcal{C}^1$).

We then consider approximations of $M$-valued semimartingales constructed as in (7) and we show that they statisfy assumptions of Definition 1.1.

THEOREM 1.8. – *If $I$ is a general interpolation rule, and $(t_k^\delta)_{k\in\mathbf{N}}$ is a subdivision of $\mathbf{R}^+$ such that* $\sup_{k\in\mathbf{N}} |t_{k+1}^\delta - t_k^\delta| \leqslant \delta$, *let us define*

$$X_t^\delta = I\left(X\left(t_{k-1}^\delta\right), X\left(t_k^\delta\right), \frac{t - t_k^\delta}{t_{k+1}^\delta - t_k^\delta}\right) \tag{9}$$

*for $t \in [t_k^\delta, t_{k+1}^\delta]$* [3] *and any $M$-valued semimartingale $X$. Then $X^\delta$ is an approximation of $X$ and if $(x^i)_{1\leqslant i\leqslant m}$ is an embedding of $M$ into $\mathbf{R}^m$ the skew-symmetric bracket is*

$$S_t^{ij} = \frac{1}{2}\int_0^t \int_0^1 \left(A_p^j \frac{\partial A_l^i}{\partial s} - A_l^i \frac{\partial A_p^j}{\partial s}\right)(s, X_u)\, ds\, d\langle X^l, X^p\rangle_u,$$

*where $A(t,x)$ is the tangent map of the partial mapping $I_{t,x}(y) = I(x,y,t)$ at point $y = x$, i.e.,*

$$A(t,x) = d_x(I_{t,x}).$$

Note that this theorem answers the questions of Examples 2 and 3 in the previous section. Clearly if $I$ satisfies Definition 1.6, $S^{ij}$ is vanishing. Actually in this instance $A(t,x) = \lambda(t)\, id_{T_xM}$ with $\lambda(1) = 1$ and $\lambda(0) = 0$.

In the same vein if we assume Definition 1.5 then $A(t,x) = t\, id_{T_xM}$. It is shown in this elementary lemma.

LEMMA 1.9. – *If $I$ is a measurable map from $M \times M \times [0,1]$ to $M$ such that*:

(i) *$I(x,x,t) = x$ , $I(x,y,0) = x$ and $I(x,y,1) = y$;*
(ii) *$t \mapsto I(x,y,t)$ is $C^2$;*
(iii) *$\frac{\partial^2}{\partial t^2} I(x,y,t) \in \mathrm{O}(d_M(x,y)^2)$ uniformly on bounded intervals then $d_x(I_{t,x}) = t\, id_{T_xM}$.*

*Proof.* – Using a first order expansion we get

$$y^i - x^i = \frac{\partial}{\partial t} I^i(x,y,0) + \mathrm{O}\big(d_M(x,y)^2\big) \tag{10}$$

where $M$ is embedded into $\mathbf{R}^m$. Then

$$I^i(x,y,t) - x^i = t\frac{\partial}{\partial t} I^i(x,y,0) + \mathrm{O}\big(d_M(x,y)^2\big)$$

[3] $X(t_{-1}^\delta)$ is set equal to $X(t_0^\delta)$ to define $X_t^\delta$ on the first interval $[t_0^\delta, t_1^\delta]$.

which in view of (i) and (10) leads to:

$$I^i(x,y,t) - I^i(x,x,t) = t(y^i - x^i) + \mathrm{O}(d_M(x,y)^2).$$

Hence Lemma 1.9 is proved. □

We now proceed to the proof of Theorem 1.8.

*Proof of Theorem 1.8.* – Since the assumptions (A0), (A1) and (A2) of Definition 1.1 are clearly satisfied it remains to check assumption (A3). Let us compute the limit of $\int_0^t (X_s^i - X_s^{\delta,i})\, dX_s^{\delta,j}$ when $\delta$ goes to 0. If we assume $t = t_n^\delta$, which means no loss in the generality since $\delta$ goes to 0, and write $t_k$ instead of $t_k^\delta$, then

$$\begin{aligned}
&\int_0^t (X_s^i - X_s^{\delta,i})\, dX_s^{\delta,j} \\
&\quad = \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} (X_s^i - X_{t_{k-1}}^i)\, dX_s^{\delta,j} - \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} (X_s^{\delta,i} - X_{t_{k-1}}^i)\, dX_s^{\delta,j} \\
&\quad = J_t^{\delta(1)} - J_t^{\delta(2)}.
\end{aligned}$$

An integration by parts in $J_t^{\delta(1)}$ gives

$$\begin{aligned}
J_t^{\delta(1)} = \sum_{k=0}^{n-1} \Bigg\{ & \big[(X_s^i - X_{t_{k-1}}^i)(X_s^{\delta,j} - X_{t_{k-1}}^j)\big]_{t_k}^{t_{k+1}} \\
& - \int_{t_k}^{t_{k+1}} (X_s^{\delta,j} - X_{t_{k-1}}^j)\, dX_s^i \Bigg\};
\end{aligned} \tag{11}$$

then

$$\begin{aligned}
J_t^{\delta(1)} = \sum_{k=0}^{n-1} \Bigg\{ & (X_{t_{k+1}}^i - X_{t_{k-1}}^i)(X_{t_k}^j - X_{t_{k-1}}^j) \\
& - \int_{t_k}^{t_{k+1}} (X_s^{\delta,j} - X_{t_{k-1}}^j)\, dX_s^i \Bigg\}.
\end{aligned} \tag{12}$$

We can split the first line of the previous equation into

$$\sum_{k=0}^{n-1} (X_{t_{k+1}}^i - X_{t_k}^i)(X_{t_k}^j - X_{t_{k-1}}^j) + \sum_{k=0}^{n-1} (X_{t_k}^i - X_{t_{k-1}}^i)(X_{t_k}^j - X_{t_{k-1}}^j)$$

and the first sum converges to 0 whereas the second one to $\langle X^i, X^j\rangle_t$.

Consequently $J_t^{\delta(1)}$ converges to $\langle X^i, X^j\rangle_t$ in probability uniformly on bounded intervals.

Introducing the interpolation rule $I$ in $J_t^{\delta(2)}$, we have

$$J_t^{\delta(2)} = \sum_{k=0}^{n-1} \int_0^1 \big(I^i(X_{t_{k-1}}, X_{t_k}, s) - I^i(X_{t_{k-1}}, X_{t_{k-1}}, s)\big)\dot{I}^j(X_{t_{k-1}}, X_{t_k}, s)\,ds.$$

Then define $\Delta_k X = X_{t_k} - X_{t_{k-1}}$ and

$$J_t^{\delta(2)} = \sum_{k=0}^{n-1} \int_0^1 \big[A(s, X_{t_{k-1}})\Delta_k X + \mathrm{O}\big(d_M(X_{t_k}, X_{t_{k-1}})^2\big)\big]^i \times \dot{I}^j(X_{t_{k-1}}, X_{t_k}, s)\,ds. \tag{13}$$

An integration by parts yields

$$\begin{aligned}
&\int_0^1 \big[A(s, X_{t_{k-1}})\Delta_k X\big]^i \dot{I}^j(X_{t_{k-1}}, X_{t_k}, s)\,ds \\
&= \Delta_k X^i \Delta_k X^j - \int_0^1 \left[\frac{\partial A}{\partial s}(s, X_{t_{k-1}})\Delta_k X\right]^i \big(I^j(X_{t_{k-1}}, X_{t_k}, s) - X^j_{t_{k-1}}\big)\,ds \\
&= \Delta_k X^i \Delta_k X^j - \int_0^1 \left[\frac{\partial A}{\partial s}(s, X_{t_{k-1}})\Delta_k X\right]^i \big[A(s, X_{t_{k-1}})\Delta_k X\big]^j\,ds \\
&\quad + \mathrm{O}\big(d_M(X_{t_k}, X_{t_{k-1}})^3\big).
\end{aligned} \tag{14}$$

Then

$$J_t^{\delta(2)} = \sum_{k=0}^{n-1} \left[\Delta_k X^i \Delta_k X^j - \int_0^1 \frac{\partial A_l^i}{\partial s}(s, X_{t_{k-1}}) A_p^j(s, X_{t_{k-1}}) \Delta_k X^l \Delta_k X^p\,ds + \mathrm{O}\big(d_M(X_{t_k}, X_{t_{k-1}})^3\big)\right]. \tag{15}$$

When $\delta$ goes to 0, since the quadratic variation of $X$ on $[0, t]$ is a.s. finite, $J_t^{\delta(2)}$ converges to

$$\langle X^i, X^j\rangle_t - \int_0^t \int_0^1 \frac{\partial A_l^i}{\partial s}(s, X_u) A_p^j(s, X_u)\,ds\,d\langle X^l, X^p\rangle_u.$$

Combining the limit of $J_t^{\delta(1)}$ and $J_t^{\delta(2)}$ concludes the proof of the theorem. □

## 2. INTEGRAL OF 1-FORMS

We again consider, in all this section, a continuous semimartingale with values in $M$ and an approximation $(X^\delta)_{\delta>0}$ of $X$ (Definition 1.1).

### 2.1. Integral of a 1-form along an approximated semimartingale

Let us recall a few words of differential geometry. A *1-form on $M$* (or a form of degree 1) is a smooth mapping $\alpha$ from $M$ to the cotangent fiber bundle $T^*M = \bigcup_{p\in M} T_p^*M$ such that for each $p \in M$, $\alpha(p)$ is a (classical) linear form on $T_pM$. An *exterior form of degree 2 on $M$* is a smooth mapping $\omega$ on $M$ such that for each $p \in M$, $\omega(p)$ is a skew-symmetric bilinear form on $T_pM$. The *exterior differentiation* maps the forms of degree 1 to forms of degree 2 by the following: for each 1-form $\alpha$, $\forall p \in M, \forall A, B \in T_pM$,

$$d^e\alpha(p)(A, B) = \big(\alpha'(p)A\big).B - \big(\alpha'(p)B\big).A.$$

So let $\alpha$ be a 1-form on $M$. The integral $\int \alpha(X^\delta)\,dX^\delta$ of $\alpha$ along the continuous with finite variation process $X^\delta$ is well defined for all $\delta > 0$, as well as the Stratonovich integral $\int \alpha \circ dX$ of $\alpha$ along the continuous semimartigale $X$ (see [3] Chapter VII). We are now interested in the limit of $\int \alpha(X^\delta)\,dX^\delta$ when $\delta$ goes to 0.

PROPOSITION 2.1. – *Suppose $(X^\delta)_{\delta>0}$ is an approximation of $X$. Then for any 1-form $\alpha$ on $M$,*

$$\lim_{\delta\to 0} \int_0^{\cdot} \alpha(X^\delta)\,dX^\delta = \int_0^{\cdot} \alpha \circ dX + \frac{1}{2}\int_0^{\cdot} (d^e\alpha)\,d\mathcal{A}(X, X) \tag{16}$$

*in probability uniformly on bounded intervals.*

*Proof.* – This proposition is proved in the next subsection as a special case of stochastic differential equation. □

The extra-term appearing in the right hand side of (16) needs some comments.

1. When you take for $\alpha$ the 1-form $df$ for some function $f \in \mathcal{C}^\infty(M)$, then the extra-term vanishes since $d^e\alpha = 0$, and Eq. (16) means that $f(X^\delta) - f(X_0^\delta)$ goes to $f(X) - f(X_0)$.

2. When $(X^\delta)_{\delta>0}$ is an approximation of $X$ obtained by interpolation as in the Examples 3 and 4 of Section 1.2, then the extra-term also vanishes since we are concerned with a symmetric approximation, and the previous proposition is nothing but Proposition 7.27 of [3].

## 2.2. A general second order calculus

In Proposition 2.1, the stochastic integration of a semimartingale approximated by $X^\delta$ has been presented by adding an extra term to the usual Stratonovich calculus. Besides this probabilistic presentation we can wonder what the geometric meaning of the right hand side of (16) is. When $M$ is embedded into $\mathbf{R}^m$, the right hand side of (16) becomes

$$\int \alpha_i(X_s)\,dX_s^i + \frac{1}{4}\left(\frac{\partial\alpha_i}{\partial x^j} + \frac{\partial\alpha_j}{\partial x^i}\right)(X_s)\,d\langle X^i, X^j\rangle_s + \frac{1}{2}\left(\frac{\partial\alpha_i}{\partial x^j} - \frac{\partial\alpha_j}{\partial x^i}\right)(X_s)\,dS_s^{ij}. \tag{17}$$

Hence we want to find the intrinsic nature of

$$\left(dX^i, \frac{1}{2}d\langle X^i, X^j\rangle, dS^{ij}\right).$$

This question has been already solved by [12,16] and [3] when the skew-symmetric bracket is vanishing and it leads to the second order calculus. In this section we present a general second order calculus that takes into account the skew-symmetric bracket. Because the second order calculus is not yet familiar to most probabilists we only give an outline of this presentation and we will not use the notations introduced in this section later in the paper. Following the spirit of Chapter VI in [3] the easiest way to give sense to the "non-symmetric tangent element of order two" of an approximated semimartingale, which we denote by "$TX$", is to define the "dual" space of forms which can be integrated against "$TX$". This dual space has been already considered among other geometric concepts in a paper from Meyer [13]. In our framework we only use the second order forms that are written

$$\theta = \theta_i d_{21}x^i + \theta_{ij} d_2x^j d_1x^i \tag{18}$$

if $(x^i)_{1\leqslant i\leqslant m}$ is an embedding of $M$ into $\mathbf{R}^m$. Although you can find the geometrical definition of $d_{21}$, $d_1$, $d_2$ in [13], a way to understand the meaning of (18) ((2) in [13]) for probabilists is to define the stochastic integral

$$\int \langle \theta, TX\rangle := \int \theta_i(X)\,dX^i + \int \theta_{ij}(X)\,d\left(\frac{1}{2}\langle X^i, X^j\rangle + S^{ij}\right), \qquad (19)$$

where $X$ is a $M$-valued semimartingale and where a skew-symmetric bracket is supposed to be defined for every semimartingale. An important feature of second order forms is the non-commutativity of $d_2x^j d_1x^i$, a fact which is clear in (19) if the approximation is not symmetric. Formula (19) can be extended to a general second order calculus following the formalism of Chapter VI of [3], but it seems more useful to stress the relationship between this general second order calculus and the approximations $X^\delta$ of semimartingales. If $\alpha$ is a 1-form on $M$, the "Leibnitz" differentiation $d_2$ is mapping 1-form to second order form:

$$d_2(\alpha_i d_1 x^i) = \alpha_i d_{21}x^i + \frac{\partial \alpha^i}{\partial x^j} d_2x^j\, d_1x^i. \qquad (20)$$

Hence (16) can be written

$$\lim_{\delta\to 0} \int_0^{\cdot} \alpha(X^\delta)\,dX^\delta = \int \langle d_2\alpha, TX\rangle. \qquad (21)$$

Since the space of second ordre form is generated as $C^2$ module by the "Leibnitz" differentiation of 1-form, it should be the most practical way to have an insight in the probabilistic meaning of the general second order calculus. Moreover every second order form can be split in a symmetric part and a skew-symmetric part. For $d_2\alpha$ given by (20) those two parts are, respectively,

$$\alpha_i d_{21}x^i + \frac{1}{2}\left(\frac{\partial \alpha^i}{\partial x^j} + \frac{\partial \alpha^j}{\partial x^i}\right) d_2x^j\, d_1x^i$$

and

$$\frac{1}{2}\left(\frac{\partial \alpha^i}{\partial x^j} - \frac{\partial \alpha^j}{\partial x^i}\right) d_2x^j\, d_1x^i,$$

which corresponds to

$$\int \langle d_2\alpha, TX\rangle = \int_0^{\cdot} \alpha \circ dX + \frac{1}{2}\int_0^{\cdot} (d^e\alpha)\, d\mathcal{A}(X, X).$$

In the next sections we consider the analogous of Stratonovich equations in our non-symmetric framework.

## 3. APPROXIMATION OF STOCHASTIC DIFFERENTIAL EQUATIONS BETWEEN MANIFOLDS

### 3.1. Differential equations between manifolds

First let us remind what ordinary differential equations between manifolds are. Let $M$ and $N$ be smooth manifolds. A general ordinary differential equation between $M$ and $N$ is described by a family $\{e(x, y);\ x \in M, y \in N\}$ where $e(x, y)$ is a linear operator from $T_xM$ to $T_yN$ which depends smoothly on $(x, y)$. Then to each $\mathcal{C}^1$-curve $(x(t); t \geqslant 0)$ on $M$ is associated the ordinary differential equation on $N$

$$\begin{cases} \dot{y}(t) = e(x(t), y(t))\dot{x}(t), \\ y(0) = y \in N. \end{cases}$$

A slight generalization of this gives a sense, for each $\delta > 0$, to

$$\begin{cases} dY_t^\delta = e\big(X_t^\delta, Y_t^\delta\big)\, dX_t^\delta, \\ Y_0^\delta = y \in N \end{cases} \tag{22}$$

and we know that (22) admits a unique solution $(Y_t^\delta; 0 \leqslant t < \eta^\delta)$ up to an explosion time $\eta^\delta$.

We now study the convergence of $(Y^\delta)_{\delta>0}$ as $\delta$ goes to 0, and prove that the limit process $Y$ is not the solution of the Stratonovich differential equation $dY_t = e(X_t, Y_t) \circ dX_t$ (as expected by the well-known result of Wong–Zakai [17]). Actually $Y$ is the solution of a stochastic differential equation which contains an extra-term along the skew-symmetric bracket $d\mathcal{A}(X, X)$. The integrant of $d\mathcal{A}(X, X)$ is introduced in the following lemma.

LEMMA 3.1. – *Let $e = \{e(x, y); x \in M, y \in N\}$ be a family of linear operators $e(x, y)$ from $T_xM$ to $T_yN$. For each $x$ in $M$ and $y$ in $N$, define a map $[e, e](x, y)$ on $T_xM \times T_xM$ by the prescription*: $\forall A, B \in T_xM$,

$$[e,e](x,y)(A,B) = \big[e(x,.)(B), e(x,.)(A)\big]_N(y) \\ + d^e e(.,y)(x)(A,B), \tag{23}$$

*where* $[.,.]_N$ *denotes the Lie product of two vector fields on* $N$ *and* $d^e$ *denotes the exterior differentiation of* 1*-degree forms.*

*Then* $[e,e](x,y)$ *is a skew-symmetric bilinear operator from* $T_xM \times T_xM$ *to* $T_yN$.

When $M$ and $N$ are embedded in $\mathbf{R}^m$ and $\mathbf{R}^n$, $[e,e](x,y)$ is described by

$$[e,e]^\alpha_{i,j}(x,y) = \left(\frac{\partial e^\alpha_i}{\partial x^j} + \frac{\partial e^\alpha_i}{\partial y^\beta} e^\beta_j - \frac{\partial e^\alpha_j}{\partial x^i} - \frac{\partial e^\alpha_j}{\partial y^\beta} e^\beta_i\right)(x,y) \tag{24}$$

for $1 \leqslant i,j \leqslant m$ and $1 \leqslant \alpha \leqslant n$.

*Proof.* – Proof of Lemma 3.1 Let us first remark that the right hand side of (23) lies in $T_yN$, for $x \in M$, $y \in N$, $A, B \in T_xM$. On one hand, for a fixed $x$ in $M$, $e(x,.)(A)$ and $e(x,.)(B)$ are both vector fields on $N$ for any pair $A, B$ in $T_xM$. On the other hand, for a fixed $y$ in $N$, $e(.,y)$ is a $T_yN$-form of degree 1 on $M$.

The skew-symmetric property of $(A,B) \mapsto [e,e](x,y)(A,B)$ is obvious. □

### 3.2. Approximated stochastic differential equations between manifolds

As in Section 3.1, two manifolds $M$ and $N$ and a continuous semimartingale $X$ with values in $M$ are given.

THEOREM 3.2. – *Let* $(X^\delta)_{\delta>0}$ *be an approximation of* $X$ *and* $e = \{e(x,y); x \in M, y \in N\}$ *be a family of linear operators* $e(x,y)$ *from* $T_xM$ *to* $T_yN$ *depending smoothly on* $(x,y)$*. For any* $y$ *in* $N$*, denote by* $(Y^\delta)_{\delta>0}$ *the family of solution on* $N$*, for all positive* $\delta$*, of the ordinary differential equation*

$$\begin{cases} dY^\delta_t = e\big(X^\delta_t, Y^\delta_t\big)\, dX^\delta_t, \\ Y^\delta_0 = y. \end{cases} \tag{25}$$

*Then* $Y^\delta$ *converges in probability uniformly on bounded intervals to the solution* $Y$ *of the following stochastic differential equation*

$$\begin{cases} dY_t = e(X_t,Y_t) \circ dX_t + \frac{1}{2}[e,e](X_t,Y_t)\, d\mathcal{A}(X,X)_t, \\ Y_0 = y. \end{cases} \tag{26}$$

*Remark.* – The solution of (26) is defined up to an explosion time, say $\eta$, as the solutions of (25), say $\eta^\delta$ for all $\delta > 0$. This stopping time $\eta$ is a.s. positive and the meaning of the convergence in probability is (see [3] p. 100): "$\eta^\delta \wedge \eta$ converges in probability toward $\eta$ and, on $[[0, \eta[[$, $Y^\delta$ converges in probability uniformly on bounded intervals to $Y$", i.e., for every $\lambda \geqslant 0$ and $k \in \mathbf{N}^*$

$$\lim_{\delta\to 0} \mathbf{P}\left(\eta^\delta \leqslant \eta - \frac{1}{k}\right) = 0$$

and

$$\lim_{\delta\to 0} \mathbf{P}\Big(\sup_{s\leqslant k\wedge(\eta-\frac{1}{k})} d_M(Y_s^\delta, Y_s) \geqslant \lambda\Big) = 0.$$

Note that the first condition implies that $Y_s^\delta$ is defined on whole $[[0, \eta - \frac{1}{k}[[$ except on an event which probability goes to 0.

Before the proof, let us mention some special cases where this theorem applies.

1. When $N = \mathbf{R}$ and $e(x, y) = \alpha(x)$ for all $(x, y) \in M \times N$ with $\alpha$ a 1-form on $M$, the theorem claims that $Y^\delta = \int \alpha(X^\delta)\, dX^\delta$ converges to $Y = \int \alpha \circ dX + \frac{1}{2}\int (d^e\alpha)\, d\mathcal{A}(X, X)$. This is nothing but Proposition 2.1.
2. When $(X^\delta)_{\delta>0}$ is a symmetric approximation of $X$, for instance approximation obtained by interpolation as in the Examples 3 and 4 of Section 1.2, then $Y^\delta$ converges to the solution $Y$ of the Stratonovich differential equation $dY = e(X, Y) \circ dX$. Theorem 3.2 is a generalization of Theorem 7.24 in [3].

*Proof.* – We reduce the problem to an embedded problem and apply a result of Gyöngy. In [4] he proved a limit theorem for approximated stochastic differential equations where both the coefficient and the driving semimartingale are approximated. In order to follow the procedure described in [4], we use similar notations and write

$$\begin{cases} dY_t^{\delta,\alpha} = \sigma_i^\alpha(\delta, t, Y_t^\delta) \circ dX_t^{\delta,i} \\ Y_0^{\delta,\alpha} = y^\alpha, \end{cases}$$

where $\sigma(\delta, t, y)$ is a $m \times n$-matrices valued semimartingale defined by

$$\sigma_i^\alpha(\delta, t, y) = e_i^\alpha(X_t^\delta, y).$$

As $\delta$ goes to 0, $\sigma(\delta,t,y)$ tends to $\sigma(t,y)=e(X_t,y)$ and $\sigma(.,y)$ has the following differential

$$d\sigma(t,y)=\sigma_{(k)}(t,y)\,dX_t^k+\sigma_{(k,l)}(t,y)\,d\langle X^k,X^l\rangle_t,$$

where

$$\sigma_{(k)}(t,y)=\frac{\partial e}{\partial x^k}(X_t,y)$$

and

$$\sigma_{(k,l)}(t,y)=\frac{1}{2}\frac{\partial^2 e}{\partial x^k\partial x^l}(X_t,y).$$

Moreover, computing the differential of $\sigma(\delta,.,y)$

$$d\sigma(\delta,t,y)=\sigma_{(k)}(\delta,t,y)\,dX_t^{\delta,k}+\sigma_{(k,l)}(\delta,t,y)\,d\langle X^{\delta,k},X^{\delta,l}\rangle_t, \tag{27}$$

where

$$\sigma_{(k)}(\delta,t,y)=\frac{\partial e}{\partial x^k}(X_t^\delta,y)$$

and

$$\sigma_{(k,l)}(\delta,t,y)=\frac{1}{2}\frac{\partial^2 e}{\partial x^k\partial x^l}(X_t^\delta,y).$$

Let us remark $\langle X^{\delta,k},X^{\delta,l}\rangle=0$, nevertheless it is useful to consider the last term in (27) since

$$\lim_{\delta\to 0}\sigma_{(k)}(\delta,t,y)=\sigma_{(k)}(t,y)$$

and

$$\lim_{\delta\to 0}\sigma_{(k,l)}(\delta,t,y)=\sigma_{(k,l)}(t,y).$$

Then by applying the Theorem 3.3 of [4] we get the convergence in probability uniformly on bounded intervals of $Y^\delta$ to the solution $Y$ of the following stochastic differential system

$$\begin{cases} dY_t^\alpha=\sigma^\alpha(t,Y_t)\circ dX_t-\frac{1}{2}[\sigma_j,\sigma_l]^\alpha(t,Y_t)\,dS_t^{jl}+\sigma_{j(k)}^\alpha(t,Y_t)\,dS_t^{jk}\\ Y_0^\alpha=y^\alpha, \end{cases}$$

where $\sigma_j$ is the $j$th column of the matrix $\sigma$ considered as a vector field on the manifold $N$ and $[\sigma_j,\sigma_l]$ denotes the Lie product of $\sigma_j$ and $\sigma_l$. Hence, at the opposite of the notation in [4], we have

$$[\sigma_j, \sigma_l]^\alpha(t, y) = \left(\frac{\partial e_l^\alpha}{\partial y^\beta} e_j^\beta - \frac{\partial e_j^\alpha}{\partial y^\beta} e_l^\beta\right)(X_t, y).$$

Using the expression of $\sigma$ and $\sigma_{(k)}$, the skew-symmetric property of $(S^{ij})$ and the coordinates of $[e, e]$ given in (24), the limit process $Y$ appears as solution of

$$\begin{cases} dY_t^\alpha = e^\alpha(X_t, Y_t) \circ dX_t + \frac{1}{2}[e, e]_{i,j}^\alpha (X_t, Y_t)\, dS_t^{ij} \\ Y_0^\alpha = y^\alpha \end{cases}$$

and proves the theorem. □

Once established the convergence of $Y^\delta$ towards $Y$, one should ask if we can compute the skew symmetric bracket $\mathcal{A}(Y, Y)$. For the sake of simplicity we assume in the following that there is no explosion time (i.e., $\forall\delta$, $\eta^\delta = +\infty$ and $\eta = +\infty$).

COROLLARY 3.3. – *With the notations of Theorem* 3.2. *if* $(Y^\delta)_{\delta>0}$ *satisfies assumption* (A2) *then it is an approximation of* $Y$ *and the following relationship between the skew-symmetric brackets* $\mathcal{A}(Y, Y)$ *and* $\mathcal{A}(X, X)$ *holds*

$$S^{\alpha\beta}(Y) = \int_0^\cdot e_i^\alpha e_j^\beta (X, Y)\, dS(X)^{ij}, \tag{28}$$

*where*

$$(S^{ij})(X)_{1\leqslant i,j\leqslant m} \quad \textit{and} \quad \big(S^{\alpha\beta}(Y)\big)_{1\leqslant \alpha,\beta\leqslant n}$$

*denote the coordinates of* $\mathcal{A}(X, X)$ *and* $\mathcal{A}(Y, Y)$, *respectively, when the manifolds* $M$ *and* $N$ *are embedded in* $\mathbf{R}^m$ *and* $\mathbf{R}^n$.

*Proof.* – Since the assumptions (A0), (A1) of Definition 1.1 are clearly satisfied it remains to check assumption (A3). Hence we have to compute the limit as $\delta$ goes to 0, of

$$\begin{aligned} &\int_0^\cdot \big(Y_s^\alpha - Y_s^{\delta,\alpha}\big)\, dY_s^{\delta,\beta} - \frac{1}{2}\langle Y^\alpha, Y^\beta\rangle \\ &\quad = \int_0^\cdot \big(Y_s^\alpha - Y_s^{\delta,\alpha}\big) e_j^\beta \big(X_s^\delta, Y_s^\delta\big)\, dX_s^{\delta,j} \\ &\qquad - \frac{1}{2}\int_0^\cdot e_i^\alpha e_j^\beta (X_s, Y_s)\, d\langle X^i, X^j\rangle_s. \end{aligned} \tag{29}$$

We consider the first term of (29) as solution of an "equation" driven by $X^\delta$ and we use Theorem 3.3 of [4] again to obtain (28). □

### 3.3. Application to Lie group and stochastic exponential

As a conclusion we apply Theorem 3.2 to a very simple case where the correction term is explicitely computed. The stochastic exponential of the Heisenberg group is taken as an example.

Hence consider a $d$-dimensional Lie group $G$ with unit $e$. The associated Lie algebra $T_eG$ is denoted by $\mathcal{G}$ and endowed with a fixed basis $(H_i)_{1\leqslant i\leqslant d}$. From [5], to each $\mathcal{G}$-valued continuous semimartingale $X$ corresponds a $G$-valued continuous semimartingale $Y=\mathcal{E}(X)$, called the stochastic exponential of $X$, which satisfies: $\forall f\in\mathcal{C}^2(G)$, $\forall t\geqslant 0$,

$$f(Y_t)=f(e)+\int_0^t H_i f(Y_s)\circ dX_s^i,$$

where $X$ is written as $X=X^iH_i$ and where $H_i$ still denotes the unique left invariant vector field on $G$ which coincides with $H_i$ at point $e$. This equation is a special case of stochastic differential equation between manifolds $\mathcal{G}$ and $G$.

A natural question is whether $\mathcal{E}(X)$ is stable if the $\mathcal{G}$-valued semimartingale $X$ is approximated. The answer is negative for a non-symmetric approximation. More precisely, Theorem 3.2 claimed that if $(X^\delta)_{\delta>0}$ is any approximation of $X$ satisfying Definition 1.1, then $\mathcal{E}(X^\delta)$ converges in probability uniformly on bounded intervals to the $G$-valued semimartingale $Y$ solution of: $\forall f\in\mathcal{C}^2(G)$, $\forall t\geqslant 0$,

$$f(Y_t)=f(e)+\int_0^t H_i f(Y_s)\circ dX_s^i+\frac{1}{2}\int_0^t [H_j,H_i]f(Y_s)\,dS_s^{ij},$$

where $[H_j,H_i]$ is the Lie product of the two left invariant vector fields $H_j$ and $H_i$ on $G$, and

$$S^{ij}=\lim_{\delta\to 0}\int \left(X^i-X^{\delta,i}\right)dX^{\delta,j}-\frac{1}{2}\langle X^i,X^j\rangle.$$

In the particular case where $G$ is the 3-dimensional Heisenberg group, an effective computation is possible. Each

$$g = \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \in G$$

is denoted by the $\mathbf{R}^3$-vector $(x, y, z)$ with the multiplicative rule $(x, y, z) \times (x', y', z') = (x + x', y + y', z + z' + xy')$. The Lie algebra associated to $G$ is

$$\mathcal{G} = \left\{ \begin{pmatrix} 0 & a & c \\ 0 & 0 & b \\ 0 & 0 & 0 \end{pmatrix} ; (a, b, c) \in \mathbf{R}^3 \right\}$$

and is identified with $\mathbf{R}^3$.

Then (see [5] p. 371) for any $\mathcal{G}$-valued semimartingale $X = (X^1, X^2, X^3)$, the stochastic exponential of $X$ is the $G$-valued semimartingale

$$\mathcal{E}(X) = \left(X^1, X^2, X^3 + \int X^1 dX^2 + \frac{1}{2}\langle X^1, X^2\rangle\right). \tag{30}$$

Let now $(X^\delta)_{\delta>0}$ be an approximation of $X$ with $X^\delta = (X^{\delta,1}, X^{\delta,2}, X^{\delta,3})$, satisfying Definition 1.1 with limit matrix process $(S^{ij})_{1\leqslant i,j\leqslant 3}$. Applying Theorem 3.2 gives the convergence of $\mathcal{E}(X^\delta)$ to the $G$-valued semimartingale

$$Y = \left(X^1, X^2, X^3 + \int X^1 dX^2 + \frac{1}{2}\langle X^1, X^2\rangle + S^{21}\right)$$

as $\delta$ goes to 0. By (30), we obtain

$$\lim_{\delta\to 0} \mathcal{E}(X^\delta) = \mathcal{E}(\widetilde{X}) \quad \text{where } \widetilde{X} = (X^1, X^2, X^3 + S^{21}).$$

## ACKNOWLEDGEMENT

We would like to thank the referee for his patient and carefull reading and for pointing out some mistakes in an earlier version.

## REFERENCES

[1] S. COHEN, Géométrie différentielle stochastique avec sauts 2 : discrétisation et applications des eds avec sauts, *Stochastics and Stochastic Rep.* 56 (1996) 205–225.

[2] S. COHEN et A. ESTRADE, Régularisation d'équations de Stratonovitch à sauts entre variétés, *Ann. Fac. Sciences de Toulouse* 7 (2) (1998) 233–265.

[3] M. EMERY, *Stochastic Calculus in Manifolds*, Springer, Berlin, Heidelberg, 1989.

[4] I. GYÖNGY, On the approximation of stochastic differential equations, *Stochastics* 23 (1988) 331–352.

[5] M. HAKIM-DOWEK et D. LÉPINGLE, L'exponentielle stochastique des groupes de Lie, in: *Sém. Probab. XX*, LNM 1204 (1986) 352–374.

[6] N. IKEDA and S. WATANABE, *Stochastic Differential Equations and Diffusion Processes*, North-Holland/Kodansha, Amsterdam/Tokyo, 1989.

[7] A. JAKUBOWSKI, J. MÉMIN et G. PAGÈS, Convergence en loi des suites d'intégrales stochastiques sur l'espace $\mathbf{D}^1$ de Skorokhod, *Probab. Theory Related Fields* 81 (1989) 111–137.

[8] T.G. KURTZ, E. PARDOUX and P. PROTTER, Stratonovich stochastic differential equations driven by general semimartingales, *Ann. Inst. Henri Poincaré* 31 (2) (1995) 351–378.

[9] T.G. KURTZ and P. PROTTER, Weak limit theorems for stochastic integrals and differential equations, *Ann. Probab.* 19 (1991) 1035–1070.

[10] V. MACKEVICIUS, $S^p$ stability of solutions of symmetric stochastic differential equations with discontinuous driving semimartingales, *Ann. Inst. Henri Poincaré* 23 (4) (1987) 575–592.

[11] E.J. MCSHANE, Stochastic differential equations and models of random processes, in: *Proc. Sixth Berkeley Symp. Math. Stat. Probab. III,* University of California Press, Berkeley, 1972, pp. 263–294.

[12] P.A. MEYER, Géométrie stochastique sans larmes, in: *Sém. Probab. XV*, LNM 850 (1981) 44–102.

[13] P.A. MEYER, Qu'est-ce qu'une différentielle d'ordre $n$?, *Expo. Math.* 7 (1989) 249–264.

[14] J. PICARD, Barycentres et martingales sur une variété, *Ann. Inst. Henri Poincaré* 30 (4) (1994) 647–702.

[15] J. PICARD, Convergence in probability for perturbed stochastic integral equations, *Probab. Theory Related Fields* 81 (1989) 383–451.

[16] L. SCHWARTZ, Géométrie différentielle du deuxième ordre, semi-martingales et équations différentielles sur une variété différentielle, in: *Sém. Probab. XVI*, LNM 921 (1982).

[17] E. WONG and M. ZAKAI, On the convergence of ordinary integrals to stochastic integrals, *Ann. Math. Stat.* 36 (1965) 1560–1564.