SÉMINAIRE DE PROBABILITÉS (STRASBOURG)

MICHEL ÉMERY

On two transfer principles in stochastic differential geometry

Séminaire de probabilités (Strasbourg), tome 24 (1990), p. 407-441

<http://www.numdam.org/item?id=SPS_1990__24__407_0>

ON TWO TRANSFER PRINCIPLES
IN STOCHASTIC DIFFERENTIAL GEOMETRY

M. Emery(*)

A well known rule of thumb in stochastic differential geometry is what Malliavin calls "the transfer principle": Geometric constructions involving manifold-valued curves can be extended to manifold-valued processes by replacing classical calculus with Stratonovich stochastic calculus. This is explained by Stratonovich differentials obeying the ordinary chain-rule, and also by an approximation result when the random process is smoothed by some convolution or a polygonal interpolation. Extending Bismut's work on Brownian diffusions [1], Schwartz [10], [11] and Meyer [9] have given a rigorous content to this principle, the former by defining intrinsic stochastic differential equations in manifolds and the latter by establishing the approximation theorem in a very general setting. On the other hand, Meyer [8] has shown how to compute intrinsic Itô integrals in a manifold endowed with a connection. This leads to another transfer principle, transforming ordinary into Itô differential equations. We shall give an approximation scheme for this principle too, generalizing at the same time the approximate construction of Itô diffusions by Bismut [1] and that of Itô integrals of first order forms due to Duncan [4] and Darling [2].

These two transfer principles don't have the same properties. Whereas the Stratonovich one respects all submanifolds (that is, every submanifold preserved by the ordinary differential equation is also by the Stratonovich one), the Itô one respects only the totally geodesic ones. On the other hand, the Itô transfer requires less smoothness and extends better to operations depending upon t and ω; but it also requires a richer geometry: every manifold must be endowed with a connection.

The Itô transfer principle explains a posteriori the discovery by Meyer [9] of a correspondence between all stochastic extensions of the equation of parallel transport of vectors and all extensions to the tangent bundle TM of the connection on M; the stochastic parallel transports studied by Meyer are exactly the Itô extensions of the deterministic parallel transport, and they depend upon the choice of the connection in TM.

In the case when the ordinary differential equation transforms geodesics into geodesics, we shall see that the approximate constructions of the Stratonovich and Itô extensions are one and the same. As a consequence, the Stratonovich and Itô equations are identical, and the Stratonovich equation, being also an Itô one, transforms

(*) This work, written while visiting UBC and McGill University, stems from stimulating conversations with J.C. Taylor.

martingales into martingales. This can be used to explain why the development in a manifold of a Brownian motion (or, more generally, a martingale) in the tangent space is itself a Brownian motion (or a martingale), even though this development is defined using a Stratonovich differential equation, simply by noticing that the development of a straight line is a geodesic.

All the manifolds considered are real, finite-dimensional, of class $C^{2,\mathrm{Lip}}$ at least (all admissible changes of chart are C^2, with locally Lipschitz second derivatives), and arcwise connected. This last assumption is quite mild: since we shall be interested in manifold-valued, continuous, adapted processes, by conditioning on $\underline{\underline{F}}_0$, everything happens in an arcwise connected component of the manifold. The word "smooth" will mean "as smooth as possible", that is, having the same regularity as the manifold itself. When using local coordinates, the Einstein summation convention on once up, once down indices is in force.

I. SECOND-ORDER GEOMETRY

This section recalls a few fundamental definitions in Schwartz' second order geometry.

If x is a point in a manifold M, the <u>second-order tangent space</u> to M at x, denoted $\tau_x M$, is the vector space of all differential operators on M, at x, of order at most 2, with no constant term. If dim M = m, $\tau_x M$ has $m + \frac{1}{2}m(m+1)$ dimensions; using local coordinates $(x^i)_{1\le i\le m}$ near x, every $L \in \tau_x M$ can be written in a unique way $L = \ell^i D_i + \ell^{ij} D_{ij}$, with $\ell^{ij} = \ell^{ji}$, where $D_i = \frac{\partial}{\partial x^i}$ and $D_{ij} = \frac{\partial^2}{\partial x^i \partial x^j}$ are differential operators at x. The elements of $\tau_x M$ are called <u>second-order tangent vectors</u> (or tangent vectors of order 2)$^{(*)}$; the elements of the dual vector space $\tau_x^* M$ are called <u>second-order forms</u> (or <u>second-order covectors</u>) <u>at</u> x; a covector field of order 2 is simply called a <u>second-order form</u> on M.

If M and N are two manifolds, and if $\phi : M \to N$ is at least C^2, tangent vectors of order 2 are pushed-forward by ϕ : for $L \in \tau_x M$, it is possible to define $\vec{\phi}_x L \in \tau_{\phi(x)} N$ by $(\vec{\phi}_x L) f = L(f \circ \phi)$; dually, for $\theta \in \tau^*_{\phi(x)} N$, one can define the pulled-back $\overleftarrow{\phi}_x \theta \in \tau_x^* M$ by $\langle \overleftarrow{\phi}_x \theta, L\rangle = \langle \theta, \vec{\phi}_x L\rangle$ for all L. If x is a point in a submanifold M of a manifold N, one says that $L \in \tau_x N$ is tangent$^{(**)}$ to M if $L \in \vec{i}_x(\tau_x M)$, where $i : M \to N$ is the identity; this is equivalent to requiring Lu = 0 for every smooth $u : N \to \mathbb{R}$ such that u = 0 on M.

If $\gamma : I \to M$ is a twice differentiable curve in M (with I an open interval in $\mathbb{R}$),

---

$^{(*)}$ A shorter, but less informative name, could be "diffusors"; and forms of order 2 could be called "codiffusors".

$^{(**)}$ This definition is not ambiguous: it agrees with the classical one when L has order one.

for $\gamma \in I$ the <u>acceleration</u> $\ddot{\gamma}(t) \in \tau_{\gamma(t)} M$ is defined as $\vec{\gamma}_t \left(\frac{d^2}{ds^2}\right)$; in other words, for $f : M \to \mathbf{R}$, $\ddot{\gamma}(t)f = \frac{d^2}{dt^2} [f(\gamma(t))]$. Using local coordinates, one sees easily that every $L \in \tau_x M$ is a linear combination of accelerations of curves (the set of accelerations linearly spans all of $\tau_x M$); if L is tangent to a submanifold, these curves can be chosen in the submanifold.

Schwartz has noticed that, if X is a continuous semimartingale in M, the Itô differentials dX^i and $\frac{1}{2} d[X^i,X^j]$ (where $(x^i)_{1\le i\le m}$ is a local chart and X^i the i-th coordinate of X in this chart) behave formally in a change of coordinates as the coefficients of a second order tangent vector: the (purely formal) stochastic differential

$$\underline{dX}_t = dX_t^i \, D_i + \frac{1}{2} d[X^i,X^j]_t \, D_{ij}$$

is a (symbolic) second order tangent vector to M at $X_t(\omega)$. This is but a heuristic statement, but it has rigorous consequences, the foremost one being the possibility of integrating second-order forms along semimartingales: If X is a continuous semimartingale and θ a second-order form on M, the real semimartingale $\int\langle\theta,\underline{dX}\rangle$ can be defined; in local coordinates, $\langle\theta,\underline{dX}_t\rangle = \theta_i(X_t)dX_t^i + \frac{1}{2}\theta_{ij}(X_t)d[X^i,X^j]_t$ (where θ_i and θ_{ij} are the coefficients of θ in those coordinates). More generally, this extends to the case where θ is not everywhere defined, but only along the path of X, and may depend predictably upon t and ω. In this case, the above integrands $\theta_i(X_t)$ and $\theta_{ij}(X_t)$ must be replaced with the coefficients $\theta_{it}(\omega)$ and $\theta_{ijt}(\omega)$ of the predictable second-order form $\theta_t(\omega) \in \mathbb{T}^*_{X_t(\omega)}M$ (see Schwartz [10] prop. 2.7, Meyer [8] 4.6 or [6] 6.24).

To each $L \in \tau_x M$, written $\ell^i D_i + \ell^{ij} D_{ij}$ in local coordinates, is canonically associated the symmetric tensor $\hat{L} = \ell^{ij} D_i \otimes D_j \in T_x M \odot T_x M$, characterized intrinsically by

$$\langle\, df \otimes dg,\ \hat{L}\,\rangle = \frac{1}{2}\,[L(fg) - fLg - gLf].$$

If you know $\hat{L}$, L is determined up to terms of order 1, so the quotient vector space $\tau_x M/T_x M$ is canonically isomorphic to $T_x M \odot T_x M$. This can be illustrated with an exact sequence

$$0 \to T_x M \to \tau_x M \to T_x M \odot T_x M \to 0.$$

<u>Definition</u> Let M and N be manifolds, x be a point in M, y a point in N. A linear mapping $f : \tau_x M \to \tau_y N$ is called a <u>Schwartz morphism</u> if

(i) fL has order at most one if L has (equivalently : $\hat{L} = 0 \Rightarrow \widehat{fL} = 0$); let $f^1 : T_x M \to T_y N$ denote the restriction of f to $T_x M$;

(ii) for every $L \in \tau_x M$, $\widehat{fL} = (f^1 \otimes f^1)\hat{L}$.

The (non-linear) space of Schwartz morphisms from $\tau_x M$ to $\tau_y N$ will be denoted $SM_{xy}(M,N)$ (the first two letters stand for "Schwartz Morphisms"; only the second M is the name of the manifold!).

Remark An attempt to merge these two conditions into one could be $\widehat{f(AB)} = f(A) \odot f(B)$ where $A \in T_xM$ and B is a (first order) vector field near x, since the only possibility for this formula to make sense is by requiring that $f(A)$ and $f(B)$ are themselves first order vectors; but of course this is cheating!

These conditions (i) and (ii) can be restated as existence of a f^1 making the following diagram commutative:

$$\begin{array}{ccccccccc} 0 & \to & T_xM & \to & \tau_xM & \to & T_xM \odot T_xM & \to & 0 \\ & & \downarrow f^1 & & \downarrow f & & \downarrow f^1 \otimes f^1 & & \\ 0 & \to & T_yN & \to & \tau_yN & \to & T_yN \odot T_yN & \to & 0. \end{array}$$

In local coordinates $((x^i)_{1\le i\le m}$ near x, $(y^\alpha)_{1\le\alpha\le n}$ near y), a linear $f : \tau_xM \to \tau_yN$ is characterized by its coefficients f_i^α, f_{ij}^α, $f_i^{\alpha\beta}$, $f_{ij}^{\alpha\beta}$ (symmetric in i,j, or α,β, wherever possible), such that, if $L = \ell^i D_i + \ell^{ij} D_{ij} \in \tau_xM$,

$$fL = (f_i^\alpha \ell^i + f_{ij}^\alpha \ell^{ij})\, D_\alpha + (f_i^{\alpha\beta}\ell^i + f_{ij}^{\alpha\beta}\ell^{ij})\, D_{\alpha\beta} \in \tau_y N;$$

and f is a Schwartz morphism if and only if

$$\text{(SM)}\qquad \begin{cases} f_i^{\alpha\beta} = 0 \\ f_{ij}^{\alpha\beta} = \frac{1}{2}\,(f_i^\alpha f_j^\beta + f_j^\alpha f_i^\beta). \end{cases}$$

PROPOSITION 1. Given $x \in M$ and $y \in N$, a mapping $f : \tau_xM \to \tau_yN$ is a Schwartz morphism if and only if there exists a smooth $\phi : M \to N$, with $\phi(x) = y$ and $f = \vec{\phi}_x$.

PROOF In local coordinates,

$$\begin{aligned}\vec{\phi}_x L &= L\phi^\alpha\, D_\alpha + \langle d\phi^\alpha \otimes d\phi^\beta, \hat{L}\rangle\, D_{\alpha\beta} \\ &= [\ell^i D_i\phi^\alpha(x) + \ell^{ij} D_{ij}\phi^\alpha(x)]D_\alpha + \ell^{ij} D_i\phi^\alpha(x) D_j\phi^\beta(x) D_{\alpha\beta},\end{aligned}$$

so $f = \vec{\phi}_x$ is given by $f_i^\alpha = D_i\phi^\alpha(x)$, $f_{ij}^\alpha = D_{ij}\phi^\alpha(x)$ and by conditions (SM); it is a Schwartz morphism.

Conversely, if f is any Schwartz morphism, the same formula shows that, for a $\phi : M \to N$ with $\phi(x) = y$, $f = \vec{\phi}_x$ if and only if $D_i\phi^\alpha(x) = f_i^\alpha$ and $D_{ij}\phi^\alpha(x) = f_{ij}^\alpha$, and the proposition holds since it is always possible to construct a function with prescribed partial derivatives up to order 2 at one given point. ∎

COROLLARY 2. Let M, N, P be manifolds and $x \in M$, $y \in N$, $z \in P$. Let $f : \tau_xM \to \tau_zP$ be a Schwartz morphism and suppose $\psi : M \to N$ is a C^2 immersion at x with $\psi(x) = y$. There exists a Schwartz morphism $g : \tau_yN \to \tau_zP$ such that $f = g \circ \vec{\psi}_x$.

Proof. By Proposition 1, there is a $\phi : M \to P$ with $f = \vec{\phi}_x$. Since ψ is an immersion at x, there are a neighbourhood V of x in M and a C^2 function $\rho : N \to P$ with $\rho(y) = z$ such that $\phi = \rho \circ \psi$ on V. This gives $f = \vec{\phi}_x = \vec{\rho}_y \circ \vec{\psi}_x$; the result follows since, by Proposition 1 again, $\tau_y\rho$ is a Schwartz morphism. ∎

DEFINITION. Let M and N be manifolds, and P be a submanifold of the product M x N. One says that a Schwartz morphism $f \in SM_{xy}(M,N)$ is <u>constrained to</u> P if $(x,y) \in P$ and if ϕ in Proposition 1 can be chosen such that $(\xi,\phi(\xi)) \in P$ for every ξ in a neighbourhood of x in M.

For $(x,y) \in P$, the set of Schwartz morphisms from $\tau_x M$ to $\tau_y N$ constrained to P will be denoted $SM_{xy}(M,N;P)$. Remark that $SM_{xy}(M,N) = SM_{xy}(M,N;M{\times}N)$.

The next proposition, a characterization of constrained Schwartz morphisms, will make use of the following notations: for $L \in \tau_x M$ and $(x,y) \in M \times N$, $(L)_M \in \tau_{(x,y)}(M{\times}N)$ will be the differential operator defined by $(L)_M u = Lv$, where $v(\xi) = u(\xi,y)$; that is, by letting L act on the first variable only, the second one being kept fixed. Similarly, for $L \in \tau_y N$, one defines $(L)_N \in \tau_{(x,y)}(M \times N)$.

PROPOSITION 3. <u>Let</u> P <u>be a submanifold of</u> M x N, $(x,y) \in P$ <u>and</u> $f \in SM_{xy}(M,N)$. <u>The Schwartz morphism</u> f <u>is constrained to</u> P <u>if and only if, for every</u> $L \in \tau_x M$, <u>the second-order vector</u>(*) $(L)_M + (fL)_N + \hat{\hat{f}}L \in \tau_{(x,y)}(M \times N)$ <u>is tangent to</u> P, <u>where</u> $\hat{f}$ <u>is the linear mapping from</u> $T_x M \odot T_x M$ <u>to</u> $\tau_{(x,y)}(M \times N)$ <u>defined by</u>

$$\hat{f}(A \odot B) = (fA)_N(B)_M + (fB)_N(A)_M.$$

In the above product $(fA)_N(B)_M$, the two first-order differential operators $(fA)_N$ and $(B)_M$ act on independent variables, so they commute, and the product is well-defined even though each of them is only defined at the point (x,y).

PROOF. Let $L^f = (L)_M + (fL)_N + \hat{\hat{f}}L$.

First, suppose f is constrained to P, so there exists a smooth $\phi : M \to N$, with $\phi(x) = y$, $\vec{\phi}_x = f$ and $\text{graph}(\phi) \subset P$. Letting $\psi(\xi) = (\xi,\phi(\xi))$ define a smooth $\psi : M \to M \times N$, we shall show that, for every $L \in \tau_x M$, $L^f = \vec{\psi}_x L$; in other words, $L^f u = L(u \circ \psi)$ for every smooth u on M x N. As both sides depend linearly on L, it suffices to see it when L is the acceleration $\ddot{\gamma}(0)$ of some curve $\gamma : \mathbb{R} \to M$ with $\gamma(0) = x$. In that case, $\vec{\psi}_x L = \ddot{\Gamma}(0)$, where $\Gamma : \mathbb{R} \to M \times N$ is the curve $\psi \circ \gamma(t) = (\gamma(t),\delta(t))$, with $\delta = \phi \circ \gamma$. So its acceleration is the vector

$$u \to \frac{d^2}{dt^2}\Big|_{t=0} u(\gamma(t),\delta(t)),$$

giving $\vec{\psi}_x L = (\ddot{\gamma}(0))_M + (\ddot{\delta}(0))_N + 2(\dot{\gamma}(0))_M(\dot{\delta}(0))_N$. But $\ddot{\gamma}(0) = L$, $\ddot{\delta}(0) = fL$ and $\hat{L} = \dot{\gamma}(0) \odot \dot{\gamma}(0)$, so the last term is precisely $\hat{\hat{f}}L$, and $\vec{\psi}_x L = L^f$.

(*) A probabilistic interpretation of this vector will be given in terms of stochastic differential equations by Proposition 5.

Taking now any u that vanishes identically on P gives $u \circ \psi \equiv 0$, hence $L^f u = L(u \circ \psi) = 0$; this shows that L^f is tangent to P at (x,y).

Conversely, if L^f is tangent to P for each L, taking $L \in T_x M$ gives a $L^f \in T_{(x,y)}(M \times N)$ tangent to P, with first projection L. So the first projection $(\xi, z) \to \xi$ from P to M is a submersion at x, and, replacing if necessary M with a neighbourhood of x, we can suppose that the first projection is onto. By replacing N with a neighbourhood of y and choosing suitably the chart (y^α), the equations of P have the form

$$y^\alpha = e^\alpha(x^1,\dots,x^m,y^1,\dots,y^q) \qquad q < \alpha \le n$$

for some functions e^α of m + q variables(*). Letting $u^\alpha(\xi,\eta) = \eta^\alpha - e^\alpha(\xi,\eta^1,\dots,\eta^q)$ gives $u^\alpha = 0$ on P for $\alpha > q$, whence $L^f u^\alpha = 0$ for every L. This can be written

$$(*) \qquad \begin{cases} f_i^\alpha = f_i^\beta \, D_\beta e^\alpha + D_i e^\alpha \\ f_{ij}^\alpha = D_{ij} e^\alpha + f_{ij}^\beta \, D_\beta e^\alpha + 2 f_i^\beta \, D_{j\beta} e^\alpha + f_i^\beta \, f_j^\gamma \, D_{\beta\gamma} e^\alpha \end{cases}$$

(with $\alpha > q$ and the summation indices β and γ ranging from 1 to q).

Now choose any $\psi : M \to N$ such that $\vec{\psi}_x = f$ (this is possible because f is a Schwartz morphism). Define $\phi : M \to N$ by

$$\phi^\alpha(\xi) = \begin{cases} \psi^\alpha(\xi) & \text{if } \alpha \le q \\ e^\alpha(\xi,\psi^1(\xi),\dots,\psi^q(\xi)) & \text{if } \alpha > q. \end{cases}$$

As the graph of ϕ is included in P by construction, the proposition will be proved if we verify that $f = \vec{\phi}_x$. But the Schwartz morphism $g = \vec{\phi}_x$ is constrained to P; so (first part of this proof) $L^g u^\alpha = 0$ for every L, and g also verifies (*).
Since these formulae give, for $\alpha > q$, the coefficients f_i^α and f_{ij}^α in terms of f_i^β and f_{ij}^β with $\beta \le q$, and since $g_i^\beta = f_i^\beta$, $g_{ij}^\beta = f_{ij}^\beta$ for $\beta \le q$ by definition of ϕ, all coefficients of f and g agree. ∎

II. INTRINSIC STOCHASTIC DIFFERENTIAL EQUATIONS

Suppose given two manifolds, M and N, a filtered probability space $(\Omega, \underline{\underline{F}}, \mathbf{P}, (\underline{\underline{F}}_t)_{t\ge 0})$ verifying the usual completeness and right-continuity conditions, a M-valued semimartingale X with continuous paths, and a $\underline{\underline{F}}_0$-measurable, N-valued random variable y_0.

We are going to deal with a stochastic differential equation of the form

$$\text{(SDE)} \qquad \underline{dY}_t(\omega) = F(Y)_t(\omega) \, \underline{dX}_t(\omega), \quad Y_0 = y_0$$

where $\underline{dX}$ and $\underline{dY}$ are the symbolic Schwartz differentials of X and Y. Since, formally,

(*) If q = n, that is, if P is open in M x N, the result is trivial.

$\underline{dX}_t(\omega) \in \tau_{X_t(\omega)}M$ and $\underline{dY}_t(\omega) \in \tau_{Y_t(\omega)}N$, the coefficient $F(Y)_t(\omega)$ should be a linear mapping from $\tau_{X_t(\omega)}M$ to $\tau_{Y_t(\omega)}N$. So, in local coordinates, it will be given by coefficients F_i^α, F_{ij}^α, $F_i^{\alpha\beta}$, $F_{ij}^{\alpha\beta}$, all depending upon Y, t and ω. Now express $\underline{dX}$ and $\underline{dY}$ in local coordinates to transform (SDE) into the system

$$\begin{cases} dY^\alpha = F_i^\alpha(Y)dX^i + \frac{1}{2} F_{ij}^\alpha(Y)\ d[X^i,X^j] \\ \frac{1}{2} d[Y^\alpha,Y^\beta] = F_i^{\alpha\beta}(Y)dX^i + \frac{1}{2} F_{ij}^{\alpha\beta}(Y)\ d[X^i,X^j]. \end{cases}$$

But this system is overdetermined: the rules of stochastic calculus make it possible to compute $\frac{1}{2} d[Y^\alpha,Y^\beta]$ from the differential dY^α and dY^β; more precisely, the first equation(s) implies

$$d[Y^\alpha,Y^\beta] = F_i^\alpha(Y)\ F_j^\beta(Y)\ d[X^i,X^j].$$

To make this compatible with the second equation(s), it is reasonable to assume that

$$F_i^{\alpha\beta} = 0\ ,\ F_{ij}^{\alpha\beta} = \frac{1}{2}\ (F_i^\alpha\ F_j^\beta + F_j^\alpha\ F_i^\beta),$$

or, equivalently, that <u>each</u> $F(Y)_t(\omega)$ <u>is a Schwartz morphism</u>.

With a Lipschitz hypothesis on F, we shall state and prove an existence and uniqueness theorem for equations of this type. Although the proof consists only in extending to manifolds results that are well-known in the vector case, it is long and boring; so it is worth trying to maximize the efficiency of the theorem by gaining generality, and we shall also take into account the case when the solution Y remains linked to X by one or more relation. Technically, this is done by considering a closed submanifold P of M x N and considering only Schwartz morphisms $F(Y)_t(\omega)$ that are constrained to P. This situation arises, for instance, when the stochastic differential equation represents a lifting of X in some fiber bundle N above M; in that case, the solution Y has to live above X, the equation is defined for those Y only, and the constraint P is the submanifold of M x N consisting of the points (x,y) such that y is above x.

THEOREM 4. <u>Given</u> M, N, Ω, $\underline{F}$, P, $(\underline{F}_t)$, X, y_0 <u>as above</u>, <u>let</u> P <u>be a closed submanifold of</u> M x N, <u>and suppose that</u> $(X_0,Y_0) \in P$. <u>For every predictable time</u> ζ <u>and every</u> N-<u>valued</u>, <u>continuous semimartingale</u> $\mathbf{Y}$ <u>with</u> $Y_0 = y_0$, <u>defined on</u> $[[0,\zeta[[$, <u>and verifying</u> $(X,Y) \in P$ <u>in this interval</u>, <u>suppose given a predictable process</u> F(Y), <u>also defined on</u> $[[0,\zeta[[$, <u>such that</u>

(i) <u>for every</u> $(\omega,t) \in [[0,\zeta[[$,

$$F(Y)_t(\omega) \in SM_{X_t(\omega)\ Y_t(\omega)}\ (M,N;P);$$

(ii) F(Y) <u>is locally bounded</u> : <u>there are stopping times</u> T_n <u>with limit</u> ζ <u>such that the image by</u> F(Y) <u>of each random interval</u> $[[0,T_n]] \cap (\{T_n>0\}\times\mathbb{R}_+)$ <u>is relatively compact</u> (<u>in the manifold</u> SM(M,N;P));

(iii) F <u>is non-anticipating</u> : <u>for any predictable time</u> T, <u>the restriction of</u> F(Y) <u>to</u> $[[0,\zeta[[\cap[[0,T]]$ <u>depends only upon the restriction of</u> Y <u>to this interval</u>;

(iv) F <u>is local</u>(*): <u>for every non-negligible</u> $A \in F$, <u>the restriction of</u> F(Y) to $[[0,\zeta[[\cap(A \times \mathbb{R}_+)$ <u>depends only upon the restriction of</u> Y <u>to this set</u>;

(v) F <u>is locally Lipschitz</u> : <u>for every compact</u> $K \subset N$, <u>there exists a measurable</u> (<u>not necessarily adapted</u>) <u>increasing process</u> $L(K,t,\omega)$ <u>such that</u>, <u>if</u> $Y'_s(\omega)$ <u>and</u> $Y''_s(\omega)$ <u>are in</u> K <u>for</u> $0 \le s \le t$,

$$d(F(Y')_t(\omega),\ F(Y'')_t(\omega)) \le L(K,t,\omega) \sup_{0\le s\le t} d(Y'_s(\omega),\ Y''_s(\omega)).$$

<u>There exists a unique pair</u> (Y,ζ) <u>as above</u>, <u>with</u> $0 < \zeta \le \infty$, <u>such that</u> Y <u>explodes at time</u> ζ <u>if</u> ζ <u>is finite</u> (<u>i.e. the path</u> $(Y_t(\omega))_{t<\zeta(\omega)}$ <u>is not relatively compact in</u> N) <u>and verifies on</u> $[[0,\zeta[[$ <u>the stochastic differential equation</u>

$$\underline{d}Y_t(\omega) = F(Y)_t(\omega)\ \underline{d}X_t(\omega)$$

(this means, for every smooth second-order form θ on N,

$$\int \langle \theta,\ \underline{d}Y \rangle = \int \langle F(Y)^*\theta(Y),\ \underline{d}X \rangle).$$

<u>Moreover</u>, <u>if</u> (Y',ζ') <u>is another solution to this equation starting from the same initial condition</u> y_0, <u>then</u> $\zeta' \le \zeta$ <u>and</u> $Y' = Y$ <u>on</u> $[[0,\zeta'[[$.

REMARKS. In hypothesis (v), d denotes any Riemannian distances on the manifolds N and SM(M,N;P); the statement does not depend on the specific choice of d since the ratio of any two Riemannian distances on a manifold is always bounded above and below on compact sets.

Hypothesis (iv) is used only once, to transform the process L(K,t) in (v) into a deterministic process. When L(K,t) does not depend on ω, that step is not necessary, and the result holds without assuming (iv).

PROOF.

<u>First step</u>: The theorem is true with the additional assumptions that $M = \mathbb{R}^m$, $N = \mathbb{R}^n$, $P = M \times N$ (that is, no constraint at all).

Taking the canonical global coordinates $(x^i)_{1\le i\le m}$ and $(y^\alpha)_{1\le\alpha\le n}$ on M and N transforms the given equation into a system

$$\begin{cases} dY^\alpha_t = F^\alpha_i(Y)_t\, dX^i_t + \frac{1}{2} F^\alpha_{ij}(Y)_t\, d[X^i,X^j]_t \\ \frac{1}{2}\, d[Y^\alpha,Y^\beta]_t = F^{\alpha\beta}_i(Y)_t\, dX^i_t + \frac{1}{2} F^{\alpha\beta}_{ij}(Y)_t\, d[X^i,X^j]_t . \end{cases}$$

As observed above, the last n^2 equations are a consequence of the first n ones and of $F(Y)_t$ being a Schwartz morphism; so we may forget about them.

(*) This hypothesis is not necessary if the increasing processes L(K,t) in (v) are deterministic.

Let $\psi_p : [0,\infty) \to [0,1]$ be compactly supported in $[0,p]$ and equal to 1 on $[0,p-1]$. Define a new system of equations

$$(*) \qquad dY_t^\alpha = G_i^{p\alpha}(Y)_t + \frac{1}{2} G_{ij}^{p\alpha}(Y)_t \, d[X^i,X^j]_t$$

by $G_i^{p\alpha}(Y)_t = \psi_p(\sup_{0\le s\le t} \|Y_s\|) F_i^\alpha(Y)_t$, $G_{ij}^{p\alpha}(Y)_t = \psi_p(\sup_{0\le s\le t} \|Y_s\|) F_{ij}^\alpha(Y)_t$.

For each p, this new system is globally Lipschitz in space. Indeed, supposing $\sup_{0\le s\le t} \|Y'_s\| \ge \sup_{0\le s\le t} \|Y''_s\|$ (else, exchange Y' and Y"),

$$\begin{aligned}
|G_i^{p\alpha}(Y')_t - G_i^{p\alpha}(Y'')_t| &\le \psi_p(\sup_{s\le t}\|Y'_s\|)\,|F_i^\alpha(Y')_t - F_i^\alpha(Y'')_t| \\
&\quad + |\psi_p(\sup_{s\le t}\|Y'_s\|) - \psi_p(\sup_{s\le t}\|Y''_s\|)|\;|F_i^\alpha(Y'')_t| \\
&\le L(\bar{B}(p),t,\omega)\, \sup_{s\le t}\|Y'_s - Y''_s\| \\
&\quad + \sup|\psi'|\,(\sup_{s\le t}\|Y'_s\| - \sup_{s\le t}\|Y''_s\|)\, L(\bar{B}(p),t,\omega) \\
&\qquad (|F_i^\alpha(0)_t| + p) \\
&\le L'(t,\omega) \sup_{0\le s\le t}\|Y'_s - Y''_s\|
\end{aligned}$$

with $L'(t,\omega) = L(\bar{B}(p),t,\omega)\ [1+\sup|\psi'|\ (p+\sup_{0\le s\le t}|F_i^\alpha(0)_s|)]$; and similarly for G_{ij}^{pk}.

For $t,q > 0$, let $\Omega_{tq} = \{\omega : L'(t,\omega) \le q\} \in \underline{\underline{F}}$. Since F, and hence also G^p, is local, it is possible to solve the globally Lipschitz system (*) on $\Omega_{tq} \times [0,t]$ with the given initial condition y_0 (see Métivier [7]); and for $q_1 < q_2$ the solutions agree on $\Omega_{tq_1} \times [0,t]$ by uniqueness. Letting $q \to \infty$ shows that (*) has a unique solution on $[[0,t]]$. Similarly, letting $t \to \infty$ and using the non-anticipation assumption gives a unique solution to (*) on $\Omega \times \mathbb{R}_+$, starting from y_0. Let Y(p) denote this solution.

If $T(p) = \inf\{t : \|Y(p)_t\| \ge p-1\}$, $Y(p)^{|T(p)}$ is a solution to (*) with X replaced by $X^{|T(p)}$, so it is also a solution to $\underline{dY} = F(Y)\ \underline{dX}^{|T(p)}$; conversely, if Y is any solution to $\underline{dY} = F(Y)\ \underline{dX}$ starting from y_0 and if S(p) is the first time when $\|Y\| \ge p-1$, then, on $[[0,T(p)\wedge S(p)]]$, Y and Y(p) are two solutions of (*), hence $Y = Y(p)$ on this interval, and $S(p) = T(p)$. This implies that $T(p) \le T(p+1)$ and $Y(p) = Y(p+1)$ on $[[0,T(p)]]$. So letting $\zeta = \sup_{p\in\mathbb{N}} T(p)$, the conclusion of Theorem 4 holds; ζ is predictable as the explosion time of the continuous, adapted process Y.

<u>Second step</u>: We still assume $M = \mathbb{R}^m$ and $N = \mathbb{R}^n$, but P is now a closed submanifold in $M \times N$ (it is not arbitrary: the very existence of the Schwartz morphism $F(Y)_t(\omega)$ constrained to P implies that the projection of P on M contains a neighbourhood of

$X_t(\omega))$.

We are now given F(Y) only for those semimartingales Y such that (X,Y) is P-valued; we shall first extend the definition of F(Y) to all N-valued continuous semimartingales. Let $\rho : M \times N \to N$ denote the second projection.

There exists an open neighbourhood Q of P in M x N such that the mapping $\pi : Q \to P$ with $\pi(z)$ the point of P closest to z (for the Euclidean distance on M x N) is well-defined and smooth on Q.(*) Let $\psi : M \times N \to [0,1]$ be smooth, with $\psi = 1$ on P and support $\psi \subset Q$. For every N-valued continuous semimartingale Y defined on some $[[0,\zeta[[$, letting $Z = (X,Y)$, define, for $(\omega,t) \in [[0,\zeta[[$

$$G_i^\alpha(Y)_t = F_i^\alpha(\rho\pi Z)_t \inf_{0\le s\le t} \psi(Z_s)$$

$$G_{ij}^\alpha(Y)_t = F_{ij}^\alpha(\rho\pi Z)_t \inf_{0\le s\le t} \psi(Z_s)$$

with the convention " undefined x 0 = 0".

This G is an extension of F to all N-valued continuous semimartingales. Each G(Y) is a locally bounded, predictable process in SM(M,N), above (X,Y); clearly, G is also non-anticipating and local. It is also locally Lipschitz for, if Y' and Y" are semimartingales in N, taking their values in a compact K, letting $C(t,\omega)$ denote the compact $\{X_s(\omega),\ 0\le s\le t\} \subset M$ and $\Gamma(t,\omega)$ the compact $[C(t,\omega) \times K] \cap \text{support}(\psi)$, one has, if for instance $\inf_{0\le s\le t} \psi(Z'_s) \le \inf_{0\le s\le t} \psi(Z''_s)$,

$$|G_i^\alpha(Y')_t - G_i^\alpha(Y'')_t| \le |F_i^\alpha(\rho\pi Z')_t - F_i^\alpha(\rho\pi Z'')_t| \inf_{s\le t} \psi(Z'_s)$$

$$+ |\inf_{s\le t} \psi(Z'_s) - \inf_{s\le t} \psi(Z''_s)|\ |F_i^\alpha(\rho\pi Z'')_t|$$

$$\le L(\rho\pi\Gamma(t,\omega),t,\omega)\ \Lambda(\Gamma(t,\omega)) \sup_{s\le t} \|Y'_s - Y''_s\|$$

$$+ \sup_{\Gamma(t,\omega)} \|\nabla\psi\| \sup_{s\le t} \|Y'_s - Y''_s\|\ |F_i^\alpha(\rho\pi Z'')_t|$$

(where $\Lambda(\Gamma)$ is a Lipschitz constant for $\rho\pi$ on the compact Γ), and the last factor $|F_i^\alpha(\rho\pi Z'')_t|$ is estimated by

$$|F_i^\alpha(\rho\pi Z'')_t| \le L(\rho\pi\Gamma(t,\omega),t,\omega)\ \Lambda(\Gamma(t,\omega))\ [\sup_{s\le t} |F_i^\alpha(y_0)_s| + \text{diam}(K)],$$

using the fact that $F_i^\alpha(y_0)$ is locally bounded.

So the first step, applied to this G, shows the existence of a unique Y in N, solution to $\underline{dY} = G(Y)\ \underline{dX}$, exploding at some time ζ. Using the hypothesis that F is constrained to P, we proceed to show that the (M x N)-valued process Z = (X,Y) spends all its life-time $[[0,\zeta[[$ in P.

If such is not the case, there is a stopping time $T < \zeta$ with $Z \in P$ on $[[0,T]]$ and

$$\mathbb{P}[T = \inf\{t : Z_t \notin P\}] > 0.$$

Without loss of generality, it is possible to suppose that T = 0 (define $\tilde{\underline{F}}_t = \underline{F}_{T+t}$,

(*) This argument corrects a mistake in Emery [6]: I erroneously assumed the existence of a neighbourhood of P diffeomorphic to P x $\mathbb{R}^q$, but there may be topological obstructions to this.

$\tilde{X}_t = X_{T+t}$ and, for a N-valued $(\underline{\underline{\tilde{F}}}_t)$-semimartingale $\tilde{Y}$ with $\tilde{Y}_0 = Y_T$,

$$\bar{Y}_t = \begin{cases} Y_t & \text{if } t \le T \\ \tilde{Y}_{t-T} & \text{if } t \ge T \end{cases}$$

and $\tilde{F}(\tilde{Y})_t = F(\bar{Y})_{T+t}$; all the hypotheses are preserved by this time-translation).

Let P' be the open subset of P consisting of the points z such that the first projection $p : P \to M$ is a submersion at z; by hypothesis (i), $SM_{X_0,Y_0}(M,N;P)$ is not empty and so (X_0,Y_0) is in P'. But P' is the union of countably many open sets, each of them diffeomorphic to a product M' x R, with M' open in M and the first projection preserved by the diffeomorphism.

Hence, for one of these open sets, say D, the $\underline{\underline{F}}_0$-event $\{Z_0 \in D \text{ and } \inf\{t : Z_t \notin P\} = 0\}$ is not negligible; conditioning on it allows us to suppose it has probability 1. Call δ the diffeomorphism from M' x R to D.

The equation $\underline{dY} = F(Y)\ \underline{dX}$ will now be transformed, using this δ, into an equation $\underline{dU} = H(U)\ \underline{dX}$ with unknown U in R. If U is a R-valued continuous semimartingale with $\delta(X_0,U_0) = (X_0,Y_0)$, define $\bar{Y}$ in N by $(X,\bar{Y}) = \delta(X,U)$ (that is, $\bar{Y} = \rho\delta(X,U)$), and, for fixed t and ω, $\phi : M \to N$ such that $(\xi,\phi(\xi)) \in D$ for ξ close enough to $X_t(\omega)$ and

$$\vec{\phi}_{X_t(\omega)} = F(\bar{Y})_t(\omega) : \tau_{X_t(\omega)}M \to \tau_{\bar{Y}_t(\omega)}N;$$

define $\psi : M' \to R$ by $\delta(\xi,\psi(\xi)) = (\xi,\phi(\xi))$ and call $H(U)_t(\omega)$ the Schwartz morphism $\vec{\psi}_{X_t(\omega)} \in SM_{X_t(\omega)U_t(\omega)}(M,R)$.

By a bicontinuous time-change, the first time when X exits from M' can be made infinite, and the first step of the proof, applied this time to equation $\underline{dU} = H(U)\ \underline{dX}$, produces a solution U on some interval $[[0,\zeta_U[[$ with $\zeta_U > 0$.

Now the N-valued $\bar{Y} = \rho\delta(X,U)$ verifies

$$\underline{d\bar{Y}} = \underline{d(\rho\delta(X,U))} = \overrightarrow{\rho\delta\circ(\mathrm{Id},\psi)}\ \underline{dX} = \vec{\phi}\ \underline{dX} = F(\bar{Y})\ \underline{dX},$$

so it is also a solution to $\underline{d\bar{Y}} = G(\bar{Y})\ \underline{dX}$ and it must agree with Y; but Y leaves P at time 0 whereas $\bar{Y}$ does not, giving the required contradiction.

Third step: Getting rid of the hypothesis that N is a vector space.

Since N is arcwise connected, it is paracompact, and hence it can be imbedded as a closed submanifold of some $\mathbf{R}^n$. So M x N is imbedded in M x $\mathbf{R}^n$, and P is a closed submanifold of M x $\mathbf{R}^n$. Denote by $i : N \to \mathbf{R}^n$ and $j : P \to M \times \mathbf{R}^n$ those imbeddings; $j(x,y) = (x,iy)$.

For each $f \in SM_{xy}(M,N;P)$, let $\mathring{f} : \tau_x M \to \tau_y \mathbf{R}^n$ be defined by $\mathring{f} : \vec{i}\circ f$. This is a Schwartz morphism constrained to jP. Indeed, there is a $\phi : M \to N$ with $(\xi,\phi(\xi)) \in P$ and $\vec{\phi} = f$; for $\psi = i\circ\phi$, $(\xi,\psi(\xi)) \in jP$ and $\mathring{f} = \vec{\psi}$. So the equation $\underline{dY} = F(Y)\ \underline{dX}$ can be transformed into an equation $\underline{dZ} = \mathring{F}(Z)\ \underline{dX}$, with unknown Z in $\mathbf{R}^n$, constrained to jP; and $Y = j^{-1}Z$ is the unique solution to the given equation; since jP is closed in M x $\mathbf{R}^n$, both Y and Z explode at the same time.

Last step: Removing the assumption $M = \mathbb{R}^m$.

By Lemma (3.5) of [6], there is an increasing sequence of predictable times T_k, with $T_0 = 0$ and $\sup_k T_k = \infty$, such that, on each interval $[[T_k, T_{k+1}]]'$, X remains in the domain of some local chart (we denote by $[[S,T]]'$ the interval $[[S,T]] \cap (\{T>S\} \times \mathbb{R}_+)$, equal to $[S,T]$ if $S < T$ and empty if $S \geq T$). By induction on k, the equation $\underline{dY} = F(Y)\ \underline{dX}$ has a unique solution on $[[0,T_k]]$ (with a possible explosion). Indeed, supposing that this holds on $[[0,T_k]]$, letting $\tilde{\Omega} = \{T_{k+1} > T_k,\ \zeta > T_k\} \in \underline{\underline{F}}_{T_k}$, $\tilde{\underline{\underline{F}}}_t = \underline{\underline{F}}_{T_k+t}$, $\tilde{X}_t = X_{T_k+t}$ (on the interval $[[0,T_{k+1}-T_k]]'$, this process lives in the domain of a local chart, so we may see it as $\mathbb{R}^m$-valued), $\tilde{F}(\tilde{Y})_t = F(Y)_{T_k+t}$, where

$$Y_t = \begin{cases} \text{the solution to } \underline{dY} = F(Y)\ \underline{dX} \text{ on } [[0,T_k]] \\ \tilde{Y}_{t-T_k} \text{ if } t \geq T_k \end{cases}$$

(this is defined only for $\tilde{Y}_0 = Y_{T_k} \in \tilde{\underline{\underline{F}}}_0$) gives a solution $\tilde{Y}$ on $[[0,T_{k+1}-T_k]]'$ (with a possible explosion); and the process equal to

$$\begin{cases} Y_t \text{ if } t \leq T_k \\ \tilde{Y}_{t-T_k} \text{ if } T_k \leq t \leq T_{k+1} \text{ and } \omega \in \tilde{\Omega} \end{cases}$$

is the unique solution to $\underline{dY} = F(Y)\ \underline{dX}$ on $[[0,T_{k+1}]]$. As its restriction to $[[0,T_k]]$ is the solution to the same equation on the latter interval, these processes can be patched up together, thus proving the theorem. ∎

Given $L \in \tau_x M$ and $f \in SM_{xy}(M,N)$, the second-order vector

$$L^f = (L)_M + (fL)_N + \hat{\hat{fL}} \in \tau_{(x,y)}(M,N)$$

introduced in Proposition 3 can be given an interpretation in terms of stochastic differential equations.

PROPOSITION 5. a) Given $f \in SM_{xy}(M,N)$, the linear mapping $\bar{f} : \tau_x M \to \tau_{(x,y)}(M \times N)$ defined by $\bar{f}L = L^f$ is a Schwartz morphism : $\bar{f} \in SM_{x,(x,y)}(M, M \times N)$.

b) Let (X,Y) be a continuous semimartingale in $M \times N$ and F be a locally bounded, predictable process in $SM(M,N)$ such that for all t and ω, $F_t(\omega) \in SM_{X_t(\omega)Y_t(\omega)}(M,N)$.
The stochastic differential equations

$$\underline{dY}_t(\omega) = F_t(\omega)\ \underline{dX}_t(\omega)$$

and

$$\underline{d(X,Y)}_t(\omega) = \bar{F}_t(\omega)\ \underline{dX}_t(\omega)$$

are equivalent : Y solves the former if and only if (X,Y) solves the latter.

PROOF. a) By Proposition 1, there is a smooth $\phi : M \to N$ with $\phi(x) = y$ and $f = \vec{\phi}_x$. Define $\psi : M \to M \times N$ by $\psi(\xi) = (\xi, \phi(\xi))$. We have seen, in the proof of Proposition 3, that $\bar{f} = \vec{\psi}_x$; so applying Proposition 1 again gives the result.

b) Of course, the rigorous meaning of $\underline{dY} = F\ \underline{dX}$ is that, for every smooth second order form θ on N,

$$\int \langle \theta, \underline{dY} \rangle = \int \langle F^*\theta(Y), \underline{dX} \rangle;$$

and similarly for $\underline{d(X,Y)} = \bar{F}\ \underline{dX}$. In local coordinates $((x^i)$ on M, (y^α) on N), using the fact that F and $\bar{F}$ are Schwartz morphisms, $\underline{dY} = F\ \underline{dX}$ is equivalent to

$$dY^\alpha = F_i^\alpha\, dX^i + \frac{1}{2} F_{ij}^\alpha\, d[X^i,X^j]$$

and $\underline{d(X,Y)} = \bar{F}\ \underline{dX}$ to

$$\begin{cases} dX^k = \bar{F}_i^k\, dX^i + \frac{1}{2} \bar{F}_{ij}^k\, d[X^i,X^j] \\ dY^\alpha = \bar{F}_i^\alpha\, dX^i + \frac{1}{2} \bar{F}_{ij}^\alpha\, d[X^i,X^j], \end{cases}$$

and if suffices to check that $\bar{F}_i^k = \delta_i^k$, $\bar{F}_{ij}^k = 0$, $\bar{F}_i^\alpha = F_i^\alpha$ and $\bar{F}_{ij}^\alpha = F_{ij}^\alpha$. These formulae are direct consequences of

$$\bar{F}(D_i) = (D_i)_M + (FD_i)_N$$
$$\bar{F}(D_{ij}) = (D_{ij})_M + (FD_{ij})_N + \frac{1}{2}\hat{F}(D_i \odot D_j)$$

and of the fact that $\hat{F}(D_i \odot D_j)$ is in the vector space spanned by the $D_{k\alpha}$'s and does not contribute to $\bar{F}_{ij}^k$ nor $\bar{F}_{ij}^\alpha$. ∎

III. ORDINARY DIFFERENTIAL EQUATIONS

Since our goal is to transform deterministic geometric constructions into stochastic ones, this section describes those deterministic operations. Everything is similar to what has been seen in the stochastic case, and may be much simpler, since only first-order geometry is involved.

Let M and N be manifolds. The vector space $T_x^*M \otimes T_yN$ of all linear maps from T_xM to T_yN will also be denoted by $L_{xy}(M,N)$; remark that Proposition 1 has no interesting analogue at order 1, since every element of $L_{xy}(M,N)$ has the form $\vec{\phi}_x$ for a smooth $\phi : M \to N$ with $\phi(x) = y$. If P is a submanifold of M x N and (x,y) a point in P, a linear $e : T_xM \to T_yN$ (that is, an element of $L_{xy}(M,N)$) is said to be constrained to P if there exists a smooth $\phi : M \to N$, with $\phi(x) = y$, $\vec{\phi}_x = e$, and $(\xi,\phi(\xi)) \in P$ for ξ close enough to x in M. The so-defined (affine) subset of $L_{xy}(M,N)$ will be denoted $L_{xy}(M,N;P)$. Of course, the analogue of Proposition 3 holds.

PROPOSITION 6. *Let P be a submanifold of M x N, (x,y) a point in P, and $e \in L_{xy}(MxN)$. The linear mapping e is constrained to P if and only if, for every $A \in T_xM$, the tangent vector (of order 1) $(A)_M + (eA)_N \in T_{xy}(MxN)$ is tangent to the submanifold P.*

The proof is quite similar to that of Proposition 3, with simpler computations since second-order terms are no longer considered; so we omit it.

THEOREM 7. *Suppose given two manifolds M and N, a closed submanifold P of MxN, a curve $(x(t))_{t\geq 0}$ of class C^1 in M and a point y_0 in N, with $(x(0),y_0) \in P$. For every*

$0 < \zeta \le \infty$ and every C^1 curve $(y(t))_{0\le t<\zeta}$ with $y(0) = y_0$ and $(x(t),y(t)) \in P$ for $t < \zeta$, suppose given a family $(e(y)_t)_{0\le t<\zeta}$ such that

(i) $e(y)_t$ is in $L_{x(t)y(t)}(M,N;P)$;

(ii) $e(y)$ is locally bounded: for $\epsilon > 0$, the set $\{e(y)_s,\ 0 \le s \le \zeta - \epsilon\}$ is relatively compact in the manifold $\bigcup_{x,y} T^*_x M \otimes T_y N$;

(iii) e is non-anticipating: for each $t < \zeta$, the restriction of $e(y)$ to $[0,t]$ depends only upon the restriction of y to $[0,t]$;

(iv) e is locally Lipschitz: for every compact $K \subset N$ there is an increasing function $L(K,t)$ such that, if $y'(s)$ and $y''(s)$ are in K for $0 \le s \le t$, then

$$d(e(y')_t, e(y'')_t) \le L(K,t) \sup_{0\le s\le t} d(y'(s),y''(s)).$$

There exists a unique pair (y,ζ) as above, with $0 < \zeta \le \infty$, such that y explodes at time ζ if $\zeta < \infty$ and verifies on $[0,\zeta)$ the ordinary differential equation

$$\dot{y}(t) = e(y)_t\ \dot{x}(t)$$

Moreover, uniqueness holds for this equation: for every $0 < \zeta' \le \infty$ and every curve $(y'(t))_{0\le t<\zeta'}$ with $y'(0) = y_0$ and $(x(t),y'(t)) \in P$ verifying $\dot{y}'(t) = e(y')_t\ \dot{x}(t)$, one has $\zeta' \le \zeta$ and $y' = y$ on $[0,\zeta')$.

The proof is very similar to that of Theorem 4 (with simpler computations since first order geometry only is involved); we omit it.

IV. THE STRATONOVICH TRANSFER PRINCIPLE

This section deals with transforming an ordinary differential equation between manifolds into a stochastic one, via Stratonovich stochastic calculus. This is, of course, quite classical and has been extensively used by many authors (a typical example is Bismut [1]), mostly in the framework of Brownian motions or diffusions; it was extended to manifold-valued semimartingales by Schwartz [10] and Meyer [8]. The setting chosen here is borrowed from [6]; the only new feature is the constraint P. Notice that the ordinary differential equations to be transferred by Theorem 8 below are much less general than those considered in Theorem 7. This is the main weakness of the Stratonovich transfer principle: it requires some smoothness(*) and the coefficients in the equation should not depend on the past values of the curves considered (though time itself can be incorporated in these curves, by the usual space-time trick).

If X is a continuous semimartingale in a manifold M, and if we are given for each t and ω a 1-form $\Psi_t(\omega)$ on M at $X_t(\omega)$ (that is, $\Psi_t(\omega) \in T^*_{X_t(\omega)}M$), such that the T^*M-valued process Ψ is a continuous semimartingale, it is possible to define the Stratonovich integral $\int_0^t \Psi_s\ \delta X_s$ as a real semimartingale (see [8] or [10]). It is characterized by the two properties, where f is an arbitrary smooth function on M

$$\int_0^t df(X_s)\ \delta X_s = f(X_t) - f(X_0);$$

(*) but T. Lyons told me that, for reversible Dirichlet processes X, the Stratonovich integal $\int_0^1 f(X)\delta X$ can be defined for any bounded, Borel f.

$$\text{if } I_t = \int_0^t \Psi_s \delta X_s, \text{ then } \int_0^t [f(X_s)\Psi_s]\delta X_s = \int_0^t f(X_s)\ \delta I_s;$$

the last integral is a Stratonovich integral of real semimartingales.

This makes it possible to give a meaning to Stratonovich stochastic differential equations of the form

$$\delta Y = e(Y)\delta X$$

where X is a given M-valued, continuous semimartingale, the unknown Y is a N-valued continuous semimartingale, and $e(Y)_t(\omega)$ is a linear mapping from $T_{X_t(\omega)}M$ to $T_{Y_t(\omega)}N$: A solution Y to this equation is a semimartingale Y such that, for every 1-form α on N, the Stratonovich integrals $\int\alpha(Y)\delta Y$ and $\int[e^*(Y)\ \alpha(Y)]\ \delta X$ exist and are equal ($e^*(Y)$ is the adjoint of $e(Y)$, so it transforms $\alpha(Y_t) \in T^*_{Y_t}N$ into an element of $T^*_{X_t}M$).

The Stratonovich transfer principle for equations between manifolds of the type considered here can now be stated.

THEOREM 8 (Stratonovich transfer principle). Let M and N be manifolds, P be a closed submanifold of MxN, and, for each $(x,y)\in P$, $e(x,y)$ be in $L_{xy}(M,N;P)$. Suppose that the mapping $e:P \to L(M,N;P)$ is of class $C^{1,\mathrm{Lip}}$.

There exists a unique family $(f(x,y))_{(x,y)\in P}$, where each $f(x,y)$ is a linear mapping from $\tau_x M$ to $\tau_y N$, such that for every curve $(x(t),y(t))$ of class C^2 in P verifying the ordinary differential equation $\dot{y}(t) = e(x(t),y(t))\ \dot{x}(t)$, one has also $\ddot{y}(t) = f(x(t),y(t))\ \ddot{x}(t)$.

Moreover, each $f(x,y)$ is a constrained Schwartz morphism : $f(x,y) \in SM_{xy}(M,N;P)$, depending in a locally Lipschitz fashion upon (x,y); and the intrinsic stochastic differential equations

$$\underline{dY}_t = f(X_t,Y_t)\ \underline{dX}_t$$

and

$$\delta Y_t = e(X_t,Y_t)\ \delta X_t$$

are equivalent: given X and Y_0, every solution Y to one of them is also a solution to the other.

Remark that Theorem 4 applies here, showing existence and uniqueness of the solution of $\underline{dY} = f(X,Y)\underline{dX}$; so these existence and uniqueness properties transfer to the Stratonovich equation $\delta Y = e(X,Y)\delta X$.

PROOF. The case when P = MxN (no constraint at all) is proved in (7.22) of [6]; so we just have to reduce the general case to that one. [Observe that the proof of this particular case consists, first in computing the f such that $\dot{y} = e\dot{x}$ implies $\ddot{y} = f\ddot{x}$, second in computing the f such that $\underline{dY} = f\underline{dX}$ and $\delta Y = e\delta X$ are equivalent, and finally in verifying that both results agree. But not only are both results the same: the computations are step by step identical; and this suggests that some computation-free proof might predict that both f agree without actually calculating them.]

In the general case, since P is closed, it is possible to extend the given family $(e(x,y))_{(x,y)\in P}$ into a family $(\tilde{e}(x,y))_{x\in M, y\in N}$ such that $\tilde{e} = e$ on P, $\tilde{e}(x,y) \in L_{xy}(M,N) = T_x^*M \otimes T_y N$, and $\tilde{e}$ is of class $C^{1,\mathrm{Lip}}$. [This can be done, for instance, using a partition of unity (ϕ_α) of some neighbourhood of P in MxN such that each ϕ_α is compactly supported in a domain D_α with the following two properties: D_α is included in a product $D'_\alpha x D''_\alpha$, with D'_α (respectively D''_α) the domain of some local chart in M (respectively N), and D_α is diffeomorphic to a vector space, with $P \cap D_\alpha$ corresponding to a linear subspace. Using these local coordinates, extend the restrictions e_α of e to D_α into some $\tilde{e}_\alpha$ defined in D_α, and set $\tilde{e} = \sum_\alpha \phi_\alpha \tilde{e}_\alpha$.]
The unconstrained theorem (P = MxN) gives a family $(\tilde{f}(x,y))_{x\in M, Y\in N}$ of unconstrained Schwartz morphisms, such that $\ddot{y} = \tilde{f}(x,y)\ddot{x}$ for every curve (x,y) verifying $\dot{y} = \tilde{e}(x,y)\dot{x}$. Denoting by f the restriction of $\tilde{f}$ to P, one has $\ddot{y} = f(x,y)\ddot{x}$ for every curve (x,y) in P verifying $\dot{y} = e(x,y)\dot{x}$, whence existence. Uniqueness stems from the fact that accelerations of curves linearly span the vector space $\tau_x M$: given a point (x,y) in P and an acceleration $a \in \tau_x M$, there is a curve x(t) in M with x(0) = x and $\ddot{x}(0) = a$; solving the differential equation

$$\dot{y}(t) = e(x(t),y(t))\dot{x}(t), \quad y(0) = y$$

gives a curve y(t); and $\ddot{y}(0)$ is the only possible value of f(x,y)a.

To verify that the Schwartz morphism f(x,y) is constrained to P, we shall use Proposition 3: it suffices to verify that, for every $L \in \tau_x M$, $\bar{L} = (L)_M + (fL)_N + \hat{f}\hat{L}$ is tangent to P. As $\bar{L}$ depends linearly on L, it suffices to verify it when L is the acceleration $\ddot{x}(0)$ of some curve x(t). In that case, fL is the acceleration $\ddot{y}(0)$ of the curve y(t) just constructed, $\hat{L}$ is just the tensor product $\dot{x}(0)\otimes\dot{x}(0)$, and, since f restricted to first order is e, $\hat{f}\hat{L} = 2(\dot{y}(0))_N(\dot{x}(0))_M$. Finally,

$$\bar{L} = (\ddot{x}(0))_M + 2(\dot{x}(0))_M(\dot{y}(0))_N + (\ddot{y}(0))_N$$

is nothing but the acceleration in MxN of the curve (x(t),y(t)). As this curve sits in P, $\bar{L}\in\tau_{xy}P$ as was to be shown.

Last, since equations $\underline{d}Y = \tilde{f}(X,Y)\underline{d}X$ and $\delta Y = \tilde{e}(X,Y)\delta X$ are equivalent, and since every solution to the former starting in P remains in P (by identification with the solution to $\underline{d}Y = f(X,Y)\underline{d}X$), it also holds for the latter, which can hence be replaced with $\delta Y = e(X,Y)\delta X$. ∎

Observe that the Stratonovich transfer principle (Theorem 8) deals with ordinary equations less general than Theorem 7. This is not a minor technical difficulty, but an essential limitation of the method itself (emphasized by Schwartz [10] p.111) : if $e(X_t,Y_t)$ is also allowed to depend on t (in a non-smooth fashion) or on ω (for instance as a functional of the past of X or Y), the construction of f from e described in the above statement does not give an intrinsic stochastic differential equation equivalent to the Stratonovich one--if it can be performed at all!

As is well known to practitioners of stochastic differential geometry, this transfer principle is as easy to apply as it is general: Write the ordinary differential equation you have to transfer in local coordinates (multiplying everything by dt if necessary to replace derivatives with differentials), then make everything random, with Stratonovich differentials δY^α and δX^i instead of dy^α and dx^i.

For the sake of a future comparison with the Itô transfer principle, this section ends with a deterministic approximation to Stratonovich equations. We shall need three definitions.

DEFINITION. An <u>interpolation rule</u> on a manifold M is a measurable mapping $I: M\times M\times[0,1]\to M$ such that

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

(ii) $I(x,y,\cdot)$ is a curve of class C^2;

(iii) for every compact $K\subset M$ and every smooth $f:M\to\mathbb{R}$, there are a constant c_K and a function ψ_K on $K\times K$, with $\lim \psi_K(x,y) = 0$ when $d(x,y)\to 0$ such that, for all s and t in [0,1] and x and y in K, the function $h(t) = f(I(x,y,t))$ verifies

$$|h''(t)| \le c_K\, d^2(x,y)$$
$$|h'(t) - h'(s) - (t-s)h''(s)| \le \psi_K(x,y)\, d^2(x,y).$$

(As above, d denotes any Riemannian distance on M; the choice of d is irrelevant since any two such distances are equivalent on compacts).

This definition is slightly more general than the one in [6]: the bound (of order 3) on the third derivative has essentially been replaced with a Lipschitz condition (with a constant of order more than 2) on the second derivative. Notice than an easy integration gives here

$$|h(t) - h(s) - (t-s)h'(s) - \tfrac{1}{2}(t-s)^2 h''(s)| \le \psi_K(x,y)\, d^2(x,y).$$

Examples of such interpolation rules are the Euclidean interpolation, if M is equal (or diffeomorphic) to $\mathbb{R}^m$, or, more generally, the geodesic interpolation, where M is endowed with a connection and $I(x,y,t)$ is the small goedesic linking x and y if x and y are close enough, and an arbitrary smooth curve if (x,y) is outside some neighbourhood of the diagonal. (See Proposition (7.13) of [6].)

DEFINITION. Given $(\Omega,\underline{\underline{F}},P,(\underline{\underline{F}}_t)_{t\ge 0})$, a <u>subdivision</u> is an increasing sequence $(T_n)_{n\ge 0}$ of stopping times such that $T_0 = 0$ and $\sup_n T_n = \infty$.

The <u>size</u> $|\sigma|$ of a subdivion $\sigma = (T_n)_{n\ge 0}$ is the number

$$\sum_{k\ge 1} 2^{-k}\mathbb{E}[1 \wedge \sup_n((T_{n+1}\wedge k) - (T_n\wedge k))]$$

(so that, for a sequence (σ^q) of subdivisions, $|\sigma^q|\to 0$ if and only if, for every compact $K \subset [0,\infty)$, the distance $\inf_{t\in\sigma}\inf_{s\in K}|t-s|$ between K and the subdivision tends to zero in probability).

DEFINITION. Let N be a manifold. On the set of all pairs (Y,ζ), with ζ a random variable in $(0,\infty]$ and Y a N-valued, continuous, measurable process defined on $[[0,\zeta[[$, the topology of <u>uniform convergence on compacts in probability</u> is defined by the following property: A sequence (Y^n,ζ^n) converges to a limit (Y,ζ) iff

$\zeta^n \wedge \zeta$ tends to ζ in probability and, for every $k > 0$, the random variable

$$\sup_{0 \le t \le k \wedge (\zeta - \frac{1}{k})} d(Y_t^n, Y_t)$$

tends to zero in probability.

Remark that this does not depend on the choice of the Riemannian distance d. Since $\zeta^n \wedge \zeta$ tends to ζ in prbability, the random variable above is well-defined except on an event whose probability tends to zero, and convergence in probability makes sense. As each point has a countable basis of neighbourhoods, the topology can be defined with sequences only.

THEOREM 9. (Stratonovich approximation). Let M,N,P,e be as in Theorem 8, X be a continuous semimartingale in M, and y_0 a $\underline{\underline{F}}_0$-measurable random variable in N such that (X_0, y_0) belongs to $\underline{P}$. Let I be an interpolation rule on M and, for each subdivion σ, let X^σ denote the (non adapted) piecewise smooth process

$$X_t^\sigma = I(X_{T_n}, X_{T_{n+1}}, \frac{t-T_n}{T_{n+1}-T_n}), \; T_n \le t \le T_{n+1}$$

When the size $|\sigma|$ tends to zero, the piecewise smooth solution Y^σ to the (pathwise ordinary) differential equation

$$\dot{Y}^\sigma = e(X^\sigma, Y^\sigma)\dot{X}^\sigma \;, \; Y_0^\sigma = y_0$$

converges uniformly on compacts in probability to the solution Y to the Stratonovich differential equation

$$\delta Y = e(X,Y)\delta X \;, \; Y_0 = y_0 \;.$$

This general form of a classical result may be found in [6], so we won't prove it. Though the definitions of an interpolation rule and of the size of a subdivision are stronger in [6] than here, it is easily verified that only the weaker properties taken here as definitions are used in the proof.

Remark that Theorem 9 easily bootstraps itself: Y^σ converges to Y, not only uniformly on compacts in probability, but also in a stronger sense: For every other Stratonovich stochastic differential equation from N to another manifold Q, $\delta Z = g(Y,Z)\delta Y$, the solution Z^σ to the equation driven by Y^σ converges to the solution Z. This is obtained by considering the process (Y,Z) in NxQ as the solution to an equation driven by X, and applying Theorem 9 to this enlarged equation.

V. THE ITÔ TRANSFER PRINCIPLE

In this section, ordinary differential equations between manifolds are transformed into stochastic ones using what Meyer calls Itô integrals on manifolds. They extend the usual Itô calculus in a flat space, not to an arbitrary manifold, but to a less general geometric structure: a manifold endowed with a connection.

Recall Meyer's interpretation of a connection in the frame of second order geometry [8]: If a manifold M is endowed with a connection, there exists for each $x \in M$ a linear mapping $\mathbf{F}: \tau_x M \to T_x M$ such that, if A and B are vector fields on M, $\mathbf{F}(A) = A$ and $\mathbf{F}(AB) = \nabla_A B - \frac{1}{2}T(A,B)$ (all these vectors are evaluated at x; AB is the second order differential operator obtained when composing the first order ones A and B; T is the torsion of ∇).This $\mathbf{F}$ does not characterize the connection, but only its torsion-free part: two connections yield the same $\mathbf{F}$ if and only if they have the same geodesics (or the same convex functions, or the same martingales). Conversely, every family of linear mappings $\mathbf{F}:\tau_x M \to T_x M$, depending smoothly on x, and such that $\mathbf{F}(A) = A$ for every first-order vector A can be obtained this way, from a unique torsion-free connection. (The letter $\mathbf{F}$ stands for "first-order part".) In local coordinates, if Γ^i_{jk} are the Christoffel symbols of the connection,

$$\mathbf{F}(\ell^{ij}D_{ij}+\ell^i D_i) = [\ell^i + \tfrac{1}{2}(\Gamma^i_{jk}+\Gamma^i_{kj})\ell^{jk}]\ D_i .$$

If M is a vector space with the flat connection, $\mathbf{F}L$ is obtained from L by keeping only the first-order terms, and killing the second-order ones (in any system of linear coordinates). If M is an arbitrary manifold (with an arbitrary connection), every $x \in M$ is the origin of a system of normal coordinates (they are linear functions of the inverse exponential map at x); the functions $\Gamma^i_{jk} + \Gamma^i_{kj}$ for this chart vanish at x, so, at x, $\mathbf{F}L$ consists simply in reading L in some normal coordinates and deleting the second-order terms.

For each $x \in M$, the linear $\mathbf{F} : \tau_x M \to T_x M$ has an adjoint $\mathbf{F}^* : T^*_x M \to \tau^*_x M$ that makes first-order forms into second-order ones. This enables Meyer [8] to define the Itô integral of a first-order form α along a continuous, M-valued semimartingale X as $\int \langle \ \mathbf{F}\alpha, \underline{dX} \ \rangle$ (the first definition of those Itô integrals is due to Duncan [4], in the Riemannian case). Clearly, this requires no regularity for α : the Itô integral $\int \langle \mathbf{F}\Psi \ , \ \underline{dX} \rangle$ can be defined if Ψ is any locally bounded, predictable, T^*M-valued process above X.

DEFINITION. Let M and N be endowed with connections, $x \in M$ and $y \in N$. A linear mapping $f : \tau_x M \to \tau_y N$ is <u>semi-affine</u> if

$$\mathbf{F}_N[f(L)] = f(\mathbf{F}_M L)$$

for every $L \in \tau_x M$.

The previous description of $\mathbf{F}$ in normal coordinates can be restated with this definition : for $x \in M$, denoting by $\phi = \exp_x$ and $\psi = \exp_x^{-1}$ the exponential mapping at x and its inverse, the inverse linear mappings

$$\vec{\phi}_0 : \tau_0 T_x M \to \tau_x M \ , \ \vec{\psi}_x : \tau_x M \to \tau_0 T_x M$$

are semi-affine (the vector space $T_x M$ is endowed with the flat connection).

The prefix semi in 'semi-affine' recalls that f does not necessarily commute with the connections themselves, but only with their torsion-free parts; nothing is said about how f carries over the torsion. This is expressed more rigorously in the next proposition, that will not be used in the sequel (so we content ourselves with a very elliptical proof, leaving the details as an exercise to the reader).

PROPOSITION 10. <u>Let</u> M <u>and</u> N <u>be endowed with connections</u> ∇_M <u>and</u> ∇_N; <u>denote by</u> $\phi : M \to N$ <u>a smooth mapping</u>. <u>The following statements are equivalent</u>:

(i) <u>for every</u> $x \in M$, <u>the push-forward</u> $\vec{\phi}_x : \tau_x M \to \tau_{\phi(x)} N$ is <u>semi-affine</u>;

(ii) <u>for every geodesic</u> $g : U \to M$, <u>with</u> U <u>an open interval in</u> $\mathbb{R}$, $\phi \circ g : U \to N$ <u>is a geodesic</u>;

(iii) ϕ <u>is affine from</u> (M, ∇'_M) <u>to</u> (N, ∇'_N), <u>where</u> ∇'_M <u>and</u> ∇'_N <u>are the torsion-free parts of</u> ∇_M <u>and</u> ∇_N.

PROOF. Denoting by $\Gamma'^{i}_{jk} = \frac{1}{2}(\Gamma^{i}_{jk}+\Gamma^{i}_{kj})$ and $\Gamma'^{\alpha}_{\beta\gamma} = \frac{1}{2}(\Gamma^{\alpha}_{\beta\gamma}+\Gamma^{\alpha}_{\gamma\beta})$ the Christoffel symbols of ∇'_M and ∇'_N respectively, it is easy to check in local coordinates ((x^i) on M, (y^α) on N) that each of the three conditions amounts to the relations

$$D_{ij}\phi^\alpha - \Gamma'^{k}_{ij} D_k\phi^\alpha + \Gamma'^{\alpha}_{\beta\gamma} D_i\phi^\beta D_j\phi^\gamma = 0$$

for all indices α, i and j. ∎

Before stating the Itô transfer principle itself, here is its geometric part. It is simpler than the Stratonovich one in that it is punctual (and not only local); on the other hand, it is more complicated to constraint it to submanifolds since they must be totally geodesic.

LEMMA 11. <u>Let</u> M <u>and</u> N <u>be endowed with connections</u>, $x \in M$ <u>and</u> $y \in N$. <u>Let</u> e <u>be a linear mapping from</u> $T_x M$ <u>to</u> $T_y N$. <u>There exists a unique Schwartz morphism</u> $f \in SM_{xy}(M,N)$ <u>such that</u>

(i) f <u>is semi-affine</u>;

(ii) e <u>is the restriction of</u> f <u>to first order vectors</u>.

<u>It is given by</u> $f = \vec{\phi}_x$ <u>where</u> ϕ, <u>defined in a neighbourhood of</u> x, <u>is the mapping</u>

$$\phi = \exp_y \circ e \circ \exp_x^{-1}.$$

<u>In local coordinates</u> ((x^i) <u>on</u> M, (y^α) <u>on</u> N), f <u>is given by the coefficients</u>

$$f^\alpha_i = e^\alpha_i \ ; \ f^\alpha_{ij} = \frac{1}{2}\,[e^\alpha_k(\Gamma^k_{ij}+\Gamma^k_{ji}) - e^\beta_i e^\gamma_j (\Gamma^\alpha_{\beta\gamma}+\Gamma^\alpha_{\gamma\beta})].$$

<u>If moreover</u> (x,y) <u>is in a totally geodesic submanifold</u> P <u>of</u> $M \times N$ (<u>for the product connection</u>), <u>and if</u> e <u>is constrained to</u> P, <u>then</u> f <u>too is constrained to</u> P.

PROOF. The coefficients of e and f are defined by

$$e(D_i) = e^\alpha_i D_\alpha \ ; \ f(D_i) = f^\alpha_i D_\alpha \ ; \ f(D_{ij}) = f^\alpha_{ij} D_\alpha + f^\alpha_i f^\beta_j D_{\alpha\beta}.$$

Condition (ii) means $e^\alpha_i = f^\alpha_i$, and (i) is equivalent to $f(\mathbf{F}_M D_{ij}) = \mathbf{F}_N f(D_{ij})$, that is, to

$$f(\tfrac{1}{2}(\Gamma^k_{ij}+\Gamma^k_{ji})D_k) = f^\alpha_{ij} D_\alpha + f^\beta_i f^\gamma_j \tfrac{1}{2}(\Gamma^\alpha_{\beta\gamma}+\Gamma^\alpha_{\gamma\beta})D_\alpha,$$

yielding $f^\alpha_k \frac{1}{2}(\Gamma^k_{ij}+\Gamma^k_{ji}) = f^\alpha_{ij} + f^\beta_i f^\gamma_j \frac{1}{2}(\Gamma^\alpha_{\beta\gamma}+\Gamma^\alpha_{\gamma\beta})$, and giving f^α_{ij} in terms of f^β_k and the connections. This proves existence, uniqueness and gives the expression in local coordinates.

To verify that $f = \vec{\phi}_x$, if suffices to check that $\vec{\phi}_x$ has properties (i) and (ii). The first one holds because the push-forward by $\exp_x^{-1}$ and $\exp_y$ are semi-affine at the centre, and e is linear (hence affine); the second one because the push-forward by

$\exp_x^{-1}$ and $\exp_y$ at the centre are the identity on first-order vectors.

If e is constrained to a totally geodesic P, let $\xi \in M$ be close enough to x and let $\eta = \phi(\xi)$. Let u(t) and v(t) denote respectively the geodesics in M and N such that $u(0) = x$, $v(0) = y$, $\dot{u}(0) = \exp_x^{-1}(\eta)$, $\dot{v}(0) = \exp_y^{-1}(\eta) = e(\dot{u}(0))$. The curve $g(t) = (u(t),v(t))$ is a geodesic in the product manifold M x N; its velocity at the origin is $\dot{g}(0) = (\dot{u}(0), \dot{v}(0)) = (\dot{u}(0), e(\dot{u}(0)))$. As e is constrained to P, $\dot{g}(0)$ is tangent to P; as P is totally geodesic, the whole geodesic $(g(t))_{0\le t\le 1}$ is in P, and $(\xi,\eta) = g(1) \in P$. So the graph of ϕ is included in P (at least, near (x,y)), and $f = \vec{\phi}_x$ is constrained to P. ∎

As the Stratonovich one, the Itô transfer principle transforms ordinary differential equations into stochastic ones, involving this time the Itô differentials $\mathbf{F}_M \underline{dX}$ and $\mathbf{F}_N \underline{dY}$ instead of the Stratonovich ones δX and δY.

DEFINITION. Let X (respectively Y) be a continuous semimartingale in a manifold M (respectively N) endowed with a connection. For each (t,ω), let $e_t(\omega)$ be a linear mapping from $T_{X_t(\omega)}M$ to $T_{Y_t(\omega)}N$; dually, $e_t^*(\omega)$ maps $T^*_{Y_t(\omega)}N$ to $T^*_{X_t(\omega)}M$. One says that Y is a solution to the <u>Itô stochastic differential equation</u>

$$\mathbf{F}_N \underline{dY} = e\mathbf{F}_M \underline{dX}$$

if, for every first order form α on N, the real semimartingales $\int \langle \alpha(Y), \mathbf{F}_N \underline{dY} \rangle$ and $\int \langle e^*\alpha(Y), \mathbf{F}_M \underline{dX} \rangle$ are equal.

The reader familiar with manifold-valued continuous martingales (see Meyer [8]) will remark immediately that Itô differential equations make martingales into martingales : if X is a M-valued martingale, every Itô integral along X is a local martingale, so by the above definition every Itô integral of a smooth first order form along Y is a local martingale, and this in turn shows that Y itself is a N-valued martingale.

THEOREM 12 (Itô transfer principle). <u>If the process</u> e <u>is predictable, the Itô stochastic differential equation</u> $\mathbf{F}_N \underline{dY} = e\, \mathbf{F}_M \underline{dX}$ <u>is equivalent to the intrinsic stochastic differential equation</u> $\underline{dY} = f\underline{dX}$, <u>where</u> $f_t(\omega) : \tau_{X_t(\omega)}M \to \tau_{Y_t(\omega)}N$ <u>is the unique semi-affine Schwartz morphism with restriction</u> $e_t(\omega)$ <u>to first order</u>.

PROOF. Given the continuous semimartingales X and Y, let T be the first time when exactly one of the equations is satisfied. If with positive probability T is neither infinite nor an explosion time for X or Y, there are local charts (x^i) on M and (y^α) on N such that X_T and Y_T are in the domains of those charts with positive probability. So, by conditioning, we may work in local coordinates.

Denoting by Γ^i_{jk} and $\Gamma^\alpha_{\beta\gamma}$ the Christoffel symbols of the connections, the Itô equation is equivalent to

$$dY^\alpha + \frac{1}{2}\Gamma^\alpha_{\beta\gamma}(Y)\ d[Y^\beta,Y^\gamma] = e^\alpha_i(dX^i + \frac{1}{2}\Gamma^i_{jk}(X)\ d[X^j,X^k]),$$

where the symmetry of the brackets $d[Y^\beta,Y^\gamma]$ and $d[X^j,X^k]$ make it possible to use the Christoffel symbols of the given connections, without having to remove the torsion. Since this implies

$$d[Y^\beta,Y^\gamma] = e^\beta_j\, e^\gamma_k\, d[X^j,X^k],$$

it is equivalent to

$$dY^\alpha = e^\alpha_i(dX^i + \tfrac{1}{2}\,\Gamma^i_{jk}(X)\ d[X^j,X^k]) - \tfrac{1}{2}\,\Gamma^\alpha_{\beta\gamma}(Y)\ e^\beta_j\, e^\gamma_k\, d[X^j,X^k];$$

direct inspections of the coefficients show that this intrinsic equation is but

$$dY^\alpha = f^\alpha_i\, dX^i + \tfrac{1}{2}\, f^\alpha_{ij}\ d[X^i,X^j],$$

with f^α_i and f^α_{ij} given by Lemma 11. ∎

COROLLARY 13. In theorem 4, assume furthermore that both M and N are endowed with connections, and that P is totally geodesic; replace the constrained Schwartz morphisms by the constrained linear mappings

$$E(Y)_t(\omega) \in L_{X_t(\omega)\ Y_t(\omega)}(M,N;P)$$

verifying the same hypotheses.

The Itô stochastic differential equation

$$\mathbf{F}_N \underline{dY} = E(Y)\ \mathbf{F}_M \underline{dX},\ Y_0 = y_0$$

has a unique maximal solution, exploding at some predictable time $\zeta \le \infty$.

PROOF. Use Theorem 12 to transform this Itô equation into $\underline{dY} = F(Y)\ \underline{dX}$, and apply Theorem 4. ∎

REMARKS. 1) As shown by the proof of 11, the Itô transfer principle is as simple to use in practice as the Stratonovich one. To transform an ordinary differential equation $\dot{y} = e(x,y)\dot{x}$ into a stochastic one, simply rewrite it $dy = e(x,y)\ dx$ and replace x, y, dx and dy by their stochastic counterparts X, Y, $\mathbf{F}\underline{dX}$ and $\mathbf{F}\underline{dY}$; in local coordinates, dx^i is to be transformed into $dX^i + \frac{1}{2}\,\Gamma^i_{jk}(X)\ d[X^j,X^k]$.

2) The stochastic differential equations we have dealt with are of three types: $\underline{dY} = f\ \underline{dX}$ (intrinsic), $\delta Y = e\ \delta X$ (Stratonovich) and $\mathbf{F}\underline{dY} = e\mathbf{F}\underline{dX}$ (Itô). Is it possible to define equations of mixed types, for instance $\delta Y = e\mathbf{F}\underline{dX}$ or $\mathbf{F}\underline{dY} = e\ \delta X$? It seems that the only mixed equations that make sense are those of the form

$$\mathbf{F}\underline{dY} = g\ \underline{dX},$$

with a coefficient $g : \tau_{X_t(\omega)}M \to T_{Y_t(\omega)}N$ and a connection on N. Two particular cases of this equation are the integral of a second-order form along X (here, $N = \mathbb{R}$) and the Itô equation $\mathbf{F}\underline{dY} = e\mathbf{F}\underline{dX}$ (here, M is endowed with a connection too and $g = e\mathbf{F}$). But besides these two examples, these equations don't look very interesting.

As is the case with Stratonovich equations, it is possible to approximate the solution of an Itô equation by discretizing time, interpolating X, and performing a deterministic operation on the so obtained piecewise smooth curve. In the Stratonovich case, much freedom was left in the choice of the interpolation rule; here, the geodesic interpolation only can be used: this is where the connection on M comes in. The connection on N will be used in the deterministic construction yielding the approximate Y: this operation will involve the construction of a geodesic with prescribed initial position and velocity. Remark that, since both connections are taken into account through their geodesics only, this construction is insensitive to a change of torsion (we already know that such is the case for Itô integrals, and hence also for the Itô equation itself).

Recall that an interpolation rule is called geodesic if there is a neighbourhood V of the diagonal in M x M such that for (x,y) in V, the curve $t \to I(x,y,t)$ is a geodesic. Given a connection, a geodesic interpolation rule always exists, and is essentially unique: any two agree on some neighbourhood of the diagonal.

THEOREM 14 (Itô approximation) Let M and N be endowed with connections, and P be a totally geodesic, closed submanifold of M x N. Let $e : \mathbb{R}_+ \times \Omega \times P \to L(M,N;P)$ be such that

(i) $e(t,\omega,x,y) \in L_{xy}(M,N,;P)$;

(ii) $\omega \mapsto e(t,\omega,x,y)$ is $\underline{\underline{F}}_t$-measurable for fixed t,x,y;

(iii) $t \mapsto e(t,\omega,x,y)$ is left-continuous with limits on the right for fixed ω,x,y;

(iv) for each compact $K \subset P$, there is a measurable (not necessarily adapted) increasing process $L(K,t,\omega)$ such that, for (x,y) and (ξ,η) in P,

$$d(e(t,\omega,x,y),e(t,\omega,\xi,\eta)) \leq L(K,t,\omega)\ d((x,y),(\xi,\eta)).$$

Let X be a continuous semimartingale in M, and y_0 a $\underline{\underline{F}}_0$-measurable, N-valued random variable with $(X_0,y_0) \in P$. Let I be a geodesic interpolation rule on M, and, for every subdivision $\sigma = (T_n)_{n\geq 0}$, define

$$X_t^\sigma = I(X_{T_n}, X_{T_{n+1}}, \frac{t-T_n}{T_{n+1}-T_n}) \text{ if } T_n \leq t \leq T_{n+1}$$

and denote by $\dot{X}_t^\sigma$ the right derivative of X_t^σ. Define inductively a continuous, N-valued process Y^σ on each interval $[[T_n,T_{n+1}]]$ by

$Y_0^\sigma = y_0$;

on each interval $[[T_n,T_{n+1}]]$, Y^σ is the geodesic with initial condition

$\dot{Y}_{T_n}^\sigma = e(T_n,\omega,X_{T_n},Y_{T_n}^\sigma)\ \dot{X}_{T_n}^\sigma$.

When the size $|\sigma|$ tends to zero, the process Y^σ, well-defined up to a random time $\zeta^\sigma \leq \infty$(*), converges uniformly on compact sets in probability(**) to the solution (Y,ζ) of

$$\mathbf{F}\underline{dY}_t(\omega) = e(t,\omega,X_t(\omega),Y_t(\omega))\ \mathbf{F}\ \underline{dX}_t(\omega),$$
$$Y_0 = y_0 .$$

REMARKS. 1) A particular case of this Itô equation is the computation of an Itô integral

$$Y_t = \int_0^t \langle \alpha_s , \mathbf{F}\underline{dX}_s \rangle ,$$

where α is a first-order form on M and N = $\mathbb{R}$. In that case, Theorem 13 reduces to

---

(*) If N is not complete, some geodesics may explode in finite time.

(**) Recall the definition before Theorem 9.

Darling's approximation result [2].

2) In the case when M is a Euclidean space, X a Brownian motion and e does not depend on ω or x (this is meaningful since, M being a vector space, all tangent spaces $T_{\hat{x}}M$ can be identified with M), this result is due to Bismut [1] (and Y is called an Ito diffusion).

PROOF. First, by induction on the interval $[[T_n, T_{n+1}]]$, remark that the process (X^σ, Y^σ) takes its values in P, so the very definition $\dot{Y}^\sigma_{T_n} = e(T_n, X_{T_n}, Y^\sigma_{T_n})\dot{X}^\sigma_{T_n}$ is meaningful. Indeed, since both curves X^σ and Y^σ are geodesics in the interval $[[T_n, T_{n+1}]]$, (X^σ, Y^σ) is also a geodesic (for the product connection) in this interval; and since P is totally geodesic and closed, the geodesic (X^σ, Y^σ) remains in P (as long as it is itself defined) provided its initial velocity $(\dot{X}^\sigma_{T_n}, \dot{Y}^\sigma_{T_n})$ is tangent to P. But this is a consequence of the definition of $\dot{Y}^\sigma_{T_n}$ and the fact that e is constrained by P.

Observe also that the Itô differential equation given by e is a particular case of those considered in Theorem 11; whence the existence and uniqueness of Y on a maximal interval $[[0, \zeta[[$.

The first step in the proof consists in replacing M and N with the vector spaces $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively. Indeed, it is possible to imbed properly M and N into such vector spaces, and by Lemma 15 below to extend to $\mathbb{R}^m$ and $\mathbb{R}^n$ the connections ∇^M and ∇^N. Since the injections $M \hookrightarrow \mathbb{R}^m$ and $N \hookrightarrow \mathbb{R}^n$ are affine, P is still (closed and) totally geodesic in the larger product $\mathbb{R}^m \times \mathbb{R}^n$, Y is still the solution to $\mathbf{F}d\underline{Y} = e(t,X,Y)\mathbf{F}d\underline{X}$ (with $\mathbf{F}$ denoting now the extended connections), and Y^σ remains the same. So no generality is lost when supposing $M = \mathbb{R}^m$ and $N = \mathbb{R}^n$; we shall freely use the global coordinates (x^i) on M and (y^α) on N, the Euclidean distances and norms, the vector-valued velocities and accelerations.

For a compact $K \subset P$ and a positive a, let $T = a \wedge \inf\{t : (X_t, Y_t) \notin K\}$. It suffices to show that Y^σ tends to Y uniformly on $[[0,T]]$, and, by replacing (X,Y) with $(X^{|T}, Y^{|T})$, we may suppose that (X,Y) takes its values in compact set, and restrict ourselves to a compact time-interval. Also, replacing $e(t,\omega,x,y)$ with $\phi(x,y)e(t,\omega,x,y)$, where ϕ is a scalar function with compact support, equal to 1 on a neighbourhood of K, we may suppose that e = 0 on a neighbourhood of infinity.

Now, it suffices to prove $Y^\sigma \to Y$ along a sequence of subdivisions with size tending to zero; and a classical argument shows that we need to prove convergence for some subsequence only; so we may suppose that $\sup_\ell (T_{\ell+1} - T_\ell)$ tends to zero a.s. This implies $\sup_\ell d(X_{T_\ell}, X_{T_{\ell+1}}) \to 0$ a.s. by uniform continuity of the paths; and, using a property of interpolation rules, $\sup_\ell \|\dot{X}^\sigma_{T_\ell}\| \to 0$ a.s. In particular, all the vectors $\dot{X}^\sigma_{T_\ell}$ remain in a (random) compact set, independent of σ and ℓ. Using now the boundedness of $e(.,\omega,.,.)$ on compacts (this is a consequence of (iii) and (iv); the existence of right limits is used), this shows that $e(T_\ell,\omega,X_{T_\ell},y)\dot{X}^\sigma_{T_\ell}$ remains in a random compact $R \subset TN$ when σ,ℓ and y vary with y in a fixed compact and $(X_{T_\ell},y)\in P$. But for y outside some compact, $e(\ldots,y) = 0$; so $e(T_\ell,X_{T_\ell},y)\dot{X}^\sigma_{T_\ell}$ remains in $R \cup$ Null (where Null is the null section in TN, the set of all null tangent vectors) when σ, ℓ and y vary, and in particular $\dot{Y}^\sigma_{T_\ell}$ is in $R \cup$ Null for all σ and ℓ such that $\dot{Y}^\sigma_{T_\ell}$ exists (that is, $\zeta^\sigma > T_\ell$).

Now, the life-duration of a geodesic γ is a lower semi-continuous function of the initial condition $\dot{\gamma}(0)$, so the set U^ϵ of all vectors $v\in TN$ such that the geodesic γ with $\dot{\gamma}(0) = v$ is well-defined on $[0,\epsilon]$, is open. As $U^\epsilon \uparrow TN$ when $\epsilon\downarrow 0$, the compact R is included in U^ϵ for some (random) $\epsilon > 0$; as $U^\epsilon \supset$ Null, $\dot{Y}^\sigma_{T_\ell}$ is in U^ϵ for all σ and ℓ such that $\dot{Y}^\sigma_{T_\ell}$ exits. Neglecting a (random) finite number of terms in the sequence of subdivisions, we have $\sup_\ell(T_{\ell+1} - T_\ell) < \epsilon$; this implies that each geodesic arc $(Y^\sigma_t)_{T_\ell \le t \le T_{\ell+1}}$ is well-defined; and the life-duration ζ^σ of Y^σ is identically infinite.

[In the Riemannian case, the above argument is not necessary since the connection on N can easily be made complete by a modification near infinity; but for arbitrary connections, geodesics may "explode" while remaining in a compact set, because their speed may become infinite.]

To simplify the notations, let $\Delta_\ell Z$ stand for the increment of the process Z from T_ℓ to $T_{\ell+1}$: $\Delta_\ell Z = Z_{T_{\ell+1}} - Z_{T_\ell}$ (all our processes are now vector-valued). As Y^σ and its velocity $\dot{Y}^\sigma$ live in a (random) compact independent of σ,
$\Delta_\ell Y^{\sigma\alpha} = (T_{\ell+1}-T_\ell)\dot{Y}^{\sigma\alpha}_{T_\ell} + \frac{1}{2}(T_{\ell+1}-T_\ell)^2\ddot{Y}^{\sigma\alpha}_{T_\ell} + \mathcal{O}(\|\Delta_\ell Y^\sigma\|^2)$ with a $\mathcal{O}$ not depending upon σ and ℓ. Using the equation of geodesics $\ddot{Y}^{\sigma\alpha}_{T_\ell} = -\Gamma^\alpha_{\beta\gamma}(Y^\sigma_{T_\ell})\dot{Y}^{\sigma\beta}_{T_\ell}\dot{Y}^{\sigma\gamma}_{T_\ell}$ and the boundedness of $\|\Delta_\ell Y^\sigma\|/\|\Delta_\ell X\|$ for $(T_{\ell+1} - T_\ell)$ small enough gives

$$\Delta_\ell Y^{\sigma\alpha} = (T_{\ell+1}-T_\ell)\dot{Y}^{\sigma\alpha}_{T_\ell} - \frac{1}{2}(T_{\ell+1}-T_\ell)^2\Gamma^\alpha_{\beta\gamma}(Y^\sigma_{T_\ell})\dot{Y}^{\sigma\beta}_{T_\ell}\,\dot{Y}^{\sigma\gamma}_{T_\ell} + \epsilon^{\sigma\alpha}_\ell$$

with $\quad \sum_\ell |\epsilon^{\sigma\alpha}_\ell| \to 0$ a.s.

Similarly,

$$\Delta_\ell X^i = (T_{\ell+1}-T_\ell)\,\dot{X}^{\sigma\alpha}_{T_\ell} - \frac{1}{2}\,(T_{\ell+1}-T_\ell)^2\,\Gamma^i_{jk}(X_{T_\ell})\,\dot{X}^{\sigma j}_{T_\ell}\,\dot{X}^{\sigma k}_{T_\ell} + \mathcal{O}(\|\Delta_\ell X\|^2),$$

whence

$$(T_{\ell+1}-T_\ell)\,\dot{X}^{\sigma i}_{T_\ell} = \Delta_\ell X^i + \frac{1}{2}\,\Gamma^i_{jk}(X_{T_\ell})\,\Delta_\ell X^j\,\Delta_\ell X^k + \epsilon^{\sigma i}_\ell$$

with another $\epsilon^{\sigma i}_\ell$ verifying also $\sum_\ell |\epsilon^{\sigma i}_\ell| \to 0$ a.s.

Now the definition $\dot{Y}^{\sigma\alpha}_{T_\ell} = e^\alpha_i\,(T_\ell, X_{T_\ell}, Y^\sigma_{T_\ell})\,\dot{X}^\sigma_{T_\ell}$ of Y^σ gives

$$\Delta_\ell Y^{\sigma\alpha} = F^\alpha_{iT_\ell}\,\Delta_\ell X^i + \frac{1}{2}\,F^\alpha_{jkT_\ell}\,\Delta_\ell X^j\,\Delta_\ell X^k + \epsilon^{\sigma\alpha}_\ell$$

with $\sum_\ell |\epsilon^{\sigma\alpha}_\ell| \to 0$ and with coefficients

$$\begin{cases} F^\alpha_{iT_\ell} = e^\alpha_i(T_\ell, X_{T_\ell}, Y^\sigma_{T_\ell}) \\ F^\alpha_{jkT_\ell} = F^\alpha_{iT_\ell}\,\Gamma^i_{jk}(X_{T_\ell}) - F^\beta_{jT_\ell}\,F^\gamma_{kT_\ell}\,\Gamma^\alpha_{\beta\gamma}(Y^\sigma_{T_\ell}). \end{cases}$$

For $T_\ell \le t < T_{\ell+1}$, let $u(t) = T_\ell$ and

$$Z^{\sigma\alpha}_t = Y^{\sigma\alpha}_{T_\ell} + F^\alpha_{iT_\ell}\,(X^i_t - X^i_{T_\ell}) + \frac{1}{2}\,F^\alpha_{jkT_\ell}\,(X^j_t - X^j_{T_\ell})\,(X^k_t - X^k_{T_\ell}).$$

The process Z^σ is a (non-continuous) semimartingale, verifying

$$Z^{\sigma\alpha}_t = H^{\sigma\alpha}_t + \int_0^t F^\alpha_{iu(s)}\,dX^i_s + \frac{1}{2}\int_0^t F^\alpha_{jku(s)}\,[d(X^jX^k)_s - X^j_{u(s)}\,dX^k_s - X^k_{u(s)}\,dX^j_s]$$

with $H^{\sigma\alpha}_t = y^\alpha_0 + \sum_{T_\ell \le t} \epsilon^{\sigma\alpha}_\ell$. Since $H^\sigma \to y_0$ uniformly a.s., we claim that Z^σ tends uniformly on compacts in probability to the (continuous) solution Z of the equation

$$Z^\alpha_t = y^\alpha_0 + \int_0^t G^\alpha_{is}\,dX^i_s + \frac{1}{2}\int_0^t G^\alpha_{jks}\,[d(X^jX^k)_s - X^j_s\,dX^k_s - X^k_s\,dX^j_s],$$

with $G^\alpha_{it} = e^\alpha_i(t, X_t, Z_t)$ and

$$G^\alpha_{jkt} = G^\alpha_{it}\,\Gamma^i_{jk}(X_t) - G^\beta_{jt}\,G^\gamma_{kt}\,\Gamma^\alpha_{\beta\gamma}(Z_t).$$

This claim is a particular instance of Theorem 1.d of [5], except that we have a family H^σ with limit H instead of just a fixed H. But it is very easy to see that 1.d remains true in that case, since the proof of 1.d consits in applying Proposition 5 of the same paper, which in turn refers to Proposition 4 where H is allowed to vary!(*)

Now, since $X^jX^k - \int X^j dX^k - \int X^k dX^j = [X^j, X^k]$, and since the coefficients G^α_i and G^α_{jk} of the equation giving Z are identical to those giving Y (see the proof of Theorem 12), Z = Y; since $Z^\sigma_{T_\ell} = Y^\sigma_{T_\ell}$,

---

(*) This shows once again that one should never give up generality for the sake of simplicity ... or laziness!

$$\|Y_t^\sigma - Y_t\| \le \|Y_t^\sigma - Y_{u(t)}^\sigma\| + \|Z_{u(t)}^\sigma - Z_{u(t)}\| + \|Y_t - Y_{u(t)}\|.$$

All three terms tend to zero uniformly in probability: the second one by what has just been seen, the third one by uniform continuity of the paths of Y, and the first one since geodesics with initial velocities in a (random, but) fixed compact and time-duration at most $\eta > 0$ have their Euclidean lengthes tending to zero uniformly when η tends to zero.

The proof of Theorem 14 is now complete, but for the next Lemma, that was admitted a moment ago.

LEMMA 15. Let M be endowed with a connection ∇^M, and $i : M \to \mathbf{R}^n$ be a proper imbedding. There exists a connection on $\mathbf{R}^n$ such that i is affine (in particular, M is totally geodesic in $\mathbf{R}^n$, for this connection).

PROOF. Since the imbedding is proper, iM is a closed submanifold of $\mathbf{R}^n$. So each point of iM has an open, relatively compact neighbourhood V in $\mathbf{R}^n$ diffeomorphic to $\mathbf{R}^m \times \mathbf{R}^{n-m}$, with $iM \cap V$ corresponding to $\mathbf{R}^m \times \{0\}$; and each point of $\mathbf{R}^n - iM$ has an open, relatively compact neighbourhood that does not meet iM. All these open sets form a covering of $\mathbf{R}^n$; there exists a partition of unity (ψ_α) subordinated to that covering (to each α is associated one of these sets, V_α, and ψ_α is compactly supported in V_α; the sum $\sum_\alpha \psi_\alpha$ is locally finite and identically equal to 1). If V_α meets iM, it is possible to endow V_α with a connection ∇^α such that i is affine from $i^{-1}(V_\alpha)$ to V_α: Using the above mentioned diffeomorphism, this amounts to extending a connection ∇^m from $\mathbf{R}^m \times \{0\}$ to $\mathbf{R}^m \times \mathbf{R}^{n-m}$, and this can be done by taking the product of ∇^m with an arbitrary connection on $\mathbf{R}^{n-m}$. If V_α does not meet iM, just endow V_α with an arbitrary connection ∇^α. Now the sum $\sum_\alpha \psi_\alpha \nabla^\alpha$ is locally finite and defines a connection ∇ on $\mathbf{R}^n$; if A and B are vector fields on M, $\vec{i}A$ and $\vec{i}B$ are the corresponding vector fields on iM, and

$$\nabla_{\vec{i}A} \vec{i}B\,(ix) = \sum \psi_\alpha(ix)\ (\nabla^\alpha_{\vec{i}A}\vec{i}B)(ix) = \sum \psi_\alpha(ix)\ \vec{i}(\nabla^M_A B(x)) = \vec{i}(\nabla^M_A B(x))$$

shows that i is affine. ∎

As in the Stratonovich case, this Itô approximation result can be bootstrapped to show that, if Y is used to direct another differential equation $F\underline{dZ} = \epsilon(Y,Z)\ F\underline{dY}$ to a third manifold, then, to construct the appromixation Z^σ of Z we may use, instead of the geodesic interpolation of Y, the Y^σ constructed in Theorem 14.

VI. COMPARING BOTH TRANSFER PRINCIPLES; APPLICATION.

Given an ordinary differential equation $\dot{y} = e(x,y)\dot{x}$ between manifolds, we have seen two ways of extending it into a stochastic one $\underline{dY} = f(X,Y)\underline{dX}$. Both agree of course at order 1, which means that, in local coordinates, the first coefficients f_i^α of the Schwartz morphism f are just the coefficients e_i^α of e. But they don't agree in general at order 2: the coefficient f^{α}_{ij} of the Stratonovich extension is given by $(D_j+e^{\beta}_j D_{\beta})e^{\alpha}_i$ (to be symmetrized in i,j) and that of the Itô extension is $e^{\alpha}_k\Gamma^k_{ij}-e^{\beta}_i e^{\gamma}_j\Gamma^{\alpha}_{\beta\gamma}$ (to by symmetrized in i,j if the connections are not torsion-free). The geometric condition, linking e with the connections, that ensures equality between both extensions, is easy to see on those formulae in local coordinates: the ordinary equation $\dot{y} = e(x,y)\dot{x}$ must make geodesics into geodesics. But all the computations have been done previously, so this result is an immediate consequence of what we already know:

COROLLARY 16. Let M and N be endowed with connections, and P be a closed, totally geodesic submanifold of M x N. Let $e : P \to L(M,N;P)$ be of class $C^{1,Lip}$ and such that $e(x,y)$ is in $L_{xy}(M,N;P)$ for all (x,y) in P.

Suppose that, if $x(t)$ is any geodesic in M and y_0 any point in N with $(x(0),y_0)\epsilon P$, the solution y to the ordinary differential equation

$$\dot{y}(t) = e(x(t),y(t))\ \dot{x}(t),\ y(0) = y_0.$$

is a geodesic too. Then, for every continuous semimartingale X in M and every $\underline{\underline{F}}_0$-measurable y_0 with $(X_0,y_0)\epsilon P$, the stochastic Itô and Stratonovich equations

$$\mathbf{F}\underline{dY} = e(X,Y)\ \mathbf{F}\underline{dX},\ Y_0 = y_0$$
$$\delta Y = e(X,Y)\ \ \delta X,\ Y_0 = y_0$$

are equivalent.

PROOF. The solutions Y^I and Y^S to these equations can be approximated by discretizing time and applying respectively Theorems 14 and 9. But the approximation Y^{σ} to Y^S is piecewise geodesic (hypothesis on e), so it coincides identically with the approximation to Y^I; finally, letting $|\sigma| \to 0$, $Y^S = Y^I$. ∎

REMARKS. 1) This result is still true if e is only once differentiable, with first-order partial derivatives locally bounded. In that case, uniqueness of Y^S is not obvious, but existence and uniqueness hold for Y^I by Corollary 13, and the equivalence between both equations can be verified directly, in local coordinates. So uniqueness holds also for Y^S; the reason is that, though the partial derivatives of e are not locally Lipschitz, some combinations of them are, namely $(D_j+e^{\beta}_j D_{\beta})e^{\alpha}_i$ (symmetrized in i,j) because, by the geometric hypothesis on e, this is precisely $e^{\alpha}_k\ \Gamma^k_{ij} - e^{\beta}_i\ e^{\gamma}_j\ \Gamma^{\alpha}_{\beta\gamma}$ (symmetrized); and these combinations are of course those appearing in the Stratonovich equation.

2) If X is a continuous semimartingale in M and Φ a continuous semimartingale in T*M, above X, the Stratonovich integral $\int\Phi\delta X$ and, if M is endowed with a connection, the Itô integral $\int\Phi\mathbf{F}\underline{dX}$ can both be defined. They are shown in Lemma (8.24) of [6] to be equal if, for every parallel transport $U \in TM$ along X, the real semimartingale $\langle U,\Phi \rangle$ has finite variation. This result does not seem to be obtainable as a consequence of Corollary 16. The reason is that the Stratonovich integral $\int\Phi\delta X$ is not a particular case of the Stratonovich stochastic differential equations considered in Theorem 8; as mentioned earlier, this is not due to our hypotheses being too

restricted, but to an essential limitation of the Stratonovich transfer principle itself.

3) In Corollary 16, the hypothesis that the ordinary differential equation e transforms geodesics into geodesics can be replaced by

for every $(\Omega, \underline{\underline{F}}, P, (\underline{\underline{F}}_t)_{t\geq 0})$, every martingale X in M and every $\underline{\underline{F}}_0$-measurable $y_0 \in N$ with $(X_0, y_0) \in P$, the solution Y to the Stratonovich equation

$$\delta Y = e(X,Y)\delta X, \quad Y_0 = y_0$$

is a martingale in N.

Indeed, this new hypothesis implies the former. For let $x : I \to M$ be a geodesic, where I is an open interval, and let $y : J \to N$ be a solution to $\dot{y} = e(x,y)\dot{x}$, with $J \subset I$ an open interval. For every J-valued continuous local martingale U, $X = x \circ U$ is a martingale in M. Denote by Y the semimartingale $y \circ U$. As the push-forward $\vec{y}$ factorizes as $e(x,y) \circ \vec{x}$ (applied to $\frac{d}{dt} \in T\mathbb{R}$, this reduces to the equation giving y), one can write for every smooth form α on N

$$\int \langle \alpha(Y), \delta Y \rangle = \int \langle \alpha(Y), \delta(y \circ U) \rangle = \int \langle \overleftarrow{y}\, \alpha(Y), \delta U \rangle$$
$$= \int \langle \overleftarrow{x}\, e^*(X,Y)\, \alpha(Y), \delta U \rangle = \int \langle e^*(X,Y)\, \alpha(Y), \delta(x \circ U) \rangle$$
$$= \int \langle e^*(X,Y)\, \alpha(Y), \delta X \rangle ;$$

so Y is also a solution to $\delta Y = e(X,Y)\delta X$, and our hypothesis implies that $Y = y \circ U$ is a martingale. Since U is arbitrary, y must be a geodesic.

But one can say a little more: this new hypothesis is in fact equivalent to the old one. This is obvious by remarking that, by Corollary 16 itself, the hypothesis that the ordinary differential equation preserves geodesics implies that the associated Stratonovich equation is in fact an Itô one, so it must preserve martingales. This is an extension to differential equations of the equivalence between preserving geodesics and preserving martingales for smooth functions between manifolds (both amount to the function being semi-affine).

4) For a given X in M, the proof and the conclusion of the corollary still hold if one does not require all geodesics of M to be made into geodesics of N by e, but only those geodesics that are needed to interpolate X (or, in the bootstrap case when X is already the result of a previous equation, those geodesics used in the approximation X^σ of X).

As an application of all this, we now turn to the problem of extending to semimartingales such geometric operations as parallel transport of vectors and rolling without slipping.

If M is endowed with a connection, the parallel transport $u(t) \in T_{x(t)}M$ of a vector $u(0) \in T_{x(0)}M$ along a curve $x(t)$ can be considered as the solution to the ordinary differential equation $\dot{u}(t) = e(x(t),u(t))\dot{x}(t)$, where $e(x,u) : T_xM \to T_uTM$, defined for $\pi u = x$ only (that is, $u \in T_xM$) is the horizontal lifting: $e(x,u)$ transforms a vector $A \in T_xM$ into the only horizontal vector in T_uTM with projection A itself: $\vec{\pi}_u e(x,u)$ is the identity on T_xM. This equation is constrained to the submanifold P of M x TM consisting of all (x,u) with $\pi u = x$. (Remark that P is trivially diffeomorphic to TM itself!) The Stratonovich transfer principle applied to this ordinary differential equation yields the Stratonovich stochastic parallel transport along semimartingales. All this is classical, save the name: The eponym 'Stratonovich' is usually omitted, so one is not tempted to worry about the possible existence of an Itô one. But this is of course what we are going to do.

To construct the "Itô stochastic parallel transport" along a semimartingale, that is, to apply the Itô transfer principle to the equation $\dot{u}(t) = e(x(t),u(t))\dot{x}(t)$ of parallel transport, we need a connection on M and a connection on TM. We already have one on M; so the Itô transfer principle explains a posteriori a phenomenon emphasized by Meyer [9]: there is a one-one correspondence between extensions of ordinary parallel transport to semimartingales and extensions of the connection in M to TM (Meyer's work is more general, TM being replaced with an artitrary vector bundle over M; this makes no essential difference).

To make the above statement a little more precise, observe first that the connection on TM cannot be completely arbitrary: since the ordinary equation is constrained to the submanifold P M x TM, we must choose the connection on TM in such a way that P is totally geodesic in the product M x TM. This is clearly equivalent to the requirement that the map $u \to (\pi u, u)$ from TM to M x TM transform geodesics into geodesics; in other words, $\pi : TM \to M$ must be semi-affine. Since, when applying the Itô transfer principle, the torsions of the connections are not taken into account, there is no loss of generality in requiring π to be affine.

[Another requirement of Meyer is that each fiber T_xM, with its flat connection of vector space, be a totally goedesic submanifold of TM, with the induced connection. With this proviso, the equation of stochastic parallel transport will be linear. This requirement is quite reasonable, but not logically necessary - and not used in the sequel.]

Now, given any such connection on the manifold TM, it is possible to define the Itô stochastic parallel transport associated with this connection; moreover, this transport can be approximated in the following way, as a direct application of Theorem 14: Given the subdivision σ and the continuous semimartingale X in M, and supposing that the approximate parallel transport U^σ along X has already been constructed up to time T_n , its restriction to the interval $[[T_n, T_{n+1}]]$ is the geodesic in TM, above the geodesic X^σ (because P is totally geodesic), starting from the previously obtained $U^\sigma_{T_n}$, with the same initial velocity $\dot{U}^\sigma_{T_n}$ as a (ordinary) parallel transport along the curve X^σ. Two particular choices of this connection on TM are specifically interesting in stochastic (and ordinary) differential geometry.

The first one, called by Yano and Ishihara [12] the horizontal lift to TM of the connection in M, can be characterized (up to a torsion term, but the Itô transfer neglects it) by the property that each parallel transport along a geodesic of M is a geodesic of TM. For a proof, see Bismut [1] page 450. As a consequence, by Corollary

16, *the Itô parallel transport* (*defined with this horizontal connection*) *is the same as the Stratonovich one*. (This can also be seen as a consequence of Meyer's Theorem 5 in [9], using the third remark following Corollary 16.) Of course, for this connection, the approximate parallel transport is exactly the ordinary parallel transport along X^σ (by the very proof of Corollary 16, the Itô and Stratonovich approximations are identical).

The other important connection on TM is the "complete lift" of Yano and Ishihara [12]. Meyer has observed in [9] that the stochastic parallel transport corresponding to this connection is (an extension to the non-Riemannian case of) the one introduced by Dohrn and Guerra [3] under the name "geodesic correction to parallel transport". It is defined by the same approximation procedure as above, but the geodesics in TM are replaced with Jacobi fields along geodesics in M. This strongly suggests the following explanation:

LEMMA 17. *Let* M *be a manifold endowed with a connection; endow* TM *with the complete lift of this connection. The geodesics of* TM *are exactly the uniform motions in each fibre and the Jacobi fields along the geodesics of* M.

PROOF. From the explicit formulae I.6.2 of [12], it is clear that removing the torsion of the connection commutes with extending it to TM. So, noticing that this does not change the geodesics or the Jacobi fields, we may and will suppose that the connections are torsion-free.

Now, in local coordinates, still using I.6.2, the equation of geodesics in TM is

$$\begin{cases} \ddot{u}^i = -u^\ell D_\ell \Gamma^i_{jk}\, \dot{x}^j \dot{x}^k - (\Gamma^i_{jk}+\Gamma^i_{kj})\dot{x}^j \dot{u}^k \\ \ddot{x}^i = -\Gamma^i_{jk}\, \dot{x}^j \dot{x}^k, \end{cases}$$

where Γ^i_{jk} are the Christoffel symbols of the connection in M. Since the connection has no torsion, the equation of a Jacobi field u along a geodesic x is

$$\nabla_{\dot{x}} \nabla_{\dot{x}} u = R(\dot{x},u)\dot{x}.$$

It is not difficult to rewrite this equation in local coordinates and to identify it with the first equation above; the computation is still simpler if one performs it at the center of some normal coordinates. There, because the connection is torsion-free, all Christoffel symbols vanish and one is left with

$$(\nabla_{\dot{x}} \nabla_{\dot{x}} u)^i = \ddot{u}^i + \dot{x}^\ell D_\ell \Gamma^i_{jk}\, \dot{x}^j u^k.$$

Using $(R(\dot{x},u)\dot{x})^i = R_j{}^i{}_{k\ell}\, \dot{x}^j \dot{x}^k u^\ell$ with $R_j{}^i{}_{k\ell} = D_k\Gamma^i_{j\ell} - D_\ell\Gamma^i_{jk}$ gives the result. ∎

So the general theorem on Itô approximations sheds a new light on the Dohrn-Guerra construction and explains in particular why the first derivative $\nabla_{\dot{x}} u$ is reset at zero at every step: this is where the ordinary equation comes in; this initial condition (common to all these procedures, whatever the connection on TM) is the one that forces the stochastic parallel transport to agree with the ordinary one on smooth curves.

A last remark on Itô stochastic transports: one might be tempted to rewrite the equation of ordinary parallel transport as

$$\frac{d}{dt} A_t f = \nabla df(\dot{x}, A_t)$$

(A_t is a parallel transport along a curve x(t), and f a smooth function on M; Af is just < A,df >). This amounts to using only functions of the type df as test-functions on TM, instead of all possible smooth functions on TM. Rewritten

$$A_t f - A_0 f = \int_0^t \langle \nabla df(\cdot, A_s), dx_s \rangle,$$

everything is M or R-valued and one may forget about TM. Applying the Stratonovich transfer gives $\delta A_t = \langle \nabla df(\cdot, A_t), \delta X_t \rangle$, which is of course equivalent to the Stratonovich parallel transport; so one might hope to define an Itô parallel transport, without any connection on TM, by

$$A_t f - A_0 f = \int_0^t \langle \nabla df(\cdot, A_s), \underline{\underline{F dX_s}} \rangle.$$

But this fails, simply by lack of intrinsicness: if this holds for some functions $f^1, \ldots, f^P$, it need not hold for a function of the form $g(f^1, \ldots, f^P)$; the reason is that both sides of this equation do not obey the same change of variable formula, the left-hand side involving the third derivatives of g, the right-hand one stopping at order 2. Another way of saying it is that the left-hand side cannot be considered as an Itô integral along A; the Hessian of the function df on TM cannot have a null projection on M simultaneously for all smooth functions f.

As an application of parallel transport, we shall now discuss <u>liftings</u> and <u>developments</u>. Recall that the lifting y(t) in $T_{x(0)}M$ of a curve x(t) in a manifold M endowed with a connection, is constructed by choosing a frame $F = (U_1, U_2, \ldots, U_m)$ of $T_{x(0)}M$, transporting it as $F(t) = (U_\alpha(t))_{1\le\alpha\le m}$ along the curve x(t), in such a way that each $U_\alpha(\cdot)$ is a parallel transport, reading the velocity $\dot{x}(t)$ in the frame F(t) (this gives $\dot{x}(t) = v^\alpha(t)U_\alpha(t)$, with $v^\alpha(t) \in \mathbb{R}$), and finally letting $y(t) = (\int_0^t v^\alpha(s)ds)U_\alpha \in T_{x(0)}M$. Since the equation of parallel transport is linear, it is very easy to see that the curve y in $T_{x(0)}M$ does not depend upon the choice of the frame F = F(0).

Using the Stratonovich transfer principle, the extension to semimartingales is straightforward: If X is a M-valued continuous semimartingale with $X_0 = x$, choose a frame F of T_xM, transport it as $F_t = (U_{\alpha t})$ along X using the Stratonovich stochastic parallel transport, define the dual frame (η_t^α) by $\eta_t^\alpha \in T^*_{X_t} M$ and $\langle U_{\alpha t}, \eta_t^\beta \rangle = \delta_\alpha^\beta$, and let $Y_t = (\int_0^t \langle \eta_s^\alpha, \delta X_s \rangle)U_\alpha$. All of this is well-known, and widely used. See Bismut [1] for a general presentation when X is a Brownian diffusion.

The development is the inverse operation: given the continuous semimartingale Y_t in T_xM with $Y_0 = 0$, find X. Of course, this is done by following the same path backwards, that is, constructing simultaneously X_t and the attached parallel moving frame F_t. We shall not go into details here, referring the reader to [6] for instance. Liftings and developments are both used in the definition(*) of rolling

---

(*) In classical mechanics, rolling and slipping are defined in terms of instantaneous rotations, and the definition given here is, for 2-manifolds imbedded in $\mathbb{R}^3$, a theorem.

without slipping: Given two manifolds of M and N with connections, two points $x \in M$ and $y \in N$, a linear bijection $\ell : T_x M \to T_y N$ and a continuous semimartingale (or smooth curve) X in M starting from x, the semimartingale in N obtained by rolling M on N along X without slipping is by definition the development in N of the curve $\ell(Z) \in T_y N$, where Z is the lifting of X in $T_x M$. So this operation of rolling without slipping is just obtained by composing a lifting, the linear mapping ℓ, and a development; lifting and development are two particular cases of rolling without slipping.

Rolling without slipping preserves geodesics (in other words, the lifting in $T_{x(0)}M$ of a curve x(t) in M is a uniform motion if and only if x is a geodesic) and manifold-valued martingales (in other words, the lifting in $T_{X_0}M$ of a continuous semimartingale X_t in M is a local martingale if and only if X is a martingale). This result, implicit in Bismut [1], is explicitly stated by Meyer [9]; but its Brownian version is much older: stochastic developments have long been used to construct manifold-valued Brownian motions from Euclidean ones. It is also a little surprising: why should such a Stratonovich procedure preserve martingales? Corollary 16 seems to imply that this can be derived from the preservation of geodesics, for in that case the Stratonovich equation is also an Itô one, therefore it preserves martingales. But it does not apply here, at least not directly, since liftings (and developments, and also a fortiori rolling without slipping) are not operations of the type considered in that corollary, but combinations of such operations. Lifting, for instance, is not constructed directly from M to $T_x M$, but needs an intermediate step in a larger manifold, the frame bundle FM. Hence, to derive rigorously martingale preservation from geodesic preservation by using Corollary 16, one needs the existence of a connection on FM such that geodesics are preserved at each step of the construction $M \to FM \to T_x M$. And in general, such a connection does not exist! The reason is that the only possible choice, the obvious extension to FM of the horizontal connection on TM described earlier, does not work. Indeed, by Proposition II.9.1 of [12], its geodesics are exactly the curves $F(t) = (U_\alpha(t))_{1 \le \alpha \le m}$ in FM such that $x(t) = \pi F(t)$ is a geodesic in M and each $U_\alpha(t)$ has the form $V_\alpha(t) + tW_\alpha(t)$, where $V_\alpha(t)$ and $W_\alpha(t)$ are parallel transports along x. But only if $W_\alpha = 0$ does the second step

$$y(t) = \left(\int_0^t \langle \eta^\alpha(s), \dot{x}(s) \rangle \, ds\right) U_\alpha(0)$$

(where (η^α) is the frame dual to (U_α)) transform the geodesic F into a straight line. The point is, of course, that those geodesics with $W_\alpha = 0$ are the only ones obtained from the first step $M \to FM$; so even though Corollary 16 does not apply, Remark 4 following it does, and gives the result.

More generally, if TM is endowed with a connection of the type considered above (that is, making π affine), and if the Stratonovich parallel transport is replaced with the corresponding Itô one in the definition of liftings, the same proof shows that martingales are still preserved. This can also be seen as a consequence of Lemma (8.24) of [6]. Indeed, in local coordinates, the equation of any stochastic parallel transport (Stratonovich or Itô) is

$$dU_t^i = -\Gamma_{jk}^i(X_t)\ dX_t^j\ U_t^k + fv,$$

where fv denotes a correction term with finite variation; hence the dual frame (η^α) to a stochastic parallel frame (U_α) is made of forms verifying

$$d\eta_{kt} = \Gamma_{jk}^i(X_t)\ dX_t^j\ \eta_{it} + fv.$$

So, if U is a Stratonovich parallel transport and η an element of the dual frame to an Itô parallel frame, the pairing $\langle U_t, \eta_t \rangle = U_t^i\ \eta_{it}$ has finite variation; and by Lemma (8.24), the Stratonovich integral $\int \langle \eta, \delta X \rangle$ is identical with the Itô one $\int \langle \eta, \mathbf{F}\underline{dX} \rangle$, yielding the result.

REFERENCES

[1] J.M. Bismut. Mécanique aléatoire. Lecture Notes in Mathematics 866, Springer 1981.

[2] R.W.R. Darling. Martingales in manifolds and geometric Itô calculus. Ph.D. Thesis, University of Warwick, 1982.

[3] D. Dohrn, F. Guerra. Nelson's stochastic mechanics on Riemannian manifolds. Lettere al nuovo cimento 22, 121-127, 1978.

[4] T.E. Duncan. Stochastic integrals in Riemannian manifolds. J. Multivariate Anal. 6, 397-413, 1976.

[5] M. Emery. Stabilité des solutions des équations différentielles stochastiques; application aux intégrales multiplicatives. Z. Wahrscheinlichkeitstheorie verw. Gebiete 41, 241-262, 1978.

[6] M. Emery, P.A. Meyer. Stochastic calculus in manifolds. Universitext, Springer 1989.

[7] M. Métivier. Semimartingales. A course on stochastic processes. de Gruyter, 1982.

[8] P.A. Meyer. Géométrie stochastique sans larmes. Séminaire de Probabilités XV, Lecture Notes in Mathematics 850, Springer 1981.

[9] P.A. Meyer, Géometrie différentielle stochastique (bis). Séminaire de

Probabilités XVI, Supplément: Géométrie différentielle stochastique, Lecture Notes in Mathematics 921, Springer 1982.

[10] L. Schwartz. Géométrie différentielle du 2^e ordre, semimartingales et équations différentielles stochastiques sur une variété différentielle. Séminaire de Probabilités XVI, Supplément: Géométrie différentielle stochastique, Lecture Notes in Mathematics 921, Springer 1982.

[11] L. Schwartz. Semimartingales and their stochastic calculus on manifolds. Presses de l'Université de Montréal, 1984.

[12] K. Yano and S. Ishihara. Tangent and cotangent bundles. Marcel Dekker, 1973.