A second-order positivity preserving scheme for semilinear parabolic problems

Hansen, Eskil; Kramer, Felix; Ostermann, Alexander

Published in:
Applied Numerical Mathematics

DOI:
10.1016/j.apnum.2012.06.003

2012

Citation for published version (APA):
A second-order positivity preserving scheme for semilinear parabolic problems

Eskil Hansen\textsuperscript{a}, Felix Kramer\textsuperscript{b}, Alexander Ostermann\textsuperscript{b}

\textsuperscript{a}Centre for Mathematical Sciences, Lund University, P.O. Box 118, SE-22100 Lund, Sweden
\textsuperscript{b}Institut für Mathematik, Universität Innsbruck, Technikerstraße 13, A-6020 Innsbruck, Austria

Abstract
In this paper we study the convergence behaviour and geometric properties of Strang splitting applied to semilinear evolution equations. We work in an abstract Banach space setting that allows us to analyse a certain class of parabolic equations and their spatial discretizations. For this class of problems, Strang splitting is shown to be stable and second-order convergent. Moreover, it is shown that exponential operator splitting methods and in particular the method of Strang will preserve positivity in certain situations. A numerical illustration of the convergence behaviour is included.

Key words: Semilinear parabolic problems, Strang splitting, stability, convergence, positivity, invariant sets.

1. Introduction

Stability, high-order consistency and preservation of geometric properties form three pillars on which numerical methods for differential equations rest. Whereas stability and consistency have already received much attention in the past, the preservation of properties of the exact flow under numerical discretization is a more recent field of research. For an impressive study of geometric integrators, we refer to the monograph [6].

Methods that preserve the positivity of the semiflow of certain parabolic differential equations and their spatial discretizations have attracted much interest in the past. Positivity is an important feature in computational biology and in reaction kinetics, for example, where the state variables represent sizes of populations, densities, temperatures or concentrations. The preservation of positivity is not at all a trivial task for standard numerical methods. As a matter of
fact, Bolley and Crouzeix have shown in [1] that the order of an unconditionally positive Runge–Kutta or multistep method, applied to an inhomogeneous linear parabolic equation

\[ \dot{u} + Au = f \]  

(1)
cannot exceed one, in general. Therefore, the backward Euler method is the only standard scheme of interest preserving positivity.

The situation is slightly better for exponential integrators. It has been shown in [12, 13] that second-order exponential Runge–Kutta and exponential multistep methods exist that preserve positivity in (1). None of these methods, however, preserves positivity of semilinear parabolic problems. There is a huge literature on positivity preservation for particular problems. As an example, we mention [2], where a particular integration scheme is devised that preserves positivity in biochemical systems.

In contrast to standard numerical schemes, exponential (operator) splitting methods make direct use of the splitted semiflows. In many situations this feature helps to preserve properties of the exact flow by the numerical discretization.

In this paper, we are concerned with the time integration of the semilinear parabolic problem

\[ \dot{u} + Au = f(u), \quad u(0) = u_0. \]  

(2)
The precise analytic framework is given in Section 2 below. As numerical scheme, we will employ a particular exponential splitting method, the well-known Strang splitting. This scheme has been analysed in several papers, see [4, 10, 11] and references therein. A convergence analysis for exponential splitting methods applied to linear evolution equations is given in [7]. For a comprehensive overview on splitting methods for parabolic equations, we refer to the monograph [9].

In the context of problem (2) the Strang splitting can be described as follows. Let \( e^{tf} \) denote the nonlinear semigroup generated by \( f \), i.e., \( v(t) = e^{tf}(v_0) \) is the solution of the ordinary differential equation \( \dot{v} = f(v) \) at time \( t \) with initial value \( v(0) = v_0 \). The numerical solution \( u_n \) that approximates the exact solution of (2) at time \( t_n = nh \) is given by

\[ u_n = S^n(u_0), \]

where \( S \) denotes the nonlinear operator

\[ S = e^{-\frac{h}{2}A}e^{hf}e^{-\frac{h}{2}A}. \]  

(3)
The properties of this operator are the main topic of this paper.

The remainder of the paper is organised as follows: We commence in Section 2 with describing the abstract setting and the employed assumptions. Consistency is treated in Section 3, and convergence is shown in Section 4. The main results are Theorems 4 and 5, where the Strang splitting is shown to be first- or second-order convergent depending on the regularity of the initial value. Section 5 deals with some geometric properties of the numerical semiflow. Finally, a numerical example illustrating the proved convergence is given in Section 6.
Throughout the paper, $X$ will denote an arbitrary Banach space with norm $\| \cdot \|$. The composition of two operators $g_1$ and $g_2$, say, will be denoted by $g_2g_1$ and consequently

$$g_2g_1(u) = g_2(g_1(u)).$$

Moreover, for an operator $g : \mathcal{D}(g) \subseteq X \to X$ we introduce the Lipschitz constant $L[g]$, which is the smallest constant satisfying

$$\|g(u) - g(v)\| \leq L\|u - v\|$$

for all $u$ and $v$ in $\mathcal{D}(g)$. Finally, $C$ will be a generic positive constant, independent of both $n$ and $h$, which may assume different values at different occurrences.

2. Problem setting

For our analysis below, we consider (2) as an abstract evolution equation in an appropriate Banach space. Our main assumption which allows us to analyse parabolic equations is the following.

**Assumption 1.** The linear operator $A : \mathcal{D}(A) \subseteq X \to X$ generates a bounded $C_0$ semigroup $e^{-tA} : X \to X$.

The theory of $C_0$ semigroups is well covered by many textbooks and monographs, see [5, 8, 14]. We recall that there exists an equivalent norm on $X$

$$\|u\|_* = \sup_{t \geq 0} \|e^{-tA}u\|$$

for which the semigroup is contractive. Without loss of generality we may thus assume that

$$\|e^{-tA}\| \leq 1.$$  \hspace{1cm} (4)

We will consider evolution equations (2) with nonlinear operators $f$ fulfilling the assumptions below.

**Assumption 2.** The operator $f : X \to X$ has the following properties:

(i) $f$ is twice continuously Fréchet differentiable for all $u \in X$ with derivatives denoted by $Df[u] : X \to X$ and $D^2f[u] : X \times X \to X$.

(ii) The subspaces $\mathcal{D}(A)$ and $\mathcal{D}(A^2) \subseteq X$ are both invariant under $f$ and $\mathcal{D}(A)$ is also invariant under $Df[u]$ for all $u \in \mathcal{D}(A)$.

Next consider the full vector field

$$F = -A + f : \mathcal{D}(A) \subseteq X \to X.$$

As the operator $f$ is continuously differentiable, it follows by a standard perturbation result, see [14, Theorem 6.1.5], that there exists a unique (mild) solution $u(t) = e^{tF}(u_0)$ to the evolution equation (2) for every $u_0 \in X$ and for sufficiently short time intervals, say $[0, t_{\text{end}}]$. The function $u(t)$ is also a classical
solution for \( u_0 \in \mathcal{D}(A) \), which implies that the subspace \( \mathcal{D}(A) \) is invariant under the solution operator \( e^{tF} \). Note that the uniqueness of the solution yields that \( e^{tF} \) is a (nonlinear) semigroup. The differentiability of \( f \) also implies that the semigroup \( e^{tf} : X \rightarrow X \) is well defined for sufficiently small times \( t \).

To assure the stability of our numerical scheme (3), we assume the following:

**Assumption 3.** The operator \( e^{tf} \) satisfies \( L[e^{tf}] \leq e^{t\omega} \), where \( \omega \in \mathbb{R} \).

**Example 1.** Consider the one-dimensional Chafee–Infante equation [3], also known as Allen–Cahn equation,

\[
\dot{u} - \Delta u = u - u^3
\]

on a bounded interval \( \Omega = (a, b) \) with homogeneous Neumann boundary conditions. If we choose \( X = C(\bar{\Omega}) \) then the operator \( A = -\Delta \), with domain

\[
\mathcal{D}(A) = \{ u \in C^2(\bar{\Omega}) : u'(a) = u'(b) = 0 \},
\]
generates a \( C_0 \) semigroup of contractions on \( X \), i.e., Assumption 1 is valid and (4) holds without rescaling the norm of \( X \). The generating properties of \( A \) are derived in [5, Section II.3.30], for example. Furthermore, \( f(u) = u - u^3 \) is well defined and twice continuously Fréchet differentiable, with

\[
Df[u]v = (1 - 3u^2)v \quad \text{and} \quad D^2f[u](v, w) = -6uvw.
\]

The remaining invariance properties of Assumption 2 follow by inspection. The nonlinear semigroup generated by \( f \) is (pointwise) given by

\[
e^{tf}(y) = \frac{y}{\sqrt{y^2 + (1 - y^2)e^{-2t}}}, \quad y \in \mathbb{R},
\]

and Assumption 3 holds for \( \omega = 1 \). The latter follows by using the mean value theorem together with the pointwise bound \( |\partial e^{tf}(y)/\partial y| \leq e^{t} \).

### 3. Consistency

We start off by deriving consistency of the splitting scheme. For this we employ a similar technique as used by Jahnke and Lubich [10] in the context of linear problems, namely, we identify the local errors generated by the splitting scheme as quadrature errors. For convenience, will use the notation \( T = e^{hF} \) henceforth.

**Lemma 2.** If Assumptions 1 and 2 are valid and \( u \in \mathcal{D}(A) \), then

\[
\| (T - S)(u) \| \leq Ch^2.
\]
Proof. A two-term Taylor expansion of e^{hf} in S gives us the expression
\[ S(u) = e^{-hA}u + he^{-\frac{h}{2}A}fe^{-\frac{h}{2}A}(u) + h^2 R_S(u), \]
where the operator \( R_S : X \to X \) is given by
\[ R_S(u) = \int_0^1 (1-s)e^{-\frac{h}{2}A}Df[e^{shf}e^{-\frac{h}{2}A}(u)]fe^{shf}e^{-\frac{h}{2}A}(u)ds. \]

To ease the forthcoming calculations, we introduce the function \( v \) defined by
\[ v(t) = \int_0^1 e^{-t(1-s)}Afe^{stF}(u)ds = \frac{1}{t} \int_0^t e^{-(t-s)}Afe^{sF}(u)ds. \]  \( (6) \)

Applying the variation-of-constants formula twice and using a one-term Taylor expansion of \( f \) then yields
\[ T(u) = e^{-hA}u + \int_0^h e^{-(h-s)A}f(e^{-sA}u + sv(s))ds = e^{-hA}u + \int_0^h e^{-(h-s)A}fe^{-sA}(u)ds + h^2 R_T(u), \]
where the remainder \( R_T : X \to X \) is defined as
\[ R_T(u) = \frac{1}{h^2} h \int_0^h s \int_0^1 e^{-(h-s)A}Df[e^{-sA}u + \tau sv(s)]v(s)d\tau ds. \]

With these expansions of \( S \) and \( T \), it is easy to identify the difference \( T - S \) as the local error related to the midpoint rule. To this end, introduce
\[ g(t) = e^{-(h-t)A}fe^{-tA}(u) \]
and we then obtain that
\[ (T-S)(u) = \int_0^h (g(s) - g(h/2))ds + h^2 R(u) = \int_{-h/2}^{h/2} s \int_0^1 g'(h/2 + \tau s)d\tau ds + h^2 R(u), \]
where \( R = R_T - R_S \). The proof is then complete if \( g'(t) \) is an element in \( X \) when \( u \in D(A) \). This follows by simply writing out \( g'(t) \) as
\[ g'(t) = e^{-(h-t)A}Af e^{-tA}(u) - e^{-(h-t)A}Df[e^{-tA}(u)]e^{-tA}Au \]
and noting that the terms on the right-hand side are well defined for \( u \in D(A) \).

\[ \square \]

Lemma 3. If Assumptions 1 and 2 are valid and \( u \in D(A^2) \), then
\[ \|(T - S)(u)\| \leq Ch^3. \]
PROOF. We proceed as in the proof of Lemma 2, with the only difference that the expansions are made up to third-order remainder terms. Starting with $S$ gives

$$S(u) = e^{-hA}u + he^{-\frac{3}{2}A}fe^{-\frac{3}{2}A}(u) + \frac{1}{2}h^2e^{-\frac{5}{2}A}Df[e^{-\frac{5}{2}A}u]fe^{-\frac{5}{2}A}(u) + h^3R_S(u),$$

where the remainder $R_S : X \rightarrow X$ is given by

$$R_S(u) = \frac{1}{2} \int_0^1 (1-s)^2e^{-\frac{3}{2}A}\left(D^2f[e^{shf}e^{-\frac{3}{2}A}(u)]\left(fe^{shf}e^{-\frac{3}{2}A}(u), fe^{shf}e^{-\frac{3}{2}A}(u)\right) + Df[e^{shf}e^{-\frac{3}{2}A}(u)]^2fe^{shf}e^{-\frac{3}{2}A}(u)\right)ds.$$

With the function $v$ defined as in (6), we can write the expansion of $T$ as

$$T(u) = e^{-hA}u + \int_0^h e^{-(h-s)A}f(e^{-sA}u + sv(s))ds$$

$$= e^{-hA}u + \int_0^h e^{-(h-s)A}f e^{-sA}(u)ds + \int_0^h e^{-(h-s)A}Df[e^{-sA}u]sv(s)ds$$

$$+ \int_0^h \frac{1}{2} h^2\int_0^1 (1-\tau)e^{-(h-s)A}D^2f[e^{-sA}u + \tau sv(s)](v(s), v(s))d\tau ds.$$

To proceed we need to expand the term $sv(s)$ in the second term from the right, which can be done as follows:

$$sv(s) = \int_0^s e^{-(s-\sigma)A}f(e^{-\sigma A}u + \sigma v(\sigma))d\sigma$$

$$= \int_0^s e^{-(s-\sigma)A}fe^{-\sigma A}(u)d\sigma + \int_0^s \int_\sigma^1 e^{-(s-\sigma)A}Df[e^{-\sigma A}u + \tau sv(\sigma)]v(\sigma)d\tau d\sigma.$$

By collecting the terms, we obtain the expansion

$$T(u) = e^{-hA}u + \int_0^h e^{-(h-s)A}f e^{-sA}(u)ds$$

$$+ \int_0^h \int_0^s e^{-(h-s)A}Df[e^{-sA}u]e^{-(s-\sigma)A}fe^{-\sigma A}(u)d\sigma ds + h^3R_T(u),$$

where the full remainder $R_T : X \rightarrow X$ is

$$R_T(u) = \frac{1}{h^3} \int_0^h \frac{1}{2}h^2\int_0^1 (1-\tau)e^{-(h-s)A}D^2f[e^{-sA}u + \tau sv(s)](v(s), v(s))d\tau ds$$

$$+ \frac{1}{h^3} \int_0^h \int_0^s \int_0^1 e^{-(h-s)A}Df[e^{-sA}u]e^{-(s-\sigma)A}Df[e^{-\sigma A}u + \tau sv(\sigma)]v(\sigma)d\tau d\sigma ds.$$

We can once more identify the difference $T - S$ as the local error of the midpoint rule, by introducing

$$g(t) = e^{-(h-t)A}fe^{-tA}(u) \quad \text{and} \quad G(t, \tau) = e^{-(h-t)A}Df[e^{-tA}u]e^{-(\tau-t)A}fe^{-\tau A}(u).$$
The local error may then be written as

\[
(T - S)(u) = \int_0^h (g(s) - g(s + \frac{h}{2}))ds + \int_0^h \int_0^s (G(s, \sigma) - G(s, \frac{h}{2} + \frac{\sigma}{2}))d\sigma ds + h^3 R(u)
\]

\[
= \int_{-h/2}^{h/2} s^2 \int_0^1 (1 - \tau)g''(h/2 + \tau s)d\tau ds
\]

\[
+ \int_{-h/2}^{h/2} \int_s^{s+h/2} \int_0^1 \left( s \frac{\partial G}{\partial s} + \sigma \frac{\partial G}{\partial \sigma} \right) (h/2 + \tau s, h/2 + \sigma) d\tau d\sigma ds
\]

\[
+ h^3 R(u),
\]

where \( R = R_T - R_S \). We can conclude that \( \|(T - S)(u)\| \leq C h^3 \), if \( g''(t) \), \( \partial G/\partial t(t, \tau) \) and \( \partial G/\partial \tau(t, \tau) \) are all well defined for \( u \in D(A^2) \). This can again be proved by inspection, for example,

\[
g''(t) = e^{-(h-t)A}A^2 f e^{-tA} + 2 e^{-(h-t)A}ADf[e^{-tA}u] e^{-tA} u
\]

\[
+ e^{-(h-t)A} D^2 f[e^{-tA}u] \left( e^{-tA} Au, e^{-tA} Au \right) + e^{-(h-t)A} Df[e^{-tA}u] e^{-tA} A^2 u.
\]

The terms involving \( G \) are in fact well defined even for \( u \in D(A) \), as they only contain first-order derivatives with respect to \( t \) and \( \tau \). \( \square \)

4. Convergence

The consistency results of the previous section and the presented Banach space framework now enable us to investigate the convergence rate of our splitting scheme.

**Theorem 4.** Consider the approximation of the solution \( u(nh) = e^{nhF}(u_0) \) to (2) by the nonlinear splitting scheme (3) with a single step given as \( S = e^{-h/2 A}e^{hA}e^{-h/2 A} \). If Assumptions 1, 2 and 3 hold and \( u_0 \in D(A) \), then the scheme is (at least) first-order convergent, i.e., the error is bounded for \( h \) sufficiently small as

\[
\|(S^n - e^{nhF}(u_0))\| \leq C h, \quad 0 \leq nh \leq t_{\text{end}}
\]

with a constant \( C \) that is uniform in \([0, t_{\text{end}}]\).

**Proof.** By expanding the error as a telescopic sum, we obtain

\[
\|(S^n - e^{nhF}(u_0))\| \leq \sum_{j=0}^{n-1} \|(S^{n-j}e^{jhF} - S^{n-j-1}e^{(j+1)hF})(u_0)\|
\]

\[
\leq \sum_{j=0}^{n-1} L \left[ S^{n-j-1} \right] \|(S - e^{hF})e^{jhF}(u_0)\|.
\]

7
As $\mathcal{D}(A)$ is invariant under $e^{tF}$ and $u_0 \in \mathcal{D}(A)$, the consistency results from Lemma 2 yield that
\[ \|(S - e^{hF})e^{jhF}(u_0)\| \leq Ch^2 \]
for all $j = 0, 1, \ldots n - 1$. Observation (4) together with Assumption 3 gives the stability bound
\[ L[S^{n-j-1}] \leq (L[e^{-h\frac{h}{2}A}]L[e^{hF}]L[e^{-h\frac{h}{2}A}])^{n-j-1} \leq e^{(n-j-1)h\omega}. \]
Taken together, these bounds imply that
\[ \|(S^n - e^{nhF})(u_0)\| \leq Ch\left(h \sum_{j=0}^{n-1} e^{(n-j-1)h\omega}\right) \leq Ch, \]
i.e., the scheme is first-order convergent for every $u_0 \in \mathcal{D}(A)$. \qed

The classical second-order convergence of the numerical scheme $S$ can also be proven whenever the exact flow $e^{tF}$ preserves the structure of $\mathcal{D}(A^2)$. The proof of second-order convergence then follows by employing the consistency result of Lemma 3 together with the very same line of reasoning as in the proof of Theorem 4.

**Theorem 5.** If the hypotheses of Theorem 4 hold and, in addition, the subspace $\mathcal{D}(A^2) \subseteq X$ is invariant under $e^{tF}$ for all $t \in [0, t_{\text{end}}]$ and $u_0 \in \mathcal{D}(A^2)$, then the scheme is second-order convergent, i.e., for $h$ sufficiently small it holds that
\[ \|(S^n - e^{nhF})(u_0)\| \leq Ch^2, \quad 0 \leq nh \leq t_{\text{end}} \]
with a constant $C$ that is uniform in $[0, t_{\text{end}}]$. \qed

**Example 6.** In the setting of Example 1, the nonlinearity $f(u) = u - u^3$ is differentiable on the Banach space $\mathcal{D}(A)$ as well. Applying [14, Theorem 6.1.5] shows that the mild solution $u(t) = e^{tF}(u_0)$ on $\mathcal{D}(A)$ is a classical solution if $u_0 \in \mathcal{D}(A^2)$. As a consequence, $\mathcal{D}(A^2)$ is invariant under $e^{tF}$, and Theorem 5 can be applied.

5. **Geometric properties**

Exponential operator splitting methods like the Strang splitting (3) are constructed by composing partial (semi)flows of the problem in an appropriate way. Due to this principle of construction, properties of the partial (semi)flows are often preserved by the numerical semiflow. We illustrate this general observation by two examples, where we discuss positivity and invariant sets.
5.1. Positivity

In order to formalise the notion of positivity in a Banach space, we need the concept of a Banach lattice. For the convenience of the reader, we briefly recall its definition. For details, we refer to the textbook [15, Chap. XII].

A real Banach space \((X, \| \cdot \|)\) with an order relation \(\leq\) is called a Banach lattice if the following three conditions hold.

(i) \((X, \leq)\) is a partially ordered set where the least upper bound \(\sup(u, v)\) and the greatest lower bound \(\inf(u, v)\) exist for any two elements \(u, v \in X\).

(ii) The vector space operations are compatible with the order relation, i.e., \(u \leq v\) implies \(u + w \leq v + w\), \(\lambda u \leq \lambda v\) for \(0 \leq \lambda \in \mathbb{R}\), and \(-v \leq -u\).

(iii) The norm in \(V\) is compatible with the absolute value \(|u| = \sup(-u, u)\), i.e., \(|u| \leq |v|\) implies \(\|u\| \leq \|v\|\).

Any element \(v \in X\) satisfying \(v \geq 0\) is called positive. A bounded operator on \(X\) is called positive if it maps positive elements to positive ones.

Example 7. The \(d\)-dimensional space \(\mathbb{R}^d\) with componentwise order (i.e., \(u \leq v\) whenever \(u_k \leq v_k\) for all \(1 \leq k \leq d\)) is an example of a Banach lattice. Other important examples are the Lebesgue spaces \(L^p(\Omega), 1 \leq p \leq \infty\), as well as their subspaces. For these spaces, the order relation is defined pointwise, i.e., we have \(u \leq v\) whenever \(u(x) \leq v(x)\) for almost all \(x \in \Omega\).

Note that the composition of positive operators is again a positive operator. Therefore, exponential operator splitting methods are positive, provided that the defining semiflows have this property. An immediate consequence of this simple observation is the following result.

Proposition 8. For the numerical solution of (2), consider the Strang splitting (3). If the semigroups \(e^{-tA}\) and \(e^{tf}\) are positive, then the Strang splitting is positive as well. \(\square\)

In Example 1, where we considered the Chafee–Infante equation, the semiflow \(e^{tF}\) as well as the partial semiflows \(e^{-tA}\) and \(e^{tf}\) are positive. Therefore, the Strang splitting (3) preserves positivity of this equation, as long as the semigroups are computed exactly.

5.2. Invariant sets

An obvious generalisation of positivity is the concept of invariant sets. Let \(g\) be a (possibly nonlinear) operator on \(X\) and \(Y \subseteq X\). The subset \(Y\) is called invariant under \(g\), if \(Y \subseteq \mathcal{D}(g)\) and \(g(Y) \subseteq Y\). Positivity is characterised by the set \(Y = \{u \in X : u \geq 0\}\).

If the set \(Y\) is invariant under two maps, it is also invariant under their composition. Therefore, exponential operator splitting methods possess an invariant set \(Y\), provided that the defining semiflows have this property.
As a simple application, we consider once more the Chafee–Infante equation (5). For this equation, the set

\[ Y = \{ u \in X ; \ u(\Omega) \subseteq [-1, 1] \} \]

is invariant under the operators \( f, e^{-tA}, e^{tf}, \) and \( e^{tF} \), respectively. In particular, it is then also invariant under the numerical semiflow (3).

![Figure 1: Numerical order of the Strang splitting applied to the Chafee–Infante equation with homogeneous Neumann boundary conditions. The errors are measured in the maximum norm. The dash-dotted reference line has slope two.](image)

6. Numerical experiment

We illustrate our convergence result with the help of Example 1. To this aim, we consider the parabolic problem (5) on the domain \( \Omega = (0, 0.5) \) for \( 0 \leq t \leq 0.1 = t_{\text{end}} \). The interval \( \Omega \) is discretized with standard finite differences using \( N = 100 \) grid points, and the initial conditions are chosen as

\[ u_0(x) = \frac{1}{10} + \frac{7}{10} \sin^2(2x + 1)\pi. \]

The errors are measured at the final time \( t_{\text{end}} \) in the maximum norm. The numerical results are displayed in Figure 1. The observed second-order convergence of Strang splitting is in line with the convergence analysis in Section 4.

References


