Milan j. math. 70 (2002) 187–207
1424-9286/02/010187-21 \$ 1.50+0.20/0

Milan Journal of Mathematics

The Systematic Derivation of Higher Order Numerical Schemes for Stochastic Differential Equations

P.E. Kloeden

1. Introduction

Deterministic calculus is much more robust to approximation than stochastic calculus because the integrand function in a Riemann sum approximating a Riemann integral can be evaluated at an arbitrary point of the discretization subinterval, whereas for a stochastic integral the integrand function must always be evaluated at a specific point in the subinterval, namely the lefthand endpoint in the case of an Ito stochastic integral. Consequently considerably care is thus needed in deriving numerical schemes for the stochastic case, in particular higher order schemes, to ensure that they are consistent with stochastic calculus. Heuristic adaptations of well known schemes for deterministic ordinary differential equations to stochastic differential equations are usually inconsistent or, when they are consistent, converge with at most a low order.

The purpose of this article is to show to higher order numerical schemes for stochastic differential equations can be derived systematically from stochastic Taylor expansions, and at the same time to bring briefly to the readers' attention various developments, implementation issues and open areas of research.

For simplicity, let us consider a scalar Ito stochastic differential equation (SDE)

$$dX(t) = a(t, X(t))\, dt + b(t, X(t))\, dW(t), \tag{1.1}$$

where $\{W_t : t \in \mathbb{R}\}$ is a standard Wiener process, i.e., with $W_0 = 0$ w.p.1 and increments $W(t) - W(s) \sim N(0; t-s)$, $t \geq s \geq 0$, which are independent on nonoverlapping subintervals. The SDE (1.1) is, in fact, only a symbolic representation for the stochastic integral equation

$$X(t) = X(t_0) + \int_{t_0}^{t} a(s, X(s))\, ds + \int_{t_0}^{t} b(s, X(s))\, dW(s), \tag{1.2}$$

where the first integral is pathwise a deterministic Riemann integral and the second an Ito stochastic integral, which looks as if it could be defined pathwise as a deterministic Riemann-Stieltjes integral, but this is not possible because the sample paths of the Wiener process, though continuous, are not differentiable or even of bounded variation on any finite subinterval. A hint of this irregularity is given by easily shown nondifferentiability in the weaker mean-square sense, which follows immediately from

$$\frac{\mathbb{E}(|W(t) - W(s)|^2)}{(t-s)^2} = \frac{t-s}{(t-s)^2} = \frac{1}{t-s}.$$

A stochastic integral, now called the *Ito stochastic integral,* of a nonanticipative mean-square integrable integrand g was proposed by the Japanese mathematician K. Ito in the 1940s. It is defined in terms of the mean-square limit, namely

$$\int_{t_0}^{T} g(s)\, dW(s) := \mathrm{qm} - \lim_{\Delta \to 0} \sum_{n=0}^{N_T - 1} g(t_n, \omega) \left\{ W(t_{n+1}, \omega) - W(t_n, \omega) \right\},$$

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 $t_{N_T} = T$. Nonanticipativeness here means, in particular, that the random variables $g(t_n)$ and $W(t_{n+1}) - W(t_n)$ are independent, from which follow two simple but very useful expectation properties of Ito integrals:

$$\begin{aligned} &\mathbb{E}\left(\int_{t_0}^{T} g(s)\, dW(s) \right) = 0, \\ &\mathbb{E}\left(\int_{t_0}^{T} g(s)\, dW(s) \right)^2 = \int_{t_0}^{T} \mathbb{E}\left(g(s)^2 \right)\, ds. \end{aligned} \tag{1.3}$$

As a consequence, the theory of Ito stochastic differential equations (1.1) is closely interrelated with the theories of diffusion processes and martingales.

The stochastic chain rule, which is known as the *Ito formula,* for scalar valued function $f(t, X(t))$ of the solution $X(t)$ of the scalar Ito SDE (1.1) is given by

$$f(t, X(t)) = f(t_0, X(t_0)) + \int_{t_0}^{t} L^0 f(s, X(s))\, ds + \int_{t_0}^{t} L^1 f(s, X(s))\, dW(s), \tag{1.4}$$

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

$$L^0 = \frac{\partial}{\partial t} + a\,\frac{\partial}{\partial x} + \frac{1}{2} b^2 \frac{\partial^2}{\partial x^2}, \qquad L^1 = b\,\frac{\partial}{\partial x}. \tag{1.5}$$

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

There is another stochastic integral, called the *Stratonovich stochastic integral,* which was proposed by the Russian physicist R.L. Stratonovich in the 1960s. It differs from the Ito stochastic integral in that its integrand function is always evaluated at the midpoint of each discretization subinterval. The Stratonovich stochastic integral no longer satisfies the convenient properties (1.3), but its chain rule coincides with the deterministic chain rule, which means integration methods and tricks from deterministic calculus can be used to find explicit solutions Stratonovich stochastic differential equations.

The availability of two stochastic calculi caused considerable confusion amongst modellers in the 1960s and 1970s. Both are mathematically correct and it is essentially a modelling rather than mathematical issue as to which is appropriate in a particular context. However, once one of the stochastic calculi has been decided upon, one can modify the drift coefficient, i.e., the $a(t, x)$ coefficient in (1.1), to find the corresponding SDE in the other calculi that has the same solutions. In this way the mathematical advantages of both stochastic calculi can be exploited. Here we focus of the Ito stochastic calculus.

There are now many excellent monographs on stochastic calculus and stochastic differential equations. A number of these are listed with the references of the author's textbook [9] with E. Platen on the numerical solution of stochastic differential equations, which also provides an introduction to these topics. The reader is referred to this textbook for a detailed treatment of the material discussed below. More advanced topics on stochastic numerics can be found in G. Milstein's monograph [12], while the recent survey paper [13] provides an up to date survey of the subject and an extensive list of recent research papers. Maple and Matlab software for stochastic differential equations is presented in [2, 6, 7, 16].

2. The stochastic Euler scheme and convergence types

The stochastic counterpart of the Euler scheme for the Ito SDE (1.1) is given by

$$X_{n+1} = X_n + a(t, X_n)\,\Delta_n + b(t, X_n)\,\Delta W_n, \tag{2.1}$$

with time and noise increments

$$\Delta_n = t_{n+1} - t_n = \int_{t_n}^{t_{n+1}} ds, \quad \Delta W_n = W(t_{n+1}) - W(t_n) = \int_{t_n}^{t_{n+1}} dW(s).$$

The noise increments ΔW_n here are Gaussian random variables with mean 0 and variance Δ_n. They can generated from uniformly distributed random (or pseudo random) numbers through the Box-Muller method, although more efficient methods are available for very long simulations. In practice one needs to simulate a large number of realizations, which can be done efficiently in parallel on a distributed network of computers or processes with a master computer coordinating the calculations of the individual realizations on the individual processes (in particular, ensuring the independence of the random variables used) and collecting the results [16].

The stochastic Euler scheme was first investigated by Maruyama in the early 1950s and is often called the *Euler-Maruyama scheme.* It seems to be consistent (and is indeed) with the Ito stochastic calculus because the noise term in (2.1) approximates the Ito stochastic integral in (1.2) over a discretization subinterval $[t_n, t_{n+1}]$ by evaluating its integrand at the lefthand end point of this interval:

$$\begin{aligned}\int_{t_n}^{t_{n+1}} b(s, X(s))\,dW(s) &\approx \int_{t_n}^{t_{n+1}} b(t_n, X(t_n))\,dW(s) \\ &= b(t_n, X(t_n)) \int_{t_n}^{t_{n+1}} dW(s).\end{aligned}$$

Convergence for numerical schemes can be defined in a number of useful different ways. It is usual to distinguish between strong and weak convergence, depending on whether the realizations or only their probability distributions are required to be close, respectively. Consider a fixed interval $[t_0, T]$ and let Δ be the maximum step size of any partition of $[t_0, T]$. Then a numerical scheme is said to converge with *strong order* γ if, for sufficiently small Δ,

$$\mathbb{E}\left(|X(T) - X_{N_T}|\right) \leq K_T \, \Delta^\gamma$$

and with *weak order* β if

$$\left|\mathbb{E}\left(g(X(T))\right) - \mathbb{E}\left(g(X_{N_T})\right)\right| \leq K_{g,T} \, \Delta^\beta$$

for each polynomial g. These are global discretization errors, and the largest possible values of γ and β give the corresponding strong and weak orders, respectively, of the scheme.

The stochastic Euler scheme (2.1) has strong order $\gamma = \frac{1}{2}$ and weak order $\beta = 1$. These orders of convergence are with respect to classes of SDEs, e.g., with continuously differentiable coefficients for which the derivatives are uniformly bounded. For restricted classes a higher order is sometimes possible, such as SDE with *additive noise,* i.e., the diffusion coefficient b is independent of the state variable x, which attain a strong order $\gamma = 1$.

The strong order $\gamma = \frac{1}{2}$ and weak order $\beta = 1$ of the stochastic Euler scheme (2.1) are quite low, particularly given the fact that a large number of realizations need to be generated for most practical applications. Thus there is a need for higher order numerical schemes.

To derive a higher order, however, one should avoid heuristic adaptations of well known deterministic numerical schemes because they are usually inconsistent with Ito calculus or, when they are consistent, then they do not improve the order of convergence. For example, the deterministic Heun scheme (a second order Runge-Kutta scheme) adapted to the Ito SDE (1.1) has the form

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

For the Ito SDE $dX(t) = X(t)\, dW(t)$ it simplifies to

$$X_{n+1} = X_n + \frac{1}{2} X_n \left(2 + \Delta W_n\right) \Delta W_n$$

so $X_{n+1} - X_n = X_n \left(1 + \frac{1}{2}\Delta W_n\right) \Delta W_n$. The conditional expectation

$$\begin{aligned}\mathbb{E}\left(\frac{X_{n+1} - X_n}{\Delta_n} \,\middle|\, X_n = x\right) &= \frac{x}{\Delta_n}\,\mathbb{E}\left(\Delta W_n + \frac{1}{2}(\Delta W_n)^2\right) \\ &= \frac{x}{\Delta_n}\left(0 + \frac{1}{2}\Delta_n\right) \\ &= \frac{1}{2}\,x.\end{aligned}$$

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

Note, however, that the adapted Heun scheme is consistent and convergent with strong order $\gamma = 1$ for SDE with additive noise. W. Rümelin [14] showed in 1982 that heuristic adaptations of deterministic Runge–Kutta schemes to SDE such as the Heun scheme, when consistent, attain a strong order of convergence no greater than $\gamma = \frac{3}{2}$ for general classes of SDE.

3. Stochastic Taylor expansions and schemes

A higher order of convergence cannot be obtained with a deterministic numerical schemes adapted to SDEs, when it happens to be consistent, because such a scheme only involves the simple increments of time Δ_n and noise ΔW_n, the latter being a poor approximation for the highly irregular Wiener process within the discretization subinterval $[t_n, t_{n+1}]$. To obtain a higher order of convergence one needs to provide more information about the Wiener process within the discretization subinterval. Such information is provided by multiple integrals of the Wiener process, which arise in stochastic Taylor expansions of the solution of an SDE. Consistent numerical schemes of arbitrarily desired higher order can be derived by truncating appropriate stochastic Taylor expansions. These expansions are themselves derived through iterated application of the stochastic chain rule, i.e., the *Ito formula* (1.4).

When $f(t,x) \equiv x$, the Ito formula (1.4) is just the integral version (1.2) of the SDE (1.1), i.e.,

$$X(t) = X(t_0) + \int_{t_0}^{t} a(s, X(s))\, ds + \int_{t_0}^{t} b(s, X(s))\, dW(s). \tag{3.1}$$

Applying the Ito formula to the integrand functions $f(t,x) = a(t,x)$ and $f(t,x) = b(t,x)$ in (3.1) gives

$$
\begin{aligned}
X(t) = X(t_0) &+ \int_{t_0}^{t} \left[a(t_0, X(t_0)) + \int_{t_0}^{s} L^0 a(u, X(u))\, du \right. \\
&\qquad \left. + \int_{t_0}^{s} L^1 a(u, X(u))\, dW(u) \right] ds \\
&+ \int_{t_0}^{t} \left[b(t_0, X(t_0)) + \int_{t_0}^{s} L^0 b(u, X(u))\, du \right. \\
&\qquad \left. + \int_{t_0}^{s} L^1 b(u, X(u))\, dW(u) \right] dW(s) \\
&= X(t_0) + a(t_0, X(t_0)) \int_{t_0}^{t} ds \\
&+ b(t_0, X(t_0)) \int_{t_0}^{t} dW(s) + R_1(t, t_0)
\end{aligned}
\tag{3.2}
$$

with the remainder

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

Replacing t_0 by t_n, $X(t_n)$ by X_n and t by t_{n+1}, $X(t)$ by X_{n+1}, and discarding the remainder gives the Euler scheme (2.1), which is the simplest nontrivial stochastic Taylor scheme.

Higher order stochastic Taylor expansions are obtained by successively applying the Ito formula to the integrand functions in the remainder, there being a number of different alternatives here. For example, applying the Ito formula to the integrand $L^1 b$ in the fourth double integral of the remainder

$R_1(t, t_0)$ gives the stochastic Taylor expansion

$$\begin{aligned} X(t) = X(t_0) + a(t_0, X(t_0)) \int_{t_0}^{t} ds + b(t_0, X(t_0)) \int_{t_0}^{t} dW(s) \\ + L^1 b(t_0, X(t_0)) \int_{t_0}^{t} \int_{t_0}^{s} dW(u)\, dW(s) + R_2(t, t_0) \end{aligned} \tag{3.3}$$

with the remainder

$$\begin{aligned} R_2(t, t_0) = & \int_{t_0}^{s} \int_{t_0}^{s} L^0 a(u, X(u))\, du\, ds \\ & + \int_{t_0}^{s} \int_{t_0}^{s} L^1 a(u, X(u))\, dW(u)\, ds \\ & + \int_{t_0}^{t} \int_{t_0}^{s} L^0 b(u, X(u))\, du\, dW(s) \\ & + \int_{t_0}^{t} \int_{t_0}^{s} \int_{t_0}^{u} L^0 L^1 b(v, X(v))\, dv\, dW(u)\, dW(s) \\ & + \int_{t_0}^{t} \int_{t_0}^{s} \int_{t_0}^{u} L^1 L^1 b(v, X(v))\, dW(v)\, dW(u)\, dW(s). \end{aligned}$$

Replacing t_0 by t_n, $X(t_n)$ by YX_n and t by t_{n+1}, $X(t)$ by X_{n+1}, and discarding the remainder gives the *Milstein scheme*

$$\begin{aligned} X_{n+1} = X_n + a(t, X_n)\, \Delta_n + b(t, X_n)\, \Delta W_n \\ + L^1 b(t_n, X(t_n)) \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} dW(u),\, dW(s), \end{aligned} \tag{3.4}$$

which was proposed by G. Milstein [11] in the early 1970s. The Milstein scheme has strong order $\gamma = 1$ and weak order β= 1. It thus has a higher order of convergence in the strong sense, but gives no improvement in the weak sense.

To obtain even higher order schemes one continues the above procedure of expanding integrand functions in appropriate remainder terms with the help of the Ito formula. There are thus many different possible stochastic Taylor expansions. The Euler and Milstein schemes already indicate the general pattern:

i) *higher order schemes achieve their higher order through the inclusion of multiple stochastic integral terms;*
ii) *a scheme 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, ….

3.1. A compact notation

A compact notation involving multi-indices was proposed by W. Wagner to describe the terms that should be included or expanded to obtain a stochastic Taylor expansion or scheme of a particular order. Writing $b^0(t,x)$ for $a(t,x)$, $dW^0(t)$ for dt, $b^1(t,x)$ for $b(t,x)$ and $dW^1(t)$ for $dW(t)$, the Ito formula (1.4) takes the compactified form

$$f(t,X(t)) = f(t_0,X(t_0)) + \sum_{j=0}^{1} \int_{t_0}^{t} L^j f(s,X(s))\, dW^j(s),$$

and reduces to the SDE (1.1)

$$X(t) = X(t_0) + \sum_{j=0}^{1} \int_{t_0}^{t} b^j(s,X(s))\, dW^j(s)$$

when $f(t,x) \equiv x$, i.e., the identity function $\mathrm{id}(x)$. With this notation in the stochastic Taylor expansions (3.2) and (3.3) the coefficients of the stochastic Euler scheme are thus $L^j\mathrm{id} \equiv b^j$ evaluated at (t_n, X_n) with associated integrals $\int_{t_n}^{t_{n+1}} dW^j(s)$ with the indices $j = 0$ and 1, while the Milstein scheme includes the additional coefficient the $L^1L^1\mathrm{id} \equiv L^1b^1$ evaluated at (t_n, X_n) and associated integral $\int_{t_n}^{t_{n+1}} \int_{t_n}^{s} dW^1(u)\, dW^1(s)$ with the multi-index $(1,1)$.

In general, a multi-index α of length $l(\alpha) = l$ is an l-dimensional row vector $\alpha = (j_1, j_2, \ldots, j_l)$ with components $j_i \in \{0,1\}$ for $i \in \{1, 2, \ldots, l\}$. Let $\mathcal{M}_2$ 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 $\alpha \in \mathcal{M}_2$ with $l(\alpha) \geq 1$, write $-\alpha$ and $\alpha-$ for the multi-index in $\mathcal{M}_2$ obtained by deleting the first and the last component, respectively, of α. Then, for such a multi-index $\alpha = (j_1, j_2, \ldots, j_l)$ with $l \geq 1$, the *multiple Ito integral* $I_\alpha[g(\cdot)]_{t_0,t}$ of a nonanticipative function g is defined recursively by

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

Similarly, the *Ito coefficient function* f_α for a deterministic function f is defined 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. Rather, the set of corresponding multi-indices must form an *hierarchical set,* i.e., a nonempty subset $\mathcal{A}$ of $\mathcal{M}_2$ with

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

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

$$\mathcal{B}(\mathcal{A}) = \{\alpha \in \mathcal{M}_2 \setminus \mathcal{A} : -\alpha \in \mathcal{A}\},$$

i.e. consisting of all of the "next following" multi-indices with respect to the given hierarchical set. Then the Ito-Taylor expansion corresponding to the hierarchical set $\mathcal{A}$ and remainder set $\mathcal{B}(\mathcal{A})$ is

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

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

Applying this to the identity function $f = \mathrm{id}$ on a subinterval $[t_n, t_{n+1}]$ at a starting point $(t_n, X_n^{\mathcal{A}})$ and truncating the remainder gives the $\mathcal{A}$-*stochastic Taylor scheme*

$$\begin{aligned} X_{n+1}^{\mathcal{A}} &= \sum_{\alpha\in\mathcal{A}} I_\alpha \left[\mathrm{id}_\alpha(t_n, X_n^{\mathcal{A}})\right]_{t_n,t_{n+1}} \\ &= X_n^{\mathcal{A}} + \sum_{\alpha\in\mathcal{A}\setminus\emptyset} \mathrm{id}_\alpha(t_n, X_n^{\mathcal{A}})\, I_\alpha\left[1\right]_{t_n,t_{n+1}}. \end{aligned} \tag{3.5}$$

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

$$\Lambda_\gamma = \left\{\alpha \in \mathcal{M}_2 :\ l(\alpha) + n(\alpha) \le 2\gamma \quad \text{or} \quad l(\alpha) = n(\alpha) = \gamma + \frac{1}{2}\right\},$$

where $n(\alpha)$ denotes the number of components of a multi-index α which are equal to 0, and the *weak order* β *stochastic Taylor scheme*, which converges with weak order β, involves the hierarchical set

$$\Gamma_\beta = \{\alpha \in \mathcal{M}_2 : l(\alpha) \leq \beta\}\,.$$

For example, the hierarchical sets $\Lambda_{1/2} = \Gamma_1 = \{\emptyset, (0), (1)\}$ give the stochastic Euler scheme, which is both strongly and weakly convergent, while the strongly convergent Milstein scheme corresponds to the hierarchical set $\Lambda_1 = \{\emptyset, (0), (1), (1,1)\}$. See Figure 3.1 for a graphical representation of the corresponding stochastic Taylor trees.

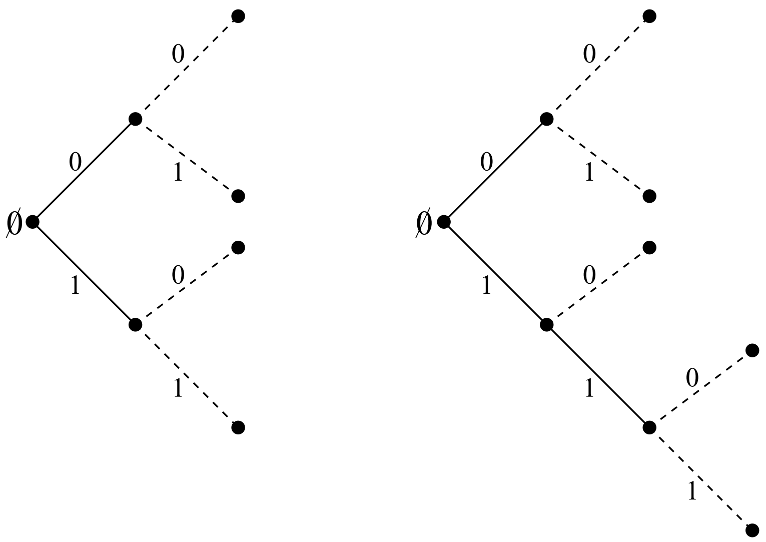

Figure 3.1: Stochastic Taylor trees for the stochastic Euler scheme (left) and the Milstein scheme (right). The multi-indices are formed by concatenating indices along a branch from the left back towards the root $\emptyset$ of the tree. Dashed line segments correspond to remainder terms.

3.2. Other types of schemes

Numerical schemes for deterministic differential equations are usually derived by some heuristic approximation of integrals or derivatives and the result is then compared with a deterministic Taylor expansion to verify its convergence and to determine its rate of convergence. In view of the lack of robustness of the definition of stochastic integrals compared to that of deterministic integrals, in the stochastic case it is wiser (and in the long run easier) to start with a stochastic Taylor expansion or scheme of the desired convergence type and order, and then to make desired modifications. Derivative-free schemes which are stochastic analogues of the deterministic Runge-Kutta schemes can be derived in this way. For example , the $L^1 b$ term can be replaced in the Milstein scheme (3.4) by the forward difference approximation

$$\frac{1}{\sqrt{\Delta_n}}\left\{b\left(t_n, X_n + a(t_n,X_n)\,\Delta_n + b(t_n,X_n)\,\sqrt{\Delta_n}\right) - b\left(t_n,X_n\right)\right\}$$

to give the *strong order* 1.0 *Platen scheme* by

$$\begin{aligned} X_{n+1} &= X_n + a(t_n,X_n)\,\Delta_n + b(t_n,X_n)\,\Delta W_n \\ &\quad + \frac{1}{\sqrt{\Delta_n}}\left\{b(t_n,\bar{\Upsilon}_n) - b(t_n,X_n)\right\} I_{(1,1)}[t_n;t_{n+1}] \end{aligned}$$

with the supporting value

$$\bar{\Upsilon}_n = X_n + a(t_n,X_n)\,\Delta_n + b(t_n,X_n)\,\sqrt{\Delta_n}.$$

Implicit schemes can also be derived in this way, but care is needed to maintain consistency with Ito calculus. This can be achieved by making implicit only the coefficients corresponding to purely deterministic multiple integrals, i.e. with multi-indices with all components $j_i = 0$. For example, the (*semi-*) *implicit stochastic Euler scheme* reads

$$X_{n+1} = X_n + a(t_{n+1},X_{n+1})\,\Delta_n + b(t_n,X_n)\,\Delta W_n$$

and retains the strong and weak orders of convergence of the explicit Euler scheme. This is fine when the stiffness in the SDE is due to the drift term $a(t,x)$, but much more work still needs to be done to construct <u>consistent</u> schemes in which the diffusion coefficient $b(t,x)$ can be made implicit.

At the time of writing, very little work has been done on multi-step schemes for stochastic differential equations, see [1] and [9, Chapter 11, Section 4].

The noise increments ΔW_n in the stochastic Euler scheme (2.1) are Gaussian random variables with mean 0 and variance Δ_n, hence can generated by the Box-Muller method from uniformly distributed random (or pseudo random) numbers. If only weak convergence is required these Gaussian random variables can be replaced by simpler two-point distributed random variables $\Delta\widehat{W}_n$ with

$$\mathbb{P}\left(\Delta\widehat{W}_n = \pm\sqrt{\Delta_n}\right) = \frac{1}{2},$$

which have the same first three moments as ΔW_n. The resulting stochastic Euler scheme is called the *simplified weak Euler scheme*. Similar approximations of the multiple stochastic integrals can be made in higher order weak schemes, but for strong schemes the multiple stochastic integrals need to be simulated more carefully. Fortunately, in some cases there exist exact formula such as

$$I_{(j,j)}[t_n;t_{n+1}] = \int_{t_n}^{t_{n+1}} \int_{t_n}^{t} dW^j(s)\, dW^j(t) = \frac{1}{2}\left\{(\Delta W_n^j)^2 - \Delta_n\right\} \tag{3.6}$$

or known probability distributions such as for the integral

$$\Delta Z_n^j = I_{(j,0)}[t_n;t_{n+1}] = \int_{t_n}^{t_{n+1}} \int_{t_n}^{t} dW^j(s)\, dt,$$

which is Gaussian with mean zero and variance $\frac{1}{3}(\Delta_n)^3$ and covariance $\mathbb{E}\left(\Delta W_n^j \Delta Z_n^j\right) = \frac{1}{2}(\Delta_n)^2$ (with $j \geq 1$ in both case here).

4. Vector valued stochastic differential equations

A similar setup holds for vector valued SDEs. For a d-dimensional solution process $X(t) = (X^1(t), \ldots, X^d(t))^\top$ and an m–dimensional Wiener process $W(t) = (W^1(t), \ldots, W^m(t))^\top$, i.e., consisting of m pairwise independent scalar Wiener processes $W^1(t)$, …, $W^m(t)$ these have the form

$$dX_t = a(t, X_t)\, dt + \sum_{j=1}^{m} b^j(t, X_t)\, dW_t^j, \tag{4.1}$$

with equivalent componentwise form

$$dX_t^i = a^i(t, X_t)\, dt + \sum_{j=1}^{m} b^{i,j}(t, X_t)\, dW_t^j, \quad i = 1, \ldots, d, \tag{4.2}$$

and the compactified vector form

$$dX(t) = \sum_{j=0}^{m} b^j(s, X(s))\, dW^j(s) \tag{4.3}$$

with $b^0(t, x) = a(t, x)$ and $W^0(t) = t$.

The multi-indices, now in a set $\mathcal{M}_m$, have components $j_i \in \{0, 1, \ldots, m\}$ and the differential operators are defined as

$$L^0 = \frac{\partial}{\partial t} + \sum_{k=1}^{d} a^k \frac{\partial}{\partial x^k} + \frac{1}{2} \sum_{k,l=1}^{d} \sum_{j=1}^{m} b^{k,j} b^{l,j} \frac{\partial^2}{\partial x^k \partial x^l} \tag{4.4}$$

and

$$L^j = \sum_{k=1}^{d} b^{k,j} \frac{\partial}{\partial x^k}, \qquad j = 1, \ldots, m, \tag{4.5}$$

i.e., there is an operator L^j for each noise component $W^j(t)$ with $j \geq 1$. These are applied to scalar valued functions, thus the Ito formula is applied componentwise to expand the solution process $X(t)$ of a vector valued SDE (4.2). For example, the Euler scheme for a vector state SDE with m independent Wiener processes reads

$$X_{n+1} = X_n + a(t, X_n)\,\Delta_n + \sum_{j=1}^{m} b^j(t, X_n)\,\Delta W_n^j \tag{4.6}$$

while the Milstein scheme reads

$$\begin{aligned} X_{n+1} \quad &= \quad X_n + a(t, X_n)\,\Delta_n + \sum_{j=1}^{m} b^j(t, X_n)\,\Delta W_n^j \\ &\qquad + \sum_{j_1, j_2 = 1}^{m} L^{j_1} b^{j_2}(t_n, X(t_n)) I_{(j_1, j_2)}\,[1]_{t_n, t_{n+1}}\,. \end{aligned} \tag{4.7}$$

Apart from the need to evaluate more terms, a vector valued state is <u>not</u> a major problem here. However, a vector valued Wiener process can give rise to serious difficulties, even when the state process is scalar. For example, the Milstein scheme (4.7) with $m \geq 2$ Wiener processes then contains the double stochastic integrals

$$I_{(j_1, j_2)}\,[1]_{t_n, t_{n+1}} = \int_{t_n}^{t_{n+1}} \int_{t_n}^{t} dW_s^{j_1}\, dW_t^{j_2} \tag{4.8}$$

with $j_1 \neq j_2$, which cannot be expressed in terms of the increments $\Delta W_n^{j_1}$ and $\Delta W_n^{j_2}$ of the components of the Wiener processes as in the case (3.6) above where $j_1 = j_2 \geq 1$. Nor is there a known tractable probability distribution for such $I_{(j_1, j_2)}$ integrals like those for the $I_{(j,0)}[t_n; t_{n+1}]$ integrals mentioned above.

4.1. SDEs with special structural properties

It is often possible to exploit the special structural properties of a given SDE to avoid having to use the double stochastic integrals (4.8), such additive noise when the $b^j(t,x) \equiv b^j(t)$ do not depend on the x variable, or *commutativity noise* when

$$L^{j_1} b^{j_2}(t,x) \equiv L^{j_2} b^{j_1}(t,x) \qquad \text{for } 1 \leq j_1, j_2 \leq m. \tag{4.9}$$

(There are Maple routines available to check for such condition as well as for setting up various schemes for given SDEs [2, 7]).

For SDEs with additive noise, coefficients of the the troublesome double integrals in the Milstein scheme disappear and the scheme essentially reduces to the stochastic Euler scheme. (This is why the Euler scheme for SDEs with additive noise seems to have strong order $\gamma = 1$). For SDEs with commutative noise, the identities

$$\int_{t_n}^{t_{n+1}} \int_{t_n}^{t} dW_s^{j_1}\, dW_t^{j_2} + \int_{t_n}^{t_{n+1}} \int_{t_n}^{t} dW_s^{j_2}\, dW_t^{j_1} = \Delta W_n^{j_1}\, \Delta W_n^{j_2}$$

for $1 \leq j_1, j_2 \leq m$ with $j_1 \neq j_2$ can be used along with the formulae (3.6) to collapse the Milstein scheme (4.7) to the *Milstein scheme with commutative noise* given by

$$\begin{aligned} X_{n+1} = X_n &+ a(t_n, X_n)\, \Delta_n + \sum_{j=1}^{m} b^j(t_n, X_n) \Delta W_n^j \\ &+ \frac{1}{2} \sum_{j=1}^{m} L^{j_1} b^j(t_n, X_n) \left\{ (\Delta W_n^j)^2 - \Delta_n \right\} \\ &+ \frac{1}{2} \sum_{\substack{j_1, j_2 = 1, \\ j_1 \neq j_2}}^{m} L^{j_1} b^{j_2}(t_n, X_n) \Delta W_n^{j_1}\, \Delta W_n^{j_2}. \end{aligned} \tag{4.10}$$

4.2. Simulating multiple stochastic integrals

In general, such multiple stochastic integrals cannot be avoided, so must be approximated somehow. One possibility is to use random Fourier series for Browian bridge processes based on the given Wiener processes. See [9, Chapter 5, Section 8].

Another way is to simulate the integrals themselves by a simpler numerical scheme. For example, double integral

$$I_{(2,1),n} = I_{(2,1)}[t_n; t_{n+1}] = \int_{t_n}^{t_{n+1}} \int_{t_n}^{t} dW_s^2\, dW_t^1$$

can be approximated by applying the (vector valued) stochastic Euler scheme to the 2-dimensional Ito SDE (with superscripts labelling components)

$$dX_t^1 = X_t^2 \, dW_t^1, \qquad dX_t^2 = dW_t^2, \tag{4.11}$$

over the discretization subinterval $[t_n, t_{n+1}]$ with a suitable step size $\delta = (t_{n+1} - t_n)/K$. The solution of the SDE (4.11) with the initial condition $X_{t_n}^1 = 0$, $X_{t_n}^2 = W_{t_n}^2$ at time $t = t_{n+1}$ is given by

$$X_{t_{n+1}}^1 = I_{(2,1)}[t_n; t_{n+1}], \qquad X_{t_{n+1}}^2 = \Delta W_n^2.$$

Writing $t'_k = t_n + k\delta$ and $\delta W_{n,k}^j = W_{t'_{k+1}}^j - W_{t'_k}^j$, the stochastic Euler scheme for the SDE (4.11) reads

$$Y_{k+1}^1 = Y_k^1 + Y_k^2 \, \delta W_{n,k}^1 \,, \quad Y_{k+1}^2 = Y_k^2 + \delta W_{n,k}^2 \,, \qquad \text{for } 0 \le k \le K-1, \tag{4.12}$$

with the initial value $Y_0^1 = 0$, $Y_0^2 = W_{t_n}^2$. The strong order of convergence of $\gamma = \frac{1}{2}$ of the stochastic Euler scheme ensures that

$$\mathbb{E}\left(\left|Y_K^1 - I_{(2,1)}[t_n; t_{n+1}]\right|\right) \le C\sqrt{\delta},$$

so $I_{(2,1)}[t_n; t_{n+1}]$ can be approximated in the Milstein scheme by Y_K^1 with $\delta \approx \Delta_n^2$ without affecting the overall order of convergence.

Other higher order multiple stochastic integrals can be simulated in a similar way, either directly as needed or off-line or on–line on a networked supporting computer.

4.3. Concluding remarks

The need to approximate multiple stochastic integrals places a practical restriction on the order of strong schemes that can be implemented for a general SDE. Wherever possible special structural properties of the SDE under investigation should be exploited to simplify strong schemes as much as possible. For weak schemes the situation is easier as the multiple integrals do not need to approximated so accurately, as we saw for the simplified Euler scheme. Moreover, the extrapolation of weak schemes is possible [15].

The moral here is: *decide first what kind of approximation one wants, strong or weak, as this will determine the type of scheme that should be used, and then exploit the structural properties of the SDE under consideration to simplify the scheme that has been chosen to be implemented.*

5. Other developments

We close this paper by briefly mentioning two new applications of stochastic Taylor schemes, in the first case to Ito SDE with affine control inputs and in the second case to finite dimensional Ito SDE obtained from parabolic stochastic PDE through a Galerkin or methods of lines approximation.

5.1. SDE with affine control inputs

Stochastic Taylor schemes for Ito stochastic differential equations with affine control inputs are investigated in [17], i.e. for equations of the form (scalar case)

$$dX(t) = a(t, X(t))\,dt + b(t, X(t))\,dW(t) + c(t, X(t))\; u(t)\,dt.$$

These involve the L^0 and L^1 operators defined in (1.5) and a new operator for the control integral

$$L^2 = c\,\frac{\partial}{\partial x}.$$

The Taylor expansions and schemes are then formally similar to those for scalar SDEs with two independent noise components, except that the integral involving the control function is a Lebesgue integral rather than a stochastic integral.

Analogous results apply for the vector valued state and noise cases.

5.2. Stochastic partial differential equations

Numerical methods for parabolic stochastic partial differential equations (SPDE) of the form

$$dU(t) = \{AU(t) + f(U(t))\}\;dt + g(U(t))dW(t), \tag{5.1}$$

where $W(t)$ is a standard scalar Wiener process, can be constructed by applying a stochastic Taylor scheme with constant time step Δ to the N-dimensional Ito SDE formed either by a Galerkin projection or by finite difference approximations for the spatial derivatives (i.e., Rothe's method of lines). To fix ideas consider a Dirichlet boundary condition on a bounded domain D in $\mathbb{R}^d$.

In the Galerkin case, one uses the eigenfunctions $\phi_j \in H_0^{1,2}(\mathcal{D})$ and corresponding eigenvalues λ_j of the operator $-A$, i.e., with $-A\phi_j = \lambda_j\phi_j$ for $j = 1, 2, \ldots$, which form an orthonormal basis in $L_2(\mathcal{D})$ with $\lambda_j \to \infty$

as $j \to \infty$. Projecting the the SPDE (5.1) onto the subspace spanned by $\{\phi_1, \ldots, \phi_N\}$ gives an N-dimensional Ito-Galerkin SDE of the form

$$dU^{(N)}(t) = \left\{A_N U^{(N)}(t) + f_N(U^{(N)}(t))\right\} dt + g_N(U^{(N)}(t))\, dW(t), \quad (5.2)$$

where A_N is the diagonal matrix diag $[\lambda_j, \ldots, \lambda_N]$.

The combined truncation and global discretization error [4] for an strong order γ stochastic Taylor scheme applied to (5.2) with constant timestep size Δ has the form

$$\mathbb{E}\left(\left\|U(k\Delta) - X_k^{(N)}\right\|\right) \le K\left(\lambda_{N+1}^{-1/2} + \lambda_N^{[\gamma+\frac{1}{2}]+1}\Delta^\gamma\right), \quad (5.3)$$

where $[x]$ denotes the integer part of the real number x and the constant K depends on the initial value and bounds on the coefficient functions f and g of the SPDE (5.1) as well as on the length of the time interval $0 \le k\Delta \le T$ under consideration. Since $\lambda_j \to \infty$ as $j \to \infty$, a very small time step is needed in high dimensions, which is needed for convergence, i.e. the Ito-Galerkin SDE (5.2) is stiff and explicit schemes such as strong stochastic Taylor schemes are not really appropriate. Obviously, an implicit scheme should be used here, but the special structure of the SDE (5.2) allows one to use a simpler linear-implicit scheme, i.e., 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. For example the *linear-implicit Euler scheme* for the SDE (5.2) is

$$X_{n+1}^{(N)} = X_n^{(N)} + A_N X_{n+1}^{(N)}\Delta_n + f_N(X_n^{(N)})\Delta_n + g_N(X_n^{(N)})\,\Delta W_n, \quad (5.4)$$

which is easily solved for $X_{n+1}^{(N)}$ because its matrix $I_N - A_N\Delta_n$ is diagonal. For a linear-implicit strong order γ stochastic Taylor schemes the combined error [10] has the form

$$\mathbb{E}\left(\left\|U(k\Delta) - X_k^{(N)}\right\|\right) \le K\left(\lambda_{N+1}^{-1/2} + \Delta^\gamma\right), \quad (5.5)$$

so the time step size can be chosen independently of the dimension N of the Ito-Galerkin SDE (5.2).

Similar error bounds to (5.3) and (5.5) hold for the method of lines. This will be illustrated here for a 1-dimensional SPDE of the form

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial t^2} + f(u) + g(u)\eta_t, \quad (5.6)$$

on the unit spacial interval $x \in (0,1)$ with a Dirichlet boundary condition. Here η_t is Gaussian white noise and the the equation can be expressed in the form of (5.1) using a scalar Wiener process. Partitioning this spatial interval $[0,1]$ by equally spaced points $x_i = ih$ for $i = 1, \ldots, N_h = h^{-1}$ and then using a central difference quotient at these points for the spatial derivative term leads to an $N = (N_h - 1)$-dimensional Ito-Rothe SDE

$$dU^{(N)}(t) = \left\{ \frac{1}{h^2} T_N U^{(N_h)}(t) + f_N(U^{(N)}(t)) \right\} dt + g_N(U^{(N)}(t)) dW(t), \tag{5.7}$$

where $U^{(N),i}(t)$ is an approximation for $U(x_i,t)$, the ith component of $f_N(U^{(N)}(t))$ is just $f(U^{(N,i)}(t))$, and similarly for g_N, and T_N is the tridiagonal matrix with diagonal components all equal to 2 and the first upper and lower diagonal components all equal to -1. The combined error bounds for the explicit strong Euler scheme applied to the Ito-Rothe SDE (5.7) is of the form

$$\mathbb{E}\left(\max_{i=1,\ldots,N_h-1} \left| U(x_i, k\Delta) - X_k^{(N,i)} \right| \right) \le K \left(h + \frac{1}{h^4}\sqrt{\Delta} \right),$$

while that for the linear-implicit strong Euler scheme is

$$\mathbb{E}\left(\max_{i=1,\ldots,N_h-1} \left| U(x_i, k\Delta) - X_k^{(N,i)} \right| \right) \le K \left(h + \sqrt{\Delta} \right).$$

Here h plays the role of $\lambda_{N+1}^{-1/2}$ in the Galerkin method. In contrast to the linear-implicit strong Euler scheme applied to the Ito-Galerkin SDE (5.2), a tridiagonal matrix needs to be inverted in the linear-implicit strong Euler scheme applied to the Ito-Rothe SDE (5.7), but this can be done efficiently. Besides the coefficient functions f_N and g_N are much easier to determine in this case than for the Ito-Galerkin SDE.

The papers [3, 4, 5, 10, 16] and the references cited therein focus on the numerical solution of stochastic partial differential equations. Considerable complications arise when the driving noise process is random both in space and time as in [3, 5]. The field is only now beginning to be researched.

References

[1] R. Horvat Bokor, Convergence and stability properties for numerical approximations of stochastic ordinary differential equations, Doctoral Dissertation, Mathematics Department, University of Zagreb, September 2000.

[2] S. Cyganowski, L. Grüne and P.E. Kloeden, Maple *for Stochastic Differential Equations*, in *Theory and Numerics of Differential Equations*, (J.F. Blowey, J.P. Coleman, A.W. Craig, eds.), Springer Verlag (2001), 127–178.

[3] A.M. Davie and J.G. Gaines, *Convergence of numerical schemes for the solution of parabolic stochastic partial differential equations*, Math. Computation **70** (2001), 121–134.

[4] W. Grecksch and P.E. Kloeden, *Time-discretized Galerkin approximations of parabolic stochastic PDEs*, Bulletin Austral. Math. Soc. **54** (1996), 79–85.

[5] I. Gyöngy and D. Nualart, *Implicit schemes for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise*, Potential Anal. **7** (1997), 725–757.

[6] D.J. Higham, *An algorithmic introduction to numerical simulation of stochastic differential equations*, SIAM Review **43** (2001), 525–546.

[7] D.J. Higham and P.E. Kloeden, Maple *and* Matlab *for Stochastic Differential Equations in Finance*, in *Programming Languages and Systems in Computational Economics and Finance*, (S.S. Nielsen, ed.), Kluwer Academic Publishers, Amsterdam, 2002.

[8] P.E. Kloeden and R.A. Pearson, *The numerical solution of stochastic differential equations*, J. Aus. Math. Soc. Series B **20** (1977), 8–12.

[9] P.E. Kloeden and E. Platen, Numerical Solution of Stochastic Differential Equations, Springer-Verlag, Heidelberg, 1992; 2nd ed., 1999.

[10] P.E. Kloeden and S. Shott, *Linear-implicit strong schemes for Ito-Galerkin approximations of stochastic PDEs*, J. Appl. Math. Stoch. Anal. **14** (2001), 47–53.

[11] G.N. Milstein, *Approximate integration of stochastic differential equations*, Theor. Prob. Appl. **19** (1974), 557–562.

[12] G.N. Milstein, Numerical Integration of Stochastic Differential Equations, Kluwer, Dordrecht, 1995.

[13] E. Platen, *An introduction to numerical methods for stochastic differential equations*, Acta Numerica **8** (1999), 197–246.

[14] W. Rümelin, *Numerical treatment of stochastic differential equations*, SIAM J. Numer. Anal. **19** (1982), 604–613.

[15] D. Talay and L. Tubaro, *Expansions of the global error for numerical schemes solving stochastic differential equations*, Stoch. Processes Appl. **8** (1990), 483–509.

[16] N. Yannios, Computational Aspects of the Numerical Solution of SDEs, Ph.D. Thesis, Deakin University, Australia, 2002.

[17] P. Zeidis, Taylor-Verfahren für stochastische affine Kontrollsysteme, Diplomarbeit, FB Mathematik, J.W. Goethe Universität, Frankfurt am Main, 2002.

P.E. Kloeden
Fachbereich Mathematik
Johann Wolfgang Goethe Universität
D-60054 Frankfurt am Main, Germany
kloeden@math.uni-frankfurt.de