Coordinate-free Stochastic Differential Equations as 2-jets

2016 Risk and Stochastics Conference, April 21, 2016, London

London School of Economics, hosted by Winton Capital

Damiano Brigo
Dept. of Mathematics, Imperial College, London
wwwf.imperial.ac.uk/∼dbrigo/
—
Joint work with *John Armstrong*
Dept. of Mathematics, King's College, London
—

Full paper in **http://arxiv.org/abs/1602.03931**

I would like to dedicate this talk to
Giovanni Battista Di Masi (1944-2016),
who passed away on April 4.

PhD at Brown University, Author in signal processing, stochastic control & filtering, probability, stochastic analysis, statistics.

Professor of Probability & Mathematical Statistics, later Head of the Department of Mathematics at the University of Padua and Assessor at the Padua Local Administration

Gianni was my Laurea dissertation supervisor (1990) and he was present at my PhD viva in Amsterdam. I learned a lot from Gianni.

"A true Gentleman, elegant, sharp, witty and generous. At the same time charismatic and amiable. I nostalgically remember our prosec--chini and espressos in the corner bar in via Paolotti. Ciao Gianni."

Elisa Nicolato

Agenda I

1. The traditional view of SDEs: Ito and Stratonovich
   - SDEs and stochastic integrals
   - Historical notes
   - Itô-Stratonovich transformation
2. Itô SDEs on manifolds: 2-Jets
   - Drawing and simulating SDEs as “fields of curves”
   - Coordinate-free converging difference scheme as SDE
   - Coordinate free Itô SDE as 2-jet scheme limit
   - Coordinate-free Itô formula
   - Commutative diagrams and rubber sheets
   - The case of vector Brownian motion as driver
   - Backward diffusion operator: Geometric interpretation
   - Weak and Strong Equivalence
   - SDEs driven by 2-dimensional Brownians. Heston model
   - Fan Diagrams

Agenda II

- Geometric interpretation of SDE coefficients and percentiles
- Itô / Stratonovich as switch between 2-jets & pairs of vector fields
- Itô & Stratonovich as different coordinates

3 Conclusions and References

Classic theory of Stochastic Differential Equations

$$dX_t = a(X_t)dt + b(X_t)dW_t, \quad X_0.$$

- W's paths have unbounded variation, and are nowhere differentiable with probability one. So the above cannot be interpreted as a differential equation directly.
- Write it as

$$X_t = X_0 + \int_0^t a(X_s)ds + \int_0^t b(X_s)dW_s, \quad X_0.$$

- Now the matter is defining the stochastic integral driven by dW
- Since W has unbounded variation, we cannot define this as an ordinary Stiltjes integral on the paths.

- $X_t = X_0 + \int_0^t a(X_s)ds + \boxed{\int_0^t b(X_s)dW_s}$??, X_0
- Traditionally, 2 main definitions of stochastic integrals, given here as generalizations of Riemann-Stiltjes sums for convenience, with convergence in mean square ($L^2(\mathbb{P})$): Initial point vs mid point

$$\int_0^T b(X_s)dW_s = \lim_n \sum_{i=1}^n b(X(t_i))(W_{t_{i+1}} - W_{t_i}) \ \text{(Itô)}$$

$$\int_0^T b(X_s)\circ dW_s = \lim_n \sum_{i=1}^n b\left(X\left(\frac{t_i + t_{i+1}}{2}\right)\right)(W_{t_{i+1}} - W_{t_i})\text{(Stratonovich)}$$

(more general def. has $[b(X(t_i)) + b(X(t_{i+1}))]/2$ in front of $W - W$) where it is understood that as n tends to infinity the mesh size of the partition $\{[0, t_1), [t_1, t_2), \ldots, [t_{n-1}, t_n = T]\}$ of $[0, T]$ tends to 0.

- Stratonovich integral looks into the future, Ito does not.

Battle of the integrals: Ito or Stratonovich?

(Doeblin)-Ito integral:

- "Does not look into the future" (good for social sciences)
- Ito integral is a martingale: many important consequences.
- Coeff. interpretation as local mean and standard deviation: "$\mathbb{E}_t[\delta X_t] \approx a(X_t)\delta t$", "$\mathrm{VAR}_t[\delta X_t] \approx b^2(X_t)\delta t$"
- More generally good probabilistically
- Does not satisfy chain rule, additional second order term

$$dX_t = a(X_t)dt + b(X_t)dW_t,$$

$$df(X_t) = ((\nabla f)(X_t))^T \, dX_t + \frac{1}{2}(dX_t)^T(Hf(X_t))(dX_t)$$

Battle of the integrals: Ito or Stratonovich?

Fisk ([11])-Stratonovich ([27]) (-McShane [21]) Integral:

- Satisfies chain rule

$$dX_t = a(X_t)dt + b(X_t) \circ dW_t, \quad df(X_t) = ((\nabla f)(X_t))^T \circ dX_t$$

- Given chain rule, SDEs in Stratonovich form behave like vector fields under changes of coordinates. Good for coordinate-free SDEs on manifolds and Stochastic Differential Geometry.
- Coefficient a no longer local mean, but related to “median” (later)
- If we take t-C^1 processes $W^{(n)}(t)$ such that $W^{(n)} \to W$ with probability 1, uniformly in t on bounded intervals (Wong Zakai)

Sol of $dX_t^{(n)} = a(X_t^{(n)})dt + b(X_t^{(n)})dW_t^{(n)} \to$ Sol of $dX = a(X)dt + b(X) \circ dW$

Battle of the integrals: Ito or Stratonovich?

- Itô good probabilistically, Stratonovich good geometrically.
- This talk: Can we make Ito integral coordinate free and good for geometry, so as to have good **probability** and **geometry** together? Yes and we'll explain how. Before that:

Historical notes and Itô - Stratonovich transformation

Itô dominates among mathematicians. Mostly Stratonovich integral managed to cut its share because of stochastic geometry. Even there, it is often stated that under the theory it is "Ito's calculus which does all the work" (Rogers & Williams [25], Chapter V.30, p. 184).

Stratonovich fared better with physicists & engineers, due to Wong Zakai & symmetry (see for example Van Kampen's "Itô vs Stratonovich" [28], where it is argued that for Langevin equations with external noise the Stratonovich version can be preferable, while for the case of internal noise any of the versions would work).

Battle of the integrals: Historical notes

Donald Fisk paper, first proposing the symmetric integral, rejected by top probability journal in mid 60's. In 1967, famous probabilist Anatoliy Skorokhod (1930-2011) reviewed [26] Stratonovich's 1966 book:

" The second part, which is the most interesting and the most rigorously treated one, *could have been shortened by one-half, if the author had used, instead of his own integral, the Itô integral.* [...] Searching out the drops of truth from this material is a thankless task, which only the rare reader would undertake. [...]"

.... CHAINSAW!!

Drop the chain rule, start the...

Does Itô-Stratonovich transform solve all problems?

There is a simple rule to re-write an Itô SDE into a Stratonovich SDE admitting the same solution (and vice versa).

This is known as Itô - Stratonovich transformation. *We will give a coordinate-free interpretation of this at the end of the talk.*

$$dX_t = a(X_t)dt + b(X_t)dW_t \quad \to \quad dX_t = \tilde{a}(X_t)dt + b(X_t) \circ dW_t$$

$$\tilde{a}_i = a_i - \frac{1}{2} \sum_{k=1}^{\dim W} \sum_{h=1}^{\dim X} \frac{\partial b_{i,k}}{\partial x_k} b_{h,k}$$

So if we need geometry and Ito, why not do the following:

- Given the initial Ito SDE, transform it in Stratonovich form with the rule above. The solution will be the same.
- Work with geometry using the Stratonovich SDE
- Once you are done, convert the equation back into Ito form

Does Itô-Stratonovich transform solve all problems?

However, the Itô Stratonovich transformation does not commute with operations related to methods for projection on submanifolds.
See for example the exponential-families assumed-density-filters (equivalent to Hellinger metric projection for Stratonovich) in stochastic filtering, B. et al [6]. In that case

- Transform full Itô SPDE for optimal filter in Stratonovich form;
- Apply the assumed density/“expected sufficient statistics matching” approximation to obtain a finite dimensional approximating Stratonovich SDE in the chosen exponential family;
- Convert the approximating Stratonovich SDE into Itô form;

VS

- Apply assumed density approximation directly to the Itô full SPDE

produce two different Itô SDEs.

Geometry & Probability: Ito calculus on manifolds

In some optimal approximation and dimensionality reduction problems it will therefore be important to be able to work directly with Itô calculus in Geometry (eg. Armstrong & B. [2]).

We now come to the main contribution of this talk: define Itô SDEs on manifolds without moving to Stratonovich.

This will be done
via the notion of jets

What are jets?

Not these!! $\rightarrow$

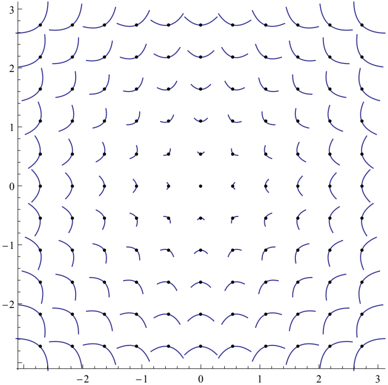
Jets and SDEs

For all $x \in \mathbb{R}^n$ consider smooth curve $\gamma_x : \mathbb{R} \to \mathbb{R}^n, \quad \gamma_x(0) = x$

Example: γ_x^E on $\mathbb{R}^2$ as follows (zero 3d-on derivatives):

$$\gamma^E_{(x_1,x_2)}(t) = (x_1, x_2) + \underbrace{t(-x_2, x_1)}_{\textit{circular counterclock}}$$

$$+ \underbrace{3t^2(x_1, x_2)}_{\textit{radially outward}}$$

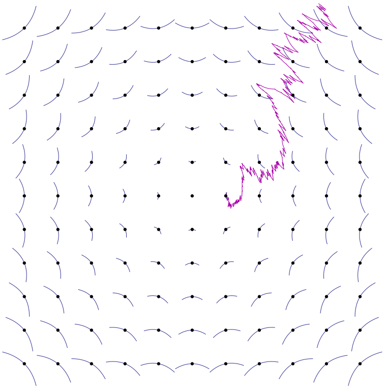


Jets and SDEs

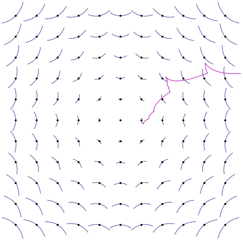
Given such a γ, a starting X_0 ($X_0 = 0$ in our example), a W_t & time step δt define discrete time stochastic process:

$$X_0 := x_0, \quad X_{t+\delta t} := \gamma^E_{X_t}(W_{t+\delta t} - W_t)$$

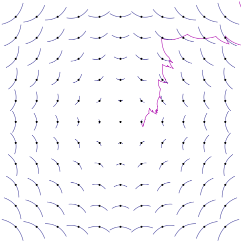
We have connected the points using the curves in $\gamma^E_{X_t}$. Notice that given W_t law we can interpret the trajectories as being randomly generated trajectories that move from X_t to $X_{t+\delta t}$ by following the curve $s \mapsto \gamma_{X_t}(s)$ from $s = 0$ to $s = \epsilon\sqrt{\delta t}$ where the ϵ's are independent std normals



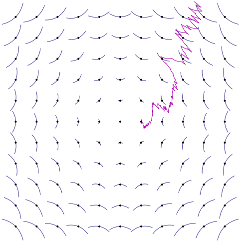
Jets and SDEs



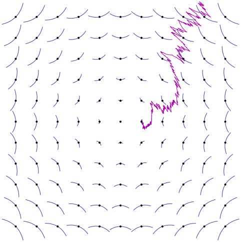
$\delta t = 0.2 \times 2^{-5}$



$\delta t = 0.2 \times 2^{-7}$



$\delta t = 0.2 \times 2^{-9}$



$\delta t = 0.2 \times 2^{-11}$

Jets and SDEs

$$X_{t+\delta t} := \gamma_{X_t}(\delta W_t) \quad \text{reads:}$$

"Follow the curve γ starting from X for a parameter increment of $\delta W_t := W_{t+\delta t} - W_t$". In this description,

Not using the $\mathbb{R}^n$ vector space structure. Coordinate Free.

These discrete time stochastic processes converge in some sense to a limit as the time step tends to zero for γ such as γ^E with sufficiently good regularity. Write the limit equation as

$$\text{Coordinate free SDE:} \quad X_t \smile\!\!\!\rightarrow \gamma_{X_t}(\mathrm{d}W_t), \quad X_0 = x_0. \tag{1}$$

Now we answer the question: How can the scheme limit be made precise and how does it relate to classic stochastic calculus?

Jets and SDEs

In a coordinate system, consider the Taylor expansion of γ_x.

$$\gamma_x(t) = x + \gamma_x'(0)t + \frac{1}{2}\gamma_x''(0)t^2 + R_x t^3, \quad R_x = \frac{1}{6}\gamma_x'''(\xi), \quad \xi \in [0,t],$$

where $R_x t^3$ is the remainder term in Lagrange form. Substituting this Taylor expansion in our scheme $X_{t+\delta t} = \gamma_{X_t}(W_{t+\delta t} - W_t)$ we obtain

$$\delta X_t = \gamma_{X_t}'(0)\delta W_t + \frac{1}{2}\gamma_{X_t}''(0)(\delta W_t)^2 + R_{X_t}(\delta W_t)^3, \quad X_0 = x_0. \tag{2}$$

Properties of Brownian motion such as "$(dW)^2 = dt$" and "$(dW)^3 = 0$" suggest we replace $(\delta W_t)^2$ with δt and $(\delta W_t)^3$ with 0. We obtain:

$$\delta \bar{X}_t = \underbrace{\gamma_{\bar{X}_t}'(0)}_{=:b(\bar{X}_t)} \delta W_t + \underbrace{\frac{1}{2}\gamma_{\bar{X}_t}''(0)}_{=:a(\bar{X}_t)} \delta t, \quad \bar{X}_0 = x_0.$$

Jets and SDEs

$$\delta \bar{X}_t = a(\bar{X}_t)\delta t + b(\bar{X}_t)\delta W_t. \tag{3}$$

This is the Euler scheme & under suitable assumptions converges in $L^2(\mathbb{P})$ to the solution to the Itô stochastic differential equation:

$$\mathrm{d}\tilde{X}_t = a(\tilde{X}_t)\,\mathrm{d}t + b(\tilde{X}_t)\mathrm{d}W_t, \quad \tilde{X}_0 = x_0. \tag{4}$$

More precisely, assume in the given coordinate system $\gamma_x(t)$ is smoothly varying in x with first & second t derivatives at 0 satisfying Lipschitz conditions in x. Assume that the third t-derivative at $t = 0$ is uniformly bounded in x. **Theorem: (Armstrong & B. 2016).** The following 3 schemes have as same $L^2(\mathbb{P})$ limit the classic Ito SDE $\tilde{X}$.

- Coordinate free γ_x scheme: $X_{t+\delta t} := \gamma_{X_t}(W_{t+\delta t} - W_t), \quad X_0$
- 2-jet scheme: $\delta \hat{X}_t = \gamma'_{\hat{X}_t}(0)\delta W_t + \frac{1}{2}\gamma''_{\hat{X}_t}(0)(\delta W_t)^2, \quad X_0$
- The classic Euler scheme: $\delta \bar{X}_t = \gamma'_{\bar{X}_t}(0)\delta W_t + \frac{1}{2}\gamma''_{\bar{X}_t}(0)\delta t, \quad X_0$

Jets and SDEs

Our earlier example $\gamma^E_{(x_1,x_2)}(t) = (x_1, x_2) + t(-x_2, x_1) + 3t^2(x_1, x_2)$

corresponds then to the SDE $d \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = 3 \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} dt + \begin{bmatrix} -X_2 \\ X_1 \end{bmatrix} dW_t$.

At this point, given that Taylor expansions occurred in a specific coordinate system, one may wonder in which sense the 2-jet scheme

$$\delta \hat{X}_t = b(\hat{X}_t)\delta W_t + a(\hat{X}_t)(\delta W_t)^2, \quad X_0$$

& its limit are **coordinate free.** It is important to note that coefficients of Itô SDE a & b only depend upon the first two derivatives of γ.
We say that two smooth curves $\gamma : \mathbb{R} \to \mathbb{R}^n$ have the same k-jet ($k \in \mathbb{N}, k > 0$) if their Taylor expansions are equal up to order $O(t^k)$ in one (and hence all) coordinate system.

Jets and SDEs

The k-jet can then be defined for example as the equivalence class of all curves that are equal up to order $O(t^k)$ in one and hence all coordinate systems, similarly to what is done to define tangent vectors, leading to a coordinate free definition.

Using this terminology, we say that the coefficients of the Ito SDE

$$\mathrm{d}\tilde{X}_t = \underbrace{a(\tilde{X}_t)}_{\frac{1}{2}\gamma''_{\tilde{X}_t}(0)} \, \mathrm{d}t + \underbrace{b(\tilde{X}_t)}_{\gamma'_{\tilde{X}_t}(0)} \mathrm{d}W_t, \quad \tilde{X}_0 = x_0$$

are determined by the 2-jet of a curve $\gamma : \mathbb{R} \to \mathbb{R}^n$ in a specific coordinate system.

Jets and SDEs

Jets are defined more generally on manifolds (eg equivalence classes of maps with the same Taylor expansion up to a given order in given – and hence all – charts).

Jets and the related notions of obstruction and symbol are powerful tools in the analysis of differential equations in a geometric setting.

In the light of the above convergence result, we can say that in our earlier pictures one should avoid interpreting any details other than the first two derivatives of the curve. One way of doing this is by insisting that we draw the quadratic curves that best fit the curves γ rather than the actual curve γ itself.

Jets and SDEs

Given a curve γ_x, we will write $j_2(\gamma_x)$ for the two jet associated with γ_x. This is formally defined to be the equivalence class of all curves which are equal to γ_x up to $O(t^2)$ included.

Given our convergence results, showing that the limit of our scheme depends only on the two-jet, we may rewrite $X_t \rightsquigarrow \gamma_{X_t}(\mathrm{d}W_t),\ X_0$ as:

Coordinate-free 2-jet SDE: $$X_t \rightsquigarrow j_2(\gamma_{X_t})(\mathrm{d}W_t), \quad X_0 = x_0. \quad (5)$$

All the theory so far extends straightforwardly to *n*-dimensional finite dimensional manifolds *M*.

Itô's formula via 2-jets

Suppose that f is a smooth mapping from $\mathbb{R}^n$ to itself and suppose that X satisfies $X_{t+\delta t} := \gamma_{X_t}(\delta W_t)$. It follows that $f(X)$ satisfies:

$$(f(X))_{t+\delta t} = (f \circ \gamma_{X_t})(\delta W_t).$$

Taking the limit as δt tends to zero we have:

Lemma (Itô's lemma — coordinate free formulation)

If the process X_t satisfies

$$X_t \smile\!\!\!\!\nearrow j_2(\gamma_{X_t})(\mathrm{d}W_t)$$

then $f(X_t)$ satisfies

$$f(X)_t \smile\!\!\!\!\nearrow j_2(f \circ \gamma_{X_t})(\mathrm{d}W_t).$$

Itô's lemma: the transformation rule for jets under a change of coordinates is given by composition of functions.

Itô's formula via 2-jets

If one prefers the more traditional coordinates SDE given in (4) we simply need to calculate the derivatives of $f \circ \gamma$. In a chosen coordinate system, let us write γ_X^i for the i-th component of γ_X with respect to the coordinates $x_1, x_2, \dots x_n$ for $\mathbb{R}^n$. Two applications of the chain rule give us:

$$(f \circ \gamma_X)'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\gamma_X(t)) \frac{\mathrm{d}\gamma_X}{\mathrm{d}t}$$

$$(f \circ \gamma_X)''(t) = \sum_{j=1}^{n} \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\gamma_X(t)) \frac{\mathrm{d}\gamma_X^i}{\mathrm{d}t} \frac{\mathrm{d}\gamma_X^j}{\mathrm{d}t} + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\gamma_X(t)) \frac{\mathrm{d}^2\gamma_X}{\mathrm{d}t^2}$$

We conclude that our lemma is equivalent to the classical Itô's lemma.

Itô's formula via 2-jets

We have illustrated a way of drawing an SDE on a rubber sheet such that if sheet is stretched, diagram transforms as per Itô's lemma.

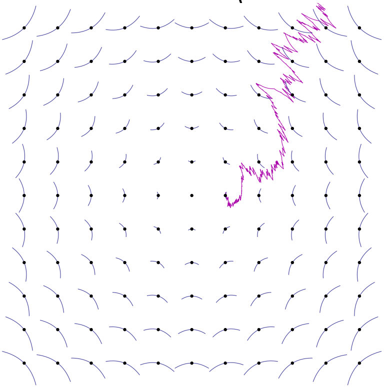
In other words given an SDE in $\mathbb{R}^n$ we have given a method of drawing SDEs such that for all $f : \mathbb{R}^n \to \mathbb{R}^n$ the following diagram commutes:

$$\begin{tikzcd} \text{SDE for X} \arrow[r, "\text{Itô's lemma}"] \arrow[d, "\text{Draw}"'] & \text{SDE for } f(X) \arrow[d, "\text{Draw}"'] \\ \text{Picture of SDE for X in } \mathbb{R}^n \arrow[r, "f"] & \begin{array}{l} f(\text{Picture of SDE for X}) \\ = \text{Picture of SDE for } f\,(\text{X}) \end{array} \end{tikzcd}$$

Itô's formula via 2-jets

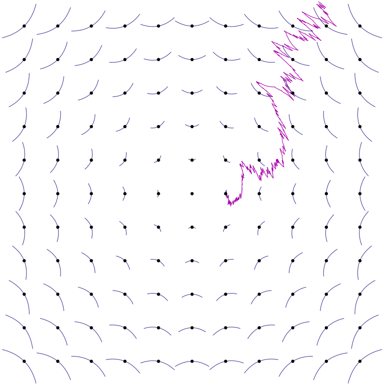
Since we now understand the geometric content of Itô's lemma, we can draw a picture to illustrate it. Consider the transformation

$$(\theta, s) = \phi(x_1, x_2) = \left(\arctan(x_2/x_1), \log(\sqrt{x_1^2 + x_2^2}) \right)$$



applied to our example process γ^E. This can be viewed as a transformation of the complex plane $\phi(z) = i \log(z)$. We use ϕ to transform our earlier picture on the left hand side in 2 ways. 1st we apply directly ϕ to each point of the left picture to obtain a new point for a new figure. This is done with *image manipulation software*: we stretch the image & don't use its mathematical structure.

Itô's formula via 2-jets



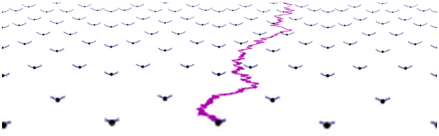
As an alternative approach, we transform our equation using Itô's lemma applied to the function ϕ. This results in the equation for θ and s:

$$d(\theta, s) = \left(0, \frac{7}{2}\right) dt + (1, 0)\, dW_t.$$

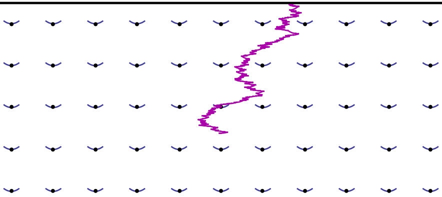
We can then use this equation to plot the process (θ, s) directly by simulating the process in discrete time as before.

The results of these two alternative approaches are now shown in the upper and lower parts of Figure 2:

Itô's formula via 2-jets



The process $j_2(\phi \circ \gamma^E)$ plotted using image manipulation software



The process $j_2(\phi \circ \gamma^E)$ plotted by applying Itô's lemma

Figure: Two plots of the process $j_2(\phi \circ \gamma^E)$ in the plane (θ, s).

Itô's formula via 2-jets

As one can see the two approaches to plotting the transformed process give essentially identical results, showing an example of our earlier commutative diagram at work.

The differences one can still see are:

- the lower quality in the upper image, obtained by transforming pixels rather than using vector graphics;
- the grid points at which the 2-jets are plotted are changed;
- small differences in the simulated path since we have only simulated discrete time paths.

SDEs driven by vector Brownians

Consider functions $\gamma_x : \mathbb{R}^d \to \mathbb{R}^n$ & as before the coordinate free

$$X_{t+\delta t} := \gamma_{X_t}\left(\delta W_t^1, \ldots, \delta W_t^d\right)$$

Again, the limiting behaviour will only depend upon the 2-jet $j_2(\gamma_x)$ and can still be denoted by $X_t \rightharpoonup j_2(\gamma_{X_t})(dW_t)$. The scheme still $\xrightarrow{L^2(\mathbb{P})}$ to the classical Itô SDE (see proof in A & B [4]) in coordinates:

$$\tilde{X}_t = \tilde{X}_0 + \int_0^t a(\tilde{X}_s)\,\mathrm{d}s + \sum_{\alpha=1}^{d} \int_0^t b_\alpha(\tilde{X}_s)\,\mathrm{d}W_s^\alpha, \quad t \in [0, T]$$

$$a(x) := \frac{1}{2}\sum_{\alpha=1}^{d} \frac{\partial^2 \gamma_x}{\partial u^\alpha \partial u^\alpha}\bigg|_{u=0}, \quad b_\alpha(x) := \frac{\partial \gamma_x}{\partial u^\alpha}\bigg|_{u=0}.$$

SDEs driven by vector Brownians

We can also write the SDE as

$$\mathrm{d}X_t^i = \frac{1}{2}\partial_\alpha\partial_\beta\gamma^i \mathrm{d}W_t^\alpha \mathrm{d}W_t^\beta + \partial_\alpha\gamma^i\, \mathrm{d}W_t^\alpha = \frac{1}{2}\partial_\alpha\partial_\beta\gamma^i g_E^{\alpha\beta} dt + \partial_\alpha\gamma^i\, \mathrm{d}W_t^\alpha \quad (6)$$

with the convention that $\mathrm{d}W_t^\alpha \mathrm{d}W_t^\beta = g_E^{\alpha\beta}\mathrm{d}t$ where g_E is a Kronecker delta with orthonormal coordinates, or in generalizations the symmetric 2-form defining the Euclidean metric on $\mathbb{R}^d$.

2-jet based definition of the SDE **backward diffusion operator**:

$$\mathcal{L}_{\gamma_x} f := \frac{1}{2}\Delta_E(f \circ \gamma_x) = \frac{1}{2}\partial_\alpha\partial_\beta(f \circ \gamma_x) g_E^{\alpha\beta}. \quad (7)$$

Here Δ_E is the Laplacian defined on $\mathbb{R}^d$. $\mathcal{L}_{\gamma_x}$ acts on functions defined on the state space manifold M. We define $\mathcal{L}^*$ to be its formal adjoint which acts on densities defined on M (Fokker Planck eq).

Weak and Strong Equivalence of SDEs through jets

Both the Itô SDE (6) & the backward diffusion operator use only part of the 2-jet: only the diagonal terms of $\partial_\alpha \partial_\beta \gamma^i$ influence the SDE and even for these terms it is only their sum that is important.

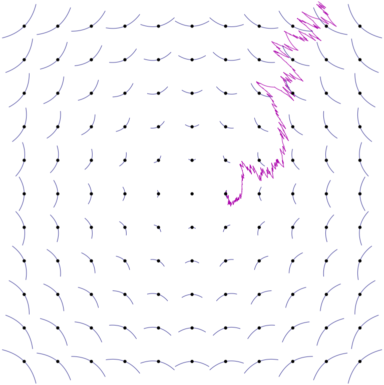
We say that two 2-jets γ_x^1 and γ_x^2 are *weakly equivalent* if $\mathcal{L}_{\gamma_x^1} = \mathcal{L}_{\gamma_x^2}$.

γ^1 and γ^2 are *strongly equivalent* if in addition $j_1(\gamma^1) = j_1(\gamma^2)$.

Strong equivalence means that given the same realization of the driving Brownian motions W_t^α the solutions of the SDEs will be almost surely the same (under assumptions ensuring pathwise uniqueness).

Weak equivalence means that the transition probability distributions are the same even though the dynamics may be different for any specific realisation of the Brownian motions.

Drawing SDEs driven by 2-dimensional Brownians



We saw previously a way to draw a SDE in $\mathbb{R}^2$, $j_2(\gamma^E)$, driven by one-dimensional Brownian motion:

$$d\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = 3\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} dt + \begin{bmatrix} -X_2 \\ X_1 \end{bmatrix} dW_t.$$

How can we draw a SDE driven by 2-dimensional Brownian motion?

Given an SDE in local coordinates $\mathrm{d}X_t = a(X_t)\mathrm{d}t + b_i(X_t)\mathrm{d}W_t^i$ (Einstein summation) with $a \in \mathbb{R}^2$ and $b_1 \in \mathbb{R}^2$, $b_2 \in R^2$, we can write down a specific representative two jet by

$$\gamma_x(t^1, t^2) = x + a g_{ij}^E t^i t^j + b_i t^i.$$

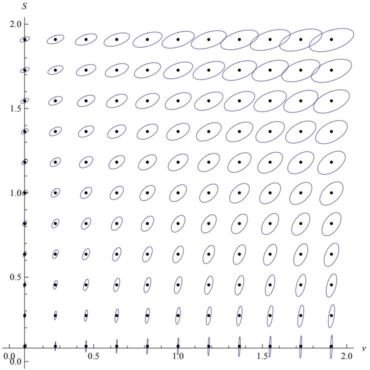
Drawing SDEs driven by 2-dimensional Brownians

$$\gamma_x(t^1, t^2) = x + a g^E_{ij} t^i t^j + b_i t^i.$$

The image of an ϵ ball under γ_x will be an ellipsoid. Moreover, if we know that γ_x is of this form, we can recover the coefficients a and b_i up to weak equivalence just from knowledge of the image of the ϵ ball.

This method of drawing an $\mathbb{R}^2$ SDE driven by 2-dim Brownian motion in local coordinates is to draw the image of an ϵ ball of (t^1, t^2) at each point.

Drawing SDEs driven by 2-dimensional Brownians



For example in this figure we show a plot of the Heston stochastic volatility model with drift (see [15]). Note that as well as plotting the ellipses, the figure indicates the exact point that each ellipse is associated with. The extent to which the centre of the ellipse differs from the associated point is a measure of the drift.

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{\nu_t} S_t dW_t^1 \\ d\nu_t &= \kappa(\theta - \nu_t)dt + \xi\sqrt{\nu_t}\left(\rho dW_t^1 + \sqrt{1-\rho^2}dW_t^2\right) \end{aligned} \tag{8}$$

Parameter values $\xi = 1$, $\theta = 0.4$, $\kappa = 1$, $\mu = 0.1$, $\rho = 0.5$. We have plotted the image of the balls for $\epsilon = 0.05$.

The one-dimensional case: Fan Diagrams

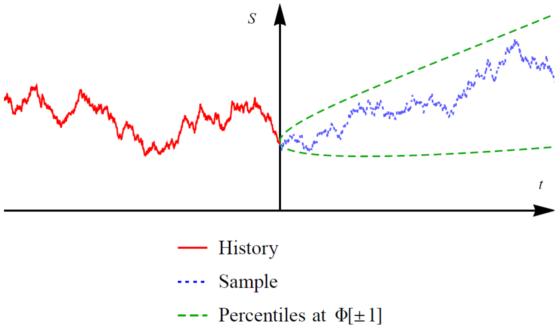
Standard statistical properties of a distribution depend upon the coordinate system.

For example $\mathbb{E}$ of a process in $\mathbb{R}^n$ involves the vector space structure of $\mathbb{R}^n$. If f is a nonlinear coordinate transition map, one has $\mathbb{E}(f(X)) \neq f(\mathbb{E}(X))$.

However, the definition of the α-percentile depends only upon the ordering of $\mathbb{R}$ and not its vector space structure.

As a result, for continuous monotonic f and X with connected state space, the median of $f(X)$ is equal to f applied to the median of X. If f is strictly increasing, the analogous result holds for the α percentile.

The one-dimensional case: Fan Diagrams



This has the implication that the trajectory of the α-percentile of an $\mathbb{R}$ valued stochastic process is invariant under smooth monotonic coordinate changes of $\mathbb{R}$. **In other words, percentiles have a coordinate free interpretation.** How can the trajectories of percentiles be related to the coefficients of the SDE?

The one-dimensional case: Fan Diagrams

Theorem (Armstrong and B.) [4]. For sufficiently small t, the α-th percentile of the solutions to

$$\mathrm{d}X_t = a(X_t, t)\,\mathrm{d}t + b(X_t, t)\mathrm{d}W_t, \qquad X_0 = x_0 \tag{9}$$

is given by: $$x_0 + b_0\sqrt{t}\Phi^{-1}(\alpha) + \left[a_0 - \frac{b_0 b_0'}{2}(1 - \Phi^{-1}(\alpha)^2)\right] t + O(t^{3/2})$$

so long as the coefficients of (9) are smooth, the diffusion coefficient b never vanishes, and sufficient conditions for the Lamperti transformed SDE and for $\mathcal{L}^* p = 0$ to have a unique regular solution hold. In this formula a_0 and b_0 denote the values of $a(x_0, 0)$ and $b(x_0, 0)$ respectively. In particular, **the median process is a straight line up to $O(t^{\frac{3}{2}})$ with tangent given by the drift of the Stratonovich version of the Itô SDE** (9)**.The $\Phi(1)$ and $\Phi(-1)$ percentiles correspond up to $O(t^{\frac{3}{2}})$ to the curves $\gamma_{X_0}(\pm\sqrt{t})$ where γ_{X_0} is any representative of the 2-jet that defines the SDE in Itô form.**

Jets & vector fields: Ito / Str as different coordinates

We have seen that, geometrically, a Str SDE is described by 2 vector fields, while a Ito SDE is described by one 2-jet. We now relate the two.

An alternative way to specify the k-jet of a curve at every point is to choose k vector fields $A_1, \ldots, A_k$ on the manifold. One can then define $\Phi^t_{A_i}$ to be the vector flow associated with the vector field A_i. This allows one to define curves at each point x as follows:

$$\gamma_x(t) = \Phi^{t^k}_{A_k}(\Phi^{t^{k-1}}_{A_{k-1}}(\ldots(\Phi^t_{A_1}(x))\ldots)) \tag{10}$$

where t^k denotes the k-th power of t. We will call this the *vector representation* for a family of k-jets.
Theorem (Armstrong B. (2016)). All k-jets of curves can be represented this way via vector fields flows.

Jets & vector fields: Ito / Str as different coordinates

Corollary (Ito Stratonovich transformation as correspondence between 2-jets and two vector fields.)

Suppose that a family of 2-jets of curves is given in the vector representation as

$$\gamma_x(t) = \Phi_A^{t^2}(\Phi_B^t(x))$$

for vector fields A and B. Choose a coordinate chart and let A^i, B^i be the components of the vector fields in this chart. Then the corresponding standard representation for the family of 2-jets is:

$$\gamma_x(t) = x + a(x)t^2 + b(x)t$$

with

$$a^i = A^i + \frac{1}{2}\frac{\partial B^i}{\partial x^j}B^j, \quad b^i = B^i.$$

Jets & vector fields: Ito / Str as different coordinates

Geometric interpretation of the Ito-Stratonovich transformation: switching between 2-jets and pairs of vector fields.

Despite Itô's 1950 paper [16] on SDEs on manifolds based on using Itô's lemma to change coordinates, a few authors have even asserted that stochastic differential geometry *requires* Stratonovich calculus.

From an extrinsic perspective (i.e. manifolds embedded in $\mathbb{R}^n$ instead of charts) Stratonovich may appear necessary since an SDE remains on a submanifold a.s. if Str-drift and Str-diffusion vector fields are tangent to the manifold.

It is easy to write down the Stratonovich SDE induced on a submanifold from a Str SDE on $\mathbb{R}^n$. However, this is simply a consequence of the curvature of the 2-jet following the curvature of the manifold, so the Itô/2-jet interpretation works as well.

From our point of view we consider these two calculi as different coordinate systems for the same underlying coordinate-free SDE.

Many notions in probability are not coordinate free however (the expected value $\mathbb{E}$ for example, but see also our earlier discussion on the assumed density principle).

Conclusions

Stratonovich SDE on manifold= 2 vector fields;   Itô SDE = one 2-jet.

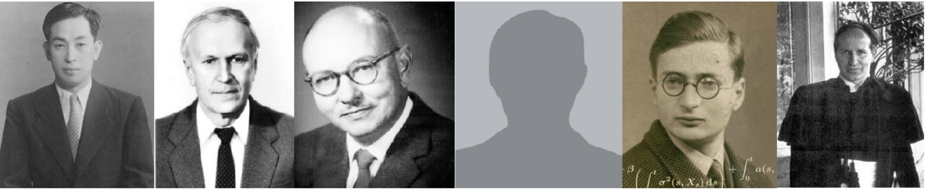
New geometric jets-based interpretation of SDE coefficients, diffusion operator, Itô-Str transformation, percentiles, & other SDE concepts.

Itô & Stratonovich can be seen as two different coordinate systems. One should choose the most convenient system for the problem at hand along the properties we highlighted above (Wong Zakai convergence, martingale, anticipative features, etc).

The most important difference between Stratonovich & Itô arises during the modelling process. It is when choosing what equation to write down in the first place that the choice is most telling.

The modelling process is not strictly mathematical: it relies upon modellers intuition. "*The ultimate goal of mathematics is to eliminate all need for intelligent thought*" ([20]) does not seem to apply

Conclusions



KYIOSI ITO | RUSLAN STRATONOVICH | EDWARD McSHANE | DONALD FISK | WOLFGANG DOEBLIN | CHARLES EHRESMANN

With thanks to the organizing committee.

Thank you for your attention.

Questions and comments welcome

References I

[1] J. Armstrong and D. Brigo.
Dimensionality reduction for SDEs with differential geometric projection methods.
Working paper, forthcoming on arxiv.org.

[2] John Armstrong and Damiano Brigo.
Extrinsic projection of Itô SDEs on submanifolds with applications to non-linear filtering.
To appear in: Nielsen, F., Critchley, F., & Dodson, K. (Eds), Computational Information Geometry for Image and Signal Processing, Springer Verlag, 2016.

References II

[3] John Armstrong and Damiano Brigo.
Nonlinear filtering via stochastic PDE projection on mixture manifolds in L^2 direct metric.
Mathematics of Control, Signals and Systems, 28(1):1–33, 2016.

[4] John Armstrong and Damiano Brigo.
Coordinate Free Stochastic Differential Equations as Jets.
`http://arxiv.org/abs/1602.03931`

[5] D Brigo, B Hanzon, and F LeGland.
A differential geometric approach to nonlinear filtering: The projection filter.
IEEE Transactions on automatic control, 43:247–252, 1998.

References III

[6] D. Brigo, B. Hanzon, and F. LeGland.
Approximate nonlinear filtering by projection on exponential manifolds of densities.
Bernoulli, 5(3):495–534, 06 1999.

[7] Ehresmann, C.
Introduction à la théorie des structures infinitésimales et des pseudo-groupes de Lie. Geometrie Differentielle, Colloq. Inter. du Centre Nat. de la Recherche Scientifique, Strasbourg, 1953, 97-127.

[8] K. D. Elworthy.
Stochastic differential equations on manifolds.
Cambridge University Press, Cambridge, New York, 1982.

References IV

[9] K. D. Elworthy.
Geometric aspects of diffusions on manifolds.
In *École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87*, pages 277–425. Springer, 1988.

[10] M. Emery.
Stochastic calculus in manifolds.
Springer-Verlag, Heidelberg, 1989.

[11] D. Fisk.
Quasi-martingales and stochastic integrals.
PhD Dissertation, Michigan State University, Department of Statistics, 1963.

References V

[12] Avner Friedman.
Partial differential equations of parabolic type.
Prentice-Hall, reprinted in 2008 by Dover-Publications, Englewood Cliffs N.J, 1964.

[13] Avner Friedman.
Stochastic differential equations and applications, vol. I and II.
Academic Press, reprinted by Dover Publications, New York, 1975.

[14] I. Gyongy.
A note on Eulers approximations.
Potential Analysis, (8):205–216, 1998.

References VI

[15] Steven L Heston.
A closed-form solution for options with stochastic volatility with applications to bond and currency options.
Review of financial studies, 6(2):327–343, 1993.

[16] K. Itô.
Stochastic differential equations in a differentiable manifold.
Nagoya Math. J., (1):35–47, 1950.

[17] P. E. Kloeden and E. Platen.
Higher-order implicit strong numerical schemes for stochastic differential equations.
Journal of statistical physics, 66(1-2):283–314, 1992.

References VII

[18] Peter E. Kloeden and Eckhard Platen.
Numerical solution of stochastic differential equations.
Applications of mathematics. Springer, Berlin, New York, Third printing, 1999.

[19] S. Kobayashi and K. Nomizu.
Foundations of differential geometry, volume 1.
New York, 1963.

[20] Graham, R.L., Knuth, D.E., and O. Patashnik (1994).
Concrete Mathematics, 2nd Ed., ADDISON-WESLEY.

[21] E.J. McShane.
Stochastic Calculus and Stochastic Models (2nd Ed.).
Academic Press, New York, 1974.

References VIII

[22] G. A. Pavliotis.
Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations.
Springer, Heidelberg, 2014.

[23] K. Pearson.
Contributions to the mathematical theory of evolution. ii. skew variation in homogeneous material.
Philosophical Transactions of the Royal Society of London. A, 186:343–414, 1895.

[24] Protter, P. (2004).
Stochastic Integration and Differential Equations. Springer Verlag.

References IX

[25] L. C. G. Rogers and D. Williams.
Diffusions, markov processes and martingales, vol 2: Ito calculus, 1987.

[26] A. V. Skorokhod (1967).
R. L. Stratonovich "Conditional Markov processes and their application to the theory of optimal control" (book review), Teor. Veroyatnost. i Primenen., 12:1 (1967), 184187.

[27] R. L. Stratonovich.
A new representation for stochastic integrals and equations.
SIAM Journal on Control, 4(2):362–371, 1966.

[28] N. G. Van Kampen.
Itô vs Stratonovich.
Journal of Statistical Physics, 24(1), 1981.