



A Unified Approach to Spurious Solutions Introduced by Time Discretisation. Part I: Basic Theory

Author(s): A. Iserles, A. T. Peplow and A. M. Stuart

Source: *SIAM Journal on Numerical Analysis*, Vol. 28, No. 6 (Dec., 1991), pp. 1723-1751

Published by: [Society for Industrial and Applied Mathematics](#)

Stable URL: <http://www.jstor.org/stable/2157957>

Accessed: 16/09/2013 06:57

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Society for Industrial and Applied Mathematics is collaborating with JSTOR to digitize, preserve and extend access to *SIAM Journal on Numerical Analysis*.

<http://www.jstor.org>

A UNIFIED APPROACH TO SPURIOUS SOLUTIONS INTRODUCED BY TIME DISCRETISATION. PART I: BASIC THEORY*

A. ISERLES[†], A. T. PELOW^{‡§}, AND A. M. STUART[‡]

Abstract. The asymptotic states of numerical methods for initial value problems are examined. In particular, spurious steady solutions, solutions with period 2 in the timestep, and spurious invariant curves are studied. A numerical method is considered as a dynamical system parameterised by the timestep h . It is shown that the three kinds of spurious solutions can bifurcate from genuine steady solutions of the numerical method (which are inherited from the differential equation) as h is varied. Conditions under which these bifurcations occur are derived for Runge–Kutta schemes, linear multistep methods, and a class of predictor-corrector methods in a PE(CE)^M implementation. The results are used to provide a unifying framework to various scattered results on spurious solutions which already exist in the literature. Furthermore, the implications for choice of numerical scheme are studied. In numerical simulation it is desirable to minimise the effect of spurious solutions. Classes of methods with desirable dynamical properties are described and evaluated.

Key words. spurious solutions, bifurcation, timestepping, regularity

AMS(MOS) subject classifications. 65L05, 65L20

1. Introduction. It is well known that a numerical method which is convergent at a fixed time does not necessarily yield the same asymptotic behaviour as the underlying differential equation for fixed values of the timestep. In many circumstances it is the asymptotic behaviour which is of interest in the differential equation. Consequently it is desirable to design discretisations which mimic the asymptotic properties of the differential equation.

The asymptotic states of a dynamical system are captured in the ω and α limit sets which may contain equilibria, periodic solutions, quasi-periodic solutions, strange attractors, etc. [Guckenheimer and Holmes (1983)]. The behaviour of a dynamical system is determined by the ω and α limit sets and the orbits connecting them to one another and to infinity. Here we analyze the structure of the ω and α limit sets for convergent discretisations of initial value problems. It is important to design schemes for which these sets (under appropriate prolongation) are close to the corresponding limit sets for the underlying differential equation. Thus it is important to understand (and hence to avoid) conditions under which spurious members of the limit sets are introduced by time discretisation. Stable spurious solutions are undesirable since they attract a certain subset of the initial states; for this subset the asymptotic behaviour of the scheme is clearly incorrect [Griffiths and Mitchell (1988), Iserles (1990)]. Unstable spurious solutions are also extremely undesirable: consider the class of differential equations whose solutions remain bounded in time; the unstable manifolds of spurious solutions introduced by discretisation are often connected to infinity [Griffiths and Mitchell (1988), Stuart (1990)] and this causes breakdown of the scheme since the sequence generated by the numerical method can diverge to infinity rather than remaining bounded.

* Received by the editors March 19, 1990; accepted for publication (in revised form) November 19, 1990.

[†] Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 9EW, United Kingdom.

[‡] School of Mathematical Sciences, University of Bath, Bath BA2 7AY, United Kingdom.

[§] Present address, Department of Civil Engineering, University of Bradford, Bradford BD7 1DP, United Kingdom. The research of this author was supported by the United Kingdom Science and Engineering Research Council.

We present a unified approach to the existence of spurious solutions in the limit sets. The numerical method is viewed as a dynamical system in which the timestep h plays the role of a bifurcation parameter. Usually, steady solutions of the underlying differential equation are inherited by the discretisation. We look for bifurcations from these genuine steady solutions into spurious solutions as h is varied.

Consider the following initial value problem. Find $y \in \mathcal{R}^p$ satisfying

$$(1.1) \quad \frac{dy}{dt} = f(y), \quad t > 0$$

and

$$(1.2) \quad y(0) = y_0.$$

Here $f \in C^2(\mathcal{R}^p, \mathcal{R}^p)$. We introduce a temporal mesh $t_n = nh$ and let Y_n denote our approximation to $y(t_n)$. Typically the sequence Y_n is derived from a k -step (implicit) map of the form

$$(1.3) \quad F(Y_n, \dots, Y_{n+k-1}, Y_{n+k}; h) = 0, \quad n = 0, 1, \dots,$$

together with k initial conditions. Thus (1.3) must be solved for Y_{n+k} given Y_n, \dots, Y_{n+k-1} . Henceforth we assume that (1.3) forms a consistent approximation to (1.1). By considering a vector of k components we may write (1.3) as a one-step map of the form

$$(1.4) \quad H(U_{n+1}, U_n; h) = 0,$$

where $U_n = [Y_n^T, \dots, Y_{n+k-1}^T]^T \in \mathcal{R}^{kp}$. For motivation we consider the case of explicit methods, in which case (1.4) takes the form

$$(1.5) \quad U_{n+1} = G(U_n; h).$$

We briefly outline the bifurcation theory that underlies the results in this paper. Let us assume that equation (1.1) has an equilibrium \hat{y} such that $f(\hat{y}) = 0$. We assume also that this equilibrium is inherited by the numerical method (1.3)—this is true of all the methods considered in this paper [Iserles (1990)]. Thus $\hat{U} = [\hat{y}, \dots, \hat{y}]^T$ is a fixed point of the map (1.5). This true equilibrium appears as a straight line in the graph of the limit sets of (1.5) against h . See Fig. 1. It is now possible to look for bifurcations from this fixed point as h varies. Bifurcation occurs (subject to various nondegeneracy conditions) when the eigenvalues of $G_U(\hat{U}; h)$ (the Jacobian of G at $U = \hat{U}$) pass through the unit circle in the complex plane as h varies. This determines critical values of h , say h_c , at which spurious solutions can bifurcate from the true equilibrium \hat{U} . See Fig. 1. Three kinds of bifurcations are of interest here:

- (i) If an eigenvalue passes through $+1$ then a fixed point of the map (1.5) bifurcates from the fixed point \hat{U} (*steady bifurcation*). This will typically occur as a transcritical bifurcation (cf. Fig. 1(a)).
- (ii) If an eigenvalue passes through -1 then a solution of the map (1.5) with period 2 in n bifurcates from the fixed point \hat{U} (*period doubling or flip bifurcation*). This will typically occur as a pitchfork bifurcation (cf. Fig. 1(b)). A period 2 solution is one for which $U_{n+2} = U_n$ for all n and for which $U_{n+1} \neq U_n$.

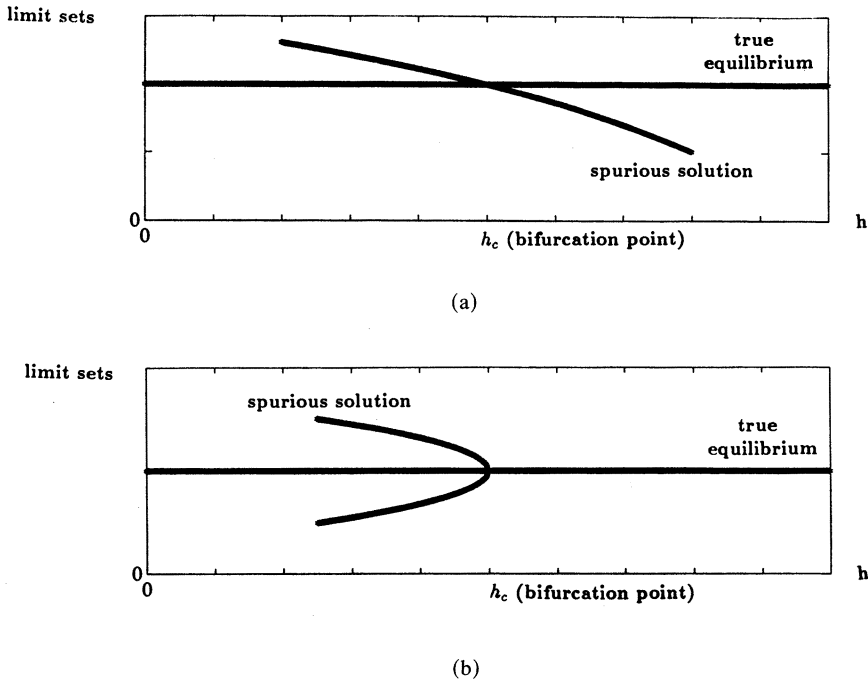


FIG. 1.1.

- (iii) If a pair of complex conjugate eigenvalues passes through the unit circle then a closed invariant curve for the map (1.5) bifurcates from the fixed point \hat{U} (Hopf bifurcation).

Precise statements of these results can be found in Guckenheimer and Holmes (1983). These bifurcations give us a handle on, respectively, spurious steady solutions, spurious solutions with period 2 in n , and spurious invariant curves. In general, we would like to prevent the existence of such spurious solutions. The local bifurcation theory gives us conditions *necessary* to prevent the existence of spurious solutions of a particular class. Rather surprisingly, these necessary conditions are sometimes *sufficient* as well: that is, the necessary conditions derived from a local theory turn out to be sufficient for a global result. (See Theorem LMM(ii).) Note that $G_U(\hat{U}; h)$ is closely related to the linear stability function for the method linearised about the equilibrium \hat{U} . This provides a natural connection between the nonlinear theory described here and classical linear numerical analysis. On a more profound level, it also points out toward a relationship between our approach, which is based on Poincaré stability, and the Lyapunov-type nonlinear stability theory of Dahlquist et al. [Burrage and Butcher (1979), Dahlquist (1978), Dekker and Verwer (1984), Lambert (1987)]. Note, however, that we can also construct spurious steady solutions of Runge–Kutta methods which do not bifurcate from the linear stability limit of the numerical method but from different critical values of h associated with the internal stages of the method. (See Theorem RK(ii), (iii).)

The existence of spurious steady solutions introduced by temporal discretisation was discovered in Iserles (1990) and analyzed further in Hairer, Iserles, and Sanz-Serna (1989). The first of these papers is concerned with Runge–Kutta, linear multistep,

and predictor-corrector methods; it contains some general theory but is mainly devoted to expository examples arising from discretisations of a Riccati equation. The second paper contains a systematic characterisation of those Runge–Kutta methods which do not possess spurious steady solutions, by means of a recursive algebraic test. Recent work by Griffiths, Sweby, and Yee (1990) analyzes the existence and stability of spurious steady solutions of explicit (up to 4-stage) Runge–Kutta methods applied to various nonlinear ordinary differential equations (ODEs) with competing equilibria. The existence of spurious period 2 solutions in various specific equations involving quadratic nonlinearities is analyzed by Griffiths and Mitchell (1988), Mitchell and Schoombie (1989), Schoombie and Mitchell (1989), and Sleeman et al. (1988). The equations studied in these papers are such that algebraic manipulations can be used to establish the existence of spurious solutions explicitly and their linearised stability properties can then be analyzed. The existence of spurious invariant curves is discussed in Brezzi, Ushiki, and Fujii (1984) where a specific example is discussed in which the spurious solution can be constructed explicitly. The paper by Lorenz (1989) also contains a brief discussion of a spurious Hopf bifurcation and its role in the onset of computationally induced chaos. Newell (1977) was the first to discuss the existence of spurious solutions by analytical techniques applicable to arbitrary nonlinearities. The bifurcation analysis outlined in his paper, where spurious period 2 solutions arising in a discretisation of the cubic Burgers equation are considered, is the precursor of the analysis in this paper. Subsequent work has generalised and extended Newell's work on period 2 solutions to a broad class of problems [Stuart (1989a), (1989b); Stuart and Peelow (1990)].

The unified treatment in this paper establishes an underlying principle in *all* the papers cited above and we will refer to these papers to illustrate our theorems. In addition to providing a unifying framework, the theory also yields many new results. We now make some definitions. These definitions generalise that in Hairer, Iserles, and Sanz-Serna (1989). Note that we assume a degree of regularity on $f(u)$. This enables a straightforward application of local bifurcation theorems and may not be necessary in certain circumstances.

DEFINITION 1.1. The consistent numerical method (1.3) is *regular of degree 1*, denoted $R^{[1]}$, if every fixed point $\hat{y} \in \mathcal{R}^p$ of the map (1.3) satisfies $f(\hat{y}) = 0$ for all $h > 0$ and all equations (1.1) with $f \in C^2(\mathcal{R}^p, \mathcal{R}^p)$. Otherwise it is *irregular of degree 1*.

DEFINITION 1.2. The consistent numerical method (1.3) is *regular of degree 2*, denoted $R^{[2]}$ if, for all $h > 0$ and all equations (1.1) with $f \in C^2(\mathcal{R}^n, \mathcal{R}^n)$, (1.3) does not admit real period 2 solutions in n . Otherwise it is *irregular of degree 2*.

DEFINITION 1.3. Let $\mathcal{S}[T]$ denote the (closed) linear stability domain of the consistent numerical method (1.3). The method is said to be *Hopf regular*, denoted $R^{[H]}$, if $\partial\mathcal{S}[T] \equiv i\mathcal{R}$.

DEFINITION 1.4. We denote by $R^{[1,2]}$ methods which are regular of degree 1 and 2. Similarly we employ the notation $R^{[1,H]}$, etc. A method which is $R^{[1,2,H]}$ is said to be *strictly regular*.

The importance of having a method which is regular of degree 1 is obvious since it is clearly undesirable to have steady solutions which are a product of the discretisation and not a feature of the underlying differential equation. Examples are given by Iserles (1990). The importance of choosing methods which are regular of degree 2 is slightly more subtle, as period 2 solutions are readily recognised as spurious since they vary on a grid scale. However, spurious period 2 solutions play a prominent

role in determining the dynamics of (1.3) and the reader should consult Griffiths and Mitchell (1988), Mitchell and Schoombie (1989), Schoombie and Mitchell (1989), Stuart (1989a)–(1989c), and Stuart and Peplow (1990) for numerous examples and discussions of their effects. Methods which are not Hopf regular can undergo spurious Hopf bifurcations. The role of spurious closed invariant curves, which arise at spurious Hopf bifurcation points, is clearly illustrated by the example in Brezzi, Ushiki, and Fujii (1984). Spurious Hopf bifurcations can occur in a method which is not $R^{[H]}$.

The following theorems summarise the results on regularity obtained in this paper (and include, for completeness, two results obtained in Iserles (1990) and Iserles and Stuart (1990)). Note that the paper contains many other results other than those concerned directly with regularity. In particular, the precise values h_c at which bifurcation of spurious solutions occurs is calculated in Theorems 2.3–2.4, 3.4, 4.3, and 5.1. These values are useful for the initiation of numerical continuation of the spurious solutions to determine how they vary with h [Stuart and Peplow, (1990)]. We emphasise that, for irregular methods, spurious solutions are a generic phenomenon in the sense that they will be observed for *all* functions f such that (1.1) has a hyperbolic equilibrium. It is important to realise that, although the spurious solutions may bifurcate from genuine solutions at values of h well above those used in practice (for example, the linear stability limit), the branches of spurious solutions can penetrate to values of h used in practice. In particular, it is possible to construct examples where the spurious solutions may be found for *arbitrarily small* h —see Iserles (1990), Stuart (1989b), and Stuart and Peplow (1990). Determining whether or not this happens for a given problem relies on knowing something about the structure of the differential equation and the form of the nonlinearity—see Iserles (1990); Stuart (1989b), (1989c); Griffiths and Mitchell (1988)—and this is an important area for further research. In terms of Fig. 1, this corresponds to determining what happens to the branches of spurious solutions as h decreases: do they exist for all h , sufficiently small, do they tend to infinity at some finite value of h , or do they turn around at some finite value of h ? In this paper, however, we take the black-box approach to numerical methods and ask how much we can say about the dynamical properties of numerical methods for initial value problems without imposing any structure on the underlying nonlinearity. Our numerical formalism follows that in Hairer, Nørsett, and Wanner (1987), but specific definitions of various relevant methods are spelt out explicitly in the sequel.

THEOREM RK. *Consider a consistent k -stage Runge–Kutta (RK) method of the form (2.1)–(2.2). Let $R(h) = 1 + hb^T(I - hA)^{-1}\mathbf{1}$ the linear stability function for the method.*

- (i) *The method is not $R^{[1]}$ if there exists a value $h \in \mathcal{R} \setminus \{0\}$ such that $(I - hA)$ is invertible and $R(h) = 1$.*
- (ii) *The method is not $R^{[1]}$ if there exists a value $h \in \mathcal{R}$ such that $(I - hA)$ is singular with one-dimensional null-space, the left and right eigenvectors are η and ξ , respectively, and $\eta^T\mathbf{1} = 0$, $b^T\xi \neq 0$. (This result holds under Technical Condition A, Theorem 2.3.)*
- (iii) *The method is not $R^{[1]}$ if there exists a value $h \in \mathcal{R}$ such that $(I - hA)$ is singular with one-dimensional null-space, the left and right eigenvectors are η and ξ , respectively, and $\eta^T\mathbf{1} \neq 0$, $b^T\xi = 0$. (This result holds under Technical Condition B, Theorem 2.3.)*
- (iv) *The method is not $R^{[2]}$ if there exists a value $h \in \mathcal{R} \setminus \{0\}$ such that $(I - hA)$ is invertible and $R(h) = -1$.*
- (v) *The highest attainable order of an $R^{[1,2]}$ RK method is 2.*

Proof. Parts (i), (ii), and (iii) are proved in Theorem 2.3. Part (iv) is proved in Theorem 2.4. See Corollary 2.5. Part (v) is proved in Theorem 2.7.

THEOREM LMM. *Consider a consistent, irreducible, linear multistep method (LMM) of the form (3.1).*

- (i) *The method is $R^{[1]}$.*
- (ii) *If the method is zero-stable and satisfies $\rho(-1) \neq 0$ then it is $R^{[2]}$ if and only if $\sigma(-1) = 0$.*
- (iii) *The method is not $R^{[2]}$ if $\rho(-1) = 0$.*
- (iv) *The highest attainable order of a convergent $R^{[1,2]}$ LMM is $2[(k+1)/2]$.*

Proof. Part (i) is proved by Iserles (1990). Part (ii) is proved in Theorems 3.4 and 3.5 and part (iii) in Theorem 3.6. Part (iv) is proved by Iserles and Stuart (1990).

THEOREM PC. *Consider the predictor-corrector method (4.1)–(4.3).*

- (i) *The method is not $R^{[1]}$ if m is odd.*
- (ii) *The method is not $R^{[2]}$ if m is even.*

Proof. Both parts are proved in Theorem 4.3. See Corollary 4.4.

THEOREM H. (i) *The theta method is $R^{[H]}$ if and only if $\theta = \frac{1}{2}$.*

- (ii) *Consider the LMM (5.1). If the method is $R^{[H]}$ then it is marginally zero-stable.*
- (iii) *Consider the RK method (2.1)–(2.2) with positive weights b_i . The method is $R^{[H]}$ if it is symplectic, that is, if $M \equiv 0$, where the matrix M is defined in (5.10).*

Proof. Part (i) is proved in Theorem 5.1, part (ii) in Theorem 5.3, and part (iii) in Theorem 5.4.

1.1. Implications for choice of schemes. Generally, we expect that for application to differential equations where trajectories converge to steady solutions, $R^{[1,2]}$ methods will be most desirable. For application to equations where trajectories converge to periodic solutions we expect that $R^{[H]}$ methods will be most desirable. (Of course, many equations will exhibit a variety of behaviours, dependent on initial conditions.)

Theorems RK and LMM show that LMMs are the most desirable $R^{[1,2]}$ methods since $R^{[1,2]}$ RK methods are limited to order 2, whilst $R^{[1,2]}$ LMMs can attain arbitrarily high order. In a companion paper [Iserles and Stuart (1990)], the class of $R^{[1,2]}$ LMMs is examined further and a modification of the backward differentiation formulae to satisfy $\sigma(-1) = 0$ is proposed and its zero-stability properties analyzed.

Theorem H shows that RK methods are the most desirable $R^{[H]}$ schemes. Marginally zero-stable LMMs are undesirable in practice. The class of RK methods which are marginally algebraically stable has been widely studied. In particular, such schemes are advantageous for the integration of Hamiltonian problems, since they preserve the symplectic structure of phase space [Lasagni (1988), Sanz-Serna (1988)]. Thus it is not surprising that they are suitable for the detection of Hopf bifurcations, since the linear problem on the centre manifold determining Hopf bifurcation is area preserving at the critical parameter value.

Theorem LMM(i), (ii) and Theorem H(i) taken together show that the trapezoidal rule is strictly regular. Moreover, it follows easily from our analysis that, amongst all the consistent and strictly (as opposed to “marginally”) zero-stable LMM and RK schemes, the trapezoidal rule is the unique—up to equivalence on linear problems—method with the aforementioned feature.

We do not advocate that the results in this paper be taken in isolation to guide the choice of scheme. There are many other issues which need to be considered. For

example, we have said nothing about the structure of the nonlinear function $f(u)$; a given method may be irregular, but for wide classes of problems with a particular nonlinear structure the spurious solutions may only exist above the linear stability limit and not affect the dynamics of practical computations. Furthermore, we have said nothing of the stability of the spurious solutions. This, too, depends on the structure of $f(u)$. Finally, we have considered only fixed timestepping schemes, whereas variable step schemes are frequently used in ODE codes. It is our belief that the analysis and development of variable timestepping schemes will benefit from the input of ideas from dynamical systems, and it is our hope that, by thorough analysis of the fixed step case, we can establish a framework within which to study the dynamics of numerical methods.

2. Runge–Kutta methods. We consider a general k -stage method which is written as

$$(2.1) \quad \xi_i = f \left(Y_n + h \sum_{j=1}^k a_{i,j} \xi_j \right), \quad i = 1, \dots, k,$$

$$(2.2) \quad Y_{n+1} = Y_n + h \sum_{i=1}^k b_i \xi_i.$$

Let A denote the matrix with entries $a_{i,j}$, let \mathbf{b} denote the vector $[b_1, \dots, b_k]^T$ and $\mathbf{1}$ be the k -vector with unit entries $[1, \dots, 1]^T$. The following lemma is needed to prove Theorems 2.3 and 2.4.

LEMMA 2.1. *Suppose we have the $(k + 1) \times (k + 1)$ bordered matrix*

$$X = \begin{bmatrix} P & \mathbf{q} \\ \mathbf{r}^T & s \end{bmatrix},$$

where P is a $k \times k$ matrix, \mathbf{q} and \mathbf{r} are vectors of length k , and s is a scalar. Then

- (i) If $\text{rank } P = k$ then X is singular if and only if $s - \mathbf{r}^T P^{-1} \mathbf{q} = 0$.
- (ii) If $\text{rank } P = k - 1$ then X is singular if and only if $\mathbf{r}^T \boldsymbol{\xi} = 0$ or $\boldsymbol{\eta}^T \mathbf{q} = 0$. Here $P \boldsymbol{\xi} = \mathbf{0}$ and $\boldsymbol{\eta}^T P = 0$.

Proof. The proof is straightforward. In the more complicated setting of linear operators between infinite-dimensional Banach spaces, it can be found in Keller (1977). \square

The following lemma is used in the proofs of Theorems 2.4, 3.4, and 4.3. It concerns the bifurcation of period 2 solutions in maps of the form (1.4).

LEMMA 2.2. *Let the function $H(a, b; h)$ satisfy $H \in C^r(\mathcal{R}^a \times \mathcal{R}^b, \mathcal{R}^a)$. Assume that the map (1.4) has a fixed point \hat{U} . Assume also that $\text{Null}(H_a(\hat{U}, \hat{U}; h_c) - H_b(\hat{U}, \hat{U}; h_c)) = \text{span}\{\theta\}$ and that $H_a(\hat{U}, \hat{U}; h_c) + H_b(\hat{U}, \hat{U}; h_c)$ is invertible. If*

$$(H_{ah}(\hat{U}, \hat{U}; h_c) - H_{bh}(\hat{U}, \hat{U}; h_c))\theta \notin \text{Range}(H_a(\hat{U}, \hat{U}; h_c) - H_b(\hat{U}, \hat{U}; h_c))$$

then, for $\epsilon \ll 1$, there exists a period 2 solution of (1.4) with the form

$$\begin{aligned} h(\epsilon) &= h_c + O(|\epsilon|), \\ U_n(\epsilon) &= \hat{U} + \epsilon(-1)^n \theta + O(|\epsilon^2|) \end{aligned}$$

which is C^{r-1} in ϵ .

Proof. Period 2 solutions of (1.4) are pairs (V, W) with $V \neq W$ satisfying

$$H(V, W; h) = 0; \quad H(W, V; h) = 0.$$

Throughout this proof, all derivatives of H are evaluated at $V = W = \hat{U}$. The Jacobian of the nonlinear system is

$$\begin{pmatrix} H_a & H_b \\ H_b & H_a \end{pmatrix}.$$

Using the fact that $H_a + H_b$ is invertible it may be shown that the Jacobian is singular at $h = h_c$ and that bifurcation occurs provided that the “range” condition is satisfied—this involves a straightforward application of the bifurcation results in Crandall and Rabinowitz (1971), taking account of the symmetry inherent in the defining equations for period 2 solutions. \square

THEOREM 2.3. *Let $f : \mathcal{R} \mapsto \mathcal{R}$ and let \hat{y} be a hyperbolic equilibrium of equation (1.1). Consider a Runge-Kutta method of the form (2.1)–(2.2). Spurious steady solutions in n bifurcate from the steady solution $Y_n = \hat{y}$ at $h = h_c$, where*

- (i) $\mathbf{b}^T(f'(\hat{y})h_c A - I)^{-1}\mathbf{1} = 0$, provided that $\mathbf{b}^T(f'(\hat{y})h_c A - I)^{-2}\mathbf{1} \neq 0$ and that $(I - h_c f'(\hat{y})A)$ is invertible; or where
- (ii) $(h_c f'(\hat{y})A - I)$ is singular with one-dimensional null-space and with left and right eigenvectors $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$, respectively, $\boldsymbol{\eta}^T \mathbf{1} = 0$, $\mathbf{b}^T \boldsymbol{\xi} \neq 0$, and Technical Condition A, below, holds; or where
- (iii) $(h_c f'(\hat{y})A - I)$ is singular with one-dimensional null-space and with left and right eigenvectors $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$, respectively, $\boldsymbol{\eta}^T \mathbf{1} \neq 0$, $\mathbf{b}^T \boldsymbol{\xi} = 0$, and Technical Condition B, below, holds.

Important Note. It is straightforward to extend this theorem to the case of general $f : \mathcal{R}^p \rightarrow \mathcal{R}^p$. It is necessary to assume that the Jacobian of f at \hat{y} is nonsingular and has a real, simple eigenvalue η . Then η replaces $f'(\hat{y})$ in the definitions of h_c . We prove the result in \mathcal{R} merely for simplicity of exposition.

Technical Conditions. We have

$$(h f'(\hat{y})A - I)\boldsymbol{\xi} = 0$$

and

$$(h f'(\hat{y})A^T - I)\boldsymbol{\eta} = 0.$$

- (i) Let $\boldsymbol{\eta}^T \mathbf{1} = 0$, $\mathbf{b}^T \boldsymbol{\xi} \neq 0$, and let $\boldsymbol{\alpha}$ be the unique vector with $\boldsymbol{\alpha}^T \mathbf{b} = 0$ satisfying

$$(h f'(\hat{y})A - I)\boldsymbol{\alpha} + f'(\hat{y})\mathbf{1} = 0.$$

Then we have Technical Condition A:

$$\boldsymbol{\eta}^T \boldsymbol{\alpha} \neq 0.$$

- (ii) Let $\boldsymbol{\eta}^T \mathbf{1} \neq 0$, $\mathbf{b}^T \boldsymbol{\xi} = 0$, and let $\boldsymbol{\beta}$ be the unique vector with $\boldsymbol{\beta}^T \mathbf{1} = 0$ satisfying

$$(h f'(\hat{y})A^T - I)\boldsymbol{\beta} + f'(\hat{y})\mathbf{b} = 0.$$

Then the Technical Condition B reads

$$\boldsymbol{\beta}^T \boldsymbol{\xi} \neq 0.$$

Proof. Clearly $\xi_i = 0$ and $Y_n = \hat{y}$ is a steady solution of (2.1)–(2.2). We wish to determine values of $h \in \mathcal{R}$ at which spurious steady solutions branch from this genuine steady solution. Steady solutions (ξ_i, Y) of (2.1)–(2.2) satisfy

$$(2.3) \quad \xi_i = f \left(Y + h \sum_{j=1}^k a_{i,j} \xi_j \right),$$

$$(2.4) \quad 0 = h \sum_{i=1}^k b_i \xi_i.$$

We seek bifurcation from the trivial solution $\xi_i = 0, Y = \hat{y}$ to other steady solutions. Bifurcation occurs at the values of h at which the Jacobian of the nonlinear system (2.3)–(2.4) is singular with simple eigenvalue h_c [Chow and Hale (1982)]. Linearising the system about the trivial solution we obtain the matrix problem

$$(2.5) \quad X\Phi = \mathbf{0}.$$

Here $\Phi = [\phi^T, u]^T$ and the matrix X is given by

$$(2.6) \quad X = \begin{bmatrix} hf'(\hat{y})A - I & f'(\hat{y})\mathbf{1} \\ h\mathbf{b}^T & 0 \end{bmatrix}.$$

It can be deduced from Lemma 2.1 that the matrix X is singular in the following two cases:

- (i) If $hf'(\hat{y})A - I$ is nonsingular then the whole matrix is singular at $h = h_c$ given by (i). This corresponds to bifurcation from the linear stability limit.
- (ii) If $hf'(\hat{y})A - I$ is singular with rank $k - 1$ then X is singular if $\mathbf{b}^T \xi = 0$ or if $\eta^T \mathbf{1} = 0$ where ξ is a right eigenvector and η^T is a left eigenvector of A , respectively, with eigenvalue $1/(f'(\hat{y})h)$. This does not correspond to bifurcation from the linear stability limit.

To establish the theorem we need only determine conditions under which these eigenvalues are simple. The eigenvalue problem (2.5) is of the form

$$(2.7) \quad hL_1\Phi = L_2\Phi.$$

Here

$$(2.8) \quad L_1 = \begin{bmatrix} f'(\hat{y})A & \mathbf{0} \\ \mathbf{b}^T & 0 \end{bmatrix}$$

and

$$(2.9) \quad L_2 = \begin{bmatrix} I & -f'(\hat{y})\mathbf{1} \\ \mathbf{0} & 0 \end{bmatrix}.$$

The eigenvalue h_c of (2.7) is simple provided that

$$(2.10) \quad (h_c L_1 - L_2)\Gamma = L_1\Phi$$

is not solvable.

In case (i), straightforward manipulations show that (2.10) is not solvable provided that $\mathbf{b}^T(f'(\hat{y})h_c A - I)^{-2}\mathbf{1} \neq 0$. In cases (ii) and (iii) similar manipulations show that (2.10) is not solvable, subject to Technical Conditions A and B, respectively. \square

Example and interesting observation. Most of the spurious solutions constructed in this paper bifurcate from the stability limit of the numerical method linearised about an equilibrium. However, the spurious steady solutions of Runge–Kutta methods which bifurcate under conditions (ii), (iii) *do not* bifurcate from the linear stability limit. We describe an example which illustrates this phenomenon. Consider the logistic equation

$$y' = y(1 - y) = f(y),$$

solved by the two-stage Runge–Kutta method

$$\begin{aligned} \xi_1 &= f(Y_n + h\xi_1), \\ \xi_2 &= f\left(Y_n + \frac{h}{2}(\xi_1 + \xi_2)\right), \\ Y_{n+1} &= Y_n + \frac{h}{2}(\xi_1 + \xi_2). \end{aligned}$$

It may be shown that $h = 1$ and $h = 2$ are candidates for fixed point bifurcations, according to Theorem 2.3(ii) and (iii), since $hA - I$ is singular at those values. A little algebra shows that, apart from true fixed points zero and 1, the method has spurious fixed points \hat{Y} satisfying

$$h^2\hat{Y}^2 + (2h - h^2)\hat{Y} + 2 - h = 0.$$

Note that these spurious solutions bifurcate from zero at $h = 2$ as predicted above. (In fact, $h = 2$ is also the linear stability limit for this method linearised about zero but, since $R(2) = -1$, “bifurcation from the linear stability limit” would lead us to expect a period 2 bifurcation there—see Theorem 2.4.)

THEOREM 2.4. *Let $f : \mathcal{R} \rightarrow \mathcal{R}$ and let \hat{y} be a hyperbolic equilibrium of (1.1). Consider a Runge–Kutta method of the form (2.1)–(2.2). Period 2 solutions in n bifurcate from the steady solution $Y_n = \hat{y}$ at $h = h_c$ where*

$$(2.11) \quad \mathbf{b}^T(I - h_c f'(\hat{y})A)^{-1}\mathbf{1} + \frac{2}{h_c f'(\hat{y})} = 0,$$

provided that

$$(2.12) \quad \mathbf{b}^T(I - h_c f'(\hat{y})A)^{-2}\mathbf{1} \neq 0$$

and that the matrix $(I - h_c f'(\hat{y})A)$ is invertible.

Proof. Our aim is to find where period 2 solutions of the discrete problem bifurcate from a genuine steady solution. Let $\mathbf{V} = ((\xi_i^v)^T, v)^T$ and $\mathbf{W} = ((\xi_i^w)^T, w)^T$. Then define

$$(2.13) \quad H(\mathbf{V}, \mathbf{W}; h) := \begin{pmatrix} f(v + h \sum_{j=1}^k a_{i,j} \xi_j^v) - \xi_i^v, & i = 1, \dots, k, \\ v + h \sum_{i=1}^k b_i \xi_i^v - w \end{pmatrix}.$$

Then period 2 solutions satisfy

$$H(\mathbf{V}, \mathbf{W}; h) = 0, \quad H(\mathbf{W}, \mathbf{V}; h) = 0, \quad \mathbf{U} \neq \mathbf{V}.$$

Since $I - h_c f'(\hat{y})A$ is invertible, application of Lemma 2.2 shows that h_c is a potential bifurcation point where the matrix Z , given by

$$Z = \begin{bmatrix} h_c f'(\hat{y})A - I & f'(\hat{y})\mathbf{1} \\ h_c \mathbf{b}^T & 2 \end{bmatrix},$$

is singular. By Lemma 2.1 this matrix is singular at h_c given by (2.11). Let $Null(Z) = \text{span}\{\Theta\}$. To complete the proof that bifurcation occurs we need only establish that the eigenvalue is simple in the sense of the “range” condition in Lemma 2.2. This requires showing that

$$Z\Phi = L_1\Theta$$

is not solvable for Φ , where L_1 is given by (2.8). Straightforward manipulation shows that the problem has no solution, and hence h_c is simple, provided that (2.12) holds. \square

COROLLARY 2.5. *Consider a general consistent k -stage Runge–Kutta method of the form (2.1)–(2.2). Let $R(h) = 1 + h\mathbf{b}^T(I - hA)^{-1}\mathbf{1}$, the linear stability function of the method.*

- (i) *The method is not $R^{[1]}$ if there exists a value $h \in \mathcal{R} \setminus \{0\}$ such that $(I - hA)$ is invertible and $R(h) = 1$.*
- (ii) *The method is not $R^{[1]}$ if there exists a value $h \in \mathcal{R}$ such that $(hA - I)$ is singular with one-dimensional null-space, the left and right eigenvectors are η and ξ , respectively, and $\eta^T\mathbf{1} = 0$, $\mathbf{b}^T\xi \neq 0$. (These results hold under Technical Condition A, Theorem 2.3.)*
- (iii) *The method is not $R^{[1]}$ if there exists a value $h \in \mathcal{R}$ such that $(hA - I)$ is singular with one-dimensional null-space, the left and right eigenvectors are η and ξ , respectively, and $\eta^T\mathbf{1} \neq 0$, $\mathbf{b}^T\xi = 0$. (This result holds under Technical Condition B, Theorem 2.3.)*
- (iv) *The method is not $R^{[2]}$ if there exists a value $h \in \mathcal{R} \setminus \{0\}$ such that $(I - hA)$ is invertible and $R(h) = -1$.*

Proof. This is a straightforward consequence of Theorems 2.3 and 2.4 with $h_c f'(\hat{y})$ replaced by h . Note that, in parts (i) and (iv), the nondegeneracy condition $\mathbf{b}^T(f'(\hat{y})h_c A - I)^{-2}\mathbf{1} \neq 0$ which ensures bifurcation at a hyperbolic equilibrium for arbitrary nonlinearities in Theorem 2.3(i) and Theorem 2.4 is not required. This is since, to prove irregularity, it is sufficient to provide a specific equation of the form (1.1) for which a spurious solution exists. For parts (i) and (iv) of the corollary we take $h = h_c$ if h_c is positive (respectively, $h = -h_c$ if h_c is negative) and $f(y) \equiv y$ (respectively, $-y$) and the existence of spurious solutions follows automatically. \square

Discussion of Corollary 2.5. It is worth noting that, in the case of a two-stage, second-order Runge–Kutta method the necessary condition (i) for regularity of degree 1 is equivalent to the necessary and sufficient condition $a_{1,1} + a_{2,2} = \frac{1}{2}$ derived by Iserles in (1990). This is proved by Peplow in (1989).

The results (i)–(iii) of Corollary 2.5 are not as complete as those in Hairer, Iserles, and Sanz-Serna (1989) where a recursive algebraic test is described which forms a complete *implicit* (recursive) characterisation of all Runge–Kutta methods which are regular of degree 1. However, the corollary describes a large class of irregular methods *explicitly* and these can be eliminated from the search for regular Runge–Kutta methods.

The condition $\mathbf{b}^T(hA - I)^{-1}\mathbf{1} \neq 0$ for all $h \in \mathcal{R}$ is necessary for regularity of degree 1, by Corollary 2.5(i). This condition is less general than the condition derived in Theorem 5 of Hairer, Iserles, and Sanz-Serna (1989) where the statement for all $h \in \mathcal{R}$ is replaced by for all $h \in \mathcal{C}$. Careful examination of the result in Hairer, Iserles, and Sanz-Serna (1989) shows an implicit definition of regularity which differs from ours in a fundamental way: the method is said to be regular *if every fixed point $\hat{y} \in \mathcal{C}$ of the map (1.3) satisfies $f(\hat{y}) = 0, \dots$* . Thus, spurious complex solutions are

considered in Hairer, Iserles, and Sanz-Serna (1989) and the argument therein is based on consideration of the linear equation $y' = \lambda y$. In a *real system* of equations complex eigenvalues λ appear in complex conjugate pairs and *under perturbation by nonlinear terms* the spurious complex fixed points bifurcate into spurious invariant curves and *not spurious equilibria*—see §5. In contrast, we show that one can expect spurious equilibria generically in nonlinear problems satisfying the mild criteria of Theorem 2.3. For this reason we believe that the class of Runge–Kutta methods exhibiting spurious equilibria that we have constructed in Theorem 2.2 should be treated with special care amongst all the irregular methods which do not satisfy the recursive algebraic test in Hairer, Iserles, and Sanz-Serna (1989). However, we note that, although it is not proved in Hairer, Iserles, and Sanz-Serna (1989), the results therein remain true if the definition of regularity is modified to coincide with the one used here. In the following we generalise the recursive test from Hairer, Iserles, and Sanz-Serna (1989) to the study of $R^{[1,2]}$ Runge–Kutta methods and show that this imposes the order barrier of 2.

We write the RK scheme (2.1)–(2.2) in the standard Butcher form

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

and define the *inflated method* by

$$\begin{array}{cc|cc} & \frac{1}{2}\mathbf{c} & \frac{1}{2}A & O \\ \frac{1}{2}\mathbf{1} + \frac{1}{2}\mathbf{c} & \frac{1}{2}\mathbf{c} & \frac{1}{2}B & \frac{1}{2}A \\ \hline & & \frac{1}{2}\mathbf{b}^T & \frac{1}{2}\mathbf{b}^T \end{array}.$$

Here

$$B = \begin{bmatrix} \mathbf{b}^T \\ \vdots \\ \mathbf{b}^T \end{bmatrix}.$$

Note that the inflated method is simply two steps of the original method, each of length $\frac{1}{2}h$.

LEMMA 2.6. *The RK method is $R^{[1,2]}$ if and only if its inflated method is $R^{[1]}$.*

Proof. Suppose that the inflated method is not $R^{[1]}$. Then there exist \hat{h} and \hat{y} such that $\mathbf{b}^T \mathbf{k} + \mathbf{b}^T \mathbf{l} = \mathbf{0}$, where \mathbf{k} and \mathbf{l} are the stages of the first and the second “half” of the method, and $f(\hat{y}) \neq 0$. There are two possibilities: either $\mathbf{b}^T \mathbf{k} = 0$, and then the original method produces false equilibria (with the step $\frac{1}{2}\hat{h}$) or $\mathbf{b}^T \mathbf{k} \neq 0$, in which case we define $\hat{u} := \hat{y}$, $\hat{v} := \hat{u} + \frac{1}{2}\hat{h}\mathbf{b}^T \mathbf{k}$. Since $\hat{u} = \hat{v} + \frac{1}{2}\hat{h}\mathbf{b}^T \mathbf{l}$, it follows that $\{\hat{u}, \hat{v}\}$ is a nontrivial 2-cycle of the original method (again, with the step $\frac{1}{2}\hat{h}$). To complete the proof, let the original method fail to be $R^{[1,2]}$. Then, proceeding as above, it is easy to construct a false equilibrium for the inflated method; hence it is not $R^{[1]}$. \square

THEOREM 2.7. *The highest attainable order of an $R^{[1,2]}$ RK method is 2.*

Proof. It is proved by Hairer, Iserles, and Sanz-Serna (1989) that a necessary condition for the RK method to be $R^{[1]}$ is that the stability function $R(z)$ must have the form

$$(2.14) \quad R(z) = 1 + \frac{z}{Q(z)},$$

for some polynomial $Q(z)$. By Lemma 2.6, for the RK scheme to be $R^{[1,2]}$ we require that

$$(2.15) \quad R^2(z) = 1 + \frac{2z}{Q^2(z)}.$$

(The factor two follows from consistency considerations.) Combining (2.14) and (2.15) we obtain $Q(z) = 1 - \frac{1}{2}z$ and the result.

3. Linear multistep methods. We now consider solving (1.1) by a general consistent linear multistep method.

DEFINITION 3.1. The k -step linear multistep method (LMM) is defined by

$$(3.1) \quad \sum_{j=0}^k \alpha_j Y_{n+j} = h \sum_{j=0}^k \beta_j f(Y_{n+j}),$$

with fixed stepsize $h > 0$.

Here Y_n approximates the exact solution of (1.1) at nh and we assume that the starting values Y_0, \dots, Y_{k-1} are given.

DEFINITION 3.2. We define the polynomials $\rho(z)$ and $\sigma(z)$ by

$$(3.2) \quad \rho(z) = \sum_{j=0}^k \alpha_j z^j, \quad \sigma(z) = \sum_{j=0}^k \beta_j z^j.$$

We can assume $\alpha_k = 1$ without loss of generality. For consistency we require $\rho(1) = 0$ and $\sigma(1) = \rho'(1)$. Therefore

$$(3.3) \quad \sum_{j=0}^k \alpha_j = 0, \quad \sum_{j=0}^k \beta_j = b,$$

where, for a consistent, zero-stable method, b is a nonzero constant.

The following lemma is needed to prove Theorem 3.4.

LEMMA 3.3. *Suppose we have the $m \times m$ companion matrix [Gantmacher (1959)]*

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{m-1} \end{bmatrix}.$$

Consider the polynomial

$$\phi(x) \equiv x^m + a_{m-1}x^{m-1} + \cdots + a_1x + a_0.$$

Then $(-1)^m \phi(x)$ is the characteristic polynomial of the matrix A . Moreover, $\det(xI - A) = \phi(x)$, $\det(I - A) = \phi(1)$, and $\det(I + A) = (-1)^m \phi(-1)$.

Proof. See Fiedler (1986, p. 158) for the proof. \square

THEOREM 3.4. *Let $f : \mathcal{R} \rightarrow \mathcal{R}$ and let $\hat{y} \in \mathcal{R}$ be a hyperbolic equilibrium of equation (1.1). Let the linear multistep method (3.1) be consistent and zero-stable*

with $\rho(-1), \sigma(-1) \neq 0$. Then period 2 solutions in n bifurcate from the steady solution $Y_n = \hat{y}$ at $h = h_c$ where

$$h_c = \frac{\rho(-1)}{f'(\hat{y})\sigma(-1)}.$$

Important Note. It is straightforward to extend this theorem to the general case of $f : \mathcal{R}^p \rightarrow \mathcal{R}^p$. It is necessary to assume that the Jacobian of f at \hat{y} is nonsingular and has a real, simple eigenvalue η . Then η replaces $f'(\hat{y})$ in the definition of h_c . We prove the result in \mathcal{R} merely for simplicity of exposition.

Proof. Let us first rewrite the k -step relation (3.1) as a 1-step vector map on vectors \mathbf{Y} of length k . Let

$$\mathbf{Y}^n = [Y_n, Y_{n+1}, \dots, Y_{n+k-1}]^T,$$

then the LMM can be written as

$$(3.4) \quad h\mathbf{F}(\mathbf{Y}^{n+1}) + \mathbf{Y}^{n+1} = B\mathbf{Y}^n + h\mathbf{G}(\mathbf{Y}^n),$$

where

$$\mathbf{F}(\mathbf{Y}^{n+1}) = [0, 0, \dots, 0, -\beta_k \mathbf{f}(\mathbf{Y}_{n+k})]^T,$$

$$\mathbf{G}(\mathbf{Y}^n) = \left[0, 0, \dots, 0, \sum_{j=0}^{k-1} \beta_j \mathbf{f}(\mathbf{Y}_{n+j}) \right]^T$$

and

$$(3.5) \quad B = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \dots & -\alpha_{k-1} \end{bmatrix}.$$

We wish to determine values of $h \in \mathcal{R}$ at which period 2 solutions branch from a steady solution of the system of differential equations (1.1). It is shown in Iserles (1990) that steady states of the differential equation (1.1) are inherited by the linear multistep method (3.1). Hence $\mathbf{Y}^n = \hat{\mathbf{Y}} = [\hat{y}, \hat{y}, \dots, \hat{y}]^T$ is a steady solution (*i.e.*, independent of n) of (3.4). Define

$$H(\mathbf{V}, \mathbf{W}; h) := h\mathbf{F}(\mathbf{V}) + \mathbf{V} - B\mathbf{W} - h\mathbf{G}(\mathbf{W}).$$

A period 2 solution of (3.4) is a pair $\{\mathbf{V}, \mathbf{W}\}$, with $\mathbf{V} \neq \mathbf{W}$ satisfying

$$H(\mathbf{V}, \mathbf{W}; h) = 0; \quad H(\mathbf{W}, \mathbf{V}; h) = 0.$$

Note that $\mathbf{V} = \mathbf{W} = \hat{\mathbf{Y}}$ is a trivial steady solution. We seek bifurcation from this trivial solution to solutions with $\mathbf{V} \neq \mathbf{W}$. Thus we apply Lemma 2.2. A straightforward calculation shows that

$$H_a(\hat{\mathbf{Y}}, \hat{\mathbf{Y}}; h) + H_b(\hat{\mathbf{Y}}, \hat{\mathbf{Y}}; h) = C,$$

where the matrix C is given by

$$(3.6) \quad \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \\ c_0 & c_1 & c_2 & \cdots & c_{k-2} & c_{k-1} \end{bmatrix}$$

and the elements along the bottom row are defined by

$$\begin{aligned} c_j &= \alpha_j - hf'(\hat{y})\beta_j, & j = 0, \dots, k-2, \\ c_{k-1} &= 1 + \alpha_{k-1} - hf'(\hat{y})\beta_{k-1} - hf'(\hat{y})\beta_k. \end{aligned}$$

Thus it can be deduced from Lemma 3.3, that the determinant of the matrix is given by

$$\begin{aligned} \det C &= 1 + (c_{k-1} - 1) + \cdots + c_1 + c_0 \\ &= \sum_{j=0}^k \alpha_j - hf'(\hat{y}) \sum_{j=0}^k \beta_j. \end{aligned}$$

From (3.3) we see that that the determinant of the matrix (3.6) is equal to $-bhf'(\hat{y})$. Since the method is zero-stable and the equilibrium is hyperbolic we deduce that (3.6) is nonsingular provided that $h \neq 0$, and we assume this henceforth.

Further calculation shows that

$$H_a(\hat{\mathbf{Y}}, \hat{\mathbf{Y}}; h) - H_b(\hat{\mathbf{Y}}, \hat{\mathbf{Y}}; h) = D,$$

where the matrix D is given by

$$\begin{bmatrix} 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 1 & 0 & \ddots \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 1 \\ d_0 & d_1 & d_2 & \cdots & d_{k-2} & d_{k-1} \end{bmatrix}$$

and the elements along the bottom row are defined by

$$\begin{aligned} d_j &= hf'(\hat{y})\beta_j - \alpha_j, & j = 0, \dots, k-2, \\ d_{k-1} &= 1 - \alpha_{k-1} + hf'(\hat{y})\beta_{k-1} - hf'(\hat{y})\beta_k. \end{aligned}$$

Thus we see from Lemma 3.3 that the determinant of the matrix D is given by

$$\begin{aligned} \det D &= (-1)^k((-1)^k + (d_{k-1} + 1)(-1)^{k-1} + \cdots - d_1 + d_0) \\ &= (-1)^k \sum_{j=0}^k (hf'(\hat{y})\beta_j - \alpha_j) (-1)^j \\ &= (-1)^k (hf'(\hat{y})\sigma(-1) - \rho(-1)). \end{aligned}$$

Consequently, the matrix D is singular at $\mathbf{V} = \mathbf{W} = \hat{\mathbf{Y}}$ and $h = h_c$. By the assumptions of the theorem, we deduce that this value of h is defined. Also, $h_c \neq 0$ so

our assumption that C was invertible is justified. To establish that bifurcation occurs it is only necessary to show that h_c is a simple eigenvalue, in the sense of the “range” condition.

Examination of D shows that at $h = h_c$ it has null-space spanned by

$$\xi = [1, -1, 1, -1, \dots]^T.$$

Thus we must establish that

$$(3.7) \quad D\mathbf{v} = (d\mathbf{F} + d\mathbf{G})\xi$$

is not solvable when evaluated at $h = h_c$. Note that

$$(d\mathbf{F} + d\mathbf{G})\xi = [0, \dots, 0, f'(\hat{y})\sigma(-1)]^T.$$

Inverting D row by row and letting v_0 denote the first element of v we obtain, from the last row,

$$v_0 \sum_{j=0}^{k-1} (-1)^j d_j = f'(\hat{y})\sigma(-1),$$

where the d_j 's are defined as above and evaluated at $h = h_c$. Simplifying we obtain

$$v_0(h_c f'(\hat{y})\sigma(-1) - \rho(-1)) = f'(\hat{y})\sigma(-1).$$

By definition of h_c the left-hand side is zero. By assumption the right-hand side is nonzero. Hence (3.7) does not have a solution and h_c is simple. This completes the proof. \square

Discussion of Theorem 3.4. The papers of Griffiths and Mitchell (1988), Mitchell and Schoombie (1989), Schoombie and Mitchell (1989), Sleeman et al. (1988), and Stuart (1989b) all contain constructions of period 2 solutions in numerical methods for ODEs with hyperbolic equilibria and parabolic partial differential equations (PDEs) with hyperbolic equilibria which are retained under spatial discretisation; Euler timestepping is employed for which $\rho(-1)$ and $\sigma(-1) \neq 0$, and hence the existence of period 2 solutions can be predicted by Theorem 3.4. The special case of Theorem 3.4 for one-step LMMs is proved in Stuart and Peelow (1990).

It is a straightforward corollary of Theorem 3.4 that the method is not regular of degree 2. A different (and substantially shorter) proof of this corollary, based on linear numerical-analytic, rather than nonlinear dynamical, considerations, is presented in the companion paper Iserles and Stuart (1990). Rather surprisingly we can obtain a converse of this result: the assumption $\sigma(-1) = 0$ implies regularity of degree 2. Thus the class of LMMs satisfying $\rho(-1) \neq 0$ and $\sigma(-1) = 0$ is $R^{[1,2]}$ (since all LMMs are $R^{[1]}$; see Iserles (1990)). This suggests that the class of LMMs with $\sigma(-1) = 0$ is a favourable class for long-time simulation of differential equations. Further examination of this class of LMMs is undertaken in Iserles and Stuart (1990).

THEOREM 3.5. *Consider a consistent, irreducible, zero-stable linear multistep method of the form (3.1), satisfying $\rho(-1) \neq 0$. The method is $R^{[2]}$ if and only if $\sigma(-1) = 0$.*

Proof. The only if part follows directly from Theorem 3.4 since we show that, if $\sigma(-1) \neq 0$, (3.1) has period 2 solutions under appropriate assumptions on f . The if

part follows from algebraic manipulations: let $\sigma(-1) = 0$. Period 2 solutions of (3.1) satisfy

$$h\mathbf{F}(\mathbf{V}) + \mathbf{V} = B\mathbf{W} + h\mathbf{G}(\mathbf{W}),$$

$$h\mathbf{F}(\mathbf{W}) + \mathbf{W} = B\mathbf{V} + h\mathbf{G}(\mathbf{V}),$$

where we require $\mathbf{V} \neq \mathbf{W}$. Furthermore, we know a lot about the properties of \mathbf{V} and \mathbf{W} . Specifically,

$$(3.8) \quad \begin{aligned} \mathbf{V} &= [v_0, \dots, v_{k-1}]^T, \\ \mathbf{W} &= [w_0, \dots, w_{k-1}]^T, \end{aligned}$$

where

$$(3.9) \quad \begin{aligned} v_j &= v \text{ for } j \text{ even, } w \text{ for } j \text{ odd,} \\ w_j &= w \text{ for } j \text{ even, } v \text{ for } j \text{ odd,} \end{aligned}$$

for two distinct numbers $v, w \in \mathcal{R}$. Subtracting, we obtain

$$(3.10) \quad (I + B)(\mathbf{W} - \mathbf{V}) = h[\mathbf{F}(\mathbf{V}) - \mathbf{F}(\mathbf{W}) + \mathbf{G}(\mathbf{V}) - \mathbf{G}(\mathbf{W})],$$

where B is defined by (3.5). Let us examine the right-hand side of (3.10). All components are zero except for the last which, from (3.8), equals

$$-h\beta_k f(v_{k-1}) + h\beta_k f(w_{k-1}) + h \sum_{j=0}^{k-1} \beta_j f(v_j) - h \sum_{j=0}^{k-1} \beta_j f(w_j).$$

Using (3.9) this simplifies to

$$h\sigma(-1)[f(v) - f(w)].$$

Hence we have, from (3.10),

$$(I + B)(\mathbf{W} - \mathbf{V}) = \mathbf{0},$$

since $\sigma(-1) = 0$. Using Lemma 3.3 we deduce that $I + B$ is invertible since $\rho(-1) \neq 0$ for an irreducible method. Hence we obtain $\mathbf{V} = \mathbf{W}$. This violates the basic requirement of period 2 solutions and the result follows. \square

THEOREM 3.6. *Consider a consistent linear multistep method of the form (3.1). The method is not $R^{[2]}$ if $\rho(-1) = 0$.*

Proof. Consider an equation of the form (1.1) where $f(y)$ has two zeros \hat{y} and \bar{y} . We show that it is possible to construct a period 2 solution of (3.1) if $\rho(-1) = 0$ and hence that the method cannot be regular of degree 2. Set

$$Y_n = \frac{\hat{y} + \bar{y}}{2} + \frac{\hat{y} - \bar{y}}{2}(-1)^n.$$

This period 2 solution satisfies (3.1): since Y_n takes on the values \hat{y} and \bar{y} at alternate steps, the right-hand side of (3.1) is annihilated. Since the method is consistent and $\rho(1) = 0$ the constant term satisfies the left-hand side of (3.1). Since $\rho(-1) = 0$ the oscillating term satisfies the left-hand side of (3.1). This completes the proof. \square

Discussion of Theorem 3.6. Note that the solution constructed is reminiscent of the *nondiagonal* solutions of the leapfrog scheme described by Sanz-Serna (1985). There is an important distinction between the period 2 solutions constructed in Theorem 3.4 ($\rho(-1) \neq 0$) and in Theorem 3.6 ($\rho(-1) = 0$). Those in Theorem 3.4 are of an intrinsically nonlinear nature and depend in a complicated way upon h ; see Griffiths and Mitchell (1988) and Stuart and Peplow (1990) for numerical examples. Those in Theorem 3.6 are essentially linear: they are independent of h and are extended to the nonlinear case $h \neq 0$ by a special choice which annihilates the effect of nonlinearity. The paper of Sleeman et al. (1988) contains an analysis of period 2 solutions in the leapfrog discretisation (for which $\rho(-1) = 0$) of the logistic equation.

4. Predictor-corrector methods. We consider the solution of equation (1.1) by a simple predictor-corrector algorithm comprising of a single Euler predictor and M corrector steps with the trapezoidal rule. As in previous sections we employ bifurcation analysis. A similar bifurcation analysis can be developed to elucidate the structure of spurious solutions for general Adams–Bashforth/Adams–Moulton predictor-corrector algorithms in a PE(CE) ^{M} implementation.

DEFINITION 4.1. We define our *Predictor-Corrector* method as

$$(4.1) \quad Y_{n+1}^{(0)} = Y_n + hf(Y_n),$$

$$(4.2) \quad Y_{n+1}^{k+1} = Y_n + \frac{h}{2} (f(Y_n) + f(Y_{n+1}^k)), \quad k = 0, \dots, M - 1,$$

$$(4.3) \quad Y_{n+1} = Y_{n+1}^M.$$

The following lemma is needed to prove Theorem 4.3.

LEMMA 4.2. *The recurrence relation*

$$A_k = \left(1 + \frac{x}{2}\right) \left(\frac{x}{2}\right)^{k-1} + A_{k-1}, \quad k = 2, \dots, M$$

has the solution

$$A_M = \left(1 + \frac{x}{2}\right) \sum_{j=1}^{M-1} \left(\frac{x}{2}\right)^j + A_1.$$

Proof. The proof follows in a straightforward fashion by induction. □

THEOREM 4.3. *Let $f : \mathcal{R} \rightarrow \mathcal{R}$ and let \hat{y} be a hyperbolic equilibrium of (1.1). Set*

$$h_c = \frac{-2}{f'(\hat{y})}.$$

The following results are valid for the predictor-corrector method (4.1)–(4.3) :

- (i) *For M odd a spurious steady solution bifurcates from the genuine steady solution $Y_n = \hat{y}$ at $h = h_c$.*
- (ii) *For M even a period 2 solution bifurcates from the genuine steady solution $Y_n = \hat{y}$ at $h = h_c$.*

Important Note. It is straightforward to extend this theorem to a general $f : \mathcal{R}^p \rightarrow \mathcal{R}^p$. It is necessary to assume that the Jacobian of f at \hat{y} is nonsingular and has a real, simple eigenvalue η . Then η replaces $f'(\hat{y})$ in the definition of h_c . We prove the result in a single dimension merely for simplicity of exposition.

Proof. Throughout both proofs we use the notation $x = hf'(\hat{y})$.

(i) **Spurious steady solutions:** Our aim is to find where steady solutions of the discrete problem bifurcate from a genuine steady solution. A steady (*i.e.*, n independent) solution of (4.1)–(4.3) satisfies

$$(4.4) \quad Y_{n+1}^0 = Y_n + hf(Y_n),$$

$$(4.5) \quad Y_{n+1}^{k+1} = Y_n + \frac{h}{2} (f(Y_n) + f(Y_{n+1}^k)), \quad k = 0, \dots, M - 2,$$

$$(4.6) \quad 0 = \frac{h}{2} (f(Y_n) + f(Y_{n+1}^{M-1})).$$

Note that $Y_n = Y_{n+1}^k = Y_{n+1} = \hat{y}$ satisfies (4.4)–(4.6). We linearise (4.4)–(4.6) about this fixed point of the map $Y_n \mapsto Y_{n+1}$. Let the $(M + 1)$ -dimensional vector of linear parts be denoted by $\xi^T = [\xi_n, \xi_{n+1}^0, \dots, \xi_{n+1}^{M-2}, \xi_{n+1}^{M-1}]$. Thus,

$$\begin{aligned} \xi_{n+1}^0 &= (1 + x)\xi_n, \\ \xi_{n+1}^{k+1} &= \left(1 + \frac{1}{2}x\right) \xi_n + \frac{1}{2}x\xi_{n+1}^k, \quad k = 0, \dots, M - 2, \\ 0 &= \frac{1}{2}x\xi_n + \frac{1}{2}x\xi_{n+1}^{M-1}. \end{aligned}$$

The linear problem can be written in matrix notation as $A\xi = 0$. To study bifurcation, we wish to determine the value of h that renders A singular. We have

$$A = \begin{bmatrix} 1 + x & -1 & 0 & \dots & 0 & 0 \\ 1 + \frac{1}{2}x & \frac{1}{2}x & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & & \ddots & \vdots \\ 1 + \frac{1}{2}x & 0 & \dots & \dots & \frac{1}{2}x & -1 \\ \frac{1}{2}x & 0 & \dots & \dots & 0 & \frac{1}{2}x \end{bmatrix}.$$

Let us define the $k \times k$ matrices A_k by

$$A_k = \begin{bmatrix} 1 + \frac{1}{2}x & -1 & 0 & \dots & 0 & 0 \\ 1 + \frac{1}{2}x & \frac{1}{2}x & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & & \ddots & \vdots \\ 1 + \frac{1}{2}x & 0 & \dots & \dots & \frac{1}{2}x & -1 \\ \frac{1}{2}x & 0 & \dots & \dots & 0 & \frac{1}{2}x \end{bmatrix}.$$

Then $\det A$ satisfies

$$(4.7) \quad \det A = (1 + x) \left(\frac{1}{2}x\right)^M + \det A_M,$$

$$\det A_k = \left(1 + \frac{1}{2}x\right) \left(\frac{1}{2}x\right)^{k-1} + \det A_{k-1}, \quad k = 3, \dots, M,$$

$$\det A_2 = \left(1 + \frac{1}{2}x\right) \frac{1}{2}x + \frac{1}{2}x.$$

From Lemma 4.2 we deduce that

$$\begin{aligned} \det A_M &= \left(1 + \frac{1}{2}x\right) \sum_{j=1}^{M-1} \left(\frac{1}{2}x\right)^j + \frac{1}{2}x \\ &= 2 \sum_{j=1}^M \left(\frac{1}{2}x\right)^j - \left(\frac{1}{2}x\right)^M. \end{aligned}$$

Therefore

$$\begin{aligned} \det A &= x \left(\frac{1}{2}x\right)^M + 2 \sum_{j=1}^M \left(\frac{1}{2}x\right)^j \\ &= 2 \sum_{j=1}^{M+1} \left(\frac{1}{2}x\right)^j \\ &= hf'(\hat{y}) \frac{1 - \left(\frac{1}{2}hf'(\hat{y})\right)^{M+1}}{1 - \frac{1}{2}hf'(\hat{y})}. \end{aligned}$$

Hence

$$\det A = 0 \text{ when } \left(\frac{1}{2}x\right)^{M+1} = 1 \text{ and } \frac{1}{2}x \neq 1.$$

Thus, for M odd, the matrix A is singular at $h = h_c$. We need only show that the eigenvalue is simple in the sense of the “range” condition to establish that bifurcation of spurious steady solutions occurs.

The eigenvalue problem $A\xi = 0$ may be written in the form

$$(4.8) \quad hL_1\mathbf{x} = L_2\mathbf{x}.$$

Here

$$L_1 = \begin{bmatrix} f'(\hat{y}) & 0 & \cdots & 0 \\ \frac{1}{2}f'(\hat{y}) & \frac{1}{2}f'(\hat{y}) & \ddots & 0 \\ \vdots & 0 & \ddots & 0 \\ \frac{1}{2}f'(\hat{y}) & 0 & \cdots & \frac{1}{2}f'(\hat{y}) \end{bmatrix}$$

and

$$L_2 = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & 0 & \vdots \\ \vdots & 0 & \ddots & \ddots & \vdots \\ 1 & 0 & \ddots & \ddots & -1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{bmatrix}.$$

Equation (4.8) is singular at $h = h_c$ at which value the matrix A simplifies to A_c where $A_c = h_cL_1 - L_2$ is given by

$$A_c = \begin{bmatrix} -1 & -1 & 0 & \cdots & 0 \\ 0 & -1 & -1 & 0 & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -1 & -1 \\ -1 & 0 & \cdots & 0 & -1 \end{bmatrix}.$$

By inspection we deduce that the eigenvector $\mathbf{x}_c = [x_0, \dots, x_M]^T$ satisfies $x_k = (-1)^k$. The condition that the eigenvector be simple is that

$$(4.9) \quad A_c \mathbf{y} = h_c L_1 \mathbf{x}_c$$

has no solution. Let $\mathbf{w} = [w_0, \dots, w_M]^T$ and $\mathbf{y} = [y_0, \dots, y_M]^T$. We set $\mathbf{w} = h_c L_1 \mathbf{x}_c$ and obtain $w_k = (-1)^{k+1} - 1$. Thus the first M rows of (4.9) give

$$y_{k+1} + y_k = 1 + (-1)^k$$

for $k = 0, \dots, M - 1$. This recursion has solution

$$(4.10) \quad y_k = a(-1)^k - k(-1)^k + \frac{1}{2}$$

for $k = 0, \dots, M$. The bottom row of (4.9) gives

$$(4.11) \quad y_M + y_0 = 1 + (-1)^M.$$

Substituting (4.10) in (4.11) and using the fact that M is odd yields a contradiction. Hence the eigenvalue of (4.8) is simple and bifurcation occurs. This completes the proof of part (i).

(ii) **Spurious period 2 solutions:** We set $W^k = Y_{2\ell}^k$ and $V^k = Y_{2\ell+1}^k$ for $k = 0, \dots, M$. Then let $\mathbf{W} = [W_0, W_1, \dots, W_M]^T$ and similarly for \mathbf{V} . Then define

$$(4.12) \quad H(\mathbf{V}, \mathbf{W}; h) := \left(\begin{array}{l} V^0 - W^M - hf(W^M) = 0, \\ V^{k+1} - W^M - \frac{h}{2}[f(W^M) + f(V^k)], \quad k = 0, \dots, M - 1. \end{array} \right).$$

Period 2 solutions of (4.1)-(4.3) are then given by solutions \mathbf{V}, \mathbf{W} with $\mathbf{V} \neq \mathbf{W}$ and

$$(4.13) \quad H(\mathbf{V}, \mathbf{W}; h) = 0; \quad H(\mathbf{W}, \mathbf{V}; h) = 0.$$

Let $\hat{\mathbf{Y}} = [\hat{y}, \dots, \hat{y}]$. Note that $\mathbf{V} = \mathbf{W} = \hat{\mathbf{Y}}$ satisfies equations (4.12)-(4.13).

We now apply Lemma 2.2 to seek bifurcations from this solution. It can be shown that

$$H_a(\hat{\mathbf{Y}}, \hat{\mathbf{Y}}; h) + H_b(\hat{\mathbf{Y}}, \hat{\mathbf{Y}}; h) = I - B - C,$$

where the matrices B and C are defined by

$$(4.14) \quad B = \begin{bmatrix} 0 & \dots & 0 & 1 + x \\ \vdots & & \vdots & 1 + \frac{1}{2}x \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 1 + \frac{1}{2}x \end{bmatrix}$$

and

$$(4.15) \quad C = \begin{bmatrix} 0 & \dots & \dots & 0 \\ \frac{1}{2}x & \ddots & 0 \dots & \\ 0 & \ddots & \ddots & \dots \\ 0 & \dots & \frac{1}{2}x & 0 \end{bmatrix}.$$

By moving the last column of $I - B - C$ into the first column we deduce that $\det(I - B - C) = \det A \neq 0$ for M even. Hence $I - B - C$ is invertible for M even. Henceforth, we assume that M is even. By Lemma 2.2, we find that bifurcation can occur where D , given by $D \equiv (I + B - C)$, is singular. Examination shows that $\det D$ satisfies the following recursion:

$$\begin{aligned} \det D &= (1 + x) \left(\frac{x}{2}\right)^M + \det D_M, \\ \det D_k &= \left(1 + \frac{x}{2}\right) \left(\frac{x}{2}\right)^{k-1} + \det D_{k-1}, \quad k = 3, \dots, M, \\ \det D_2 &= 2 + \frac{x}{2} + \left(1 + \frac{x}{2}\right) \frac{x}{2}. \end{aligned}$$

Comparing with the recursion (4.7) we deduce that

$$\det D = \det A + 2 = \frac{4 - 4 \left(\frac{1}{2}hf'(\hat{y})\right)^{M+2}}{2 - hf'(\hat{y})}.$$

Thus D is singular at $h = h_c$ for M even. Let $Null(D) = \text{span}\{\xi\}$. By inspection, it is clear that

$$(4.16) \quad \xi = [1, -1, 1, \dots, -1, 1]^T.$$

To establish that bifurcation actually occurs we need to check the ‘‘range’’ condition in Lemma 2.2. This involves showing that the matrix equation

$$(4.17) \quad D\mathbf{v} = (B_h - C_h)\xi$$

is not solvable at $h = h_c$. At this value of h we obtain

$$D = \begin{bmatrix} 1 & 0 & \dots & -1 \\ 1 & \ddots & 0 & \ddots \\ 0 & \ddots & \ddots & 0 \\ 0 & \dots & 1 & 1 \end{bmatrix}$$

and

$$B_h - C_h = \begin{bmatrix} 0 & \dots & \dots & 2 \\ -1 & \ddots & 0 & 1 \\ 0 & \ddots & \ddots & \vdots \\ 0 & \dots & -1 & 1 \end{bmatrix}.$$

Thus $(B_h - C_h)\xi_v = [x_0, \dots, x_M]^T$ where

$$x_k = (-1)^k + 1.$$

Letting $\mathbf{v} = [v_0, \dots, v_M]^T$ we obtain from the last M rows of (4.17)

$$v_k + v_{k-1} = (-1)^k + 1,$$

for $k = 1, \dots, M$. This recursion has solution

$$(4.18) \quad v_k = a(-1)^k + k(-1)^k + \frac{1}{2}.$$

The first row of (4.17) gives

$$(4.19) \quad v_0 = v_M.$$

Substituting (4.18) in (4.19) yields a contradiction ($M = 0$) and hence the eigenvalue is simple. This completes the proof of (ii). \square

Discussion of Theorem 4.3. The effect of the parity of M on the spurious equilibria of the predictor-corrector map (4.1)–(4.3) was first observed in Iserles (1990) for the specific case of the Riccati equation. Theorem 4.3 provides an explanation and a generalisation of that result. An intuitive explanation of the result is as follows: let $P_M(h)$ be the linear stability function for the method (4.1)–(4.3). Then, if M is odd, $P_M(h_c) = 1$ and a spurious steady solution can be found in the linear problem at $h = h_c$. (Under perturbation by nonlinear terms, this becomes a branch of spurious steady solutions, locally continuous in h .) If M is even, $P_M(h_c) = -1$ and a period 2 solution can be found in the linear problem at $h = h_c$. (Under perturbation by nonlinear terms, this becomes a branch of spurious period 2 solutions, locally continuous in h .)

COROLLARY 4.4. *Consider the predictor-corrector method (4.1)–(4.3). The method is not $R^{[1]}$ if M is odd and it is not $R^{[2]}$ if M is even. \square*

Important Note. For the linear multistep method we were able to show that the necessary condition for regularity of degree 2 derived from bifurcation theory was also sufficient. This is not the case for the predictor-corrector method considered here. We present two examples to illustrate that Corollary 4.4 *cannot* be extended to a necessary and sufficient result. Note that the example concerns a function $f(y)$ which is not C^2 on the whole of \mathcal{R} . This is not crucial to the analysis—a C^2 extension of f can be made to eliminate the singularity—the spurious solutions constructed are then only valid for h sufficiently far from zero.

(i) Consider the case $f(u) = -1/u$. A fixed point of the predictor-corrector method with $M = 2$ satisfies

$$\begin{aligned} v &= u - \frac{h}{u}, \\ w &= u - \frac{h}{2} \left(\frac{1}{u} + \frac{1}{v} \right), \\ 0 &= -\frac{1}{u} - \frac{1}{w}. \end{aligned}$$

A simple calculation affirms that these equations have the solution

$$u^2 = \frac{3 \pm \sqrt{5}}{4} h.$$

Hence the method is not regular of degree 1 for M even.

(ii) Consider the case $f(u) = -1/u$. A period 2 solution of the predictor-corrector method with $M = 1$ satisfies

$$\begin{aligned} W^0 &= V^1 - \frac{h}{V^1}, \\ W^1 &= V^1 - \frac{h}{2} \left(\frac{1}{V^1} + \frac{1}{W^0} \right), \\ V^0 &= W^1 - \frac{h}{W^1}, \\ V^1 &= W^1 - \frac{h}{2} \left(\frac{1}{W^1} + \frac{1}{V^0} \right). \end{aligned}$$

These equations admit solutions of the form $W^0 = -V^0 = a$ and $W^1 = -V^1 = b$ with

$$b^2 = \frac{3 \pm \sqrt{5}}{4}h,$$

and $a = (h - b^2)/b$. Hence the method is not regular of degree 2 for M odd.

5. Hopf regularity. In this section we study the effect of discretisation on systems which undergo Hopf bifurcation. Our arguments here will be purely formal. Specifically, we shall not examine in detail the two degenerate cases of Hopf bifurcation occurring (i) at an eigenvalue with algebraic multiplicity greater than one; and (ii) zero crossing speed for the eigenvalues at criticality. Fully rigorous analysis of these cases requires examination of higher-order nonlinear terms and obscures the essence of the argument which centres on a linear analysis. The arguments can be made rigorous by application of Theorem 3.5.2 in Guckenheimer and Holmes (1983) (which can be generalised to other problems in higher dimensions by means of centre manifold reduction).

Consider (1.1), parameterised by s and written in the form

$$(5.1) \quad \frac{dy}{dt} = A_s y + g_s(y),$$

where the matrix A is continuous in s , the spectrum of A_s includes a complex conjugate pair $\{\lambda_s, \bar{\lambda}_s\}$, with nontrivial imaginary part, $\text{Re } \lambda_0 = 0$, $(d/ds)\text{Re } \lambda_s \neq 0$ at $s = 0$, and $g_s(y) = \mathcal{O}(\|y\|^2)$ for small $\|y\|$. Equation (5.1) undergoes a Hopf bifurcation at $s = 0$, $y = 0$ and a branch of periodic solutions is born. Analogously, a map of the form

$$(5.2) \quad C_s U_{n+1} = D_s U_n + g_s(U_{n+1}, U_n)$$

can undergo a Hopf bifurcation at the values of s at which the eigenvalues μ_s of $\mu_s C_s - D_s$ pass through the unit circle as a complex conjugate pair (provided that $g_s(x, y) = \mathcal{O}(\|x\|^2, \|xy\|, \|y\|^2)$). A branch of invariant curves for the map (5.2) is born.

We consider the application of multistep and Runge–Kutta methods to the Hopf bifurcation problem. Close to an equilibrium, these can always be written in the form (5.2). The motivation for the Definition 1.3 of Hopf regularity is that we would like (5.2) to undergo a Hopf bifurcation if and only if (5.1) does. By considering the matrix A_s in Jordan normal form, this is equivalent to studying the effect of discretisation on the equation $y' = \lambda y$. Note that our arguments, based only on the linear part of the problem, are not sufficient to *prove* that Hopf bifurcation occurs in the map if and only if it occurs in the differential equation. To do this requires certain technical details which, we feel, obscure the discussion.

To introduce ideas we apply the *theta method* to

$$(5.3) \quad \frac{dy}{dt} = \lambda y.$$

This yields

$$(5.4) \quad Y_{n+1} = Y_n + h\lambda\{(1 - \theta)Y_n + \theta Y_{n+1}\}.$$

Here $\theta \in [0, 1]$. Thus the method is a convex linear combination of forward and backward Euler methods. Rearranging we obtain $CY_{n+1} = DY_n$, where $C = 1 - h\lambda\theta$ and $D = 1 + h\lambda(1 - \theta)$. It is well known that the boundary of the linear stability function (values of λh where $|D/C| = 1$) coincides with the imaginary axis if and only if $\theta = \frac{1}{2}$. Thus we have Theorem 5.1.

THEOREM 5.1. *The theta method (5.4) is Hopf regular if and only if $\theta = \frac{1}{2}$.*

Discussion of Theorem 5.1. Brezzi, Ushiki, and Fujii (1984) contains an explicit example of a spurious invariant curve introduced by Euler discretisation of the complex equation $z_t = z(i + s - |z|^2)$. This method is not R^H and it may be shown that: (i) the true Hopf bifurcation in the map occurs for small $\mathcal{O}(h)$ nonzero s whereas, in the differential equation, Hopf bifurcation occurs at $s = 0$; (ii) a spurious invariant curve bifurcates from zero at a spurious Hopf bifurcation point at large s of $\mathcal{O}(1/h)$. This spurious bifurcation is introduced by the discretisation.

We now consider the application of a k -step LMM (ρ, σ) to the Hopf bifurcation problem, assuming that ρ and σ are irreducible, ρ obeys the ‘‘root condition’’ (is zero-stable), and the method is consistent. Let

$$(5.5) \quad T(w, z) := \rho(w) - z\sigma(w)$$

and denote by $S[T]$ the (closed) linear stability domain of the method. Thus,

$$S[T] = \{z \in \mathcal{C} : \text{The polynomial } T(w, z) \text{ obeys the root condition}\}.$$

Hence a necessary condition for Hopf regularity is that, for every $\theta \in [-\pi, \pi]$,

$$T(e^{i\theta}, z) = 0 \quad \Rightarrow \quad z \in i\mathcal{R}.$$

Thus, from (5.5), we require that for every $\theta \in [-\pi, \pi]$

$$\operatorname{Re} \frac{\rho(e^{i\theta})}{\sigma(e^{i\theta})} = 0.$$

The last identity yields at once the following lemma.

LEMMA 5.2. *The LMM (3.1) is Hopf-regular only if*

$$(5.6) \quad \operatorname{Re} \rho(e^{i\theta})\sigma(e^{-i\theta}) \equiv 0$$

identically in θ . \square

THEOREM 5.3. *An $R^{[H]}$ LMM of the form (3.1) is marginally zero-stable, provided that $k \geq 2$.*

Proof. Assume that

$$\begin{aligned} \rho(z) &= z^{k_0} \prod_{j=k_0+1}^k (z - r_j), & k_0 &\geq 0, \\ \sigma(z) &= Cz^{\ell_0} \prod_{j=\ell_0+1}^\ell (z - s_j), & \ell_0 &\geq 0, \end{aligned}$$

where $r_{k_0+1}, \dots, r_k, s_{\ell_0+1}, \dots, s_\ell \neq 0$. Note that $\ell \leq k$ and that irreducibility implies that $k_0\ell_0 = 0$. Thus,

$$\begin{aligned} \rho(z^{-1}) &= C_\rho z^{-k} \prod_{j=k_0+1}^k (z - r_j^{-1}), & C_\rho &:= (-1)^{k-k_0} \prod_{j=k_0+1}^k r_j, \\ \sigma(z^{-1}) &= C_\sigma z^{-\ell} \prod_{j=\ell_0+1}^\ell (z - s_j^{-1}), & C_\sigma &= (-1)^{\ell-\ell_0} C \prod_{j=\ell_0+1}^\ell s_j. \end{aligned}$$

Consequently, equation (5.6) implies that, for all $|z| = 1$,

$$CC_\rho z^{\ell_0-k} \prod_{j=\ell_0+1}^{\ell} (z-s_j) \prod_{j=k_0+1}^k (z-r_j^{-1}) - C_\sigma z^{k_0-\ell} \prod_{j=\ell_0+1}^{\ell} (z-s_j^{-1}) \prod_{j=k_0+1}^k (z-r_j) = 0.$$

We now compare coefficients and factorizations. First we note that $k - \ell + k_0 - \ell_0 = 0$. Assume that $k_0 \geq 1$. Then $\ell_0 = 0$ and $\ell = k + k_0 > k$, which cannot be true. Consequently, $k_0 = 0$ and $\ell_0 = k - \ell$. Moreover, $C_\rho = C_\sigma$.

Irreducibility implies that $s_i \neq r_j$ for all i and j . Thus, it follows that

- (i) For every $j = \ell_0 + 1, \dots, \ell$ there exists $m \in \{\ell_0 + 1, \dots, \ell\} \setminus \{j\}$ such that $s_j s_m = 1$; and
- (ii) For every $j = 1, \dots, k$ there exists $m \in \{1, \dots, k\} \setminus \{j\}$ such that $r_j r_m = 1$.

Note that ρ obeys the root condition. Thus, $|r_j| \leq 1$ for all $j = 1, \dots, k$ and (ii) is equivalent to the condition $|r_j| = 1$ for all $j = 1, \dots, k$. The result follows at once. \square

Theorem 5.3 shows that Hopf regular multistep methods are not very useful in practice. Hence we now consider the effect of discretisation by Runge–Kutta methods on systems undergoing Hopf bifurcation. Consider the method (2.1)–(2.2) applied to the equation (5.3). The method then satisfies

$$Y_{n+1} = R(\lambda h)Y_n,$$

where R is the linear stability function defined in Theorem RK.

As for the LMM we denote by $\mathcal{S}[T]$ the (closed) linear stability domain of the method. Thus

$$\mathcal{S}[T] = \{z \in \mathcal{C} : |R(z)| \leq 1\}.$$

To show that a method is Hopf regular requires establishing that

$$\lambda h \in i\mathcal{R} \iff |R(\lambda h)| = 1.$$

In the following we shall require the matrix M whose entries are defined by

$$(5.7) \quad m_{i,j} = b_i a_{i,j} + b_j a_{j,i} - b_i b_j.$$

This matrix was introduced by Burrage and Butcher (1979). Recall that an algebraically stable Runge–Kutta method satisfies $b_i \geq 0$ and $m_{i,j} \geq 0$ [Burrage and Butcher (1979)]. In Sanz-Serna (1988) it is proved that the Runge–Kutta method (2.1)–(2.2) is symplectic if $M \equiv 0$, and a similar result has been announced by Lasagni (1988). Since the linear problem on the centre manifold governing Hopf bifurcation is symplectic, it is not surprising that the schemes identified by Sanz-Serna and Lasagni are appropriate for the integration of equations undergoing Hopf bifurcations. Such a result is contained in the following theorem. Henceforth we use \bar{a} to denote the complex conjugate of a .

THEOREM 5.4. *The Runge–Kutta method (2.1)–(2.2) with $b_i \geq 0$ for all i is $R^{[H]}$ if $M \equiv 0$.*

Proof. We start by establishing that the RK method (2.1)–(2.2) applied to (5.3) satisfies

$$(5.8) \quad |Y_{n+1}|^2 = |Y_n|^2 + \frac{2h\text{Re}\lambda}{|\lambda|^2} \sum_{i=1}^k b_i |\xi_i|^2 - h^2 \sum_{i,j=1}^k m_{i,j} \xi_i \bar{\xi}_j.$$

Taking inner products, we obtain from (2.2)

$$|Y_{n+1}|^2 = |Y_n|^2 + h \sum_{i=1}^k b_i (\xi_i \bar{Y}_n + \bar{\xi}_i Y_n) + h^2 \sum_{i,j=1}^k b_i b_j \xi_i \bar{\xi}_j.$$

From (2.1) with $f(y) = \lambda y$ we have that

$$|\xi_i|^2 = \lambda Y_n \bar{\xi}_i + \lambda h \sum_{j=1}^k a_{ij} \bar{\xi}_i \xi_j.$$

Hence

$$\begin{aligned} |Y_{n+1}|^2 &= |Y_n|^2 + \frac{h}{\lambda} \sum_{i=1}^k b_i |\xi_i|^2 - h^2 \sum_{i,j=1}^k b_i a_{ij} \xi_i \bar{\xi}_j \\ (5.9) \quad &+ \frac{h}{\lambda} \sum_{i=1}^k b_i |\xi_i|^2 - h^2 \sum_{i,j=1}^k b_i a_{ij} \bar{\xi}_i \xi_j + h^2 \sum_{i,j=1}^k b_i b_j \xi_i \bar{\xi}_j. \end{aligned}$$

Collecting terms and using (5.7) we obtain equation (5.8) as required. Now assume that $M \equiv 0$. Then (5.9) gives

$$|Y_{n+1}|^2 = |Y_n|^2 + \frac{2h \operatorname{Re} \lambda}{|\lambda|^2} \sum_{i=1}^k b_i |\xi_i|^2.$$

If $\operatorname{Re} \lambda = 0$ then $|Y_{n+1}|^2 = |Y_n|^2$ so that $|R(\lambda h)| = 1$. Conversely, if $|R(\lambda h)| = 1$ then $|Y_{n+1}|^2 = |Y_n|^2$ and, since the $b_i \geq 0$, this implies that $\operatorname{Re} \lambda = 0$. Thus the method is Hopf regular. \square

A natural question to ask is whether Theorem 5.4 can be extended to an “if and only if” result. Unfortunately, this is not possible: from Theorem 5.1 the trapezoidal rule is Hopf regular, but it does not satisfy $M \equiv 0$. The problem is that the matrix M is, in some ways, an unnatural object to study in the context of Hopf bifurcation. Hopf bifurcation is governed by a linear differential equation and there are many methods, which when applied to linear problems are identical, but which have different M matrices. For example, the implicit midpoint rule ($M \equiv 0$) and the trapezoidal rule ($M \neq 0$) are identical when applied to linear problems. For linear problems a Runge–Kutta method is determined by its stability function $R(z)$ and it is to this that we now turn our attention.

DEFINITION 5.5. The Runge–Kutta method (2.1)–(2.2) is said to be *symmetric* if $R(z)R(-z) = 1$ for all $z \in \mathcal{C}$.

THEOREM 5.6. *Subject to consistency, the Runge–Kutta method (2.1)–(2.2) is $R^{[n]}$ if and only if the linear stability function $R(z)$ is symmetric and A -stable.*

Proof. Symmetry implies that $i\mathcal{R} \subseteq \partial\mathcal{S}[T]$. A -stability means that no points of $\mathcal{S}[T]$ may reside in the left half plane and symmetry excludes such points in the right half plane. Hence, A -stability and symmetry imply Hopf regularity.

Next, we prove that Hopf regularity implies both A -stability and symmetry. Since $R(it)R(-it) = |R(it)| \equiv 1$, symmetry can be deduced at once by analytic continuation. To prove A -stability we suppose that the method is not A -stable, hence there exists $z_0, \operatorname{Re}(z_0)$, such that $|R(z_0)| > 1$. Because of consistency, for sufficiently small and real $\varepsilon > 0$ the point $-\varepsilon$ lies in $\mathcal{S}[T]$, since $R(-\varepsilon) = 1 - \varepsilon + \mathcal{O}(\varepsilon^2)$. Consider the

line segment, Γ say, linking $-\varepsilon$ with z_0 . If R is continuous along Γ then there exists on Γ a point, z_1 say, such that $|R(z_1)| = 1$. Thus, $z_1 \in \partial\mathcal{S}[T]$, $\text{Re}(z_0)$, contradicting the assumption that the method is Hopf-regular. Finally, if R fails to be continuous along Γ then it must have a pole there. Since poles of rational functions are isolated, it follows that, travelling from $-\varepsilon$ along Γ , there will be point, z_1^* say, that is reached *before* the first pole and such that $|R(z_1^*)| > 1$. We replace z_1 by z_1^* and repeat the proof. Either way, we reach a contradiction to the method being Hopf regular. \square

Acknowledgments. We thank J. M. Sanz-Serna for pointing out an error in an early version of Theorem 5.4 and for suggesting its resolution, Theorem 5.6. We also thank T. Eirola for helpful discussions and an anonymous referee for his careful reading of the manuscript.

REFERENCES

- F. BREZZI, S. USHIKI, AND H. FUJII (1984), *Real and ghost bifurcations dynamics in difference schemes for ordinary differential equations*, in Numerical Methods for Bifurcation Problems, T. Kupper, et al., eds, Birkhauser-Verlag, Boston.
- K. BURRAGE AND J. C. BUTCHER (1979), *Stability criteria for implicit Runge–Kutta methods*, SIAM J. Numer. Anal., 19, pp. 46–57.
- S. -N. CHOW AND J. HALE (1982), *Methods of Bifurcation Theory*, Springer-Verlag, Berlin.
- M. G. CRANDALL AND P. H. RABINOWITZ (1971), *Bifurcation from simple eigenvalues*, J. Funct. Anal., 8, pp. 321–340.
- G. DAHLQUIST (1978), *G-stability is equivalent to A-stability*, BIT, 18, pp. 384–401.
- K. DEKKER AND J. G. VERWER (1984), *Stability of Runge–Kutta Methods for Stiff Nonlinear Equations*, North-Holland, Amsterdam.
- M. FIEDLER (1986), *Special Matrices and Their Applications in Numerical Mathematics*, Nijhoff, Lancaster.
- F. R. GANTMACHER (1959), *The Theory of Matrices*, Chelsea, New York.
- D. F. GRIFFITHS AND A. R. MITCHELL (1988), *Stable periodic bifurcations of an explicit discretisation of a nonlinear partial differential equation in reaction diffusion*, IMA J. Numer. Anal., 8, pp. 435–454.
- D. F. GRIFFITHS, P. K. SWEBY, AND H. C. YEE (1990), *Spurious steady state solutions of explicit Runge–Kutta schemes*, IMA J. Numer. Anal., to appear.
- J. GUCKENHEIMER AND P. HOLMES (1983), *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields*, Applied Mathematical Sciences 42, Springer-Verlag, New York.
- E. HAIRER, S. P. NØRSETT, AND G. WANNER (1987), *Solving Ordinary Differential Equations I: Nonstiff Problems*, Springer-Verlag, Berlin.
- E. HAIRER, A. ISERLES, AND J. M. SANZ-SERNA (1990), *Equilibria of Runge–Kutta methods*, Numer. Math., 58, pp. 243–254.
- A. ISERLES (1990), *Stability and dynamics of numerical methods for nonlinear ordinary differential equations*, IMA J. Numer. Anal., 10, pp. 1–30.
- A. ISERLES AND A. M. STUART (1990), *A unified approach to spurious solutions introduced by time-discretisation. Part II: BDF-like methods*, IMA J. Numer. Anal., submitted.
- H. B. KELLER (1977), *Numerical solution of bifurcation and nonlinear eigenvalue problems*, in Application of Bifurcation Theory, P. H. Rabinowitz, ed., Academic Press, New York.
- J. D. LAMBERT (1987), *Developments in stability theory for ordinary differential equations*, in State of the Art in Numerical Analysis, A. Iserles and M. J. D. Powell, eds., Oxford University Press, Oxford.
- F. M. LASAGNI (1988), *Canonical Runge–Kutta methods*, J. Appl. Maths. Phys., 39, pp. 952–953.
- E. N. LORENZ (1989), *Computational chaos—a prelude to computational instability*, Phys. D, 35, pp. 299–317.
- A. R. MITCHELL AND S. SCHOOMBIE (1989), *Nonlinear diffusion and stable period 2 solutions of a discrete reaction-diffusion model*, J. Comput. Appl. Math. 25, pp. 363–372.
- A. C. NEWELL (1977), *Finite amplitude instabilities of partial difference equations*, SIAM J. Appl. Math., 33, pp. 133–160.
- A. T. PEPLOW (1989), *Dynamics of Numerical Methods for Initial Value Problems*, M.Sc. Thesis, University of Bath, Bath, U.K.

- J. M. SANZ-SERNA (1985), *Studies in numerical nonlinear stability I. Why do leapfrog schemes go unstable?*, SIAM J. Sci. Statist. Comput., 6, pp. 923–938.
- , (1988), *Runge–Kutta schemes for Hamiltonian systems*, BIT, 28, pp. 877–883.
- S. W. SCHOOMBIE AND A. R. MITCHELL (1989), *Numerical studies of a discretised version of Fisher's equation beyond the linearised stability limit*, University of Dundee Tech. Report, Dundee, U.K.
- B. D. SLEEMAN, D. F. GRIFFITHS, A. R. MITCHELL, AND P. D. SMITH (1988), *Stable periodic solutions in nonlinear difference equations*, SIAM J. Sci. Statist. Comput., 9, pp. 543–557.
- A. M. STUART (1989a), *Nonlinear instability in dissipative finite difference schemes*, SIAM Rev., 31, pp. 191–220.
- , (1989b), *Linear instability implies spurious periodic solutions*, IMA J. Numer. Anal., 9, pp. 465–486.
- , (1990), *The global attractor under discretisation*, in Continuation and Bifurcations: Numerical Techniques and Applications, D. Roose, B. De Dier, and A. Spence, eds., Kluwer, Dordrecht, the Netherlands.
- A. M. STUART AND A. T. PELOW (1991), *The dynamics of the theta method*, SIAM J. Sci. Statist. Comput., 12, pp. 1351–1372.
- P. K. SWEBY, H. C. LEE, AND D. F. GRIFFITHS (1990), *Spurious steady state solutions of explicit Runge–Kutta schemes*, IMA J. Numer. Anal., submitted.