NUMERICAL ANALYSIS PROJECT
MANUSCRIPT NA-92-16

NOVEMBER 1992

# Model Problems in Numerical Stability Theory for Initial Value Problems

by

A.M. Stuart
A.R. Humphries

NUMERICAL ANALYSIS PROJECT
COMPUTER SCIENCE DEPARTMENT
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305

# MODEL PROBLEMS IN NUMERICAL STABILITY THEORY FOR INITIAL VALUE PROBLEMS

A.M. Stuart[1] and A.R. Humphries[2,3]

**Abstract.** In the past numerical stability theory for initial value problems in ordinary differential equations has been dominated by the study of problems with essentially trivial dynamics. Whilst this has resulted in a coherent and self-contained body of knowledge, it has not thoroughly addressed the problems of real interest in applications. Recently there have been a number of studies of numerical stability for wider classes of problems admitting more complicated dynamics. This on-going work is unified and possible directions for future work are outlined. In particular, striking similarities between this new developing stability theory and the classical non-linear stability theory are emphasised.

The classical theories of *A*, B, and algebraic stability for Runge-Kutta methods are briefly reviewed, and it is emphasised that the classes of equations to which these theories apply – linear decay and contractive problems – only admit trivial dynamics. Four other categories of equations – gradient, dissipative, conservative and Hamiltonian systems – are considered. Relationships and differences between the possible dynamics in each category, which range from multiple competing equilibria to fully chaotic solutions, are highlighted and it is stressed that the wide range of possible behaviour allows a large variety of applications. Runge-Kutta schemes which preserve the dynamical structure of the underlying problem are sought, and indications of a strong relationship between the developing stability theory for these new categories and the classical existing stability theory for the older problems are given. Algebraic stability, in particular, is seen to play a central role. The effects of error control are considered, and multi-step methods are discussed briefly. Finally, various open problems are described.

**KEY WORDS: Numerical Stability, Runge-Kutta Methods, Linear Decay, Contractivity, Gradient Systems, Dissipativity, Conservative Systems, Hamiltonian Systems.**

**AMS SUBJECT CLASSIFICATIONS: 34C35, 34D05, 65L07, 65L20**

***Preprint; January 27, 1993***

[1]Division of Applied Mechanics, Durand Building 252, Stanford University, CA94305-4040, USA.
[2]Program in Scientific Computing and Computational Mathematics, Building 460, Stanford University, CA94305-2140, USA.
[3]School of Mathematical Sciences, University of Bath, Bath, Avon BA2 2AY, UK.

## 1. Introduction

Many problems of interest in the physical sciences and engineering require the understanding of dynamical features which evolve over long-time periods. Examples include the process of coarsening in solid phase separation, where metastability causes extremely long time-scales, turbulence in fluid mechanics, where statistical measures (such as Liapunov exponents) require averages over long time intervals, and the simulation of planetary interactions in the solar system. Thus the numerical approximation of evolution equations over long time intervals is of some importance.

For simplicity we concentrate here on the system of ordinary differential equations

$$\frac{du}{dt} = f(u,t), u(0) = u_0, \tag{1.1}$$

where u $\in \mathcal{C}^p$ and $f(\bullet, t) : \mathcal{C}^p \to \mathcal{C}^p$ for each $t \in \mathcal{R}^+$. We will assume that $f$ is, at least, continuously differentiable with respect to its arguments. The large time dynamics of (1.1) can exhibit a variety of behaviour ranging from very simple, such as reaching steady state, through moderately complex periodic or quasi-periodic behaviour, to the extremely complex chaotic behaviour observed in, for example, the Lorenz equations. Throughout the following we will denote the inner product on $\mathcal{C}^p$ by $\langle \bullet, \bullet \rangle$ with corresponding norm $\| \bullet \|$ denoted by $\|u\|^2 = \langle u,$ u). The precise inner-product used will be that which appears in the structural assumptions made on $f$.

A fundamental question in the numerical analysis of initial value problems is to determine how closely, and in what sense, the numerical approximation relates to the underlying continuous problem. If we let $U_n$ denote an approximation to the true solution $u(t_n)$, where $t_n = n\Delta t$ and the time-step At is typically chosen to be small relative to an appropriate time-scale in the problem, then standard analysis on sufficiently smooth problems of the form (1.1) shows that the error satisfies

$$\|u(t_n) - U_n\| \leq c_1 e^{c_2 T} \Delta t^r, \tag{1.2}$$

for $0 \leq n\Delta t \leq T$. Here $r > 0$ is the order of the method and, typically, $c_1$ and $c_2$ are positive constants. Notice that, for fixed $T$, letting At $\to 0$ results in a proof of convergence of the numerical scheme on finite time intervals. However, fixing At and letting $T \to \infty$ gives no error bound; thus standard error analysis tells us nothing about the relationship between the long-time dynamics of the discrete and continuous problems. Understanding the behaviour of algorithms for fixed At as $T \to \infty$ is what we shall term ***numerical stability*** for the purposes of this paper. In contrast to the question of convergence on fixed time intervals, it is necessary to impose structural assumptions on $f(\bullet)$ to make substantial progress with the question of numerical stability. These structural assumptions confer certain dynamical properties on the underlying equations and ***numerical stability is the question of whether, and in what sense, these dynamical properties are inherited by the numerical approximation.***

The purposes of the paper are: 1) to unify the classical and the currently evolving numerical stability theories as far as possible; 2) emphasise the ***extremely restrictive*** nature of the problems covered by classical stability theories and to stress other, more realistic,

categories motivated by applications in science, engineering and the theory of differential equations; 3) to show that there are strong relationships between the classical and modern theories and, in particular, to emphasise the unifying role of algebraic stability; 4) to highlight the importance of interactions between the theory if dynamical systems and the numerical analysis of intitial value problems; 5) to discuss open problems.

For the purposes of this paper it is possible to think of the numerical methods which approximate (1 .1) as mappings of the form

$$U_{n+1} = \phi(U_n; \Delta t). \tag{1.3}$$

We shall only study Runge-Kutta methods in detail here and, for the purposes of this review, it is sufficient to be aware only of the following facts concerning these approximation met hods:

(i) whilst the numerical solution sequence $\{ U_0, U_1, U_2, \ldots \}$ remains in a compact set $\mathcal{B}$ there is At(B) such that the Runge-Kutta method may be though of as a mapping of the form (1.3) for 0 < At < At(B).

(ii) Runge-Kutta methods satisfy a local approximation property which may be expressed as

$$\|\phi(u(t_n); \Delta t) - u(t_{n+1})\| \leq C\Delta t^{r+1}$$

where $u(t_n)$ satsifies (1.1); this approximation property implies an estimate of the form (1.2).

(iii) Runge-Kutta methods depend on certain parameters (see below) which form a matrix $\boldsymbol{A}$ and vector $\boldsymbol{b}$. In particular, the matrices $M$ and $B$ formed from $\boldsymbol{A}$ and $\boldsymbol{b}$ will be important in framing our stability results. The parameters in $\boldsymbol{A}$ and $\boldsymbol{b}$ are generally adjusted to achieve many different, sometimes conflicting, goals. An example is the integer $r$ in (1.2) which depends upon the choice of $\boldsymbol{A}$, $\boldsymbol{b}$. In this paper we shall concentrate on the choices of $\boldsymbol{A}$ and $\boldsymbol{b}$ which ensure important stability properties, in the sense alluded to earlier. We shall not discuss the important question of how these choices interact with other chocies (such as the determination of $r$.)

The notation used for Runge-Kutta methods is now described:

**Runge-Kutta Methods.** Given a sequence of points $t_n = n\Delta t$ and approximations $U_n \approx u(t_n)$ to the solution of (1.1) we define a general L-stage Runge-Kutta Method (RKM) by

$$\eta_i = U_n + \Delta t \sum_{j=1}^{k} a_{ij} f(\eta_j, t_n + c_j \Delta t), \; i = 1, \ldots, k,$$

$$U_{n+1} = U_n + \Delta t \sum_{i=1}^{k} b_i f(\eta_i, t_n + c_j \Delta t), \quad U_0 = u_0.$$

The following notation will be used throughout: let $\boldsymbol{A}$, $\boldsymbol{I}$ denote the $\boldsymbol{k} \times \boldsymbol{k}$ matrices with entries

$$\{A\}_{ij} = a_{ij}, \qquad \{I\}_{ij} = \delta_{ij},$$

let

$$c = [c_1, \ldots, c_k]^T, \qquad \textit{where} \qquad c_i = \sum_{j=1}^{k} a_{ij},$$

let

$$b = [b_1, \ldots, b_k]^T, \qquad 1 = [1, \ldots, 1]^T,$$

let $B$ denote the $k \times k$ matrix

$$B := diag(b_1, b_2, \ldots, b_k) \tag{1.4}$$

and let $M$ denote the $k \times k$ matrix

$$M := BA + A^T B - bb^T. \tag{1.5}$$

We use the notation

$$m_{ij} = \{M\}_{ij} = b_i a_{ij} + b_j a_{ji} - b_i b_j.$$

Note that, assuming the solvability of the equations for the $\eta_i$, the RKM defines a map from $C^p$ into $C^p$. For any given $U_n$, the solvability of the Runge-Kutta equations is ensured for sufficiently small At [5]. However, the question of solvability for a complete sequence $\{U_n\}_{n=0}^{\infty}$ and given arbitrary At and $u_0$ is non-trivial and we will return to it throughout the paper when particular structural assumptions on j(u) allow us to make more detailed comments. However, all general statements about the large n behaviour of the RKMs are made on the implicit assumption that a solution sequence exists. □

Here we discuss stability theory for the numerical solution of (1.1) by Runge-Kutta methods: the behaviour of numerical algorithms is studied for fixed time-steps over arbitrarily long time intervals under structural assumptions on the underlying differential equations which guarantee certain asymptotic properties for large time. Ensuring stability usually boils down to certain constraints on the co-efficients in the matrix $A$ and vector $b$ which define the Runge-Kutta method.

The classical theories of ***A-*** and ***AN-*** stability (for ***linear decay*** problems) and ***B-*** and algebraic stability (for ***contractive nonlinear problems)*** are reviewed with emphasis placed on the implications of the structural assumptions for the dynamics of the underlying equations. It is stressed that the dynamic possibilities are very ***limited*** for linear decay and contractive problems; consequently ***the range of applications is also limited.*** Various other classes of problem are discussed, motivated by real applications. All these classes admit ***very complicated dynamics*** and hence ***the range of application is immense.*** Specifically ***gradient, dissipative, conservative*** and ***Hamiltonian*** equations are considered in turn. (Note that contractive problems are often referred to as dissipative in the numerical analysis literature; this conflicts with the more widely used and widely applicable definition of dissipativity in differential equations [29], which is essentially one of dissipation at large amplitude – we defer to the usage in the theory of differential equations; see section 5.) For most of these problems numerical stability theory is far from complete and is currently developing. Nonet heless, we make it clear that there are striking relationships with the classical theory.

Sections 2-7 go through a sequence of model problems relevant to numerical stability, starting with linear decay and ending with Hamiltonian systems. In section 8 we discuss briefly analogous problems for linear multi-step and one-leg methods. Section 9 contains a brief description of the effect of error control on numerical stability and section 10 contains the conclusions and open problems.

In summary we find the following important role played by the matrices $M$ and $B$ in numerical stability theory; the precise meaning of stability in each case can be found by reference to the appropriate section. The symbols $\geq$ and $> 0$ in the context of matrices mean ***positive semi-definite*** and ***positive definite.***

- Contractive Problems (section 3);

$$M \geq 0,\ B > 0 \Rightarrow \quad \textit{stability.}$$

- Dissipative Gradient Problems (sections 4 and 5);

$$M \geq 0,\ B > 0 \Rightarrow \quad \textit{stability.}$$

- Dissipative Problems (section 5);

$$M \geq 0,\ B > 0 \Rightarrow \quad \textit{stability.}$$

- Conservative Problems (section 6);

$$M \equiv 0 \Rightarrow \quad \textit{stability.}$$

- Liapunov Exponent Calculations (section 6);

$$M \equiv 0 \Rightarrow \quad \textit{stability.}$$

- Hamiltonian Problems (section 7);

$$M \equiv 0 \Rightarrow \quad \textit{stability.}$$

Throughout we illustrate the various categories of equations by considering the following partial differential equation:

**Example** The Ginzburg-Landau equation for a complex function $u(x, t)$ satisfys

$$u_t = (\hat{a} + i\hat{b})u_{xx} - (\hat{c} + i\hat{d})|u|^2 u + \hat{e}u, \quad x \in (0, 1), \tag{1.6}$$

$$u(0, t) = u(1,t),\ u_x(0,t) = u_x(1, t). \tag{1.7}$$

Here $\hat{a}, \hat{b}, \hat{c}, \hat{d}, \hat{e} \in R$. In this context we introduce the complex inner product

$$(u, v) = \int_0^1 Re(u\bar{v})dx$$

and corresponding $L_2$ norm

$$\|u\|^2 = \int_0^1 |u|^2 dx.$$

Provided that $\hat{a}$ and $\hat{c}$ are positive this equation has a unique bounded solution for all time $t \geq 0$ given arbitrary initial data in $L_2$ [49].

Under spatial discretisation this equation yields a system of ordinary differential equations in the form (1.1). Thus all statements about the complex partial differential equation have natural analogues for related systems of ordinary differential equations provided that the spatial discretisation confers those properties from the infinite dimensional problem to the finite dimensional one. For simplicity of exposition we shall discuss (1.6), (1.7) directly as an illustrative example and ignore the (important) issue of appropriate spatial discretisation. □

## 2. Linear Decay

The analysis of the large-time behaviour of numerical methods for initial value problems begins with the study of the linear, constant coefficient, test problem (1.1) together with the assumption of ***linear decay***

$$f(u) = \lambda u, \; Re(\lambda) \leq 0, \; p = 1, \tag{2.1}$$

where u $\in$ C and $p$ is the dimension of the problem. See [14, 19] and the references cited therein. In this section we use the standard norm $\|u\| = u\bar{u}$ on C. The following solution behaviour may be easily established:

**Result** 2.1 ***Any two solutions u(t), w(t) of*** (1.1), (2.1) ***satisfy***

$$\|u(t) - v(t)\| \leq \|u(0) - v(0)\|,$$

***for all*** $t \geq 0$. ***Furthermore, if inequality in (2.1) is strict then***

$$\lim_{t \to \infty} u(t) = 0$$

***for any*** $u \in C$. □

Numerical stability analysis focuses on determining conditions under which the numerical method replicates these properties. This is the motivation behind the following definition [7]:

**Definition** 2.2 A RKM is said to be ***A-stable*** provided that the function

$$S(z) = 1 + zb^T(I - zA)^{-1}1$$

satisfies $|S(z)| < 1$ for all $z : Re(z) < 0$. □

It is worth noting that there are also algebraic characterisations of A-stability; see [41]. Straightforward analysis shows that, for a RKM applied to (1.1), (2.1), $U_{n+1} = S(\Delta t \lambda) U_n$ and hence [7]:

**Result** 2.3 ***Any two solution sequences*** $\{U_n\}_{n=0}^{\infty}$ ***and*** $\{V_n\}_{n=0}^{\infty}$ ***of an A-stable RKM applied to the problem*** (1.1), ***(2.1) satisfy***

$$\|U_{n+1} - V_{n+1}\| \leq \|U_n - V_n\| \tag{2.2}$$

***for all*** $n \geq 0$. ***Furthermore, if the inequality in*** (2.1) ***is strict then***

$$\lim_{n\to\infty} \|U_n\| = 0 \tag{2.3}$$

***for all*** At > 0 ***and any*** $U_0 \in C$. 0

**Remark** For A-stable RKMs applied to (1.1), (2.1) the unique solvability of the defining equations is guaranteed for all At > 0 if $I - zA$ be invertible for any $z = \lambda\Delta t$ in the left-half plane. Typically $I - zA$ will be invertible in the left-half plane since, where it is not, poles occur in the stability function and A-stability cannot hold. However, cancellation of factors in the stability function can lead to methods which are A-stable but not invertible for certain isolated values of $z = \lambda\Delta t$ in the left-half plane; the scheme

$$\eta_1 = U_n + \Delta t f(\eta_1),$$
$$\eta_2 = U_n + 2\Delta t f(\eta_1) - \Delta t f(\eta_2),$$
$$U_{n+1} = U_n + 2\Delta t f(\eta_1) - \Delta t f(\eta_2),$$

has a linear stability function which is equivalent to backward Euler (which is A-stable) but $I - zA$ is non-invertible for $z = \lambda\Delta t = -1$.

It is possible to generalise this theory into a conditional theory where the properties of Result 2.1 are inherited for sufficiently small At. This leads to the following result.

**Result** ***2.4 The region of absolute stability S for a RKM is the open set in the complex plane for which*** $z \in S \leftrightarrow |S(z)| < 1$. ***If*** $z = \lambda\Delta t \in \bar{S}$ ***then any two solution sequences*** $\{U_n\}_{n=0}^{\infty}$ ***and*** $\{V_n\}_{n=0}^{\infty}$ ***of a RKM applied to the problem*** (1.1), ***(2.1) satisfy (2.2) and if*** $z \in S$ ***then (2.3) holds.*** 0

**Remark** Remarks analogous to those following Result 2.3 also apply in this case.

There is an important point to raise about Results 2.3 and 2.4 in the context we are considering: since the problem is linear, conditions for ensuring this correct large time behaviour are ***independent of the amplitude of initial conditions.*** As we shall show, in general, dependence on initial data is a barrier to complete conditional theories for nonlinear problems.

Non-autonomous analogues of (2.1), with $\lambda$ depending on $t$, are considered in [4]. Hence we briefly consider the problem

$$u_t = \lambda(t)u, \quad Re(\lambda(t)) \leq 0, \; p = 1. \tag{2.4}$$

For such problems it is clear that:

**Result** ***2.5 Any two solutions u(t), v(t) of*** (l.l), ***(2.4) satisfy the same conclusions as Result 2.1.*** □

In [4] the notion of an A-stable scheme was generalised to cope with the non-autonomous problem and consequently AN-stability was defined:

**Definition 2.6** Given any RKM, let

$$\Gamma := diag(\gamma_1, \gamma_2, \ldots, \gamma_k)$$

where $Re(\gamma_i) < 0$ and $\gamma_i = \gamma_j$ if $c_i = c_j$. The RKM is said to be ***AN-stable*** if, for all such $\Gamma$, the matrix $I - A\Gamma$ is non-singular and

$$R(\Gamma) := 1 + b^T\Gamma(I - A\Gamma)^{-1}1$$

satisfies $|R(\Gamma)| < 1$. **0**

The motivation for this definition is to ensure that the numerical solution decays on a step-by-step basis, to mimic the behaviour of the differential equation [4]:

**Result 2.7** *Any two solution sequences $\{U_n\}_{n=0}^{\infty}$ and $\{V_n\}_{n=0}^{\infty}$ of an AN-stable RKM applied to the problem (1.1), (2.4) satisfy the conclusions of Result 2.3* □

As we shall see in the next section, this nonautonomous linear theory is best discussed in the context of a general class of nonlinear problems.

## 3. Contractive Nonlinear Problems

Clearly the linear problems of section **2** are very restrictive and naturally attempts were made to study nonlinear problems. The first class of nonlinear problems studied in any systematic way were *contractive* problems (1.1). For simplicity of exposition we will consider the case where (1.1) is autonomous and real for which $f(u, t) \equiv f(u) \in C^1(\mathbb{R}^p, \mathbb{R}^p)$ and the condition for contractivity becomes

$$\langle f(u) - f(v), u - v\rangle \leq 0 \quad \forall u, v \in \mathbb{R}^p, u \neq v. \tag{3.1}$$

This class of problem was introduced by Dahlquist [15, 16] and the motivation was to generalise the error propagation results from linear decay theory. A simple example of an equation satisfying (3.1) is the following:

**Example** For $p = 1$ and $f(u) = -u^3$ we have

$$(f(u) - f(v), u - v) = -(u^2 + uv + v^2)(u - v)^2 = -\frac{1}{2}[(u + v)^2 + u^2 + v^2](u - v)^2 \leq 0. \ 0$$

**Example** Consider equation (1.6), (1.7) with $\hat{b} = \hat{d} = \hat{e} = 0$ and $\hat{a} = \hat{c} = 1$ and $u(x, t) \in \mathbb{R}$. This gives the scalar reaction-diffusion equation

$$u_t = u_{xx} - u^3$$

together with periodic boundary conditions on the unit interval. Then, taking the right-hand side of this equation as $f(u)$ we can show that (3.1) holds, using integration by

parts:

$$\langle u_{xx} - u^3 - v_{xx} + v^3, u - v\rangle = \int_0^1 \{(u - v)(u - v)_{xx} - (u^3 - v^3)(u - v)\}dx$$

$$= -\int_0^1 \{(u_x - v_x)^2 + \frac{1}{2}[(u + v)^2 + u^2 + v^2](u - v)^2\}dx \leq 0.$$

Thus the problem is contractive and satisfies an infinite dimensional analog of (3.1). 0

Throughout we will use the following definition for the distance between a point x $\in \mathbb{R}^p$ and a set $\mathcal{B} \subset \mathbb{R}^p$ :

$$dist(x, \mathcal{B}) = \inf_{y \in \mathcal{B}} \|x - y\|.$$

For problems satisfying (3.1) it may be shown that:

**Result** 3.1 ***Any two solutions u(t), v(t) of*** (1.1), (3.1) ***satisfy***

$$\|u(t) - v(t)\| \leq \|u(0) - v(0)\|,$$

***for all t $\geq$ 0. Furthermore, the set of steady states of the system define a closed convex set $\mathcal{E}$ and, if the inequality (3.1) is strict for all u, v : v $\in \mathcal{E}$, u $\notin \mathcal{E}$ then***

$$\lim_{t \to \infty} dist(u(t), \mathcal{E}) \to 0.$$

***Finally, if the inequality in (3.1) is strict and $\exists \bar{u}$ : f $(\bar{u})$ = 0 then $\bar{u}$ is a unique equilibrium point and***

$$\lim_{t \to \infty} u(t) = ii.$$

**Proof** A calculation shows that

$$\frac{1}{2}\frac{d}{dt}\|u - v\|^2 = \langle u - v, f(u) - f(v)\rangle \leq 0. \tag{3.2}$$

Thus the first result follows.

To prove that the steady states of the system define a convex set it is sufficient to show that any convex combination of zeros of *f* is also a zero of *f*. Let *z* = Xx + (1 $-$ $\lambda$)*y* where *f(x)* = *f(y)* = 0, $\lambda \in (0,1)$ and define $z' = z + \delta f(z), \delta > 0$. Then $z' -$ x = $\delta f(z)$ + $(\lambda - 1)(x -$ y). Now (3.1) implies that

$$\langle f(z'), z' - x\rangle \leq 0$$

and hence

$$\langle f(z'), \delta f(z)\rangle \leq (1 - \lambda)\langle f(z'), x - y\rangle.$$

Similarly

$$\langle f(z'), z' - y\rangle \leq 0$$

implies that

$$\langle f(z'), \delta f(z)\rangle \leq -\lambda\langle f(z'), x - y\rangle.$$

Notice that, since $(1-\lambda)$ and $-\lambda$ have opposite signs, we have $\langle f(z'), \delta f(z)\rangle \leq 0$ which is equivalent to $(f(z+\delta f(z)), f(z)) \leq 0$ since $\delta > 0$; letting $\delta \to 0$ and using the continuity of $f$, we obtain $\|f(z)\|^2 \leq 0$, and thus $f(z) = 0$. Convexity follows. Let $u_i \to u^*$ be such that $f(u_i) = 0$ for each $i$. By the continuity of $f(o)$ it also follows that $f(u^*) = 0$ and hence that $\mathcal{E}$ is closed.

Now assume that (3.1) is strict for v $\in \mathcal{E}$ and u $\notin \mathcal{E}$. Define the set $U$ by

$$U = \{u \in \mathbb{R}^p : r \leq dist(u, \mathcal{E}) \leq R\}.$$

Then, if $u(0) \in U$ it follows that there exists $\bar{u} \in \mathcal{E}$ for which

$$\|u(0) - \bar{u}\| \leq R.$$

Thus it follows that, for all $t \geq 0$,

$$dist(u(t), \mathcal{E}) \leq \|u(t) - \bar{u}\| \leq \|u(0) - \bar{u}\| \leq R. \tag{3.3}$$

Now assume, for the purposes of contradiction, that $dist(u(t), \mathcal{E}) > r$ for all $t \geq 0$. Let

$$\tilde{U} = U \bigcap \{u \in \mathbb{R}^p : \|u - \bar{u}\| \leq R\}$$

and then define

$$\epsilon := \epsilon(r, R) = \inf_{u \in \tilde{U}} \langle f(u), \bar{u} - u\rangle.$$

Note that $\tilde{U}$ is compact since it is formed as the intersection of a compact set with a closed set. Clearly $\epsilon > 0$ since $\tilde{U}$ is compact, and since strict inequality holds in (3.1) as $\bar{u} \in \mathcal{E}$, u $\notin \mathcal{E}$. Thus, by assumption and by (3.3) we have $u(t) \in \tilde{U}$ for all $t \geq 0$ and hence

$$\frac{1}{2}\frac{d}{dt}\|u - \bar{u}\|^2 \leq -\epsilon \ \forall t \geq 0.$$

Hence, as $t \to \infty$

$$\|u - \bar{u}\|^2 \to -00,$$

a contradiction. Thus there exists a time $t^*(r, R)$ for which $dist(u(t), \mathcal{E}) \leq r$. Replacing $R$ by $r$ we deduce from (3.3) that $dist(u(t), \mathcal{E}) \leq r$ for all $t \geq t^*(r, R)$. Since $r$ is arbitrary the result follows.

For the case of strict inequality for all u, v : u $\neq$ v, uniqueness of $\bar{u}$ follows automatically since otherwise we have a contradiction. Thus $\mathcal{E} = \{\bar{u}\}$ and the preceding argument establishes that

$$\lim_{t \to \infty} u(t) = \text{ii. } \square$$

The original motivation for the study of these problems was to generalise the notion of **contractivity** from linear to nonlinear problems since this notion is fundamental in understanding certain hinds of error propagation for numerical methods. However, as a result, the large time behaviour of (1.1), (3.1) is very closely related to that of the model linear problem (1.1), (2.1) (cornpare Results 2.1 and 3.1) and is essentially trivial. Runge-Kutta methods for (1.1), (3.1) were studied in [6, 4]. These studies resulted in the following definitions:

**Definition 3.2** An RKM is said to be ***algebraically stable*** if the matrices $B$ and $M$ defined by (1.4), (1.5) are non-negative definite. An RKM is said to be ***B-stable*** if, when applied to (1.1),(3.1), any two solution sequences $\{U_n\}_{n=0}^{\infty}$, $\{V_n\}_{n=0}^{\infty}$ satisfy

$$\|U_{n+1} - V_{n+1}\| \leq \|U_n - V_n\| \tag{3.4}$$

for any $U_0, V_0 \in \mathbb{R}^p$ and any $\Delta t \geq 0$.

**Example** The simplest example of an algebraically stable scheme is the ***Backward Euler method***

$$U_{n+1} = U_n + \Delta t f(U_{n+1}).$$

There exist arbitrarily high-order schemes which are algebraically stable, but all of them are implicit – that is, they involve the solution of nonlinear equations at each step. 0

Once again, numerical stability is the requirement that a certain qualitative property of the differential equation is inherited by the numerical method; the following result shows that the purely algebraic criterion of [4] and Definition 3.2 is important in this context:

**Result** *3.3 Any two solution sequences $\{U_n\}_{n=0}^{\infty}$ and $\{V_n\}_{n=0}^{\infty}$ of an algebraically stable RKM applied to the problem (1.1), (3.1) satisfy*

$$\|U_{n+1} - V_{n+1}\| \leq \|U_n - V_n\|$$

*for all $n \geq 0$. Hence*

***algebraic stability $\Rightarrow$ B-stability.***

*Furthermore, if the inequality (3.1) is strict for all $u, v : v \in \mathcal{E}, u \notin \mathcal{E}$ then the set of fixed points of the RKM is equivalent to the set $\mathcal{E}$ of equilibrium points for (1.1), (3.1) and*

$$\lim_{n\to\infty} dist(U_n, \mathcal{E}) \to 0.$$

*Finally, if the inequality in (3.1) is strict and $\exists \bar{u} : f(\bar{u}) = 0$ then $\bar{u}$ is a unique equilibrium point of the RKM and*

$$\lim_{n\to\infty} U_n = \bar{u}$$

*for all $\Delta t > 0$ and any $U_0 \in \mathbb{R}^p$.* 0

**Proof** Let the sequence $V_n$ satisfy

$$\xi_i = V_n + \Delta t \sum_{j=1}^{k} a_{ij} f(\xi_j), \quad i = 1, \ldots, k,$$

$$V_{n+1} = V_n + \Delta t \sum_{i=1}^{k} b_i f(\xi_i), \quad U_0 = u_0.$$

We also define

$$D_n = U_n - V_n, \quad E_i = \eta_i - \xi_i, \quad F_i = f(\eta_i) - f(\xi_i).$$

Then

$$D_{n+1} = D_n + \Delta t \sum_{j=1}^{k} b_j F_j$$

and

$$E_i = D_n + \Delta t \sum_{j=1}^{k} a_{ij} F_j$$

and it follows that

$$\|D_{n+1}\|^2 = \|D_n\|^2 + 2\Delta t \sum_{j=1}^{k} b_j \langle D_n, F_j \rangle + \Delta t^2 \sum_{i,j=1}^{s} b_i b_j \langle F_i, F_j \rangle.$$

$$= \|D_n\|^2 + \text{At} \sum_{j=1}^{k} b_j \langle D_n, F_j \rangle + \Delta t \sum_{i=1}^{k} b_i \langle D_n, F_i \rangle + \Delta t^2 \sum_{i,j=1}^{s} b_i b_j \langle F_i, F_j \rangle.$$

Using the fact that

$$\langle D_n, F_i \rangle = \langle E_i, F_i \rangle - \Delta t \sum_{j=1}^{k} a_{ij} \langle F_i, F_j \rangle$$

and that

$$\langle D_n, F_j \rangle = \langle E_j, F_j \rangle - \Delta t \sum_{i=1}^{k} a_{ji} \langle F_i, F_j \rangle$$

we obtain, since the scheme is algebraically stable,

$$\|D_{n+1}\|^2 = \|D_n\|^2 + 2\Delta t \sum_{j=1}^{k} b_j \langle E_j, F_j \rangle - \Delta t^2 \sum_{i,j=1}^{k} m_{ij} \langle F_i, F_j \rangle$$

$$\leq \|D_n\|^2 + 2\Delta t \sum_{j=1}^{k} b_j \langle E_j, F_j \rangle.$$

Thus we have

$$\|U_{n+1} - V_{n+1}\|^2 \leq \|U_n - V_n\|^2 + 2\Delta t \sum_{j=1}^{k} b_j \langle \eta_j - \xi_j, f(\eta_j) - f(\xi_j) \rangle. \tag{3.5}$$

Using (3.1) it follows that

$$\|U_{n+1} - V_{n+1}\| \leq \|U_n - V_n\|$$

and B-stability is established.

Now we assume that (3.1) holds with strict inequality for v $\in \mathcal{E}$ and $u \notin \mathcal{E}$. First we consider the case of a DJ-irreducible method (see remarks after proof). To obtain a contradiction assume that there exists $\bar{w} \notin \mathcal{E}$ which is a fixed point of the Runge-Kutta method and let $\bar{u} \in \mathcal{E}$. Since $f(\bar{u}) = 0$ the Runge-Kutta equations have a solution $\eta_i = \bar{u}, i = 1, \ldots, k$ and $\bar{u}$ is also a fixed point of the Runge-Kutta method. Thus from (3.5), setting $U_n = \bar{u}$ and $V_n = \bar{w}$,

$$\|\bar{u} - \bar{w}\|^2 \leq \|\bar{u} - \bar{w}\|^2 + 2\Delta t \sum_{j=1}^{k} b_j \langle \bar{u} - \xi_j, f(\bar{u}) - f(\xi_j) \rangle. \tag{3.6}$$

In addition it is not possible for all the $\xi_j$ to be contained in $\mathcal{E}$ for, if they were, then $f(u)) = f(\xi_i) = 0$ which implies that $\bar{w} \in \mathcal{E}$ and this is not possible. Hence there exists j such that

$$\langle \bar{u} - \xi_j, f(\bar{u}) - f(\xi_j) \rangle < 0$$

and furthermore, it is known from general theory [28] that D J-irreducible, algebraically stable Runge-Kutta methods satisfy $b_i > 0$ Vi. Thus, from (3.1), (3.6) we have that

$$\|\ddot{u} - \bar{w}\|^2 < \|\bar{u} - \bar{w}\|^2,$$

a contradiction, and hence such a $\bar{w}$ cannot exist. Now note that a DJ-reducible method cannot possess a fixed point $\bar{w} \notin \mathcal{E}$, since $\bar{w}$ would also be a fixed point of the reduced method, and we have just proved that a DJ-irreducible method cannot possess such a fixed point.

We now prove that $\mathcal{E}$ is attracting; the same notation is employed as for the proof of Result 3.1. Since the solution sequence for a DJ-reducible method will be the same as for some DJ-irreducible method, we need only establish the result for DJ-irreducible methods. Let $U_0 \in U$. Then, by (3.1) and (3.5) it follows that there exists $\bar{u} \in \mathcal{E}$ such that

$$dist(U_n, \mathcal{E}) \leq \|U_n - \bar{u}\| \leq \|U_0 - \bar{u}\| \leq R.$$

Assume for the purposes of contradiction that $U_n \in \tilde{U}\ \forall n \geq 0$. Now notice that, if $U_n \notin \mathcal{E}$ then $\exists j : \eta_j \notin \mathcal{E}$ since otherwise $U_n = \eta_i \in \mathcal{E}$. Thus

$$\epsilon(r, R) = \inf_{u \in \tilde{U}} \max_{1 \leq j \leq k} \langle f(\eta_j), \bar{u} - \eta_j \rangle > 0$$

and

$$b_{\min} = \min_{1 \leq j \leq k} b_j > 0.$$

From (3.5) we have that

$$\|U_{n+1} - V_{n+1}\|^2 \leq \|U_n - V_n\|^2 - 2\Delta t b_{min} \max_{1 \leq j \leq k} \langle \bar{u} - \eta_j, f(\eta_j)) \rangle.$$

so that, since $\bar{u} \in \tilde{U}$

$$\|U_{n+1} - \bar{u}\|^2 \leq \|U_n - \bar{u}\|^2 - 2\Delta t b_{min} \epsilon(r, R)\ \forall n \geq 0.$$

Letting n $\rightarrow \infty$ gives a contradiction and hence we deduce that $\exists n^*(r, R)$ for which $dist(U_n, \mathcal{E}) \leq r$. Since $r$ is arbitrary the result follows as for Result 3.1.

Finally assume that (3.1) holds with strict inequality for all u $\neq$ v and that $f(\bar{u}) = 0$. Thus $\bar{u}$ is a fixed point of the RKM [35]. In Result 3.1 we established that $\bar{u}$ is the unique fixed point of (1.1), (3.1) and hence that $\mathcal{E}$ = ii. Applying the previous part of this result to the case where the inequality (3.1) is strict for all u, $v$: v $\in \mathcal{E}$, u $\notin \mathcal{E}$ proves that $\bar{u}$ is the unique fixed point of the RKM. Since $\mathcal{E} = \bar{u}$ the convergence result from the previous case can also be applied to show that $U_n \rightarrow \bar{u}$ as $n \rightarrow \infty$. □

**Remark** (i) The role of algebraic stability in the proof is to enable a certain quadratic form, which is defined by the matrix $M$, to be bounded above when manipulating inequalities and yielding (3.5). This basic idea, and variants on it, will recur throughout the paper.

(ii) The first part of this Theorem, showing that algebraic stability implies B-stability is proved in [4].
(iii) DJ-irreducibility is a technical property defined precisely in [28]. Roughly speaking a method is ***DJ-reducible*** if one or more stages have no effect on the solution. Since such a method can be simplified by deleting the irrelevant stages, DJ-reducible methods are not used in practice.
(iv) In Result 3.3 we have not considered the solubility of the implicit Runge-Kutta equations. The existence of unique solutions under (3.1), for any $U_n$ and any $\Delta t \geq 0$, has been established for many classes of algebraically stable methods including those based on Gauss-Legendre quadrature, for which $M \equiv 0$, the Radau IA, IIA and Lobatto IIIC methods; see [19], [28].
(v) Notice that, in the course of the proof, we ruled out the existence of spurious fixed points of the algebraically stable RKM applied to contractive problems. The possible existence of such spurious fixed points was first observed in [35] and, in [27], the class of ***regular*** methods which do not possess spurious fixed points was identified and various order barriers established. Result 3.3 shows that a wider class of methods exists which do not have spurious fixed points, provided that structure is imposed on the function $f$.

Butcher [6] took B-stability as a basic definition and it was only later that the significance of algebraic stability was discovered in [4]. This was achieved through the study of AN-stability as defined in section 2. It is clear from Results 2.1, 2.5 and 3.1 that the classes of problems (1.1),(2.1) and (1.1), (2.4) and (1.1),(3.1) are very closely related and this is reflected in the close relationship between the stability theories The following remarkable result proved in [34] is an extension of results proved in [4] and [13].

***Result*** 3.4 ***For S-irreducible RKMs***

***algebraic stability $\Leftrightarrow$ AN-stability $\Leftrightarrow$ B-stability $\Rightarrow$ A-stability.*** □

**Remark** S-irreducibility is a technical property defined in [28]. We will not reproduce the definition here, but restrict ourselves to noting that these methods are widely occurring in practice, and in particular that every non-confluent RKM ($c_i \neq c_j$ for $i \neq j$) is $S$-irreducible.

This is only a brief overview of the theories for linear decay and contractive nonlinear problems; for further details see [19] and [28]. These theories are very complete but, as is transparent from Results 2.1, 2.5 and 3.1, they only apply to problems with a ***very limited range of dynamical behaviour.*** The theory of nonlinear contractive problems has been extended to contraction in norms other than those induced by an inner product [43], [45] and a very clear account can be found in [36]. However, these theories are similarly restrictive in terms of the range of dynamical behaviour admitted by the underlying model problems.

In contrast to the linear decay problem, any conditional theory of numerical contractivity (a generalisation of absolute stability) will involve dependence of the allowable time-step on the initial data and is hence much harder to develop. The following example illustrates this point:

**Example** Consider the equation $u_t = -u^3$ which is strictly contractive in the sense of (3.1) for all u $\neq$ v and also satisfies $f(0) = 0$. Hence, by Result 3.1, $u(t) \to 0$ as $t \to \infty$. The explicit Euler scheme gives the map $U_{n+1} = (1 - \Delta t U_n^2)U_n$ from which it is possible [47] to deduce that $U_n \to 0$ as $n \to \infty$ if and only if At < $2/(U_0^2)$. For general nonlinear problems and general non algebraically stable methods it is very difficult to isolate explicitly such initial data dependent bounds on At. □

To make progress with a conditional theory for nonlinear problems it is necessary to impose still further restrictions on the class of problems. Indeed much of the generalised theory of contractivity reviewed in [36] is further restricted by employing the ***circle condition*** [17]

$$\|f(u) - f(v) + \rho(u - v)\| \leq \rho\|u - v\|, \forall u, v \in \mathbb{R}^p \tag{3.7}$$

and some $\rho > 0$. This unnatural condition can only be satisfied by globally Lipschitz functions which ***limits even further the range of direct applications.*** Of course the motivation behind (3.7) is to combine it with some ***a priori*** analysis of the underlying equation and its numerical approximations which enables the vector field defining the differential equation to be replaced by a globally Lipschitz one satisfying (3.7). However this important step is rarely addressed in the literature.

**Result 3.5** ***A function $f(\bullet)$ satisfying (3.7) is necessarily globally Lipschitz.***

**Proof** Assume to the contrary. Then, for any $K > 0$, there exist u, v $\in \mathbb{R}^p$ with u $\neq$ v such that

$$\|f(u) - f(v) + \rho(u - v)\| \geq \|f(u) - f(v)\| - \rho\|u - v\| \geq (K - \rho)\|u - v\|.$$

Now choosing $K > 2\rho$ contradicts (3.7). This completes the proof. 0

It is worth noting that the circle condition (3.7) is implied by the assumption

$$\exists \alpha > 0 : \langle f(u) - f(v), u - v \rangle \leq -\alpha\|f(u) - f(v)\|^2 \ \forall u, v \in \mathbb{R}^p, u \neq v.$$

The circle condition (3.7) then holds with $p = 1/\alpha$. In this sense it can be seen that (3.7) is a very special case of the contractivity condition (3.1).

## 4. Gradient Systems

As in section 3, for simplicity of exposition we will consider the case where (1.1) is autonomous and real and for which $f(u, t) \equiv f(u) \in C^1(\mathbb{R}^p, \mathbb{R}^p)$. It is clear from sections 2 and 3 that linear decay and contractivity are such strong conditions that they rule out interesting dynamics and hence it is natural to relax the notion of contractivity to allow some expansion of trajectories. The function $f$ is said to satisfy a ***one-sided Lipschitz condition*** if there exists a constant c > 0 such that

$$\langle f(u) - f(v), u - v \rangle \leq c\|u - v\|^2, \quad \forall u, v \in \mathbb{R}^p. \tag{4.1}$$

This allows exponential separation of trajectories and specifically it is straightforward to show that

**Result** 4.1 *Any two solutions* $u(t), v(t)$ *of* (1.1), (4.1) *satisfy*

$$\|u(t) - v(t)\| \leq e^{ct}\|u(0) - v(0)\|,$$

*for all* $t \geq 0$. 0

Numerical counterparts of Result 4.1 have been studied and these may be useful in establishing continuity of the numerical solution with respect to initial data – see Butcher [7] and [28]. Solvability of the Runge-Kutta equations in this context is discussed in [10]. The importance of continuity with respect to initial data will become apparent from Result 4.4.

Since exponential separation of trajectories allows the possibility of exponential growth of the solutions themselves, (4.1) alone is far too broad a class of problems to work with and make substantial progress; for this reason it is sensible to add further structure to the problem. Both linear decay and strictly contractive nonlinear problems are characterised by the property that ***u(t)*** approaches a unique equilibrium as time increases. This can be relaxed to the notion that ***u(t)*** approaches an equilibrium as time increases but that it is ***not necessarily unique.*** This leads naturally to the class of ***gradient systems*** for which there exists $F \in C^2(\mathbb{R}^p, \mathbb{R})$ such that

$$\left.\begin{array}{c} f(u) = -\nabla F(u),\ \forall u \in \mathbb{R}^p \\ F(u) \geq 0,\ \forall u \in \mathbb{R}^p, \\ F(u) \to \infty \text{ as } \|u\| \to \infty. \end{array}\right\} \tag{4.2}$$

For gradient systems it follows that

$$\frac{d}{dt}(F(u)) = \langle \nabla F(u), u_t \rangle = -\langle f(u), u_t \rangle = -\|u_t\|^2. \tag{4.3}$$

Hence, arguing loosely, that u will be driven to the critical points of ***F,*** which are the equilibria of (1.1). If ***F*** is convex so that

$$\langle \nabla F(u) - \nabla F(v), u - v \rangle \geq 0 \ \forall u, v \in \mathbb{R}^p.$$

then (4.2) is a contractive problem and the analysis of section 2 applies; in particular the set of equilibria define a convex set. However, for non-convex ***F*** equation (1.1), (4.2) may have multiple isolated equilibria. A simple example is the following:

**Example** Consider equation (1.1) in dimension $p = 1$ with $f(u) = u - u^3$. This is in gradient form with

$$F(u) = \frac{1}{4}(u^2 - 1)^{\text{“}}.$$

Notice the three equilibria 0, 1, -1. □

**Example** Consider equation (1.6), (1.7) with $\hat{b} = \hat{d} = 0$ and $\hat{a} = \gamma, \hat{c} = \hat{e} = 1$ and $u(x, t) \in \mathbb{R}$. Then, defining

$$F(u) = \int_0^1 \frac{\gamma}{2}u_x^2 + \frac{1}{4}(u^2 - 1)^2 dx,$$

the equation may be written as

$$u_t = -\nabla F(u)$$

where V is now interpreted as the variational derivative of ***F(u)*** with respect to changes in u, confined to an appropriate function space satisfying the boundary conditions. □

Gradient systems arise in a variety of applications; in particular, many phenomenological models of phase transitions such as the solid/solid Cahn-Hilliard equations [22] and the super/normal conducting Ginzburg-Landau equations [9] are in gradient form. Furthermore, gradient systems have been fundamental in the development of many important concepts in the theory of ordinary differential equations and are important for this reason alone; see [29] and the references therein. As suggested by (4.3) gradient systems are characterised by the following behaviour, proved in [31]:

**Result** 4.2 ***For any solution u(t) of*** $(1.1),(4.2)$ ***and any sequence*** $t_i \to \infty$ ***for which the w-limit point***

$$x := \lim_{i\to\infty} u(t_i) \tag{4.4}$$

***exists, it follows that*** $x \in \mathcal{E}$, ***the set of zeros of f. Furthermore, if all members of*** $\mathcal{E}$ ***are isolated then, for each*** $u(0)$ ***there exists*** $\bar{u} := \bar{u}(u(0)) \in \mathcal{E}$ ***such that***

$$\lim_{t\to\infty} u(t) = ii.$$

**Proof** Given $u(0) \in \mathbb{R}^p$ let $w(u(0))$ be the union of points such that (4.4) is defined for some sequence $t_i$. Then $\omega(u(0))$ is known as the w-limit set and is a closed, invariant (under forward evolution of the differential equation) set which is connected if compact [2].

Let $x, y \in \omega(u(0))$. Then $F(y) = F(x)$ for otherwise we obtain a contradiction to (4.3). Now consider the solution $u(t)$ of (l.l), (4.2) with $u(0)$ = x $\in \omega(u(0))$; since the w-limit set is invariant it follows that $u(t) \in \omega(u(0))$ and hence that $F(u(t)) = F(u(0)) \ \forall t \geq 0$. By (4.3) this implies that $u_t \equiv 0 \ \forall t \geq 0$ and hence that $u(0)$ = x $\in \mathcal{E}$.

Finally note that since $0 \leq F(u(t)) \leq F(u(0))$ it follows from (4.2) that all trajectories are uniformly bounded as $t \to \infty$. Thus $\omega(u(0))$ is compact since it is closed and we deduce that it is also connected. Since the equilibria are isolated it follows that the w-limit set must be a single point $\bar{u} \in \mathcal{E}$. Since the closure of the trajectory is compact it follows that

$$\lim_{t\to\infty} u(t) = \bar{u}. \ 0$$

For gradient systems it is natural to ask that a numerical approximation replicates the property (4.3) that there is a Liapunov function which drives the solution to equilibrium. Even if the additional constraint (4.1) is imposed it is unlikely to be possible to find a stability theory which holds for arbitrary At, since the problems under consideration admit both contractive and divergent behaviour. However, it is both feasible and desirable to find restrictions which are ***independent of initial data.*** This motivates the following definition.

**Definition** 4.3 A RKM is said to be ***gradient stable*** if, when applied to $(1.1)$, $(4.1)$, (4.2) there exists At, > 0 and a function $F_{\Delta t}(\bullet) : \mathbb{R}^p \to \mathbb{R}$ such that, for all At $\in$ (0, At,) :

(i) $F_{\Delta t}(U) \geq 0$ $\forall U \in \mathbb{R}^p$;

*(ii)* $F_{\Delta t}(U) \to \infty$ as $\|U\| \to \infty$.

(iii) $F_{\Delta t}(U_{n+1}) \leq F_{\Delta t}(U_n)$ $\forall U_n \in \mathbb{R}^p$;

(iv) if $F_{\Delta t}(U_n) \equiv F_{\Delta t}(U_0)$ Qn $\geq 0$ then $U_0 \in \mathcal{E}$, the set of equilibrium points for (1.1),(4.2). □

Such a definition was implicit in the work of Elliott [22] where discrete gradient systems were used in the analysis of numerical approximations of the Cahn-Hilliard equation. A theorem closely related to the following result is proved in [24].

**Result 4.4** ***Assume that, given initial data in*** $\mathbb{R}^p$, ***the RKM generates a unique*** $C^1$ ***mapping from*** $\mathbb{R}^p$ ***into itself which depends continuously on initial data. Then, for any solution of a gradient stable RKM applied to*** (1.1), (4.2) ***and any sequence*** $n_i \to \infty$ ***for which the w-limit point***

$$x := \lim_{i \to \infty} U_{n_i}$$

***exists, it follows that*** $x \in \mathcal{E}$, ***the set of zeros of f. Furthermore, if all members of*** $\mathcal{E}$ ***are isolated then, for each*** $u(0)$ ***there exists*** $\bar{u} := \bar{u}(u(0)) \in \mathcal{E}$ ***such that***

$$\lim_{n \to \infty} U_n = \bar{u}.$$

**Proof** As for the differential equation, the w-limit set, $\omega(u(0))$ is defined as the union of all possible limit points corresponding to given initial data. A similar argument to that in the Result 4.2 shows that $U_n$ is uniformly bounded in n. From Lemma 2.1.2 in [29] it follows that, since the RKM defines a unique sequence, continuously dependent upon initial data, $\omega(U_0)$ is non-empty, compact and invariant and an argument identical to that in the proof of Result 4.2 shows that x $\in \mathcal{E}$.

However, for dynamical systems defined by mappings it does not follow that $\omega(U_0)$ is connected if compact, and a different argument is needed for the last part of the result. Now assume that the members of $\mathcal{E}$ are isolated. Since the solution sequence is bounded it is contained in a compact set $B$ and this implies that there are a finite number of possible equilibria contained in $\omega(U_0)$, say $x_j$, j = 1,. . ., J, in $\mathcal{E}$. Let $B_j = B(x_j, \delta) := \{u \in \mathbb{R}^p : \|u - x_j\| < \delta\}$, $B^+ = \bigcup_{j=1,\ldots,J} B_j$ and $B^- = B \backslash B^+$; note that $B^-$ is closed by construction. Assume that $\delta$ is sufficiently small that $dist(x, B_k) \geq$ A > 0 Qx $\in B_j$, j $\neq k$. Note that $\omega(U_0)$ is non-empty. Assume for the purposes of contradiction that $x_1 \in \omega(U_0)$ and that it is not the unique member of $\omega(U_0)$. Then for all $\delta > 0$ there exists a sequence $n_i \to \infty$ such that $U_{n_i} \in B_1$ and $U_{n_i} \to x_1$ as $n_i \to$ co. Since $x_1$ is not the unique limit point there is an infinite sequence of integers $m_j$ such that $U_{m_j} \in B_1$ and $U_{m_j+1} \notin B_1$. Since the mapping defined by the RKM is $C^1$, it is Lipschitz with constant $L$ on $B_1$ and since $x_1$ is a fixed point, we deduce that

$$\|U_{m_j+1} - x_1\| \leq L\|U_{m_j} - x_1\| \leq L\delta.$$

Hence, if $L\delta <$ A we deduce that $U_{m_j+1} \in B^-$ for each $j$. But $B^-$ is compact and hence the infinite sequence $U_{m_j+1}$ must have a limit point; such a limit point cannot be contained

in $\mathcal{E}$ by definition of $B^-$ and this contradicts the first part of the result. This completes the proof, since the sequence is bounded. **0**

**Remark** The assumption that the RKM generates a unique continuously dependent solution sequence is often made, and is usually a very strong assumption. However it is not an unreasonable assumption to make for a system that satisfies (4.1): the one-sided Lipschitz condition implies unique solvability of the Runge-Kutta equations for many classes of implicit methods, if At is sufficiently small (but independent of $U_n$), including those based on Gauss-Legendre quadrature, the Radau IA, IIA and Lobatto IIIA, IIB and IIIC methods; see [19], [28]. Continuous dependence on initial data can be similarly established.

Further studies of gradient stability may found in [20] where one-step methods for the Cahn-Hilliard equation are examined. Here we present a proof that the theta method

$$U_{n+1} = U_n + \Delta t[(1-\theta)f(U_n) + \theta f(U_{n+1})] \tag{4.5}$$

is gradient stable for $\theta \in [\frac{1}{2}, 1]$. This illustrates some of the issues involved in establishing gradient stability. Note that the condition on $\theta$ is equivalent to the condition that the method be A-stable. We require a preliminary lemma.

**Lemma 4.5** ***For a gradient system,*** (4.1) ***implies that***

$$F(u) - F(v) \leq \langle f(u),\ v - u\rangle + c\|u - v\|^2 \tag{4.6}$$

*for any* $u, v \in \mathbb{R}^p$.

**Proof** Let G(x): [0, 1] → IR be defined by

$$G(x) = F(v + x[u - v]).$$

Then we have

$$\begin{aligned} G'(x) &= \langle \nabla F(v + x[u - v]),\ u - v\rangle \\ &= \langle f(v + x[u - v]),\ v - u\rangle. \end{aligned}$$

Now by the mean value theorem G(1) − G(0) = G'(x) for some x ∈ [0, 1]. Hence writing $\xi = v + x[u - v]$ implies that

$$F(u) - F(v) = \langle f(\xi),\ v - u\rangle$$

and since

$$\frac{v - u}{\|v - u\|} = \frac{\xi - u}{\|\xi - u\|}$$

it follows that

$$F(u) - F(v) = \frac{\|v - u\|}{\|\xi - u\|}\langle f(\xi), \xi - u\rangle$$

Now substituting $\xi$ for v in (4.1) implies that

$$\langle f(\xi), \xi - u\rangle \leq \langle f(u), \xi - u\rangle + c\|u - \xi\|^2$$

and hence

$$\begin{aligned} F(u) - F(v) &\leq \frac{\|v-u\|}{\|\xi-u\|}\langle f(u), \xi - u\rangle + c\|\xi - u\|.\|v-u\| \\ &\leq \langle f(u), v-u\rangle + c\|v-u\|^2 \end{aligned}$$

as required. □

**Result** 4.6 *The theta method (4.5) is gradient stable for* $\theta \in [\frac{1}{2}, 1]$ *with* $\Delta t = 1/c$, *where* $c$ *is the constant in* (4.1), *and*

$$F_{\Delta t}(U) = F(U) + \frac{\Delta t}{2}(1-\theta)\|f(U)\|^2.$$

**Proof** By (4.6)

$$\begin{aligned} F(U_{n+1}) - F(U_n) &\leq \langle f(U_{n+1}), U_n - U_{n+1}\rangle + c\|U_{n+1} - U_n\|^2 \\ &= \langle \frac{1}{\Delta t}\left[U_{n+1} - U_n - \Delta t(1-\theta)f(U_n) - \Delta t\theta f(U_{n+1})\right], U_n - U_{n+1}\rangle \\ &\quad + \langle f(U_{n+1}), U_n - U_{n+1}\rangle + c\|U_{n+1} - U_n\|^2 \\ &= \left(c - \frac{1}{\Delta t}\right)\|U_{n+1} - U_n\|^2 + (1-\theta)\langle f(U_n) - f(U_{n+1}), U_{n+1} - U_n\rangle \\ &= \left(c - \frac{1}{\Delta t}\right)\|U_{n+1} - U_n\|^2 + \frac{\Delta t}{2}(1-\theta)\left[\|f(U_n)\|^2 - \|f(U_{n+1})\|^2\right] \\ &\quad - \frac{\Delta t}{2}(1-\theta)(2\theta - 1)\|f(U_n) - f(U_{n+1})\|^2. \end{aligned}$$

Hence, for $\theta \in [\frac{1}{2}, 1]$,

$$F_{\Delta t}(U_{n+1}) - F_{\Delta t}(U_n) \leq \left(c - \frac{1}{\Delta t}\right)\|U_{n+1} - U_n\|^2.$$

Clearly $F_{\Delta t}$ is bounded below for $\theta \leq 1$ and, since $F_{\Delta t}(U) \geq F(U)$ for all $U \in \mathbb{R}^p$ *(ii)* of Definition 4.3 follows. It is also clear that $F_{\Delta t}(U)$ is non-increasing for $\Delta t \in (0, 1/c)$. Furthermore, if $F_{\Delta t}(U_{n+1}) = F_{\Delta t}(U_n)$ then $U_{n+1} = U_n$. The equilibrium points for (4.5) coincide with those of (1.1), (4.2) [35] and so gradient stability has been established. *0*

A similar method of proof establishes that the one-leg counterpart of the theta method (4.5) is also gradient stable; see [32].

As can be seen a complete theory of gradient stability is not yet developed. However, it is worth observing that, if the additional assumption (5.3) (a form of dissipativity) is appended to (4.2) and the equilibria are isolated then the conclusion of Result 4.4 follows provided $M$ is positive semi-definite and $B$ is positive definite [33].

## 5. Dissipative Systems

As in sections 3 and 4, for simplicity of exposition we will consider the case where (1.1) is autonomous and real so that $f(u, t) \equiv f(u) \in C^1(\mathbb{R}^p, \mathbb{R}^p)$. Even the one-sided Lipschitz

condition which we introduced in the previous section is far too restrictive for many interesting applications and so we relax this condition in our study of dissipative problems. Furthermore, gradient systems only allow solutions to approach equilibria for large time so that periodic, quasi-periodic or chaotic behaviour is not admitted; the dissipative problems we study will admit such behaviour.

The notion of dissipativity is an important one in many physical applications and naturally there is a mathematical abstraction of this idea in the theory of differential equations; see, for example, [29] and [49]. Roughly speaking an initial value problem is said to be ***dissipative*** if there is a bounded set, in an appropriate function space for the problem, which all solutions enter after a finite time and thereafter remain inside: thus some measure of energy is dissipated outside the bounded set.

To motivate the study of dissipative problems consider first the equation **(1.1)** under (3.1), together with the assumption that $f(0) = 0$. Taking v = 0 in (3.1) we then deduce that

$$\langle f(u), u\rangle \leq 0. \tag{5.1}$$

It is straightforward to prove from this that

$$\|u(t)\|^2 \leq \|u(0)\|^2 \ \forall t \geq 0. \tag{5.2}$$

This property is often termed ***montonicity*** or ***weak contractivity.*** The numerical analogue of property (5.2), under the assumption (3.1) together with $f(u) = 0$, is studied by a number of authors including [12], [46], [44]. A straightforward application of the theory in section 3 shows that algebraic stability is sufficient for a numerical analogue of (5.2) to hold for all step-sizes $\Delta t > 0$ and all initial data. In fact, a wider class of methods suffices in this context as described in [12].

The monotonicity induced by (5.1) can be weakened to enforce monotonicity only outside a certain bounded region of phase space. This corresponds to a notion of dissipation at sufficiently large amplitude. In this section we will concentrate on a particular class of problems where dissipativity is induced by the structural assumption

$$\langle f(u), u\rangle \leq \gamma - \omega\|u\|^2, \forall u \in \mathbb{R}^p \tag{5.3}$$

Under (5.3) monotnicity is induced outside the set $\mathcal{B}$ = {u $\in \mathbb{R}^p$ : $\|u\|^2 \leq$ y/w}. An example of a system satisfying (5.3) is the Lorenz equations, after translation of the origin. Many other examples exist; in particular, infinite dimensional systems such as the complex Ginzburg-Landau equations (see below) and the Navier-Stokes equation (in two dimensions) satisfy generalisations of (5.3) (see [49]) and, under appropriate spatial discretisation, the resulting system of ordinary differential equations satisfy (5.3). (Note that the contractive problems of section 3 are sometimes referred to as dissipative in the numerical analysis literature; since this conflicts with the terminology in the theory of differential equations we have deferred to the usage in the differential equations literature.)

**Example** Consider equation (1.6), (1.7) with $\hat{a} = \hat{b} = \hat{c} = \hat{d} = \hat{e} = 1$. Then we obtain

$$u_t = (1 + i)u_{xx} - (1 + i)|u|^2u + u, \ x \in (0, 1),$$

together with periodic boundary conditions (1.7). Taking ***f(u)*** as the right-hand side of this equation and employing the standard $L_2$ norm and inner-product we obtain

$$\langle (1+i)u_{xx} - (1+i)|u|^2u + u, u\rangle = -\int_0^1 |u_x|^2 dx + \int_0^1 |u|^2 - |u|^4 dx \leq \int_0^1 1 - |u|^2 dx = 1 - \|u\|^2.$$

Thus an infinite dimensional analog of (5.3) is satisfied with $\gamma = 1, \omega = 1$. ◊

**Result 5.1** ***For*** (1.1), (5.3), ***any*** $u(0) \in I\!R^p$ ***and any*** $p > 0$ ***there exists*** $t^* := t^*(\rho, u(0))$ ***such that***

$$\|u(t)\|^2 \leq \frac{\gamma}{\omega} + \rho$$

***for all*** $t \geq t^*$.

**Proof** Taking the inner product of (1.1) with u gives

$$\frac{1}{2}\frac{d}{dt}\|u\|^2 = \langle u, u_t\rangle \leq \gamma - \omega\|u\|^2.$$

Thus

$$\frac{d}{dt}(e^{2\omega t}\|u\|^2) \leq 2\gamma e^{2\omega t}$$

$$\Rightarrow \|u(t)\|^2 \leq \frac{\gamma}{\omega} + e^{-2\omega t}\left[|u(0)\|^2 - \frac{\gamma}{\omega_I}\right.$$

The result follows. □

Thus all the information about the asymptotic behaviour for (1.1), (5.3) is captured in a bounded set; within this set the dynamics may be very complicated, for example chaotic. It is important to note that problems in the class ***(5.3) do not necessarily satisfy a one-sided Lipschita condition,*** as the following example shows:

**Example** Consider the 2-dimensional problem

$$\begin{aligned} \dot{x} &= -x + xy \\ \dot{y} &= -y - x^2. \end{aligned}$$

. We will show that this problem is dissipative in the sense of (5.3) but that the system does not satisfy aone-sided Lipschitz condition. Let u = (x, $y)^T$ and ***f(u)*** = $(-x+xy, -y-x^2)^T$ then

$$\begin{aligned} \langle f(u), u\rangle &= -x^2 - y^2 \\ &= -\|u\|^2. \end{aligned}$$

Thus (5.3) is satisfied with $\gamma$ = 0 and w = 1, and Result 5.1 implies that $\|u\| \to 0$ as $t \to \infty$. Now, to show that a one-sided Lipschitz condition is not satisfied, let v = (x', $y')^T$ so that

$$\textit{(f(u) - f(v), u - v)} = -(x - x')^2 + (x - x')(xy - x'y') - (y - y')^2 + (y - y')(x'^2 - x^2)$$

Suppose that **(4.1)** holds and let u = $(\beta, \alpha)^T$ and v = $(\alpha, \beta)^T$ where the constants $\alpha$ and $\beta$ are to be specified below. Notice that $\|u - v\|^2 = 2(\beta - \alpha)^2$ and observe that

$$\begin{aligned}\langle f(u) - f(v),\, u - v\rangle &= -2(\beta-\alpha)^2 + (\beta-\alpha)(\beta^2-\alpha^2) \\ &= \left[\frac{1}{2}(\alpha+\beta) - 1\right]\|u-v\|^2.\end{aligned}$$

Choose $\alpha + \beta > 2(c + 1)$ to obtain a contradiction. Thus this system does not satisfy a one-sided Lipschitz condition for any c > 0, even though this system is dissipative in the sense of (5.3) and, in fact, the origin is globally attracting. 0

This is not an isolated example. In [33] it is shown that the Lorenz equations do not satisfy a one-sided Lipschitz condition and there are many other examples within the class of dissipative systems. Because of this, we will ***not*** assume that (4.1) holds for systems in the class (1.1),(5.3).

It is natural to ask for a property analogous to Result 5.1 for the numerical method. However, in the light of the example above it perhaps seems too much to ask for a stability theory which is independent of initial data for problems satisfying (5.3) since there is not even a one-sided Lipschitz constant for these problems. However this view is overly pessimistic as we now show. First we make a definition:

**Definition** 5.2 A RKM is said to be ***dissipative stable*** if, when applied to (1.1),(5.3), there exists At,, $R$ > 0 both independent of $U_0$ such that for all At $\in$ (0, At,) and any $U_0 \in \mathbb{R}^p$ there exists n* := $n^*(U_0$, At) for which any sequence $\{U_n\}_{n=0}^{\infty}$ generated by the RKM satisfies

$$\|U_n\|^2 \le R$$

for all $n \ge n^*$. 0

The following result shows a remarkable correspondence between the contractive nonlinear stability theories and the appropriate theory for problems satisfying (5.3): algebraically stable RKMs are again seen to have desirable stability properties.

**Result** *5.3 Consider a DJ-irreducible, algebraically stable RKM applied to* (1.1), (5.3) *with any* At $>$ *0. Then the RKM is dissipative stable with* At, = $\infty$ *and hence*

***algebraic stability*** $\Rightarrow$ ***dissipative stability.*** 0

**Proof** From the definition of the Runge-Kutta method it follows that

$$\|U_{n+1}\|^2 \le \|U_n\|^2 + 2\Delta t \sum_{j=1}^{k} b_j \langle U_n, f_j\rangle + \Delta t^2 \sum_{i,j=1}^{k} b_i b_j \langle f_i, f_j\rangle$$

where $f_i := f(\eta_i)$. Using the equation for the $\eta_i$ we have

$$\langle U_n, f_i\rangle = \langle \eta_i, f_i\rangle - \Delta t \sum_{j=1}^{k} a_{ij} \langle j_i, f_j\rangle$$

and this gives

$$\|U_{n+1}\|^2 \leq \|U_n\|^2 + 2\Delta t \sum_{j=1}^{k} b_j \langle \eta_j, fj) - \Delta t^2 \sum_{i,j=1}^{k} m_{ij} \langle f_i, f_j \rangle.$$

Using algebraic stability and (5.3) we deduce that

$$\|U_{n+1}\|^2 \leq \|U_n\|^2 + 2\Delta t \sum_{j=1}^{k} b_j [\gamma - \omega \|\eta_j\|^2].$$

Thus we have, for any given $\epsilon > 0$, that either

$$\|U_{n+1}\|^2 \leq \|U_n\|^2 - 2\Delta t \epsilon \tag{5.4}$$

or

$$\sum_{j=1}^{k} b_j [\gamma - \omega \|\eta_j\|^2] \geq -\epsilon$$

$$\Rightarrow \sum_{j=1}^{k} b_j \|\eta_j\|^2 \leq \frac{\gamma + \epsilon}{w}. \tag{5.5}$$

Since the method is D J-irreducible it follows that the $b_i > 0$ [28] and thus (5.5) implies that

$$\|\eta_j\|^2 \leq \frac{\gamma + \epsilon}{\omega b_j}. \tag{5.6}$$

However, using the bound (5.6) it is possible to deduce a bound on $U_{n+1}$ simply by noting that

$$U_{n+1} = \eta_i + \Delta t \sum_{j=1}^{k} [b_j - a_{ij}] f(\eta_j).$$

Squaring both sides of this expression we obtain

$$\|U_{n+1}\|^2 \leq \|\eta_i\|^2 + K\Delta t \tag{5.7}$$

where $K$ is independent of $U_0$ and depends only on the bounds (5.6) on the $\eta_j$. Performing a sum weighted by the $b_i$, and recalling that $\sum_{j=1}^{k} b_j = 1$ by consistency, we obtain from (5.5), (5.7)

$$\|U_{n+1}\|^2 \leq \sum_{j=1}^{k} b_i [\|\eta_i\|^2 + K\Delta t] \leq \frac{\gamma + \epsilon}{w} + K\Delta t. \tag{5.8}$$

Thus either (5.4) or (5.8) holds. An induction based on these quantities yields the desired result with

$$R = \frac{\gamma + \epsilon}{w} + K\Delta t. \quad \bullet I$$

**Remarks** (i) In [33] it is shown by use of the Brouwer fixed point theorem that under (5.3), for a DJ-irreducible algebraically stable method with invertible $A$, the Runge-Kutta

equations have *a* solution for all $t \geq 0$ and any $U_n \in \mathbb{R}^p$. However, uniqueness cannot be established under (5.3) alone.
(ii) Notice that the bound $R$ on $U_n$ obtained for sufficiently large n is very close to the bound (y/w) + $\rho$ for the differential equation; thus the set into which the large time dynamics are confined is also closely related to the equivalent set for the differential equation.
(iii) Notice again the role of algebraic stability: it enables us to determine the sign of the quadratic form defined by $M$, just as in the proof of Result 3.3.

## 6. Conservative Systems

We shall start this section, as in sections 3,4 and 5 by considering the case where (1.1) is autonomous and real for which $f(u, t) \equiv f(u) \in C^1(\mathbb{R}^p, \mathbb{R}^p)$. We will then go further and, look at certain complex *matrix* systems of differential equations.

In many physical applications, no energy-loss mechanism is present and conservative systems result. As a simple example of a conservative system, which arises naturally from the limit $\gamma$, w $\to 0$ of the dissipative systems considered in section 5, we take the structural assumption

$$\langle f(u), u \rangle = 0, \quad \forall u \in \mathbb{R}^p. \tag{6.1}$$

**Example** The equations

$$x_t = -xy^2, \qquad y_t = x^2 y$$

satisfy (6.1) □.

**Example** The nonlinear Schrodinger equation, which is a non-dissipative limit of the complex Ginzburg-Landau equation, satisfies an infinite-dimensional analogue of (6.1) and arises throughout mathematical physics. Specifically we take $\hat{a} = \hat{c} = \hat{e} = 0$ and $\hat{b} = \hat{d} = 1$ in (1.6), (1.7) and we obtain

$$u_t = iu_{xx} - i|u|^2 u. \tag{6.2}$$

Note that, using integration by parts,

$$\langle iu_{xx} - i|u|^2 u, u \rangle = -\int_0^1 Re\{i|u_x|^2 + i|u|^4\} dx = 0$$

and so we have an infinite dimensional analog of (6.1). *0*

By following the proof of Result 5.1 it is straightforward to see that:

**Result 6.1** *The solution u(t) of* (1.1), **(6.1)** *satisfies*

$$\|u(t)\| = \|u(0)\|$$

*for all* $t \geq 0$. 0

Again it is natural to ask for numerical schemes which mimic this property. This approach was taken by Cooper [11] and a modification of the classical theory of [4] and the use of ideas from the proof of Result 5.3 enables proof of the following result, which

shows a remarkable correspondence with the classical theories of sections 2 and 3 and the new theory described in section 5.

**Result** 6.2 ***Consider the numerical solution of*** (1.1), (6.1) ***by a RKM. If the RKM satisfies*** $M \equiv 0$ ***where*** $M$ ***is defined by*** (1.5) ***then***

$$\|U_n\| = \|U_0\|$$

***for all*** $n \geq 0$.

**Proof** By definition of the RKM we have

$$\|U_{n+1}\|^2 = \|U_n\|^2 + 2\Delta t \sum_{i=1}^{k} b_i \langle U_n, f(\eta_i)\rangle + \Delta t^2 \sum_{i,j=1}^{k} b_i b_j \langle f(\eta_i), f(\eta_j)\rangle.$$

Using the defining equation for the $\eta_i$ gives

$$\|U_{n+1}\|^2 = \|U_n\|^2 + 2\Delta t \sum_{i=1}^{k} b_i \langle \eta_i, f(\eta_i)\rangle - \Delta t^2 \sum_{i,j=1}^{k} m_{ij} \langle f(\eta_i), f(\eta_j)\rangle.$$

Using the fact that $M \equiv 0$, and the structural assumption (6.1), the result follows. *0*

**Remark** (i) Algebraically stable methods of arbitrarily high order which satisfy $M \equiv 0$ do exist: they are those schemes based on Gauss-Legendre quadrature and discussed in [6], [3]. In particular, the implicit mid-point rule

$$U_{n+1} = U_n + \Delta t f\left(\frac{U_{n+1} + U_n}{2}\right)$$

is algebraically stable and satisfies $M \equiv 0$.
(ii) Again the role of the matrix $M$ is crucial; in this case not only is a bound on the quadratic form important but it is necessary to remove its contribution. Setting $M \equiv 0$ does this.
(iii) The solvability of the RK equations has not been investigated for RKMs under (6.1).
(iv) In the related case where $(f(u), u) \leq 0$ there is relevant numerical analysis contained in the work of Spijker [46].

We now consider a stronger kind of conservation: we consider the matrix system of differential equations (with * denoting Hermitian transpose)

$$Q_t = S(Q)Q, \quad Q^*(0)Q(0) = I, \tag{6.3}$$

where $Q(t)$ is a time-dependent p x $p$ complex-valued matrix, S(Q) is a skew-Hermitian matrix-valued function of Q which satisfies

$$S^*(Q) = -S(Q) \quad QQ \in \mathbb{R}^{p\times p} \tag{6.4}$$

and $I$ is the $p \times p$ identity. Equation (6.3) arises in applications such as the continuous SVD and closely related problems arise in the computation of Liapunov exponents for systems

of ordinary differential equations. The system is conservative in a very strong sense: the orthonormality of the columns of the matrix Q are preserved with time evolution.

**Result 6.3** ***The solution Q(t) of*** $(6.3)$, ***(6.4) satisfies***

$$Q^*(t)Q(t) = I$$

*for all* $t \geq 0$.

**Proof** Clearly

$$\frac{d}{dt}(Q^*Q) = Q^*Q_t + Q_t^*Q.$$

But

$$Q^*Q_t = Q^*S(Q)Q$$

and hence, by $(6.4)$,

$$\frac{d}{dt}(Q^*Q) = Q^*[S(Q) \text{ t } S^*(Q)]Q = 0.$$

Thus $Q^*(t)\&(t) = Q^*(0)Q(0) = I$ as required □.

Applying the standard Runge-Kutta method to the matrix system (6.3) gives, for $Qn \approx Q(n\Delta t)$,

$$Q_{n+1} = Q_n + \text{At} \sum_{j=1}^{k} b_j S(\Gamma_j)\Gamma_j,$$

$$\Gamma_i = Q_n + \Delta t \sum_{j=1}^{k} a_{ij} S(\Gamma_j)\Gamma_j, \; i = 1, \ldots, k$$

where $\Gamma_i$ is a complex-valued $p$ x $p$ matrix. We will employ the notation $S_i$ := S( I';).

It is important in some contexts to find numerical methods which will automatically enforce the orthonormality of the columns of $Q(t)$ during numerical simulation. This was realised in [18] where the following result is proved:

**Result 6.4** ***The solution of*** $(6.3)$, ***(6.4) by a RKM with*** $M \equiv 0$ ***where*** $M$ ***is defined by*** ( 1.5) ***satisfies***

$$Q_n^*Q_n = I$$

***for all*** $n \geq 0$.

**Proof** From the definition of the RKM applied to (6.3) we obtain

$$Q_{n+1}^*Q_{n+1} = [Q_n^* + \Delta t \sum_{i=1}^{k} b_i \Gamma_i^* S_i^*][Q_n + \Delta t \sum_{j=1}^{k} b_j S_j \Gamma_j]$$

$$= Q_n^*Q_n + \Delta t \sum_{i=1}^{k} b_i \Gamma_i^* S_i^* Q_n + \Delta t \sum_{j=1}^{k} b_j Q_n^* S_j \Gamma_j + \Delta t^2 \sum_{i,j=1}^{k} b_i b_j \Gamma_i^* S_i^* S_j \Gamma_j.$$

Now, from the defining equation for the $\Gamma_i$,

$$\Gamma_i^* S_i^* Q_n = \Gamma_i^* S_i^* \Gamma_i - \Delta t \sum_{j=1}^{k} a_{ij} \Gamma_i^* S_i^* S_j \Gamma_j$$

and

$$Q_n^* S_j \Gamma_j = \Gamma_j{}^* S_j \Gamma_j - \Delta t \sum_{i=1}^{k} a_{ji} \Gamma_i{}^* S_i^* S_j \Gamma_j .$$

Combining these three expressions we find that

$$Q_{n+1}^* Q_{n+1} = Q_n^* Q_n \text{ t } \Delta t \sum_{i=1}^{k} b_i \Gamma_i{}^* [S_i^* + S_i] \Gamma_i - \Delta t^2 \sum_{i,j=1}^{k} m_{ij} \Gamma_i{}^* S_i^* S_j \Gamma_j .$$

Setting $M \equiv 0$ and employing (6.4) we obtain

$$Q_{n+1}^* Q_{n+1} = Q_n^* Q_n$$

and the desired result follows. □

**Remarks** (i) The proof presented in [18] is considerably more elegant than the proof given here, employing the theory of symplectic integrators as outlined in the next section and also yielding an "if-and only if" result. However, the proof given here once again makes clear the role of the positive definite quadratic form defined by $M$ and its annihilation by the choice $M \equiv 0$. Recall again that algebraically stable schemes satisfying $M \equiv 0$ exist and so, once again, the importance of algebraic stability is apparent.

(ii) The solvability of the Runge-Kutta equations has not been addressed here. However, in [18] an explicit iteration scheme is constructed which, if iterated to convergence, satisfies the Runge Kutta equations but which also retains the orthonormality of the system regardless of the number of iterations used. This then corresponds to a linearly implicit numerical method which is "stable" in an appropriate sense.

## 7. Hamiltonian Systems

The class of conservative systems induced by the inner product structure (6.1) is clearly a somewhat restrictive one and it is natural to broaden the scope somewhat to include more general schemes with conservation properties. To this end we consider the case where (1.1) is a real autonomous Hamiltonian system of even dimension with $f(u, t) \equiv$ f(u) $\in C^1(\mathbb{R}^p, \mathbb{R}^p)$ and p = 2N. To establish a connection with section 6 we consider first the linear problem

$$u_t = JAu, \tag{7.1}$$

where $A$ is positive definite symmetric and where $J$ is a skew-symmetric matrix satisfying

$$J^T = J^{-1} = -J. \tag{7.2}$$

Then we may define a norm based on $A$ by

$$\|u\|^2 = \frac{1}{2}u^T Au.$$

It follows that

$$\frac{d}{dt}\|u\|^2 = \frac{1}{2}[u_t^T Au + u^T Au_t]$$

$$= \frac{1}{2}[u^T AJ^T Au + u^T AJAu] = 0,$$

using $J^T$ = -J. This is equivalent to Result 6.1 and shows conservation of the Hamiltonian

$$H(u) := \frac{1}{2}u^T Au.$$

However, for nonlinear Hamiltonian systems this equivalence does not hold.

Given $H \in C^2(\mathbb{R}^{2N}, \mathbb{R})$, general Hamiltonian systems are of the form (1.1), where

$$f(u) = J\nabla H(u) \tag{7.3}$$

and $J$ is a skew-symmetric matrix satisfying (7.2). We shall mainly consider the case where

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

and $I$ is the $N$ x $N$ identity. Equation (1.1), (7.3) then takes the familiar form

$$Pt = \nabla_q H(u), \qquad q_t = -\nabla_p H(u)$$

where $u^T = (p^T, q^T)$ for p, q $\in \mathcal{R}^N$ and $\nabla_p$ (resp. $\nabla_q$) denotes the gradient with respect to the p (resp. q) variables.

**Example** A simple example is the system

$$p_t = p^2 q, \; q_t = -pq^2$$

which corresponds to the Hamiltonian $\frac{p^2q^2}{2}$. □

**Example** The nonlinear Schrodinger equation, (6.2), (1.7) is Hamiltonian with conjugation replacing the transpose and i playing the role of the skew symmetric operator ***J :***

$$u_t = -i\nabla F(u)$$

where

$$F(u) = \int_0^1 \frac{1}{2}|u_x|^2 + \frac{1}{4}|u|^4 dx,$$

and V represents the variational derivative with respect to changes in u, confined to an appropriate function space. □

Two important properties of Hamiltonian systems are described in Result 7.2. In order to explain the result we need to define the following:

**Definition** 7.1 A mapping G(U) $\in C(\mathbb{R}^{2N}, \mathbb{R}^{2N})$ is said to be ***symplectic*** if

$$DG(U)^T J DG(U) = J \quad \forall U \in \mathbb{R}^{2N}.$$

Here ***DG*** denotes the Jacobian of the mapping G with respect to the variable ***U.*** We will use an analogous notation for mappings other that G throughout this section.

**Result 7.2** ***Solutions of*** (1.1), (7.3) ***satisfy***

***(i) H(u(t))*** = $H(u(0))$ $\forall t \geq 0$;

***(ii) if the solution operator*** $G(U; t)$ ***is defined by u(t)*** = $G(u(0); t)$ ***for given initial data*** $u(0)$ ***then*** $G(\bullet, t)$ ***is a symplectic mapping for each*** $t \in \mathcal{R}^+$. □

**Proof** The first fact follows in a straightforward way since

$$\frac{d}{dt}H(u(t)) = \frac{1}{2}[\nabla H(u)^T u_t + u_t^T \nabla H(u)]$$

$$= \frac{1}{2}[\nabla H(u)^T J \nabla H(u) + \nabla H(u)^T J^T \nabla H(u)] = 0,$$

since $J^T = -J$. The result follows.

For the second part, let ***R(t)*** denote ***DG( U; t),*** where ***D*** denotes the Jacobian with respect to ***U.*** Then ***R(t)*** satisfies the matrix differential equation

$$R_t = JA(t)R, \quad R(0) = I$$

where ***A(t)*** is the Hessian of ***H(u)*** evaluated at u = ***u(t)*** and is hence symmetric. Now let ***V(t)*** = $R^T JR$ and note that ***V(0)*** = ***J.*** Clearly

$$V_t = R_t^T JR + R^T JR_t = R^T AJ^T JR + R^T JJAR.$$

Now, using (7.2) we obtain

$$V_t = R^T AR - R^T AR = 0$$

and hence $V(t) = J$ for all $t$. By definition of $V$ the result follows. □

Clearly (i) is a conservation property; since $H$ is in general not a positive-definite quadratic form this property is not equivalent to Result 6.1 except for the linear problem (7.1) with positive definite $A$. Although it is heavily disguised, (ii) is also a conservation property: it states that the area of the projection of any set in $\mathbb{R}^{2N}$ onto certain distinguished planes in $\mathcal{R}^2$ is preserved under the solution operator G [1]. Again it is natural to ask that the conservation properties (i) and (ii) are inherited by any numerical approximations. In this context the following result of Sanz-Serna [38] and of Lasagni [37] is of interest since it again shows a close relationship with the classical theory of section 3 and in particular the role of the matrix $M$ from algebraic stability theory in preserving (ii).

**Result** *7.3 Solutions of* (l.l), (7.3) *by the RKM with* $M \equiv 0$ *where* $M$ *is defined by* (1.5) *define a symplectic mapping for each* At $\geq$ 0. 0

**Proof** The Runge-Kutta method defines a mapping $U \rightarrow W$ determined implicitly by the equations

$$W = U + \Delta t \sum_{j=1}^{k} b_j f(\eta_j),$$

$$\eta_j = U + \Delta t \sum_{j=1}^{k} a_{ij} f(\eta_j).$$

We let $R = DW(U)$ and $\Gamma_j = D\eta_j(U)$ and denote the Jacobian of $f(\eta)$ with respect to $\eta$ evaluated at $\eta = \eta_i$ by $D f_i = D f(\eta_i)$. Then, differentiating the mapping with respect to $U$ gives

$$R = I + \Delta t \sum_{j=1}^{k} b_j D f_j \Gamma_j,$$

$$\Gamma_i = I + \Delta t \sum_{j=1}^{k} a_{ij} D f_j \Gamma_j.$$

Thus we obtain

$$R^T J R = [I + \Delta t \sum_{i=1}^{k} b_i \Gamma_i^T D f_i^T] J [I + \Delta t \sum_{j=1}^{k} b_j D f_j \Gamma_j]$$

so that

$$R^T J R = J + \Delta t \sum_{i=1}^{k} b_i \Gamma_i^T D f_i^T J + \Delta t \sum_{j=1}^{k} b_j J D f_j \Gamma_j + \Delta t^2 \sum_{i,j=1}^{k} b_i b_j \Gamma_i^T D f_i^T J D f_j^T \Gamma_j. \quad (7.4)$$

Now, from the defining equations for the $\Gamma_i$ ,

$$\Gamma_i^T D f_i^T J = \Gamma_i^T D f_i^T J \Gamma_i - \Delta t \sum_{j=1}^{k} a_{ij} \Gamma_i^T D f_i^T J D f_j \Gamma_j$$

and

$$JDf_j\Gamma_j = \Gamma_j{}^T JDf_j\Gamma_j - \Delta t\sum_{i=1}^{k} a_{ji}\Gamma_i{}^T Df_i^T JDf_j\Gamma_j.$$

Combining these expression with (7.4) we obtain

$$R^T JR = J + \Delta t\sum_{j=1}^{k} b_i\Gamma_i{}^T[Df_i^T J + JDf_i]\Gamma_i - \Delta t^2\sum_{i,j=1}^{k} m_{ij}\Gamma_i{}^T Df_i^T JDf_j\Gamma_j.$$

Since the method satisfies $M \equiv 0$ we obtain

$$R^T JR = J + \Delta t\sum_{j=1}^{k} b_i\Gamma_i{}^T[Df_i^T J + JDf_i]\Gamma_i.$$

**Now,** $Df_i = JA_i$ where $A_i$, the Hessian of $H$ evaluated at $\eta_i$, is symmetric. Hence, using (7.2),

$$Df_i^T J + JDf_i = A_i^T J^T J + JJA_i \; = A; -- A; \; = O.$$

Thus $R^T JR = J$ so that the RKM defines a symplectic mapping for each At $\geq 0$. • I

**Remarks** (i) Again this result assumes the solvability of the Runge-Kutta equations. This matter has not been investigated in detail for Hamiltonian systems and little is known in general.

(ii) Again the role of the matrix $M$ is clear: a certain quadratic form is annihilated by setting $M \equiv 0$.

(iii) Recall from section 6 that there exist methods of arbitrarily high order satisfying $M \equiv 0$; see [5].

For general nonlinear, nonintegrable Hamiltonian problems it is not possible to enforce both properties (i) and (ii) from Result 7.2 onto a numerical scheme since it would then have to be exact; see [25]. Thus it is an open and interesting question to determine the relative merits of preserving the two properties under discretisation; see for example [42] where energy-momentum conserving methods are shown to be superior to symplectic momentum conserving methods for an application in elasto-dynamics.

To discuss Hamiltonian systems in detail is well beyond the scope of this review. Here our purpose is merely to emphasise connections with other classes of problems. For a complete overview of the numerical analysis of Hamiltonian systems see [40].

## 8. Remarks on Multi-Step Methods

Throughout the paper we have concentrated on Runge-Kutta methods; this has allowed a unified exposition and the theme of algebraic stability has run throughout. Nonetheless, much of the theory for RKMs was developed in tandem with that for Linear Multi-Step and One-Leg Methods (LMMs and OLMs) and indeed in the 1960s and 1970s the theory for RKMs was often pre-dated by that for multi-step methods. Thus it is in order to briefly sketch how the theory for LMMs and OLMs fits in to that described here. An important point to appreciate is that LMMs and OLMs naturally define a dynamical system on a

space of higher dimension than the original problem — specifically in $\mathbb{R}^{pk}$ for a k-step method — and this is a source of some difficulty.

For linear decay problems the properties of A-stability and absolute stability have natural analogues for multi-step methods and indeed this was the starting point for numerical stability theory [14]. The importance of contractive problems in numerical analysis was recognized by Dahlquist in [15] in the context of multi-step methods; the notion of $G$-stability was defined for multi-step methods applied to (1. 1), (3.1) and inheriting a notion of contractivity; subsequently a remarkable equivalence Theorem, analogous to Result 3.4, was proved: G-stability is equivalent to A-stability [16]. For gradient systems there has been a little work on multi-step methods; in particular in [23] it is proved that the first three backward differentiation formulae are gradient stable, employing a natural generalization of Definition 4.3. The concept of dissipative stability is unstudied for multi-step methods. Ideas relating to conservation properties and symplectic structure for multi-step methods are considered in [21]. The preservation of orthonormality properties in matrix differential equations are studied in [18].

## 9. The Effect of Error Control

An important question which we briefly analyse here is whether the variation of time-step according to local error control will automatically enforce some form of numerical stability, even for explicit schemes. Such results are conjectured in [39] based on an illuminating study of a particular example. The question now becomes whether it is possible to find stability theories which hold for a wide range of the ***error tolerance*** $\tau$, given arbitrary initial data in $\mathbb{R}^p$. Thus $\tau$ takes on the role played by At in the remainder of the paper.

However, it is not immediately clear that such results should be true since local error control is an ***accuracy requirement*** whilst we are seeking ***stability*** results. In a notable paper, Hall [30] established a remarkable connection between accuracy and stability for error control schemes. We illustrate this with a simple example modified from [30] and [26]: consider (1.1) with $p = 1$ and

$$f(u) = -u.$$

If we apply the explicit Euler scheme with variable time-step then we obtain

$$U_{n+1} = U_n - \Delta t_n U_n$$

where the time-step At, now varies with $n$. This is a second-order accurate approximation to the true solution over a step At; that is the error is proportional to $\Delta t^2$. A third-order accurate approximation is formed with error of $\mathcal{O}(\Delta t^3)$ by calculating the first step of a Trapezoidal rule correction:

$$V_{n+1} = U_n - \frac{\Delta t_n}{2}[U_{n+1} + U_n].$$

A simple error estimate for $U_{n+1}$ is then formed as the difference between $U_{n+1}$ and $V_{n+1}$ on the assumption that At is small. The ***error per unit step*** strategy requires that At, is

chosen so that

$$\|U_{n+1} - V_{n+1}\| \leq \frac{1}{2}\Delta t_n \tau, \tag{9.1}$$

where $\tau << 1$ is an error tolerance. Under this local error control we deduce that, since the standard Euclidean norm $\|\bullet\|$ is equivalent to $|\bullet|$ in dimension $p = 1$,

$$|U_n - U_{n+1}| \leq \tau$$

is required for the step to be acceptable and hence that

$$\Delta t_n \leq \frac{\tau}{|U_n|}$$

is required. If we choose the largest time-step compatible with this error control then we obtain

$$U_{n+1} = U_n(1 - \frac{\tau}{|U_n|}), \text{ A t , } = \frac{\tau}{|U_{nI}|}. \tag{9.2}$$

Straightforward analysis shows that

$$|U_n| > \frac{\tau}{2} \Rightarrow |U_{n+1}| < |U_n|$$

whilst

$$|U_n| \leq \frac{\tau}{2} \Rightarrow |U_{n+1}| \leq \tau.$$

Using this it is possible to show that the local error control forces iterates to enter and remain in interval $[-\tau, \tau]$ about the origin; during this process the time-step approaches the linear stability limit.

This kind of desirable behaviour can be generalised to the dissipative and gradient systems studied in sections 4 and 5 – see [48] for details. Here we outline the key to that analysis which revolves around the fact that certain error control mechanisms force the RKM to behave like an algebraically stable RKM even if the underlying method is not algebraically stable in a fixed time-step implementation.

One of the simplest error control strategies for the solution of (1.1) is to take the explicit Euler scheme

$$U_{n+1} = U_n + \Delta t_n f(U_n) \tag{9.3}$$

and then form the more accurate approximation

$$V_{n+1} = U_n + \frac{\Delta t_n}{2}[f(U_n) + f(U_{n+1})]. \tag{9.4}$$

This genera&es what we did for the linear problem above. Thus the difference of $U_{n+1}$ and $V_{n+1}$ is an estimate of the error incurred in (9.3) and the error per unit step strategy then requires that At, is chosen so that (9.1) is satisfied. This implies that

$$\|f(U_n) - f(U_{n+1})\| \leq \tau \tag{9.5}$$

and hence that, under error control, the explicit scheme (9.3) is never far from the backward Euler scheme (9.3). Specifically we have that

$$U_{n+1} = U_n + \Delta t_n f\,(U_{n+1}) + \Delta t_n E,$$

where $\|E\| \leq \tau$ by (9.5). The backward Euler scheme is algebraically stable and for this reason we might expect that the error control confers desirable stability properties on the explicit scheme. This intuition is placed on a firm mathematical foundation in [48] for the contractive, gradient and dissipative problems studied here in sections 3,4 and 5.

As a second example we consider the Fehlberg (2,3) method given by

$$\eta_1 = U_n,$$

$$\eta_2 = U_n + \Delta t_n f(\eta_1),$$

$$\eta_3 = U_n + \frac{\Delta t_n}{4}[f(\eta_1) + f(\eta_2)],$$

$$U_{n+1} = U_n + \frac{\Delta t_n}{2}[f(\eta_1) + f(\eta_2)],$$

$$V_{n+1} = U_n + \frac{\Delta t_n}{6}[f(\eta_1) + f(\eta_2)] + \frac{2\Delta t_n}{3} f(\eta_3).$$

We define $\tilde{f}$ by

$$\tilde{f}(U_n; \Delta t_n) = \frac{1}{2}[f(\eta_1) + f(\eta_2)]$$

and thus

$$U_{n+1} = U_n + \Delta t_n \tilde{f}(U_n; \Delta t_n);$$

note that

$$V_{n+1} = U_n + \frac{\Delta t_n}{3}\tilde{f}(U_n; \Delta t_n) + \frac{2\Delta t_n}{3} f(\eta_3).$$

Thus the error control

$$\|U_{n+1} - V_{n+1}\| \leq \frac{2}{3}\Delta t_n \tau$$

implies that

$$\|\tilde{f}(U_n) - f(\eta_3)\| \leq \tau.$$

Then, since

$$\eta_3 = \frac{1}{2}[U_{n+1} + U_n],$$

we deduce that the error control ensures that the explicit Fehlberg (2,3) pair is close to the implicit mid-point rule. Specifically we have that

$$U_{n+1} = U_n + \Delta t_n f\left(\frac{U_{n+1} + U_n}{2}\right) + \Delta t_n E$$

where $\|E\| \leq \tau$. The mid-point rule is also an algebraically stable scheme and, in fact satisfies $M \equiv 0$. Again, desirable stability properties follow for dissipative problems (see [48]) and we might expect relatively good behaviour for the conservative and Hamiltonian problems considered in sections 6 and 7. The effect of error control on numerical methods for Hamiltonian systems is studied in [8].

These two examples are incorporated in a general framework of algebraically stable RKM pairs in [48]. This facilitates an anlysis of the the behaviour of variable time step Runge-Kutta methods on contractive, dissipative and gradient systems.

## 10. Conclusions and Open Problems

It will be clear from reading this article that the numerical stability theory for problems in sections 4-9 is far from complete. Nonetheless, it should also be clear that the classes of problems in sections 4-9 all form a natural progression from the simple problems in sections 2 and 3. Furthermore, the problems in sections 4-9 arise in many applications and admit a variety of ***interesting and complicated dynamical features*** ranging from multiple competing equilibria, through dissipative chaos to conservative systems and finally Hamiltonian systems which can exhibit both integrability and Hamiltonian chaos. In contrast, the classical problems discussed in sections 2 and 3 admit only ***trivial dynamics.*** Our main aim is to emphasise this point and at the same time to motivate and encourage study of the broader classes of problems suggested. An important point is that there are clear indications of connections in the numerical stability theory for the problems admitting trivial dynamics and those admitting complicated dynamics. In particular, algebraic stability plays a fundamental role. With this in mind we make a subjective list of open problems:

- To explore the class of numerical methods which are gradient stable in the sense of Definition 4.2. Currently only the theta method ($\theta \in [\frac{1}{2}, 1]$ – see section 4), the first three BDFs [23] and a modification of Crank-Nicolson ([22], [20]) have been identified. There are many open questions for RKMs, LMMs and OLMs.

- To determine whether A(o)-stability is sufficient for gradient stability; this is a reasonable conjecture to make since the Jacobian of a gradient system is symmetric.

- To determine whether Definition **4.2** is appropriate; there may be RKMs which do not satisfy part (ii) of the definition on a step-by-step basis, but may satisfy an analogous property over several steps or for all sufficiently large $n$. In any case the crucial point is that conditions are required under which convergence to an equilibrium is guaranteed.

- To explore the class of LMMs which are dissipative stable for the problems defined by (1.1),(5.3). To frame a suitable definition of dissipative stable for LMMs and OLMs and to identify schemes in this class.

- To extend the analysis of dissipative problems induced by the inner product structure (5.3) to others fitting into the precise notion of ***point dissipative*** as defined in [29]. The work of [43], [45] and [36] on contractivity in non-inner-product norms is of interest in this context.

- To examine the affect of numerical approximation on Liapunov functionals other than those arising in sections 4 and 5.

- To assess the relative merits of Hamiltonian conserving algorithms which preserve the property of Result 7.1(i), and symplectic algorithms which inherit the property of Result 7.2(ii). In particular it is of interest to determine what can be said about the behaviour of the Hamiltonian for symplectic schemes.

- To impose structural assumptions on the Hamiltonian $H$ in an attempt to further assess

symplectic and conserving algorithmss. For example, the convexity of $H$ is important in some applications and its implications for numerical algorithms should be studied in detail. As another example, consider the class of Hamiltonian systems satisfying the following property: for any $c_1$ there exists $c_2$ such that $H(u) = c_1 \Rightarrow \|u\| \leq c_2$. This implies global existence and boundedness of trajectories and again, its implications for numerical algorithms should be investigated.

- To develop rational numerical stability theories for variable time-stepping algorithms, in particular for the classes of problems outlined in this paper. Some work in this direction may be found in [48], [8]. In particular it is of interest that typical software codes including local error control lead to discontinuous dynamical systems since the time-step sequence chosen may change discontinuously as a function of the initial data. Thus new mathematical machinery is required to analyse this problem.

- To identify other classes of problems motivated by either by real applications or by a need for theoretical understanding of the differential equations, for which it is beneficial to develop numerical stability theories. The need for applications to differential equations should be directing numerical stability theory; to date this has not always been the case.

Finally we conclude with a disclaimer: it is not our purpose to completely review the subject of numerical stability theory for initial value problems. We have concentrated on the mathematical properties of the underlying problems and this has been our unifying theme. For this reason there are numerous references to related work in the numerical analysis literature that have not been made here.

## Acknowledgements

The work of AMS is funded by the Office of Naval Research under grant N00014-92-J-1876 and by the National Science Foundation under grant DMS-9201727. ARH is grateful to the Science and Engineering Research Council, UK, and to Stanford University, USA for financial support. Both authors are grateful to Luca Deici, Kjell Gustafsson, Arieh Iserles, Bob Russell, Juan Simo and Marc Spijker for helpful conversations.

## References

[1] V.I. Arnold. ***Mathematical Methods of Classical Mechanics.*** Springer, New York, 1978.

[2] N.P. Bhatia and G.P. ***Szego. Stability Theory of Dynamical Systems.*** Springer, New York, 1970.

[3] K. Burrage. High order algebraically stable Runge-Kutta methods. ***BIT,*** 18:373–383, 1978.

[4] K. Burrage and J.C. Butcher. Stability criteria for implicit Runge-Kutta processes. ***SIAM J. Num. Anal.,*** 16:46–57, 1979.

[5] J.C. Butcher. Implicit Runge-Kutta processes. ***Math. Comp.,*** 18:50–64, 1964.

[6] J.C. Butcher. A stability property of implicit Runge-Kutta methods. ***BIT,*** 15:358–361, 1975.

[7] J.C. Butcher. ***The Numerical Analysis of Ordinary Differential Equations.*** Wiley, Chichester, 1987.

[8] M.P. Calvo and J.M. Sanz-Serna. The development of variable-step symplectic integrators, with applications to the two-body problem. ***In Proceedings of the 14th Dundee Conference on Numerical Analysis,*** London, 1991. Pitman.

[9] J. Chapman, S.H. Howison, and J.R. Ockendon. Macroscopic models for superconductivity. ***SIAM Review,*** To appear, 1993.

[10] G. Cooper. On the existence of algebraically stable Runge-Kutta methods. ***IMA J. Num. Anal.,*** 6:325–330, 1986.

[11] G. Cooper. Stability of Runge-Kutta methods for trajectory problems. ***IMA J. Num. Anal.,*** 7:1–13, 1987.

[12] G. Cooper. Weak nonlinear stability of implicit Runge-Kutta methods. ***IMA J. Num. Anal.,*** 12:57–68, 1992.

[13] M. Crouziex. Sur la B-stabilite des methodes de Runge-Kutta. ***Numer. Math., 32:75-*** 82, 1979.

[14] G. Dahlquist. A special stability problem for linear multistep methods. ***BIT,*** 3:27–43, 1963.

[15] G. Dahlquist. Error analysis for a class of methods for stiff non-linear initial value problems. ***In Numerical Analysis, Dundee 1975,*** pages 60-74. Springer, 1975.

[16] G. Dahlquist. G-stability is equivalent to A-stability. *BIT,* 18:384–401, 1978.

[17] G. Dahlquist and R. Jeltsch. Generalized disks of contractivity for explicit and implicit Runge-Kutta methods. ***TRITA-NA-7906, Dept. Numer. Anal. and Comp. Sci., Stokholm.,*** 1979.

[18] L. Deici, R.D. Russell, and E. van Vleick. Title. Where., Submitted:-, 1992.

[19] K. Dekker and J.G. Verwer. ***Stability of Runge-Kutta Methods for Stiff Nonlinear Equations.*** North Holland, Amsterdam, 1984.

[20] Q. Du and R.A. Nicolaides. Numerical analysis of a continuum model of phase transition. ***SIAM J. Num. Anal.,*** 28:1310–1322, 1991.

[21] T. Eirola and J.M. Sanz-Serna. Conservation of integrals and symplectic structure of differential equations by multistep methods. ***Numer. Math.,*** 61:281–290, 1992.

[22] C.M. Elliott. The Cahn-Hilliard model for the kinetics of phase seperation. In J.F. Rodrigues, editor, ***Mathematical models for phase change problems.*** Birkhauser, 1988.

[23] C.M. Elliott and A.M. Stuart. The global dynamics of semilinear parabolic equations. ***Submitted to SIAM J. Num. Anal., 1992.***

[24] D. French and S. Jensen. Long-time behaviour of arbitrary order continuous time galerkin schemes for some on-dimensional phase transition problems. ***Submitted to IMA J. Num. Anal.,*** ?:?, 1992.

[25] Z. Ge and J.E. Marsden. Lie-Poisson Hamilton-Jacobi theory. ***Phys. Lett. A.,*** 133:134-139, 1988.

[26] D.F. Griffiths. The dynamics of some linear multistep methods with step-size control. In ***Proc. 12 Biennial Dundee Conference on Numerical Analysis.*** Pitman, London, 1987.

[27] E. Hairer, A. Iserles, and J.M. Sanz-Serna. Equilibria of Runge-Kutta methods. ***Numer. Math.,*** 58:243–254, ***1990.***

[28] E. Hairer and G. Wanner. ***Solving Ordinary Differential Equations II: Stiff Problems.*** Springer-Verlag, Berlin, 1991.

[29] J.K. Hale. ***Asymptotic behaviour of dissipative systems.*** American Mathematical Society, Providence, 1988.

[30] G. Hall. Equilibrium states of Runge-Kutta schemes. ***ACM Trans on Math. Software,*** 11:289–301, 1985.

[31] M.W. Hirsch and S. Smale. ***Differential Equations, Dynamical Systems and Linear Algebra.*** Academic Press, London, 1974.

[32] A.R.. Humphries. ***The Dynamics of Numerical Methods for Dynamical Systems.*** PhD thesis, University of Bath, 1993.

[33] A.R. Humphries and A.M. Stuart. Runge-Kutta methods for dissipative and gradient dynamical systems, 1992. In Preparation.

[34] W.H. Hundsdorfer and M.N. Spijker. A note on B-stability of Runge-Kutta methods. ***Numer. Math.,*** 36:319–331, 1981.

[35] A. Iserles. Stability and dynamics of numerical methods for nonlinear ordinary differential equations. *IMA J. Num. Anal.,* lO:l-30, 1990.

[36] J.F.B.M. Kraaijevanger. Contractivity of Runge-Kutta methods. Preprint TW-89-12, University of Leiden, 1989.

[37] F.M. Lasagni. Canonical runge-kutta methods. ***Journal of Applied Mathematics and Physics.,*** 39:951–953, 1988.

[38] J.M. Sanz-Serna. Runge-Kutta schemes for Hamiltonian systems. ***BIT,*** 28:877–883, 1988.

[39] J.M. Sanz-Serna. Runge-Kutta schemes for Hamiltonian systems. ***BIT,*** 28:877–883, 1988.

[40] J.M. Sanz-Serna. Symplectic integrators for Hamiltonian problems: an overview. ***Acta Numerica, 1, 1992.***

[41] R. Schere and H. Turke. Albegraic charaterization of *a-* stable Runge-Kutta methos. ***Appl. Num. Math.,*** 5:133–144, 1989.

[42] J.C. Simo and N. Tarnow. The discrete energy-momentum method. Part I. conserving algorithms for elastodynamics, 1992. Stanford University Division of Applied Mechanics Report 92-3.

[43] M.N. Spijker. Contractivity in the numerical solution of initial value problems. ***Numer. Math.,*** 42:271–290, 1983.

[44] M.N. Spijker. Monotonicity and boundedness in implicit Runge-Kutta methods. ***Numer. Math.,*** 42:271–290, 1983.

[45] M.N. Spijker. Stepsize restrictions for stability of one-step methods in the numerical solution of initial value problems. ***Math. Comp.,*** 45:377–392, 1985.

[46] M.N. Spijker. A note on contractivity in the numerical solution of initial value problems. ***BIT,*** 27:424–437, 1987.

[47] A.M. Stuart. Linear instability implies spurious periodic solutions. ***IMA J. Num. Anal.,*** 9:465–486, 1989.

[48] A.M. Stuart and A.R. Humphries. An analysis of local error control for dissipative and gradient dynamical systems, 1992. In Preparation.

[49] R. Temam. ***Infinite Dimensional Dynamical Systems in Mechanics and Physics.*** Springer, New York, 1989.