*IMA Journal of Numerical Analysis* (1987) **7**, 1–13

# Stability of Runge–Kutta Methods for Trajectory Problems

G. J. COOPER

*School of Mathematical and Physical Sciences, The University of Sussex, Brighton, BN1 9QH.*

[Received 3 May 1983 and in final revised form 14 July 1986]

A solution of a system of $m$ autonomous differential equations defines a trajectory in $m$-dimensional space and, in particular, may give a closed orbital path. Typical trajectories are described by a model nonlinear problem introduced in this article. For this problem, a trajectory lies on a surface characterized by a real symmetric matrix. It is shown that some Runge–Kutta methods possess a property which ensures that, for this model problem, the numerical solution lies on the same surface as the trajectory. When $m = 2$, the numerical solution lies on the trajectory. This property is related to algebraic stability. A weaker property suffices for normalized differential systems.

## 1. Introduction

A SOLUTION $\boldsymbol{x}$ of the autonomous system of ordinary differential equations

$$\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x}), \qquad \boldsymbol{f}: \mathbb{R}^m \to \mathbb{R}^m, \tag{1.1}$$

defines a curve, or trajectory, in $\mathbb{R}^m$. The function $\boldsymbol{z}$ defined by $\boldsymbol{z}(t) = \boldsymbol{x}(t + \alpha)$, where $\alpha$ is an arbitrary constant, is also a solution of (1.1) and defines the same trajectory. For a given initial value $\boldsymbol{x}(t_0) = \boldsymbol{v}$, each value of the independent variable $t$ specifies a point on the trajectory but, for some problems, the main interest is to trace the trajectory rather than to compute accurately the solution as a function of $t$. This situation arises in the determination of the solution fields of nonlinear equations, a problem which has been discussed, in particular, by Keller (1979) and by Rheinboldt (1980). This situation also occurs directly in the computation of orbits. Steifel & Bettis (1969) and Lambert & McLeod (1979) designed methods suitable for this purpose, and some extensions have been described by Laurie (1980) and by McLeod & Sanz-Serna (1982). These methods, which are of multistep type, were designed primarily for the calculation of circular orbits.

To indicate the type of result of interest here, consider a linear system

$$\boldsymbol{x}' = J(\boldsymbol{x} - \boldsymbol{a})$$

where, for some real nonsingular matrix $S$, the matrix $\bar{J} = SJS^{-1}$ is skew-symmetric. Let $Q = S^{\mathrm{T}}S$. Since

$$[(\boldsymbol{x} - \boldsymbol{a})^{\mathrm{T}} Q(\boldsymbol{x} - \boldsymbol{a})]' = 2[S(\boldsymbol{x} - \boldsymbol{a})]^{\mathrm{T}} \bar{J}[S(\boldsymbol{x} - \boldsymbol{a})]$$

and since $2\boldsymbol{u}^{\mathrm{T}}\bar{J}\boldsymbol{u} = \boldsymbol{u}^{\mathrm{T}}(\bar{J}^{\mathrm{T}} + \bar{J})\boldsymbol{u} = 0$ for all real $\boldsymbol{u}$, it follows that any solution of the system satisfies $(\boldsymbol{x} - \boldsymbol{a})^{\mathrm{T}} Q(\boldsymbol{x} - \boldsymbol{a}) = c$. The constant $c$ is defined by an initial value

$x(t_0) = v$. The trajectory corresponding to this particular solution lies on the surface of dimension $m - 1$ defined by the set of values of $u$ such that

$$(u - a)^T Q(u - a) = c.$$

Since $Q$ is positive definite, this is the surface of an ellipsoid.

Suppose that the trapezoidal rule is applied to $x' = J(x - a)$, with an initial value $x(t_0) = v$ and an arbitrary step length $h$, to obtain $y$, an approximation to $x(t_0 + h)$. The method gives

$$\bar{y} - \bar{v} = \tfrac{1}{2}hJ(\bar{y} + \bar{v}),$$

where $\bar{y} = y - a$ and $\bar{v} = v - a$, and therefore

$$\begin{aligned}(\bar{y} + \bar{v})^T Q(\bar{y} - \bar{v}) &= \tfrac{1}{2}h(\bar{y} + \bar{v})^T QJ(\bar{y} + \bar{v}) \\ &= \tfrac{1}{2}h[S(\bar{y} + \bar{v})]^T \bar{J}[S(\bar{y} + \bar{v})] = 0.\end{aligned}$$

Hence $\bar{y}^T Q\bar{y} = \bar{v}^T Q\bar{v}$, so that $y$ and $v$ lie on the same ellipsoidal surface. When $m = 2$ it follows immediately that the numerical solution lies on the trajectory but this is not true in general. To see this, observe that $y$ lies on the trajectory if and only if there is a $t$ such that $y = x(t)$. This is equivalent to the requirement that, for some $t$,

$$e^{(t-t_0)J} = (I - \tfrac{1}{2}hJ)^{-1}(I + \tfrac{1}{2}hJ).$$

Since $J$ is similar to a real skew-symmetric matrix, it is similar to a diagonal matrix and it follows that this equation can be satisfied only when $J$ has at most one distinct pair of complex (imaginary) conjugate eigenvalues.

Now consider a nonlinear system $x' = f(x)$ and suppose that there is a trajectory such that $(x - a)^T Q(x - a) = c$ where $Q$ is a given positive definite matrix. On this trajectory,

$$[(x - a)^T Q(x - a)]' = 2(x - a)^T Qf(x) = 0.$$

In order to describe methods for this problem, sometimes it is convenient first to transform the independent variable, as proposed by Keller (1979). Let $z(\tau) = x(t)$ be given by

$$z' = g(z) = \frac{1}{\|f(z)\|} f(z), \tag{1.2}$$

where differentiation is with respect to the new independent variable, and where the norm is defined by $\|f\|^2 = f^T Qf$. It is supposed that $f(z) \neq 0$ on that part of the trajectory to be integrated. This is referred to as a normalized system.

Consider the two-step method

$$y_{n+2} = y_n + 2h_n g(y_{n+1}), \tag{1.3}$$

where the step length is given by $h_n = (y_{n+1} - y_n)^T Qg(y_{n+1})$. This method, with $Q = I$, was derived by Lambert & McLeod (1979) for phase-plane problems and extended by Laurie (1980) to problems in $\mathbb{R}^m$. Suppose, for the moment, that $y_n$ and $y_{n+1}$ lie on the trajectory. This implies, in particular, that $\bar{y}_{n+1}^T Qg(y_{n+1}) = 0$, where $\bar{y}_{n+1} = y_{n+1} - \boldsymbol{a}$, and hence

$$\begin{aligned}\bar{y}_{n+2} &= \bar{y}_n + 2[(y_{n+1} - y_n)^{\mathrm{T}} Q g(y_{n+1})] g(y_{n+1}) \\ &= \bar{y}_n - 2[\bar{y}_n^{\mathrm{T}} Q g(y_{n+1})] g(y_{n+1}).\end{aligned}$$

It follows that

$$\bar{y}_{n+2}^{\mathrm{T}} Q \bar{y}_{n+2} = \bar{y}_n^{\mathrm{T}} Q \bar{y}_n - 4[\bar{y}_n^{\mathrm{T}} Q g(y_{n+1})]^2 [1 - g(y_{n+1})^{\mathrm{T}} Q g(y_{n+1})],$$

and, since $\|g(y_{n+1})\| = 1$, we have $\bar{y}_{n+2}^{\mathrm{T}} Q \bar{y}_{n+2} = \bar{y}_n^{\mathrm{T}} Q \bar{y}_n$. Hence $y_{n+2}$ and the trajectory lie on the same ellipsoidal surface $(\boldsymbol{u} - \boldsymbol{a})^{\mathrm{T}} Q(\boldsymbol{u} - \boldsymbol{a}) = c$. In general, $y_{n+2}$ does not lie on the trajectory and it cannot be inferred that

$$\bar{y}_{n+2}^{\mathrm{T}} Q g(y_{n+2}) = 0.$$

That is, even when $y_0$ and $y_1$ lie on the trajectory, the sequence $(y_n)$ may not lie on the surface. (This problem does not arise when $m = 2$.)

To overcome this problem, suppose that

$$(\boldsymbol{u} - \boldsymbol{a})^{\mathrm{T}} Q \boldsymbol{f}(\boldsymbol{u}) = 0 \quad \forall\, \boldsymbol{u} \in \mathbb{R}^m. \tag{1.4}$$

Then any solution of the nonlinear system $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$ must satisfy

$$(\boldsymbol{x} - \boldsymbol{a})^{\mathrm{T}} Q(\boldsymbol{x} - \boldsymbol{a}) = c,$$

where $c$ is defined by an initial value $\boldsymbol{x}(t_0) = \boldsymbol{v}$. Now suppose merely that $y_0$ and $y_1$ lie on the surface $(\boldsymbol{u} - \boldsymbol{a})^{\mathrm{T}} Q(\boldsymbol{u} - \boldsymbol{a}) = c$. Then the argument shows that the sequence $(y_n)$ lies on this surface.

Implicit methods may be obtained also. Suppose again that (1.4) holds, so that each trajectory lies on a surface of the form $(\boldsymbol{u} - \boldsymbol{a})^{\mathrm{T}} Q(\boldsymbol{u} - \boldsymbol{a}) = c$. The trapezoidal rule, applied to the normalized system (1.2), gives

$$\bar{y}_{n+1} - \tfrac{1}{2} h g(y_{n+1}) = \bar{y}_n + \tfrac{1}{2} h g(y_n)$$

and a simple calculation shows that $\bar{y}_{n+1}^{\mathrm{T}} Q \bar{y}_{n+1} = \bar{y}_n^{\mathrm{T}} Q \bar{y}_n$. McLeod & Sanz-Serna (1982) combine the two-step method (1.3) and the trapezoidal rule, with $Q = I$, to yield a predictor–corrector algorithm.

These arguments may be adapted to develop other linear $k$-step methods. In general, the step lengths are not free parameters, but are calculated by forming inner products, and the approach has some other disadvantages. It is not clear how best to proceed for surfaces where $Q$ is merely symmetric (and perhaps positive definite) and the type of surface needs to be known in advance. In practice, a method with $Q = I$ might be used in the hope that, locally, the surface is approximately spherical. This may not be the case throughout a numerical integration. Finally, the computation of starting values may present problems. The starting values need to lie on the surface and they determine the step lengths used throughout the integration. (A procedure for changing the step length is given by McLeod & Sanz-Serna (1982).)

For Runge–Kutta methods the position is rather different. The analysis given in this article shows that, if a method satisfies a particular algebraic property, then the method is suitable for integration on diverse surfaces. The step length is not specified and remains a parameter of the method. Even though this property (which is related to the property of algebraic stability) is restrictive, some methods can be obtained. In particular, the $s$-stage method of order $2s$ possesses the property. A weaker algebraic property suffices for the normalized system (1.2) and additional methods are available, although, in this case, the type of surface needs to be known. It seems to be very difficult to obtain methods which can be implemented efficiently.

## 2. The test problem

Dahlquist (1975) used a nonlinear test problem to examine certain stability properties of linear multistep methods and Butcher (1975) used the corresponding autonomous problem to study the stability of Runge–Kutta methods. Burrage & Butcher (1979) generalized this analysis to include non-autonomous problems and introduced the concept of algebraic stability. More recently, Burrage & Butcher (1980) extended the concept of algebraic stability to include general linear methods and showed that various other stability definitions are contained in this framework. In this article, a similar nonlinear test problem is used to examine the behaviour of Runge–Kutta methods when they are used to determine certain types of trajectory.

The test problem that is considered is a nonlinear system

$$\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x}), \quad \boldsymbol{f}: \mathbb{R}^m \to \mathbb{R}^m; \qquad (\boldsymbol{u}-\boldsymbol{a})^{\mathrm{T}} Q \boldsymbol{f}(\boldsymbol{u}) = 0 \quad \forall\, \boldsymbol{u} \in \mathbb{R}^m, \tag{2.1}$$

where $\boldsymbol{a} \in \mathbb{R}^m$ is given and $Q$ is a real symmetric matrix. It follows that each trajectory lies on a surface of the form $(\boldsymbol{u}-\boldsymbol{a})^{\mathrm{T}} Q(\boldsymbol{u}-\boldsymbol{a}) = c$. The requirement that $(\boldsymbol{u}-\boldsymbol{a})^{\mathrm{T}} Q\boldsymbol{f}(\boldsymbol{u}) = 0$ for all $\boldsymbol{u}$ seems to be necessary to make real progress.

The type of surface is characterized by the signs of the eigenvalues of $Q$ and the case where $Q$ is positive definite is of principal interest. Let $Q$ be positive definite and define a norm on $\mathbb{R}^m$ by $\|\boldsymbol{u}\|^2 = \boldsymbol{u}^{\mathrm{T}} Q \boldsymbol{u}$. In this case, in particular, it is of interest to examine stability concepts which arise in the qualitative theory of ordinary differential equations, since it is reasonable to request that a numerical method has some stability properties similar to those possessed by the test problem (2.1). In the present context, the concept of orbital stability, discussed by Hahn (1967), is of interest. Let $\boldsymbol{x}$ be the solution of the test problem corresponding to a given initial value $\boldsymbol{x}(t_0)$. The solution is orbitally stable if, for each $\varepsilon > 0$, there is a $\delta > 0$ such that

$$\inf_{t > t_0} \|\boldsymbol{z}(\tau) - \boldsymbol{x}(t)\| \leqslant \varepsilon \quad \forall\, \tau > \tau_0,$$

for any solution $\boldsymbol{z}$ of (2.1) which satisfies $\|\boldsymbol{z}(\tau_0) - \boldsymbol{x}(t_0)\| \leqslant \delta$. That is, each point on the trajectory $\boldsymbol{z}$ is close to some point on the trajectory $\boldsymbol{x}$.

It might be supposed that, when $Q$ is positive definite, any solution of (2.1) is orbitally stable. To see that this is not so, consider the problem where

$$\boldsymbol{f}(\boldsymbol{u}) = \begin{bmatrix} -u_2 - u_1 u_3 u_4/\rho \\ u_1 - u_2 u_3 u_4/\rho \\ \rho u_4 \\ 0 \end{bmatrix}, \qquad \rho = (u_1^2 + u_2^2)^{\frac{1}{2}}.$$

Since $\boldsymbol{u}^T\boldsymbol{f}(\boldsymbol{u})=0$, each trajectory of $\boldsymbol{x}'=\boldsymbol{f}(\boldsymbol{x})$ lies on the surface of a ball in $\mathbb{R}^4$, centred at the origin. The system has the solution

$$\begin{aligned} x_1(t) &= (\cos at + b\sin at)(\alpha\cos t+\beta\sin t),\\ x_2(t) &= (\cos at + b\sin at)(\alpha\sin t-\beta\cos t),\\ x_3(t) &= (\alpha^2+\beta^2)^{\frac{1}{2}}(\sin at - b\cos at),\\ x_4(t) &= a, \end{aligned}$$

and the particular solution defined by $\boldsymbol{x}(0)=[1, 0, 0, 0]^T$ lies on the unit circle in the $(x_1, x_2)$ plane. On the other hand, a solution defined by

$$\boldsymbol{z}(0)=[\alpha, \beta, 0, a]^T$$

has a component $z_3$ which has a maximum value $(\alpha^2+\beta^2)^{\frac{1}{2}}$. The initial values may be chosen so that both $\|\boldsymbol{z}(0)-\boldsymbol{x}(0)\|$ and $|\alpha^2+\beta^2-1|$ are arbitrarily small, and even so that $\boldsymbol{z}$ also lies on the surface of the unit ball. This implies that the solution $\boldsymbol{x}$ is not orbitally stable.

It is possible to impose on the test problem additional constraints that guarantee orbital stability. In practice, the more obvious constraints also imply Lyapunov stability, and this is an appreciably stronger property than desired. An alternative procedure, adopted here, is to consider a stability property which is weaker than orbital stability. This property, treated by Hahn (1967), deals with the stability of the set of trajectories lying on a given surface. Let $Q$ be positive definite and let $\mathscr{S}(c)$ be the surface defined by $(\boldsymbol{u}-\boldsymbol{a})^TQ(\boldsymbol{u}-\boldsymbol{a})=c$. The distance of a point $\boldsymbol{y}$ from the surface is defined by

$$d(\boldsymbol{y}, \mathscr{S}(c)) = \inf_{\boldsymbol{u}\in\mathscr{S}(c)} \|\boldsymbol{y}-\boldsymbol{u}\|. \tag{2.2}$$

A surface $\mathscr{S}(c)$ is stable if, for each $\varepsilon>0$, there is a $\delta<0$ such that

$$d(\boldsymbol{z}(t), \mathscr{S}(c)) \leqslant \varepsilon \quad \forall\, t>t_0,$$

for any solution $\boldsymbol{z}$ of (2.1) which satisfies $d(\boldsymbol{z}(t_0), \mathscr{S}(c))\leqslant\delta$. To show that $\mathscr{S}(c)$ is stable for the test problem suppose that $\boldsymbol{z}\in\mathscr{S}(k)$. Then

$$\boldsymbol{u}=\boldsymbol{a}+(c/k)^{\frac{1}{2}}(\boldsymbol{z}-\boldsymbol{a})$$

lies on $\mathscr{S}(c)$ and $\|\boldsymbol{z}-\boldsymbol{u}\|=|k^{\frac{1}{2}}-c^{\frac{1}{2}}|$. Note that, in the example just discussed, the unit circle is not a stable trajectory but the surface of the unit ball is a stable surface.

This stability property is equivalent to orbital stability when $m=2$. Some Runge–Kutta methods possess a similar property, in that computed values lie on the same surface as the initial value. Suppose that $Q$ is positive definite and that the initial value $\boldsymbol{x}(t_0)$ is perturbed. Then each computed value is close to some point on the surface which contains the trajectory through $\boldsymbol{x}(t_0)$. In this article, this is the sense in which numerical methods are considered to be stable. In order to indicate the type of problem for which such methods may be suitable, the numerical property is referred to as orbital stability. A precise definition is given in the next section.

When $Q$ is not positive definite, there is no natural measure, corresponding to (2.2), of distance from a surface. Nevertheless, it is of some interest to know that, when the initial value is perturbed, the computed values lie on a surface of the same type as a surface $\mathcal{S}$ containing the trajectory and 'close' to this surface in the sense that, for any computed value $\boldsymbol{y}$,

$$\inf_{\boldsymbol{u}\in\mathcal{S}} |(\boldsymbol{y}-\boldsymbol{u})^{\mathrm{T}}Q(\boldsymbol{y}-\boldsymbol{u})|$$

is small. The following examples give some indication of the types of surface that might be considered.

Although linear problems suffice to indicate the types of surface that are covered, it seems worthwhile to give one example of a nonlinear problem. Consider the problem given by

$$\boldsymbol{f}(\boldsymbol{x}) = \begin{bmatrix} -1/(x_1+x_2) \\ (2x_1+x_2)/(x_1+x_2)^2 \end{bmatrix}, \qquad Q = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

Since $\boldsymbol{x}^{\mathrm{T}}Q\boldsymbol{f}(\boldsymbol{x}) = 0$ for all $\boldsymbol{x}$ and since $Q$ is positive definite, the trajectories are a family of ellipses $\boldsymbol{x}^{\mathrm{T}}Q\boldsymbol{x} = c$.

The next example is a linear problem $\boldsymbol{x}' = J\boldsymbol{x}$, where

$$SJS^{-1} = \begin{bmatrix} \lambda & 0 \\ 0 & -\lambda \end{bmatrix}, \qquad Q = S^{\mathrm{T}}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}S,$$

for some nonsingular matrix $S$. For this problem, $\boldsymbol{x}^{\mathrm{T}}QJ\boldsymbol{x} = 0$ for all $\boldsymbol{x}$, and the trajectories are the hyperbolas $\boldsymbol{x}^{\mathrm{T}}Q\boldsymbol{x} = c$.

A more elaborate example is a linear system given, in canonical form, by

$$\boldsymbol{x}' = J\boldsymbol{x}, \qquad J = \begin{bmatrix} \lambda & -1 \\ 1 & \lambda \end{bmatrix},$$

whose solutions are logarithmic spirals. To cope with this situation, consider the augmented problem

$$\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix}' = \begin{bmatrix} J & 0 \\ 0 & -J^{\mathrm{T}} \end{bmatrix}\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix}, \qquad Q = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}.$$

For this problem, each trajectory lies on a surface in $\mathbb{R}^4$ of the form $\boldsymbol{u}^{\mathrm{T}}Q\boldsymbol{u} = c$.

## 3. Orbital stability

The concept of orbital stability of a Runge–Kutta method is defined in this section. The aim is to obtain conditions which ensure that, for a given initial value, the numerical solution and the solution of the test problem lie on the same surface. Despite the severity of this requirement, it happens that such conditions can be obtained and that there are methods which satisfy these conditions. The conditions may be weakened for normalized systems.

Consider the system $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$ and suppose that an approximation $\boldsymbol{v}$ to $\boldsymbol{x}(t_0)$ is given. For given $h$, we may compute an approximation $\boldsymbol{y}$ to estimate $\boldsymbol{x}(t_0+h)$, using an $s$-stage Runge–Kutta method defined by the equations

$$\begin{aligned} y_i &= v + h\sum_{j=1}^{s} a_{ij} f(y_j) \quad (i = 1, 2, \ldots, s), \\ y &= v + h\sum_{i=1}^{s} b_i f(y_i). \end{aligned} \tag{3.1}$$

A particular method may be represented by an array

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

where $A$ is a real $s \times s$ matrix $[a_{ij}]$ and $b^{\mathrm{T}} = [b_i]$ is a real row vector, and where $c^{\mathrm{T}} = [c_i]$ is defined by $c = Ae$, with $e^{\mathrm{T}} = [1, 1, \ldots, 1]$. In the following, $B$ denotes the diagonal matrix such that $Be = b$.

It is convenient to express the Runge–Kutta equations in a more compact form using tensor-product notation. The tensor product $M \otimes R$, of an arbitrary $p \times q$ matrix $M$ and an arbitrary $r \times s$ matrix $R$, is defined by

$$M \otimes R = \begin{bmatrix} m_{11}R & m_{12}R & \ldots & m_{1q}R \\ m_{21}R & m_{22}R & \ldots & m_{2q}R \\ \vdots & \vdots & & \vdots \\ m_{p1}R & m_{p2}R & \ldots & m_{pq}R \end{bmatrix}.$$

An account of the properties of this product is given by Lancaster (1969). In particular, $(M \otimes R)^{\mathrm{T}} = M^{\mathrm{T}} \otimes R^{\mathrm{T}}$, and $(M \otimes R)(N \otimes S) = (MN) \otimes (RS)$ for conforming matrices. Using this notation, the method is given by the equations

$$Y = e \otimes v + (A \otimes I)F, \tag{3.2}$$

$$y = v + (b^{\mathrm{T}} \otimes I)F, \tag{3.3}$$

where $I$ is the $m \times m$ identity matrix and where

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_s \end{bmatrix}, \qquad F = h\begin{bmatrix} f(y_1) \\ f(y_2) \\ \vdots \\ f(y_s) \end{bmatrix}.$$

The definition of orbital stability is given next. This is followed immediately by a theorem which gives sufficient algebraic conditions for orbital stability.

DEFINITION 1 Suppose that a Runge–Kutta method is applied to an arbitrary differential system $x' = f(x)$ where

$$(u - a)^{\mathrm{T}} Q f(u) = 0 \quad \forall\, u \in \mathbb{R}^m, \tag{3.4}$$

for some real symmetric matrix $Q$. The method is *orbitally stable* if, for any step length $h$ and any initial value $v$, $(y - a)^{\mathrm{T}} Q (y - a) = (v - a)^{\mathrm{T}} Q (v - a)$.

THEOREM 1 *A Runge–Kutta method is orbitally stable if* $BA + A^{\mathrm{T}}B - bb^{\mathrm{T}} = 0$, *where* $B$ *is the diagonal matrix given by* $Be = b$.

*Proof.* Let $\bar{y} = y - a$ and $\bar{v} = v - a$. Since $Q$ is symmetric, (3.3) gives

$$\bar{y}^T Q \bar{y} - \bar{v}^T Q \bar{v} = [\bar{v}^T + F^T(b \otimes I)]Q[\bar{v} + (b^T \otimes I)F] - \bar{v}^T Q \bar{v},$$

$$\bar{y}^T Q \bar{y} - \bar{v}^T Q \bar{v} = 2F^T(b \otimes Q)\bar{v} + F^T(bb^T \otimes Q)F. \tag{3.5}$$

Let $\bar{Y} = Y - e \otimes a$ so that (3.2) may be written as $\bar{Y} = e \otimes \bar{v} + (A \otimes I)F$. Since $B$ and $Q$ are symmetric,

$$\begin{aligned} F^T(B \otimes Q)\bar{Y} &= F^T[(Be) \otimes Q]\bar{v} + F^T[(BA) \otimes Q]F \\ &= F^T[(Be) \otimes Q]\bar{v} + F^T[(A^T B) \otimes Q]F. \end{aligned}$$

Noting that $Be = b$ and that $F^T(B \otimes Q)\bar{Y} = \bar{Y}^T(B \otimes Q)F$, addition gives

$$F^T[(BA + A^T B) \otimes Q]F = -2F^T(b \otimes Q)\bar{v} + 2\bar{Y}^T(B \otimes Q)F. \tag{3.6}$$

The two results (3.5) and (3.6) may be combined to yield

$$\bar{y}^T Q \bar{y} - \bar{v}^T Q \bar{v} = -F^T[(BA + A^T B - bb^T) \otimes Q]F - 2\bar{Y}^T(B \otimes Q)F. \tag{3.7}$$

Since $BA + A^T B - bb^T = 0$, and since (3.4) gives

$$\bar{Y}^T(B \otimes Q)F = \sum_{i=1}^{s} b_i(y_i - a)^T Q f(y_i) = 0,$$

it follows that $\bar{y}^T Q \bar{y} = \bar{v}^T Q \bar{v}$.

This condition for orbital stability is closely related to the property of algebraic stability. A Runge–Kutta method is algebraically stable if $BA + A^T B - bb^T$ is non-negative definite and if the elements of $b$ are non-negative. There are orbitally stable methods with some elements of $b$ negative but otherwise orbital stability may be regarded as a borderline case of algebraic stability. Methods with stronger stability properties appear to be unsuitable for the computation of trajectories.

Now suppose that $(x - a)^T Q f(x) = 0$ only on the trajectory being computed. It is of interest to know if, nevertheless, $\bar{y}^T Q \bar{y} - \bar{v}^T Q v$ is small for an orbitally stable method. For any Runge–Kutta method, the conditions

$$\sum_{j=1}^{s} a_{ij} c_j^{k-1} = \frac{1}{k} c_i^k \qquad (i = 1, 2, \ldots, s;\ k = 1, 2, \ldots, \eta)$$

hold for some $\eta \geq 1$. Let $v = x(t_0)$ and suppose that $x$ is analytic on $\mathbb{R}$ and that $f$ satisfies a Lipschitz condition on $\mathbb{R}^m$. Then it can be shown that the computed values satisfy

$$y_i = x(t_0 + hc_i) + O(h^{\eta+1}) \qquad (i = 1, 2, \ldots, s),$$

and it follows that $\bar{Y}^T(B \otimes Q)F = O(h^{\eta+2})$. Since $\eta = s$ for a collocation method, this suggests that an orbitally stable collocation method is particularly suitable for trajectory problems. In the next section, it is shown that the only such method is the method of order $2s$.

It seems to be difficult to construct orbitally stable methods that can be implemented efficiently. Partly for this reason, it is of interest to consider normalized differential systems in order to obtain weaker algebraic conditions.

The following definition does not require $Q$ to be positive definite, though this is the likely situation.

DEFINITION 2 Suppose that a Runge–Kutta method is applied to an arbitrary differential system $\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x})$ where

$$(\boldsymbol{u}-\boldsymbol{a})^{\mathrm{T}}Q\boldsymbol{f}(\boldsymbol{u}) = 0, \quad \boldsymbol{f}(\boldsymbol{u})^{\mathrm{T}}Q\boldsymbol{f}(\boldsymbol{u}) = 1, \qquad \forall \boldsymbol{u} \in \mathbb{R}^m, \tag{3.8}$$

for some real symmetric matrix $Q$. The method is *orbitally stable for normalized systems* if, for any step length $h$ and any initial value $\boldsymbol{v}$,

$$(\boldsymbol{y}-\boldsymbol{a})^{\mathrm{T}}Q(\boldsymbol{y}-\boldsymbol{a}) = (\boldsymbol{v}-\boldsymbol{a})^{\mathrm{T}}Q(\boldsymbol{v}-\boldsymbol{a}).$$

To consider the effect of normalization, let $D$ be a diagonal matrix with diagonal elements $d_1, d_2, \ldots, d_s$. Then

$$\boldsymbol{F}^{\mathrm{T}}(D \otimes Q)\boldsymbol{F} = \sum_{i=1}^{s} d_i \boldsymbol{f}(\boldsymbol{y}_i)^{\mathrm{T}}Q\boldsymbol{f}(\boldsymbol{y}_i) = \sum_{i=1}^{s} d_i,$$

and this combined with (3.7) yields the following result.

COROLLARY *A Runge–Kutta method is orbitally stable for normalized systems if* $BA + A^{\mathrm{T}}B - \boldsymbol{b}\boldsymbol{b}^{\mathrm{T}} = D$ *where $D$ is a diagonal matrix such that* $\boldsymbol{e}^{\mathrm{T}}D\boldsymbol{e} = 0$.

## 4. Orbitally stable methods

Some results, concerning the existence of methods, are given in this section. In particular, it is shown that the $s$-stage method of order $2s$ is orbitally stable. It is shown that, for methods of order $p > s$ with $c_1, c_2, \ldots, c_s$ distinct, the algebraic conditions for orbital stability, i.e.

$$BA + A^{\mathrm{T}}B - \boldsymbol{b}\boldsymbol{b}^{\mathrm{T}} = 0, \qquad B\boldsymbol{e} = \boldsymbol{b}, \qquad B \text{ diagonal}, \tag{4.1}$$

and the conditions for normalized systems, i.e.

$$BA + A^{\mathrm{T}}B - \boldsymbol{b}\boldsymbol{b}^{\mathrm{T}} = D, \qquad \boldsymbol{e}^{\mathrm{T}}D\boldsymbol{e} = 0, \qquad D \text{ diagonal}, \tag{4.2}$$

are equivalent. Some negative results, dealing with singly implicit methods, are given. Particular semi-implicit methods are obtained.

No consistent explicit Runge–Kutta method can satisfy conditions (4.2). To see this, observe that if $A$ is strictly lower triangular then the condition $\boldsymbol{e}^{\mathrm{T}}D\boldsymbol{e} = 0$ gives $b_1 = b_2 = \cdots = b_s = 0$. This conflicts with the consistency condition $b_1 + b_2 + \cdots + b_s = 1$.

To investigate the orbital stability of implicit methods, it is convenient to use the notation

$$\sigma(k) = \sum_{i=1}^{s} b_i c_i^{k-1}, \qquad \sigma(k, l) = \sum_{i=1}^{s}\sum_{j=1}^{s} b_i c_i^{k-1} a_{ij} c_j^{l-1}.$$

The following statements, given by Butcher (1964), are required.

$$\boldsymbol{B}(p): \qquad \sigma(k) = \frac{1}{k} \qquad (k = 1, 2, \ldots, p),$$

$$\boldsymbol{C}(\eta): \qquad \sum_{j=1}^{s} a_{ij} c_j^{k-1} = \frac{1}{k} c_i^k \qquad (i = 1, 2, \ldots, s;\ k = 1, 2, \ldots, \eta),$$

$$\boldsymbol{D}(\xi): \qquad \sum_{i=1}^{s} b_i c_i^{k-1} a_{ij} = \frac{1}{k} b_j (1 - c_j^k) \qquad (j = 1, 2, \ldots, s;\ k = 1, 2, \ldots, \xi),$$

$$\boldsymbol{E}(\xi, \eta): \qquad \sigma(k, l) = \frac{1}{l(k+l)} \qquad (k = 1, 2, \ldots, \xi;\ l = 1, 2, \ldots, \eta).$$

Butcher proved that if $\boldsymbol{B}(p)$, $\boldsymbol{C}(\eta)$, and $\boldsymbol{D}(\xi)$, where $p \leqslant \xi + \eta + 1$ and $p \leqslant 2\eta + 2$, then the method is of order $p$. Butcher also proved that, if a method is of order $\xi + \eta$, then $\boldsymbol{B}(\xi + \eta)$ and $\boldsymbol{E}(\xi, \eta)$.

Various results can be deduced from the next theorem. This theorem connects the order conditions with the stability conditions (4.1) and (4.2). This is useful for the construction of methods. The proof is similar to an argument given by Burrage (1978a).

THEOREM 2 *Consider a method of order $p$ with $c_1, c_2, \ldots, c_s$ distinct. Condition* (4.2) *holds if and only if there exist $d_1, d_2, \ldots, d_s$ such that*

$$\sum_{i=1}^{s} d_i c_i^{k-1} = 0 \quad (k = 1, 2, \ldots, p) \tag{4.3}$$

$$\sum_{i=1}^{s} d_i c_i^{l+m-1} = \sigma(l, m) + \sigma(m, l) - \sigma(l)\sigma(m) \quad (l + m > p;\ 1 \leqslant l, m \leqslant s). \tag{4.4}$$

*Proof.* Let $M = D - BA - A^{\mathrm{T}}B + \boldsymbol{b}\boldsymbol{b}^{\mathrm{T}}$ where $D$ is a diagonal matrix such that $\boldsymbol{e}^{\mathrm{T}} D \boldsymbol{e} = d_1 + d_2 + \cdots + d_s = 0$. Let $V$ be the matrix,

$$V = \begin{bmatrix} 1 & c_1 & \ldots & c_1^{s-1} \\ 1 & c_2 & \ldots & c_2^{s-1} \\ \vdots & \vdots & & \vdots \\ 1 & c_s & \ldots & c_s^{s-1} \end{bmatrix}$$

which is nonsingular since $c_1, c_2, \ldots, c_s$ are distinct. The elements of $V^{\mathrm{T}}MV = [r_{lm}]$ are given by

$$r_{lm} = \sum_{i=1}^{s} d_i c_i^{l+m-1} - \sigma(l, m) - \sigma(m, l) + \sigma(l)\sigma(m),$$

and (4.2) holds if and only if $\boldsymbol{e}^{\mathrm{T}} D \boldsymbol{e} = 0$ and $V^{\mathrm{T}}MV = 0$. Since the method is of order $p$, the statements $\boldsymbol{B}(p)$ and $\boldsymbol{E}(\xi, p - \xi)$ $(\xi = 1, 2, \ldots, p - 1)$ give

$$\sigma(l, m) + \sigma(m, l) = \sigma(l)\sigma(m) \quad \text{for } l + m \leqslant p.$$

Hence (4.2) holds if and only if (4.3) and (4.4) hold.

Suppose that $p > s$ and $c_1, c_2, \ldots, c_s$ are distinct. The theorem shows that, if (4.2) holds, then $d_1 = d_2 = \cdots = d_s = 0$. In this case, conditions (4.1) and (4.2) are equivalent.

For the method of order $2s$, the values $c_1, c_2, \dots, c_s$ are distinct and $\boldsymbol{B}(2s)$ and $\boldsymbol{E}(s, s)$. Hence this method, a collocation method, is orbitally stable. The next theorem shows that this is the only collocation method; indeed the only method with either $\boldsymbol{C}(s)$ or $\boldsymbol{D}(s)$, which satisfies the algebraic condition (4.1) for orbital stability. On the other hand, the trapezoidal rule, represented by the array

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

is an example of a collocation method which satisfies condition (4.2) for normalized systems.

THEOREM 3 *If* $\boldsymbol{C}(\eta)$ *or* $\boldsymbol{D}(\eta)$, *a consistent Runge–Kutta method satisfies condition* (4.1) *only if it is of order* $2\eta$.

*Proof.* Consider the matrix $V^{\mathrm{T}}MV = [r_{lm}]$, where now $D = 0$. Either $\boldsymbol{C}(\eta)$ or $\boldsymbol{D}(\eta)$ gives

$$r_{lm} = \sigma(l)\sigma(m) - \left(\frac{1}{l} + \frac{1}{m}\right)\sigma(l+m) \quad (1 \leqslant l, m \leqslant \eta).$$

It is necessary that $V^{\mathrm{T}}MV = 0$ and so, in particular, it is necessary that $r_{lm} = 0$ for $l, m \leqslant \eta$. Since the method is consistent, $\sigma(1) = 1$ and it follows that $\sigma(k) = 1/k$ for $k = 1, 2, \dots, 2\eta$. This is $\boldsymbol{B}(2\eta)$. The equality $BA + A^{\mathrm{T}}B = \boldsymbol{b}\boldsymbol{b}^{\mathrm{T}}$ gives

$$\sum_{i=1}^{s} b_i c_i^{l-1} a_{ij} + \sum_{i=1}^{s} b_j a_{ji} c_i^{l-1} = b_j \sigma(l) \quad (j = 1, 2, \dots, s;\ l = 1, 2, 3, \dots).$$

Since $\boldsymbol{B}(2\eta)$, it follows that either of $\boldsymbol{C}(\eta)$ and $\boldsymbol{D}(\eta)$ implies the other. Hence the method is of order $2\eta$.

Cooper & Butcher (1983) have described an iterative scheme for implementing the $s$-stage method of order $2s$. In the hope that there may be orbitally stable methods that can be implemented in a more efficient manner, it is of interest to examine singly implicit methods, where the coefficient matrix $A$ has one eigenvalue of multiplicity $s$. Suppose, for the moment, that $c_1, c_2, \dots, c_s$ are distinct and that $\boldsymbol{C}(s-1)$ holds. These methods have been investigated by Burrage (1978b) who showed that the maximum attainable order is $s+1$. On the other hand, if such a method satisfies the algebraic conditions for orbital stability, the order must be at least $2s-2$. Hence $s \leqslant 3$.

The case $s = 3$ will be considered in more generality. The cases $s < 3$ are dealt with first. The only one-stage method satisfying (4.1) is the implicit midpoint rule, of order $p = 2$. Suppose that $s = 2$ and $p > 2$. Then it can be shown that the order conditions and condition (4.1) imply that the eigenvalues of $A$ are not real. When $s = p = 2$, singly implicit methods are obtained easily. An example is the semi-implicit method represented by

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

This method is just the implicit midpoint rule applied over two steps of length $\frac{1}{2}h$ and any number of compositions of this rule give a method with the property $BA + A^{\mathrm{T}}B - bb^{\mathrm{T}} = 0$.

The case $s = 3$ is considered now. It is not assumed that $c_1, c_2, \ldots, c_s$ are distinct and $\boldsymbol{C}(s-1)$ is not assumed. There is no loss of generality in assuming that $b_i \neq 0$ $(i = 1, 2, \ldots, s)$, because condition (4.1), i.e.

$$b_i a_{ij} + b_j a_{ji} = b_i b_j \quad (i, j = 1, 2, \ldots, s),$$

shows that, if $b_k = 0$, then $b_i a_{ik} = 0$ $(i = 1, 2, \ldots, s)$ and so the method is equivalent to a method with fewer stages. Let $s = 3$. Then (4.1) gives

$$A = \begin{bmatrix} \frac{1}{2}b_1 & (b_2/b_1)(b_1 - a_{21}) & (b_3/b_1)(b_1 - a_{31}) \\ a_{21} & \frac{1}{2}b_2 & (b_3/b_2)(b_2 - a_{32}) \\ a_{31} & a_{32} & \frac{1}{2}b_3 \end{bmatrix},$$

and $\det(\lambda I - A) = \lambda^3 - \frac{1}{2}\lambda^2 + a_1\lambda - a_0$. Suppose $p \geqslant 3$. Then a lengthy calculation shows that $12a_1 = 24a_0 + 1$ and $b_1 b_2 b_3 a_0 \geqslant 0$. It follows that, if $b_1 b_2 b_3 > 0$, the eigenvalues of $A$ are not all real and, if $b_1 b_2 b_3 < 0$, all eigenvalues cannot be equal. An example of an orbitally stable method with real eigenvalues and $s = p = 3$ is given by the array

$$\begin{array}{c|ccc} \frac{1}{2}a & \frac{1}{2}a & 0 & 0 \\ \frac{3}{2}a & a & \frac{1}{2}a & 0 \\ \frac{1}{2}+a & a & a & \frac{1}{2}-a \\ \hline & a & a & 1-2a \end{array}$$

where $a \approx 1.351207$ is the real root of $6x^3 - 12x^2 + 6x - 1$. It is not known if there are higher-order methods with real eigenvalues.

To obtain singly implicit methods with $s = p = 3$ it is necessary to consider orbital stability for normalized systems. Although the calculations are laborious it happens that the one extra degree of freedom given by (4.2) suffices. In the following array

$$a = \tfrac{1}{2}[b^2 + c^2 + (1 - b - c)^2], \qquad b = \tfrac{1}{2}(1 - 2c)/p(c), \qquad p(c) = 6c^3 - 6c^2 + 1,$$

and $c$ is a real zero of

$$4p(x)^4 - 12x(1-2x)^2 p(x)^2 - 6(1-x)(1-2x)^2 p(x) + 3(1-2x)^3.$$

The array

$$\begin{array}{c|ccc} a & a & 0 & 0 \\ a+b & b & a & 0 \\ a+b+c & b & c & a \\ \hline & b & c & 1-b-c \end{array}$$

represents a method with $s = p = 3$ which is orbitally stable for normalized systems. The root $c \simeq 0.5383640$ gives $b \simeq -0.1945356$ and $a \simeq 0.3791205$.

**Acknowledgements**

The author thanks Dr. F. E. Clifford for drawing his attention to this problem, and the referees for many helpful comments.

REFERENCES

BURRAGE, K. 1978a High order algebraically stable Runge–Kutta methods. *BIT* **18,** 373–383.

BURRAGE, K. 1978b A special family of Runge–Kutta methods for solving stiff differential equations. *BIT* **18,** 22–41.

BURRAGE, D., & BUTCHER, J. C. 1979 Stability criteria for implicit Runge–Kutta methods. *SIAM J. num. Anal.* **16,** 46–57.

BURRAGE, K., & BUTCHER, J. C. 1980 Non-linear stability of a general class of differential equation methods. *BIT* **20,** 185–203.

BUTCHER, J. C. 1975 A stability property of implicit Runge–Kutta methods. *BIT* **15,** 358–361.

BUTCHER, J.C. 1964 Implicit Runge–Kutta processes. *Maths. Comput.* **18,** 50–64.

COOPER, G. J., & BUTCHER, J. C. 1983 An iteration scheme for implicit Runge–Kutta methods. *IMA J. num. Anal.* **3,** 127–140.

DAHLQUIST, G. 1975 On stability and error analysis for stiff non-linear problems. Dept. of Information Processing, Royal Institute of Technology, Stockholm, *Report* NA 75.08.

HAHN, W. 1967 *Stability of Motion.* Pp. 166–172. Berlin: Springer.

KELLER, H. B. 1979 Constructive methods for bifurcation and non-linear eigenvalue problems. *Computing Methods in Applied Sciences and Engineering.* Berlin: Springer.

LAMBERT, J.D., & MCLEOD, R. J. Y. 1980 Numerical methods for phase plane problems in ordinary differential equations. *Numerical Analysis Proceedings* (*Dundee 1979*). Berlin: Springer.

LANCASTER, P. 1969 *Theory of Matrices.* Pp. 255–260. New York: Academic Press.

LAURIE, D. P. 1980 Equispacing numerical methods for trajectory problems. *Proceedings of the Sixth South African Symposium in Numerical Analysis* (*Durban*).

MCLEOD, R. J. Y., & SANZ-SERNA, J. M. 1982 Geometrically derived difference formulae for the numerical integration of trajectory problems. *IMA J. num. Anal.* **2,** 357–370.

RHEINBOLDT, W. C. 1980 Solution fields of nonlinear equations and continuation methods. *SIAM J. num. Anal.* **17,** 221–239.

STEIFEL, E., & BETTIS, D. G. 1969 Stabilization of Cowell's method. *Numer. Math.* **13,** 154–175.