ELSEVIER
Applied Numerical Mathematics 24 (1997) 425–438
APPLIED NUMERICAL MATHEMATICS

Dynamics of constant and variable stepsize methods for a nonlinear population model with delay

Tasneem K. Sardar [1], Desmond J. Higham [*,2]
Department of Mathematics and Computer Science, The University of Dundee, Dundee DD1 4HN, Scotland, UK

Abstract

Hutchinson's equation is a reaction-diffusion model where the quadratic reaction term involves a delay. It is a natural extension of the logistic equation (no diffusion, no delay) and Fisher's equation (no delay), both of which have been used to illustrate the potential for spurious long-term dynamics in numerical methods. For the case where initial conditions and periodic boundary conditions are supplied, we look at the use of central differences in space and either Euler's method or the Improved Euler method in time. Our aim is to investigate the impact of the delay on the long-term behaviour of the scheme. After studying the fixed points of the methods in constant stepsize mode, we consider an adaptive time-stepping approach, using a standard local error control strategy. Applying ideas of Hall (1985) we are able to explain the fine detail of the time-step selection process.

1. Introduction

Long-term numerical simulations are often performed on nonlinear evolutionary problems, even when there is little theoretical justification of the "correctness" of the answers. To redress the balance, numerical analysts are currently building up a theory of "numerical dynamics" by studying classes of nonlinear problems and/or certain types of long-term behaviour [14]. The simplest object to study is the fixed point, or constant steady-state. In the case of constant stepsize implementations of ordinary differential equation methods, fundamental results about fixed points were obtained by Iserles [10] and Humphries [9]. An important tool for analyzing variable stepsize methods was developed by Hall [4].

Other authors have gone on to investigate the issue of fixed points under discretization for various types of evolutionary problem; see, for example, [1,3,7,11]. In this work we investigate methods for

* Corresponding author. Current address: Department of Mathematics, University of Strathclyde, Glasgow, G1 1XH. E-mail: na.dhigham@na-net.ornl.gov.

[1] The work of this author is supported by an Earmarked Research Studentship from the UK Engineering and Physical Sciences Research Council.

[2] The work of this author is supported by the UK Engineering and Physical Sciences Research Council under grants GR/H94634 and GR/K80228.

PII S0168-9274(97)00037-8

a model arising in population dynamics. The model can be regarded as a generalization of a number of test-problems that have been used in the study of numerical dynamics. The material presented here is unusual in that it looks at both fixed and variable stepsize algorithms for a problem that has a diffusion term and a delayed reaction term. Our interests lie (a) in the impact of the delay term on the behaviour of the algorithm, and (b) in the effect of the error control process. Further results, including an analysis of spurious solutions, appear in [13]. Reference [13] also contains technical details and illustrative figures that have been omitted from this paper in the interest of brevity.

This work forms a natural extension to the earlier paper [8]. In [8], the same equation, subject to homogeneous Dirichlet boundary conditions, was studied, and a constant stepsize discretization was used. Emphasis was placed on the existence and stability of spurious solutions for stepsizes inside or outside the interval of linearized stability. In this work, we look at the case of periodic boundary conditions, and concentrate on the stability of the true fixed point. We study two simple time-stepping methods—first we look at the individual formulas in constant stepsize mode, and then we study adaptive, variable stepsize methods based on local error control. We use the ideas of Hall [4] to investigate the long-term dynamics of the adaptive method. (This approach was used in [5,6] for scalar delay differential equations.)

The material is organized as follows. In Section 2 we introduce the differential equation and the discretizations. We also look at the stability of the true fixed points of the semi-discrete system. Sections 3 and 4 are concerned with fixed point stability when Euler's method and the Improved Euler method are used to advance the solution in time, respectively. In Section 5 we turn our attention to adaptive algorithms in which the two time-stepping formulas are combined to give a local error estimate. We identify equilibrium states that are seen to be relevant to the behaviour of the adaptive algorithms in practice. The theoretical analysis, and the numerical tests, suggest that the error-controlled algorithm chooses stepsizes that correspond to the largest stepsize that would be acceptable in constant-stepsize mode. The relevant equilibrium state can be computed exactly if it is stable with respect to small perturbations. We show by example that stability in this sense is a function of the problem parameters as well as the numerical method.

2. Discretizations of Hutchinson's equation

Consider Hutchinson's equation

$$\begin{cases} \dfrac{\partial}{\partial t}u(x,t) = \varepsilon\dfrac{\partial^2}{\partial x^2}u(x,t) + u(x,t)\big[1 - u(x,t-\tau)\big], & t > 0, \quad 0 \leqslant x \leqslant 1; \\ u(x,t) = \Psi(x,t), & t \in [-\tau, 0], \end{cases} \tag{2.1}$$

subject to periodic boundary conditions, where $\varepsilon > 0$ is the diffusion coefficient, $\tau > 0$ is the amount of delay and the "initial value function" $\Psi(x,t)$ is continuous.

The scalar partial delay differential equation (2.1) is initially transformed into a system of ordinary delay differential equations by discretizing the space variable x into $(N+1)$ discrete values $(N \geqslant 1)$, with a constant stepsize in space, $\Delta x = 1/N$, so that $x_j = j\Delta x$, $j = 0, 1, \ldots, N$.

Let $\Phi_j(t) = \Psi(x_j, t)$ and $U_j(t)$ denote the approximation to $u(x_j, t)$. Using the standard central difference operator to approximate the Laplacian we obtain the system

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}U(t) = \dfrac{\varepsilon}{(\Delta x)^2}MU(t) + U(t)\circ\big[e - U(t-\tau)\big], & t > 0; \\ U(t) = \Phi(t), & t \in [-\tau, 0], \end{cases} \tag{2.2}$$

where $U(t-\tau)$, $U(t)$, $\Phi(t) \in \mathbb{R}^N$; $e = [1, 1, \ldots, 1]^{\mathrm{T}} \in \mathbb{R}^N$ and the matrix $M \in \mathbb{R}^{N\times N}$ is of the form

$$M = \begin{cases} [0], & \text{if } N = 1, \\ \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix}, & \text{if } N = 2, \\ \begin{bmatrix} -2 & 1 & & & 1 \\ 1 & -2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ 1 & & & 1 & -2 \end{bmatrix}, & \text{if } N \geqslant 3. \end{cases}$$

The symbol "$\circ$" denotes componentwise multiplication, so that $(U \circ W)_j = U_j W_j$.

2.1. *Constant solutions of partially discretized system and linear stability*

The system (2.2) has two constant solutions that are period one in space; that is, of the form $U(t) \equiv U^* e \in \mathbb{R}^N$. These are $U(t) \equiv 0$ and $U(t) \equiv e$. After linearization, the zero fixed point can be shown to be unstable for all $\tau > 0$. For the nonzero fixed point, extensive manipulation leads to the linear stability condition $0 < \tau < \pi/2$.

2.2. *Full discretizations of Hutchinson's equation*

Consider the partially discretized system (2.2). Let

$$f\big(t, U(t), U(t-\tau)\big) := \big(\varepsilon N^2\big)MU(t) + U(t)\circ\big[e - U(t-\tau)\big] \in \mathbb{R}^N.$$

To solve the system $U'(t) = f(t, U(t), U(t-\tau))$, either Euler's method or the Improved Euler method is applied to advance the solution in time and a linear Lagrange interpolant, $q(t-\tau)$, is used to approximate the delayed values, $U(t-\tau)$.

Setting $t_n = n\Delta t$, where Δt is a positive constant stepsize in time, let U^n denote the approximation to $U(t_n)$. If $t_n - \tau \leqslant 0$ then we can take $q(t_n - \tau) = \Phi(t_n - \tau)$. Otherwise, let m be the smallest integer such that $m\Delta t \geqslant \tau$, so that $t_n - \tau$ lies in the interval $[t_{n-m}, t_{n-m+1})$, that is, $(m-1)\Delta t < \tau \leqslant m\Delta t$. Let $\sigma\Delta t = (m\Delta t - \tau)$, so that

$$\sigma = m - \frac{\tau}{\Delta t}, \quad \sigma \in [0, 1).$$

The linear Lagrange interpolant based on the values U^{n-m+1} and U^{n-m} is then given by $q(t_n - \tau) = \sigma U^{n-m+1} + (1-\sigma)U^{n-m}$. Therefore, $k_1 := f(t_n, U^n, q(t_n - \tau))$ results in

$$k_1 = \frac{1}{\Delta t}[B_r - I]U^n - U^n \circ \big[\sigma U^{n-m+1} + (1-\sigma)U^{n-m}\big]$$

when $t_n - \tau > 0$, where $B_r := [(\Delta t + 1)I + rM] \in \mathbb{R}^{N\times N}$ and $r := \varepsilon\Delta t/(\Delta x)^2 = \varepsilon N^2\Delta t$.

If $t_n + \Delta t - \tau \leqslant 0$ then we can take $q(t_n + \Delta t - \tau) = \Phi(t_n + \Delta t - \tau)$. Otherwise $t_n + \Delta t - \tau$ lies in the interval $[t_{n-m+1}, t_{n-m+2})$ since $(t_n - \tau) \in [t_{n-m}, t_{n-m+1})$. The linear Lagrange interpolant based on the values U^{n-m+2} and U^{n-m+1} is given by

$$q(t_n + \Delta t - \tau) = \sigma U^{n-m+2} + (1-\sigma)U^{n-m+1}.$$

Therefore,

$$k_2 := f\big(t_n + \Delta t, U^n + \Delta t k_1, q(t_n + \Delta t - \tau)\big)$$

results in

$$\begin{aligned} k_2 = {} & \frac{1}{\Delta t}[B_r - I]B_r U^n - [B_r - I]\big(U^n \circ \big[\sigma U^{n-m+1} + (1-\sigma)U^{n-m}\big]\big) \\ & - \big[B_r U^n\big] \circ \big[\sigma \widehat{U}^{n-m+2} + (1-\sigma)U^{n-m+1}\big] \\ & + \Delta t U^n \circ \big[\sigma U^{n-m+1} + (1-\sigma)U^{n-m}\big] \circ \big[\sigma \widehat{U}^{n-m+2} + (1-\sigma)U^{n-m+1}\big] \end{aligned}$$

when $t_n + \Delta t - \tau > 0$. Note that the quantity $\widehat{U}^{n-m+2}$ appears in the expression for k_2. This is done because when $m = 1$ the linear interpolant involves the *unknown* value U^{n+1}. In this case, to keep the algorithm simple, we use the approximation $U^{n+1} \approx U^n + \Delta t k_1$. Hence, we define $\widehat{U}^{n-m+2} = U^{n-m+2}$ for $m > 1$ and $\widehat{U}^{n-m+2} = U^n + \Delta t k_1$ for $m = 1$.

Let

$$v^n := \begin{bmatrix} U^n \\ U^{n-1} \\ \vdots \\ U^{n-m+1} \\ U^{n-m} \end{bmatrix} \in \mathbb{R}^{N(m+1)}.$$

It follows that, using linear Lagrange interpolation to approximate the delayed values:

(E) Euler's method $U^{n+1} = U^n + \Delta t k_1$ applied to (2.2) on a general step with $n > \tau/\Delta t$ gives the nonlinear recurrence $U^{n+1} = g(v^n) \in \mathbb{R}^N$, where

$$g(v^n) := B_r U^n - U^n \circ \big[(m\Delta t - \tau)U^{n-m+1} - (m\Delta t - \tau - \Delta t)U^{n-m}\big].$$

(IE) The Improved Euler method $U^{n+1} = U^n + \frac{1}{2}\Delta t(k_1 + k_2)$ applied to (2.2) on a general step with $n > \tau/\Delta t$, gives the nonlinear recurrence $U^{n+1} = g(v^n) \in \mathbb{R}^N$, where

$$\begin{aligned} g(v^n) := {} & \tfrac{1}{2}U^n + \tfrac{1}{2}B_r^2 U^n - \tfrac{1}{2}B_r\big(U^n \circ \big[(m\Delta t - \tau)U^{n-m+1} - (m\Delta t - \tau - \Delta t)U^{n-m}\big]\big) \\ & - \tfrac{1}{2}\big[B_r U^n\big] \circ \big[(m\Delta t - \tau)\widehat{U}^{n-m+2} - (m\Delta t - \tau - \Delta t)U^{n-m+1}\big] \\ & + \tfrac{1}{2}U^n \circ \big[(m\Delta t - \tau)U^{n-m+1} - (m\Delta t - \tau - \Delta t)U^{n-m}\big] \\ & \circ \big[(m\Delta t - \tau)\widehat{U}^{n-m+2} - (m\Delta t - \tau - \Delta t)U^{n-m+1}\big]. \end{aligned}$$

3. Euler's method: constant stepsize

It is easily seen that Euler's method combined with linear Lagrange interpolation is *regular*, that is, it does not admit spurious constant solutions. The period 1 in space and time fixed points of the iteration are therefore $U^n \equiv 0$ and $U^n \equiv e$. It is simple to show that the rest state $U^n \equiv 0 \in \mathbb{R}^N$ (for all m) is linearly unstable for all $\tau, \varepsilon, \Delta t > 0$.

For $m = 1$, the linear stability region for the fixed point $U^n \equiv e \in \mathbb{R}^N$ can be shown to have the form $0 < \tau < 1$ and $\tau \leqslant \Delta t < 2(1+\tau) - 4\widehat{r}$, where

$$\widehat{r} := \begin{cases} r[1-\sin^2(\pi/2N)], & \text{if } N \text{ is odd,} \\ r, & \text{if } N \text{ is even.} \end{cases} \tag{3.1}$$

Note that the linear stability regions corresponding to the cases odd $N > 1$ and even N become identical as $N \to \infty$. For even N, in the limit as $\tau \to 0$ for $m = 1$, the linear stability interval of $U^n \equiv e$ tends to $0 < \Delta t < 2 - 4r$ (which requires $0 < r < 1/2$). This agrees with the stability of the nonzero solution for the zero delay case, as established for even N in [2]. For odd N, in the limit as $\tau \to 0$ for $m = 1$, the linear stability interval of $U^n \equiv e$ tends to

$$0 < \Delta t < 2 - 4r\left[1 - \sin^2\left(\frac{\pi}{2N}\right)\right]. \tag{3.2}$$

It can be shown that (3.2) is the linear stability of $U^n \equiv e$, corresponding to the zero delay case, for odd N.

For $m > 1$, analysis shows that the nonzero fixed point is stable whenever the polynomial $\prod_{j=1}^{N} p_j^m(\lambda)$ is Schur, where

$$p_j^m(\lambda) := \lambda^{m+1} - \lambda^m\left(1 - 4r\sin^2\left(\frac{j\pi}{N}\right)\right) + \lambda(m\Delta t - \tau) - (m\Delta t - \tau - \Delta t).$$

Note that the factor $p_N^m(\lambda)$ in the characteristic polynomial is independent of r and N. Further, when $r \equiv 0$, $p_j^m(\lambda) = p_N^m(\lambda)$ for all $j = 1, 2, \ldots, N-1$. It follows that the region in the $(\Delta t, \tau)$-plane defined by $\{(\Delta t, \tau)\colon\ p_N^m(\lambda) \text{ is Schur}\}$ represents:

- linear stability of $U^n \equiv e \in \mathbb{R}^N$ in the scalar case $N = 1$,
- linear stability of $U^n \equiv e \in \mathbb{R}^N$ for all N when $r \equiv 0$.

Also note that for a fixed positive r, the linear stability region of the nonzero constant solution $U^n \equiv e \in \mathbb{R}^N$ must be a subset of the linear stability region for $N = 1$ (because the product $\prod_{j=1}^{N} p_j^m(\lambda)$ can never be Schur for any $(\Delta t, \tau)$-value for which the factor $p_N^m(\lambda)$ is not Schur).

The linear stability region of $U^n \equiv e$ for large m can be determined numerically. Stability regions are shown in Figs. 1 (using "•") and 2 (using "$*$") for various values of $\tau < 1$ and $\tau > 1$. Note that, as predicted in the theoretical analysis of case $m = 1$, when $\tau < 1$ the fixed point is linearly stable in a part of the region where $\Delta t \geqslant \tau$ but as soon as $1 < \tau$ $(< \pi/2)$ it becomes unstable for any $\Delta t \geqslant \tau$. This abrupt change in the stability of the fixed point as the delay passes over one can be seen when the delay from $\tau = 0.99$ in the last diagram of Fig. 1 is slightly increased to $\tau = 1.01$ in the first diagram of Fig. 2. Also note that in Fig. 1, a small increase in τ improves stability in the sense that $U^n \equiv e$ is stable for a larger set of values $(\Delta t, \widehat{r})$.

Based on experience with simple non-delayed parabolic partial differential equations, it is reasonable to refine the space-time discretization while keeping $\Delta t/(\Delta x)^2$ fixed [12]. To analyze the stability

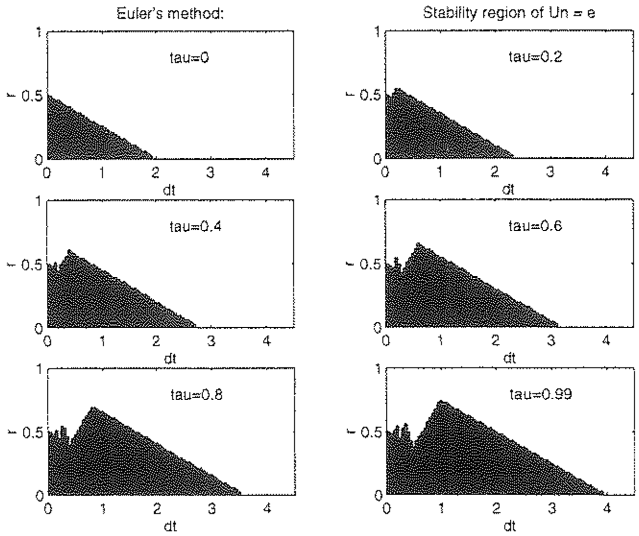

Fig. 1. Stability region of $U^n \equiv e \in \mathbb{R}^N$, for various values of $\tau < 1$, in the $(\Delta t, \widehat{r})$-plane.

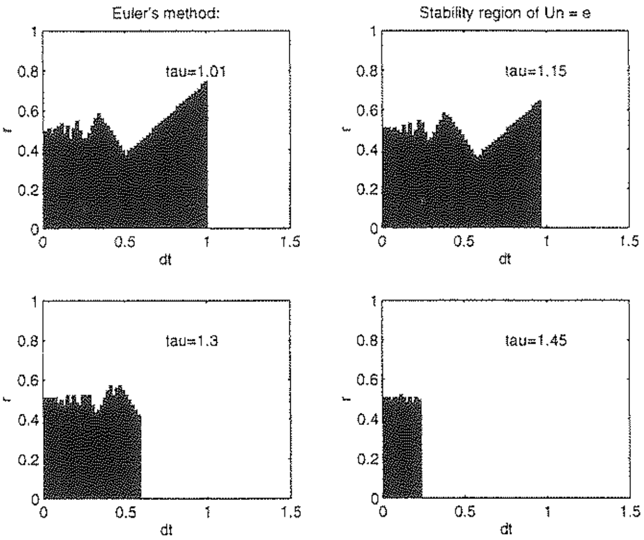

Fig. 2. Stability region of $U^n \equiv e \in \mathbb{R}^N$, for various values of $\tau > 1$, in the $(\Delta t, \widehat{r})$-plane.

properties when Δt, Δx tend to zero with $\Delta t/(\Delta x)^2$ $(= r/\varepsilon)$ constant, consider the characteristic equation $\prod_{j=1}^{N} p_j^m(\lambda) = 0$, with the polynomial $p_j^m(\lambda)$ re-written in the form

$$p_j^m(\lambda) = \lambda^m \left[\lambda - \left(1 - 4r\sin^2\left(\frac{j\pi}{N}\right)\right)\right] + \frac{\tau}{(m-\sigma)}(\sigma\lambda - \sigma + 1).$$

If Δt and Δx decrease to zero (so that m and N are large) and such that r/ε remains constant, then $\tau(\sigma\lambda - \sigma + 1)/(m - \sigma)$ is negligible, so the roots of the polynomial $p_j^m(\lambda) = 0$ with $j \in [1, 2, \ldots, N-1]$, will be close to 0 and $1-4r\sin^2(x_j\pi)$ with $x_j \in (0,1)$. The nonzero root close to $1-4r\sin^2(x_j\pi)$ is less than one in modulus when $0<r<1/(2\sin^2(x_j\pi))$. These roots are all less than one in modulus when

$$0<r<\frac{1}{2}\left\{\min_{1\leqslant j<N}\left(\frac{1}{\sin^2(x_j\pi)}\right)\right\},$$

which reduces to $0<\widehat{r}<1/2$.

The above analysis cannot be applied to the polynomial $p_N^m(\lambda)=0$, which is independent of r and Δx, since $\sin^2(j\pi/N)$ is zero when $j=N$. If Δt is small (so that m is large) then the roots of the polynomial $p_N^m(\lambda)=0$ will be close to zero and one. To proceed with the analysis, assume that the root close to one has the expansion

$$\lambda = 1+\tau\gamma/m+\mathrm{O}(1/m^2)$$

for some complex constant γ. Substituting this expansion into $p_N^m(\lambda)=0$ and noting that $1/(m-\sigma)=1/m+\mathrm{O}(1/m^2)$ leads to

$$\left[1+\frac{\tau\gamma}{m}\right]^m\tau\gamma = -\tau+\mathrm{O}\left(\frac{1}{m}\right).$$

Letting $m\to\infty$ and multiplying by $\mathrm{e}^{-\tau\gamma}$ results in the relation $\tau\gamma+\tau\mathrm{e}^{-\tau\gamma}=0$, so the polynomial $p_N^m(\lambda)$ is Schur for sufficiently large m (that is γ has negative real parts) when $0<\tau<\pi/2$.

Thus, if the expansion above is valid, the nonzero fixed point $U^n\equiv e\in\mathbb{R}^N$ is linearly stable—that is, the characteristic polynomial $\prod_{j=1}^N p_j^m(\lambda)$ is Schur—for sufficiently large m and N (so that Δt and Δx are small) and such that $\Delta t/(\Delta x)^2 (=r/\varepsilon)$ remains constant when $0<\tau<\pi/2$ and $0<\widehat{r}<1/2$. Numerical tests were in agreement with this analysis.

4. The Improved Euler method: constant stepsize

The Improved Euler method inherits the fixed points $U^n\equiv 0\in\mathbb{R}^N$ and $U^n\equiv e\in\mathbb{R}^N$.

It can be deduced easily that the fixed point $U^n\equiv 0\in\mathbb{R}^N$ (for all m) is linearly unstable for all $\tau,\varepsilon,\Delta t>0$.

Results of a lengthy linear stability analysis of $U^n\equiv e\in\mathbb{R}^N$ when $m=1$ are given in Table 1. When $m\geqslant 2$, linear stability of e is determined by the roots of the polynomial $\prod_{j=1}^N p_j^m(\lambda)$, where

Table 1
Linear stability of $U^n\equiv e\in\mathbb{R}^N$ for the Improved Euler method when $m=1$

N	Delay	Stepsizes for which $U^n\equiv e\in\mathbb{R}^N$ is linearly stable when $m=1$
$N=1$	$\tau<2$	$\tau\leqslant\Delta t<\tau+2$
$N=1$	$\tau\geqslant 2$	$\frac{1}{\tau}(\tau-1)(\tau+2)<\Delta t<\frac{1}{2}(3\tau+2-\sqrt{(\tau-2)(\tau+6)})$
even N	$\tau<2$	$\tau\leqslant\Delta t<\tau+2-4r$
even N	$\tau\geqslant 2$	$\frac{1}{\tau}(\tau-1)(\tau+2)<\Delta t<\frac{1}{2}(3\tau+2-\sqrt{(\tau-2)(\tau+6)})-4r$
odd N	$\tau<2$	$\tau\leqslant\Delta t<\tau+2-4r[1-\sin^2(\frac{1}{2}\pi/N)]$
odd N	$\tau\geqslant 2$	$\frac{1}{\tau}(\tau-1)(\tau+2)<\Delta t<\frac{1}{2}(3\tau+2-\sqrt{(\tau-2)(\tau+6)})-4r[1-\sin^2(\frac{1}{2}\pi/N)]$

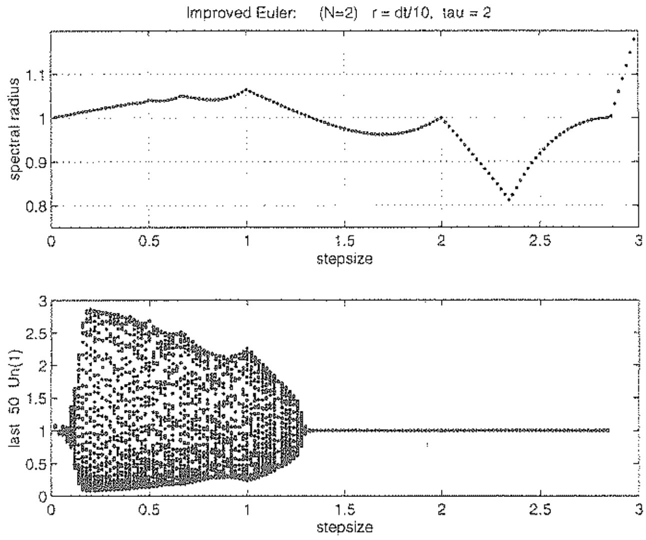

Fig. 3. Behaviour of the Improved Euler method when $\tau = 2$ and $\widehat{r} = \Delta t/10$.

$$p_j^m(\lambda) := \lambda^{m+1} - \frac{1}{2}\left(1 + \left[1 - 4r\sin^2\left(\frac{j\pi}{N}\right)\right]^2\right)\lambda^m + \frac{1}{2}(m\Delta t - \tau)\lambda^2 + \frac{1}{2}\left[\Delta t - 4r(m\Delta t - \tau)\sin^2\left(\frac{j\pi}{N}\right)\right]\lambda - \frac{1}{2}(m\Delta t - \tau - \Delta t)\left[1 - 4r\sin^2\left(\frac{j\pi}{N}\right)\right].$$

Note that only the stepsizes $\Delta t \in [\tau/m, \tau/(m-1))$ are relevant for each $m \geqslant 2$.

We computed stability regions, for various fixed values of τ and found that, with $0 < \tau < \pi/2$:

- $U^n \equiv e \in \mathbb{R}^N$ is linearly stable for all positive stepsizes $\Delta t < \tau$ (that is, all $m > 1$) when $0 < \widehat{r} < 1/2$ (cf. the smaller region $0 < \widehat{r} < \frac{1}{2} - \frac{1}{4}(\Delta t - \tau)$ for $\Delta t \geqslant \tau$).
- A small positive delay improves the stability of $U^n \equiv e \in \mathbb{R}^N$ in the sense that it remains stable for larger stepsizes, namely $2(1 - 2\widehat{r}) \leqslant \Delta t < \tau + 2(1 - 2\widehat{r})$ and $0 < \widehat{r} < 1/2$.
- There is a critical value of $\tau \geqslant \pi/2$ beyond which the fixed point becomes unstable for small stepsizes but continues to be stable for larger stepsizes. This behaviour with the Improved Euler method is distinct from that for Euler's method which ceased to be stable for all stepsizes with a delay larger than $\pi/2$. Recall that, for the continuous partially discretized system (2.2), the constant solution $U(t) \equiv e \in \mathbb{R}^N$ is linearly stable only when $0 < \tau < \pi/2$.

The technique for analyzing the stability as $\Delta t \to 0$ and $\Delta x \to 0$ with $\Delta t/(\Delta x)^2$ fixed that was developed for Euler's method can be applied to the Improved Euler method, leading to the same conclusion.

A numerical example to illustrate the stability of the fixed point $U^n \equiv e \in \mathbb{R}^N$ is presented in Fig. 3. Here $\tau = 2$ and $\widehat{r} = \Delta t/10$ for stepsizes in the range $0 < \Delta t < 3$. Note that ε remains constant for all stepsizes and, in particular, $N = 2$ gives $\varepsilon = 0.025$. The top diagram shows the largest root (in modulus) of the characteristic polynomial at the solution $U^n \equiv e$. The bottom diagram shows the last 50 values $\{U_1^n\}_{n=451}^{500}$ resulting after executing 500 time steps using initial values of the form $\Phi(t) \equiv [0.9, 1.1]^{\mathrm{T}}$. It can be seen that the iterates converge to the nonzero fixed point when the polynomial is Schur.

5. Dynamics of variable stepsize algorithms

We now investigate the behaviour of a variable stepsize algorithm, based on local error control. We consider Euler's method and the Improved Euler method, regarded as a Runge–Kutta pair. This allows us to compare the variable stepsize results with the constant stepsize results derived in the previous sections.

To analyze the adaptive algorithm our approach is to assume that the error control process advances the numerical solution to the vicinity of the true fixed point $U(t) \equiv e \in \mathbb{R}^N$.

We then examine the behaviour of the algorithm around this solution. Linearizing the partially discretized Hutchinson's equation about the fixed point, by substituting $U(t) = e + \delta(t) \in \mathbb{R}^N$ in (2.2) and ignoring second and higher order terms, leads to the system

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\delta(t) = (\varepsilon N^2)M\delta(t) - \delta(t-\tau), & t > 0, \\ \delta(t) = \psi(t), & t \in [-\tau, 0], \end{cases}$$

where $\psi(t) := [\Phi(t) - e] \in \mathbb{R}^N$. Using the linear Lagrange interpolant $q(t-\tau)$ to approximate the delayed values $\delta(t-\tau)$, let $f(t, \delta(t), q(t-\tau)) := (\varepsilon N^2)M\delta(t) - q(t-\tau) \in \mathbb{R}^N$.

We outline the linear Lagrange interpolation procedure when the stepsize is no longer constant. If $(t_n - \tau) \leqslant 0$ then we can take $q(t_n - \tau) = \psi(t_n - \tau)$. Otherwise, suppose that $t_n - \tau$ lies in the interval $[t_{n-m}, t_{n-m+1})$. That is, m is the smallest integer such that $\Delta t_{n-m} + \Delta t_{n-m+1} + \cdots + \Delta t_{n-1} \geqslant \tau$, where Δt_i denotes the (non-constant) stepsize taken from t_i to t_{i+1}. Let

$$\theta_n := \frac{1}{\Delta t_{n-m}} \left[\sum_{i=1}^{m} \Delta t_{n-i} - \tau \right] \in [0, 1)$$

so that the remainder $\Delta t_{n-m} + \Delta t_{n-m+1} + \cdots + \Delta t_{n-1} - \tau$ is denoted by $\theta_n \Delta t_{n-m}$.

Both Euler's method and the Improved Euler method require the approximation $\delta(t_n - \tau) \approx q(t_n - \tau)$, where $(t_n - \tau) \in [t_{n-m}, t_{n-m+1})$. The linear interpolation over the interval $[t_{n-m}, t_{n-m+1})$ gives $q(t_n - \tau) = \theta_n \delta^{n-m+1} + (1 - \theta_n)\delta^{n-m}$.

Further, the Improved Euler method also requires the approximation $\delta(t_n + \Delta t_n - \tau) \approx q(t_n + \Delta t_n - \tau)$. If $(t_n + \Delta t_n - \tau) \leqslant 0$ then we can take $q(t_n + \Delta t_n - \tau) = \psi(t_n + \Delta t_n - \tau)$. Otherwise, suppose that $t_n + \Delta t_n - \tau$ lies in the interval $[t_{n-m+l}, t_{n-m+l+1})$. That is, l is the smallest integer such that $\Delta t_{n-m+l} + \Delta t_{n-m+l+1} + \cdots + \Delta t_n \geqslant \tau$. Let

$$\widehat{\theta}_n := \frac{1}{\Delta t_{n-m+l}} \left[\sum_{i=0}^{m-l} \Delta t_{n-i} - \tau \right] \in [0, 1)$$

so that the remainder $\Delta t_{n-m+l} + \Delta t_{n-m+l+1} + \cdots + \Delta t_n - \tau$ is denoted by $\widehat{\theta}_n \Delta t_{n-m+l}$. The formula for $q(t_n + \Delta t_n - \tau) = q(t_{n-m+l} + \widehat{\theta}_n \Delta t_{n-m+l})$, which interpolates linearly over the interval $[t_{n-m+l}, t_{n-m+l+1})$, is given by

$$q(t_n + \Delta t_n - \tau) = \widehat{\theta}_n \widehat{\delta}^{n-m+l+1} + \left(1 - \widehat{\theta}_n\right)\delta^{n-m+l},$$

where (to keep the variable stepsize algorithm simple)

$$\widehat{\delta}^{n-m+l+1} := \begin{cases} \delta^n + \Delta t_n f(t_n, \delta^n, q(t_n - \tau)) & \text{for } m = l, \\ \delta^{n-m+l+1} & \text{for } m \neq l. \end{cases}$$

The current nth step of the variable stepsize algorithm begins by computing

$$\begin{aligned} k_1 &= f\big(t_n, \delta^n, q(t_n - \tau)\big), \\ k_2 &= f\big(t_n + \Delta t_n, \delta^n + \Delta t_n k_1, q(t_n + \Delta t_n - \tau)\big), \\ \mathrm{est}_{n+1} &= \tfrac{1}{2}\Delta t_n (k_2 - k_1), \\ \Delta t_{\mathrm{new}} &= \left(\frac{\theta\,\mathrm{TOL}}{\|\mathrm{est}_{n+1}\|}\right)^{1/2} \Delta t_n, \end{aligned}$$

where TOL is a parameter supplied by the user and the fixed safety factor $\theta \in (0,1)$ is used to reduce the likelihood of step rejections.

If $\|\mathrm{est}_{n+1}\| \leqslant \mathrm{TOL}$, then the step is accepted and the current values are updated to

$$\begin{aligned} \delta^{n+1} &= \begin{cases} \delta^n + \Delta t_n k_1 & \text{for Euler's method, or} \\ \delta^n + \frac{1}{2}\Delta t_n (k_1 + k_2) & \text{for the Improved Euler method,} \end{cases} \\ t_{n+1} &= t_n + \Delta t_n, \\ \Delta t_{n+1} &= \Delta t_{\mathrm{new}}. \end{aligned}$$

Otherwise, $\|\mathrm{est}_{n+1}\| > \mathrm{TOL}$, so the step is rejected. In this case, we set $\Delta t_n = \Delta t_{\mathrm{new}}$ and repeat the process until $\|\mathrm{est}_{n+1}\| \leqslant \mathrm{TOL}$. Note that, either the Improved Euler method or Euler's method can be used to advance the solution in time. The behaviour of the Improved Euler method or Euler's method can therefore be analyzed in a variable stepsize context, enabling comparisons to be made with the corresponding constant stepsize method.

Let $w_j(M) \in \mathbb{R}^N$, with $j \in [1, 2, \ldots, N]$, be an eigenvector corresponding to the eigenvalue $\lambda_j(M) = -4\sin^2(j\pi/N)$ of the matrix $M \in \mathbb{R}^{N\times N}$. Note that the eigenvalues satisfy $\lambda_{N-j}(M) = \lambda_j(M)$. Also, let $r_n := \varepsilon N^2 \Delta t_n$.

5.1. Euler's method: variable stepsize

Consider advancing the solution in time using Euler's method, so that $\delta^{n+1} = \delta^n + \Delta t_n k_1 \in \mathbb{R}^N$. Following the analysis in [6] we look for period two equilibrium states, where the stepsize remains constant while the numerical solution alternates in sign, of the form

$$\Delta t_n \equiv \Delta t_D \quad \text{and} \quad \delta^{n\pm k} \equiv (-1)^k \delta_D \in \mathbb{R}^N.$$

In such a state, $r_n \equiv r_D := \varepsilon N^2 \Delta t_D$, $l = 1$ and $\theta_D = \widehat{\theta}_D \equiv m - \tau/\Delta t_D$, where m is the smallest integer such that $m\Delta t_D \geqslant \tau$. Imposing these conditions, we find that such solutions occur with

$$\begin{cases} \Delta t_{D_j} = \dfrac{2}{(2m-1)}\left[\tau - (-1)^m + (-1)^m 2 r_{D_j} \sin^2\left(\dfrac{j\pi}{N}\right)\right], \\ \delta_{D_j} = \pm\left(\dfrac{\theta\,\mathrm{TOL}}{2}\right) \dfrac{w_j(M)}{\|w_j(M)\|} \in \mathbb{R}^N, \end{cases}$$

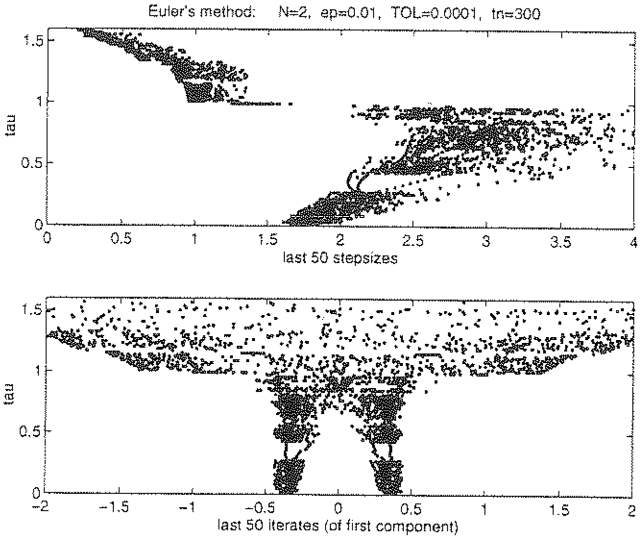

Fig. 4. Behaviour of the variable stepsize Euler's method when $N = 2$ and $\varepsilon = 0.01$.

for $j = 1, 2, \ldots, N$. Each jth equilibrium state is only relevant when $\Delta t_{D_j} > 0$ and the relation $(m-1)\Delta t_{D_j} < \tau \leqslant m\Delta t_{D_j}$ is satisfied, implying that

$$0 < r_{D_j} \sin^2\left(\frac{j\pi}{N}\right) \leqslant \frac{1}{4}(2+\tau) \quad \text{if } m = 1,$$

or

$$\frac{1}{2}\left[(-1)^m - \frac{\tau}{2m}\right] \leqslant (-1)^m r_{D_j} \sin^2\left(\frac{j\pi}{N}\right) < \frac{1}{2}\left[(-1)^m + \frac{\tau}{2(m-1)}\right] \quad \text{if } m \neq 1.$$

Consider the equilibrium state $j = N/2$ (if N is even) or $j = (N+1)/2$ (if N is odd), with stepsize written compactly in the form

$$\Delta t_n \equiv \Delta t_D := \frac{2}{(2m-1)}\left[\tau - (-1)^m + (-1)^m 2\widehat{r}_D\right], \tag{5.1}$$

where $\widehat{r}_D$ is as defined by (3.1) with $\widehat{r}$ and r respectively replaced by $\widehat{r}_D$ and r_D. Numerical tests indicated that the stepsize given by (5.1) for the variable stepsize Euler's method agrees with the largest stepsize Δt for which the fixed point $U^n \equiv e \in \mathbb{R}^N$ is linearly stable for the corresponding constant stepsize Euler's method.

A numerical example showing the long-time behaviour of the variable stepsize Euler's method applied to the nonlinear system (2.2) is shown in Fig. 4 when $N = 2$ and $\varepsilon = 0.01$. It was derived using $\theta = 0.8$, TOL $= 1.0\text{e}{-04}$ and initial values $\Phi(t) = [1 + \theta\,\text{TOL}, 1 - \theta\,\text{TOL}]^{\mathrm{T}} \in \mathbb{R}^2$. The algorithm is executed up to $t = 300$. As τ is varied from 0 to 1.6, the last 50 stepsizes Δt_n are plotted in the top diagram and the last 50 $(U_1^n - 1)/(\theta\,\mathrm{TOL})$ are plotted in the bottom diagram. Note that in the case $m = 1$ the equilibrium state with $j = N/2$ for the linearized problem is

$$\begin{cases} \Delta t_n \equiv \dfrac{2(1+\tau)}{(1+4\varepsilon N^2)} \approx 1.72(1+\tau), \\ \delta^n \equiv \pm(-1)^n \dfrac{\theta\,\mathrm{TOL}}{2\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \approx \pm(-1)^n 0.35(\theta\,\mathrm{TOL}) \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \end{cases}$$

using the l_2 norm. The long-time numerical solution in Fig. 4 appears to be oscillating around this equilibrium state for $0 < \tau < 1$. This illustrates a typical example of an unstable equilibrium state. The fixed point $U^n \equiv e \in \mathbb{R}^N$ becomes linearly unstable for all $\tau \geqslant 1$ (when $m = 1$) for the corresponding constant stepsize case and the algorithm therefore chooses more appropriate stepsizes for a delay larger than one.

5.2. The Improved Euler method: variable stepsize

Consider advancing the solution in time using the Improved Euler method, so that $\delta^{n+1} = \delta^n + \frac{1}{2}\Delta t_n(k_1 + k_2) \in \mathbb{R}^N$. Following the analysis in [6] we look for period one equilibrium states, where the stepsize and numerical solution both remain constant from step to step, giving

$$\Delta t_n \equiv \Delta t_D \quad \text{and} \quad \delta^{n\pm k} \equiv \delta_D \in \mathbb{R}^N.$$

In such a state, $r_n \equiv r_D := \varepsilon N^2 \Delta t_D$, $l = 1$ and $\theta_D = \widehat{\theta}_D \equiv m - \tau/\Delta t_D$, where m is such that $(m-1)\Delta t_D < \tau \leqslant m\Delta t_D$. The two individual cases $m = 1$ and $m \neq 1$ are analyzed separately.

In the case $m = 1$, there are period one equilibrium states of the form

$$\begin{cases} \Delta t_{D_j} = (\tau+2) - 4r_{D_j}\sin^2\left(\dfrac{j\pi}{N}\right), \\ \delta_{D_j} = \pm\left(\dfrac{\theta\mathrm{TOL}}{\tau+2}\right)\dfrac{w_j(M)}{\|w_j(M)\|} \in \mathbb{R}^N, \end{cases}$$

for $j = 1, 2, \ldots, N$. This solution is only relevant when $\Delta t_{D_j} > 0$ and $\Delta t_{D_j} \geqslant \tau$, implying that

$$0 < r_{D_j}\sin^2\left(\frac{j\pi}{N}\right) \leqslant \frac{1}{2}.$$

The smallest Δt_{D_j} occurs when either $j = N/2$ (if N is even) or $j = (N+1)/2$ (if N is odd), say $\Delta t_n \equiv \Delta t_D := (\tau+2) - 4\widehat{r}_D$. This is only relevant when $0 < \widehat{r}_D \leqslant 1/2$. Recall that with a delay in the range $0 < \tau < \pi/2$ the fixed point $U^n \equiv e \in \mathbb{R}^N$ for the corresponding constant stepsize Improved Euler method was determined to be linearly stable when $\tau \leqslant \Delta t < (\tau+2) - 4\widehat{r}$. It follows that in the case $m = 1$, $\Delta t_n \equiv \min_{1\leqslant j\leqslant N}(\Delta t_{D_j})$ for the variable stepsize Improved Euler method agrees with the largest stepsize Δt for which the fixed point $U^n \equiv e \in \mathbb{R}^N$ is linearly stable for the corresponding constant stepsize Improved Euler method.

In the case $m \neq 1$, there are period one equilibrium states of the form

$$\begin{cases} \Delta t_{D_j} = \dfrac{1}{2\varepsilon N^2 \sin^2(j\pi/N)}, \\ \delta_{D_j} = \pm\left(\dfrac{2\varepsilon N^2\sin^2(j\pi/N)\theta\,\mathrm{TOL}}{4\varepsilon N^2\sin^2(j\pi/N)+1}\right)\dfrac{w_j(M)}{\|w_j(M)\|} \in \mathbb{R}^N, \end{cases}$$

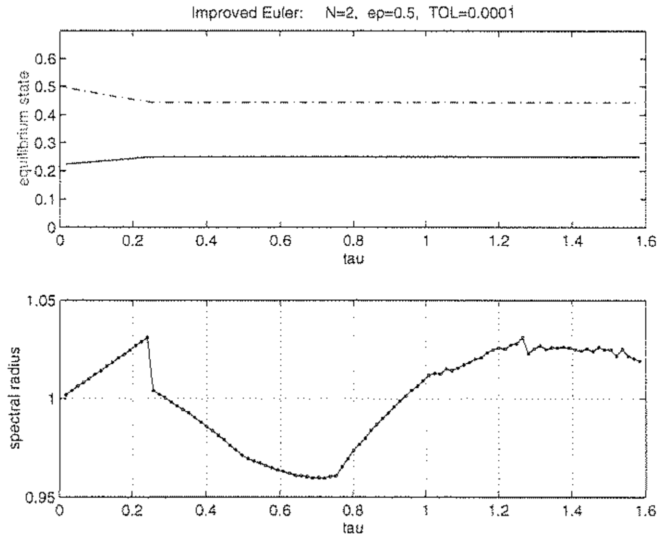

Fig. 5. Stability of equilibrium states for the Improved Euler method.

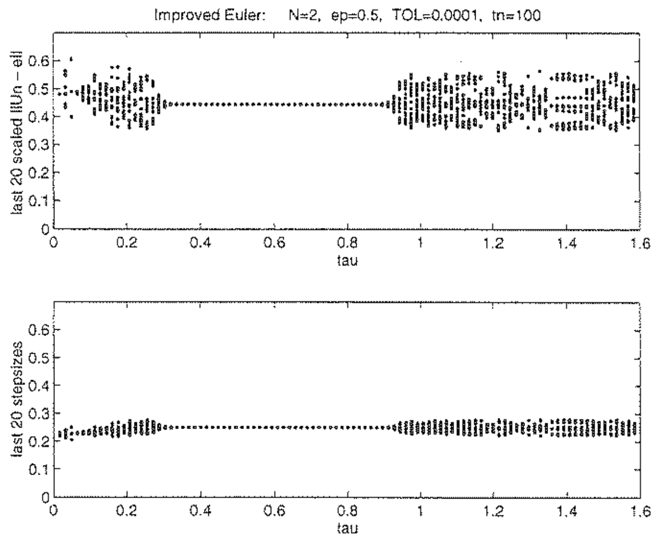

Fig. 6. Long-time behaviour of the variable stepsize Improved Euler method.

for $j = 1, 2, \ldots, N-1$. This solution is only relevant when $\Delta t_{D_j} > 0$ and $(m-1)\Delta t_{D_j} < \tau \leqslant m\Delta t_{D_j}$, implying that

$$\frac{m-1}{2\varepsilon N^2 \sin^2(j\pi/N)} < \tau \leqslant \frac{m}{2\varepsilon N^2 \sin^2(j\pi/N)}.$$

As in the $m = 1$ case above, the stepsize $\Delta t_n \equiv \min_j(\Delta t_{D_j})$ for the variable stepsize Improved Euler method agrees with the largest constant stepsize for which the Improved Euler method is stable.

The stability of the equilibrium state $j = N/2$ when $N = 2$ and $\varepsilon = 0.5$ (using $\theta = 0.8$ and TOL $= 1.0\text{e}{-04}$) is shown in Figs. 5 and 6 for $0 < \tau < 1.6$. In Fig. 5, Δt_{D_1} and $\|\delta_{D_1}\|_2/(\theta\,\mathrm{TOL})$ are shown respectively by the lines "—" and "–·" on the top diagram. The values of the spectral radius of the Jacobian at this equilibrium state are shown in the bottom diagram. These were derived using Maple. The long-time behaviour of the variable stepsize algorithm applied to the nonlinear problem (2.2), using initial values of the form $\Phi(t) = [1 + \theta\,\mathrm{TOL}, 1 - \theta\,\mathrm{TOL}]^{\mathrm{T}} \in \mathbb{R}^2$, is shown in Fig. 6. The algorithm is executed up to $t = 100$. As τ is varied from 0 to 1.6, the last 20 values of $\|U^n - e\|_2/(\theta\mathrm{TOL})$ are plotted in the top diagram and the last 20 stepsizes Δt_n are plotted in the bottom diagram. It can be seen that the long-time numerical solution is attracted to the equilibrium state where the spectral radius is less than one and oscillates around this equilibrium state where the spectral radius is greater than one.

References

[1] M.A. Aves, D.F. Griffiths and D.J. Higham, Does error control suppress spuriosity?, *SIAM J. Numer. Anal.* 34 (1997) 756–778.

[2] D.F. Griffiths and A.R. Mitchell, Stable periodic bifurcations of an explicit discretization of a nonlinear partial differential equation in reaction diffusion, *IMA J. Numer. Anal.* 8 (1988) 435–454.

[3] D.F. Griffiths, P.K. Sweby and H.C. Yee, On spurious asymptotic numerical solutions of explicit Runge–Kutta methods, *IMA J. Numer. Anal.* 12 (1992) 319–338.

[4] G. Hall, Equilibrium states of Runge–Kutta schemes, *ACM Trans. Math. Software* 11 (1985) 289–301.

[5] D.J. Higham, The dynamics of a discretized nonlinear delay differential equation, in: D.F. Griffiths and G.A. Watson, eds., *Numerical Analysis 1993* (Longman Scientific and Technical, 1994) 167–179.

[6] D.J. Higham and I.T. Famelis, Equilibrium states of adaptive algorithms for delay differential equations, *J. Comput. Appl. Math.* 58 (1995) 151–169.

[7] D.J. Higham and B. Owren, Nonnormality effects in a discretized nonlinear reaction-convection–diffusion equation, *J. Comput. Phys.* 124 (1996) 309–323.

[8] D.J. Higham and T. Sardar, Existence and stability of fixed points for a discretized nonlinear reaction–diffusion equation with delay, *Appl. Numer. Math.* 18 (1995) 155–173.

[9] A.R. Humphries, Spurious solutions of numerical methods for initial value problems, *IMA J. Numer. Anal.* 13 (1993) 263–290.

[10] A. Iserles, Stability and dynamics of numerical methods for nonlinear ordinary differential equations, *IMA J. Numer. Anal.* 10 (1990) 1–30.

[11] Z. Jackiewicz, R. Vermiglio and M. Zennaro, Regularity properties of Runge–Kutta methods for delay differential equations, *Appl. Numer. Math.* 24 (1997) 265–278 (this issue).

[12] A.R. Mitchell and D.F. Griffiths, *The Finite Difference Method in Partial Differential Equations* (Wiley, New York, 1985).

[13] T.K. Sardar, Dynamics of constant and variable stepsize methods for a nonlinear reaction-diffusion model with delay, Ph.D. Thesis, Department of Mathematics and Computer Science, University of Dundee (1996).

[14] A.M. Stuart and A.R. Humphries, *Dynamical Systems and Numerical Analysis* (Cambridge University Press, 1996).