# NUMERICAL ANALYSIS OF SEMILINEAR PARABOLIC PROBLEMS

STIG LARSSON

ABSTRACT. In these lectures I discuss error analysis techniques for finite element methods for systems of reaction-diffusion equations with applications in dynamical systems theory. The emphasis is on pedagogical aspects and analysis techniques rather than on results. The list of techniques discussed include: analytic semigroup, parabolic smoothing, non-smooth data error estimate, a priori error estimate, a posteriori error estimate, exponential dichotomy, shadowing.

## CONTENTS



## 1. THE CONTINUOUS PROBLEM

We consider the following initial-boundary value problem for a system of reaction-diffusion equations,

$$\begin{aligned} u_t - \epsilon\Delta u &= \tilde{f}(u), && x \in \Omega,\ t > 0, \\ u &= 0, && x \in \partial\Omega,\ t > 0, \\ u(\cdot, 0) &= u_0, && x \in \Omega, \end{aligned} \tag{1.1}$$

Lectures delivered at the EPSRC VIII-th Summer School in Numerical Analysis, University of Leicester, July 12–17, 1998. Published in "The Graduate Student's Guide to Numerical Analysis '98", editors M. Ainsworth, J. Levesley, and M. Marletta, SSCM volume 26, Springer-Verlag, 1999, pp. 83–117.

where $\Omega$ is a bounded domain in $\mathbf{R}^d$, $d = 1,2,3$, $u = u(x,t) \in \mathbf{R}^p$, $u_t = \partial u/\partial t$, $\Delta u = \sum_{i=1}^d \partial^2 u/\partial x_i^2$, $\epsilon = \operatorname{diag}(\epsilon_1, \dots, \epsilon_p)$ is a diagonal matrix of constant coefficients $\epsilon_j > 0$, and $\tilde f : \mathbf{R}^p \to \mathbf{R}^p$ is twice continuously differentiable. If $d = 2,3$ we assume, in addition, that the derivatives of $\tilde f$ satisfy the growth condition

$$|\tilde f^{(l)}(\xi)| \le C(1 + |\xi|^{\delta+1-l}), \quad \xi \in \mathbf{R}^p,\ l = 1,2, \tag{1.2}$$

where $|\cdot|$ denotes the Euclidean norm on $\mathbf{R}^p$ and the induced operator norms, and where $\delta = 2$ if $d = 3$, $\delta \in [1,\infty)$ if $d = 2$.

We assume that $\Omega$ is a convex polygon, so that we have access to the elliptic regularity theory and so that finite element meshes can be fitted exactly to the domain.

In the sequel we use the Hilbert space $H = (L_2(\Omega))^p$, with its standard norm and inner product

$$\|v\| = \Big( \int_\Omega |v|^2\,dx \Big)^{1/2}, \quad (v,w) = \int_\Omega v \cdot w\,dx. \tag{1.3}$$

The norms in the Sobolev spaces $(H^m(\Omega))^p$, $m \ge 0$, are denoted by

$$\|v\|_m = \left( \sum_{|\alpha| \le m} \|D^\alpha v\|^2 \right)^{1/2}. \tag{1.4}$$

The space $V = (H_0^1(\Omega))^p$, with norm $\|\cdot\|_1$, consists of the functions in $(H^1(\Omega))^p$ that vanish on $\partial\Omega$. $V^* = (H^{-1}(\Omega))^p$ is the dual space of $V$ with norm

$$\|v\|_{-1} = \sup_{\chi \in V} \frac{|(v,\chi)|}{\|\chi\|_1}. \tag{1.5}$$

If $X, Y$ are Banach spaces, then $\mathcal{L}(X,Y)$ denotes the space of bounded linear operators from $X$ into $Y$, $\mathcal{L}(X) = \mathcal{L}(X,X)$, and $B_X(x,R)$ denotes the closed ball in $X$ with center $x$ and radius $R$. In particular, we let $B_R = B_V(0,R)$ denote the the closed ball of radius $R$ in $V$:

$$B_R = \{v \in V : \|v\|_1 \le R\}.$$

We also use the notation

$$\|v\|_{L_\infty([0,T],X)} = \sup_{t \in [0,T]} \|v(t)\|_X.$$

We set the problem up in the framework of [6]. We define the unbounded operator $A = -\epsilon\Delta$ on $H$ with domain of definition $\mathcal{D}(A) = (H^2(\Omega) \cap H_0^1(\Omega))^p$. Then $A$ is a closed, densely defined, and self-adjoint positive definite operator in $H$ with compact inverse. Moreover, our assumption (1.2) guarantees that the mapping $\tilde f$ induces a nonlinear operator $f : V \to H$ through $f(v)(x) = \tilde f(v(x))$, see Lemma 1.1 below. The initial-boundary value problem (1.1) may then be formulated as an initial value problem in $V$: find $u(t) \in V$ such that

$$u' + Au = f(u),\ t > 0; \quad u(0) = u_0. \tag{1.6}$$

The operator $-A$ is the infinitesimal generator of the analytic semigroup $E(t) = \exp(-tA)$ defined by

$$E(t)v = \sum_{j=1}^\infty e^{-t\lambda_j}(v,\varphi_j)\varphi_j, \quad v \in H, \tag{1.7}$$

where $\lambda_j$ and $\varphi_j$ denote the eigenvalues and an orthonormal basis of eigenvectors of $A$. The semigroup $E(t)$ is the solution operator of the initial value problem for the homogeneous equation,

$$u' + Au = 0,\ t > 0; \quad u(0) = u_0.$$

Its solution is thus given by $u(t) = E(t)u_0$. By Duhamel's principle it follows that solutions of (1.6) satisfy

$$u(t) = E(t)u_0 + \int_0^t E(t-s)f(u(s))\,ds, \quad t \geq 0. \tag{1.8}$$

Conversely, appropriately defined solutions of the nonlinear integral equation (1.8) are solutions of the differential equation (1.6); see Theorem 1.3 below. We shall mainly work with (1.8) and discretized variants of it.

An important ingredient in our framework is that $f$ can be controlled by fractional powers of $A$. We define the fractional powers of $A$ by means of the spectral theorem and we have

$$\begin{aligned} \|A^\alpha v\| &= \Big( \sum_{j=1}^{\infty} \Big( \lambda_j^\alpha (v, \varphi_j) \Big)^2 \Big)^{1/2}, \quad \alpha \in \mathbf{R}, \\ \mathcal{D}(A^\alpha) &= \Big\{ v : \|A^\alpha v\| < \infty \Big\}, \quad \alpha \in \mathbf{R}. \end{aligned} \tag{1.9}$$

Using also the elliptic regularity estimate,

$$\|v\|_2 \leq C\|Av\|, \quad v \in (H^2(\Omega) \cap H_0^1(\Omega))^p, \tag{1.10}$$

and the trace inequality,

$$\|v\|_{L_2(\partial\Omega)} \leq C\|v\|_1, \quad v \in (H^1(\Omega))^p,$$

we obtain $\mathcal{D}(A^{l/2}) = (H^l(\Omega) \cap H_0^1(\Omega))^p$, $l = 1, 2$, with the equivalence of norms

$$c\|v\|_l \leq \|A^{l/2}v\| \leq C\|v\|_l, \quad v \in \mathcal{D}(A^{l/2}),\ l = 1, 2; \tag{1.11}$$

cf. [12]. A simple exercise using the special case $l = 1$ of (1.11) shows

$$c\|v\|_{-1} \leq \|A^{-1/2}v\| \leq C\|v\|_{-1}; \tag{1.12}$$

cf. (2.7) below.

The analyticity of the semigroup is reflected in the inequalities ($D_t = \partial/\partial t$)

$$\|D_t^l E(t)v\| = \|A^l E(t)v\| \leq C_l t^{-l}\|v\|, \quad t > 0,\ v \in H,\ l \geq 0. \tag{1.13}$$

These follow easily from (1.7) and Parseval's identity. Using also the norm equivalences (1.11), (1.12) we obtain the smoothing property

$$\begin{aligned} \|D_t^l E(t)v\|_\beta \leq C_l t^{-l-(\beta-\alpha)/2}\|v\|_\alpha, \quad & t > 0,\ v \in \mathcal{D}(A^{\alpha/2}), \\ & -1 \leq \alpha \leq \beta \leq 2,\ l = 0, 1. \end{aligned} \tag{1.14}$$

We need the fact that the operator $f : V \to H$ satisfies a local Lipschitz condition. This is contained in our first lemma. We also prove that the Fréchet derivative $f' : V \to \mathcal{L}(V, H)$ satisfies a local Lipschitz condition. The proof is based on the assumption (1.2) and the Sobolev inequality (where $p = 6$ if $d \leq 3$, $p < \infty$ if $d = 2$, and $p = \infty$ if $d = 1$)

$$\|v\|_{L_p} \leq C\|v\|_1, \tag{1.15}$$

and on Hölder's inequality

$$\|v^\delta w\|_{L_r} \le \|v\|_{L_q}^\delta \|w\|_{L_p}, \quad \tfrac{\delta}{q} + \tfrac{1}{p} = \tfrac{1}{r},\ \delta > 0. \tag{1.16}$$

**Lemma 1.1.** *For each nonnegative number $R$ there is a constant $C(R)$ such that, for all $u, v, w \in B_R$, $l = 0, 1$,*

$$\|f'(u)\|_{\mathcal{L}(V,H)} \le C(R), \tag{1.17}$$

$$\|f'(u)\|_{\mathcal{L}(H,V^*)} \le C(R), \tag{1.18}$$

$$\|f(u) - f(v)\| \le C(R)\|u - v\|_1, \tag{1.19}$$

$$\|f(u) - f(v)\|_{-1} \le C(R)\|u - v\|, \tag{1.20}$$

$$\|f'(u) - f'(v)\|_{\mathcal{L}(V,H)} \le C(R)\|u - v\|_1, \tag{1.21}$$

$$\|f'(u) - f'(v)\|_{\mathcal{L}(H,V^*)} \le C(R)\|u - v\|_1, \tag{1.22}$$

$$\|f(u) - f(v) - f'(w)(u - v)\|_{l-1} \le C(R)\big(\|u - w\|_1 + \|v - w\|_1\big)\|u - v\|_l. \tag{1.23}$$

*If, in addition, $u \in (H^2(\Omega))^p$ and $z \in (H^2(\Omega))^p \cap V$, then*

$$\|A^{1/2}(f'(u)z)\| \le C(R)\big(\|z\|_2 + \|u\|_2\|z\|_1\big). \tag{1.24}$$

*Proof.* We have in view of (1.2) and the Hölder and Sobolev inequalities, for $z \in V$,

$$\begin{aligned} \|f'(u)z\|_{L_2} &\le C\big(1 + \|u\|_{L_q}^\delta\big)\|z\|_{L_p} \\ &\le C\big(1 + \|u\|_1^\delta\big)\|z\|_1, \end{aligned}$$

where $\frac{1}{p} + \frac{\delta}{q} = \frac{1}{2}$ with $p = q = 6$ if $d = 3$, and with arbitrary $p \in (1, \infty)$ if $d \le 2$. This proves (1.17) and (1.19) follows. Similarly, for any $z \in V$,

$$\begin{aligned} \|(f'(u) - f'(v))z\|_{L_2} &\le C\big(1 + \|u\|_{L_q}^{\delta-1} + \|v\|_{L_q}^{\delta-1}\big)\|u - v\|_{L_p}\|z\|_{L_q} \\ &\le C\big(1 + \|u\|_1^{\delta-1} + \|v\|_1^{\delta-1}\big)\|u - v\|_1\|z\|_1, \end{aligned}$$

with the same $p$ and $q$ as before. This proves (1.21).

Moreover, for any $z, \chi \in V$,

$$\begin{aligned} (f'(u)z, \chi) &\le C\big(1 + \|u\|_{L_q}^\delta\big)\|z\|_{L_2}\|\chi\|_{L_p} \\ &\le C\big(1 + \|u\|_1^\delta\big)\|z\|\,\|\chi\|_1, \end{aligned}$$

and

$$\begin{aligned} ((f'(u) - f'(v))z, \chi) &\le C\big(1 + \|u\|_{L_q}^{\delta-1} + \|v\|_{L_q}^{\delta-1}\big)\|u - v\|_{L_2}\|z\|_{L_q}\|\chi\|_{L_p} \\ &\le C\big(1 + \|u\|_1^{\delta-1} + \|v\|_1^{\delta-1}\big)\|u - v\|\|z\|_1\|\chi\|_1, \end{aligned}$$

where $\frac{\delta}{q} + \frac{1}{2} + \frac{1}{p} = 1$, i.e., with the same $p$ and $q$ as before. This proves (1.18) and (1.22); (1.20) follows from (1.18).

Next, (1.23) is obtained by applying (1.21) and (1.22) to the identity

$$f(u) - f(v) - f'(w)(u - v) = \int_0^1 \Big(f'(su + (1 - s)v) - f'(w)\Big)\, ds\, (u - v).$$

Finally, (1.24) is proved in a similar way, by using the equivalence of norms (1.11) and computing the first order partial derivatives of $f'(u)z$. Note that this uses the fact that $f'(u)z \in V$, i.e., $f'(u)z = 0$ on $\partial\Omega$. □

We may now prove local existence of solutions of (1.8) and hence of (1.6).

**Theorem 1.2.** *For any $R_0 > 0$ there is $\tau = \tau(R_0)$ such that (1.8) has a unique solution $u \in C([0,\tau], V)$ for any initial value $u_0 \in V$ with $\|u_0\|_1 \le R_0$. Moreover, there is $c$ such that $\|u\|_{L_\infty([0,\tau],V)} \le cR_0$.*

*Proof.* Let $u_0 \in B_{R_0}$, define

$$\mathcal{S}(u)(t) = E(t)u_0 + \int_0^t E(t-s)f(u(s))\,ds,$$

and note that (1.8) is a fixed point equation $u = \mathcal{S}(u)$. We shall choose $\tau$ and $R$ such that we can apply Banach's fixed point theorem (the contraction mapping theorem) in the closed ball

$$\mathcal{B} = \{u \in C([0,\tau], V) : \|u\|_{L_\infty([0,\tau],V)} \le R\}$$

in the Banach space $C([0,\tau], V)$.

We must show (i) that $\mathcal{S}$ maps $\mathcal{B}$ into itself, (ii) that $\mathcal{S}$ is a contraction on $\mathcal{B}$. In order to prove (i) we take $u \in \mathcal{B}$ and first note that the Lipschitz condition (1.19) implies that

$$\begin{aligned} \|f(u(t))\| &\le \|f(0)\| + \|f(u(t)) - f(0)\| \\ &\le \|f(0)\| + C(R)\|u(t)\|_1 \\ &\le \|f(0)\| + C(R)R, \quad 0 \le t \le \tau. \end{aligned} \tag{1.25}$$

Hence, using also (1.14), we get

$$\begin{aligned} \|\mathcal{S}(u)(t)\|_1 &\le \|E(t)u_0\|_1 + \int_0^t \|E(t-s)f(u(s))\|_1\,ds \\ &\le c_0\|u_0\|_1 + c_1\int_0^t (t-s)^{-1/2}\|f(u(s))\|\,ds \\ &\le c_0R_0 + 2c_1\tau^{1/2}\big(\|f(0)\| + C(R)R\big), \quad 0 \le t \le \tau. \end{aligned}$$

This implies

$$\|\mathcal{S}(u)\|_{L_\infty([0,\tau],V)} \le c_0R_0 + 2c_1\tau^{1/2}\big(\|f(0)\| + C(R)R\big).$$

Choose $R = 2c_0R_0$ and $\tau = \tau(R_0)$ so small that

$$2c_1\tau^{1/2}\big(\|f(0)\| + C(R)R\big) \le \tfrac{1}{2}R. \tag{1.26}$$

Then $\|\mathcal{S}(u)\|_{L_\infty([0,\tau],V)} \le R$ and we conclude that $\mathcal{S}$ maps $\mathcal{B}$ into itself.

To show (ii) we take $u, v \in \mathcal{B}$ and note that

$$\|f(u(t)) - f(v(t))\| \le C(R)\|u - v\|_{L_\infty([0,\tau],V)}, \quad 0 \le t \le \tau.$$

Hence

$$\begin{aligned} \|\mathcal{S}(u)(t) - \mathcal{S}(v(t))\|_1 &\le \int_0^t \|E(t-s)(f(u(s)) - f(v(s)))\|_1\,ds \\ &\le c_1\int_0^t (t-s)^{-1/2}\|f(u(s)) - f(v(s))\|\,ds \\ &\le 2c_1\tau^{1/2}C(R)\|u - v\|_{L_\infty([0,\tau],V)}, \quad 0 \le t \le \tau, \end{aligned}$$

so that

$$\|\mathcal{S}(u) - \mathcal{S}(v)\|_{L_\infty([0,\tau],V)} \le 2c_1\tau^{1/2}C(R)\|u - v\|_{L_\infty([0,\tau],V)}.$$

It follows from (1.26) that $2c_1\tau^{1/2}C(R) \le \frac{1}{2}$ and we conclude that $\mathcal{S}$ is a contraction on $\mathcal{B}$. Hence $\mathcal{S}$ has a unique fixed point $u \in \mathcal{B}$. □

The initial value problem (1.8) (or (1.6)) thus has a unique local solution for any initial datum $u_0 \in V$. We denote by $S(t,\cdot)$ the corresponding (local) solution operator, so that $u(t) = S(t,u_0)$ is the solution of (1.8).

The following theorem provides regularity estimates for solutions $u$ of (1.8). These will be used in our error analysis, but the theorem also shows that $u'(t) \in H$ and $Au(t) \in H$ for $t > 0$, so that any solution of the integral equation (1.8) is also a solution of the differential equation (1.6).

**Theorem 1.3.** *Let $R \geq 0$ and $\tau > 0$ be given and let $u \in C([0,\tau],V)$ be a solution of* (1.8). *If $\|u(t)\|_1 \leq R$ for $t \in [0,\tau]$, then*

$$\|u(t)\|_2 \leq C(R,\tau)t^{-1/2}, \qquad t \in (0,\tau], \tag{1.27}$$

$$\|u_t(t)\|_s \leq C(R,\tau)t^{-1-(s-1)/2}, \qquad t \in (0,\tau],\ s = 0,1,2. \tag{1.28}$$

*Proof.* In view of (1.19) we have $\|f(u(t))\| \leq C(R)$ for $t \in [0,\tau]$, cf. (1.25). The proof of (1.27), (1.28) can now obtained by tracing the constants in [6, Theorem 3.5.2]. The argument is based on a generalization of Gronwall's lemma, Lemma 1.4 below. The case $s = 2$ of (1.28) requires that we differentiate (1.6) with respect to $t$ and use (1.24). □

**Lemma 1.4.** (Generalized Gronwall lemma.) *Let the function $\varphi(t,\tau) \geq 0$ be continuous for $0 \leq \tau < t \leq T$. If*

$$\varphi(t,\tau) \leq A\,(t-\tau)^{-1+\alpha} + B\int_\tau^t (t-s)^{-1+\beta}\varphi(s,\tau)\,ds, \quad 0 \leq \tau < t \leq T,$$

*for some constants $A, B \geq 0$, $\alpha,\beta > 0$, then there is a constant $C = C(B,T,\alpha,\beta)$ such that*

$$\varphi(t,\tau) \leq CA\,(t-\tau)^{-1+\alpha}, \quad 0 \leq \tau < t \leq T.$$

*Proof.* Iterating the given inequality $N-1$ times, using the identity

$$\int_\tau^t (t-s)^{-1+\alpha}(s-\tau)^{-1+\beta}\,ds = C(\alpha,\beta)\,(t-\tau)^{-1+\alpha+\beta}, \quad \alpha,\beta > 0, \tag{1.29}$$

(Abel's integral) and estimating $(t-\tau)^\beta$ by $T^\beta$, we obtain

$$\varphi(t,\tau) \leq C_1 A\,(t-\tau)^{-1+\alpha} + C_2\int_\tau^t (t-s)^{-1+N\beta}\varphi(s,\tau)\,ds, \quad 0 \leq \tau < t \leq T,$$

where $C_1 = C_1(B,T,\alpha,\beta,N)$, $C_2 = C_2(B,\beta,N)$. We now choose the smallest $N$ such that $-1+N\beta \geq 0$, and estimate $(t-s)^{-1+N\beta}$ by $T^{-1+N\beta}$. If $-1+\alpha \geq 0$ we obtain the desired conclusion by the standard version of Gronwall's lemma. Otherwise we set $\psi(t,\tau) = (t-\tau)^{1-\alpha}\varphi(t,\tau)$ to obtain

$$\psi(t,\tau) \leq C_1 A + C_3\int_\tau^t (s-\tau)^{-1+\alpha}\psi(s,\tau)\,ds, \qquad 0 \leq \tau < t \leq T,$$

and the standard Gronwall lemma yields $\psi(t,\tau) \leq CA$ for $0 \leq \tau < t \leq T$, which is the desired result. □

Note that the constant in Gronwall's lemma grows exponentially with the length $T$ of the time interval. Hence, results derived by means of this lemma are often useful only for short time intervals. There is also a discrete version of Lemma 1.4; see [3, Lemma 7.1].

We conclude this section with a remark about the abstract framework. It is tempting to start by saying "Let $X$ be a Banach space and $A$ the generator of an analytic semigroup in X." This approach goes a rather long way with the continuous problem (see [6]) and with pure time discretization. But when it comes to spatial discretization we soon need to know that, for example, $X = L_2(\Omega)$ and $A$ is an elliptic operator of second order. "PDE theory is not a branch of functional analysis" as L. C. Evans puts it in the preface of his textbook [5].

## 2. Local a priori error estimates

In this section we first discretize (1.1) with respect to the spatial variables by means of a standard piecewise linear finite element method. We then briefly discuss completely discrete approximation by means of the backward Euler time-stepping. This material is taken from [7]. A general reference is [12].

2.1. **The spatially semidiscrete problem.** The weak formulation of (1.1) is: find $u(t) \in V$ such that

$$\begin{aligned} (u', v) + a(u, v) &= (f(u), v), \quad \forall v \in V,\ t > 0, \\ u(0) &= u_0, \end{aligned} \tag{2.1}$$

where $a(u, v) = (\epsilon\nabla u, \nabla v) = (-\epsilon\Delta u, v) = (Au, v)$ is the bilinear form associated with $A$.

Let $\{V_h\}_{0<h<1}$ be a family of finite dimensional subspaces of $V$, where each $V_h$ consists of continuous piecewise polynomials of degree $\leq 1$ with respect to a triangulation $\mathcal{T}$ of $\Omega$ with maximal mesh size $h$, see [2]. The approximate solution $u_h(t) \in V_h$ of (1.1) is defined by

$$\begin{aligned} (u_h', \chi) + a(u_h, \chi) &= (f(u_h), \chi), \quad \forall \chi \in V_h,\ t > 0, \\ u_h(0) &= u_{h,0}, \end{aligned} \tag{2.2}$$

where $u_{h,0} \in V_h$ is an approximation of $u_0$.

Introducing the linear operator $A_h : V_h \to V_h$ and the orthogonal projection $P_h : H \to V_h$, defined by

$$(A_h\psi, \chi) = a(\psi, \chi), \quad (P_h g, \chi) = (g, \chi) \quad \forall \psi, \chi \in V_h,\ g \in H, \tag{2.3}$$

we may write (2.2) as

$$u_h' + A_h u_h = P_h f(u_h),\ t > 0; \quad u_h(0) = u_{h,0}. \tag{2.4}$$

The operator $A_h$ is self-adjoint positive definite (uniformly in $h$); the corresponding semigroup $E_h(t) = \exp(-tA_h) : V_h \to V_h$ therefore satisfies inequalities analogous to (1.13) (uniformly in $h$):

$$\|D_t^l E_h(t)v\| = \|A_h^l E_h(t)v\| \leq C_l t^{-l}\|v\|, \quad t > 0,\ v \in V_h,\ l \geq 0.$$

Moreover, for the operator $A_h$ we have the equivalence of norms (cf. (1.11))

$$c\|v\|_1 \leq \|A_h^{1/2}v\| = \sqrt{a(v, v)} = \|A^{1/2}v\| \leq C\|v\|_1, \quad v \in V_h. \tag{2.5}$$

We also have

$$\|P_h f\| \leq \|f\|, \quad f \in H,$$

and (cf. (1.12))

$$\|A_h^{-1/2}P_h f\| \leq C\|f\|_{-1}, \quad f \in H, \tag{2.6}$$

which is obtained by using (2.5) in the calculation

$$
\begin{aligned}
\|A_h^{-1/2}P_hf\| &= \sup_{v_h\in V_h}\frac{|(A_h^{-1/2}P_hf,v_h)|}{\|v_h\|} = \sup_{v_h\in V_h}\frac{|(f,A_h^{-1/2}v_h)|}{\|v_h\|}\\
&= \sup_{w_h\in V_h}\frac{|(f,w_h)|}{\|A_h^{1/2}w_h\|} \le C\sup_{w_h\in V_h}\frac{|(f,w_h)|}{\|w_h\|_1}\\
&\le C\sup_{w\in V}\frac{|(f,w)|}{\|w\|_1} = C\|f\|_{-1}.
\end{aligned}
\tag{2.7}
$$

Using these inequalities we easily prove the smoothing property of $E_h(t)P_h$:

$$
\begin{aligned}
\|D_t^lE_h(t)P_hf\|_\beta \le Ct^{-l-(\beta-\alpha)/2}\|f\|_\alpha, \quad t>0,\ f\in\mathcal{D}(A^{\alpha/2}),\\
-1\le\alpha\le\beta\le 1,\ l=0,1.
\end{aligned}
\tag{2.8}
$$

Note that the upper limit to $\beta$ is 1, while it is 2 in the continuous case (1.14).

The initial-value problem (2.4) is equivalent to the integral equation

$$
u_h(t) = E_h(t)u_{h,0} + \int_0^t E_h(t-s)P_hf(u_h(s))\,ds, \quad t\ge 0. \tag{2.9}
$$

Because of (2.8) the proof of Theorem 1.2 carries over verbatim to the semidiscrete case. We thus have:

**Theorem 2.1.** *For any $R_0>0$ there is $\tau=\tau(R_0)$ such that* (2.9) *has a unique solution $u_h\in C([0,\tau],V)$ for any initial value $u_{h,0}\in V_h$ with $\|u_{h,0}\|_1\le R_0$. Moreover, there is $c$ such that $\|u_h\|_{L_\infty([0,\tau],V)}\le cR_0$.*

We denote by $S_h(t,\cdot)$ the corresponding (local) solution operator, so that $u_h(t)=S_h(t,u_{h,0})$ is the solution of (2.9).

We may now estimate the difference between the local solutions $u(t)=S(t,u_0)$ and $u_h(t)=S_h(t,u_{h,0})$ that we have obtained. We refer to the following result as a local a priori error estimate. It is local because the constant $C(R,\tau)$ grows exponentially with the length $\tau$ of the time interval and also because it grows with the size $R$ of the solutions. It is a priori because the error is evaluated in terms of derivatives of $u$, which are estimated a priori in Theorem 1.3.

**Theorem 2.2.** *Let $R\ge 0$ and $\tau>0$ be given. Let $u(t)$ and $u_h(t)$ be solutions of* (2.1) *and* (2.2) *respectively, such that $u(t),u_h(t)\in B_R$ for $t\in[0,\tau]$. Then*

$$
\|u_h(t)-u(t)\|_1 \le C(R,\tau)t^{-1/2}\big(\|u_{h,0}-P_hu_0\|+h\big), \qquad t\in(0,\tau], \tag{2.10}
$$

$$
\|u_h(t)-u(t)\| \le C(R,\tau)\big(\|u_{h,0}-P_hu_0\|+h^2t^{-1/2}\big), \qquad t\in(0,\tau]. \tag{2.11}
$$

*Proof.* We introduce the elliptic projection operator $R_h: V\to V_h$ defined by

$$
a(R_hv,\chi) = a(v,\chi), \quad \forall\chi\in V_h. \tag{2.12}
$$

Under the usual regularity assumptions on the triangulation and the elliptic regularity estimate (1.10), we have the error bounds (see [2])

$$
\|R_hv-v\| + h\|R_hv-v\|_1 \le Ch^s\|v\|_s, \quad v\in V\cap(H^s(\Omega))^p,\ s=1,2, \tag{2.13}
$$

Following a standard practice we divide the error into two parts:

$$
e(t)\equiv u_h(t)-u(t) = \big(u_h(t)-R_hu(t)\big) + \big(R_hu(t)-u(t)\big) \equiv \theta(t)+\rho(t).
$$

In view of (2.13) and (1.27), (1.28) we have, for $j = 0, 1$ and $s = 1, 2$,

$$\|\rho(t)\|_j \le Ch^{s-j}\|u(t)\|_s \le C(R,\tau)h^{s-j}t^{-(s-1)/2}, \qquad t \in (0,\tau], \tag{2.14}$$

$$\|\rho_t(t)\| \le Ch^s\|u_t(t)\|_s \le C(R,\tau)ht^{-1-(s-1)/2}, \qquad t \in (0,\tau]. \tag{2.15}$$

It remains to estimate $\theta(t)$, which belongs to $V_h$. In view of (2.2), the identity $A_hR_h = P_hA$ (which follows easily from (2.3) and (2.12)), and (1.6), we find that

$$\theta_t + A_h\theta = P_h\big(f(u_h) - f(u) - \rho_t\big). \tag{2.16}$$

Hence

$$\theta(t) = E_h(t)\theta(0) + \int_0^t E_h(t-\sigma)P_h\Big(f(u_h(\sigma)) - f(u(\sigma)) - D_\sigma\rho(\sigma)\Big)\,d\sigma, \tag{2.17}$$

where $D_\sigma = \partial/\partial\sigma$. Integration by parts yields

$$\begin{aligned}-\int_0^{t/2} E_h(t-\sigma)P_hD_\sigma\rho(\sigma)\,d\sigma &= E_h(t)P_h\rho(0) - E_h(t/2)P_h\rho(t/2)\\ &\quad + \int_0^{t/2}\big(D_\sigma E_h(t-\sigma)\big)P_h\rho(\sigma)\,d\sigma.\end{aligned}$$

Hence

$$\begin{aligned}\theta(t) &= E_h(t)P_he(0) - E_h(t/2)P_h\rho(t/2) + \int_0^{t/2}\big(D_\sigma E_h(t-\sigma)\big)P_h\rho(\sigma)\,d\sigma\\ &\quad - \int_{t/2}^t E_h(t-\sigma)P_hD_\sigma\rho(\sigma)\,d\sigma + \int_0^t E_h(t-\sigma)P_h\Big(f(u_h(\sigma)) - f(u(\sigma))\Big)\,d\sigma.\end{aligned}$$

Using the smoothing property (2.8) of $E_h(t)P_h$, the error estimates (2.14), (2.15) with $j = 0$, $s = 1$, and the Lipschitz condition (1.19) we obtain

$$\begin{aligned}\|\theta(t)\|_1 &\le Ct^{-1/2}\big(\|P_he(0)\| + \|\rho(t/2)\|\big) + C\int_0^{t/2}(t-\sigma)^{-3/2}\|\rho(\sigma)\|\,d\sigma\\ &\quad + C\int_{t/2}^t(t-\sigma)^{-1/2}\|D_\sigma\rho(\sigma)\|\,d\sigma\\ &\quad + C\int_0^t(t-\sigma)^{-1/2}\|f(u_h(\sigma)) - f(u(\sigma))\|\,d\sigma\\ &\le C(R,\tau)t^{-1/2}\big(\|P_he(0)\| + h\big)\\ &\quad + C(R,\tau)h\bigg(\int_0^{t/2}(t-\sigma)^{-3/2}\,d\sigma + \int_{t/2}^t(t-\sigma)^{-1/2}\sigma^{-1}\,d\sigma\bigg)\\ &\quad + C(R)\int_0^t(t-\sigma)^{-1/2}\|e(\sigma)\|_1\,d\sigma\\ &\le C(R,\tau)t^{-1/2}\big(\|P_he(0)\| + h\big) + C(R)\int_0^t(t-\sigma)^{-1/2}\|e(\sigma)\|_1\,d\sigma,\end{aligned}$$

for $t \in (0,\tau]$. Together with (2.14), (2.15) and $e = \theta + \rho$ this yields

$$\|e(t)\|_1 \le C(R,\tau)t^{-1/2}\big(\|P_he(0)\| + h\big) + C(R)\int_0^t(t-\sigma)^{-1/2}\|e(\sigma)\|_1\,d\sigma, \ t \in (0,\tau],$$

and the desired bound follows by the generalized Gronwall lemma, cf. Lemma 1.4. This proves (2.10), because $P_he(0) = u_{h,0} - P_hu_0$.

To prove (2.11) we use the Lipschitz condition (1.20) instead of (1.19). We omit the details. □

We now formulate a consequence of the previous theorem, which is the form in which we will apply it later on. (It is convenient to change the notation so that the length of the local interval of existence is now $2\tau$.)

**Theorem 2.3.** *For any $R_0 \geq 0$ there are numbers $\tau = \tau(R_0) > 0$ and $C(R_0)$ such that if $v \in B_{R_0}$, $v_h \in V_h \cap B_{R_0}$, then $S(t, v)$ and $S_h(t, v_h)$ exist for $t \in [0, 2\tau]$. Moreover, for $l = 0, 1$,*

$$\|S_h(t, v_h) - S(t, v)\|_l \leq C(R_0) t^{-1/2} \big(\|v_h - P_h v\| + h^{2-l}\big), \quad t \in (0, 2\tau],$$

*and*

$$\|S_h(t, v_h) - S(t, v)\|_l \leq C(R_0) \big(\|v_h - P_h v\| + h^{2-l}\big), \quad t \in [\tau, 2\tau].$$

*Proof.* The local existence theorems 1.2 and 2.1 give $S(t, v), S_h(t, v_h) \in B_R$, $t \in [0, 2\tau]$, for some $\tau = \tau(R_0)$ and with $R = cR_0$. We now apply Theorem 2.2. □

2.2. **A completely discrete scheme.** In this subsection we show that the above program can be carried through also for a completely discrete scheme based on the backward Euler method. We replace the time derivative in (2.2) by a backward difference quotient $\partial_t U_j = (U_j - U_{j-1})/k$, where $k$ is a time step and $U_j$ is the approximation of $u_j = u(t_j)$ and $t_j = jk$. The discrete solution $U_j \in V_h$ thus satisfies

$$\partial_t U_j + A_h U_j = P_h f(U_j), \ t_j > 0; \quad U_0 = u_{h,0}. \tag{2.18}$$

Duhamel's principle yields

$$U_j = E_{kh}^j u_{h,0} + k \sum_{l=1}^{j} E_{kh}^{j-l-1} P_h f(U_l), \quad t_j \geq 0, \tag{2.19}$$

where $E_{kh} = (I + kA_h)^{-1}$. Since $A_h$ is self-adjoint positive definite (uniformly in $h$), we have the inequality

$$\|\partial_t^l E_{kh}^j v\| = \|A_h^l E_{kh}^j v\| \leq C_l t_j^{-l} \|v\|, \quad t_j \geq t_l, \ v \in V_h, \ l \geq 0,$$

which is the discrete analogue of (1.13). In view of the inequalities (2.5) and (2.6) this leads to a smoothing property analogous to (2.8):

$$\begin{aligned} \|\partial_t^l E_{kh}^j P_h f\|_\beta \leq C t_j^{-l-(\beta-\alpha)/2} \|f\|_\alpha, \quad & t_j > 0, \ f \in \mathcal{D}(A^{\alpha/2}), \\ & -1 \leq \alpha \leq \beta \leq 1, \ l = 0, 1. \end{aligned} \tag{2.20}$$

Again the proof of Theorem 1.2 carries over verbatim to the discrete case. We thus have:

**Theorem 2.4.** *For any $R_0 > 0$ there is $\tau = \tau(R_0)$ such that (2.19) has a unique solution $U_j$, $t_j \in [0, \tau]$, for any initial value $u_{h,0} \in V_h$ with $\|u_{h,0}\|_1 \leq R_0$. Moreover, there is $c$ such that $\max_{t_j \in [0,\tau]} \|U_j\|_1 \leq cR_0$.*

We denote by $S_{hk}(t_j, \cdot)$ the corresponding (local) solution operator, so that $U_j = S_{hk}(t_j, u_{h,0})$ is the solution of (2.19).

We may now estimate the difference between the local solutions $u(t) = S(t, u_0)$ and $U_j = S_{hk}(t_j, u_{h,0})$ that we have obtained. We refer to the following result as a local a priori error estimate. The proof can be found in [7]. It follows the lines of the proof of Theorem 2.2 but is more technical.

**Theorem 2.5.** *Let $R \geq 0$ and $\tau > 0$ be given. Let $u(t)$ and $U_j$ be solutions of* (2.1) *and* (2.19) *respectively, such that $u(t), U_j \in B_R$ for $t, t_j \in [0, \tau]$. Then, for $k \leq k_0(R)$, we have*

$$\|U_j - u(t_j)\|_1 \leq C(R,\tau)\big(\|u_{h,0} - P_h u_0\| t_j^{-1/2} + h t_j^{-1/2} + k t_j^{-1}\big), \qquad t_j \in (0,\tau],$$
$$\|U_j - u(t_j)\| \leq C(R,\tau)\big(\|u_{h,0} - P_h u_0\| + h^2 t_j^{-1/2} + k t_j^{-1/2}\big), \qquad t_j \in (0,\tau].$$

There is also an analogue of Theorem 2.3.

**Theorem 2.6.** *For any $R_0 \geq 0$ there are numbers $\tau = \tau(R_0) > 0$, $k_0(R_0) > 0$ and $C(R_0)$ such that if $v \in B_{R_0}$, $v_h \in V_h \cap B_{R_0}$, then $S(t, v)$ and $S_{hk}(t_j, v_h)$ exist for $t, t_j \in [0, 2\tau]$. Moreover, for $k \leq k_0(R_0)$, we have*

$$\|S_{hk}(t_j, v_h) - S(t_j, v)\|_l \leq C(R_0)\big(\|v_h - P_h v\| + h^{2-l} + k\big), \quad t_j \in [\tau, 2\tau],\ l = 0, 1.$$

## 3. Shadowing—first approach

In this section we present a numerical shadowing result. It is taken essentially from [10] but we present it in a slightly more general form.

### 3.1. Linearization.

Let $\bar{u} \in C([0,T], V)$ with $\|\bar{u}(t)\|_1 \leq R$, $t \in [0,T]$, for some $T$ and $R$. In order to linearize around $\bar{u}$ we rewrite the differential equation in (1.6),

$$u' + Au + B(t)u = F(t, u), \tag{3.1}$$

where $B(t)$ is the linear part of $f$ and $F(t, \cdot)$ is the nonlinear remainder:

$$\begin{aligned} B(t) &= -f'(\bar{u}(t)) \in \mathcal{L}(V, H), \\ F(t, v) &= f(v) - f'(\bar{u}(t))v. \end{aligned} \tag{3.2}$$

Let $L(t, s)$, $0 \leq s \leq t \leq T$, denote the solution operator of the linearized homogeneous problem

$$v' + Av + B(t)v = 0,\ t > s; \quad v(s) = \phi. \tag{3.3}$$

Thus, $v(t) = L(t, s)\phi$ is the solution of (3.3). The operator $w \mapsto -B(t)w = f'(\bar{u}(t))w$ satisfies a local Lipschitz condition from $V$ into $H$ (uniformly in $t$); cf. Lemma 1.1. We therefore obtain local existence in $C([0,\tau], V)$ by the same argument as for (1.6). The following lemma provides a global a priori bound, which implies that any local solution can be extended to a global solution on the whole interval $[0, T]$. The lemma will also be useful in our analysis below. The bound grows exponentially with the length $T$ of the time interval; this is realistic when the linearized problem is unstable, otherwise it is an overestimate. Note, by the way, that the linearized operator $A + B(t)$ is not self-adjoint in general even though $A$ is.

**Lemma 3.1.** *Assume that $\bar{u} \in C([0,T], V)$ with $\|\bar{u}(t)\|_1 \leq R$, $t \in [0,T]$, for some $T$ and $R$. Then any solution $v(t) = L(t,s)\phi$ of* (3.3) *satisfies*

$$\|L(t,s)\phi\|_m \leq C(R,T)(t-s)^{-(m-l)/2}\|\phi\|_l, \quad 0 \leq s < t \leq T,\ 0 \leq l \leq m \leq 1.$$

*Proof.* By Duhamel's principle we have

$$v(t) = E(t-s)\phi + \int_s^t E(t-\sigma) f'(\bar{u}(\sigma)) v(\sigma)\, d\sigma,$$

so that, in view of (1.14), (1.17) and (1.18),

$$\begin{aligned}\|v(t)\|_m &\leq C(t-s)^{-(m-l)/2}\|\phi\|_l + C\int_s^t (t-\sigma)^{-1/2}\|f'(\bar{u}(\sigma))v(\sigma)\|_{m-1}\,d\sigma \\ &\leq C(t-s)^{-(m-l)/2}\|\phi\|_l + C(R)\int_s^t (t-\sigma)^{-1/2}\|v(\sigma)\|_m\,d\sigma.\end{aligned}$$

The desired bound now follows by Gronwall's lemma 1.4. □

In order to conveniently take advantage of the smoothing properties of our problems, as reflected in the error bounds of Theorems 2.3 and 2.6, we replace the continuous dynamical system $u(t) = S(t, u_0)$, $t \geq 0$, by a discrete dynamical system $u_n = u(n\tau) = S(\tau, u_{n-1})$, $n = 1, 2, \ldots$, with an appropriately chosen step $\tau$.

Let $\|\bar{u}(t)\|_1 \leq R$, $t \in [0, T]$, and let $\tau = \tau(R+1)$ be half the time of local existence as in Theorems 2.3 and 2.6, adjusted so that $N\tau = T$ for some positive integer $N$. This implies that the operator $S = S(\tau, \cdot)$ is defined on the ball $D = B_{R+1}$. We may thus think of $S$ as a nonlinear operator $S : D \subset X \to X$ with both $X = V$ and $X = H$; its domain $D = B_{R+1}$ is a ball in $V$ in both cases.

Let $T_n = n\tau$ and $L_n = L(T_{n+1}, T_n)$. Duhamel's principle applied to (3.1) gives

$$u(t) = S(t - T_n, v) = L(t, T_n)v + \int_{T_n}^t L(t,s)F(s, u(s))\,ds. \tag{3.4}$$

With $t = T_{n+1}$ we get a decomposition of $S = S(\tau, \cdot)$,

$$S = L_n + N_n,$$

where

$$N_n(v) = \int_{T_n}^{T_{n+1}} L(t,s)F(s, S(s - T_n, v))\,ds.$$

We think of $L_n$ as the linear part of the nonlinear time-$\tau$ evolution $S$ at time level $T_n$. In the following lemma we estimate the nonlinear remainder $N_n$.

**Lemma 3.2.** *Let* $\rho \leq 1$. (a) *If*

$$\|S(t - T_n, v_i) - \bar{u}(t)\|_1 \leq \rho, \quad t \in [T_n, T_{n+1}],\ i = 1, 2, \tag{3.5}$$

*then*

$$\|N_n(v_1) - N_n(v_2)\|_1 \leq C(R, \tau)\rho\|v_1 - v_2\|_l, \quad l = 0, 1. \tag{3.6}$$

(b) *If* $\bar{u}$ *is a solution of* (1.6) *and* $\|v_i - \bar{u}(T_n)\|_1 \leq \rho$, $i = 1, 2$, *then* (3.6) *holds.*

*Proof.* Let $u_i(t) = S(t - T_n, v_i)$ and $w_i(t) = \int_{T_n}^t L(t,s)F(s, u_i(s))\,ds$ for $i = 1, 2$. Then $N_n(v_i) = w_i(T_{n+1})$. (Note that $u_i(t)$ exists for $t \in [T_n, T_{n+1}]$ because $v_i \in D = B_{R+1}$.) Applying Lemma 3.1 over the (short!) time interval $[T_n, T_{n+1}]$ we get, with $e = w_1 - w_2$,

$$\begin{aligned}\|e(t)\|_1 &\leq \int_{T_n}^t \Big\|L(t,s)\Big(F(s, u_1(s)) - F(s, u_2(s))\Big)\Big\|_1\,ds \\ &\leq C(R, \tau)\int_{T_n}^t (t-s)^{-1/2}\|F(s, u_1(s)) - F(s, u_2(s))\|\,ds.\end{aligned}$$

In view of (1.23) and (3.5) we have here (note that $u_i(t) \in B_{R+1}$)

$$\|F(s,u_1) - F(s,u_2)\| \le C(R) \sum_{i=1}^{2} \|u_i - \bar u\|_1 \|u_1 - u_2\|_1 \le C(R)\rho\|u_1 - u_2\|_1,$$

so that

$$\|e(t)\|_1 \le C(R,\tau)\rho \int_{T_n}^{t} (t-s)^{-1/2} \|u_1(s) - u_2(s)\|_1 \, ds.$$

Here, by (3.4) and Lemma 3.1,

$$\begin{aligned}\|u_1(s) - u_2(s)\|_1 &= \|L(s,T_n)(v_1 - v_2) + w_1(s) - w_2(s)\|_1 \\ &\le C(R,\tau)(s-T_n)^{-(1-l)/2}\|v_1 - v_2\|_l + \|e(s)\|_1.\end{aligned}$$

Hence, since

$$\int_{T_n}^{t} (t-s)^{-1/2}(s-T_n)^{-(1-l)/2} \, ds \le C\tau^{l/2},$$

see (1.29), we get

$$\|e(t)\|_1 \le C(R,\tau)\rho\|v_1 - v_2\|_l + C(R,\tau)\rho \int_{T_n}^{t} (t-s)^{-1/2}\|e(s)\|_1 \, ds,$$

and by Gronwall's lemma 1.4

$$\|e(T_{n+1})\|_1 \le C(R,\tau)\rho\|v_1 - v_2\|_l,$$

which is (3.6) and part (a) is proved. If $\bar u$ is a solution of (1.6), then a similar Gronwall argument gives

$$\|u_i(t) - \bar u(t)\|_1 \le C(R,\tau)\|v_i - \bar u(T_n)\|_1, \quad t \in [T_n, T_{n+1}].$$

Combining this with the result of part (a) we obtain the second statement of the lemma. □

We have now decomposed the operator $S$ as $S = L_n + N_n$, where, according to Lemma 3.1, the $L_n$ are bounded linear operators on both $V$ and $H$,

$$\|L_n\|_{\mathcal{L}(X)} \le C(R,\tau), \quad n = 0, \dots, N-1, \ X = V, H,$$

and where the nonlinear remainder $N_n : D \subset X \to V$ has a Lipschitz constant that can be rendered arbitrarily small (uniformly in $n$) by restricting to a small neighborhood of $\bar u$; see Lemma 3.2.

Note that $L_n$ is the linearization of $S$ at $\bar u(T_n)$. In fact, the mapping $S : D \subset V \to V$ is Fréchet differentiable with $L_n = S'(\bar u(T_n)) \in \mathcal{L}(V)$. However, the mapping $S : D \subset H \to H$ is not differentiable, because $D_n = B_V(\bar u(T_n), \rho)$ is not a neighborhood of $\bar u(T_n)$ with respect to the topology of $H$.

3.2. **Exponential dichotomies.** We need a condition under which we can prove existence of global solutions of (1.6). Let $X$ denote either $V$ or $H$. We assume that the linear evolution operator $L(t,s)$ has an exponential dichotomy in $X$ on the interval $J = [0,T]$; see [6]. This means that there are projections $P(t) \in \mathcal{L}(X)$, $t \in J$, and constants $M \ge 1$, $\beta > 0$ such that the following conditions are satisfied for $s, t \in J$, $t \ge s$.

1. $L(t,s)P(s) = P(t)L(t,s)$.
2. The restriction $L(t,s)|_{\mathcal{R}(I-P(s))} : \mathcal{R}(I-P(s)) \to \mathcal{R}(I-P(t))$ is an isomorphism. We define $L(s,t) : \mathcal{R}(I-P(t)) \to \mathcal{R}(I-P(s))$ to be its inverse.

3. $\|L(t,s)P(s)\|_{\mathcal{L}(X)} \le Me^{-\beta(t-s)}$.
4. $\|L(s,t)(I-P(t))\|_{\mathcal{L}(X)} \le Me^{-\beta(t-s)}$.

The ranges $\mathcal{R}(P(t))$ of the projections $P(t)$ are called stable subspaces; the complementary spaces $\mathcal{R}(I-P(t)) = \mathcal{N}(P(t))$ (null spaces of $P(t)$) are called unstable subspaces. Note that, due to parabolic smoothing, $L(t,s)$ is not invertible. Condition (2) means that $L(t,s)$ can be inverted on the unstable subspaces, so that we can evolve the solution backwards there. Condition (1) implies that the stable/unstable subspaces are invariant. Conditions (3) and (4) mean that solutions decay exponentially: forwards in time on the stable subspaces, backwards in time on the unstable subspaces. The projections are not orthogonal in general.

*Example* 3.1. (Hyperbolic fixed point.) Let $\bar{u} \in V$ be a stationary solution of (1.6); $A\bar{u} = f(\bar{u})$. Assume, for simplicity, that $B = -f'(\bar{u})$ is self-adjoint so that the linearized operator $A+B$ is self-adjoint, too, and has real spectrum. Let $\mu_j$, $\phi_j$ be the eigenvalues and an orthonormal basis of eigenvectors of $A+B$. The evolution operator $L(t,s) = \exp(-(t-s)(A+B))$ is then given by

$$e^{-t(A+B)}v = \sum_{j=1}^{\infty} e^{-t\mu_j}(v,\phi_j)\phi_j.$$

The fixed point $\bar{u}$ is said to be hyperbolic if the spectrum does not intersect the imaginary axis, i.e., in this case, if all eigenvalues $\mu_j$ are non-zero. We then have an unbounded sequence

$$\mu_1 \le \cdots \le \mu_q < 0 < \mu_{q+1} \le \cdots < \infty.$$

It is now easy to show that we have a constant exponential dichotomy with constant stable and unstable projections

$$Pv = \sum_{j=q+1}^{\infty} (v,\phi_j)\phi_j, \quad I - P = \sum_{j=1}^{q} (v,\phi_j)\phi_j.$$

When we linearize around a time-dependent solution, $\bar{u}(t)$, then we cannot use spectral theory in this way. This is why we need the more general concept of exponential dichotomy.

If $L(t,s)$ has an exponential dichotomy on the interval $J$, then the inhomogeneous boundary value problem, with $u_0 \in \mathcal{R}(P_0)$, $u_T \in \mathcal{R}(I-P_N)$,

$$\begin{aligned} &u' + Au + B(t)u = g(t), \quad t \in J, \\ &P(0)u(0) = u_0, \ (I-P(T))u(T) = u_T, \end{aligned} \tag{3.7}$$

is well-posed. We leave this as an exercise; see Exercise 3.1 below. Instead we switch the discussion to the discrete case, where we shall prove an analogous result.

Recall the sequence of operators $L_n = L(T_{n+1}, T_n)$, $n = 0, \dots, N-1$. Define

$$L_{mn} = L_{m-1} \dots L_n, \ m > n; \quad L_{nn} = I$$

(note $L_n = L_{n+1,n}$). If $L(t,s)$ has an exponential dichotomy on $J$, then $\{L_n\}_{n=0}^{N-1}$ has a discrete exponential dichotomy; see [6]. This means that there are projections $P_n$, $n = 0, \dots, N$, and constants $M \ge 1$ and $\theta \in (0,1)$ such that the following conditions are satisfied for $n = 0, \dots, N-1$.

1. $L_n P_n = P_{n+1} L_n$.

2. The restriction $L_n|_{\mathcal{R}(I-P_n)} : \mathcal{R}(I-P_n) \to \mathcal{R}(I-P_{n+1})$ is an isomorphism. We define $L_{n,n+1} : \mathcal{R}(I-P_{n+1}) \to \mathcal{R}(I-P_n)$ to be its inverse.
3. $\|L_{mn}P_n\|_{\mathcal{L}(X)} \le M\theta^{m-n}$, $m \ge n$.
4. $\|L_{nm}(I-P_m)\|_{\mathcal{L}(X)} \le M\theta^{m-n}$, $m \ge n$.

Note that (1) implies $L_{mn}P_n = P_m L_{mn}$, $m \ge n$.

Let $X_0^S = \mathcal{R}(P_0)$ and $X_N^U = \mathcal{R}(I-P_N)$. Consider the following linear boundary value problem with data $\{y_n\}_{n=0}^{N+1}$, where $y_0 \in X_0^S$, $y_{N+1} \in X_N^U$, and $y_n \in X$ for $n = 1, \dots, N$.

$$
\begin{aligned}
&x_{n+1} = L_n x_n + y_{n+1}, \quad n = 0, \dots, N-1, \\
&P_0 x_0 = y_0, \; (I-P_N)x_N = y_{N+1}.
\end{aligned}
\tag{3.8}
$$

A simple calculation gives (Duhamel's principle)

$$x_n = L_{n0}x_0 + \sum_{j=1}^{n} L_{nj}y_j.$$

Hence, by condition (1),

$$P_n x_n = L_{n0}P_0 x_0 + \sum_{j=1}^{n} L_{nj}P_j y_j.$$

Similarly, evolving from $T_n$ to $T_N$,

$$(I-P_N)x_N = L_{Nn}(I-P_n)x_n + \sum_{j=n+1}^{N} L_{Nj}(I-P_j)y_j,$$

where, by condition (2), $L_{Nn}|_{\mathcal{R}(I-P_n)}$ is invertible, so that

$$(I-P_n)x_n = L_{nN}(I-P_N)x_N - \sum_{j=n+1}^{N} L_{nj}(I-P_j)y_j.$$

Hence, using also the boundary conditions, we get

$$
\begin{aligned}
x_n &= P_n x_n + (I-P_n)x_n \\
&= L_{n0}P_0 y_0 + \sum_{j=1}^{n} L_{nj}P_j y_j \\
&\quad + L_{nN}(I-P_N)y_{N+1} - \sum_{j=n+1}^{N} L_{nj}(I-P_j)y_j.
\end{aligned}
\tag{3.9}
$$

Using (3) and (4) we get

$$
\begin{aligned}
\|x_n\|_X &\le M\Big(\theta^n + \sum_{j=1}^{n}\theta^{n-j} + \theta^{N-n} + \sum_{j=n+1}^{N}\theta^{j-n}\Big)\max_j \|y_j\|_X \\
&\le M\Big(1 + \sum_{j=n+1}^{N}\theta^{j-n} + \theta^n + \sum_{j=1}^{n}\theta^{n-j}\Big)\max_j \|y_j\|_X \\
&\le M\Big(\sum_{j=n}^{\infty}\theta^{j-n} + \sum_{j=-\infty}^{n}\theta^{n-j}\Big)\max_j \|y_j\|_X = \frac{2M}{1-\theta}\max_j \|y_j\|_X.
\end{aligned}
$$

It follows that (3.8) is uniquely solvable and the solution satisfies

$$\max_{0\le n\le N}\|x_n\|_X \le \frac{2M}{1-\theta}\max_{0\le n\le N+1}\|y_n\|_X. \tag{3.10}$$

We formulate this as a lemma.

**Lemma 3.3.** *Assume that $\{L_n\}_{n=0}^{N-1}$ has a discrete exponential dichotomy in the Banach space $X$ with projections $P_n$ and constants $M\ge 1$ and $\theta\in(0,1)$. Set $\mathbf{X}=X^{N+1}$, $\mathbf{Y}=X_0^S\times X^N\times X_N^U$, and define a linear operator $\mathbf{A}:\mathbf{X}\to\mathbf{Y}$ by $\mathbf{A}:(x_0,\dots,x_N)\mapsto(y_0,\dots,y_{N+1})$, where*

$$\begin{aligned} y_{n+1}&=x_{n+1}-L_nx_n,\quad n=0,\dots,N-1,\\ y_0&=P_0x_0,\ y_{N+1}=(I-P_N)x_N. \end{aligned}$$

*Then $\mathbf{A}$ is invertible and*

$$\|\mathbf{A}^{-1}\|_{\mathcal{L}(\mathbf{Y},\mathbf{X})}\le\frac{2M}{1-\theta}.$$

For future reference we note that the solution formula (3.9) may be written

$$x_n=G_{n0}y_0-G_{nN}y_{N+1}+\sum_{j=1}^{N}G_{nj}y_j, \tag{3.11}$$

by means of the Green's operator

$$G_{nj}=\begin{cases} L_{nj}P_j, & 0\le j\le n,\\ -L_{nj}(I-P_j), & n<j\le N. \end{cases} \tag{3.12}$$

*Exercise* 3.1. Show that the solution of (3.7) may be written

$$u(t)=G(t,0)u_0-G(t,T)u_T+\int_0^T G(t,s)g(s)\,ds, \tag{3.13}$$

where

$$G(t,s)=\begin{cases} L(t,s)P(s), & 0\le s\le t,\\ -L(t,s)(I-P(s)), & t<s\le T. \end{cases} \tag{3.14}$$

Use this to derive an estimate of $\|u(t)\|_X$ similar to (3.10).

3.3. **Shadowing.** Let $\bar u\in C([0,T],V)$ be a solution of (1.6) with $\|\bar u(t)\|_1\le R$, $t\in[0,T]$, for some $T$ and $R$. Assume that we have a numerical solution $u_h(t)=S_h(t,u_{h,0})$, $t\in[0,T]$, which belongs to a neighborhood of $\bar u$ in the sense that, for some $\rho_0\le 1$,

$$\|u_h(t)-\bar u(t)\|_1\le\rho_0,\quad t\in[0,T]. \tag{3.15}$$

With $T_n=n\tau$ chosen as in §3.1 we form the sequence $u_{h,n}=S_h(n\tau,u_{h,0})$. This means that $u_{h,n}$ is an orbit of the nonlinear mapping $S_h=S_h(\tau,\cdot)$, i.e., $u_{h,n+1}=S_h(u_{h,n})$. Comparing with the mapping $S=S(\tau,\cdot)$ we get, in view of Theorem 2.3,

$$\begin{aligned} \|u_{h,n+1}-S(u_{h,n})\|_l&=\|S_h(\tau,u_{h,n})-S(\tau,u_{h,n})\|_l\\ &\le C(R+1)h^{2-l}=:\delta,\quad n=0,\dots,N-1,\ l=0,1. \end{aligned} \tag{3.16}$$

(A similar result holds for a completely discrete sequence via Theorem 2.6.) With the terminology used in shadowing theory this means that the sequence $\{u_{h,n}\}_{n=0}^N$

is a $\delta$-pseudo-orbit of $S$ with $\delta = C(R+1)h^{2-l}$. Our aim is to find an orbit $\{u_n\}_{n=0}^N$ of $S$, i.e., $u_{n+1} = S(u_n)$, such that, for some $\sigma$,

$$\max_{0\le n\le N} \|u_n - u_{h,n}\|_l \le \sigma\delta =: \epsilon, \quad l = 0, 1.$$

This means that $\{u_{h,n}\}_{n=0}^N$ is $\epsilon$-shadowed by the true orbit $\{u_n\}_{n=0}^N$.

In order to achieve this we shall prove the following theorem which is set in a more general context.

**Theorem 3.4.** (Shadowing theorem.) *Consider a nonlinear operator $S : D \subset X \to X$, where $X$ is a Banach space and $D$ a nonempty subset of $X$. Assume that $S$ can be decomposed in the form*

$$S = L_n + N_n \tag{3.17}$$

*where $\{L_n\}_{n=0}^{N-1}$ has a discrete exponential dichotomy on $X$ with projections $P_n$ and constants $M \ge 1$, $\theta \in (0,1)$, and where the operators $N_n : D_n \subset X \to X$ are Lipschitz continuous with Lipschitz constants that satisfy*

$$\operatorname{Lip}(N_n) \le \frac{1-\theta}{4M}, \quad n = 0, \dots, N-1. \tag{3.18}$$

*Set*

$$\sigma = \frac{4M}{1-\theta}. \tag{3.19}$$

(a) *Let $\{\tilde{x}_n\}_{n=0}^N$ be a sequence with $\tilde{x}_n \in D_n$, $n = 0, \dots, N-1$. If $\{x_n\}_{n=0}^N$ satisfies $x_n \in D_n$, $n = 0, \dots, N-1$, and*

$$\begin{aligned} &x_{n+1} = S(x_n), \quad n = 0, \dots, N-1, \\ &P_0 x_0 = P_0 \tilde{x}_0,\ (I - P_N)x_N = (I - P_N)\tilde{x}_N, \end{aligned} \tag{3.20}$$

*then*

$$\max_{0\le n\le N} \|x_n - \tilde{x}_n\|_X \le \sigma \max_{0\le n\le N-1} \|\tilde{x}_{n+1} - S(\tilde{x}_n)\|_X. \tag{3.21}$$

(b) *Assume that $D_n$ contains a closed ball $B_X(z_n, \rho)$ for $n = 0, \dots, N-1$, where $\{z_n\}_{n=0}^N$ is a sequence that satisfies*

$$\|z_{n+1} - S(z_n)\|_X \le \rho/\sigma, \quad n = 0, \dots, N-1. \tag{3.22}$$

*Set $\rho_0 = \rho/(\sigma M)$. Then, for any data $\tilde{x}_0 \in B_X(z_0, \rho_0)$, $\tilde{x}_N \in B_X(z_N, \rho_0)$, the boundary value problem* (3.20) *has unique solution $\{x_n\}_{n=0}^N$ with $x_n \in B_X(z_n, \rho)$.*

The proof is based on the following simple lemmas.

**Lemma 3.5.** *Let $\mathbf{X}$ and $\mathbf{Y}$ be Banach spaces and let $\mathbf{A} : \mathbf{X} \to \mathbf{Y}$ be a linear bijection with bounded inverse $\mathbf{A}^{-1}$. Assume that the mapping $\mathbf{N} : \mathbf{D} \subset \mathbf{X} \to \mathbf{Y}$, defined in a nonempty subset $\mathbf{D}$ of $\mathbf{X}$, is Lipschitz continuous with*

$$\alpha := \|\mathbf{A}^{-1}\|_{\mathcal{L}(\mathbf{Y},\mathbf{X})} \operatorname{Lip}(\mathbf{N}) < 1.$$

*Define $\mathbf{F} = \mathbf{A} + \mathbf{N}$ and set $\sigma := \|\mathbf{A}^{-1}\|_{\mathcal{L}(\mathbf{Y},\mathbf{X})}/(1-\alpha)$. Then*

$$\|\mathbf{x}_1 - \mathbf{x}_2\|_{\mathbf{X}} \le \sigma \|\mathbf{F}(\mathbf{x}_1) - \mathbf{F}(\mathbf{x}_2)\|_{\mathbf{Y}}, \quad \textit{for all } \mathbf{x}_1, \mathbf{x}_2 \in \mathbf{D}. \tag{3.23}$$

*Proof.* The bound (3.23) follows readily from the identity

$$\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{A}^{-1}\big(\mathbf{F}(\mathbf{x}_1) - \mathbf{F}(\mathbf{x}_2)\big) - \mathbf{A}^{-1}\big(\mathbf{N}(\mathbf{x}_1) - \mathbf{N}(\mathbf{x}_2)\big).$$

□

**Lemma 3.6.** *In addition to the hypotheses of Lemma 3.5, assume that the domain* $\mathbf{D}$ *of* $\mathbf{N}$ *contains a closed ball* $B_{\mathbf{X}}(\mathbf{z},\rho)$. *Then, for each* $\mathbf{y}\in\mathbf{Y}$ *with*

$$\|\mathbf{y}-\mathbf{F}(\mathbf{z})\|_{\mathbf{Y}} \le \rho/\sigma, \tag{3.24}$$

*the equation* $\mathbf{F}(\mathbf{x})=\mathbf{y}$ *has a unique solution* $\mathbf{x}\in B_{\mathbf{X}}(\mathbf{z},\rho)$.

*Proof.* This is a simple consequence of the contraction mapping theorem. In fact, if we define $\mathbf{T}(\mathbf{x})=\mathbf{A}^{-1}(\mathbf{y}-\mathbf{N}(\mathbf{x}))$, then $\mathrm{Lip}(\mathbf{T})=\alpha<1$. Moreover, the identity

$$\mathbf{T}(\mathbf{x})-\mathbf{z}=\mathbf{A}^{-1}\big(\mathbf{y}-\mathbf{F}(\mathbf{z})\big)-\mathbf{A}^{-1}\big(\mathbf{N}(\mathbf{x})-\mathbf{N}(\mathbf{z})\big)$$

and (3.24) imply that $\mathbf{T}$ maps $B_{\mathbf{X}}(\mathbf{z},\rho)$ into itself. □

*Proof of Theorem* 3.4. Set $X_0^S=\mathcal{R}(P_0)$, $X_N^U=\mathcal{R}(I-P_N)$, $\mathbf{X}=X^{N+1}$, $\mathbf{Y}=X_0^S\times X^N\times X_N^U$, $\mathbf{D}=D_0\times\cdots\times D_{N-1}\times X$, and define $\mathbf{y}\in\mathbf{Y}$, $\mathbf{A}:\mathbf{X}\to\mathbf{Y}$, $\mathbf{N}:\mathbf{D}\subset\mathbf{X}\to\mathbf{Y}$ by

$$\mathbf{y}=\begin{bmatrix} y_0\\ y_{n+1}\\ y_{N+1}\end{bmatrix} := \begin{bmatrix} P_0\tilde{x}_0\\ 0\\ (I-P_N)\tilde{x}_N\end{bmatrix},$$

$$\mathbf{A}\mathbf{x} := \begin{bmatrix} P_0x_0\\ x_{n+1}-L_nx_n\\ (I-P_N)x_N\end{bmatrix},\quad \mathbf{N}(\mathbf{x}) := \begin{bmatrix} 0\\ -N_n(x_n)\\ 0\end{bmatrix},$$

where $n=0,\dots,N-1$, and set $\mathbf{F}=\mathbf{A}+\mathbf{N}$. The boundary value problem (3.20) then becomes

$$\mathbf{y}=\mathbf{F}(\mathbf{x}).$$

The assumption (3.18) implies that $\mathrm{Lip}(\mathbf{N})\le(1-\theta)/(4M)$. We may therefore apply Lemma 3.5 with $\alpha=1/2$; this yields a value of $\sigma=\|\mathbf{A}^{-1}\|_{\mathcal{L}(\mathbf{Y},\mathbf{X})}/(1-\alpha)=4M/(1-\theta)$ that coincides with (3.19). The error bound (3.21) follows from (3.23) since

$$\mathbf{F}(\tilde{\mathbf{x}})-\mathbf{F}(\mathbf{x})=\begin{bmatrix} 0\\ \tilde{x}_{n+1}-S(\tilde{x}_n)\\ 0\end{bmatrix}.$$

This proves part (a). To prove part (b) we apply Lemma 3.6. Set $\mathbf{z}=(z_0,\dots,z_N)$. Our assumption implies that $B_{\mathbf{X}}(\mathbf{z},\rho)\subset\mathbf{D}$. Moreover, with $\mathbf{y}$ as above,

$$\mathbf{y}-\mathbf{F}(\mathbf{z})=\begin{bmatrix} P_0(\tilde{x}_0-z_0)\\ -(z_{n+1}-S(z_n))\\ (I-P_N)(\tilde{x}_N-z_N)\end{bmatrix}.$$

Here, by conditions (3) and (4) in the definition of discrete exponential dichotomy, we have $\|P_0\|_{\mathcal{L}(X)}\le M$, $\|I-P_N\|_{\mathcal{L}(X)}\le M$, so that (3.24) follows by our assumptions. □

We now return to the problem of shadowing a numerical solution $u_h(t)$ in a neighborhood of $\bar{u}$.

**Theorem 3.7.** (Numerical shadowing theorem 1.) *Let* $\bar{u}\in C([0,T],V)$ *be a solution of* (1.6) *with*

$$\max_{t\in[0,T]}\|\bar{u}(t)\|_1\le R, \tag{3.25}$$

*for some* $T$ *and* $R$. *Assume that the solution operator* $L(t,s)$ *of the linearized problem* (3.3) *has an exponential dichotomy in* $V$ *and* $H$ *on the interval* $[0,T]$.

*Then there are numbers $\rho_0$ and $C$ such that, for each solution $u_h(t) = S_h(t, u_{h,0})$, $t \in [0,T]$, of* (2.2) *with*

$$\max_{t\in[0,T]} \|u_h(t) - \bar{u}(t)\|_1 \le \rho_0, \tag{3.26}$$

*there is a solution $u$ of* (1.6) *such that*

$$\|u_h(t) - u(t)\|_l \le C\big(1 + t^{-1/2}\big)h^{2-l}, \quad t \in (0,T],\ l = 0,1. \tag{3.27}$$

*Proof.* Let $u_{h,n} = u_h(T_n)$, $T_n = n\tau$, and $S = L_n + N_n$ be as above. We apply Theorem 3.4 with $\tilde{x}_n = u_{h,n}$. Set $D = B_{R+1} = B_V(0, R+1)$, $z_n = \bar{u}(T_n)$, and

$$D_n = \{v \in V : \|v - \bar{u}(T_n)\|_1 \le \rho\} = B_V(z_n, \rho).$$

Lemma 3.2 (b) shows that $N_n : D_n \to X$ has Lipschitz constant

$$\mathrm{Lip}(N_n) \le C(R)\rho$$

for both $X = V$ and $X = H$. We fix $\rho$ so that (3.18) holds and set $\rho_0 = \rho/(\sigma M)$.

In order to prove existence we use $X = V$. Then $D_n$ is a closed ball in $X$. Moreover, $z_{n+1} - S(z_n) = 0$ so that (3.22) holds, and (3.26) implies that $u_{h,n} = \tilde{x}_n \in B_V(z_n, \rho_0) \subset D_n$, $n = 0, \dots, N$. Theorem 3.4 (b) now gives the existence of an orbit $u_n$, $n = 0, \dots, N$, of $S$, satisfying (3.20), and hence, by part (a), (3.21) and (3.16) give the special case $l = 1$ of the inequality

$$\max_{0\le n\le N} \|u_{h,n} - u_n\|_l \le \sigma \max_{0\le n\le N-1} \|u_{h,n+1} - S(u_{h,n})\|_l \le Ch^{2-l}, \quad l = 0,1. \tag{3.28}$$

Another application of Theorem 3.4 (a), now with $X = H$, proves the case $l = 0$ of (3.28). Note that $D_n$ does not contain any ball defined using the norm of $X = H$ so part (b) does not apply with $X = H$.

Using the sequence $u_n$ we define $u(t) = S(t - T_n, u_n)$ for $t \in [T_n, T_{n+1}]$. By existence and uniqueness of solutions (Theorem 2.3) this is a solution of (1.6). Error bounds at intermediate times are obtained by combining (3.28) with Theorem 2.3 as follows. For $t \in [0, T_1]$ we have

$$\|u_h(t) - u(t)\|_1 \le C(R)t^{-1/2}\Big(\|u_{h,0} - P_h u_0\|_1 + h\Big) \le Ct^{-1/2}h.$$

For $t \in [T_{n+1}, T_{n+2}]$, $n \ge 0$, we have

$$\|u_h(t) - u(t)\|_1 \le C(R)\Big(\|u_{h,n} - u_n\|_1 + h\Big) \le Ch.$$

This proves the special case $l = 1$ of (3.27). The case $l = 0$ is obtained similarly. □

*Exercise* 3.2. State and prove a similar theorem without the assumption that $\bar{u}$ is a solution. (Hint: (3.22) means that $z_n$ needs only be a pseudo-orbit of $S$.)

*Exercise* 3.3. State and prove a similar theorem for the backward Euler discretization in §2.2.

Note that the shadowing property is a property of the evolution operator $S$; properties of $S_h$ (or $S_{hk}$) enter only through (3.16).

Shadowing in a neighborhood of a hyperbolic fixed point (Example 3.1) was proved in [9] and [10] by the above technique. There we also proved the converse: every exact solution near the fixed point is shadowed by a numerical solution. The proof involves showing that the exponential dichotomy of the evolution operator $L$ carries over to its numerical counterpart and then letting $S$ and $S_h$ change roles.

This work was inspired by the paper [1], where similar results were proved in the context of ordinary differential equations.

In [8] we obtained shadowing in a neighborhood of the whole attractor of $S$. Here we assume that the attractor has the Morse-Smale property, which means, among other things, that $S$ has a finite number of fixed points, all of them are hyperbolic, and the attractor consists of the union of their unstable manifolds. After reduction to finite dimension by means of an inertial manifold, we apply a more complicated, global shadowing theorem of Pilyugin. Again properties of $S_h$ (or $S_{hk}$) enter mainly through (3.16).

## 4. A posteriori error estimates

In this section we express the error in terms of an appropriately defined residual of the computed solution and the solution of a linearized adjoint problem. The residual is estimated in terms of a posteriori computable residuals. The presentation here is inspired by [4].

4.1. **The error equation.** We will derive an equation for the error by means of a duality argument. We therefore begin by recalling the weak form of (1.1): find $u(t) \in V$ such that

$$\begin{aligned} &(u', v) + a(u, v) - (f(u), v) = 0, \quad \forall v \in V,\ t > 0, \\ &u(0) = u_0, \end{aligned} \tag{4.1}$$

where $a(u, v) = (\epsilon \nabla u, \nabla v) = (-\epsilon \Delta u, v) = (Au, v)$ is the bilinear form associated with $A$. To allow for time-discretization by means of discontinuous piecewise polynomials in $t$, we introduce a mesh on the time interval $[0, T]$,

$$0 = t_0 < \cdots < t_{j-1} < t_j < \cdots < t_N = T,$$

with time intervals $I_j = (t_{j-1}, t_j)$ of lengths $k_j = t_j - t_{j-1}$, and we define function spaces

$$\begin{aligned} &\mathcal{V} = C([0, T], V), \\ &\tilde{\mathcal{V}} = \{v : v \in C(\bar{I}_j, V) \cap C^1(\bar{I}_j, V^*),\ j = 1, \ldots, N\}. \end{aligned} \tag{4.2}$$

Functions in $\tilde{\mathcal{V}}$ are piecewise smooth with respect to the mesh, but may be discontinuous at the nodes $t_j$.

Let $U \in \tilde{\mathcal{V}}$ be arbitrary. (In the end $U$ will be a numerical approximation of $u$, but this is not essential right now.) We define the local residual of $U$,

$$\begin{aligned} r_{t_{j-1},t}(U, v) = &\int_{t_{j-1}}^{t} \Big( (U', v) + a(U, v) - (f(U), v) \Big)\, d\sigma \\ &+ ([U]_{t_{j-1}}, v(t_{j-1}^+)), \quad v \in \tilde{\mathcal{V}},\ t \in (t_{j-1}, t_j], \end{aligned} \tag{4.3}$$

where $[v]_{t_j} = v(t_j^+) - v(t_j^-)$ denotes the jump at $t_j$.

Let $s, t \in [0, T]$ with $s < t$. We define a punctured interval and two sets of nodes: if $s \in [t_m, t_{m+1})$, $t \in (t_{n-1}, t_n]$, then

$$\begin{aligned} &J_{s,t} = (s, t_{m+1}) \cup I_{m+2} \cup \ldots I_{n-1} \cup (t_{n-1}, t), \\ &J_{s,t}^- = \{t_j : s \le t_j < t\}, \quad J_{s,t}^+ = \{t_j : s < t_j \le t\}. \end{aligned}$$

We can now define a global residual

$$
\begin{aligned}
R_{s,t}(U,v) = \int_{J_{s,t}} \Big( (U',v) + a(U,v) - (f(U),v) \Big)\, d\sigma \\
+ \sum_{J^-_{s,t}} ([U]_{t_j}, v(t_j^+)), \quad v \in \tilde{\mathcal{V}}.
\end{aligned} \tag{4.4}
$$

Any solution $u$ of (4.1) satisfies

$$
R_{s,t}(u,v) = 0, \quad \forall v \in \tilde{\mathcal{V}}. \tag{4.5}
$$

Hence

$$
\begin{aligned}
R_{s,t}(U,v) &= R_{s,t}(U,v) - R_{s,t}(u,v) \\
&= \int_{J_{s,t}} \Big( (U'-u',v) + a(U-u,v) - (f(U)-f(u),v) \Big)\, d\sigma \\
&\quad + \sum_{J^-_{s,t}} ([U-u]_{t_j}, v(t_j^+)), \quad \forall v \in \tilde{\mathcal{V}}.
\end{aligned}
$$

We linearize around $U$. Thus, let

$$
\begin{aligned}
e &:= U - u, \\
B(t) &:= -f'(U(t)) \in \mathcal{L}(V,H), \\
\eta = \eta(U,u) &:= f(U) - f(u) - f'(U)(U-u);
\end{aligned} \tag{4.6}
$$

cf. (3.2) with $\bar{u}$ replaced by $U$. Then

$$
\begin{aligned}
R_{s,t}(U,v) = \int_{J_{s,t}} \Big( (e',v) + a(e,v) + (Be,v) - (\eta,v) \Big)\, d\sigma \\
+ \sum_{J^-_{s,t}} ([e]_{t_j}, v(t_j^+)), \quad \forall v \in \tilde{\mathcal{V}}.
\end{aligned} \tag{4.7}
$$

We define the bilinear form

$$
\begin{aligned}
B_{s,t}(w,v) &= \int_{J_{s,t}} \Big( (w',v) + a(w,v) + (Bw,v) \Big)\, d\sigma \\
&\quad + \sum_{J^-_{s,t}} ([w]_{t_j}, v(t_j^+)) + (w(s^-), v(s^+)) \\
&= \int_{J_{s,t}} \Big( (w,-v') + a(w,v) + (w,B^*v) \Big)\, d\sigma \\
&\quad + \sum_{J^+_{s,t}} (w(t_j^-), -[v]_{t_j}) + (w(t^-), v(t^+)), \quad w,v \in \tilde{\mathcal{V}},
\end{aligned} \tag{4.8}
$$

where the second variant is obtained by integration by parts. When restricted to smooth test functions $v \in C_0^\infty$ the relevant terms in (4.8) coincide with the definition of the derivative $w'$ in the sense of distributions. The point is that we extend this definition to piecewise smooth test functions $v \in \tilde{\mathcal{V}}$.

Now (4.7) becomes

$$
e \in \tilde{\mathcal{V}}; \quad B_{s,t}(e,v) = (e(s^-), v(s^+)) + R_{s,t}(U,v) + \int_s^t (\eta, v)\, d\sigma, \quad \forall v \in \tilde{\mathcal{V}}. \tag{4.9}
$$

We argue by duality. We therefore introduce the linearized adjoint problem with data $\psi \in H$:

$$-z'(s) + Az(s) + B^*(s)z(s) = 0, \ s < t; \quad z(t) = \psi. \tag{4.10}$$

This is the adjoint of the linearized problem (3.3) (but linearized around $U(t)$ instead of $\bar{u}(t)$). Using the second form of (4.8) we write (4.10) as:

$$z \in \mathcal{V}; \quad B_{s,t}(w, z) = (w(t^-), \psi), \quad \forall w \in \tilde{\mathcal{V}}. \tag{4.11}$$

Let $K(s,t)$ be the solution operator of (4.10), so that its solution is $z(s) = K(s,t)\psi$. Global existence on $[0,T]$ is obtained in the same way as for the evolution operator $L(t,s)$ in §3.1. In fact, the solution $u(t) = L(t,s)\phi$ of (3.3) with data $\phi \in H$ satisfies

$$u \in \mathcal{V}; \quad B_{s,t}(u, v) = (\phi, v(s^+)), \quad \forall v \in \tilde{\mathcal{V}}. \tag{4.12}$$

With $v = z$ in (4.12) and $w = u$ in (4.11) we get

$$(L(t,s)\phi, \psi) = B(u, z) = (\phi, K(s,t)\psi).$$

We conclude that $K(s,t) = L(t,s)^*$.

We now return to the equation (4.9) for $e$. With $v = z$ in (4.9) and $w = e$ in (4.11) we get

$$(e(t^-), \psi) = (e(s^-), z(s^+)) + R_{s,t}(U, z) + \int_s^t (\eta, z)\, d\sigma, \quad \forall \psi \in H, \tag{4.13}$$

or, with $z(s) = K(s,t)\psi$,

$$\begin{aligned}(e(t^-), \psi) &= (e(s^-), K(s,t)\psi) \\ &\quad + \int_{J_{s,t}} \Big( (U' - f(U), K(\sigma,t)\psi) + a(U, K(\sigma,t)\psi) \Big)\, d\sigma \\ &\quad + \sum_{J_{s,t}^-} ([U]_{t_j}, K(t_j,t)\psi) + \int_s^t (\eta, K(\sigma,t)\psi)\, d\sigma, \quad \forall \psi \in H.\end{aligned} \tag{4.14}$$

Formally, since $K(s,t) = L(t,s)^*$, we may write

$$\begin{aligned} e(t^-) &= L(t,s)e(s^-) + \int_s^t L(t,\sigma)\big(U' + AU - f(U)\big)\, d\sigma \\ &\quad + \int_s^t L(t,\sigma)\eta\, d\sigma. \end{aligned} \tag{4.15}$$

This is the variation of constants formula (Duhamel's principle) for the equation

$$e' + Ae + B(t)e = U' + AU - f(U) + \eta.$$

The first term in the right side of (4.15) describes the evolution of the initial error $e(s^-)$; the second term is the contribution to $e(t^-)$ from the residual $U' + AU - f(U)$. The third term is the remainder from linearization. However, the derivatives $U'(\sigma)$ and $AU(\sigma) = -\epsilon\Delta U(\sigma)$, where $U(\sigma)$ is only in $V$, have to be understood in the sense of distributions. This is why we argue by duality and use the weak formulation (4.13).

We finish this subsection by stating an estimate of the remainder $\eta$ defined in (4.6). It is an immediate consequence of (1.23).

**Lemma 4.1.** *Assume that* $\|U(t)\|_1 \leq R$, $\|u(t)\|_1 \leq R$, *for* $t \in [0,T]$. *Then, for* $s,t \in [0,T]$, $l = 0,1$,

$$\Big| \int_s^t (\eta, v)\, d\sigma \Big| \leq C(R) \|e\|_{L_\infty([s,t],H^1)} \|e\|_{L_\infty([s,t],H^l)} \|v\|_{L_1([s,t],H^{1-l})}.$$

4.2. **Local estimates of the residual.** We now assume that $U$ is defined by the discontinuous Galerkin finite element method; see [12]. We define a space of functions that are discontinuous, piecewise constant in time, and piecewise linear in space:

$$\tilde{\mathcal{V}}_h = \{ v \in \tilde{\mathcal{V}} : v|_{I_j}(t) = v(t_j^-) \in V_h,\ j = 1, \dots, N \}.$$

Then $U \in \tilde{\mathcal{V}}_h$ is defined by

$$\int_{I_j} \Big( (U', v) + a(U,v) - (f(U), v) \Big)\, d\sigma + ([U]_{t_{j-1}}, v(t_{j-1}^+)) = 0, \quad \forall v \in \tilde{\mathcal{V}}_h, \tag{4.16}$$

for $j = 1, \dots, N$. An initial value $U(t_0^-) = u_{h,0}$ must be provided. Here, of course, $U' = 0$, since $U$ is constant on $I_j$. Setting $U_j = U(t_j^-)$, $v(t) = \chi \in V_h$ on $I_j$, we see that (4.16) becomes

$$k_j a(U_j, \chi) - k_j (f(U_j), \chi) + (U_j - U_{j-1}, \chi) = 0, \quad \forall \chi \in V_h, \tag{4.17}$$

or

$$A_h U_j - P_h f(U_j) + k_j^{-1}(U_j - U_{j-1}) = 0,$$

which is the backward Euler method (2.18).

The equation (4.16) means that

$$r_{t_{j-1},t_j}(U,v) = 0, \quad \forall v \in \tilde{\mathcal{V}}_h. \tag{4.18}$$

As a consequence of this “orthogonality” we have the following a posteriori estimate of the local residual.

**Lemma 4.2.** *Let* $U \in \tilde{\mathcal{V}}_h$ *be a solution of* (4.16) *and assume that* $v \in \mathcal{V}$ *is sufficiently smooth. Then, for* $l = 0,1$, $m = 1,2$, $\tilde{I}_j = (t_{j-1}, t) \subset I_j$,

$$\begin{aligned} |r_{t_{j-1},t}(U,v)| &\leq C \|k R_t\|_{L_\infty(\tilde{I}_j,H^l)} \min\Big( \|v'\|_{L_1(\tilde{I}_j,H^{-l})}, \|v\|_{L_\infty(\tilde{I}_j,H^{-l})} \Big) \\ &\quad + C \|h^m R_x\|_{L_\infty(\tilde{I}_j,H)} \Big\| \int_{\tilde{I}_j} v\, d\sigma \Big\|_m, \end{aligned}$$

*where the functions* $h = h(x)$, $k = k(t)$, $R_t = R_t(x,t)$, $R_x = R_x(x,t)$, *are defined piecewise by*

$$\begin{aligned} h|_K &= h_K = \operatorname{diam}(K), \quad k|_{I_j} = k_j, \\ R_t|_{K \times I_j} &= k_j^{-1} [U]_{t_{j-1}} = k_j^{-1}\big(U_j - U_{j-1}\big), \\ R_x|_{K \times I_j} &= |k_j^{-1}[U]_{t_{j-1}} - f(U_j)| + |K|^{-1/2} h_K^{-1/2} \|\epsilon [\partial U_j / \partial n]\|_{L_2(\partial K)}. \end{aligned} \tag{4.19}$$

*Here* $[\partial U_j/\partial n]$ *is the jump across* $\partial K$ *in the exterior normal derivative* $\partial U_j / \partial n$.

*Proof.* Since $U$ is constant on $I_j$ we have by (4.17) and (4.3), for any $\chi \in V_h$,

$$\begin{aligned} r_{t_{j-1},t}(U,v) &= a\big(U,V(t)\big) + \big(U' - f(U), V(t)\big) + \big([U]_{t_{j-1}}, v(t_{j-1}^+)\big) \\ &= a\big(U_j,V(t)\big) + \big(k_j^{-1}[U]_{t_{j-1}} - f(U_j), V(t)\big) \\ &\quad + \big([U]_{t_{j-1}}, v(t_{j-1}^+) - k_j^{-1}V(t)\big) \\ &= \Big(a\big(U_j,V(t)-\chi\big) + \big(k_j^{-1}[U]_{t_{j-1}} - f(U_j), V(t)-\chi\big)\Big) \\ &\quad + \big([U]_{t_{j-1}}, v(t_{j-1}^+) - k_j^{-1}V(t)\big) \equiv \mathrm{I} + \mathrm{II}, \end{aligned}$$

where $V(t) = \int_{t_{j-1}}^t v\,d\sigma = \int_{\tilde{I}_j} v\,d\sigma$. Here, by the definition of $R_t$ and a simple estimation,

$$\begin{aligned} |\mathrm{II}| &\le \|k_jR_t\|_l \|v(t_{j-1}^+) - k_j^{-1}V(t)\|_{-l} \\ &\le \|k_jR_t\|_l \min\Big(\|v'\|_{L_1(\tilde{I}_j,H^{-l})}, 2\|v\|_{L_\infty(\tilde{I}_j,H^{-l})}\Big). \end{aligned}$$

We now estimate I. We write $g = -f(U_j) + k_j^{-1}[U]_{t_{j-1}}$ for short. Elementwise integration by parts gives, since $\Delta U_j = 0$ on $K$,

$$\begin{aligned} a(U_j,w) + (g,w) &= \sum_K \Big((\epsilon\nabla U_j, \nabla w)_K + (g,w)_K\Big) \\ &= \sum_K \Big((-\epsilon\Delta U_j + g, w)_K + (\epsilon\partial U_j/\partial n, w)_{\partial K}\Big) \\ &= \sum_K \Big((g,w)_K - \tfrac12(\epsilon[\partial U_j/\partial n], w)_{\partial K\setminus\partial\Omega}\Big), \quad \forall w \in V. \end{aligned}$$

With $w = V(t) - \chi$ we get

$$\begin{aligned} |\mathrm{I}| \le \sum_K \Big(&\|g\|_{L_2(K)}\|V(t)-\chi\|_{L_2(K)} \\ &+ \tfrac12\|\epsilon[\partial U_j/\partial n]\|_{L_2(\partial K\setminus\partial\Omega)}\|V(t)-\chi\|_{L_2(\partial K)}\Big). \end{aligned}$$

We choose $\chi = \Pi V(t)$, where $\Pi : V \to V_h$ is an interpolation operator such that

$$\|D^l(w - \Pi w)\|_{L_2(K)} \le Ch_K^{m-l}\|D^m w\|_{L_2(S_K)}, \quad l = 0,1,\ m = 1,2, \tag{4.20}$$

where $S_K$ is the union of all simplices adjacent to $K$; see [11]. Using also the (scaled) trace inequality

$$\|w\|_{L_2(\partial K)} \le C\Big(h_K^{-1/2}\|w\|_{L_2(K)} + h_K^{1/2}\|Dw\|_{L_2(K)}\Big), \tag{4.21}$$

we obtain

$$\|V(t) - \Pi V(t)\|_{L_2(\partial K)} \le Ch_K^{m-1/2}\|D^mV(t)\|_{L_2(S_K)}, \quad m = 1,2.$$

Hence

$$\begin{aligned} |\mathrm{I}| &\le C\sum_K h_K^m\Big(\|g\|_{L_2(K)} \\ &\qquad + h_K^{-1/2}\|\epsilon[\partial U_j/\partial n]\|_{L_2(\partial K\setminus\partial\Omega)}\Big)\|D^mV(t)\|_{L_2(S_K)} \\ &\le C\|h^mR_x\|_{L_2}\,\|D^mV(t)\|_{L_2} \le C\|h^mR_x\|_{L_2}\Big\|\int_{\tilde{I}_j} v\,d\sigma\Big\|_m, \end{aligned}$$

where the constant depends only on the constants in (4.20) and (4.21). □

Note that the quantities $R_t$ and $R_x$, in the previous theorem are computable a posteriori from the numerical solution $U$.

We now show that the local residuals tend to zero with $h$ and $k$ by using local a priori estimates from §2. Essentially, the idea is to compare $U$ with an auxiliary local solution $\tilde{u}(t) = S(t - T_n, U(T_n))$ (we use the notation $T_n = n\tau$ from §3). In order to present the idea with a minimum of technical details we make a very strong assumption about the mesh. For a more general result we refer to [4].

**Theorem 4.3.** *Assume that the family of spatial meshes is quasi-uniform and that the the time step is constant with $h^2 k^{-1} \le C$, where now $h = \max_K h_K$, $k = k_j$. Let $U \in \tilde{\mathcal{V}}_h$ be a solution of* (4.16) *with $\|U(t)\|_1 \le R$ for $t \in [0,T]$. Let $\tau$ be as in Theorems 2.3 and 2.6 and assume that $t_{j-1} \ge \tau$. Then, for $l = 0, 1$, we have*

$$\|kR_t\|_{L_\infty(I_j,H^l)} + \|h^{2-l}R_x\|_{L_\infty(I_j,H)} \le C(R)\big(h^{2-l} + k\big).$$

*Proof.* Since $t_{j-1} \ge \tau$ we have $I_j \subset [T_{n+1}, T_{n+2}]$ for some $n \ge 0$ (we use the notation $T_n = n\tau$ from §3). Let $\tilde{u}(t) = S(t - T_n, U(T_n))$ be the local exact solution starting with $\tilde{u}(T_n) = U(T_n)$. Theorem 2.6 gives for the error $\tilde{e} = U - \tilde{u}$ the bound

$$\|\tilde{e}\|_{L_\infty((T_{n+1},T_{n+2}),H^l)} \le C(R)\big(h^{2-l} + k\big).$$

Hence

$$\begin{aligned}\|kR_t\|_{L_\infty(I_j,H^l)} &= \|[U]_{t_{j-1}}\|_l = \|[\tilde{e}]_{t_{j-1}}\|_l \le 2\|\tilde{e}\|_{L_\infty((T_{n+1},T_{n+2}),H^l)}\\ &\le C(R)\big(h^{2-l} + k\big).\end{aligned}$$

Similarly, for the first term in $R_x$,

$$\|h^{2-l}k_j^{-1}[U]_{t_{j-1}}\|_{L_\infty(I_j,H)} \le C(R)h^{2-l}k^{-1}\big(h^{2-l} + k\big) \le C(R)\big(h^{2-l} + k\big).$$

Also, by Lemma 1.1,

$$\|h^{2-l}f(U)\|_{L_\infty(I_j,H)} \le C(R)h^{2-l}.$$

Finally, for the last contribution to $R_x$, we use (4.21) and Theorem 1.3,

$$\begin{aligned}h^{2-l}\Big(\sum_K h^{-1}\|\epsilon[\partial U_j/\partial n]\|^2_{L_2(\partial K)}\Big)^{1/2} &= h^{2-l}\Big(\sum_K h^{-1}\|\epsilon[\partial \tilde{e}/\partial n]\|^2_{L_2(\partial K)}\Big)^{1/2}\\ &\le Ch^{2-l}\Big(\sum_K h^{-1}\|D\tilde{e}\|^2_{L_2(\partial K)}\Big)^{1/2}\\ &\le Ch^{2-l}\Big(\sum_K h^{-1}\big(h^{-1/2}\|D\tilde{e}\|_{L_2(K)} + h^{1/2}\|D^2\tilde{e}\|_{L_2(K)}\big)^2\Big)^{1/2}\\ &\le Ch^{2-l}\Big(h^{-1}\|D\tilde{e}\|_{L_\infty((T_{n+1},T_{n+2}),H)} + \|D^2\tilde{e}\|_{L_\infty((T_{n+1},T_{n+2}),H)}\Big)\\ &\le C\Big(h^{1-l}\|\tilde{e}\|_{L_\infty((T_{n+1},T_{n+2}),H^1)} + h^{2-l}\|D^2\tilde{u}\|_{L_\infty((T_{n+1},T_{n+2}),H)}\Big)\\ &\le C(R)\big(h^{2-l} + k\big).\end{aligned}$$

□

4.3. **A global error estimate.** We now combine the results of §4.1 and §4.2 to obtain a global a posteriori error estimate. We estimate both the $H$ and $V$ norms of the error ($l = 0, 1$). To achieve this we use a duality argument:

$$\|e\|_l \le C\|A^{l/2}e\| = C \sup_{\|\varphi\|=1} (A^{l/2}e, \varphi) = C \sup_{\|\varphi\|=1} (e, A^{l/2}\varphi).$$

We thus take $\psi = A^{l/2}\varphi$, $\|\varphi\| = 1$, in (4.13) and (4.10), so that $z(s) = K(s,t)A^{l/2}\varphi$. Using also Lemma 4.2 with $m = 2 - l$ and Lemma 4.1 we get the following.

**Theorem 4.4.** *Let $U$ and $u$ be solutions of* (4.16) *and* (4.1)*, respectively. Assume that $\|U(t)\|_1 \le R$, $\|u(t)\|_1 \le R$, for $t \in [0,T]$. Then, for $t \in [0,T]$, $l = 0,1$, with $n$ such that $t \in [t_n, t_{n+1})$,*

$$\begin{aligned} \|e(t^-)\|_l \le\ & CS_l^0(0,t)\|e(0^-)\|_l \\ & + CS_l^1(0,t)\|kR_t\|_{L_\infty((0,t),H^l)} \\ & + CS_l^2(0,t)\|h^{2-l}R_x\|_{L_\infty((0,t),H)} \\ & + C(R)S_l^3(0,t)\|e\|_{L_\infty((0,t),H^1)}\,\|e\|_{L_\infty((0,t),H^l)}, \end{aligned}$$

*where the so called stability factors are defined by*

$$\begin{aligned} S_l^0(0,t) &= \sup_{\|\varphi\|=1}\max_{\tau\in[0,t]}\|A^{-l/2}K(0,\tau)A^{l/2}\varphi\|, \\ S_l^1(0,t) &= \sup_{\|\varphi\|=1}\Big(\int_0^{t_{n-1}}\|A^{-l/2}D_sK(s,t)A^{l/2}\varphi\|\,ds \\ &\qquad + \max_{s\in[t_{n-1},t]}\|A^{-l/2}K(s,t)A^{l/2}\varphi\|\Big), \\ S_l^2(0,t) &= \sup_{\|\varphi\|=1}\Big(\int_0^{t_{n-1}}\|A^{(2-l)/2}K(s,t)A^{l/2}\varphi\|\,ds \\ &\qquad + \Big\|\int_{t_{n-1}}^t A^{(2-l)/2}K(s,t)A^{l/2}\varphi\,ds\Big\|\Big), \\ S_l^3(0,t) &= \sup_{\|\varphi\|=1}\int_0^t\|A^{(1-l)/2}K(s,t)A^{l/2}\varphi\|\,ds. \end{aligned}$$

Note in this connection that

$$\|A^{-l/2}K(s,t)A^{l/2}\|_{\mathcal{L}(H)} = \|A^{l/2}L(t,s)A^{-l/2}\|_{\mathcal{L}(H)}.$$

The special treatment of the last interval $(t_{n-1},t)$ is necessitated by the expected singular behavior (cf. Lemma 3.1)

$$\|A^{(2-l)/2}K(s,t)A^{l/2}\varphi\| \sim \|A^{-l/2}D_sK(s,t)A^{l/2}\varphi\| \sim (t-s)^{-1}, \quad \text{as } s\to t.$$

The stability factors grow fast with $t$ if the linearized problem is unstable.

Of course, the error estimate of the the previous theorem is not "closed" because the square of the error occurs on the right hand side. However, this square remainder can be removed by means of the contraction mapping principle, if we compute with a sufficiently small tolerance. We demonstrate this in a slightly different context in the next section.

## 5. Shadowing—second approach

We finish these lectures by describing another approach to shadowing based on a posteriori error analysis. This is work in progress together with M. G. Larson, and we only sketch a simplified version.

The error estimate in Theorem 4.4 is based on the stability of the adjoint (4.10) of the linearized initial value problem (3.3). We now use the stability of the boundary value problem (3.7). Let $t \in (0,T)$, $\psi \in H$, and let $z$ be a solution of

$$-z' + Az + B^*(s)z = 0, \ \ s \in (0,T), \ \ s \neq t; \quad [z]_t = -\psi. \tag{5.1}$$

Thus $z$ has a jump at $s = t$. This equation does not determine $z$ uniquely; we must provide appropriate boundary conditions at $s = 0, T$. But for the time being we simply assume that we have some solution operator, denoted $z(s) = G(t,s)^*\psi$, of (5.1).

Multiplying (from the left) by a test function $w \in \tilde{\mathcal{V}}$ and integrating by parts as in (4.8), taking into account that $z$ is discontinuous at $t$, we get

$$\begin{aligned} 0 &= \int_0^T (w, -z' + Az + B^*z)\, d\sigma \\ &= \int_{J_{0,T}} \Big( (w', z) + a(w,z) + (Bw, z) \Big)\, d\sigma \\ &\quad + \sum_{J_{0,T}^-} ([w]_{t_j}, z(t_j)) + (w(0^-), z(0)) - (w(T^-), z(T)) + (w(t^-), [z]_t). \end{aligned}$$

Thus $z$ satisfies

$$\tilde{B}_{0,T}(w,z) = (w(t^-), \psi), \quad \forall w \in \tilde{\mathcal{V}}, \tag{5.2}$$

where now, instead of (4.8),

$$\begin{aligned} \tilde{B}_{0,T}(w,v) = &\int_{J_{0,T}} \Big( (w', v) + a(w,v) + (Bw, v) \Big)\, d\sigma \\ &+ \sum_{J_{0,T}^-} ([w]_{t_j}, v(t_j^+)) + (w(0^-), v(0^+)) - (w(T^-), v(T^+)). \end{aligned} \tag{5.3}$$

Combining this definition with (4.7) yields, instead of (4.9),

$$\begin{aligned} e \in \tilde{\mathcal{V}}; \quad \tilde{B}_{0,T}(e,v) &= (e(0^-), v(0^+)) - (e(T^-), v(T^+)) \\ &\quad + R_{0,T}(U,v) + \int_0^T (\eta, v)\, d\sigma, \quad \forall v \in \tilde{\mathcal{V}}. \end{aligned} \tag{5.4}$$

Together with (5.2) this gives, instead of (4.13),

$$\begin{aligned} (e(t^-), \psi) &= (e(0^-), z(0)) - (e(T^-), z(T)) \\ &\quad + R_{0,T}(U,z) + \int_0^T (\eta, z)\, d\sigma, \quad \forall \psi \in H. \end{aligned} \tag{5.5}$$

Note in this connection that $z(s)$ may have an additional jump at $s = t$ in which case $z \notin \tilde{\mathcal{V}}$. However, we may still use $v = z$ in (5.4) because $R_{0,T}(U,v)$ does not involve $v'$; see (4.4).

The error representation (5.5) is similar to (3.13). In fact, if the linearized problem has an exponential dichotomy, then we choose $z(s) = G(t,s)^*\psi$, where $G(t,s)$ is the Green's operator in (3.14).

With $U$ given we want to solve equation (5.5) for $e$. We then obtain a shadow solution $u = U - e$. In order to achieve this we set the boundary data $(e(0^-), z(0)) = (e(T^-), z(T)) = 0$. With $U$ fixed we consider $\eta = \eta(U,u) = \eta(U, U-e)$ as a function of $e$ and define the nonlinear operator

$$N_{0,T}(e,v) = \int_0^T (\eta(U, U-e), v)\, d\sigma.$$

We then write (5.5) as a fixed point equation

$$(e(t^-), \psi) = R_{0,T}(U, z) + N_{0,T}(e, z), \quad \forall \psi \in H. \tag{5.6}$$

As before we take $\psi = A^{l/2}\varphi$, $\|\varphi\| = 1$, in (5.6), so that $z(s) = G(t,s)^* A^{l/2}\varphi$. Define the ball

$$\mathcal{B}_\rho = \{e \in L_\infty((0,T), V) : \|e\|_{L_\infty((0,T),H^1)} \le \rho\}.$$

Similarly to Lemma 4.1 it follows that $N_{0,T}$ has a small Lipschitz constant in $\mathcal{B}_\rho$: if $\psi = A^{1/2}\varphi$, $\|\varphi\| = 1$, then

$$|N_{0,T}(e_1, z) - N_{0,T}(e_2, z)| \le S_1^3 C(R)\rho \|e_1 - e_2\|_{L_\infty((0,T),H^1)}, \quad e_1, e_2 \in \mathcal{B}_\rho. \tag{5.7}$$

By Lemma 4.2 we get an error bound similar to that in Theorem 4.4. It is of the form

$$|R_{0,T}(U, z)| \le CS_1^1 \|kR_t\|_{L_\infty((0,T),H^1)} + CS_1^2 \|hR_x\|_{L_\infty((0,T),H)} =: \mathcal{R}(U). \tag{5.8}$$

Here the stability factors $S_1^i = S_1^i(0,T)$, $i = 1, 2, 3$, are defined as in Theorem 4.4, but using $G(t,s)^*$ instead of $K(s,t) = L(t,s)^*$ and giving special treatment to an interval near $s = t$ instead of the endpoint $s = T$.

We can now apply the contraction mapping principle to the equation (5.6) in the ball $\mathcal{B}_\rho$ and obtain the following.

**Theorem 5.1.** (Numerical shadowing theorem 2.) *Assume that*

$$\|U\|_{L_\infty((0,T),H^1)} \le R.$$

*Let $\rho$ be so small that*

$$S_1^3 C(R)\rho \le \tfrac{1}{2}. \tag{5.9}$$

*Let $h$ and $k$ be so small that*

$$\mathcal{R}(U) \le \tfrac{1}{2}\rho. \tag{5.10}$$

*Then $U$ has a shadow $u$ with*

$$\|U - u\|_{L_\infty((0,T),H^1)} \le \rho. \tag{5.11}$$

*Proof.* The assumption (5.9) together with (5.7) guarantees that the right side of (5.6) defines a contraction on $\mathcal{B}_\rho$. The assumption (5.10) together with (5.8) guarantees that $\mathcal{B}_\rho$ is mapped into itself. Hence there is a unique solution $e \in \mathcal{B}_\rho$ of (5.6). The shadow solution is given by $u = U - e$. $\square$

As in Theorem 4.3 it follows from local a priori estimates that the quantity $\mathcal{R}(U)$ tends to zero with $h$ and $k$, so that (5.10) can be achieved if the stability factors occurring in its definition are not too large. This would be the case if the linearized problem has an exponential dichotomy on the interval $(0,T)$.

Note that, in addition to the existence of $u$, we also get an error bound (5.11) in the $V$ norm. By an additional application of the contraction property, now with respect to $H$, we get an error bound also in $H$, cf. the proof of Theorem 3.7.

The radius $\rho$ plays the role of a computational tolerance; see (5.10) and (5.8). Thus, we have shadowing if we compute with a sufficiently small tolerance.

The fixed point iteration that we employ in the proof is a quasi-Newton iteration for the equation

$$u' + Au - f(u) = 0,$$

namely

$$e'_{n+1} + Ae_{n+1} - f'(U(t))e_{n+1} = -(u'_n + Au_n - f(u_n)), \quad u_n = U - e_n.$$

However, the derivatives that occur here must be understood in a weak sense. That is why we argue by duality as in (5.6).

## References


1. W.-J. Beyn, *On the numerical approximation of phase portraits near stationary points*, SIAM J. Numer. Anal. **24** (1987), 1095–1113.
2. P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.
3. C. M. Elliott and S. Larsson, *Error estimates with smooth and nonsmooth data for a finite element method for the Cahn-Hilliard equation*, Math. Comp. **58** (1992), 603–630, S33–S36.
4. D. Estep, M. G. Larson, and R. D. Williams, *Estimating the error of numerical solutions of systems of nonlinear reaction-diffusion equations*, Mem. Amer. Math. Soc., to appear.
5. L. C. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 1998.
6. D. Henry, *Geometric Theory of Semilinear Parabolic Equations*, Lecture Notes in Mathematics, vol. 840, Springer-Verlag, 1981.
7. S. Larsson, *Nonsmooth data error estimates with applications to the study of the long-time behavior of finite element solutions of semilinear parabolic problems*, preprint 1992–36, Department of Mathematics, Chalmers University of Technology, 1992.
8. S. Larsson and S. Yu. Pilyugin, *Numerical shadowing near the global attractor for a semilinear parabolic equation*, preprint 1998–21, Department of Mathematics, Chalmers University of Technology and Göteborg University, 1998.
9. S. Larsson and J.-M. Sanz-Serna, *The behavior of finite element solutions of semilinear parabolic problems near stationary points*, SIAM J. Numer. Anal. **31** (1994), 1000–1018.
10. ______, *A shadowing result with applications to finite element approximation of reaction-diffusion equations*, Math. Comp. **68** (1999), 55–72.
11. L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp. **54** (1990), 483–493.
12. V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, Springer Series in Computational Mathematics, vol. 25, Springer-Verlag, Berlin, 1997.



Department of Mathematics, Chalmers University of Technology and Göteborg University, SE–412 96 Göteborg, Sweden

*E-mail address*: `stig@math.chalmers.se, http://www.math.chalmers.se/~stig`