Diffusions on manifolds

Feller diffusions and stochastic differential equations

Let K be a compact subset of $\mathbb{R}^n$, and assume that K is the closure of its interior. By definition, $\mathcal{C}^2(K)$ is the collection of all real functions on K that can be extended to a function in $\mathcal{C}^2(\mathbb{R}^n)$. For such functions f, partial derivatives $\frac{\partial}{\partial x_i}f$ and $\frac{\partial^2}{\partial x_i \partial x_j}f$ are defined first on the interior of K in the usual way, and then by continuous extension on all of K. Consider the operator L, with domain $\mathcal{C}^2(K)$, given by

$$(Lf)(x) := \sum_i b_i(x)\tfrac{\partial}{\partial x_i}f(x) + \tfrac{1}{2}\sum_{ij} a_{ij}(x)\tfrac{\partial^2}{\partial x_i \partial x_j}f(x),$$

where b, a are continuous functions from K into the n-dimensional real vectors and the $n \times n$ positive symmetric real matrices, respectively. We call b the *drift* vector and a the *diffusion matrix*. A stochastic process $(X(t))_{t\geq 0}$ with continuous sample paths is said to solve the martingale problem for L iff. for every $f \in \mathcal{C}^2(K)$, the following is a martingale with respect to the filtration $(\mathcal{F}^X_t)_{t\geq 0}$ generated by X

$$M(t) := f(X(t)) - \int_0^t (Lf)(X(s))ds.$$

If for every initial condition $x \in K$ the martingale problem for L has a unique solution $(X^x(t))_{t\geq 0}$ (in distribution), then the collection of processes $(X^x)^{x\in K}$ is a Feller process, i.e. the formula

$$(S_t f)(x) := E[f(X^x(t))]$$

defines a Feller semigroup on $\mathcal{C}(K)$ whose generator is an extension of L.

If the space K is not compact, for example if $K = \mathbb{R}^n$, the picture is technically a bit more complicated. Now, L is defined on $\mathcal{C}^2$ functions with compact support, and the Feller semigroup acts on bounded continuous functions. The process M above is in general not a martingale, but it is a martingale if we introduce stopping times and stop the process M when X leaves a compact set $C \subset \mathbb{R}^n$. This leads to the so-called local martingale problem. We will not go into this any further.

As is well-known, solutions to the martingale problem for L can be constructed as solutions to a stochastic differential equation. For this, we have to chose a root of the diffusion matrix, i.e. a function σ from K into the $n \times m$ real matrices such that

$$a_{ij}(x) = \sum_u \sigma_{iu}(x)\sigma_{ju}(x).$$

Here the sum runs over $u = 1, \ldots, m$. Now let $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t\geq 0}, B, X)$ be a filtered probability space with m-dimensional Brownian motion B and a solution X of the stochastic differential equation

$$dX_i(t) = b_i(X(t))dt + \sum_u \sigma_{iu}(X(t))dB_u(t) \qquad (i = 1, \ldots, d).$$

Then a simple application of Itô's formula shows that for any $f \in \mathcal{C}^2(K)$

$$df(X) = (Lf)(X)dt + \sum_{iu} \sigma_{iu}(X)(\tfrac{\partial}{\partial x_i} f)(X)dB_u.$$

We should read this equation in integral form: $f(X(t)) = f(X(0)) + \int_0^t (Lf)(X(s))ds +$ *stochastic integral*, where the stochastic integral yields an $\mathcal{F}_t$-martingale. This means that the process $M := f(X) - \int_0^t (Lf)(X)ds$ is an $\mathcal{F}_t$-martingale, and since it is $\mathcal{F}_t^X$-adapted it is also an $\mathcal{F}_t^X$-martingale, i.e. X solves the martingale problem for L.

There is a deep converse to this observation. Whenever X is a solution to the martingale problem for L and σ a continuous root of the diffusion matrix a, the process X is equal in law to some solution of the stochastic differential equation above, on an appropriate filtered probability space equipped with m-dimensional Brownian motion. Here we still have freedom in the choice of our root σ. If $m \geq n$, and a is a continuous diffusion matrix, then it is alway possible to find a root σ that is continuous too. For example, if $m = n$ we can take for σ the unique positive symmetric root of a; there are however many other possibilities.

Every different choice of the root σ leads to a different way of *coupling* solutions X of the martingale problem for L to m-dimensional Brownian motion. The two processes are coupled in such a way that if B makes an infinitesimal step $dB(t)$, then X makes an infinitesimal step

$$dX(t) = b(X(t))dt + \sigma(X(t))dB(t),$$

where $b(x)$ is the vector $(b_1(x), \ldots, b_n(x))$, and $\sigma(x) : \mathbb{R}^m \to \mathbb{R}^n$ is the linear map

$$(\sigma(x)z)_i := \sum_u \sigma_{iu}(x)z_u.$$

To stress the fact that the spaces $\mathbb{R}^m$ and $\mathbb{R}^n$ need not be the same, we use indices u, v, w for coordinates referring to $\mathbb{R}^m$, and indices i, j, k for coordinates referring to $\mathbb{R}^n$.

Transformations of the space

Let $\phi : K \to K' \subset \mathbb{R}^n$ be an invertible $\mathcal{C}^2$ map, and consider the K'-valued process $X' := \phi(X) = (\phi_1(X), \ldots, \phi_n(X))$. By Itô's formula

$$d\phi_i(X) = (L\phi_i)(X)dt + \sum_{ju} \sigma_{ju}(X)(\tfrac{\partial}{\partial x_j}\phi_i)(X)dB_u.$$

Thus, if we define continuous functions b', σ' on K' by

$$\begin{aligned} b'_i(\phi(x)) &:= (L\phi_i)(x) \\ \sigma'_{iu}(\phi(x)) &:= \sum_j \sigma_{ju}(x)(\tfrac{\partial}{\partial x_j}\phi_i)(x), \end{aligned}$$

then the process X' solves the stochastic differential equation

$$dX' = b'(X)'dt + \sigma'(X')dB.$$

Transformation rules take a simpler form if we consider Stratonovich stochastic differential equations, instead of the normal stochastic differential equations with the Itô integral.

Stratonovich stochastic differential equations

Recall that the Itô integral $\int_0^t XdY$ of one continuous semimartingale X with respect to another, Y, is defined as the limit of the sums:

$$\sum_{i=1}^{n} X(t_{i-1})(Y(t_i) - Y(t_{i-1})),$$

where the partitions $0 = t_0 < t_1 < \cdots < t_n = t$ are such that $\sup_k |t_k - t_{k-1}| \to 0$ in the limit, and the integrant X is evaluated at the left end of the interval $[t_{i-1}, t_i)$. The Stratonovich integral arises if instead we take the average of X at the left and right end of the interval, i.e., we take the limit of sums

$$\sum_{i=1}^{n} \tfrac{1}{2}(X(t_{i-1}) + X(t_i))(Y(t_i) - Y(t_{i-1})).$$

We write $\int_0^t XdY$ for the Itô intergral and $\int_0^t X * dY$ for the Stratonovich integral.

Recall that Itô's formula says that if X is an n-dimensional semimartingale and $f \in \mathcal{C}^2(\mathbb{R}^n)$, then $f(X)$ is a semimartingale and

$$df(X) = \sum_i (\tfrac{\partial}{\partial x_i} f)(X)dX_i + \tfrac{1}{2}\sum_{ij}(\tfrac{\partial^2}{\partial x_i \partial x_j} f)(X)dX_i dX_j,$$

where $dXdY := d[X,Y]$, with $[X,Y]$ the quadratic covariance process of X and Y. The analogue formula for the Stratonivich integral is the Itô-Stratonovich formula:

$$df(X) = \sum_i (\tfrac{\partial}{\partial x_i} f)(X) * dX_i,$$

where this time we assume $f \in \mathcal{C}^3(\mathbb{R}^n)$. This has the following consequence. Consider continuous maps $A_0, \ldots, A_m$ from K into $\mathbb{R}^n$, and write $A_u(x) = (A_{1u}(x), \ldots, A_{nu}(x))$. If X solves the Stratonovich stochastic differential equation

$$dX_i = A_{i0}(X)dt + \sum_{u=1}^{m} A_{iu}(X) * dB_u \qquad (i = 1, \ldots, n)$$

then for any $f \in \mathcal{C}^3(K)$

$$df(X) = (A_0 f)(X)dt + \sum_{u=1}^{m}(A_u f)(X) * dB_u,$$

where we associate the maps $A_0, \ldots, A_m$ from K to $\mathbb{R}^n$ with the first order differential operators $A_0, \ldots, A_m$ given by

$$(A_u f)(x) := \sum_i A_{iu}(x)(\tfrac{\partial}{\partial x_i} f)(x).$$

Note that $(A_u f)(x)$ is the derivative of f in the direction $A_u(x)$:

$$(A_u f)(x) := \lim_{\varepsilon \to 0} \varepsilon^{-1}\Big(f(x + \varepsilon A_u(x)) - f(x)\Big)$$

for x in the interior of K.

We can now deduce the transformation rule for Stratonovich stochastic differential equations. Let $\phi : K \to K' \subset \mathbb{R}^n$ be a smooth (at least $\mathcal{C}^3$) invertible transformation of the space, then

$$d\phi_i(X) = (A_0\phi_i)(X)dt + \sum_{u=1}^{m}(A_u\phi_i)(X) * dB_u,$$

where, for $u = 0, \ldots, m$

$$(A_u\phi_i)(x) = \sum_j A_{ju}(x)(\tfrac{\partial}{\partial x_j}\phi_i)(x).$$

Thus, if we write $X' := \phi(X)$ for the transformed process and we define A'_u on K' by

$$A'_{iu}(\phi(x)) := \sum_j A_{ju}(x)(\tfrac{\partial}{\partial x_j}\phi_i)(x), \tag{1}$$

then X' solves the Stratonovich stochastic differential equation

$$dX' = A'_0(X')dt + \sum_{u=1}^{m} A'_u(X') * dB_u.$$

Relation between Itô and Stratonovich integrals

Let X and Y be continuous semimartingales. The relation between the Itô integral and the Stratonovich integral can formally be denoted as follows:

$$X * dY = XdY + \tfrac{1}{2}dXdY.$$

Exercise: Use this formula to deduce the Itô-Stratonovich formula from Itô's formula.

For us, the following is important: if X solves the Stratonovich stochastic differential equation

$$dX = A_0(X)dt + \sum_{u=1}^{m} A_u(X) * dB_u,$$

and the vector fields A_u are at least continuously differentiable, then X solves the (Itô) stochastic differential equation

$$dX = b(X)dt + \sigma(X)dB,$$

with

$$\begin{aligned} b_i(x) &= A_{i0}(x) + \tfrac{1}{2}\sum_{u=1}^{m}\sum_j A_{ju}(x)(\tfrac{\partial}{\partial x_j}A_{iu})(x) \\ \sigma_{iu}(x) &= A_{iu}(x). \end{aligned}$$

Proof: For f sufficiently smooth we have

$$df(X) = (A_0f)(X)dt + \sum_{u=1}^{m}(A_uf)(X) * dB_u, \tag{2}$$

where
$$(A_u f)(X) * dB_u = (A_u f)(X) dB_u + \tfrac{1}{2} d(A_u f)(X) dB_u.$$
Applying (2) to $A_u f$ we get

$$d(A_u f)(X) = (A_0 A_u f)(X) dt + \sum_{v=1}^{m} (A_v A_u f)(X) * dB_v.$$

Using the rules $dt\, dB_u = 0$ and $dB_v dB_u = \delta_{vu} dt$ we get

$$d(A_u f)(X) dB_u = (A_u^2 f)(X) dt.$$

Thus, we arrive at

$$df(X) = (A_0 f)(X) dt + \tfrac{1}{2} \sum_{u=1}^{m} (A_u^2 f)(X) dt + \sum_{u=1}^{m} (A_u f)(X) dB_u. \tag{3}$$

If we insert into this formula the coordinate functions $f(x) = x_i$, we arrive at the formulas for b and σ. In particular

$$\sum_{u=1}^{m} \Big(\sum_j A_{ju}(x) \tfrac{\partial}{\partial x_j} \Big) \Big(\sum_k A_{ku}(x) \tfrac{\partial}{\partial x_k} \Big) x_i = \sum_{u=1}^{m} \Big(\sum_j A_{ju}(x) \tfrac{\partial}{\partial x_j} \Big) A_{iu}(x),$$

which leads to the formula for b. ∎

Note that formula (3) tells us immediately that X solves the martingale problem for the operator
$$L = A_0 + \tfrac{1}{2} \sum_{u=1}^{m} A_u^2.$$
This is called a differential operator in Hörmander form. If L is given in the form

$$(Lf)(x) = \sum_i b_i(x) \tfrac{\partial}{\partial x_i} f(x) + \tfrac{1}{2} \sum_{ij} a_{ij}(x) \tfrac{\partial^2}{\partial x_i \partial x_j} f(x),$$

and if σ is a root of a that is at least once continuously differentiable, then we can write L in Hörmander form by defining

$$\begin{aligned} A_{iu}(x) &:= \sigma_{iu}(x) \qquad (u = 1, \ldots, m) \\ A_{i0}(x) &:= b_i(x) - \tfrac{1}{2} \sum_{u=1}^{m} \sum_j \sigma_{ju}(x) (\tfrac{\partial}{\partial x_j} \sigma_{iu})(x). \end{aligned}$$

Note, however, that it may not always be possible to choose a differentiable root σ, even is the diffusion matrix itself is smooth (for example $\mathcal{C}^\infty$). Note furthermore that there is not one good way to write L in Hörmander form; different roots σ lead to different ways of writing L.

Manifolds

Let M be a second countable Hausdorff space. We say that M is equipped with an n-dimensional differentiable structure if we are given an n-dimensional real vectorspace E and a collection $\mathcal{D}$ of pairs (U, h), where $U \subset M$ is an open set and h is a homeomorfism from U to an open set $h(U) \subset E$. A pair $(U, h) \in \mathcal{D}$ is called a *local chart*. We assume that the collection $\mathcal{D}$ of local charts satisfies the following requirements:

1. For each (U, h) and (U', h') in $\mathcal{D}$, the map $h' \circ h^{-1} : h(U \cap U') \to h'(U \cap U')$ is $\mathcal{C}^\infty$, i.e., the local charts (U, h) and (U', h') 'change smoothly' with each other.
2. If $U \subset M$ is open, $h : U \to h(U) \subset E$ is a homeomorfism, and (U, h) changes smoothly with all local charts in $\mathcal{D}$, then (U, h) is itself in $\mathcal{D}$.

A collection $\mathcal{D}'$ that satisfies property 1 and $2'$. $\bigcup\{U : (U, h) \in \mathcal{D}'\} = M$ is called an atlas for M. Any atlas can be extended to an n-dimensional differentiable structure. Of course, different atlases can define the same differentiable structure. By definition, an n-dimensional differentiable manifold is a second countable Hausdorff space, equipped with an n-dimensional differentiable structure which is generated by a countable atlas.

By definition, $\mathcal{C}^\infty(M)$ is the class of all real functions f on M such that $f \circ h^{-1}$ is $\mathcal{C}^\infty$ on $h(U)$ for each local chart (U, h). As the name suggest, we can define derivatives of functions $f \in \mathcal{C}^\infty(M)$.

A parameter curve in M is a pair (I, α) where $I \subset \mathbb{R}$ is an open neighbourhood of 0 and $\alpha : I \to M$. By definition, α is differentiable if $h \circ \alpha$ is continuously differentiable for all local charts. We now want to give meaning to the speed $\dot{\alpha}(0) := \frac{\partial}{\partial t}\alpha(t)|_{t=0}$ of α in the point $t = 0$. We start by saying when two curves have the same speed. Let $\alpha, \tilde{\alpha}$ be parameter curves with $\alpha(0) = x = \tilde{\alpha}(0)$. We write

$$\alpha \sim \tilde{\alpha},$$

if $(h \circ \alpha)^{\cdot}(0) = (h \circ \tilde{\alpha})^{\cdot}(0)$ for one and then for every local chart around x. This is an equivalence relation, and we denote the set of all equivalence classes $[\alpha]$ by $T_x M$, the tangent space to M at x. We now simply consider $[\alpha]$ as the speed of α at $t = 0$. The space $T_x M$ can in a natural way be equipped with the structure of an n-dimensional linear space. To see this, we use the fact that

$$\alpha \sim \beta \Leftrightarrow (f \circ \alpha)^{\cdot}(0) = (f \circ \beta)^{\cdot}(0) \qquad \forall f \in \mathcal{C}^\infty(M).$$

Thus, we can associate an equivalence class $[\alpha] \in T_x M$ with the map $f \mapsto [\alpha]f$, given by

$$[\alpha]f := (f \circ \alpha)^{\cdot}(0).$$

Linear combinations of such maps are defined in the obvious way. Any linear combination $\lambda[\alpha] + \mu[\beta]$ can be associated with some equivalence class $[\gamma]$ of parameter curves, and in this way $T_x M$ is equipped with the structure of an n-dimensional linear space.

A vector field on M is a collection $A = (A(x))_{x \in M}$ with $A(x) \in T_x M$ for all $x \in M$. We define, for $f \in \mathcal{C}^\infty(M)$

$$(Af)(x) := A(x)f,$$

where $A(x) \in T_x M$ is associated with a map of the form $f \mapsto [\alpha]f$ as described above.

By definition, a system of local coordinates is a local chart (U, h) plus a basis $(e^1, \ldots, e^n)$ for E. Our local chart provides us, for each $x \in U$, with a natural identification between the space T_xM and E, when we identify $[\alpha] \in T_xM$ with $(h \circ \alpha)^{\cdot}(0) \in E$. We use this identification to transport our basis of E to T_xM. Thus, for each $x \in U$ we define a basis $(e^1(x), \ldots, e^n(x))$ for T_xM in such a way that the speed of $e^i(x) \in T_xM$ is the equivalence class all α such that $(h \circ \alpha)^{\cdot}(0) = e^i \in E$.

When we are given a system of local coordinates, we can write a vector field A on U as

$$A(x) = \sum_i A_i(x) e^i(x).$$

This is called A, expressed in local coordinates. Let $A(x) = [\alpha]$, for some parameter curve α; then through our identification between T_xM and E the equation above reads $(h \circ \alpha)^{\cdot}(0) = \sum_i A_i(x) e^i$. Thus $A_i(x) = (h \circ \alpha)^{\cdot}(0) = [\alpha] h_i = (Ah_i)(x)$, and

$$A_i(x) = (Ah_i)(x).$$

We say that a vector field is smooth if $A_1, \ldots, A_n$ are $\mathcal{C}^\infty$ on U, in each system of local coordinates. This is equivalent to $Af \in \mathcal{C}^\infty(M)$ for all $f \in \mathcal{C}^\infty(M)$.

For any $f \in \mathcal{C}^\infty(M)$, let $\overline{f} \in \mathcal{C}^\infty(h(U))$ be defined as

$$\overline{f}(h(x)) := f(x).$$

In this notation, we have

$$\overline{Af}(z) = \sum_i \overline{A}_i(z) \tfrac{\partial}{\partial z_i} \overline{f}(z).$$

A function f or a vector field A are smooth if and only if $\overline{f}$, or all the $\overline{A}_i$, respectively, are $\mathcal{C}^\infty$ on $h(U)$. *From now on, we reserve the word vector field for smooth vector fields.*

Now imagine that we are given two systems of local coordinates, (U, h) with a basis $(e^1, \ldots, e^n)$, and (U', h') with a basis $(e'^1, \ldots, e'^n)$. Write $A_i(x)$ for A expressed in the first system of local coordinates, and A'_i for the second system of coordinates. For any $f \in \mathcal{C}^\infty(M)$, write $\overline{f}(h(x)) := f(x) =: \hat{f}(h'(x))$. Let $\phi := h' \circ h^{-1}$, and write $\phi(z) = \sum_i \phi_i(z) e'^i$. Then $\hat{A}'_i(h'(x)) = A'_i(x) = (Ah'_i)(x) = (A(\phi_i \circ h))(x) = \overline{A(\phi_i \circ h)}(h(x)) = \sum_j \overline{A}_j(h(x)) (\frac{\partial}{\partial z_j} \overline{\phi_i \circ h})(h(x)) = \sum_j \overline{A}_j(h(x)) (\frac{\partial}{\partial z_j} \phi_i)(h(x))$ and hence

$$\hat{A}'_i(\phi(z)) = \sum_j \overline{A}_j(z) (\tfrac{\partial}{\partial z_j} \phi_i)(z). \tag{4}$$

This is called the transformation rule for vectorfields; compare formula (1).

Stratonovich stochastic differential equations on manifolds

Let M be a n-dimensional differentiable manifold and let X be a continuous M-valued process on a filtered probability space, equipped with an m-dimensional Brownian motion. For simplicity, we restrict ourselves to M compact. What we say will also be true for processes taking values in a compact subset of a general manifold, or for general manifolds, if the right technicalities involving the local martingale problem etc. are added.

Let $A_0, \ldots, A_m$ be vector fields on M. By definition, we say that X solves the Stratonovich stochastic differential equation

$$dX = A_0(X)dt + \sum_{u=1}^{m} A_u(X) * dB_u,$$

if for every $f \in \mathcal{C}^\infty(M)$

$$df(X) = (A_0 f)(X)dt + \sum_{u=1}^{m} (A_u f)(X) * dB_u.$$

Note that by (2), this is consistent with our previous definition of a Stratonovich stochastic differential equation. Again (compare (3)), we can check that $f(X)$ solves the (Itô) stochastic differential equation

$$df(X) = (A_0 f)(X)dt + \tfrac{1}{2}\sum_{u=1}^{m} (A_u^2 f)(X)dt + \sum_{u=1}^{m} (A_u f)(X)dB_u,$$

and therefore X solves the martingale problem for the operator L, acting on $\mathcal{C}^\infty(M)$, written in the Hörmander form

$$L = A_0 + \tfrac{1}{2}\sum_{u=1}^{m} A_u^2.$$

We now check how our equation looks in local coordinates. Let (U, h) be a local chart and write $h(x) = \sum_i h_i(x)e^i$, where $(e^1, \ldots, e^n)$ is a basis for E. We use the notation

$$\overline{X}_i := h_i(X)$$

for the image of our process X under h. (Clearly, we must restrict the process in some way to U. For example, we can start it in U and only consider it up to a stopping time τ such that $P[X(t) \in U \ \forall t \leq \tau] = 1$. I will not be too precise on this.) Then

$$\overline{X}_i = \overline{A}_{i0}(\overline{X})dt + \sum_{u=1}^{m} \overline{A}_{iu}(\overline{X}) * dB_u,$$

where we have defined $\overline{A}_u$ in the usual way. To see this, just note that

$$d\overline{X}_i = dh_i(X) = (A_0 h_i)(X)dt + \sum_{u=1}^{m} (A_u h_i)(X) * dB_u,$$

where $(A_u h_i)(x) = \overline{A}_{iu}(x)$.

As on $\mathbb{R}^n$, we can ask for existence and uniqueness of solutions of stochastic differential equations. The following can be proved either with the use of local coordinates, or with the help of Whitney's imbedding theorem, which says that any n-dimensional manifold can be imbedded in $\mathbb{R}^k$, for k sufficiently large (for example $k = 2n + 1$).

Theorem 1 *Assume that $A_0, \dots, A_m$ are smooth vectorfields and that M is compact. Then, for each M-valued initial condition $X(0)$, independent of the Brownian motion, the Stratonovich stochastic differential equation*

$$dX = A_0(X)dt + \sum_{u=1}^{m} A_u(X) * dB_u$$

has a unique strong solution.

Here strong and weak uniqueness for stochastic differential equations on manifolds are defined as for ordinary stochastic differential equations.

Examples

We are going to construct Brownian motion on the unit circle. Let $(B(t))_{t\geq 0}$ be standard one-dimensional Brownian motion. Define a 2-dimensional process X by

$$X(t) = (X_1(t), X_2(t)) := (\sin(B(t)), \cos(B(t))).$$

Then, by Itô's formula

$$\begin{aligned} dX_1(t) &= \cos(B(t))dB(t) - \tfrac{1}{2}\sin(B(t))dt = X_2(t)dB(t) - \tfrac{1}{2}X_1(t)dt \\ dX_2(t) &= -\sin(B(t))dB(t) - \tfrac{1}{2}\cos(B(t))dt = -X_1(t)dB(t) - \tfrac{1}{2}X_2(t)dt. \end{aligned}$$

Thus, if, for $x \in \mathbb{R}^2$, we define vectors (or, if you want, 2×1-matrices) $\sigma(x)$ by

$$\sigma_1(x) := x_2 \qquad \sigma_2(x) := -x_1,$$

then X solves the stochastic differential equation

$$dX_i = -\tfrac{1}{2}X_i dt + \sigma_i(X)dB.$$

Likewise, by the Itô-Stratonovich formula

$$\begin{aligned} dX_1(t) &= \cos(B(t)) * dB(t) = X_2(t) * dB(t) \\ dX_2(t) &= -\sin(B(t)) * dB(t) = -X_1(t) * dB(t), \end{aligned}$$

and hence X solves the Stratonovich stochastic differential equation

$$dX_i = \sigma_i(X) * dB.$$

So far, we are considering X as a process in $\mathbb{R}^2$. This is of course not very elegant, since X only takes values in the space $M := \{x \in \mathbb{R}^2; |x| = 1\}$. It is therefore a bit odd to define σ on all of $\mathbb{R}^2$, when we only need it on M. We will now do this more correctly.

M inherits the structure of a (1-dimensional) manifold from its imbedding in $\mathbb{R}^2$. More precisely, we equip M with the induced topology and an atlas which consists of all (U, h) where $U \subset M$ is open and $h : U \to h(U) \subset \mathbb{R}$ is a homeomorfism, such that we can find an open set $\overline{U} \subset \mathbb{R}^2$ with $\overline{U} \cap M = U$, and a local chart $\overline{h} : \overline{U} \to \overline{h}(\overline{U}) \subset \mathbb{R}^2$, with $\overline{h}(x) = (h(x), 0)$ for all $x \in U$.

It turns out that $\mathcal{C}^\infty(M) = \{f|_M : f \in \mathcal{C}^\infty(\mathbb{R}^2)\}$, i.e., the $\mathcal{C}^\infty$ functions on M are just the restrictions of $\mathcal{C}^\infty$ functions on $\mathbb{R}^2$ to M.

The imbedding $M \subset \mathbb{R}^2$ also provides us with a natural imbedding of T_xM in $\mathbb{R}^2$. The differentiable parameter curves in M are just the differentiable parameter curves in $\mathbb{R}^2$ that take values in M only. If α is such a parameter curve, then we denote its speed (in $\mathbb{R}^2$) by $\dot{\alpha}(t) := (\frac{\partial}{\partial t}\alpha_1(t), \frac{\partial}{\partial t}\alpha_2(t))$. Now if α, β are differentiable parameter curves in M with $\alpha(0) = \beta(0) = x$, then $\alpha \sim \beta$ if and only if $\dot{\alpha}(0) = \dot{\beta}(0)$ in $\mathbb{R}^2$. Thus, we can identify T_xM with the 1-dimensional linear subspace of $\mathbb{R}^2$ consisting of all speeds of M-valued differentiable parameter curves $\alpha : I \to \mathbb{R}^2$ with $\alpha(0) = x$. Write, for $x \in M$

$$E(x) := (x_2, -x_1).$$

Then it is not hard to see that

$$T_xM = \operatorname{span}\{E(x)\}.$$

Now E is a vectorfield on M, and we see that X solves the Stratonovich stochastic differential equation

$$dX = E(X) * dB.$$

Thus, X solves the martingale problem for the operator (in Hörmander form)

$$Lf := \tfrac{1}{2}E^2 f.$$

Here $f \in \mathcal{C}^\infty(M)$ and the differential operator E is defined in terms of differentiation along curves α that lie entirely in M. Therefore we now have a formulation that truly considers X as a process in M, instead of a process in $\mathbb{R}^2$ that just happens to take values in M.

A solution X to the martingale problem for L we call a Brownian motion on the unit circle. As remarked before, there is in general more than one way to write a differential operator in Hörmander form. Each different way of writing L in Hörmander form leads to a different Stratonovich stochastic differential equation, which corresponds to a different way of coupling X to a Brownian motion. Here we give an example of one more way. This time, we couple X to a 2-dimensional Brownian motion. The corresponding Stratonovich stochastic differential equation arises in a natural way from the imbedding of M in $\mathbb{R}^2$.

Write P_x for the orthogonal projection from $\mathbb{R}^2$ to T_xM imbedded in $\mathbb{R}^2$ in the natural way. Note that $|E(x)| = 1$ (in $\mathbb{R}^2$) and hence $P_x z = \langle E(x), z\rangle E(x)$. (Here $\langle \cdot, \cdot \rangle$ is the standard inner product on $\mathbb{R}^2$ and $|\cdot| = \langle \cdot, \cdot \rangle^{1/2}$.) Thus, if we define vector fields A_1, A_2 on M by

$$\begin{aligned} A_1(x) &:= x_2 E(x) \\ A_2(x) &:= -x_1 E(x), \end{aligned}$$

then $P_x z = \sum_{u=1}^2 A_u(x) z_u$. Now consider the Stratonovich stochastic differential equation

$$dX = \sum_{u=1}^2 A_u(X) * dB_u,$$

where $B = (B_1, B_2)$ is 2-dimensional Brownian motion. We have coupled X and B in such a way that $X(t)$ always makes an infinitesimal step $dX(t)$ in the same direction as the projection of $dB(t)$ on $T_{X(t)}M$.

Solutions to this equation solve the martingale problem for the operator

$$L = \tfrac{1}{2} \sum_{u=1}^{2} A_u^2.$$

Exercise: Show that this is the same L as the one we had before, and hence solutions to our new Stratonovich stochastic differential equation are again Brownian motion on the circle. Hint: Let Y be a solution to $dY = sin(Y)dB_1 + \cos(Y)dB_2$ and put $X := (\sin(Y), \cos(Y))$.

Remark: Although solutions X to the Stratonovich stochastic differential equations

$$\begin{aligned} &(1)\ dX = E(X) * dB \\ &(2)\ dX = \sum_{u=1}^{2} A_u(X) * dB_u \end{aligned}$$

are in both cases Brownian motion on the circle, they differ in the way X is coupled to a Brownian motion. This has the following consequence. Let X^x, X^y be Brownian motions on the circle starting in the points x, y, respectively. We can construct these processes as solutions to equation (1), adapted to the same 1-dimensional Brownian motion B, or as solutions to equation (2), adapted to the same 2-dimensional Brownian motion (B_1, B_2). In both cases, the result is a joint process (X^x, X^y), where the equations (1) and (2) provide us with two different ways of coupling X^x and X^y. In order to demonstrate that these couplings are very different, we remark that, for solutions X^x and X^y to (1), with the same B, one has

$$|X^x(t) - X^y(t)| = |x - y| \quad \forall t \geq 0 \quad \text{a.s.}$$

On the other hand, if X^x and X^y are solutions to (2), with the same (B_1, B_2), then

$$\lim_{t\to\infty} |X^x(t) - X^y(t)| = 0 \quad \text{a.s.}$$

Here $|\cdot|$ denotes the Euclidian norm on $\mathbb{R}^2$.

Exercise: Let M be the surface of the unit sphere in $\mathbb{R}^3$ and let X be Broanian motion on M. Construct M by means of a Stratonovich stochastic differential equation which couples X to 3-dimensional Brownian motion in such a way that $X(t)$ always makes an infinitesimal step $dX(t)$ in the same direction as the projection of $dB(t)$ on $T_{X(t)}M$.

Tensor fields

Denote by T_x^*M the dual of the space T_xM, i.e. T_x^*M is the space of all linear forms $l : T_xM \to \mathbb{R}$. A covariant vector field on M is a collection $B = (B(x))_{x\in M}$ with $B(x) \in T_x^*M$ for all $x \in M$.

Assume that we are given local coordinates, i.e., a local chart (U, h) and a basis $(e^1, \ldots, e^n)$ for E, and let $(e_1(x), \ldots, e_n(x))$ be the basis for T_xM defined before. Let $(e_1(x), \ldots, e_n(x))$ be the associated basis for T_x^*M, i.e., $e_i(x)(e^j(x)) := \delta_{ij}$. Then B, expressed in local coordinates, reads

$$B(x) = \sum_i B^i(x) e_i(x).$$

We say that B is smooth if the B^i are $\mathcal{C}^\infty$ on U, in each system of local coordinates. We adopt the convention of using upper indices for coordinates of covectors and under indices for normal vectors. If B is a covariant vectorfield, and A a vectorfield, then

$$(BA)(x) := B(x)(A(x)) = \sum_i B^i(x)A_i(x).$$

Covariant vectorfields transform according to the rule

$$\hat{B}'^i(y) = \sum_j \overline{B}^j(\phi^{-1}(y))(\tfrac{\partial}{\partial y_i}\phi_j^{-1})(y), \tag{5}$$

where we use notation as in (4).

A tensorfield on M is a collection $A = (A(x))_{x\in M}$ with $A(x) \in E_x$ for all $x \in M$, where E_x is the tensor product of a number of copies of T_xM and T_x^*M. For example, a one time contravariant and two times covariant tensorfield is a collection $A = (A(x))_{x\in M}$ with $A(x) \in T_xM \otimes T_x^*M \otimes T_x^*M$ for all $x \in M$. Expressed in local coordinates:

$$A(x) = \sum_{ijk} A_i{}^{jk}(x)\, e^i(x) \otimes e_j(x) \otimes e_k(x).$$

A tensorfield is smooth if it is $\mathcal{C}^\infty$ in local coordinates. *From now on, a tensorfield always means a smooth tensorfield.* Special cases are vectorfields, covariant vectorfields and even functions $f \in \mathcal{C}^\infty(M)$, when we consider $\mathbb{R}$ as the tensor product of no copies of T_xM and no copies of T_x^*M. Smooth functions are sometimes called scalar fields.

We mention the contraction of a tensorfield. If A is a one time contravariant and two times covariant tensorfield, then the contraction of its first and third coordinate yields a covariant vector field B, which in local coordinates takes the following form:

$$B^i(x) = \sum_j A_j{}^{ij}(x).$$

For brevity, we often drop the x in the notation and adopt the summation convention, i.e., when an index occurs twice (below and above) it is understood that we sum over it. The formula above thus reads $B^i = A_j{}^{ij}$.

We give one example of a covariant vector field. For any fixed $f \in \mathcal{C}^\infty(M)$ and $x \in M$, the map $[\alpha] \mapsto [\alpha]f$ defined on page 6 is an element of T_x^*M. We denote this map by $\partial f(x)$. In local coordinates:

$$\partial f(x) = \sum_i \partial^i f(x) e_i(x).$$

Now the collection $\partial f = (\partial f(x))_{x\in M}$ is a (smooth!) covariant vectorfield on M. Note that in local coordinates

$$\sum_i \overline{\partial^i f}(h(x))\overline{A}_i(h(x)) = \sum_i \partial^i f(x)A_i(x) = (Af)(x) = \sum_i \overline{A}_i(h(x))(\tfrac{\partial}{\partial z_i}\overline{f})(h(x)),$$

so that

$$\overline{\partial^i f}(z) = (\tfrac{\partial}{\partial z_i}\overline{f})(z).$$

One may wonder if there exists a two times covariant tensorfield $\partial^2 f$ such that, in local coordinates

$$\overline{\partial^2 f}^{ij}(z) \stackrel{?}{=} (\frac{\partial^2}{\partial z_i \partial z_j}\overline{f})(z).$$

This is *not* the case. We can easily understand this as follows. Assume that such a tensorfield $\partial^2 f$ exists, and let A be a two times contravariant tensorfield. Then the second order differentiable operator

$$(Lf)(x) := \sum_{ij} A_{ij}(x)\partial^2 f^{ij}(x)$$

consists, in any system of local coordinates, of second order derivatives only. We know, however, (think about Itô's formula) that certain transformations of the space produce first order derivatives in the expression for L, and hence we arrive at a contradiction.

In order to give a proper definition of the second derivative of f, we need on our manifold the additional structure of a connection. In particular, we can define a second derivative if we are on a Riemannian manifold.

Connections

If $f \in \mathcal{C}^\infty(M)$, then we saw in the last section that there exists a covariant vectorfield ∂f with components $\partial^i f$. However, if A is a vectorfield on M, there does not exist a one time covariant, one time contravariant tensorfield whose components are $\partial^i A_j$.

We now define derivatives of vectorfields. By definition, a linear map ∇ from the space of vectorfields on M into the space of one time covariant, one time contravariant tensorfields on M is called a covariant derivative, if ∇ satisfies the product rule

$$\nabla(fA)(x) = \partial f(x) \otimes A(x) + f(x)(\nabla A)(x).$$

If ∇ is a covariant derivative, then using local coordinates we can write

$$\nabla A(x) = \nabla\Big(\sum_k A_k(x)e^k(x)\Big) = \sum_k \partial A_k(x) \otimes e^k(x) + \sum_k A_k(x)\nabla e^k(x).$$

We use the notation

$$\nabla A(x) =: \sum_{ij}(\nabla^i A)_j(x),$$

so that

$$(\nabla^i A)_j(x) = (\partial^i A)_j(x) + \sum_k A_k(x)(\nabla^i e^k)_j(x).$$

Here the functions $\Gamma^i{}_j{}^k(x) := (\nabla^i e^k)_j(x)$ are the *Christoffel symbols.* They clearly determine the covariant derivative. Using summation convention and dropping the x in our notation:

$$(\nabla^i A)_j = \partial^i A_j + \Gamma^i{}_j{}^k A_k.$$

Note that the $(\nabla^i A)_j$ are the coordinates of a tensor, but the $\partial^i A_j$ are not and hence neither are the $\Gamma^i{}_j{}^k$.

Once we have chosen Christoffel symbols and hence a covariant derivative, we say that we have equipped our manifold with a connection. A connection can also be characterised in terms of parallel transport. If α is a differentiable parameter curve and A a vectorfield, we say that A is parallel along α if

$$\dot\alpha(t)_i(\nabla^i A(t))_j = 0.$$

We will not explore this concept any further.

Once we have fixed the covariant derivative of vectorfields, there is a natural way to define the covariant derivative of covariant vectorfields. There is a unique linear map $\hat\nabla$ from the space of covariant vectorfields on M into the space of two times covariant tensorfields on M, satisfying the product rule

$$\hat\nabla(fA)(x) = \partial f(x) \otimes A(x) + f(x)(\hat\nabla A)(x).$$

and such that in addition the following 'contraction product rule' holds:

$$\partial^i(B^j A_j) = (\hat\nabla^i B)^j A_j + B^j(\nabla^i A)_j.$$

To see that this is so, first note that $\hat\nabla$ can be written in the form

$$(\hat\nabla^i B)^j = \partial^i B^j + \Gamma^{ij}{}_k B^k,$$

where by definition the $\Gamma^{ij}{}_k$ are the Christoffel symbols of $\hat\nabla$. (We can distinguish them in our notation from the Christoffel symbols for ∇ by the different order of lower and upper indices.) Now the 'contraction product rule' says that

$$(\partial^i B^j)A_j + B^j(\partial^i A_j) = (\partial^i B^j)A_j + \Gamma^{ij}{}_k B^k A_j + B^j(\partial^i A_j) + B^j \Gamma^i{}_j{}^k A_k,$$

and hence $B^k\Gamma^{ij}{}_k A_j = -B^k \Gamma^i{}_k{}^j A_j$. This is true for all A, B if and only if

$$\Gamma^{ij}{}_k = -\Gamma^i{}_k{}^j.$$

Let X, Y be vectorfields, and let $f \in \mathcal{C}^\infty(M)$. Then the commutator (or 'Lie product') $[X, Y]$ of X and Y is defined as

$$[X, Y]f := XYf - YXf.$$

In coordinates

$$[X, Y]f = (X_i\partial^i)(Y_j\partial^j)f - (Y_i\partial^i)(X_j\partial^j)f = \Big(X_i(\partial^i Y_j) - Y_i(\partial^i X_j)\Big)\partial^j f.$$

One can check that $[X, Y]$ is a vectorfield; in particular, for each x there is a $[\alpha] \in T_x M$ such that $[X, Y]f(x) = (f \circ \alpha)^{\cdot}(0)$.

If ∇ is a covariant derivative and X, Y are vectorfields, then we use the notation

$$\nabla_X Y := X_i(\nabla^i Y)_j.$$

By definition, a connection is torsion-free if for all vectorfields X, Y, $\nabla_X Y - \nabla_Y X = [X, Y]$. In coordinates

$$X_i(\nabla^i Y)_j - Y_i(\nabla^i X)_j = X_i\partial^i Y_j - Y_i\partial^i X_j.$$

For the Christoffel symbols this means that

$$X_i \Gamma^i{}_j{}^k Y_k - Y_i \Gamma^i{}_j{}^k X_k = 0.$$

This is true for all X, Y if and only if

$$\Gamma^i{}_j{}^k = \Gamma^k{}_j{}^i.$$

Equivalent statements are:

$$\Gamma^{ij}{}_k = \Gamma^{ji}{}_k \quad \Leftrightarrow \quad (\hat\nabla^i \partial f)^j = (\hat\nabla^j \partial f)^i.$$

Riemannian manifolds

Let M be a differentiable manifold. By definition, a metric on M is an object that gives, for each $x \in M$, an inner product $\langle \cdot, \cdot \rangle_x$ on T_xM. An inner product is a bilinear form on T_xM and therefore in a natural way an element of $T_x^*M \otimes T_x^*M$. We denote this tensor by $g(x)$. Thus, if A and B are vector fields, then in local coordinates:

$$\langle A, B \rangle (x) := \langle A(x), B(x) \rangle_x = \sum_{ij} A_i(x) g^{ij}(x) B_i(x).$$

We note that an inner product is symmetric, $\langle x, y \rangle = \langle y, x \rangle$, and hence in local coordinates

$$g^{ij} = g^{ji}.$$

We associate g with the linear map from T_xM to T_x^*M that maps a vector $a \in T_xM$ into the linear form $g(a) \in T_x^*M$ given by $(g(a))(b) = \langle a, b \rangle_x$. In coordinates:

$$g(a)^i = g^{ij} a_j.$$

Now $g^{-1} : T_x^*M \to T_xM$ can be associated with a tensor in $T_xM \otimes T_xM$. In coordinates:

$$g^{-1}(b)_i = g^{-1}{}_{ij} b^j.$$

The maps g and g^{-1} provide us with a natural identification between vectors and covariant vectors, and it is customary to identify them and denote them by the same symbol. Thus, we denote a and $g(a)$ both by the symbol a, and distinguish their coordinates a_i and $g(a)^i =: a^i$ (which are not the same!) only by the position of the index. In the same way, we can raise and lower the indices of any tensor with the help of g and g^{-1}. If we apply this convention to g itself, then we see that

$$g_{ij} = g^{-1}{}_{ik} g^{-1}{}_{jl} g^{kl} = g^{-1}{}_{ij}.$$

Thus, our identification between T_xM and T_x^*M leads us to identify $g(x) \in T_x^*M \otimes T_x^*M$ with $g^{-1}(x) \in T_xM \otimes T_xM$. We therefore denote them both simply by g.

We say that a metric is smooth if the tensor field $(g(x))_{x \in M}$ is smooth. By definition, a Riemannian manifold is an n-dimensional differentiable manifold, equipped with a smooth metric.

If we are on a Riemannian manifold, there is a canonical way to choose the Christoffel symbols and hence the covariant derivative. There exists a unique torsion-free covariant derivative ∇ on M such that, for every vectorfield A

$$(\nabla^i A)_j = g_{jk} (\hat{\nabla}^i A)^k.$$

This means that ∇A and $\hat{\nabla} A$ are just the same thing, in the usual identification between T_xM and T_x^*M brought about by g. The 'contraction product rule' now takes the form

$$\partial^i (X_j g^{jk} Y_k) = (\nabla^i X)_j g^{jk} Y_k + X_j g^{jk} (\nabla^i Y)_k,$$

which implies that for each X, Y, Z

$$Z \langle X, Y \rangle = \langle \nabla_Z X, Y \rangle + \langle X, \nabla_Z Y \rangle,$$

the so-called Ricci-identity.

To see that our covariant derivative is indeed uniquely determined, use the 'contraction product rule' to see that (the terms with derivatives of vector fields cancel)

$$\begin{aligned} X_j(\partial^i g^{jk})Y_k &= \Gamma^i{}_j{}^l X_l g^{jk} Y_k + X_j g^{jk} \Gamma^i{}_k{}^l Y_l \\ &= X_j \Gamma^i{}_l{}^j g^{lk} Y_k + X_j g^{jl} \Gamma^i{}_l{}^k Y_k, \end{aligned}$$

which is true for all X, Y if and only if

$$\partial^i g^{jk} = g^{kl}\Gamma^i{}_l{}^j + g^{jl}\Gamma^i{}_l{}^k.$$

Write $G^{ijk} := g^{jl}\Gamma^i{}_l{}^k$. Then the requirement that the connection be torsion-free leads to $G^{ijk} = g^{jl}\Gamma^i{}_l{}^k = g^{jl}\Gamma^k{}_l{}^i = G^{kji}$, and therefore

$$\begin{aligned} \partial^i g^{jk} &= G^{ijk} + G^{ikj} \\ G^{ijk} &= G^{kji}. \end{aligned}$$

We can solve these equations by writing

$$\begin{aligned} &\partial^i g^{jk} + \partial^j g^{ki} - \partial^k g^{ij} = G^{ijk} + G^{ikj} + G^{jki} + G^{jik} - G^{kij} - G^{kji} \\ &= (G^{ikj} + G^{jki}) + (G^{ijk} - G^{kji}) + (G^{jik} - G^{kij}) = 2G^{ikj}. \end{aligned}$$

Which leads to the formula (Levi-Civita connection)

$$\Gamma^i{}_j{}^k = \tfrac{1}{2} g_{jl}\Big(\partial^i g^{kl} + \partial^k g^{li} - \partial^l g^{ik}\Big).$$

An equivalent formula is

$$\Gamma^{ij}{}_k = \tfrac{1}{2} g_{kl}\Big(\partial^l g^{ij} - \partial^i g^{jl} - \partial^j g^{il}\Big).$$

We can now identify ∇ and $\hat{\nabla}$, and use the symbol ∇ for both of them.

Itô stochastic differential equations and connections

Let M be a compact n-dimensional differentiable manifold, equipped with a connection. For any $f \in \mathcal{C}^\infty(M)$, denote by Df be the two times covariant tensorfield

$$D^{ij} f = (\hat{\nabla}^i \partial f)^j.$$

Note that $D^{ij} = D^{ji}$ if the connection is torsion-free. Let $b, \sigma_1, \ldots, \sigma_m$ be vectorfields on M, and define a two times contravariant tensorfield a by

$$a_{ij} = \sum_u \sigma_{iu}\sigma_{ju}.$$

On other words, $a(x) = \sum_u \sigma_u(x) \otimes \sigma_u(x)$. Let L be the operator

$$Lf = b_i \partial^i f + \tfrac{1}{2} a_{ij} D^{ij} f.$$

Let us call this a differential operator in standard form. Given the connection, there is only one way to write a given L in standard form, i.e. the coefficients b, a are unique. They depend, however, on the choice of the connection!

We say that a continuous M-valued process X on a filtered probability space, equipped with m-dimensional Brownian motion, solves the Itô stochastic differential equation

$$dX = b(X)dt + \sum_u \sigma_u(X)dB_u,$$

if for every $f \in \mathcal{C}^\infty(M)$

$$df(X) = (Lf)(X)dt + \sum_u (\sigma_u f)(X)dB_u.$$

Note that this definition of a solution to an (Itô) stochastic differential equation is consistent with our earlier definition.

We now investigate the relation between solutions to Stratonovich and Itô stochastic differential equations. Let $A_0, \ldots, A_m$ be vectorfields on M and let X be a solution to the Stratonovich stochastic differential equation

$$dX = A_0(X)dt + \sum_{u=1}^m A_u(X) * dB_u.$$

Then, as we have seen, for any $f \in \mathcal{C}^\infty(M)$, the process $f(X)$ solves the Itô stochastic differential equation

$$df(X) = (Lf)(X)dt + \sum_{u=1}^m (A_u f)(X)dB_u,$$

where L is the operator in Hörmander form

$$L = A_0 + \tfrac{1}{2}\sum_{u=1}^m A_u^2.$$

Thus, by our definition, X solves the Itô stochastic differential equation

$$dX = b(X)dt + \sum_u \sigma_u(X)dB_u,$$

if the coefficients b and σ are such that $\sigma_u = A_u$ $(u = 1, \ldots, m)$ and

$$Lf = b_i \partial^i f + \tfrac{1}{2} a_{ij} D^{ij} f.$$

In order to see what b is, we must bring L, which is given in Hörmander form, in standard form. Note that

$$\begin{aligned}
A_u^2 f &= \left(A_{iu}\partial^i\right)\left(A_{ju}\partial^j\right) f = A_{iu}A_{ju}\partial^i\partial^j f + A_{iu}(\partial^i A_{ju})\partial^j f \\
&= A_{iu}A_{ju}(D^{ij}f) + A_{iu}(\partial^i A_{ju})\partial^j f - A_{iu}A_{ju}\Gamma^{ij}{}_k \partial^k f \\
&= A_{iu}A_{ju}(D^{ij}f) + A_{iu}\left(\partial^i A_{ju} - A_{ku}\Gamma^{ik}{}_j\right)\partial^j f \\
&= A_{iu}A_{ju}(D^{ij}f) + A_{iu}\left(\partial^i A_{ju} + \Gamma^i{}_j{}^k A_{ku}\right)\partial^j f \\
&= A_{iu}A_{ju}(D^{ij}f) + A_{iu}(\nabla^i A_u)_j \partial^j f,
\end{aligned}$$

which leads to the formulas

$$\begin{aligned}
\sigma_u &= A_u & (u = 1, \ldots, m) \\
b &= A_0 + \tfrac{1}{2}\sum_u \nabla_{A_u} A_u.
\end{aligned}$$

This is true for any connection, although in applications we often have the canonical connection on a Riemannian manifold.

Conclusions

We have seen that it is possible to define Stratonovich stochastic differential equations on an arbitrary manifold, but in order to define ordinary (Itô) stochastic differential equations, we need the additional structure of a connection.

Note that, although we have defined solutions of stochastic differential equations on manifolds, we have not defined stochastic integrals on manifolds! Let $A_0, \dots, A_m$ be vectorfields on M and let X be a solution to the Stratonovich stochastic differential equation

$$dX = A_0(X)dt + \sum_{u=1}^{m} A_u(X) * dB_u.$$

Then is seems natural to read this equation in integral form:

$$X(t) = X(0) + \int_0^t A_0(X(s))ds + \sum_{u=1}^{m} \int_0^t A_u(X(s)) * dB_u(s).$$

However, it is not at all clear what the meaning of an individual integral on the right-hand side should be. Consider, for example, the process

$$Y(t) = \int_0^t A_1(X(s)) * dB_1(s).$$

Presumably, this integral should be defined by some approximating sheme, so that in the limit $dt \to 0$

$$Y(t+dt) - Y(t) = \tfrac{1}{2}(A_1(X(t+dt)) + A(X(t)))(B_1(t+dt) - B_1(t)).$$

Here the left-hand side is an element of $T_{Y(t)}M$, while the right-hand side is an element of $T_{X(t)}M$. On a manifold, we do not have a natural identification between these spaces, and hence the process Y is ill-defined!

Although we cannot read stochastic differential equations on a manifold in integral form as we are used to, there exist several different ways of defining M-valued stochastic integrals. If the manifold is equipped with a connection, then there exists also a concept of an M-valued martingale, and one can check that solutions to Itô stochastic differential equations with zero drift are martingales.

The examples revisited

Let $M \subset \mathbb{R}^2$ be the unit circle, equipped with a 1-dimensional differentiable structure through its imbedding. The space M inherits in a natural way a metric from its imbedding in $\mathbb{R}^2$, when we use the natural imbedding $T_xM \subset \mathbb{R}^2$. Thus, for $a, b \in T_xM$ we define

$$\langle a, b\rangle_x := \langle a, b\rangle,$$

where the inner product on the right-hand side is in $\mathbb{R}^2$. Write ∇ for the canonical covariant derivative belonging to this metric.

An easy way to characterise ∇ with the help of our imbedding is the following. Let X, Y be vectorfields on M, let $x \in M$ and $t, \lambda \in \mathbb{R}$ be fixed and let α be a smooth parameter curve, defined in a neigbourhood of t, such that $x = \alpha(t)$, $X(x) = \lambda\dot{\alpha}(t)$ and

$$\langle \ddot{\alpha}(t), z\rangle = 0 \quad \forall z \in T_xM.$$

Then

$$(\nabla_X Y)(x) = \lambda P_x \tfrac{\partial}{\partial t} Y(\alpha(t)),$$

where P_x is the orthogonal projection on $T_x M$.
Exercise: Prove this formula.

Consider the Stratonovich stochstic differential equations

$$(1)\ dX = E(X) * dB$$
$$(2)\ dX = \sum_{u=1}^{2} A_u(X) * dB_u$$

introduced on page 11. We will rewrite these equations, using the connection on M, as Itô stochastic differential equations. We start with (1). We put

$$\sigma := E$$
$$b := \tfrac{1}{2}\nabla_E E.$$

To see what E is, define a smooth parameter curve α in M by

$$\alpha(t) := (\sin(t), \cos(t)).$$

It is easily checked that, for all $t \in \mathbb{R}$,

$$\dot{\alpha}(t) = E(\alpha(t))$$
$$\langle \ddot{\alpha}(t), z\rangle = 0 \quad \forall z \in T_{\alpha(t)}M.$$

It therefore follows that, for all $t \in \mathbb{R}$,

$$(\nabla_E E)(\alpha(t)) = P_{\alpha(t)} \tfrac{\partial}{\partial t} E(\alpha(t)) = 0,$$

where we used that $P_{\alpha(t)}\ddot{\alpha}(t) = 0$. We conclude that $b = 0$ and hence X solves the Itô stochastic differential equation

$$(1') \quad dX = E(X)dB.$$

We now rewrite (2). This time, we put

$$\sigma_u := A_u \qquad (u = 1, 2)$$
$$b := \tfrac{1}{2}\sum_{u=1}^{2} \nabla_{A_u} A_u.$$

Here $A_1(x) = x_2 E(x)$, so that $A_1(\alpha(t)) = \cos(\alpha(t))\dot{\alpha}(t)$ and

$$\begin{aligned}(\nabla_{A_1} A_1)(\alpha(t)) &= \cos(\alpha(t)) P_{\alpha(t)} \tfrac{\partial}{\partial t} \cos(\alpha(t))\dot{\alpha}(t) \\ &= \cos(\alpha(t)) P_{\alpha(t)}\Big(-\sin(\alpha(t))\dot{\alpha}(t) + \cos(\alpha(t))\ddot{\alpha}(t)\Big) \\ &= -\cos(\alpha(t))\sin(\alpha(t)).\end{aligned}$$

A similar calculation shows that

$$(\nabla_{A_2} A_2)(\alpha(t)) = \cos(\alpha(t))\sin(\alpha(t)),$$

and hence $\nabla_{A_1} A_1 + \nabla_{A_2} A_2 = 0$ so that again $b = 0$. This time, X solves the Itô stochastic differential equation

$$(2') \quad dX = \sum_{u=1}^{2} A_u(X) dB_u.$$

Recommeded literature

Good, but in German: W. Hackenbroch and A. Thalmaier, *Stochastische Analysis*, Teubner, 1994, Stuttgart.

Technical, treats a lot of stochastic analysis but not stochastic differential equations: M. Emery, *Stochastic calculus in Manifolds*, Springer-Verlag 1989 Berlin.

Contains a section on manifolds: N. Ikeda and S. Watanebe, *Stochastic differential equations and diffusion processes*, North Holland Publ. Comp. 1989 Amsterdam.