Acta Numerica (1994), *pp.* 467–572

Numerical analysis of dynamical systems

Andrew M. Stuart
Program in Scientific Computing and Computational Mathematics
Division of Applied Mechanics
Stanford University
California, CA94305-4040, USA
E-mail: stuart@sccm.stanford.edu

This article reviews the application of various notions from the theory of dynamical systems to the analysis of numerical approximation of initial value problems over long-time intervals. Standard error estimates comparing individual trajectories are of no direct use in this context since the error constant typically grows like the exponential of the time interval under consideration.

Instead of comparing trajectories, the effect of discretization on various sets which are invariant under the evolution of the underlying differential equation is studied. Such invariant sets are crucial in determining long-time dynamics. The particular invariant sets which are studied are equilibrium points, together with their unstable manifolds and local phase portraits, periodic solutions, quasi-periodic solutions and strange attractors.

Particular attention is paid to the development of a unified theory and to the development of an existence theory for invariant sets of the underlying differential equation which may be used directly to construct an analogous existence theory (and hence a simple approximation theory) for the numerical method.

CONTENTS

1. Introduction

In this article we study the numerical approximation of the ordinary differential equation

$$u_t = f(u), \quad u(0) = U, \tag{1.1}$$

for $u(t) \in C^1(\mathbb{R}^+, \mathbb{R}^p)$. We introduce a time discretization through the points $t_n = n\Delta t$ and study the approximation of (1.1) by one-step numerical methods of the form

$$U_{n+1} = \mathcal{F}(U_n; \Delta t), \quad U_0 = U. \tag{1.2}$$

Here $U_n \in \mathbb{R}^p$ is considered as an approximation to $u(t_n)$. All Runge–Kutta methods, for example, can be considered in this form, provided solvability of the defining equations has been established.

The classical error bound for the approximation of (1.1) by (1.2) is of the form

$$\|U_n - u(t_n)\| \leq ce^{kT}\Delta t^r \tag{1.3}$$

for $0 \leq t_n \leq T$. Such error bounds can be derived purely under the assumptions which yield existence, uniqueness, smoothness and continuous dependence of a solution to (1.1) and no further understanding of the behaviour of solutions is required. Typically $k > 0$ reflecting the fact that different solutions of (1.1) may diverge exponentially over certain parts of phase space. Consequently the error bound (1.3) is of little direct use in studying the long-time behaviour of approximations (1.2) for the equation (1.1) since it yields no information for fixed Δt as $T \to \infty$.

To understand the behaviour of the approximation (1.2) of (1.1) over long-time intervals requires a deeper knowledge of the behaviour of solutions of (1.1) and, in particular, an understanding of how these solutions behave over long-time intervals. In particular, the study of a variety of sets invariant under the evolution generated by the equation (1.1) is crucial. Such knowledge, combined with the standard error estimate (1.3) or a truncation error bound, can provide very powerful results about the long-time behaviour of numerical methods. The purpose of this review is to describe such results within the unified framework of dynamical systems.

Section 2 contains background and motivational material. In particular we state what we aim to show in this article and describe the types of problems that we have in mind. In so doing we also make it clear that a number of important issues will *not be covered* and give appropriate references to existing literature in these areas. We also describe, by means of simple examples, the types of theorems whose statements and detailed proofs we consider in the remainder of the article.

In Section 3 we formulate the basic notions from the theory of dynamical systems that are relevant to this article; in particular the concept of semigroup $S(t)$ for (1.1) is introduced. In addition, the basic assumptions concerning relationships between the true semigroup $S(t)$ and approximate semigroup $S^n_{\Delta t}$ for (1.2) are spelled out and certain convergence results proved for individual trajectories.

In the remaining three sections we study the existence and convergence properties of a variety of objects under discretization. Since the focus of the article is on convergence, and a rather general framework for this question is considered, we do not distinguish between the practical value of different methods. Rather we address the question 'what meaning can be attached to computations performed with arbitrary finite time convergent numerical methods when used over very long-time intervals?'

In all three sections the basic format is the same: the introduction is concerned with the development of an existence theory for the invariant sets of (1.1) itself, whilst the remaining sections contain modifications of this theory for the numerical approximation and the derivation of error bounds. Wherever possible, the existence theory is developed in such a way as to be directly applicable to both the equation (1.1) and its approximation (1.2). For this reason the existence theory is formulated in terms of the time Δt evolution of the equation (1.1).

Consequently it is true that, in many cases, much of the work involved in proving results about numerical approximation is concerned with formulating existing theories from continuous dynamical systems in a form amenable to the study of discrete maps arising in numerical analysis; this involves a fair amount of rehashing of well known theories in dynamical systems but is a fruitful process since the approximation theory for the numerical method then falls out in a relatively straightforward manner. Note that, in some cases, we will develop several approaches to the same question. In particular we provide two alternative constructions and convergence proofs for the stable and unstable manifolds, for uniformly asymptotically stable sets and for attractors.

In Section 4 we examine the behaviour of approximations $S^n_{\Delta t}$ in the neighbourhood of an equilibrium point of $S(t)$. We are led to study the existence and convergence of an approximate equilibrium point, the convergence of stable and unstable manifolds of the equilibrium point and the convergence of phase portraits near to the equilibrium point.

In Section 5 we study periodic solutions of $S(t)$ under approximation by $S^n_{\Delta t}$. We show that, under a condition which ensures that the periodic solution is isolated, the semigroup $S^n_{\Delta t}$ has a closed invariant curve which converges, in the sense of sets, to the periodic solution of $S(t)$. We also discuss briefly the effect of discretization on quasi-periodic solutions (the sum of two irrationally related periodic solutions).

In Section 6 we study uniformly asymptotically stable sets and attractors; these objects may include, for example, strange attractors such as those observed in the Lorenz equations. Again we study existence and convergence of approximations of these objects found in the numerical scheme.

2. Background and motivation

In order to motivate the material in the remainder of the article, we start by relating the approach taken here to the classical theory of numerical analysis of initial value problems. Broadly speaking, the two fundamental issues in the classical study of the approximation of (1.1) by (1.2) are *convergence* and *stability*; we consider these two issues in turn and discuss how they might be generalized to the consideration of nonlinear dynamical systems over long-time intervals.

2.1. Convergence

As mentioned in the introduction, standard error bounds relating (1.1) and (1.2) are of the form (1.3). This bound reflects the exponential divergence of trajectories that may be present in well posed problems of the form (1.1). Since most problems do not exhibit exponential divergence throughout the whole of phase space, the bound (1.3) can sometimes be improved upon in a number of ways: (a) for equations (1.1) exhibiting exponential contraction of trajectories throughout phase space, or asymptotically as $t \to \infty$, k may be negative yielding uniform convergence of trajectories for $t \in [0, \infty)$; (b) for equations with conserved quantities, such as Hamiltonian systems, the error bound (1.3) can sometimes be weakened to

$$\|U_n - u(t_n)\| \leq cT^{\alpha}\Delta t^r$$

for $0 \leq t_n \leq T$, for some $\alpha > 0$; (c) for equations whose solution and approximation ultimately lie in a bounded set B the error estimate (1.3) is clearly pessimistic as $t \to \infty$ and can trivially be replaced by

$$\|U_n - u(t_n)\| \leq \text{diam}(B)$$

for t sufficiently large, where diam(B) denotes the largest distsnce between any two points in B.

Possibility (a) is of *minor interest* since it admits only convergence to a stable equilibrium point as $t \to \infty$ and therefore rules out many applications involving interesting dynamical behaviour; it is discussed briefly in Section 3. Possibility (b) *is of interest* and some results in this direction are described in Calvo and Sanz-Serna (1992; 1993a,b); however, the techniques of use in that case are rather specialized to Hamiltonian and other conservative systems, an area which is extensively reviewed in, for example, Sanz-Serna (1992a). It is possibility (c) with which we shall concern ourselves in this article.

Of course, there are important application areas where unbounded solutions are of relevance. Furthermore, in many cases it is not just the asymptotic behaviour which is of interest but also the transient behaviour. However, there *are many applications* in which bounded asymptotic behaviour is of paramount importance and here we concentrate on such situations; we do not study the approximation of unbounded trajectories nor do we study the approximation of transients in any detail. Thus we suppose that solutions of both (1.1) and (1.2) ultimately lie in some bounded set B. Within B the solution will typically approach an *ω-limit set* as $t \to \infty$. This might be, for example, an equilibrium point, a periodic solution, a quasi-periodic solution or a strange attractor. Examples of the four possibilities are given in Figure 1. The four objects observed as $t \to \infty$ are all examples of ω-limit sets; a precise definition is given in Section 3 but, roughly, they are objects observed for large t in (1.1).

A natural generalization of the standard convergence question to this situation is to ask *are ω-limit sets of* (1.1) *well approximated by ω-limit sets of* (1.2)? It is predominantly this question, and others closely related to it, that we address in this article. To introduce ideas concerned with convergence of limit sets, and other related sets, we describe six examples.

Examples

(i) *Equilibrium points.* The explicit Euler scheme for the approximation of (1.1) is

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

Its fixed points satisfy

$$\bar{U} = \bar{U} + \Delta t f(\bar{U}) \Leftrightarrow f(\bar{U}) = 0, \quad \forall \Delta t > 0.$$

Hence they coincide with the equilibrium points of (1.1). Thus convergence of these limit sets (equilibrium solutions) as $\Delta t \to 0$ is trivial. It is worth noting, however, that general Runge–Kutta methods may produce *spurious fixed points* which are not close to the true equilibria as $\Delta t \to 0$ – this point was first observed in Iserles (1990); see Theorem 4.11 and the example preceding it. In Section 4 we shall consider the questions of convergence of fixed points to true equilibria, and the existence of spurious solutions, under fairly weak hypotheses on the nature of the approximation – see Theorems 4.10 and 4.11.

(ii) *Unstable manifolds.* Consider the pair of equations

$$p_t = p, \quad q_t = -q + p^2. \tag{2.1}$$

These equations have the equilibrium point $p = q = 0$. Now consider the curve $q = p^2/3$. If we define the variable $z = q - p^2/3$ then

$$z_t = q_t - \tfrac{2}{3}pp_t = -q + p^2 - \tfrac{2}{3}p^2 = -z. \tag{2.2}$$

Hence, if $z(0) = 0$ then $z(t) = 0 \; \forall t \in \mathbb{R}$. Thus the curve

$$q = \tfrac{1}{3}p^2 \tag{2.3}$$

is invariant for the equations – solutions starting on the curve remain on it. Furthermore, using (2.1), we find that if (2.3) holds then

$$p(t) = Ae^t, \quad q = \tfrac{1}{3}A^2e^{2t};$$

thus $p(t), q(t) \to 0$ as $t \to -\infty$ for solutions on the invariant curve; this curve is referred to as *the unstable manifold* of the origin.

Now consider the Euler approximation

$$p_{n+1} = (1 + \Delta t)p_n, \quad q_{n+1} = (1 - \Delta t)q_n + \Delta t p_n^2. \tag{2.4}$$

It is natural to seek an invariant curve of the same form as for the differential equation. Specifically, we seek an $a \in \mathbb{R}$ such that

$$q_n = ap_n^2 \Leftrightarrow q_{n+1} = ap_{n+1}^2.$$

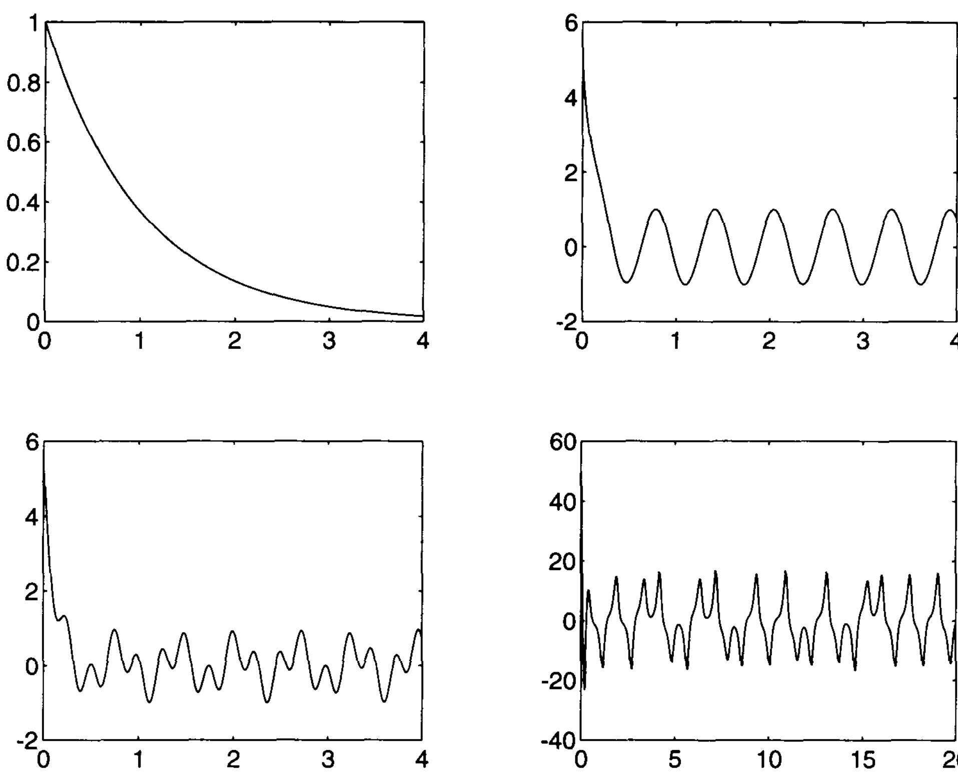

Fig. 1. Plot of a component of $u(t)$ (vertical axis) against t illustrating four different limit sets; from left to right, top to bottom: an equilibrium point, a periodic solution, a quasi-periodic solution, a chaotic solution.

From (2.4) we see that this implies that

$$(1-\Delta t)ap_n^2 + \Delta t p_n^2 = q_{n+1} = ap_{n+1}^2 = a(1+\Delta t)^2 p_n^2.$$

Hence, since this must be satisfied for all p_n, we obtain

$$(1-\Delta t)a + \Delta t = a(1 + 2\Delta t + \Delta t^2)$$

which yields $a = (3+\Delta t)^{-1}$. Thus the curve

$$q = \frac{1}{3+\Delta t}p^2 \tag{2.5}$$

is invariant for the numerical method. Furthermore, on (2.5),

$$p_n = (1+\Delta t)^n p_0, \quad q_n = \frac{(1+\Delta t)^{2n}}{3+\Delta t}p_0^2$$

so that $p_n, q_n \to 0$ as $n \to -\infty$. This demonstrates that (2.5) is the unstable manifold of the origin for the Euler approximation.

The important observation to make from this example is that both the underlying equations and their approximation have an invariant manifold and that, furthermore, by virtue of (2.3) and (2.5), these manifolds are close as $\Delta t \to 0$. Generalizations of this idea will be proved in Theorem 4.12 and Corollary 4.13.

(iii) *Phase portraits.* Consider approximation of the equation

$$x_t = -x, \quad x(0) = x_0,$$

by the explicit Euler scheme

$$X_{n+1} = (1-\Delta t)X_n, \quad X_0 = x_0.$$

It is straightforward to show that there exists $C > 0$ so that the error satisfies

$$|X_n - x(t_n)| \leq C\Delta t[1-(1-\Delta t)^n] \leq C\Delta t \quad \forall n, \Delta t : 0 \leq n\Delta t < \infty,$$

for Δt sufficiently small. Thus, due to the exponential contraction of the true solution, this error bound is uniform in T for $0 \leq t_n \leq T$.

The equations

$$x_t = -x, \quad y_t = y,$$

are illustrative of the behaviour of trajectories of (1.1) in the neighbourhood of an equilibrium point. Because of the exponential divergence present in the y-component of the solution, standard error bounds are not uniform with respect to T on the time interval $0 \leq t \leq T$. However, if we consider numerical trajectories with different initial conditions from the true trajectories *it is possible* to find approximations of solutions which are uniform with respect to the length of time interval T. The key point is that the y-equation *is contractive backwards in time.* We choose the initial condition in y so that the true and numerical solutions agree at $t_n = T$ and then exploit backwards contractivity in y. As an example, consider solving the differential equations subject to the boundary conditions

$$x(0) = \eta, \quad y(T) = \xi, \quad T = N\Delta t.$$

Thus the true solution to the problem is

$$x(0) = \eta \mathrm{e}^{-t}, \; y(t) = \xi \mathrm{e}^{t-T}.$$

The backward Euler scheme scheme

$$X_{n+1} = (1 - \Delta t)X_n, \quad X_0 = \eta$$

$$Y_{n+1} = (1 + \Delta t)Y_n, \quad Y_N = \xi$$

gives the solution

$$X_n = \eta(1 - \Delta t)^{-n}, \; Y_n = (1 + \Delta t)^{n-N}\xi.$$

It can then be shown that there exists $C > 0$ such that the errors satisfy

$$|X_n - x(t_n)| \leq C\Delta t[1 - (1 - \Delta t)^n] \leq C\Delta t \quad \forall n, \Delta t : 0 \leq n\Delta t \leq T$$

and

$$|Y_n - y(t_n)| \leq C\Delta t[1 - (1 + \Delta t)^{n-N}] \leq C\Delta t \quad \forall n, \Delta t : 0 \leq n\Delta t \leq T$$

for Δt sufficiently small. Again C is independent of T. Thus, by comparing suitably chosen solutions, it is possible to find error bounds which *do not depend on the length of the time interval* for problems exhibiting exponential divergence of trajectories. The basic idea described here was described for a general class of linear problems in Enquist (1969). The idea can be generalized to a wide class of nonlinear problems in the neighbourhood of an equilibrium point; this we show in Section 4 – see Theorem 4.14 and Corollary 4.15.

(iv) *Periodic solutions.* Consider the complex equation ($\mathrm{i}^2 = -1$)

$$z_t = (\alpha\,\mathrm{i} + 1 - |z|^2)z$$

with periodic solution $z(t) = \mathrm{e}^{\alpha\,\mathrm{i}t}$. As a set of points in $\mathbb{C}$ the periodic solution is given by

$$\mathcal{P} := \{z \in \mathbb{C} : |z| = 1\}. \tag{2.6}$$

The explicit Euler approximation yields the map

$$Z_{n+1} = Z_n + \Delta t(\alpha\,\mathrm{i} + 1 - |Z_n|^2)Z_n.$$

The analogue of the periodic solution of the differential equation is to seek a circle in the complex plane which is *invariant* under the maps – that is a circle with the property that points starting on the circle remain on the circle. (A precise definition of invariant will be given in Section 3.) Thus we seek fixed points of the map $|Z_n|^2 \mapsto |Z_{n+1}|^2$. A little algebra shows that there is a circle of fixed points with the form

$$|Z_n| = R_-(\alpha)$$

where

$$R_\pm(\alpha)^2 = 1 + \frac{1 \pm [1 - \alpha^2 \Delta t^2]^{1/2}}{\Delta t}. \tag{2.7}$$

Hence the mapping has the invariant circle

$$\mathcal{P}_{\Delta t} := \{z \in \mathbb{C} \colon |z| = R_-(\alpha)\}. \tag{2.8}$$

Noting that $R_-(\alpha) = 1 + \mathcal{O}(\Delta t)$, we deduce from (2.6) and (2.8) that the set $\mathcal{P}_{\Delta t}$ converges to $\mathcal{P}$ as $\Delta t \to 0$. Such a convergence result for periodic solutions is true under much more general circumstances and this is investigated further in Section 5 – see Theorem 5.7 and Corollary 5.8. Note, however, that the numerical method also has a spurious limit set in the form of the invariant circle $|Z_n| = R_+(\alpha) = \mathcal{O}(\Delta t^{-1/2})$. The example constructed here was introduced in Brezzi *et al.* (1984).

We introduce a brief note of caution concerning the approximation of periodic solutions by numerical methods. In order to generalize the example considered to other periodic solutions and other numerical methods, it is necessary to assume that the periodic solution is isolated in phase space (no other periodic solutions arbitrarily close to it). To illustrate why this is necessary, consider the equations

$$x_t = -y, \quad y_t = x$$

with periodic solutions

$$x(t) = A\cos(t), \;\; y(t) = A\sin(t), \quad A \in \mathbb{R}$$

Since A is arbitrary these solutions are not isolated. Furthermore, for the explicit Euler scheme

$$X_{n+1} = X_n - \Delta t Y_n, \quad Y_{n+1} = Y_n + \Delta t X_n,$$

all solutions satisfy

$$[X_n^2 + Y_n^2] = (1 + \Delta t^2)^n [X_0^2 + Y_0^2];$$

thus, unless the initial data are at the origin,

$$X_n^2 + Y_n^2 \to \infty, \quad n \to \infty.$$

Thus no closed invariant curves approximating a periodic solution can exist. Issues of this nature are encountered frequently in the approximation of Hamiltonian and other conservative systems; as mentioned earlier, this is a somewhat separate subject area which we will not address in this article.

(v) *Quasi-periodic solutions.* Consider the coupled complex equations

$$z_t = (\mathrm{i} + 1 - |w|^2)z,$$

$$w_t = (\sqrt{2}\,\mathrm{i} + 1 - |z|^2)w.$$

Note that the equations admit the solution $z(t) = \mathrm{e}^{\mathrm{i}t}, w(t) = \mathrm{e}^{\sqrt{2}\,\mathrm{i}t}$. This is a quasi-periodic solution of the coupled system and, as a set, it may be written

$$\mathcal{Q} = \{z, w \in \mathbb{C}\colon |z| = |w| = 1\}. \tag{2.9}$$

The explicit Euler scheme for these equations is

$$Z_{n+1} = Z_n + \Delta t(\mathrm{i} + 1 - |W_n|^2)Z_n,$$

$$W_{n+1} = W_n + \Delta t(\sqrt{2}\,\mathrm{i} + 1 - |Z_n|^2)W_n.$$

By analogy with the continuous solution, we seek an invariant set with the form

$$|Z_{n+1}| = |Z_n|, \quad |W_{n+1}| = |W_n|, \quad \forall n \geq 0.$$

A calculation shows that such an invariant set may be found with the form

$$|Z_n| = R_-(\sqrt{2}), \quad |W_n| = R_-(1),$$

where $R_-(\alpha)$ is given by (2.7). Noting that $R_-(\alpha) = 1 + \mathcal{O}(\Delta t)$ we deduce that the numerical method has an invariant set

$$\mathcal{Q}_{\Delta t} = \{z, w \in \mathbb{C}\colon |z| = R_-(\sqrt{2}), |w| = R_-(1)\},$$

which converges to the true invariant set $\mathcal{Q}$ given by (2.9) as $\Delta t \to 0$. Note that the example constructed here is simply a modification of that described for periodic solutions in the previous example. Spurious invariant sets can be constructed by choosing the root $R_+(\cdot)$ in the construction of the invariant set.

Once again, in order to generalize this example, it is crucial to require that the quasi-periodic solution be isolated. The convergence of quasi-periodic solutions is discussed in Section 5.

(vi) *Strange attractors.* Consider the Lorenz equations (Lorenz, 1963)

$$\begin{aligned} x_t &= \sigma(y - x), \\ y_t &= rx - y - xz, \\ z_t &= xy - bz. \end{aligned} \tag{2.10}$$

Figure 2 shows solutions of the equations, with parameters set at $\sigma = 10, r = 28$ and $b = \frac{8}{3}$, for four entirely different initial conditions. Note that, in all cases, the solutions are attracted to a very complicated set in $\mathbb{R}^3$ and this set is an example of a *strange attractor.* The same set is observed in all four cases.

This complicated set is observed for almost all initial conditions chosen and forms the ω-limit set for equations (2.10) for almost all initial data. The issues involved in proving convergence of such strange attractors are far more complicated than for equilibrium points and periodic solutions; no simple illustrations can be constructed. Indeed one of the stumbling blocks in the numerical analysis of such objects is that the existing theory of perturbations to such strange attractors is itself far from fully developed. The convergence of such attractors, and other related objects, is considered in Section 6. See Theorems 6.12, 6.20, 6.21, 6.22, 6.26 and Corollaries 6.18 and 6.30.

2.2. Stability

In the previous section we discussed the notion of convergence and described a particular generalization that is useful in the study of dynamical systems – namely to look at the existence of limit sets (and other related objects) and then study their convergence as $\Delta t \to 0$. Such a convergence study will

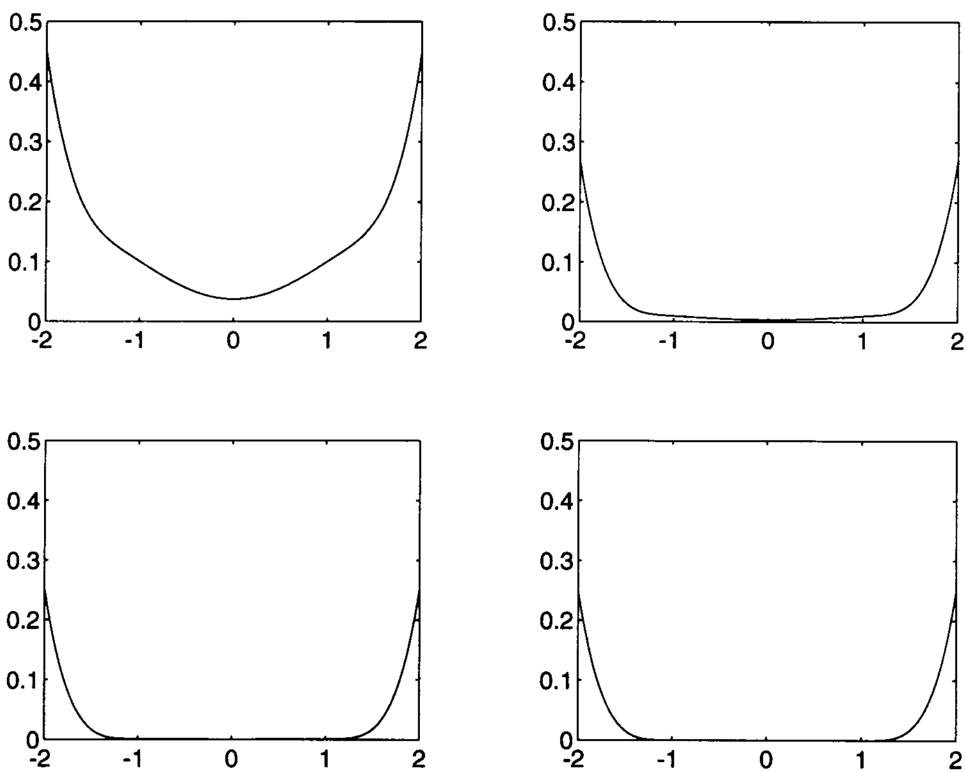

Fig. 2. Plot of x (vertical axis) against y for four different solutions of the Lorenz equations, with $\sigma = 10, r = 28, b = \frac{8}{3}$.

form the majority of the remainder of the article. However, to set things in context, we briefly discuss notions of stability appropriate to the study of dynamical systems.

Classical linear stability theory is concerned with the analysis of approximating the equation

$$u_t = \lambda u, \quad u(0) = U, \quad \mathrm{Re}\,(\lambda) < 0. \tag{2.11}$$

(For the study of linear conservative or Hamiltonian problems it is important to include the effect of approximating this problem for purely imaginary values of λ – the equation $u_t = iu$ is the archetypal example of a Hamiltonian equation. However, such problems are not our concern here and so we restrict attention to the case where $\mathrm{Re}\,(\lambda) < 0$.)

The approximation (1.2) to (2.11) is typically a rational function of $\lambda\Delta t$; for a given numerical method applied to (2.11), the *region of absolute stability* $\mathcal{S} \subseteq \mathbb{C}$ is defined to be the set with the property that

$$\lambda\Delta t \in \mathcal{S} \Leftrightarrow |U_n| \to 0 \text{ as } n \to \infty.$$

Thus, if $\lambda\Delta t \in \mathcal{S}$, the numerical solution replicates the behaviour of the underlying equation (2.11). In many circumstances it is important that this occurs without restriction on Δt. Hence the following definition is useful: the numerical method is said to be *A-stable* if

$$\{z \in \mathbb{C}: \mathrm{Re}\,(z) < 0\} \subseteq \mathcal{S}.$$

One abstraction of the above concepts of stability (which we can generalize to other problems) is that they yield *conditions under which an important qualitative property of the equation is inherited by the discretization.* (This abstraction misses the important connection between such practical stability conditions and error propagation in stiff problems, but is nonetheless a useful notion.) With this abstraction in mind, let us consider nonlinear problems.

A fundamental difference between linear and nonlinear problems is that, in the latter case, the stability notions become *initial data dependent.* Consider the approximation (1.2) of a nonlinear problem of the form (1.1) with the property that all solutions tend to the origin as $t \to \infty$. In this case we could define the *region of absolute stability* of a given method to be the set $\mathcal{S} \subseteq \mathbb{R} \times \mathbb{R}^p$ with the property that

$$(\Delta t, U_0) \in \mathcal{S} \Leftrightarrow |U_n| \to 0 \text{ as } \quad t \to \infty.$$

Defining the *basin of attraction* of the origin to be the set of initial data which yield an asymptote at the origin as t or n tends to infinity, we see that this notion of stability is nothing more than seeking conditions on Δt which preserve the basin of attraction. We have assumed that the basin of attraction of the origin for the differential equation is the whole of $\mathbb{R}^p$. Thus it might be natural to seek numerical methods which replicate this property for an interval of Δt; that is to seek methods with the property that there exists $\Delta t_c > 0$ such that

$$\{(\Delta t, U) \in \mathbb{R} \times \mathbb{R}^p : \Delta t \in (0, \Delta t_c)\} \subseteq \mathcal{S}.$$

Under such conditions, all choices of $\Delta t \in (0, \Delta t_c)$ will yield a numerical solution with the correct asymptotic behaviour, *independently of initial data.* An even stronger constraint on the numerical method would be to ask that no upper bound on Δt is required either; that is to seek methods with the property that

$$\{(\Delta t, U) \in \mathbb{R} \times \mathbb{R}^p : \Delta t \in (0, \infty)\} \subseteq \mathcal{S}.$$

Such notions of nonlinear stability are contained in the literature although they are not framed in this way. In particular, there has been a great deal of work devoted to problems of the form (1.1) where $f(\cdot)$ satisfies

$$\langle f(u) - f(v), u - v \rangle \leq -\alpha \|u - v\|^2 \tag{2.12}$$

for some $\alpha > 0$. Under (2.12), equation (1.1) has a unique steady solution (without loss of generality at the origin) which all solutions approach exponentially as $t \to \infty$. The study of numerical stability for such problems was initiated in Dahlquist (1975; 1978) where linear multistep methods were considered and generalized to Runge–Kutta methods by Burrage and Butcher (Butcher, 1975; Burrage and Butcher, 1979).

Whilst the nonlinear stability theory developed for (1.1) satisfying (2.12) has been very important in unifying the linear and nonlinear theories of error propagation, its range of applicability is somewhat limited since the condition (2.12) rules out nontrivial dynamical behaviour. Nonetheless, analogous stability theories can be developed under other hypotheses on $f(\cdot)$. For instance, the assumption that

$$\exists a, b > 0 : \langle f(u), u \rangle \leq a - b\|u\|^2 \tag{2.13}$$

is of interest in this context. Under (2.13) all solutions of (1.1) eventually enter the ball

$$\{u \in \mathbb{R}^p : \|u\|^2 \leq (a + \epsilon)/b\} \text{ for any } \epsilon > 0.$$

Hence this asymptotic bound on the solution is *independent of the initial data.* It is natural to examine numerical methods which replicate this property and this is done in Humphries and Stuart (1994) for Runge–Kutta methods and in Hill (1994) for linear multistep methods. It is interesting that the stability conditions required to make numerical methods replicate the qualitative behaviour of the underlying equation (1.1) under (2.11), (2.12) or (2.13) are very closely related. An overview of the area of nonlinear stability for (1.1) under a variety of different structural assumptions, including those described here, may be found in Stuart and Humphries (1992a). Since that article is self-contained we will not pursue stability issues any further in this article.

2.3. Summary

There are two important obervations to make concerning the discussion in this section. First *one possible* generalization of the notion of convergence to include dynamical systems is to consider the convergence of ω-limit sets – the objects observed for large time in (1.1). Second *one possible* generalization of the notion of stability to dynamical systems is to ask for preservation of certain qualitative properties of the underlying differential equation (1.1) under numerical approximation; in particular it is desirable to have such preservation occurring for a wide range of time step Δt and initial data U.

These two separate questions of convergence and stability are bridged by the question of the convergence of basins of attraction of ω-limit sets. This is because numerical instability is often manifested in a blow-up of the scheme so that the basins of attraction of limit sets are affected. Consider the following example.

Example This example shows relationships between the convergence of limit sets and basins of attraction, and numerical stability. All solutions of the equation

$$u_t = -u^3, \quad u(0) = U \tag{2.14}$$

tend to the origin 0 as $t \to \infty$. Thus $\{0\}$ is the only ω-limit set and its basin of attraction is $\mathbb{R}$. Now consider the Euler approximation

$$U_{n+1} = (1 - \Delta t U_n^2)U_n, \quad U_0 = U.$$

Analysis of this map given in Stuart (1991) shows that the origin $\{0\}$ is an ω-limit point (which trivially converges to the true limit set $\{0\}$ as $\Delta t \to 0$.) The origin has basin of attraction $(-\sqrt{(2/\Delta t)}, \sqrt{(2/\Delta t)})$ and thus the basin of attraction converges to the true basin of attraction $\mathbb{R}$ as $\Delta t \to 0$. To see the connection with numerical stability, consider initial data outside the basin of attraction: if $|U|^2 > \sqrt{(2/\Delta t)}$ then it may be shown that $|U_n| \to \infty$ as $n \to \infty$, a form of numerical instability. Thus the convergence of basins of attraction is closely related to the determination of conditional numerical stability questions.

The question of convergence of basins of attraction is little studied and there are many open questions in the area; see Humphries (1994) for some analysis in this direction.

As stated before, in this article we concentrate solely on the convergence of ω-limit sets and other related objects. This in itself is an enormous subject area but, as we hope to show, comprises a cohesive body of knowledge. In particular we have striven to put the results already contained in the literature in a unified framework, paying particular attention to the development of an existence theory for the objects of interest which can also be used to study analogous questions for numerical approximations.

2.4. Bibliography

Interest in the subject of the interaction between numerical analysis and dynamical systems has been growing steadily over the past decade. In particular two major international conferences have been held concerning the subject – the first was in Bristol, UK, in 1990 (see Broomhead and Iserles, (1992), Budd (1990) and Sanz-Serna (1992a)) and the second in Geelong, Australia, in 1993 (see Kloeden and Palmer (1994)). Furthermore, a series of lectures at the IVth SERC Numerical Analysis Summer School in Lancaster, UK, was given in 1991 – see Beyn (1992).

3. Semigroups and their approximation

3.1. Notation

We shall not define a specific norm in this article except in a few special circumstances. However, the norm should be taken as fixed throughout any given argument used and, furthermore, all matrix norms are those subordinate to the underlying vector norm. For simplicity it is sufficient to consider the Euclidean norm unless otherwise stated.

It will be important to have an appropriate definition of the distance between sets. Let A and B be sets in $\mathbb{R}^p$ and u a point in $\mathbb{R}^p$; we introduce the following notation:

$$\begin{aligned} \operatorname{dist}(u, A) &= \inf_{v \in A} \|u - v\|, \\ \operatorname{dist}(B, A) &= \sup_{u \in B} \operatorname{dist}(u, A), \\ \mathcal{N}(A, \epsilon) &= \{u \in \mathbb{R}^p : \operatorname{dist}(u, A) < \epsilon\}, \\ \partial\mathcal{N}(A, \epsilon) &= \{u \in \mathbb{R}^p : \operatorname{dist}(u, A) = \epsilon\}. \end{aligned} \tag{3.1}$$

Notice that, if $\operatorname{dist}(B, A) < \epsilon$ it follows that $\bar{B} \subseteq \mathcal{N}(\bar{A}, \epsilon)$ so that

$$\operatorname{dist}(B, A) = 0 \Longrightarrow \bar{B} \subseteq \bar{A}.$$

Hence 'dist' only defines a *semidistance* – the asymmetric Hausdorff semidistance as distinct from the *Hausdorff distance* between two sets A and B which is defined by

$$d_{\mathrm{H}}(A, B) = \max\{\operatorname{dist}(A, B), \operatorname{dist}(B, A)\}. \tag{3.2}$$

We also employ the following notation for open balls:

$$\begin{aligned} B(v, \epsilon) &:= \{u \in \mathbb{R}^p : \|u - v\| < \epsilon\}, \\ \partial B(v, \epsilon) &:= \{u \in \mathbb{R}^p : \|u - v\| = \epsilon\}. \end{aligned} \tag{3.3}$$

Thus $B(v, \epsilon) = \mathcal{N}(v, \epsilon)$ and $\partial B(v, \epsilon) = \partial\mathcal{N}(v, \epsilon)$.

3.2. The differential equation

Let us assume for the moment that a unique solution of (1.1) exists for all $t \geq 0$ and any $U \in \mathbb{R}^p$. Consequently we may define a semigroup $S(t) : \mathbb{R}^p \to \mathbb{R}^p$ in such a way that the solution $u(t)$ of (1.1) is given by

$$u(t) = S(t)U.$$

The one-parameter mapping $S(t)$ satisfies the usual semigroup properties

(i) $S(0) = I$, the identity on $\mathbb{R}^p$;
(ii) $S(t+s) = S(t)S(s) \quad \forall t, s \in \mathbb{R}^+$.

Under the assumption that f is differentiable on $\mathbb{R}^p$ it follows that the semigroup $S(t)U$ is continuous in both $t \in \mathbb{R}^+$ and $U \in \mathbb{R}^p$. We make this assumption throughout. We denote the Jacobian of $S(t)U$ with respect to $U \in \mathbb{R}^p$, evaluated at a point $V \in \mathbb{R}^p$, by $\mathrm{d}S(V;t)$.

Example To illustrate the semigroup $S(t)$ we consider the equation (2.14). Since solutions exist for all positive time, a semigroup may be defined. Solving the equation explicitly gives

$$S(t)U = \frac{U}{(1+2tU^2)^{1/2}}. \tag{3.4}$$

Properties (i) and (ii) are easily verified. Differentiation with respect to U shows that

$$dS(V;t) = \frac{1}{(1+2tV^2)^{3/2}}. \quad \square \tag{3.5}$$

Frequently we shall we require the action of $S(t)$ on a set of points $E \subset \mathbb{R}^p$. We define

$$S(t)E = \bigcup_{x \in E} S(t)x. \tag{3.6}$$

When analysing dynamical systems it is of great importance to study sets with the property that trajectories starting in a given set remain within that set. This motivates the following definition:

Definition 3.1 A set E is said to be *invariant* (respectively *positively invariant*) if, for any $t \geq 0$, $S(t)E \equiv E$ (respectively $S(t)E \subseteq E$).

Example Consider the equation

$$u_t = u(1-u^2), \quad u(0) = U. \tag{3.7}$$

This has solution

$$u(t) = \frac{U}{[U^2 + (1-U^2)\mathrm{e}^{-2t}]^{1/2}}. \tag{3.8}$$

Let $E = [-1,1]$. We show that $S(t)E = E$. If $U \in [-1,1]$ then it follows from (3.8) that $S(t)U \in [-1,1]$ and hence that $S(t)E \subseteq E$. Furthermore reversing time in (3.8) shows that, if $V = S(t)U$, then

$$U = \frac{V}{[V^2 + (1-V^2)\mathrm{e}^{2t}]^{1/2}}.$$

Thus, if $V \in [-1,1]$ then $U \in [-1,1]$ and hence $E \subseteq S(t)E$. Thus we have shown that $S(t)E \equiv E$ and so E is invariant.

Now consider the interval $B = [-a,a], a > 1$. From (3.7) we have

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|u|^2 = u^2 - u^4 < 0, \quad |u| > 1.$$

Thus solutions starting on the boundary of B enter B and hence no solutions can leave B. Thus $S(t)B \subseteq B$ and B is positively invariant. □

The behaviour as $t \to \infty$ of the dynamical system defined by $S(t)$ is captured by its ω-limit sets. Roughly these are sets of accumulation points at $t = \infty$ for subsequences in time extracted from a solution $u(t)$, $t \geq 0$. Precisely we have:

Definition 3.2 *The* ω-limit set *of a point* U *is defined by*

$$\omega(U) = \{x \in \mathbb{R}^p | \exists \{t_i\}, t_i \to \infty : S(t_i)U \to x \text{ as } t_i \to \infty\}.$$

An equivalent definition is

$$\omega(U) = \bigcap_{s\geq 0} \overline{\bigcup_{t\geq s} S(t)U}. \tag{3.9}$$

Similary we may define the ω-limit set *of a set* E *by*

$$\omega(E) = \{x \in \mathbb{R}^p | \exists \{t_i\}, \{U_i\}, t_i \to \infty, U_i \in E : S(t_i)U_i \to x \text{ as } t_i \to \infty\}.$$

An equivalent definition is

$$\omega(E) = \bigcap_{s\geq 0} \overline{\bigcup_{t\geq s} S(t)E}. \tag{3.10}$$

Examples Typical examples of ω-limit sets of individual points are equilibrium points, periodic solutions, quasi-periodic solutions and strange attractors. We illustrate these objects by example.

(i) *Equilibrium point.* Consider the equation

$$u_t = -u, \quad u(0) = U. \tag{3.11}$$

Since $u(t) = \mathrm{e}^{-t}U$ it follows that $\omega(U) = \{0\}$ for all $U \in \mathbb{R}^p$. The point $\{0\}$ is simply the equilibrium point for the equation.

(ii) *Periodic solutions.* Consider the equations

$$\begin{aligned} x_t &= x + y - x(x^2+y^2), \quad x(0) = x_0, \\ y_t &= -x + y - y(x^2+y^2), \quad y(0) = y_0. \end{aligned} \tag{3.12}$$

If we change variables to polar coordinates by introducing R and ϕ given by

$$x = R\cos\phi, \quad y = R\sin\phi$$

then we obtain the equations

$$R_t = R(1 - R^2), \quad \phi_t = -1. \tag{3.13}$$

Thus $\phi(t) = \phi(0) - t$. Explicit use of the solution (3.8) shows that, for all $R(0) > 0$, $R(t) \to 1$ as $t \to \infty$. Thus

$$x(t) \to \cos(\phi(0) - t), \;\; y(t) \to \sin(\phi(0) - t).$$

The solution rotates clockwise in the plane and asymptotically approaches a solution with radius 1; we have found a periodic solution of equations (3.12). Thus, for every solution with $R(0) > 0$ the ω-limit set is simply

$$\mathcal{P} = \{(x, y)^T \in \mathbb{R}^2 : x^2 + y^2 = 1\},$$

that is the set of all points on the periodic solution.

(iii) *Quasi-periodic solutions.* Consider equation (1.1) in $\mathbb{R}^p$, $p \geq 2$, with solution $u(t) = (u_1(t), u_2(t), \ldots, u_p(t))^T$. Assume that the first two solution components satisfy

$$u_1(t) = \mathrm{e}^{-t} + \sin(t), \quad u_2(t) = \mathrm{e}^{-t} + \cos(\delta t)$$

and that the remaining solution components approach 0 as $t \to \infty$. Then, if δ is irrational, as $t \to \infty$ the limiting solution is quasi-periodic. The ω-limit set is given by

$$\mathcal{Q} = \{(u \in \mathbb{R}^p : -1 \leq u_1 \leq 1, -1 \leq u_2 \leq 1, u_j = 0, \; j = 3, \ldots, p\}.$$

(iv) *Strange attractors.* Consider the *Lorenz equations* given by (2.10). Figure 2 shows solutions of the equations, with parameters set at $\sigma = 10, r = 28$ and $b = \frac{8}{3}$, for four entirely different initial conditions. As described in Section 2, the solutions eventually lie on a very complicated set in $\mathbb{R}^3$ and this set is an example of a *strange attractor.* It is the ω-limit set for equations (2.10) for almost all initial data. □

It is important to note that, in general,

$$\omega(E) \neq \bigcup_{x \in E} \omega(x).$$

The following example illustrates this.

Example Consider the equation (3.7) with solution (3.8). From the explicit solution it is clear that

$$\omega(U) = 1, \; U > 0; \quad \omega(U) = -1, \; U < 0; \quad \omega(0) = 0.$$

Thus for any interval $E = [-a, a]$, $a > 0$ we have

$$\bigcup_{x \in E} \omega(x) = \{-1, 0, 1\}.$$

It may also be shown that

$$\omega(E) = [-1, 1];$$

this follows from showing that

$$S(t)E = [-S(t)a, S(t)a],$$

with $S(t)$ determined by (3.8), and noting that $S(t)a \to 1$ as $t \to \infty$.

The following property of ω limit sets is very useful:

Theorem 3.3 The ω-limit set of any bounded set $E \subset \mathbb{R}^p$, $\omega(E)$, is a closed positively invariant set. Furthermore, if $\exists T > 0 : S(t)E$ is bounded for $t \geq T$, then $\omega(E)$ is invariant. Finally, if $\omega(U)$ is bounded for some $U \in \mathbb{R}^p$ then it is connected.

Proof. We first establish closure: consider a sequence of ω-limit points $x_k \to x$, with each $x_k \in \omega(E)$. We wish to show that $x \in \omega(E)$. By Definition 3.2, for each k there are sequences $\{v_i^k\}, \{t_i^k\}$ such that

$$S(t_i^k)v_i^k \to x_k \text{ as } i \to \infty.$$

Hence, without loss of generality, let

$$\|x_k - S(t_i^k)v_i^k\| \leq k^{-1} \quad \text{and} \quad t_i^k \geq k \text{ for } i \geq k.$$

Now define $v_i^* : = v_i^i$ and $t_i^* := t_i^i$. By construction $t_i^* \to \infty$ as $i \to \infty$. Now

$$\begin{aligned} \|x - S(t_i^*)v_i^*\| &\leq \|x - x_i\| + \|S(t_i^*)v_i^* - x_i\| \\ &\leq \|x - x_i\| + i^{-1}. \end{aligned}$$

Taking $i \to \infty$ we find $S(t_i^*)v_i^* \to x$. Hence, by Definition 3.2, $x \in \omega(E)$. We now show positive invariance. Assume that $x \in \omega(E)$. If

$$S(t_i)v_i \to x$$

then by continuity of $S(t)\cdot$,

$$S(t + t_i)v_i = S(t)S(t_i)v_i \to S(t)x \quad \forall t \geq 0.$$

Thus $S(t + t_i)v_i \to S(t)x$ and, hence $S(t)x \in \omega(E)$ by Definition 3.2. Thus we deduce positive invariance of $\omega(E)$.

Now we establish negative invariance. We assume that $S(t)E$ is bounded for $t \geq T$ and assume that $x \in \omega(E)$. We aim to show that, for any $t > 0$, $\exists y$: $S(t)y = x$ and $y \in \omega(E)$. Let $S(t_i)v_i \to x$, where, without loss of generality, we may choose $t_1 \geq 1 + T + t$. Now consider the sequence $S(t_i - t)v_i$. Since

$S(t)E$ is bounded for $t \geq T$ and $t_i - t \geq T+1$ for all $i \geq 1$ it follows that there exists a convergent subsequence

$$S(t_{i_j} - t)v_{i_j} \to y.$$

Now

$$\begin{aligned} x &= \lim_{j\to\infty} S(t_{i_j})v_{i_j} = \lim_{j\to\infty} S(t)S(t_{i_j} - t)v_{ij} \\ &= S(t)\lim_{j\to\infty} S(t_{i_j} - t)v_{i_j} = S(t)y. \end{aligned}$$

Hence the negative invariance is proved.

Finally we show that boundedness of $\omega(U)$ implies connectedness. Assume for contradiction that $\omega(U)$ comprises two disjoint components P and Q with $\mathcal{N}(P,\epsilon) \cap \mathcal{N}(Q,\epsilon) = \emptyset$, for some $\epsilon > 0$. Then there exist sequences $t_i \to \infty$ and $\tau_i \to \infty$ such that $S(t_i)U \to x \in P$ and $S(\tau_i)U \to y \in Q$. Without loss of generality we may assume that $t_i < \tau_i$ and that $S(t_i)U \in \mathcal{N}(P,\epsilon), S(\tau_i) \in \mathcal{N}(Q,\epsilon)$ $\forall i \geq 1$. By continuity of $S(\cdot)U$ it follows that there exists $T_i \in (t_i, \tau_i)$ such that $S(T_i)U \in \partial\mathcal{N}(P,\epsilon)$. But the set $\partial\mathcal{N}(P,\epsilon)$ is closed and bounded, since P is bounded, and hence compact. Thus there exists a convergent subsequence $S(T_{i_j})U \to z \in \partial\mathcal{N}(P,\epsilon)$. But this is a contradiction since then, by definition, $z \in \omega(U)$ but $z \notin P \cup Q$. This completes the proof. □

The behaviour of a dynamical system is very well understood if all the ω-limit sets can be determined together with knowledge of which initial data are associated with a given limit set; this motivates the following:

Definition 3.4 *An ω-limit set $\mathcal{W}$ has* basin of attraction $\mathcal{B}$ *if*

$$\{\omega(U) = \mathcal{W}\} \Leftrightarrow \{U \in \mathcal{B}\}.$$

Examples For equation (3.11) it is is clear that the basin of attraction of the ω-limit set $\mathcal{Q} = \{0\}$ is the whole of $\mathbb{R}$. For equation (3.12) the basin of attraction of the periodic solution $\mathcal{P}$ is $\mathbb{R}^2\backslash\{0\}$. □

3.3. Approximating semigroups

The numerical method (1.2) generates a semigroup $S^n_{\Delta t} : \mathbb{R}^p \to \mathbb{R}^p$ in such a way that the solution U_n of (1.2) is given by

$$U_n = S^n_{\Delta t}U_0.$$

Here the subscript Δt is used simply to emphasize the dependence of the numerical method on Δt. This semigroup satisfies properties analogous to those for $S(t)$:

(i) $S^0_{\Delta t} = I$, the identity on $\mathbb{R}^p$;

(ii) $S^{n+m}_{\Delta t} = S^n_{\Delta t}S^m_{\Delta t} \quad \forall n,m \in \mathbb{Z}^+$.

If f is differentiable then standard numerical methods yield a differentiable semigroup $S^1_{\Delta t}U$. Throughout we denote the Jacobian of $S^1_{\Delta t}U$ with respect to $U \in \mathbb{R}^p$ evaluated at a point $V \in \mathbb{R}^p$ by $\mathrm{d}S^1_{\Delta t}(V)$.

Example We consider equation (2.14) under approximation by the explicit Euler scheme; this yields the mapping

$$U_{n+1} = U_n - \Delta t U_n^3, \tag{3.14}$$

so that

$$S^1_{\Delta t}U = U - \Delta t U^3 \tag{3.15}$$

and $S^n_{\Delta t}$ is an n-fold composition of $S^1_{\Delta t}$. Again Properties (i) and (ii) are straightforward to check. Differentiation with respect to U gives

$$\mathrm{d}S^1_{\Delta t}(V) = 1 - 3\Delta t V^2. \quad \square \tag{3.16}$$

Important remark It is straightforward to define concepts of invariance, ω-limit set and basin of attraction for the semigroup $S^n_{\Delta t}$. Indeed the only change necessary to Definitions 3.1, 3.2 and 3.4 is to replace t_i by a sequence of integers $n_i \to \infty$. Furthermore, Theorem 3.3 has a discrete analogue for $S^n_{\Delta t}$ with the caveat that the last part of the theorem, concerning connectedness, does not hold.

Example Consider the explicit Euler approximation of (3.11). This yields the map

$$U_{n+1} = (1 - \Delta t)U_n, \quad U_0 = U,$$

so that $S^1_{\Delta t}U = (1 - \Delta t)U$. If $\Delta t \in (0, 2)$ then $\omega(U) = 0$ for any $U \in \mathbb{R}$, replicating the behaviour of the differential equation. Thus $\{0\}$ has basin of attraction $\mathbb{R}$ if $\Delta t \in (0, 2)$. However, if $\Delta t = 2$ then $\omega(U) = \{U, -U\}$ and the basin of attraction of $\omega(U)$ is $\omega(U)$ itself. Note that $\omega(U)$ is not connected in this case. $\square$

The truncation error is now defined to be the error committed by the approximation (1.2) over one time step of length Δt.

Definition 3.5 The *truncation error* for the map (1.2) as an approximation to the ordinary differential equation (1.1) at a point $U \in \mathbb{R}^p$ is defined by

$$T(U; \Delta t) := S(\Delta t)U - S^1_{\Delta t}U.$$

The Jacobian of $T(U; \Delta t)$ with respect to $U \in \mathbb{R}^p$ evaluated at $V \in \mathbb{R}^p$ is denoted by

$$\mathrm{d}T(V; \Delta t) := \mathrm{d}S(V; \Delta t) - \mathrm{d}S^1_{\Delta t}(V).$$

Example We illustrate this definition by considering the approximation of (2.14) by (3.14). First notice that from (3.15)

$$|S^1_{\Delta t}U - S^1_{\Delta t}V| \leq |U - V| + \Delta t|U^3 - V^3| \leq [1 + \Delta t K_1(U, V)]|U - V|,$$

where $K_1(U,V) = |U^2+UV+V^2|$. This shows the local Lipschitz continuity of $S^1_{\Delta t}$. By (3.4) and (3.15) we see that the truncation error satisfies

$$T(U;\Delta t) = \frac{U}{(1+2\Delta tU^2)^{1/2}} - (U - \Delta tU^3).$$

Taylor expansion of $S(\Delta t)U$ shows that there exists $\Delta t_c = \Delta t_c(U) > 0$ and $K_2 = K_2(U) > 0$ such that

$$|T(U;\Delta t)| \leq K_2\Delta t^2 \quad \forall \Delta t \in (0, \Delta t_c].$$

This shows that the truncation error is bounded above by a quantity proportional to Δt^2. Furthermore

$$\mathrm{d}T(U;\Delta t) = \frac{1}{(1+2\Delta tU^2)^{3/2}} - (1 - 3\Delta tU^2).$$

Again, a Taylor series expansion shows that there exists $\Delta t_c = \Delta t_c(U) > 0$ and $K_3 = K_3(U) > 0$ such that

$$|\mathrm{d}T(U;\Delta t)| \leq K_3\Delta t^2 \quad \forall \Delta t \in (0, \Delta t_c],$$

possibly by further reduction of Δt_c. This shows that the Jacobian of the truncation error is also bounded above by a constant of $\mathcal{O}(\Delta t^2)$.

There are two remaining properties of $S^1_{\Delta t}$ worth emphasizing by example. Note from (3.16) that

$$|\mathrm{d}S^1_{\Delta t}(V)| \leq 1 + 3V^2\Delta t$$

and

$$|\mathrm{d}S^1_{\Delta t}(U) - \mathrm{d}S^1_{\Delta t}(V)| \leq 3\Delta t|U+V||U-V|. \quad \square$$

The example illustrates a number of basic properties of the approximate semigroup that we will need throughout this article. It is clear that four important properties of the approximate semigroup are: (i) a Lipschitz condition on $S^1_{\Delta t}$ (or, relatedly, a bound on the Jacobian $\mathrm{d}S^1_{\Delta t}$) of size $1+\mathcal{O}(\Delta t)$; (ii) a Lipschitz condition on $\mathrm{d}S^1_{\Delta t}$ of size $\mathcal{O}(\Delta t)$; (iii) an $\mathcal{O}(\Delta t^2)$ closeness of $S^1_{\Delta t}$ to $S(\Delta t)$ uniformly in any bounded set $B(0,R)$; (iv) an $\mathcal{O}(\Delta t^2)$ closeness of $\mathrm{d}S^1_{\Delta t}(U)$ to $\mathrm{d}S(U;\Delta t)$ uniformly in any bounded set $B(0,R)$.

We will assume that $S^1_{\Delta t}$ satisfies generalizations of these four conditions throughout the remainder of the article. However, our analysis will be greatly streamlined if estimates for the size of the truncation error in terms of Δt are uniform across the whole of $\mathbb{R}^p$. Since our interest is primarily in the local behaviour of $S(t)$ and $S^n_{\Delta t}$ near to bounded limit sets it is sufficient to consider vector fields which are globally bounded. Specifically we make the following assumption concerning the vector field f :

Assumption 3.6 The vector field f in (1.1) satisfies $f \in C^\infty(\mathbb{R}^p, \mathbb{R}^p)$ and f and all of its derivatives are uniformly bounded for all $u \in \mathbb{R}^p$.

In fact this assumption also yields both forwards and backwards in time global existence and uniqueness for the equation (1.1) so that the semigroup may be extended to a group, but we do not pursue this further. Assumption 3.6 will be made throughout the remainder of this section and throughout Sections 4–6; it will not be stated explicitly in the theorems. The assumption is made for simplicity and, since the results described are local in nature, is not necessary – the vector field $f(\cdot)$ can always be modified outside a compact set to yield Assumption 3.6.

From Assumption 3.6 it is possible to prove that many one-step approximations $S^1_{\Delta t}$ to the semigroup $S(t)$ (including all Runge–Kutta methods) generated by (1.1) satisfy the following uniform continuity and approximation properties:

Assumption 3.7 *There exists constants $K > 0, \Delta t_c > 0$ and an integer $r \geq 1$ such that for all $\Delta t \in [0, \Delta t_c)$ the semigroups $S(t)$ and $S^1_{\Delta t}$ satisfy*

(i) $\|\mathrm{d}S^1_{\Delta t}(U)\| \leq (1 + K\Delta t) \quad \forall U \in \mathbb{R}^p$;
(ii) $\|\mathrm{d}S^1_{\Delta t}(U) - \mathrm{d}S^1_{\Delta t}(V)\| \leq K\Delta t\|U - V\| \quad \forall U, V \in \mathbb{R}^p$;
(iii) $\|T(U; \Delta t)\| \leq K\Delta t^{r+1} \quad \forall U \in \mathbb{R}^p$;
(iv) $\|\mathrm{d}T(U; \Delta t)\| \leq K\Delta t^{r+1} \quad \forall U \in \mathbb{R}^p$.

We will not make this assumption explicit in the statement of the theorems but will assume it throughout the remainder of this section and throughout Sections 4–6. Assumption 3.7 is satisfied by standard Runge–Kutta methods applied to (1.1) under Assumption 3.6.

We now prove certain results concerning the closeness of the semigroups $S^n_{\Delta t}$ and $S(t_n)$ over fixed time intervals $0 \leq n\Delta t = t_n \leq T$. These results follow directly from Assumption 3.7. We need one preliminary observation. Note that, if $u(t) = S(t)U$ then $w(t) = \mathrm{d}S(U; t)v$ satisfies the equation

$$w_t = \mathrm{d}f(u(t))w, \quad w(0) = v$$

where $\mathrm{d}f(\cdot)$ denotes the Jacobian of f. By Assumption 3.6 we may assume, without loss of generality, that $\|\mathrm{d}f(u)\| \leq K \; \forall u \in \mathbb{R}^p$ so that

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|w\|^2 = \langle w, w_t\rangle = \langle w, \, \mathrm{d}f(u)w\rangle \leq K\|w\|^2.$$

Hence

$$\|w(t)\| \leq \mathrm{e}^{Kt}\|v\| \Rightarrow \|\mathrm{d}S(U, t)\| \leq \mathrm{e}^{Kt}. \tag{3.17}$$

We can now prove the following theorem concerning the closeness of both the semigroups $S(t)$ and $S^1_{\Delta t}$ and their derivatives.

Theorem 3.8 *Consider the semigroups $S^1_{\Delta t}U$ and $S(\Delta t)U$. It follows that, if $t_n = n\Delta t, 0 \leq t_n \leq T$ and $\Delta t \in (0, \Delta t_c]$, then*

$$e_n := \|S^n_{\Delta t}U - S(t_n)U\| \leq [\mathrm{e}^{KT} - 1]\Delta t^r$$

and

$$E_n := \|\mathrm{d}S^n_{\Delta t}(U) - \mathrm{d}S(U;t_n)\| \le [\mathrm{e}^{3KT} - \mathrm{e}^{2KT}]\Delta t^r.$$

Furthermore

$$\|S^n_{\Delta t}U - S(t_n)V\| \le (\mathrm{e}^{KT} - 1)\Delta t^r + \mathrm{e}^{KT}\|U - V\|.$$

Proof. Clearly Assumption 3.7(i) yields

$$\|S^1_{\Delta t}V - S^1_{\Delta t}W\| \le (1 + K\Delta t)\|V - W\| \quad \forall V, W \in \mathbb{R}^p. \tag{3.18}$$

Thus, by (i) and (iii) of Assumption 3.7, we have

$$\begin{aligned} e_{m+1} &= \|S^{m+1}_{\Delta t}U - S(t_{m+1})U\| = \|S^1_{\Delta t}S^m_{\Delta t}U - S(\Delta t)S(t_m)U\| \\ &\le \|S^1_{\Delta t}S^m_{\Delta t}U - S^1_{\Delta t}S(t_m)U\| + \|S(\Delta t)S(t_m)U - S^1_{\Delta t}S(t_m)U\| \\ &\le (1 + K\Delta t)e_m + K\Delta t^{r+1}. \end{aligned}$$

Thus, by induction, we obtain

$$e_n \le [(1 + K\Delta t)^n - 1]\Delta t^r + (1 + K\Delta t)^n e_0. \tag{3.19}$$

Since $e_0 = 0$ and $(1 + K\Delta t)^n \le \mathrm{e}^{KT}$ the desired result follows.

Now we consider E_n. Note that from (3.17) we have

$$\mathrm{d}S(U;t_m) \le \mathrm{e}^{KT}, \quad \forall t_m \in [0,T]. \tag{3.20}$$

The quantity E_m satisfies

$$\begin{aligned} E_{m+1} &= \|\mathrm{d}S^{m+1}_{\Delta t}(U) - \mathrm{d}S(U;t_{m+1})\| \\ &= \|\mathrm{d}\{S^1_{\Delta t}S^m_{\Delta t}(U)\} - \mathrm{d}\{S(\Delta t)S(t_m)U\}\| \\ &= \|\mathrm{d}S^1_{\Delta t}(S^m_{\Delta t}U)\,\mathrm{d}S^m_{\Delta t}(U) - \mathrm{d}S(S(t_m)U;\Delta t)\,\mathrm{d}S(U;t_m)\| \\ &\le \|\mathrm{d}S^1_{\Delta t}(S^m_{\Delta t}U)\,\mathrm{d}S^m_{\Delta t}(U) - \mathrm{d}S^1_{\Delta t}(S^m_{\Delta t}U)\,\mathrm{d}S(U;t_m)\| \\ &\quad + \|\mathrm{d}S^1_{\Delta t}(S^m_{\Delta t}U)\,\mathrm{d}S(U;t_m) - \mathrm{d}S(S(t_m)U;\Delta t)\,\mathrm{d}S(U;t_m)\|. \end{aligned}$$

Hence, by Assumption 3.7(i),

$$\begin{aligned} E_{m+1} &\le (1 + K\Delta t)E_m \\ &\quad + \|\mathrm{d}S^1_{\Delta t}(S^m_{\Delta t}U)\,\mathrm{d}S(U;t_m) - \mathrm{d}S^1_{\Delta t}(S(t_m)U)\,\mathrm{d}S(U;t_m)\| \\ &\quad + \|\mathrm{d}S^1_{\Delta t}(S(t_m)U)\,\mathrm{d}S(U;t_m) - \mathrm{d}S(S(t_m)U;\Delta t)\,\mathrm{d}S(U;t_m)\|. \end{aligned}$$

Assumption 3.7(ii),(iv) and (3.19), (3.20) thus give us

$$E_{m+1} \le (1 + K\Delta t)E_m + K\mathrm{e}^{KT}\Delta t e_m + K\mathrm{e}^{KT}\Delta t^{r+1}.$$

Using the known bound on e_m we obtain

$$E_{m+1} \le (1 + K\Delta t)E_m + K\mathrm{e}^{KT}\Delta t(\mathrm{e}^{KT} - 1)\Delta t^r + K\mathrm{e}^{KT}\Delta t^{r+1}.$$

Hence

$$E_{m+1} \le (1 + K\Delta t)E_m + K\mathrm{e}^{2KT}\Delta t^{r+1}.$$

By induction we obtain

$$E_n \leq [(1+K\Delta t)^n - 1]\mathrm{e}^{2KT}\Delta t^r + (1+K\Delta t)^n E_0.$$

Since $E_0 = 0$ and $(1+K\Delta t)^n \leq \mathrm{e}^{KT}$ the desired result follows.

To prove the final result, note that

$$\|S^n_{\Delta t}U - S(t_n)V\| \leq \|S^n_{\Delta t}V - S(t_n)V\| + \|S^n_{\Delta t}U - S^n_{\Delta t}V\|.$$

Using the error bound for e_n and (3.18) we obtain

$$\|S^n_{\Delta t}U - S(t_n)V\| \leq [\mathrm{e}^{KT} - 1]\Delta t^r + (1+K\Delta t)^n\|U - V\|$$

and, since $0 \leq n\Delta t \leq T$, the required result follows. □

Remarks

(i) Note that the derivation of the error e_n is obtained by using two facts: the Lipshitz continuity of the numerical method together with a uniform truncation error bound. An alternative method of proof is to exploit the Lipschitz continuity of the underlying differential equation. Assume that

$$\|S(t)V - S(t)W\| \leq \mathrm{e}^{Kt}\|V - W\|. \tag{3.21}$$

(The choice of a constant K in this bound can be made without loss of generality.) The error equation can now be studied thus:

$$\begin{aligned} e_{m+1} &= \|S^{m+1}_{\Delta t}U - S(t_{m+1})U\| = \|S^1_{\Delta t}S^m_{\Delta t}U - S(\Delta t)S(t_m)U\| \\ &\leq \|S(\Delta t)S^m_{\Delta t}U - S(\Delta t)S(t_m)U\| + \|S(\Delta t)S^m_{\Delta t}U - S^1_{\Delta t}S^m_{\Delta t}U\|. \end{aligned}$$

Using the Lipschitz continuity of $S(t)$ and Assumption 3.7(iii) we obtain

$$e_{m+1} \leq \mathrm{e}^{K\Delta t}e_m + K\Delta t^{r+1}. \tag{3.22}$$

By induction we obtain

$$e_n \leq [\mathrm{e}^{Kt_n} - 1]\frac{K\Delta t^{r+1}}{\mathrm{e}^{K\Delta t} - 1} + \mathrm{e}^{Kt_n}e_0.$$

Noting that

$$\mathrm{e}^x - 1 \geq x/2$$

and that $e_0 = 0$ we obtain

$$e_n \leq 2[\mathrm{e}^{KT} - 1]K\Delta t^r, \quad 0 \leq t_n \leq T$$

an analogous bound to that obtained in Theorem 3.8.

(ii) It is worth observing that the constant K appearing in the error bounds derived here is far from optimal. This is since the straightforward bound on the Lipschitz constant used to obtain (3.17) is often pessimistic in estimating the divergence between solutions of the differential equation. Smaller constants can be obtained by working with the so-called logarithmic norm which, roughly speaking, captures the possible rate of divergence more accurately. See Dekker and Verwer (1984).

As briefly mentioned in Section 2 it is possible to obtain uniformly valid error bounds in the case where $u(t)$ approaches an exponentially stable equilibrium point. Specifically we assume that (2.12) holds in a neighbourhood of an equilibrium point $\bar{u}$ from which it follows that

$$\exists \alpha, R > 0 : \quad \|S(t)v - S(t)w\| \leq \mathrm{e}^{-\alpha t}\|v - w\| \quad \forall v, w \in B(\bar{u}, R). \tag{3.23}$$

The following result is a precursor for the remainder of the article: the proof contains two essential components namely (i) the use of a standard truncation error estimate or resulting finite-time error bound together with (ii) exploitation of a structural property of the underlying solutions of (1.1). In this case the structural property is the exponential stability of $\bar{u}$ manifest in (3.23).

Theorem 3.9 Consider the semigroups $S^1_{\Delta t}U$ and $S(\Delta t)U$ under Assumption 3.7. Assume further that $S(t)U \to \bar{u}$ as $t \to \infty$ and that (3.23) holds. It follows that, if $t_n = n\Delta t, 0 \leq t_n \leq T$ and $\Delta t \in (0, \Delta t_c]$, then there exists $C > 0$ such that

$$e_n := \|S^n_{\Delta t}U - S(t_n)U\| \leq C\Delta t^r \quad \forall n, \Delta t : 0 \leq n\Delta t < \infty.$$

Proof. Assume that $T = N\Delta t$ is a time chosen so that $\|S(T)U - \bar{u}\| < R/2$. By assumption it follows that, for $\tau > 0$ and whilst $S(\tau + T)U \in B(\bar{u}, R)$,

$$\begin{aligned} \|S(\tau + T)U - S(\tau + T)\bar{u}\| &= \|S(T)S(\tau)U - S(T)S(\tau)\bar{u}\| \\ &= \|S(T)S(\tau)U - S(\tau)\bar{u}\| \\ &\leq \mathrm{e}^{-\alpha\tau}\|S(T)U - \bar{u}\| < R/2. \end{aligned}$$

Hence $S(t)U \in B(\bar{u}, R/2)$ for all $t \geq T$. From Theorem 3.8 we have that, for $\Delta t \in (0, \Delta t_c]$,

$$e_n \leq [\mathrm{e}^{KT} - 1]\Delta t^r, \quad 0 \leq t_n \leq T. \tag{3.24}$$

Thus we can ensure that (possibly by further reduction of Δt_c) that

$$\|S^n_{\Delta t}U - S(t_n)U\| \leq R/2 \quad 0 \leq t_n \leq T$$

and hence that $S^N_{\Delta t}U \in B(\bar{u}, R)$. Now assume, for the purposes of induction, that for some $m \geq N$,

$$e_m \leq \frac{[1 - \mathrm{e}^{-\alpha(t_m - T)}]}{1 - \mathrm{e}^{-\alpha\Delta t}}K\Delta t^{r+1} + \mathrm{e}^{-\alpha(t_m - T)}[\mathrm{e}^{KT} - 1]\Delta t^r. \tag{3.25}$$

We also assume that, again possibly by further reduction of Δt_c,

$$(1 - \mathrm{e}^{-\alpha\Delta t})^{-1} \leq \frac{2}{\Delta t} \quad \text{and} \quad (2K + \mathrm{e}^{KT} - 1)\Delta t^r \leq \tfrac{1}{2}R. \tag{3.26}$$

Note that (3.26) and (3.25) yield $e_m \leq R/2$ and hence $S^m_{\Delta t}U \in B(\bar{u}, R)$.

Clearly (3.25) holds for $m = N$ by (3.24). By an identical argument to that yielding (3.22) but with $K = -\alpha$ in (3.21) we have

$$e_{m+1} \le e^{-\alpha\Delta t} e_m + K\Delta t^{r+1}.$$

If (3.25) holds then

$$\begin{aligned} e_{m+1} &\le \frac{e^{-\alpha\Delta t}}{1-e^{-\alpha\Delta t}}[1-e^{-\alpha(t_M-T)}]K\Delta t^{r+1} + K\Delta t^{r+1} \\ &\quad + e^{-\alpha\Delta t}e^{-\alpha(t_M-T)}[e^{KT}-1]\Delta t^r \\ &= \frac{1-e^{-\alpha(t_{M+1}-T)}}{1-e^{-\alpha\Delta t}}K\Delta t^{r+1} + e^{-\alpha(t_{M+1}-T)}[e^{KT}-1]\Delta t^r. \end{aligned}$$

Thus (3.25) holds with $m \mapsto m+1$ and the induction is complete. Hence (3.25) and (3.26) give

$$e_n \le [2K + e^{KT} - 1]\Delta t^r \quad \forall n, \Delta t : T \le n\Delta t < \infty.$$

Combining this estimate and (3.24) we have the desired result by choosing $C = 2K + e^{KT} - 1$. □

3.4. Bibliography

The subject of dynamical systems is discussed in numerous text books and monographs. In the context of ordinary differential equations see, for example, Arrowsmith and Place (1990), Bhatia and Szego (1970), Devaney (1989), Drazin (1992), Guckenheimer and Holmes (1983) and Hale and Kocak (1991); in the context of partial differential equations see, for example, Babin and Vishik (1992), Hale *et al.* (1988), Hale (1984), Ladyzhenskaya (1991) and Temam (1988). The text by Bhatia and Szego (1970) is closest in outlook to the presentation of results in Section 3.1.

The pivotal point in Section 3.2 is Assumption 3.7. It is worth noting that in Section 4 we only require (i), (iii) and (iv) of Assumption 3.7, in Section 5 we require all four points and in Section 6 we use only (i) and (iii). Assumption 3.7 hold for all consistent Runge–Kutta methods: point (iii) is a standard truncation error bound proved in, for example, Butcher (1987) whilst point (iv) is proved in Stoffer (1994); points (i) and (ii) are readily established and, indeed, use of point (i) is implicit in all convergence proofs for Runge–Kutta methods. The situation for s-step multistep methods considered as dynamical systems is somewhat more complicated since the natural phase space for the problem is $\mathbb{R}^{ps}$. However, it has been shown in Kirchgraber (1986) that, under Assumption 3.6, for all strictly stable multistep methods there exists a consistent one-step method which is an attractive invariant manifold for the multistep method. In essence this means that there is a linear combination of $s-1$ successive steps of the method whose behaviour is governed by a one-step method – after a large number of iterations. This work has been generalized in Stoffer (1993) and a related question considered in Eirola and Nevanlinna (1988). These results enable the use of Assumption 3.7 in the study of multistep methods as well as one-step methods.

The first error bound in Theorem 3.8 is standard for Runge–Kutta methods and proofs may be found in (for example) Butcher (1987), Hairer *et al.* (1987), Lambert (1991) and Stetter (1973). The second error bound, concerning the C^1 closeness of the true and approximate semigroups, is not in the literature to the best of our knowledge in this general form; however, for the specific case of approximations of reaction-diffusion equations, such results are proved in Alouges and Debussche (1991) and Hale *et al.* (1988). The uniform in time error bound of Theorem 3.9 may be found in Stetter (1973). Generalizations of this idea may be found in Heywood and Rannacher (1986) for finite-element approximations of the Navier–Stokes equation, in Larsson (1989) for finite element approximations of nonsmooth solutions to reaction-diffusion equations and in Sanz-Serna and Stuart (1992) for finite difference approximations of smooth solutions to reaction-diffusion equations.

4. Neighbourhood of an equilibrium point

4.1. Background theory

In this section we study the affect of approximation on equilibrium points, their stable and unstable manifolds and their local phase portraits. In all cases we employ the contraction mapping theorem to develop an existence theory and exploit this to prove convergence results. This means that the basic existence theory for $S(t)$ takes the longest to develop whilst the existence and approximation theory for $S^n_{\Delta t}$ follows simply.

An *equilibrium point* for (1.1) is a point $\bar{u} \in \mathbb{R}^p$ satisfying

$$f(\bar{u}) = 0. \tag{4.1}$$

Consequently such a point $\bar{u}$ also satisfies the defining equation

$$\bar{u} = S(t)\bar{u} \quad \forall t \in \mathbb{R} \tag{4.2}$$

Thus $\bar{u}$ is a *fixed point* of the mapping $S(t)$ for every $t \in \mathbb{R}^p$ – see (4.5) below. The equilibrium point is said to be *hyperbolic* if none of the eigenvalues of the Jacobian of f at $\bar{u}$, $\mathrm{d}f(\bar{u})$, lies on the imaginary axis. A hyperbolic equilibrium point is said to be *stable* if all eigenvalues of $\mathrm{d}f(\bar{u})$ lie in the left-half plane. It is *unstable* if at least one eigenvalue of $\mathrm{d}f(\bar{u})$ lies in the right-half plane.

It may be shown that $\mathrm{d}S(\bar{u}; t) = \exp[\mathrm{d}f(\bar{u})t]$ and hence, that $\mathrm{d}S(\bar{u}; t)$ has no eigenvalues on the unit circle if $\bar{u}$ is hyperbolic. Thus we can define

$$D(t) = [I - \mathrm{d}S(\bar{u}; t)]^{-1}. \tag{4.3}$$

Using the fact that 0 is not in the spectrum of $\mathrm{d}f(\bar{u})$ and writing $\mathrm{d}S(\bar{u};t)$ as an infinite series in $\mathrm{d}f(\bar{u})$ it may be shown that $\exists\beta > 0, t_c > 0$ such that

$$\|D(t)\| \leq \beta/t \quad \forall t \in (0, t_c). \tag{4.4}$$

Note that stability may also be formulated in terms of the eigenvalues of $\mathrm{d}S(\bar{u};t)$; this can be done by modifying the definitions that we are about to make for fixed points of $S^1_{\Delta t}$.

A *fixed point* $\bar{U}$ of the semigroup $S^1_{\Delta t}$ satisfies the equation

$$\bar{U} = S^1_{\Delta t}\bar{U} \tag{4.5}$$

or, equivalently for (1.2),

$$\bar{U} = \mathcal{F}(\bar{U}, \Delta t).$$

The fixed point is *hyperbolic* if $\mathrm{d}S^1_{\Delta t}(\bar{U})$ has no eigenvalues on the unit circle; such a hyperbolic fixed point is said to be stable if all eigenvalues of $\mathrm{d}S^1_{\Delta t}(\bar{U})$ lie inside the unit circle and unstable if at least one eigenvalue lies outside the unit circle.

Throughout this article we will use the following notation for the set of fixed points of $S(t)$ and $S^1_{\Delta t}$ together with their neighbourhoods:

$$\begin{aligned} \mathcal{E} &= \{v \in \mathbb{R}^p : f(v) = 0\}, \\ \mathcal{E}(\epsilon) &= \{v \in \mathbb{R}^p : \|f(v)\| \leq \epsilon\}, \\ \mathcal{E}_{\Delta t} &= \{v \in \mathbb{R}^p : v = S^1_{\Delta t}v\}. \end{aligned} \tag{4.6}$$

We will need the following definitions; illustrative examples will be given later on.

Definition 4.1 The *unstable manifold* of an equilibrium point $\bar{u}$ of (1.1) is the set

$$W^u(\bar{u}) := \{U \in \mathbb{R}^p : u(t) \to \bar{u} \;\; as \;\; t \to -\infty\}.$$

The *local unstable manifold* of $\bar{u}$ is the set

$$W^{u,\epsilon}(\bar{u}) := \{U \in W^u(\bar{u}) : \|u(t) - \bar{u}\| \leq \epsilon \; \forall t \leq 0\}.$$

The *stable manifold* of an equilibrium point $\bar{u}$ of (1.1) is the set

$$W^s(\bar{u}) := \{U \in \mathbb{R}^p : u(t) \to \bar{u} \;\; as \;\; t \to \infty\}.$$

The *local stable manifold* of $\bar{u}$ is the set

$$W^{s\epsilon}(\bar{u}) := \{U \in W^s(\bar{u}) : \|u(t) - \bar{u}\| \leq \epsilon \; \forall t \geq 0\}$$

The following facts concerning unstable manifolds are of interest:

Lemma 4.2 The unstable manifold $W^u(\bar{u})$ is invariant and, furthermore,

$$W^u(\bar{u}) = W^{u,\epsilon}(\bar{u}) \cup \bigcup_{t>0} S(t)\Gamma$$

where

$$\Gamma = W^{u,\epsilon}(\bar{u}) \cap \partial B(\bar{u}, \epsilon). \tag{4.7}$$

Proof. It follows from the definition that, if $u \in W^u(\bar{u})$ then, for every $\tau > 0$ there exists $v^\tau \in \mathbb{R}^p$ such that

$$S(\tau)v^\tau = u \quad v^\tau \to \bar{u} \quad \text{as} \quad \tau \to \infty. \tag{4.8}$$

The converse is also true: if (4.8) holds for every $\tau > 0$ then $u \in W^u(\bar{u})$.

Thus $S(\tau + t)v^\tau = S(t)u$ and, since (4.8) holds, we deduce that $S(t)u \in W^u(\bar{u})$ so that $S(t)W^u(\bar{u}) \subseteq W^u(\bar{u})$. Furthermore, since $S(t)v^t = u$ we have that, for every $t > 0$, $S(\tau - t)v^\tau = v^t$. Thus, from (4.8), we deduce that $v^t \in W^u(\bar{u})$. Thus $W^u(\bar{u}) \subseteq S(t)W^u(\bar{u})$ and the first part of the proof is complete.

We now establish (4.7). First we show that

$$W^u(\bar{u}) \subseteq W^{u,\epsilon}(\bar{u}) \cup \bigcup_{t>0} S(t)\Gamma.$$

Let $u \in W^u(\bar{u}) \backslash W^{u,\epsilon}(\bar{u})$. If $u \notin B(\bar{u}, \epsilon)$ then, since (4.8) holds it follows that $v^0 = u$ and, by continuity, there exists $t > 0$ such that $S(t)v^t = u$ and $v^t \in \Gamma$. On the other hand, if $u \in B(\bar{u}, \epsilon)$ then $\exists t > 0, v^t \in \mathbb{R}^p : S(t)v^t = u$ with $v^t \in \Gamma$ since otherwise we have $u \in W^u, \epsilon(\bar{u})$.

Now we show that

$$W^u(\bar{u}) \supset W^{u,\epsilon}(\bar{u}) \cup \bigcup_{t>0} S(t)\Gamma.$$

If

$$u \in \bigcup_{t \geq 0} S(t)\Gamma$$

then there exists $w \in \Gamma$ such that $S(t)w = u$. Furthermore, since $w \in W^{u,\epsilon}(\bar{u})$ it follows that $u \in W^{u,\epsilon}(\bar{u})$ and the result is proved. □

Important remark It is straightforward to generalize Definition 4.1 to the discrete semigroup $S^1_{\Delta t}$. We employ the notation

$$W^u_{\Delta t}(\bar{U}), \quad W^{u,\epsilon}_{\Delta t}(\bar{U}) \tag{4.9}$$

to denote the unstable and local unstable manifolds of a fixed point $\bar{U}$ respectively. Similar notation is employed for the stable manifold. With this notation it is also straightforward to generalize Theorem 4.2 to the discrete semigroup.

We now discuss the behaviour of trajectories in the neighbourhood of hyperbolic equilibrium points. If we introduce the variable

$$v(t) = u(t) - \bar{u}$$

and change variables in (1.1) then we find that $v(t)$ satisfies the equation

$$\begin{aligned} v_t &= Av + g(v), v(0) = V := U - \bar{u}, \\ A &= \mathrm{d}f(\bar{u}), g(v) = [f(v+\bar{u}) - \mathrm{d}f(\bar{u})v]. \end{aligned} \tag{4.10}$$

Furthermore we introduce the notation $\bar{S} : \mathbb{R}^p \to \mathbb{R}^p$ to denote the semigroup constructed so that $v(t) = \bar{S}v_0$. Hence

$$\bar{S}(t)v = S(t)(\bar{u} + v) - \bar{u}. \tag{4.11}$$

Since $f(\bar{u}) = 0$ it is clear that $g(v) = \mathcal{O}\|v\|^2$ and hence it is reasonable to expect that, for hyperbolic equilibria, the properties of the linear equation $w_t = Aw$ describe the dynamics of solutions to (4.10) in the neighbourhood of $v = 0$. Our aim is to put this intuition on a firm mathematical basis and then to understand analogous behaviour for the numerical method. For the purposes of comparison with the numerical method we now formulate the solution of (4.10) as a mapping over time interval Δt. Since $\bar{u}$ is hyperbolic we can split the space

$$\mathbb{R}^p = X \oplus Y$$

where X (respectively Y) is an l (respectively m) dimensional subspace of $\mathbb{R}^p$ spanned by the generalized eigenspace of A corresponding to eigenvalues with positive (respectively negative) real parts so that $p = l + m$.

We denote by $\mathcal{P}$ and $\mathcal{Q}$ the spectral projections $\mathcal{P} : \mathbb{R}^p \to X$ and $\mathcal{Q} : \mathbb{R}^p \to Y$. Using the variation of constants formula we write the solution of (4.10) as

$$v(t) = \mathrm{e}^{At}v(0) + \int_0^t \mathrm{e}^{A(t-s)} g(v(s))\,\mathrm{d}s. \tag{4.12}$$

Hence we may write

$$v(t) = L(t)v(0) + G(v(0), t) \tag{4.13}$$

where

$$L(t) := \mathrm{e}^{At}, \quad G(v,t) := \int_0^t L(t-s)g(\bar{S}(s)v)\,\mathrm{d}s. \tag{4.14}$$

Now we define $t_n = n\Delta t$, $v_n = v(t_n)$ and write (4.13) as

$$v_{n+1} = Lv_n + G(v_n) \tag{4.15}$$

where $L := L(\Delta t)$ and $G(\cdot) := G(\cdot, \Delta t)$. Using the spectral projections $\mathcal{P}$ and $\mathcal{Q}$ we can decompose v_n as $v_n = p_n + q_n$ where $p_n = \mathcal{P}v_n$ and $q_n = \mathcal{Q}v_n$ to obtain

$$\begin{aligned} p_{n+1} &= Lp_n + \mathcal{P}G(p_n + q_n), \\ q_{n+1} &= Lq_n + \mathcal{Q}G(p_n + q_n). \end{aligned} \tag{4.16}$$

This splitting of the variable v will be particularly useful to us both in studying stable and unstable manifolds and in the examination of phase portraits. Finally note that because of the spectral properties of A on X and Y it follows that there exist norms $\|\cdot\|_u$ and $\|\cdot\|_s$ on $\mathbb{R}^p$ and an $\alpha > 0$ such that, for all $t \geq 0$

$$\begin{aligned} \|L(-t)v\|_u &\leq \mathrm{e}^{-\alpha t}\|v\|_u \quad \forall v \in X, \\ \|L(t)v\|_s &\leq \mathrm{e}^{-\alpha t}\|v\|_s \quad \forall v \in Y. \end{aligned} \tag{4.17}$$

Equivalently

$$\begin{aligned} \|L^{-1}v\|_u &\leq \mathrm{e}^{-\alpha \Delta t}\|v\|_u \quad \forall v \in X, \\ \|Lv\|_s &\leq \mathrm{e}^{-\alpha \Delta t}\|v\|_s \quad \forall v \in Y. \end{aligned} \tag{4.18}$$

For the remainder of this section, whenever estimating v or its numerical counterpart, we employ the norm on $\mathbb{R}^p$ given by

$$\|v\| = \max\{\|\mathcal{P}v\|_u, \|\mathcal{Q}v\|_s\}. \tag{4.19}$$

Here the subscripts 'u' and 's' denote 'unstable' and 'stable' respectively. This choice of norm simplifies the exposition considerably since the majority of the estimation takes place either in X (respectively Y) where $\|\cdot\| \equiv \|\cdot\|_u$ (respectively $\|\cdot\| \equiv \|\cdot\|_s$.)

Using the formulation (4.16) we prove several results concerning the behaviour of solutions of equation (1.1) in the neighbourhood of a hyperbolic equilibrium point $\bar{u}$ – we study stable and unstable manifolds and phase portraits. In subsequent sections we examine the effect of discretization error on these objects. The following two examples serve to illustrate the type of results which interest us here.

Example (Unstable manifolds) Consider the equations (2.1) with the equilibrium point $p = q = 0$. Linearizing about the equilibrium point gives the system

$$p_t = p, \quad q_t = -q. \tag{4.20}$$

Thus, in the notation (4.10), the matrix A has eigenvalues ± 1.

The system (4.20) has an *unstable manifold* given by $q = 0$: on the unstable manifold solutions tend to the origin as $t \to -\infty$ since $q = 0$ and $p(t) = \exp(t)p(0)$. Notice also that the unstable manifold is attractive in the sense that solutions approach it as $t \to \infty$.

Since (2.1) is, in a neighbourhood of the origin, a small perturbation of the linear system (4.20) we expect that it should have an unstable manifold (on which solutions tend to the origin as $t \to -\infty$) close to $q = 0$. This is indeed the case – recall that in Section 2 we showed that the curve $q = p^2/3$ is the unstable manifold for the nonlinear system.

The important point to take away from this example is that the linear system has unstable manifold $q = \Phi_l(p) := 0$ whilst the true equation has unstable manifold $q = \Phi(p) := p^2/3$. It follows that

$$\sup_{|p|\le\epsilon} |\Phi(p) - \Phi_l(p)| \le \tfrac{1}{3}\epsilon^2$$

so that the true and linearized unstable manifolds are close in a neighbourhood of the equilibrium point. Indeed they are tangential at the equilibrium point itself. Our aim is to generalize this to the system (4.10).

The second example broadens our study from the unstable manifold, which comprises certain subclass of trajectories in the neighbourhood of equilibria, to the complete local phase portrait, which comprises all trajectories in the neighbourhood of equilibria.

Example (Phase portrait) We return to the system (2.1) and its linearization (4.20). It is clear that, for any $T > 0$, equations (4.20) can be solved subject to the boundary conditions

$$p(T) = \xi, \quad q(0) = \eta \tag{4.21}$$

and this yields a solution of the form

$$p(t) = \mathrm{e}^{t-T}\xi, \quad q(t) = \mathrm{e}^{-t}\eta.$$

By choosing different ξ and η and different values of T the complete phase portrait for the equation can be constructed. Thus we might expect that a local phase portrait for the nonlinear system can be constructed analogously. We show this to be so.

Explicit solution of (2.1) subject to (4.21) yields

$$p(t) = \mathrm{e}^{t-T}\xi, \quad q(t) = \mathrm{e}^{-t}\eta + \tfrac{1}{3}(\mathrm{e}^{-2T}\xi^2)(\mathrm{e}^{2t} - \mathrm{e}^{-t}).$$

Notice that, if ξ and η are small, then the perturbation from the linear solution is uniformly small in $(0, T)$. Specifically

$$\sup_{0\le t\le T} |\tfrac{1}{3}(\mathrm{e}^{-2T}\xi^2)(\mathrm{e}^{2t} - \mathrm{e}^{-t})| \le \tfrac{1}{3}\xi^2.$$

Thus the perturbation is bounded independently of the time of flight T between boundary conditions.

We now proceed to extend what we have observed in these two examples to the more general case. In the following note that, from (4.14) and the properties of projections, there exists a constant $\kappa > 1$ and $\beta > 0$ such that

$$\|L(t)\| \le \kappa \mathrm{e}^{\beta t} \quad \forall t \ge 0 \quad \text{and} \quad \|\mathcal{P}\|, \|\mathcal{Q}\| \le \kappa. \tag{4.22}$$

Since the function $g(v)$ defined in (4.10) is $\mathcal{O}(\|v\|^2)$ it follows that $\exists C > 0$:

$$\begin{aligned} \|\mathcal{R}g(v)\| &\le (C/2\kappa)\epsilon^2, \quad \mathcal{R} = I, \mathcal{P}, \mathcal{Q} \\ \|\mathcal{R}(g(v) - g(w))\| &\le (C/2\kappa)\epsilon\|v - w\|, \quad \mathcal{R} = I, \mathcal{P}, \mathcal{Q} \end{aligned} \tag{4.23}$$

for all $v, w \in B(0, \epsilon)$. Since our interest in this section is focused on a small neighbourhood of the equilibrium point, smooth modification of the function g outside such a neighbourhood will not affect the results. Thus we assume:

Assumption 4.3 The function g in equation (4.10) satisfies (4.23) for all $v, w \in \mathbb{R}^p$.

After proving results under Assumption 4.3 we will derive corollaries concerning the original, unmodified flow generated by (4.10). Under Assumption 4.3 we have

$$\begin{aligned} \|G(v)\| &\leq (1/2\kappa)C\epsilon^2 \int_0^{\Delta t} \kappa e^{\beta(\Delta t - s)}\, ds \\ &\leq C\epsilon^2[-e^{\beta(\Delta t - s)}/2\beta]_0^t \\ &= [(e^{\beta\Delta t} - 1)/2\beta]C\epsilon^2 \leq \Delta t C\epsilon^2 \end{aligned}$$

for Δt sufficiently small.

Using similar analysis, it is possible to show that $\exists \Delta t_c > 0$ such that

$$\begin{gathered} \|\mathcal{R}G(v)\| \leq \Delta t C\epsilon^2, \quad \mathcal{R} = I, \mathcal{P}, \mathcal{Q} \\ \|\mathcal{R}(G(v) - G(w))\| \leq \Delta t C\epsilon\|v - w\|, \quad \mathcal{R} = I, \mathcal{P}, \mathcal{Q} \end{gathered} \tag{4.24}$$

for all $v, w \in \mathbb{R}^p$ and all $\Delta t \in (0, \Delta t_c]$.

Important remark To simplify notation we denote the norms in the v-coordinates and the u-coordinates in the same way ($\|\cdot\|$). However, the norm in v is always defined through (4.19) whereas the norm in u is defined differently (e.g. the Euclidean norm). This should not cause confusion but it is important to note that, in the remainder of this section, $B(0, \delta)$ denotes a ball in the v-coordinate with norm (4.19) whilst $B(\bar{u}, \delta)$ denotes a ball in the u-coordinates with a different norm.

We start by considering unstable manifolds. Our aim now is to prove the existence of an invariant manifold for the mapping (4.16). Specifically we seek a function $\Phi : X \to Y$ which satisfies the following:

$$q_n = \Phi(p_n) \Leftrightarrow q_{n+1} = \Phi(p_{n+1}). \tag{4.25}$$

We shall look for Φ lying in the space

$$\begin{aligned} \Gamma = \{&\Phi \in C(X, Y) : \|\Phi\|_C = \sup_{p \in X} \|\Phi(p)\| \leq \epsilon, \\ &\|\Phi(p_1) - \Phi(p_2)\| \leq \|p_1 - p_2\| \quad \forall p_1, p_2 \in X\}. \end{aligned}$$

The subscript C in the norm on $C(X, Y)$ is simply to denote the space of functions in which Φ lies.

In the following note that, if $16C\epsilon \leq \alpha$ and $8C\epsilon\Delta t \leq 1$, we have that, for all $\gamma \in [1,2]$,

$$\begin{aligned}(1-\tfrac{1}{2}\Delta t\alpha)\epsilon+\gamma\Delta t C\epsilon^2 &\le \epsilon;\\ 1-\tfrac{1}{2}\Delta t\alpha+2\gamma\Delta t C\epsilon &\le (1-2\gamma\Delta t C\epsilon);\\ (1-\tfrac{1}{2}\Delta t\alpha)+2\gamma\Delta t C\epsilon &\le 1-\tfrac{1}{4}\Delta t\alpha\\ 2\gamma\Delta t C\epsilon &\le \tfrac{1}{2}.\end{aligned} \tag{4.26}$$

Furthermore, note that if $\Phi \in \Gamma$ then

$$\|\mathcal{R}[G(\xi_1+\Phi(\xi_1))-G(\xi_2+\Phi(\xi_2))]\| \le 2\Delta t C\epsilon\|\xi_1-\xi_2\|, \mathcal{R}=I,\mathcal{P},\mathcal{Q}, \tag{4.27}$$

by (4.24). That the estimates (4.26) are robust for $\gamma \in [1,2]$ will be used in Section 4.3 where the same method of proof used to construct an invariant manifold for (4.15) will be employed to construct an invariant manifold for the approximate semigroup; hence C will be enlarged by a factor of 2 to incorporate the effect of the truncation error.

Theorem 4.4 Assume that Δt_c is chosen so that

$$\mathrm{e}^{-\alpha\Delta t} \le 1-\tfrac{1}{2}\alpha\Delta t \,\forall \Delta t \in (0,\Delta t_c],$$

that ϵ is chosen so that $16C\epsilon \le \alpha$ and that $8C\epsilon\Delta t \le 1$. Then, under Assumption 4.3, there exists a function $\Phi \in \Gamma$ so that solutions of (4.16) satisfy (4.25). Furthermore, the graph of Φ is attractive in the sense that

$$\|q_{n+1}-\Phi(p_{n+1})\| \le (1-\tfrac{1}{4}\alpha\Delta t)\|q_n-\Phi(p_n)\|. \tag{4.28}$$

Proof. We use the contraction mapping theorem. Given a function $\Phi \in \Gamma$ consider the construction of a new function $M\Phi : X \mapsto Y$ defined by

$$\begin{aligned}p &= L\xi+\mathcal{P}G(\xi+\Phi(\xi))\\ (M\Phi)(p) &= L\Phi(\xi)+\mathcal{Q}G(\xi+\Phi(\xi)).\end{aligned} \tag{4.29}$$

We show that this new function $M\Phi$ is well defined, lies in Γ and that M contracts on Γ. Thus we construct a fixed point of M; comparison with (4.16) shows that this fixed point is an invariant manifold for (4.16) so that (4.25) is satisfied. Exponential attractivity will then be shown.

To show that $M\Phi$ is well defined we must show that, for every $p \in X$ $\exists \xi \in X$ such that (4.29) is satisfied. To do this consider the mapping

$$\xi^{k+1} = L^{-1}p - L^{-1}\mathcal{P}G(\xi^k+\Phi(\xi^k)).$$

A fixed point of this mapping satisfies (4.29) and will provide the requisite ξ. If η^k also satisfies

$$\eta^{k+1} = L^{-1}p - L^{-1}\mathcal{P}G(\eta^k+\Phi(\eta^k))$$

then, provided that $\Phi \in \Gamma$, (4.18), (4.26) and (4.27) give

$$\|\xi^{k+1}-\eta^{k+1}\| \le 2\mathrm{e}^{-\alpha\Delta t}C\epsilon\Delta t\|\xi^k-\eta^k\| \le \tfrac{1}{2}\|\xi^k-\eta^k\|.$$

Thus $\exists \xi \in X$:

$$\xi = L^{-1}p - L^{-1}\mathcal{P}G(\xi + \Phi(\xi)).$$

Since L is invertible we deduce that $\exists \xi \in X$ so that the first equation in (4.29) can be satisfied for any $p \in X$; hence $M\Phi : X \mapsto Y$ is well defined if $\Phi \in \Gamma$.

Now we show that $M : \Gamma \to \Gamma$. From (4.29) we obtain, using (4.18), (4.24) and (4.26)

$$\begin{aligned} \|M\Phi(p)\| &\le \mathrm{e}^{-\alpha\Delta t}\|\Phi(\xi)\| + \|\mathcal{Q}G(\xi + \Phi(\xi))\| \\ &\le (1 - \tfrac{1}{2}\alpha\Delta t)\epsilon + \Delta t C\epsilon^2 \le \epsilon \end{aligned}$$

as required. Since this is true for every $p \in X$ we have

$$\|M\Phi\|_C \le \epsilon.$$

Also, by considering (4.29) with $\xi \to \{\xi_i\}_{i=1}^2$ and $p \to \{\mathbf{p}_i\}_{i=1}^2$ we obtain, using (4.18) and (4.27),

$$\begin{aligned} \|(M\Phi)(p_1) - (M\Phi)(p_2)\| &\le \mathrm{e}^{-\alpha\Delta t}\|\Phi(\xi_1) - \Phi(\xi_2)\| + 2\Delta t C\epsilon\|\xi_1 - \xi_2\| \\ &\le [\mathrm{e}^{-\alpha\Delta t} + 2\Delta t C\epsilon]\|\xi_1 - \xi_2\|. \end{aligned}$$

But, also from (4.29),

$$\xi_1 - \xi_2 = L^{-1}(p_1 - p_2) - L^{-1}[\mathcal{P}G(\xi_1 + \Phi(\xi_1)) - \mathcal{P}G(\xi_2 + \Phi(\xi_2))]$$

so that by (4.18) and (4.27)

$$\|\xi_1 - \xi_2\| \le \|p_1 - p_2\| + 2\Delta t C\epsilon\|\xi_1 - \xi_2\|.$$

Combining these two estimates we obtain, using (4.26)

$$\|(M\Phi)(p_1) - (M\Phi)(p_2)\| \le \frac{(1 - \frac{1}{2}\alpha\Delta t) + 2\Delta t C\epsilon}{1 - 2\Delta t C\epsilon}\|\xi_1 - \xi_2\| \le \|\xi_1 - \xi_2\|$$

concluding the proof that $M : \Gamma \to \Gamma$.

We now show that $M : \Gamma \to \Gamma$ contracts. Consider (4.29) with $\Phi \to \{\Phi_i\}_{i=1}^2$ and $\xi \to \{\xi_i\}_{i=1}^2$. Now

$$\|M\Phi_1 - M\Phi_2\|_c = \sup_{p \in X}\|M\Phi_1(p) - M\Phi_2(p)\|.$$

Using (4.18) and (4.24) we obtain from (4.16) that, for any $p \in X$

$$\begin{aligned} &\|M\Phi_1(p) - M\Phi_2(p)\| \\ &\le \mathrm{e}^{-\alpha\Delta t}\|\Phi_1(\xi_1) - \Phi_2(\xi_2)\| + \Delta t C\epsilon\|\xi_1 - \xi_2 + \Phi_1(\xi_1) - \Phi_2(\xi_2)\| \\ &\le (\mathrm{e}^{-\alpha\Delta t} + \Delta t C\epsilon)\|\Phi_1(\xi_1) - \Phi_2(\xi_1) + \Phi_2(\xi_1) - \Phi_2(\xi_2)\| + \Delta t C\epsilon\|\xi_1 - \xi_2\| \\ &\le (\mathrm{e}^{-\alpha\Delta t} + \Delta t C\epsilon)\|\Phi_1(\xi_1) - \Phi_2(\xi_1)\| + [\mathrm{e}^{-\alpha\Delta t} + 2\Delta t C\epsilon]\|\xi_1 - \xi_2\| \\ &\le (\mathrm{e}^{-\alpha\Delta t} + \Delta t C\epsilon)\|\Phi_1 - \Phi_2\|_C + [\mathrm{e}^{-\alpha\Delta t} + 2\Delta t C\epsilon]\|\xi_1 - \xi_2\|. \end{aligned}$$

But we also know, by similar reasoning, that

$$\|\xi_1 - \xi_2\| \le \Delta t C\epsilon \|\Phi_1(\xi_1) - \Phi_2(\xi_1)\| + 2\Delta t C\epsilon \|\xi_1 - \xi_2\|$$

so that

$$\|\xi_1 - \xi_2\| \le \frac{\Delta t C\epsilon \|\Phi_1(\xi_1) - \Phi_2(\xi_1)\|}{1 - 2C\epsilon\Delta t} \le \frac{\Delta t C\epsilon \|\Phi_1 - \Phi_2\|_C}{1 - 2C\epsilon\Delta t}.$$

Combining the two estimates and using (4.26) we obtain

$$\begin{aligned}
\|M\Phi_1(p) - M\Phi_2(p)\| &\le (\mathrm{e}^{-\alpha\Delta t} + \Delta t C\epsilon)\|\Phi_1 - \Phi_2\|_C \\
&\quad + \frac{\mathrm{e}^{-\alpha\Delta t} + 2\Delta t C\epsilon}{1 - 2\Delta t C\epsilon}\Delta t C\epsilon \|\Phi_1 - \Phi_2\|_C \\
&\le (\mathrm{e}^{-\alpha\Delta t} + 2\Delta t C\epsilon)\|\Phi_1 - \Phi_2\|_C \\
&\le (1 - \tfrac{1}{4}\alpha\Delta t)\|\Phi_1 - \Phi_2\|_C.
\end{aligned}$$

Since this is true for any $p \in X$ it follows that

$$\|M\Phi_1 - M\Phi_2\|_C \le (1 - \tfrac{1}{4}\alpha\Delta t)\|\Phi_1 - \Phi_2\|_C.$$

Thus the mapping is a contraction and the existence of an invariant manifold satisfying (4.25) follows.

Finally we show that the manifold is attracting. Let

$$\begin{aligned}
p &= Lp_n + \mathcal{P}G(p_n + \Phi(p_n)), \\
\Phi(p) &= L\Phi(p_n) + \mathcal{Q}G(p_n + \Phi(p_n)).
\end{aligned}$$

Subtracting this from (4.16) yields, by (4.26),

$$\begin{aligned}
\|q_{n+1} - \Phi(p_{n+1})\| &\le \|q_{n+1} - \Phi(p)\| + \|\Phi(p) - \Phi(p_{n+1})\| \\
&\le (\mathrm{e}^{-\alpha\Delta t} + \Delta t C\epsilon)\|q_n - \Phi(p_n)\| + \|p - p_{n+1}\| \\
&\le (\mathrm{e}^{-\alpha\Delta t} + 2\Delta t C\epsilon)\|q_n - \Phi(p_n)\| \\
&\le (1 - \tfrac{1}{2}\alpha\Delta t + 2\Delta t C\epsilon)\|q_n - \Phi(p_n)\| \\
&\le (1 - \tfrac{1}{4}\alpha\Delta t)\|q_n - \Phi(p_n)\|
\end{aligned}$$

and the desired result follows. □

Using Theorem 4.4 we may prove:

Corollary 4.5 (**Local unstable manifolds**) Assume that ϵ is chosen so that $16C\epsilon \le \alpha$. Then, there exists a function $\Phi \in \Gamma$ such that, if $v(t)$ satisfies (4.10) and $v(t) \in B(0, \epsilon)$ for $t \in [t_n, t_{n+1}]$ then $v_n = v(t_n)$ satisfies (4.25) and (4.28). Furthermore, there exists $c > 0$ such that the set of points

$$\{u \in \mathbb{R}^p | \mathcal{P}(u - \bar{u}) = p, \mathcal{Q}(u - \bar{u}) = \Phi(p)),\ p \in X\} \cap B(\bar{u}, c\epsilon)$$

is the local unstable manifold of the equilibrium point $\bar{u}$ of (1.1).

Proof. Note that if we are considering a solution $v(t) \in B(0, \epsilon)$ then we can modify g outside $B(0,\epsilon)$ to make Assumption 4.3 valid, without affecting the solution. Thus to establish the first part of the corollary it is sufficient to show that, under Assumption 4.3, Φ is indeed an invariant manifold for the equation (4.10) as well as for the map (4.16). To do this it is sufficient to show that the function Φ constructed in Theorem 4.4 is independent of the choice of $\Delta t \in (0, \Delta t_c]$ used in the construction. To this end we denote Φ by $\Phi(\Delta t)$ and the mapping $M : \Gamma \mapsto \Gamma$ by $M(\Delta t)$ defined by (4.29).

A little work shows that, since M is constructed using the semigroup $S(t)$,

$$M(t) \cdot M(s) \equiv M(s) \cdot M(t).$$

Now consider $t, s \in (0, \Delta t_c]$ and assume that $8C\epsilon\Delta t_c \leq 1$ and that $\mathrm{e}^{-\alpha\Delta t_c} \leq 1 - \alpha\Delta t_c/2$. Then, by Theorem 4.4, both $\Phi(s)$ and $\Phi(t)$ lie in Γ. Furthermore, $M(s)\Phi(t) \in \Gamma$. Now, by definition

$$M(t)\Phi(t) = \Phi(t).$$

Thus

$$M(s) \cdot M(t)\Phi(t) = M(s)\Phi(t) \Rightarrow M(t) \cdot M(s)\Phi(t) = M(s)\Phi(t).$$

Hence, since $M(s)\Phi(t) \in \Gamma$, and since $M(s)\Phi(t)$ is a fixed point of $M(t)$ we deduce that $M(s)\Phi(t) = \Phi(t)$. But this shows that $\Phi(t)$ is a fixed point of $M(s)$ and, since it lies in Γ, we deduce that $\Phi(t) \equiv \Phi(s)$. Hence the manifold $\Phi(\Delta t)$ is independent of Δt and the result follows.

Now we show that the set

$$\mathcal{M} := \{v \in \mathbb{R}^p | \mathcal{P}v = p, \mathcal{Q}v = \Phi(p)\ p \in X\} \cap B(0,\epsilon)$$

defines the local unstable manifold of the equilibrium point 0 of (4.10). We must show that if $v(0) \in \mathcal{M}$ then $v(t) \in \mathcal{M}\ \forall t \leq 0$ and that $v(t) \to 0$ as $t \to -\infty$.

First note that, since $v = 0$ is an equilibrium point for (4.13), analysis of the mapping (4.29) shows that the fixed point Φ satisfies $\Phi(0) = 0$. Now, on the invariant manifold we have from (4.16), that

$$\begin{aligned} p_{n+1} &= Lp_n + \mathcal{P}G(p_n + \Phi(p_n)), \\ 0 &= 0 + \mathcal{P}G(0 + \Phi(0)). \end{aligned}$$

Hence, by (4.18) and (4.27),

$$\begin{aligned} \|p_n\| &\leq \mathrm{e}^{-\alpha\Delta t}\|p_{n+1}\| + \|\mathcal{P}[G(p_n + \Phi(p_n)) - G(0 + \Phi(0))]\| \\ &\leq (1 - \tfrac{1}{2}\alpha\Delta t)\|p_{n+1}\| + 2C\Delta t\epsilon\|p_n\|. \end{aligned}$$

Thus

$$\|p_n\| \leq \frac{1 - \frac{1}{2}\alpha\Delta t}{1 - 2C\Delta t\epsilon}\|p_{n+1}\|$$

$$\leq \frac{1-\frac{1}{2}\alpha\Delta t}{1-\frac{1}{4}\alpha\Delta t}\|p_{n+1}\|$$
$$\leq (1-\tfrac{1}{4}\alpha\Delta t)\|p_{n+1}\|$$

by (4.26).

From this it is clear that, if $v_0 \in \mathcal{M}$ so that $\|p_0\| \leq \epsilon$ then $\|p_n\| \leq \epsilon \; \forall n \leq 0$. Since $\|q_n\| = \|\Phi(p_n)\| \leq \epsilon$ it follows that $v_n \in \mathcal{M} \; \forall n \leq 0$. Since $\Delta t \in (0, \Delta t_c]$ is arbitrary this shows that $v(t) \in \mathcal{M} \; \forall t \leq 0$ as required. Furthermore it is clear that $\|p_n\| \to 0$ as $n \to -\infty$ so that, since $q_n = \Phi(p_n)$ for $v_n \in \mathcal{M}$ and $\Phi(0) = 0$ it follows that $q_n \to 0$ as $n \to -\infty$. Thus $v_n \to 0$ as $n \to -\infty$ and hence that $v(t) \to 0$ as $t \to -\infty$. Thus we have constructed the local unstable manifold for (4.10). Converting back to the u variables from the v variables and changing norms introduces the constant c and completes the proof. □

It should be noted that the methodology used here can be extended to construct the stable manifold of (4.10).

Motivated by the example described above concerning local phase portraits, we now examine the phase portrait of the nonlinear equation (4.10) near to the origin, again using the mapping formulation (4.16). In the particular example above, the result is established easily because the p and q equations decouple. In the general case they do not decouple but, nonetheless, a result of this type still holds. Specifically we seek a solution of (4.16) which satisfies the boundary conditions

$$p_N = \xi \in X, \quad q_0 = \eta \in Y \tag{4.30}$$

where $\|\xi\|, \|\eta\| \leq \epsilon/2$. Noting that $v_n = p_n + q_n$ induction on (4.16) yields

$$\begin{aligned} p_n &= L^{n-N}p_N - \textstyle\sum_{j=n}^{N-1} L^{n-j-1}\mathcal{P}G(v_j), \\ q_n &= L^n q_0 + \textstyle\sum_{j=0}^{n-1} L^{n-1-j}\mathcal{Q}G(v_j). \end{aligned} \tag{4.31}$$

Thus it is our purpose to solve (4.31) subject to (4.30) for arbitrary $N > 0$. Such a solution corresponds to solving the equation (4.10) with boundary conditions specified in X at $t = N\Delta t$ and in Y at $t = 0$ rather than the initial condition $v(0) = V$; since N is arbitrary, the time of flight between these points is arbitrary. Finding all such solutions with ϵ small corresponds to constructing the local phase portrait near $v = 0$.

In the following we let $\mathcal{V} = \{v_n\}_{n=0}^N$ denote an element of the product space $\Psi = \{\mathbb{R}^p\}^N$ and define

$$\|\mathcal{V}\|_\infty = \max_{0\leq n\leq N} \|v_n\|.$$

We consider the set

$$\Psi_\epsilon = \{\mathcal{V} \in \Psi : \|\mathcal{V}\|_\infty \leq \epsilon\}.$$

To study (4.31), (4.30) we use the contraction mapping theorem in Ψ_ϵ. We generate iterates $V^k = \{v_n^k\}_{n=0}^N$ through the definition

$$M\mathcal{V} = \{Mp_n + Mq_n\}_{n=0}^N$$

where $Mp_n \in X$ and $Mq_n \in Y$ are defined by

$$\begin{aligned} Mp_n &= L^{n-N}\xi - \textstyle\sum_{j=n}^{N-1} L^{n-j-1}\mathcal{P}G(v_j), \\ Mq_n &= L^n\eta + \textstyle\sum_{j=0}^{n-1} L^{n-1-j}\mathcal{Q}G(v_j). \end{aligned} \tag{4.32}$$

Clearly a fixed point of M is a solution of (4.30), (4.31).

Now it is straightforward to show that, under the conditions on Δt imposed in Theorem 4.4 and as a result of the bounds (4.18),

$$\begin{aligned} \sum_{j=n}^{N-1} \|L^{n-j-1}v_j\| &\leq \sum_{j=n}^{N-1} \mathrm{e}^{\alpha(n-j-1)\Delta t}\|v_j\| \\ &\leq \left[\frac{1-\mathrm{e}^{-\alpha(N-n)\Delta t}}{1-\mathrm{e}^{-\alpha\Delta t}}\right]\|\mathcal{V}\|_\infty \leq [2/\alpha\Delta t]\|\mathcal{V}\|_\infty \quad \forall v \in X, \\ \sum_{j=0}^{n-1} \|L^{n-1-j}v\| &\leq \sum_{j=0}^{n-1} \mathrm{e}^{-\alpha(n-j-1)\Delta t}\|v_j\| \\ &\leq \left[\frac{1-\mathrm{e}^{-\alpha n\Delta t}}{1-\mathrm{e}^{-\alpha\Delta t}}\right]\|\mathcal{V}\|_\infty \leq [2/\alpha\Delta t]\|\mathcal{V}\|_\infty \quad \forall v \in Y. \end{aligned} \tag{4.33}$$

We may now prove:

Theorem 4.6 Assume that Δt_c is chosen so that

$$\mathrm{e}^{-\alpha\Delta t} \leq 1 - \tfrac{1}{2}\alpha\Delta t \;\forall \Delta t \in (0, \Delta t_c]$$

and that ϵ is chosen so that $8C\epsilon \leq \alpha$. Then, under Assumption 4.3, for any $N > 0$ and any $\xi \in X$, $\eta \in Y$ with $\|\xi\|, \|\eta\| \leq \frac{1}{2}\epsilon$ $\exists$ a solution of (4.16) subject to (4.30) satisfying

$$\max_{0\leq n\leq N} \|\mathbf{v}_n\| \leq \epsilon.$$

Proof. Since (4.16) implies (4.31) we examine (4.31), (4.30). In the following we will use the fact that

$$2\gamma C\epsilon/\alpha \leq \tfrac{1}{2} \tag{4.34}$$

for all $\gamma \in [1, 2]$. Again the factor of γ is incorporated to allow an analogous proof for the numerical method where C is enlarged by a factor of 2 to incorporate the truncation error.

To prove the result we show that $M : \Psi_\epsilon \mapsto \Psi_\epsilon$ and is a contraction. From

(4.32) we have, using (4.18), (4.24) and (4.33), that

$$\begin{aligned}\|Mp_n\| &\leq \|L^{n-N}\xi\| + \sum_{j=n}^{N-1} \|L^{n-j-1}\mathcal{P}G(v_j)\| \\ &\leq \|\xi\| + \frac{2}{\alpha\Delta t}\Delta t C\epsilon^2.\end{aligned}$$

Hence, by the assumptions on ϵ and (4.34) it follows that

$$\|Mp_n\| \leq \epsilon \quad \forall n : 0 \leq n \leq N.$$

Likewise it may be shown that

$$\|Mq_n\| \leq \epsilon \quad \forall n : 0 \leq n \leq N$$

and hence that $M\mathcal{V} \in \Psi_\epsilon$.

To show that the mapping contracts, consider (4.32) with $p_n \mapsto x_n$, $q_n \mapsto y_n$, $v_n \mapsto w_n$, define $w_n = x_n + y_n$ and set $\Omega = \{w_n\}_{n=0}^N$. Then, using (4.18), (4.24) and (4.33) we obtain from (4.32)

$$\|Mp_n - Mx_n\| \leq \frac{2}{\alpha\Delta t}\Delta t C\epsilon\|\mathcal{V} - \Omega\|_\infty \quad \forall n : 0 \leq n \leq N$$

and

$$\|Mq_n - Mz_n\| \leq \frac{2}{\alpha\Delta t}\Delta t C\epsilon\|\mathcal{V} - \Omega\|_\infty \quad \forall n : 0 \leq n \leq N.$$

Thus it follows from (4.34) that

$$\|\mathcal{V}^{k+1} - \Omega^{k+1}\|_\infty \leq \frac{2C\epsilon}{\alpha}\|\mathcal{V}^k - \Omega^k\|_\infty \leq \frac{1}{2}\|\mathcal{V}^k - \Omega^k\|_\infty.$$

Hence $M : \Psi_\epsilon \mapsto \Psi_\epsilon$ is a contraction and the result follows. □

We may now remove the Assumption 4.3 from Theorem 4.6. Our aim is to solve the equation (4.10) subject to specified boundary conditions:

$$v_t = Av + g(v), \quad \mathcal{P}v(T) = \xi,\ \mathcal{Q}v(0) = \eta. \tag{4.35}$$

This is equivalent to solving

$$u_t = f(u), \quad \mathcal{P}(u(T) - \bar{u}) = \xi,\ \mathcal{Q}(u(0) - \bar{u}) = \eta. \tag{4.36}$$

Recall the constant $\kappa > 1$ from (4.22).

Corollary 4.7 **(Phase portraits)** Assume that ϵ is chosen so that $8C\epsilon \leq \alpha$. Then for any $T > 0$ and any $\xi \in X$, $\eta \in Y$ with $\|\xi\|, \|\eta\| \leq \frac{1}{2}\epsilon$ there exists a constant $c > 0$ and a unique solution $u(t)$ of (4.36) satisfying $u(t) \in B(\bar{u}, c\epsilon)$ for all $t \in [0, T]$.

Proof. We consider (4.35). The simple change of variable $u(t) = \bar{u} + v(t)$ will then yield the required result; the constant c is introduced since the norms used to measure $u \in \mathbb{R}^p$ and $v \in \mathbb{R}^p$ may differ. We first show that if (4.35) has a solution in $B(0,\epsilon)$ then it is unique. Let $v^i(t), i = 1,2$ denote two solutions of (4.35) and decompose them as

$$v^i(t) = p^i(t) + q^i(t), \quad p^i(t) \in X,\ q^i(t) \in Y,\ i = 1,2.$$

Projecting the solution appropriately and using the variation of constants formula we obtain

$$\begin{aligned} p^i(t) &= L(t-T)\xi + \textstyle\int_T^t \mathcal{P}L(t-s)g(v^i(s))\,\mathrm{d}s, \\ q^i(t) &= L(t)\eta + \textstyle\int_0^t \mathcal{Q}L(t-s)g(v^i(s))\,\mathrm{d}s. \end{aligned}$$

Thus, by subtracting and using (4.17), (4.23) we obtain

$$\|p^1(t) - p^2(t)\| \le \int_t^T \mathrm{e}^{-\alpha(s-t)}\frac{C}{2\kappa}\epsilon\|v^1(s) - v^2(s)\|\,\mathrm{d}s$$

and

$$\|q^1(t) - q^2(t)\| \le \int_0^t \mathrm{e}^{-\alpha(t-s)}\frac{C}{2\kappa}\epsilon\|v^1(s) - v^2(s)\|\,\mathrm{d}s.$$

Thus it follows that, since $\kappa > 1$,

$$\sup_{0\le t\le T}\|v^1(t) - v^2(t)\| \le \frac{C\epsilon}{2\alpha}\sup_{0\le s\le T}\|v^1(s) - v^2(s)\|.$$

Since $8C\epsilon \le \alpha$ it follows that $\|v^1(t) - v^2(t)\| = 0$ for $t \in [0,T]$ as required.

To establish the existence of a solution in $B(0,\epsilon)$ we simply use Theorem 4.6. For any choice of N, Δt such that $N\Delta t = T$ this gives a solution of (4.35) if Assumption 4.3 holds and, by uniqueness, this solution is independent of the choice of $\Delta t \in (0, \Delta t_c]$. It remains to establish that the solution is in $B(0,\epsilon)$ for all $t \in [0,T]$ so that Assumption 4.3 is not needed. Assume to the contrary that $\exists \tau \in [0,T]$ such that $\|v(\tau)\| = \epsilon + \eta, \eta > 0$. Clearly, for any $\Delta t > 0$ $\exists m \in \mathbb{Z}^+ : \tau \in [m\Delta t, (m+1)\Delta t]$ and, by Theorem 4.6, $\|v_m\|, \|v_{m+1}\| \le \epsilon$. Now, by the boundeness of f it follows that $\exists L > 0 : \|v(t)\| \le \epsilon + L\Delta t\ \forall t \in [m\Delta t, (m+1)\Delta t]$. Since Δt may be chosen arbitrarily small the choice of Δt so that $L\Delta t < \eta$ yields a contradiction. □

It is possible to modify the analysis of phase portraits to prove the existence of stable and unstable manifolds. For brevity we consider stable manifolds. Essentially the stable manifold is constructed by solving (4.30) and (4.31) in the limit $N \to \infty$ whilst asking that $\|p_n\|$ remain uniformly bounded in $n \ge 0$; this yields the problem

$$\begin{aligned} p_n &= -\textstyle\sum_{j=n}^{\infty} L^{n-j-1}\mathcal{P}G(v_j), \\ q_n &= L^n q_0 + \textstyle\sum_{j=0}^{n-1} L^{n-1-j}\mathcal{Q}G(v_j), \\ q_0 &= \eta, \quad \exists\delta > 0 : \|p_n\| \le \delta\ \forall n \ge 0. \end{aligned} \tag{4.37}$$

Recall that $v_n = p_n + q_n$. The following theorem may be proved identically to Theorem 4.6.

Theorem 4.8 Assume that Δt_c is chosen so that $e^{-\alpha\Delta t} \leq 1 - \frac{1}{2}\alpha\Delta t \, \forall \Delta t \in (0, \Delta t_c]$ and that ϵ is chosen so that $8C\epsilon \leq \alpha$. Then, under Assumption 4.3, any $\eta \in Y$ with $\|\eta\| \leq \frac{\epsilon}{2}$ $\exists$ a solution of (4.37) satisfying

$$\max_{0 \leq n < \infty} \|\mathbf{v}_n\| \leq \epsilon.$$

□

We are now in a position to show that the stable manifold has been constructed.

Corollary 4.9 (**Stable manifolds**) Assume that ϵ is chosen so that

$$8C\epsilon \leq \alpha \quad \text{and} \quad \epsilon < \alpha.$$

Then there exists a function $\Phi : X \mapsto Y$ and $c > 0$ such that the set of points

$$\{u \in \mathbb{R}^p | \mathcal{P}(u - \bar{u}) = \Phi(q), \; \mathcal{Q}(u - \bar{u}) = q, \; q \in B(0, \epsilon/2)\}$$

is the local unstable manifold $W^{s,c\epsilon}(\bar{u})$ of the equilibrium point $\bar{u}$ of (1.1).

Proof. It is possible to prove that the solution of (4.37) is independent of Δt as in the proof of Corollary 4.7. To construct Φ solve (4.37) for all $\eta : \|\eta\| \leq \epsilon/2$ and set $\Phi(\eta) = p_0$. It is straightforward to show that

$$\|\Phi(\eta^1) - \Phi(\eta^2)\| \leq 2\|\eta^1 - \eta^2\| \tag{4.38}$$

by considering the Lipschitz properties of solutions to (4.37) with respect to the data η.

Now note that the graph of Φ is positively invariant: given any solution p_n^1, q_n^1 of (4.37) with $p_0^1 = \Phi(\eta_1)$ we can construct another solution p_n^2, q_n^2 of (4.37) by setting $p_n^2 = p_{n+m}^1$, $q_n^2 = q_{n+m}^1$ and imposing the boundary condition that $q_0^2 = \eta^2 = q_m^1$. Hence $\Phi(\eta^2) = p_0^2$; this construction can be done for any $m > 0$ and, since $p_0^2 = p_m^1$, we deduce that $p_m^1 = \Phi(q_m^1)$ so that the graph of Φ is positively invariant.

Thus any solution of (4.37) satisfies

$$q_n = L^n q_0 + \sum_{j=0}^{n-1} L^{n-j-1} \mathcal{Q}G(\Phi(q_j) + q_j).$$

Hence, by (4.33) and (4.38), we have

$$\|q_n\| \leq e^{-\alpha n\Delta t}\|q_0\| + \sum_{j=0}^{n-1} 2C\epsilon\Delta t e^{-\alpha(n-j-1)\Delta t}\|q_j\|.$$

Application of the Gronwall lemma gives

$$\|q_n\| \leq \|q_0\| e^{2(\epsilon - \alpha)n\Delta t}$$

so that, since $\epsilon < \alpha$, $\|q_n\| \to 0$ as $n \to \infty$. Hence $\|p_n\| = \|\Phi(q_n)\| \to 0$ as $n \to \infty$ and the proof is complete. □

This concludes our analysis of equation (1.1) in the neighbourhood of a hyperbolic equilibrium point. We now proceed to study the effect of numerical approximation.

4.2. Equilibrium points and stability

A natural first question to ask about the approximation (1.2) under Assumption 3.7 is whether or not the equilibrium points of (1.1) are inherited by (1.2) and, furthermore, to study the stability of the approximate fixed points. Recall $D(t)$ given by (4.3) and satisfying (4.4).

Theorem 4.10 (Equilibrium points under approximation) Let $\bar{u}$ be a hyperbolic equilibrium point of (1.1). Then there exists $\Delta t_c > 0$ such that the numerical approximation (1.2) has a fixed point $\bar{U} \in B(\bar{u}; 2K\beta\Delta t^r)$ for all $\Delta t \in (0, \Delta t_c]$. Furthermore $\bar{U}$ is stable (respectively unstable) if $\bar{u}$ is stable (respectively unstable).

Proof. For simplicity we consider initially the case where $r > 1$ in Assumption 3.7, returning to $r = 1$ at the end of the proof. The proof is a modification of the proof of the implicit function theorem. Consider the mapping

$$W^{k+1} = W^k - D[W^k - S^1_{\Delta t} W^k] \tag{4.39}$$

where $D = D(\Delta t)$ is given by (4.3). To prove existence of a fixed point of (4.39) we show that the iteration maps $B(\bar{u}; 2K\beta\Delta t^r)$ into itself and is a contraction on that set. Clearly a fixed point of this mapping is necessarily a fixed point of $S^1_{\Delta t}$ and hence of (1.2).

To show that the mapping is into, note that by Definition 3.5, (4.39) may be written as

$$W^{k+1} = W^k - D[W^k - S(\Delta t)W^k + T(W^k; \Delta t)]. \tag{4.40}$$

Also, from (4.2) it follows that

$$\bar{u} = \bar{u} - D[\bar{u} - S(\Delta t)\bar{u}].$$

Let $W^k \in B(\bar{u}; 2K\beta\Delta t^r)$ and set $\mathrm{e}^k = W^k - \bar{u}$. Then (4.40) yields, upon appplication of the mean value theorem,

$$\|\mathrm{e}^{k+1}\| = \|\mathrm{e}^k - D[(I - \mathrm{d}S(\bar{u}, t))\mathrm{e}^k + Q_1 + T(W^k; \Delta t)]\|$$

where

$$\|Q_1\| \leq C_1 \|\mathrm{e}^k\|^2,$$

for some $C_1 > 0$. Thus, using (4.4) and Assumption 3.7(iii),

$$\|\mathrm{e}^{k+1}\| \leq \|D\|\|Q_1\| + \|D\|\|T(W^k; \Delta t)\|$$

$$\begin{aligned} &\leq \frac{\beta C_1 4K^2\beta^2\Delta t^{2r}}{\Delta t} + \beta K \Delta t^r \\ &\leq 2\beta K \Delta t^r \end{aligned}$$

for Δt sufficiently small. Hence the mapping takes $B(\bar{u}, 2K\beta\Delta t^r)$ into itself for Δt sufficiently small.

To show that the mapping is a contraction, let V^k satisfy (4.40) with $W^k \to V^k$ and and define $d^k = W^k - V^k$. A similar manipulation to that used in showing that the mapping is 'into' yields

$$\|d^{k+1}\| \leq \|d^k - D[(I - \mathrm{d}S(\bar{u},t))d^k + Q_2 + T(W^k;\Delta t) - T(V^k;\Delta t)]\|$$

where

$$\|Q_2\| \leq C_2\|d^k\|^2,$$

for some constant $C_2 > 0$. Hence, by (4.4) and Assumption 3.7(iv),

$$\begin{aligned} \|d^{k+1}\| &\leq \|D\|\|Q_2\| + \|D\|\|T(W^k;\Delta t) - T(V^k;\Delta t)\| \\ &\leq \frac{\beta C_2}{\Delta t}\|d^k\|^2 + \beta K \Delta t^r \|d^k\|. \end{aligned}$$

Since $V^k, W^k \in B(\bar{u}; 2K\beta\Delta t^r)$ it follows that $\|d^k\| \leq 4K\beta\Delta t^r$ and hence that

$$\|d^{k+1}\| \leq \tfrac{1}{2}\|d^k\|$$

for Δt sufficiently small. The existence of a fixed point $\bar{U}$ of $S^1_{\Delta t}$ follows for Δt sufficiently small.

To deal with the case $r = 1$ it is sufficient to show that $C_1, C_2 \to 0$ as $\Delta t \to 0$: this holds since $S(\Delta t)$ and $S^1_{\Delta t}$ yield the identity for $\Delta t = 0$.

The stability of $\bar{U}$ follows from the spectral properties of $\mathrm{d}S^1_{\Delta t}\bar{U}$. By Assumption 3.7(iv) the eigenvalues of $\mathrm{d}S^1_{\Delta t}\bar{U}$ converge to those of $\mathrm{d}S(\bar{U};\Delta t)$ as $\Delta t \to 0$; furthermore, the eigenvalues of $\mathrm{d}S(\bar{U};\Delta t)$ converge to those of $\mathrm{d}S(\bar{u};\Delta t)$ as $\Delta t \to 0$ by standard finite-dimensional spectral theory since $\bar{U} \to \bar{u}$. Since $\mathrm{d}S(\bar{u};\Delta t)$ has no eigenvalues on the unit circle it follows that, for Δt sufficiently small, $\mathrm{d}S^1_{\Delta t}\bar{U}$ has the same number of eigenvalues inside and outside the unit circle as $\mathrm{d}S(\bar{u};\Delta t)$ and the result follows. □

Consideration of standard Runge–Kutta methods shows that all equilibrium points of (1.1) become fixed points of the Runge–Kutta method for any $\Delta t > 0$. However, not all fixed points of the Runge–Kutta method are equilibrium points of (1.1) as the following example shows:

Example Consider the scalar equation (1.1) with

$$f(u) = -\lambda u/(1 + u^2)$$

and the Runge–Kutta method

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

Notice that the differential equation has a single equilibrium solution $\bar{u} = 0$. If $\Delta t > 1/\lambda$ then the Runge–Kutta method has the fixed points

$$U = \pm(\lambda \Delta t - 1)^{1/2}$$

in addition to the true fixed point $U = 0$. □

However, it is possible to show that such spurious fixed points cannot exist for Δt sufficiently small; recall (4.6):

Theorem 4.11 **(Spurious solutions as $\Delta t \to 0$)** For any $\epsilon > 0$ $\exists \Delta t_c > 0$ such that $\mathcal{E}_{\Delta t} \in \mathcal{E}(\epsilon)$ $\forall \Delta t \in (0, \Delta t_c]$.

Proof. Let $v \in \mathbb{R}^p \backslash \mathcal{E}(\epsilon)$ so that $\|f(v)\| > \epsilon$. We prove that there exists $\Delta t_c > 0$ such that $v \notin \mathcal{E}_{\Delta t}$ for $\Delta t \in (0, \Delta t_c]$. Note that, from (1.1),

$$S(t)v = v + \int_0^t f(S(s)v)\, \mathrm{d}s. \tag{4.41}$$

Thus, by Assumption 3.6 $\exists L > 0$ such that $\|S(t)v - v\| \leq tL$ and hence, for any $\delta > 0$, $\exists \Delta t_c > 0$:

$$\|S(t)v - v\| \leq \delta \quad \forall t \in (0, \Delta t_c].$$

Thus, by (4.41) and continuity of f,

$$\|S(\Delta t)v - v\| \geq \|\int_0^{\Delta t} f(v) ds\| - \|\int_0^{\Delta t} [f(S(s)v) - f(v)] ds\| \geq \frac{\Delta t \epsilon}{2} \quad \forall \Delta t \in (0, \Delta t_c]$$

possibly by further reduction of Δt_c. Now, by Assumption 3.7(iii),

$$\begin{aligned} \|S^1_{\Delta t} v - v\| &= \|S(\Delta t) - v + S^1_{\Delta t} v - S(\Delta t) v\| \\ &\geq \|S(\Delta t)v - v\| - K \Delta t^{r+1} \\ &\geq \tfrac{1}{2} \Delta t \epsilon - K \Delta t^{r+1} \geq \tfrac{1}{4} \Delta t \epsilon, \end{aligned}$$

possibly by further reduction of Δt_c. Hence $v \notin \mathcal{E}_{\Delta t}$ and the result follows. □

The strength of this result relies heavily on Assumption 3.6. Without Assumption 3.6 Theorem 4.11 can be used to show that, given any ball $B(0, R)$ there exists $\Delta t_c = \Delta t_c(R) > 0$ such that no spurious steady solutions can be found in $B(0, R)$ for all $\Delta t \in (0, \Delta t_c]$.

4.3. Unstable manifolds

In this section we show that the unstable manifolds for (1.1) constructed in Corollary 4.5 persist under numerical approximation and that, furthermore, the numerical unstable manifold is close to the true unstable manifold.

Recall that $U_n \approx u(t_n)$ and define

$$V_n = U_n - \bar{u}.$$

Thus V_n is our numerical approximation to $v(t_n) = u(t_n) - \bar{u}$. Recall also (4.11) and define

$$\bar{S}^1_{\Delta t} v = S^1_{\Delta t}(\bar{u} + v) - \bar{u}.$$

Thus

$$\bar{S}^1_{\Delta t} v = \bar{S}(t)v - S(t)(\bar{u} + v) + S^1_{\Delta t}(\bar{u} + v).$$

Hence

$$\bar{S}^1_{\Delta t} v - \bar{S}(t)v = -T(\bar{u} + v; \Delta t).$$

Using the Definition 3.5 of truncation error and (4.15) we deduce that

$$V_{n+1} = LV_n + G(V_n) - T(\bar{u} + V_n; \Delta t).$$

Defining

$$\tilde{G}(v) = G(v) - T(\bar{u} + v; \Delta t) \tag{4.42}$$

we obtain

$$V_{n+1} = LV_n + \tilde{G}(V_n). \tag{4.43}$$

If we let $P_n = \mathcal{P}V_n$ and $Q_n = \mathcal{Q}V_n$ then (4.43) can be written as

$$\begin{aligned} P_{n+1} &= LP_n + \mathcal{P}\tilde{G}(V_n), \\ Q_{n+1} &= LQ_n + \mathcal{Q}\tilde{G}(V_n). \end{aligned} \tag{4.44}$$

Our aim is to prove that, as for (4.15), the mapping (4.43) has an attractive invariant manifold $\Phi_{\Delta t} : X \mapsto Y$ satisfying

$$Q_n = \Phi_{\Delta t}(P_n) \Leftrightarrow Q_{n+1} = \Phi_{\Delta t}(P_{n+1}). \tag{4.45}$$

and, in addition, to show that Φ and $\Phi_{\Delta t}$ are close.

Using Assumption 3.7(iii) and (iv) it follows from (4.42) that

$$\begin{aligned} \|\tilde{G}(v) - G(v)\| &\leq K\Delta t^{r+1}, \\ \|\tilde{G}(v) - \tilde{G}(w)\| &\leq \|G(v) - G(w)\| + K\Delta t^{r+1}\|v - w\| \end{aligned} \tag{4.46}$$

and hence from (4.24) that under Assumption 4.3,

$$\begin{aligned} \|\mathcal{R}\tilde{G}(v)\| &\leq 2\Delta t C\epsilon^2, \quad \mathcal{R} = I, \mathcal{P}, \mathcal{Q} \\ \|\mathcal{R}(\tilde{G}(v) - \tilde{G}(w))\| &\leq 2\Delta t C\epsilon\|v - w\|, \quad \mathcal{R} = I, \mathcal{P}, \mathcal{Q} \end{aligned} \tag{4.47}$$

for all $v, w \in \mathbb{R}^p$ and all $\Delta t \in (0, \Delta t_c]$. We now exploit this to prove:

Theorem 4.12 Assume that Δt_c is chosen so that

$$e^{-\alpha\Delta t} \leq 1 - \tfrac{1}{2}\alpha\Delta t \;\forall \Delta t \in (0, \Delta t_c]$$

and that ϵ is chosen so that $16C\epsilon \leq \alpha$ and that $8C\epsilon\Delta t \leq 1$. Then, under Assumption 4.3, there exists a function $\Phi_{\Delta t} \in \Gamma$ so that solutions of (4.44) satisfy (4.45). Furthermore, the graph of $\Phi_{\Delta t}$ is attractive in the sense that

$$\|Q_{n+1} - \Phi_{\Delta t}(P_{n+1})\| \leq (1 - \tfrac{1}{4}\alpha\Delta t)\|Q_n - \Phi_{\Delta t}(P_n)\|. \tag{4.48}$$

Finally the graph $\Phi_{\Delta t}$ is close to Φ given in Theorem 4.4 in the sense that

$$\|\Phi - \Phi_{\Delta t}\|_c \leq 8K\Delta t^r/\alpha.$$

Proof. The existence of $\Phi_{\Delta t}$ is proved precisely as for Φ in Theorem 4.4, except that $C \mapsto 2C$ since conditions (4.24) have been replaced by (4.47), by considering the fixed point mapping

$$\begin{aligned} P &= L\xi + \mathcal{P}\tilde{G}(\xi + \Phi_{\Delta t}(\xi)), \\ (M_{\Delta t}\Phi_{\Delta t})(P) &= L\Phi_{\Delta t}(\xi) + \mathcal{Q}\tilde{G}(\xi + \Phi_{\Delta t}(\xi)). \end{aligned} \tag{4.49}$$

Note that the conditions (4.26) employed in the proof of Theorem 4.4 are sufficiently robust to admit essentially the same proof with C enlarged by a factor of 2; indeed this is why they were constructed that way. Hence the existence and attractivity of $\Phi_{\Delta t} \in \Gamma$ follows. It remains to estimate the closeness of $\Phi_{\Delta t}$ to Φ. To do this we use an argument which is essentially the *uniform contraction principle.*

Now, since Φ and $\Phi_{\Delta t}$ both lie in Γ, are fixed points of M and $M_{\Delta t}$ and M has contraction constant $(1 - \frac{1}{4}\alpha\Delta t)$ on Γ it follows that

$$\begin{aligned} \|\Phi - \Phi_{\Delta t}\| &= \|M\Phi - M_{\Delta t}\Phi_{\Delta t}\| \\ &\leq \|M\Phi - M\Phi_{\Delta t}\| + \|M\Phi_{\Delta t} - M_{\Delta t}\Phi_{\Delta t}\|, \\ &\leq (1 - \tfrac{1}{4}\alpha\Delta t)\|\Phi - \Phi_{\Delta t}\| + \|M\Phi_{\Delta t} - M_{\Delta t}\Phi_{\Delta t}\|. \end{aligned}$$

Hence

$$\|\Phi - \Phi_{\Delta t}\| \leq \frac{4}{\alpha\Delta t}\|M\Phi_{\Delta t} - M_{\Delta t}\Phi_{\Delta t}\|.$$

Thus it remains to estimate $M - M_{\Delta t}$.

Clearly

$$\begin{aligned} \|(M_{\Delta t}\Phi_{\Delta t})(P) - (M\Phi_{\Delta t})(P)\| &\leq \|(M_{\Delta t}\Phi_{\Delta t})(P) - (M\Phi_{\Delta t})(p)\| \\ &\quad + \|(M\Phi_{\Delta t})(p) - (M\Phi_{\Delta t})(P)\|. \end{aligned}$$

Since $M\Phi_{\Delta t} \in \Gamma$ we deduce that

$$\|(M_{\Delta t}\Phi_{\Delta t})(P) - (M\Phi_{\Delta t})(P)\| \leq \|(M_{\Delta t}\Phi_{\Delta t})(P) - (M\Phi_{\Delta t})(p)\| + \|p - P\|.$$

Now consider (4.29) with $\Phi \mapsto \Phi_{\Delta t}$ and (4.49); subtracting and using (4.46) we obtain

$$\|(M_{\Delta t}\Phi_{\Delta t})(P) - (M\Phi_{\Delta t})(P)\| \leq 2K\Delta t^{r+1}.$$

Thus, in summary we have that

$$\|\Phi - \Phi_{\Delta t}\| \leq 8K\Delta t^r/\alpha$$

and the proof is complete. □

We can use this result to prove convergence of the local unstable manifold of the map (1.2) to the local unstable manifold of the equation (1.1). Recall that by Theorem 4.10 the map (1.2) has a fixed point $\bar{U}$ close to the equilibrium solution $\bar{u}$ of (1.1). Recall Definition 4.1 and the analogous notation (4.9) for the unstable manifolds of the map (1.2). We may now prove:

Corollary 4.13 (Local unstable manifolds under approximation) Let $W^{u,\epsilon}(\bar{u})$ denote the local unstable manifold of an equilibrium point $\bar{u}$ of (1.1), and $W^{u,\epsilon}_{\Delta t}(\bar{U})$, the local unstable manifold of the fixed point $\bar{U}$ of (1.2) given by Theorem 4.10. Then there exists $C, \Delta t_c, \epsilon_c > 0$ such that for any $\epsilon \in (0, \epsilon_c]$ and $u \in W^{u,\epsilon}(\bar{u})$ there exists $\epsilon' > 0$ and $U \in W^{u,\epsilon'}_{\Delta t}(\bar{U})$ such that

$$\|u - U\| \leq C\Delta t^r \quad \forall \Delta t \in (0, \Delta t_c].$$

That is

$$\mathrm{dist}(W^{u,\epsilon}(\bar{u}), W^{u,\epsilon'}_{\Delta t}(\bar{U})) \leq C\Delta t^r \quad \forall \Delta t \in (0, \Delta t_c].$$

Proof. The existence of a local invariant manifold for (1.2) follows from Theorem 4.12. That it is in fact the unstable manifold of $\bar{U}$ may be proved analogously to the proof of Corollary 4.5; it is necessary to use the fact that $\Phi_{\Delta t}(\mathcal{P}\bar{U}) = \mathcal{Q}\bar{U}$, that is that the fixed point lies on the invariant manifold. The closeness of the two local unstable manifolds follows from the closeness of the graphs Φ and $\Phi_{\Delta t}$ given in Theorem 4.12. The fact that ϵ' may differ from ϵ occurs when changing from a global to a local result since nearby points on the global graphs constructed under Assumption 4.3 may lie in balls of slightly different radii when localizing the result and removing Assumption 4.3. The change in norms yields the constant $C > 0$. □

4.4. Phase portraits and stable manifolds

Our aim in this section is to show the existence of a solution to (4.44) subject to the boundary conditions

$$P_N = \xi, \quad Q_0 = \eta \tag{4.50}$$

for any $N > 0$ and sufficiently small ξ, η. Furthermore, we then show that this solution is $\mathcal{O}(\Delta t^r)$ close to the analogous solution of (4.16) subject to (4.30). Since N is arbitrary this result yields convergence of the approximate trajectories to true trajectories over arbitrarily long-time intervals in the neighbourhood of equilibrium points and *could not be obtained by standard error analysis.* As we shall see, the key to the uniform in time convergence result is that the initial condition for the two trajectories is not the same.

Let $\mathcal{V}_{\Delta t}$ denote the sequence $\{V_n\}_{n=0}^N$ where $V_n = P_n + Q_n$, as in Section 4.3. To solve (4.44), (4.50) we consider finding fixed points $\mathcal{V}_{\Delta t}$ of the mapping $M_{\Delta t} : \Psi \mapsto \Psi$ defined by setting $M_{\Delta t}\mathcal{V}_{\Delta t} = \{M_{\Delta t}P_n + M_{\Delta t}Q_n\}_{n=0}^N$ where

$$\begin{aligned} M_{\Delta t}P_n &= L^{n-N}\xi - \textstyle\sum_{j=n}^{N-1} L^{n-j-1}\mathcal{P}\tilde{G}(V_j), \\ M_{\Delta t}Q_n &= L^n\eta + \textstyle\sum_{j=0}^{n-1} L^{n-1-j}\mathcal{Q}\tilde{G}(V_j). \end{aligned} \tag{4.51}$$

This map should be compared with its continuous counterpart (4.32). To prove existence of a solution to (4.50), (4.51) we follow the method of proof employed for the differential equation. To prove closeness of the solution $\mathcal{V}_{\Delta t}$ to $\mathcal{V}$ we use the uniform contraction principle in a similar manner to the proof of Theorem 4.12.

Theorem 4.14 Assume that Δt_c is chosen so that

$$e^{-\alpha\Delta t} \leq 1 - \tfrac{1}{2}\alpha\Delta t \forall \Delta t \in (0, \Delta t_c]$$

and that ϵ is chosen so that $8C\epsilon \leq \alpha$. Then, under Assumption 4.3, for any $N > 0$ and any $\xi \in X$, $\eta \in Y$ with $\|\xi\|, \|\eta\| \leq \frac{1}{2}\epsilon$ $\exists$ a solution of (4.44) subject to (4.50) satisfying

$$\max_{0\leq n\leq N} \|V_n\| \leq \epsilon.$$

Furthermore the following error estimate holds between the solution v_n of (4.16) and (4.30) and V_n:

$$\max_{0\leq n\leq N} \|\mathbf{v}_n - V_n\| \leq 4K\Delta t^r/\alpha.$$

Proof. To show the existence of $\mathcal{V}_{\Delta t} \in \Psi_\epsilon$ we show that $M_{\Delta t} : \Psi_\epsilon \mapsto \Psi_\epsilon$ is a contraction. This may be achieved by the following the proof of Theorem 4.6; note that (4.47) holds and so it is sufficient to enlarge C by a factor of 2. Since the estimate (4.34) was constructed to be robust under enlargement of C by a factor of 2 the existence of $\mathcal{V}_{\Delta t} \in \Psi_\epsilon$ follows, giving the bound on $\|V_n\|$.

To show convergence of $\mathcal{V}$ to $\mathcal{V}_{\Delta t}$ note that

$$\begin{aligned} \|\mathcal{V} - \mathcal{V}_{\Delta t}\| &= \|M\mathcal{V} - M_{\Delta t}\mathcal{V}_{\Delta t}\| \\ &\leq \|M\mathcal{V} - M\mathcal{V}_{\Delta t}\| + \|M\mathcal{V}_{\Delta t} - M_{\Delta t}\mathcal{V}_{\Delta t}\| \\ &\leq \tfrac{1}{2}\|\mathcal{V} - \mathcal{V}_{\Delta t}\| + \|M\mathcal{V}_{\Delta t} - M_{\Delta t}\mathcal{V}_{\Delta t}\|. \end{aligned}$$

Hence

$$\|\mathcal{V} - \mathcal{V}_{\Delta t}\| \leq 2\|M\mathcal{V}_{\Delta t} - M_{\Delta t}\mathcal{V}_{\Delta t}\|.$$

Now, by consideration of (4.32) with $\mathcal{V} \mapsto \mathcal{V}_{\Delta t}$ (that is $v_n \mapsto V_n$) and (4.51) we deduce that $\|M\mathcal{V}_{\Delta t} - M_{\Delta t}\mathcal{V}_{\Delta t}\|$ can be bounded above by

$$\max\left\{\sum_{j=n}^{N-1} L^{n-j-1}[\mathcal{P}G(V_j) - \mathcal{P}\tilde{G}(V_j)], \sum_{j=0}^{n-1} L^{n-1-j}[\mathcal{Q}G(V_j) - \mathcal{Q}\tilde{G}(V_j)]\right\}.$$

Using (4.33) and (4.46) we thus find that

$$\|M\mathcal{V}_{\Delta t} - M_{\Delta t}\mathcal{V}_{\Delta t}\| \leq \frac{2}{\alpha\Delta t}K\Delta t^{r+1}$$

and hence that

$$\|\mathcal{V} - \mathcal{V}_{\Delta t}\| \leq 4K\Delta t^r/\alpha. \quad \square$$

□

Using Theorem 4.14 we are able to state an interesting result concerning error bounds for approximate solutions of (1.1) near to an equilibrium point.

Consider the boundary value problems (4.36) and the discrete analogue

$$U_{n+1} = \mathcal{F}(U_n; \Delta t), \quad \mathcal{P}(U_N - \bar{u}) = \xi, \; \mathcal{Q}(U_0 - \bar{u}) = \eta. \tag{4.52}$$

We can now prove:

Corollary 4.15 **(Phase portrait under approximation)** There exist $C, \Delta t_c, \epsilon_c > 0$ such that, if $\epsilon \in (0, \epsilon_c]$ then for any $T > 0$ and any $\xi \in X$, $\eta \in Y$ with $\|\xi\|, \|\eta\| \leq \frac{1}{2}\epsilon$ $\exists$ a solution of (4.36) and, if $N\Delta t = T$, a solution of (4.52) satisfying

$$\max_{0 \leq n \leq N} \|U_n - u(t_n)\| \leq C\Delta t^r \quad \forall \Delta t \in (0, \Delta t_c].$$

Proof. This is simply a restatement of Theorem 4.14 in the original variables $u(t)$ and U_n. The change of variables introduces a change in the error constant through the change of norms. □

The important point here is that the error bound is independent of T. Thus it improves upon the standard estimate (1.3) which contains a constant growing exponentially with T. Note that this is achieved by comparing two solutions of (1.1) and (1.2) which *do not share the same initial condition*; specifically, only the projection of the initial condition into the subspace Y is identical at $t = 0$.

We now consider the existence and convergence of stable manifolds under approximation. The local stable manifold for the map (1.2) can be constructed by solving

$$\begin{aligned} P_n &= -\textstyle\sum_{j=n}^{\infty} L^{n-j-1}\mathcal{P}\tilde{G}(V_j), \\ Q_n &= L^n Q_0 + \textstyle\sum_{j=0}^{n-1} L^{n-1-j}\mathcal{Q}\tilde{G}(V_j), \\ Q_0 &= \eta, \quad \exists\delta > 0 : \|P_n\| \leq \delta \; \forall n \geq 0. \end{aligned} \tag{4.53}$$

where $V_n = P_n + Q_n$ and P_n, Q_n and $\tilde{G}(\cdot)$ are defined by (4.44), (4.42). The stable manifold is formed from solutions of (4.53) as the graph $\Theta : Y \mapsto X$ given by $\Theta(\eta) = P_0$.

Since the proof of Theorem 4.8 is robust to enlargement of C by a factor of 2 it follows that

Theorem 4.16 Assume that Δt_c is chosen so that

$$\mathrm{e}^{-\alpha\Delta t} \leq 1 - \tfrac{1}{2}\alpha\Delta t \forall \Delta t \in (0, \Delta t_c]$$

and that ϵ is chosen so that $8C\epsilon \leq \alpha$. Then, under Assumption 4.3, for any $\eta \in Y$ with $\|\eta\| \leq \frac{1}{2}\epsilon$ $\exists$ a solution of (4.53) satisfying

$$\max_{0 \leq n < \infty} \|\mathbf{V}_n\| \leq \epsilon.$$

We are now in a position to show that the stable manifold is well approximated numerically.

Corollary 4.17 (Local stable manifolds under approximation) Let $W^{s,\epsilon}(\bar{u})$ denote the local unstable manifold of an equilibrium point $\bar{u}$ of (1.1), and $W^{s,\epsilon}_{\Delta t}(\bar{U})$, the unstable manifold of the fixed point $\bar{U}$ of (1.2) given by Theorem 4.10. Then there exists $C, \Delta t_c, \epsilon_c > 0$ such that for any $\epsilon \in (0, \epsilon_c]$ and $u \in W^{s,\epsilon}(\bar{u})$ there exists $\epsilon' > 0$ and $U \in W^{s,\epsilon'}_{\Delta t}(\bar{U})$ such that

$$\|u - U\| \le C\Delta t^r \quad \forall \Delta t \in (0, \Delta t_c].$$

That is

$$\text{dist}(W^{s,\epsilon}(\bar{u}), W^{s,\epsilon'}_{\Delta t}(\bar{U})) \le C\Delta t^r \quad \forall \Delta t \in (0, \Delta t_c].$$

Proof. The existence of a local invariant manifold for (1.2) follows from Theorem 4.16 by setting $\Phi_{\Delta t}(\eta) = P_0$ for every $\eta : \|\eta\| \le \epsilon/2$. That it is in fact the stable manifold of $\bar{U}$ may be proved analogously to the proof of Corollary 4.9; it is necessary to use the fact that $\Phi_{\Delta t}(\mathcal{Q}\bar{U}) = \mathcal{P}\bar{U}$, that is that the fixed point lies on the invariant manifold. The closeness of the two local stable manifolds follows from the closeness of the solution v_n and V_n of (4.37) and (4.53) given in Theorems 4.8 and 4.16. The change in norms yields the constant $C > 0$. □

4.5. Bibliography

For background material concerning equilibria, fixed points, hyperbolicity and stability see Hale and Kŏcak (1991) and Wiggins (1990). For discussion of unstable manifolds see Babin and Vishik (1992), Hale (1988), Hale and Kŏcak (1991) and Wiggins (1990). The construction of unstable manifolds described in Section 4.1 is based on an approach known as the *Hadamard graph transform*; this transform technique can also be used to construct stable and centre manifods. The construction of the phase portrait in Section 4.1 is closely related to the *Hartman–Grobman* theorem which states that there is a 1:1 correspondence between solutions of (1.2) and its linearization in the neighbourhood of a hyperbolic equilibrium point. See Hartman (1982). The construction of the stable manifold in Section 4.1 is based on the *Lyapunov–Perron* technique, here modified from differential equations to mappings. For a discussion of this technique see, for example, Carr (1982) and Medved (1991). Again, this technique can be modifed to construct unstable and centre manifolds.

Since most numerical methods for ordinary differential equations replicate exactly all the equilibria of the underlying equation as fixed points of the numerical method, Theorem 4.10 may seem a little pointless. However, the method of proof employed there can be used to study approximation of partial differential equations where exact preservation of equilibria under spatial approximation does not occur; see Crouziex and Rappaz (1990). Analysis of the existence of spurious solutions introduced by discretization using techniques from dynamical systems can be traced back to the article by Newell (1977) and the subsequent related work undertaken in Mitchell and Griffiths (1986) and Stuart (1989a,b); all these articles concerned spurious solutions oscillating on a grid scale in time. The article by Brezzi *et al.* (1984) considered the existence of spurious invariant curves introduced by time discretization. However, it was not until the work of Iserles (1990) that the more interesting question of the possible existence of spurious equilibrium solutions was investigated. He showed that Runge–Kutta methods could admit spurious equilibria whilst linear multistep methods could not. Subsequent analysis of this phenomena can be found in Hairer *et al.* (1990), Yee *et al.* (1991) and Griffiths *et al.* (1992). It is fair to say that the area of spurious solutions introduced by time discretization is now very well understood – the article by Iserles *et al.* (1991) puts the subject in a unified framework whilst in Humphries (1993) it is proved that spurious solutions must either converge to true solutions or become unbounded as $\Delta t \to 0$; such a result can also be deduced from Theorem 4.11. There is probably little of interest remaining to do in the area of spurious solutions introduced by fixed time-step time discretization. Note also that it is reasonable to expect that, under many circumstances, codes which vary the time-step to control the local error will also prevent spurious solutions. Such a result was conjectured in Sanz-Serna (1992b) and is proved for certain error control schemes applied to (1.1) under a variety of a structural assumptions in Stuart and Humphries (1992b). The effect of spurious solutions introduced by spatial discretization is an area in which there are still many open questions. For representative work in this area see Beyn and Doedel (1981), Budd (1991), Murdoch and Budd (1990), Elliott and Stuart (1993) and Stephens and Shubin (1987).

The first proof of convergence of local stable and unstable manifolds, together with phase portraits, was contained in the article by Beyn (1987b). This article employed a very clean presentation, involving use of a Lipschitz inverse mapping theorem. The approach presented there can be extended to multistep methods. We have chosen to present a more transparent, if lengthier, proof of the convergence of local unstable manifolds; it is based on a similar proof for centre-unstable manifolds contained in Beyn and Lorenz (1987) involving the Hadamard graph transform. It should be noted that Corollary 4.13 can easily be extended to show that the distance $d_H(\cdot, \cdot)$ between the local unstable manifolds is small, rather than just the semi-distance described. A thorough study of the behaviour of discretizations near equilibria may be found in Garay (1993) which unifies and extends much of the work described here.

The proof we employ to construct the phase portrait follows the approach taken in Larrson and Sanz-Serna (1993) very closely (see also Sanz-Serna and Larsson (1993)) where finite-element approximations of reaction-diffusion equations are studied; their approach is closely related to the construction of stable and unstable manifold by the Lyapunov–Perron method in Henry (1981). Hence the Lyapunov–Perron method underlies our result concerning the convergence of stable manifolds. The approach of Beyn (1987b) has been generalized to study the numerical approximation of certain partial differential equations in Alouges and Debussche (1991).

Whilst on the subject of unstable manifolds, it is relevant to mention the literature concerning the effect of numerical approximation on *inertial manifolds.* These attractive invariant manifolds for partial differential equations on a Hilbert space $\mathcal{H}$ may be represented as graphs relating a certain projection of the space $\mathcal{H}$ to its complement. See Foais *et al.* (1988), Mallet-Paret and Sell (1988) and Constantin *et al.* (1989) for the background theory. The original construction of the inertial manifold in Foais *et al.* (1988) uses the Lyapunov–Perron approach and contains a convergence result concerning the effect of Galerkin approximation on the inertial manifold; a related method of analysis was employed in Demengel and Ghidaglia (1989) to study the effect of a particular time discretization on the problem. In Jones and Stuart (1993) the inertial manifold is constructed by use of a technique similar to that employed to prove Theorem 4.4 (the Hadamard graph transform) and a convergence proof, sufficiently general to include a variety of numerical approximations and similar to the proof of Corollary 4.13, is given.

5. Periodic solutions and invariant tori

5.1. Background theory

In this section we study the effect of discretization on periodic solutions of (1.1); we shall not describe the theory for quasi-periodic solutions but give some references to the literature in the final section. The methods employed are very similar to those we describe for the study of periodic solutions.

For simplicity we assume that the periodic solution is stable and hyperbolic – we shall be precise about the meaning of this later on – see (5.6). Let us assume that (1.1) has the periodic solution $\bar{u}(t)$ with period T :

$$\{\bar{u}(t) \in C^1(\mathbb{R}, \mathbb{R}^p) | \bar{u}(t+T) = \bar{u}(t)\ \forall t \in \mathbb{R}\}. \tag{5.1}$$

In order to facilitate study of the periodic solution, we introduce new coordinates $r \in \mathbb{R}^{p-1}$ and $\theta \in \mathbb{R}$ where, roughly, r measures the coordinates normal to the tangent space of the periodic solution and θ measures an angular coordinate in the tangent space of the periodic solution. Letting $v = (r^T, \theta)^T \in \mathbb{R}^p$ it may be shown that there exists a C^3 diffeomorphism $\chi : \mathbb{R}^p \mapsto \mathbb{R}^p$ under which the transformation $u = \chi(v)$ renders (1.1) in a very useful form. Specifically it may be shown that we obtain

$$\begin{aligned} r_t &= A(\theta)r + g(r,\theta), \quad r(0) = \xi, \\ \theta_t &= 1 + h(r,\theta), \quad \theta(0) = \phi. \end{aligned} \tag{5.2}$$

Here A, g and h are C^2 in a neighbourhood of the periodic solution and satisfy the following conditions for all $\theta \in \mathbb{R}$:

$$g(0,\theta) = 0, \quad g_r(0,\theta) = 0, \quad h(0,\theta) = 0, \tag{5.3}$$

$$A(\theta + T) = A(\theta), \quad g(r,\theta+T) = g(r,\theta), \quad h(r,\theta+T) = h(r,\theta). \tag{5.4}$$

Thus the periodic solution is simply $r = 0$ and $\theta = t$ in this coordinate system and A, g and h are defined for all $\theta \in \mathbb{R}$ and all $r \in B(0,\epsilon)$, for some ϵ sufficiently small.

For simplicity consider the case $p = 2$ so that $r(t) \in \mathbb{R}$. If we define

$$\begin{aligned} B(s,t;\phi) &:= \exp[\int_s^t A(\phi+\tau)\mathrm{d}\tau], \\ B &:= B(0,T;\phi), \end{aligned} \tag{5.5}$$

then, using the fact that the periodic solution is hyperbolic and stable, it follows that there exists a norm on $\mathbb{R}^{p-1}$ in which the following property holds for B:

$$\|B\| \le \alpha < 1 \quad \forall \phi \in \mathbb{R} \tag{5.6}$$

In dimension $p > 2$ a bound similar to (5.6) holds where $B(s,t;\phi)$ is replaced by the solution operator for the non-autonomous equation

$$r_t = A(\phi + t)r, \quad r(s) = \psi.$$

Whenever we require use of $B(s,t;\phi)$ in this chapter we will consider the case $p = 2$ and refer to the representation (5.5) for $B(s,t;\phi)$; however, by using the more general definition of $B(s,t;\phi)$ instead of (5.5) arbitrary $p > 2$ may be considered similarly. We will employ the norm on $\mathbb{R}^{p-1}$ given in (5.6) throughout the remainder of our discussion of periodic solutions. The following example illustrates the transformation of variables just described.

Example Consider equations (3.12). We modify the change to polar coordinates used to study these equations and introduce the variables

$$r = R - 1, \quad \theta = -\phi.$$

Thus

$$x = (1+r)\cos\theta, \quad y = -(1+r)\sin\theta. \tag{5.7}$$

Then, from (3.13), we obtain

$$r_t = -2r - (3r^2 + r^3), \quad \theta_t = 1. \tag{5.8}$$

Hence $A(\theta) = -2$, $g(r,\theta) = -(3r^2 + r^3)$ and $h(r,\theta) = 0$. Note that (5.3) and (5.4) are trivially satisfied. Furthermore

$$B(s,t;\phi) = \mathrm{e}^{-2(t-s)}, \quad B = \mathrm{e}^{-4\pi}$$

since the period $T = 2\pi$. This shows that (5.6) is satisfied. □

As we have observed, the existence of a periodic solution $r = 0, \theta = t$ in (5.2) is trivial under (5.3), (5.4). However, since our aim is to develop an existence theory which is sufficiently robust to incorporate the effect of numerical approximation at a later point, we must relax (5.3). The crucial consequence of (5.3) (which is also shared by equations generated by applying $u = \chi(v)$ to equations (1.1) found from smooth perturbations of a vector field yielding (5.2)) is that $\exists C_1 > 0$:

$$\begin{gathered} \|B(0,t;\theta)\| \le C_1 \ \forall t \in [0,T], \quad \|g(r,\theta)\|, \|g_\theta(r,\theta)\| \le C_1\epsilon^2, \\ \|h(r,\theta)\|, \|h_\theta(r,\theta)\|, \|g_r\| \le C_1\epsilon, \quad \|h_r(r,\theta)\| \le C_1 \end{gathered} \tag{5.9}$$

for all $r \in B(0,\epsilon)$, $t \in \mathbb{R}$ where the subscripts r and θ denote appropriate derivatives. Thus, by considering (5.9) instead of (5.3) we are considering the effect of small perturbations of the vector field $f(\cdot)$ on the periodic solution $\bar{u}(t)$.

The mappings generated by (5.2) under (5.9) are sufficiently general to enable us to incorporate the effect of numerical approximations within the same framework.

Since our interest in now focused on a small neighbourhood of the periodic solution, smooth modification of the functions g and h outside such a neighbourhood will not affect the results. Thus we assume:

Assumption 5.1 The functions

$$A(\theta) \in C^2(\mathbb{R}^{p-1}, \mathbb{R}^{p-1}), \quad g(r,\theta) \in C^2(\mathbb{R}^{p-1} \times \mathbb{R} \quad \mathbb{R}^{p-1})$$

and

$$h(r,\theta) \in C^2(\mathbb{R}^{p-1} \times \mathbb{R}, \mathbb{R}).$$

Furthermore, (5.9) and (5.4) are satisfied for all $r \in \mathbb{R}^{p-1}$ and $\theta \in \mathbb{R}$

Under Assumption 5.1 it is straightforward to prove that (5.2) has a unique solution for all $t \ge 0$. The following example illustrates that periodic solutions persist under perturbations to the vector field $f(\cdot)$ in (1.1) such that Assumption 5.1 holds.

Example If, instead of considering equations (5.8) we modify g and h to obtain

$$\begin{gathered} r_t = -2r + \epsilon^2\{\cos((1+\epsilon)\theta) + 2\sin\theta\}, \\ \theta_t = 1 + \epsilon, \end{gathered} \tag{5.10}$$

then we obtain a solution in the form

$$r(t) = \epsilon^2 \sin((1+\epsilon)t), \quad \theta(t) = (1+\epsilon)t. \tag{5.11}$$

Recalling the transformation (5.7) it is clear that the solution (5.11) yields a periodic solution in x, y coordinates. □

After proving results under Assumption 5.1 we will derive a corollary concerning the original, unmodified flow generated by (1.1). We shall employ the theory of attractive invariant manifolds to construct a periodic solution for (5.2) under Assumption 5.1 and to study the effect of numerical approximation. The basic idea of the proof is that if $g, h \equiv 0$ then (5.2) has solution

$$r(t) = B(0,t;\phi)\xi, \quad \theta(t) = \phi + T.$$

Hence $\|r(mT)\| \leq \alpha^m \xi$ by (5.6). It is this contractivity that we wish to exploit in the case where g and h are small, but not identically zero. Hence we integrate (5.2), using the integrating factor $\{B(0,t;\phi)\}^{-1}$ and the variation of constants formula to obtain

$$\begin{aligned} r(t) &= B(0,t;\phi)\xi + \int_0^t B(s,t;\phi)Z(s)\,\mathrm{d}s, \\ \theta(t) &= \phi + t + \int_0^t h(r(s),\theta(s))\,\mathrm{d}s, \end{aligned} \tag{5.12}$$

where

$$Z(s) = [A(\theta(s)) - A(\phi+s)]r(s) + g(r(s),\theta(s)).$$

In order to exploit the contractivity induced by (5.6) it will be convenient to consider the solution of (5.12) at time $t = T$. Denoting the solutions of (5.2) by $r(s)$ and $\theta(s)$ and noting that these vectors are functions of ξ and ϕ we obtain from (5.12)

$$\begin{aligned} r(T) &= B\xi + G(\xi,\phi), \\ \theta(T) &= \phi + T + H(\xi,\phi), \end{aligned} \tag{5.13}$$

where $G : \mathbb{R}^{p-1} \times \mathbb{R} \mapsto \mathbb{R}^{p-1}$ and $H : \mathbb{R}^{p-1} \times \mathbb{R} \mapsto \mathbb{R}$ are defined in the following way

$$\begin{aligned} G(\xi,\phi;\tau) &:= \int_0^\tau B(s,\tau;\phi)Z(s)\,\mathrm{d}s, \\ H(\xi,\phi;\tau) &:= \int_0^\tau h(r(s),\theta(s))\,\mathrm{d}s, \\ G(\xi,\phi) &:= G(\xi,\phi;T), \\ H(\xi,\phi) &:= H(\xi,\phi;T). \end{aligned} \tag{5.14}$$

We now give an example to illustrate the preceding definitions.

Example Consider the equations

$$r_t = -2r + \epsilon^2, \quad \theta_t = 1 + \epsilon.$$

Then it follows that

$$G(\xi,\phi) = \epsilon^2(1-\mathrm{e}^{-2T})/2, \quad H(\xi,\phi) = \epsilon T.$$

The following space of functions will be useful in the succeeding analysis:

$$\begin{aligned}\Gamma = \{\Psi \in C(\mathbb{R},\mathbb{R}^{p-1}) : \|\Psi\|_P = \sup\nolimits_{\theta\in[0,T)} \|\Psi(\theta)\| \leq K\epsilon^2,\\ \|\Psi(\theta_1)-\Psi(\theta_2)\| \leq \epsilon|\theta_1-\theta_2|,\ \Psi(\theta_1+T)=\Psi(\theta_1)\ \forall\theta_1,\theta_2\in\mathbb{R}\}.\end{aligned} \tag{5.15}$$

The subscript P on the norm simply denotes the space of periodic functions in $C(\mathbb{R},\mathbb{R}^{p-1})$ with period T. After Lemma 5.2 we will require a particular value for K and use

$$K = 4C_1T/(1-\alpha). \tag{5.16}$$

Our aim is to find an invariant manifold for (5.13), namely $\Phi \in \Gamma$ such that

$$\xi = \Phi(\phi) \Leftrightarrow r(T) = \Phi(\theta(T)). \tag{5.17}$$

Comparison with (5.13) shows that to do this is equivalent to finding a fixed point of the mapping $\mathcal{T}$ defined by

$$\begin{aligned}(\mathcal{T}\Phi)(\theta) = B\Phi(\phi) + G(\Phi(\phi),\phi),\\ \theta = \phi + T + H(\Phi(\phi),\phi).\end{aligned} \tag{5.18}$$

Our proof of existence of such a fixed point is closely related to our proof of the existence of an unstable manifold of a fixed point, given in Section 4. We first establish certain 'smallness' properties of G, H and their Lipschitz constants; we then show that $\mathcal{T}$ maps Γ into itself and finally that $\mathcal{T}$ is a contraction on Γ. The details are more complicated than for the unstable manifold and hence we break up the proof into a sequence of lemmas. The proof of Lemma 5.2, in particular, is very technical and may be omitted without disrupting the flow.

Lemma 5.2 Assume that Assumption 5.1 holds. Let $\Phi^i \in \Gamma$, $i = 1,2$. Then, for all $\phi^i \in \mathbb{R}$, $i = 1,2$ it follows that there exists $C_2 = C_2(T)$ and $\epsilon^* > 0$ such that, for all $\epsilon \in (0,\epsilon^*]$,
(i) $|H(\Phi^1(\phi^1),\phi^1)| \leq C_1T\epsilon$;
(ii) $\|G(\Phi^1(\phi^1),\phi^1)\| \leq 2C_1T\epsilon^2$;
(iii) $|H(\Phi^1(\phi^1),\phi^1) - H(\Phi^1(\phi^2),\phi^2)| \leq C_2\epsilon|\phi^1-\phi^2|$;
(iv) $\|G(\Phi^1(\phi^1),\phi^1) - G(\Phi^1(\phi^2),\phi^2)\| \leq C_2\epsilon^2|\phi^1-\phi^2|$;
(v) $|H(\Phi^1(\phi),\phi) - H(\Phi^2(\phi),\phi)| \leq C_2\|\Phi^1-\Phi^2\|_P$;
(vi) $\|G(\Phi^1(\phi),\phi) - G(\Phi^2(\phi),\phi)\| \leq C_2\epsilon\|\Phi^1-\Phi^2\|_P$.

Proof. Throughout the proof we use $C(T)$ to denote a global constant independent of ϵ.
(i), (ii) Consider equations (5.12) with $\xi = \Phi(\phi)$ and let Assumption 5.1 hold. It follows that

$$|\theta(t) - \phi - t| \leq C_1t\epsilon \leq C_1T\epsilon,\ \forall t\in[0,T] \tag{5.19}$$

and hence the result (i) follows. From (5.12), using Assumption 5.1, (5.15), 5.19 and the assumption that $r(0) = \Phi(\phi)$ we have that

$$\|r(t)\| \leq C_1 K\epsilon^2 + \int_0^t [C(T)\epsilon\|r(s)\| + C_1\epsilon^2]\,\mathrm{d}s \quad \forall t \in [0,T]. \tag{5.20}$$

Application of the Gronwall lemma yields

$$\|r(t)\| \leq C_1(K+T)\epsilon^2 \mathrm{e}^{C(T)\epsilon t}, \quad \forall t \in [0,T]. \tag{5.21}$$

Thus (5.12) gives, using (5.20), (5.21) and Assumption 5.1,

$$\|r(T) - B\Phi(\phi)\| \leq \int_0^T [C(T)\epsilon^3 + C_1\epsilon^2]\,\mathrm{d}s \leq 2C_1T\epsilon^2,$$

for ϵ sufficiently small. This yields (ii).

We also derive a related estimate used in the proof of Corollary 5.6. By an argument similar to that giving (5.20) we may show that

$$\|r(t) - \xi\| \leq \|B(0,t;\phi) - I\|K\epsilon^2 + \int_0^t [C(T)\epsilon\|r(s)\| + C_1\epsilon^2]\,\mathrm{d}s.$$

Using (5.5) and (5.21) it follows that, for t sufficiently small, there exists $C_3 > 0$ such that

$$\|r(t) - \xi\| \leq C_3 tK\epsilon^2 \Rightarrow \|r(t)\| \leq (1 + C_3 t)K\epsilon^2. \tag{5.22}$$

(iii), (iv) Consider the equations

$$\begin{array}{rclr} r_t^i &=& A(\theta^i)r^i + g(r^i,\theta^i), & r^i(0) = \Phi(\phi^i) \\ \theta_t^i &=& 1 + h(r^i,\theta^i), & \theta^i(0) = \phi^i \end{array} \tag{5.23}$$

for $i = 1,2$. Define $\delta = \phi^1 - \phi^2,\ \rho = r^1 - r^2, \gamma = \theta^1 - \theta^2$ and

$$\|\rho\|_\infty = \sup_{0 \leq t \leq T} \|\rho(t)\|.$$

Note that, since $\Phi \in \Gamma$, we have that

$$\gamma(0) = \delta,\ \|\rho(0)\| \leq \epsilon|\delta|. \tag{5.24}$$

Now, by the mean value theorem,

$$\gamma_t = \bar{h}(r^1,r^2,\theta^1)\rho + h_\theta(r^2,\beta)\gamma \tag{5.25}$$

where

$$\bar{h}(r^1,r^2,\theta) := \int_0^1 h_r(zr^1 + (1-z)r^2,\theta)\,\mathrm{d}z$$

and $\beta = \beta(t) := (1-\zeta)\theta^1 + \zeta\theta^2$ for some $\zeta = \zeta(t) \in [0,1]$.

Thus, by (5.24),

$$\gamma(t) = \delta + \int_0^t \{\bar{h}(r^1(s),r^2(s),\theta^1(s))\rho(s) + h_\theta(r^2(s),\beta(s))\gamma(s)\}\,\mathrm{d}s. \tag{5.26}$$

Hence, by Assumption 5.1, we obtain

$$|\gamma(t)-\delta| \le C(T)\|\rho\|_\infty + \int_0^t C_1\epsilon|\gamma(s)|\,\mathrm{d}s \quad \forall t\in[0,T] \tag{5.27}$$

and application of the Gronwall lemma yields

$$|\gamma(t)| \le C(T)[|\delta| + \|\rho\|_\infty], \quad \forall t\in[0,T]. \tag{5.28}$$

Also, we have that

$$\rho_t = A(\theta^1)\rho + [A(\theta^1)-A(\theta^2)]r^2 + \bar{g}(r^1,r^2,\theta^1)\rho + g_\theta(r^2,\alpha)\gamma,$$

where

$$\bar{g}(r^1,r^2,\theta) := \int_0^1 g_r(zr^1+(1-z)r^2,\theta^1)\,\mathrm{d}s$$

and $\alpha=\alpha(t) := (1-\zeta)\theta^1+\zeta\theta^2$ for some $\zeta=\zeta(t)\in[0,1]$.

Hence integration yields

$$\begin{aligned}\rho(t) = \;& B(0,t;\phi^1)\rho(0) + \int_0^t B(s,t;\phi^1)\{[A(\theta^1(s)) - A(\phi^1+s)]\rho(s)\\ &+[A(\theta^1(s))-A(\theta^2(s))]r^2(s) + \bar{g}(r^1,r^2,\theta^1(s))\rho(s)\\ &+g_\theta(r^2(s),\alpha(s))\gamma(s)\}\,\mathrm{d}s.\end{aligned} \tag{5.29}$$

From 5.19, (5.21), (5.24), (5.28) and Assumption 5.1 we deduce that

$$\begin{aligned}\|\rho(t)\| &\le C_1\epsilon|\delta| + \int_0^t C(T)\epsilon[\|\rho(s)\| + \epsilon|\gamma(s)|]\,\mathrm{d}s\\ &\le C_1\epsilon|\delta| + C(T)\epsilon[\epsilon|\delta| + \|\rho\|_\infty].\end{aligned} \tag{5.30}$$

Hence we deduce that, for ϵ sufficiently small,

$$\|\rho\|_\infty \le C(T)\epsilon|\delta| \quad \forall t\in[0,T]. \tag{5.31}$$

Thus (5.28) gives

$$|\gamma(t)| \le C(T)(1+\epsilon)|\delta| \tag{5.32}$$

and (5.27) yields, for ϵ sufficiently small,

$$|\gamma(T)-\delta| \le \epsilon C_2(T)|\delta|$$

as required for (iii). Now let $B^i := B(0,T;\phi)$ and note that

$$\|B^1-B^2\| \le c|\delta|.$$

Returning to (5.29) we obtain, by (5.31), (5.32)

$$\|\rho(T)-B^1\rho(0)\| \le \int_0^T C(T)\epsilon[\|\rho(s)\|+\epsilon|\gamma(s)|]\,\mathrm{d}s \le \epsilon^2 C(T)|\delta|.$$

Hence

$$\|\rho(T) - (B^1r^1(0)-B^2r^2(0))\| \;\le\; \|\rho(T)-B^1\rho(0)\| + \|(B^2-B^1)r^2(0)\|$$

$$\begin{aligned} &\le \epsilon^2 C(T)|\delta| + CK\epsilon^2\|\delta| \\ &\le \epsilon^2 C_2(T)|\delta| \end{aligned}$$

as required for (iv).

Again we derive a related estimate used in the proof of Corollary 5.6. From (5.29) an argument similar to that yielding (5.30) gives

$$\|\rho(t) - \rho(0)\| \le \|B(0,t;\phi^1) - I\|\|\rho(0)\| + \int_0^t C(T)\epsilon[\|\rho(s)\| + \epsilon|\gamma(s)|]\,\mathrm{d}s.$$

Thus (5.24), (5.30), (5.31) and (5.32) yield the existence of $C_4 > 0$:

$$\|\rho(t)\| \le (1 + C_4 t)\epsilon|\delta|, \tag{5.33}$$

for t and ϵ sufficiently small. Hence, by (5.31) and (5.32), we have from (5.26)

$$\begin{aligned} |\delta| &\le |\gamma(t)| + C_1 t\|\rho\|_\infty + \int_0^t C_1\epsilon|\gamma(s)|\,\mathrm{d}s \\ &\le |\gamma(t)| + C_4 t\epsilon|\delta|. \end{aligned}$$

Thus

$$|\delta| \le \frac{|\gamma(t)|}{1 - C_4 t\epsilon}.$$

Hence (5.33) gives us

$$\|\rho(t)\| \le (1 + C_3 t)\epsilon|\gamma(t)| \tag{5.34}$$

for t sufficiently small. (Note that we have chosen the same constant C_3 as appears in (5.22) without loss of generality.)

(v), (vi) Consider the equations

$$\begin{aligned} r_t^i &= A(\theta^i)r^i + g(r^i,\theta^i), \quad r^i(0) = \Phi^i(\phi) \\ \theta_t^i &= 1 + h(r^i,\theta^i), \theta^i(0) = \phi \end{aligned} \tag{5.35}$$

for $i = 1, 2$. Define ρ, γ and $\|\rho\|_\infty$ as in the proof of (iii) and (iv). Note that

$$\|\rho(0)\| \le \|\Phi^1 - \Phi^2\|, \quad \gamma(0) = 0.$$

As in cases (iii) and (iv), (5.25) holds and, since $\gamma(0) = 0$, it follows from (5.28) that

$$|\gamma(t)| \le C(T)\|\rho\|_\infty \quad \forall t \in [0, T]. \tag{5.36}$$

Then (5.29), together with Assumption 5.1, gives

$$\begin{aligned} \|\rho(t)\| &\le C_1\|\Phi^1 - \Phi^2\| + \int_0^t C(T)\epsilon[\|\rho(s)\| + \epsilon|\gamma(s)|]\,\mathrm{d}s \\ &\le C_1\|\Phi^1 - \Phi^2\| + \int_0^t C(T)\epsilon\|\rho\|_\infty\,\mathrm{d}s; \end{aligned}$$

hence, for ϵ sufficiently small,

$$\|\rho\|_\infty \leq C_2(T)\|\Phi^1 - \Phi^2\| \quad \forall t \in [0, T].$$

Application to (5.36) gives (v) as required; application to (5.29) gives

$$\|\rho(T) - B\rho(0)\| \leq \int_0^T C(T)\epsilon[\|\rho(s)\| + \epsilon|\gamma(s)|]\,\mathrm{d}s \leq \epsilon C_2(T)\|\Phi^1 - \Phi^2\|$$

as required for (vi). □

Armed with the bounds on G and H proved in Lemma 5.2 we can proceed to show that the map $\mathcal{T}$ given in (5.18) is well defined. Note that we have already assumed ϵ sufficiently small in the proof of Lemma 5.2. Unlike the case of unstable manifolds considered in Section 3 we will not detail the bounds on ϵ required in the sequel – we simply observe that our arguments hold for ϵ sufficiently small. Without loss of generality we will use the same constant ϵ^* to denote the upper bound on ϵ sufficient for all our arguments to work.

Lemma 5.3 Let Assumption 5.1 hold. If $\Phi \in \Gamma$ then there exists $\epsilon^* > 0$, such that if $\epsilon \in (0, \epsilon^*], \mathcal{T}\Phi$ given by (5.18) is well defined.

Proof. We show that, for every $\theta \in \mathbb{R}$, there is a unique $\phi \in \mathbb{R}$ so that (5.18) is well defined. That is we solve the equation

$$\theta = \phi + T + H(\Phi(\phi), \phi)$$

for ϕ, given $\theta \in \mathbb{R}$ and $\Phi \in \Gamma$. To do this we use the contraction mapping theorem. Consider the iterates

$$\begin{aligned}\phi^{k+1} &= \theta - T - H(\Phi(\phi^k), \phi^k),\\ \psi^{k+1} &= \theta - T - H(\Phi(\psi^k), \psi^k).\end{aligned}$$

Clearly $\phi^k \in \mathbb{R}$ yields $\phi^{k+1} \in \mathbb{R}$ and furthermore, by Lemma 5.2(iii),

$$|\phi^{k+1} - \psi^{k+1}| \leq C_2\epsilon|\phi^k - \psi^k| \leq \tfrac{1}{2}|\phi^k - \psi^k|$$

for ϵ sufficiently small; thus the mapping is a contraction on $\mathbb{R}$ and the existence of ϕ given θ follows. □

We may now show that $\mathcal{T}$ maps Γ to itself. We assume throughout the remainder of this section that K is given by (5.16).

Lemma 5.4 Let Assumption 5.1 hold. Then there exists ϵ^* such that, if $\epsilon \in (0, \epsilon^*]$ then the mapping $\mathcal{T} : \Gamma \mapsto \Gamma$.

Proof. From (5.18), (5.6) and Lemma 5.2(ii) we obtain, for ϵ sufficiently small and $\Phi \in \Gamma$,

$$\|(\mathcal{T}\Phi)(\theta)\| \leq \alpha K\epsilon^2 + 2C_1T\epsilon^2.$$

Noting that K is given by (5.16) and that $2C_1 < 4C_1$ we deduce that

$$\|(\mathcal{T}\Phi)(\theta)\| \leq \frac{4\alpha C_1 T\epsilon^2}{1-\alpha} + 4C_1 T\epsilon^2 \leq \frac{4C_1 T\epsilon^2}{1-\alpha} = K\epsilon^2.$$

(This argument has been constructed to be robust under an increase of C_1 by a factor of 2.)

By Lemma 5.2(iv) it follows that

$$\begin{aligned}\|(\mathcal{T}\Phi)(\theta^1) - (\mathcal{T}\Phi)(\theta^2)\| &\leq \alpha\|\Phi(\phi^1) - \Phi(\phi^2)\| + \epsilon^2 C_2|\phi^1 - \phi^2| \\ &\leq [\alpha + \epsilon C_2]\epsilon|\phi^1 - \phi^2|.\end{aligned}$$

By Lemma 5.2(iii) and (5.18) it follows that

$$|\theta^1 - \theta^2| \geq |\phi^1 - \phi^2| - C_2\epsilon|\phi^1 - \phi^2|.$$

Thus, for ϵ sufficiently small,

$$|\phi^1 - \phi^2| \leq \frac{|\theta^1 - \theta^2|}{1 - C_2\epsilon}.$$

Combining we find that

$$\|(\mathcal{T}\Phi)(\theta^1) - (\mathcal{T}\Phi)(\theta^2)\| \leq \frac{[\alpha + \epsilon C_2]\epsilon}{1 - \epsilon C_2}|\theta^1 - \theta^2|.$$

We deduce that, for $\epsilon : 2\epsilon C_2 \leq 1 - \alpha$,

$$\|(\mathcal{T}\Phi)(\theta^1) - (\mathcal{T}\Phi)(\theta^2)\| \leq \epsilon|\theta^1 - \theta^2|$$

as required.

It remains to establish that $(\mathcal{T}\Phi)(\theta)$ is periodic. If we set $\phi \mapsto \phi + T$ in (5.18) we obtain

$$\begin{aligned}(\mathcal{T}\Phi)(\psi) &= B\Phi(\phi + T) + G(\Phi(\phi + T), \phi + T), \\ \psi &= \phi + 2T + H(\Phi(\phi + T), \phi + T).\end{aligned}$$

Since A, g, h and Φ are T-periodic in θ it follows that $\psi = \theta + T$ and $(\mathcal{T}\Phi)(\psi) = (\mathcal{T}\Psi)(\theta)$ and thus periodicity of $(\mathcal{T}\Phi)(\theta)$ follows. □

Finally we may prove:

Theorem 5.5 Let Assumption 5.1 hold. Then there exists $\epsilon^* > 0$ such that, for all $\epsilon \in (0, \epsilon^*]$, equations (5.13) have an invariant manifold $\Phi \in \Gamma$ satisfying (5.17).

Proof. To show the existence of the invariant manifold for (5.13) it is sufficient, by Lemma 5.4, to show that $\mathcal{T} : \Gamma \mapsto \Gamma$ is a contraction. Consider the equations

$$\begin{aligned}(\mathcal{T}\Phi^i)(\theta) &= B\Phi^i(\phi^i) + G(\Phi^i(\phi^i), \phi^i), \\ \theta &= \phi^i + T + H(\Phi^i(\phi^i), \phi^i)\end{aligned} \tag{5.37}$$

for $i = 1, 2$. It follows that

$$\|(\mathcal{T}\Phi^1)(\theta) - (\mathcal{T}\Phi^2)(\theta)\| \\ \leq \|B\Phi^1(\phi^1) - B\Phi^2(\phi^2)\| + \|G(\Phi^1(\phi^1), \phi^1) - G(\Phi^2(\phi^2), \phi^2)\|. \quad (5.38)$$

Hence

$$\|(\mathcal{T}\Phi^1)(\theta) - (\mathcal{T}\Phi^2)(\theta)\| \leq \|B\Phi^1(\phi^1) - B\Phi^2(\phi^1)\| + \|B\Phi^2(\phi^1) - B\Phi^2(\phi^2)\|$$

$$+\|G(\Phi^1(\phi^1), \phi^1) - G(\Phi^2(\phi^1), \phi^1)\| + \|G(\Phi^2(\phi^1), \phi^1) - G(\Phi^2(\phi^2), \phi^2)\|.$$

Using Assumption 5.1 and Lemma 5.2, it follows that

$$\|(\mathcal{T}\Phi^1)(\theta) - (\mathcal{T}\Phi^2)(\theta)\| \leq (\alpha + C_2\epsilon)\|\Phi^1 - \Phi^2\|_P + (\alpha + C_2\epsilon)\epsilon|\phi^1 - \phi^2|.$$

Also, by similar manipulations, it follows that

$$|\phi^1 - \phi^2| \leq C_2\|\Phi^1 - \Phi^2\|_P + C_2\epsilon|\phi^1 - \phi^2|.$$

Combining these two estimates we obtain

$$\|(\mathcal{T}\Phi^1)(\theta) - (\mathcal{T}\Phi^2)(\theta)\| \leq (\alpha + C_2\epsilon)\|\Phi^1 - \Phi^2\|_P + \frac{(\alpha + C_2\epsilon)}{(1 - C_2\epsilon)}\epsilon C_2\|\Phi^1 - \Phi^2\|_P$$

Again, choosing ϵ sufficiently small, we obtain

$$\|(\mathcal{T}\Phi^1)(\theta) - (\mathcal{T}\Phi^2)(\theta)\| \leq \frac{(}{1} + \alpha)\|\Phi^1 - \Phi^2\|_P.$$

Since this holds for all θ it follows that the mapping $\mathcal{T}$ is contractive on Γ with constant $(1 + \alpha)/2 < 1$ and existence of an invariant manifold follows. □

Given ϵ^* and Φ from Theorem 5.5, it is a corollary that equations (1.1) have a periodic solution:

Corollary 5.6 (**Periodic solutions**) Assume that there exists a C^3 diffeomorphism $\chi : \mathbb{R}^p \mapsto \mathbb{R}^p$ which renders (1.1) in the form (5.2) under $u = \chi(v)$; assume further that there exists $\epsilon \in (0, \epsilon^*/2]$ such that A, g and h satisfy (5.4) and (5.9) for all $r \in B(0, \epsilon)$, all $\theta \in \mathbb{R}$ and all $t \in \mathbb{R}$. Then (1.1) has a periodic solution $\bar{u}(t)$ comprising the sets of points

$$\{u \in \mathbb{R}^p | u = \chi(v),\ v = (\Phi(\theta)^T, \theta)^T : \theta \in \mathbb{R},\ \Phi \in \Gamma\}.$$

Proof. By virtue of the coordinate tranformation, it is sufficient to show that the invariant manifold Φ constructed in Thoerem 5.5 yields a periodic solution of (5.2). To do this we introduce some notation. Let $\bar{S}(t)$ denote the semigroup for $v(t) = (r(t)^T, \theta(t))^T$ given by the solution of (5.2). Thus $v(t) = \bar{S}(t)(\xi^T, \phi)^T$. Let

$$\mathcal{M} := \{(r^T, \theta)^T \in \mathbb{R}^p : r = \Phi(\theta)\}.$$

Now consider the set

$$\mathcal{M}(t) := \bar{S}(t)\mathcal{M}.$$

Thus $\mathcal{M}(t)$ is obtained by taking every point in $\mathcal{M}$ and evolving it forward t time units under equation (5.2), that is under $\bar{S}(t)$. It is our aim to show that $\mathcal{M}(t) \equiv \mathcal{M}$ for all t sufficiently small; then we deduce that the closed curve given by the invariant manifold Φ of Theorem 5.5, which is invariant under the time T evolution of the differential equation (5.2), is actually invariant under the evolution of (5.2) for any time t sufficiently small. This shows that $\mathcal{M}$ is a periodic solution, since it is a closed curve.

Note that, from (5.12) and (5.14), every point $(r^T, \theta)^T \in \mathcal{M}(t)$ satisfies

$$\begin{aligned} r &= B(0,t;\phi)\Phi(\phi) + G(\Phi(\phi),\phi;t), \\ \theta &= \phi + t + H(\Phi(\phi),\phi;t) \end{aligned} \tag{5.39}$$

for some $\phi \in \mathbb{R}$. Analogously to the proof of Lemma 5.3, we may show that there is a graphical relationship between r and θ so that $r = \Psi(\theta)$ for some $\Psi \in C(\mathbb{R}, \mathbb{R}^{p-1})$. Furthermore, use of (5.22) and (5.34) show that there exists $C_3 > 0$ such that

$$\|r\| = \|\Psi(\theta)\| \leq (1 + C_3 t)K\epsilon^2$$

and

$$\|\Psi(\theta_1) - \Psi(\theta_2)\| \leq (1 + C_3 t)\epsilon|\theta_1 - \theta_2|.$$

Hence, if t is chosen sufficiently small that $(1 + C_3 t) \leq 2$ and ϵ sufficiently small that $\epsilon \in (0, \epsilon^*/2]$, then we deduce that $\Psi \in \Gamma$.

Finally, observe that

$$\bar{S}(T)\mathcal{M}(t) = \bar{S}(T)\bar{S}(t)\mathcal{M} = \bar{S}(t)\bar{S}(T)\mathcal{M}.$$

But, $\mathcal{M}$ is invariant under $\bar{S}(T)$ by construction and hence we have

$$\bar{S}(T)\mathcal{M}(t) = \bar{S}(t)\mathcal{M} = \mathcal{M}(t).$$

Hence $\mathcal{M}(t)$ is invaraint under $\bar{S}(T)$. Since all points in $\mathcal{M}(t)$ may be represented by means of a graph Ψ lying in Γ the uniqueness implied by Theorem 5.5 gives $\Psi = \Phi$. Hence $\mathcal{M}(t) = \mathcal{M}$ and the proof is complete. □

5.2. Periodic solutions

We now modify the analysis of Section 1 to prove the existence and convergence of a closed invariant curve for the numerical method (1.2) which lies close to the periodic solution $\bar{u}(t)$ of (1.1). Let $m\Delta t = T$ and

$$U_m = S^m_{\Delta t}U, \quad u(T) = S(T)U$$

denote the solutions of (1.2), (1.1) respectively subject to the same initial condition U. We know, from Theorem 3.8, that there exists $C_5 = C_5(T)$ such that

$$\begin{aligned} \|S^m_{\Delta t|\Delta t}U - S(T)U\| &\leq C_5\Delta t^r, \quad \forall U \in \mathbb{R}^p, \\ \|\mathrm{d}S^m_{\Delta t}(U) - \mathrm{d}S(U;T)\| &\leq C_5\Delta t^r, \quad \forall U \in \mathbb{R}^p. \end{aligned} \tag{5.40}$$

We define $v(t)$ and V_n by $u(t) = \chi(v(t))$ and $U_n = \chi(V_n)$ where χ is the C^3 diffeomorphism given in Corollary 5.6. We also set

$$V = (\xi^T, \phi)^T \in \mathbb{R}^p, \quad V_m = (R^T, \Theta)^T \in \mathbb{R}^p$$

We may define semigroups appropriate for the variables $v(t)$ and V_n from those in the original variables by

$$\begin{aligned} \bar{S}(t)v &:= \chi^{-1}(S(t)\chi(v)), \\ \bar{S}^n_{\Delta t}v &:= \chi^{-1}(S^n_{\Delta t}\chi(v)). \end{aligned} \tag{5.41}$$

Then, since χ is a C^3 diffeomorphism it follows from (5.40) and (5.41) that there exists $C_6 = C_6(T)$ such that

$$\begin{aligned} \|\bar{S}^m_{\Delta t}V - \bar{S}(T)V\| &\le C_6\Delta t^r, \quad \forall V \in \mathbb{R}^p : \xi \in B(0,\epsilon) \\ \|\mathrm{d}\bar{S}^m_{\Delta t}(V) - \mathrm{d}\bar{S}(V;T)\| &\le C_6\Delta t^r, \quad \forall V \in \mathbb{R}^p : \xi \in B(0,\epsilon). \end{aligned} \tag{5.42}$$

(Here the derivatives of the semigroups $\bar{S}^m_{\Delta t}$ and $\bar{S}(t)$ are defined analogously to those for $S^m_{\Delta t}$ and $S(t)$.)

We now exploit the existence theory derived for the periodic solutions of (5.2) to study the numerical method. From (5.13) and (5.42) we obtain

$$\begin{aligned} R &= B\xi + \tilde{G}(\xi,\phi), \\ \Theta &= \phi + T + \tilde{H}(\xi,\phi), \end{aligned} \tag{5.43}$$

where

$$\begin{aligned} \|\tilde{G}(\xi,\phi) - G(\xi,\phi)\| &\le C_6\Delta t^r, \\ \|\tilde{H}(\xi,\phi) - H(\xi,\phi)\| &\le C_6\Delta t^r, \\ \|\mathrm{d}\tilde{G}(\xi,\phi) - \mathrm{d}G(\xi,\phi)\| &\le C_6\Delta t^r, \\ \|\mathrm{d}\tilde{H}(\xi,\phi) - \mathrm{d}H(\xi,\phi)\| &\le C_6\Delta t^r. \end{aligned} \tag{5.44}$$

Here $\mathrm{d}G$, $\mathrm{d}\tilde{G}$, $\mathrm{d}H$ and $\mathrm{d}\tilde{H}$ denote the Jacobians of G and $\tilde{G}$ with respect to V.

For simplicity we assume that (5.44) holds for all $v \in \mathbb{R}^p$; since all the analysis takes place within the ball $\xi \in B(0,\epsilon)$ the results hold for any numerical approximation satisfying (5.42). Using (5.44), together with Lemma 5.2, it follows that for Δt sufficiently small, $\tilde{G}$ and $\tilde{H}$ satisfy estimates analogous to those appearing in Lemma 5.2 for G and H but with $C_1 \mapsto 2C_1$ and $C_2 \mapsto 2C_2$. Hence, following the proofs of Lemmas 5.3, 5.4 and Theorem 5.5 with $C_1 \mapsto 2C_1$ and $C_2 \mapsto 2C_2$, we can prove the existence of an invariant manifold $\Phi_{\Delta t} \in \Gamma$ for (5.43). Specifically we seek a $\Phi_{\Delta t}$ with the property that

$$\xi = \Phi_{\Delta t}(\phi) \Leftrightarrow R = \Phi_{\Delta t}(\Theta). \tag{5.45}$$

Comparison with (5.39) shows that to do this is equivalent to finding a fixed point of the mapping $\mathcal{T}_{\Delta t}$ defined by

$$\begin{aligned}(\mathcal{T}_{\Delta t}\Phi)(\Theta) &= B\Phi(\phi) + \tilde{G}(\Phi(\phi), \phi), \\ \Theta &= \phi + T + \tilde{H}(\Phi(\phi), \phi).\end{aligned} \tag{5.46}$$

Using this formulation we can prove the following theorem:

Theorem 5.7 Let Assumption 5.1 hold. Then there exists $\Delta t_c > 0$ and $\epsilon^{**} > 0$ such that, for all $\Delta t \in (0, \Delta t_c]$ and $\epsilon \in (0, \epsilon^{**}]$ equations (5.43) have an invariant manifold $\Phi_{\Delta t} \in \Gamma$, satisfying (5.45). Furthermore, $\Phi_{\Delta t}$ is close to the graph Φ constructed in Theorem 5.5 in the sense that there exists $C_7 = C_7(T)$ such that

$$\|\Phi - \Phi_{\Delta t}\| \leq C_7 \Delta t^r.$$

Proof. The existence of such a manifold follows precisely as in the proof of Theorem 5.5, except that C_2 is enlarged by a factor of 2. Hence further reduction of ϵ is necessary in the proof.

To prove closeness of the manifolds we again use, essentially, the uniform contraction principle; note that both Φ and $\Phi_{\Delta t}$ lie in Γ. Thus, using the contractivity of $\mathcal{T}_{\Delta t}$ we obtain

$$\begin{aligned}\|\Phi(\theta) - \Phi_{\Delta t}(\theta)\| &= \|\mathcal{T}\Phi(\theta) - \mathcal{T}_{\Delta t}\Phi_{\Delta t}(\theta)\| \\ &\leq \|\mathcal{T}\Phi(\theta) - \mathcal{T}\Phi_{\Delta t}(\theta)\| + \|\mathcal{T}\Phi_{\Delta t}(\theta) - \mathcal{T}_{\Delta t}\Phi_{\Delta t}(\theta)\| \\ &\leq \tfrac{1}{2}(1+\alpha)\|\Phi - \Phi_{\Delta t}\|_P + \|\mathcal{T}\Phi_{\Delta t}(\theta) - \mathcal{T}_{\Delta t}\Phi_{\Delta t}(\theta)\|.\end{aligned}$$

Now, from (5.18), (5.46) and (5.44) and the Lipschitz properties of $\mathcal{T}_{\Delta t}\Phi \in \Gamma$, we deduce that

$$E := \|\mathcal{T}\Phi_{\Delta t}(\theta) - \mathcal{T}_{\Delta t}\Phi_{\Delta t}(\theta)\|$$

satisfies

$$E \leq \|\mathcal{T}\Phi_{\Delta t}(\theta) - \mathcal{T}_{\Delta t}\Phi_{\Delta t}(\Theta)\| + \|\mathcal{T}_{\Delta t}\Phi_{\Delta t}(\theta) - \mathcal{T}_{\Delta t}\Phi_{\Delta t}(\Theta)\|.$$

Hence, since $\mathcal{T}_{\Delta t}\Phi_{\Delta t} \in \Gamma$, we have from (5.18), (5.43) and (5.44) that

$$E \leq C_6(1+\epsilon)\Delta t^r.$$

Thus, putting these estimates together we obtain

$$\|\Phi - \Phi_{\Delta t}\|_P \leq \frac{2C_6(1+\epsilon)}{1-\alpha}\Delta t^r.$$

This completes the proof. □

As a corollary of Theorem 5.7 we consider the approximation of periodic solutions in (1.1) by closed invariant curves in (1.2). Recall the notions (3.1), (3.2) of distance.

Corollary 5.8 (Periodic solutions under approximation) Assume that (1.1) has a hyperbolic, stable periodic solution $\bar{u}(t)$ comprising the set of points $\mathcal{P}$. Then (1.2) has closed invariant curve comprising the set of points $\mathcal{P}_{\Delta t}$ and, furthermore, there exists a constant $C > 0$:

$$d_{\mathrm{H}}(\mathcal{P}_{\Delta t}, \mathcal{P}) \leq C\Delta t^r.$$

Proof. If (1.1) has a hyperbolic, stable periodic solution then a transformation χ exists rendering (1.1) in the form to which Corollary 5.6 applies. Application of Theorem 5.7 yields the required closeness result. □

5.3. Bibliography

The standard local construction of periodic solutions uses the Hopf bifurcation which facilitates the construction of a periodic solution branching from an equilibrium solution; see Guckenheimer and Holmes (1983) and Drazin (1992). For background material describing global questions concerning the existence of periodic solutions in (1.1) see, for example, Hale (1969), Hartman (1982) and Guckenheimer and Holmes (1983). In particular, Hale (1969) presents the full details of the coordinate transformation $u = \chi(v)$ which is central to the analysis described here.

The first article to study the effect of numerical approximation on periodic solutions in a general context was Braun and Hershenov (1977). They studied stable hyperbolic periodic solutions and employed the time T map to perform the analysis (where T is the period). The next article to address such questions was Brezzi *et al.* (1984) where the existence of Hopf bifurcation points, and the resulting closed invariant curves close to the Hopf point, was studied for numerical methods (1.2) operating in a parameter regime close to that in which a Hopf bifurcation occurs in (1.1). The work of Braun and Hershenov (1977) was generalized in Doan (1985) to encompass multistep methods and the general case of hyperbolic periodic solutions which are not necessarily stable. Neither of the articles (Braun and Hershenov, 1977) nor (Doan, 1985) obtained precise orders of convergence for the approximate invariant curve. A little later Beyn (1987a) generalized the work of Braun and Hershenov (1977) to encompass arbitrary hyperbolic periodic solutions, obtaining precise orders of convergence; his approach is similar to that in Braun and Hershenov (1977) but, rather than employing the time T map he uses the time Δt map of the differential equation. The whole subject area was put in a very clear setting in Eirola (1988; 1989) where results similar to those of Beyn derived but in the stronger C^k topology, the value of k depending on the smoothness of the vector field $f(\cdot)$ in (1.1). In turn the work of Eirola can be viewed in a very general setting concerning the stability of invariant circles of mappings in $\mathbb{R}^p$; see Pugh and Shub (1988). The proof given in this article is closely related to the proofs in Braun and Hershenov (1977) and Beyn (1987a); it is far from optimal in the sense that only stable periodic solutions are considered and the result is in the C^0 topology. However, this presentation has been chosen because it is self-contained, relatively simple and is similar to the method of analysis used to construct unstable manifolds in Section 4. A recent article (Alouges and Debussche, 1993) is concerned with extensions of the work referenced here to partial differential equations. In this context the work of Titi (1991) is also of interest.

We have not described here an analogous theory for the behaviour of invariant tori under numerical approximation. Such a theory has recently been developed in Lorenz (1994) using the approach of Fenichel (1971) to determine an appropriate coordinate system analogous to the coordinate system described in Hale (1969) used to study periodic solutions.

6. Uniform asymptotic stability and attractors

6.1. Background theory

In this section we consider the effect of numerical approximation on general objects which are attracting, in certain senses to be made precise, for solutions of (1.1). In most cases we do not make specific statements about the nature of the dynamics within the attracting object so that our framework will be sufficiently general to include, for example, the strange attractors observed in the Lorenz equations (2.10). Thus our assumptions will be the existence of an arbitrary compact set possessing some form of attractivity. Up to this point we have considered the numerical approximation of equilibrium points, invariant sets in the neighbourhood of equilibrium points and periodic solutions. In all cases our methodology has been the same: we have employed the contraction mapping theorem to develop an existence theory for the object in question and then used the uniform contraction principle to incorporate the effect of numerical approximation. In this section a different approach will be necessary since there is no known existence theory based on the contraction mapping theorem which we can exploit.

We start by defining the type of objects of interest to us, together with certain of their properties which we require. In this section, all of our definitions and theorems concern the continuous semigroup $S(t)$. At the end of the section we detail which definitions and theorems can be generalized from $S(t)$ to $S^n_{\Delta t}$.

Definition 6.1 A set A *attracts* a set B under $S(t)$ if, for any $\epsilon > 0$, there exists $t^* = t^*(\epsilon, B, A)$ such that $S(t)B \subset \mathcal{N}(A, \epsilon)\ \forall t \geq t^*$. A compact invariant set A is said to be a an *attractor* if A attracts an open neighbourhood of itself. A *global attractor* is an attractor which attracts every bounded set in $\mathbb{R}^p$.

Example Define the set $\mathcal{A} := (x, y) \in [-1, 1] \times \{0\}$ and note that $\mathcal{A}$ is clearly compact; by arguments similar to those used to establish the example following Definition 3.1, we deduce that $\mathcal{A}$ is invariant under the differential equations

$$x_t = x - x^3, \quad x(0) = x_0$$

$$y_t = \lambda y, \quad y(0) = y_0.$$

Denote the semigroup for the x-equation by $S_x(t)$ and for the y-equation by $S_y(t)$. By (3.8), the solution of these equations has the property that $S_x(t)[-a,a] \to [-1,1]$ as $t \to \infty$, for any $a > 0$; furthermore, $S_x(t)0 = 0$. Clearly $S_y(t)y_0 = e^{\lambda t}y_0$.

It follows that, for any $\lambda \geq 0$, $\mathcal{A}$ attracts any set $(x,y) \in [-a,a] \times \{0\}, a \geq 0$. Furthermore, if $\lambda < 0$ then $\mathcal{A}$ is an attractor – since then $[-a,a] \times [-\epsilon,\epsilon]$ is attracted to $\mathcal{A}$ for any $a, \epsilon \geq 0$. Indeed, since $a, \epsilon \geq 0$ are arbitrary, it is a global attractor. □

Attractors are often constructed by applying the following theorem.

Theorem 6.2 Assume that $B \subset \mathbb{R}^p$ is a bounded open set such that $S(t)\bar{B} \subset B \; \forall t > 0$. Then $\omega(B)$ is an attractor which attracts B.

Proof. Since $S(t)\bar{B} \subset B$ it follows that

$$\omega(B) = \bigcap_{s \geq 0} \overline{\bigcup_{t \geq s} S(t)B} \subset \bigcap_{s \geq 0} \overline{\bigcup_{t \geq s} B} = \bar{B}. \tag{6.1}$$

Thus $\omega(B)$ is bounded. Furthermore, $\omega(B)$ is closed and invariant by Theorem 3.3 and it follows that $\omega(B)$ is a compact invariant set.

We now show that $\mathcal{A} := \omega(B)$ attracts B. Assume that it does not. Then, for all sufficiently small $\epsilon > 0$, there does not exist t^* such that $S(t)B \subset \mathcal{N}(\mathcal{A},\epsilon) \; \forall \; t \geq t^*$. Thus there exists $x_k \in B$ and $t_k \to \infty$ such that $S(t_k)x_k \not\subset \mathcal{N}(\mathcal{A},\epsilon)$. But $S(t_k)x_k$ is a bounded sequence contained in B and hence has a convergent subsequence $S(t_{k_i})x_{k_i} \to y \in \overline{B}$. By Definition 3.2 $y \in \mathcal{A}$ and this is a contradiction.

Note that $\omega(B) \subset \bar{B}$ by (6.1). We show that, in fact, $\omega(B) \subset B$. Assume for the purposes of contradiction that $\exists y \in \omega(B) \cap \partial B$ (where $\partial B = \bar{B} \backslash B$). Since $\omega(B)$ is invariant it follows that, for any $t > 0 \; \exists x \in \omega(B) : S(t)x = y$. But, since $\omega(B) \subset \bar{B}$ we have $x \in \bar{B}$ and hence, by assumption, $y = S(t)x \in B$. This is a contradiction and thus no such y exists. Thus $\omega(B) \subset B$.

Now, since $\omega(B) \subset B$ is closed it follows that, for ϵ sufficiently small, $\mathcal{N}(\omega(B),\epsilon) \subset B$. Since $\omega(B)$ attracts B it follows that $\omega(B)$ attracts an open neighbourhood of itself and the proof is complete. □

Example Consider the equation $u_t = u - u^3$. Let $B = (-a,a), a > 1$. Then $\omega(B) = [-1,1]$ as shown in the second example following Definition 3.2; furthermore, $\omega(B)$ is an attractor: on $|u| = a$ we have $\mathrm{d}|u|^2/\mathrm{d}t < 0$ and hence $S(t)\bar{B} \subset B \; \forall t > 0$. Theorem 6.2 gives the desried result. □

A second attracting object of interest to us is now defined:

Definition 6.3 A compact set Λ is *uniformly stable* if for each $\epsilon > 0$ $\exists \delta = \delta(\epsilon) > 0$ such that $\text{dist}(U, \Lambda) < \delta \Rightarrow \text{dist}(S(t)U, \Lambda) < \epsilon$ $\forall t \geq 0$; a compact set Λ is *asymptotically stable* if there exists $\delta_0 > 0$ and for each ϵ, a $T = T(\epsilon)$ such that $\text{dist}(U, \Lambda) < \delta_0$ implies that $\text{dist}(S(t)U, \Lambda) < \epsilon$ $\forall t \geq T$. A compact set Λ is *uniformly asymptotically stable* if it is uniformly stable and asymptotically stable.

Theorem 6.4 The following properties hold for uniformly asymptotically stable sets:

(i) uniformly stable sets are positively invariant;

(ii) an attractor is uniformly asymptotically stable;

(iii) if Λ is uniformly asymptotically stable then $\mathcal{A} = \omega(\Lambda) \subseteq \Lambda$ is an attractor.

Proof. We first prove (i). For contradiction assume that Λ is uniformly stable and not positively invariant. Then $\exists \tau > 0, \epsilon > 0$ and $U \in \Lambda$ such that $\text{dist}(S(\tau)U, \Lambda) > \epsilon$. But, since the set is uniformly stable it follows that $\exists \delta = \delta(\epsilon)$ such that $\text{dist}(U, \Lambda) < \delta \Rightarrow \text{dist}(S(\tau)U, \Lambda) < \epsilon$. Since $\text{dist}(U, \Lambda) = 0$ this gives the required contradiction.

Now consider (ii). It is automatic that an attractor is asymptotically stable since it attracts a neighbourhood of itself. Thus it suffices to show uniform stability. Assume for contradiction that $\mathcal{A}$ is an attractor, attracting a neighbourhood W, but it is not uniformly stable. Thus, for any $\epsilon > 0$ there exists a sequence of times $\{t_j\}_{j=1}^{\infty}$ and a sequence $\{x_j\}_{j=1}^{\infty}$ with $x_j \in W$ for each j and $x_j \to x \in \mathcal{A}$, such that $S(t)x_j \in \mathcal{N}(A, \epsilon), t \in [0, t_j)$ and $S(t_j)x_j \notin \mathcal{N}(\mathcal{A}, \epsilon)$. Now let $H = \{x, \{x_j\}_{j=1}^{\infty}\}$ and note that, since $x_j \in W$ (a bounded set) and since $x_j \to x$, we have that $H \subset W$ is compact. Since $\mathcal{A}$ attracts W it follows that $\mathcal{A}$ attracts H and hence that $\omega(H) \subset \mathcal{A}$. We deduce that, for any $T > 0$, the sequence $S(t_j - T)x_j \to z \in \omega(H) \subset \mathcal{A}$. But $S(T)S(t_j - T)x_j = S(t_j)x_j \to S(T)z$. By the invariance of ω-limit sets (see Theorem 3.3) it follows that $S(T)z \in \mathcal{A}$. However, $S(T)z \notin \mathcal{N}(\mathcal{A}, \epsilon)$ since $S(t_j)x_j \notin \mathcal{N}(\mathcal{A}, \epsilon)$ and this gives the required contradiction.

Finally we prove (iii). Since Λ is postively invariant it follows that $\mathcal{A} = \omega(\Lambda) \subseteq \Lambda$ and hence $\mathcal{A}$ is bounded; $\mathcal{A}$ is closed by Theorem 3.8 and hence compact. Let $H = \mathcal{N}(\Lambda, \delta_0)$ where δ_0 is given by the definition of asymptotic stability. To show that $\mathcal{A}$ attracts H it is sufficient to show that $\omega(H) \subseteq \mathcal{A}$. Clearly $\omega(H) \subseteq \Lambda$ since Λ is asymptotically stable. Let $y \in \omega(H)$. Since $\omega(H)$ is invariant (by Theorem 3.8) it follows that, for each $t_k > 0$ there exists $x_k \in \omega(H) \subseteq \Lambda$ such that $S(t_k)x_k = y$. Thus the sequence $S(t_k)x_k \to y$ as $k \to \infty$. But, since $x_k \in \Lambda$ it follows that $y \in \omega(\Lambda) = \mathcal{A}$ and the result follows. □

Thus the important fact distinguishing attractors and uniformly asymptotically stable sets is that the former are necessarily invariant whilst the latter need only be positively invariant. The following example illustrates this.

Examples Consider the equation $u_t = -u$, $u(0) = U$. Any interval $[-a, a], a \geq 0$ is a uniformly asymptotically stable set. However, only the point 0 is an attractor since $[-a, a]$ is not invariant for $a > 0$.

It is well known that the existence of stable equilibrium points can be deduced from construction of appropriate Lyapunov functions. Converse results are also available and the following result, deducing the existence of a Lyapunov function from the uniform asymptotic stability of a compact set, will be extremely useful.

Theorem 6.5 Given a compact uniformly asymptotically stable set Λ there exists $R_0 > 0$ and a Lyapunov function $V : \mathcal{N}(\Lambda : R_0) \to \mathbb{R}^+$ satisfying the following three properties:

(i) there exists $L > 0 : |V(x) - V(y)| \leq L\|x - y\| \quad \forall x, y \in \mathcal{N}(\Lambda, R_0)$;
(ii) there exist continuous, strictly increasing functions α, β: $[0, R_0) \to \mathbb{R}^+$ with $\alpha(0) = \beta(0) = 0$ and $\alpha(r) < \beta(r)$ for $r > 0$ such that

$$\alpha(\mathrm{dist}(x, \Lambda)) \leq V(x) \leq \beta(\mathrm{dist}(x, \Lambda));$$

(iii) there exists a constant $C > 0$ such that

$$V(S(t)U) \leq \mathrm{e}^{-Ct} V(U), \; 0 \leq t \leq T$$

provided $S(t)U \in \mathcal{N}(\Lambda, R_0), \; 0 \leq t \leq T$.

Proof. This theorem is proved in Theorem 22.5 of Yoshizawa (1966) with a slightly different conclusion in part (iii); that our point (iii) holds may be deduced from Theorem 4.1 of Yoshizawa (1966). □

Corollary 6.6 Let $\beta(R_1) = \alpha(R_0)$ and assume that $\mathrm{dist}(U, \Lambda) < R_1$. Then

$$S(t)U \in \mathcal{N}(\Lambda, R_0) \quad \text{and} \quad V(S(t)U) \leq \mathrm{e}^{-Ct} V(U) \quad \forall t \in [0, \infty). \tag{6.2}$$

Proof. For the purpose of contraction assume that there exists $T > 0$:

$$\mathrm{dist}(S(T)U, \Lambda) < R_0, \quad t \in [0, T) \quad \text{and} \quad \mathrm{dist}(S(T))U, \Lambda) = R_0.$$

By Theorem 6.5(iii) it follows that $V(S(t)U) \leq \mathrm{e}^{-Ct} V(U), t \in [0, T]$. From Theorem 6.5(ii) we have that

$$\begin{aligned} \alpha(R_0) &= \alpha(\mathrm{dist}(S(T)U, \Lambda)) \\ &\leq V(S(T)U) \leq \mathrm{e}^{-CT} V(U) \\ &\leq \mathrm{e}^{-CT} \beta(\mathrm{dist}(U, \Lambda)) \\ &\leq \mathrm{e}^{-CT} \beta(R_1) = \mathrm{e}^{-CT} \alpha(R_0). \end{aligned}$$

This is a contradiction, hence no such T exists and $S(t)U \in \mathcal{N}(\Lambda, R_0)$ $\forall t \geq 0$. Theorem 6.5(iii) then completes the proof. □

We now discuss structural assumptions on the vector field $f(\cdot)$ in (1.1) which yield the existence of attractors and uniformly asymptotically stable sets.

First we consider the assumption

$$\exists \epsilon, R > 0 : \langle f(u), u \rangle \leq -\epsilon \quad \forall u : \|u\| = R. \tag{6.3}$$

(Such an assumption holds, for example, for (1.1) under (2.13) for any $R \geq (a + \epsilon)/b$ although (2.13) cannot hold for globally bounded vector fields.) We may now prove:

Theorem 6.7 Consider (1.1) under (6.3) and let $B = \{u \in \mathbb{R}^p : \|u\|^2 < R\}$. Then the semigroup $S(t)$ has an attractor given by $\mathcal{A} = \omega(B)$.

Proof. By Theorem 6.2 it is sufficient to show that $S(t)\bar{B} \subset B \;\; \forall t > 0$. From (6.3) we have,

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u\|^2 = \langle f(u), u \rangle, \quad \forall u \in \mathbb{R}^p \tag{6.4}$$

and hence, for $\partial B = \bar{B} \backslash B$,

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u\|^2 = \langle f(u), u \rangle \leq -\epsilon, \quad \|u\| \in \partial B.$$

This shows that trajectories on ∂B point into B and establishes the required property of B; the result follows. □

A second class of systems of interest to us are gradient systems. Recall the definition $\mathcal{E} = \{v \in \mathbb{R}^p : f(v) = 0\}$ from Section 1.

Definition 6.8 The dynamical system $S(t)$ generated by (1.1) is said to define a *gradient system* if $\exists F \in C(\mathbb{R}^p, \mathbb{R})$, called a *Lyapunov function*, satisfying

(i) $F(u) \geq 0$ for all $u \in \mathbb{R}^p$;
(ii) $F(u) \to \infty$ as $\|\mathbf{u}\| \to \infty$;
(iii) for a solution of (1.1) $F(S(t)U)$ is nonincreasing in t;
(iv) if $F(S(t)U) = F(U)$ for $t > 0$ then $U \in \mathcal{E}$.

A particular case where gradient systems arise is when $\mathbf{f}$ is a gradient vector field, so that

$$f(u) = -\nabla F(u). \tag{6.5}$$

With this assumption, taking the inner-product in (1.1) with u_t yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\{F(u(t))\} = -\|u_t(t)\|^2. \tag{6.6}$$

Hence Properties (iii) and (iv) of Definition 6.8 hold; thus, if F satisfies (i) and (ii) then (1.1), (6.5) defines a gradient system. The following theorem shows that the dynamics of a gradient system must be relatively simple.

Theorem 6.9 If (1.1) is a gradient system then $\omega(U) \subseteq \mathcal{E}$. If, furthermore, the zeros of f are isolated and $\omega(U)$ is bounded then $\omega(U) = x$ for some $x \in \mathcal{E}$.

Proof. Now let x, y be two points in $\omega(U)$. Thus, without loss of generality, there exist sequences t_i and τ_i with $\tau_{i-1} < t_i < \tau_i$ such that

$$S(t_i)U \to x, \quad S(\tau_i)U \to y.$$

By Definition 6.8(iii) we have

$$F(S(\tau_i)U) \leq F(S(t_i)U) \leq F(S(\tau_{i-1})U).$$

Hence, by continuity, we deduce that $F(x) = F(y)$. Now, since $\omega(U)$ is positively invariant by Theorem 3.3 we have that for any $x \in \omega(U)$ and $t > 0$, $y = S(t)x \in \omega(U)$. But $F(x) = F(y)$ and, by Definition 6.8(iv), we deduce that $x \in \mathcal{E}$ yielding the first result.

By Theorem 3.3 we know that, if $\omega(U)$ is bounded it is connected. Since $\omega(U) \in \mathcal{E}$ and $\mathcal{E}$ comprises isolated points it follows that $\omega(U)$ is a single point $x \in \mathcal{E}$. □

Theorem 6.10 Consider a gradient dynamical system $S(t)$ generated by (1.1), 6.5. Assume that $F(u)$ has the property that

$$\exists \xi > 0 : v \in \mathcal{E} \Rightarrow F(v) \leq \xi.$$

Then the set $\overline{B(R)}$, where

$$B(R) := \{u \in \mathbb{R}^p : F(u) < R\},$$

is uniformly asymptotically stable for any $R > \xi$ and, furthermore, $\mathcal{A} = \omega(B(R))$ is a global attractor for any $R > \xi$.

Proof. Let $R > \xi$. First observe that $\overline{B(R)}$ is a closed bounded set because of Definition 6.8(ii); hence it is compact. Note that, for any $\epsilon > 0$, we may choose $r > 0$ such that

$$F(x) < r \Rightarrow \operatorname{dist}(x, B(R)) < \epsilon. \tag{6.7}$$

We may also choose $\delta > 0$ such that

$$\operatorname{dist}(x, B(R)) < \delta \Rightarrow F(x) < r. \tag{6.8}$$

Now, by (6.8), if $\operatorname{dist}(U, B(R)) < \delta$ then $F(U) < r$. By Definition 6.8(iii) we have $F(S(t)U) \leq F(U) < r$. Hence, by (6.7) we have $\operatorname{dist}(S(t)U, B(R)) < \epsilon$ as required to establish uniform stability (see Definition 6.3). Furthermore, choose δ_0 such that

$$\text{dist}(x, B(R)) < \delta_0 \Rightarrow F(x) < R + l.$$

Then let

$$\eta(l) := \inf_{\{x : R \leq F(x) \leq R+l\}} |f(x)| > 0; \tag{6.9}$$

the strict positivity follows since there are no equilibria with $F(x) \geq R$ and the set

$$\{x \in \mathbb{R}^p : R \leq F(x) \leq R + l\}$$

is compact. If $\text{dist}(U, B(R)) < \delta_0$ then $F(S(t)U) < R+l \; \forall t \geq 0$ by Definition 6.8(iii). Assume, for contradiction, that $F(S(t)U) \geq R \; \forall t > 0$. Then, from (6.6), we obtain

$$F(S(t)U) - (R + l) \leq -t\eta(l)^2 \tag{6.10}$$

yielding a contradiction for $t > l/\eta(l)^2$. Hence there exists $T \in (0, l/\eta(l)^2]$ such that $F(S(t)U) \leq R$ for all $t \geq T$. Thus $\text{dist}(S(t)U, B(R)) = 0 \; \forall t \geq T$ and asymptotic stability (see Definition 6.3) is proved, establishing uniform asymptotic stability.

Finally, (6.6) and (6.9) show that

$$F(S(t)U) = R \Rightarrow \frac{\mathrm{d}}{\mathrm{d}t}\{F(S(t)U)\} \leq -\eta(0)^2$$

so that trajectories on the boundary of $B(R)$ point into the set and, hence, by Theorem 6.2, $\omega(B(R))$ is an attractor. Since l is arbitrary and $B(R)$ is attracted to $\mathcal{A}$, we deduce from (6.9) and (6.10) that the bounded set $B(R+l)$ is attracted to $B(R)$ in finite time $T \leq l/\eta(l)^2$. Since l is arbitrary this gives the required global attraction. □

Recall Definition 4.1 of the unstable manifold given in Section 3. The following theorem elucidates the structure of the attractor for gradient systems.

Theorem 6.11 If $\mathcal{A}$ is a compact global attractor then it comprises all solutions of (1.1) which exist and are uniformly bounded for all $t \in \mathbb{R}$ Under the same assumptions as Theorem 6.10, $S(t)$ has a global attractor $\mathcal{A}$ given by

$$\mathcal{A} = \{U \in \mathbb{R}^p : \text{dist}(u(t), \mathcal{E}) \to 0 \quad \text{as} \quad t \to -\infty\}.$$

Furthermore, if $\mathcal{E}$ comprises only hyperbolic equilibrium points then

$$\mathcal{A} = \bigcup_{v \in \mathcal{E}} W^u(v).$$

Proof. The proof of the first part of this result may be found in Hale *et al.* (1984) and the remainder in Hale (1988). □

Example Consider the equation $u_t = u - u^3$. This is a gradient system with Lyapunov function $F(u) = (1-u^2)^2/4$. The equilibria are the points $\{0, +1, -1\}$ and all are hyperbolic since $f'(0) = 1$, $f'(\pm 1) = -2$. Furthermore, the points ± 1 are stable whilst the point 0 is unstable; by (3.8) it follows that the unstable manifold of 0 is $(-1, 1)$. Thus, by Theorem 6.11, the global attractor is the set $\mathcal{A} = [-1, 1]$. This result is an agreement with the construction of $\mathcal{A}$ in the second example following Definition 3.2. □

We will show in Section 6.3 that if $S(t)$ has an attractor $\mathcal{A}$ then $S^n_{\Delta t}$ has an attractor $\mathcal{A}_{\Delta t}$ satisfying

$$\text{dist}(\mathcal{A}_{\Delta t}, \mathcal{A}) \to 0 \quad \text{as} \quad \Delta t \to 0.$$

This shows that, in the limit as $\Delta t \to 0$, every point on the numerical attractor is close to a point on the true attractor and is known as *upper semicontinuity.* We will also show that, in general, the converse is not true – we will only be able to prove *lower-semicontinuity*, namely that

$$\text{dist}(\mathcal{A}, \mathcal{A}_{\Delta t}) \to 0 \quad \text{as} \quad \Delta t \to 0,$$

under very special assumptions on the nature of the flow on the attractor; these essentially amount to the system being in gradient form, or something closely related to it – see Sections 6.4 and 6.5. The following example illustrates the essential difficulty in trying to derive lower-semicontinuity results:

Example Consider equation (1.1) in dimension $p = 1$. It follows that, for $F(u) : \mathbb{R} \mapsto \mathbb{R}$ defined so that $F'(u) = -f(u)$, we have that (6.6) holds and the system is in gradient form provided F satisfies (i) and (ii) of Definition 6.8. In this case, all solutions will have their ω-limit sets contained in the set of equilibrium points.

Now consider (1.1) with $f(u) \mapsto f_\epsilon(u)$ given by

$$f_\epsilon(u) = \left\{ \begin{array}{cc} -(u+1)^3 + \epsilon, & u \le -1, \\ \epsilon(u^3/2 - 3u/2), & -1 < u < 1, \\ -(u-1)^3 - \epsilon, & u \ge 1, \end{array} \right\} \tag{6.11}$$

This vector field is $C^1(\mathbb{R}, \mathbb{R})$ and satisfies (i) and (ii) of Definition 6.8 for each $\epsilon \ge 0$. Hence, by Theorem 6.10 the system has a global attractor $\mathcal{A}_\epsilon$, say, for each $\epsilon \ge 0$.

The potential $F_\epsilon(u)$ satisfying $F'_\epsilon(u) = -f_\epsilon(u)$ and $F_\epsilon(0) = 0$ is shown in Figure 3 for $\epsilon > 0$ and $\epsilon = 0$ respectively.

Examination of these figures, together with application of Theorem 6.11, shows that, for every $\epsilon > 0$, the attractor of (1.1), (6.11) is given by

$$\mathcal{A}_\epsilon = \{0\}, \quad \epsilon > 0,$$

a single point; a similar analysis for $\epsilon = 0$ shows that

$$\mathcal{A}_0 = [-1, 1],$$

an entire interval. Thus the perturbed attractors with $\epsilon > 0$ are contained in the unperturbed attractor at $\epsilon = 0$ but not the other way around. This shows that the attractor $\mathcal{A}_0$ is *upper semicontinuous* with respect to $\epsilon > 0$ but it is not *lower semicontinuous.* Although the perturbation induced by ϵ in this example is not directly analogous to a numerical approximation, it nonetheless indicates an important point – without strong assumptions it may be difficult to prove lower semicontinuity of attractors with respect to perturbations of any kind, including those induced by numerical approximation.

The difficulty observed in the example is a consequence of the fact that certain portions of the attractor may attract very slowly – specifically slower than exponentially – and hence disappear under perturbation. In order to get around this difficulty it is natural to consider a slightly enlarged object which does have a form of exponential attraction. This is one motivation for the consideration of the weaker concept of a uniformly asymptotically stable set for which it is possible to prove both upper and lower semicontinuity.

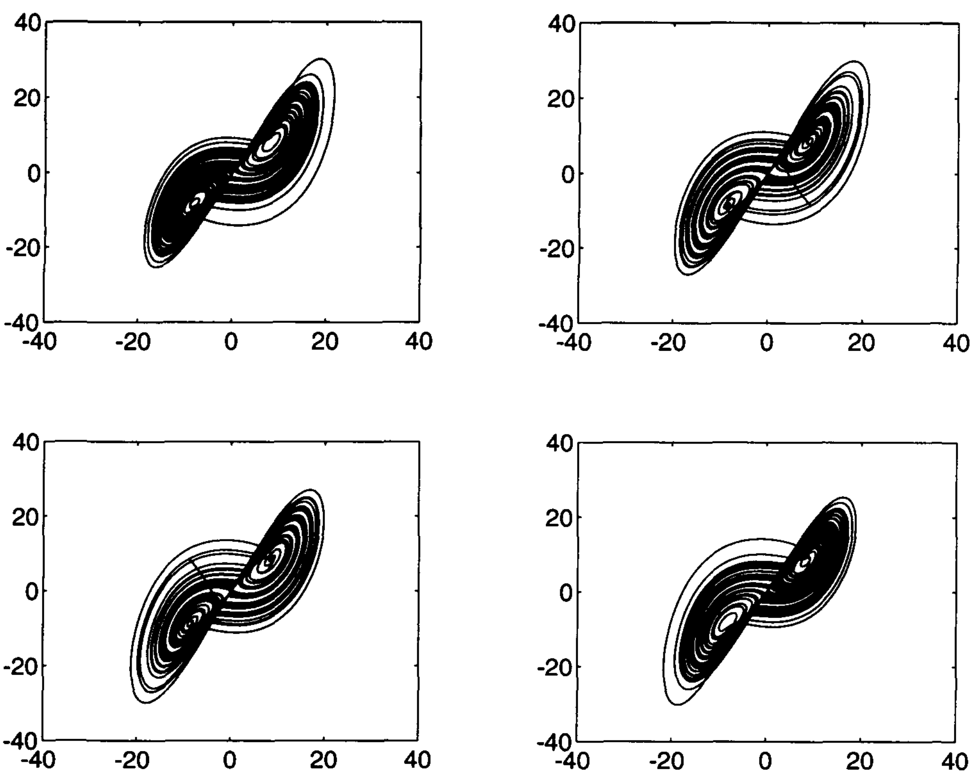

Fig. 3. The potential $F_\epsilon(u)$ (vertical axis) against u for $\epsilon = 0.01$ (top) and $\epsilon = 0.0$ (bottom).

Important remark All our definitions and theorems concern the continuous semigroup $S(t)$. However, the Definitions 6.1,6.3 and 6.8 and Theorems 6.2, 6.4 and 6.11 can all be extended to a discrete semigroup $S^n_{\Delta t}$ simply by replacing t by n in the definitions and using the continuity of $S^n_{\Delta t}$. Theorem 6.9 can also be extended but the extension is slightly less trivial – see Humphries and Stuart (1994).

6.2. Continuity of uniformly asymptotically stable attracting sets

We now consider (1.1) under the assumption that there exists a compact set Λ which is uniformly asymptotically stable.

Theorem 6.12 (Uniformly asymptotically stable sets under approximation) Assume that the semigroup $S(t)$ for (1.1) has a compact, uniformly asymptotically stable set Λ. Then there exists $\Delta t_c > 0$ such that, for all $\Delta t \in (0, \Delta t_c]$, the approximating semigroup $S^n_{\Delta t}$ for (1.2) has a compact, uniformly asymptotically stable set $\Lambda_{\Delta t} \supset \Lambda$ which satisfies

$$d_{\mathrm{H}}(\Lambda_{\Delta t}, \Lambda) \to 0 \quad \text{as} \quad \Delta t \to 0.$$

Furthermore there exists a set $B \supset \Lambda_{\Delta t}$ and $T = T(\Delta t)$ with the property that

$$S^n_{\Delta t} B \subset \Lambda_{\Delta t} \;\; \forall n : n\Delta t \geq T.$$

The proof will be performed through a sequence of lemmas. Recall the definition of R_1 from the proof of Corollary 6.6.

Lemma 6.13 There exists $r_0 \in (0, R_1)$ and $\Delta t_0 > 0$ such that

$$U \in \mathcal{N}(\Lambda, r_0) \Rightarrow S^1_{\Delta t} U \in \mathcal{N}(\Lambda, R_0) \quad \forall \Delta t \in (0, \Delta t_0].$$

Furthermore it then follows that

$$V(S^1_{\Delta t} U) \leq \mathrm{e}^{-C\Delta t} V(U) + KL\Delta t^{r+1}.$$

Proof. Define r_0 and Δt_0 by

$$\beta(r_0) = \alpha(R_0/2) \quad \text{and} \quad K\Delta t_0^{r+1} = R_0/2. \tag{6.12}$$

Note that $r_0 < R_1 < R_0$, so that, by Corollary 6.6,

$$U \in \mathcal{N}(\Lambda, r_0) \Rightarrow S(t)U \in \mathcal{N}(\Lambda, R_0) \quad \forall t > 0 \quad .$$

Theorem 6.5(iii) shows that we have

$$V(S(\Delta t)U) \leq \mathrm{e}^{-C\Delta t} V(U) < V(U). \tag{6.13}$$

But

$$\alpha(\mathrm{dist}(S(\Delta t)U, \Lambda)) \leq V(S(\Delta t)U) \tag{6.14}$$

and

$$V(U) \leq \beta(\mathrm{dist}(U,\Lambda)) < \beta(r_0) = \alpha(R_0/2). \tag{6.15}$$

Equations (6.13)–(6.15) imply that

$$\alpha(\mathrm{dist}(S(\Delta t)U,\Lambda)) \leq \alpha(R_0/2) \Rightarrow \mathrm{dist}(S(\Delta t)U,\Lambda) \leq R_0/2\ . \tag{6.16}$$

Now,

$$\begin{aligned} \mathrm{dist}(S^1_{\Delta t}U,\Lambda) &= \inf_{x\in\Lambda} \|S^1_{\Delta t}U - x\| \\ &\leq \inf_{x\in\Lambda} \|S(t)U - x\| + \|S(t)U - S^1_{\Delta t}U\| \\ &\leq \mathrm{dist}(S(\Delta t)U,\Lambda) + \|S(\Delta t)U - S^1_{\Delta t}U\|. \end{aligned} \tag{6.17}$$

Thus (6.16), (6.18), (6.12) and Assumption 3.7(iii) gives

$$\mathrm{dist}(S^1_{\Delta t}U,\Lambda) \leq \tfrac{1}{2}R_0 + K\Delta t^{r+1} \leq R_0 \forall \Delta t \in (0,\Delta t_0].$$

Now, by Theorem 6.5(i) and Assumption 3.7(iii) we have

$$|V(S^1_{\Delta t}U) - V(S(\Delta t)U)| \leq L\|S^1_{\Delta t}U - S(\Delta t)U\| \leq KL\Delta t^{r+1}.$$

Hence, by Theorem 6.5(iii),

$$\begin{aligned} V(S^1_{\Delta t}U) &\leq V(S(\Delta t)U) + KL\Delta t^{r+1} \\ &\leq \mathrm{e}^{-C\Delta t}V(U) + KL\Delta t^{r+1}, \end{aligned}$$

by (6.13). □

The previous lemma gives a bound on U which ensures that $V(S^1_{\Delta t}U)$ is defined. Using this fact we now construct a positively invariant set B for $S^1_{\Delta t}$.

Lemma 6.14 The set $B = \{x \in \mathcal{N}(\Lambda, R_0) : V(x) < \alpha(r_0)\}$ is open and $\Lambda \subset B \subset \mathcal{N}(\Lambda, r_0)$. Furthermore, there exists Δt_1 such that B is positively invariant under $S^1_{\Delta t}$ for all $\Delta t \in (0, \Delta t_1]$.

Proof. The set B is open since V is continuous and $V(x) < \alpha(r_0)$ for all $x \in B$. The set B contains Λ since $V(x) = 0$ if $x \in \Lambda$. Let $U \in B$; then

$$\alpha(\mathrm{dist}(U,\Lambda)) \leq V(U) < \alpha(r_0) \tag{6.18}$$

by the properties of $\alpha(\cdot)$ and B. Thus $\mathrm{dist}(U,\Lambda) < r_0$ which implies that $U \in \mathcal{N}(\Lambda, r_0)$ as required. Hence $B \subset \mathcal{N}(\Lambda, r_0)$. Now we define Δt_1, the largest number in $(0, \Delta t_0]$ such that

$$\frac{KL\Delta t^{p+1}}{1 - \mathrm{e}^{-C\Delta t}} < \tfrac{1}{4}\alpha(r^*) \quad \forall \Delta t \in (0, \Delta t_1], \tag{6.19}$$

where

$$r^* = \tfrac{1}{2}\beta^{-1}(\alpha(r_0)). \tag{6.20}$$

Note that

$$r^* < \tfrac{1}{2}\alpha^{-1}(\alpha(r_0)) = \tfrac{1}{2}r_0 < r_0.$$

Thus

$$\alpha(r^*) < \alpha(r_0). \tag{6.21}$$

Now, if $U \in B \subset \mathcal{N}(\Lambda, r_0)$ then by Lemma 6.13, (6.18) and (6.19) we have

$$\begin{aligned} V(S^1_{\Delta t}U) &\leq \mathrm{e}^{-C\Delta t}V(U) + KL\Delta t^{p+1} \\ &< \mathrm{e}^{-C\Delta t}\alpha(r_0) + \tfrac{1}{4}\alpha(r^*)(1-\mathrm{e}^{-C\Delta t}) \\ &\leq \tfrac{1}{4}\alpha(r_0) + \tfrac{3}{4}\mathrm{e}^{-C\Delta t}\alpha(r_0) < \alpha(r_0). \end{aligned}$$

In addition, $S^1_{\Delta t}U \in \mathcal{N}(\Lambda, R_0)$ by Lemma 6.13 and so we have $S^1_{\Delta t}B \subset B$ as required. □

We now construct the approximate uniformly asymptotically stable set $\Lambda_{\Delta t}$.

Lemma 6.15 Define

$$\eta(\Delta t) = \frac{2KL\Delta t^{r+1}}{(1-\mathrm{e}^{-C\Delta t})}, \quad \Delta t \in (0, \Delta t_1]$$

and define

$$\Lambda_{\Delta t} = \{x \in \mathcal{N}(\Lambda, R_0) : V(x) \leq \eta(\Delta t)\}.$$

Then $\Lambda_{\Delta t}$ is compact, positively invariant, contains Λ in its interior and satisfies $d_{\mathrm{H}}(\Lambda_{\Delta t}, \Lambda) \to 0$ as $\Delta t \to 0$.

Proof. $\Lambda_{\Delta t}$ is bounded since Λ is bounded and V is continuous; $\Lambda_{\Delta t}$ is closed by construction. Hence it is a compact set. Clearly, by (6.19) and (6.21) it follows that

$$\Lambda_{\Delta t} \subset B \subset \mathcal{N}(\Lambda, r_0), \quad \Delta t \in (0, \Delta t_1], \tag{6.22}$$

since $\alpha(r^*) < \alpha(r_0)$. Thus by Lemma 6.13, if $U \in \Lambda_{\Delta t}$ we have

$$\begin{aligned} V(S^1_{\Delta t}U) &\leq \mathrm{e}^{-C\Delta t}V(U) + KL\Delta t^{r+1} \\ &\leq \mathrm{e}^{-C\Delta t}\eta(\Delta t) + \tfrac{1}{2}\eta(\Delta t)(1-\mathrm{e}^{-C\Delta t}) \leq \eta(\Delta t). \end{aligned}$$

Since $S^1_{\Delta t}U \in \mathcal{N}(\Lambda, R_0)$ by (6.22) and Lemma 6.13, we have $S^1_{\Delta t}U \in \Lambda_{\Delta t}$ as required.

It is clear that Λ is contained in the interior of $\Lambda_{\Delta t}$ since V is continuous and $\eta(\Delta t) > 0$. Thus $\mathrm{dist}(\Lambda, \Lambda_{\Delta t}) = 0$. Also $\mathrm{dist}(\Lambda_{\Delta t}, \Lambda) = \sup_{x \in \Lambda_{\Delta t}} \mathrm{dist}(x, \Lambda)$. But, for every $x \in \Lambda_{\Delta t}$ we have that $\mathrm{dist}(x, \Lambda) \leq \alpha^{-1}(\eta(\Delta t))$. Since $\eta(\Delta t) \to 0$ as $\Delta t \to 0$ it follows that $\mathrm{dist}(\Lambda_{\Delta t}, \Lambda) \to 0$ as $\Delta t \to 0$. □

We now show that iterates starting in B are absorbed into $\Lambda_{\Delta t}$ in a finite number of steps under $S^1_{\Delta t}$, giving asymptotic stability.

Lemma 6.16 There exists $\Delta t_2 \in (0, \Delta t_1]$ such that, for any $U \in B\backslash\Lambda_{\Delta t}$ and any $\Delta t \in (0, \Delta t_2]$,

$$V(S^1_{\Delta t}U) \leq \mathrm{e}^{-C\Delta t/4}V(U).$$

Furthermore, if

$$T(\Delta t) = \frac{4}{C}\ln\{\alpha(r_0)/\eta(\Delta t)\},$$

then there exists a $\delta_0 > 0$ such that, $\mathcal{N}(\Lambda_{\Delta t}, \delta_0) \subset B$ and, furthermore, if $U \in \mathcal{N}(\Lambda_{\Delta t}, \delta_0)$ then $S^n_{\Delta t}U \in \Lambda_{\Delta t}$ $\forall n : n\Delta t \geq T(\Delta t)$.

Proof. Let $U \in B\backslash\Lambda_{\Delta t}$. It follows that

$$\eta(\Delta t) \leq V(U) \leq \alpha(r_0). \tag{6.23}$$

Then, by Lemma 6.13 and (6.23) it follows that, for $\Delta t \in (0, \Delta t_1]$

$$\begin{aligned} V(S^1_{\Delta t}U) &\leq \mathrm{e}^{-C\Delta t}V(U) + KL\Delta t^{r+1} \\ &= \mathrm{e}^{-C\Delta t}V(U) + \tfrac{1}{2}\eta(\Delta t)(1 - \mathrm{e}^{-C\Delta t}) \\ &\leq \tfrac{1}{2}(1 + \mathrm{e}^{-C\Delta t})V(U). \end{aligned}$$

Now define $\Delta t_2 = \min\{\Delta t_1, \gamma\}$ where $\mathrm{e}^{-C\gamma} + 1 = 2\mathrm{e}^{-C\gamma/4}$. Then

$$V(S^1_{\Delta t}U) < \mathrm{e}^{-C\Delta t/4}V(U)$$

as required. Iterating this bound shows that, if we assume that $S^j_{\Delta t} \in B\backslash\Lambda_{\Delta t}$ for $j = 0, \ldots, n-1$ then

$$V(S^n_{\Delta t}U) < \mathrm{e}^{-C\eta\Delta t/4}V(U). \tag{6.24}$$

From (6.23) it follows that

$$V(S^n_{\Delta t}U) < \mathrm{e}^{-C\eta\Delta t/4}\alpha(r_0)\ .$$

If $n\Delta t \geq T(\Delta t)$ then

$$\mathrm{e}^{-Cn\Delta t/4} \leq \eta(\Delta t)/\alpha(r_0)$$

so that

$$V(S^n_{\Delta t}U) < \eta(\Delta t).$$

Hence $S^n_{\Delta t}U \in \Lambda_{\Delta t}$, a contradiction. It follows that $S^n_{\Delta t}$ enters $\Lambda_{\Delta t}$ for some $n > 0 : n\Delta t \leq T(\Delta t)$.

Finally, to exhibit asymptotic stability we need to find an appropriate δ_0 such that $\mathcal{N}(\Lambda_{\Delta t}, \delta_0) \subset B$. Let $\delta_0 = \frac{1}{2}\beta^{-1}(\alpha(r_0)) > 0$. Then, if $U \in \mathcal{N}(\Lambda_{\Delta t}, \delta_0)$ we wish to show that $U \in B$. Now, for any $z \in \mathbb{R}^p$,

$$\mathrm{dist}(U, \Lambda) = \inf_{y\in\Lambda} \|U - y\| \leq \|U - z\| + \inf_{y\in\Lambda} \|z - y\|.$$

Choose z: $\|U - z\| = \inf_{y\in\Lambda_{\Delta t}} \|U - y\|$. Then, by Lemmas 6.15 and (6.19), (6.20) we obtain

$$\begin{aligned}
\mathrm{dist}(U,\Lambda) &\le \mathrm{dist}(U,\Lambda_{\Delta t}) + \mathrm{dist}(\Lambda_{\Delta t},\Lambda)\\
&< \delta_0 + \alpha^{-1}(\eta(\Delta t))\\
&= \tfrac{1}{2}\beta^{-1}(\alpha(r_0)) + \alpha^{-1}\left(\frac{2KL\Delta t^{r+1}}{1-\mathrm{e}^{-C\Delta t}}\right)\\
&< \tfrac{1}{2}\beta^{-1}(\alpha(r_0)) + \alpha^{-1}(\tfrac{1}{2}\alpha(r_*))\\
&< \tfrac{1}{2}\beta^{-1}(\alpha(r_0)) + r_* = \beta^{-1}(\alpha(r_0)).
\end{aligned}$$

Hence $\mathrm{dist}(U,\Lambda) < \beta^{-1}(\alpha(r_0))$ which implies that $V(U) \le \beta(\mathrm{dist}(U,\Lambda)) < \alpha(r_0)$. Hence $U \in B$ and the result follows. □

Finally we show that the stability is uniform.

Lemma 6.17 Let $\Delta t \in (0, \Delta t_2]$. Then, for each $\epsilon > 0$ there exists $\delta = \delta(\epsilon, \Delta t) > 0$ such that

$$U \in \mathcal{N}(\Lambda_{\Delta t}, \delta) \Rightarrow S^n_{\Delta t}U \in \mathcal{N}(\Lambda_{\Delta t}, \epsilon) \quad \forall n \ge 0\ .$$

Proof. By Assumption 3.7(i)

$$\|S^1_{\Delta t}U - S^1_{\Delta t}V\| \le (1 + K\Delta t)\|U - V\|.$$

Let

$$\delta = \min\{\delta_0, \tfrac{1}{2}\epsilon(1 + K\Delta t)^{-T(\Delta t)/\Delta t}\}\ .$$

Let $U \in \mathcal{N}(\Lambda_{\Delta t}, \delta)$. If $U \in \Lambda_{\Delta t}$ then positive invariance implies the result automatically, by Lemma 6.15. Hence suppose $U \notin \Lambda_{\Delta t}$. Choose $V \in \Lambda_{\Delta t}$ such that $\mathrm{dist}(U, \Lambda_{\Delta t}) = \|U - V\|$; then $\|S^n_{\Delta t}U - S^n_{\Delta t}V\| \le (1 + K\Delta t)^n\|U - V\|$. Also, since $V \in \Lambda_{\Delta t}$ we have that $S^n_{\Delta t}V \in \Lambda_{\Delta t}$. Thus

$$\begin{aligned}
\mathrm{dist}(S^n_{\Delta t}U, \Lambda_{\Delta t}) &\le \inf_{y\in\Lambda_{\Delta t}} \|S^n_{\Delta t}U - y\| \le \|S^n_{\Delta t}U - S^n_{\Delta t}V\|\\
&\le (1 + K\Delta t)^n\|U - V\|\\
&= (1 + K\Delta t)^n \mathrm{dist}(U, \Lambda_{\Delta t})\\
&= (1 + K\Delta t)^n\delta. \qquad (6.25)
\end{aligned}$$

The bound (6.25) and the fact that $\delta \le \delta_0$ shows that

$$\mathrm{dist}(S^n_{\Delta t}U, \Lambda_{\Delta t}) \le \tfrac{1}{2}\epsilon \quad \forall n : n\Delta t \le T(\Delta t).$$

Also, since

$$S^n_{\Delta t}U \in \Lambda_{\Delta t} \quad \forall n : n\Delta t \ge T(\Delta t)$$

we have

$$\mathrm{dist}(S^n_{\Delta t}U, \Lambda_{\Delta t}) = 0 \le \tfrac{1}{2}\epsilon \quad \forall n \ge 0.$$

□

Proof of Theorem 6.12. Lemma 6.15 establishes existence of a positive invariant set $\Lambda_{\Delta t}$ and its convergence properties. Lemmas 6.16 and 6.17 show uniform asymptotic stability and Lemma 6.16 gives the required absorbtion property. □

In the following section we will give a different construction of a set $\Lambda_{\Delta t}$ which has the same properties as the set $\Lambda_{\Delta t}$ constructed in Theorem 6.12. We conclude this section with a corollary of Theorem 6.12 which relates to the existence of attractors. Note that, whilst uniformly asymptotically stable sets have been proven both upper and lower semicontinuous, the following result establishes only upper semiconinuity of attractors.

Corollary 6.18 **(Attractors under approximation)** Assume that the semigroup $S(t)$ for (1.1) has an attractor $\mathcal{A}$. Then there exists $\Delta t_c > 0$ such that, for all $\Delta t \in (0, \Delta t_c]$, the approximating semigroup $S^n_{\Delta t}$ for (1.2) has a compact, uniformly asymptotically stable set $\Lambda_{\Delta t} \supset \mathcal{A}$ and an attractor $\mathcal{A}_{\Delta t} = \omega(\Lambda_{\Delta t})$ satisfying

$$\mathrm{dist}(\mathcal{A}_{\Delta t}, \mathcal{A}) \to 0 \quad \text{as} \quad \Delta t \to 0.$$

Proof. Note that $\mathcal{A}$ is uniformly asymptotically stable by Theorem 6.4(ii). Hence, by Theorem 6.12 with $\Lambda = \mathcal{A}$, we deduce that (1.2) has a uniformly asymptotically stable set $\Lambda_{\Delta t} \supset \mathcal{A}$. Furthermore, by Theorem 6.4(iii) it follows that $\mathcal{A}_{\Delta t} = \omega(\Lambda_{\Delta t})$ (the limit set being defined through $S^n_{\Delta t}$) is an attractor for $S^n_{\Delta t}$. Note that $\mathcal{A}_{\Delta t} \subseteq \Lambda_{\Delta t}$ and that, by Theorem 6.12,

$$\mathrm{dist}(\Lambda_{\Delta t}, \mathcal{A}) \to 0 \quad \text{as} \quad \Delta t \to 0.$$

Hence

$$\mathrm{dist}(\mathcal{A}_{\Delta t}, \mathcal{A}) \to 0 \quad \text{as} \quad \Delta t \to 0$$

and the proof is complete. □

6.3. Upper semicontinuity of attractors

In this section we consider the numerical approximation of (1.1) satisfying (6.3). Under this assumption it follows from Theorem 6.7 that (1.1) has a local attractor $\mathcal{A} = \omega(B)$ and it is thus natural to study the effect of the numerical approximation of (1.1) on $\mathcal{A}$. It would be possible to use (6.3) to deduce the exsitence of a uniformly asymptotically stable set for (1.1) and then apply the methods of the previous section to deduce all the results in this section. However, we present here an entirely different approach to the problem. Our first result proves that an approximate attractors exists.

Theorem 6.19 **(Existence of an approximate attractor)** Consider the approximation of (1.1) under (6.3) by (1.2) and let $B = \{u \in \mathbb{R}^p : \|u\| < R\}$. Then there exists $\Delta t_c > 0$ such that the semigroup $S^n_{\Delta t}$ has an attractor given by $\mathcal{A}_{\Delta t} = \omega(B)$ for all $\Delta t \in (0, \Delta t_c]$.

Proof. Recall that $\partial B = \bar{B}\backslash B$. By continuity of f and (6.3) it is possible to choose $\delta > 0$ sufficiently small so that

$$\langle f(u), u\rangle \leq -\tfrac{1}{2}\epsilon \quad \forall u : R-\delta \leq \|u\| \leq R. \tag{6.26}$$

Recall also that, by Assumption 3.6, there exists $L > 0$ such that $\|f(u)\| \leq L$ for all $u \in \mathbb{R}^p$. Now choose $\Delta t_c > 0$ so that

$$L\Delta t \leq \tfrac{1}{4}\delta,\ K\Delta t^{r+1} \leq R,\ RK\Delta t^r \leq \tfrac{1}{10}\epsilon,\ K\Delta t^{r+1} \leq \tfrac{1}{8}\delta\ \forall \Delta t \in (0, \Delta t_c]. \tag{6.27}$$

In the remainder of this proof we assume that $\Delta t \in (0, \Delta t_c]$. We deduce that

$$\|u(t+s) - u(t)\| \leq \int_t^{t+s} \|f(u(\tau)\|\,\mathrm{d}\tau \leq L\Delta t < \tfrac{1}{2}\delta \quad \forall s \in (0, \Delta t_c]. \tag{6.28}$$

Thus, if $R - \delta/2 \leq \|U\| \leq R$ we have, from the positive invariance of B under $S(t)$ and (6.28), that

$$R - \delta \leq \|S(t)U\| \leq R, \quad t \in (0, \Delta t_c].$$

It then follows from (6.4) and (6.26) that, for all $\Delta t \in (0, \Delta t_c]$,

$$\|S(\Delta t)U\|^2 - \|U\|^2 = 2\int_0^{\Delta t} \langle f(S(\tau)U), S(\tau)U\rangle\,\mathrm{d}\tau \leq -\epsilon\Delta t.$$

Hence by Assumption 3.7(iii), (6.26) and (6.27) we have

$$\begin{aligned}
\|S^1_{\Delta t}U\|^2 - \|U\|^2 &\leq \|S(\Delta t)U\|^2 - \|U\|^2 \\
&\quad + [2\|S^1_{\Delta t}U\| + \|S(\Delta t)U - S^1_{\Delta t}U\|] \\
&\quad \|S(\Delta t)U - S^1_{\Delta t}U\| \\
&\leq -\epsilon\Delta t + [2\|S(\Delta t)U\| + 3\|S(\Delta t)U - S^1_{\Delta t}U\|] \\
&\quad \times \|S(\Delta t)U - S^1_{\Delta t}U\| \\
&\leq -\epsilon\Delta t + [2R + 3K\Delta t^{r+1}]K\Delta t^{r+1} \\
&\leq -\epsilon\Delta t + 5RK\Delta t^{r+1} \\
&\leq -\tfrac{1}{2}\epsilon\Delta t.
\end{aligned}$$

Thus we have proved that

$$R - \tfrac{1}{2}\delta \leq \|U\| \leq R \quad \Rightarrow \quad \|S^1_{\Delta t}U\|^2 - \|U\|^2 \leq -\tfrac{1}{2}\epsilon\Delta t. \tag{6.29}$$

Now, if $\|U\| < R - \frac{1}{2}\delta$ we have, by (6.28), (6.27) and Assumption 3.7(iii),

$$\begin{aligned}
\|S^1_{\Delta t}U\| &\leq \|S(\Delta t)U\| + \|S(\Delta t) - S^1_{\Delta t}U\| \\
&\leq \|U\| + \|S(\Delta t)U - U\| + \|S(\Delta t) - S^1_{\Delta t}U\| \\
&< R - \tfrac{1}{2}\delta + \tfrac{1}{4}\delta + K\Delta t^{r+1} \\
&\leq R - \tfrac{1}{8}\delta.
\end{aligned}$$

Hence

$$\|U\| < R - \tfrac{1}{2}\delta \quad \Rightarrow \quad \|S^1_{\Delta t}U\| \leq R - \tfrac{1}{8}\delta. \tag{6.30}$$

It is clear that (6.29) and (6.30) imply that

$$S^1_{\Delta t}\bar{B} \subset B \tag{6.31}$$

and hence, by the generalization of Theorem 6.2 from $S(t)$ to $S^n_{\Delta t}$, we deduce that $\mathcal{A}_{\Delta t} = \omega(B)$ is an attractor for $S^n_{\Delta t}$. □

As the example in Section 6.1 illustrates we cannot expect to obtain lower semicontinuity of $\mathcal{A}$ without imposing further conditions on the dynamical system in question. However, it is possible to derive upper semicontinuity results without further assumptions.

Theorem 6.20 **(Upper semicontinuity of attractors)** Consider the approximation of (1.1) under (6.3) by (1.2). Then the attractors of the semigroups $S(t)$ and $S^n_{\Delta t}$, $\mathcal{A}$ and $\mathcal{A}_{\Delta t}$ respectively, satisfy

$$\text{dist}(\mathcal{A}_{\Delta t}, \mathcal{A}) \to 0 \qquad \text{as} \qquad \Delta t \to 0.$$

Proof. Note that, if $\mathcal{A}_{\Delta t} \subseteq \overline{\mathcal{N}}(\mathcal{A}, \epsilon)$, then $\text{dist}(\mathcal{A}_{\Delta t}, \mathcal{A}) \leq \epsilon$; thus, if we show that for any $\epsilon > 0$ there exists $\Delta = \Delta(\epsilon)$ such that

$$\mathcal{A}_{\Delta t} \in \overline{\mathcal{N}}(\mathcal{A}, \epsilon) \quad \forall \Delta t \in (0, \Delta], \tag{6.32}$$

then the result follows. Thus we aim to prove (6.32).

First we estimate the attraction of B to a neighbourhood of $\mathcal{A}$ under $S^n_{\Delta t}$ by using the attractivity of B to $\mathcal{A}$ under $S(t)$ and the truncation error bound. We have that

$$\begin{aligned} \text{dist}(S^n_{\Delta t}U, \mathcal{A}) &= \inf_{x \in \mathcal{A}} \|S^n_{\Delta t}U - x\| \\ &\leq \|S^n_{\Delta t}U - S(n\Delta t)U\| + \inf_{x \in \mathcal{A}} \|S(n\Delta t)U - \mathbf{x}\| \\ &\leq \|S^n_{\Delta t}U - S(n\Delta t)U\| + \text{dist}(S(n\Delta t)U, \mathcal{A}). \end{aligned}$$

Hence we have that

$$\text{dist}(S^n_{\Delta t}U, \mathcal{A}) \leq \text{dist}(S(n\Delta t)U, \mathcal{A}) + \|S^n_{\Delta t}U - S(n\Delta t)U\|. \tag{6.33}$$

Now, since $\mathcal{A}$ attracts B under $S(t)$ there exists $T = T(\epsilon) > 0$ such that, for all $U \in B$,

$$S(t)U \in \mathcal{N}(\mathcal{A}, \epsilon/2) \quad \forall t \geq T. \tag{6.34}$$

Without loss of generality we may choose $T = N\Delta t$ for some integer N. By Theorem 3.8 it follows that, for any $U \in B$, there exists $\Delta = \Delta(\epsilon) > 0$ such that for any $\Delta t \in (0, \Delta]$

$$\|S^n_{\Delta t}U - S(n\Delta t)U\| \leq \epsilon/2 \quad \forall n, \Delta t : 0 \leq n\Delta t \leq 2T. \tag{6.35}$$

Hence it follows from (6.33), (6.34) and (6.35) that, for all $\Delta t \in (0, \Delta]$

$$\text{dist}(S^n_{\Delta t}B, \mathcal{A}) < \epsilon \quad \forall n, \Delta t : T \leq n\Delta t \leq 2T. \tag{6.36}$$

We now proceed to use induction. Suppose that, for some integer $k \geq 2$

$$\operatorname{dist}(S^n_{\Delta t}B, \mathcal{A}) < \epsilon \quad \forall n, \Delta t : T \leq n\Delta t \leq kT. \tag{6.37}$$

Note that this has been proved true for $k = 2$. Now consider integer n such that $kT \leq n\Delta t \leq (k+1)T$. Choose m and p such that $n = m + p$, $T \leq m\Delta t \leq 2T$ and $p\Delta t = (k-1)T$; thus $p = (k-1)N$. Then $S^n_{\Delta t}B = S^m_{\Delta t}S^p_{\Delta t}B$ and by (6.31) it follows that $S^p_{\Delta t}B \subset B$. This implies that

$$S^n_{\Delta t}B = S^m_{\Delta t}S^p_{\Delta t}B \subset S^m_{\Delta t}B;$$

since $T \leq mh \leq 2T$ it may be shown, from (6.36) with $n \mapsto m$, that

$$S^m_{\Delta t}B \subset \mathcal{N}(\mathcal{A}, \epsilon)$$

and hence that

$$S^n_{\Delta t}B \subset \mathcal{N}(\mathcal{A}, \epsilon) \quad \forall n, \Delta t : kT \leq n \leq (k+1)T.$$

This, together with (6.37), completes the inductive step and we deduce that

$$\operatorname{dist}(S^n_{\Delta t}B, \mathcal{A}) < \epsilon \quad \forall n, \Delta t : T \leq n\Delta t < \infty. \tag{6.38}$$

Finally recall that

$$\mathcal{A}_{\Delta t} = \bigcap_{m \geq 0} \overline{\bigcup_{n \geq m} S^n_{\Delta t}B}$$

and since (6.38) holds it follows that

$$\overline{\bigcup_{n \geq N} S^n_{\Delta t}B} \subseteq \overline{\mathcal{N}}(\mathcal{A}, \epsilon),$$

where $N\Delta t = T$, and hence $\mathcal{A}_{\Delta t} \subseteq \overline{\mathcal{N}}(\mathcal{A}, \epsilon)$ as required. □

Note that Theorem 6.20 does not give a rate of convergence for the quantity $\operatorname{dist}(\mathcal{A}_{\Delta t}, \mathcal{A})$. This is since nothing is assumed about the *rate of attraction* of the attractor. If the rate is assumed exponential then a stronger result can be proved and we obtain the error bound given in Theorem 6.21 below. Note that the bound is less than the rate of convergence of individual trajectories and reflects the competition between the exponential attraction to $\mathcal{A}$ (which determines α) and the exponential divergence of trajectories on $\mathcal{A}$ (which determines K).

Theorem 6.21 (**Rate of convergence of attractors**) Consider the approximation of (1.1) satisfying (6.3) by (1.2) and assume that the attractor of the semigroup $\mathcal{A}$ is exponentially attracting in the sense that there exist $C_1 > 0, \alpha > 0$ such that

$$\operatorname{dist}(S(t)B, \mathcal{A}) \leq C_1 e^{-\alpha t}. \tag{6.39}$$

Then there exist $\Delta t_c, C_2 > 0$ such that the the global attractors $\mathcal{A}$ and $\mathcal{A}_{\Delta t}$ of $S(t)$ and $S^n_{\Delta t}$ respectively, satisfy

$$\mathrm{dist}(\mathcal{A}_{\Delta t}, \mathcal{A}) \le C_2 \Delta t^\beta,$$

where $\beta = \alpha r/(K+\alpha)$, for all $\Delta t \in (0, \Delta t_c]$.

Proof. Using (6.33) and the arguments following it in the proof of Theorems 6.20, 6.39 and using Theorem 3.8, we obtain, for any $U \in B$,

$$\mathrm{dist}(S^n_{\Delta t} U, \mathcal{A}) \le C_1 \mathrm{e}^{-\alpha T} + \mathrm{e}^{KT} \Delta t^r \quad \text{for} \quad n, \Delta t : T \le n\Delta t \le 2T. \tag{6.40}$$

We can balance the contributions in (6.40) to find the relationship between T and Δt which optimizes the error. We find that

$$\Delta t \propto \mathrm{e}^{-(\alpha+K)T/r}$$

is the appropriate choice. This shows that there exists $C_2 > 0$ such that, for any $U \in B$,

$$\mathrm{dist}(S^n_{\Delta t} U, \mathcal{A}) \le C_2 \Delta t^\beta \quad \text{for} \quad n, \Delta t : T \le n\Delta t \le 2T.$$

Proceeding with an induction argument as in Theorem 6.20 we obtain the required result. □

We can use the construction of Theorem 6.20 to prove a result closely related to Theorem 6.12.

Theorem 6.22 **(Uniformly asymptotically stable sets under approximation)** Consider the semigroup $S(t)$ for (1.1) under (6.3) with attractor $\mathcal{A}$. Then there exists $\Delta t_c > 0$ such that, for all $\Delta t \in (0, \Delta t_c]$, the approximating semigroup $S^n_{\Delta t}$ for (1.2) has a compact, uniformly asymptotically stable set $\Lambda_{\Delta t} \supseteq \mathcal{A}$ which satisfies

$$d_{\mathrm{H}}(\Lambda_{\Delta t}, \mathcal{A}) \to 0 \quad \text{as} \quad \Delta t \to 0.$$

Proof. Let $\Delta t \in (0, \Delta t_c]$ where Δt_c is given by Theorem 6.20. To prove this result define

$$\Lambda_{\Delta t} = \mathcal{A}_{\Delta t} \cup \bigcup_{n=0}^{\infty} S^n_{\Delta t} \mathcal{A}$$

where $\mathcal{A} = \omega(B)$ and, by Theorem 6.19, $\mathcal{A}_{\Delta t} = \omega(B)$ (the limit sets being defined through $S(t)$ and $S^n_{\Delta t}$ respectively.) Note that, by Theorem 6.19, both $\mathcal{A}$ and $\mathcal{A}_{\Delta t}$ are contained in B so that $\Lambda_{\Delta t} \subseteq B$. Since $\Lambda_{\Delta t}$ contains $\mathcal{A}_{\Delta t}$ which is an attractor for $S^n_{\Delta t}$, it follows that $\Lambda_{\Delta t}$ is asymptotically stable for $S^n_{\Delta t}$. Note also that $\Lambda_{\Delta t}$ is positively invariant under $S^n_{\Delta t}$ since $\mathcal{A}_{\Delta t}$ is invariant by Theorem 3.3 and $\cup_{n=0}^{\infty} S^n_{\Delta t}\mathcal{A}$ is positively invariant by construction; thus it follows by using a similar argument to that used in establishing Theorem 6.4(ii), that $\Lambda_{\Delta t}$ is uniformly stable.

It remains to establish the error bound. Since $\Lambda_{\Delta t} \supseteq \mathcal{A}$ we have

$$\operatorname{dist}(\mathcal{A}, \Lambda_{\Delta t}) = 0 \quad \forall \Delta t \in (0, \Delta t_c].$$

Note that, since $\mathcal{A} \in B$ we have from (6.38) that, for any $\epsilon > 0$, there exists $T = T(\epsilon) > 0$ such that

$$\operatorname{dist}(S^n_{\Delta t}\mathcal{A}, \mathcal{A}) < \epsilon \quad \forall n, \Delta t : T \leq n\Delta t < \infty.$$

Furthermore, since $S(t)\mathcal{A} = \mathcal{A}$ it follows from (6.35) that there exists $\Delta = \Delta(\epsilon) > 0$ such that, if $\Delta t \in (0, \Delta]$ then

$$\operatorname{dist}(S^n_{\Delta t}\mathcal{A}, \mathcal{A}) \leq \tfrac{1}{2}\epsilon \quad \forall n, \Delta t : 0 \leq n\Delta t \leq 2T.$$

Also Theorem 6.20 yields

$$\operatorname{dist}(\mathcal{A}_{\Delta t}, \mathcal{A}) < \epsilon$$

for $\Delta < \Delta(\epsilon)$. Combining these estimates shows that

$$\operatorname{dist}(\Lambda_{\Delta t}, \mathcal{A}) < \epsilon$$

for $\Delta < \Delta(\epsilon)$ as required. □

6.4. Lower semicontinuity of attractors

As we have seen, lower semicontinuity results are not true in general. However, if we make assumptions about the nature of the flow on the attractor $\mathcal{A}$ then it is possible to prove lower semicontinuity with respect to numerical perturbations. One important case where this is possible is when the dynamical system $S(t)$ is in gradient form and the set $\mathcal{E}$ is a bounded set containing only hyperbolic equilibria. We assume this henceforth. The method of proof is to decompose the attractor $\mathcal{A}$ according to the value of $F(\cdot)$ and build up the nearby approximate attractor $\mathcal{A}_{\Delta t}$ starting from the smallest value of F on the attractor.

We make the following assumption throughout this section:

Assumption 6.23 The dynamical system (1.1) has vector field f given by (6.5) and is a gradient system; furthermore, the set $\mathcal{E}$ of equilibrium points is bounded and comprises only hyperbolic equilibria.

Thus we may enumerate the set of equilibrium points of $S(t)$ as

$$\mathcal{E} = \{x_1, \ldots, x_M\}. \tag{6.41}$$

Let

$$v_1 > v_2 > \ldots > v_N$$

be the distinct points of $\{F(x_1), \ldots, F(x_M)\}$. Since $\mathcal{E}$ is bounded the assumptions of Theorem 6.10 hold. Thus the set $B = B(R)$ given in that theorem is uniformly asymptotically stable for any $R > \xi$. Hence we may define

$$E^k = \{x \in \mathcal{E}: F(x) = v_k\}, \quad U^k = \{x \in B: F(x) < v_k\},$$
$$W^k = \bigcup_{x\in E^k} W^u(x), \quad \mathcal{A}^k = \bigcup_{j=k}^N W^k. \tag{6.42}$$

We will require the following lemma concerning these sets:

Lemma 6.24 Consider a dynamical system (1.1) under Assumption 6.23. Then $S(t) : U^k \mapsto U^k$, $k = 1, \ldots, N$ and $\mathcal{A}^k$ attracts all compact sets in U^{k-1}.

Proof. See Hale (1988), Theorem 3.8.7. □

Since $\mathcal{A}^1$ contains the unstable manifolds of all equilibria, it follows from Theorem 6.11, noting that all equilibria are assumed hyperbolic, that $\mathcal{A}^1 = \mathcal{A}$. On the other hand, since all points in E^N must be stable as there are no equilibria with lower values of F than those in E^N and (6.6) holds, we deduce that $\mathcal{A}^N = E^N$. It is straightforward to show that E^N is close to its discrete counterpart; we use this as the basis of an inductive proof to build up the properties of the approximate attractor. We require the following lemma – recall the definition $\mathcal{E}_{\Delta t}$ of the fixed points of $S^1_{\Delta t}$.

Lemma 6.25 Under Assumption 6.23 there exist $C, \Delta t_c > 0$ such that, for all $\Delta t \in (0, \Delta t_c]$, the semigroup $S^1_{\Delta t}1$ has M fixed points, $X_j \in \mathcal{E}_{\Delta t} \subset B$ $j = 1, \ldots, M$, all of which are hyperbolic and satisfy $\|x_j - X_j\| \leq C\Delta t^r$. Furthermore, for any $\epsilon > 0$ there exists $\epsilon' > 0$ such that

$$\text{dist}(W^{u,\epsilon}(x_j), W^{u,\epsilon'}_{\Delta t}(X_j)) \leq C\Delta t^r.$$

Proof. Let Δt_c be given by the minimum Δt_c found in Theorems 4.10 and 4.11 and Corollary 4.13. By Theorem 4.10 the existence and closeness of the approximate fixed points follows and, by Theorem 4.11 we deduce that no others exist. Corollary 4.13 gives the required bound for the local unstable manifolds. □

Thus we may set

$$\mathcal{E}_{\Delta t} = \{X_1, \ldots, X_M\}. \tag{6.43}$$

and let

$$V_1 > V_2 > \ldots > V_N$$

be the distinct value of $F(\cdot)$ on the members of $\mathcal{E}_{\Delta t}$. Definitions analogous to (6.42) can be made for the dynamical system generated by (1.2). Thus we define

$$E^k_{\Delta t} = \{x \in \mathcal{E}_{\Delta t}: F(x) = V_k\}\text{and}U^k_{\Delta t} = \{x \in B: F(x) < V_k\},$$
$$W^k_{\Delta t} = \bigcup_{x\in E^k_{\Delta t}} W^u_{\Delta t}(x)\text{and}\mathcal{A}^k_{\Delta t} = \bigcup_{j=k}^N W^k_{\Delta t}. \tag{6.44}$$

Notice that since the global attractor must include the union of all unstable manifolds of fixed points by Theorem 6.11, we have

$$\mathcal{A}^k_{\Delta t} \subseteq \mathcal{A}_{\Delta t}, k = 1, \dots, N \tag{6.45}$$

and this is the only property required of the $\mathcal{A}^k_{\Delta t}$. We now use the decomposition to prove lower semicontinuity for the numerical method.

Theorem 6.26 **(Lower semicontinuity of attractors)** Consider (1.1) under Assumption 6.23 with attractor $\mathcal{A}$. Then there exists $\Delta t_c > 0$ such that for $\Delta t \in (0, \Delta t_c]$ the numerical solution (1.2) possesses an attractor $\mathcal{A}_{\Delta t}$ which satisfies

$$d_{\mathrm{H}}(\mathcal{A}_{\Delta t}, \mathcal{A}) \to 0 \qquad \text{as} \qquad \Delta t \to 0.$$

We postpone the proof of the theorem until after the following lemma, which is fundamental in the proof.

Lemma 6.27 Consider (1.1) under Assumption 6.23. Assume that there exists Δ_k such that

$$\operatorname{dist}(\mathcal{A}^k, \mathcal{A}^k_{\Delta t}) \leq \frac{\epsilon}{2^k} \quad \forall \Delta t \in (0, \Delta_k]. \tag{6.46}$$

Then there exists Δ_{k-1} such that

$$\operatorname{dist}(\mathcal{A}^{k-1}, \mathcal{A}^{k-1}_{\Delta t}) \leq \frac{\epsilon}{2^{k-1}} \quad \forall \Delta t \in (0, \Delta_{k-1}]. \tag{6.47}$$

Proof. In the proof it is useful to observe from (6.42) and (6.44) that

$$\begin{aligned} \mathcal{A}^{k-1} &= W^{k-1} \cup \mathcal{A}^k \\ \mathcal{A}^{k-1}_{\Delta t} &= W^{k-1}_{\Delta t} \cup \mathcal{A}^k_{\Delta t}. \end{aligned} \tag{6.48}$$

Suppose that (6.46) holds. Now, since $\mathcal{A}^{k-1}_{\Delta t} \supseteq \mathcal{A}^k_{\Delta t}$ by (6.48), it follows that

$$\operatorname{dist}(\mathcal{A}^k, \mathcal{A}^{k-1}_{\Delta t}) \leq \operatorname{dist}(\mathcal{A}^k, \mathcal{A}^k_{\Delta t}) = \frac{\epsilon}{2^k}.$$

Furthermore, by (6.48) we deduce that

$$\operatorname{dist}(\mathcal{A}^{k-1}, \mathcal{A}^{k-1}_{\Delta t}) = \max(\operatorname{dist}(W^{k-1}, \mathcal{A}^{k-1}_{\Delta t}), \operatorname{dist}(\mathcal{A}^k, \mathcal{A}^{k-1}_{\Delta t})).$$

Hence, to establish (6.47), it is sufficient to show that

$$\operatorname{dist}(W^{k-1}, \mathcal{A}^{k-1}_{\Delta t}) \leq \frac{\epsilon}{2^{k-1}}. \tag{6.49}$$

Recall the notation (3.3) and let

$$\Gamma^{k-1} = \bigcup_{x \in E^{k-1}} \Big(W^{u,\delta}(x) \cap \partial B(x, \delta) \Big),$$

for some $\delta > 0$. The set Γ^{k-1} is compact and, by Lemma 4.2, we have

$$W^{k-1} = \bigcup_{x \in E^{k-1}} \left\{ W^{u,\delta}(x) \right\} \cup \bigcup_{t>0} S(t)\Gamma^{k-1}. \tag{6.50}$$

To establish that (6.49) holds we consider three separate cases corresponding to a breakdown of W^{k-1} into three different subsets.

(a) Note that $\Gamma^{k-1} \subset U^{k-1}$ by (6.6). By Lemma 6.24 $\mathcal{A}^k$ attracts all compact subsets of U^{k-1} and so there exists t_{k-1} such that

$$\operatorname{dist}(S(t)\Gamma^{k-1}, \mathcal{A}^k) \le \frac{\epsilon}{2^k} \quad \forall t \ge t_{k-1}.$$

But by the inductive hypothesis (6.46) and (6.48) we have

$$\begin{aligned} \operatorname{dist}\Big(\bigcup_{t \ge t_{k-1}} S(t)\Gamma^{k-1}, \mathcal{A}^{k-1}_{\Delta t} \Big) &\le \operatorname{dist}\Big(\bigcup_{t \ge t_{k-1}} S(t)\Gamma^{k-1}, \mathcal{A}^{k}_{\Delta t} \Big) \\ &\le \operatorname{dist}\Big(\bigcup_{t \ge t_{k-1}} S(t)\Gamma^{k-1}, \mathcal{A}^{k} \Big) + \operatorname{dist}(\mathcal{A}^k, \mathcal{A}^k_{\Delta t}) \qquad (6.51) \\ &\le 2\epsilon/2^k = \epsilon/2^{k-1}. \end{aligned}$$

(b) Recall the time t_{k-1} given in (a), the constant $\Delta_k > 0$ from (6.46) and the constant K from Assumption 3.7.

Now, given $x \in E^{k-1}$, let $X \in E^{k-1}_{\Delta t}$ be the approximate fixed point given by Lemma 6.25. By Lemma 6.25 it follows that there exist $\Delta^1 > 0$ and $\delta' > 0$ such that

$$\operatorname{dist}(W^u_\delta(x), W^u_{\delta',\Delta t}(X)) < \frac{\epsilon}{2^k e^{K(t_{k-1}+\Delta t_k)}} < \frac{\epsilon}{2^{k-1}}, \quad \forall \Delta t \in (0, \Delta^1]. \tag{6.52}$$

Hence

$$\operatorname{dist}(W^u_\delta(x), \mathcal{A}^{k-1}_{\Delta t}) < \frac{\epsilon}{2^{k-1}}, \tag{6.53}$$

since $W^u_{\delta',\Delta t}(x) \subseteq \mathcal{A}^{k-1}_{\Delta t}$.

(c) Now we show that $\operatorname{dist}(S(t)\Gamma^{k-1}, \mathcal{A}^{k-1}_{\Delta t}) \le \epsilon/2^{k-1}$ for $t \in (0, t_{k-1}]$. By Theorem 4.11 we can choose $\Delta t^2 > 0$ such that for $\Delta t \in (0, \Delta t^2]$ and any $v, w \in \mathbb{R}^p$ satisfying

$$\|v - w\| \le \frac{\epsilon}{2^k e^{K(t_{k-1}+\Delta_k)}} \tag{6.54}$$

we have

$$\|S(n\Delta t)v - S^n_{\Delta t} w\| \le \epsilon/2^{k-1} \quad \text{for} \quad n\Delta t \le t_{k-1} + \Delta_k. \tag{6.55}$$

Now let $\Delta_{k-1} = \min(\Delta t^1, \Delta t^2, \Delta_k)$. Suppose $\Delta t < \Delta_{k-1}$ and that $u \in S(t)\Gamma^{k-1}$ for $t \in (0, t_{k-1}]$. Then, by Lemma 4.2, there exists v so that

$$v \in \overline{B}(x, \delta) \cap W^u(x)$$

for some $x \in E^{k-1}$, such that $S(n\Delta t)v = u$ and $n\Delta t \in [0, t_{k-1} + \Delta t_k]$. By (6.52) there exists $w \in W^u_{\Delta t}(X)$ such that (6.54) holds. Now, by (6.55),

$$\|u - S^n_{\Delta t} w\| = \|S(n\Delta t)v - S^n_{\Delta t} w\| \leq \epsilon/2^{k-1} \quad \forall \Delta t \in (0, \Delta_{k-1}].$$

Since the unstable manifold is invariant by Theorem 4.2 and is contained in the attractor by Theorem 6.11, we have $S^n_{\Delta t} w \in \mathcal{A}^{k-1}_{\Delta t}$; hence it follows that

$$\operatorname{dist}(u, \mathcal{A}^{k-1}_{\Delta t}) \leq \epsilon/2^{k-1}.$$

But u is an arbitrary point in $\bigcup_{t \in (0, t_{k-1}]} S(t)\Gamma^{k-1}$ and hence

$$\operatorname{dist}\left(\bigcup_{t \in (0, t_{k-1}]} S(t)\Gamma^{k-1}, \mathcal{A}^{k-1}_{\Delta t}\right) \leq \frac{\epsilon}{2^{k-1}}. \tag{6.56}$$

By the definition (6.50) of W^{k-1}, the estimates (6.51), (6.53) and (6.56) together establish that

$$\operatorname{dist}(W^{k-1}, \mathcal{A}^{k-1}_{\Delta t}) \leq \epsilon/2^{k-1}$$

and complete the proof of the lemma. □

Proof of Theorem 6.26. First note that, since $S(t)$ has an attractor $\mathcal{A}$, it follows from Corollary 6.18 that $S^1_{\Delta t}$ has a nearby uniformly asymptotically stable set $\Lambda_{\Delta t}$ and an attractor $\mathcal{A}_{\Delta t} = \omega(\Lambda_{\Delta t})$ satisfying

$$\operatorname{dist}(\mathcal{A}_{\Delta t}, \mathcal{A}) \to 0 \quad \text{as} \quad \Delta t \to 0.$$

Thus it remains to establish the lower semicontinuity result that

$$\operatorname{dist}(\mathcal{A}, \mathcal{A}_{\Delta t}) \to 0 \quad \text{as} \quad \Delta t \to 0.$$

It is sufficient to prove that given any $\epsilon > 0$ there exists $\Delta = \Delta(\epsilon)$ such that if $\Delta t < \Delta$ then $\operatorname{dist}(\mathcal{A}, \mathcal{A}_{\Delta t}) \leq \epsilon$.

Recall the notation and decomposition of $\mathcal{A}$ and $\mathcal{A}_{\Delta t}$ given in (6.42) and (6.44). Note that, as described above, $\mathcal{A}^N = E^N$ and $\mathcal{A}^N_{\Delta t} = E^N_{\Delta t}$. Applying Lemma 6.25 we deduce that there exists $\Delta_N > 0$ such that

$$\operatorname{dist}(E^N, E^N_{\Delta t}) = \operatorname{dist}(\mathcal{A}^N, \mathcal{A}^N_{\Delta t}) \leq \epsilon/2^N \quad \forall \Delta t \in (0, \Delta_N].$$

By induction, using Lemma 6.27, we deduce that there exists Δ_k such that, for $k = 1, \ldots, N$

$$\operatorname{dist}(\mathcal{A}^k, \mathcal{A}^k_{\Delta t}) \leq \epsilon/2^k \quad \forall \Delta t \in (0, \Delta_k].$$

In particular, since $\mathcal{A}^1 = \mathcal{A}$ and $\mathcal{A}^1_{\Delta t} \subseteq \mathcal{A}_{\Delta t}$ by (6.45) we deduce that

$$\operatorname{dist}(\mathcal{A}, \mathcal{A}_{\Delta t}) \leq \epsilon \quad \forall \Delta t \in (0, \Delta_1].$$

This completes the proof. □

6.5. Lower semicontinuity of global unstable manifolds

In this section we examine the lower semicontinuity of the global unstable manifolds of (1.1) with respect to numerical perturbation. Recall that we have already studied a related question for the local unstable manifold in Section 4 – see Corollary 4.13. Since the global unstable manifold of a fixed point is necessarily contained in the global attractor by Theorem 6.11, it is natural to study them in the context of attractors. As a corollary we shall obtain a simpler proof of Theorem 6.26 since, for gradient systems with only hyperbolic equilibria, the attractor comprises unstable manifolds of equilibrium points.

Theorem 6.28 (Global unstable manifold under approximation) Assume that (1.1) has an equilibrium point v and that V is the equilibrium point of (1.2) which converges to v as $\Delta t \to 0$. Then if the unstable manifold $W^u(v)$ is bounded it follows that

$$\mathrm{dist}(\overline{W^u(v)}, \overline{W^u_{\Delta t}(V)}) \to 0 \qquad \text{as} \qquad \Delta t \to 0.$$

Proof. It is sufficient to prove that, given any $\epsilon > 0$, there exists $\Delta > 0$ such that for every $y \in W^u(v)$ there exists $Y \in W^u(V)$ with the property that $\|y - Y\| \leq 2\epsilon$ for $\Delta t \in (0, \Delta]$.

Recall $\partial B(v; r)$ and Γ given by (3.3) and (4.7). Now set

$$\mathcal{W} = W^u(v) \backslash W^{u,\epsilon}(v). \tag{6.57}$$

Then, for ϵ sufficiently small,

$$\mathcal{W} = \bigcup_{t>0} S(t)\Gamma.$$

Since $W^u(v)$ is bounded it follows that $\overline{\mathcal{W}}$ is compact. It may be noted that $\{B(x; \epsilon): x \in \mathcal{W}\}$ is an ϵ-cover for $\overline{\mathcal{W}}$ and hence, since $\overline{\mathcal{W}}$ is compact, we may extract a finite subcover. Denote this subcover by $\{B_i(\epsilon)\}_{i=1}^I$ and note that each $B_i(\epsilon)$ contains a point $y_i \in \mathcal{W}$, where $B_i(\epsilon) = B(y_i, \epsilon)$. By construction there exists $x_i \in \Gamma$ and $T_i > 0$ such that $S(T_i)x_i = y_i$ for each $y_i \in \mathcal{W}$. Now, by Corollary 4.13, it follows that there exists $X_i \in W^u(V)$ and $\Delta(i) > 0$ such that

$$\|x_i - X_i\| \leq \epsilon/(2\mathrm{e}^{KT_i}) \quad \forall \Delta t \in (0, \Delta(i)]; \tag{6.58}$$

by the invariance of the unstable manifold (see Lemma 4.2) it follows that $Y_i = S^n_{\Delta t} X_i \in W^u(V)$.

It now follows from Theorems 3.8 and (6.58) that

$$\|y_i - Y_i\| \leq (\mathrm{e}^{KT} - 1)\Delta t^r + \tfrac{1}{2}\epsilon$$

for $\Delta t \in (0, \Delta(i)]$. Thus, by further reduction of $\Delta(i)$ if necessary, we find that

$$\|y_i - Y_i\| \leq \epsilon, \quad \forall \Delta t \in (0, \Delta(i)].$$

Since I is finite, we deduce that there exists $\{Y_i\}_{i=1}^I$ each lying in $W^u(V)$ and $\Delta > 0$ such that

$$\max_{1\le i\le I} \|y_i - Y_i\| \le \epsilon \quad \forall \Delta t \in (0, \Delta].$$

Thus, since y_i is the centre of $B_i(\epsilon)$, we deduce that for every $y \in B_i(\epsilon)$ and $i : 1 \le i \le I$ there exists $Y_i \in W^u(V)$ such that

$$\|y - Y_i\| \le 2\epsilon \quad \forall \Delta t \in (0, \Delta].$$

Since the $B_i(\epsilon), i = 1, \ldots, I$ form a cover of $\overline{\mathcal{W}}$ we deduce that

$$\text{dist}(\overline{\mathcal{W}}, W^u(V)) \le 2\epsilon \quad \forall \Delta t \in (0, \Delta]. \tag{6.59}$$

Now, by Corollary 4.13 there exists $\delta' > 0$ such that

$$\text{dist}(W^{u,\delta}(v), W^{u,\delta'}_{\Delta t}(V)) \le 2\epsilon \quad \forall \Delta t \in (0, H], \tag{6.60}$$

possibly by further reduction of Δ. Putting (6.59) and (6.60) the result follows by (6.57). □

We now use this result to study lower semicontinuity of attractors. We make the following assumption:

Assumption 6.29 The dynamical system (1.1) has a global attractor $\mathcal{A}$ where

$$\mathcal{A} = \bigcup_{x\in\mathcal{E}'} \overline{W^u(x)}$$

and $\mathcal{E}'$ comprises a finite number of hyperbolic equilibrium points of (1.1).

Note that this assumption is a consequence of Assumption 6.23 but that Assumption 6.29 is weaker. For example, Assumption 6.29 admits the equations (3.12) with global attractor given by the disc $x^2 + y^2 \le 1$. Under Assumption 6.29 we may prove lower semicontinuity of the attractor, yielding a simpler proof of Theorem 6.26.

Corollary 6.30 **(Lower semicontinuity of attractors)** Consider (1.1) under Assumption 6.29. Then there exists $\Delta > 0$ such that for $\Delta t \in (0, \Delta]$ the numerical solution (1.2) possesses an attractor $\mathcal{A}_{\Delta t}$ which satisfies

$$d_{\text{H}}(\mathcal{A}_{\Delta t}, \mathcal{A}) \to 0 \qquad \text{as} \qquad \Delta t \to 0.$$

Proof. It follows from Corollary 6.18 that there exists an approximate attractor $\mathcal{A}_{\Delta t}$ satisfying

$$\text{dist}(\mathcal{A}_{\Delta t}, \mathcal{A}) \to 0 \qquad \text{as} \qquad \Delta t \to 0.$$

Thus it remains to establish the lower semicontinuity result that

$$\text{dist}(\mathcal{A}, \mathcal{A}_{\Delta t}) \to 0 \qquad \text{as} \qquad \Delta t \to 0. \tag{6.61}$$

Let $v \in \mathcal{E}'$. By Theorem 6.28 we deduce that there exists a fixed point V of $S^1_{\Delta t}$ such that

$$\text{dist}(\overline{W^u(v)}, \overline{W^u_{\Delta t}(V)}) \to 0 \qquad \text{as} \qquad \Delta t \to 0.$$

But, by Theorem 6.11, $\overline{W^u_{\Delta t}(V)} \subseteq \mathcal{A}_{\Delta t}$ and hence it follows that

$$\text{dist}(\overline{W^u(v)}, \mathcal{A}_{\Delta t}) \to 0 \qquad \text{as} \qquad \Delta t \to 0.$$

Since Assumption 6.29 holds it is clear that (6.61) follows and the proof is complete. □

6.6. *Bibliography*

The material in Section 6.1 can be found in several books, notably Hale (1988), Bhatia and Szego (1970), Hirsch and Smale (1974) and Yoshizawa (1966). For Theorem 6.2 and related results see Hale (1988). For Theorem 6.5 see Yoshizawa (1966); Bhatia and Szego (1970) also contains similar results on converse theorems for Lyapunov functions. The importance of gradient systems is that they enable an explicit decomposition of the dynamics into equilibrium points together with a Lyapunov function decreasing along all trajectories connecting equilibrium points at $t = \pm\infty$; this idea can be generalized to allow the equilibrium points to be replaced by more general limit sets. For results concerning gradient systems see Hirsch and Smale (1974) and Hale (1988).

The first article concerning a detailed analysis of the effect of numerical approximation on sets possessing some general form of attractivity is Kloeden and Lorenz (1986). They essentially proved Theorem 6.12, which concerns upper and lower semicontinuity of uniformly asymptotically stable sets Corollary 6.18, concerning upper semicontinuity of attractors, is a consequence of their work. The approach of Kloden and Lorenz, using Lyapunov functions, was generalized to partial differential equations in Kloeden and Lorenz (1989); it is extended to multistep methods in (Kloeden and Lorenz, 1990). Theorem 6.20 is due to Hale and Raugel (1989) although their result is more general, concerning an arbitrary attractor in a Banach space. A similar result may also be found in Temam (1988). As can be seen by comparing the work required to prove Corollary 6.18 (via Theorem 6.12) and Theorem 6.20, the approach of Hale *et al.* (1988) is considerably shorter than that of Kloeden and Lorenz (1986) if interest is focused only on attractors. Furthermore, as pointed out in Hill and Suli (1993), results strongly related to those of Kloeden and Lorenz (1986) concerning uniformly asymptotically stable sets can be deduced from the approach of Hale *et al.* (1988) – this is then proved in Theorem 6.22. Theorem 6.21, concerning the rate of convergence of the attractor when it is exponentially attracting, has not appeared in the literature; however, the basic idea for the proof is contained in the study of exponential attractors in Babin and Vishik (1992). The issues concerning the derivation of upper semicontinuity results for partial differential equations are considerably more subtle than for ordinary differential equations since, frequently, the spaces in which the attractors lie are not sufficiently regular to apply standard smooth initial data error bounds; Larsson (1989) contains a self-contained and clear presentation of this issue in the context of the finite-element approximation of the reaction-diffusion equation and that work is generalized to cover the Cahn–Hilliard equation in Elliott and Larsson (1992). In Yin-Yan (1993) similar issues are considered for finite difference approximations of the Navier–Stokes equation and in Lord and Stuart (1994) for finite difference approximations of the Ginzburg–Landau equation. The whole question of the existence of global attractors under approximation is reviewed in Humphries *et al.* (1994).

Section 6.4 contains an exposition of the work of Hale and Raugel (1989) concerning the lower semicontinuity of attractors for gradient systems. The presentation given here is closely related to that given in Humphries and Stuart (1994) where Runge–Kutta methods are studied in this context. As can be seen, the proof is not at all straightforward and for this reason more accessible proofs have been sought. An alternative, more accessible approach is described in Section 6.5, culminating in Corollary 6.30. This approach is due to Humphries (1994); it is interesting to note that, whilst a more general class of problems is considered in Section 6.5 than in Section 6.6, Humphries' proof of lower semicontinuity is more straightforward than that of Hale and Raugel – compare Theorems 6.26 and Corollary 6.30. Furthermore, the approach of Humphries also yields further information about the global unstable manifolds and can be trivially modifed to obtain upper semicontinuity of global unstable manifolds. The approach of Humphries is extended to partial differential equations in Humphries *et al.* (1994).

Note that we have described a variety of results in this section some of which supersede others, at least on a superficial level. However, since it is not clear in which direction these results can be generalized, we feel it worthwhile to document in detail the various approaches to these problems.

7. Conclusions

In this article we have concentrated on the convergence of limit sets and invariant sets of a time continuous semigroup, under numerical approximation of the evolution semigroup by a time discrete semigroup. It should be clear that a fairly full picture of this subject has now been developed and that, furthermore, there are a variety of approaches to some of the questions studied. It is natural to ask at this point what future directions are likely to be of scientific interest in this subject area. We give a purely subjective answer by describing areas likely to be fruitful for future development.

7.1. Convergence of attractors

As the work of Section 6 shows, it is not in general possible to prove lower semicontinuity of attractors. In practice this means that numerical computations may 'miss' part of the true attractor. The only situations in which it is currently possible to prove both upper and lower semicontinuity are those in which something is known about the flow on the attractor – specifically, lower semicontinuity has been proved for certain hyperbolic gradient systems (Section 6.4) and for systems whose attractors are the union of the closure of unstable manifolds of equilibria (see Section 6.5). The important point about the assumptions made in Sections 6.4 and 6.5 is that they amount to a form of *hyperbolicity* of the flow on the attractor. There are two directions that the study of attractors can be taken.

The first is to investigate further the hyperbolicity conditions on the flow on the attractor which yield lower semicontinuity results. However, since such questions are by no means fully understood even in the context of smooth perturbations of the vector field in (1.1), this question appears quite difficult. To pursue this avenue will require a parallel development of the general theory of structural stability of attractors. The work of Pliss and Sell (1991) is of interest in this context.

The second is to weaken the concept of attractor to obtain an object which is, for example, exponentially attractive and retains favourable properties under perturbation. In a sense this is what the concept of uniformly asymptotically stable sets (see Section 6.2) does. However, other generalizations are possible. The inertial manifold is an enlargement of the global attractor to obtain an exponentially attracting object and it is possible to prove both upper and lower semicontinuity for the inertial manifold – see Foais *et al.* (1988), Demengel and Ghidaglia (1989) and Jones and Stuart (1993). If this work could be combined with, for example, the approach of Pliss and Sell (1991) it might be possible to make useful deductions about the relationship between the flows on the true and numerical attractors. The concept of *intertial sets* is of also interest (see Eden *et al.* (1990)). The inertial set is an enlargment of the attractor to a positively invariant set which is contained in an inertial manifold and is exponentially attracting. It is plausible that this object is both upper and lower semicontinuous with respect to numerical perturbations of the semigroup.

7.2. Shadowing

We have not explicitly described the subject of shadowing at all in this article although it will almost certainly play an increasingly important role in making statements about the meaning of long-time computations. The general area of shadowing is enormous and there is not room in this article to do it justice. We briefly mention some existing literature in this area.

Seminal work in this area can be found in Hammel *et al.* (1987; 1988) where the effect of round-off error is studied on the computer iteration of certain chaotic maps, such as the quadratic map and the Henon map. Subsequently this work was generalized to consider the effect of numerical approximation and the following references in the area are representative of this growing field: Chow and Palmer (1990a,b), Chow and Van-Vleck (1993), Corless (1992), Corless and Pilyugin (1993) and Sauer and Yorke (1991).

The concept of shadowing is closely related to the notion of *backward error analysis*, familiar to numerical analysts. The approach of Beyn (1987b) to the approximation of phase portraits is a form of backward error analysis and the idea has been taken further in Eirola (1993), Corless and Corliss (1991) and Elliott and Stuart (1994).

7.3. Direct numerical approximation of invariant sets

In this article we have been mainly concerned with numerical approximation of the semigroup $S(t)$ generated by (1.1). This approach, where the invariant sets of the differential equation are observed indirectly as corresponding invariant sets in the numerical method, is sometimes termed the *indirect approach* (Beyn, 1992). An alternative is the *direct approach* where a numerical method is constructed to compute a given invariant set directly. To do this it is necessary to set-up defining equations, typically a boundary value problem, for the invariant set of interest. This is an area in which there is some existing work but in which there is much room for further development, especially as increases in computational power mean that computations previously prohibitively expensive, in comparison with standard indirect simulations, are now straightforward. We briefly describe some of the existing literature.

The simplest invariant objects (1.1) are, of course, equilibrium solutions and much literature exists concerning solution of equilibrium problem exists; indeed, the development of the subject is such that excellent packages now exist – for example the package PITCON; see Rheinboldt (1986). It is also true in the case of periodic solutions arising from Hopf bifurcations that excellent packages exist – see for example the package AUTO described in Doedel and Kervenez (1986). The computation of quasi-periodic solutions (invariant tori) has not yet evolved to the extent where automatic software is available. However, considerable advances have been made and the following references summarise the existing literature: Aronson *et al.* (1987), Van Veldhuizen (1988), Dieci *et al.* (1991) and Dieci and Lorenz (1993). The article by Moore (1993) contains a unified treatment of computational techniques for periodic solutions and invariant tori.

It appears that it is not possible to formulate the question of the existence of a general compact invariant set as a boundary value problem. However, there are a number of direct computational methods of relevance to the study of general, possibly strange, invariant sets. In Beyn (1990) the concept of computing *connecting orbits* directly was introduced and analysed. Connecting orbits are solutions of (1.1) which connect together two limit sets as $t \to \pm\infty$; they are of importance in understanding dynamical systems in many different contexts, including the existence of chaos. The work of Beyn has been taken further in Bai *et al.* (1993), Moore (1993) and Liu *et al.* (1993). Connecting orbits may be viewed simply as the intersection of stable and unstable manifolds; the direct computation of stable and unstable manifolds is considered in Homburg *et al.* (1993) and Hubert (1993). In many circumstances the charactersitics of a strange attractor can be well understood by investigating the Lyapunov exponents; an article containing recent developments in this area, together with a survey of the existing literature, is Dieci *et al.* (1993).

7.4. Generalization to partial differential equations

Much of the numerical analysis described in Sections 3, 4 and 5 has only been fully developed for ordinary differential equations and there are many interesting remaining questions concerning extensions to partial differential equations. Some of these have been addressed for specific equations (typically in reaction-diffusion) and specific methods.

Acknowledgements This work was supported by the Office of Naval Research, contract number N00014-92-J-1876 and by the National Science Foundation, contract number DMS-9201727. I am greatly indebted to the people who helped in the task of checking this article: Fengshan Bai, Chris Budd, Adrian Hill, Arieh Iserles, Yunkang Liu, Gerald Moore and Alastair Spence.

REFERENCES

F. Alouges and A. Debussche (1991), 'On the qualitative behaviour of the orbits of a parabolic partial differential equation and its discretization in the neighbourhood of a hyperbolic fixed point', *Numer. Funct. Anal. Opt.* **12**, 253–269.

F. Alouges and A. Debussche (1993), 'On the discretization of a partial differential equation in the neighbourhood of a periodic orbit', *Numer. Math.* **65**, 143–175.

D.G. Aronson, E.J. Doedel and H.G. Othmer (1987), 'An analytical and numerical study of the bifurcations in a system of linearly coupled oscillators', *Physica* D, **25**, 20–104.

D.K. Arrowsmith and C.M. Place (1990), *An Introduction to Dynamical Systems*, Cambridge Univeristy Press (Cambridge).

A. Babin and M.I. Vishik (1992), 'Attractors of evolution equations', in *Studies in Mathematics and its Applications*, North-Holland (Amsterdam).

F. Bai, A. Spence, A.M. Stuart (1993), 'The numerical computation of heteroclinic connections in systems of gradient partial differential equations', *SIAM J. Appl. Math.* **53**, 743–769.

W.-J. Beyn (1987a) 'On invariant closed curves for one-step methods', *Numer. Math.* **51**, 103–122.

W.-J. Beyn (1987b) 'On the numerical approximation of phase portraits near stationary points', *SIAM J. Numer. Anal.* **24**, 1095–1113.

W.-J. Beyn (1990), 'The numerical computation of connecting orbits in dynamical systems', *IMA J. Numer. Anal.* **9**, 379–405.

W.-J. Beyn (1992), 'Numerical methods for dynamical systems', in *Numerical Analyis; Proceedings of the SERC Summer School, Lancaster, 1990* (W.A. Light, ed.), Oxford University Press (Oxford).

W.-J. Beyn and E. Dodel (1981), 'Stability and multiplicity of solutions to discretizations of nonlinear ordinary differential equations', *SIAM J. Sci. Stat. Comput.* **2**, 107–120.

W.-J. Beyn and J. Lorenz (1987), 'Center manifolds of dynamical systems under discretization', *Numer. Func. Anal. Opt.* **9**, 381–414.

N.P. Bhatia and G.P. Szego (1970), *Stability Theory of Dynamical Systems*, Springer (New York).

M. Braun and J. Hershenov (1977), 'Periodic solution of finite difference equations', *Quart. Appl. Maths* **35**, 139–147.

F. Brezzi, S. Ushiki and H. Fujii (1984), 'Real and ghost bifurcation dynamics in difference schemes for ordinary differential equations', in *Numerical Methods for Bifurcation Problems* (T. Kupper, H.D. Mittleman and H. Weber, eds), Birkhauser (Boston).

D. Broomhead and A. Iserles (1992), *Proceedings of the IMA Conference on the Dynamics of Numerics and the Numerics of Dynamics, 1990*, Oxford University Press (Oxford).

C.J. Budd (1990), 'The dynamics of numerics and the numerics of dynamics', Bristol Univerisity Applied Mathematics Report AM-90-13.

C.J. Budd (1991), 'IMA Conference on the dynamics of numerics and the numerics of dynamics. Report of the meeting', *Bull. IMA* **27**, 56–58.

K. Burrage and J. Butcher (1979), 'Stability criteria for implicit Runge–Kutta processes,' *SIAM J. Numer. Anal.* **16**, 46–57.

J.C. Butcher (1975), 'A stability property of implicit Runge–Kutta methods', *BIT* **15**, 358–361.

J.C. Butcher (1987), *The Numerical Analysis of Ordinary Differential Equations: Runge-Kutta Methods and General Linear Methods*, Wiley (Chichester).

M.P. Calvo and J.M. Sanz-Serna (1992), 'Variable steps for symplectic integrators', in *Numerical Analysis, 1991* (D.F. Griffiths and G.A. Watson, eds), Longman (London).

M.P. Calvo and J.M. Sanz-Serna (1993a), 'Reasons for a failure. The integration of the two-body problem with a symplectic Runge–Kutta–Nystrom code with stepchanging facilities', in *Equadiff-91* (C. Perello, C. Simo and J. de Sola-Marales, eds), World Scientific (Singapore).

M.P. Calvo and J.M. Sanz-Serna (1993b), 'The development of variable-step symplectic integrators, with applications to the two-body problem', *SIAM J. Sci. Comput.* **14**, 936–952.

J. Carr (1982), *Applications of Centre Manifold Theory*, Springer (New York).

S.N. Chow and K.J. Palmer (1989), 'The accuracy of numerically computed orbits of dynamical systems', in *Equadiff 1989, Prague*.

S.N. Chow and K.J. Palmer (1990a), 'The accuracy of numerically computed orbits of dynamical systems in $\mathbb{R}^k$', Georgia Inst. Tech. Report CDSNS90–28.

S.N. Chow and K.J. Palmer(1990b), 'On the numerical computation of orbits of dynamical systems: the higher dimensional case', Georgia Inst. Tech. Report CDSNS90–32.

S.N. Chow and E. Van Vleck (1993), 'A shadowing lemma approach to global error analysis for initial value ODEs', to appear in *SIAM J. Sci. Stat. Comput.*

P. Constantin, C. Foias, B. Nicolaenko and R. Temam (1989), *Integral Manifolds and Inertial Manifolds for Dissipative Partial Differential Equations, Applied Mathematical Sciences*, Springer (New York).

R.M. Corless (1992), 'Defect-controlled numerical methods and shadowing for chaotic differential equations', *Physica* D **60**, 323–334.

R.M. Corless and G.F. Corliss (1991), 'Rationale for guaranteed ODE defect control', Argonne Nat. Lab. Preprint MCS-P273–1191.

R.M. Corless and S.Y. Pilyugin (1993), 'Approximate and real trajectories for generic dynamical systems', to appear in *J. Math. Anal. Appl.*

M. Crouziex and J. Rappaz (1990), *Numerical Approximation in Bifurcation Theory*, Masson–Springer (Paris).

G. Dahlquist (1975), 'Error analysis for a class of methods for stiff non-linear initial value problems', in *Numerical Analysis, Dundee 1975* (G.A. Watson, ed.), Springer (New York), 60–74.

G. Dahlquist (1978), 'G-stability is equivalent to A-stability' *BIT* **18**, 384–401.

K. Dekker and J.G. Verwer (1984), *Stability of Runge-Kutta Methods for Stiff Nonlinear Differential Equations*, North-Holland (Amsterdam).

F. Demengel and J.M. Ghidaglia (1989), 'Time-discretization and inertial manifolds', *Math. Mod. Numer. Anal.* **23**, 395-404.

R.L. Devaney (1989), *An Introduction to Chaotic Dynamical Systems*, Addison-Wesley (New York).

L. Dieci and J. Lorenz (1993), 'Computation of invariant tori by the method of characteristics', to be submitted.

L. Dieci, J. Lorenz and R.D. Russell (1991), 'Numerical calculation of invariant tori', *SIAM J. Sci. Stat. Comput.* **12**, 607–647.

L. Dieci, R.D. Russell and E. Van Vleck (1993), 'On the computation of Lyapunov exponents for continuous dynamical systems', submitted to *SIAM J. Numer. Anal.*

H.T. Doan (1985), 'Invariant curves for numerical methods', *Quart. Appl. Math.* **3**, 385–393.

E.J. Doedel and J.P. Kervenez (1986), 'AUTO: Software for continuation of bifurcation problems in ordinary differential equations', Appl. Maths Tech. Rep., California Institute of Technology.

P.G. Drazin (1992), *Nonlinear Dynamics*, Cambridge University Press (Cambridge).

A. Eden, C. Foias, B. Nicolaenko and R. Temam (1990), 'Inertial sets for dissipative evolution equations', IMA Preprint Series, #694.

T. Eirola (1988), 'Invariant curves for one-step methods', *BIT* **28**, 113–122.

T. Eirola (1989), 'Two concepts for numerical periodic solutions of ODEs', *Appl. Math. Comput.* **31**, 121–131.

T. Eirola (1993), 'Aspects of backward error analysis of numerical ODEs', *J. Comput. Appl. Math.* **45**, 65–73.

T. Eirola and O. Nevanlinna (1988), 'What do multistep methods approximate?', *Numer. Math.* **53**, 559–569.

C.M. Elliott and S.Larsson (1992), 'Error estimates with smooth and nonsmooth data for a finite element method for the Cahn–Hilliard equation', *Math. Comput.* **58**, 603–630.

C.M. Elliott and A.M. Stuart (1993), 'Global dynamics of discrete semilinear parabolic equations.' to appear in *SIAM J. Numer. Anal.*

C.M. Elliott and A.M. Stuart (1994), 'Error analysis for gradient groups with applications', in preparation.

B. Enquist (1969), 'On difference equations approximating linear ordinary differential equations,' Report 21, Department of Computer Science, Uppsala University.

N. Fenichel (1971), 'Persistence and smoothness of invariant manifolds for flows', *Indiana Univ. Math. J.* **21**, 193–226.

C. Foias, G. Sell and R. Temam (1988), 'Inertial manifolds for nonlinear evolutionary equations', *J. Diff. Eqns* **73**, 309-353.

B.M. Gacay (1993), 'Discretization and some qualitative properties of ordinary differential equations about equilibria', Budapest University of Technology, Technical Report.

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

J. Guckenheimer and P. Holmes (1983), *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields*, Springer (New York).

E. Hairer, A. Iserles and J.M. Sanz-Serna (1990), 'Equilibria of Runge–Kutta methods', *Numer. Math.* **58**, 243–254.

E. Hairer, S.P. Norsett and G. Wanner (1987), *Solving Ordinary Differential Equations, Parts I & II*, Springer (New York).

J.K. Hale (1969), *Ordinary Differential Equations*, Wiley (New York).

J.K. Hale(1988), *Asymptotic Behaviour of Dissipative Systems*, AMS Mathematical Surveys and Monographs 25, American Mathematical Society (Providence, RI).

J.K. Hale and H. Kŏcak (1991), *Dynamics and Bifurcations*, Springer (New York).

J.K. Hale, L. Magalhaes and W. Oliva (1984), *An Introdcution to Infinite Dimensional Dynamical Systems*, Springer (New-York).

J.K. Hale and G. Raugel (1989), 'Lower semicontinuity of attractors of gradient systems and applications', *Annali di Mat. Pura. Applic.* **CLIV**, 281–326.

J.K. Hale, X.-B. Lin and G. Raugel (1988), 'Upper semicontinuity of attractors for approximations of semigroups and partial differential equations,' *Math. Comput.* **50**, 89–123.

S. Hammel, J.A. Yorke and C. Grebogi (1987), 'Do numerical orbits of chaotic dynamical processes represent true orbits?', *J. Complexity* **3**, 136–145.

S. Hammel, J.A. Yorke and C. Grebogi (1988), 'Numerical orbits of chaotic processes represent true orbits', *Bull. AMS* **19**, 465–470.

P. Hartman (1982), *Ordinary Differential Equations*, Birkhauser (Boston).

D. Henry (1981), *Geometric Theory of Semilinear Parabolic Equations. Lecture Notes in Mathematics*, Springer (New York).

J.G. Heywood and R. Rannacher (1986), 'Finite element approximations of the Navier-Stokes problem. Part II: Stability of solutions and error estimates uniform in time,' *SIAM J. Numer. Anal.* **23**, 750–777.

A.J. Homburg, H.M. Osinga and G. Vegter (1993), 'On the numerical computation of invariant manifolds', University of Groningen Technical Report.

A.T. Hill (1994), 'Global dissipativity for multistep methods,' in preparation.

A.T. Hill and E. Suli (1993), 'Upper semicontinuity of attractors for linear multistep methods approximating sectorial evolution equations,' submitted to *Math. Comput.*

A.T. Hill and E. Suli (1994), 'Set convergence for discretizations of the attractor,' submitted to *Numer. Math.*

M.W. Hirsch and S. Smale (1974), *Differential Equations, Dynamical Systems and Linear Algebra*, Academic Press (London).

E. Hubert (1993), 'Computation of stable manifolds', MSc Thesis, Imperial College (London).

A.R. Humphries (1993), 'Spurious solutions of numerical methods for initial value problems', *IMA J. Numer. Anal.* **13**, 262–290.

A.R. Humphries (1994), 'Approximation of attractors and invariant sets by Runge–Kutta methods', in preparation, 1994.

A.R. Humphries and A.M. Stuart (1994), 'Runge–Kutta methods for dissipative and gradient dynamical systems', to appear in *SIAM J. Numer. Anal.*

A.R. Humphries, D.A. Jones and A.M. Stuart (1994), 'Approximation of dissipative partial differential equations over long-time intervals', in *Numerical Analysis, Dundee, 1993* (D.F. Griffiths and G.A. Watson, eds), Longman (New York).

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

A. Iserles, A.T. Peplow and A.M. Stuart (1991) 'A unified approach to spurious soluions introduced by time discretization. Part I: basic theory', *SIAM J. Numer. Anal.* **28**, 1723–1751.

D.A. Jones and A.M. Stuart (1993), 'Attractive invariant manifolds under approximation', submitted to *J. Diff. Eqns*

U. Kirchgraber (1986), 'Multi-step methods are essentially one-step methods', *Numer. Math.* **48**, 85–90.

P. Kloeden and J. Lorenz (1986), 'Stable attracting sets in dynamical systems and their one-step discretizations,' *SIAM J. Numer. Anal.* **23**, 986–995.

P. Kloeden and J. Lorenz (1989), 'Liapunov stability and attractors under discretization', in *Differential Equations, Proceedings of the Equadiff Conference*

(C.M. Dafermos, G. Ladas and G. Papanicolaou, eds), Marcel-Dekker (New York).
P. Kloeden and J. Lorenz (1990), 'A note on multistep methods and attracting sets of dynamical systems', *Num. Math.* **56**, 667–673.
P. Kloeden and K.J. Palmer (1994), 'Chaotic numerics: the approximation and computation of complicated dynamics', Proceedings of the Conference on Chaotic Numerics, Geelong, Australia, July 12th–16th 1993.
O. Ladyzhenskaya (1991), *Attractors for Semigroups and Evolution Equations*, Cambridge University Press (Cambridge).
J. Lambert (1991), *Numerical Methods for Ordinary Differential Equations*, Wiley (Chichester).
S. Larsson (1989), 'The long-time behaviour of finite element approximations of solutions to semilinear parabolic problems', *SIAM J. Numer. Anal.* **26**, 348–365.
S. Larsson (1992), 'Non-smooth data error estimates with applications to the study of long-time behaviour of finite element solutions of semilinear parabolic problems', Preprint, Chalmers University (Sweden).
S. Larsson and J.M. Sanz-Serna (1993), 'The behaviour of finite element solutions of semilinear parabolic problems near stationary points', submitted to *SIAM J. Numer. Anal.*
L. Liu, G. Moore and R.D. Russell (1993), 'Computing connecting orbits between steady solutions', Technical Report, Mathematics Department, Simon Fraser University.
G.J. Lord and A.M. Stuart (1994), 'Existence and convergence of attractors and inertial manifolds for a finite difference approximation of the Ginzburg–Landau equation,' in preparation.
E.N. Lorenz (1963), 'Deterministic noperiodic flow', *J. Atmos. Sci.* **20**, 130–141.
J. Lorenz (1994), 'Convergence of invariant tori under numerical approximation', in Proceedings of the Conference on Chaotic Numerics, Geelong, Australia, July 12th–16th 1993.
J. Mallet-Paret and G.R. Sell (1988), 'Inertial manifolds for reaction-diffusion equations in higher space dimensions', *J. Amer. Math. Soc.* **1**, 805–864.
M. Medved (1991), *Fundamentals of Dynamical Systems and Bifurcation Theory*, Adam-Hilger (Bristol).
A.R. Mitchell and D.F. Griffiths (1986), 'Beyond the linearized stability limit in nonlinear problems', in *Numerical Analysis* (D.F. Griffiths and G.A. Watson, eds), Pitman (Boston).
G. Moore (1993), 'Computation and paramterization of connecting and periodic orbits', submitted to *IMA J. Numer. Anal.*
G. Moore (1993), 'Computation and paramterization of invariant curves and tori', in preparation.
T. Murdoch and C.J. Budd (1990), 'Convergent and spurious solutions of nonlinear elliptic equations', *IMA J. Numer. Anal.* **12**, 365–386.
A.C. Newell (1977), 'Finite amplitude instabilities or partial difference schemes', *SIAM J. Appl. Math.* **33**, 133–160.
V. Pliss (1966), *Nonlocal Problems in the Theory of Oscillations*, Academic Press (New York).

V.A. Pliss and G.R. Sell (1991), 'Perturbations of attractors of differential equations,' *J. Diff. Eqns* **92**, 100-124.

C. Pugh and M. Shub (1988), 'C^r stability of periodic solutions and solution schemes', *Appl. Math. Lett.* **1**, 281–285.

W.C. Rheinboldt (1986), *Numerical Analysis of Parameterised Nonlinear Equations*, Wiley (New York).

J.M. Sanz-Serna (1992a), 'Symplectic integrators for Hamiltonian problems: an overview', *Acta Numerica*, Cambridge University Press (Cambridge), 244–286.

J.M. Sanz-Serna (1992b), 'Numerical ordinary differential equations vs. dynamical systems', in *Proceedings of the IMA Conference on the Dynamics of Numerics and the Numerics of Dynamics, 1990* (D. Broomhead and A. Iserles, eds), Cambridge University Press (Cambridge).

J.M. Sanz-Serna and S.Larsson (1993), 'Shadows, chaos and saddles', to appear in *Appl. Numer. Math.*

J.M. Sanz-Serna and A.M. Stuart (1992), 'A note on uniform in time error estimates for approximations to reaction-diffusion equations', *IMA J. Numer. Anal.* **12**, 457–462.

T. Sauer and J.A. Yorke (1991), 'Rigorous verification of trajectories for the computer simulation of dynamical systems', *Nonlinearity* **4**, 961–979.

A.B. Stephens and G.R. Shubin (1987), 'Multiple solutions and bifurcations of finite difference approximations to some steady problems of fluid mechanics.' *SIAM J. Sci. Stat. Comput.* **2**, 404–415.

H. Stetter (1973), '*Analysis of Discretization Methods for Ordinary Differential Equations*, Springer (New York).

D. Stoffer (1993), 'General linear methods: connection to one-step methods and invariant curves', *Numer. Math.* **64**.

D. Stoffer (1994) 'Averaging for almost identical maps and weakly attractive tori', submitted.

A.M. Stuart (1989a) 'Nonlinear instability in dissipative finite difference schemes', *SIAM Rev.* **31**, 191–220.

A.M. Stuart (1989b), 'Linear instability implies spurious periodic solutions', *IMA J. Numer. Anal.* **9**, 465–486.

A.M. Stuart (1991), 'The global attractor under discretization', in *Continuation and Bifurcations: Numerical Techniques and Applications* (D. Roose, B. De Dier and A. Spence, eds), Kluwer (Dordrecht).

A.M. Stuart and A.R. Humphries (1992a), 'Model problems in numerical stability theory for initial value problems', submitted to *SIAM Rev.*

A.M. Stuart and A.R. Humphries (1992b), 'The essential stability of local error control for dynamical systems', submitted to *SIAM J. Numer. Anal.*

R. Temam (1988), *Infinite Dimensional Dynamical Systems in Mechanics and Physics*, Springer (New York).

E.S. Titi (1991) 'Un critère pour l'approximation des solutions périodiques des équations de Navier–Stokes', *C.R. Acad. Sci. Paris*, **312**, 41–43.

M. Van Veldhuizen (1988), 'Convergence results for invariant curve algorithms', *Math. Comput.* **51**, 677–697.

S. Wiggins (1990), *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, Springer (New York).

H. C. Yee, D. F. Griffiths and P. K. Sweby (1991), 'Dynamical approach to the study of spurious steady-state numerical solutions for nonlinear differential equations, I: The dynamics of time discretization and its implications for algorithm development in CFD', *J. Comput. Phys.* **97**, 249–310.

Yin-Yan (1993), 'Attractors and error estimates for discretizations of incompressible Navier–Stokes equations', submitted to *SIAM J. Numer. Anal.*

T. Yoshizawa (1966), *Stability Theory by Lyapunov's Second Method*, Publ. Math. Soc. Japan (Tokyo).