BIT 32 (1992), 620–633.

A-STABILITY OF RUNGE-KUTTA METHODS FOR SYSTEMS WITH ADDITIVE NOISE

DIEGO BRICIO HERNANDEZ[1] and RENATO SPIGLER

CIMAT, Apdo. 402
36000 Guanajuato, Gto., Mexico

DMMMSA, Università di Padova
Via Belzoni 7, 35131 Padova, Italy

Abstract.

Numerical stability of both explicit and implicit Runge-Kutta methods for solving ordinary differential equations with an additive noise term is studied. The concept of numerical stability of deterministic schemes is extended to the stochastic case, and a stochastic analogue of Dahlquist's A-stability is proposed. It is shown that the discretization of the drift term alone controls the A-stability of the whole scheme. The quantitative effect of implicitness upon A-stability is also investigated, and stability regions are given for a family of implicit Runge-Kutta methods with optimal order of convergence.

1980 AMS Subject Classification: 65L20 (primary), 60H10,34F05, 65L07, 93E15.

Keywords and phrases: Numerical stability, Runge-Kutta methods, implicit methods, stochastic differential equations, stochastic stability.

1. Introduction: Systems with additive noise.

Ordinary differential equations depending on time-varying parameters, such as $\dot{x} = F(t, x, \xi(t))$, $t \geq 0$, appear frequently while modelling natural phenomena, as well as engineering and economic systems. Here $x \in \mathbb{R}^d$ is the *state vector*, whereas $\xi := \{\xi(t), t \geq 0\}$ (with $\xi(t) \in \mathbb{R}^p$) models some *input* variable, e.g. an exogenous economic variable, or a control variable that can be manipulated, or a noncontrollable input, inaccessible to the modeller. In this second case the input can be called *noise*, and it is often assumed to be a trajectory of a stochastic process ξ defined on a probability space $(\Omega, \mathcal{A}, P)$[2]. A stochastic process such as ξ will have infinitely many possible trajectories – one for each $\omega \in \Omega$, say $t \mapsto \xi(t, \omega)$ – trajectories being chosen according to the probability law P: $\mathcal{A} \to [0, 1]$.

If the noise amplitude is "small", then F above can be linearized with respect to ξ, yielding $\dot{x} = f(t, x) + g(t, x)\xi(t)$. In the physical literature [18] noise is called *additive*

[1] This author was partially supported by the Italian Consiglio Nazionale delle Ricerche.

[2] Recall that a stochastic (vector) process ξ over $(\Omega, \mathcal{A}, P)$ is a collection $\{\xi(t), t \geq 0\}$ of random vectors defined on Ω, see [1]. Dependence on $\omega \in \Omega$ is usually omitted.

Received July 1991. Revised March 1992.

if g does not depend on x; otherwise it is called *multiplicative*. Hereafter we shall confine ourselves to the case of additive noise, i.e. to systems decribed by

$$\dot{x} = f(t,x) + G(t)\xi. \tag{1.1}$$

Here $G(t)$ is a real $d \times p$ matrix for each $t \geq 0$. An important class is represented by linear equations such as

$$\dot{x} = F(t)x + G(t)\xi, \quad x(0) \text{ given} \tag{1.2}$$

(where each $F(t)$ is a $d \times d$ real matrix).

If $G(\cdot)$ is also constant (say $G(t) \equiv \sigma \in \mathbb{R}^{d \times p}$), then (1.2) becomes

$$\dot{x} = Ax + \sigma\xi, \quad x(0) \text{ given}, \tag{1.3}$$

which admits the explicit solution

$$x(t) = e^{tA}x(0) + \int_0^t e^{(t-s)A}\sigma\xi(s)\,ds, \tag{1.4}$$

provided each trajectory of ξ is at least piecewise continuous. However, in many applications ξ is standard Gaussian *white noise*, and in such case its trajectories are not even functions, but rather *distributions*. Integrals like those appearing above should be understood in the sense of Wiener [1]. If ψ: $[0, \infty) \to \mathbb{R}^{d \times p}$ is smooth and ξ is Gaussian white noise, then

$$\rho(t) := \int_0^t \psi(s)\xi(s)\,ds, \quad t \geq 0$$

is a Gaussian process with continuous trajectories, satisfying [1][3]

$$E\rho(t) = 0, \tag{1.5}$$

$$E\,\|\rho(t)\|^2 = \int_0^t \|\psi(s)\|^2\,ds. \tag{1.6}$$

Hereafter, the symbol $\|\cdot\|$ will denote the euclidean norms in $\mathbb{R}^d$, $\mathbb{R}^p$, $\mathbb{R}^{d \times p}$. In particular,

$$W(t) := \int_0^t \xi(s)\,ds$$

is a well defined stochastic process with continuous trajectories, called the standard p-dimensional *Wiener process* or *Brownian motion*. It can be characterized by the following three properties: a) $W(0) = 0$, b) each increment $W(t) - W(s)$ is a Gaussian random vector, with zero mean and covariance matrix $(t - s)I$, and c) for each choice of $n \geq 1$ and $0 \leq t_0 < t_1 < \ldots < t_n$, the increments $W(t_i) - W(t_{i-1})$, $i = 1, \ldots, n$ are independent random vectors.

---

[3] E denotes *mathematical expectation*, i.e. integration on Ω against the probability measure P.

In view of the relationship between white noise and the Wiener process (the latter is a primitive of the former), Wiener integrals are often written in the form

$$\rho(t) = \int_0^t \psi(s)\, dW(s), \quad t \geq 0.$$

Correspondingly, differential equations such as (1.2) are usually rewritten as

$$dx = F(t)x\, dt + G(t)\, dW.$$

More generally, equation (1.1) is rewritten as

$$dx = f(t, x)\, dt + G(t)\, dW, \tag{1.7}$$

which is a shorthand for the integral equation

$$x(t) = x(0) + \int_0^t f(s, x(s))\, ds + \int_0^t G(s)\, dW(s). \tag{1.8}$$

Notice that the first integral is ordinary, whereas the second one should be understood in the sense of Wiener. In what follows, the term *stochastic differential equation* (SDE) will stand for a differential equation such as (1.7).

The solutions of (1.7) are stochastic processes over $(\Omega, \mathcal{A}, P)$. A stochastic process x is a *strong* solution of (1.7) over $[t_0, T]$ if (1.8) holds for each $t \in [t_0, T]$. A unique strong solution is known to exist (see [1]) provided that: i) $x(0)$ and $\{W(t), t_0 \leq t \leq T\}$ are independent, and ii) the drift term $f(t, x)$ satisfies both a Lipschitz and a linear growth condition in x, uniformly in t.

Explicitly solvable SDE's are rather rare, and this circumstance calls for effective methods to approximate the solutions of (1.7). In particular, *numerical* methods can be used in order to construct approximations to the trajectories of the stochastic process x. For generalities, see [9].

In this paper, *general* stochastic Runge-Kutta (SRK) methods are investigated, with special emphasis on the *implicit* schemes, applied to the integration of ODE's containing *additive* white noise terms. SRK methods and implicitness are introduced in sections 2 and 3; A-stability for these Runge-Kutta schemes is defined in section 4, and examples of A-stability regions for a four-parameter implicit scheme are worked out in section 5.

2. Numerical treatment.

A numerical solution of (1.7) is obtained by specifying, for each positive integer n,

a) a partition $\tau^{(n)}$ of $[t_0, T]$, say

$$t_0 = t_0^{(n)} \leq t_1^{(n)} \leq \ldots \leq t_{k_n}^{(n)} = T,$$

where k_n increases monotonically with n, with time increments

$$\Delta t_i^{(n)} := t_{i+1}^{(n)} - t_i^{(n)}, \quad i = 0, 1, \ldots, k_n - 1,$$

and *norm*

$$\|\tau^{(n)}\| := \max_{0 \le i \le k_n} \Delta t_i^{(n)};$$

b) a finite sequence $\{x_0 =: x_0^{(n)}, x_1^{(n)}, \ldots, x_{k_n}^{(n)}\}$ of d-dimensional random vectors, where $x_i^{(n)}$ is independent of the future increments $W_{t_{i+1}} - W_{t_i}$, $i = 0, 1, \ldots, k_n - 1$ of the Brownian motion W, and $x_i^{(n)}$ "approximates" $x(t_i^{(n)})$, $i = 0, 1, \ldots, k_n$ in some sense, and

c) an interpolation procedure mapping the finite sequence of random vectors

$$\{x_0^{(n)}, x_1^{(n)}, \ldots, x_{k_n}^{(n)}\}$$

into a stochastic process $x^{(n)} := \{x^{(n)}(t), t_0 \le t \le T\}$ with continuous sample paths.

Typically, one would construct equispaced partitions, $t_{i+1}^{(n)} = t_i^{(n)} + \Delta t_i^{(n)}$, $i = 0, 1, \ldots, k_n - 1$, though varying the step size (perhaps randomly) may lead to better results at greater computational cost.

The finite sequence $\{x_0^{(n)}, x_1^{(n)}, \ldots, x_{k_n}^{(n)}\}$ is typically generated by first replacing equation (1.7) by its integral version (1.8), applied to the generic subinterval $[t_i^{(n)}, t_{i+1}^{(n)}]$,

$$x(t_{i+1}^{(n)}) - x(t_i^{(n)}) = \int_{t_i^{(n)}}^{t_{i+1}^{(n)}} f(s, x(s))\, ds + \int_{t_i^{(n)}}^{t_{i+1}^{(n)}} G(s)\, dW(s).$$

Then, numerical quadrature rules can be used to approximate the two integrals over $[t_i, t_{i+1}]$ in terms of values of $x(t)$ for $t \in \tau^{(n)}$.

A typical example of this procedure is obtained when the integration rule

$$\int_a^b \varphi(s)\, ds = \varphi(a)(b - a) + \text{error} \tag{2.1}$$

is chosen for both integrals, thus resulting in the *Euler method*

$$x_{i+1}^{(n)} = x_i^{(n)} + f(t_i^{(n)}, x_i^{(n)})\Delta t_i^{(n)} + G(t_i^{(n)})\Delta W_i^{(n)}, \tag{2.2}$$

where

$$\Delta W_i^{(n)} := W_{t_{i+1}^{(n)}} - W_{t_i^{(n)}}. \tag{2.3}$$

Every other numerical integration rule yields a different algorithm for the generation of $\{x_0^{(n)}, x_1^{(n)}, \ldots, x_{k_n}^{(n)}\}$. Convergence and stability depend, of course, on the choice of the rule. As to c), a simple *piecewise linear interpolation* $x^{(n)}(t)$ is perhaps the best choice, because the process to be approximated has *continuous* trajectories.

The *uniform error in quadratic mean*,

$$\bar{R}_n := \sup_{t_0 \le t \le T} E\, \|x^{(n)}(t) - x(t)\|^2 \tag{2.4}$$

can be used to assess the performance of such an approximation procedure. Setting

$$R_n := \max_{t \in \tau^{(n)}} E \, \|x^{(n)}(t) - x(t)\|^2, \tag{2.5}$$

it follows that

$$R_n \leq \bar{R}_n \leq 8(c \, \|\tau^{(n)}\| + R_n),$$

for some constant $c > 0$, cf. [17].

DEFINITION 2.1. A numerical algorithm for the solution of (1.7) is said to *converge uniformly in quadratic mean* if $\bar{R}_n \to 0$ (and hence $R_n \to 0$) as $n \to \infty$ for any drift term f and initial condition x_0 guaranteeing existence and uniqueness of strong solutions. The algorithm is said to have *global order of convergence*[4] ν if $R_n = O(h^{2\nu})$ as $h \to 0$, where $\|\tau^{(n)}\| \leq h$, $n \geq 1$.

As is known, the *Runge-Kutta* (RK) family results from a combined choice of numerical integration rules over each subinterval $[t_i^{(n)}, t_{i+1}^{(n)}]$ [3]. Two RK methods are required: one for the drift, the other for the noise term. Such a pair of RK methods can be denoted by a pair of "Butcher's tableaux", i.e. by a pair of block matrices like

$$\begin{array}{c|c} c & A \\ \hline 1 & b^T \end{array} \qquad \begin{array}{c|c} c & U \\ \hline 1 & v^T \end{array}, \tag{2.6}$$

where b, c, v are real m-vectors, A, U are real $m \times m$-matrices,

$$A := \|a_{jk}\|_{j,k=1}^m, \; b := (b_1, \ldots, b_m)^T,$$

respectively. These vectors and matrices are restricted by the conditions

$$b^T \mathbf{1} = 1 = v^T \mathbf{1} \tag{2.7}$$

$$A\mathbf{1} = c = U\mathbf{1} \tag{2.8}$$

where $\mathbf{1} := (1, \ldots, 1)^T$. Each of the two block matrices in (2.6) represents a *component* of the method, and corresponds to a given RK integration rule for deterministic ODE's, see [3]. The two components in (2.6) will be termed the *deterministic* and the *stochastic* part, respectively, of the SRK method. For instance, the SRK method described in (2.6) is

$$x_{i+1} = x_i + \left\{ \sum_{j=1}^m b_j f(t_i^j, x_i^j) \right\} \Delta t_i + \left\{ \sum_{j=1}^m v_j G(t_i^j) \right\} \Delta W_i$$

$$x_i^j = x_i + \left\{ \sum_{k=1}^m a_{jk} f(t_i^k, x_i^k) \right\} \Delta t_i + \left\{ \sum_{k=1}^m u_{jk} G(t_i^k) \right\} \Delta W_i$$

[4] Note that this concept of *order of convergence* differs from e.g. that in [17]: ours is Rümelin's divided by 2. Moreover, our error is *global*, while other authors consider one step iteration only.

$j = 1, 2, \ldots, m$, where x_i^j is an approximation to $x(t_i^j)$, the choice of $t_i^j \in [t_i, t_{i+1}]$ depending on the chosen RK method.

Notice that the discretization of an SDE like (1.7) involves *two* possibly different RK methods, one for the deterministic integral and another for the stochastic integral. An RK method for (deterministic) ODE's is *explicit* if its coefficient matrix A is strictly lower triangular, i.e. if $a_{jk} = 0, \quad j \leq k, \quad 1 \leq k \leq m$.

DEFINITION 2.2 The ordered pair of RK methods in (2.6), (2.7) and (2.8) will be termed a *stochastic Runge-Kutta method (SRK) with m intermediate stages.* An SRK is *explicit* if both of its components are explicit, otherwise it is *implicit.*

The convergence of some special *explicit* Runge-Kutta methods for the numerical integration of systems affected by additive white noise has been investigated in [4]. *Explicit* Runge-Kutta methods in the general stochastic context were discussed earlier by Rümelin in [17]. Special consideration was devoted to their convergence properties in the sense of Definition 2.1.

In this paper, *general* SRK methods are studied, with special emphasis on *implicit* schemes for the integration of ODE's affected by additive white noise. Below, the superindex, (n), will be omitted for simplicity.

3. Implicit methods for stochastic equations.

The choice of the *initial point rule* in (2.1) led to Euler's method (2.2), which is explicit. On the other hand, the implicit *backward Euler scheme*

$$x_{i+1} = x_i + f(t_{i+1}, x_{i+1})\Delta t_i + G(t_{i+1})\Delta W_i \tag{3.1}$$

results when the so called *final point rule*

$$\int_a^b \varphi(s)\, ds = \varphi(b)(b - a) + \text{error} \tag{3.2}$$

is applied on the subinterval $[t_i, t_{i+1}]$. Another example is provided by the *trapezoidal rule*

$$\int_a^b \varphi(s)\, ds = \frac{1}{2}(\varphi(a) + \varphi(b))(b - a) + \text{error}, \tag{3.3}$$

leading to the numerical integration scheme

$$x_{i+1} = x_i + \frac{1}{2}[f(t_i, x_i) + f(t_{i+1}, x_{i+1})]\Delta t_i + \frac{1}{2}[G(t_i) + G(t_{i+1})]\Delta W_i \tag{3.4}$$

which is implicit too.

In practice, applying one of these implicit methods, for instance (3.4), requires

a) *sampling* each ΔW_i, thereby obtaining vectors $\sqrt{(\Delta t_i)}\xi_{ir}$, $r = 1, \ldots, N$, $i = 0, 1, \ldots, k-1$, where N is the number of trajectories of x to be generated. Here, each ξ_{ir} is an independent variate drawn from a standard Gaussian p-dimensional random vector,

b) *solving* the set of kN vector algebraic equations in z,

$$z = x_i + \frac{1}{2}[f(t_i, x_i) + f(t_{i+1}, z)]\Delta t_i + \frac{1}{2}[G(t_i) \cdot G(t_{i+1})]\sqrt{(\Delta t_i)}\xi_{ir}$$

thus obtaining N samples drawn from $(x_0, x_1, \ldots, x_{k-1})$, and

c) *interpolating* each of these N sets of data, in order to obtain N samples of the desired approximate solution.

The key step of these three is clearly b), as in the deterministic case (see e.g. [2], [19] and the references therein). For instance, in the linear case (1.3) the kN algebraic systems to be solved for z are

$$\left[I - \frac{\Delta t_i}{2}A\right]z = \left[I + \frac{\Delta t_i}{2}A\right]x_i + \sqrt{(\Delta t_i)}\sigma\xi_{ir}.$$

Substantial computational economy can be achieved in this case because of the nonrandom character of the matrices $I \pm \frac{1}{2}\Delta t_i A$, $0 \leq i < k$: the LU decomposition $L_i U_i$ of each $I - \frac{1}{2}\Delta t_i A$ *is computed only once* and stored. If, in addition, the time step is constant, the computational savings will be even greater.

In general, however, applying a given implicit method requires solving a set of kN systems of *nonlinear algebraic* equations in *several* unknowns. Nevertheless, the computational effort involved may be justified whenever the better stability properties of implicit methods are essential, as suggested e.g. in [10], [14].

The convergence of *explicit* SRK methods was studied in [17]. See also [4] for the treatment of systems with additive noise. The following result is applicable to *arbitrary* (i.e. explicit as well as implicit) SRK methods:

THEOREM 3.1. *Let f in* (1.7) *have bounded partial derivatives, and suppose G is continuously differentiable. Then, the numerical solution of* (1.7) *obtained by using a general SRK method converges uniformly in quadratic mean to the solution of* (1.7).

PROOF. Let $U \doteq V$ stand for $U = V + \eta$, where $\sqrt{(E|\eta|^2)} = O(\Delta t)$. Recall that $(\Delta t)^r(\Delta W)^s \doteq 0$ for all nonnegative integers r and s with $r + s > 0$ unless either $r = 1$, $s = 0$ or $r = 0$, $s = 2$. Fix i, and let $[a, b]$ be the ith subinterval. Write t_j and x^j instead of t_i^j and x_i^j, $j = 1, \ldots, m$. Let $F := (F_1, \ldots, F_m)$, $S := (S_1, \ldots, S_m)$, with $F_j := f(t_j, x_j)$, $S_j := G(t_j)$, $j = 1, \ldots, m$.

Use x_a and x_b to denote the approximations to $x(a)$ and $x(b)$, respectively. Then $x_b - x_a = b^T F + v^T S$. Expand both $f(t_j, x^j)$ and $G(t_j)$ about (a, x_a), to obtain $F \doteq f(a, x_a)\Delta t\mathbf{1} + (\partial f/\partial x)(a, x_a)[AF + US]\Delta t$, $S \doteq G(a)\Delta W\mathbf{1}$. Taking (2.7), (2.8) into

account, we obtain $AF + US \doteq [f(a, x_a)\Delta t + G(a)\Delta W]c + (\partial f/\partial x)(a, x_a)\Delta t A \times (AF + US)$, hence

$$AF + US \doteq \left[I - \frac{\partial f}{\partial x}(a, x_a)\Delta t A\right]^{-1} [f(a, x_a)\Delta t + G(a)\Delta W]c$$

$$\doteq [f(a, x_a)\Delta t + G(a)\Delta W]\left[I + \frac{\partial f}{\partial x}(a, x_a)\Delta t A\right]c$$

$$\doteq [f(a, x_a)\Delta t + G(a)\Delta W]c.$$

Therefore, $(AF + US)\Delta t \doteq 0$ and $F \doteq f(a, x_a)\Delta t\mathbf{1}$. Finally, $x_b - x_a = f(a, x_a)\Delta t + G(a)\Delta W + \eta$.

Now consider the Euler algorithm $y_b - y_a = f(a, y_a)\Delta t + G(a)\Delta W$. By Maruyama's 1955 results [11], this algorithm converges in quadratic mean to the solution of (1.7). On the other hand, by direct subtraction, $x_b - y_b = x_a - y_a + [f(a, x_a) - f(a, y_a)]\Delta t + \eta$. Let L denote a bound for $\partial f/\partial x$. Square and take expectations to get $E(x_b - y_b)^2 \leq 3\{(1 + L^2\Delta t^2)E(x_a - y_a)^2 + E|\eta|^2\}$. An easy induction argument shows that $E(x_{kh} - y_{kh})^2 \leq 3(1 + L^2\Delta t^2)^{k-1}E|\eta|^2$, hence $E(x_{kh} - y_{kh})^2 \to 0$ as $\Delta t \to 0$ for fixed k. This proves the desired convergence. ■

The performance of an SRK method will be measured by the *order of convergence*, in the sense of Def. 2.1. Even if all SRK methods converge, they do differ in performance. In this connection, according to the well known result of [5], whenever the computation relies on the past and present Brownian increments *only* – as is the case here – the highest possible global order of convergence is 1. Therefore, an SRK method will be optimal with respect to order of convergence if and only if it has global order of convergence 1. A classical result of McShane [12] states that Heun's method

(3.5)

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

is optimal in this sense.

On the other hand, consider the four-parameter family of SRK methods

(3.6)

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 1-\gamma & \gamma \\ \hline 1 & 1-\mu & \mu \end{array} \qquad \begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 1-\beta & \beta \\ \hline 1 & 1-\lambda & \lambda \end{array}$$

for $\beta, \lambda, \gamma, \mu \in [0, 1]$. This family corresponds to the discretization scheme

(3.7)

$$x_{i+1} - x_i = [(1 - \mu)f(t_i, x_i) + \mu f(t_{i+1}, x_{i+1}^p)]\Delta t_i + [(1 - \lambda)G(t_i) + \lambda G(t_{i+1})]\Delta W_i$$

$$x_{i+1}^p - x_i = [(1-\gamma)f(t_i, x_i) + \gamma f(t_{i+1}, x_{i+1}^p)]\Delta t_i + [(1-\beta)G(t_i) + \beta G(t_{i+1})]\Delta W_i, \tag{3.8}$$

where the superscript p stands for "predictor". The classical Heun's method (3.5) is obtained when

$$\mu = \lambda = \frac{1}{2}, \qquad \beta = \gamma = 0.$$

A natural extension of McShane's proof of the optimality of (3.5) (see also Thm. 4 of [17], for general *explicit* SRK methods) yields the following:

THEOREM 3.2. *Let f be as in Theorem* 3.1. *Then the generalized Heun family* (3.7)–(3.8) *achieves optimal order of convergence if and only if* $\lambda = \frac{1}{2}$.

PROOF. Follows by the same "first order Taylor expansion" technique used in the proof of Theorem 3.1 above. It suffices to observe that the correction terms involving γ have order $E(\Delta W)^3$, and hence can be neglected. ■

In (3.6), the choice $\beta = \gamma = 0$ corresponds to *explicit* RK algorithms, whereas $|\beta| + |\gamma| > 0$ yields *implicit* RK schemes.

4. *A*-stability of SRK methods.

The stability of Runge-Kutta methods for ODE's has been extensively studied, see e.g. [7]. As to SRK methods, some results concerning Euler's method are reported in [8], see also [13], [15]. It seems desirable to extend to the stochastic case at least some of the various stability concepts applied to RK methods for deterministic ODE's.

The classical concept of *A-stability*, first introduced by Dahlquist [6], seems to be the natural starting point. It requires that all exponentially decaying modes be well approximated by the numerically computed solution, hence

$$\dot{z} = qz, \qquad \mathrm{Re}(q) < 0 \tag{4.1}$$

(with $z(t) \in \mathbb{C}$ for each $t \in \mathbb{R}$) is the natural test equation to adopt. Applied to (4.1), an RK method such as

$$\begin{array}{c|c} c & A \\ \hline 1 & b^T \end{array}, \tag{4.2}$$

leads to the difference equation

$$z_{i+1} = R(qh)z_i, \tag{4.3}$$

with the *growth factor* $R(\cdot)$ given by

$$R(z) := 1 + zb^T(I - zA)^{-1}\mathbf{1}. \tag{4.4}$$

Then, the RK method (4.2) is said to be *A-stable* if R maps the left complex half-plane into the open unit disk.

The simplest test equation that generalizes (4.1) to the stochastic case is

$$dz = qz\,dt + \sigma\,dW, \qquad \mathrm{Re}(q) < 0. \tag{4.5}$$

Such an equation is obtained by linearizing the autonomous SDE $dx = f(x)\,dt + \sigma\,dW$ (obtained by superimposing Gaussian white noise on the autonomous dynamical system $\dot{x} = f(x)$), about an equilibrium state of the *unperturbed* system, say $\bar{x}$ with $f(\bar{x}) = 0$. Again, we shall require the preservation of the decaying modes upon discretization.

In what follows, we shall assume that $q := -\alpha + i\omega$ ($\alpha > 0$), and $\sigma := \sqrt{2}(\rho + i\tau)$ are complex constants, $z := x + iy$, and $W := (U + iV)/\sqrt{2}$. Here x and y are real-valued stochastic processes, and U and V are independent real-valued Wiener processes. Using the 2×2-matrix representation of complex numbers, the single complex SDE (4.5) can be expressed in real matrix form, namely

$$d\zeta = Q\zeta\,dt + \Sigma\,dB, \tag{4.6}$$

where

$$Q := \begin{pmatrix} -\alpha & -\omega \\ \omega & -\alpha \end{pmatrix}; \quad \Sigma := \begin{pmatrix} \rho & -\tau \\ \tau & \rho \end{pmatrix}$$

and $\zeta := (x, y)^T$, $B := (U, V)^T$. Assume $\zeta(0)$ is Gaussian, independent of the Wiener process W. Then, by (1.4) ζ is a two-dimensional Gaussian process given by

$$\zeta(t) = e^{tQ}\zeta(0) + \int_0^t e^{(t-s)Q}\Sigma\,dB(s), \qquad t \geq 0.$$

Its mean $m(t) := E\zeta(t)$ and its covariance matrix

$$K(t) := E(\zeta(t) - m(t))(\zeta(t) - m(t))^T$$

are given by

$$m(t) = e^{tQ}E\zeta(0)$$

and

$$K(t) = e^{tQ}\,\mathrm{Cov}(\zeta(0))e^{tQ^T} + \int_0^t e^{(t-s)Q}\Sigma\Sigma^T e^{(t-s)Q^T}\,ds,$$

respectively, as follows from (1.5) and (1.6). Moreover, K satisfies the Lyapunov differential equation

$$\dot{K} = QK + KQ^T + \Sigma\Sigma^T, \quad K(0) = \mathrm{Cov}(\zeta(0)) \tag{4.7}$$

as can be verified directly. Observe that, because $\alpha > 0$, (4.7) has a unique equilibrium solution $\bar{K}$, which satisfies the Lyapunov algebraic equation

$$Q\bar{K} + \bar{K}Q^T + \Sigma\Sigma^T = 0. \tag{4.8}$$

An easy calculation shows that the steady state covariance matrix associated with equation (4.6) is $\bar{K} = (|\sigma|^2/2\alpha)I$, which represents a *noise-to-damping ratio*. Similarly, the stability of Q (i.e. the fact that $\alpha > 0$) implies that $m(t) \to 0$ and $K(t) \to \bar{K}$ as $t \to \infty$. Thus, the test equation (4.5) has a unique equilibrium solution, which is a stationary complex Gaussian random process z with zero mean and variance $|\sigma|^2/2\alpha$. Moreover, any other solution decays exponentially to this equilibrium solution.

Applying the SRK method (2.6)–(2.8) to (4.5), an easy calculation leads to the difference equation

$$z_{i+1} = R(qh)z_i + S(qh)\sigma\Delta W_i, \tag{4.9}$$

where R is given by (4.4) and

$$S(z) := 1 + zb^T(I - zA)^{-1}c. \tag{4.10}$$

The solution to (4.9) with the initial condition $z(0)$ is given by

$$z_i = R(qh)^i z(0) + \sigma S(qh) \sum_{j=0}^{i-1} R(qh)^{i-j-1}\Delta W_j. \tag{4.11}$$

A further elementary calculation finally shows that

$$Ez_i = R(qh)^i Ez(0), \tag{4.12}$$

$$\mathrm{Cov}(z_i) = R(qh)^{2i}\,\mathrm{Cov}(z(0)) + \frac{S(qh)^2\sigma^2}{1 - R(qh)^2}(1 - R(qh)^{2i}), \tag{4.13}$$

and a limiting distribution of the solution to the difference equation (4.9) exists if and only if $|R(qh)| < 1$. Such a limiting distribution is Gaussian, with zero mean and variance $S(qh)^2\sigma^2/(1 - R(qh)^2)$.

DEFINITION 4.1. The SRK algorithm (2.6)–(2.8) is said to be *A-stable* if R (defined in (4.4)) maps every complex z with $\mathrm{Re}(z) < 0$ into the open unit disk.

In terms of this concept, the foregoing developments can be expressed as follows:

THEOREM 4.1. *The SRK method* (2.6)–(2.8) *is A-stable if and only if so is its deterministic component.*

REMARK. Notice that the discretization of the drift term alone controls the numerical stability. This is so, simply because the diffusion term in the test equation (4.5) *does not* depend on z.

It is a classical result that *no explicit RK method is A-stable* [6], although suitable

stability regions can be given for them. Similar regions can also be given for SRK methods, see section 5 below.

5. The effect of implicitness upon A-stability.

By Theorem 4.1, neither λ nor β (i.e. the parameters in the second component of (3.6)), will affect the A-stability of such an SRK method. They influence only convergence, cf. Theorem 3.2. The parameter γ measures the *degree of implicitness* of (3.6). In fact, $\gamma = 0$ corresponds to explicit methods, and the larger γ the greater the weight attached to the (still unknown) value of x_{i+1} when computing it from x_i using (3.6). Therefore, γ (and also μ) are the A-stability determining parameters.

To find out when and how this happens, consider the test equation (4.5) and apply (3.6) to it, thereby obtaining (4.9), with

$$R(z) = \frac{1 + (1-\gamma)z + (\mu - \gamma)z^2}{1 - \gamma z}. \tag{5.1}$$

Clearly, $|R(z)| \to \infty$ as $|z| \to \infty$ if either $\mu \neq \gamma$ or $\gamma = 0$ (for any μ), hence A-stability cannot hold unless $\mu = \gamma$, $\gamma \neq 0$. Under such conditions, the growth factor (5.1) simplifies to

$$R(z) = \frac{1 + (1-\gamma)z}{1 - \gamma z}, \tag{5.2}$$

which can be viewed as a complex linear fractional transformation $R: z \mapsto w$, which transforms the imaginary axis $x = 0$ in the z-plane into a curve Γ in the w-plane. Under this transformation, the left complex halfplane Σ transforms into the interior of the region bounded by Γ, when the imaginary axis is traversed in the direction of increasing y.

An elementary analysis shows that *for every* $\gamma > 0$, Γ *is a circle, with center at* $(1 - 1/2\gamma, 0)$ *and radius* $(2\gamma)^{-1}$. *Therefore,* Σ *transforms, under* R, *into the open disk bounded by* Γ. As $\gamma \downarrow 0$, $R(\Sigma)$ *becomes the halfplane* $\mathrm{Re}(w) < 1$. *On the other hand, as* $\gamma \uparrow 1$, Γ *tends to the circle of radius* $\frac{1}{2}$ *and center at* $(\frac{1}{2}, 0)$. *Notice that* $\gamma = \frac{1}{2}$ *is the smallest value of* γ *for which the SRK method* (3.6) *is A-stable.*

In fact, it is easy to see that

$$\max_{-\infty < y < +\infty} |R(iy)| = \begin{cases} \dfrac{1}{\gamma} - 1, & \text{if } \gamma < \frac{1}{2} \\ 1, & \text{if } \gamma \geq \frac{1}{2}, \end{cases}$$

which shows that the whole of Σ is mapped into the interior of the unit disk under R if and only if $\gamma \geq \frac{1}{2}$. We say that *A-stability improves* as the area of $R(\Sigma)$ decreases. Then, we can summarize with the following

THEOREM 5.1. *The SRK method* (3.6) *is A-stable if and only if* $\gamma = \mu \geq \frac{1}{2}$. *In such a case the A-stability of* (3.6) *improves as* $\gamma \uparrow 1$.

When $\gamma = \mu < \frac{1}{2}$ the SRK methods (3.6) are certainly *not* A-stable, but they still can be used for the numerical integration of SDE's provided the integration step is sufficiently small, as $|R(qh)|$ can still be less than 1 in this case. Naming any subset S of the left complex halfplane $\mathrm{Re}(z) < 0$ a *stability region* for (3.6) if $|R(z)| < 1$ for every $z \in S$, A-stable methods are those for which the whole complex halfplane is a stability region and stepsize is limited only by convergence requirements. Yet, the stability region of a method which is not A-stable may be large enough to allow for reasonably large integration steps. When the stability region of an integration method is a proper subset of the whole complex halfplane, the method is said to be *conditionally stable.*

Stability regions having the form of an open disk S_a with radius a and center at $(-a, 0)$, for $a > 0$, can be readily found for SRK methods like (3.6). All we must guarantee is that the points of ∂S_a are mapped into the closed unit disk (and thus S_a is mapped into the open unit disk) for such a. An elementary calculation shows that *if* $\gamma \geq 1/2$, *then* S_a *is a stability region for any* $a > 0$. On the contrary, *if* $0 \leq \gamma < 1/2$, *then* S_a *is a stability region only for* $a \leq (1 - 2\gamma)^{-1}$. In other words,

PROPOSITION 5.1. *A stability region for the SRK method* (3.6) *with* $0 \leq \gamma < 1/2$ *is given by any open disk* S_a *with* $a < (1 - 2\gamma)^{-1}$. *On the other hand, the method is unconditionally stable if* $\gamma \geq 1/2$.

In practice, the integration stepsize will be limited by the stability requirement if $\gamma < \frac{1}{2}$: it must be chosen in such a way that $hq \in S$, where S is any stability region and q is an estimate of the spectrum of the local Jacobian $\partial f/\partial x$. However, there are other factors limiting the integration step even when $\gamma \geq \frac{1}{2}$. Besides the convergence requirements, it is necessary to ensure the well-posedness of the method: *the algebraic equations to be solved at every integration step must have a unique solution,* at least with a high probability. Using, e.g., the SRK method (3.6) to solve (1.7), at the ith stage it is required to solve the algebraic system

$$z = F_i(z), \tag{5.3}$$

for $z := x_{i+1}^p$. Here $F_i(z) := a_i + \gamma f(t_{i+1}, z)\Delta t_i$, and

$$a_i := x_i + (1 - \gamma) f(t_i, x_i)\Delta t_i + [(1 - \beta)G(t_i) + \beta G(t_{i+1})]\Delta W_i.$$

If f satisfies a Lipschitz condition with constant L, then

$$\|F_i(z_1) - F_i(z_2)\| \leq \gamma L \Delta t_i \, \|z_1 - z_2\|$$

and therefore F_i is a contraction provided that

$$\gamma L \Delta t_i < 1. \tag{5.4}$$

Consequently, equation (5.3) has a unique solution $z := x_{i+1}^p$, for each $i \geq 0$. This shows that the method is well posed under condition (5.4). Note that each a_i is random, but this randomness does not affect the solvability of (5.3). Of course, this stems from the fact that the SDE being solved is affected by *additive* rather than *multiplicative* noise.

REFERENCES

1. L. Arnold, *Stochastic Differential Equations*, Wiley, New York, 1974.
2. J. C. Butcher, *On the implementation of implicit Runge-Kutta methods*, BIT, 16, 237–240 (1976).
3. J. C. Butcher, *The Numerical Analysis of Ordinary Differential Equations*, Wiley, Chichester, 1987.
4. C. C. Chang, *Numerical solution of stochastic differential equations with constant diffusion coefficients*, Math. Comp. 49, 523–542 (1987).
5. J. M. C. Clark, and R. J. Cameron, *The maximum rate of convergence of discrete approximations for stochastic differential equations*, in B. Grigelionis (ed.), *Stochastic differential systems*, Lecture Notes in Control and Information Systems, 25, Springer Verlag, Berlin, 1980.
6. G. Dahlquist, *A special stability problem for linear multistep methods*, BIT, 3, 27–43 (1963).
7. K. Dekker and J. G. Verwer, *Stability of Runge-Kutta Methods for Stiff Nonlinear Differential Equations*, North Holland, Amsterdam, 1984.
8. R. Janssen, *Diskretisierung Stochasticher Differentialgleichungen*, Preprint Nr. 51, FB Mathematik, Universität Kaiserslautern, 1982.
9. P. E. Kloeden and E. Platen, *The Numerical Solution of Stochastic Differential Equations*, Springer Verlag, Berlin, 1991.
10. H. Liske and E. Platen, *Simulation studies on time discrete diffusion approximations*, Math. Comp. Simulation, 29, 253–260 (1987).
11. G. Maruyama, *Continuous Markov processes and stochastic equations*, Rend. Circ. Mat. Palermo, 4, 48–90 (1955).
12. E. J. McShane, *Stochastic Calculus and Stochastic Models*, Academic Press, New York, 1974.
13. G. N. Mil'shtein, *The Numerical Integration of Stochastic Differential Equations* (in Russian), Urals University Press, Sverdlovsk, 1988.
14. E. Pardoux and D. Talay, *Discretization and simulation of stochastic differential equations*, Acta Appl. Math., 3, 23–47 (1985).
15. W. P. Petersen, *Stability and accuracy of simulations for stochastic differential equations*, IPS Research Report No. 90-02, ETH-Zentrum, Zürich, January 1990.
16. E. Platen, *An approximation method for a class of Ito equations*, Litovsk. Matem. Sb. 21 (1981), 121–133.
17. W. Rümelin, *Numerical treatment of stochastic differential equations*, SIAM J. Numer. Anal, 19, 604–613 (1982).
18. J. M. Sancho, M. San Miguel, S. L. Katz and J. D. Gunton, *Analytical and numerical studies of computational noise*, Phys. Rev. A 26, 1589–1609 (1982).
19. J. M. Varah, *On the efficient implementation of implicit Runge-Kutta schemes*, Math. Comp., 33, 557–561 (1979).