CBMS Lecture Series

Recent Advances in the Numerical Approximation of Stochastic Partial Differential Equations

or more accurately

Taylor Approximations of Stochastic Partial Differential Equations

A. Jentzen

Department of Mathematics
Universität Bielefeld
Bielefeld, Germany
Email: jentzen@math.uni-bielefeld.de

P.E. Kloeden

Institut für Mathematik
Goethe Universität
Frankfurt am Main, Germany
Email: kloeden@math.uni-frankfurt.de

10 Lectures on Taylor Expansions!!

Taylor expansions are a very basic tool in numerical analysis and other areas of mathematics that require approximations.

They enable the derivation of one–step numerical schemes for differential equations of arbitrary high order.

1. Ordinary Differential Equations

2. Random Ordinary Differential Equations

3. Stochastic Ordinary Differential Equations

4. Stochastic Ordinary Differential Equations: Nonstandard Assumptions

5. Stochastic Partial Differential Equations: Introduction

6. Stochastic Partial Differential Equations: Numerical Methods

7. & 8. Stochastic Partial Differential Equations: Additive Noise I & II

9. & 10. Stochastic Partial Differential Equations: Multiplicative Noise I & II

Lecture 1: Ordinary Differential Equations

The Taylor expansion of a $p+1$ times continuously differentiable function $x : \mathbb{R} \to \mathbb{R}$ is given by

$$x(t) = x(t_0) + x'(t_0)\ h + \ldots + \frac{1}{p!}x^{(p)}(t_0)\ h^p + \frac{1}{(p+1)!}x^{(p+1)}(\theta)\ h^{p+1} \quad (1)$$

with the remainder term evaluated at some intermediate value $\theta \in [t_0, t]$, which is usually unknown. Here $h = t - t_0$.

1 Taylor Expansions for ODEs

Let $x(t) = x(t, t_0, x_0)$ be the solution of a scalar ODE

$$\frac{dx}{dt} = f(t, x), \quad (2)$$

with the initial value $x(t_0) = x_0$ and define the differential operator L by

$$Lg(t, x) := \frac{\partial g}{\partial t}(t, x) + f(t, x)\frac{\partial g}{\partial x}(t, x),$$

i.e., $Lg(t, x(t))$ is the total derivative of $g(t, x(t))$ with respect to a solution $x(t)$ of the ODE (2), since

$$\frac{d}{dt}g(t, x(t)) = \frac{\partial g}{\partial t}(t, x(t)) + \frac{\partial g}{\partial x}(t, x(t))\ x'(t) = Lg(t, x(t))$$

by the chain rule.

In particular, for any such solution

$$\begin{aligned} x'(t) &= f(t, x(t)) \\ x''(t) &= \frac{d}{dt}x'(t) = \frac{d}{dt}f(t, x(t)) = Lf(t, x(t)) \\ x'''(t) &= \frac{d}{dt}x''(t) = \frac{d}{dt}Lf(t, x(t)) = L^2 f(t, x(t)) \ , \end{aligned}$$

and, in general,

$$\boxed{x^{(j)}(t) = L^{j-1} f(t, x(t)), \qquad j = 1, 2, \dots \ ,}$$

provided f is smooth enough.

For notational convenience, define $L^0 f(t, x) \equiv f(t, x)$.

If f is p times continuously differentiable, then the solution $x(t)$ of the ODE (2) is $p+1$ times continuously differentiable and has a Taylor expansion (1), which can be rewritten as

$$\begin{aligned} x(t) &= x(t_0) + \sum_{j=1}^{p} \frac{1}{j!} L^{j-1} f(t_0, x(t_0)) \ (t - t_0)^j \\ &\qquad + \frac{1}{(p+1)!} L^p f(\theta, x(\theta)) \ (t - t_0)^{p+1}. \end{aligned}$$

On a subinterval $[t_n, t_{n+1}]$ with $h = t_{n+1} - t_n > 0$, the Taylor expansion is

$$\boxed{x(t_{n+1}) = x(t_n) + \sum_{j=1}^{p} \frac{1}{j!} L^{j-1} f(t_n, x(t_n)) \ h^j + \frac{1}{(p+1)!} L^p f(\theta_n, x(\theta_n)) \ h^{p+1} \qquad (3)}$$

for some $\theta_n \in [t_n, t_{n+1}]$, which is usually unknown.

Nevertheless, the error term can be estimated and is of order $0(h^{p+1})$, since

$$\frac{h^{p+1}}{(p+1)!}\,|L^p f(\theta_n, x(\theta_n; t_n, x_n))| \le \frac{h^{p+1}}{(p+1)!} \max_{\substack{t_0 \le t \le T \\ x \in D}} |L^p f(t,x)| \le C_{p,T,D}\, h^{p+1},$$

where D is some sufficiently large compact subset of $\mathbb{R}$ which contains the solution over a bounded time interval $[t_0, T]$ containing the subintervals under consideration.

The maximum can be used here since $L^p f$ is continuous on $[t_0, T] \times D$.

Note that $L^p f$ contains the partial derivatives of f of up to order p.

2 Taylor schemes for ODEs

The Taylor scheme of order p for the ODE (2),

$$x_{n+1} = x_n + \sum_{j=1}^{p} \frac{h^j}{j!}\, L^{j-1} f(t_n, x_n), \qquad (4)$$

is obtained by discarding the remainder term in the Taylor expansion (3) and replacing $x(t_n)$ by x_n.

The Taylor scheme (4) is an example of one–step explicit scheme which has the general form

$$x_{n+1} = x_n + hF(h, t_n, x_n) \qquad (5)$$

with an increment function F defined by

$$F(h,t,x) := \sum_{j=1}^{p} \frac{1}{j!}\, L^{j-1} f(t,x)\; h^{j-1}.$$

A one–step explicit scheme is said to have order p if its global discretisation error

$$G_n(h) := |x(t_n, t_0, x_0) - x_n|, \qquad n = 0, 1, \ldots, N_h := \frac{T - t_0}{h},$$

converges with order p, i.e., if

$$\max_{0 \le n \le N_h} G_n(h) \le C_{p,T,D}\, h^p.$$

A basic result in numerical analyis says that a one–step explicit scheme converges with order p if its local discretisation error converges with order $p + 1$. This is defined by

$$L_{n+1}(h) := |x(t_{n+1}) - x(t_n) - hF(h, t_n, x(t_n))|,$$

i.e., the error on each subinterval taking one interation of the scheme starting at the exact value of the solution $x(t_n)$ at time t_n.

- Thus, the Taylor scheme of order p is indeed a pth order scheme.

The simplest nontrivial Taylor scheme is the Euler scheme

$$x_{n+1} = x_n + hf(t_n, x_n),$$

which has order $p = 1$.

The higher coefficients $L^{j-1}f(t,x)$ of a Taylor scheme of order $p > 1$ are, however, very complicated.

For example,

$$\begin{aligned} L^2 f &= L[Lf] = \frac{\partial}{\partial t}[Lf] + f\frac{\partial}{\partial x}[Lf] \\ &= \frac{\partial}{\partial t}\left\{\frac{\partial f}{\partial t} + f\frac{\partial f}{\partial x}\right\} + f\frac{\partial}{\partial x}\left\{\frac{\partial f}{\partial t} + f\frac{\partial f}{\partial x}\right\} \\ &= \frac{\partial^2 f}{\partial t^2} + \frac{\partial f}{\partial t}\,\frac{\partial f}{\partial x} + f\frac{\partial^2 f}{\partial t \partial x} + f\frac{\partial^2 f}{\partial x \partial t} + f\left(\frac{\partial f}{\partial x}\right)^2 + f^2\frac{\partial^2 f}{\partial x^2} \end{aligned}$$

and this is just the scalar case!!

Taylor schemes are thus rarely used in practice, but they are very useful for theoretical purposes,

e.g., for determining by comparison the local discretization order of other numerical schemes derived by heuristic means such as the *Heun scheme*

$$x_{n+1} = x_n + \frac{1}{2}h\left[f(t_n, x_n) + f(t_{n+1}, x_n + hf(t_n, x_n))\right],$$

which is a Runge-Kutta scheme of order 2.

Symbolic manipulators now greatly facilitate the use of Taylor schemes. Indeed,

B. Coomes, H. Koçak and K. Palmer, *Rigorous computational shadowing of orbits of ordinary differential equations*, Numerische Mathematik 69 (1995), no. 4, 401–421.

applied a Taylor scheme of order 31 to the 3-dimensional Lorenz equations.

3 Integral representation of Taylor expansion

Taylor expansions of a solution $x(t) = x(t, t_0, x_0)$ of an ODE (2) also have an integral derivation and representation.

These are based on the integral equation representation of the initial value problem of the ODE,

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))\, ds. \tag{6}$$

By the Fundamental Theorem of Calculus, the integral form of the total derivative is

$$\boxed{g(t, x(t)) = g(t_0, x_0) + \int_{t_0}^{t} Lg(s, x(s))\, ds.} \tag{7}$$

Note that (7) reduces to the integral equation (6) with $g(t, x) = x$, since $Lg = f$ in this case.

Applying (7) with $g = f$ over the interval $[t_0, s]$ to the integrand of the integral equation (6) gives

$$\begin{aligned} x(t) &= x_0 + \int_{t_0}^{t} \left[f(t_0, x_0) + + \int_{t_0}^{s} Lf(\tau, x(\tau))\, d\tau \right] ds \\ &= x_0 + f(t_0, x_0) \int_{t_0}^{t} ds + \int_{t_0}^{t} \int_{t_0}^{s} Lf(\tau, x(\tau))\, d\tau\, ds, \end{aligned}$$

which is the first order Taylor expansion.

Then, applying (7) with $g = Lf$ over the interval $[t_0, \tau]$ to the integrand in the double integral remainder term leads to

$$\begin{aligned} x(t) &= x_0 + f(t_0, x_0) \int_{t_0}^{t} ds + Lf(t_0, x_0) \int_{t_0}^{t} \int_{t_0}^{s} d\tau \, ds \\ &\quad + \int_{t_0}^{t} \int_{t_0}^{s} \int_{t_0}^{\tau} L^2 f(\rho, x(\rho)) \, d\rho \, d\tau \, ds. \end{aligned}$$

In this way one obtains the Taylor expansion in integral form

$$\begin{aligned} x(t) &= x(t_0) + \sum_{j=1}^{p} L^{j-1} f(t_0, x_0) \int_{t_0}^{t} \int_{t_0}^{s_1} \cdots \int_{t_0}^{s_{j-1}} ds_j \cdots ds_1 \\ &\quad + \int_{t_0}^{t} \int_{t_0}^{s_1} \cdots \int_{t_0}^{s_j} L^p f(s_{j+1}, x(s_{j+1})) \, ds_{j+1} \cdots ds_1, \end{aligned}$$

where for $j = 1$ there is just a single integral over $t_0 \leq s_1 \leq t$.

This is equivalent to the differential form of the Taylor expansion (3) by the Intemediate Value Theorem for Integrals and the fact that

$$\int_{t_0}^{t} \int_{t_0}^{s_1} \cdots \int_{t_0}^{s_{j-1}} ds_j \cdots ds_1 = \frac{1}{j!} (t - t_0)^j, \quad j = 1, 2, \ldots.$$

Lecture 2: Random Ordinary Differential Equations

Random ordinary differential equations (RODEs) are pathwise ODEs that contain a stochastic process in their vector field functions.

Typically, the driving stochastic process has at most Hölder continuous sample paths, so the sample paths of the solutions are certainly continuously differentiable. However, the derivatives of the solution sample paths are at most Hölder continuous in time.

Thus, after insertion of the driving stochastic process, the resulting vector field is at most Hölder continuous in time, no matter how smooth the vector field is in its original variables.

Consequently, although classical numerical schemes for ODEs can be used pathwise for RODEs, they rarely attain their traditional order and new forms of higher order schemes are required.

Let

- $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space
- $(\zeta_t)_{t\in[0,T]}$ be an $\mathbb{R}^m$-valued stochastic process with continuous sample paths
- $f : \mathbb{R}^m \times \mathbb{R}^d \to \mathbb{R}^d$ be a continuous function.

A *random ordinary differential equation* in $\mathbb{R}^d$,

$$\boxed{\frac{dx}{dt} = f(\zeta_t(\omega), x), \qquad x \in \mathbb{R}^d, \qquad (1)}$$

is a nonautonomous ordinary differential equation (ODE)

$$\boxed{\frac{dx}{dt} = F_\omega(t, x) := f(\zeta_t(\omega), x) \qquad (2)}$$

for almost every $\omega \in \Omega$.

A simple example of a scalar RODE is

$$\boxed{\frac{dx}{dt} = -x + \sin W_t(\omega),}$$

where W_t is a scalar Wiener process.

Here $f(z, x) = -x + \sin z$ and $d = m = 1$.

- Other kinds of noise such as *fractional Brownian motion* can also be used.

It will be assumed that f is infinitely often continuously differentiable in its variables, although k-times continuously differentiable with k sufficiently large would suffice.

Then f is locally Lipschitz in x and the initial value problem

$$\boxed{\frac{d}{dt}x_t(\omega) = f(\zeta_t(\omega), x_t(\omega)), \qquad x_0(\omega) = X_0(\omega), \quad (3)}$$

where the initial value X_0 is a $\mathbb{R}^d$-valued random variable, has a unique pathwise solution $x_t(\omega)$ for every $\omega \in \Omega$, which will be assumed to exist on the finite time interval $[0, T]$ under consideration.

• Sufficient conditions guaranteeing the existence and uniqueness of solutions of (3) are similar to those for ODEs.

The solution of the RODE (3) is a stochastic process $(x_t)_{t\in[0,T]}$, which is nonanticipative if the driving process ζ_t is nonanticipative and independent of the initial condition X_0.

Important: The sample paths $t \to x_t(\omega)$ of a RODE are continuously differentiable. They need not be further differentiable, since the vector field $F_\omega(t, x)$ of the nonautonomous ODE (2) is usually only continuous, but not differentiable in t, no matter how smooth the function f is in its variables.

1 Equivalence of RODEs and SODEs

RODEs with Wiener processes can be rewritten as SODEs, so results for one can be applied to the other. For example, the scalar RODE

$$\boxed{\frac{dx}{dt} = -x + \sin W_t(\omega)}$$

can be rewritten as the 2-dimensional SODE

$$\boxed{d\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} -X_t + \sin Y_t \\ 0 \end{pmatrix} dt + \begin{pmatrix} 0 \\ 1 \end{pmatrix} dW_t.}$$

On the other hand, any finite dimensional SODE can be transformed to a RODE. This is the famous Doss–Sussmann result.

It is easily illustrated for a scalar SODE with additive noise: the SODE

$$\boxed{dX_t = f(X_t)\, dt + dW_t}$$

is equivalent to the RODE

$$\boxed{\frac{dz}{dt} = f(z + O_t) + O_t, \qquad (4)}$$

where $z(t) := X_t - O_t$ and O_t is the Ornstein–Uhlenbeck stochastic stationary process satisfying the linear SDE

$$\boxed{dO_t = -O_t\, dt + dW_t.}$$

To see this, subtract integral versions of both SODEs and substitute to obtain

$$z(t) = z(0) + \int_0^t [f(z(s) + O_s) + O_s]\ ds.$$

It follows by continuity and the Fundamental Theorem of Calculus that z is pathwise differentiable. In particular, deterministic calculus can be used pathwise for SODEs via RODEs.

This greatly facilitates the investigation of dynamical behaviour and other qualitative properties of SODEs. For example, suppose that f in the RODE (4) satisfies a one–sided dissipative Lipschitz condition ($L > 0$),

$$\langle x - y, f(x) - f(y) \rangle \leq -L|x - y|^2, \qquad \forall x, y \in \mathbb{R}.$$

Then, for any two solutions $z_1(t)$ and $z_2(t)$ of the RODE (4),

$$\begin{aligned} \frac{d}{dt}|z_1(t) - z_2(t)|^2 &= 2\left\langle z_1(t) - z_2(t), \frac{dz_1}{dt} - \frac{dz_2}{dt} \right\rangle \\ &= 2\left\langle z_1(t) - z_2(t), f(z_1(t) + O_t) - f(z_2(t) + O_t) \right\rangle \\ &\leq -L\,|z_1(t) - z_2(t)|^2, \end{aligned}$$

from which it follows that

$$|z_1(t) - z_2(t)|^2 \leq e^{-2Lt}|z_1(0) - z_2(0)|^2 \to 0 \quad \text{as } t \to \infty \qquad \text{(pathwise)}.$$

By the theory of random dynamical systems, there thus exists a pathwise asymptotically stable stochastic stationary solution. Transforming back to the SODE, one concludes that the SODE also has a pathwise asymptotically stable stochastic stationary solution.

2 Simple Numerical schemes for RODEs

The rules of deterministic calculus apply pathwise to RODEs, but the vector field function in $F_\omega(t,x)$ in (2) is not smooth in t.
It is at most Hölder continuous in time like the driving stochastic process ζ_t and thus lacks the smoothness needed to justify the Taylor expansions and the error analysis of traditional numerical methods for ODEs.
Such methods can be used, but will attain at best a low convergence order, so new higher order numerical schemes must be derived for RODEs.

For example, let $t \to \zeta_t(\omega)$ be pathwise Hölder continuous with Hölder exponent $\frac{1}{2}$.

The *Euler scheme*

$$Y_{n+1}(\omega) = (1 - \Delta_n)\, Y_n(\omega) + \zeta_{t_n}(\omega)\, \Delta_n$$

for the RODE

$$\frac{dx}{dt} = -x + \zeta_t(\omega),$$

attains the pathwise order $\frac{1}{2}$.

One can do better by using the pathwise *averaged Euler scheme*

$$Y_{n+1}(\omega) = (1 - \Delta_n)\, Y_n(\omega) + \int_{t_n}^{t_{n+1}} \zeta_t(\omega)\, dt,$$

which was proposed in

L. Grüne and P.E. Kloeden, *Pathwise approximation of random ordinary differential equations*, BIT 41 (2001), 710–721.

It attains the pathwise order 1 provided the integral is approximated with Riemann sums

$$\int_{t_n}^{t_{n+1}} \zeta_t(\omega)\, dt \approx \sum_{j=1}^{J_{\Delta_n}} \zeta_{t_n+j\delta}(\omega)\, \delta$$

with the step size δ satisfying $\delta^{1/2} \approx \Delta_n$ and $\delta \cdot J_{\Delta_n} = \Delta_n$.

In fact, this was proved for RODEs with an affine structure, i.e., of the form

$$\boxed{\frac{dx}{dt} = g(x) + H(x)\zeta_t,}$$

where $g : \mathbb{R}^d \to \mathbb{R}^d$ and $H : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^m$ and ζ_t is an m-dimensional process. The explicit averaged Euler scheme here is

$$\boxed{Y_{n+1} = Y_n + \left[g\left(Y_n\right) + H\left(Y_n\right)\ I_n\right]\ \Delta_n,}$$

where

$$\boxed{I_n(\omega) := \frac{1}{\Delta_n} \int_{t_n}^{t_{n+1}} \zeta_s(\omega)\, ds.}$$

For the general RODE (1) this suggests that one could pathwise average the vector field, i.e.,

$$\frac{1}{\Delta_n} \int_{t_n}^{t_{n+1}} f\left(\zeta_s(\omega), Y_n(\omega)\right)\ ds,$$

which is computationally expensive even for low dimensional systems.

An better alternative is to use the averaged noise within the vector field, which gives the general *averaged Euler scheme*

$$\boxed{Y_{n+1} = Y_n + f\left(I_n, Y_n\right)\ \Delta_n.}$$

3 Taylor–like expansions for RODEs

Taylor–like expansions were used to derive systematically higher order numerical schemes for RODEs in

P.E. Kloeden and A. Jentzen, *Pathwise convergent higher order numerical schemes for random ordinary differential equations*, Proc. Roy. Soc. London, Series A 463 (2007), 2929–2944.

A. Jentzen and P. E. Kloeden, *Pathwise Taylor schemes for random ordinary differential equations*, BIT 49 (1) (2009), 113–140.

even though the solutions of RODES are at most continuously differentiable.

To emphasize the role of the sample paths of the driving stochastic process ζ_t, a canonical sample space $\Omega = C(\mathbb{R}^+, \mathbb{R}^m)$ of continuous functions from $\omega : \mathbb{R}^+ \to \mathbb{R}^m$ will be used, so $\zeta_t(\omega) = \omega(t)$ for $t \in \mathbb{R}^+$.

The RODE (1) will henceforth be written

$$\frac{dx}{dt} = f(\omega(t), x),$$

Since f is assumed to be infinitely often continuously differentiable in its variables, the initial value problem (3) has a unique solution, which will be assumed to exist on a finite time interval $[t_0, T]$ under consideration.

By the continuity of the solution $x(t) = x(t; t_0, x_0, \omega)$ on $[t_0, T]$, there exists an $R = R(\omega, T) > 0$ such that

$$|x(t)| \leq R(\omega, T) \qquad \text{for all} \quad t \in [t_0, T].$$

Hölder continuity of the noise

It will be assumed that the sample paths of ζ_t are locally Hölder continuous with the same Hölder exponent,

i.e., there is a $\gamma \in (0, 1]$ such that for $\mathbb{P}$ almost all $\omega \in \Omega$ and each $T > 0$ there exists a $C_{\omega,T} > 0$ such that

$$|\omega(t) - \omega(s)| \leq C_{\omega,T} \cdot |t - s|^{\gamma} \quad \text{for all} \quad 0 \leq s, t \leq T. \qquad (5)$$

For such sample paths ω define

$$\|\omega\|_{\infty} := \sup_{t \in [t_0,T]} |\omega(t)|, \qquad \|\omega\|_{\gamma} := \|\omega\|_{\infty} + \sup_{\substack{s \neq t \in [t_0,T] \\ |t-s| \leq 1}} \frac{|\omega(t) - \omega(s)|}{|t - s|^{\gamma}},$$

so

$$|\omega(t) - \omega(s)| \leq \|\omega\|_{\gamma} \cdot |t - s|^{\gamma} \quad \text{for all} \quad s, t \in [t_0, T].$$

Let θ be the supremum of all γ with this property.

• Two cases will be distinguished: Case A in which (5) also holds for θ itself and Case B when it does not.

The Wiener process with $\theta = \frac{1}{2}$ is an example of Case B.

4 Multi–index notation

Let $\mathbb{N}_0$ denote the nonnegative integers. For a multi–index $\alpha = (\alpha_1, \alpha_2) \in \mathbb{N}_0^2$ define

$$|\alpha| := \alpha_1 + \alpha_2 \ , \qquad \alpha! := \alpha_1!\, \alpha_2!$$

For a given $\gamma \in (0, 1]$ define the weighted magnitude of a multi-index α by

$$|\alpha|_\gamma := \gamma\, \alpha_1 + \alpha_2$$

For each $K \in \mathbb{R}_+$ with $K \geq |\alpha|_\gamma$ define $|\alpha|_\gamma^K := K - |\alpha|_\gamma$.

For $f \in C^\infty(\mathbb{R} \times U, \mathbb{R}\,)$ denote

$$f_\alpha := \partial^\alpha f := (\partial_1)^{\alpha_1} (\partial_2)^{\alpha_2} f$$

with $\partial^{(0,0)} f = f$ and $(0, 0)! = 1$.

Let $R(\omega, T) > 0$ be an upper bound on the solution of the initial value problem (3) corresponding to the sample path ω on a fixed interval $[t_0, T]$ and define

$$\|f\|_k := \max_{|\alpha| \leq k} \sup_{\substack{|y| \leq \|\omega\|_\infty \\ |z| \leq R}} |f_\alpha(y, z)|.$$

For brevity, write $\|f\| := \|f\|_0$.

Note that the solution of the initial value problem (3) is Lipschitz continuous with

$$|x(t) - x(s)| \leq \|f\|\, |t - s| \qquad \text{for all} \quad s, t \in [t_0, T].$$

5 Taylor expansions of the vector field

The solution $x(t)$ of the initial value problem (3) is only once differentiable, so the usual Taylor expansion cannot be continued beyond the linear term.
Nevertheless, the special structure of a RODE and smoothness of f in both of its variables allows one to derive implicit Taylor–like expansions of arbitrary order for the solution.

Fix $\omega \in \Omega$ and an arbitrary $\hat{t} \in [t_0, T)$ and write

$$\Delta\omega_s := \omega(s) - \hat{\omega}, \qquad \Delta x_s := x(s) - \hat{x},$$

where $\hat{\omega} := \omega(\hat{t})$ and $\hat{x} := x(\hat{t})$.

Then, for a fixed $k \in \mathbb{Z}^+$, the usual Taylor expansion for f in both variables gives

$$f(\omega(s), x(s)) = \sum_{|\alpha| \le k} \frac{1}{\alpha!} \partial^\alpha f(\hat{\omega}, \hat{x}) \; (\Delta\omega_s)^{\alpha_1} (\Delta x_s)^{\alpha_2} \; + \; R_{k+1}(s)$$

with remainder term

$$R_{k+1}(s) = \sum_{|\alpha| = k+1} \frac{1}{\alpha!} \partial^\alpha f(\hat{\omega} + \xi_s \Delta\omega_s, \; \hat{x} + \xi_s \Delta x_s) \; (\Delta\omega_s)^{\alpha_1} (\Delta x_s)^{\alpha_2}$$

for some $\xi_s \in [0, 1]$.

Substituting this into the integral equation representation of the solution of (3),

$$\boxed{x(t) = \hat{x} + \int_{\hat{t}}^{t} f(\omega(s), x(s))\, ds}$$

gives

$$\Delta x_t = \underbrace{\sum_{|\alpha|\leq k} \frac{1}{\alpha!}\partial^{\alpha} f(\hat{\omega}, \hat{x}) \int_{\hat{t}}^{t} (\Delta\omega_s)^{\alpha_1}(\Delta x_s)^{\alpha_2}\, ds}_{Taylor-like\,approximation} + \underbrace{\int_{\hat{t}}^{t} R_{k+1}(s)\, ds}_{remainder},$$

or, more compactly, as

$$\boxed{\Delta x_t = \sum_{|\alpha|\leq k} T_{\alpha}(t;\hat{t}) + \int_{\hat{t}}^{t} R_{k+1}(s)\, ds\,. \qquad (6)}$$

where

$$\boxed{T_{\alpha}(t;\hat{t}) := \frac{1}{\alpha!}\, f_{\alpha}(\hat{\omega}, \hat{x}) \int_{\hat{t}}^{t} (\Delta\omega_s)^{\alpha_1}(\Delta x_s)^{\alpha_2}\, ds.}$$

- The expression (6) is implicit in Δx_s, so is not a standard Taylor expansion.

- Nevertheless, it can be used as the basis for constructing higher order numerical schemes for the RODE (1).

6 RODE–Taylor schemes

The RODE–Taylor schemes are a family of explicit one–step schemes for RODEs on subintervals $[t_n, t_{n+1}]$ of $[t_0, T]$ with step size, which are derived from the Taylor–like expansion (6).

The simplest case is for $k = 0$. Then (6) reduces to

$$\begin{aligned} x(t) &= \hat{x} + \frac{1}{(0,0)!}\partial^{(0,0)} f(\hat{\omega}\,, \hat{x}) \int_{\hat{t}}^{t} (\Delta\omega_s)^0 (\Delta x_s)^0 \, ds + \int_{\hat{t}}^{t} R_1(s) \, ds \\ &= \hat{x} + f(\hat{\omega}\,, \hat{x}) \, \Delta t + \int_{\hat{t}}^{t} R_1(s) \, ds, \end{aligned}$$

which leads to the well known Euler scheme

$$y_{n+1} = y_n + h_n \, f(\omega(t_n), y_n).$$

- To derive higher order schemes, the Δx_s terms inside the integrals must also be approximated.

This can be done with a numerical scheme of one order lower than that of the scheme to be derived.

Higher order schemes can thus be built up recursively for sets of multi–indices of the form

$$\mathcal{A}_K := \left\{\alpha = (\alpha_1, \alpha_2) \in \mathbb{N}_0^2 \ : \ |\alpha|_\theta = \theta\alpha_1 + \alpha_2 < K\right\},$$

where $K \in \mathbb{R}_+$ and $\theta \in (0, 1]$ is specified by the noise process in the RODE.

Fix $K \in \mathbb{R}_+$ and consider the first step

$$y_1^{K,h} \ = \ \hat{y} + \Delta y_h^{(K)}(\hat{t}, \hat{y})$$

of a numerical approximation at the time instant $\hat{t} + h$ for a step size $h \in (0, 1]$ and initial value $(\hat{t}, \hat{y})$, where the increments $\Delta y_h^{(K)}$ are defined recursively as follows

$$\Delta y_h^{(0)} := 0, \qquad \Delta y_h^{(K)}(\hat{t}, \hat{y}) := \sum_{|\alpha|_\theta < K} N_\alpha^{(K)}(\hat{t} + h, \hat{t}, \hat{y}),$$

where

$$N_\alpha^{(K)}(\hat{t} + h, \hat{t}, \hat{y}) := \frac{1}{\alpha!} f_\alpha(\hat{\omega}, \hat{y}) \int_{\hat{t}}^{\hat{t}+h} (\Delta\omega_s)^{\alpha_1} \left(\Delta y_{\Delta s}^{(|\alpha|_\theta^K)}(\hat{t}, \hat{y})\right)^{\alpha_2} \, ds$$

with $\Delta s = s - \hat{t}$ and $|\alpha|_\theta^K = K - \theta \cdot \alpha_1 - \alpha_2$, i.e., in nontrivial cases, the $N_\alpha^{(K)}$ are evaluated in terms of previously determined $\Delta y_{\Delta s}^{(L)}$ with $L := K - \theta \cdot \alpha_1 - \alpha_2 < K$. This procedure is repeated for each time step to give the $\mathcal{A}_K$-RODE–Taylor scheme

$$y_{n+1}^{K,h} := y_n^{K,h} + \sum_{\mathcal{A}_K} N_\alpha^{(K)}\left(t_{n+1}, t_n, y_n^{K,h}\right), \qquad n = 0, 1, 2 \ldots, N_T - 1,$$

on discretisation subintervals $[t_n, t_{n+1}]$ with step sizes $h_n = t_{n+1} - t_n > 0$.

7 Discretisation error

The increment function of an $\mathcal{A}_K$-RODE–Taylor scheme or, simply, K-RODE–Taylor scheme

$$F^{(K)}(h,\hat{t},\hat{y}) := \frac{1}{h}\sum_{|\alpha|_\theta<K}\frac{1}{\alpha!}\,f_\alpha(\hat{\omega},\hat{y})\int_{\hat{t}}^{\hat{t}+h}(\Delta\omega_s)^{\alpha_1}\left(\Delta y_{\Delta s}^{(|\alpha|_\theta^K)}(\hat{t},\hat{y})\right)^{\alpha_2}\,ds$$

is continuous in its variables as well as continuously differentiable, hence locally Lipschitz, in the $\hat{y}$ variable with

$$\lim_{h\to 0+} F^{(K)}(h,\hat{t},\hat{y}) = f(\hat{\omega},\hat{y}).$$

RODE–Taylor schemes are thus consistent and hence convergent for each $K > 0$.

Moreover, the classical theorem for ODEs on the loss of a power from the local to global discretisation errors also holds for the RODE–Taylor schemes. Consequently, it suffices to estimate the local discretisation error

$$L_h^{(K)}(\hat{t},\hat{y}) := \left|x(\hat{t}+h,\hat{t},\hat{y}) - y_1^{K,h}(\hat{t},\hat{y})\right|,$$

where $x(\hat{t}+h,\hat{t},\hat{y})$ is the value of the solution of the RODE at time $\hat{t}+h$ with initial value $(\hat{t},\hat{y})$ and $y_1^{K,h}(\hat{t},\hat{y})$ is the first step of the numerical scheme with step size h for the same initial value.

Define $\tilde{R}_0 := 0$ and for $K > 0$ define

$$\tilde{R}_K := \sup_{0<L\leq K} \max_{\substack{(h,t,x)\in \\ [0,1]\times[t_0,T]\times[-R,R]^d}} |F^{(L)}(h,t,x)|.$$

In addition, let

$$k = k_K := \left\lfloor \frac{K}{\theta} \right\rfloor \qquad (\lfloor\cdot\rfloor \quad \text{integer part})$$

and define

$$R_K := \max\left\{\tilde{R}_K, \|f\|_{k+1}\right\}.$$

It is necessary to distinguish two cases, Case A in which the Hölder estimate (5) also holds for the supremum θ of the admissible exponents itself and Case B when it does not.

Theorem 1. *The local discretisation error for a RODE–Taylor scheme in Case A satisfies*

$$\left|L_h^{(K)}(\hat{t},\hat{x})\right| \leq C_K\, h^{K+1}$$

for each $0 \leq h \leq 1$, *where*

$$C_K := \left(e^{\|\omega\|_\theta + 2R_K}\right)^{K+1}.$$

In Case B it satisfies

$$\left|L_h^{(K)}(\hat{t},\hat{x})\right| \leq C_K^\varepsilon \cdot h^{K+1-\varepsilon}$$

for $\varepsilon > 0$ *arbitrarily small, where*

$$C_K^\varepsilon := \left(e^{\|\omega\|_{\gamma_\varepsilon} + 2R_K}\right)^{K+1}, \qquad \gamma_\varepsilon := \theta - \frac{\varepsilon}{(k+1)^2}.$$

8 RODE–Taylor schemes: Wiener process

A Wiener process as the driving process falls into Case B with $\theta = \frac{1}{2}$.

The $\mathcal{A}_K$-RODE–Taylor schemes have pathwise global convergence order $K - \varepsilon$.

$$\mathcal{A}_K := \{\ \alpha\ :\ |\alpha|_{\frac{1}{2}} < K\ \} = \{\ \alpha\ :\ \alpha_1\ +\ 2\alpha_2 \leq 2K - 1\ \}$$

Example 1. *The* 0*-RODE–Taylor scheme corresponding to* $\mathcal{A}_0 = \emptyset$ *is* $y_n \equiv y_0$*, which is an* inconsistent *scheme.*

Example 2. *The* 0.5*-RODE–Taylor scheme corresponding to* $\mathcal{A}_{0.5} = \{\ (0,0)\ \}$ *is the classical Euler scheme*

$$y_{n+1} = y_n\ +\ hf(\omega(t_n), y_n), \tag{7}$$

which has order $0.5 - \varepsilon$*.*

Example 3. *The* 1.0*-RODE–Taylor scheme corresponding to* $\mathcal{A}_{1.0} = \{(0,0),\ (1,0)\}$ *is the "improved" Euler scheme,*

$$y_{n+1} = y_n\ +\ hf(\omega(t_n), y_n)\ +\ f_{(1,0)}(\omega(t_n), y_n) \int_{t_n}^{t_{n+1}} \Delta\omega_s\, ds.$$

Its order $1 - \varepsilon$ *is comparable to that of the Euler scheme for smooth ODEs.*

In the following schemes the coefficient functions on the right side are evaluated at $(\omega(t_n), y_n)$.

Example 4. *The* 1.5*-RODE–Taylor scheme corresponding to* $\mathcal{A}_{1.5} = \{(0,0),\ (1,0),\ (2,0),\ (0,1)\}$ *is*

$$y_{n+1} = y_n \ + \ hf \ + \ f_{(1,0)} \int_{t_n}^{t_{n+1}} \Delta\omega_s \, ds + \ \frac{f_{(2,0)}}{2} \int_{t_n}^{t_{n+1}} (\Delta\omega_s)^2 \, ds + f_{(0,1)} f \, \frac{h^2}{2}$$

Here $|(0,1)|_{\frac{1}{2}}^{1.5} = 1.5{-}1 = 0.5$ *and the last term is obtained from* $f_{(0,1)} \int_{t_n}^{t_{n+1}} (\Delta y_{\Delta s}^{(0.5)})\, ds$ *with* $\Delta y_{\Delta s}^{(0.5)} = (s - t_n) f$ *coming from the Euler scheme (7).*

Example 5. *In the next case* $\mathcal{A}_{2.0} = \{(0,0),(1,0),(2,0),\ (3,0),(0,1),(1,1)\}$ *and* $|(1,1)|_{\frac{1}{2}}^2 = 0.5$, $|(0,1)|_{\frac{1}{2}}^2 = 1$, *so the terms* $\Delta y_{\Delta s}^{(0.5)}$, $\Delta y_{\Delta s}^{(1.0)}$ *corresponding to the* 0.5*-RODE–Taylor scheme and* 1.0*-RODE–Taylor scheme are required in the right hand side of the new scheme. The resulting* 2.0*-RODE–Taylor scheme is then*

$$\begin{aligned} y_{n+1} \ = \ & y_n \ + \ hf \ + \ f_{(1,0)} \int_{t_n}^{t_{n+1}} \Delta\omega_s \, ds + \ \frac{f_{(2,0)}}{2} \int_{t_n}^{t_{n+1}} (\Delta\omega_s)^2 \, ds \\ & + \frac{f_{(3,0)}}{6} \int_{t_n}^{t_{n+1}} (\Delta\omega_s)^3 \, ds + f_{(0,1)} f \frac{h^2}{2} \\ & + f_{(0,1)} f_{(1,0)} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} \Delta\omega_v \, dv \, ds + \ f_{(1,1)} f \int_{t_n}^{t_{n+1}} \Delta\omega_s \Delta s \, ds \ . \end{aligned}$$

Remark: The above K-RODE–Taylor schemes are not necessarily optimal in the sense of involving the minimum number of terms for the given order.

9 Fractional Brownian motion

Fractional Brownian motion with Hurst exponent $H = \frac{3}{4}$ also falls into Case B with $\theta = \frac{3}{4}$. The RODE–Taylor schemes generally contain fewer terms than the schemes of the same order with a Wiener process or attain a higher order when they contain the same terms.

Example 6. *For $\mathcal{A}_K = \{\ (0,0)\ \}$ with $K \in (0, \frac{3}{4}]$ the RODE–Taylor scheme is the classical Euler scheme in Example 2, but now the order is $\frac{3}{4} - \varepsilon$.*

Example 7. *$\mathcal{A}_K = \{(0,0),(1,0)\}$ for $K \in (\frac{3}{4}, 1]$: the RODE–Taylor scheme is the same as that in Example 3 and also has order $1 - \varepsilon$.*

Example 8. *, For $\mathcal{A}_K = \{(0,0),(1,0),(0,1)\}$ with $K \in (1, \frac{3}{2}]$ the RODE–Taylor scheme,*

$$y_{n+1} = y_n \ + \ hf \ + \ f_{(1,0)} \int_{t_n}^{t_{n+1}} \Delta\omega_s + f_{(0,1)} f \, \frac{h^2}{2},$$

which has order $1.5 - \varepsilon$, omits one of the terms in the RODE–Taylor scheme of the same order for a Wiener process given in Example 4.

Example 9. *For $\mathcal{A}_K = \{(0,0),(1,0),(0,1),(2,0)\}$ with $K \in (\frac{3}{2}, \frac{7}{4}]$ the RODE–Taylor scheme is the same as that for a Wiener process case in Example 4, but now has order $\frac{7}{4} - \varepsilon$ instead of order order $\frac{3}{2} - \varepsilon$.*

Lecture 3: Stochastic Ordinary Differential Equations

Deterministic calculus is much more robust to approximation than Itô stochastic calculus because the integrand function in a Riemann sum approximating a Riemann integral can be evaluated at an arbitrary point of the discretisation subinterval, whereas for an Itô stochastic integral the integrand function must always be evaluated at the left hand endpoint.

Consequently, considerable care is needed in deriving numerical schemes for stochastic ordinary differential equations (SODEs) to ensure that they are consistent with Itô calculus.

In particular, stochastic Taylor schemes are the essential starting point for the derivation of consistent higher order numerical schemes for SODEs.

Other types of schemes for SODEs, such as derivative–free schemes, can then be obtained by modifying the corresponding stochastic Taylor schemes.

1 Itô SODEs

Consider a scalar Itô stochastic differential equation (SODE)

$$dX_t = a(t, X_t)\, dt + b(t, X_t)\, dW_t, \qquad (1)$$

where $(W_t)_{t\in\mathbb{R}^+}$ is a standard Wiener process, i.e., with $W_0 = 0$, w.p.1, and increments $W_t - W_s \sim N(0; t-s)$ for $t \geq s \geq 0$, which are independent on nonoverlapping subintervals.

The SODE (1) is, in fact, only symbolic for the stochastic integral equation

$$X_t = X_{t_0} + \int_{t_0}^{t} a(s, X_s)\, ds + \int_{t_0}^{t} b(s, X_s)\, dW_s, \qquad (2)$$

where the first integral is pathwise a deterministic Riemann integral and the second is an Itô stochastic integral — it is not a pathwise Riemann–Stieltjes integral !

The Itô stochastic integral of a nonanticipative mean–square integrable integrand g is defined in terms of the mean–square limit, namely,

$$\int_{t_0}^{T} g(s)\, dW_s := \text{m.s.} - \lim_{\Delta\to 0} \sum_{n=0}^{N_T-1} g(t_n, \omega) \left\{ W_{t_{n+1}}(\omega) - W_{t_n}(\omega) \right\},$$

taken over partitions of $[t_0, T]$ of maximum step size $\Delta := \max_n \Delta_n$, where $\Delta_n = t_{n+1} - t_n$ and $t_{N_T} = T$.

Two very useful properties of Itô integrals are

$$\mathbb{E}\left(\int_{t_0}^{T} g(s)\, dW_s \right) = 0, \qquad \mathbb{E}\left(\int_{t_0}^{T} g(s)\, dW_s \right)^2 = \int_{t_0}^{T} \mathbb{E} g(s)^2\, ds.$$

2 Existence and uniqueness of strong solutions

The following is a standard existence and uniqueness theorem for SODEs. The vector valued case is analogous.

Theorem 1. *Supose that* a, $b : [t_0, T] \times \mathbb{R} \to \mathbb{R}$ *are continuous in* (t, x) *and satisfy the global Lipschitz condition*

$$|a(t,x) - a(t,y)| + |b(t,x) - b(t,y)| \leq L|x - y|$$

uniformly in $t \in [t_0, T]$ *and suppose that the random variable* X_0 *is nonanticipative with respect to the Wiener process* W_t *with*

$$\mathbb{E}(X_0^2) < \infty.$$

Then the Itô SODE

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

has a unique strong solution on $[t_0, T]$ *with initial value* $X_{t_0} = X_0$.

Alternatively, one could assume just a local Lipschitz condition. Then, the additional linear growth condition

$$|a(t,x)| + |b(t,x)| \leq K(1 + |x|)$$

ensures the existence of the solution on the entire time interval $[0, T]$, i.e., prevents explosions of solutions in finite time.

P.E. Kloeden and E. Platen, *Numerical Solutions of Stochastic Differential Equations*, Springer-Verlag, Berlin Heidelberg New York, 1992.

3 Simple numerical schemes for SODEs

The counterpart of the Euler scheme for the SODE (1) is the Euler–Maruyama scheme

$$Y_{n+1} = Y_n + a(t_n, Y_n)\,\Delta_n + b(t_n, Y_n)\,\Delta W_n, \qquad (3)$$

with time and noise increments

$$\Delta_n = t_{n+1} - t_n = \int_{t_n}^{t_{n+1}} ds, \qquad \Delta W_n = W_{t_{n+1}} - W_{t_n} = \int_{t_n}^{t_{n+1}} dW_s.$$

It seems to be (and is) consistent with the Itô stochastic calculus because the noise term in (3) approximates the Itô stochastic integral in (2) over a discretisation subinterval $[t_n, t_{n+1}]$ by evaluating its integrand at the left hand end point:

$$\int_{t_n}^{t_{n+1}} b(s, X_s)\,dW_s \approx \int_{t_n}^{t_{n+1}} b(t_n, X_{t_n})\,dW_s = b(t_n, X_{t_n}) \int_{t_n}^{t_{n+1}} dW_s.$$

It is usual to distinguish between strong and weak convergence, depending on whether the realisations or only their probability distributions are required to be close.

Let Δ be the maximum step size of a given partition of a fixed interval $[t_0, T]$. A numerical scheme is said to converge with strong order γ if

$$\mathbb{E}\left(\left|X_T - Y_{N_T}^{(\Delta)}\right|\right) \le K_T\,\Delta^{\gamma} \qquad (4)$$

and with weak order β if, for each polynomial g,

$$\left|\mathbb{E}\left(g(X_T)\right) - \mathbb{E}\left(g(Y_{N_T}^{(\Delta)})\right)\right| \le K_{g,T}\,\Delta^{\beta} \qquad (5)$$

- The Euler–Maruyama scheme (3) has strong order $\gamma = \frac{1}{2}$, weak order $\beta = 1$.

To obtain a higher order, one should avoid heuristic adaptations of well known deterministic numerical schemes because they are usually inconsistent with Itô calculus or, when consistent, do not improve the order of convergence.

Example: The deterministic Heun scheme adapted to the Itô SODE (1) has the form

$$\begin{aligned} Y_{n+1} &= Y_n + \frac{1}{2}\left[a(t_n, Y_n) + a(t_{n+1}, Y_n + a(t_n, Y_n)\Delta_n + b(t_n, Y_n)\Delta W_n)\right]\,\Delta_n \\ &\quad + \frac{1}{2}\left[b(t_n, Y_n) + b(t_{n+1}, Y_n + a(t_n, Y_n)\Delta_n + b(t_n, Y_n)\Delta W_n)\right]\,\Delta W_n. \end{aligned}$$

For the Itô SODE $dX_t = X_t\, dW_t$ it simplifies to

$$Y_{n+1} = Y_n + \frac{1}{2} Y_n\,(2 + \Delta W_n)\,\Delta W_n.$$

The conditional expectation

$$\mathbb{E}\left(\frac{Y_{n+1} - Y_n}{\Delta_n}\,\middle|\, Y_n = x\right) = \frac{x}{\Delta_n}\,\mathbb{E}\left(\Delta W_n + \frac{1}{2}(\Delta W_n)^2\right) = \frac{x}{\Delta_n}\left(0 + \frac{1}{2}\Delta_n\right) = \frac{1}{2}\,x$$

should approximate the drift term $a(t, x) \equiv 0$ of the SODE. The adapted Heun scheme is thus not consistent with Itô calculus and does not converge in either the weak or strong sense.

To obtain a higher order of convergence one needs to provide more information about the Wiener process within the discretisation subinterval than that provided by the simple increment ΔW_n.

Such information is provided by multiple integrals of the Wiener process, which arise in stochastic Taylor expansions of the solution of an SODE.

- Consistent numerical schemes of arbitrarily desired higher order can be derived by truncating appropriate stochastic Taylor expansions.

4 Itô–Taylor expansions

Itô–Taylor expansions or stochastic Taylor expansions of solutions of Itô SODEs are derived through an iterated application of the stochastic chain rule, the Itô formula.

The nondifferentiability of the solutions in time is circumvented by using the integral form of the Itô formula.

The Itô formula for scalar valued function $f(t, X_t)$ of the solution X_t of the scalar Itô SODE (1) is

$$f(t, X_t) = f(t_0, X_{t_0}) + \int_{t_0}^{t} L^0 f(s, X_s)\, ds + \int_{t_0}^{t} L^1 f(s, X_s)\, dW_s, \qquad (6)$$

where the operators L^0 and L^1 are defined by

$$L^0 = \frac{\partial}{\partial t} + a\,\frac{\partial}{\partial x} + \frac{1}{2} b^2 \frac{\partial^2}{\partial x^2}, \qquad L^1 = b\,\frac{\partial}{\partial x}.$$

This differs from the deterministic chain rule by the additional third term in the L^0 operator, which is due to the fact that $\mathbb{E}\,(|\Delta W|^2) = \Delta t$.

When $f(t, x) \equiv x$, the Itô formula (6) is just the integral version (2) of the SODE (1) in the integral form

$$X_t = X_{t_0} + \int_{t_0}^{t} a(s, X_s)\, ds + \int_{t_0}^{t} b(s, X_s)\, dW_s. \qquad (7)$$

Iterated application of the Itô formula

Applying the Itô formula to the integrand functions

$$f(t,x) = a(t,x), \qquad f(t,x) = b(t,x)$$

in the SODE (7) gives

$$\begin{aligned} X_t &= X_{t_0} + \int_{t_0}^{t} \left[a(t_0, X_{t_0}) + \int_{t_0}^{s} L^0 a(u, X_u)\, du + \int_{t_0}^{s} L^1 a(u, X_u)\, dW_u \right] ds \\ &\quad + \int_{t_0}^{t} \left[b(t_0, X_{t_0}) + \int_{t_0}^{s} L^0 b(u, X_u)\, du + \int_{t_0}^{s} L^1 b(u, X_u)\, dW_u \right] dW_s. \\ &= X_{t_0} + a(t_0, X_{t_0}) \int_{t_0}^{t} ds + b(t_0, X_{t_0}) \int_{t_0}^{t} dW_s + R_1(t, t_0) \end{aligned} \tag{8}$$

with the remainder

$$\begin{aligned} R_1(t,t_0) &= \int_{t_0}^{s} \int_{t_0}^{s} L^0 a(u, X_u)\, du\, ds + \int_{t_0}^{s} \int_{t_0}^{s} L^1 a(u, X_u)\, dW_u\, ds \\ &\quad + \int_{t_0}^{t} \int_{t_0}^{s} L^0 b(u, X_u)\, du\, dW_s + \int_{t_0}^{t} \int_{t_0}^{s} L^1 b(u, X_u)\, dW_u\, dW_s. \end{aligned}$$

Discarding the remainder gives the simplest stochastic Taylor approximation

$$\boxed{X_t \approx X_{t_0} + a(t_0, X_{t_0}) \int_{t_0}^{t} ds + b(t_0, X_{t_0}) \int_{t_0}^{t} dW_s,}$$

which has strong order $\gamma = 0.5$ and weak order $\beta = 1$.

- Higher order stochastic Taylor expansions are obtained by applying the Itô formula to selected integrand functions in the remainder.

e.g., to the integrand $L^1 b$ in the fourth double integral of the remainder $R_1(t, t_0)$ gives the stochastic Taylor expansion

$$\begin{aligned} X_t &= X_{t_0} + a(t_0, X_{t_0}) \int_{t_0}^{t} ds + b(t_0, X_{t_0}) \int_{t_0}^{t} dW_s \\ &\quad + L^1 b(t_0, X_{t_0}) \int_{t_0}^{t} \int_{t_0}^{s} dW_u \, dW_s + R_2(t, t_0) \end{aligned} \tag{9}$$

with the remainder

$$\begin{aligned} R_2(t, t_0) &= \int_{t_0}^{s} \int_{t_0}^{s} L^0 a(u, X_u) \, du \, ds + \int_{t_0}^{s} \int_{t_0}^{s} a(u, X_u) \, dW_u \, ds \\ &\quad + \int_{t_0}^{t} \int_{t_0}^{s} L^0 b(u, X_u) \, du \, dW_s + \int_{t_0}^{t} \int_{t_0}^{s} \int_{t_0}^{u} L^0 L^1 b(v, X_v) \, dv \, dW_u \, dW_s \\ &\quad + \int_{t_0}^{t} \int_{t_0}^{s} \int_{t_0}^{u} L^1 L^1 b(v, X_v) \, dW_v \, dW_u \, dW_s. \end{aligned}$$

Discarding the remainder gives the stochastic Taylor approximation

$$\begin{aligned} X_t &\approx X_{t_0} + a(t_0, X_{t_0}) \int_{t_0}^{t} ds + b(t_0, X_{t_0}) \int_{t_0}^{t} dW_s \\ &\quad + b(t_0, X_{t_0}) \frac{\partial b}{\partial x}(t_0, X_{t_0}) \int_{t_0}^{t} \int_{t_0}^{s} dW_u \, dW_s, \end{aligned}$$

since $L^1 b = b \dfrac{\partial b}{\partial x}$. This has strong order $\gamma = 1$ (and also weak order $\beta = 1$).

Remark: These two examples already indicate the general pattern of the schemes:
i) *they achieve their higher order through the inclusion of multiple stochastic integral terms;*
ii) *an expansion may have different strong and weak orders of convergence;*
iii) *the possible orders for strong schemes increase by a fraction* $\frac{1}{2}$*, taking values* $\frac{1}{2}$, 1, $\frac{3}{2}$, 2, ..., *whereas possible orders for weak schemes are whole numbers* 1, 2, 3, ...

5 General stochastic Taylor expansions

Multi–indices provide a succinct means to describe the terms that should be included or expanded to obtain a stochastic Taylor expansion of a particular order as well as for representing the iterated differential operators and stochastic integrals that appear in the terms of such expansions.

A multi–index α of length $l(\alpha) = l$ is an l-dimensional row vector $\alpha = (j_1, j_2, \ldots, j_l) \in \{0,1\}^l$ with components $j_i \in \{0,1\}$ for $i \in \{1, 2, \ldots, l\}$.

Let $\mathcal{M}_1$ be the set of all multi–indices of length greater than or equal to zero, where a multi–index $\emptyset$ of length zero is introduced for convenience.

Given a multi–index $\alpha \in \mathcal{M}_1$ with $l(\alpha) \geq 1$, write $-\alpha$ and $\alpha-$ for the multi–index in $\mathcal{M}_1$ obtained by deleting the first and the last component, respectively, of α.

For a multi–index $\alpha = (j_1, j_2, \ldots, j_l)$ with $l \geq 1$, define the multiple Itô integral recursively by (here $W_t^0 = t$ and $W_t^1 = W_t$)

$$I_\alpha[g(\cdot)]_{t_0,t} := \int_{t_0}^{t} I_{\alpha-}[g(\cdot)]_{t_0,s}\, dW_s^{j_l}, \qquad I_\emptyset[g(\cdot)]_{t_0,t} := g(t)$$

and the Itô coefficient function for a deterministic function f recursively by

$$f_\alpha := L^{j_1} f_{-\alpha}, \qquad f_\emptyset = f.$$

The multiple stochastic integrals appearing in a stochastic Taylor expansion with constant integrands cannot be chosen completely arbitrarily.

The set of corresponding multi–indices must form an hierarchical set, i.e., a nonempty subset $\mathcal{A}$ of $\mathcal{M}_1$ with

$$\sup_{\alpha\in\mathcal{A}} l(\alpha) < \infty \qquad \text{and} \qquad -\alpha \in \mathcal{A} \quad \text{for each} \quad \alpha \in \mathcal{A} \setminus \{\emptyset\}.$$

The multi–indices of the remainder terms in a stochastic Taylor expansion for a given hierarchical set $\mathcal{A}$ belong to the corresponding *remainder set* $\mathcal{B}(\mathcal{A})$ of $\mathcal{A}$ defined by

$$\mathcal{B}(\mathcal{A}) = \{\alpha \in \mathcal{M}_1 \setminus \mathcal{A} : -\alpha \in \mathcal{A}\},$$

i.e., consisting of all of the "next following" multi–indices with respect to the given hierarchical set.

The Itô–Taylor expansion for the hierarchical set $\mathcal{A}$ and remainder set $\mathcal{B}(\mathcal{A})$ is

$$f(t, X_t) = \sum_{\alpha\in\mathcal{A}} f_\alpha(t_0, X_{t_0})\ I_\alpha\left[1\right]_{t_0,t} + \sum_{\alpha\in\mathcal{B}(\mathcal{A})} I_\alpha\left[f_\alpha(\cdot, X_\cdot)\right]_{t_0,t},$$

i.e., with constant integrands (hence constant coefficients) in the first sum and time dependent integrands in the remainder sum.

6 Itô–Taylor Numerical Schemes for SODEs

Itô–Taylor numerical schemes are obtained by applying Itô–Taylor expansion to the identity function $f = \mathrm{id}$ on a subinterval $[t_n, t_{n+1}]$ at a starting point (t_n, Y_n) and discarding the remainder term.

The Itô-Taylor expansion (8) gives the Euler–Maruyama scheme (3), which is the simplest nontrivial stochastic Taylor scheme. It has strong order $\gamma = 0.5$ and weak order $\beta = 1$.

The Itô–Taylor expansion (9) gives the *Milstein scheme*

$$\begin{aligned} Y_{n+1} &= Y_n + a(t_n, Y_n)\,\Delta_n + b(t_n, Y_n)\,\Delta W_n \\ &\quad + b(t_n, Y_n)\frac{\partial b}{\partial x}(t_n, X_n)\int_{t_n}^{t_{n+1}}\int_{t_n}^{s} dW_u\, dW_s. \end{aligned}$$

Here the coefficient functions are $f_{(0)} = a$, $f_{(1)} = b$, $f_{(1,1)} = b\dfrac{\partial b}{\partial x}$ and the iterated integrals

$$I_{(0)}(t_n, t_{n+1}) = \int_{t_n}^{t_{n+1}} dW_s^0 = \Delta_n, \qquad I_{(1)}(t_n, t_{n+1}) = \int_{t_n}^{t_{n+1}} dW_s^1 = \Delta W_n,$$

and

$$I_{(1,1)}(t_n, t_{n+1}) = \int_{t_n}^{t_{n+1}}\int_{t_n}^{s} dW_\tau^1\, dW_s^1 = \frac{1}{2}\left[(\Delta W_n)^2 - \Delta_n\right]. \tag{10}$$

The Milstein scheme converges with strong order $\gamma = 1$ and weak order $\beta = 1$.

It has a higher order of convergence in the strong sense than the Euler–Maruyama scheme, but gives no improvement in the weak sense.

Applying this idea to the Itô–Taylor expansion corresponding to the hierarchical set $\mathcal{A}$ gives the $\mathcal{A}$-stochastic Taylor scheme

$$Y_{n+1}^{\mathcal{A}} = \sum_{\alpha \in \mathcal{A}} I_\alpha \left[\mathrm{id}_\alpha(t_n, Y_n^{\mathcal{A}}) \right]_{t_n, t_{n+1}} = Y_n^{\mathcal{A}} + \sum_{\alpha \in \mathcal{A} \setminus \emptyset} \mathrm{id}_\alpha(t_n, Y_n^{\mathcal{A}})\, I_\alpha \left[1\right]_{t_n, t_{n+1}}. \tag{11}$$

The strong order γ stochastic Taylor scheme, which converges with strong order γ, involves the hierarchical set

$$\Lambda_\gamma = \left\{ \alpha \in \mathcal{M}_m : \; l(\alpha) + n(\alpha) \leq 2\gamma \quad \text{or} \quad l(\alpha) = n(\alpha) = \gamma + \frac{1}{2} \right\},$$

where $n(\alpha)$ denotes the number of components of a multi–index α equal to 0.

The weak order β stochastic Taylor scheme, which converges with weak order β, involves the hierarchical set

$$\Gamma_\beta = \{ \alpha \in \mathcal{M}_m : l(\alpha) \leq \beta \}.$$

For example, the hierarchical sets $\Lambda_{1/2} = \Gamma_1 = \{\emptyset, (0), (1)\}$ give the stochastic Euler–Maruyama scheme, which is both strongly and weakly convergent, while the strongly convergent Milstein scheme corresponds to the hierarchical set $\Lambda_1 = \{\emptyset, (0), (1), (1,1)\}$.

- The Milstein scheme does not correspond to a stochastic Taylor scheme for an hierarchical set Γ_β.

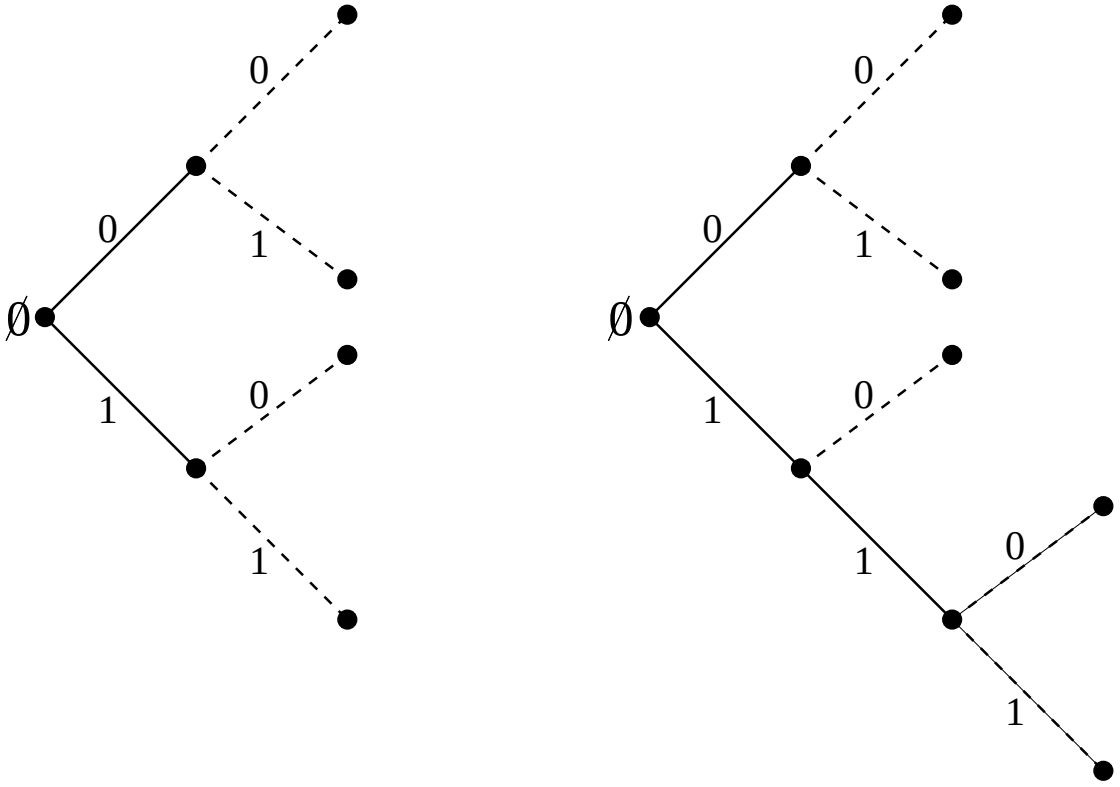

Figure 1: Stochastic Taylor trees for the Euler–Maruyama scheme (left) and the Milstein scheme (right). The multi–indices are formed by concatenating indices along a branch from the right back towards the root $\emptyset$ of the tree. Dashed line segments correspond to remainder terms.

Remark: Convergence theorems for stochastic Taylor schemes (Theorems 10.6.3 and 14.5.1 in Kloeden and Platen) assume that the coefficients of the SODE are sufficiently often differentiable so that all terms in these schemes make sense, with all of these partial derivatives being uniformly bounded, hence globally Lipschitz, in the case of strong convergence, and with globally Lipschitz coefficients and all of the required partial derivatives satisfying a linear growth bound in the case of weak convergence.

These will be called the standard assumptions. They are obviously not satisfied by some SODEs that arise in many interesting applications .

7 Pathwise convergence

Pathwise convergence was already considered for RODEs and is also interesting for SODEs because numerical calculations of the random variables Y_n in the numerical scheme above are carried out path by path.

Itô stochastic calculus is, however, an L^2 or a mean–square calculus and not a pathwise calculus.

Given that the sample paths of a Wiener process are Hölder continuous with exponent $\frac{1}{2} - \epsilon$ one may ask:

Is the convergence order $\frac{1}{2} - \epsilon$ "sharp" for pathwise approximation?

The answer is no! An arbitrary pathwise convergence order is possible.

P.E. Kloeden and A. Neuenkirch, *The pathwise convergence of approximation schemes for stochastic differential equations*, LMS J. Comp. Math. 10 (2007)

Theorem 2. *Under the standard assumptions the Itô–Taylor scheme of strong order $\gamma > 0$ converges pathwise with order $\gamma - \epsilon$ for all $\epsilon > 0$, i.e.,*

$$\sup_{i=0,\ldots,N_T} \left| X_{t_n}(\omega) - Y_n^{(\gamma)}(\omega) \right| \leq K_{\epsilon,T}^{(\gamma)}(\omega) \cdot \Delta^{\gamma-\epsilon}$$

for almost all $\omega \in \Omega$.

Note that the error constant here depends on ω, so it is in fact a random variable. The nature of its statistical properties is an interesting question, about which little is known theoretically so far and requires further investigation.

The proof of Theorem 2 is based on the Burkholder–Davis–Gundy inequality

$$\mathbb{E} \sup_{s\in[0,t]} \left| \int_0^s X_\tau \, dW_\tau \right|^p \leq C_p \cdot \mathbb{E} \left| \int_0^t X_\tau^2 \, d\tau \right|^{p/2}$$

and a Borel–Cantelli argument in the following lemma.

Lemma 1. *Let $\gamma > 0$ and $c_p \geq 0$ for $p \geq 1$. If $\{Z_n\}_{n\in\mathbb{N}}$ is a sequence of random variables with*

$$(\mathbb{E}|Z_n|^p)^{1/p} \leq c_p \cdot n^{-\gamma}$$

for all $p \geq 1$ and $n \in \mathbb{N}$, then for each $\epsilon > 0$ there exists a non–negative random variable K_ϵ such that

$$|Z_n(\omega)| \leq K_\epsilon(\omega) \cdot n^{-\gamma+\varepsilon}, \qquad a.s.,$$

for all $n \in \mathbb{N}$.

8 Restrictiveness of the standard assumptions

Proofs in the literature of the convergence orders of Itô–Taylor schemes assume that the coefficient functions f_α are uniformly bounded on $\mathbb{R}^1$, i.e., the partial derivatives of appropriately high order of the SODE coefficient functions a and b are uniformly bounded on $\mathbb{R}^1$.

This assumption is not satisfied for many SODEs in important applications such as:

- the stochastic Ginzburg–Landau equation

$$dX_t = \left(\left(\nu + \frac{1}{2}\sigma^2 \right) X_t - \lambda X_t^3 \right) \, dt + \sigma X_t \, dW_t.$$

Matters are even worse for

- the Fisher–Wright equation

$$dX_t = \left[\kappa_1 (1 - X_t) - \kappa_2 X_t \right] \, dt + \sqrt{X_t (1 - X_t)} \, dW_t$$

- the Feller diffusion with logistic growth SODE

$$dX_t = \lambda X_t \left(K - X_t \right) \, dt + \sigma \sqrt{X_t} \, dW_t$$

- the Cox–Ingersoll–Ross equation

$$dV_t = \kappa \left(\lambda - V_t \right) \, dt + \theta \sqrt{V_t} \, dW_t,$$

since the square root function is not differentiable at zero and requires the expression under it to remain non–negative for the SODE to make sense.

9 Counterexample: Euler–Maruyama scheme

The scalar SODE

$$dX_t = -X_t^3\,dt + dW_t$$

with the cubic drift and additive noise has a globally pathwise asymptotically stable stochastic stationary solution. Its solution on the time interval $[0, 1]$ for initial value $X_0 = 0$ satisfies the stochastic integral equation

$$X_t = -\int_0^t X_s^3\,ds + W_t \tag{12}$$

and has finite first moment $\mathbb{E}\,|X_1| < \infty$ at $T = 1$.

The corresponding Euler–Maruyama scheme with constant step size $\Delta = \frac{1}{N}$ is given by

$$Y_{k+1}^{(N)} = Y_k^{(N)} - \left(Y_k^{(N)}\right)^3 \Delta + \Delta W_k(\omega).$$

This scheme does not converge either strongly or weakly.

M. Hutzenthaler, A. Jentzen and P.E. Kloeden, *Strong and weak divergence in finite time of Euler's method for SDEs with non-globally Lipschitz coefficients,* (submitted).

Theorem 3. *The solution X_t of* (12) *and its Euler–Maruyama approximation $Y_k^{(N)}$ satisfy*

$$\lim_{N\to\infty} \mathbb{E}\left|X_1 - Y_N^{(N)}\right| = \infty. \tag{13}$$

Outline of proof

Let $N \in \mathbb{N}$ be arbitrary, define $r_N := \max\{3N, 2\}$ and consider the event

$$\Omega_N := \left\{ \omega \in \Omega \, : \, \sup_{k=1,\ldots,N-1} |\Delta W_k(\omega)| \leq 1, \quad |\Delta W_0(\omega)| \geq r_N \right\}.$$

Then, it follows by induction that

$$\left| Y_k^{(N)}(\omega) \right| \geq r_N^{2^{k-1}}, \qquad \forall \ \omega \in \Omega_N,$$

for every $k = 1, 2, \ldots, N$.

$$\begin{aligned}
\mathbb{P}[\Omega_N] &= \mathbb{P}\left[\sup_{k=1,\ldots,N-1} |\Delta W_k| \leq 1 \right] \cdot \mathbb{P}\left[|\Delta W_0| \geq r_N \right] \\
&\geq \mathbb{P}\left[\sup_{0 \leq t \leq 1} |W_t| \leq \frac{1}{2} \right] \cdot \mathbb{P}\left[|\Delta W_0| \geq r_N \right] \\
&= \mathbb{P}\left[\sup_{0 \leq t \leq 1} |W_t| \leq \frac{1}{2} \right] \cdot \mathbb{P}\left[\sqrt{N} \left| W_{1/N} \right| \geq \sqrt{N} r_N \right] \\
&\geq \mathbb{P}\left[\sup_{0 \leq t \leq 1} |W_t| \leq \frac{1}{2} \right] \cdot \frac{1}{4} \sqrt{N} r_N e^{-(\sqrt{N} r_N)^2} \\
&\geq \frac{1}{4} \cdot \mathbb{P}\left[\sup_{0 \leq t \leq 1} |W_t| \leq \frac{1}{2} \right] \cdot e^{-N r_N^2}
\end{aligned}$$

for every $N \in \mathbb{N}$. It follows that

$$\begin{aligned}
\lim_{N \to \infty} \mathbb{E}\left| Y_N^{(N)} \right| &\geq \frac{1}{4} \cdot \mathbb{P}\left[\sup_{0 \leq t \leq 1} |W_t| \leq \frac{1}{2} \right] \cdot \lim_{N \to \infty} e^{-N r_N^2} \cdot 2^{2^{N-1}} \\
&= \frac{1}{4} \cdot \mathbb{P}\left[\sup_{0 \leq t \leq 1} |W_t| \leq \frac{1}{2} \right] \cdot \lim_{N \to \infty} e^{-9N^3} \cdot 2^{2^{N-1}} = \infty.
\end{aligned}$$

Finally, since $\mathbb{E}|X_1|$ is finite,

$$\lim_{N \to \infty} \mathbb{E}\left| X_1 - Y_N^{(N)} \right| \geq \lim_{N \to \infty} \mathbb{E}\left| Y_N^{(N)} \right| - \mathbb{E}|X_1| = \infty. \qquad \square$$

Lecture 4: SODEs: Nonstandard Assumptions

There are various ways to overcome the problems caused by nonstandard assumptions on the coefficients of an SODE.

One way is to restrict attention to SODEs with special dynamical properties such as ergodicity, e.g., by assuming that the coefficients satisfy certain dissipativity and nondegeneracy conditions.

This yields the appropriate order estimates without bounded derivatives of coefficients.

However, several type of SODEs and in particular SODEs with square root coefficients remain a problem.

Many numerical schemes do not preserve the domain of the solution of the SODE and hence may crash when implemented, which has led to various ad hoc modifications to prevent this from happening.

Pathwise and Monte Carlo convergences often have to be used instead of strong and weak convergences.

1 SODEs without uniformly bounded coefficients

A localisation argument was used by

A. Jentzen, P.E. Kloeden and A. Neuenkirch, *Convergence of numerical approximations of stochastic differential equations on domains: higher order convergence rates without global Lipschitz coefficients*, Numerische Mathematik 112 (2009), no. 1, 41–64.

to show that the convergence theorem for strong Taylor schemes remains true for an SODE

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

when the coefficients satisfy

$$a, b \in C^{2\gamma+1}(\mathbb{R}^1;\mathbb{R}^1),$$

i.e., they do not necessarily have uniformly bounded derivatives.

The convergence obtained is pathwise. This is a special case of Theorem 1 below.

Pathwise convergence with order $\gamma - \epsilon$ where $\epsilon > 0$:

$$\sup_{i=0,\ldots,N} \left| X_{t_n}(\omega) - Y_n^{(\gamma)}(\omega) \right| \leq K_\epsilon^{(\gamma)}(\omega) \cdot \Delta^{\gamma-\epsilon}$$

for almost all $\omega \in \Omega$.

Whether a strong convergence rate can always be derived under these assumptions remains an open problem.

It applies, for example, to the stochastic Landau–Ginzburg equation to give convergence is pathwise of the Taylor schemes.

Empirical distributions for the random error constants of the Euler–Maruyama and Milstein schemes applied to the SODE

$$dX_t = -(1 + X_t)(1 - X_t^2)\, dt + (1 - X_t^2)\, dW_t, \qquad X(0) = 0,$$

on the time interval $[0, T]$ for $N = 10^4$ sample paths and time step $\Delta = 0.0001$ are plotted in Figure 1.

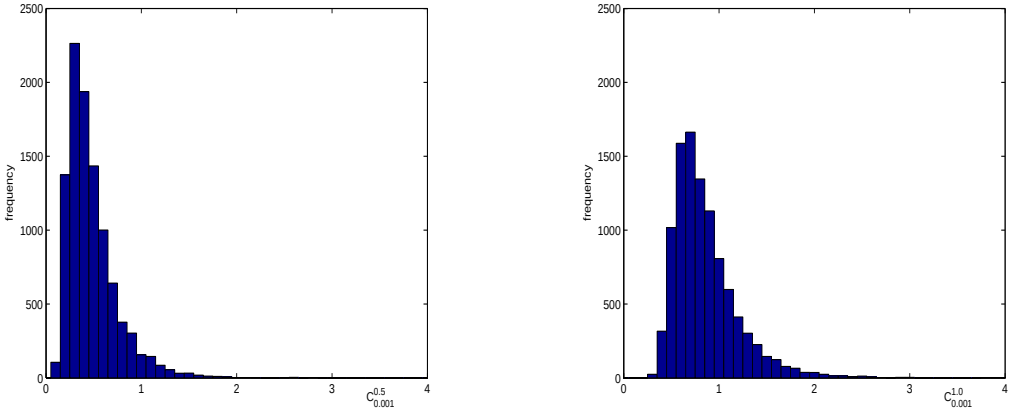

Figure 1: Empirical distributions of $K^{(0.5)}_{0.001}$ and $K^{(1.0)}_{0.001}$ (sample size: $N = 10^4$).

2 SODE on restricted regions

The Fisher–Wright SODE, the Feller diffusion with logistic growth SODE and the Cox–Ingersoll–Ross SODE have square root coefficients, which require the solutions to remain in the region where the expression under the square root is non–negative.

However, numerical iterations may leave this restricted region, in which case the algorithm will terminate.

One way to avoid this problem is to use appropriately modified Itô–Taylor schemes.

Consider an SODE

$$dX_t = a(X_t)\,dt + b(X_t)\,dW_t, \tag{1}$$

where X_t takes values in a domain $D \subset \mathbb{R}^1$ for $t \in [0,T]$. Suppose that the coefficients a and b are r-times continuously differentiable on D and that the SODE (1) has a unique strong solution. Define

$$E := \{x \in \mathbb{R}^1 \,:\, x \notin \overline{D}\}.$$

Then choose auxiliary functions f, $g \in C^s(E;\mathbb{R}^1)$ for $s \in \mathbb{N}$ and define

$$\begin{aligned} \widetilde{a}(x) &= a(x)\cdot \mathbb{I}_D(x) + f(x)\cdot \mathbb{I}_E(x), \qquad x \in \mathbb{R}^1, \\ \widetilde{b}(x) &= b(x)\cdot \mathbb{I}_D(x) + g(x)\cdot \mathbb{I}_E(x), \qquad x \in \mathbb{R}^1. \end{aligned}$$

In addition, for $x \in \partial D$ define

$$\widetilde{a}(x) = \lim_{y\to x;\, y\in D} \widetilde{a}(y), \qquad \widetilde{b}(x) = \lim_{y\to x;\in y\in D} \widetilde{b}(y),$$

and the "modified" derivative of a function $h : \mathbb{R}^1 \to \mathbb{R}^1$ by

$$\partial_{x^l} h(x) = \frac{\partial}{\partial x^l} h(x), \quad x \in D \cup E,$$

$$\partial_{x^l} h(x) = \lim_{y \to x;\, y \in D} \partial_{x^l} h(x), \qquad l = 1, \ldots, d, \quad x \in \partial D.$$

If the above limits do not exist, set $\widetilde{a}(x) = 0$, $\widetilde{b}(x) = 0$ and $\partial_{x^l} h(x) = 0$ for $x \in \partial D$.

A modified Itô–Taylor scheme is the corresponding Itô–Taylor scheme for the SODE with modified coefficients

$$dX_t = \widetilde{a}(X_t)\, dt + \widetilde{b}(X_t)\, dW_t, \tag{2}$$

The purpose of the auxiliary functions is twofold:

- to obtain a well defined approximation scheme
- to "reflect" the numerical scheme back into D after it has left D.

Theorem 1. *Assume that* $\widetilde{a}$, $\widetilde{b} \in C^{2\gamma+1}(D; \mathbb{R}^1) \bigcap C^{2\gamma-1}(E; \mathbb{R}^1)$.
Then, for every $\epsilon > 0$ *and* $\gamma = \frac{1}{2}, 1, \frac{3}{2}, \ldots$ *there exists a non–negative random variable* $K_{\gamma,\epsilon}^{(f,g)}$ *such that*

$$\sup_{n=0,\ldots,N_T} \left| X_{t_n}(\omega) - Y_n^{(mod,\gamma)}(\omega) \right| \leq K_{\gamma,\epsilon}^{(f,g)}(\omega) \cdot \Delta^{\gamma-\epsilon}$$

for almost all $\omega \in \Omega$ *and all* $n = 1, \ldots, N_T$, *where* $\Delta = T/N_T$ *and the* $Y_n^{(mod,\gamma)}$ *correspond to the modified Itô–Taylor scheme applied to the SODE (2).*

The convergence rate does not depend on the choice of the auxiliary functions, but the random constant in the error bound clearly does.

3 Examples

Example 1

Consider the Cox–Ingersoll–Ross SODE

$$\boxed{dX_t = \kappa\,(\lambda - X_t)\,dt + \theta\sqrt{X_t}\,dW_t}$$

with $\kappa\lambda \geq \theta^2/2$ and $x(0) = x_0 > 0$.

Here $D = (0, \infty)$ and the coefficients

$$a(x) = \kappa\,(\lambda - x)\,, \qquad b(x) = \theta\sqrt{x}, \qquad x \in D,$$

satisfy a, $b \in C^\infty(D; \mathbb{R}^1)$.

As auxiliary functions on $E = = (-\infty, 0)$ choose, e.g.,

$$f(x) = g(x) = 0, \qquad x \in E,$$

or

$$f(x) = \kappa\,(\lambda - x)\,, \qquad g(x) = 0, \qquad x \in E.$$

The first choice of auxiliary functions "kills" the numerical approximation as soon as it reaches a negative value.

However, the second is more appropriate, since if the scheme would take a negative value, the auxiliary functions force the numerical scheme to be positive again after the next steps, which recovers better the positivity of the exact solution.

Example 2

A general version of the Wright–Fisher SODE

$$dX_t = f(X_t)\,dt + \sqrt{X_t(1-X_t)}\,dW_t \qquad (3)$$

typically has a polynomial drift coefficient f.

Its solution should take values in the interval $[0,1]$, but, in general, depending on the structure of f, the solution can attain the boundary $\{0,1\}$ in finite time, i.e.,

$$\tau_{\{0,1\}} = \inf\{t \in [0,T] :\ X_t \notin (0,1)\} < T$$

almost surely. Thus, Theorem 1 cannot be applied directly to the SODE (3).

But a modified Itô–Taylor method $Y_n^{(mod,\gamma)}$ with the auxiliary functions $f = 0$ and $g = 0$ to (3) can be used up to the first hitting time of the boundary of D.

The error bound then takes the form

$$\sup_{i=0,\ldots,n} \left| Y_{t_i}(\omega) - Y_i^{(mod,\gamma)}(\omega) \right| \leq \eta_{\gamma,\epsilon}(\omega) \cdot n^{-\gamma+\epsilon}$$

for almost all $\omega \in \Omega$ and all $n \in \mathbb{N}$, where

$$Y_{t_i}(\omega) = X_t(\omega)\,\mathbb{I}_{\{X_t(\omega)\in(0,1)\}}, \qquad t \geq 0,\ \omega \in \Omega.$$

4 Monte Carlo convergence

The concepts of strong and weak convergence of a numerical scheme are theoretical discretisation concepts.

In practice, one has to estimate the expectations by a finite random sample.

In the weak case, neglecting roundoff and other computational errors, one uses in fact the convergence

$$\lim_{N,M\to\infty}\left|\mathbb{E}\Big[g\big(X_T\big)\Big]-\frac{1}{M}\sum_{k=1}^{M}g\big(Y_N^{(N)}(\omega_k)\big)\right|=0 \tag{4}$$

for smooth functions $g:\mathbb{R}\to\mathbb{R}$ with derivatives having at most polynomial growth.

By the triangle inequality

$$\begin{aligned}&\left|\mathbb{E}\Big[g\big(X_T\big)\Big]-\frac{1}{M}\sum_{k=1}^{M}g\big(Y_N^{(N)}(\omega_k)\big)\right|\\ &\quad\leq\ \left|\mathbb{E}\Big[g\big(X_T\big)\Big]-\mathbb{E}\Big[g\big(Y_N^{(N)}\big)\Big]\right|+\left|\mathbb{E}\Big[g\big(Y_N^{(N)}\big)\Big]-\frac{1}{M}\sum_{k=1}^{M}g\big(Y_N^{(N)}(\omega_k)\big)\right|.\end{aligned}$$

where the first and second summands on the right hand side are, respectively,

- the weak discretisation error due to approximating the exact solution with numerical

- the statistical error due to approximating an expectation with the arithmetic average of finitely many independent samples.

Thus, if the numerical scheme converges weakly, then it also converges in the above sense (4), which was called Monte Carlo convergence in

M. Hutzenthaler and A. Jentzen, *Convergence of the stochastic Euler scheme for locally Lipschitz coefficients,* (submitted).

Monte Carlo convergence often holds for the Euler–Maruyama scheme applied to a scalar SODE such as

$$dX_t = -X_t^3\,dt + dW_t, \qquad X_0 = 0,$$

for which neither strong nor weak convergence holds.

The sample mean $\frac{1}{M}\sum_{k=1}^{M} g\big(Y_N^{(N)}(\omega_k)\big)$ in (4) is, in fact, a random variable, so the Monte Carlo convergence (4) should be interpreted as holding almost surely.

To formulate this in a succinct way, the sample paths of M independent Wiener processes $W_t^{(1)}(\omega)$, …, $W_t^{(M)}(\omega)$ for the same ω will be considered instead of M different sample paths $W_t(\omega_1)$, …, $W_t(\omega_M)$ of the same Wiener process W_t.

• The choice $M = N^2$ ensures that Monte Carlo convergence for the Euler–Maruyama scheme has the same order as that for weak convergence under standard assumptions, namely 1, since the Monte Carlo simulation of $\mathbb{E}\big[g\big(Y_N^{(N)}\big)\big]$ with M independent Euler approximations has convergence order $\frac{1}{2} - \epsilon$ for an arbitrarily small $\epsilon > 0$.

With these modifications, Monte Carlo convergence takes the form

$$\lim_{N\to\infty} \left| \mathbb{E}\Big[g\big(X_T\big)\Big] - \frac{1}{N^2}\sum_{k=1}^{N^2} g\big(Y_N^{(N,k)}(\omega)\big) \right| = 0, \qquad \text{a.s.,} \qquad (5)$$

where $Y_N^{(N,k)}(\omega)$ is the ω-realisation of the Nth iterate of the Euler–Maruyama scheme applied to the SODE with the ω-realisation of the kth Wiener process $W_t^{(k)}(\omega)$.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

Theorem 2. *Suppose that a, b, $g : \mathbb{R} \to \mathbb{R}$ are four times continuously differentiable functions with derivatives satisfying*

$$\left|a^{(n)}(x)\right| + \left|b^{(n)}(x)\right| + \left|g^{(n)}(x)\right| \leq L\left(1 + |x|^r\right), \quad \forall\, x \in \mathbb{R},$$

for $n = 0,1,\ldots, 4$, where $L \in (0,\infty)$ and $r \in (1,\infty)$ are fixed constants. Moreover, suppose that the drift coefficient a satisfies the global one–sided Lipschitz condition

$$(x-y)\cdot(a(x)-a(y)) \leq L\,(x-y)^2, \quad \forall\, x, y \in \mathbb{R},$$

and that the diffusion coefficient satisfies the global Lipschitz condition

$$|b(x)-b(y)| \leq L\,|x-y|, \quad \forall\, x, y \in \mathbb{R}.$$

Then, there is $\mathcal{F}$-measurable mappings $C_\varepsilon : \Omega \to [0,\infty)$ for each $\varepsilon \in (0,1)$ and an event $\tilde{\Omega} \in \mathcal{F}$ with $\mathbb{P}[\tilde{\Omega}] = 1$ such that

$$\left|\mathbb{E}\Big[g(X_T)\Big] - \frac{1}{N^2}\left(\sum_{m=1}^{N^2} g(Y_N^{(N,k)}(\omega))\right)\right| \leq C_\varepsilon(\omega)\cdot\frac{1}{N^{1-\varepsilon}}$$

for every $\omega \in \tilde{\Omega}$, $N \in \mathbb{N}$ and $\varepsilon \in (0,1)$, where X_t is the solution of the SODE

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

and $Y_N^{(N,k)}$ is the Nth iterate of the Euler–Maruyama scheme applied to this SODE with the Wiener process $W_t^{(k)}$ for $k = 1, \ldots, N^2$.

Lecture 5: Stochastic Partial Differential Equations

The stochastic partial differential equations (SPDEs) considered here are stochastic evolution equations of the parabolic or hyperbolic types.

There is an extensive literature on SPDEs.

P.L. Chow, *Stochastic Partial Differential Equations*, Chapman & Hall/CRC, Boca Raton, 2007.

G. Da Prato and G. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Cambridge University Press, Cambridge, 1992.

W. Grecksch and C. Tudor, *Stochastic Evolution Equations. A Hilbert Space Approach*, Akademie–Verlag, Berlin, 1995.

N.V. Krylov and B.L. Rozovskii, *Stochastic Evolution Equations*, World Sci. Publ., Hackensack, N.J., 2007.

C. Prévot and M. Röckner, *A Concise Course on Stochastic Partial Differential Equations*, Springer–Verlag, Berlin, 2007.

The theory of such SPDEs is complicated by different types of solution concepts and function spaces depending on the spatial regularity of the driving noise process.

1 Random and stochastic PDEs

As with RODEs and SODEs, one can distinguish between random and stochastic partial differential equations.

Attention is restricted here to parabolic reaction–diffusion type equations on a bounded spatial domain $\mathcal{D}$ in $\mathbb{R}^d$ with smooth boundary $\partial\mathcal{D}$ with a Dirichlet boundary condition.

An example of a random PDE (RPDE) is

$$\frac{\partial u}{\partial t} = \Delta u + f(\zeta_t, u), \qquad u\big|_{\partial\mathcal{D}} = 0,$$

where ζ_t is a stochastic process (possibly infinite dimensional). This is interpreted and analysed pathwise as a deterministic PDE.

An example of an Itô stochastic PDE (SPDE) is

$$dX_t = [\Delta X_t + f(X_t)]\, dt + g(X_t)\, dW_t, \qquad X_t\big|_{\partial\mathcal{D}} = 0, \quad (1)$$

where W_t an infinite dimensional Wiener process of the form

$$W_t(x,\omega) = \sum_{k=1}^{\infty} c_k W_t^k(\omega)\phi_k(x), \qquad t \geq 0, x \in \mathcal{D},$$

with pairwise independent scalar Wiener processes W_t^k, $k \in \mathbb{N}$, and an orthonormal basis system $(\phi_k)_{k\in\mathbb{N}}$ of some function space, e.g., $L^2(\mathcal{D})$.

As for SODEs, the theory of SPDEs is a mean–square theory and requires an infinite dimensional version of Itô stochastic calculus.

The Doss–Sussmann theory is not as well developed for SPDEs as for SODEs, but in simple cases an SPDE can be transformed to an RPDE.

For example, the SPDE (1) with additive noise

$$dX_t = [\Delta X_t + f(X_t)]\ dt + \ dW_t, \qquad X_t\big|_{\partial\mathcal{D}} = 0,$$

is equivalent to the RPDE

$$\frac{\partial v}{\partial t} = \Delta v + f(v + O_t) + O_t$$

with $v(t) = X_t - O_t$, $t \geq 0$, where O_t is the (infinite-dimensional) Ornstein–Uhlenbeck stochastic stationary solution of the linear SPDE

$$dO_t = [\Delta O_t - O_t]\ dt + dW_t, \qquad O_t\big|_{\partial\mathcal{D}} = 0.$$

As a specific example, the RPDE with a scalar Ornstein–Uhlenbeck process,

$$\frac{\partial v}{\partial t} = \frac{\partial^2 v}{\partial x^2} - v - (v + O_t)^3$$

on the spatial interval $0 \leq x \leq 1$ with Dirichlet boundary conditions is equivalent to the SPDE with additive noise

$$dX_t = \left[\frac{\partial^2}{\partial x^2} X_t - X_t - X_t^3\right]\ dt + dW_t$$

on $\mathcal{D} = [0, 1]$ with Dirichlet boundary conditions and a scalar Wiener process.

2 Mild solutions of SPDEs

An Itô stochastic partial differential equation (SPDE)

$$dX_t = [AX_t + F(X_t)]\, dt + B(X_t)\, dW_t \tag{2}$$

on a Hilbert space H, where

- A is, in general, an unbounded linear operator, e.g., the Laplace operator Δ with the Dirichlet boundary condition
- $(W_t)_{t\in\mathbb{R}^+}$ is an infinite dimensional cylindrical Wiener process,

is a stochastic integral equation

$$X_t = x_0 + \int_0^t [AX_s + F(X_s)]\, ds + \int_0^t B(X_s)\, dW_s$$

on H, where the first integral is pathwise a deterministic integral and the second an Itô stochastic integral in H.

There are several different interpretations of the stochastic integral equation (2) in the literature.

The mild form is used here since it is better suited for the derivation of Taylor expansions and numerical schemes.

The Itô stochastic integrals here and below are defined analogously to the finite dimensional case.

The mild form of the SPDE (2) is also a stochastic integral equation in H

$$X_t = e^{At}x_0 + \int_0^t e^{A(t-s)}F(X_s)\,ds + \int_0^t e^{A(t-s)}B(X_s)\,dW_s, \quad \text{a.s.,} \qquad (3)$$

where $\left(e^{At}\right)_{t\geq 0}$ is a semigroup of solution operators of the deterministic ODE/PDE

$$\frac{dX}{dt} = AX \qquad \Leftrightarrow \qquad \frac{\partial u}{\partial t} = \Delta u, \quad u\big|_{\partial\mathcal{D}} = 0,$$

on H, i.e., $e^{At} = S_t$ for $t \geq 0$, where $X(t) = S_t(x_0)$ is the solution with Dirichlet boundary conditions for the initial value $X(0) = x_0$.

In the finite dimensional case, $H = \mathbb{R}^d$, the SPDE is an SODE, A is a $d \times d$ matrix and e^{At} is a matrix exponential.

Define the L^q-norm of of a random variable $Z : \Omega \to U$, where U is a Hilbert space, for $q \geq 1$ by

$$|Z|_{L^q} := (\mathbb{E}|Z|_U^q)^{\frac{1}{q}}$$

The following is a version of the Burkholder–Davis–Gundy inequality in infinite dimensions

Lemma 1. *Let $(\Gamma_t)_{t\in[0,T]}$ be a predictable stochastic process, whose values are Hilbert–Schmidt operators from U to H with $\mathbb{E}\int_0^T \|\Gamma_s\|_{HS}^2\,ds < \infty$. Then,*

$$\left|\int_0^t \Gamma_s\,dW_s\right|_{L^q} \leq q\left(\int_0^t \left|\|\Gamma_s\|_{HS}\right|_{L^q}^2 ds\right)^{\frac{1}{2}}$$

for every $t \in [0,T]$ and every $q \geq 2$. (Both sides could be infinite).

3 Function space setting

Let $(H, \langle \cdot, \cdot \rangle)$ and $(U, \langle \cdot, \cdot \rangle)$ be two separable Hilbert spaces with norms $|\cdot| = |\cdot|_H$ and $|\cdot|_U$.

Let $(D, |\cdot|_D)$ be a separable Banach space with $H \subset D$ continuously.

Let $L(U, D)$ be the space of all bounded linear operators Γ from U to D. Then $L(U, D)$ is a *Banach space* with the *operator norm* $\| \cdot \|$.

The space $L_{HS}(U, D)$ of Hilbert–Schmidt operators Γ from U to D is the subspace of $L(U, D)$ consisting of bounded linear operators with the finite Hilbert–Schmidt norm

$$\|\Gamma\|_{HS} := \left(\sum_{k=1}^{\infty} |\Gamma u_k|_D^2 \right)^{1/2} < \infty,$$

where $(u_k)_{k \in \mathbb{N}}$ is a complete orthonormal basis of U.

The space $L_N(U, D)$ of nuclear operators Γ from U to D is the subspace of $L_{HS}(U, D)$ consisting of bounded linear operators with finite trace norm

$$\|\Gamma\|_N := \operatorname{Tr} \Gamma^* \Gamma = \sum_{k=1}^{\infty} \langle \Gamma^* \Gamma u_k, u_k \rangle_U < \infty,$$

where Γ^* is the adjoint of Γ and $(u_k)_{k \in \mathbb{N}}$ is a complete orthonormal basis of U.

4 Infinite dimensional Wiener processes

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a normal filtration $(\mathcal{F}_t)_{t\geq 0}$. and let W_t^k, $k \geq 1$, be pairwise independent scalar Wiener processes that are all adapted to the filtration $\mathcal{F}_t$.

Let $Q \in L(U, U)$ be a symmetric and non–negative operator with $\operatorname{Tr} Q < \infty$.

$\Rightarrow$ there exists a complete orthonormal basis system $(u_k)_{k\in\mathbb{N}}$ of U and a bounded sequence of non–negative real numbers λ_k such that $Qu_k = \lambda_k u_k$ for $k \in \mathbb{N}$.

Then the infinite sequence

$$W_t := \sum_{k=1}^{\infty} \sqrt{\lambda_k} W_t^k u_k \qquad (4)$$

converges in the space $(U, |\cdot|_U)$ and has the properties

$$\mathbb{E}W_t = 0, \qquad \operatorname{Cov} W_t = t\,Q, \qquad t \geq 0.$$

It is called a Q-Wiener process or cylindrical Q-Wiener process on U with respect to the filtration $\mathcal{F}_t$ and Q is called its covariance operator.

W_t is called a cylindrical I-Wiener process on U with respect to $\mathcal{F}_t$ when the covariance operator $Q = I$, the identity operator on U.

Then $\operatorname{Tr} Q = \infty$ and the infinite series (4) does not converge in U, but it it does converge in a larger space with a weaker topology.

5 Assumptions

The coefficient terms A, F and B of the SPDE (2) and the stochastic integral equation (3) are assumed to satisfy the following assumptions.

Assumption 1. (Linear Operator A) *Let $(\lambda_i)_{i\in\mathcal{I}}$ be a family of positive real numbers with $\inf_{i\in\mathcal{I}} \lambda_i > 0$ and let $(e_i)_{i\in\mathcal{I}}$ be an orthonormal basis of H, where $\mathcal{I}$ is a finite or countable set.*
The linear operator $A : D(A) \subset H \to H$ is given by

$$Av = \sum_{i\in\mathcal{I}} -\lambda_i \langle e_i, v\rangle e_i$$

for all $v \in D(A) = \left\{v \in H : \sum_{i\in\mathcal{I}} |\lambda_i|^2 |\langle e_i, v\rangle|^2 < \infty\right\}$.

Assumption 2. (Drift F) *The mapping $F : H \to H$ is global Lipschitz continuous with respect to $|\cdot|_H$.*

Assumption 3. (Diffusion B) *The embedding $D \subset D((-A)^{-r})$ is continuous for some $r \geq 0$ and $B : H \to L(U, D)$ is a measurable mapping such that $e^{At}B(v)$ is a Hilbert–Schmidt operator from U to H and*

$$\left\|e^{At}B(v)\right\|_{HS} \leq L\left(1 + |v|_H\right) t^{\varepsilon-\frac{1}{2}}, \quad \left\|e^{At}\left(B(v) - B(w)\right)\right\|_{HS} \leq L|v-w|_H t^{\varepsilon-\frac{1}{2}}$$

for all v, $w \in H$ and $t \in (0, T]$, where $L > 0$ and $\varepsilon > 0$ are given constants.

Here $D((-A)^r)$ with $r \in \mathbb{R}$ is the interpolation space of powers of the operator $-A$ and $\|\cdot\|_{HS}$ the Hilbert–Schmidt norm for Hilbert–Schmidt operators from U to H.

Assumption 4. (Initial value) *The initial value is an $\mathcal{F}_0$-measurable random variable $x_0 : \Omega \to H$ with $\mathbb{E}|x_0|_H^p < \infty$ for some $p \geq 2$.*

6 Existence and uniqueness of mild solutions

The literature contains many existence and uniqueness theorems for mild solutions of SPDEs, such as Theorems 7.4 and 7.6 in Da Prato & Zabczyk which treat the space–time white noise (cylindrical I-Wiener process) and trace–class noise (cylindrical Q-Wiener process) cases separately under different assumptions.

For example, something is assumed about e^{At} (see equations (7.27) and (7.28) in DP & Z) in the space–time white noise case and on B (see equation (7.5) in DP & Z) in the trace–class noise case.

In contrast, Assumption 3, which postulates something on the mapping $v \mapsto e^{At}B(v)$ satisfies a linear growth bound and a global Lipschitz condition with respect to the Hilbert–Schmidt norm for each $t > 0$ with the constants depending of a fractional power of t.

This allows the space–time white noise and trace–class noise cases to be combined in a single setting.

The theorems here are taken from

A. Jentzen and P. E. Kloeden, *A unified existence and uniqueness theorem for stochastic evolution equations*, Bull. Austral. Math. Soc. **100** (2010), 33–46.

Theorem 1. *Let Assumptions 1–4 be satisfied and let* $(W_t)_{t\in[0,T]}$ *be a cylindrical* I*-Wiener process in* U*.*

Then there is a unique (up to modifications) predictable stochastic process $X : \Omega \times [0,T] \to H$ *with* $\sup_{0\le t\le T} \mathbb{E}|X_t|_H^p < \infty$*, where* $p \ge 2$ *is given in Assumption 4, such that*

$$\mathbb{P}\left[X_t = e^{At}x_0 + \int_0^t e^{A(t-s)}F(X_s)\,ds + \int_0^t e^{A(t-s)}B(X_s)\,dW_s\right] = 1 \tag{5}$$

for all $t \in [0,T]$*.* X *is the unique mild solution of the SPDE* (2) *in this sense.*

The integrals in equation (5) are well defined under the Assumptions 1–4.

The following regularity property of the solution holds if further assumptions are made on $e^{At}B$.

Theorem 2. *Let Assumptions 1–4 be satisfied and let* $\gamma \in (0,1)$ *be such that* $\mathbb{E}\,|(-A)^\gamma x_0|_H^p < \infty$*. Furthermore, suppose that* $(-A)^\gamma e^{At}B(v)$ *is a Hilbert–Schmidt operator from* U *to* H *with*

$$\left\|(-A)^\gamma e^{At}B(v)\right\|_{HS} \le L\,(1+|v|_H)\,t^{\varepsilon-\frac{1}{2}}$$

for all $v \in H$ *and all* $t \in (0,T]$ *with constants* $L > 0$ *and* $\varepsilon > 0$*. Then the unique solution process* $X : \Omega \times [0,T] \to H$ *of the SPDE* (2) *given by Theorem 1 satisfies* $\sup_{0\le t\le T} \mathbb{E}\,|(-A)^\gamma X_t|_H^p < \infty$*.*

Sketch proof of Theorem 1

Introduce the real vector space $\mathcal{V}_p$ of all equivalence classes of predictable stochastic processes $X : \Omega \times [0,T] \to H$ with $\sup_{0\le t\le T} |X_t|_{L^p} < \infty$.

Equip this space with the norm

$$\boxed{\|X\|_\mu := \sup_{0\le t\le T} e^{\mu t} |X_t|_{L^p}}$$

for every $X \in \mathcal{V}_p$ and some $\mu \in \mathbb{R}$.

Note that the pair $\left(\mathcal{V}_p, \|\cdot\|_\mu\right)$ is a Banach space for any $\mu \in \mathbb{R}$.

Define the mapping $\Phi : \mathcal{V}_p \to \mathcal{V}_p$ by

$$\boxed{(\Phi X)_t := e^{At}u_0 + \int_0^t e^{A(t-s)}F(X_s)ds + \int_0^t e^{A(t-s)}B(X_s)dW_s}$$

for every $t \in [0,T]$ and $X \in \mathcal{V}_p$.

First, it needs to be shown that Φ is well defined and then that it is a contraction with respect to $\|\cdot\|_\mu$ for an appropriate $\mu \in \mathbb{R}$. Now

$$\begin{aligned}
\|\Phi X - \Phi Y\|_\mu &\le \frac{K}{|\mu|}\|X-Y\|_\mu + Lp\left(\int_0^t s^{2\varepsilon-1}e^{2\mu s}\,ds\right)^{\frac{1}{2}}\|X-Y\|_\mu \\
&\le \left(\frac{K}{|\mu|} + Lp\sqrt{\int_0^T s^{2\varepsilon-1}e^{2\mu s}\,ds}\right)\|X-Y\|_\mu
\end{aligned}$$

for $\mu < 0$.

Hence Φ is a contraction with respect to $\|\cdot\|_\mu$ for $\mu \ll 0$ and it has a unique fixed point $X \in \mathcal{V}_p$, which is the desired solution. $\square$

7 Examples: common setup

Two examples in which the above assumptions hold will be considered.

These have the following common set up for the linear operator A and the drift coefficient F, but have different types of noise.

Let $\mathcal{D} := (0,1)^d \subset \mathbb{R}^d$ with $d \geq 1$ and let $H = L^2(\mathcal{D}, \mathbb{R})$ be the Hilbert space of all square integrable functions from $\mathcal{D}$ to $\mathbb{R}$ with the scalar product and the norm

$$\langle v, w \rangle = \int_{\mathcal{D}} v(x) w(x)\, dx, \qquad |v|_H = \left(\int_{\mathcal{D}} v(x)^2\, dx \right)^{\frac{1}{2}}$$

for all v, $w \in H$. Also define $U := H$.

Let $A = \vartheta \Delta$, i.e., a constant $\vartheta > 0$ times the Laplacian with Dirichlet boundary conditions.

The eigenfunctions and eigenvalues of linear operator $-A$ are

$$e_i(x) = 2^{\frac{d}{2}} \sin(i_1 \pi x_1) \cdots \sin(i_d \pi x_d), \qquad \lambda_i = \vartheta \pi^2 \left(i_1^2 + \ldots + i_d^2 \right)$$

for all $x = (x_1, \ldots, x_d) \in \mathcal{D}$ and all $i = (i_1, \ldots, i_d) \in \mathcal{I} := \mathbb{N}^d$.

The linear operator A has the representation

$$Af = \sum_{i \in \mathcal{I}} -\lambda_i \langle e_i, f \rangle e_i$$

for all $f \in D(A) = \left\{ f \in H \, : \, \sum_{i \in \mathcal{I}} \lambda_i^2 \left| \langle e_i, f \rangle \right|^2 < \infty \right\}$.

Thus Assumption 1 holds.

Furthermore, let f, $g : \mathbb{R} \to \mathbb{R}$ be globally Lipschitz continuous functions, i.e.,

$$|g(x) - g(y)| \leq L|x - y|, \qquad |f(x) - f(y)| \leq L|x - y|$$

for all x, $y \in \mathbb{R}$ with a constant $L > 0$.

Define the corresponding Nemytskii operators

$$F : H \to H, \qquad F(v)(x) = f(v(x)), \qquad x \in (0,1)^d, \tag{6}$$

$$G : H \to H, \qquad G(v)(x) = g(v(x)), \qquad x \in (0,1)^d, \tag{7}$$

for all $v \in H$.

Hence, F and G are also globally Lipschitz continuous functions on H, i.e.,

$$|G(v) - G(w)|_H \leq L\,|v - w|_H\,, \qquad |F(v) - F(w)|_H \leq L\,|v - w|_H$$

for all v, $w \in H$, and Assumption 2 holds.

Assumption 3 will be verified separately for the two cases of space–time white noise and trace–class noise.

8 Example 1: Space–time white noise

Remark: It was shown by WALSH that mild solutions do not exist for space–time white noise in spatial domains of dimension higher than one.

Let $d = 1$, so $\mathcal{D} := (0, 1)$ and $D = L^1(0, 1)$. Define B by

$$B : H \to L(H, D), \qquad (B(v)(w))(x) := (G(v))(x) \cdot w(x) \tag{8}$$

for every $x \in (0, 1)$ and $v, w \in H$.

Then B is well defined, since by the Cauchy–Schwarz inequality

$$\begin{aligned} |B(v)(w)|_D &= \int_0^1 |G(v)(x) \cdot w(x)| \, dx \\ &\leq \left(\int_0^1 |G(v)(x)|^2 \, dx \right)^{\frac{1}{2}} \left(\int_0^1 |w(x)|^2 \, dx \right)^{\frac{1}{2}} = |G(v)|_H \cdot |w|_H \end{aligned}$$

for all v, $w \in H$.

In particular, $B(v)$ is a bounded linear operator from H to D with the property

$$\|B(v)\|_{L(H,D)} \leq |G(v)|_H$$

for all $v \in H$. In the same way, it follows that

$$\|B(v) - B(u)\|_{L(H,D)} \leq |G(v) - G(u)|_H \leq L \, |v - u|_H$$

for all $v, u \in H$, since G is globally Lipschitz continuous.

Hence B is also a globally Lipschitz continuous function from H to $L(H, D)$ and thus measurable.

Combining the definitions in equations (6)–(7) and (8) gives for every $v, w \in H$

$$B : H \to L(H, D), \quad (B(v)(w))(x) := g(v(x)) \cdot w(x) \quad \text{for } x \in (0,1).$$

In the next step, let $\gamma \geq 0$. Then $(-A)^\gamma e^{At} B(v)$ is a bounded linear operator from H to H for every $v \in H$ and $t \in (0, T]$, since

$$\begin{aligned} \left\|(-A)^\gamma e^{At} B(v)\right\|_{HS} &= \left(\sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{I}} \left(\lambda_j^{2\gamma} e^{-2\lambda_j t} \left| \langle e_j, B(v) e_i \rangle \right|^2 \right) \right)^{\frac{1}{2}} \\ &\leq \left(\sum_{j \in \mathcal{I}} \left(\lambda_j^{2\gamma} e^{-2\lambda_j t} \, 2 \left| G(v) \right|_H^2 \right) \right)^{\frac{1}{2}} = \sqrt{2} \left| G(v) \right|_H \left\|(-A)^\gamma e^{At}\right\|_{HS} \end{aligned}$$

for all $v \in H$, $t \in (0, T]$.

Now suppose that $\gamma \in [0, \frac{1}{4})$. Then,

$$\begin{aligned} \left\|(-A)^\gamma e^{At}\right\|_{HS}^2 &= \sum_{j=1}^{\infty} \lambda_j^{2\gamma} e^{-2\lambda_j t} \\ &= \sum_{j=1}^{\infty} (\vartheta j^2 \pi^2)^{2\gamma} e^{-2\vartheta j^2 \pi^2 t} \leq \left(\frac{2(T+2)}{\min(\vartheta, 1)} \right)^2 t^{-\frac{1}{2} - 2\gamma} \end{aligned}$$

from which it follows that

$$\left\|(-A)^\gamma e^{At} B(v)\right\|_{HS} \leq \left(\frac{4(T+2)}{\min(\vartheta, 1)} \right) \left| G(v) \right|_H \, t^{-\frac{1}{4} - \gamma}$$

for every $v \in H$, $t \in (0, T]$ and $\gamma \in [0, \frac{1}{4})$.

In the same way, it can be shown that

$$\left\|(-A)^{\gamma} e^{At} \left(B(v) - B(u)\right)\right\|_{HS} \leq \left(\frac{4(T+2)}{\min(\vartheta, 1)}\right) |G(v) - G(u)|_H \; t^{-\frac{1}{4}-\gamma}$$

for every v, $u \in H$, $t \in (0, T]$ and $\gamma \in [0, \frac{1}{4})$. Thus Assumption 3 holds.

Finally, if the initial value satisfies Assumption 4, then the SPDE

$$\boxed{dX_t = [\Delta X_t + f(X_t)] \; dt + g(X_t) \, dW_t,}$$

on the domain $\mathcal{D} = (0, 1)$ has a unique mild solution by Theorem 1.

In addition, if the initial value satisfies $\mathbb{E}\,|(-A)^{\gamma} x_0|_H^p < \infty$ with $\gamma \in [0, \frac{1}{4})$, then the solution has almost surely values in $D((-A)^{\gamma})$ in the sense of Theorem 2.

In particular, the SPDE with additive noise

$$dX_t = [\Delta X_t + f(X_t)] \; dt + dW_t, \qquad g(y) \equiv 1,$$

and the stochastic heat equation with linear multiplicative noise

$$dX_t = \Delta X_t dt + X_t \, dW_t, \qquad g(y) \equiv y, \; f(y) \equiv 0,$$

both have unique mild solutions.

9 Example 2: Trace–class noise

Let $d \geq 1$, $D = H$ and let $(f_i)_{i\in\mathcal{I}}$ be another orthonormal basis in H with the property that $f_i : \bar{\mathcal{D}} \to \mathbb{R}$ are continuous functions, which satisfy

$$\sup_{i\in\mathcal{I}} \sup_{x\in\bar{\mathcal{D}}} |f_i(x)| < \infty.$$

Moreover, let $\sqrt{Q} : H \to H$ be a bounded linear operator given by

$$\sqrt{Q}v = \sum_{i\in\mathcal{I}} \alpha_i f_i \langle f_i, v\rangle$$

for real numbers $(\alpha_i)_{i\in\mathcal{I}}$ satisfying $\sum_{i\in\mathcal{I}} \alpha_i^2 < \infty$.

Then $\sqrt{Q}v : \bar{\mathcal{D}} \to H$ is a continuous mapping with

$$\begin{aligned}
\sup_{x\in\bar{\mathcal{D}}} \left| \left(\sqrt{Q}v\right)(x) \right| &\leq \sum_{i\in I} \left(|\alpha_i| \cdot |\langle f_i, v\rangle| \cdot \sup_{x\in\bar{\mathcal{D}}} |f_i(x)| \right) \\
&\leq \left(\sum_{i\in I} \alpha_i^2 \right)^{\frac{1}{2}} \left(\sum_{i\in I} |\langle f_i, v\rangle|^2 \right)^{\frac{1}{2}} \left(\sup_{i\in\mathcal{I}} \sup_{x\in\bar{\mathcal{D}}} |f_i(x)| \right) \\
&\leq \left(\sum_{i\in I} \alpha_i^2 \right)^{\frac{1}{2}} \left(\sup_{i\in\mathcal{I}} \sup_{x\in\bar{\mathcal{D}}} |f_i(x)| \right) |v|_H = c\,|v|_H .
\end{aligned}$$

In the next step, define B by

$$B : H \to L(H, D), \qquad (B(v)(w))(x) := (G(v))(x) \cdot \left(\sqrt{Q}w\right)(x) \tag{9}$$

for every $x \in (0,1)^d$ and v, $w \in H$.

Then B is indeed well defined, since

$$\begin{aligned} |B(v)(w)|_D &= \left(\int_{(0,1)^d} \left| G(v)(x) \cdot (\sqrt{Q}w)(x) \right|^2 dx \right)^{\frac{1}{2}} \\ &\leq \left(\int_{(0,1)^d} |G(v)(x)|^2 \, dx \right)^{\frac{1}{2}} \left(\sup_{x \in \bar{\mathcal{D}}} \left| \left(\sqrt{Q}w \right)(x) \right| \right) \\ &= |G(v)|_H \left(\sup_{x \in \bar{\mathcal{D}}} \left| \left(\sqrt{Q}w \right)(x) \right| \right) \leq c \, |G(v)|_H \, |w|_H \end{aligned}$$

for all v, $w \in H$.

Hence, $B(v)$ is a bounded linear operator from H to $H = D$ with the property

$$\|B(v)\|_{L(H,D)} \leq c \, |G(v)|_H \,, \qquad \forall \, v \in H.$$

In the same way,

$$\|B(v) - B(u)\|_{L(H,D)} \leq c \, |G(v) - G(u)|_H \leq cL \, |v - u|_H$$

for all v, $u \in H$, since G is global Lipschitz continuous.

Hence B is also a global Lipschitz continuous function from H to $L(H, H)$.

Combining the definitions in equation (6)–(7) and (9), then shows that the operator $B : H \to L(H, D)$ is defined by

$$\left(B(v)(w) \right)(x) := g(v(x)) \cdot (\sqrt{Q}w)(x), \qquad x \in (0,1)^d,$$

for every v, $w \in H$.

Let $\gamma \in [0,1)$. Hence $(-A)^\gamma e^{At} B(v)$ is a linear bounded operator from H to H and

$$\begin{aligned}
\left\|(-A)^\gamma e^{At} B(v)\right\|_{HS} &\leq \left\|(-A)^\gamma e^{At}\right\| \|B(v)\|_{HS} \leq t^{-\gamma} \|B(v)\|_{HS} \\
&= t^{-\gamma} \left(\sum_{i \in \mathcal{I}} \alpha_i^2 \left| B(v) \left(\frac{f_i}{\alpha_i} \right) \right|_H^2 \right)^{\frac{1}{2}} \\
&\leq t^{-\gamma} \left(\sum_{i \in \mathcal{I}} \alpha_i^2 \right)^{\frac{1}{2}} |G(v)|_H \left(\sup_{i \in \mathcal{I}} \sup_{x \in \bar{\mathcal{D}}} |f_i(x)| \right) = c\, |G(v)|_H\, t^{-\gamma}
\end{aligned}$$

for every $t > 0$ and $v \in H$. In the same way,

$$\left\|(-A)^\gamma e^{At} \left(B(v) - B(u)\right)\right\|_{HS} \leq c\, |G(v) - G(u)|_H\, t^{-\gamma}$$

for every $t > 0$ and v, $u \in H$, which verifies that Assumption 3.

Finally, if the initial value satisfies Assumption 4, then the SPDE

$$\boxed{dX_t = \left[\Delta X_t + f(X_t)\right] dt + g(X_t) \sqrt{Q}\, dW_t}$$

on the domain $\mathcal{D} = (0,1)^d$ has a unique solution by Theorem 1. Moreover, if $\mathbb{E}\,|(-A)^\gamma u_0|^p < \infty$ with $\gamma \in [0, \frac{1}{2})$, then the solution has almost surely values in $D((-A)^\gamma)$ in the sense of Theorem 2.

In particular, the SPDE with additive noise

$$dX_t = \left[\Delta X_t + f(X_t)\right] dt + \sqrt{Q}\, dW_t, \qquad g(y) \equiv 1,$$

and the stochastic heat equation with linear multiplicative noise

$$dX_t = \Delta X_t dt + X_t \sqrt{Q}\, dW_t, \qquad g(y) \equiv y,\ f(y) \equiv 0,$$

both have unique mild solutions.

Lecture 6: Numerical Methods for SPDEs

The numerical approximation of stochastic partial differential equations (SPDEs), specifically stochastic evolution equations of the parabolic or hyperbolic type, encounters all of the difficulties that arise in the numerical solution of both deterministic PDEs and finite dimensional stochastic ordinary differential equations (SODEs) as well as many more due to the infinite dimensional nature of the driving noise process.

The state of development of numerical schemes for SPDEs compares with that for SODEs in the early 1970s.

However, most of the numerical schemes that have been proposed to date have a low order of convergence, especially in terms of an overall computational effort.

Only recently has it been shown how to construct higher order schemes.

The breakthrough for SODEs started with the Milstein scheme and continued with the systematic derivation of stochastic Taylor expansions and the numerical schemes based on them.

These stochastic Taylor schemes are based on an iterated application of the Itô formula.

The crucial point is that the multiple stochastic integrals which they contain provide more information about the noise processes within discretisation subintervals and this allows an approximation of higher order to be obtained.

There is, however, NO general Itô formula for the solutions of SPDEs in Hilbert spaces or Banach spaces.

Nevertheless, it has recently been shown that Taylor expansions for the solutions of such equations can be constructed by taking advantage of the mild form representation of the solutions.

1 An early result

Consider a parabolic SPDE with a Dirichlet boundary condition on a bounded domain $\mathcal{D}$ in $\mathbb{R}^d$

$$dX_t = [AX_t + F(X_t)]\, dt + G(X_t)\, dW_t. \qquad (1)$$

Suppose that the coefficients satisfy the assumptions of the previous lecture. In particular, assume that the eigenvalues λ_j and the corresponding eigenfunctions $\phi_j \in H_0^{1,2}(\mathcal{D})$ of the linear operator $-A$, i.e., with

$$-A\phi_j = \lambda_j \phi_j, \qquad j = 1, 2, \ldots,$$

form an orthonormal basis in $L_2(\mathcal{D})$ with $\lambda_j \to \infty$ as $j \to \infty$.

- Assume (for now) also that W_t is a standard scalar Wiener process.

Projecting the the SPDE (1) onto the N-dimensional subspace H_N of $L_2(\mathcal{D})$ spanned by $\{\phi_1, \ldots, \phi_N\}$ gives an N-dimensional Itô–Galerkin SODE in $\mathbb{R}^N$ of the form

$$dX_t^{(N)} = \left[A_N X_t^{(N)} + F_N(X_t^{(N)})\right] dt + G_N(X_t^{(N)})\, dW_t, \qquad (2)$$

where $X^{(N)}$ is written synonomously for $(X^{N,1}, \cdots, X^{N,N})^\top \in \mathbb{R}^N$ or $\sum_{j=1}^N X^{N,j}\phi_j \in H_N$ according to the context.

Moreover, $F_N = P_N F\big|_{H_N}$ and $G_N = P_N G\big|_{H_N}$, where F and G are now interpreted as mappings of $L_2(\mathcal{D})$ or $H_0^{1,2}(\mathcal{D})$ into itself, and P_N is the projection of $L_2(\mathcal{D})$ or $H_0^{1,2}(\mathcal{D})$ onto H_N, while $A_N = P_N A\big|_{H_N}$ is the diagonal matrix $\operatorname{diag}[\lambda_j, \ldots, \lambda_N]$.

W. Grecksch and P.E. Kloeden, *Time–discretised Galerkin approximations of parabolic stochastic PDEs*, Bulletin Austral. Math. Soc. 54 (1996), 79–84,

showed that the combined truncation and global discretisation error for an strong order γ stochastic Taylor scheme applied to (2) with time step Δ has the form

$$\max_{k=0,1,\ldots,N_T} \mathbb{E}\left(\left|X_{k\Delta} - Y_k^{(N,\Delta)}\right|_{L_2(\mathcal{D})}\right) \le K_T\left(\lambda_{N+1}^{-1/2} + \lambda_N^{\lfloor\gamma+\frac{1}{2}\rfloor+1}\Delta^\gamma\right),$$

where $\lfloor x \rfloor$ is the integer part of the real number x and K_T is a constant.

Since $\lambda_j \to \infty$ as $j \to \infty$, a very small time step is needed in high dimensions to ensure convergence, i.e., the Itô–Galerkin SODE (2) is stiff and explicit schemes such as strong stochastic Taylor schemes are not really appropriate.

An implicit scheme should be used here, but the special structure of the SODE (2) suggests a simpler linear–implicit scheme, since it is the matrix A_N in the linear part of the drift coefficient that causes the troublesome growth with respect to the eigenvalues, so only this part of the drift coefficient needs to be made implicit, e.g., as in the linear–implicit Euler scheme

$$Y_{k+1}^{(N)} = Y_k^{(N)} + A_N Y_{k+1}^{(N)}\,\Delta + F_N(Y_k^{(N)})\Delta + G_N(Y_k^{(N)})\,\Delta W_k,$$

These results are of limited value because W_t is only one dimensional and the proofs of the convergence of Taylor schemes for SODE require the partial derivatives of the coefficient functions of the SODE to be uniformly bounded on $\mathbb{R}^N$.

2 Overall convergence rate

The overall convergence rate with respective to the temporal and spatial discretisation is usually expressed in terms of the computational cost of the scheme.

For one dimensional domains this is defined by $K = N \cdot M$, where

- N is the number of arithmetical operations, random number and function evaluations per time step to calculate the next iterate $Y_k^{(N,M)}$ of the scheme (this is related to the dimension in a Galerkin approximation) with
- M time steps of constant length $\Delta = \dfrac{T}{M}$.

Write Y_k as $Y_k^{(N,M)}$ to emphasize the dependence on the number of time steps M.

If the scheme has error bound

$$\sup_{k=0,\dots,M} \left(\mathbb{E} \left| X_{t_k} - Y_k^{(N,M)} \right|_{L_2(\mathcal{D})}^2 \right)^{\frac{1}{2}} \leq K_T \left(\frac{1}{N^\alpha} + \frac{1}{M^\beta} \right) \tag{3}$$

for α, $\beta > 0$, then the optimal overall rate, i.e.,

$$\frac{\alpha\beta}{\alpha + \beta}.$$

with respect to the computational cost is

$$\max_{k=0,\dots,M} \left(\mathbb{E} \left| X_{t_k} - Y_k^{(N,M)} \right|_{L_2(\mathcal{D})}^2 \right)^{\frac{1}{2}} \leq K_T \cdot K^{-\frac{\alpha\beta}{\alpha + \beta}}.$$

- e.g., if $\alpha = \dfrac{1}{2}$ and $\beta = 1$, then obtain the overall rate is $\dfrac{1}{3}$.

3 Other early results

Much of the literature is concerned with a semilinear stochastic heat equation with additive space–time white noise $\dot{W}_t$ on the one dimensional domain $\mathcal{D} = (0, 1)$ with the Dirichlet boundary condition over the time interval $[0, T]$, i.e.,

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(u) + \dot{W}_t \tag{4}$$

The following papers are a representative selection of many others in the literature dealing with the SPDE (4).

In 1998 and 1999 GYÖNGY applied finite differences to an SPDE driven by space–time white noise and then used several temporal implicit and explicit schemes, in particular, the linear–implicit Euler scheme. He showed that these schemes converge with order $\frac{1}{2}$ in the space and with order $\frac{1}{4}$ in time — hence with an overall convergence rate of $\frac{1}{6}$ with respect to the computational cost in space and time.

In 1999 SHARDLOW also applied finite differences to the SPDE (4) to obtain a spatial discretisation, which he then discretised in time with a θ-method. This had an overall convergence rate $\frac{1}{6}$ with respect to the computational cost.

A.M. DAVIE AND J.G. GAINES, *Convergence of numerical schemes for the solution of parabolic stochastic partial differential equations*, Math. Comput. 70 (2001), 123-134,

showed that any numerical scheme applied to the SPDE (4) with $f = 0$ which uses only values of the noise W_t cannot converge faster than the rate of $\frac{1}{6}$ with respect to the computational cost.

Müller–Gronbach & Ritter showed in 2007 that this is also a lower bound for the convergence rate.

They even showed that one cannot improve this rate of convergence by choosing nonuniform time steps.

Higher rates were obtained for smoother types of noise

For example, in 2003 Hausenblas applied the linear–implicit and explicit Euler schemes and the Crank–Nicolson scheme to an SPDE of the form (4).

For trace–class noise she obtained the order $\frac{1}{4}$ with respect to the computational cost, but in the general case of space–time white noise the convergence rate was no better than the Davie–Gaines barrier rate $\frac{1}{6}$.

Similarly, in 2004 Lord & Rougemont discretised in time the Galerkin–SODE obtained from the SPDE (4) with the numerical scheme

$$Y_{k+1}^{(N,M)} = e^{A_N h}\left(Y_k^{(N,M)} + F_N(Y_k^{(N,M)}\Delta_k) + \Delta W_k^N\right) \tag{5}$$

which they showed to be useful when the noise is very smooth in space, in particular with Gévrey regularity.

However, in the general case of space–time white noise the scheme (5) converges at Davie–Gaines barrier rate $\frac{1}{6}$.

4 The exponential Euler scheme

A. Jentzen and P. E. Kloeden, *Overcoming the order barrier in the numerical approximation of SPDEs with additive space-time noise*, Proc. Roy. Soc. London, Series A 465 (2009), no. 2102, 649–667.

Consider a parabolic SPDE with additive noise

$$dX_t = [AX_t + F(X_t)]\, dt + dW_t, \qquad X_0 = x_0, \tag{6}$$

in a Hilbert space $(H, |\cdot|)$ with inner product $<\cdot,\cdot>$, where A is an in general unbounded linear operator, $F : H \to H$ is a nonlinear continuous function and W_t is a cylindrical Wiener process.

Interpet the SPDE (6) in the mild sense

$$X_t = e^{At}x_0 + \int_0^t e^{A(t-s)}F(X_s)\, ds + \int_0^t e^{A(t-s)}\, dW_s. \tag{7}$$

Use the fact that the solution of the N-dimensional Itô–Galerkin SODE in the space $H_N := P_N H$ (or, equivalently, in $\mathbb{R}^N$)

$$dX_t^N = \left(A_N X_t^N + F_N(X_t^N)\right)\, dt + dW_t^N \tag{8}$$

has an analogous "mild" representation

$$X_t^N = e^{A_N t}u_0^N + \int_0^t e^{A_N(t-s)}F_N(X_s^N)\, ds + \int_0^t e^{A_N(t-s)}\, dW_s^N. \tag{9}$$

This motivated what JENTZEN & KLOEDEN called the exponential Euler scheme

$$\boxed{\begin{aligned} Y_{k+1}^{(N,M)} &= e^{A_N\Delta} Y_k^{(N,M)} + A_N^{-1}\left(e^{A_N\Delta} - I\right) F_N(Y_k^{(N,M)}) \\ &\quad + \int_{t_k}^{t_{k+1}} e^{A_N(t_{k+1}-s)}\, dW_s^N \end{aligned} \tag{10}}$$

with the time step $\Delta = \frac{T}{M}$ for some $M \in \mathbb{N}$ and discretisation times $t_k = k\Delta$ for $k = 0, 1, \ldots, M$.

The exponential Euler scheme is easier to simulate than may seem on the first sight.

Denoting the components of $Y_k^{(N,M)}$ and F_N by

$$Y_{k,i}^{(N,M)} = \left\langle e_i, Y_k^{(N,M)} \right\rangle, \qquad F_N^i = \langle e_i, F_N \rangle, \qquad i = 1, \ldots, N,$$

the numerical scheme (10) can be rewritten as

$$\boxed{\begin{aligned} Y_{k+1,1}^{(N,M)} &= e^{-\lambda_1\Delta} Y_{k,1}^{(N,M)} + \frac{(1-e^{-\lambda_1\Delta})}{\lambda_1} F_N^1(Y_k^{(N,M)}) + \left(\frac{q_1}{2\lambda_1}(1-e^{-2\lambda_1\Delta})\right)^{\frac{1}{2}} R_k^1 \\ &\vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ Y_{k+1,N}^{(N,M)} &= e^{-\lambda_N\Delta} Y_{k,N}^{(N,M)} + \frac{(1-e^{-\lambda_N\Delta})}{\lambda_N} F_N^N(Y_k^{(N,M)}) + \left(\frac{q_N}{2\lambda_N}(1-e^{-2\lambda_N\Delta})\right)^{\frac{1}{2}} R_k^N, \end{aligned}}$$

where the R_k^i for $i = 1,\ldots, N$ and $k = 0, 1, \ldots, M-1$ are independent, standard normally distributed random variables.

5 Assumptions

Assumption 1. (Linear Operator A) *There exist sequences of real eigenvalues* $0 < \lambda_1 \leq \lambda_2 \leq \dots$ *and orthonormal eigenfunctions* $(e_n)_{n\geq 1}$ *of* A *such that the linear operator* $A : D(A) \subset H \to H$ *is given by*

$$Av = \sum_{n=1}^{\infty} -\lambda_n \langle e_n, v\rangle e_n$$

for all $v \in D(A)$ *with* $D(A) = \left\{v \in H : \sum_{n=1}^{\infty} |\lambda_n|^2 |\langle e_n, v\rangle|^2 < \infty\right\}$.

Assumption 2. (Nonlinearity F) *The nonlinearity* $F : H \to H$ *is two times continuously Fréchet differentiable and its derivatives satisfy*

$$\|F'(x) - F'(y)\| \leq L|x-y|_H, \qquad \left|(-A)^{(-r)}F'(x)(-A)^r v\right|_H \leq L|v|_H$$

for all x, $y \in H$, $v \in D((-A)^r)$ *and* $r = 0, \frac{1}{2}, 1$ *and*

$$\left|A^{-1}F''(x)(v,w)\right|_H \leq L\left|(-A)^{-\frac{1}{2}}v\right|_H \left|(-A)^{-\frac{1}{2}}w\right|_H$$

for all v,w,$x \in H$, *where* $L > 0$ *is a positive constant.*

Assumption 3. (Cylindrical Q-Wiener process W_t) *There exist a sequence* $(q_n)_{n\geq 1}$ *of positive real numbers and a real number* $\gamma \in (0,1)$ *such that*

$$\sum_{n=1}^{\infty} \lambda_n^{2\gamma-1} q_n < \infty$$

and pairwise independent, scalar $\mathcal{F}_t$*-adapted Wiener processes* $(W_t^n)_{t\geq 0}$ *for* $n \geq 1$. *The cylindrical* Q*-Wiener process* W_t *is given formally by*

$$W_t = \sum_{n=1}^{\infty} \sqrt{q_n} W_t^n \, e_n.$$

Assumption 4. (Initial value) *The random variable* $x_0 : \Omega \to D((-A)^\gamma)$ *satisfies* $\mathbb{E}\,|(-A)^\gamma x_0|_H^4 < \infty$, *where* $\gamma > 0$ *is given in Assumption 3.*

Under Assumptions 1–4 the SPDE (6) has a unique mild solution X_t on the time interval $[0, T]$, where X_t is the predictable stochastic process in $D((-A)^\gamma)$ given by the mild form equation (7).

Since Assumption 2 also applies to F_N, the Itô–Galerkin SODE (8) also has a unique solution on $[0, T]$, which is given implicitly by the mild form equation (9).

The formalism here includes space–time white noise (in one dimensional domains) as well as trace–class noise.

Theorem 1. *Suppose that Assumptions 1–4 are satisfied. Then, there is a constant* $C_T > 0$ *such that*

$$\sup_{k=0,\ldots,M} \left(\mathbb{E} \left| X_{t_k} - Y_k^{(N,M)} \right|_H^2 \right)^{\frac{1}{2}} \leq C_T \left(\lambda_N^{-\gamma} + \frac{\log(M)}{M} \right)$$

holds for all N, $M \in \mathbb{N}$, *where* X_t *is the solution of SPDE* (6), $Y_k^{(N,M)}$ *is the numerical solution given by* (10), $t_k = T\dfrac{k}{M}$ *for* $k = 0, 1, \ldots, M$, *and* $\gamma > 0$ *is the constant given in Assumption (A2).*

In fact, the exponential Euler scheme (10) converges in time with a strong order $1 - \varepsilon$ for an arbitrary small $\varepsilon > 0$ since $\log(M)$ can be estimated by M^ε, so

$$\frac{\log(M)}{M} \approx \frac{1}{M^{1-\varepsilon}}.$$

An essential point is that the integral $\int_{t_k}^{t_{k+1}} e^{A_N(t_{k+1}-s)}\, dW_s^N$ includes includes more information about the noise on the discretisation interval.

6 Restrictiveness of the assumptions

Assumptions 1–4 in Theorem 1 are typical of those used in the mathematical literature and are interesting for when they are valid and not valid.

> Assumptions 1–4 are quiet restrictive. and Theorem 1 has several serious shortcomings.

Firstly, the eigenvalues and eigenfunctions of the operator A are rarely known except in very simple domains — finite element methods are a possible way around this difficulty.

More seriously, Assumption 2 on the nonlinearity F, which is understood as the Nemytskii operator of some function $f : \mathbb{R} \to \mathbb{R}$, is very restrictive and excludes Nemytskii operators for functions like

$$f(u) = \frac{u}{1+u^2}, \qquad f(u) = u - u^3, \qquad u \in \mathbb{R}.$$

A particular problem is the Fréchet differentiability of the function F when considered as a mapping on the Hilbert space H into itself and the boundedness of the derivatives of F as expressed in Assumption 2.

The other problem is the global Lipschitz estimate on F — this difficulty also arises for finite dimensional Itô SODE for which it can be overcome by using pathwise convergence rather than strong convergence.

Lectures 7 & 8: SPDEs with Additive Noise

Taylor expansions of solutions of SPDE in Banach spaces are the basis for deriving higher order numerical schemes for SPDE, just as for SODE.

There is, however, a major difficulty for SPDE. Although the driving Wiener process is a martingale, the solution process is usually not even a semi–martingale.

In particular, in general, an Itô formula does not exist for the solutions of SPDEs, just special cases.

Hence stochastic Taylor expansions for the solutions of an SPDE cannot be derived as for the solutions of finite dimensional SODE.

In this and the next lecture we consider the derivation of robust Taylor expansions of solutions of SPDE with additive noise of the form

$$dX_t = [AX_t + f(X_t)]\, dt + B\, dW_t, \qquad (1)$$

which will be interpreted in mild form

$$X_t = S_t\xi + \int_0^t S_{t-s}\, F(X_s)\, ds + O_t$$

for an appropriate semigroup of operators $(S_t)_{t\geq 0}$.

• The Taylor expansions will be derived in essential the same way as for RODEs.

1 Assumptions

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(V, |\cdot|)$ be separable real Banach space.

Let $(L(V), \|\cdot\|)$ denote the real Banach space of all bounded linear operator from V to V and let $\left(L^{(n)}(V,V), \|\cdot\|\right)$ with $n \in \mathbb{N}$ be the real Banach space of all bounded n-multilinear operators from V^n to V.

Finally, let $\mathcal{B}(V)$ be the Borel σ-algebra on V and fix $T > 0$.

A mapping $Y : [0,T] \times \Omega \to V$ is said to be a stochastic process if the mappings

$$Y_t : \Omega \to V, \qquad \omega \mapsto Y_t(\omega) := Y(t,\omega), \qquad \omega \in \Omega,$$

are $\mathcal{F}/\mathcal{B}(V)$-measurable for every $t \in [0,T]$.

Such a stochastic process $Y : [0,T] \times \Omega \to V$ is said to have continuous sample paths if the mappings

$$[0,T] \to V, \qquad t \mapsto Y_t(\omega), \qquad t \in [0,T],$$

are continuous for every $\omega \in \Omega$.

It is said to have θ-Hölder continuous sample paths where $\theta \in (0,1]$ if

$$\boxed{\sup_{0 \le t_1 < t_2 \le T} \frac{|Y_{t_2}(\omega) - Y_{t_1}(\omega)|}{(t_1 - t_2)^{\theta}} < \infty}$$

holds for every $\omega \in \Omega$.

Assumption 1 (Semigroup S)**.** *The mapping $S : [0,\infty) \to L(V)$ satisfies*

$$S_0 = I, \qquad S_{t_1} S_{t_2} = S_{t_1+t_2},$$

$$\sup_{0\le t\le T} \|S_t\| < \infty, \quad \sup_{0\le s<t\le T} \frac{\|S_t - S_s\| \cdot s}{t-s} < \infty \tag{2}$$

for every $t_1, t_2 \in [0,\infty)$, where I is the identity operator on V.

Comments on Assumption 1:

For greater generality the semigroup $S : [0,\infty) \to L(V)$ in Assumption 1 is not assumed to be strongly continuous as is usual in the literature.

The last condition in (2) implies that S_t is locally Lipschitz continuous on $(0,T]$ and also indicates the size of the local Lipschitz constant.

If $S_t = e^{At}$, $t \ge 0$, is a strongly continuous semigroup on V with the generator $A : D(A) \subset V \to V$, then the last condition in (2) follows from the estimate

$$\left\|Ae^{At}\right\| \le c/t, \qquad t \in (0,T],$$

and some constant $c > 0$, since

$$\begin{aligned} \left|e^{At_2}v - e^{At_1}v\right| &= \left|\left(e^{A(t_2-t_1)} - I\right) e^{At_1}v\right| = \left|\int_0^{t_2-t_1} \left(e^{As}Ae^{At_1}v\right)\, ds\right| \\ &\le \int_0^{t_2-t_1} \left|e^{As}Ae^{At_1}v\right|\, ds \le \left(\sup_{0\le t\le T} \left\|e^{At}\right\|\right) \frac{c}{t_1}\,(t_2-t_1)\,|v| \end{aligned}$$

for every $0 < t_1 \le t_2 \le T$ and every $v \in V$.

Assumption 2 (Nonlinearity F)**.** *The mapping $F : V \to V$ is infinitely often Fréchet differentiable.*

Comments on Assumption 2: The smoothness of the drift coefficient provided by Assumption 2 is required for the Taylor expansions that will be derived below.

The infinitely often differentiablility is assumed only for convenience and it would suffice to assume that F is k-times Fréchet differentiable for some $k \in \mathbb{N}$, which is sufficiently large.

Assumption 3 (Stochastic process O)**.** *The stochastic process $O : [0,T] \times \Omega \to V$ is θ-Hölder continuous sample paths with $\theta \in (0,1)$.*

Comments on Assumption 3: This means that the mappings

$$O_t : \Omega \to V, \qquad \omega \mapsto O_t(\omega) := O(t,\omega), \qquad \omega \in \Omega,$$

are $\mathcal{F}/\mathcal{B}(V)$-measurable for every $t \in [0,T]$ and that

$$\sup_{\substack{t_1,t_2 \in [0,T] \\ t_1 \neq t_2}} \frac{|O_{t_2}(\omega) - O_{t_1}(\omega)|}{|t_2 - t_1|^\theta} < \infty$$

holds for every $\omega \in \Omega$.

Pathwise Hölder continuity is typical of noise processes in applications.

Assumption 3 does not require the noise process to be generated by a Wiener process — fractional Brownian motion is also possible.

Assumption 4 (Initial value ξ)**.** *The initial value $\xi : \Omega \to V$ is $\mathcal{F}/\mathcal{B}(V)$-measurable with*

$$\sup_{t\in(0,T]} \frac{1}{t} \left|S_t\xi(\omega) - \xi(\omega)\right| < \infty$$

for every $\omega \in \Omega$.

Comments on Assumption 4: The condition in Assumption 4 states that the initial random variable $\xi : \Omega \to V$ is smooth in some sense. This may not be satisfied in some applications, but can be overcome in some cases.

Assumption 5 (Existence of a solution)**.** *There exists a stochastic process $X : [0,T] \times \Omega \to V$ with continuous sample paths, which satisfies the integral equation*

$$X_t(\omega) = S_t\xi(\omega) + \int_0^t S_{t-s}\, F(X_s(\omega))\, ds + O_t(\omega) \tag{3}$$

for every $t \in [0,T]$ and every $\omega \in \Omega$.

(The integral in (3) *is a V-valued Bochner integral).*

Comments on Assumption 5: The existence of a solution of the SPDE provided by Assumption 5 is a minimal assumption to do numerical analysis.

It is provided by more specific assumptions on the linear operator A in the SPDE.

The explicit use of the semigroup allows greater generality and simpler notation.

2 Properties of the solutions

The pathwise uniqueness and Hölder continuity of the solution process, the existence of which is provided by Assumption 5, follow from Assumptions 1–4.

Lemma 1. *Suppose that Assumptions 1–5 are satisfied and let* X, $Y : \Omega \times [0,T] \to V$ *be two stochastic processes with continuous sample paths, which satisfy* (3)*, i.e.,*

$$X_t(\omega) = S_t\xi(\omega) + \int_0^t S_{t-s}\, F(X_s(\omega))\, ds + O_t(\omega)$$

$$Y_t(\omega) = S_t\xi(\omega) + \int_0^t S_{t-s}\, F(Y_s(\omega))\, ds + O_t(\omega)$$

for all $t \in [0,T]$ *and* $\omega \in \Omega$*. Then*

$$X_t(\omega) = Y_t(\omega)$$

for all $t \in [0,T]$ *and* $\omega \in \Omega$*.*

The sample paths of the unique solution process are not only continuous, but in fact Hölder continuous.

Lemma 2. *Suppose that Assumptions 1–5 hold. Then the unique solution process* $X : [0,T] \times \Omega \to V$ *has* θ*-Hölder continuous sample paths, i.e.,*

$$\sup_{0 \le t_1 < t_2 \le T} \frac{|X_{t_2}(\omega) - X_{t_1}(\omega)|}{(t_2 - t_1)^\theta} < \infty$$

holds for every $\omega \in \Omega$*, where* $\theta \in (0,1)$ *is given in Assumption 3.*

Subtracting the noise process O_t from the solution process X_t in integral equation (3) gives the equation

$$X_t - O_t = S_t\xi + \int_0^t S_{t-s}\, F(X_s)\, ds \tag{4}$$

for every $t \in [0,T]$.

Proposition 1. *Suppose that Assumptions 1 and 2 are sastisfied and let* $X : [0,T] \times \Omega \to V$ *be a stochastic process with continuous sample paths. Then* $Y : [0,T] \times \Omega \to V$ *given by*

$$Y_t(\omega) := \int_0^t S_{t-s}\, F(X_s(\omega))\, ds$$

for every $t \in [0,T]$ *and* $\omega \in \Omega$ *is a well defined stochastic process with continuous sample paths.*

The process $X_t - O_t$ is smoother than the original solution process X_t. This additional regularity of the stochastic process $X_t - O_t$ will play an important role in the derivation Taylor expansions.

Lemma 3. *Suppose that Assumptions 1–5 hold. Then the stochastic process* $X - O : [0,T] \times \Omega \to V$ *has Lipschitz continuous sample paths, i.e.,*

$$\sup_{0 \le t_1 < t_2 \le T} \frac{|(X_{t_2}(\omega) - O_{t_2}(\omega)) - (X_{t_1}(\omega) - O_{t_1}(\omega))|}{t_2 - t_1} < \infty$$

holds for every $\omega \in \Omega$*, where* O *and* X *are given in Assumptions 3 and 5.*

3 Examples

Let $H = L^2\left((0,1)^d, \mathbb{R}\right)$ be the real Hilbert space of equivalence classes of square integrable functions from $\mathcal{D} = (0,1)^d$ to $\mathbb{R}$ for some $d \in \mathbb{N}$ with the scalar product and the norm

$$\langle u, v \rangle_H = \int_{(0,1)^d} u(x)\, v(x)\, dx, \qquad |u|_H = \left(\int_{(0,1)^d} u(x)^2\, dx \right)^{\frac{1}{2}}$$

Let $V = C([0,1]^d, \mathbb{R})$ be the real Banach space of continuous functions from $[0,1]^d$ to $\mathbb{R}$ with the norm is given by

$$|v|_V = \sup_{x \in [0,1]^d} |v(x)|$$

- Note that $V \subset H$ densely and continuously.

$$e_i(x) = 2^{\frac{d}{2}} \sin(i_1 \pi x_1) \ldots \sin(i_d \pi x_d), \qquad x \in [0,1]^d, \tag{5}$$

$$\lambda_i = \pi^2 \left(i_1^2 + \ldots + i_d^2 \right) \tag{6}$$

with indices $i = (i_1, \ldots, i_d) \in \mathcal{I} = \mathbb{N}^d$ with the Euclidean norm $\|i\|_2$.

These are eigenfunctions and eigenvalues of the minus Laplace operator on $(0,1)^d$ with the Dirichlet boundary condition.

The $e_i \in V$ for $i \in \mathbb{N}^d$ and form an orthonormal basis of H.

3.1 Semigroup generated by the Laplacian

The next proposition gives an example of a semigroup satisfying Assumption 1.

Proposition 2. *The mapping $S : [0, \infty) \to L(V)$ defined by*

$$S_0 v = v, \qquad S_t v = \sum_{i \in \mathbb{N}^d} e^{-\lambda_i t} \langle e_i, v \rangle_H \, e_i = \lim_{N \to \infty} \sum_{\substack{i \in \mathbb{N}^d \\ \|i\|_2 \leq N}} e^{-\lambda_i t} \langle e_i, v \rangle_H \, e_i$$

for every $t \in (0, \infty)$ and $v \in V$ satisfies Assumption 1.
The functions $(e_i)_{i \in \mathbb{N}^d}$ and the real numbers $(\lambda_i)_{i \in \mathbb{N}^d}$ are given by (5) and (6).

This is simply the semigroup generated by the Laplace operator with Dirichlet boundary conditions. Other boundary conditions could also be considered here.

It is easy to see that the semigroup given in Proposition 2 is not strongly continuous. The mapping

$$[0, T] \to V, \qquad t \mapsto S_t v, \qquad t \in [0, T],$$

is not continuous at $t = 0$ if, e.g., $v(x) = 1$, although the mapping

$$[0, T] \to H, \qquad t \mapsto S_t v, \qquad t \in [0, T],$$

is continuous for every $v \in V$ with respect to the H-norm $|\cdot|_H$.

3.2 The drift as a Nemytskii operator

The drift mapping $F : V \to V$ can often be defined as the Nemytskii operator of a real valued function of real variables.

Let $f : [0,1]^d \times \mathbb{R} \to \mathbb{R}$ be a continuous function such that the mappings

$$\mathbb{R} \to \mathbb{R}, \qquad y \mapsto f(x,y), \qquad y \in \mathbb{R}, \tag{7}$$

are infinitely often differentiable for each $x \in [0,1]^d$ and let

$$\frac{\partial^n}{\partial y^n} f : [0,1]^d \times \mathbb{R} \to \mathbb{R}$$

be the nth derivative of f with respect to its last variable in $[0,1]^d \times \mathbb{R} \subset \mathbb{R}^{d+1}$.

Proposition 3. *Let $f : [0,1]^d \times \mathbb{R} \to \mathbb{R}$ be a continuous function which is smooth in its last variable in the sense of (7) such that that the partial derivatives $\frac{\partial^n}{\partial y^n} f : [0,1]^d \times \mathbb{R} \to \mathbb{R}$ are continuous for every $n \in \mathbb{N}$.*
Then the corresponding Nemytskii operator

$$F : V \to V, \qquad (F(v))(x) := f(x, v(x)), \qquad x \in [0,1]^d, \qquad v \in V,$$

is infinitely often Fréchet differentiable and its Fréchet derivatives $F^{(n)} : V \to L^{(n)}(V,V)$ are given by

$$\left(F^{(n)}(v)(v_1, \ldots, v_n)\right)(x) = \left(\frac{\partial^n}{\partial y^n} f\right)(x, v(x)) \cdot v_1(x) \cdot \ldots \cdot v_n(x)$$

for every $x \in [0,1]^d$ and v, v_1, …, $v_n \in V$ for each $n \in \mathbb{N}$.
Moreover, these derivatives satisfy

$$\left\|F^{(n)}(v)\right\| = \sup_{x \in [0,1]^d} \left| \left(\frac{\partial^n}{\partial y^n} f\right)(x, v(x)) \right|$$

for every $v \in V$ and $n \in \mathbb{N}$.

3.3 Stochastic process as stochastic convolution

In principle, O could be an arbitrary stochastic process with Hölder continuous sample paths such as a fractional Brownian motion

The following example is a very important case involving stochastic convolutions of the semigroup S constructed in Proposition 2 and a cylindrical Wiener process.

Proposition 4. *Let $\rho > 0$, let $(W_t^i)_{t\geq 0}$, $i \in \mathbb{N}^d$, be a family of pairwise independent standard scalar Wiener processes and let $b : \mathbb{N}^d \to \mathbb{R}$ be a given function with $\sum_{i\in\mathbb{N}^d} \|i\|_2^{2\rho-2} |b(i)|^2 < \infty$.*
Then, there exists a stochastic process $O : [0,T] \times \Omega \to V$ which satisfies

$$\sup_{0\leq t_1<t_2\leq T} \frac{|O_{t_2}(\omega) - O_{t_1}(\omega)|}{(t_2 - t_1)^\theta} < \infty$$

for every $\omega \in \Omega$ and $\theta \in (0, \min(1,\rho))$ such that

$$\mathbb{P}\left[\lim_{N\to\infty} \sup_{0\leq t\leq T} \left| O_t - \sum_{i\in\{1,\dots,N\}^d} b(i) \left(-\lambda_i \int_0^t e^{-\lambda_i(t-s)} W_s^i \, ds + W_t^i\right) e_i \right| = 0\right] = 1,$$

where the functions $(e_i)_{i\in\mathbb{N}^d}$ and the real numbers $(\lambda_i)_{i\in\mathbb{N}^d}$ are given by (5) and (6). In particular, the stochastic process O satisfies Assumption 3.

For an appropriate cylindrical I-Wiener process W_t on H

$$O_t = \sum_{i\in\mathbb{N}^d} b(i) \left(\int_0^t e^{-\lambda_i(t-s)} \, dW_s^i\right) e_i = \int_0^t S_{t-s} \, B \, dW_s, \qquad \mathbb{P}-\text{a.s.},$$

where the bounded linear operator $B : H \to H$ is given by

$$Bv = \sum_{i\in\mathbb{N}^d} b(i) \langle e_i, v\rangle_H \, e_i, \qquad \forall\, v \in H.$$

4 Concrete examples

As above consider

$$X_t = S_t\xi + \int_0^t S_{t-s}\, F(X_s)\, ds + O_t, \qquad t \in [0,T],$$

on the real Banach space $V = C\left([0,1]^d, \mathbb{R}\right)$.

There are a many possible choices for the state space V for evolutionary SPDEs.

The choice $V = C\left([0,1]^d, \mathbb{R}\right)$ with its supremum norm $|v| = \sup_{x\in[0,1]^d} |v(x)|$ for $v \in V$ yields very strong results.

In principle, the Dirichlet boundary condition could be incorporated in the space V so that the semigroup becomes strongly continuous.

This may, however, restrict the nonlinearity in some sense.

Example 1.)

Let $d = 1$ and $T = 1$. In adddition, let $f(x,y) = y - y^3$ for all $x \in [0,1]$ and $y \in \mathbb{R}$, let $b(i) = \frac{1}{2}$ for each $i \in \mathbb{N}$ and, finally, let $\xi = \frac{1}{4}e_1$.

The SPDE

$$dX_t = \left[\frac{\partial^2}{\partial x^2} X_t + X_t - X_t^3\right] dt + \frac{1}{2}\, dW_t \tag{8}$$

with boundary and initial conditions

$$X_t(0) = X_t(1) = 0, \qquad X_0(x) = \frac{\sqrt{2}}{4}\sin(\pi x)$$

satisfies Assumption 3 on $C\left([0,1], \mathbb{R}\right)$ for every $\theta \in (0, \frac{1}{4})$.

Example 2.)

As before set $d = 1$ and $T = 1$, but now choose the functions

$$f(x, y) = \frac{y}{1 + y^2}, \qquad x \in [0, 1],\ y \in \mathbb{R},$$

and $b(i) = 1$ for all $i \in \mathbb{N}$ and the initial value $\xi = 0$.

The SPDE

$$\boxed{dX_t = \left[\frac{\partial^2}{\partial x^2} X_t + \frac{X_t}{1 + X_t^2}\right] dt + dW_t \qquad (9)}$$

with boundary and initial conditions

$$X_t(0) = X_t(1) = 0, \quad X_0 = 0$$

satisfies Assumption 3 on $C([0, 1], \mathbb{R})$ for all $\theta \in (0, \frac{1}{4})$.

Example 3.)

Now let $d = 2$ with $T = 1$ and $\xi = 0$ as before and consider the function

$$f(x, y) = -y^3 \qquad x \in [0, 1]^2,\ y \in \mathbb{R},$$

and

$$b(i) = \frac{1}{\sqrt{i_1^2 + i_2^2}}, \qquad i = (i_1, i_2) \in \mathbb{N}^2.$$

The SPDE

$$\boxed{dX_t = \left[\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}\right) X_t - X_t^3\right] dt + B\, dW_t \qquad (10)}$$

with boundary and initial conditions

$$X_t|_{\partial(0,1)^2} = 0, \qquad X_0 = 0$$

satisfies Assumption 3 on $C([0, 1]^2, \mathbb{R})$ for every $\theta \in (0, \frac{1}{2})$.

5 Taylor Expansions

Recall that we are considering the pathwise SPDE with additive noise in mild form

$$X_t(\omega) = S_t\xi(\omega) + \int_0^t S_{t-s}\, F(X_s(\omega))\, ds + O_t(\omega)$$

Taylor expansions of the solution of the SPDE (3) will be derived here in much the same way as was done for RODEs,

i.e., using Taylor expansions of the drift function and inserting lower order Taylor expansions of the right hand side of higher order ones to obtain a closed form expression for the solution.

This will involve special integral operators acting on a space $\mathcal{C}$ of all stochastic processes $Y : [t_0, T] \times \Omega \to V$ with continuous sample paths defined by

$$\mathcal{C} := \Big\{ Y : [t_0,T] \times \Omega \to V \;\Big|\; Y_t : \Omega \to V \text{ is } \mathcal{F}/\mathcal{B}(V)\text{-measurable } \forall t \in [t_0,T] \text{ and } [t_0,T] \ni t \mapsto Y_t(\omega) \text{ is continuous } \forall\, \omega \in \Omega \Big\}.$$

This is a real vector space since V is assumed to be separable.

Define $\Delta t := t - t_0$ for $t \in [t_0, T]$ and define the stochastic processes

$$\Delta X,\, \Delta O \,:\, [t_0, T] \times \Omega \,\to\, V \,\in \mathcal{C}$$

by

$$\boxed{\Delta X_t(\omega) := X_t(\omega) - X_{t_0}(\omega), \qquad \Delta O_t(\omega) := O_t(\omega) - O_{t_0}(\omega)}$$

for $t \in [t_0, T]$ and $\omega \in \Omega$, where O_t is the stochastic process given in Assumption 3 and X_t is the unique solution process of the SPDE (3).

Since

$$\begin{aligned}
X_t &= S_t\xi + \int_0^t S_{t-s}\, F(X_s)\, ds + O_t \\
&= S_{\Delta t}\,(S_{t_0}\xi) + \int_0^{t_0} S_{t-s}\, F(X_s)\, ds + \int_{t_0}^t S_{t-s}\, F(X_s)\, ds + O_t \\
&= S_{\Delta t}\left(S_{t_0}\xi + \int_0^{t_0} S_{t_0-s}\, F(X_s)\, ds\right) + \int_{t_0}^t S_{t-s}\, F(X_s)\, ds + O_t \\
&= S_{\Delta t}\,(X_{t_0} - O_{t_0}) + \int_{t_0}^t S_{t-s}\, F(X_s)\, ds + O_t
\end{aligned}$$

for every $t \in [t_0, T]$, the <u>increment</u> ΔX of the solution is given by

$$\boxed{\Delta X_t = (S_{\Delta t} - I)\,(X_{t_0} - O_{t_0}) + \int_{t_0}^t S_{t-s}\, F(X_s)\, ds + \Delta O_t \qquad (11)}$$

for all $t \in [t_0, T]$.

The basic formula (11) is the <u>starting point</u> for derivation of Taylor expansions of the solution process X and its increment ΔX.

6 Integral operators

Define the stochastic processes I^0 and $I^0_* \in \mathcal{C}$ by

$$\begin{aligned} I^0(t,\omega) &:= (S_{\Delta t} - I)\,(X_{t_0}(\omega) - O_{t_0}(\omega)) + \left(\int_0^{\Delta t} S_s\,ds\right) F(X_{t_0}(\omega)) \\ I^0_*(t,\omega) &:= (S_{\Delta t} - I)\,(X_{t_0}(\omega) - O_{t_0}(\omega)) + \int_{t_0}^{t} S_{t-s}\,F(X_s(\omega))\,ds \end{aligned}$$

for every $t \in [t_0, T]$ and $\omega \in \Omega$.

Note that the subscript $*$ means that the solution process X is included in the integrand.

The absence of this subscript means that only the constant value X_{t_0} is present.

For each $n \in \mathbb{N}$ define n-multilinear symmetric mappings I^n and $I^n_* : \mathcal{C}^n \to \mathcal{C}$ by

$$I^n[g_1,\ldots,g_n](t,\omega) := \frac{1}{n!}\int_{t_0}^{t} S_{t-s}\,F^{(n)}(X_{t_0}(\omega))\,(g_1(s,\omega),\ldots,g_n(s,\omega))\;ds$$

and

$$I^n_*[g_1,\ldots,g_n](t,\omega) := \int_{t_0}^{t} S_{t-s}\left(\int_0^1 F^{(n)}(X_{t_0}(\omega) + r\Delta X_s(\omega))\,(g_1(s,\omega),\ldots,g_n(s,\omega))\,\frac{(1-r)^{n-1}}{(n-1)!}\,dr\right) ds$$

for every $t \in [t_0, T]$, $\omega \in \Omega$ and $g_1, \ldots, g_n \in \mathcal{C}$.

It follows as in Proposition 1 from Assumptions 1–5 that the stochastic processes I^0, $I^0_* \in \mathcal{C}$ and the mappings I^n, $I^n_* : \mathcal{C}^n \to \mathcal{C}$ for $n \in \mathbb{N}$ are well defined.

A Taylor approximation of X_t with $t \in (t_0, T]$ about X_{t_0} can depend only on the value of the process X at time t_0, i.e., only on X_{t_0}.

The stochastic processes I^0 and $I^n[g_1, \ldots, g_n] \in \mathcal{C}$ for $g_1, \ldots, g_n \in \mathcal{C}$, $n \in \mathbb{N}$, depend on the solution at time t_0 only and are therefore useful Taylor approximations for the solution process X and its increment ΔX.

The stochastic processes I_*^0 and $I_*^n[g_1, \ldots, g_n] \in \mathcal{C}$ for $g_1, \ldots, g_n \in \mathcal{C}$, $n \in \mathbb{N}$, depend on the whole solution X_s with $s \in [t_0, t]$ and can thus represent remainder terms, which can then be further expanded to give a better approximation using the following proposition.

Proposition 5. *Let Assumptions 1–5 hold. Then,*

$$I_*^0 = I^0 + I_*^1[I_*^0] + I_*^1[\Delta O], \tag{12}$$

and

$$\begin{aligned} I_*^n[g_1, \ldots, g_n] &= I^n[g_1, \ldots, g_n] + I_*^{n+1}[I_*^0, g_1, \ldots, g_n] \\ &\quad + I_*^{n+1}[\Delta O, g_1, \ldots, g_n] \end{aligned} \tag{13}$$

for every $g_1, \ldots, g_n \in \mathcal{C}$ and $n \in \mathbb{N}$.

The basic formula (11) for ΔX can be written in terms of these integral operators as $\Delta X_t = I_*^0(t) + \Delta O_t$ t for every $t \in [t_0, T]$ and symbolically in the space $\mathcal{C}$ as

$$\Delta X = I_*^0 + \Delta O.$$

Proof. Consider (12). Applying the Fundamental Theorem of Calculus for Banach space valued functions to the function $[0,1] \to V$ defined by

$$r \;\mapsto\; F\left(X_{t_0} + r\left(X_s - X_{t_0}\right)\right)$$

yields

$$\begin{aligned} F(X_s) &= F(X_{t_0}) + \int_0^1 F'(X_{t_0} + r(X_s - X_{t_0}))(X_s - X_{t_0})\, dr \\ &= F(X_{t_0}) + \int_0^1 F'(X_{t_0} + r\Delta X_s)(\Delta X_s)\, dr \\ &= F(X_{t_0}) + \int_0^1 F'(X_{t_0} + r\Delta X_s)(I_*^0(s))\, dr + \int_0^1 F'(X_{t_0} + r\Delta X_s)(\Delta O_s)\, dr \end{aligned}$$

for every $s \in [t_0, T]$, where (15) has been used. Hence,

$$\begin{aligned} I_*^0(t) &= \left(S_{\Delta t} - I\right)\left(X_{t_0} - O_{t_0}\right) + \int_{t_0}^t S_{t-s}\, F(X_s)\, ds \\ &= \left(S_{\Delta t} - I\right)\left(X_{t_0} - O_{t_0}\right) + \int_{t_0}^t S_{t-s}\, F(X_{t_0})\, ds \\ &\quad + \int_{t_0}^t S_{t-s} \int_0^1 F'(X_{t_0} + r\Delta X_s)(I_*^0(s))\, dr\, ds \\ &\quad + \int_{t_0}^t S_{t-s} \int_0^1 F'(X_{t_0} + r\Delta X_s)(\Delta O_s)\, dr\, ds, \end{aligned}$$

which implies that

$$I_*^0(t) = I^0(t) + I_*^1[I_*^0](t) + I_*^1[\Delta O](t)$$

for all $t \in [t_0, T]$. □

An iterated application of Proposition 5 then allows Taylor expansions to be derived step by step, as will be seen below.

The corresponding Taylor approximation is obtained by omitting the remainder terms in the Taylor expansion.

Write $Y_t = O((\Delta t)^r)$ for a stochastic process $Y \in \mathcal{C}$ and some real number $r \in (0, \infty)$ to denote that

$$\sup_{t \in (t_0, T]} \frac{|Y_t(\omega)|}{(t - t_0)^r} < \infty$$

holds for every $\omega \in \Omega$.

The following proposition will allow the order of a Taylor approximation to be estimated.

Proposition 6. *Let Assumptions 1–5 hold. Then* $I^0(t) = O(\Delta t)$ *and* $I^0_*(t) = O(\Delta t)$. *Moreover, if* $g_1, \ldots, g_n \in \mathcal{C}$ *for* $n \in \mathbb{N}$ *satisfy* $g_1(t) = O((\Delta t)^{\alpha_1})$, …, $g_n(t) = O((\Delta t)^{\alpha_n})$ *with* $\alpha_1, \ldots, \alpha_n \in (0, \infty)$, *then*

$$I^n[g_1, \ldots, g_n](t) = O\left((\Delta t)^{1+\alpha_1+\ldots+\alpha_n}\right), \quad I^n_*[g_1, \ldots, g_n](t) = O\left((\Delta t)^{1+\alpha_1+\ldots+\alpha_n}\right). \quad (14)$$

Proof. Define the $\mathcal{F}/\mathcal{B}([0, \infty))$-measurable mapping $R : \Omega \to [0, \infty)$ by

$$\begin{aligned} R(\omega) \;:=\; & 1 + \sup_{0 \le t \le T} |F(X_t(\omega))| + \sup_{0 \le t \le T} \|S_t\| \\ & + \sup_{0 \le t_1 < t_2 \le T} \frac{|(X_{t_2}(\omega) - O_{t_2}(\omega)) - (X_{t_1}(\omega) - O_{t_1}(\omega))|}{t_2 - t_1} \end{aligned}$$

for every $\omega \in \Omega$, which is indeed finite and hence well defined by Assumptions 1–5 and Lemma 3.

From the previous lecture

$$(X_{t_2} - O_{t_2}) - (X_{t_1} - O_{t_1})$$

$$\begin{aligned}
&= (S_{t_2} - S_{t_1})\,\xi + \int_0^{t_2} S_{t_2-s}\,F(X_s)\,ds - \int_0^{t_1} S_{t_1-s}\,F(X_s)\,ds \\
&= (S_{t_2-t_1} - I)\left(S_{t_1}\xi + \int_0^{t_1} S_{t_1-s}\,F(X_s)\,ds\right) + \int_{t_1}^{t_2} S_{t_2-s}\,F(X_s)\,ds \\
&= (S_{t_2-t_1} - I)\,(X_{t_1} - O_{t_1}) + \int_{t_1}^{t_2} S_{t_2-s}\,F(X_s)\,ds
\end{aligned}$$

for all $0 \le t_1 < t_2 \le T$.

Hence

$$|(S_{t_2-t_1} - I)\,(X_{t_1} - O_{t_1})|$$

$$\begin{aligned}
&= \left|(X_{t_2} - O_{t_2}) - (X_{t_1} - O_{t_1}) - \int_{t_1}^{t_2} S_{t_2-s}\,F(X_s)\,ds\right| \\
&\le |(X_{t_2} - O_{t_2}) - (X_{t_1} - O_{t_1})| + \left|\int_{t_1}^{t_2} S_{t_2-s}\,F(X_s)\,ds\right| \\
&\le R\,(t_2 - t_1) + R^2\,(t_2 - t_1) \le 2R^2\,(t_2 - t_1)
\end{aligned}$$

for all $0 \le t_1 < t_2 \le T$.

It follows that

$$\begin{aligned}
\left|I^0(t)\right| &= \left|\left(S_{t-t_0} - I\right)\left(X_{t_0} - O_{t_0}\right) + \int_{t_0}^{t} S_{t-s}\, F(X_{t_0})\; ds\right| \\
&\leq \left|\left(S_{t-t_0} - I\right)\left(X_{t_0} - O_{t_0}\right)\right| + \left|\int_{t_0}^{t} S_{t-s}\, F(X_{t_0})\, ds\right| \\
&\leq 2R^2\left(t - t_0\right) + \int_{t_0}^{t} \|S_{t-s}\|\, |F(X_{t_0})|\; ds \\
&\leq 2R^2\left(t - t_0\right) + R^2\left(t - t_0\right) \leq 3R^2\left(t - t_0\right)
\end{aligned}$$

and

$$\begin{aligned}
\left|I_*^0(t)\right| &= \left|\left(S_{t-t_0} - I\right)\left(X_{t_0} - O_{t_0}\right) + \int_{t_0}^{t} S_{t-s}\, F(X_s)\, ds\right| \\
&\leq \left|\left(S_{t-t_0} - I\right)\left(X_{t_0} - O_{t_0}\right)\right| + \left|\int_{t_0}^{t} S_{t-s}\, F(X_s)\, ds\right| \\
&\leq 2R^2\left(t - t_0\right) + \int_{t_0}^{t} \|S_{t-s}\|\, |F(X_s)|\; ds \\
&\leq 2R^2\left(t - t_0\right) + R^2\left(t - t_0\right) \leq 3R^2\left(t - t_0\right)
\end{aligned}$$

for every $t \in [t_0, T]$, which shows

$$I^0(t) = O\left(\Delta t\right) \qquad \text{and} \qquad I_*^0(t) = O\left(\Delta t\right).$$

The rest of the proof is by induction. □

7 Abstract examples of Taylor expansions

Proposition 5 will be used here to derive various Taylor expansions of the solution X of the SPDE (3) and Proposition 6 will be used to determine their orders of convergence.

Taylor expansion of order 1

The simplest Taylor expansion of the increment ΔX is given by the basic formula (15), i.e.,

$$\Delta X = I_*^0 + \Delta O. \tag{15}$$

or $\Delta X_t = I_*^0(t) + \Delta O_t$ for $t \in [t_0, T]$.

If the solution X is known at time t_0 only, the expression $I_*^0(t)$ cannot be used to determine $\Delta X_t = X_t - X_{t_0}$ for $t \in (t_0, T]$ since it contains the unknown solution path X_s with $s \in [t_0, t]$.

Omitting I_*^0 gives the approximation $\Delta X \approx \Delta O$, which can also be written as

$$X_t \approx X_{t_0} + O_t - O_{t_0}, \qquad t \in [t_0, T]. \tag{16}$$

By Proposition 6 the remainder term I_*^0 of the Taylor approximation (16) can be estimated by $I_*^0(t) = O(\Delta t)$, so

$$X_t = X_{t_0} + O_t - O_{t_0} + O(\Delta t) \tag{17}$$

gives the simplest Taylor approximation of the solution process X.

Taylor expansion of order $1 + \theta$

Proposition 5 can now be used to derive a higher order expansion of the basic formula (15).

More precisely, formula (12) is inserted to the remainder stochastic process I^0_*, i.e.,

$$I^0_* = I^0 + I^1_*[I^0_*] + I^1_*[\Delta O]$$

into the basic formula (15) to give

$$\Delta X = \left(I^0 + I^1_*[I^0_*] + I^1_*[\Delta O]\right) + \Delta O,$$

which can also be written as

$$\Delta X = I^0 + \Delta O + \left(I^1_*[I^0_*] + I^1_*[\Delta O]\right). \tag{18}$$

Omitting the double integral terms $I^1_*[I^0_*]$ and $I^1_*[\Delta O]$ gives the Taylor approximation

$$\Delta X \approx I^0 + \Delta O, \tag{19}$$

which, using the definition of the stochastic process I^0, is equal to

$$\Delta X_t \approx (S_{\Delta t} - I)\,(X_{t_0} - O_{t_0}) + \left(\int_0^{\Delta t} S_s\, ds\right) F(X_{t_0}) + \Delta O_t$$

for $t \in [t_0, T]$.

Hence

$$\boxed{X_t \approx S_{\Delta t} X_{t_0} + \left(\int_0^{\Delta t} S_s \, ds \right) F(X_{t_0}) + (O_t - S_{\Delta t} O_{t_0}) \quad (20)}$$

is another Taylor approximation for the solution of the SPDE (3).

By Proposition 6, $I_*^0(t) = O(\Delta t)$ and, by Assumption 3, $\Delta O_t = O\left((\Delta t)^\theta\right)$. Hence from (14) it follows that

$$I_*^1[I_*^0](t) = O\left((\Delta t)^2\right), \qquad I_*^1[\Delta O](t) = O\left((\Delta t)^{1+\theta}\right).$$

Since $1 + \theta < 2$, these imply that

$$I_*^1[I_*^0](t) + I_*^1[\Delta O](t) = O\left((\Delta t)^{1+\theta}\right).$$

Hence, from (18),

$$\Delta X = I^0 + \Delta O + O\left((\Delta t)^{1+\theta}\right)$$

It then follows from the definition of the stochastic process $I^0 \in \mathcal{C}$ that

$$\boxed{X_t = S_{\Delta t} X_{t_0} + \left(\int_0^{\Delta t} S_s \, ds \right) F(X_{t_0}) + (O_t - S_{\Delta t} O_{t_0}) + O\left((\Delta t)^{1+\theta}\right). \quad (21)}$$

The Taylor approximation (21) plays an analogous role to the strong order $\gamma = 0.5$ Taylor expansion giving the Euler–Maruyama scheme for finite dimensional SODEs.

It will be called the exponential Euler approximation since it gives the exponential Euler scheme for SPDES with additive noise.

Taylor expansion of order $1 + \min(1, 2\theta)$

Further expansions of the remainder terms in a Taylor expansion give a Taylor expansion of higher order.

The remainder term of (18) consists of two terms: $I^1_*[I^0_*]$ and $I^1_*[\Delta O]$.

There are now two possibilities for obtaining a higher order Taylor expansion: expand either

$$I^1_*[I^0_*](t) = O\left((\Delta t)^2\right) \qquad \text{or} \qquad I^1_*[\Delta O](t) = O\left((\Delta t)^{1+\theta}\right).$$

Since $1 + \theta < 2$ by Assumption 3, the stochastic process $I^1_*[\Delta O]$ is of lower order than $I^1_*[I^0_*]$.

Hence the term $I^1_*[\Delta O]$ should be expanded to improve on the approximation order $1 + \theta$ of the Taylor approximation (21).

More precisely, from (13),

$$I^1_*[\Delta O] = I^1[\Delta O] + I^2_*[I^0_*, \Delta O] + I^2_*[\Delta O, \Delta O],$$

which is inserted into (18) to yield

$$\Delta X = \left(I^0 + \Delta O + I^1[\Delta O]\right) + R, \qquad (22)$$

where the remainder term $R \in \mathcal{C}$ is given by

$$R = I^1_*[I^0_*] + I^2_*[I^0_*, \Delta O] + I^2_*[\Delta O, \Delta O]. \qquad (23)$$

By Proposition 6

$$I_*^1[I_*^0](t) = O\left((\Delta t)^2\right), \qquad I_*^2[I_*^0, \Delta O](t) = O\left((\Delta t)^{2+\theta}\right),$$

$$I_*^2[\Delta O, \Delta O](t) = O\left((\Delta t)^{1+2\theta}\right).$$

Since $\min(2, 2+\theta, 1+2\theta) = 1 + \min(1, 2\theta)$, it follows that

$$R = O\left((\Delta t)^{1+\min(1,2\theta)}\right).$$

Thus

$$\Delta X_t = I^0(t) + \Delta O_t + I^1[\Delta O](t) + O\left((\Delta t)^{1+\min(1,2\theta)}\right).$$

This can also be written as

$$\begin{aligned} X_t &= S_{\Delta t} X_{t_0} + \left(\int_0^{\Delta t} S_s \, ds\right) F(X_{t_0}) + (O_t - S_{\Delta t} O_{t_0}) \\ &+ \int_{t_0}^{t} S_{t-s} F'(X_{t_0}) \Delta O_s \, ds + O\left((\Delta t)^{1+\min(1,2\theta).}\right) \end{aligned} \tag{24}$$

This is a Taylor approximation of order $1 + \min(1, 2\theta)$.

Taylor expansion of order $1 + \min(1, 3\theta)$

The remainder term (23) consists of three parts, namely

$$I_*^1[I_*^0](t) = O\left((\Delta t)^2\right), \; I_*^2[I_*^0, \Delta O](t) = O\left((\Delta t)^{2+\theta}\right),$$

$$I_*^2[\Delta O, \Delta O](t) = O\left((\Delta t)^{1+2\theta}\right).$$

Let $\theta < \frac{1}{2}$. Then $\min(2, 2+\theta, 1+2\theta) = 1+2\theta$ and the stochastic process $I_*^2[\Delta O, \Delta O]$ will be expanded here.

Applying Propostion 5 to this term yields

$$I_*^2[\Delta O, \Delta O] = I^2[\Delta O, \Delta O] + I_*^3[I_*^0, \Delta O, \Delta O] + I_*^3[\Delta O, \Delta O, \Delta O]$$

and inserting this into (22) then gives

$$\Delta X = I^0 + \Delta O + I^1[\Delta O] + I^2[\Delta O, \Delta O] + R$$

with remainder

$$R = I_*^1[I_*^0] + I_*^2[I_*^0, \Delta O] + I_*^3[I_*^0, \Delta O, \Delta O] + I_*^3[\Delta O, \Delta O, \Delta O].$$

Then $R = O\left((\Delta t)^{1+\min(1,3\theta)}\right)$, since by Proposition 6

$$I_*^3[I_*^0, \Delta O, \Delta O](t) = O\left((\Delta t)^{2+2\theta}\right), \qquad I_*^3[\Delta O, \Delta O, \Delta O](t) = O\left((\Delta t)^{1+3\theta}\right),$$

because here $\min(2, 2+\theta, 2+2\theta, 1+3\theta) = 1 + \min(1, 3\theta)$.

The resulting Taylor expansion is

$$\begin{aligned} X_t &= S_{\Delta t} X_{t_0} + \left(\int_0^{\Delta t} S_s \, ds\right) F(X_{t_0}) + (O_t - S_{\Delta t} O_{t_0}) + \int_{t_0}^t S_{t-s} F'(X_{t_0}) \Delta O_s \, ds \\ &\quad + \frac{1}{2} \int_{t_0}^t S_{t-s} F''(X_{t_0}) \left(\Delta O_s, \Delta O_s\right) \, ds + O\left((\Delta t)^{1+\min(1,3\theta)}\right). \end{aligned} \tag{25}$$

Taylor expansion of order $1 + \min(2, 1 + 2\theta, 6\theta)$

The expansions become quite complicated and depend on the values taken by θ.

For example, if $\theta < \frac{1}{4}$, then the Taylor expansion of order $1 + \min(2, 1 + 2\theta, 6\theta)$ is

$$
\begin{aligned}
X_t \;=\; & S_{\Delta t} X_{t_0} + \left(\int_0^{\Delta t} S_s \, ds \right) F(X_{t_0}) + (O_t - S_{\Delta t} O_{t_0}) \\
& + \int_{t_0}^{t} S_{t-s} F'(X_{t_0}) \left((S_{\Delta s} - I)(X_{t_0} - O_{t_0}) + \left(\int_0^{\Delta s} S_u \, du \right) F(X_{t_0}) \right) ds \\
& + \int_{t_0}^{t} S_{t-s} F'(X_{t_0}) \Delta O_s \, ds + \frac{1}{2} \int_{t_0}^{t} S_{t-s} F''(X_{t_0}) (\Delta O_s, \Delta O_s) \, ds \\
& + \frac{1}{6} \int_{t_0}^{t} S_{t-s} F^{(3)}(X_{t_0}) (\Delta O_s, \Delta O_s, \Delta O_s) \, ds \\
& + \frac{1}{24} \int_{t_0}^{t} S_{(t-s)} F^{(4)}(X_{t_0}) (\Delta O_s, \Delta O_s, \Delta O_s, \Delta O_s) \, ds \qquad (26) \\
& + \frac{1}{120} \int_{t_0}^{t} S_{t-s} F^{(5)}(X_{t_0}) (\Delta O_s, \Delta O_s, \Delta O_s, \Delta O_s, \Delta O_s) \, ds \\
& + \int_{t_0}^{t} S_{t-s} F''(X_{t_0}) \left((S_{\Delta s} - I)(X_{t_0} - O_{t_0}) + \left(\int_0^{\Delta s} S_u \, du \right) F(X_{t_0}), \Delta O_s \right) ds \\
& + \int_{t_0}^{t} S_{t-s} F'(X_{t_0}) \left(\int_{t_0}^{s} S_{(s-u)} F'(X_{t_0}) \Delta O_u \; du \right) ds + O\left((\Delta t)^{1 + \min(2, 1+2\theta, 6\theta)} \right).
\end{aligned}
$$

The Taylor approximations clearly become increasingly cumbersome.

The terms that need to be included can be characterised succinctly with concepts of stochastic trees and stochastic woods of appropriate indices.

The corresponding notation will be explained later in the more general setting of SPDEs with multiplicative noise.

8 Examples: Space–time white noise

Assumption 3 is satisfied here with $\theta = \frac{1}{4} - \varepsilon$ for an arbitrarily small $\varepsilon \in (0, \frac{1}{4})$ and Taylor approximations (21), (24) , (25) and (26) with this parameter are

$$X_t = S_{\Delta t} X_{t_0} + \left(\int_0^{\Delta t} S_s \, ds \right) F(X_{t_0}) + (O_t - S_{\Delta t} O_{t_0}) + O\left((\Delta t)^{\frac{5}{4} - \varepsilon} \right)$$

$$\begin{aligned} X_t &= S_{\Delta t} X_{t_0} + \left(\int_0^{\Delta t} S_s \, ds \right) F(X_{t_0}) + (O_t - S_{\Delta t} O_{t_0}) \\ &\quad + \int_{t_0}^{t} S_{t-s} F'(X_{t_0}) \Delta O_s \, ds + O\left((\Delta t)^{\frac{3}{2} - \varepsilon} \right) \end{aligned}$$

$$\begin{aligned} X_t &= S_{\Delta t} X_{t_0} + \left(\int_0^{\Delta t} S_s \, ds \right) F(X_{t_0}) + (O_t - S_{\Delta t} O_{t_0}) + \int_{t_0}^{t} S_{t-s} F'(X_{t_0}) \Delta O_s \, ds \\ &\quad + \frac{1}{2} \int_{t_0}^{t} S_{t-s} F''(X_{t_0}) \left(\Delta O_s, \Delta O_s \right) ds + O\left((\Delta t)^{\frac{7}{4} - \varepsilon} \right) \end{aligned}$$

$$\begin{aligned} X_t &= S_{\Delta t} X_{t_0} + \left(\int_0^{\Delta t} S_s \, ds \right) F(X_{t_0}) + (O_t - S_{\Delta t} O_{t_0}) \\ &\quad + \int_{t_0}^{t} S_{t-s} F'(X_{t_0}) \left((S_{\Delta s} - I)(X_{t_0} - O_{t_0}) + \left(\int_0^{\Delta s} S_u \, du \right) F(X_{t_0}) \right) ds \\ &\quad + \int_{t_0}^{t} S_{t-s} F'(X_{t_0}) \Delta O_s \, ds + \frac{1}{2} \int_{t_0}^{t} S_{t-s} F''(X_{t_0}) \left(\Delta O_s, \Delta O_s \right) ds \\ &\quad + \frac{1}{6} \int_{t_0}^{t} S_{t-s} F^{(3)}(X_{t_0}) \left(\Delta O_s, \Delta O_s, \Delta O_s \right) ds \\ &\quad + \frac{1}{24} \int_{t_0}^{t} S_{t-s} F^{(4)}(X_{t_0}) \left(\Delta O_s, \Delta O_s, \Delta O_s, \Delta O_s \right) ds \\ &\quad + \frac{1}{120} \int_{t_0}^{t} S_{t-s} F^{(5)}(X_{t_0}) \left(\Delta O_s, \Delta O_s, \Delta O_s, \Delta O_s, \Delta O_s \right) ds \\ &\quad + \int_{t_0}^{t} S_{t-s} F''(X_{t_0}) \left((S_{\Delta s} - I)(X_{t_0} - O_{t_0}) + \left(\int_0^{\Delta s} S_u du \right) F(X_{t_0}), \Delta O_s \right) ds \\ &\quad + \int_{t_0}^{t} S_{t-s} F'(X_{t_0}) \left(\int_{t_0}^{s} S_{s-u} F'(X_{t_0}) \Delta O_u \, du \right) ds + O\left((\Delta t)^{\frac{5}{2} - \varepsilon} \right) \end{aligned}$$

9 Examples: a nonlinear SPDE

Consider the above Taylor approximations for the the concrete nonlinear SPDE

$$\boxed{dX_t = \left[\frac{\partial^2}{\partial x^2} X_t + X_t - X_t^3\right] dt + \frac{1}{2}\, dW_t,}$$

in the function space $V = C([0,1], \mathbb{R})$ with the supremum norm. Here ε is always an arbitrarily small real number in $(0, \frac{1}{4})$.

$$\boxed{X_t = S_{\Delta t} X_{t_0} + \left(\int_0^{\Delta t} S_s\, ds\right)(X_{t_0} - X_{t_0}^3) + (O_t - S_{\Delta t} O_{t_0}) + O\left((\Delta t)^{\frac{5}{4}-\varepsilon}\right) \quad (27)}$$

$$\boxed{\begin{aligned} X_t &= S_{\Delta t} X_{t_0} + \left(\int_0^{\Delta t} S_s\, ds\right)(X_{t_0} - X_{t_0}^3) + (O_t - S_{\Delta t} O_{t_0}) \\ &\quad + \int_{t_0}^t S_{t-s}\left(\left(1 - 3X_{t_0}^2\right)\Delta O_s\right) ds + O\left((\Delta t)^{\frac{3}{2}-\varepsilon}\right) \end{aligned} \quad (28)}$$

$$\boxed{\begin{aligned} X_t &= S_{\Delta t} X_{t_0} + \left(\int_0^{\Delta t} S_s\, ds\right)(X_{t_0} - X_{t_0}^3) + (O_t - S_{\Delta t} O_{t_0}) \\ &\quad + \int_{t_0}^t S_{t-s}\left((1 - 3X_{t_0}^2)\Delta O_s\right)\, ds - 3\int_{t_0}^t S_{t-s}\left(X_{t_0}(\Delta O_s)^2\right)\, ds + O\left((\Delta t)^{\frac{7}{4}-\varepsilon}\right) \end{aligned} \quad (29)}$$

$$\boxed{\begin{aligned} X_t &= S_{\Delta t} X_{t_0} + \int_{t_0}^t S_{t-s}\left(X_{t_0} + \Delta O_s - (X_{t_0} + \Delta O_s)^3\right)\, ds + (O_t - S_{\Delta t} O_{t_0}) \\ &\quad + \int_{t_0}^t S_{t-s}\left(\left(1 - 3X_{t_0}^2\right)\left(\left(S_{\Delta s} - I\right)(X_{t_0} - O_{t_0})\right)\right)\, ds \\ &\quad + \int_{t_0}^t S_{t-s}\left(\left(1 - 3X_{t_0}^2\right)\left(\left(\int_0^{\Delta s} S_u\, du\right)\left(X_{t_0} - X_{t_0}^3\right)\right)\right)\, ds \\ &\quad - 6\int_{t_0}^t S_{t-s}\left(X_{t_0}\Delta O_s\left(\left(S_{\Delta s} - I\right)(X_{t_0} - O_{t_0})\right)\right)\, ds \\ &\quad - 6\int_{t_0}^t S_{t-s}\left(X_{t_0}\Delta O_s\left(\left(\int_0^{\Delta s} S_u\, du\right)\left(X_{t_0} - X_{t_0}^3\right)\right)\right)\, ds \\ &\quad + \int_{t_0}^t S_{t-s}\left(1 - 3X_{t_0}^2\right)\left(\int_{t_0}^s S_{s-u}\left((1 - 3X_{t_0}^2)\Delta O_u\right)\, du\right)\, ds + O\left((\Delta t)^{\frac{5}{2}-\varepsilon}\right) \end{aligned} \quad (30)}$$

10 Examples: Commonly used approximations

The linear–implicit Euler and linear–implicit Crank–Nicolson approximations are commonly used approximations in the literature.

For the above nonlinear SPDE they are, respectively,

$$X_t \approx (I - \Delta t A)^{-1} \left(X_{t_0} + \Delta t \left(X_{t_0} - X_{t_0}^3 \right) \right) + \int_{t_0}^{t} (I - \Delta t A)^{-1} \, dW_s \quad (31)$$

and

$$\begin{aligned} X_t \;\approx\; & \left(I + \frac{1}{2}\Delta t A \right) \left(I - \frac{1}{2}\Delta t A \right)^{-1} X_{t_0} + \Delta t \left(I - \frac{1}{2}\Delta t A \right)^{-1} \left(X_{t_0} - X_{t_0}^3 \right) \\ & + \int_{t_0}^{t} \left(I - \frac{1}{2}\Delta t A \right)^{-1} dW_s, \end{aligned} \quad (32)$$

where W_t is a cylindrical I-Wiener process and A is the Laplacian with Dirichlet boundary conditions.

The approximations (31) and (32) approximate the exact solution locally in time with the order $\frac{1}{4} - \varepsilon$.

In contrast, the above Taylor approximations approximate it locally in time with the orders 1, $\frac{5}{4} - \varepsilon$, $\frac{3}{2} - \varepsilon$, $\ldots$, $\frac{5}{2} - \varepsilon$.

Computational burden versus higher order

The iterated integrals in the Taylor approximations need to be computed and it is natural to ask if the extra computational burden is really worthwhile.

To answer this, note firstly that Taylor approximations are not only useful for computational purposes, but are also important from a theoretical point of view, i.e., for understanding local properties of the solution process.

In addition, for example, the Taylor approximation in equation (27) approximates the exact solution with the order $\frac{5}{4} - \varepsilon$, but with the same computational effort! (up to a constant) as the Euler scheme.

Moreover, it is quiet easy to derive other approximations, which attain a higher order and involve just a few more terms.

Finally, Taylor approximations such as (27)–(30) provide a theoretical basis for deriving various higher order one–step numerical schemes for reaction–diffusion like SPDEs with additive noise.

11 Numerical Schemes from Taylor Expansions

The numerical approximation of the SPDE (3) requires the discretisation of the solution on both the time interval $[0,T]$ and the infinite dimensional space V.

The above Taylor expansions provide the temporal discretisation in the space V, while the space discretisation is introduced in the following assumption, which is motivated by Galerkin approximations.

Assumption 6 (Projection operators)**.** *For each $N \in \mathbb{N}$, let $P_N : V \to V$ be a bounded linear operator with $(P_N)^2 = P_N$ and $P_N S_t = S_t P_N$ for all $t \in [0,\infty)$.*

- Denote the image of P_N in V by $V_N := P_N(V) = \mathrm{im}(P_N) \subset V$.

Since the operator $P_N : V \to V$ is linear, $(V_N, |\cdot|)$ is a real normed linear subspace of V for each $N \in \mathbb{N}$ and, by Assumption 6, $P_N(v) = v$ for all $v \in V_N$.

In addition, V_N is separable and its Borel σ-algebra satisfies $\mathcal{B}(V_N) = \mathcal{B}(V) \cap V_N$.

Moreover, the spaces V_N are invariant with respect to the semigroup S_t, i.e., $S_t(V_N) \subset V_N$ for all $t \in [0,\infty)$ by to Assumption 6.

Usually the spaces V_N are finite dimensional. Then P_N projects the infinite dimensional SPDE (3) down onto the finite dimensional space V_N, where numerical computations can be done.

The semigroup S, the nonlinearity F, the driving process O and the initial value ξ will be approximated by mappings $S^N : [0,\infty) \to L(V_N)$ and $F^N : V_N \to V_N$, by a stochastic process $O^N : [0,T] \times \Omega \to V_N \subset V$ and a $\mathcal{F}/\mathcal{B}(V_N)$-measurable random variable $\xi : \Omega \to V_N$ defined, respectively, by

$$\begin{aligned} S_t^N(v) &:= S_t(v), \qquad F_N(v) := P_N\left(F(v)\right), \\ O_s^N(\omega) &:= P_N\left(O_s(\omega)\right), \qquad x_0^N(\omega) := P_N\left(\xi(\omega)\right) \end{aligned}$$

for all $v \in V_N$, $t \in [0, \infty)$, $s \in [0, T]$, $\omega \in \Omega$ and each $N \in \mathbb{N}$.

These truncated mappings S^N satisfy

$$S_0^N = I, \quad S_{t_1}^N S_{t_1}^N = S_{t_1+t_2}^N, \quad \sup_{0\le t\le T} \left\|S_t^N\right\| < \infty, \quad \sup_{0\le s<t\le T} \frac{\left\|S_t^N - S_s^N\right\| s}{t-s} < \infty$$

for all t_1, $t_2 \in [0, \infty)$ by Assumption 1, so are a semigroup on V_N for each $N \in \mathbb{N}$.

In addition, by Assumption 2, $F_N : V_N \to V_N$ is infinitely often Fréchet differentiable and for all $v_0, \ldots, v_n \in V_N$ for every n and N in $\mathbb{N}$

$$F_N^{(n)}(v_0)\left(v_1, \ldots, v_n\right) = P_N\left(F^{(n)}(v_0)\left(v_1, \ldots, v_n\right)\right).$$

Finally, $O^N : [0, T] \times \Omega \to V_N$ has θ-Hölder continuous sample paths for every $N \in \mathbb{N}$, where $\theta \in (0, 1)$ is given in Assumption 3.

Using this set up one–step numerical schemes in the spaces V_N, $N \in \mathbb{N}$, can be derived from the Taylor approximations in the previous sections to approximate the solution process X_t of the SPDE (3).

The iterates of these schemes will be denoted by the $\mathcal{F}/\mathcal{B}(V_N)$-measurable mappings

$$Y_k^{N,M} : \Omega \to V_N \qquad \text{with} \qquad Y_0^{N,M} := \xi^N + O_0^N$$

for $k = 0, 1, \ldots, M$ and N, $M \in \mathbb{N}$.

The exponential Euler scheme

The global convergence order of the Taylor approximation (17) is too low to give a consistent scheme. The next higher order Taylor approximation (21), the exponential Euler approximation, leads to the exponential Euler scheme

$$Y_{k+1}^{N,M} = S_h^N Y_k^{N,M} + \left(\int_0^h S_s^N \, ds \right) F_N(Y_k^{N,M}) + \left(O_{(k+1)h}^N - S_h^N O_{kh}^N \right) \qquad (33)$$

for $k = 0, 1, \ldots, M-1$ for fixed $N, M \in \mathbb{N}$.

The integral $\int_0^h S_s^N \, ds \in L(V_N, V)$ is understood as an $L(V_N, V)$-valued Bochner integral, where $L(V_N, V)$ is the real Banach space of all bounded linear operators from V_N to V. Now

$$P_N \left(\int_0^h S_s^N \, ds \right) = \int_0^h P_N \left(S_s^N \right) \, ds = \int_0^h S_s^N \, ds, \qquad \forall \, N, M \in \mathbb{N},$$

which shows that $\int_0^h S_s^N \, ds$ is in fact in $L(V_N) = L(V_N, V_N)$ for each $N \in \mathbb{N}$. and hence the iterates of the numerical approximation (33) are well defined.

A higher order Taylor scheme

In a similar way, the Taylor approximation (24) leads to the numerical scheme

$$\begin{aligned} Y_{k+1}^{N,M} &= S_h^N Y_k^{N,M} + \left(\int_0^h S_s^N \, ds \right) F_N \left(Y_k^{N,M} \right) + \left(O_{(k+1)h}^N - S_h^N O_{kh}^N \right) \\ &\quad + \int_{kh}^{(k+1)h} S_{(k+1)h-s}^N F_N' \left(Y_k^{N,M} \right) \left(O_s^N - O_{kh}^N \right) \, ds \end{aligned} \qquad (34)$$

for $k = 0, 1, \ldots, M-1$ with fixed N and M in $\mathbb{N}$. The well definedness of the terms here can be shown in the same way as above.

A Runge–Kutta scheme for SPDEs

In principle, the next Taylor approximations can be used to derive numerical schemes of higher order.

These schemes are, however, of limited practical use due to cost and difficulty of computing the higher iterated integrals as well as the higher order derivatives in the Taylor approximations.

This situation is already well known for ODEs, where derivative–free schemes, known as Runge–Kutta schemes, are often used.

Taylor schemes are then used for theoretical puroposes such as for determining the convergence order.

To show how a Runge–Kutta scheme can be derived for SPDEs consider the Taylor approximation (24) on the time subinterval $[kh, (k+1)h]$, namely

$$\begin{aligned}
X_{(k+1)h} &\approx S_h X_{kh} + \left(\int_0^h S_s\, ds\right) F(X_{kh}) + \left(O_{(k+1)h} - S_h O_{kh}\right) \\
&\quad + \int_{kh}^{(k+1)h} S_{(k+1)h-s} F'(X_{kh})\,(O_s - O_{kh})\; ds \\
&\approx S_h \left(X_{kh} + h\, F(X_{kh}) + \int_{kh}^{(k+1)h} F'(X_{kh})\,(O_s - O_{kh})\; ds \right) \\
&\quad + \left(O_{(k+1)h} - S_h O_{kh}\right) \\
&= S_h \left(X_{kh} + h\, F(X_{kh}) + h\, F'(X_{kh}) \left(\frac{1}{h} \int_{kh}^{(k+1)h} (O_s - O_{kh})\, ds \right) \right) \\
&\quad + \left(O_{(k+1)h} - S_h O_{kh}\right)
\end{aligned}$$

for $k = 0,1, \ldots, M-1$ for a fixed $M \in \mathbb{N}$.

The expression in the second last line is the classical Taylor approximation

$$F\left(X_{kh} + \frac{1}{h}\int_{kh}^{(k+1)h} (O_s - O_{kh})\,ds\right) \approx F(X_{kh}) + F'(X_{kh})\left(\frac{1}{h}\int_{kh}^{(k+1)h} (O_s - O_{kh})\;ds\right).$$

Using this in the approximation above gives

$$\begin{aligned} X_{(k+1)h} &\approx S_h\left(X_{kh} + h\,F\left(X_{kh} + \frac{1}{h}\int_{kh}^{(k+1)h} (O_s - O_{kh})\;ds\right)\right) \\ &\quad + \left(O_{(k+1)h} - S_h O_{kh}\right), \end{aligned}$$

and hence the Runga–Kutta scheme

$$\begin{aligned} Y_{k+1}^{N,M} &= S_h^N\left(Y_k^{N,M} + h\,F_N\left(Y_k^{N,M} + \frac{1}{h}\int_{kh}^{(k+1)h} \left(O_s^N - O_{kh}^N\right)\;ds\right)\right) \\ &\quad + \left(O_{(k+1)h}^N - S_h^N O_{kh}^N\right) \end{aligned} \tag{35}$$

for $k = 0, 1, \ldots, M-1$ and fixed N, $M \in \mathbb{N}$.

Reference

S. Becker, A. Jentzen and P. E. Kloeden, *Taylor expansions for stochastic reaction–diffusion equations*, (submitted)

A. Jentzen and P. E. Kloeden, *Taylor-expansions of solutions of stochastic partial differential equations with additive noise*, Annals Probab. (to appear)

Lectures 9 & 10: SPDEs with Multiplicative Noise

The key idea in deriving a higher order Taylor expansion of the solution of a SPDE with non–additive noise is the same as for SPDE with additive noise.

Simply said, one uses classical Taylor expansions of the drift and diffusion coefficients of the SPDE in the mild form and then recursively inserts lower order Taylor expansions of the solution of this SPDE to obtained a closed form with remainder.

Iterating this procedure yields Taylor expansions of the solution of a SPDE of arbitrarily high orders.

These Taylor expansions contain multiple stochastic integrals involving the infinite dimensional Wiener process and the semigroup, which provide more information about the SPDE solution and the noise and, hence, allow an approximation of higher order to be obtained.

This method for deriving Taylor expansions depends strongly on the semigroup approach, i.e., interpreting the SPDE as a mild integral equation.

The multiplicative noise case here is technically more demanding than the additive noise case and is presented for analytical semigroups rather than general ones.

The convergence obtained is strong convergence, not pathwise convergence.

1 Heuristic derivation of Taylor expansions

The underlying idea for deriving Taylor approximations for SPDEs will be sketched briefly again. For simplicity the remainder terms are omitted and assumptions will not be fully stated just now.

Let U and H be real separable Hilbert spaces and let $A : D(A) \subset H \to H$ be a linear operator which generates an analytic semigroup on H., i.e., $e^{At} : H \to H$ for $t \geq 0$.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a normal filtration $\mathcal{F}_t$ and let W_t be a cylindrical I-Wiener process on U with respect to $\mathcal{F}_t$.

Consider the SPDE

$$dX_t = [AX_t + F(X_t)]\,dt + B(X_t)\,dW_t,$$

which is understood in the mild form

$$X_t = e^{At}X_0 + \int_0^t e^{A(t-s)}F(X_s)\,ds + \int_0^t e^{A(t-s)}B(X_s)\,dW_s. \qquad (1)$$

The aim is to derive Taylor approximations of X_t about the pivot time $t_0 \in [0, T)$.

A first simple Taylor approximation of X_t is obtained by omitting the second term in (1) and by using the classical Taylor approximation $B(X_s) \approx B(X_{t_0})$

$$X_t \approx e^{A(t-t_0)}X_{t_0} + \int_{t_0}^t e^{A(t-s)}B(X_{t_0})\,dW_s. \qquad (2)$$

The Taylor approximation (2) is obviously an approximation of low order.

To derive a Taylor approximation of higher order consider the classical Taylor approximations

$$F(X_s) \approx F(X_{t_0}) \tag{3}$$

$$B(X_s) \approx B(X_{t_0}) + B'(X_{t_0})\,(X_s - X_{t_0})\,, \tag{4}$$

where B' denotes the Fréchet derivative of B.

Inserting (3) and (4) into (1) and rearranging terms yields

$$\begin{aligned} X_t \;\approx\; & e^{A(t-t_0)}X_{t_0} + \int_{t_0}^{t} e^{A(t-s)}F(X_{t_0})\,ds + \int_{t_0}^{t} e^{A(t-s)}B(X_{t_0})\,dW_s \\ & + \int_{t_0}^{t} e^{A(t-s)}B'(X_{t_0})\,(X_s - X_{t_0})\;dW_s. \end{aligned} \tag{5}$$

The right hand side of (5) is not a Taylor approximation since the integral

$$\int_{t_0}^{t} e^{A(t-s)}B'(X_{t_0})\,(X_s - X_{t_0})\;dW_s, \qquad t \in [t_0, T],$$

contains the unknown solution X_s for $s \in (t_0, t]$, which is what is to be approximated.

The trick is to replace X_s in the integral in (5) by the lower order Taylor approximation (2).

This yields

$$
\begin{aligned}
X_t \;\approx\; & e^{A(t-t_0)}X_{t_0} + \int_{t_0}^{t} e^{A(t-s)}F(X_{t_0})\,ds + \int_{t_0}^{t} e^{A(t-s)}B(X_{t_0})\,dW_s \\
& + \int_{t_0}^{t} e^{A(t-s)}B'(X_{t_0})\left(e^{A(s-t_0)}X_{t_0} + \int_{t_0}^{s} e^{A(s-u)}B(X_{t_0})\,dW_u - X_{t_0}\right)\,dW_s,
\end{aligned}
$$

which after rearranging gives

$$
\begin{aligned}
X_t \;\approx\; & e^{A(t-t_0)}X_{t_0} + \int_{t_0}^{t} e^{A(t-s)}F(X_{t_0})\,ds + \int_{t_0}^{t} e^{A(t-s)}B(X_{t_0})\,dW_s \\
& + \int_{t_0}^{t} e^{A(t-s)}B'(X_{t_0})\left(\left(e^{A(s-t_0)} - I\right)X_{t_0}\right)\,dW_s \\
& + \int_{t_0}^{t} e^{A(t-s)}B'(X_{t_0})\int_{t_0}^{s} e^{A(s-u)}B(X_{t_0})\,dW_u\,dW_s.
\end{aligned}
\tag{6}
$$

Now only the right hand side of (6) contains the solution X at time t_0, so this is an appropriate Taylor approximation. Moreoever, it has a higher order in the mean–square sense.

The method for deriving the higher order Taylor approximation (6) has two steps:

1) *first, use classical Taylor approximations for the coefficients F and B in the mild integral equation of the SPDE, see* (5)*;*

2) *second, insert recursively a lower order Taylor approximation, see* (2)*.*

The essential ingredients of such Taylor approximations are iterated derivatives of the coefficients and iterated stochastic integrals involving the semigroup generated by the dominant linear operator of the SPDE.

2 Setting and assumptions

Let $(H, \langle \cdot, \cdot \rangle_H, |\cdot|_H)$ and $(U, \langle \cdot, \cdot \rangle_U, |\cdot|_U)$ be two real separable Hilbert spaces.

Let $(D, |\cdot|_D)$ be a real separable Banach space with $H \subset D$ continuously.

Recall that

- $\|\cdot\|_{L^{(n)}(H,H)}$ is the operator norm in the real Banach space $L^{(n)}(H,H)$ of n-multilinear bounded operators from $H^n := H \times \cdots \times H$ to H for each $n \in \mathbb{N}$.

- $\left(L(U,D), \|\cdot\|_{L(U,D)}\right)$ denotes the real Banach space of all bounded linear operators from U to D.

- $(L_{HS}(U,H), \langle \cdot, \cdot \rangle_{HS}, \|\cdot\|_{HS})$ denotes the real Hilbert space of all Hilbert–Schmidt operators from U to H.

Fix $T > 0$ and a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a normal filtration $\mathcal{F}_t$ and let W_t be a cylindrical I-Wiener process on U with respect to $\mathcal{F}_t$.

Assumption 1 (Linear Operator A)**.** *Let $\mathcal{I}$ be a finite or countable set. Moreover, let $(\lambda_i)_{i\in\mathcal{I}}$ be a family of real numbers with $\inf_{i\in\mathcal{I}} \lambda_i > -\infty$ and let $(e_i)_{i\in\mathcal{I}}$ be an orthonormal basis of H. The linear operator $A : D(A) \to H$ is given by*

$$Av = \sum_{i\in\mathcal{I}} -\lambda_i \langle e_i, v \rangle_H e_i$$

for all $v \in D(A)$ with $D(A) = \left\{v \in H : \sum_{i\in\mathcal{I}} |\lambda_i|^2 |\langle e_i, v\rangle_H|^2 < \infty\right\} \subset H$.

The linear operator $A : D(A) \to H$ is closed, densely defined and generates an analytic semigroup $e^{At} : H \to H$, $t \geq 0$.

Assumption 2 (Drift F)**.** *Let $F : H \to H$ be an infinitely often Fréchet differentiable mapping with $\sup_{v\in H} \left\|F^{(n)}(v)\right\|_{L^{(n)}(H,H)} < \infty$ for all $n \in \mathbb{N}$.*

Assumption 3 (Diffusion B)**.** *The embedding $D \subset D((\kappa - A)^{-r})$ is continuous for a given $r \in [0, \infty)$, the operator $B : H \to L(U, D)$ is infinitely often Fréchet differentiable and the operators $e^{At}B^{(n)}(v)(w_1, \ldots, w_n)$ and $(\kappa - A)^{\gamma}e^{At}B(v)$ are Hilbert–Schmidt operators in $L_{HS}(U, H)$. In addition, there exist a family of real positive numbers $(L_n)_{n\in\mathbb{N}}$ and real numbers θ, $\rho \in (0, \frac{1}{2}]$ and $\gamma \in (0, 1)$ such that*

$$\left\|e^{At}B^{(n)}(v)(w_1, \ldots, w_n)\right\|_{HS} \leq L_n \cdot (1 + |v|_H) \cdot |w_1|_H \cdot \ldots \cdot |w_n|_H \cdot t^{\theta-\frac{1}{2}},$$

$$\left\|e^{At}\left(B(v) - B(w)\right)\right\|_{HS} \leq L_0 \cdot |v - w|_H \cdot t^{\rho-\frac{1}{2}},$$

$$\left\|(\kappa - A)^{\gamma} e^{At}B(v)\right\|_{HS} \leq L_0 \cdot (1 + |v|_H) \cdot t^{\rho-\frac{1}{2}}$$

for all v, w, w_1,…,$w_n \in H$, $n \in \{0, 1, 2, \ldots\}$ and all $t \in (0, T]$.

Assumption 4 (Initial value)**.** *Let $x_0 : \Omega \to D((\kappa - A)^{\gamma})$ be an $\mathcal{F}_0$ / $\mathcal{B}\left(D((\kappa - A)^{\gamma})\right)$-measurable mapping with $\mathbb{E}\left|(\kappa - A)^{\gamma} x_0\right|_H^p < \infty$ for some $p \in [1, \infty)$, where $\gamma \in (0, 1)$ is given in Assumption 3.*

Consider the SPDE

$$dX_t = \Big[AX_t + F(X_t) \Big] dt + B(X_t)\, dW_t, \qquad X_0 = x_0 \quad (6)$$

for $t \in [0, T]$.

Under Assumptions 1–4 the SPDE (6) has a unique mild solution.

Proposition 1. *Let Assumptions 1–4 be satisfied and let $\gamma \in (0,1)$ be given by Assumption 3. Then there is an up to modifications unique predictable stochastic process $X : [0,T] \times \Omega \to D((\kappa - A)^\gamma)$ with $\sup_{0 \le t \le T} \mathbb{E}\,|(\kappa - A)^\gamma X_t|_H^p < \infty$ such that*

$$\mathbb{P}\left[X_t = e^{At} x_0 + \int_0^t e^{A(t-s)} F(X_s)\, ds + \int_0^t e^{A(t-s)} B(X_s)\, dW_s \right] = 1 \qquad (7)$$

for all $t \in [0,T]$.

Moreover, X is the unique mild solution of the SPDE (6) in the sense of equation (7).

Reference

A. Jentzen, *Taylor expansions of solutions of stochastic partial differential equations*, Discrete & Cont. Dyn. Systems Series B **14** (2010), 515–557.

3 Example: Stochastic heat equation

The stochastic heat equation with multiplicative space–time white noise on a one dimensional spatial domain provides a simple example of an SPDE, which satisfies Assumptions 1–4.

$$dX_t = \left(\vartheta \frac{\partial^2}{\partial x^2} X_t\right) dt + (\sigma X_t)\, dW_t \tag{8}$$

with

$$X_t(0) = X_t(1) = 0, \qquad X_0 = x_0,$$

where W_t is a cylindrical I-Wiener process on $L^2((0,1),\mathbb{R})$.

Let $H = U = L^2((0,1),\mathbb{R})$ and $D = L^1((0,1),\mathbb{R})$.

The operator $A = \vartheta \dfrac{\partial^2}{\partial x^2}$ with Dirichlet boundary conditions on the one dimensional domain $(0,1)$ has eigenvalues and eigenfunctions

$$e_i(x) = \sqrt{2}\sin(i\pi x), \qquad \lambda_i = \vartheta \pi^2 i^2, \qquad i \in \mathbb{N}.$$

It reduces to

$$Av = \sum_{i=1}^{\infty} -\vartheta \pi^2 i^2 \langle e_i, v\rangle_H e_i, \qquad v \in D(A),$$

with $D(A) = \left\{w \in H \ : \ \sum_{i\in\mathcal{I}} i^4 \left|\langle e_i, w\rangle_H\right|^2 < \infty\right\}$. Thus Assumption 1 is fulfilled.

The drift $F : H \to H$ given by $F(v) \equiv 0$ for all $v \in H$ satisfies Assumption 2. In addition, the operator $B : H \to L(H, D)$ given by

$$(B(v)(w))(x) = \sigma \cdot v(x) \cdot w(x, \qquad x \in (0,1),$$

is infinitely often Fréchet differentiable with the derivatives $B'(v) = B$ and $B^{(i)}(v) \equiv 0$. Assumption 3 is satisfied with the parameters $\gamma = \frac{1}{4} - \varepsilon$ and $\theta = \frac{1}{4}$ for every arbitrarily small $\varepsilon \in (0, \frac{1}{4})$.

Assumption 4 is also fulfilled for the initial value, e.g., $x_0 = \sum_{i=1}^{\infty} i^{-1} e_i$.

4 Taylor expansions for SPDEs

Fix $t_0 \in [0, T)$ and let Π denote the real vector space of all equivalence classes of predictable stochastic processes $Y : [t_0, T] \times \Omega \to H$ with

$$\sup_{t_0 \le t \le T} |Y_t|_{L^p(\Omega;H)} := \sup_{t_0 \le t \le T} \left(\mathbb{E}\,|Y_t|_H^p\right)^{1/p} < \infty$$

for $p \in [1, \infty)$.

Write

$$\Delta X_t := X_t - X_{t_0}, \qquad \Delta t := t - t_0,$$

for $t \in [t_0, T] \subset [0, T]$, where X is the unique solution process of the SPDE (6).

4.1 Integral operators

Let $j \in \{0, 1, 1^*, 2, 2^*\}$, where the indices 0, 1 and 2 label expressions containing only a constant value of the SPDE solution, while 1^* and 2^* label certain integrals with time dependent values of the solution in the integrand.

Define the stochastic processes $I_j^0 \in \Pi$ by

$$I_j^0(t) := \begin{cases} \left(e^{A\Delta t} - I\right) X_{t_0}, & j = 0 \\ \int_{t_0}^t e^{A(t-s)} F(X_{t_0})\, ds, & j = 1 \\ \int_{t_0}^t e^{A(t-s)} F(X_s)\, ds, & j = 1^* \\ \int_{t_0}^t e^{A(t-s)} B(X_{t_0})\, dW_s, & j = 2 \\ \int_{t_0}^t e^{A(t-s)} B(X_s)\, dW_s, & j = 2^* \end{cases}$$

for each $t \in [t_0, T]$.

Let $i \in \mathbb{N}$ and $j \in \{1, 1^*, 2, 2^*\}$. Then, for all $t \in [t_0, T]$ and all $g_1, \ldots, g_i \in \Pi$, define the i-multilinear symmetric mappings $I_j^i : \Pi^i := \underbrace{\Pi \times \cdots \times \Pi}_{i\text{-times}} \to \Pi$ by

$$I_j^i[g_1, \ldots, g_i](t) := \frac{1}{i!} \int_{t_0}^{t} e^{A(t-s)} F^{(i)}(X_{t_0}) \, (g_1(s), \ldots, g_i(s)) \, ds$$

when $j = 1$, by

$$I_j^i[g_1, \ldots, g_i](t) := \int_{t_0}^{t} e^{A(t-s)} \left(\int_0^1 F^{(i)}(X_{t_0} + r\Delta X_s) \, (g_1(s), \ldots, g_i(s)) \, \frac{(1-r)^{i-1}}{(i-1)!} \, dr \right) ds$$

when $j = 1^*$, by

$$I_j^i[g_1, \ldots, g_i](t) := \frac{1}{i!} \int_{t_0}^{t} e^{A(t-s)} B^{(i)}(X_{t_0}) \, (g_1(s), \ldots, g_i(s)) \, dW_s$$

when $j = 2$, and by

$$I_j^i[g_1, \ldots, g_i](t) := \int_{t_0}^{t} e^{A(t-s)} \left(\int_0^1 B^{(i)}(X_{t_0} + r\Delta X_s) \, (g_1(s), \ldots, g_i(s)) \, \frac{(1-r)^{i-1}}{(i-1)!} \, dr \right) dW_s$$

when $j = 2^*$.

These are all well defined by Assumptions 1–4.

The increment ΔX_t of the solution X of the SPDE (6) obviously satisfies

$$\Delta X_t = \left(e^{A\Delta t} - I\right) X_{t_0} + \int_{t_0}^{t} e^{A(t-s)} F(X_s)\, ds + \int_{t_0}^{t} e^{A(t-s)} B(X_s)\, dW_s$$

or, in terms of the above integral operators,

$$\Delta X_t = I_0^0(t) + I_{1^*}^0(t) + I_{2^*}^0(t),$$

for all $t \in [t_0, T]$, which can be written symbolically in the space Π as

$$\Delta X = I_0^0 + I_{1^*}^0 + I_{2^*}^0. \tag{9}$$

The stochastic processes I_0^0, $I_1^i[g_1, \ldots, g_i]$ and $I_2^i[g_1, \ldots, g_i]$ for $g_1, \ldots, g_i \in \Pi$ and $i \in \mathbb{N}$ depend on the solution only at time $t = t_0$ and are therefore useful approximations for the solution X_t, $t \in [t_0, T]$. On the other hand, the stochastic processes $I_{1^*}^i[g_1, \ldots, g_i]$ and $I_{2^*}^i[g_1, \ldots, g_i]$ for $g_1, \ldots, g_i \in \Pi$ and $i \in \mathbb{N}$ depend on the whole solution process X_t for $t \in [t_0, T]$, i.e., on what is to be approximated.

Lemma 1. *Suppose that Assumptions 1–4 hold. Then,*

$$\begin{aligned} I_{1^*}^i[g_1, \ldots, g_i] &= I_1^i[g_1, \ldots, g_i] + I_{1^*}^{i+1}\left[I_0^0, g_1, \ldots, g_i\right] \\ &\quad + I_{1^*}^{i+1}\left[I_{1^*}^0, g_1, \ldots, g_i\right] + I_{1^*}^{i+1}\left[I_{2^*}^0, g_1, \ldots, g_i\right] \end{aligned} \tag{10}$$

and

$$\begin{aligned} I_{2^*}^i[g_1, \ldots, g_i] &= I_2^i[g_1, \ldots, g_i] + I_{2^*}^{i+1}\left[I_0^0, g_1, \ldots, g_i\right] \\ &\quad + I_{2^*}^{i+1}\left[I_{1^*}^0, g_1, \ldots, g_i\right] + I_{2^*}^{i+1}\left[I_{2^*}^0, g_1, \ldots, g_i\right] \end{aligned} \tag{11}$$

for every $g_1, \ldots, g_i \in \Pi$ and every $i \in \mathbb{N}$.

These also hold for $i = 0$, in which case the $g_1, \ldots, g_i$ are omitted.

5 Derivation of simple Taylor expansions

A further expansion of (9) can be obtained by applying formula (10) to the stochastic process $I^0_{1^*}$ and formula (11) to the stochastic process $I^0_{2^*}$, i.e.,

$$\begin{aligned} I^0_{1^*} &= I^0_1 + I^1_{1^*}[I^0_0] + I^1_{1^*}[I^0_{1^*}] + I^1_{1^*}[I^0_{2^*}], \\ I^0_{2^*} &= I^0_2 + I^1_{2^*}[I^0_0] + I^1_{2^*}[I^0_{1^*}] + I^1_{2^*}[I^0_{2^*}] \end{aligned}$$

and inserting these into (9) to obtain

$$\begin{aligned} \Delta X &= I^0_0 + \left(I^0_1 + I^1_{1^*}[I^0_0] + I^1_{1^*}[I^0_{1^*}] + I^1_{1^*}[I^0_{2^*}]\right) \\ &\quad + \left(I^0_2 + I^1_{2^*}[I^0_0] + I^1_{2^*}[I^0_{1^*}] + I^1_{2^*}[I^0_{2^*}]\right). \end{aligned}$$

This can also be written as

$$\Delta X = I^0_0 + I^0_1 + I^0_2 + R \tag{12}$$

with the remainder

$$R = I^1_{1^*}[I^0_0] + I^1_{1^*}[I^0_{1^*}] + I^1_{1^*}[I^0_{2^*}] + I^1_{2^*}[I^0_0] + I^1_{2^*}[I^0_{1^*}] + I^1_{2^*}[I^0_{2^*}].$$

Since the double integral terms $I^1_{1^*}[I^0_j]$, $I^1_{2^*}[I^0_j]$ for $j \in \{0, 1^*, 2^*\}$ in R can be shown to be sufficient small, this gives the approximation

$$\Delta X \approx I^0_0 + I^0_1 + I^0_2.$$

i.e.,

$$X_t \approx e^{A\Delta t}X_{t_0} + \left(\int_0^{\Delta t} e^{As}\,ds\right) F(X_{t_0}) + \int_{t_0}^{t} e^{A(t-s)}B(X_{t_0})\,dW_s, \tag{13}$$

which is a Taylor approximation for the solution of SPDE (6) since the right hand side of (13) depends on the solution only at time $t = t_0$.

Recall that $|Y|_{L^p(\Omega;H)} := (\mathbb{E}\,|Y|_H^p)^{\frac{1}{p}}$ for a $\mathcal{F}/\mathcal{B}(H)$-measurable mapping $Y : \Omega \to H$ and a real number $p \in [1, \infty)$.

Write $Y_t = Z_t + O\left((\Delta t)^r\right)$ for two stochastic processes Y, $Z \in \Pi$ and a real number $r > 0$ if

$$\sup_{t\in(t_0,T]} \left(\frac{|Y_t - Z_t|_{L^p(\Omega;H)}}{(\Delta t)^r} \right) < \infty.$$

It will be shown in Theorem 1 below the above remainder term can be estimated by

$$|R(t)|_{L^p(\Omega;H)} \leq C_p\,(\Delta t)^{\theta+\min(\gamma,\theta)}$$

for every $t \in [t_0, T]$ with constant $C_p \geq 0$ for $p \in [1, \infty)$, which implies that

$$\begin{aligned} X_t \;=\; & e^{A\Delta t}X_{t_0} + \left(\int_0^{\Delta t} e^{As}\,ds\right) F(X_{t_0}) + \int_{t_0}^{t} e^{A(t-s)}B(X_{t_0})\,dW_s \\ & + O\left((\Delta t)^{\theta+\min(\gamma,\theta)}\right). \end{aligned} \tag{14}$$

The approximation (14) has order $\theta + \min(\gamma, \theta)$ in the above strong sense.

It is called the exponential Euler approximation.

Higher order Taylor expansions

Further expansions of the remainder terms in a Taylor expansion give a Taylor expansion of higher order. To illustrate this the terms $I^1_{2^*}[I^0_0]$ and $I^1_{2^*}[I^0_{2^*}]$ in the remainder R in (12) will be expanded.

The expansion formulas (10) and (11) yield

$$I^1_{2^*}[I^0_0] = I^1_2[I^0_0] + I^2_{2^*}[I^0_0, I^0_0] + I^2_{2^*}[I^0_{1^*}, I^0_0] + I^2_{2^*}[I^0_{2^*}, I^0_0]$$

and

$$\begin{aligned} I^1_{2^*}[I^0_{2^*}] &= I^1_2[I^0_{2^*}] + I^2_{2^*}[I^0_0, I^0_{2^*}] + I^2_{2^*}[I^0_{1^*}, I^0_{2^*}] + I^2_{2^*}[I^0_{2^*}, I^0_{2^*}] \\ &= I^1_2[I^0_2] + I^1_2[I^1_{2^*}[I^0_0]] + I^1_2[I^1_{2^*}[I^0_{1^*}]] + I^1_2[I^1_{2^*}[I^0_{2^*}]] \\ &\quad + I^2_{2^*}[I^0_0, I^0_{2^*}] + I^2_{2^*}[I^0_{1^*}, I^0_{2^*}] + I^2_{2^*}[I^0_{2^*}, I^0_{2^*}], \end{aligned}$$

which are inserted into (12) to give

$$\boxed{\Delta X = I^0_0 + I^0_1 + I^0_2 + I^1_2[I^0_0] + I^1_2[I^0_2] + R.}$$

The remainder term can be shown to satisfy $R(t) = O\left((\Delta t)^{\theta+2\min(\gamma,\theta)}\right)$. Thus

$$\begin{aligned} X_t &= e^{A\Delta t}X_{t_0} + \left(\int_0^{\Delta t} e^{As}\,ds\right) F(X_{t_0}) + \int_{t_0}^t e^{A(t-s)}B(X_{t_0})\,dW_s \\ &\quad + \int_{t_0}^t e^{A(t-s)}B'(X_{t_0})\left(e^{A\Delta s} - I\right)X_{t_0}\,dW_s \\ &\quad + \int_{t_0}^t e^{A(t-s)}B'(X_{t_0})\int_{t_0}^s e^{A(s-r)}B(X_{t_0})\,dW_r\,dW_s + O\left((\Delta t)^{\theta+2\min(\gamma,\theta)}\right). \end{aligned}$$

This approximation is of order $\theta + 2\min(\gamma, \theta)$. It was derived heuristically earlier.

6 Stochastic trees and woods

To construct Taylor approximations of arbitrarily high order a succinct way is required to determine and label systematically the integral operators that need to be used.

This is provided by the combinatorial concepts of rooted trees and woods.

Let $N \in \mathbb{N}$ be a natural number and let

$$\mathbf{t}' : \{2, \dots, N\} \to \{1, \dots, N-1\}, \qquad \mathbf{t}'' : \{1, \dots, N\} \to \{0, 1, 2, 1^*, 2^*\}$$

be two mappings with the property that $\mathbf{t}'(j) < j$ for all $j \in \{2, \dots, N\}$.

The pair of mappings $\mathbf{t} = (\mathbf{t}', \mathbf{t}'')$ is called a stochastic tree (S-tree) with $\mathrm{l}(\mathbf{t}) := N$ nodes.

The set of all stochastic trees is denoted by $\mathbf{ST}$.

Every S-tree can be represented as a graph whose nodes are given by the set $\mathrm{nd}(\mathbf{t}) := \{1, \dots, N\}$ and whose arcs are described by the mapping $\mathbf{t}'$ in the sense that there is an edge from j to $\mathbf{t}'(j)$ for every node $j \in \{2, \dots, N\}$.

The mapping $\mathbf{t}''$ is an additional labelling of the nodes with $\mathbf{t}''(j) \in \{0, 1, 2, 1^*, 2^*\}$ indicating the type of node j for every $j \in \mathrm{nd}(\mathbf{t})$.

The set of stochastic woods (S-woods) is defined by

$$\mathbf{SW} := \bigcup_{n=1}^{\infty} (\mathbf{ST})^n.$$

Of course, the embedding $\mathbf{ST} \subset \mathbf{SW}$ holds.

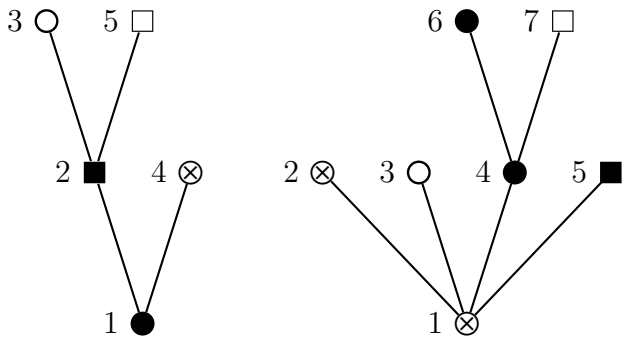

Figure 1: Two examples of stochastic trees

A node $j \in \mathrm{nd}(\mathbf{t}_1)$ is represented by $\otimes$ if $\mathbf{t}_1''(j) = 0$, by $\bullet$ if $\mathbf{t}_1''(j) = 1$, by $\circ$ if $\mathbf{t}_1''(j) = 2$, by $\blacksquare$ if $\mathbf{t}_1''(j) = 1^*$ and, finally, by $\square$ if $\mathbf{t}_1''(j) = 2^*$.

The left picture in Figure 1 corresponds to the tree $\mathbf{t}_1 = (\mathbf{t}_1', \mathbf{t}_1'')$ with $\mathrm{nd}(\mathbf{t}_1) = \{1, 2, 3, 4, 5\}$ given by

$$\mathbf{t}_1'(5) = 2, \qquad \mathbf{t}_1'(4) = 1, \qquad \mathbf{t}_1'(3) = 2, \qquad \mathbf{t}_1'(2) = 1$$

and

$$\mathbf{t}_1''(1) = 1, \qquad \mathbf{t}_1''(2) = 1^*, \qquad \mathbf{t}_1''(3) = 2, \qquad \mathbf{t}_1''(4) = 0, \qquad \mathbf{t}_1''(5) = 2^*.$$

The root is always represented by the lowest node.

The number on the left of a node in Figure 1 is the number of the node of the corresponding tree, while the type of the node is given by $\mathbf{t}_2''$.

The right picture in Figure 1 corresponds to the tree $\mathbf{t}_2 = (\mathbf{t}_2', \mathbf{t}_2'')$ with $\mathrm{nd}(\mathbf{t}_2) = \{1, \ldots, 7\}$ given by

$$\mathbf{t}_2'(7) = 4, \;\; \mathbf{t}_2'(6) = 4, \;\; \mathbf{t}_2'(5) = 1, \;\; \mathbf{t}_2'(4) = 1, \;\; \mathbf{t}_2'(3) = 1, \;\; \mathbf{t}_2'(2) = 1$$

and

$$\mathbf{t}_2''(1) = 0, \;\; \mathbf{t}_2''(2) = 0, \;\; \mathbf{t}_2''(3) = 2, \;\; \mathbf{t}_2''(4) = 1, \;\; \mathbf{t}_2''(5) = 1^*, \;\; \mathbf{t}_2''(6) = 1, \;\; \mathbf{t}_2''(7) = 2^*.$$

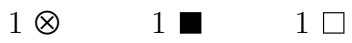

Figure 2: The stochastic wood $\mathbf{w}_0$ in $\mathbf{SW}$

A simple example of an S-wood required later is $\mathbf{w}_0 = (\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3) \in \mathbf{SW}$ with $\mathbf{t}_1$, $\mathbf{t}_2$ and $\mathbf{t}_3$ in $\mathbf{ST}$ given by $\mathrm{l}(\mathbf{t}_1) = \mathrm{l}(\mathbf{t}_2) = \mathrm{l}(\mathbf{t}_3) = 1$ and $\mathbf{t}_1''(1) = 0$, $\mathbf{t}_2''(1) = 1^*$, $\mathbf{t}_3''(1) = 2^*$.

This is illustrated in Figure 2, where the left tree corresponds to $\mathbf{t}_1$, the middle one to $\mathbf{t}_2$ and the right tree corresponds to $\mathbf{t}_3$.

Construction of stochastic trees and woods

Certain stochastic woods in $\mathbf{SW}$ are used to represent Taylor expansions of the solution X of the SPDE (6).

An operator on the set $\mathbf{SW}$ enables the step by step construction of an appropriate stochastic woods.

Let $\mathbf{w} = (\mathbf{t}_1, \ldots, \mathbf{t}_n) \in \mathbf{SW}$ with $n \in \mathbb{N}$ be an S-wood with $\mathbf{t}_i = (\mathbf{t}_i', \mathbf{t}_i'') \in \mathbf{ST}$ for $i \in \{1, \ldots, n\}$.

In addition, let $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, \mathrm{l}(\mathbf{t}_i)\}$ be given and suppose that either $\mathbf{t}_i''(j) = 1^*$ or $\mathbf{t}_i''(j) = 2^*$. In this case the pair (i, j) is called an active node of $\mathbf{w}$.

Denote the set of all active nodes of $\mathbf{w}$ by $\mathrm{acn}(\mathbf{w}) \subset \mathbb{N}^2$.

In Figures of woods and trees, e.g., in Figure 1, active nodes are represented by a square (a filled square) ■ for 1^* and a simple square □ for 2^*.

Introduce the trees $\mathbf{t}_{n+1} = (\mathbf{t}'_{n+1}, \mathbf{t}''_{n+1})$, $\mathbf{t}_{n+2} = (\mathbf{t}'_{n+2}, \mathbf{t}''_{n+2})$, $\mathbf{t}_{n+3} = (\mathbf{t}'_{n+3}, \mathbf{t}''_{n+3})$ $\in \mathbf{ST}$ by $\mathrm{nd}(\mathbf{t}_{n+m}) = \{1, \ldots, \mathrm{l}(\mathbf{t}_i), \mathrm{l}(\mathbf{t}_i) + 1\}$ and

$$\mathbf{t}'_{n+m}(k) = \mathbf{t}'_i(k), \;\; k = 2, \ldots, \mathrm{l}(\mathbf{t}_i), \qquad \mathbf{t}''_{n+m}(k) = \mathbf{t}''_i(k), \;\; k = 1, \ldots, \mathrm{l}(\mathbf{t}_i),$$

$$\mathbf{t}'_{n+m}\left(\mathrm{l}(\mathbf{t}_i) + 1\right) = j, \qquad \mathbf{t}''_{n+m}\left(\mathrm{l}(\mathbf{t}_i) + 1\right) = \begin{cases} 0, & m = 1 \\ 1^*, & m = 2 \\ 2^*, & m = 3 \end{cases}$$

for $m \in \{1, 2, 3\}$.

Further, consider the S-tree $\tilde{\mathbf{t}} = \left(\tilde{\mathbf{t}}', \tilde{\mathbf{t}}''\right) \in \mathbf{ST}$ given by $\tilde{\mathbf{t}}' = \mathbf{t}'_i$, but with $\tilde{\mathbf{t}}'' : \{1, \ldots, \mathrm{l}(\mathbf{t}_i)\} \to \{0, 1, 2, 1^*, 2^*\}$ given by $\tilde{\mathbf{t}}''(k) = \mathbf{t}''_i(k)$ for every $k \in \mathrm{nd}(\mathbf{t}_i) \setminus \{j\}$ and by $\tilde{\mathbf{t}}''(j) = 1$ if $\mathbf{t}''_i(j) = 1^*$ and $\tilde{\mathbf{t}}''(j) = 2$ if $\mathbf{t}''_i(j) = 2^*$.

Then define

$$\mathrm{E}_{(i,j)}(\mathbf{w}) = \mathrm{E}_{(i,j)}\left((\mathbf{t}_1, \ldots, \mathbf{t}_n)\right) := \left(\mathbf{t}_1, \ldots, \mathbf{t}_{i-1}, \tilde{\mathbf{t}}, \mathbf{t}_{i+1}, \ldots \mathbf{t}_{n+3}\right) \in \mathbf{SW}$$

and consider the set of all woods that can be constructed iteratively by applying the $\mathrm{E}_{(i,j)}$ operations, i.e.,

$$\mathbf{SW'} := \{\mathbf{w}_0\} \bigcup \left\{ \mathbf{w} \in \mathbf{SW} \;\; : \;\; \begin{array}{l} \exists\, n \in \mathbb{N}, i_1, \ldots, i_n, j_1, \ldots, j_n \in \mathbb{N} : \forall\, k \in \{1, \ldots, n\} \\ (i_k, j_k) \in \mathrm{acn}\left(\mathrm{E}_{(i_{k-1}, j_{k-1})} \ldots \mathrm{E}_{(i_1, j_1)} \mathbf{w}_0\right), \\ \mathbf{w} = \mathrm{E}_{(i_n, j_n)} \ldots \mathrm{E}_{(i_1, j_1)} \mathbf{w}_0 \end{array} \right\}$$

for the $\mathbf{w}_0$ introduced above in Figure 2.

The following examples using the initial stochastic wood $\mathbf{w}_0$ given in Figure 2 illustrate these definitions.

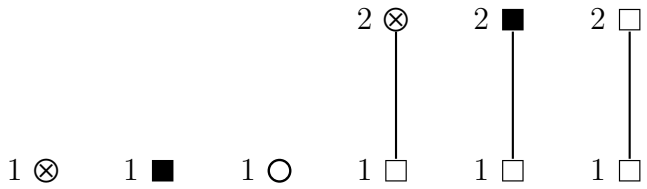

Figure 3: The stochastic wood $\mathbf{w}_1$ in **SW**

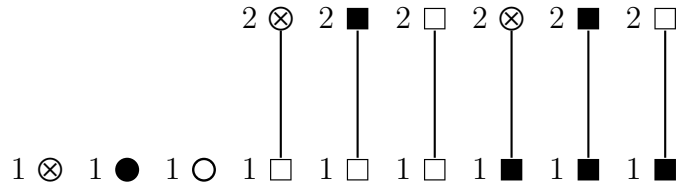

Figure 4: The stochastic wood $\mathbf{w}_2$ in **SW**

Firstly, the active nodes of $\mathbf{w}_0$ are $\mathrm{acn}(\mathbf{w}_0) = \{(2,1),(3,1)\}$, since the first node in the second tree and the first node in the third tree are represented by squares.

Hence, $\mathrm{E}_{(3,1)}\mathbf{w}_0$ is well defined and the resulting stochastic wood $\mathbf{w}_1 = \mathrm{E}_{(3,1)}\mathbf{w}_0$ contains six trees shown in Figure 3. Writing $\mathbf{w}_1 = (\mathbf{t}_1, \ldots, \mathbf{t}_6)$, the left tree in Figure 3 corresponds to $\mathbf{t}_1$, the second tree to $\mathbf{t}_2$ and so on. The active nodes of $\mathbf{w}_1$ are

$$\mathrm{acn}(\mathbf{w}_1) = \{(2,1),(4,1),(5,1),(5,2),(6,1),(6,2)\}$$

so $\mathbf{w}_2 = \mathrm{E}_{(2,1)}\mathbf{w}_1$ is also well defined. See Figure 4.

Figure 5 shows the stochastic wood $\mathbf{w}_3 = \mathrm{E}_{(4,1)}\mathbf{w}_2$, which is well defined since

$$\mathrm{acn}(\mathbf{w}_2) = \{(4,1),(5,1),(5,2),(6,1),(6,2),(7,1),(8,1),(8,2),(9,1),(9,2)\}.$$

For the S-wood $\mathbf{w}_3$

$$\mathrm{acn}(\mathbf{w}_3) = \left\{ \begin{array}{c} (5,1),(5,2),(6,1),(6,2),(7,1),(8,1),(8,2), \\ (9,1),(9,2),(10,1),(11,1),(11,3),(12,1),(12,3) \end{array} \right\}.$$

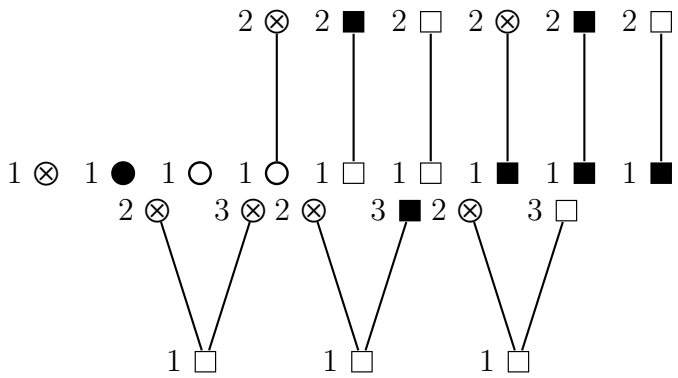

Figure 5: The stochastic wood $\mathbf{w}_3$ in **SW**

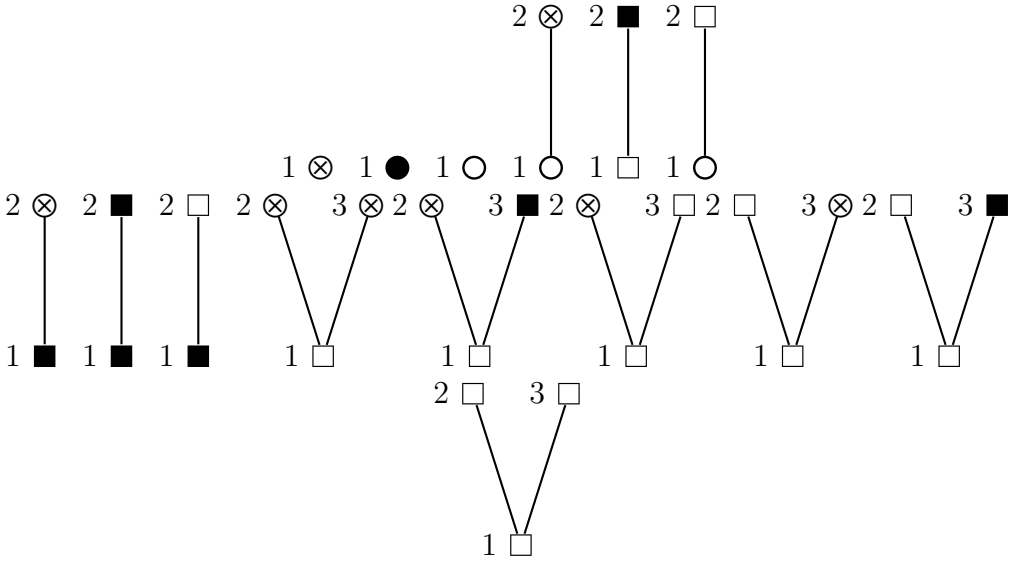

Figure 6: The stochastic wood $\mathbf{w}_4$ in **SW**

Since $(6,1) \in \mathrm{acn}(\mathbf{w}_3)$, the stochastic wood $\mathbf{w}_4 = \mathrm{E}_{(6,1)}\mathbf{w}_3$ given in Figure 6, is also well defined. Its active nodes are

$$\mathrm{acn}(\mathbf{w}_4) = \left\{ \begin{array}{c} (5,1),(5,2),(6,2),(7,1),(8,1),(8,2),(9,1),(9,2), \\ (10,1),(11,1),(11,3),(12,1),(12,3),(13,1), \\ (13,2),(14,1),(14,2),(14,3),(15,1),(15,2),(15,3) \end{array} \right\}.$$

By definition the S-woods $\mathbf{w}_0$, $\mathbf{w}_1$, ..., $\mathbf{w}_5$ are in **SW'**, but the stochastic wood given by Figure 1 is not in **SW'**.

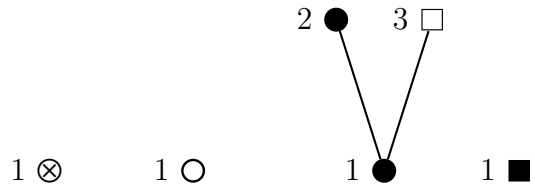

Figure 7: Subtrees of the right tree in Figure 1

Subtrees

Let $\mathbf{t} = (\mathbf{t}', \mathbf{t}'') \in \mathbf{SW}$ be a given S-tree with cardinality $|\mathrm{nd}(\mathbf{t})| \geq 2$.

For two nodes k, $l \in \mathrm{nd}(\mathbf{t})$ with $k \leq l$ the node at l is said to be a grandchild of k if there exists a sequence $k_1 = k < k_2 < \ldots < k_n = l$ of nodes for some $n \in \mathbb{N}$ such that $\mathbf{t}'(k_{v+1}) = k_v$ for every $v \in \{1, \ldots, n-1\}$.

Suppose that $j_1 < \ldots < j_n$ with $n \in \mathbb{N}$ are the nodes of $\mathbf{t}$ such that $\mathbf{t}'(j_i) = 1$ for every $i \in \{1, \ldots, n\}$.

For a given $i \in \{1, \ldots, n\}$ suppose that $j_{i,1}$, …, $j_{i,l_i} \in \mathrm{nd}(\mathbf{t})$ with $l_i \in \mathbb{N}$ and $j_i = j_{i,1} < j_{i,2} < \ldots < j_{i,l_i} \leq \mathrm{l}(\mathbf{t})$ are the grandchildren of $j_i \in \mathrm{nd}(\mathbf{t})$.

- Define the trees $\mathbf{t}_i = (\mathbf{t}_i', \mathbf{t}_i'') \in \mathbf{ST}$ with $\mathrm{l}(\mathbf{t}_i) := l_i$ by $j_{i,\mathbf{t}_i'(k)} = \mathbf{t}'(j_{i,k})$, $\mathbf{t}_i''(k) = \mathbf{t}''(j_{i,k})$ for $k \in \{2, \ldots, l_i\}$ and by $\mathbf{t}_i''(1) = \mathbf{t}''(j_i)$ for every $i \in \{1, \ldots, n\}$.

The trees $\mathbf{t}_1$, …, $\mathbf{t}_n \in \mathbf{ST}$ defined in this way are called subtrees of $\mathbf{t}$, e.g., the subtrees of the right tree in Figure 1 are given in Figure 7.
A node is called a leaf if it has no grandchildren except itself.

$$\mathbf{ST'} := \{\mathbf{t} = (\mathbf{t}', \mathbf{t}'') \in \mathbf{ST} \,:\, (\mathbf{t}''(j) = 0 \Rightarrow j \text{ is a leaf}) \; \forall\, j \in \{1, 2, \ldots, \mathrm{l}(\mathbf{t})\}\}\,.$$

Of course, if $\mathbf{w} = (\mathbf{t}_1, \ldots, \mathbf{t}_n) \in \mathbf{SW'}$ is a stochastic wood in **SW'**, then $\mathbf{t}_1$, …, $\mathbf{t}_n$ $\in$ **ST'**.

Order of stochastic trees and woods

The order of an S-tree $\mathbf{t} \in$ **ST'** is given by the function ordt : **ST'** $\to [0, \infty)$

$$\begin{aligned} \mathrm{ordt}(\mathbf{t}) \;:=\; & |\{j \in \mathrm{nd}(\mathbf{t}) \,:\, \mathbf{t}''(j) = 1 \text{ or } \mathbf{t}''(j) = 1^*\}| \\ & +\gamma\,|\{j \in \mathrm{nd}(\mathbf{t}) \,:\, \mathbf{t}''(j) = 0\}| + \theta\,|\{j \in \mathrm{nd}(\mathbf{t}) \,:\, \mathbf{t}''(j) = 2 \text{ or } \mathbf{t}''(j) = 2^*\}| \end{aligned}$$

for every S-tree $\mathbf{t} = (\mathbf{t}', \mathbf{t}'') \in$ **ST'**.

Example The order of the left tree in Figure 1 is $2+\gamma+2\theta$ since the left tree has one node of type 0, two nodes of type 1 and type 1^*, and also two nodes of type 2 and type 2^*.

A tree $\mathbf{t} = (\mathbf{t}', \mathbf{t}'') \in$ **ST'** is said to be active if there is a $j \in \mathrm{nd}(\mathbf{t})$ such that $\mathbf{t}''(j) = 1^*$ or $\mathbf{t}''(j) = 2^*$, i.e., an S-tree is active if it has an active node.

The order of an S-wood $\mathbf{w} \in$ **SW'** is given by the function ord : **SW'** $\to [0, \infty)$

$$\mathrm{ord}(\mathbf{w}) := \min\{\,\mathrm{ordt}(\mathbf{t}_i) \in [0,\infty) \,:\, i \in \{1,\ldots,n\} \text{ such that } \mathbf{t}_i \text{ is active}\,\}$$

for every S-wood $\mathbf{w} = (\mathbf{t}_1, \ldots, \mathbf{t}_n) \in$ **SW'** with $n \in \mathbb{N}$.

$$\begin{aligned} \mathrm{ord}(\mathbf{w}_0) &= \min\{1, \theta\} = \theta \\ \mathrm{ord}(\mathbf{w}_1) &= \min\{1, \theta+\gamma, \theta+1, \} = \theta + \min(\gamma, \theta) \\ \mathrm{ord}(\mathbf{w}_2) &= \min\{\theta+\gamma, \theta+1, 2\theta, 1+\gamma, 2, 1+\theta\} = \theta + \min(\theta, \gamma) \\ \mathrm{ord}(\mathbf{w}_3) &= \theta + \min(\theta, 2\gamma) = \mathrm{ord}(\mathbf{w}_4) \\ \mathrm{ord}(\mathbf{w}_5) &= \mathbf{w}_5\theta + 2\min(\theta, \gamma) \end{aligned}$$

Stochastic woods and Taylor expansions

Each stochastic wood in $\mathbf{w} \in$ **SW'** characterises a Taylor expansion of the solution X of the SPDE (6).

In particular, every stochastic tree $\mathbf{t}_i \in$ **ST'** for $i \in \{1, \ldots, n\}$ of a stochastic wood $\mathbf{w} = (\mathbf{t}_1, \ldots, \mathbf{t}_n) \in$ **SW'** with $n \in \mathbb{N}$ represents a certain summand in the Taylor expansion.

Consider the mappings ϕ, ψ : **ST'** $\to \Pi$, which are defined recursively as follows:

For a given S-tree $\mathbf{t} = (\mathbf{t}', \mathbf{t}'') \in$ **ST'** define

$$\begin{aligned} \phi(\mathbf{t}) &:= I^0_{\mathbf{t}''(1)} \qquad \text{when} \quad \mathrm{l}(\mathbf{t}) = 1 \\ \phi(\mathbf{t}) &:= I^n_{\mathbf{t}''(1)}\left[\phi(\mathbf{t}_1), \ldots, \phi(\mathbf{t}_n)\right] \qquad \text{when} \quad \mathrm{l}(\mathbf{t}) \geq 2 \end{aligned}$$

where $\mathbf{t}_1, \ldots, \mathbf{t}_n \in$ **ST'** are the subtrees of $\mathbf{t}$ when $\mathrm{l}(\mathbf{t}) \geq 2$.

Note that if $\mathrm{l}(\mathbf{t}) \geq 2$ for a tree $\mathbf{t} = (\mathbf{t}', \mathbf{t}'') \in$ **ST'**, then $\mathbf{t}''(1) \neq 0$.

In addition, for an arbitrary $\mathbf{t} \in$ **ST'** define

$$\begin{aligned} \psi(\mathbf{t}) &:= 0 \quad : \quad \mathbf{t} \text{ an active tree} \\ \psi(\mathbf{t}) &:= \phi(\mathbf{t}) \quad : \quad \text{otherwise} \end{aligned}$$

Finally, define mappings Φ, Ψ : **SW'** $\to \Pi$ by

$$\Phi(\mathbf{w}) = \phi(\mathbf{t}_1) + \ldots + \phi(\mathbf{t}_n), \qquad \Psi(\mathbf{w}) = \psi(\mathbf{t}_1) + \ldots + \psi(\mathbf{t}_n)$$

for every S-wood $\mathbf{w} = (\mathbf{t}_1, \ldots, \mathbf{t}_n) \in$ **SW'** with $n \in \mathbb{N}$.

For example, for the initial stochastic wood $\mathbf{w}_0$ (see Figure 2)

$$\Phi(\mathbf{w}_0) = I_0^0 + I_{1^*}^0 + I_{2^*}^0, \qquad \Psi(\mathbf{w}_0) = I_0^0. \tag{15}$$

Hence, $\Phi(\mathbf{w}_0) = \Delta X$ from (15) and (9).

Furthermore,

$$\Phi(\mathbf{w}_1) = I_0^0 + I_{1^*}^0 + I_2^0 + I_{2^*}^1[I_0^0] + I_{2^*}^1[I_{1^*}^0] + I_{2^*}^1[I_{2^*}^0] \tag{16}$$

and

$$\Psi(\mathbf{w}_1) = I_0^0 + I_2^0 \tag{17}$$

Each stochastic wood $\mathbf{w} \in$ **SW'** thus represents a Taylor expansion

$$X_{t_0} + \Phi(\mathbf{w})(t)$$

and the corresponding Taylor approximation

$$X_{t_0} + \Psi(\mathbf{w})(t)$$

for $t \in [t_0, T]$ of the solution process X of the SPDE (6).

Theorem 1. *Let Assumptions 1–4 be fulfilled, let* $\boldsymbol{w} \in \boldsymbol{SW'}$. *Then there exists a constant* $C_p > 0$ *such that*

$$\mathbb{P}\left[X_t = X_{t_0} + \Phi(\boldsymbol{w})\,(t)\right] = 1,$$

and

$$\left(\mathbb{E}\left[\left|X_t - X_{t_0} - \Psi(\boldsymbol{w})(t)\right|_H^p\right]\right)^{\frac{1}{p}} \leq C_p\,(t-t_0)^{ord(\boldsymbol{w})}$$

holds for every $t \in [t_0, T]$, *where* X *is the up to modifications unique solution of the SPDE* (6).

The constant C_p *here only depends on the S-wood* $\boldsymbol{w}$, p *and the coefficients of the SPDE.*

The Taylor approximation in Theorem 1 can also be written as

$$X_t = X_{t_0} + \Psi(\mathbf{w}) + O\left((\Delta t)^{\mathrm{ord}(\mathbf{w})}\right) \tag{18}$$

for every stochastic wood $\mathbf{w} \in$ **SW'**.

The following lemma shows that there are woods in **SW'** with arbitrarily high orders.

Lemma 2. *Let Assumptions 1–4 be fulfilled. Then*

$$\sup_{\boldsymbol{w} \in \boldsymbol{SW'}} ord\,(\boldsymbol{w}) = \infty.$$

Thus Taylor approximations of arbitrarily high orders can be constructed by successively applying the $\mathrm{E}_{(i,j)}$ operator.

7 Examples of Taylor approximations

Taylor approximations $X_{t_0}+\Psi(\mathbf{w})(t)$ in Theorem 1 and (18) based on the stochastic woods $\mathbf{w} \in \{\mathbf{w}_0, \mathbf{w}_1, \ldots, \mathbf{w}_5\}$ are presented here. Theorem 1 gives

$$\boxed{\left(\mathbb{E}\left[\left|X_t - X_{t_0} - \Psi(\mathbf{w})(t)\right|_H^p\right]\right)^{\frac{1}{p}} \leq C_p\,(\Delta t)^{\mathrm{ord}(\mathbf{w})}}$$

for every $t \in [t_0, T]$ and appropriate constant $C_p > 0$, where $p \in [1, \infty)$.

The orders of the Taylor approximations here depend on the two parameters $\theta \in (0, \frac{1}{2}]$ and $\gamma \in (0, 1)$ in Assumption 3.

Taylor approximation of order θ

The first Taylor expansion of the solution is given by the initial stochastic wood $\mathbf{w}_0$ (see Figure 2), i.e., $\Phi(\mathbf{w}_0) = \Delta X$ is approximated by $\Psi(\mathbf{w}_0)$ with order $\mathrm{ord}(\mathbf{w}_0)$. Specifically, by (15),

$$\boxed{\Psi(\mathbf{w}_0)(t) = (e^{A\Delta t} - I)X_{t_0},}$$

and

$$\boxed{\Phi(\mathbf{w}_0)(t) = (e^{A\Delta t} - I)X_{t_0} + \int_{t_0}^{t} e^{A(t-s)}F(X_s)\,ds + \int_{t_0}^{t} e^{A(t-s)}B(X_s)\,dW_s}$$

for every $t \in [t_0, T]$.

Since $\mathrm{ord}(\mathbf{w}_0) = \theta$, the corresponding Taylor approximation

$$\boxed{X_t = e^{A\Delta t}X_{t_0} + O\left((\Delta t)^{\theta}\right) \qquad (19)}$$

has order θ.

Two Taylor approximations of order $\theta + \min(\gamma, \theta)$

Consider the Taylor approximation given by the S-wood $\mathbf{w}_1$ (see Figure 3) is based on $\Phi(\mathbf{w}_1)$ and $\Psi(\mathbf{w}_1)$, which were given in (16) and (17).

Since $\mathrm{ord}(\mathbf{w}_1) = \theta + \min(\gamma, \theta)$, the Taylor approximation

$$X_t = e^{A\Delta t} X_{t_0} + \int_{t_0}^{t} e^{A(t-s)} B(X_{t_0})\, dW_s + O\left((\Delta t)^{\theta+\min(\gamma,\theta)}\right) \tag{20}$$

has order $\theta + \min(\gamma, \theta)$.

The stochastic wood $\mathbf{w}_2$ (see Figure 4) has order $\theta + \min(\gamma, \theta)$ and the Taylor approximation $\Psi(\mathbf{w}_2)$ of $\Phi(\mathbf{w}_2) = \Delta X$ is given by $\Psi(\mathbf{w}_2) = I_0^0 + I_1^0 + I_2^0$, i.e.,

$$\begin{aligned} X_t &= e^{A\Delta t} X_{t_0} + \left(\int_0^{\Delta t} e^{As}\, ds\right) F(X_{t_0}) + \int_{t_0}^{t} e^{A(t-s)} B(X_{t_0})\, dW_s \\ &\quad + O\left((\Delta t)^{\theta+\min(\gamma,\theta)}\right). \end{aligned} \tag{21}$$

This approximation of order $\theta + \min(\gamma, \theta)$ is the exponential Euler approximation.

It is natural to ask is why one should use the Taylor approximation (21) if it is of the same order as the Taylor approximation (20).

Although both of these Taylor approximations have the same local approximation order, for numerical schemes it is the global approximation order and behaviour that are of primary importance, as will be discussed in subsection 9.

It turns out that the exponential Euler scheme based on the Taylor approximation (21) has very good global approximation properties.

Taylor approximation of order $\theta + \min(2\gamma, \theta)$

The stochastic wood $\mathbf{w}_3$ (see Figure 5) has order $\theta + \min(2\gamma, \theta)$ and

$$\Psi(\mathbf{w}_3) = I_0^0 + I_1^0 + I_2^0 + I_2^1[I_0^0].$$

The corresponding Taylor approximation

$$\begin{aligned} X_t &= e^{A\Delta t} X_{t_0} + \left(\int_0^{\Delta t} e^{As}\, ds \right) F(X_{t_0}) + \int_{t_0}^{t} e^{A(t-s)} B(X_{t_0})\, dW_s \\ &\quad + \int_{t_0}^{t} e^{A(t-s)} B'(X_{t_0}) \left(\left(e^{A\Delta s} - I \right) X_{t_0} \right)\, dW_s + O\left((\Delta t)^{\theta + \min(2\gamma, \theta)} \right) \end{aligned} \tag{22}$$

thus has order $\theta + \min(2\gamma, \theta)$.

Note that the stochastic woods $\mathbf{w}_3$ and $\mathbf{w}_4$ give the same Taylor approximation, i.e., $\Psi(\mathbf{w}_4) = \Psi(\mathbf{w}_3)$.

Taylor approximation of order $\theta + 2\min(\gamma, \theta)$

The Taylor approximation corresponding to the S-wood $\mathbf{w}_5$ is

$$\begin{aligned} X_t &= e^{A\Delta t} X_{t_0} + \left(\int_0^{\Delta t} e^{As}\, ds \right) F(X_{t_0}) + \int_{t_0}^{t} e^{A(t-s)} B(X_{t_0})\, dW_s \\ &\quad + \int_{t_0}^{t} e^{A(t-s)} B'(X_{t_0}) \int_{t_0}^{s} e^{A(s-r)} B(X_{t_0})\, dW_r\, dW_s \\ &\quad + \int_{t_0}^{t} e^{A(t-s)} B'(X_{t_0}) \left(\left(e^{A\Delta s} - I \right) X_{t_0} \right)\, dW_s + O\left((\Delta t)^{\theta + 2\min(\gamma, \theta)} \right) \end{aligned} \tag{23}$$

and thus has order $\theta + 2\min(\gamma, \theta)$.

7.1 Application to the stochastic heat equation

The stochastic heat equation (8) satisfies the Assumptions 1–4 with the parameters $\theta = \frac{1}{4}$ and $\gamma = \frac{1}{4} - \varepsilon$ for arbitrarily small $\varepsilon \in (0, \frac{1}{4})$.

The Taylor approximations below exploit the special structure of the stochastic heat equation, i.e., its drift term vanishes and its diffusion term is linear.

The Taylor approximations (20), (22) and (23) reduce to

$$X_t = e^{A\Delta t} X_{t_0} + \int_{t_0}^{t} e^{A(t-s)} B(X_{t_0})\, dW_s + O\left((\Delta t)^{\frac{1}{2}-\varepsilon}\right).$$

$$\begin{aligned} X_t &= e^{A\Delta t} X_{t_0} + \int_{t_0}^{t} e^{A(t-s)} B(X_{t_0})\, dW_s \\ &\quad + \int_{t_0}^{t} e^{A(t-s)} B((e^{A\Delta s} - I) X_{t_0})\, dW_s + O\left((\Delta t)^{\frac{1}{2}}\right) \\ &= e^{A\Delta t} X_{t_0} + \int_{t_0}^{t} e^{A(t-s)} B(e^{A\Delta s} X_{t_0})\, dW_s + O\left((\Delta t)^{\frac{1}{2}}\right). \end{aligned}$$

$$\begin{aligned} X_t &= e^{A\Delta t} X_{t_0} + \int_{t_0}^{t} e^{A(t-s)} B(e^{A\Delta s} X_{t_0})\, dW_s \\ &\quad + \int_{t_0}^{t} e^{A(t-s)} B\left(\int_{t_0}^{s} e^{A(s-r)} B(X_{t_0})\, dW_r\right) dW_s + O\left((\Delta t)^{\frac{3}{4}-\varepsilon}\right) \\ &= e^{A\Delta t} X_{t_0} + \int_{t_0}^{t} e^{A(t-s)} B\left(e^{A\Delta s} X_{t_0} + \int_{t_0}^{s} e^{A(s-r)} B(X_{t_0})\, dW_r\right) dW_s \\ &\quad + O\left((\Delta t)^{\frac{3}{4}-\varepsilon}\right). \end{aligned}$$

7.2 Finite dimensional SODEs

The abstract setting for SPDE here includes finite dimensional SODEs.

Let $U = H = \mathbb{R}$ and let W_t be a standard scalar Wiener process.

Suppose that the eigenfunctions and the eigenvalues of the linear operator $A \equiv 0$ in Assumption 1 are given by $e_1 = 1 \in H$ and $\lambda_1 = 0$ with the index set $\mathcal{I} = \{1\}$. Thus $D(A) = D((\kappa - A)^r) = H = \mathbb{R}$, $r \in \mathbb{R}$. Define $D = H = \mathbb{R}$, so $L(U, D) = L_{HS}(U, D) = \mathbb{R}$.

The SPDE (6) in this setup is just the scalar SODE

$$dX_t = F(X_t)\,dt + B(X_t)\,dW_t.$$

The appropriate parameters in Assumption 3 here are $\theta = \frac{1}{2} - \varepsilon$ and $\gamma = 1 - \varepsilon$ for arbitrarily small $\varepsilon \in (0, 1)$.

The exponential Euler approximation (21) here becomes

$$X_t = X_{t_0} + F(X_{t_0}) \cdot \Delta t + B(X_{t_0}) \cdot (W_t - W_{t_0}) + O(\Delta t),$$

which is just the Euler–Maruyama scheme.

Finally, the Taylor approximation (23) reduces to

$$\begin{aligned} X_t &= X_{t_0} + F(X_{t_0}) \cdot \Delta t + B(X_{t_0}) \cdot (W_t - W_{t_0}) \\ &\quad + B'(X_{t_0}) B(X_{t_0}) \int_{t_0}^{t} \int_{t_0}^{s} dW_r\, dW_s + O\left((\Delta t)^{\frac{3}{2}}\right), \end{aligned}$$

which is the Milstein scheme.

8 Numerical schemes for SPDEs

Numerical approximations of the SPDE (6) require the discretisation of both the infinite dimensional space H and the time interval $[0, T]$.

The Galerkin discretisation set up of the previous lectures holds here too.

8.1 The exponential Euler scheme

The global convergence orders of the Taylor approximation (19) and the Taylor approximation (20) are too low to be consistent.

The Taylor approximation (21) gives the consistent exponential Euler scheme

$$\begin{aligned} Y_{k+1}^{N,M,L} &= e^{A_N h}\, Y_k^{N,M,L} + \left(\int_0^h e^{A_N s}\, ds\right) F_N\left(Y_k^{N,M,L}\right) \\ &\quad + \int_{kh}^{(k+1)h} e^{A_N((k+1)h-s)} B_{N,L}\left(Y_k^{N,M,L}\right)\, dW_s, \end{aligned} \tag{24}$$

for $k \in \{0, 1, \ldots, M-1\}$ and N, M, $L \in \mathbb{N}$.

The conditional distribution with respect to $\mathcal{F}_{kh}$ of the Itô integrals

$$\int_{kh}^{(k+1)h} e^{A_N((k+1)h-s)} B_{N,L}\left(Y_k^{N,M,L}\right)\, dW_s$$

in this numerical scheme is the normal distribution.

8.2 Another Taylor scheme

Similarly, the Taylor approximation (22) gives the numerical scheme

$$\begin{aligned} Y_{k+1}^{N,M,L} &= e^{A_N h}\, Y_k^{N,M,L} + \left(\int_0^h e^{A_N s}\, ds\right) F_N\Big(Y_k^{N,M,L}\Big) \\ &\quad + \int_{kh}^{(k+1)h} e^{A_N((k+1)h-s)} B_{N,L}\Big(Y_k^{N,M,L}\Big)\, dW_s \\ &\quad + \int_{kh}^{(k+1)h} e^{A_N((k+1)h-s)} B'_{N,L}\Big(Y_k^{N,M,L}\Big)\Big(\big(e^{A_N(s-kh)} - I\big)\, Y_k^{N,M,L}\Big)\, dW_s \end{aligned}$$

for $k \in \{0, 1, \ldots, M-1\}$ and N, M, $L \in \mathbb{N}$.

The conditional distributions with respect to $\mathcal{F}_{kh}$ of the integrals

$$\int_{kh}^{(k+1)h} e^{A_N((k+1)h-s)} B_{N,L}\Big(Y_k^{N,M,L}\Big)\, dW_s,$$

$$\int_{kh}^{(k+1)h} e^{A_N((k+1)h-s)} B'_{N,L}\Big(Y_k^{N,M,L}\Big)\Big(\big(e^{A_N(s-kh)} - I\big)\, Y_k^{N,M,L}\Big)\, dW_s$$

are also the normal distribution.

However, it is much more complicated to compute the covariance matrix of these normal distributed random variables than those used in the exponential Euler scheme.

8.3 An infinite dimensional analog of Milstein's scheme

The Taylor approximation (23) gives the one–step numerical scheme

$$
\begin{aligned}
Y_{k+1}^{N,M,L} &= e^{A_N h}\, Y_k^{N,M,L} + \left(\int_0^h e^{A_N s} ds \right) F_N\Big(Y_k^{N,M,L} \Big) \\
&+ \int_{kh}^{(k+1)h} e^{A_N((k+1)h-s)} B_{N,L}\Big(Y_k^{N,M,L} \Big)\, dW_s \\
&+ \int_{kh}^{(k+1)h} e^{A_N((k+1)h-s)} B'_{N,L}\Big(Y_k^{N,M,L} \Big) \Big(\big(e^{A_N(s-kh)} - I \big)\, Y_k^{N,M,L} \Big)\, dW_s \\
&+ \int_{kh}^{(k+1)h} e^{A_N((k+1)h-s)} B'_{N,L}\Big(Y_k^{N,M,L} \Big) \int_{kh}^{s} e^{A_N(s-r)} B_{N,L}\Big(Y_k^{N,M,L} \Big)\, dW_r\, dW_s
\end{aligned}
$$

for $k \in \{0, 1, \ldots, M-1\}$ and N, M, $L \in \mathbb{N}$.

This is an infinite dimensional analog of the Milstein scheme for SODEs.

8.4 Linear–implicit Euler and Crank–Nicolson schemes

Two representative numerical schemes used in the literature for the SPDE (6) are the linear–implicit Euler scheme

$$\begin{aligned} Y_{k+1}^{N,M,L} &= (I - hA_N)^{-1} \left(Y_k^{N,M,L} + h\, F_N\Big(Y_k^{N,M,L}\Big) \right) \\ &\quad + \int_{kh}^{(k+1)h} (I - hA_N)^{-1}\, B_{N,L}\Big(Y_k^{N,M,L}\Big)\, dW_s \end{aligned}$$

and the linear–implicit Crank–Nicolson scheme

$$\begin{aligned} Y_{k+1}^{N,M,L} &= \left(I - \frac{h}{2}A_N\right)^{-1} \left(\left(I + \frac{h}{2}A_N\right) Y_k^{N,M,L} + h\, F_N\Big(Y_k^{N,M,L}\Big) \right) \\ &\quad + \int_{kh}^{(k+1)h} \left(I - \frac{h}{2}A_N\right)^{-1} B_{N,L}\Big(Y_k^{N,M,L}\Big)\, dW_s \end{aligned}$$

for $k \in \{0, 1, \ldots, M-1\}$ and N, M, $L \in \mathbb{N}$

Here it is necessary to assume that $\lambda_i \geq 0$ for all $i \in \mathcal{I}$ in Assumption 1 in order to ensure that $(I - hA)$ is invertible for every $h \geq 0$.

9 Global and local convergence orders

Taylor expansions provide a fundamental instrument for understanding the local approximation properties of the solution of a differential equation.

For deterministic ODEs it is a general result (under suitable assumptions) that a numerical scheme for an ODE has global convergence order p if it has local order $p+1$ with $p > 0$.

Here local errors accumulate to give a larger global error.

For SODEs the situation is different.

A general result of MILSTEIN that says that a numerical scheme for an SODE with local order $r > \frac{1}{2}$ has global convergence order $r - \frac{1}{2}$.

The reason is that the errors in every step do not accumulate so rapidly since they are centered and independent random variables. Hence usually a half order only is lost.

For SPDEs the situation is different again. For example, the linear–implicit Euler scheme and the linear–implicit Crank–Nicolson scheme converge in time to the exact solution of the heat SPDE with local and global order $\frac{1}{4} - \varepsilon$.

Hence, the temporal discretisation error does not accumulate at all !

A similar situation (but with a higher convergence order) holds for many Taylor schemes introduced above, e.g., the exponential Euler scheme (24) converges with local and global order $\frac{1}{2} - \varepsilon$.

A combination of independent errors and the parabolic regularisation effect of the semigroup seems to be the reason for this.

However, such global phenomena have been barely investigated.
At present there is no general result for SPDEs like Milstein's Theorem for SODEs.