MATHEMATICS OF COMPUTATION
VOLUME 64, NUMBER 211
JULY 1995, PAGES 1097-1122

UPPER SEMICONTINUITY OF ATTRACTORS FOR LINEAR MULTISTEP METHODS APPROXIMATING SECTORIAL EVOLUTION EQUATIONS

ADRIAN T. HILL AND ENDRE SÜLI

ABSTRACT. This paper sets out a theoretical framework for approximating the attractor $\mathscr{A}$ of a semigroup $S(t)$ defined on a Banach space X by a q-step semidiscretization in time with constant steplength k. Using the one-step theory of Hale, Lin and Raugel, sufficient conditions are established for the existence of a family of attractors $\{\mathscr{A}_k\} \subset X^q$, for the discrete semigroups S_k^n defined by the numerical method. The convergence properties of this family are also considered. Full details of the theory are exemplified in the context of strictly $A(\alpha)$-stable linear multistep approximations of an abstract dissipative sectorial evolution equation.

1. INTRODUCTION

In recent years there has been great interest in the numerical approximation of the asymptotic structures and invariant sets of dynamical systems. For autonomous dissipative systems, the global attractor is characterized as the largest compact set invariant under the action of the evolution operator, and the set on which all the asymptotic dynamics take place. It is therefore of paramount importance to show that numerical methods can approximate the global attractor of such systems.

The multiplicity of behavior possible in a global attractor poses a severe test in approximation, and thus in general it cannot be hoped that an arbitrary numerical method will reflect either qualitatively or quantitatively all the features of the underlying dynamical system. Indeed, even when the numerical method itself possesses an attractor, it may differ in several ways from that of the original system. It is well known that if a single trajectory is approximated over a long time, then the corresponding numerical trajectory can be expected to diverge from the true trajectory. However, if there are qualitative differences between the numerical and true attractors, a numerical trajectory may reproduce behavior entirely without counterpart in the underlying system. Examples of this phenomenon are considered, for instance, in the papers of Griffiths and Mitchell [9] and Stuart [33].

Received by the editor December 2, 1993 and, in revised form, August 3, 1994.
1991 *Mathematics Subject Classification*. Primary 65J05, 58F13, 65L06, 65M12.
Key words and phrases. Numerical analysis in abstract spaces, approximation of attractors, multistep methods, error bounds and convergence of numerical methods.
The work of the first author was completed at the Oxford University Computing Laboratory, Oxford, U.K., and was funded by SERC U.K.

An appreciation of the difficulties identified in the last paragraph has led to efforts, principally in the context of ODEs, to identify particular structures in the attractor, such as fixed points and periodic orbits, and to test whether such structures are preserved by common time discretizations, such as Runge-Kutta and linear multistep methods. This work has been developed by Iserles et al. in the articles [10], [20–22]. The inheritance of stability properties for periodic orbits has also been studied by investigating the circumstances under which trajectories are approximated in a C^1 sense. For the literature in this area, we refer to [1] for one-step methods, and [6] for multistep methods. In the special case of gradient systems whose attractor decomposes into a finite union of hyperbolically intersecting manifolds, Hale and Raugel [15] have shown, under certain practically meaningful conditions, that the entire attractor may be quantitatively approximated.

The investigations outlined above deal with the more fundamental difficulties of asymptotic approximation—that of showing precisely what structure of the attractor is inherited under discretization. Complementary to this work has been a movement to prove weaker but more global results relating to attractor approximation. Kloeden and Lorenz [25, 26] coined the notion of an attracting set, which is compact and asymptotically stable—an object more general than an attractor—and showed in various contexts that these sets are well approximated under discretization. The construction of the approximating attracting sets is simplified and extended to a Banach space setting in [17]. However, the landmark paper in this area was that of Hale, Lin and Raugel [14], which derived a general approximation result for the attractor itself. Below, we outline this result in the context of time discretizations.

Consider a C^0-semigroup $S(t)$, defined on a Banach space X, possessing a local attractor $\mathscr{A}$. (A global attractor is a fortiori a local attractor, see Definition 4.4; we note that other terminology from the theory of dynamical systems is also defined in §4.) Next, take a family of one-step time discretizations of $S(t)$, generating a family of maps $\{S_k\}$ parametrized by the time step k. Then, for all sufficiently small k, under relatively weak conditions on the approximation and compactness properties of the method, there exists a local attractor $\mathscr{A}_k$ for the semigroup S_k^n (the latter is defined by composition of S_k, for $n \in \mathbb{N}$). In addition, the following convergence result was shown. Given $\varepsilon > 0$, there exists $k_0(\varepsilon)$ such that for all $k \in (0, k_0]$, $\mathscr{A}_k$ is within an ε-neighborhood of $\mathscr{A}$. This is otherwise called the **upper semicontinuity** of the sets $\mathscr{A}_k$ with respect to the parameter k at the value $k = 0$.

As remarked in [14], the upper semicontinuity of the family $\{\mathscr{A}_k\}_{k>0}$ does not show that it is a good approximation to $\mathscr{A}$, and we again note that in such generality and under such weak conditions this is not to be hoped for. However, the theory does guarantee that this family will be close to $\mathscr{A}$. This implies that, provided the trajectories of S_k^n remain in a bounded region, their quantitative deviation from some true behavior of the trajectories of $S(t)$ is restricted as k tends to 0. However, on a small scale, the qualitative behavior of the $\{\mathscr{A}_k\}$ may continue to deviate; also whole regions of $\mathscr{A}$ may fail to be approximated in any sense by the attractors of S_k^n.

We conclude that whilst the theory described above does not penetrate into the more thorny problems of attractor approximation, it does provide a unifying structure for other efforts, and yields a good deal of general and practically useful information. This is our motivation for extending the theory of Hale, Lin and Raugel to the case of multistep time discretizations of $S(t)$.

There are several ways in which one might try to do this, and in the case of backward difference approximations to a Galerkin discretization of certain gradient systems, a method has already been described by Elliott and Stuart [7]. In the context of ordinary differential equations, Kirchgraber, Stoffer and others [23, 24, 32] have shown that, given a consistent linear multistep approximation generating a trajectory $\{U^n\}_{n=0}^{\infty}$, there exists a one-step method with trajectory $\{Y^n\}_{n=0}^{\infty}$, and constants $C>0$, $\theta \in (0, 1)$, such that

$$\|U^n - Y^n\| \leq C\theta^n \quad \text{for all } n \geq 0.$$

Here, C and θ are independent of both n and k. Furthermore, all possible trajectories for the one-step method may be generated by the multistep method. Hence, the asymptotic behavior of the multistep method is completely characterized by that of the one-step method. If such a theory were extended to the infinite-dimensional case we study here, then the convergence of the attractors of an asymptotically equivalent one-step method could be considered using the results of [14]. We shall not use either of these approaches here.

Our approach to this problem is to consider the continuous map S_k defined on some subset of the product space X^q, where q is the number of steps in the method. This mapping, originally considered by Butcher [2] and Skeel [31], is given by

$$S_k(U^1, \dots, U^q)^T \equiv (U^2, \dots, U^q, \psi(k; U^1, \dots, U^q))^T, \tag{1.1}$$

where $\psi(k; U^n, \dots, U^{n+q-1})$ is the mapping that generates U^{n+q} from the previous q time steps. Repeated compositions of S_k generate a semigroup S_k^n on X^q. One expects S_k^n to approximate $S(t)$ in some sense, and one also expects that if $S(t)$ possesses an attractor, then so does S_k^n, at least for sufficiently small k. However, problems of commensurability arise because these two semigroups, and therefore their attractors, are defined on different spaces for $q>1$.

Conventionally, when a multistep method is used to approximate a single trajectory $S(t)u_0$, for given u_0, the mapping defined by the numerical method is described by

$$\mathbf{y}^T S_k^n \mathbf{v} : X \to X, \tag{1.2}$$

where $\mathbf{v}: X \to X^q$ is a starting method, and $\mathbf{y}^T: X^q \to X$ is a finishing method. Typically, $\mathbf{v}$ is a complicated function mapping u_0 to an approximation of $[u_0, S(k)u_0, \dots, S((q-1)k)u_0]^T$, whilst $\mathbf{y}^T$ is merely the projection onto the last coordinate. However, whatever $\mathbf{v}$ and $\mathbf{y}^T$ are taken to be, (1.2) cannot represent a semigroup, unless

$$\mathbf{y}^T S_k^n \mathbf{v} \mathbf{y}^T S_k^m \mathbf{v} = \mathbf{y}^T S_k^{n+m} \mathbf{v} \quad \text{for all } n, m \geq 0.$$

In the simple case where S_k is linear and nonsingular, a nonzero relation of this form is possible only when for a certain eigenvalue of S_k, $\mathbf{y}^T$ and $\mathbf{v}$ belong to the respective left and right eigenspaces. The possibilities when S_k is nonlinear are less clear, and we do not pursue this approach.

Rather than modifying S_k^n to act on X, we instead propose to modify $S(t)$ to act on X^q. To do this, we consider $\mathbf{v}: X \to X^q$ and $\mathbf{y}^T: X^q \to X$ such that $\mathbf{y}^T\mathbf{v} = I$. In general, we allow $\mathbf{v}$ and $\mathbf{y}^T$ to depend on any parameter of the two semigroups, including k, but excluding n and t. We observe that

$$\mathbf{v}S(t)\mathbf{y}^T : X^q \to X^q, \quad \mathbf{v}S(t+s)\mathbf{y}^T = \mathbf{v}S(t)\mathbf{y}^T\mathbf{v}S(s)\mathbf{y}^T. \tag{1.3}$$

We note however that this construction has the drawback that $\mathbf{v}S(0)\mathbf{y}^T$ is not the identity on X^q, and hence the family $\{\mathbf{v}S(t)\mathbf{y}^T\}_{t\geq 0}$ is not a conventional semigroup. Nevertheless, it is still a closed associative set with an identity element— $\mathbf{v}\mathbf{y}^T$—and hence, we term it a **monoid** (see Definition 4.3). Furthermore, all the available theory relating to an arbitrary semigroup $\mathcal{S}(t)$ acting on a given Banach space $\mathcal{X}$, which does not need the axiom that $\mathcal{S}(0) = I_{\mathcal{X}}$ applies in an obvious way to the monoid case. Both, attractor existence theory and the attractor approximation theory of [14], are in this category.

If $\mathcal{A}$ is the attractor of $S(t)$, then the attractor of the monoid we have defined in (1.3) is $\mathbf{v}(k)\mathcal{A}$. Our aim is to show that there exist local attractors $\mathcal{A}_k$ for S_k^n in a neighborhood of this set converging to $\mathbf{v}(0)\mathcal{A}$ in Hausdorff semidistance (see Definition 4.2), as $k \to 0$. For small k, one may deduce that $S_k\mathbf{v}(k) \approx \mathbf{v}(k)$, and since S_k maps the second coordinate to the first, and so on, then, provided the method is strictly stable,

$$\lim_{k\to 0} \mathbf{v}(k) = \mathbf{1} \equiv [1, 1, \dots, 1]^T.$$

In fact, for linear multistep methods approximating sectorial evolution equations, we will show that one may choose $\mathbf{v}(k) \equiv \mathbf{1}$, and $\mathbf{y}^T$ to be another constant vector in $\mathbb{R}^q$, which we shall describe later.

Assuming that, for k sufficiently small and $n \geq 0$, S_k^n is well defined in a neighborhood of $\mathbf{1}\mathcal{A}$; that is, the trajectories $\{U^n\}_{n\geq 0}$ of S_k^n always remain in the domain of definition of ψ in (1.1), we may apply the theory of [14] to demonstrate, under appropriate conditions on the compactness of S_k^n and its approximation properties with respect to $\mathbf{1}S(t)\mathbf{y}^T$, that there exist local attractors $\mathcal{A}_k$ for S_k^n converging in Hausdorff semidistance to $\mathbf{1}\mathcal{A}$, as $k \to 0$.

We have outlined above a theoretical framework for the approximation of the attractor using a multistep method. However, there are a number of matters that we have glossed over in the last paragraph, which form the principal difficulty in applications. These are verifying that:

(i) S_k^n is well defined in a neighborhood of $\mathbf{1}\mathcal{A}$;

(ii) S_k^n approximates $\mathbf{1}S(t)\mathbf{y}^T$ uniformly in the sense of [14] (see Definition 4.6);

(iii) S_k^n is asymptotically smooth in the sense of [14] (see Definition 4.5).

It is these matters, and especially establishing (i) and (ii) by a suitable error estimate, that constitute the main work of this article.

We apply our theory in the context of *linear* multistep methods approximating sectorial evolution equations of the form

$$u_t + Au = f(u) \tag{1.4}$$

on X. Here, A is a sectorial operator generating an analytic semigroup e^{-At} on X, and for $\gamma \in [0, 1)$, $f : \mathcal{D}(A^\gamma) \to X$ is a locally Lipschitz function. (Note that $\mathcal{D}(A^\gamma)$ is the domain of a fractional power of A, otherwise viewed as an interpolation space intermediate between X and $\mathcal{D}(A)$, the domain of A.) We remark on the wide applicability of our results since, as is shown by Henry

[16], systems of nonlinear reaction-convection-diffusion equations, including the incompressible Navier-Stokes equation, may be written in the form (1.4). The basic local existence, uniqueness and regularity results for (1.4) are stated in Lemma 3.1. For more detailed information see [16].

The outline of the remainder of the paper is as follows. In §2, error estimates are first obtained for linear multistep approximations of the linear semigroup e^{-At}. The corresponding discrete semigroup is $T^n(kA)$, where $T(z) \in \mathscr{L}(\mathbb{C}^q)$ is the companion matrix for the method, whose domain of stability is required to contain $\sigma(A)$, the spectrum of A. $T^n(kA)$ has previously been considered by Le Roux [27] and Crouzeix [3], and more recently by Lubich and Nevanlinna [28] and Palencia [29]. The principal linear error bound, for the case $\Re e[\sigma(A)] \geq a \geq 0$, is stated in Theorem 1 as

$$(1.5) \qquad \|T^n(kA) - \mathbf{1}e^{-Ank}\mathbf{y}^T\|_{\mathscr{L}(X^q)} \leq C_0 e^{-ank/2}(ak + 1/n),$$

for C_0 independent of a, k and n. Here, $\mathbf{y}^T$ is the left eigenvector of $T(0)$ corresponding to 1, the principal eigenvalue.

It is of interest that this estimate also implies that

$$\|T^n(kA)\| \leq Ce^{-ank/2},$$

which improves on the best bound of which we are aware, obtained in the pioneering work of Le Roux [27].

Section 3 addresses the linear multistep approximation of (1.4). In Theorem 2, the associated mapping S_k^n is shown to be well defined in a ball $B(0, R) \subset \mathscr{D}(A^\gamma)^q$, for sufficiently small k. The main nonlinear error estimate is then established by Theorem 3, in association with a proof that S_k^n is well defined in a ball $B(0, R) \subset \mathscr{D}(A^\gamma)^q$, for $nk \in [0, T]$, and some $T > 0$, provided k is sufficiently small. This last result is somewhat more powerful than condition (i) above, that S_k^n be well defined in a neighborhood of $\mathbf{1}\mathscr{A}$, and implies that the attractors $\mathscr{A}_k$ have as domains of attraction arbitrary balls in $\mathscr{D}(A^\gamma)^q$, rather than sufficiently thin neighborhoods of the pencil space $\mathbf{1}\mathscr{D}(A^\gamma)$. We remark here that the idea for our constructions in this section and the method of proof for the main error bound were inspired by the paper of Eirola and Nevanlinna [6] in the context of ordinary differential equations.

In §4, we outline the attractor approximation theory of [14], and state in Theorem 4 a variation of their main attractor convergence theorem for the monoid case introduced above. The approximation properties of S_k^n required by the hypotheses of this theorem are given by the results of Theorem 3. We require additional assumptions on A, in particular that it has compact resolvent, in order to show the compactness properties of S_k^n required by the hypotheses of Theorem 4. Finally, in Theorem 5, we apply Theorem 4 to demonstrate the existence and upper semicontinuity of a family of local attractors $\{\mathscr{A}_k\}_{k>0}$ for the family of semigroups $\{S_k^n\}_{k>0}$.

2. Linear estimates

We begin by defining a sectorial operator A and by stating the assumptions we place on a linear multistep method to ensure that it is a dissipative approximation to the equation

$$(2.1) \qquad u_t + Au = 0.$$

For X a Banach space, let $\mathscr{D}(A)$ be a dense subspace of X. We consider a closed linear operator $A : \mathscr{D}(A) \to X$, such that for constants $a \geq 0$, $M \geq 1$ and $\theta \in (0, \pi/2)$,

$$\|(\lambda I - A)^{-1}\| \leq \frac{M}{|\lambda - a| + a} \quad \text{for all } \lambda \in \mathbb{C}\backslash S_{a,\theta}, \tag{2.2}$$

where $S_{a,\theta} = \{a\} \cup \{\lambda \in \mathbb{C}\backslash\{a\} \mid \theta \geq |\arg(\lambda - a)|\}$ (we also consider $\overline{S}_{a,\theta}$, the closure of $S_{a,\theta}$ in the Riemann sphere topology). An operator A satisfying these conditions is known as **sectorial**. Thanks to Friedman [8], A is the generator of an analytic semigroup e^{-At} on X, for which there exists a constant C such that

$$\|e^{-At}\| \leq Ce^{-at} \quad \text{for all } t \geq 0. \tag{2.3}$$

Definition 2.1. For the equation $u_t = F(t, u)$, and given initial data $\{U_i\}_{i=0}^{q-1}$, a q-step **linear multistep method**, parametrized by its time steplength k, generates $\{U^{n+q}\}_{n\geq 0}$, an approximation to $\{u((n+q)k)\}_{n\geq 0}$, from the equation

$$\sum_{i=0}^{q} \alpha_i U^{n+i} = k \sum_{i=0}^{q} \beta_i F((n+i)k, U^{n+i}) \quad \text{for all } n \geq 0, \tag{2.4}$$

for constants $\{\alpha_i\}_{i=0}^{q}$, $\{\beta_i\}_{i=0}^{q}$ normalized so that $\alpha_q = 1$.

The method is known as **consistent** if

$$\sum_{i=0}^{q} \alpha_i = 0. \tag{2.5}$$

It is known to have pth order, for $p = 1, 2, \ldots$, if, in addition,

$$\sum_{i=0}^{q} \alpha_i i^s = s \sum_{i=0}^{q} \beta_i i^{s-1} \quad \text{for all } s = 1, 2, \ldots, p. \tag{2.6}$$

As is well known [11], an equivalent condition to (2.5, 2.6) is that

$$\rho(e^{-z}) + z\sigma(e^{-z}) = O(z^{p+1}) \tag{2.7}$$

in a neighborhood of the origin, where

$$\rho(z) \equiv \sum_{i=0}^{q} \alpha_i z^i, \quad \sigma(z) \equiv \sum_{i=0}^{q} \beta_i z^i. \tag{2.8}$$

Definition 2.2. For a linear q-step method, we define

$$\delta_i(z) \equiv \frac{\alpha_i + \beta_i z}{\alpha_q + \beta_q z} \quad \text{for } i = 0, 1, 2, \ldots, q.$$

The **companion matrix** for the method is given by

$$T(z) \equiv \left[\begin{array}{c|ccc} 0 & & & \\ \vdots & & I & \\ 0 & & & \\ \hline -\delta_0(z) & -\delta_1(z) & \cdots & -\delta_{q-1}(z) \end{array}\right]. \tag{2.9}$$

Definition 2.3. We call a linear q-step method **strictly** $A(\alpha)$**-stable**, for some $\alpha \in (0, \pi/2)$, if the following assumptions hold: The spectrum of $T(z)$, $\sigma(T(z)) = \{\lambda_1(z), \dots, \lambda_q(z)\}$, ordered so that $|\lambda_1(z)| \geq |\lambda_2(z)| \geq \cdots \geq |\lambda_q(z)|$, satisfies

$$|\lambda_1(z)| < 1 \text{ for all } z \in \overline{S}_{0,\alpha}\backslash\{0\}; \ \lambda_1(0) = 1;$$
$$|\lambda_i(z)| < 1 \text{ for } i = 2, 3, \dots, q, \text{ for all } z \in \overline{S}_{0,\alpha}.$$

Henceforth, we shall assume that (2.1) is approximated by a pth-order, q-step, strictly $A(\alpha)$-stable linear multistep method, where $p, q \geq 1$, $\alpha \in [\theta, \pi/2)$, with θ the sectorial angle characterizing A in (2.2).

In order to obtain the linear estimate (1.5), it is necessary to study the behavior of $T^n(z)$ for all $z \in \overline{S}_{0,\theta}$. Because of the very strong conditions imposed in Definition 2.3, $T^n(z)$, like e^{-nz}, tends uniformly exponentially to 0 in $\overline{S}_{0,\theta}\backslash N_0$, for N_0 any neighborhood of the origin. Hence, it is only the behavior of $T^n(z)$ in a region of the form $N_0 \cap \overline{S}_{0,\theta}$ that prevents the method from having an infinite order of accuracy, and so it is in this region that we must study its behavior most closely.

Lemma 2.1. *There exists a neighborhood N_0 of the origin such that $\lambda_1(z)$, the principal eigenvalue of $T(z)$, is an analytic function with*

$$\lambda_1(z) - e^{-z} = O(z^{p+1}), \quad z \in N_0. \tag{2.10}$$

Furthermore, the corresponding left and right eigenvectors $\mathbf{y}^T(z)$ and $\mathbf{v}(z)$ may be analytically defined in N_0, so that $\mathbf{v}(z) = [1, \lambda_1(z), \dots, \lambda_1^{q-1}(z)]^T$ and $\mathbf{y}^T\mathbf{v} = 1$.

Proof. As is well known [27], the eigenvalues of $T(z)$ are the solutions of the w-equation $P(w; z) = 0$, where

$$P(w; z) \equiv \rho(w) + z\sigma(w)$$

is an analytic function from $\mathbb{C}^2$ to $\mathbb{C}$. At $z = 0$, $w = 1$,

$$\frac{\partial P}{\partial w} = \rho'(1) \neq 0,$$

where the last inequality is a consequence of $\lambda_1(0)$ being a simple root. Furthermore, it is known from (2.7) that

$$P(e^{-z}; z) = O(z^{p+1}).$$

The analytic implicit function theorem [4] implies the existence of N_0, a z-neighborhood of 0, in which $\lambda_1(z)$, the branch of the solution to $P(w; z) = 0$ beginning at $z = 0$, $w = 1$, is analytic and possesses a convergent Taylor series expansion about 0, agreeing with e^{-z} up to and including $O(z^p)$.

The form of $\mathbf{v}(z)$ may be verified by direct substitution. The individual components of $\mathbf{y}^T(z)$ satisfy

$$-\delta_0(z)y_q(z) = \lambda_1(z)y_1(z),$$
$$y_i(z) - \delta_i(z)y_q(z) = \lambda_1(z)y_{i+1}(z) \quad \text{for } i = 1, \dots, q-1.$$

Since $y_q(z) = 0$ implies $y_i(z) = 0$ for all i, we may provisionally set $\hat{y}_q(z) = 1$, and the above equations may be solved recursively to find analytic $\hat{y}_i(z)$, $i = 1, 2, \dots, q-1$, using the fact that $\lambda_1(z)$ is bounded away from 0 in a refined neighborhood N_0. At $z = 0$, an argument of Eirola and Nevanlinna [6] implies that

$$\begin{aligned}\hat{\mathbf{y}}^T(0)\mathbf{v}(0) &= \hat{\mathbf{y}}^T(0)\mathbf{1} = \sum_{i=1}^{q} \hat{y}_i \\ &= q\hat{y}_q + \sum_{i=1}^{q-1} i(\hat{y}_i - \hat{y}_{i+1}) = \hat{y}_q \sum_{i=0}^{q} i\alpha_i \\ &= \hat{y}_q \sum_{i=0}^{q} \beta_i = \hat{y}_q \rho'(1) \neq 0.\end{aligned} \tag{2.11}$$

Hence, in a refined neighborhood N_0, $\hat{\mathbf{y}}^T(z)\mathbf{v}(z) \neq 0$ by continuity. Thus, the required analytic normalization for $\mathbf{y}(z)$ exists. □

Considering the operator $(T(z) - \lambda_1(z)\mathbf{v}(z)\mathbf{y}^T(z))^n$ in N_0, we see that its eigenvalues are

$$\{0, \lambda_2^n(z), \dots, \lambda_q^n(z)\},$$

and that, therefore, it tends to 0 exponentially fast with n. Hence, for large n, $T^n(z) \sim \lambda_1^n(z)\mathbf{v}(z)\mathbf{y}^T(z)$ in N_0. One knows from Lemma 2.1 that $\lambda_1(z)$ is closely approximated by e^{-z}, so one expects $e^{-nz}\mathbf{v}(z)\mathbf{y}^T(z)$ to be a good approximation of $T^n(z)$ for large n. However, in this paper we will only consider the suboptimal approximation $e^{-nz}\mathbf{1}\mathbf{Y}^T$, where $\mathbf{1} = \mathbf{v}(0)$ and $\mathbf{Y}^T = \mathbf{y}^T(0)$. The following lemma establishes some of the properties of that approximation in a region of the complex plane associated with a contour integral appearing later.

Lemma 2.2. *There exist* N_0, *a neighborhood of* 0, *and constants* $C > 0$ *and* $\Lambda \in (0, 1)$, *such that*

$$\|T^n(z) - \mathbf{1}e^{-nz}\mathbf{Y}^T\| \le C(\Lambda^n + |ze^{-nz/2}|), \tag{2.12}$$

for all $z \in N_0 \cap (\overline{B}(0, (\operatorname{cosec}\theta)/n) \cup \overline{S}_{0,\theta})$ *and all* $n \in \mathbb{N}$.

Proof. Let N_0 be as in the statement of Lemma 2.1. For $z \in N_0$ and $n \in \mathbb{N}$,

$$\begin{aligned}\|T^n(z) - \mathbf{1}e^{-nz}\mathbf{Y}^T\| &\le \|T^n(z) - \mathbf{v}(z)\lambda_1^n(z)\mathbf{y}^T(z)\| \\ &\quad + |e^{-nz}|\,\|\mathbf{v}(z)\mathbf{y}^T(z) - \mathbf{1}\mathbf{Y}^T\| \\ &\quad + \|\mathbf{v}(z)\mathbf{y}(z)^T\|\,|\lambda_1^n(z) - e^{-nz}| \\ &\equiv I_1 + I_2 + I_3.\end{aligned} \tag{2.13}$$

In the analysis below, let C denote a generic constant independent of k and z. The matrix $T(z) - \lambda_1(z)\mathbf{v}(z)\mathbf{y}^T(z)$ has eigenvalues $\{0, \lambda_2(z), \dots, \lambda_q(z)\}$. We recall the well-known construction of Isaacson and Keller [19], which shows that for a given invertible $M \in \mathscr{L}(\mathbb{C}^q)$ and $\varepsilon > 0$ one may find an invertible $H \in \mathscr{L}(\mathbb{C}^q)$ such that $\|H^{-1}MH\| \le \rho(M) + \varepsilon$, where $\rho(M)$ is the spectral radius of M. Since $\rho(T(z) - \lambda_1(z)\mathbf{v}(z)\mathbf{y}^T(z)) = |\lambda_2(z)| < 1$, there exist invertible matrices $H(z)$ and $H^{-1}(z)$ in a neighborhood of 0, continuously dependent on z, such that

$$\|H^{-1}(z)(T(z) - \lambda_1(z)\mathbf{v}(z)\mathbf{y}^T(z))^n H(z)\| \le \left(\frac{|\lambda_2(z)| + 1}{2}\right)^n.$$

Hence, there is a continuous function $C(z)$, defined on N_0, such that

$$\|(T(z) - \lambda_1(z)\mathbf{v}(z)\mathbf{y}^T(z))^n\| \leq C(z)\left(\frac{|\lambda_2(z)|+1}{2}\right)^n. \tag{2.14}$$

The functions $C(z)$ and $\frac{1}{2}(|\lambda_2(z)|+1)$ may be uniformly bounded above in N_0 by constants $C > 0$ and $\Lambda \in (0, 1)$, respectively. Noting the identity

$$(T(z) - \lambda_1(z)\mathbf{v}(z)\mathbf{y}^T(z))^n = T^n(z) - \lambda_1^n(z)\mathbf{v}(z)\mathbf{y}^T(z),$$

we deduce the existence of constants $C > 0$ and $\Lambda \in (0, 1)$ such that

$$I_1 \leq C\Lambda^n \quad \text{for all } z \in N_0 \text{ and } n \in \mathbb{N}. \tag{2.15}$$

By the continuity of $d\mathbf{v}/dz$ and $d\mathbf{y}/dz$ in N_0, there is a constant C such that

$$I_2 \leq C|ze^{-nz}| \quad \text{for all } z \in N_0 \text{ and } n \in \mathbb{N}. \tag{2.16}$$

We now consider $|\lambda_1^n(z) - e^{-nz}|$ in N_0. By Lemma 2.1, $\lambda_1(z) = e^{-z} + C(z)z^{p+1}$, for some analytic $C(z)$ and $z \in N_0$, which implies

$$\log\lambda_1(z) = -z + C_1(z)z^{p+1}$$

for some analytic $C_1(z)$ in a possibly refined neighborhood N_0. Hence,

$$|\lambda_1^n(z) - e^{-nz}| \leq A(z)B(z)|ze^{-nz/2}|,$$

where

$$A(z) = |e^{-nz/4}|\left|\frac{1-\exp(nC_1(z)z^{p+1})}{nC_1(z)z^{p+1}}\right|,$$
$$B(z) = |nC_1(z)z^pe^{-nz/4}|.$$

An application of the maximum modulus principle implies that both $A(z)$ and $B(z)$ may be bounded independently of n in the regions $N_0 \cap \overline{B}(0, (\operatorname{cosec}\theta)/n)$ and $N_0 \cap \overline{S}_{0,\theta}$, provided $\operatorname{diam}(N_0)$ is sufficiently small. Hence, for a possibly refined neighborhood N_0,

$$I_3 \leq C|ze^{-nz/2}| \quad \text{for all } z \in N_0 \cap (\overline{B}(0, (\operatorname{cosec}\theta)/n) \cup \overline{S}_{0,\theta}) \text{ and } n \in \mathbb{N}. \tag{2.17}$$

By a similar argument, an identical bound may be obtained for I_2, and hence (2.12) follows. □

The following two lemmas, due to Le Roux [27, Lemmes 3 and 4], characterize the behavior of the operator $T^n(z)$ away from N_0.

Lemma 2.3. *For any bounded neighborhood N_0 of 0 there exist constants $C > 0$ and $\Lambda \in (0, 1)$, independent of n, such that for all $z \in \overline{S}_{0,\theta}\backslash N_0$ and $n \in \mathbb{N}$*

$$\|T^n(z) - \mathbf{1}e^{-nz}\mathbf{Y}^T\|_{\mathscr{L}(\mathbb{C}^q)} \leq C\Lambda^n. \tag{2.18}$$

Lemma 2.4. *For any bounded neighborhood N_0 of 0 there exist constants $C > 0$ and $\Lambda \in (0, 1)$, independent of n, such that for all $z \in \overline{S}_{0,\theta}\backslash N_0$ and $n \in \mathbb{N}$*

$$\|T^n(z) - T^n(\infty)\|_{\mathscr{L}(\mathbb{C}^q)} \leq \frac{C\Lambda^n}{|z|}. \tag{2.19}$$

We are now in a position to state and prove the main result of this section.

Theorem 1. *Suppose that A is a sectorial operator satisfying* (2.2) *for constants $M \geq 1$, $a \geq 0$ and $\theta \in (0, \pi/2)$. In addition, suppose that for $p, q \geq 1$, the equation $u_t + Au = 0$ is discretized by a pth-order q-step linear multistep method, strongly $A(\alpha)$-stable for some $\alpha \in [\theta, \pi/2)$, with companion matrix $T(z)$.*

Then, the linear operator $T^n(kA) : X^q \to X^q$ is well defined for all $n \geq 0$, and there exists a constant C_0, independent of n, a and k, such that for k sufficiently small,

$$\|T^n(kA) - \mathbf{1}e^{-nkA}\mathbf{Y}^T\|_{\mathscr{L}(X^q)} \leq C_0 e^{-ank/2}(ak + 1/n) \quad \textit{for all } n \in \mathbb{N}. \tag{2.20}$$

Proof. Let R be the largest radius such that $B(0, 2R) \subset N_0$, where N_0 is as in the statement of Lemma 2.2. Let us choose the maximum of the three constants Λ in the statements of Lemmas 2.2–2.4 and $\exp(-(R/4)\cos\theta)$, and again denote it Λ. For $a > 0$, we require that $k \leq |\log \Lambda|/a$ (no restriction applies if $a = 0$). This implies that $e^{ak/2}\Lambda \leq \Lambda^{1/2} < 1$.

We consider first the case $n \geq 2/R$. The function $f(z) \equiv T^n(z) - T^n(\infty) - \mathbf{1}e^{-nz}\mathbf{Y}^T$ is analytic in the region $\overline{S}_{0,\theta} \cup N_0$ and vanishes at ∞. Hence, one may use the Dunford-Taylor integral formulation [5, Section VII.9]:

$$\begin{aligned} &e^{ank/2}(T^n(kA) - \mathbf{1}e^{-nkA}\mathbf{Y}^T) \\ &\quad = e^{ank/2}T^n(\infty) + \frac{e^{ank/2}}{2\pi i}\int_{\Gamma + ak} [T^n(z) - T^n(\infty) - \mathbf{1}e^{-nz}\mathbf{Y}^T](zI - kA)^{-1}\,dz. \end{aligned} \tag{2.21}$$

Here, $\Gamma + ak$ is the positively oriented curve, lying in the resolvent set of kA, equal to the union of the three portions

$$\begin{aligned} \Gamma_1 + ak &= ak + \frac{e^{i\phi}}{n}, && \theta \leq \phi \leq 2\pi - \theta, \\ \Gamma_2 + ak &= ak + re^{\pm i\theta}, && 1/n \leq r \leq R, \\ \Gamma_3 + ak &= ak + re^{\pm i\theta}, && R \leq r \leq \infty. \end{aligned}$$

Geometric considerations show that $\Gamma + ak \subset \overline{B}(0, (\operatorname{cosec}\theta)/n) \cup \overline{S}_{0,\theta}$ for all $ak \geq 0$. Hence, we may apply Lemma 2.2 to bound $\|T^n(z) - e^{-nkz}\mathbf{1}\mathbf{Y}^T\|$ in $(\Gamma_1 + ak) \cup (\Gamma_2 + ak)$.

Below, we show that the right-hand side of (2.21) may be bounded by $C(ak + 1/n)$, for some C independent of n, a and k, and hence deduce the theorem for $n \geq 2/R$. Inequality (2.2) and Lemmas 2.2–2.4 are used tacitly. Note that $\|e^{ank/2}T^n(\infty)\| \leq \Lambda^{n/2} \leq C/n$. Hence,

$$\begin{aligned} &\left\| \frac{e^{ank/2}}{2\pi i}\int_{\Gamma_1 + ak} [T^n(z) - T^n(\infty) - \mathbf{1}e^{-nz}\mathbf{Y}^T](zI - kA)^{-1}\,dz \right\| \\ &\quad \leq \frac{M}{2\pi}\int_{\Gamma_1 + ak} \frac{C((e^{ak/2}\Lambda)^n + |ze^{-n(z-ak)/2}|)}{|z - ak|}\,dz \\ &\quad \leq \frac{M}{2\pi}\int_{\Gamma_1} \frac{C(\Lambda^{n/2} + (|y| + ak)|e^{-ny/2}|)}{|y|}\,dy \\ &\quad \leq MC(\Lambda^{n/2} + (ak + 1/n)) \\ &\quad \leq C(ak + 1/n), \end{aligned}$$

$$\begin{aligned}
&\left\| \frac{e^{ank/2}}{2\pi i} \int_{\Gamma_2+ak} [T^n(z) - T^n(\infty) - \mathbf{1}e^{-nz}\mathbf{Y}^T](zI - kA)^{-1}\, dz \right\| \\
&\quad \le \frac{M}{2\pi} \int_{\Gamma_2+ak} \frac{C((e^{ak/2}\Lambda)^n + |ze^{-n(z-ak)/2}|)}{|z-ak|}\, dz \\
&\quad \le \frac{M}{2\pi} \int_{\Gamma_2} \frac{C(\Lambda^{n/2} + (|y| + ak)|e^{-ny/2}|)}{|y|}\, dy \\
&\quad \le \frac{CM}{\pi} \int_{1/n}^{R} \left(\frac{\Lambda^{n/2}}{r} + e^{-(nr/2)\cos\theta}(1 + ak/r) \right) dr \\
&\quad \le C(\Lambda^{n/2} \log n + 2\sec\theta(ak + 1/n)) \\
&\quad \le C(ak + 1/n), \\
&\left\| \frac{e^{ank/2}}{2\pi i} \int_{\Gamma_3+ak} [T^n(z) - T^n(\infty) - \mathbf{1}e^{-nz}\mathbf{Y}^T](zI - kA)^{-1}\, dz \right\| \\
&\quad \le \frac{M}{2\pi} \int_{\Gamma_3+ak} \frac{C(e^{ak/2}\Lambda)^n}{|z|\,|z-ak|}\, dz \\
&\quad \le \frac{M}{2\pi} \int_{\Gamma_3} \frac{C\Lambda^{n/2}}{|y+ak|\,|y|}\, dy \\
&\quad \le \frac{CM\Lambda^{n/2}}{\pi} \int_R^\infty \frac{dr}{r^2} \\
&\quad \le C/n.
\end{aligned}$$

For $n \in [1, 2/R]$ we may similarly bound the right-hand side of (2.21) by $C(1 + ak)$, for C independent of n and k, by considering $\Gamma + ak$ where $\Gamma = \Gamma_3 \cup \Gamma_4$,

$$\Gamma_4 = Re^{i\phi}, \qquad \theta \le \phi \le 2\pi - \theta.$$

Hence, we may choose C_0 so that (2.20) holds for all $n \in \mathbb{N}$. Since e^{-nkA} is a bounded linear operator, so is $T^n(kA)$, and hence the latter is well defined. □

Note that combining (2.3) and (2.20), we may deduce that there exists a constant C, independent of k and n, such that, for sufficiently small k,

$$\|T^n(kA)\| \le Ce^{-ank/2}. \tag{2.22}$$

We remark that if A obeys (2.2) with strictly positive a, then (2.2) is also obeyed with a taken to be 0, for the same M and θ. Hence, the following result holds for the same C_0 as in (2.20), without restriction on the step size k:

$$\|T^n(kA) - \mathbf{1}e^{-nkA}\mathbf{Y}^T\|_{\mathscr{L}(X^q)} \le \frac{C_0}{n} \quad \text{for all } n \in \mathbb{N}. \tag{2.23}$$

3. Nonlinear approximation

Let A be a sectorial operator of the type considered in the last section, and suppose that A^γ is a fractional power of A, for $\gamma \in [0, 1)$. The domain of A^γ, $\mathscr{D}(A^\gamma)$, may be considered as a Banach subspace of X with norm $\|A^\gamma(\cdot)\|$, where $\|\cdot\|$ is the norm of X. In this section we consider equations of the type

$$u_t + Au = f(u), \tag{3.1}$$

where $f : \mathscr{D}(A^\gamma) \to X$ is a locally Lipschitz continuous function. The theory of fractional powers of sectorial operators and of equations of the form (3.1) is considered in [16] and [30]. As is shown in [16], systems of reaction-convection-diffusion equations and the incompressible Navier-Stokes equations may be put into the form (3.1), and hence the range of applicability of our results is wide.

To approximate (3.1), we use a strongly $A(\alpha)$-stable linear pth-order, q-step method, for $p, q \geq 1$ and $\alpha \in [\theta, \pi/2)$:

$$\sum_{i=0}^{q} \alpha_i U^{n+i} = k \sum_{i=0}^{q} \beta_i(-AU^{n+i} + f(U^{n+i})) \quad \text{for } n \geq 0, \tag{3.2}$$

with initial data $\{U^i\}_{i=0}^{q-1}$ given in a bounded set of $\mathscr{D}(A^\gamma)$.

Our aim in the present section is to prove, in a nonlinear context, an error bound of the same kind as was given by Theorem 1 for the linear monoid $\mathbf{1}e^{-At}\mathbf{Y}^T$. We proceed initially to outline the known existence and regularity theory for (3.1). Next, we show that, for sufficiently small k, there exists a unique solution U^{n+q} to the implicit equation defined by the numerical method (3.2). This result implies that $S_k : \mathscr{D}(A^\gamma)^q \to \mathscr{D}(A^\gamma)^q$, the mapping associated with the method, given by

$$S_k\mathbf{U}^n \mapsto \mathbf{U}^{n+1} \quad \text{for } \mathbf{U}^n \equiv [U^n, \dots, U^{n+q-1}]^T,$$

is well defined.

Subsequently, given the semigroup $S(t) : \mathscr{D}(A^\gamma) \to \mathscr{D}(A^\gamma)$, defined on $u_0 \in X$ by $S(t)u_0 = u(t)$, the solution to (3.1), we use the same principle as was applied to e^{-At} in the last section, to transform $S(t)$ into the monoid

$$\mathbf{1}S(t)\mathbf{Y}^T : \mathscr{D}(A^\gamma)^q \to \mathscr{D}(A^\gamma)^q,$$

where $\mathbf{Y}^T$ is again the left eigenvector of the matrix $T(0)$ for the eigenvalue 1.

Following this construction, the semigroup S_k^n and the monoid $\mathbf{1}S(t)\mathbf{Y}^T$ are commensurable—both are defined on $\mathscr{D}(A^\gamma)^q$. This enables us to consider and bound the error $\|(S_k^n - \mathbf{1}S(t)\mathbf{Y}^T)\mathbf{U}^0\|$ in the norm of $\mathscr{D}(A^\gamma)^q$, which we do in Theorem 3, the main result of this section. When a local attractor $\mathscr{A}$ exists for $S(t)$, this error bound is a key element in establishing sufficient conditions for the existence of local attractors for S_k^n in a neighborhood of $\mathbf{1}\mathscr{A}$, the attractor of $\mathbf{1}S(t)\mathbf{Y}^T$.

Besides giving the existence and regularity of solutions of (3.1), the following lemma states a Lipschitz property of $f(u(t))$, needed for our main error bound.

Lemma 3.1. *Let A be a sectorial operator obeying* (2.2) *for some $M \geq 1$, $a > 0$ and $\theta \in (0, \pi/2)$, and suppose that, for $\gamma \in [0, 1)$, $f : \mathscr{D}(A^\gamma) \to X$ is locally Lipschitz continuous. Then, given initial data u_0 in a ball $B(\mathscr{D}(A^\gamma); 0, r)$, for some $r > 0$, there exists a time $T(r) > 0$ and a unique solution $u(t)$, defined on $[0, T]$, such that*

1. *$u \in C[[0, T]; \mathscr{D}(A^\gamma)]$, and $u(0) = u_0$;*
2. *$u_t \in C[(0, T]; \mathscr{D}(A^\gamma)]$, $u \in C[(0, T]; \mathscr{D}(A)]$, $f(u(\cdot)) \in C[[0, T]; X]$, and so for $t > 0$ every term in* (3.1) *is an element of X.*

In addition, for $t \in (0, T]$, the solution may be written as

$$u(t) = e^{-At}u_0 + \int_0^t e^{-A(t-s)} f(u(s))\, ds. \tag{3.3}$$

Furthermore, if such a solution, as described above, exists on a compact time interval $[0, T]$, *then there exists* $C > 0$ *such that if* $T \geq t + h > t > h > 0$, *then*

$$\|f(u(t)) - f(u(t+h))\| \leq C\frac{h}{t}. \tag{3.4}$$

Proof. All results, except the last sentence of the lemma, are given explicitly in [16, Theorems 3.3.3 and 3.5.2]. The last part also follows easily from these results. □

The following lemma summarizes several results on norms of simple functions of A which we shall need in the remainder of this section.

Lemma 3.2. *Let* A *be a sectorial operator, obeying* (2.2) *for some* $a > 0$. *Then, for* $\beta \in [0, 1]$, *there exist generic constants* $C = C(M, a, \theta, \beta)$ *such that*

$$\|A^\beta(\lambda I - A)^{-1}\| \leq C|\lambda|^{-(1-\beta)} \quad \textit{for } \lambda \in \mathbb{C}\backslash S_{a,\theta},$$
$$\|A^\beta e^{-At}\| \leq Ct^{-\beta}e^{-at} \quad \textit{for } t \in (0, \infty),$$
$$\|A^{-\beta}[I - e^{-Ak}]\| \leq Ck^\beta \quad \textit{for } k \in (0, 1].$$

Furthermore, if a linear multistep method is strongly $A(\alpha)$*-stable, for some* $\alpha \in (0, \pi/2)$, *then there exists a constant* $C_1 \geq 1$, *independent of* k, *such that*

$$\|\delta_i(kA)\| \leq C_1 \quad \textit{for all } i \in [0, q].$$

Proof. We note that the conditions given in Definition 2.3 for strict $A(\alpha)$-stability are sufficient to imply $\beta_q > 0$ (see [12, p. 277]). All the results now follow easily from those of [16, §1.4]. □

The following theorem establishes the existence of a locally unique solution U^{n+q} to (3.2) for all sufficiently small k, using the contraction mapping theorem.

Theorem 2. *Let the assumptions of Lemma* 3.1 *hold with respect to* A *and* f, *and suppose that* $\{U^i\}_{i=0}^{q-1}$ *is given data for* (3.2) *in* $B(\mathscr{D}(A^\gamma); 0, r)$, *for* $r > 0$. *Then, there exists a constant* $R_0(r) \geq r$ *such that for all* $R \geq R_0$ *there is a* $k_0(R)$ *such that, for* $n = 0$ *and* $k \in (0, k_0]$, *there is exactly one solution of* (3.2) *in the ball* $B(\mathscr{D}(A^\gamma); 0, R)$.

Proof. We rewrite (3.2), for $n = 0$, as the implicit equation $U^q = G_k(U^q)$, where

$$\begin{aligned} G_k(v) \equiv \sum_{i=0}^{q-1}\{-\delta_i(kA)U^i + k(I + \beta_q kA)^{-1}\beta_i f(U^i)\} \\ + k(I + \beta_q kA)^{-1}\beta_q f(v). \end{aligned} \tag{3.5}$$

Let us define the closed convex $V \equiv \{u \in \mathscr{D}(A^\gamma) \mid \|A^\gamma(G_k(0) - u)\| \leq 1\}$. We shall show that G_k is an endomorphism of V, for all sufficiently small k.

We first show that $G_k(0)$, and hence V, can be bounded in $\mathscr{D}(A^\gamma)$ independently of k:

$$\begin{aligned} \|A^\gamma G_k(0)\| &\leq \sum_{i=0}^{q-1}\|\delta_i(kA)\|\,\|A^\gamma U^i\| + k\|A^\gamma(I + \beta_q kA)^{-1}\|\,|\beta_i|\,\|f(U^i)\| \\ &\quad + k\|A^\gamma[I + \beta_q kA]^{-1}\|\beta_q\|f(0)\| \\ &\leq (q+1)C_1(r + Ck^{1-\gamma}L), \end{aligned} \tag{3.6}$$

where $C_1 \geq 1$ and C are constants from Lemma 3.2, and L is a uniform bound for $\|f(u)\|$, for $u \in B(\mathscr{D}(A^\gamma); 0, r)$.

Now, we assume that $k \leq 1$ and fix $R_0 = (q+1)C_1(r + CL) + 1$. We note that $C_1 \geq 1$ implies $R_0 > r$; thus, $V \subset B(\mathscr{D}(A^\gamma); 0, R_0)$. Suppose that $u \in V$; then

$$\begin{aligned}\|A^\gamma(G_k(u) - G_k(0))\| &\leq k\|A^\gamma(I + \beta_q kA)^{-1}\| \, \|\beta_q(f(u) - f(0))\| \\ &\leq k^{1-\gamma}C(R_0, \gamma)\|A^\gamma u\| \leq k^{1-\gamma}R_0 C(R_0, \gamma),\end{aligned}$$

where $C(R_0, \gamma)$ depends on the local Lipschitz constant for f in $B(\mathscr{D}(A^\gamma); 0, R_0)$.

Hence, G_k is an endomorphism of V, provided that

$$k^{1-\gamma} < \frac{1}{R_0 C(R_0, \gamma)}. \tag{3.7}$$

We now show that G_k is a contraction on V. For $u, v \in V$,

$$\begin{aligned}\|A^\gamma[G_k(u) - G_k(v)]\| &= k\|A^\gamma(I + \beta_q kA)^{-1}\beta_q(f(u) - f(v))\| \\ &\leq k^{1-\gamma}C(R_0, \gamma)\|A^\gamma(u - v)\|,\end{aligned}$$

where $C(R_0, \gamma)$ is as above. Hence, (3.7) is sufficient to imply that G_k is both an endomorphism and a contraction of V, and we may apply the contraction mapping theorem to deduce the existence of a fixed point for G_k, unique in V.

Given some $R > 0$, the function $\|f(\cdot)\|$ is uniformly bounded on $B_R = B(\mathscr{D}(A^\gamma); 0, R)$. Hence, provided k is sufficiently small, any solution to (3.2) lying in B_R is also in V, and thus there is at most one solution in B_R. On the other hand, if $R \geq R_0$, then $V \subset B_R$. Hence, for $R \geq R_0$ and k small enough, there exists a unique U^q in B_R. □

To compare the quantities $\mathbf{U}(nk) = \mathbf{1}S(nk)\mathbf{Y}^T\mathbf{U}^0$ and $\mathbf{U}^n = S_k^n\mathbf{U}^0$, for $\mathbf{U}^0 = [U^0, \dots, U^{q-1}]^T$, we consider integrated forms of equations (3.1) and (3.2), which are written in terms of their respective vector variables in such a way as to make explicit the linear operators $\mathbf{1}e^{-At}\mathbf{Y}^T$ and $T^n(kA)$. Considering S_k^n first, we use a construction for $f(\cdot)$, due to Eirola and Nevanlinna [6] in the context of ODEs, to write (3.2) in a vector form which mimics (3.3), the Volterra integral equation form of the solution of (3.1):

$$\mathbf{U}^n = T^n(kA)\mathbf{U}^0 + k\sum_{i=0}^{n-1} T^{n-1-i}(kA)[I + \beta_q kA]^{-1}\mathbf{F}(\mathbf{U}^i, \mathbf{U}^{i+1}), \tag{3.8}$$

where we define

$$\mathbf{F}(\mathbf{U}, \mathbf{V}) \equiv \left(0, \dots, 0, \beta_q f(V_q) + \sum_{j=0}^{q-1} \beta_j f(U_{j+1})\right)^T. \tag{3.9}$$

We recall from (2.11) that $1 = \mathbf{Y}^T\mathbf{1} = Y_q\sum_{i=0}^q \beta_i$. Hence,

$$\mathbf{Y}^T\mathbf{F}(u\mathbf{1}, u\mathbf{1}) = Y_q\sum_{i=0}^{q} \beta_i f(u) = f(u).$$

Now if $u(t)$ is the solution of (3.1) for initial data $\mathbf{Y}^T\mathbf{U}^0$, then from (3.3) we have

$$u(t) = e^{-At}\mathbf{Y}^T\mathbf{U}^0 + \int_0^t e^{-A(t-s)}\mathbf{Y}^T\mathbf{F}(\mathbf{1}u(s), \mathbf{1}u(s))\,ds.$$

Multiplying the last equation by the vector $\mathbf{1}$, we note that $\mathbf{U}(t) = \mathbf{1}S(t)\mathbf{Y}^T\mathbf{U}^0 = \mathbf{1}u(t)$, and hence

$$\mathbf{U}(t) = [\mathbf{1}e^{-At}\mathbf{Y}^T]\mathbf{U}^0 + \int_0^t [\mathbf{1}e^{-A(t-s)}\mathbf{Y}^T]\mathbf{F}(\mathbf{U}(s), \mathbf{U}(s))\,ds. \tag{3.10}$$

The representations (3.8) and (3.10) put us in a position to derive an error bound for $\|A^\gamma(\mathbf{U}(nk) - \mathbf{U}^n)\| = \|A^\gamma(\mathbf{1}S(nk)\mathbf{Y}^T - S_k^n)\mathbf{U}^0\|$.

Theorem 3. *Suppose that $T(z)$ obeys the conditions of Theorem 1; that sectorial A obeys (2.2) for $a > 0$; that $f : \mathscr{D}(A^\gamma) \to X$ is locally Lipschitz continuous, and that initial data $\mathbf{U}^0 \in B(\mathscr{D}(A^\gamma)^q; 0, r)$ is chosen for (3.10), for some $r > 0$.*

Then, there exists a constant $\tau(r) > 0$ and a maximal time $T^* \in [\tau, \infty]$, for which a solution $\mathbf{U} \in C[[0, T^*]; \mathscr{D}(A^\gamma)^q] \cap C[(0, T^*]; \mathscr{D}(A)^q]$ exists. Furthermore, if $T \in (0, T^*)$, then for all sufficiently small $k > 0$, and all $nk \in [0, T]$, a solution $\mathbf{U}^n$ exists for (3.8) with initial data $\mathbf{U}^0$ such that

$$\|A^\gamma(\mathbf{U}(nk) - \mathbf{U}^n)\| \le \left[\frac{2C_0 r}{n^{1-\gamma}} + K(T, r)k^\delta\right], \tag{3.11}$$

where δ is an arbitrary constant in $(0, 1-\gamma)$, C_0 is the constant in the statement of Theorem 1, and $K(\cdot, \cdot)$ is a positive function, nondecreasing in its arguments, dependent on $(1-\gamma-\delta)^{-1}$, but independent of k.

Note. If $g(A)$ is a function of A, defined on $\mathscr{D}(A)$, we use the notation $g(A)$ to represent the q-dimensional matrix $\operatorname{diag}(g(A), \dots, g(A))$ defined on $\mathscr{D}(A)^q$.

Proof. Since $\mathbf{U}^0 \in B(\mathscr{D}(A^\gamma)^q; 0, r)$, we have $u_0 = \mathbf{Y}^T\mathbf{U}^0 \in B(\mathscr{D}(A^\gamma); 0, \|\mathbf{Y}\|r)$. Lemma 3.1 implies the existence of a solution $u(t)$ to (3.1), for $u_0 = \mathbf{Y}^T\mathbf{U}^0$, on an interval $[0, T]$ for some $T \ge \tau(r) > 0$. The derivation leading to (3.10) demonstrates that $\mathbf{1}u(t)$ is a solution to (3.10) if and only if $u(t)$ is a solution of (3.1). Hence, existence and uniqueness of a solution $\mathbf{U}(t)$ for (3.10), with $\mathbf{U}(0) = \mathbf{U}^0$, follows on setting $\mathbf{U}(t) = \mathbf{1}u(t)$.

We now take k to be an arbitrary, but fixed, positive number. Defining

$$M_1 = \sup_{s\in[0, T]} \|A^\gamma\mathbf{U}(s)\| + 2C_0 r + 1,$$

we consider the following inductive hypothesis.

$\mathbf{H}^n$: There exists exactly one solution $\{\mathbf{U}^i\}_{i=0}^n$ for (3.8) with the property that $\|A^\gamma\mathbf{U}^i\| \le M_1$ for all $i = 0, 1, \dots, n$.

We will show the existence of k_0, independent of n, such that if $k \in (0, k_0]$,

$$\mathbf{H}^{n-1} \Rightarrow \mathbf{H}^n, \tag{3.12}$$

for integer $n \in [1, T/k]$. When this has been achieved, induction will imply that $\mathbf{H}^n$ is true for all $n \in [0, T/k]$ because $\mathbf{H}^0$ is clearly true regardless of k. The required existence of solutions to (3.8) will then follow. The error bound (3.11) will be derived in the course of establishing the implication (3.12).

Let us assume now that $\mathbf{H}^{n-1}$ *is true for some* $n \geq 1$.

Applying Theorem 2, with $\mathbf{U}^0$ replaced by $\mathbf{U}^{n-1}$, and r by M_1, we conclude that, provided $k \in (0, k_0(R_0(M_1))]$, there is exactly one solution $\mathbf{U}^n$ to (3.2) such that $\|A^\gamma U^{n+q-1}\| \leq M_2$, where $M_2 = R_0(M_1) > M_1$ and R_0 is as in the statement of Theorem 2.

For $t = nk$ and $t_i = ik$, we may write (3.10) as

$$\mathbf{U}(t) = \mathbf{1}e^{-Ank}\mathbf{Y}^T + \sum_{i=0}^{n-1}\int_{ik}^{(i+1)k} \mathbf{1}e^{-A(t-s)}\mathbf{Y}^T\mathbf{F}(\mathbf{U}(s), \mathbf{U}(s))\,ds.$$

Subtracting (3.8) and taking norms in $\mathscr{D}(A^\gamma)$, we obtain

(3.13)

$$\begin{aligned}
&\|A^\gamma(\mathbf{U}(t) - \mathbf{U}^n)\| \\
&\quad\leq \|A^\gamma[e^{-At}\mathbf{1}\mathbf{Y}^T - T^n(kA)]\mathbf{U}^0\| \\
&\qquad+ \sum_{i=0}^{n-1}\int_{ik}^{(i+1)k} \|A^\gamma e^{-A(t-s)}\mathbf{1}\mathbf{Y}^T[\mathbf{F}(\mathbf{U}(s), \mathbf{U}(s)) - \mathbf{F}(\mathbf{U}(t_i), \mathbf{U}(t_{i+1}))]\|\,ds \\
&\qquad+ \sum_{i=0}^{n-1}\left\|\int_{ik}^{(i+1)k} A^\gamma e^{-A(t-s)}(I - [I + \beta_q kA]^{-1})\mathbf{1}\mathbf{Y}^T\mathbf{F}(\mathbf{U}(t_i), \mathbf{U}(t_{i+1}))\,ds\right\| \\
&\qquad+ \sum_{i=0}^{n-1}\int_{ik}^{(i+1)k} \|A^\gamma e^{-A(t-s)}[I + \beta_q kA]^{-1}\mathbf{1}\mathbf{Y}^T \\
&\qquad\qquad\qquad\cdot[\mathbf{F}(\mathbf{U}(t_i), \mathbf{U}(t_{i+1})) - \mathbf{F}(\mathbf{U}^i, \mathbf{U}^{i+1})]\|\,ds \\
&\qquad+ \sum_{i=0}^{n-1}\int_{ik}^{(i+1)k} \|A^\gamma[I+\beta_q kA]^{-1}[e^{-A(t-s)} - e^{-(n-i-1)kA}]\mathbf{1}\mathbf{Y}^T\mathbf{F}(\mathbf{U}^i, \mathbf{U}^{i+1})\|\,ds \\
&\qquad+ \sum_{i=0}^{n-1}\int_{ik}^{(i+1)k} \|A^\gamma[I + \beta_q kA]^{-1} \\
&\qquad\qquad\qquad\cdot[\mathbf{1}e^{-(n-i-1)kA}\mathbf{Y}^T - T^{n-1-i}(kA)]\mathbf{F}(\mathbf{U}^i, \mathbf{U}^{i+1})\|\,ds \\
&\quad\equiv I_1 + I_2 + I_3 + I_4 + I_5 + I_6.
\end{aligned}$$

We consider the terms $\{I_i\}_{i=1}^6$ separately below. For I_1, we may apply (2.23) to deduce that

$$\begin{aligned}
\|A^\gamma[e^{-Ank}\mathbf{1}\mathbf{Y}^T - T^n(kA)]\mathbf{U}_0\| &\leq \|\mathbf{1}e^{-Ank}\mathbf{Y}^T - T^n(kA)\|\,\|A^\gamma\mathbf{U}^0\| \\
&\leq \frac{C_0 r}{n}.
\end{aligned} \tag{3.14}$$

Below, $K(\cdot)$ and $K(\cdot, \cdot)$ are taken to be generic continuous positive functions nondecreasing with respect to their arguments on their respective domains: $[0, \infty)$ and $[0, \infty) \times [0, \infty)$. In addition, they depend on k and n only through t. However, they are allowed to depend on the bound M_2 for $\|A^\gamma\mathbf{U}^n\|$, whereas this is not true of C_0 in inequality (3.14). It is also assumed that $K(T, r) \geq 1$. Below, C will denote a generic positive constant independent of k, n, t and r.

Taking the integrand of the ith term of I_2, for $i \geq 1$, we obtain the following for $s \in [t_i, t_{i+1})$, using Lemma 3.1 and the definition of $\mathbf{F}$:

$$\begin{aligned}
&\|A^\gamma e^{-A(t-s)}\mathbf{1}\mathbf{Y}^T(\mathbf{F}(\mathbf{U}(s), \mathbf{U}(s)) - \mathbf{F}(\mathbf{U}(t_i), \mathbf{U}(t_{i+1}))\| \\
&\quad \le \|A^\gamma e^{-A(t-s)}\| \, \|\mathbf{1}\mathbf{Y}^T\| \, \|\mathbf{F}(\mathbf{U}(s), \mathbf{U}(s)) - \mathbf{F}(\mathbf{U}(t_i), \mathbf{U}(t_{i+1}))\| \\
&\quad \le C(t-s)^{-\gamma}\left(\left(\sum_{j=0}^{q-1}|\beta_j| \, \|f(u(s)) - f(u(t_i))\|\right) + \beta_q\|f(u(s)) - f(u(t_{i+1}))\|\right) \\
&\quad \le K(T, r)(t-s)^{-\gamma}\left(\frac{(t_{i+1}-s)}{s} + \frac{(s-t_i)}{s}\right) \\
&\quad = \frac{K(T, r)k}{s(t-s)^\gamma}.
\end{aligned}$$

Considering the integrand corresponding to $i = 0$, and using only the boundedness of $\|f(u(t))\|$, we instead obtain

$$\|A^\gamma e^{-A(t-s)}\mathbf{1}\mathbf{Y}^T[\mathbf{F}(\mathbf{U}(s), \mathbf{U}(s)) - \mathbf{F}(\mathbf{U}(0), \mathbf{U}(k))]\| \le K(r)(t-s)^{-\gamma}.$$

Gathering these terms together, we deduce that

$$\begin{aligned}
I_2 &\le K(T, r)\left(kt^{-\gamma} + \int_k^{t/2} ks^{-1}(t/2)^{-\gamma}\, ds + \int_{t/2}^t (t/2)^{-1}k(t-s)^{-\gamma}\, ds\right) \\
&\le K(T, r)t^{-\gamma}k(1 + \log n) \\
&\le K(t, r)k^{1-\gamma-\varepsilon}, \qquad \varepsilon \in (0, 1-\gamma).
\end{aligned} \tag{3.15}$$

Considering the ith term in the sum I_3, for $i \in [0, n-2]$, we evaluate the integral, and bound it above using Lemma 3.2:

$$\begin{aligned}
&\left\|\int_{ik}^{(i+1)k} A^\gamma e^{-A(t-s)}(I - [I + \beta_q kA]^{-1})\mathbf{1}\mathbf{Y}^T\mathbf{F}(\mathbf{U}(t_i), \mathbf{U}(t_{i+1}))\, ds\right\| \\
&\quad = \|A^\gamma e^{-A(n-i-1)k}A^{-1}[I - e^{-Ak}](\beta_q kA[I + \beta_q kA]^{-1})\mathbf{1}\mathbf{Y}^T\mathbf{F}(\mathbf{U}(t_i), \mathbf{U}(t_{i+1}))\| \\
&\quad \le K(T, r)k^{1-\gamma-\varepsilon}\|A^{1-\varepsilon}e^{-A(n-i-1)k}A^{-1}[I - e^{-Ak}]\beta_q(kA)^{\gamma+\varepsilon}[I + \beta_q kA]^{-1}\| \\
&\quad \le K(T, r)k^{2-\gamma-\varepsilon}(k(n-i-1))^{-(1-\varepsilon)},
\end{aligned}$$

for $\varepsilon \in (0, 1-\gamma)$. For the term corresponding to $i = n-1$, the integrand may be bounded by $K(T, r)(nk-s)^{-\gamma}$ for $s \in [(n-1)k, nk)$, and hence the integral of this term is bounded by $K(T, r)k^{1-\gamma}$. Thus, taking the terms in reverse order, we get

$$\begin{aligned}
I_3 &\le K(T, r)\left[k^{1-\gamma} + k^{2-\gamma-\varepsilon}\sum_{i=1}^{n-1}(ki)^{-(1-\varepsilon)}\right] \\
&\le K(T, r)(t^\varepsilon/\varepsilon)k^{1-\gamma-\varepsilon}.
\end{aligned} \tag{3.16}$$

Considering part of the integrand in the ith term of I_4, we may use the Lipschitz property of f on the ball $B(\mathscr{D}(A^\gamma); 0, M_2)$ to obtain

$$\begin{aligned}
&\|\mathbf{F}(\mathbf{U}(t_i), \mathbf{U}(t_{i+1})) - \mathbf{F}(\mathbf{U}^i, \mathbf{U}^{i+1})\| \\
&\quad \le \left(\sum_{j=0}^{q-1}|\beta_j| \, \|f(u(t_i)) - f(U^{i+j})\|\right) + \beta_q\|f(u(t_{i+1}) - f(U^{i+q})\| \\
&\quad \le K(T, r)(\Delta_i + \Delta_{i+1}),
\end{aligned}$$

where $\Delta_i = \|A^\gamma(\mathbf{U}(t_i) - \mathbf{U}^i)\|$. We also note the following inequality: If $a, b \in (0, 1)$, then

$$1 - (1-a)^b \le ab(1-a)^{b-1}.$$

Combining these two results below, we obtain

$$\begin{aligned}
I_4 &\le \sum_{i=0}^{n-1} \int_{ik}^{(i+1)k} C(t-s)^{-\gamma} K(T, r)(\Delta_i + \Delta_{i+1})\, ds \\
&\le K(T, r) k^{1-\gamma} \sum_{i=0}^{n-1} (\Delta_i + \Delta_{i+1}) \left[1 - \left(1 - \frac{1}{n-i}\right)^{1-\gamma}\right] (n-i)^{1-\gamma} \\
&\le K(T, r) k^{1-\gamma} \sum_{i=0}^{n-1} (n-i)^{-\gamma} (\Delta_i + \Delta_{i+1}) \\
&\le 2K(T, r) k^{1-\gamma} \sum_{i=0}^{n-1} (n-i)^{-\gamma} \|A^\gamma(\mathbf{U}(t_i) - \mathbf{U}^i)\| \\
&\quad + K(t, r) k^{1-\gamma} \|A^\gamma(\mathbf{U}(t_n) - \mathbf{U}^n)\|.
\end{aligned} \tag{3.17}$$

Considering the ith term in I_5 for $i \in [0, n-2]$, and taking ε as for I_3, we obtain the following:

$$\begin{aligned}
&\int_{ik}^{(i+1)k} \|A^\gamma [I + \beta_q kA]^{-1} [e^{-A(t-s)} - e^{-(n-i-1)kA}] \mathbf{1} \mathbf{Y}^T \mathbf{F}(\mathbf{U}^i, \mathbf{U}^{i+1})\|\, ds \\
&\quad \le C \|\mathbf{F}(\mathbf{U}^i, \mathbf{U}^{i+1})\| \\
&\qquad \cdot \int_0^k \|A^\gamma [I + k\beta_q A]^{-1} A^{1-\varepsilon} e^{-A(n-i-1)k} A^{-1+\varepsilon} [e^{-A(k-s)} - I]\|\, ds \\
&\quad \le K(T, r) \int_0^k ((n-i-1)k)^{-1+\varepsilon} k^{1-\gamma-\varepsilon}\, ds \\
&\quad = K(T, r) k^{2-\gamma-\varepsilon} ((n-i-1)k)^{-1+\varepsilon}.
\end{aligned}$$

The term $i = n-1$ in I_5 may be bounded by $K(t, R)k^{1-\gamma}$. Using a comparison integral, we deduce that

$$\begin{aligned}
I_5 &\le K(t, r) k^{1-\gamma-\varepsilon} \left\{ k \sum_{i=0}^{n-1} ((n-i-1)k)^{-1+\varepsilon} \right\} \\
&\le K(T, r)(t^\varepsilon/\varepsilon) k^{1-\gamma-\varepsilon}.
\end{aligned} \tag{3.18}$$

For ε as for I_3, the ith term of I_6 may be bounded as follows, for $i = 0, \dots, n-2$, using Theorem 1 (the term for $i = n-1$ is $\le K(T, r)k^{1-\gamma}$):

$$\begin{aligned}
&\int_{ik}^{(i+1)k} \|A^\gamma [I + \beta_q kA]^{-1} [\mathbf{1} e^{-A(n-1-i)k} \mathbf{Y}^T - T^{n-1-i}(kA)] \mathbf{F}(\mathbf{U}^i, \mathbf{U}^{i+1})\|\, ds \\
&\quad \le K(T, r) k^{1-\gamma} (n-i)^{-1} \le K(T, r) k^{2-\gamma-\varepsilon} (k(n-i))^{-1+\varepsilon}.
\end{aligned}$$

Using a comparison integral, we obtain

$$\begin{aligned}
I_6 &\le K(T, r) k^{1-\gamma-\varepsilon} k \sum_{i=0}^{n-1} ((n-i)k)^{-1+\varepsilon} \\
&\le K(T, r)(t^\varepsilon/\varepsilon) k^{1-\gamma-\varepsilon}.
\end{aligned} \tag{3.19}$$

Assembling the terms $\{I_i\}_{i=1}^{6}$ from (3.15)–(3.20), we fix $\varepsilon \in (0, 1-\gamma)$, and take $\delta = 1-\gamma-\varepsilon$, to deduce that

$$\begin{aligned} \|A^\gamma(\mathbf{U}(t) - \mathbf{U}^n)\| &\le I_1 + I_2 + I_3 + I_4 + I_5 + I_6 \\ &\le \frac{C_0 r}{n} + K(T, r)k^\delta + K(T, r)k^\delta \\ &\quad + K(T, r)k^{1-\gamma}\sum_{i=0}^{n-1}(n-i)^{-\gamma}\|A^\gamma(\mathbf{U}(t_i) - \mathbf{U}^i)\| \\ &\quad + K(t, r)k^{1-\gamma}\|A^\gamma(\mathbf{U}(t) - \mathbf{U}^n)\| \\ &\quad + K(t, r)k^\delta + K(T, r)k^\delta. \end{aligned} \tag{3.20}$$

The inequality remains implicit however, and it is necessary that k be small enough for the coefficient of $\|A^\gamma(\mathbf{U}(t) - \mathbf{U}^n)\|$, on the right-hand side of (3.20), to be less than $1/2$; that is, we require

$$K(T, r)k^{1-\gamma} \le \frac{1}{2}. \tag{3.21}$$

After some rearrangement, we obtain

$$\begin{aligned} \|A^\gamma(\mathbf{U}(t) - \mathbf{U}^n)\| \le \frac{2C_0 r k^{1-\gamma}}{(nk)^{1-\gamma}} &+ K(T, r)k^\delta \\ &+ K(T, r)k^{1-\gamma}\sum_{i=0}^{n-1}(n-i)^{-\gamma}\|A^\gamma(\mathbf{U}(ik) - \mathbf{U}^i)\| \end{aligned} \tag{3.22}$$

for $n \ge 1$. As shown in [18], by defining the extension

$$e(s) = \|A^\gamma(\mathbf{U}(mk) - \mathbf{U}^m)\| \quad \text{for } s \in [mk, (m-1)k), \tag{3.23}$$

one may apply the Gronwall Lemma, described in [16, p. 188], to the inequality

$$e(t) \le 2C_0 r k^{1-\gamma} t^{\gamma-1} + K(T, r)k^\delta + K(T, r)\int_0^t (t-s)^{-\gamma} e(s)\, ds,$$

and, reversing the piecewise constant extension (3.23), deduce that

$$\|A^\gamma(\mathbf{U}(nk) - \mathbf{U}^n)\| \le 2C_0 r n^{\gamma-1} + K(T, r)k^\delta. \tag{3.24}$$

Fixing $K(T, r)$ to be the constant on the right-hand side of (3.24), we impose the further condition,

$$k \le K(T, r)^{-1/\delta}, \tag{3.25}$$

which will generally tend to be more restrictive than (3.21). For such k,

$$\|A^\gamma(\mathbf{U}(nk) - \mathbf{U}^n)\| \le 2C_0 r + 1 = M_1. \tag{3.26}$$

In combination with $\mathbf{H}^{n-1}$, (3.26) implies $\mathbf{H}^n$. Hence, we conclude $\mathbf{H}^{n-1} \Rightarrow \mathbf{H}^n$, and the theorem follows by induction and (3.24). □

4. Attractor approximation

As noted in the Introduction, the theory of attractor approximation for one-step time discretizations was considered by Hale, Lin and Raugel in [14]. Below, we introduce their terminology and state their main convergence theorem. However, we substitute the term monoid for semigroup on every occasion that it refers to the continuous semigroup. This makes no difference to the proofs.

Subsequently, we show, using the error bound derived in §3, that the hypotheses required by the convergence theorem are satisfied by the monoid $\mathbf{1}S(t)\mathbf{Y}^T$ and the semigroup S_k^n, corresponding to (3.10) and (3.8), respectively, provided that k is sufficiently small. Thus, we establish the existence of a family of attractors for the family of discrete semigroups, upper semicontinuous with respect to k at $k = 0$.

Definition 4.1. If $B \subset \mathscr{X}$, $\mathscr{X}$ a Banach space, we denote an r-neighborhood of B by $N(\mathscr{X}; B, r)$ or by $N(B, r)$.

Definition 4.2. If A and B are two sets in $\mathscr{X}$, then we denote the nonsymmetric **Hausdorff semidistance** from A to B by

$$\delta(A, B) \equiv \sup_{x\in A} \inf_{y\in B} \|x - y\|.$$

Definition 4.3. For $V \subseteq \mathscr{X}$, the family of maps $\{\mathscr{S}(t)\}_{t\geq 0}$, $\mathscr{S}(t) : V \to \mathscr{X}$, is a C^0**-monoid** if

(i) $\mathscr{S}(t)x$ is a continuous function for all $(x, t) \in V \times [0, \infty)$;

(ii) $\mathscr{S}(t)\mathscr{S}(s) = \mathscr{S}(t+s)$ for all $t, s \geq 0$.

If $\mathscr{S}(0) = I_{\mathscr{X}}$, then $\mathscr{S}(t)$ is a **semigroup.**

Definition 4.4. Suppose $\mathscr{S}(t)$ is a C^0-monoid defined on an open subset V of $\mathscr{X}$. A bounded subset $P \subset V$ is called **absorbing** in V, or an absorbing set in V, if for all bounded $B \subset V$ there is a time $t_0(B)$ such that $S(t)B \subset P$ for all $t \geq t_0$.

A compact set $\mathscr{A} \subset V$, for which $\mathscr{S}(t)\mathscr{A} = \mathscr{A}$ for $t \geq 0$, is said to be a **local attractor** in V, if $N(\mathscr{A}, \varepsilon)$ is an absorbing set in V for all $\varepsilon > 0$.

Discrete semigroups etc. are defined in an analogous fashion, see [13].

Definition 4.5. A discrete semigroup $\mathscr{S}^n$ defined on $V \subset \mathscr{X}$ is said to be **asymptotically smooth** if, for every nonempty closed bounded set $B \subset V$, there is a compact set $J(B)$ with the following property: given $\varepsilon > 0$, there exists an integer $n_\varepsilon > 0$ such that, for all $n > n_\varepsilon$, one has $\mathscr{S}^n L(B) \subseteq N(J(B), \varepsilon)$, where

$$L(B) \equiv \{x \in B \,|\, \mathscr{S}^n x \in B \quad \text{for } n \geq 0\}.$$

In the sequel, we shall use the following result concerning asymptotic smoothness of discrete semigroups, the proof of which is given in [13, p. 13].

Lemma 4.1. *Suppose that the discrete semigroup $\mathscr{S}^n$ is defined on $V \subseteq \mathscr{X}$ and that for each $n \in \mathbb{N}$ there exist two mappings P_n and Q_n from V to $\mathscr{X}$ such that*

(i) $\mathscr{S}^n = P_n + Q_n$;

(ii) Q_n *maps bounded sets to sets with compact closure*;

(iii) *There exists a function* $\mu : \mathbb{N} \times \mathbb{R}^+ \to \mathbb{R}^+$ *such that for all* $r > 0$, $\|P_n x\| \leq \mu(n, r)$ *for all* $x \in B(0, r) \cap V$, *and* $\lim_{n\to\infty} \mu(n, r) = 0$.

Then, $\mathcal{S}^n$ *is asymptotically smooth.*

Before quoting the abstract attractor convergence theorem, we first define the concept of conditional uniform approximation.

Definition 4.6. A monoid $\mathcal{S}(t)$, defined on $V \subset \mathcal{X}$, is said to be **conditionally approximated** on U, a bounded subset of V, **uniformly** on a compact subinterval I of $(0, \infty)$ by the family of discrete semigroups $\{\mathcal{S}_k^n\}_{k\in(0,k_0]}$, for some $k_0 > 0$, if there exists a function $\eta(k, I, U)$ such that

(i) $\lim_{k\to 0} \eta(k, I, U) = 0$;

(ii) When, for $x \in U$, $k \in (0, k_0]$ and $nk \in I$, both $\mathcal{S}_k^n x$ and $\mathcal{S}(nk)x$ are defined, then $\|\mathcal{S}(nk)x - \mathcal{S}_k^n x\| \leq \eta(k, I, U)$.

Theorem 4 (Hale-Lin-Raugel). *Suppose that* $\mathcal{S}(t)$ *is a* C^0*-monoid defined on* V, *a bounded subset of a Banach space* $\mathcal{X}$; *that* $\mathcal{A}$ *is a local attractor for* $\mathcal{S}(t)$ *in a bounded open set* $U \subset V$; *that* $\mathcal{S}(t)U \subset V$ *for all* $t \geq 0$, *and that there exists* $\delta > 0$ *such that* $N(U, \delta) \subset V$. *Suppose, in addition, that there exist constants* $t_0, k_0 > 0$ *such that* $\mathcal{S}_k^n N(U, \delta) \subset V$ *for all* $k \in (0, k_0]$ *and* $n \leq t_0/k$. *Suppose, moreover, that the family* $\{\mathcal{S}_k^n\}_{k\in(0,k_0]}$ *conditionally approximates* $\mathcal{S}(t)$ *on* V, *uniformly on compact subintervals of* $[t_0, \infty)$.

Then, given $\varepsilon > 0$, *there exist* $k_1(\varepsilon)$ *and* $\tau_0 > t_0$ *such that*

$$\mathcal{S}_k^n U \subset N(\mathcal{A}, \varepsilon) \quad \text{for all } nk \geq \tau_0 \text{ and } k \in (0, k_1]. \tag{4.1}$$

If, in addition, $\mathcal{S}_k^n$ *is asymptotically smooth on* U *for all sufficiently small* k, *the discrete semigroup* $\mathcal{S}_k^n$ *possesses a local attractor* $\mathcal{A}_k$ *in* U. *Furthermore, if* $k \in (0, k_1(\varepsilon)]$, *then* $N(\mathcal{A}, \varepsilon) \supset \mathcal{A}_k$. *Otherwise said,*

$$\lim_{k\to 0} \delta(\mathcal{A}_k, \mathcal{A}) = 0. \tag{4.2}$$

Proof. The proof is similar to [14, Proposition 2.2, Theorem 2.4]. □

We remark that once the attractors $\mathcal{A}_k$ have been shown to exist, it is the convergence of these sets to $\mathcal{A}$ in Hausdorff semidistance (see Definition 4.2) that shows their upper semicontinuity at $k = 0$. Lower semicontinuity is equivalent to the convergence of $\delta(\mathcal{A}, \mathcal{A}_k)$ to 0; however, as is discussed in [13] and [15], far more information is required to ensure this latter property.

The next two lemmas verify that the hypotheses of Theorem 4 are satisfied by the monoid $\mathbf{1}S(t)\mathbf{Y}^T$ and the semigroup S_k^n, defined in §3, provided that $S(t)$, defined by (3.1), possesses a local attractor.

Lemma 4.2. *Suppose that for the semigroup* $S(t)$, *defined on* $\mathcal{D}(A^\gamma)$ *by* (3.1), *there exist* $\delta_0 > 0$ *and bounded open sets* $U, W \subset \mathcal{D}(A^\gamma)$ *such that, for all* $t \geq 0$, $S(t)N(U, \delta_0) \subseteq W$. *Suppose also that* $S(t)$ *has a local attractor* $\mathcal{A}$ *for* $N(U, \delta_0)$. *Suppose, furthermore, that* S_k^n *is the semigroup corresponding to* (3.8), *and* $\mathbf{1}S(t)\mathbf{Y}^T$ *is the monoid corresponding to* (3.10).

Then there exist a bounded set $V \subset \mathcal{D}(A^\gamma)^q$ *and constants* $\delta_1 > 0$ *and* $k_0 > 0$ *such that the following statements are true*:

(i) $\mathbf{1}\mathcal{A}$ *is a local attractor for* $\mathbf{1}S(t)\mathbf{Y}^T$ *in* $N(\mathcal{D}(A^\gamma)^q; \mathbf{1}U, \delta_1)$;

(ii) $\mathbf{1}S(t)\mathbf{Y}^T N(\mathcal{D}(A^\gamma)^q; \mathbf{1}U, \delta_1) \subset V$, *for all* $t \geq 0$;

(iii) *There exists* $t_0 > 0$, *such that* $S_k^n N(\mathscr{D}(A^\gamma)^q; \mathbf{1}U, \delta_1) \subset V$ *for all* $n \leq t_0/k$, *for all* $k \in (0, k_0]$;

(iv) *The family* $\{S_k^n\}_{k\in(0,k_0]}$ *conditionally approximates* $\mathbf{1}S(t)\mathbf{Y}^T$ *on* V, *uniformly on all compact subintervals* $I \subset (0, \infty)$.

Proof. For (i), (ii) and (iii), let us fix $r > 0$, such that $B(0, r) \supseteq N(\mathbf{1}U, \delta_0)$, and let $\delta_1 = \delta_0/\|\mathbf{Y}\|$, $V = N(\mathscr{D}(A^\gamma)^q; \mathbf{1}W, 2C_0r + 1)$, where C_0 is the constant in the statement of Theorems 1 and 3. Since $\mathbf{Y}^T N(\mathscr{D}(A^\gamma)^q; \mathbf{1}U, \delta_1) \subseteq N(\mathscr{D}(A^\gamma); U, \delta_0)$, (i) and (ii) follow from the hypotheses.

Given some arbitrary $t_0 \in (0, \infty)$, the proof of Theorem 3 implies that, provided k is sufficiently small, S_k^n is defined on $B(0, r)$, and $\|A^\gamma(\mathbf{U}(nk) - \mathbf{U}^n)\| \leq 2C_0r + 1$, for all $nk \in [0, t_0]$. This implies property (iii).

To prove (iv), we now take r to be such that $B(0, r) \supseteq V$, $T = \sup(I)$ and $\eta(k, I, V)$ to be the right-hand side of (3.11). Theorem 3 shows that, when k is sufficiently small, S_k^n is defined and the norm of the error in approximating $\mathbf{1}S(nk)\mathbf{Y}^T$ on $V \times I$ is bounded by η. Since $I \subset (0, \infty)$, we have $0 < t_1 \equiv \inf(I)$, which implies that if $nk \in I$, then $1/n \leq k/t_1$. So, all the terms in the expression for η tend to 0, as k tends to 0. □

In the following lemma, we show that S_k^n is asymptotically smooth by showing that the hypotheses of Lemma 4.1 are satisfied.

Lemma 4.3. *Suppose now that, in addition to the conditions of Lemma* 4.2, *the sectorial operator* A *has compact resolvent, and obeys* (2.2) *for a strictly positive constant* a. *Then, for all sufficiently small* k, *the semigroup* S_k^n *defined by* (3.8) *is asymptotically smooth on a neighborhood of* $\mathbf{1}U$.

Proof. We recall from (3.8) that we may write

$$S_k^n\mathbf{U}^0 = \mathbf{U}^n = T^n(kA)\mathbf{U}^0 + \sum_{i=0}^{n-1} T^{n-i-1}(kA)[I + \beta_q kA]^{-1}\mathbf{F}(\mathbf{U}^i, \mathbf{U}^{i+1}).$$

Using the terminology of Lemma 4.1, we define

$$P_n\mathbf{U}^0 \equiv T^n(kA)\mathbf{U}^0,$$
$$Q_n\mathbf{U}^0 \equiv \sum_{i=0}^{n-1} T^{n-i-1}(kA)[I + \beta_q kA]^{-1}\mathbf{F}(\mathbf{U}^i, \mathbf{U}^{i+1}).$$

We will proceed to show that P_n and Q_n possess the required properties.

Inequality (2.22) implies that $\|T^n(kA)\| \leq Ce^{-ank/2}$, so P_n satisfies condition (iii) of Lemma 4.1.

Lemma 4.2 implies that, for sufficiently small k, $S_k^n\mathbf{1}U \subset V$ for all $n \in \mathbb{N}$, where V is as defined in Lemma 4.2. Hence, $\mathbf{U}^n$ is contained in a bounded subset of $\mathscr{D}(A^\gamma)^q$.

Recall from (3.1) that f maps bounded sets of $\mathscr{D}(A^\gamma)$ to bounded sets of X. This implies that $\mathbf{F}$ maps bounded sets of $\mathscr{D}(A^\gamma)^q \times \mathscr{D}(A^\gamma)^q$ to bounded sets of X^q. The mapping $\operatorname{diag}([I + \beta_q A]^{-1})$ maps bounded sets of X^q to bounded sets of $\mathscr{D}(A)^q$. Finally, for $m \geq 0$, $T^m(kA)$ is subject to the same bound (2.22), considered as a member of $\mathscr{L}(\mathscr{D}(A)^q)$, as it is considered as a member of $\mathscr{L}(X^q)$, and so is a bounded map on $\mathscr{D}(A)^q$.

Hence, Q_n maps bounded sets of $\mathscr{D}(A^\gamma)^q$ to bounded subsets of $\mathscr{D}(A)^q$. The domain $\mathscr{D}(A)$ is compactly imbedded in $\mathscr{D}(A^\gamma)$, if A has compact resolvent [16, Theorem 1.4.8], and therefore Q_n is a completely continuous mapping for all $n \in \mathbb{N}$, as required by condition (ii) of Lemma 4.1. □

Theorem 5. *Suppose that U is a bounded open subset of $\mathscr{D}(A^\gamma)$, that $\delta > 0$, and that the following hypotheses are satisfied:*

(i) *A is a sectorial operator with compact resolvent on X satisfying* (2.2) *for some constants $M \geq 1$, $a > 0$ and $\theta \in (0, \pi/2)$;*

(ii) *For some $\gamma \in [0, 1)$, $f : \mathscr{D}(A^\gamma) \to X$ is locally Lipschitz continuous;*

(iii) *The semigroup $S(t)$, defined on $\mathscr{D}(A^\gamma)$ by* (3.1), *possesses a local attractor $\mathscr{A}$ in $N(U, \delta)$;*

(iv) *There exists W, a bounded subset of $\mathscr{D}(A^\gamma)$, such that $S(t)N(U, \delta) \subseteq W$ for all $t \geq 0$;*

(v) *For $p, q \geq 1$, the pth-order q-step linear multistep method* (3.2) *is strongly $A(\alpha)$-stable for some $\alpha \in [\theta, \pi/2)$.*

Then, for sufficiently small k, the semigroup S_k^n, defined on bounded subsets of $\mathscr{D}(A^\gamma)^q$ by (3.2, 3.8), *possesses a local attractor $\mathscr{A}_k$, attracting in a neighborhood of $\mathbf{1}U$. Moreover, given $\varepsilon > 0$, there exists $k_1 = k_1(\varepsilon)$ such that for all $k \in (0, k_1]$,*

$$\mathscr{A}_k \subset N(\mathscr{D}(A^\gamma)^q ; \mathbf{1}\mathscr{A}, \varepsilon). \tag{4.3}$$

Otherwise stated,

$$\lim_{k \to 0} \delta(\mathscr{A}_k, \mathbf{1}\mathscr{A}) = 0. \tag{4.4}$$

Proof. Lemmas 4.2 and 4.3 show that all the hypotheses of Theorem 4 hold for S_k^n and the monoid $\mathbf{1}S(t)\mathbf{Y}^T$, defined by (3.8) and (3.10), respectively. Hence, Theorem 5 follows from Theorem 4. □

We remark that if $\mathbf{y} \in \mathbb{R}^q$ is such that $\mathbf{y}^T\mathbf{1} = 1$, then

$$\lim_{k \to 0} \delta(\mathbf{y}^T \mathscr{A}_k, \mathscr{A}) = 0.$$

5. Summary and conclusions

We review here how we arrived at our main result, Theorem 5, and discuss some of the implications of our work. The difficulties of directly applying the theory of [14] to multistep methods were considered in the Introduction, together with the motivation for the constructions defining the commensurable families of operators S_k^n and $\mathbf{v}S(t)\mathbf{y}^T$ on the q-fold product space for a q-step method. These constructions having been defined, a slight generalization of the main result of [14] was necessary to take account of the fact that $\mathbf{v}\mathbf{y}^T$, the identity of $\mathbf{v}S(t)\mathbf{y}^T$, is not the same as the identity of the space on which it operates, as is required by the usual definition of a semigroup. Such deficient semigroups were termed monoids in Definition 4.3, and our modification of the main result of [14] was stated as Theorem 4.

From an abstract point of view, one might have chosen to stop once this background was established. However, as pointed out in the Introduction, applying the theory—that is, verifying the hypotheses of Theorem 4—poses more difficulties than is generally true in the one-step case. Hence, a broad example was chosen—linear multistep methods approximating sectorial evolution equations—and the theory was applied in this context.

The main points needing to be verified to apply Theorem 4 were that, in a neighborhood of $\mathbf{1}\mathscr{A}$,

(i) S_k^n is well defined,

(ii) that it approximates $\mathbf{1}S(t)\mathbf{Y}^T$ uniformly, and

(iii) that it is asymptotically smooth.

The proof of Lemma 4.2 established that the abstract conditions (i) and (ii) may be restated in terms of vectors $\mathbf{U}^n$ and $\mathbf{U}(t)$ representing, respectively, discrete and continuous trajectories from a common initial data point $\mathbf{U}^0$, taken from a neighborhood of $\mathbf{1}\mathscr{A}$. This in turn depended on Theorem 3, which showed that the trajectory $\{\mathbf{U}^n\}_{n\geq 0}$ is well defined, for sufficiently small k, and bounded the error made in approximating $\mathbf{U}(t)$.

Although the proof of asymptotic smoothness, established through Lemmas 4.1 and 4.3, was relatively short, it did require two restrictive conditions:

(i) That A obeys (2.2), with strictly positive a;

(ii) That the operator A has compact resolvent.

The first, (i), was necessary to conclude, from Theorem 1, that $\lim_{n\to\infty} \|T^n(kA)\| = 0$. On the other hand, (ii) was needed to show that $S_k^n - T^n(kA)$ is compact. These two conditions were essential requirements of Lemma 4.1. We note that the error bound given by Theorem 3 might have been obtained under weaker assumptions on the method, but the full force of Theorem 1, for which strong $A(\alpha)$-stability was necessary, was also required for Lemma 4.3.

Potentially of some interest is the style of error bound considered in Theorems 1 and 3, despite the fact that they are singular at $t = 0$ and suboptimal for $p > 1$. Theoretical necessity forced us to consider a numerical method in the absence of a starting method. The error bounds derived may be interpreted variously. They still imply a conventional error bound for the approximation of a particular trajectory starting at u_0, if $\mathbf{U}^0$ is chosen so that $\mathbf{Y}^T\mathbf{U}^0 = u_0$. On the other hand, if one considers the trajectories generated by a linear multistep method from fixed initial data $\mathbf{U}^0$, for various values of k, then Theorem 3 shows that for positive t, the numerical trajectories converge towards the continuous trajectory $u(t)$ given by the solution of (3.1) for initial data $u_0 = \mathbf{Y}^T\mathbf{U}^0$, as k tends to 0.

In asymptotic approximation one is frequently concerned not with approximating a specific trajectory, but rather with approximating a set of continuous trajectories by a set of discrete trajectories. With this interpretation, the bounds of Theorem 1 and 3 are quite natural, since it is clearly unnecessary in such a case to expend effort approximating a specific trajectory closely at $t = 0$ using a starting method. We also remark that whilst the singularity at $t = 0$ is an essential property of the bounds of Theorems 1 and 3, we believe that suboptimality is not. In a future paper, we will consider optimal-order error estimates for multistep approximations of e^{-At} in the absence of a start-up procedure. However, currently we do not know how to obtain an optimal bound in the nonlinear case for $p > 1$.

Acknowledgments

The authors are grateful for helpful discussions with Arieh Iserles and Tony Ware.

Bibliography

1. W.-J. Beyn, *On invariant closed curves for one-step methods*, Numer. Math. **51** (1987), 103–122.
2. J. C. Butcher, *On the convergence of numerical solutions to ordinary differential equations*, Math. Comp. **20** (1966), 1–10.
3. M. Crouzeix, *On multistep approximation of semigroups in Banach space*, J. Comput. Appl. Math. **20** (1987), 25–35.
4. J. Dieudonné, *Treatise on analysis,* Vol. III, Academic Press, New York, 1972.
5. N. Dunford and J. T. Schwartz, *Linear operators Part* I: *General theory*, Wiley-Interscience, New York, 1958.
6. T. Eirola and O. Nevanlinna, *What do multistep methods approximate*?, Numer. Math. **53** (1988), 559–569.
7. C. M. Elliott and A. M. Stuart, *The global dynamics of discrete semilinear parabolic equations*, SIAM J. Numer. Anal. **30** (1993), 1622–1663.
8. A. Friedman, *Partial differential equations*, Holt, New York, 1969.
9. D. F. Griffiths and A. R. Mitchell, *Stable periodic bifurcations of an explicit discretization of a nonlinear partial differential equation in reaction-diffusion*, IMA J. Numer. Anal. **8** (1988), 435–454.
10. E. Hairer, A. Iserles, and J. M. Sanz-Serna, *Equilibria of Runge-Kutta methods*, Numer. Math. **58** (1990), 243–254.
11. E. Hairer, S. P. Nørsett, and G. Wanner, *Solving ordinary differential equations,* Vol. I, Springer, Berlin, 1987.
12. E. Hairer and G. Wanner, *Solving ordinary differential equations,* Vol. II, Springer, Berlin, 1991.
13. J. K. Hale, *Asymptotic behaviour of dissipative systems*, Math. Surveys Monographs, vol. 25, Amer. Math. Soc., Providence, RI, 1988.
14. J. K. Hale, X.-B. Lin, and G. Raugel, *Upper semicontinuity of attractors for approximations of semigroups and partial differential equations*, Math. Comp. **50** (1988), 89–123.
15. J. K. Hale and G. Raugel, *Lower semicontinuity of attractors of gradient systems and applications*, Ann. Mat. Pura Appl. **154** (1989), 281–326.
16. D. Henry, *Geometric theory of semilinear parabolic equations*, Lecture Notes in Math., vol. 840, Springer, Berlin, 1981.
17. A. T. Hill, *Attractors for nonlinear convection-diffusion equations and their numerical approximation*, D. Phil. Thesis, Oxford, 1992.
18. A. T. Hill and E. Süli, *Upper semicontinuity of attractors for linear multistep methods*, University of Bath Mathematics Preprint 94/11.
19. E. Isaacson and H. B. Keller, *Analysis of numerical methods*, Wiley, New York, 1966.
20. A. Iserles, *Stability and dynamics for nonlinear ordinary differential equations*, IMA J. Numer. Anal. **10** (1990), 1–30.
21. A. Iserles, A. T. Peplow, and A. M. Stuart, *A unified approach to spurious solutions introduced by time discretization. Part* I: *basic theory*, SIAM J. Numer. Anal. **28** (1991), 1723–1751.
22. A. Iserles and A. M. Stuart, *Unified approach to spurious solutions introduced by time-discretization. Part* II: *BDF-like methods*, IMA J. Numer. Anal. **12** (1992), 487–502.

23. U. Kirchgraber, *Multistep methods are essentially one-step methods*, Numer. Math. **48** (1986), 85–90.

24. U. Kirchgraber, F. Lasagni, K. Nipp, and D. Stoffer, *On the application of the invariant manifold theory, in particular to numerical analysis*, Internat. Ser. Numer. Math. **97** (1991), 189–197.

25. P. E. Kloeden and J. Lorenz, *Stable attracting sets in dynamical systems and in their one-step discretizations*, SIAM J. Numer. Anal. **23** (1986), 986–995.

26. ——, *A note on multistep methods and attracting sets of dynamical systems*, Numer. Math. **56** (1990), 667–673.

27. M.-N. Le Roux, *Semi-discrétisation en temps pour les équations d'évolution paraboliques lorsque l'opérateur dépend du temps*, RAIRO Anal. Numér. **13** (1979), 119–137.

28. C. Lubich and O. Nevanlinna, *On resolvent conditions and stability estimates*, BIT **31** (1991), 293–313.

29. C. Palencia, *Stability of rational multistep approximations of holomorphic semigroups*, Math. Comp. **64** (1995), 591–599.

30. A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Appl. Math. Sci., vol. 44, Springer, Berlin, 1983.

31. R. Skeel, *Analysis of fixed-stepsize methods*, SIAM J. Numer. Anal. **13** (1976), 664–685.

32. D. Stoffer, *General linear methods: connection to one step methods and invariant curves*, Numer. Math. **64** (1993), 395–407.

33. A. M. Stuart, *Nonlinear instability in dissipative finite difference schemes*, SIAM Rev. **31** (1989), 191–220.

School of Mathematical Sciences, University of Bath, Claverton Down, Bath BA2 7AY, United Kingdom

E-mail address: `ath@maths.bath.ac.uk`

Oxford University Computing Laboratory, Numerical Analysis Group, Wolfson Building, Parks Road, Oxford OX1 3QD, United Kingdom

E-mail address: `Endre.Suli@comlab.ox.ac.uk`