ELSEVIER

Applied Numerical Mathematics 22 (1996) 81–101

APPLIED NUMERICAL MATHEMATICS

High strong order explicit Runge–Kutta methods for stochastic ordinary differential equations

K. Burrage *, P.M. Burrage

Department of Mathematics, University of Queensland, Brisbane 4072, Australia

Abstract

The pioneering work of Runge and Kutta a hundred years ago has ultimately led to suites of sophisticated numerical methods suitable for solving complex systems of deterministic ordinary differential equations. However, in many modelling situations, the appropriate representation is a stochastic differential equation and here numerical methods are much less sophisticated. In this paper a very general class of stochastic Runge–Kutta methods is presented and much more efficient classes of explicit methods than previous extant methods are constructed. In particular, a method of strong order 2 with a deterministic component based on the classical Runge–Kutta method is constructed and some numerical results are presented to demonstrate the efficacy of this approach.

1. Introduction

Until recently many models that have been developed to describe physical phenomena have ignored stochastic effects because of the difficulty in solution both in terms of the lack of suitable numerical methods and also the non-availability of sufficiently powerful computers. However, recently there has been much interest in developing numerical methods for the numerical solution of stochastic differential equations (SDEs) and this has meant that more realistic models are capable of being solved. For example, it is natural to use SDEs in models of investment finance, while SDEs have long been used to model turbulent diffusion. It is also natural to consider stochastic counterparts of, for example, chemical kinetic models as typified by the Brusselator equations. In order to solve an SDE numerically a large number of sample paths have to be computed so that various statistical measures can be applied appropriately. Hence the availability of supercomputer resources will have a significant impact on the practical implementation of numerical SDE schemes.

* Corresponding author. E-mail: kb@maths.uq.oz.au.

PII S0168-9274(96)00027-X

This paper is partly a tribute to the pioneering work of Runge and Kutta, by extending their approaches to the numerical solution of SDEs. It was Runge [19] who extended the simple Euler method for solving the deterministic ordinary differential equation (DODE)

$$y'(t) = f\big(y(t)\big), \qquad y(t_0) = y_0, \qquad y \in \mathbb{R}^m \tag{1}$$

by allowing for a multiplicity of function evaluations within a single numerical step. This approach was extended by Kutta [12] who characterized the set of Runge–Kutta methods of order 4 which led to the famous classical Runge–Kutta method

$$\begin{aligned} Y_1 &= y_n, \\ Y_2 &= y_n + \tfrac{1}{2}hf(Y_1), \\ Y_3 &= y_n + \tfrac{1}{2}hf(Y_2), \\ Y_4 &= y_n + hf(Y_3), \\ y_{n+1} &= y_n + \tfrac{1}{6}h\big(f(Y_1) + 2f(Y_2) + 2f(Y_3) + f(Y_4)\big). \end{aligned} \tag{2}$$

This basic idea has led to a plethora of methods (both explicit and implicit) based on a general class of Runge–Kutta methods but it was not until Butcher [3] developed a very general methodology (which followed on from the work of Gill [7], Huťa [8] and Merson [14]) which allowed the analysis of the order of a general class of s-stage methods that efficient methods were capable of being developed.

In the case of numerical methods for SODEs (Stochastic Ordinary Differential Equations), the work is much less advanced than is the case for DODEs. (A survey paper by Burrage and Platen [1] discusses some of the current differences between numerical methods for DODEs and SODEs and outlines possible future directions in which work in the stochastic arena should proceed in order to generate robust and efficient software.) The autonomous stochastic version of (1) can be written in differential form as

$$\mathrm{d}y = f(y)\,\mathrm{d}t + g(y)\,\mathrm{d}W, \qquad y(t_0) = y_0, \qquad y \in \mathbb{R}^m. \tag{3}$$

Here f is an m-vector-valued function, g is an $m \times p$ matrix-valued function and $W(t)$ is a p-dimensional process having independent scalar Wiener process components ($t \geqslant 0$), and the solution $y(t)$ is an m-vector process. This is called an Itô stochastic ordinary differential equation and consists of a varying continuous component called the drift and a rapidly varying continuous component called the diffusion. The integral formulation of (3) can be written as

$$y(t) = y_0 + \int_{t_0}^{t} f\big(y(s)\big)\,\mathrm{d}s + \int_{t_0}^{t} g\big(y(s)\big)\,\mathrm{d}W(s), \tag{4}$$

where the second integral in (4) is an Itô stochastic integral with respect to the Wiener process $W(t)$. In order to simplify some notational details, for the rest of this paper it will be assumed without loss of generality that $m = p = 1$.

It is now known (see [10], for example) that it is not possible to merely translate a deterministic numerical method such as (2) to a SODE. Instead a very detailed analysis of order, stability and error behaviour is needed in order to construct suitably appropriate methods. Perhaps the simplest such

method is the stochastic generalization of the explicit Euler method [13] which takes the form (when $m = 1$)

$$y_{n+1} = y_n + h_n f(y_n) + \Delta W_n g(y_n), \tag{5}$$

where

$$h_n = t_{n+1} - t_n, \qquad \Delta W_n = W(t_{n+1}) - W(t_n).$$

Here the ΔW_n are independent $\mathrm{N}(0, h_n)$ normally distributed random variables which can be generated from independent, uniformly distributed random variables on $[0, 1]$.

This method is a very inaccurate one (order has a number of different meanings in the case of SODEs) and so more efficient methods are needed. One possible approach is to use truncated forms of the stochastic Taylor series formula (see [9,17,20], for example). By adding more and more stochastic integral terms from the stochastic Taylor expansion more accurate methods are obtained. But this is at the cost of requiring more and more partial derivatives of f and g. Consequently, a great deal of attention has been recently paid to developing derivative-free schemes. One approach is to replace the derivatives in the stochastic Taylor approximations [16] by differences, and this leads to the derivative-free version of the Milstein method ($m = 1$)

$$\begin{aligned} Y_1 &= y_n + \sqrt{h_n} g(y_n), \\ y_{n+1} &= y_n + h_n f(y_n) + h_n \Delta W_n g(y_n) + \frac{\sqrt{h_n}}{2}\left(\left(\frac{\Delta W_n}{\sqrt{h_n}}\right)^2 - 1\right)\big(g(Y_1) - g(y_n)\big). \end{aligned} \tag{6}$$

Rümelin [18] has shown that methods such as (6) converge strongly with strong order at most 1 (see Section 2 for definitions of order), in comparison with a strong order of 0.5 for method (5). More general Runge–Kutta type schemes can be constructed but it is possible to show that a strong order of 1.5 cannot be surpassed if just the increments ΔW_n of the Wiener process are used. Thus as far as the authors are aware there has not appeared in the literature any general stochastic Runge–Kutta type methods (implicit or explicit) of strong order 2.0 or more. (Note that Chang [5] has constructed an explicit order 2.0 strong scheme in which the stochastic term is constant.)

The purpose of this paper is to show how this order bound can be overcome, by presenting a very general class of stochastic Runge–Kutta methods. In particular a family of methods of strong order 2.0 will be constructed in which the deterministic component of the method is the classical Runge–Kutta method given in (2). This general methodology will, in principle, allow the construction of even higher order methods in a systematic fashion. For this present paper the focus will only be on explicit methods—implicit methods will be considered in later works.

Thus the outline of the paper is as follows: in Section 2 a discussion on strong order and weak order along with the stochastic Taylor series formula will be given. In Section 3 the class of Runge–Kutta methods based on the Wiener increment will be introduced and general order conditions will be analysed to show that such methods cannot have a strong order greater than 1.5. In particular, a new explicit method of order 1.5 with minimal local error truncation coefficients will be constructed. In Section 4 new classes of Runge–Kutta methods will be constructed based on additional multiple stochastic integral terms from the stochastic Taylor series, and, in particular, an explicit 4-stage method of strong order 2.0 will be constructed. In Section 5 a number of numerical results will be presented which shows the efficacy of this new approach compared with extant methods of strong order 1.5, and the paper will finish with some conclusions and outlines for future work.

2. Stochastic differential equations

In the integral formulation of an SODE given in (4), $W(t)$ is a Wiener process and can be interpreted in such a way that the derivative of W is the Gaussian white noise process (so that $W(t)$ is in fact not differentiable). The process W is indeed a Gaussian process with independent increments and satisfies

$$\mathrm{E}\big(W(t)\big) = 0, \quad \mathrm{E}\big(W(t)^2\big) = t, \qquad \forall t \geqslant 0. \tag{7}$$

Since the sample paths of a Wiener process are not of bounded variation, the second integral in (4) cannot be interpreted in the Riemann–Stieltjes sense (see [6], for example). In fact, if different choices of τ_i are made from the subintervals $[t_{i-1}, t_i]$ the natural approximating sums

$$\sum_{i=1}^{N} g\big(y(\tau_i)\big)\big(W(t_i) - W(t_{i-1})\big) \tag{8}$$

converge in the mean square sense to different values of this integral. (Note that a sequence of random variables $\{X_n\}$ is said to have mean square convergence to a random variable X if $(\mathrm{E}(|X_n - X|^2))^{1/2} \to 0$ as $n \to \infty$.) In particular, if $\tau_i = t_{i-1}$ this gives the Itô integral, and this leads to a calculus based on Itô's chain rule.

A nice feature of the Itô integral is that it can be defined for a general class of non-anticipating random functions in such a way as to preserve various Wiener process properties as well as allowing easy calculation of moments of the solution of an SODE. In particular, the Itô integral forms a martingale which is a natural extension of the fact that $W(t)$ is a martingale, and, in addition, the Itô integral satisfies

$$\mathrm{E}\left[\int_a^b g(t)\,\mathrm{d}W(t)\right] = 0,$$

$$\mathrm{E}\left[\left|\int_a^b g(t)\,\mathrm{d}W(t)\right|^2\right] = \int_a^b \mathrm{E}\big[|g(t)|^2\big]\,\mathrm{d}t. \tag{9}$$

There are of course an infinite number of possibilities for the τ_i in (8) of the form

$$\tau_i = \theta t_i + (1-\theta)t_{i-1}$$

for which (8) gives

$$\int_a^b W(t)\,\mathrm{d}W(t) = \frac{1}{2}\big(W^2(b) - W^2(a)\big) + \left(\theta - \frac{1}{2}\right)(b-a). \tag{10}$$

Only for the non-anticipating Itô case ($\theta = 0$) does the martingale property (and (9)) hold, although another very important case ($\theta = 1/2$) leads to the Stratonovich calculus and the Stratonovich integral then satisfies the usual rules of calculus. The choice of which interpretation (Itô or Stratonovich) should be used depends on the type of analysis required for an SODE and in this paper the Stratonovich form will be used. In order to avoid any confusion in notation, henceforth the symbol $\circ$ will be used to denote the Stratonovich form (i.e., $\circ\,\mathrm{d}W(t)$).

It is important therefore to understand the relationship between the Itô and Stratonovich representations of a stochastic differential equation. Thus if the nonautonomous version of an SODE is given by

$$\mathrm{d}y = f(t,y)\,\mathrm{d}t + g(t,y)\,\mathrm{d}W \tag{11}$$

then the related Stratonovich SODE is given by

$$\mathrm{d}y = \bar{f}(t,y)\,\mathrm{d}t + g(t,y)\circ\mathrm{d}W, \tag{12}$$

where

$$\bar{f}(t,y) = f(t,y) - \frac{1}{2}\frac{\partial g}{\partial y}(t,y)g(t,y). \tag{13}$$

In other words the two equations (11) and (13), under different rules of calculus, have the same solution.

Finally, in this brief introduction to the theory of SODEs, the Itô form of the chain rule is given (see [10], for example). Thus for a given function F and with certain smoothness, measurability and boundedness properties on f and g in (11) to guarantee the existence, pathwise uniqueness and bounded second moments then

$$\mathrm{d}F(t,y) = \frac{\partial F}{\partial t}\,\mathrm{d}t + f^{\mathrm{T}}\frac{\partial F}{\partial y}\,\mathrm{d}t + \frac{1}{2}\,\mathrm{trace}\left(gg^{\mathrm{T}}\frac{\partial^2 F}{\partial y^2}\right)\mathrm{d}t + \frac{\partial F^{\mathrm{T}}}{\partial y}g\,\mathrm{d}W. \tag{14}$$

Some numerical time-discretization methods for the numerical solution of SODEs have already been discussed in Section 1. In order to evaluate the efficacy of such methods two ways of measuring accuracy are used: strong convergence and weak convergence.

For problems involving direct simulation it is important that the trajectories of the numerical approximation be close to the exact solution. Thus let $\bar{y}_N$ be the numerical approximation to $y(t_N)$ after N steps with constant step size $h = (t_N - t_0)/N$; then $\bar{y}$ is said to converge strongly to y with order p if $\exists C > 0$ (independent of h) and $\delta > 0$ such that

$$\mathrm{E}\big(|\bar{y}_N - y(t_N)|\big) \leqslant Ch^p, \quad h \in (0,\delta). \tag{15}$$

Note that in the deterministic case this reduces to the standard order condition. Here p can be fractional since the root mean square order of the Wiener process is $h^{1/2}$. Indeed the Euler–Maruyama scheme has strong order of convergence 0.5.

However, in some cases it is not necessary to find an accurate pathwise approximation of an Itô process. Instead, only some of the moments may be of interest or, more generally, $\mathrm{E}(g(\bar{y}))$ for some function g. This is a much weaker condition. Thus the discrete time approximation is said to converge weakly with order p to y if for each polynomial g (which is $2(p+1)$ times continuously differentiable), $\exists C > 0$ (independent of h) and $\delta > 0$ such that

$$\big|\mathrm{E}\big(g(\bar{y}_N)\big) - \mathrm{E}\big(g\big(y(t_N)\big)\big)\big| \leqslant Ch^p, \quad h \in (0,\delta). \tag{16}$$

Milstein [15] showed that the Euler–Maruyama scheme has weak order 1.0. In this present paper the focus will be on the strong order of explicit Runge–Kutta methods.

In the case of Runge–Kutta methods for deterministic problems, the order of accuracy is found by comparing the computed solution with the exact solution over one step assuming exact initial values. This necessitates the use of the Taylor series expansion, and a similar situation holds for SODEs in

which a numerical method is compared with the stochastic Taylor series expansion using either the Itô or Stratonovich calculus. Furthermore, because of the simplified nature of the Stratonovich calculus only the Stratonovich form of the stochastic Taylor series will be used here.

Thus consider the autonomous, one-dimensional Stratonovich SODE

$$\mathrm{d}y(t) = f\big(y(t)\big)\,\mathrm{d}t + g\big(y(t)\big)\circ \mathrm{d}W(t) \tag{17}$$

which can be written in integral form as

$$y(t) = y(t_0) + \int_{t_0}^{t} f\big(y(s)\big)\,\mathrm{d}s + \int_{t_0}^{t} g\big(y(s)\big)\circ \mathrm{d}W(s). \tag{18}$$

Itô's formula states that a given function a of the solution y can be written as

$$a\big(y(t)\big) = a\big(y(t_0)\big) + \int_{t_0}^{t} L^0 a\big(y(s)\big)\,\mathrm{d}s + \int_{t_0}^{t} L^1 a\big(y(s)\big)\circ \mathrm{d}W(s), \tag{19}$$

where in the Itô form

$$L^0 = f\frac{\partial}{\partial y} + \frac{1}{2}g^2\frac{\partial^2}{\partial y^2}, \qquad L^1 = g\frac{\partial}{\partial y},$$

while in the Stratonovich form

$$L^0 = f\frac{\partial}{\partial y}, \qquad L^1 = g\frac{\partial}{\partial y}. \tag{20}$$

Applying (18) and (20) with $a(y) \equiv y$ and $m = 1$ and writing $y(t_0) = y_0$ gives

$$\begin{aligned}
y(t) = {} & y_0 + f(y_0)J_0 + g(y_0)J_1 + f'(y_0)\big(f(y_0)\big)J_{00} + f'(y_0)\big(g(y_0)\big)J_{10} \\
& + g'(y_0)\big(f(y_0)\big)J_{01} + g'(y_0)\big(g(y_0)\big)J_{11} + f''(y_0)\big(f(y_0), f(y_0)\big)J_{000} \\
& + f'(y_0)\big(f'(y_0)\big(f(y_0)\big)\big)J_{000} + f''(y_0)\big(f(y_0), g(y_0)\big)J_{100} \\
& + f'(y_0)\big(f'(y_0)\big(g(y_0)\big)\big)J_{100} + f''(y_0)\big(g(y_0), f(y_0)\big)J_{010} \\
& + f'(y_0)\big(g'(y_0)\big(f(y_0)\big)\big)J_{010} + f''(y_0)\big(g(y_0), g(y_0)\big)J_{110} \\
& + f'(y_0)\big(g'(y_0)\big(g(y_0)\big)\big)J_{110} + g''(y_0)\big(f(y_0), f(y_0)\big)J_{001} \\
& + g'(y_0)\big(f'(y_0)\big(f(y_0)\big)\big)J_{001} + g''(y_0)\big(f(y_0), g(y_0)\big)J_{101} \\
& + g'(y_0)\big(f'(y_0)\big(g(y_0)\big)\big)J_{101} + g''(y_0)\big(g(y_0), f(y_0)\big)J_{011} \\
& + g'(y_0)\big(g'(y_0)\big(f(y_0)\big)\big)J_{011} + g''(y_0)\big(g(y_0), g(y_0)\big)J_{111} \\
& + g'(y_0)\big(g'(y_0)\big(g(y_0)\big)\big)J_{111} + R.
\end{aligned} \tag{21}$$

Here R is a remainder term and $J_{j_1 j_2 \cdots j_k}$ represents the Stratonovich multiple integral, where integration is with respect to $\mathrm{d}s$ if $j_i = 0$ or $\circ\,\mathrm{d}W(s)$ if $j_i = 1$. Thus, for example, in one dimension

$$J_{101} = \int_{t_0}^{t}\int_{t_0}^{s_2}\int_{t_0}^{s} \circ\,\mathrm{d}W(s_1)\,\mathrm{d}s\circ \mathrm{d}W(s_2).$$

Remarks.

(1) Eq. (21) is the generalization of the Taylor series expansion for deterministic equations with $g \equiv 0$.

(2) Just as there is a rooted tree expansion in terms of elementary differentials for the Taylor series expansion of the solution of (1) (see [4]), so there is a similar expansion for the Stratonovich Taylor series expansion of the solution of (18).

(3) Kloeden and Platen [10] give these stochastic Taylor series expansions in terms of a hierarchical set concept but do not relate it to the rooted tree theory introduced by Butcher [3] for the deterministic case. For the purposes of unifying the theory for both deterministic and stochastic problems this unification is now given (see [2,11]).

This unification is based on the consideration of the set of bi-coloured rooted trees, T, in which $\bullet$ (τ for a deterministic node) and $\circ$ (σ for a stochastic node) play the crucial role. Thus if $t_1, \ldots, t_m$ are bi-coloured trees then $[t_1, \ldots, t_m]$ and $\{t_1, \ldots, t_m\}$ are trees in which $t_1, \ldots, t_m$ are each joined by a single branch to $\bullet$ or $\circ$, respectively. In a similar manner to the deterministic case [4], an elementary differential can be associated with any $t \in T$ such that

$$F(\tau)(y) = f(y), \qquad F(\sigma)(y) = g(y),$$

$$F(t)(y) = \begin{cases} f^{(m)}(y)\big[F(t_1)(y), \ldots, F(t_m)(y)\big], & t = [t_1, \ldots, t_m], \\ g^{(m)}(y)\big[F(t_1)(y), \ldots, F(t_m)(y)\big], & t = \{t_1, \ldots, t_m\}. \end{cases} \tag{22}$$

In addition, an elementary weight can be associated with each elementary differential (see (21)) by associating the integer 0 with a deterministic node ($\bullet$) and the integer 1 with a stochastic node ($\circ$). These elementary weights are in fact integrals. With $y(t)$ the solution of (18) and for any integrable function F of y, define

$$J_0(F) = \int_{t_0}^{t} F\big(y(s)\big)\,\mathrm{d}s, \qquad J_1(F) = \int_{t_0}^{t} F\big(y(s)\big) \circ \mathrm{d}W(s), \tag{23}$$

then these elementary weights can be written as

$$\theta(\tau) = J_0(1), \qquad \theta(\sigma) = J_1(1),$$

$$\theta(t) = \begin{cases} J_0\left(\prod_{j=1}^{m} \theta(t_j)\right), & t = [t_1, \ldots, t_m], \\ J_1\left(\prod_{j=1}^{m} \theta(t_j)\right), & t = \{t_1, \ldots, t_m\}. \end{cases} \tag{24}$$

As in [4], let $\alpha(t)$ be the number of ways of labelling t with a set of $\rho(t)$ (the number of nodes of t) ordered symbols such that along each outwardly directed arc the labels increase; then it can be shown that the Stratonovich Taylor series is given by

$$y(t) = \sum_{t \in T} \alpha(t) F(t)\big(y(t_0)\big)\theta(t). \tag{25}$$

Note here that isomorphic trees are regarded as identical.

While for lower order Runge–Kutta methods it is sufficient to obtain the order conditions by comparing directly the Runge–Kutta scheme with the stochastic Stratonovich Taylor expansion (see, for example, [10]), the development of the order conditions via tree theory allows the structure of more general Runge–Kutta methods to be described.

3. Runge–Kutta methods for SODEs

Certain classes of one-step multistage methods have been introduced for solving the general SODE problem (3) (see [6], for example). Perhaps the most general class of methods considered so far takes the form

$$\begin{aligned} Y_i &= y_n + h\sum_{j=1}^{s} a_{ij} f(Y_j) + J_1 \sum_{j=1}^{s} b_{ij} g(Y_j), \quad i = 1,\dots,s, \\ y_{n+1} &= y_n + h\sum_{j=1}^{s} \alpha_j f(Y_j) + J_1 \sum_{j=1}^{s} \gamma_j g(Y_j). \end{aligned} \tag{26}$$

Here $A = (a_{ij})$ and $B = (b_{ij})$ are $s \times s$ matrices of real elements while $\alpha^{\mathrm{T}} = (\alpha_1,\dots,\alpha_s)$ and $\gamma^{\mathrm{T}} = (\gamma_1,\dots,\gamma_s)$ are row vectors $\in \mathbb{R}^s$. If both A and B are strictly lower triangular then (26) is said to be explicit, otherwise it is implicit. The stochastic component comes from the J_1 integral ($J_1 = \int_{t_n}^{t_{n+1}} \mathrm{d}W$) associated with B and γ. Most authors [6,18] only consider explicit methods.

Rümelin [18] has shown that if f and g and the necessary partial derivatives of f and g are bounded then (26) converges uniformly on $[t_0, T]$ in the quadratic mean sense to the Itô solution of

$$\mathrm{d}y = f(y) + \lambda \frac{\partial g}{\partial y} g(y) + g(y)\,\mathrm{d}W,$$

where

$$\lambda = \gamma^{\mathrm{T}} B e.$$

Furthermore, if $\lambda = 1/2$ then (26) converges to the solution of the corresponding Stratonovich equation.

In particular Rümelin [18] has proven

Theorem 1. *If f and g are arbitrary functions and have continuous and bounded partial derivatives up to the sixth order then the strong order of* (26) *cannot exceed* 1.5.

Examples of a Runge–Kutta method of the form (26) with strong order 1.5 include the method of Platen which can be written with $s = 2$ as

$$A = B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad \alpha^{\mathrm{T}} = (1,0), \qquad \gamma^{\mathrm{T}} = \left(\tfrac{1}{2}, \tfrac{1}{2}\right). \tag{27}$$

Thus if higher strong order methods are required, (26) needs to be modified in some way so as to include other stochastic elements apart from just J_1. This will be done by the introduction of an arbitrary matrix Z and vector z^{T} whose elements are themselves random variables. Hence the general family of s-stage stochastic Runge–Kutta methods (SRKs) will be given by

$$\begin{aligned} Y_i &= y_n + h\sum_{j=1}^{s} a_{ij} f(Y_j) + \sum_{j=1}^{s} Z_{ij} g(Y_j), \quad i = 1, \ldots, s, \\ y_{n+1} &= y_n + h\sum_{j=1}^{s} \alpha_j f(Y_j) + \sum_{j=1}^{s} z_j g(Y_j). \end{aligned} \tag{28}$$

Clearly (26) is now a specific case of (28).

By studying the general strong order properties of (28) for arbitrary random variable elements within Z and z, we will show how the strong order barrier given in Theorem 1 can be broken by constructing a class of 4-stage explicit SRKs with strong order 2 which relies on both the random variables J_1 and J_{10}. At the same time these order conditions will be used to construct a two-stage method of the form (26) which is optimal in terms of minimising the local truncation error coefficients.

Of course (28) is a very general representation, and so a simplifying assumption will be placed on the Z_{ij} and z_j in that it will be assumed that each of these random variables can be written as a linear combination of p different random variables $\theta_1, \ldots, \theta_p$, where the time dependence is implicitly assumed. Thus the Z_{ij} and z_j will be written as

$$\begin{aligned} Z_{ij} &= \sum_{l=1}^{p} b_{ij}^{(l)} \theta_l, \quad i, j = 1, \ldots, s, \\ z_j &= \sum_{l=1}^{p} \gamma_j^{(l)} \theta_l, \quad j = 1, \ldots, s, \end{aligned} \tag{29}$$

and (28) can be written as

$$\begin{aligned} Y_i &= y_n + h\sum_{j=1}^{s} a_{ij} f(Y_j) + \sum_{l=1}^{p} \left(\sum_{j=1}^{s} b_{ij}^{(l)} g(Y_j) \right) \theta_l, \quad i = 1, \ldots, s, \\ y_{n+1} &= y_n + h\sum_{j=1}^{s} \alpha_j f(Y_j) + \sum_{l=1}^{p} \left(\sum_{j=1}^{s} \gamma_j^{(l)} g(Y_j) \right) \theta_l. \end{aligned} \tag{30}$$

This family of methods can be characterized by the tableau

$$\begin{array}{c|ccc} A & B^{(1)} & \ldots & B^{(p)} \\ \hline \alpha & \gamma^{(1)} & \ldots & \gamma^{(p)} \end{array}. \tag{31}$$

In order to study the order conditions associated with (30) the approach of Butcher [4] will be used in which by writing $t_n = t_0$ and for a given $t = t_0 + h$, y_{n+1} will be calculated as $Y(t)$ with intermediate values $Y_1(t), \ldots, Y_s(t)$ given by

$$\begin{aligned} Y_i(t) &= y(t_0) + (t - t_0)\sum_{j=1}^{s} a_{ij} f\big(Y_j(t)\big) + \sum_{l=1}^{p} \theta_l \sum_{j=1}^{s} b_{ij}^{(l)} g\big(Y_j(t)\big), \quad i = 1, \ldots, s, \\ Y(t) &= y(t_0) + (t - t_0)\sum_{i=1}^{s} \alpha_i f\big(Y_i(t)\big) + \sum_{l=1}^{p} \theta_l \sum_{i=1}^{s} \gamma_i^{(l)} g\big(Y_j(t)\big). \end{aligned} \tag{32}$$

Note that it will be assumed that $\theta_l(t_0) = 0,\ l = 1, \ldots, p$.

By substituting for $Y_i(t)$ in the expression for $Y(t)$ in (32), the $f(Y_i(t))$ can be expanded using the Taylor series expansion

$$f\big(Y_i(t)\big) = f\big(Y(t_0)\big) + \sum_{k=1}^{\infty} \frac{(t-t_0)^k}{k!} L^k f\big(Y_i(t_0)\big), \tag{33}$$

where L is the differential operator given by

$$L = \frac{\partial}{\partial t} + \sum_{j=1}^{p} \frac{\partial}{\partial \theta_j}, \tag{34}$$

and where it is assumed that f (and g) are sufficiently differentiable.

Thus, for example,

$$f\big(Y(t_0)\big) = f\big(y(t_0)\big)$$

$$Lf\big(Y_i(t)\big) = f'\big(Y_i(t)\big)\left(\sum_{j=1}^{s} a_{ij} f\big(Y_j(t)\big) + (t-t_0)\sum_{j=1}^{s} f'\big(Y_j(t)\big)\frac{\partial Y_j}{\partial t} + \sum_{l=1}^{p}\theta_l \sum_{j=1}^{s} b_{ij}^{(l)} g'\big(Y_j(t)\big)\frac{\partial Y_j}{\partial t}\right) + \sum_{l=1}^{p}\theta_l f'\big(Y_i(t)\big)\sum_{j=1}^{s} b_{ij}^{(l)} g\big(Y_j(t)\big),$$

so that

$$Lf\big(Y_i(t_0)\big) = \sum_{j=1}^{s} a_{ij} f'\big(y(t_0)\big) f\big(y(t_0)\big) + \sum_{l=1}^{p}\theta_l \sum_{j=1}^{s} b_{ij}^{(l)} f'\big(y(t_0)\big) g\big(y(t_0)\big).$$

Similarly $L^2 f(Y_i(t)) = L(Lf(Y_i(t))$, etc. (see [2] for more details).

In fact if the Z_{ij} and z_j satisfy a linear relationship between the p different random variables as described in (29) then it can be shown that the Taylor series expansion for the numerical method can be written as

$$Y(t) = \sum_{t\in T} \Phi(t) F(t)\big(y(t_0)\big)\frac{h^{\rho_1(t)}}{\rho(t)!}. \tag{35}$$

Here $\Phi(t)$ is defined recursively by

$$k(\phi) = e,$$

$$\Phi(t) = \begin{cases} \rho(t)\alpha^{\mathrm{T}} \displaystyle\prod_{l=1}^{m} k(t_l), & t = [t_1, \ldots, t_m], \\ \rho(t) z^{\mathrm{T}} \displaystyle\prod_{l=1}^{m} k(t_l), & t = \{t_1, \ldots, t_m\}, \end{cases} \tag{36}$$

where

$$k(t) = \rho(t) A \prod_{l=1}^{m} k(t_l), \quad t = [t_1, \ldots, t_m],$$

$$k(t) = \rho(t) Z \prod_{l=1}^{m} k(t_l), \quad t = \{t_1, \ldots, t_m\}, \tag{37}$$

and where multiplication of vectors is considered componentwise. In addition $\rho(t)$ is the number of nodes of t while $\rho_1(t)$ is the number of $\bullet$ (deterministic) nodes of t.

Hence the local truncation error at $t = t_n$ of an SRK method can be written as

$$L_n = \sum_{t \in T} \left(\alpha(t)\theta(t) - \Phi(t) \frac{h^{\rho_1(t)}}{\rho(t)!} \right) F(t)\big(y(t_n)\big). \tag{38}$$

Thus if

$$\sqrt{\mathrm{E}\big(|L_n|\big)^2} \leqslant Ch^p$$

then a method will have strong order p.

Writing L_n as

$$L_n = \sum_{t \in T} e(t) F(t)\big(y(t_n)\big) \tag{39}$$

and letting

$$c = Ae, \qquad \lambda = Ze \tag{40}$$

then Table 1 gives $e(t)$ for all trees with $\rho(t) \leqslant 3$.

We are now in a position to study general order conditions. The first point to note is that for every tree t in the deterministic case there are $2^{\rho(t)}$ trees that must be considered in the stochastic case. Table 2 illustrates how quickly the number of trees grows.

We will now study the order properties of the class of methods given by (26) in which

$$Z = J_1 B, \qquad z = J_1 \gamma^{\mathrm{T}}, \qquad b = Be.$$

From Table 1 a necessary condition for strong order 2 is from condition 4

$$\mathrm{E}(J_{10} - hJ_1\psi)^2 = \mathrm{O}\big(h^4\big), \tag{41}$$

where

$$\psi = \alpha^{\mathrm{T}} b.$$

Now the lefthand side of (41) is given by

$$\mathrm{E}\big[J_{10}^2\big] - 2\psi h\, \mathrm{E}[J_{10}J_1] + \psi^2 h^2\, \mathrm{E}\big[J_1^2\big]. \tag{42}$$

But $J_1 \sim \mathrm{N}(0, h)$ and from [10]

$$\mathrm{E}\big[J_1^{2k+1}\big] = 0, \qquad \mathrm{E}\big[J_1^{2k}\big] = \frac{2k!}{k!2^k} h^k, \qquad \mathrm{E}\big[J_{10}J_1\big] = \frac{1}{2}h^2, \qquad \mathrm{E}[J_{10}^2] = \frac{1}{3}h^3 \tag{43}$$

and so (42) is

$$h^3 \left(\frac{1}{3} - \psi + \psi^2 \right) \neq 0.$$

Table 1
Local error coefficients

#	t	$e(t)$
1	$\bullet$	$J_0 - h\alpha^{\mathrm{T}}e$
2	$\circ$	$J_1 - z^{\mathrm{T}}e$
3	$[\bullet]$	$J_{00} - h^2\alpha^{\mathrm{T}}c$
4	$[\circ]$	$J_{10} - h\alpha^{\mathrm{T}}\lambda$
5	$\{\bullet\}$	$J_{01} - hz^{\mathrm{T}}c$
6	$\{\circ\}$	$J_{11} - z^{\mathrm{T}}\lambda$
7	$[\bullet,\bullet]$	$J_{000} - \frac{1}{2}h^3\alpha^{\mathrm{T}}c^2$
8	$[[\bullet]]$	$J_{000} - h^3\alpha^{\mathrm{T}}Ac$
9	$[[\circ]]$	$J_{100} - h^2\alpha^{\mathrm{T}}A\lambda$
10	$[\bullet,\circ]$	$J_{100} - \frac{1}{2}h^2\alpha^{\mathrm{T}}c\lambda$
11	$[\circ,\bullet]$	$J_{010} - \frac{1}{2}h^2\alpha^{\mathrm{T}}c\lambda$
12	$[\{\bullet\}]$	$J_{010} - h^2\alpha^{\mathrm{T}}Zc$
13	$[\{\circ\}]$	$J_{110} - h\alpha^{\mathrm{T}}Z\lambda$
14	$[\circ,\circ]$	$J_{110} - \frac{1}{2}h\alpha^{\mathrm{T}}\lambda^2$
15	$\{[\bullet]\}$	$J_{001} - h^2z^{\mathrm{T}}Ac$
16	$\{[\circ]\}$	$J_{101} - hz^{\mathrm{T}}A\lambda$
17	$\{\bullet,\bullet\}$	$J_{001} - \frac{1}{2}h^2z^{\mathrm{T}}c^2$
18	$\{\bullet,\circ\}$	$J_{101} - \frac{1}{2}hz^{\mathrm{T}}c\lambda$
19	$\{\circ,\bullet\}$	$J_{011} - \frac{1}{2}hz^{\mathrm{T}}c\lambda$
20	$\{\circ,\circ\}$	$J_{111} - \frac{1}{2}z^{\mathrm{T}}\lambda^2$
21	$\{\{\bullet\}\}$	$J_{011} - hz^{\mathrm{T}}Zc$
22	$\{\{\circ\}\}$	$J_{111} - z^{\mathrm{T}}Z\lambda$

Table 2
Number of trees

$\rho(t)$	1	2	3	4
deterministic	1	1	2	4
stochastic	2	4	16	64

In fact the minimum of the quadratic occurs when $\psi = 1/2$ in which case the minimum value is 1/12. This leads us to construct the complete class of explicit SRK methods of the form (26) with strong order 1.5 and minimum principal local truncation error.

Now trees 1, 2 and 6 are of order h, $\sqrt{h}$ and h, respectively so that it is necessary for

$$\alpha^{\mathrm{T}}e = 1, \qquad \gamma^{\mathrm{T}}e = 1 \tag{44}$$

and

$$\mathrm{E}\left(J_{11} - J_1^2\gamma^{\mathrm{T}}b\right)^2 = 0. \tag{45}$$

Since

$$J_{1\ldots1} = \frac{J_1^p}{p!},$$

(43) and (45) imply

$$\frac{3}{4} - 3\gamma^{\mathrm{T}}b + 3(\gamma^{\mathrm{T}}b)^2 = 0$$

or

$$\gamma^{\mathrm{T}}b = \frac{1}{2}. \tag{46}$$

The terms corresponding to the $h^{1.5}$ terms arise from trees 4, 5, 20 and 22. These give, from (43),

$$\begin{aligned}
&\mathrm{E}\left(J_{10}^2 - hJ_1\alpha^{\mathrm{T}}b\right)^2 = \left(\tfrac{1}{3} - \alpha^{\mathrm{T}}b + (\alpha^{\mathrm{T}}b)^2\right)h^3,\\
&\mathrm{E}\left(J_{01}^2 - hJ_1\gamma^{\mathrm{T}}c\right)^2 = \left(\tfrac{1}{3} - \gamma^{\mathrm{T}}c + (\gamma^{\mathrm{T}}c)^2\right)h^3,\\
&\mathrm{E}\left(J_{111} - \tfrac{1}{2}J_1^3\gamma^{\mathrm{T}}b^2\right)^2 = \left(\tfrac{1}{9} - \tfrac{2}{3}\gamma^{\mathrm{T}}b^2 + (\gamma^{\mathrm{T}}b^2)^2\right)\tfrac{15}{4}h^3,\\
&\mathrm{E}\left(J_{111} - J_1^3\gamma^{\mathrm{T}}Bb\right)^2 = \left(\tfrac{1}{36} - \tfrac{1}{3}\gamma^{\mathrm{T}}Bb + (\gamma^{\mathrm{T}}Bb)^2\right)15h^3.
\end{aligned} \tag{47}$$

These four equations are minimized if

$$\alpha^{\mathrm{T}}b = \frac{1}{2}, \qquad \gamma^{\mathrm{T}}c = \frac{1}{2}, \qquad \gamma^{\mathrm{T}}b^2 = \frac{1}{3}, \qquad \gamma^{\mathrm{T}}Bb = \frac{1}{6} \tag{48}$$

in which case the (respective) minima are

$$\frac{h^3}{12}, \qquad \frac{h^3}{12}, \qquad 0, \qquad 0. \tag{49}$$

Note if $s = 2$, $\gamma^{\mathrm{T}}Bb = 0$ and the principal error constants are

$$\frac{h^3}{12}, \qquad \frac{h^3}{12}, \qquad 0, \qquad \frac{5h^3}{12}. \tag{50}$$

The $s = 2$ and $s = 3$ methods that lead to these optimal methods have tableaus, respectively,

$$\begin{array}{c|cc} 0 & 0 & \\ \frac{2}{3} & 0 & \\ \hline & \frac{1}{4} & \frac{3}{4} \end{array} \qquad \begin{array}{c|cc} 0 & 0 & \\ \frac{2}{3}J_1 & 0 & \\ \hline & \frac{1}{4}J_1 & \frac{3}{4}J_1 \end{array}$$

$$\begin{array}{c|ccc} 0 & 0 & 0 \\ c_2 & 0 & 0 \\ c_3 - \theta & \theta & 0 \\ \hline & \alpha_1 & \alpha_2 \quad \alpha_3 \end{array} \qquad \begin{array}{c|ccc} 0 & 0 & 0 \\ b_2J_1 & 0 & 0 \\ (b_3 - \psi)J_1 & \psi J_1 & 0 \\ \hline & \gamma_1 & \gamma_2 \quad \gamma_3 \end{array} \tag{51}$$

where

$$\psi = \frac{1}{6b_2\gamma_3}, \qquad \gamma_2 = \frac{\frac{1}{3} - \frac{1}{2}b_2}{b_2(b_2 - b_3)}, \qquad \gamma_3 = \frac{\frac{1}{3} - \frac{1}{2}b_3}{b_3(b_3 - b_2)}, \qquad \gamma_1 = 1 - \gamma_2 - \gamma_3,$$

$$\alpha_1 = 1 - \alpha_2 - \alpha_3, \qquad \alpha_2 = \frac{1}{2b_2} - \alpha_3\frac{b_3}{b_2}, \qquad \frac{c_3}{b_3} - \frac{c_2}{b_2} = \frac{3}{2}\big(c_3 - c_2 - (b_3 - b_2)\big).$$

Note that the Platen method

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 0 & 0 \\ \hline & 1 & 0 \end{array} \qquad \begin{array}{c|cc} 0 & 0 & 0 \\ J_1 & 0 & 0 \\ \hline & \frac{1}{2}J_1 & \frac{1}{2}J_1 \end{array} \tag{52}$$

has principal error constants

$$\frac{h^3}{3}, \qquad \frac{h^3}{3}, \qquad \frac{h^3}{36}, \qquad \frac{5}{12}h^3. \tag{53}$$

4. An explicit SRK with strong order 2

In this section an explicit SRK method with strong order 2 will be constructed based on (31) with $p = 2$ and with

$$\theta_1 = J_1, \qquad \theta_2 = J_{10}/h.$$

Thus the method can be written as

$$\begin{aligned} Y_i &= y_n + h\sum_{j=1}^{i-1} a_{ij} f(Y_j) + \sum_{j=1}^{i-1}\left(b_{ij}^{(1)}J_1 + b_{ij}^{(2)}\frac{J_{10}}{h}\right)g(Y_j), \quad i = 1,\dots,s, \\ y_{n+1} &= y_n + h\sum_{j=1}^{s}\alpha_j f(Y_j) + \sum_{j=1}^{s}\left(\gamma_j^{(1)}J_1 + \gamma_j^{(2)}\frac{J_{10}}{h}\right)g(Y_j). \end{aligned} \tag{54}$$

For convenience it will be assumed that

$$c = Ae, \qquad b = B^{(1)}e, \qquad d = B^{(2)}e, \qquad \lambda = bJ_1 + d\frac{J_{10}}{h}.$$

In passing it should be noted that in constructing an SRK method based on only the first and second order Stratonovich multiple integrals only J_1 and J_{10}/h need to be considered, since the other multiple integrals of order 2 or less can be expressed in terms of these integrals and h by the formula

$$J_{00} = \frac{h^2}{2}, \qquad J_{11} = \frac{J_1^2}{2}, \qquad \frac{J_{01}}{h} = J_1 - \frac{J_{10}}{h}$$

(see [10], for example).

In order to construct a method with strong order 2, tree conditions 1, 2, 4, 5, 6, 20 and 22 must be considered, since $E(e(t)^2)$ for all of these trees is $O(h^p)$, $p \leqslant 3$. Each of these terms will now be

Table 3
Expectation values

Expectation	value
J_1^2	h
J_{10}^2/h^2	$\frac{1}{3}h$
$J_1 J_{10}/h$	$\frac{1}{2}h$
J_1^4	$3h^2$
$J_1^3 J_{10}/h$	$\frac{3}{2}h^2$
$J_1^2(J_{10}/h)^2$	$\frac{5}{6}h^2$
$J_1(J_{10}/h)^3$	$\frac{1}{2}h^2$
$(J_{10}/h)^4$	$\frac{1}{3}h^2$
J_1^6	$15h^3$
$J_1^5 J_{10}/h$	$\frac{15}{2}h^3$
$J_1^4(J_{10}/h)^2$	$4h^3$
$J_1^3(J_{10}/h)^3$	$\frac{9}{4}h^3$
$J_1^2(J_{10}/h)^4$	$\frac{4}{3}h^3$
$J_1(J_{10}/h)^5$	$\frac{5}{6}h^3$
$(J_{10}/h)^6$	$\frac{5}{9}h^3$

considered in turn but first a table of various expected values will be given (Table 3) which will prove helpful in this analysis.

1. $\mathrm{E}\big[(J_0 - h\alpha^{\mathrm{T}}e)^2\big] = (1 - \alpha^{\mathrm{T}}e)h^2 \Longrightarrow \alpha^{\mathrm{T}}e = 1.$

2. $$\mathrm{E}\big[(J_1 - z^{\mathrm{T}}e)^2\big] = \mathrm{E}\big[\big(J_1\big(1 - \gamma^{(1)\mathrm{T}}e\big) - (J_{10}/h)\gamma^{(2)\mathrm{T}}e\big)^2\big]$$
$$= (g_1, g_2)\begin{pmatrix} \mathrm{E}(J_1)^2 & -\mathrm{E}(J_1 J_{10}/h) \\ -\mathrm{E}(J_1 J_{10}/h) & \mathrm{E}(J_{10}/h)^2 \end{pmatrix}\begin{pmatrix} g_1 \\ g_2 \end{pmatrix}$$
$$= (g_1, g_2)\begin{pmatrix} \frac{1}{3} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix}\begin{pmatrix} g_1 \\ g_2 \end{pmatrix}h^2 \geqslant 0$$
$$\Longrightarrow g_1 = g_2 = 0$$
$$\Longrightarrow \gamma^{(1)\mathrm{T}}e = 1,\ \gamma^{(2)\mathrm{T}}e = 0.$$

4. $$\mathrm{E}\big[\big(J_{10} - h\big(\alpha^{\mathrm{T}}bJ_1 + \alpha^{\mathrm{T}}dJ_{10}/h\big)\big)^2\big] = \mathrm{E}\big[\big(J_{10}(1 - \alpha^{\mathrm{T}}d) - h\alpha^{\mathrm{T}}bJ_1\big)^2\big]$$
$$\Longrightarrow \alpha^{\mathrm{T}}d = 1,\ \alpha^{\mathrm{T}}b = 0 \quad \text{(case 2)}.$$

5. $$\mathrm{E}\big[\big(J_{01} - h\big(\gamma^{(1)\mathrm{T}}cJ_1 + \gamma^{(2)\mathrm{T}}cJ_{10}/h\big)\big)^2\big] = h^2\mathrm{E}\big[\big(J_1\big(1 - \gamma^{(1)\mathrm{T}}c\big) - (J_{10}/h)\big(1 + \gamma^{(2)\mathrm{T}}c\big)\big)^2\big]$$
$$\Longrightarrow \gamma^{(1)\mathrm{T}}c = 1,\ \gamma^{(2)\mathrm{T}}c = -1 \quad \text{(case 2)}.$$

6. $\mathrm{E}\big[(J_{11} - z^{\mathrm{T}}\lambda)^2\big] = \mathrm{E}\big[\big(J_1^2/2 - z^{\mathrm{T}}\lambda\big)^2\big]$
$= \mathrm{E}\big[\big(J_1^2\big(\frac{1}{2} - \gamma^{(1)\mathrm{T}}b\big) - J_1(J_{10}/h)\big(\gamma^{(1)\mathrm{T}}d + \gamma^{(2)\mathrm{T}}b\big) - \big(J_{10}/h\big)^2\gamma^{(2)\mathrm{T}}d\big)^2\big]$
$= g^{\mathrm{T}}Xg,$

where X is the 3×3 matrix with elements

$x_{ij} = \mathrm{E}\big[J_1^{6-(i+j)}(J_{10}/h)^{i+j-2}\big].$

So condition 6

$$= h^2(g_1, g_2, g_3)\begin{pmatrix} 3 & \frac{3}{2} & \frac{5}{6} \\ \frac{3}{2} & \frac{5}{6} & \frac{1}{2} \\ \frac{5}{6} & \frac{1}{2} & \frac{1}{3} \end{pmatrix}\begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix} \geqslant 0$$

$$\Longrightarrow \gamma^{(1)\mathrm{T}}b = \tfrac{1}{2},\ \gamma^{(1)\mathrm{T}}d + \gamma^{(2)\mathrm{T}}b = 0,\ \gamma^{(2)\mathrm{T}}d = 0.$$

22. $\mathrm{E}\big[\big(J_{111} - \frac{1}{2}z^{\mathrm{T}}\lambda^2\big)^2\big] = \mathrm{E}\big[\big(J_1^3/6 - \frac{1}{2}z^{\mathrm{T}}\lambda^2\big)^2\big] = g^{\mathrm{T}}Xgh^3,$

where

$$g^{\mathrm{T}} = \big(\tfrac{1}{6} - \gamma^{(1)\mathrm{T}}b^2, -\tfrac{1}{2}\big(\gamma^{(2)\mathrm{T}}b^2 + 2\gamma^{(1)\mathrm{T}}bd\big), -\tfrac{1}{2}\big(\gamma^{(1)\mathrm{T}}d^2 + 2\gamma^{(2)\mathrm{T}}bd\big), -\tfrac{1}{2}\gamma^{(2)\mathrm{T}}d^2\big),$$

$$x_{ij} = \mathrm{E}\big[J_1^{8-(i+j)}(J_{10}/h)^{i+j-2}\big].$$

Thus condition 22 becomes

$$g^{\mathrm{T}}\begin{pmatrix} 15 & \frac{15}{2} & 4 & \frac{9}{4} \\ \frac{15}{2} & 4 & \frac{9}{4} & \frac{4}{3} \\ 4 & \frac{9}{4} & \frac{4}{3} & \frac{5}{6} \\ \frac{9}{4} & \frac{4}{3} & \frac{5}{6} & \frac{5}{9} \end{pmatrix} g \geqslant 0$$

and this implies

$$\gamma^{(1)\mathrm{T}}b^2 = \tfrac{1}{3}, \qquad \gamma^{(2)\mathrm{T}}b^2 + 2\gamma^{(1)\mathrm{T}}bd = 0, \qquad \gamma^{(1)\mathrm{T}}d^2 + 2\gamma^{(2)\mathrm{T}}bd = 0, \qquad \gamma^{(2)\mathrm{T}}d^2 = 0.$$

22. $\mathrm{E}\big[\big(J_{111} - zZ^{\mathrm{T}}\lambda\big)^2\big] = g^{\mathrm{T}}Xgh^3 \geqslant 0,$

where X is as above and $g = 0$ implies

$$\gamma^{(1)\mathrm{T}}B^{(1)}b = \tfrac{1}{6}, \qquad \gamma^{(2)\mathrm{T}}B^{(1)}b + \gamma^{(1)\mathrm{T}}\big(B^{(2)}b + B^{(1)}d\big) = 0,$$

$$\gamma^{(2)\mathrm{T}}B^{(2)}d = 0, \qquad \gamma^{(1)\mathrm{T}}B^{(2)}d + \gamma^{(2)\mathrm{T}}\big(B^{(2)}b + B^{(1)}d\big) = 0.$$

Thus a method of the form given in (54) will have strong order 2 if and only if

$$\begin{aligned}
&\alpha^{\mathrm{T}}[e,\ d,\ b] = [1,\ 1,\ 0], \\
&\gamma^{(1)\mathrm{T}}[e,\ d,\ b,\ c] = \big[1,\ -\gamma^{(2)\mathrm{T}}b,\ \tfrac{1}{2},\ 1\big], \\
&\gamma^{(2)\mathrm{T}}[e,\ d,\ c] = [0,\ 0,\ -1], \\
&\gamma^{(1)\mathrm{T}}\big[b^2,\ B^{(1)}b,\ d^2,\ B^{(2)}d\big] = \big[\tfrac{1}{3},\ \tfrac{1}{6},\ -2\gamma^{(2)\mathrm{T}}bd,\ -\gamma^{(2)\mathrm{T}}\big(B^{(2)}b + B^{(1)}d\big)\big], \\
&\gamma^{(2)\mathrm{T}}\big[b^2,\ B^{(1)}b,\ d^2,\ B^{(2)}d\big] = \big[-2\gamma^{(1)\mathrm{T}}bd,\ -\gamma^{(1)\mathrm{T}}\big(B^{(2)}b + B^{(1)}d\big),\ 0,\ 0\big].
\end{aligned} \tag{55}$$

It is now necessary to construct a family of methods satisfying (55). Some simple analysis shows that this is not possible with $s = 3$, since in this case, with d_1, d_2 and d_3 distinct,

$$\gamma^{(2)\mathrm{T}} = 0.$$

However, it is possible to satisfy (55) if $s = 4$ with a large number of free parameters. In principle it is possible to choose these free parameters in such a way so as many as possible of the h^5 principal error terms are minimized. However, since this paper is partially written to honour the pioneering work of Runge and Kutta, most of these free parameters will be chosen so that the deterministic component of (54) is the classical Runge–Kutta method of Kutta [12] given by

$$\begin{array}{c|cccc} 0 & 0 & & & \\ \frac{1}{2} & \frac{1}{2} & & & \\ \frac{1}{2} & 0 & \frac{1}{2} & & \\ 1 & 0 & 0 & 1 & \\ \hline & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} \end{array}$$

Since the general 4-stage method (54) has 30 free parameters and there are 18 equations to be solved, the above choice for the deterministic component still leaves 3 free parameters. The remaining 17 conditions (one is already satisfied by $\alpha^{\mathrm{T}} e = 1$) were solved using MAPLE and in order to simplify the solution a choice of $b_4 = 1$ was made. This leads to a number of possible methods and the following method was selected:

$$A = \begin{pmatrix} 0 & & & \\ \frac{1}{2} & 0 & & \\ 0 & \frac{1}{2} & 0 & \\ 0 & 0 & 1 & 0 \end{pmatrix}, \qquad \alpha^{\mathrm{T}} = \left(\tfrac{1}{6}, \tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{6}\right),$$

$$B^{(1)} = \begin{pmatrix} 0 & & & \\ -0.7242916356 & & & \\ 0.4237353406 & -0.1994437050 & & \\ -1.578475506 & 0.840100343 & 1.738375163 & 0 \end{pmatrix},$$

$$B^{(2)} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 2.702000410 & 0 & 0 & 0 \\ 1.757261649 & 0 & 0 & 0 \\ -2.918524118 & 0 & 0 & 0 \end{pmatrix}, \tag{56}$$

$$\gamma^{(1)\mathrm{T}} = (-0.7800788474,\ 0.07363768240,\ 1.486520013,\ 0.2199211524),$$

$$\gamma^{(2)\mathrm{T}} = (1.693950844,\ 1.636107882,\ -3.024009558,\ -0.3060491602).$$

5. Numerical results and conclusions

In this section, numerical results from the implementation of three methods are presented. The methods are the Platen method (52), the new two-stage method of strong order 1.5 (51) and the four-stage method of strong order 2 (56). These methods will be denoted by M1, M2 and M3, respectively. They will be implemented with constant step size on two problems taken from Kloeden and Platen [10] for which the exact solution in terms of a Wiener process is known.

When implementing these methods, the same sequence of random numbers for the Wiener increment J_1 is used for the step size under consideration. The random variable $J_{10} = \int\int \circ dW\, ds$ is approximated by the formula

$$J^p_{(1,0)} = \frac{1}{2}\Delta\left(\sqrt{\Delta}\xi_1 + a_{1,0}\right)$$

(see [10]) with truncation index $p = 5$, where

$$a_{1,0} = \frac{-\sqrt{2\Delta}}{\pi}\sum_{r=1}^{p}\frac{1}{r}\zeta_{1,r} - 2\sqrt{\Delta}\sqrt{\rho_p}\mu.$$

Here $\zeta_{1,r}$ and μ are Gaussian random variables, and

$$\rho_p = \frac{1}{12} - \frac{1}{2\pi^2}\sum_{r=1}^{p}\frac{1}{r^2}.$$

For both problems and all methods, 25 trajectories are computed for each step size. While in practice this sample size would be considered unrealistically small, for the purposes of this implementation 25 trajectories are sufficient to show the relative performances of the methods in question. The implementation determines the average error for each step size at the end of the interval of integration for each method, and the results appear in Tables 4, 5 and 6. The errors are presented in the form $a.b(-p)$ where p is the exponent.

Test Problem 1 (See [10, Eq. 4.4.31]).

$$dy = -a^2y(1-y^2)\,dt + a(1-y^2)\,dW, \qquad y(0) = y_0, \qquad t \in [0,1],$$

with solution

$$y(t) = \tanh\bigl(aW(t) + \operatorname{arctanh}(y_0)\bigr).$$

In Stratonovich form, the SDE becomes

$$dy = a(1-y^2)\circ dW.$$

In the results below, the parameter a was set to 1.0, and the initial value was $y_0 = 0.0$.

Test Problem 2 (See [10, Eq. 4.4.46]).

$$dy = -(\alpha + \beta^2 y)(1-y^2)\,dt + \beta(1-y^2)\,dW, \qquad y(0) = y_0, \qquad t \in [0,1],$$

Table 4
Global errors for Problem 1

h	$\frac{1}{25}$	$\frac{1}{50}$	$\frac{1}{100}$	$\frac{1}{200}$	$\frac{1}{400}$	$\frac{1}{800}$
M1	1.7(−2)	1.3(−2)	6.7(−3)	3.3(−3)	1.9(−3)	6.9(−4)
M2	9.7(−3)	6.2(−3)	3.4(−3)	2.0(−3)	1.0(−3)	2.8(−4)
M3	6.5(−3)	4.5(−3)	1.6(−3)	3.0(−4)	1.5(−4)	7.0(−5)

Table 5
Global errors for Problem 2, $\beta = 2.0$

h	$\frac{1}{25}$	$\frac{1}{50}$	$\frac{1}{100}$	$\frac{1}{200}$	$\frac{1}{400}$	$\frac{1}{800}$
M1	2.5(−1)	1.4(−1)	5.0(−2)	3.1(−2)	1.3(−2)	7.8(−3)
M2	2.1(−1)	1.1(−1)	4.1(−2)	1.9(−2)	1.0(−2)	6.5(−3)
M3	3.0(−1)	1.1(−1)	3.1(−2)	1.1(−2)	3.4(−3)	1.9(−3)

Table 6
Global errors for Problem 2, $\beta = 0.01$

h	$\frac{1}{25}$	$\frac{1}{50}$	$\frac{1}{100}$	$\frac{1}{200}$	$\frac{1}{400}$	$\frac{1}{800}$
M1	7.4(−3)	3.7(−3)	1.8(−3)	9.1(−4)	4.6(−4)	2.3(−4)
M2	1.1(−4)	2.7(−5)	7.0(−6)	1.8(−6)	4.6(−7)	1.3(−7)
M3	1.9(−6)	7.6(−7)	2.8(−7)	1.5(−7)	8.2(−8)	3.9(−8)

with solution

$$y(t) = \frac{(1+y_0)\exp(-2\alpha t + 2\beta W(t)) + y_0 - 1}{(1+y_0)\exp(-2\alpha t + 2\beta W(t)) + 1 - y_0}.$$

In Stratonovich form, the SDE has the form

$$\mathrm{d}y = -\alpha(1-y^2)\,\mathrm{d}t + \beta(1-y^2)\circ \mathrm{d}W.$$

This problem was solved numerically twice, first with parameters $\alpha = 1.0$, $\beta = 2.0$ and secondly with choices $\alpha = 1.0$, $\beta = 0.01$. This demonstrates the variation in emphasis of the stochastic and deterministic parts of the SDE. The initial value was $y_0 = 0.0$.

While further tests are needed to obtain definite conclusions, some trends from these tables can be deduced.

- For problems in which the deterministic term dominates (Problem 2 with $\beta = 0.01$) M3 is by far the best with M2 performing significantly better than Platen's method. This is because the deterministic component of M3 is the fourth order classical RK method, while the deterministic component of M2 has order 2.
- For problems with a moderately large stochastic component the global error of M2 ranges between 1.5 times and 7 times as large as that for M3, while the global error of M1 is at worst approximately

twice as large as that for M2. In all cases, the improvement becomes more noticeable as the step size is reduced.

The aim of this paper has been to show that more effective stochastic Runge–Kutta methods could be constructed than was previously the case. A strong order 2 SRK method with the fourth order classical RK method as the deterministic component was constructed, and this seems to perform substantially better than Platen's method. However, as can be seen from the numerical tests a variable step size implementation is needed. It is planned to embed the method M2 in a 4-stage method with strong order 2 (not M3) and then perform local extrapolation. It is also planned to run the number of simulations adaptively. This will be the subject of further papers. But in the meantime it is encouraging to note that the pioneering work of Runge and Kutta 100 years ago still has relevance in the relatively new area of stochastic differential equations.

References

[1] K. Burrage and E. Platen, Runge–Kutta methods for stochastic differential equations, *Ann. Numer. Math.* 1 (1–4) (1994) 63–78.

[2] P.M. Burrage, Numerical methods for stochastic differential equations, Ph.D. Thesis, Department of Mathematics, University Queensland, Australia (1997).

[3] J.C. Butcher, Coefficients for the study of Runge–Kutta integration processes, *J. Austral. Math. Soc.* 3 (1963) 185–201.

[4] J.C. Butcher, *The Numerical Analysis of Ordinary Differential Equations* (Wiley, London, 1987).

[5] C.C. Chang, Numerical solution of stochastic differential equations with constant diffusion coefficients, *Math. Comp.* 49 (1987) 523–542.

[6] T.C. Gard, *Introduction to Stochastic Differential Equations* (Marcel Dekker, New York, 1988).

[7] S. Gill, A process for the step-by-step integration of differential equations in an automatic digital computing machine, *Proc. Cambridge Philos. Soc.* 47 (1951) 96–108.

[8] A. Huťa, Une amélioration de la méthode de Runge–Kutta–Nyström pour la résolution numérique des équations différentielles du premier ordre, *Acta Math. Univ. Comenian.* 1 (1956) 201–224.

[9] P.E. Kloeden and E. Platen, Stratonovich and Ito Taylor expansions, *Math. Nachr.* 151 (1991) 33–50.

[10] P.E. Kloeden and E. Platen, *The Numerical Solution of Stochastic Differential Equations* (Springer, 1992).

[11] Y. Komori, H. Sugiura and T. Mitsui, Rooted tree analysis of the order conditions of ROW-type schemes for stochastic differential equations, Report No. 7, Mathematical Sciences, Nagoya University, Nagoya, Japan (1994).

[12] W. Kutta, Beitrag zür näherungsweisen Integration totaler Differentialgleichungen, *Z. Math. Phys.* 46 (1901) 435–453.

[13] G. Maruyama, Continuous Markov processes and stochastic equations, *Rend. Circ. Mat. Palermo* 4 (1955) 48–90.

[14] R.H. Merson, An operational method for the study of integration processes, in: *Proceedings Symposium on Data Processing*, Weapons Research Establishment, Salisbury, Australia (1957).

[15] G.N. Milstein, A method of second-order accuracy integration of stochastic differential equations, *Theor. Probab. Appl.* 23 (1978) 396–401.

[16] E. Platen, Zur zeitdiskreten Approximation von Itoprozessen, Diss. B., IMath, Akademie der Wissenschaften der DDR, Berlin (1984).

[17] E. Platen and W. Wagner, On a Taylor formula for a class of Ito processes, *Probab. Math. Statist.* 3 (1982) 37–51.

[18] W. Rümelin, Numerical treatment of stochastic differential equations, *SIAM J. Numer. Anal.* 19 (1982) 604–613.
[19] C. Runge, Über die numerische Auflösung von Differentialgleichungen, *Math. Ann.* 46 (1895) 167–178.
[20] W. Wagner and E. Platen, Approximation of Ito integral equations, Preprint ZIMM, Akademie der Wissenschaften der DDR, Berlin (1978).