European Society of Computational Methods in Sciences and Engineering (ESCMSE)

Journal of Numerical Analysis, Industrial and Applied Mathematics (JNAIAM)
vol. 6, no. 1-2, 2011, pp. 15-27
ISSN 1790–8140

Error Estimation via Defect Computation and Reconstruction: Some Particular Techniques [1]

W. Auzinger[2]

Institute for Analysis and Scientific Computing,
Vienna University of Technology,
1040 Vienna, Austria

Received 20 February, 2011; accepted in revised form 10 March, 2011

Abstract: The well-known technique of defect correction can be used in various ways for estimating local or global errors of numerical approximations to differential or integral equations. In this paper we describe the general principle in the context of linear and nonlinear problems and indicate the interplay between the auxiliary scheme involved and a correct definition of the defect. Applications discussed include collocation approximations to first and second order boundary value problems for nonlinear ODEs and, in particular, exponential splitting approximations for linear evolution equations. We describe the design of error estimators and their essential properties and give numerical examples. The theoretical tools for the analysis of the asymptotical correctness of such estimators are described, and references to original research papers are given where the complete analysis is provided.

Keywords: error estimation, defect correction, collocation, exponential splitting

Mathematics Subject Classification: 65L05, 65L10, 65L70, 65M15

PACS: 02.60.Lj, 02.70.Bf, 02.70.Jn

1 Introduction and general setting

We consider several realizations of a posteriori error estimators based on the principle of defect correction. Based on earlier work from [16], the underlying principle was first described in [15] in a general, abstract framework. The idea is to combine the defect of a given numerical approximation with a 'reconstruction' (backsolving) procedure based on a simple auxiliary scheme, with the goal to produce an asymptotically correct local or global error estimate in an efficient and stable way. Such estimates are the basis for adaptive mesh control, see for instance [3, 4, 5] for the case of singular boundary value problems and resulting implementation in the code `sbvp`, see [6].

In concrete applications, the success of this approach depends on the proper choice of algorithmic components, in particular on the interplay between the way how to define the defect and the

[1] Published electronically May 15, 2011
[2] Corresponding author. E-mail: w.auzinger@tuwien.ac.at

auxiliary scheme for computing the error estimate. Since the main purpose of an error estimate is adaptation, an essential requirement is that it has to be robust with respect to arbitrary meshes, and we will exemplify how this can be achieved. We illustrate the general principle (Section 1.1) and discuss application of the method for some problem classes, namely boundary value problems for first and second order ODEs (Section 2), and, in particular, a new way of realizing the idea, namely in the context of exponential splitting schemes for linear evolution equations (Section 3). Here, the simplest case of first order Lie-Trotter splitting is considered in more detail. Numerical examples are also provided.

1.1 Design of defect-based error estimators, representation of the error

The typical design a successful defect-based error estimator can be conveniently described in an abstract, setting, with a leading differential operator $\mathcal{L}$. Let

$$\mathcal{L}u = \mathcal{F}(u) \tag{1}$$

represent a nonlinear differential equation with linear leading part $\mathcal{L}u$ of order $k \geq 1$. The exact solution for given initial or boundary data is denoted by u_* and is assumed to be locally unique. We assume that a numerical approximation $\tilde{u}$ for u_* has been computed,[3] and we wish to estimate its global error $\tilde{e} := \tilde{u} - u_*$. If $\tilde{u}(x)$ is a continuous function, e.g., a collocation polynomial, it is natural to consider the *defect* (or residual) of $\tilde{u}$ with respect to (1),

$$\tilde{d} := \mathcal{L}\tilde{u} - \mathcal{F}(\tilde{u}) \tag{2}$$

as a natural measure for the quality of $\tilde{u}$. The classical idea due to [16] is to consider the 'neighboring problem' $\mathcal{L}u = \mathcal{F}(u) + \tilde{d}$ with exact solution $\tilde{u}$, compute its numerical approximation and use its error $\tilde{\varepsilon}$ as an estimate for $\tilde{e}$. In other words: an improved approximation $\tilde{u} - \tilde{\varepsilon} \approx u_*$ is obtained via reconstruction, i.e., backsolving using the defect information. This procedure is often continued in an iterative fashion, cf., e.g. [1].

For the purpose of estimating $\tilde{e}$ (without iterating) it was proposed in [15] to use a simple, low order scheme in the backsolving process. We are adopting this point of view but realize it in a modified way. To this end, let

$$Lu = F(u) \tag{3}$$

represent your favorite low order, stable method, e.g., a simple finite difference scheme defined on the mesh used in the computation of $\tilde{u}$. We call (3) the 'auxiliary scheme'.

Furthermore we assume that the given problem (1) can recast into a form where the leading term has the same form as in (3), i.e.,

$$\mathcal{L}u = \mathcal{F}(u) \quad \rightsquigarrow \quad Lu = G(u) \tag{4}$$

with G appropriately defined, such that $Lu_* - G(u_*) = 0$. With this reformulation, we define the defect of $\tilde{u}$ as

$$\tilde{d} := L\tilde{u} - G(\tilde{u}) \tag{5}$$

instead of (2). In practice, exact evaluation of G will not be possible and needs to be approximated, $G \approx \hat{G}$, typically involving a sufficiently accurate quadrature scheme. With this numerically evaluated defect,

$$\hat{d} := L\tilde{u} - \hat{G}(\tilde{u})$$

[3] Here we are not very specific in denoting side conditions as well as mesh transfer and interpolation operators involved in the process.

we consider the solutions u_{bas}, u_{def} of

$$Lu_{bas} = F(u_{bas}), \qquad Lu_{def} = F(u_{def}) + \hat{d}, \tag{6}$$

satisfying

$$L(u_{def} - u_{bas}) = F(u_{def}) - F(u_{bas}) + \hat{d}. \tag{7}$$

On the other hand, for $\tilde{u}$ and u_* we have

$$L(\tilde{u} - u_*) = \hat{G}(\tilde{u}) - \hat{G}(u_*) + \Delta G(u_*) + \hat{d}, \tag{8}$$

with $\Delta G := \hat{G} - G$ representing the quadrature error. Note that only the quadrature error with respect to the exact solution u_* is involved; this means that no higher order smoothness properties of the numerical approximation $\tilde{u}$ are required for the analysis of the estimator.

Since F is an approximation for G, this suggests the a posteriori error estimate

$$\hat{\varepsilon} := u_{def} - u_{bas} \approx \tilde{u} - u_* = \tilde{e}$$

From (7),(8) we see that the *deviation* of $\hat{\varepsilon}$, i.e., the error $\hat{\sigma} := \hat{\varepsilon} - \tilde{e}$ of the estimate satisfies

$$L\hat{\sigma} = (F(u_{def}) - F(u_{bas})) - (F(\tilde{u}) - F(u_*)) + (\Delta F(\tilde{u}) - \Delta F(u_*)) - \Delta G(u_*), \tag{9}$$

with $\Delta F := F - \hat{G}$.

For linear problems, the estimate $\hat{\varepsilon}$ can be directly obtained by solving for $\hat{\varepsilon}$ in (6), i.e.,

$$L\hat{\varepsilon} = F_{hom}(\hat{\varepsilon}) + \hat{d},$$

where F_{hom} denotes the homogeneous part of the affine operator F. Here, (9) simplifies to

$$L\hat{\sigma} = F_{hom}(\hat{\sigma}) + \Delta F(\tilde{e}) - \Delta G(u_*). \tag{10}$$

Linear systems of this type are also solved when the method is applied to nonlinear problems in combination with Newton iteration. In practice it is favorable first to compute the basic approximation u_{bas} and use it as an initial guess for computing $\tilde{u}$ and, hereafter, of u_{def}.

From (9) (or (10), respectively) it can be seen that the deviation $\hat{\sigma}$ is not influenced by any approximation error concerning the leading part $\mathcal{L}u$. If we would use (2) instead of (5), a term involving $L - \mathcal{L}$ would appear, which typically depends on higher derivatives of u_*. This is essential for the robustness of the estimate, in particular on variable meshes (cf. the discussion and examples in [1]). The quality of our estimate depends only on

(i) the asymptotic behavior of $\Delta F(\tilde{u}) - \Delta F(u_*)$,

(ii) the quadrature error $\Delta G(u_*)$,

(iii) and the stability of the auxiliary scheme (3).

In applications of the method, the analysis of (i) is problem-specific, while (ii) and (iii) are usually standard.

2 Boundary value problems for ODEs

2.1 First order problems

The approach sketched in Section 1.1 has been realized and analyzed in [3, 7] in the following context:

- $\mathcal{L}u = \mathcal{F}(u)$ is a system of explicit first order ODEs,

$$u'(x) = f(x, u(x)), \quad x \in (a, b),$$

together with boundary conditions on an interval $[a, b]$,

- $\tilde{u} \in C^0[a, b]$ is a collocation polynomial of degree m,
- $Lu = F(u)$ is a simple difference scheme over the collocation mesh $\{x_j\}$, e.g., an Euler or midpoint scheme.

Recasting the problem according to (5) is straightforward via integrating u' between successive collocation points, such that $Lu = G(u)$ is a locally weighted version of the ODE. An appropriate approximation $\hat{G}(u)$ of $G(u)$ is obtained via quadrature over collocation subintervals, and the outcome is closely related to a higher order Runge-Kutta scheme defining the defect of $\tilde{u}$. The analysis of the resulting error estimate $\tilde{e}$ is based on the asymptotic properties of the collocation error, namely $\|\tilde{e}\| = \mathcal{O}(h^m)$, $\|\tilde{e}'\| = \mathcal{O}(h^m)$. This permits an estimate of the critical quantity $\Delta F(\tilde{u}) - \Delta F(u_*)$, leading to the conclusion that the estimate $\hat{\varepsilon}$ is asymptotically correct, i.e., the deviation $\hat{\sigma} = \hat{\varepsilon} - \tilde{e}$ satisfies $\|\hat{\sigma}\| = \mathcal{O}(h^{m+1})$. In [4, 12] this analysis was extended to the case of singular boundary value problems.

For related results concerning implicit first order systems, in particular regular and singular index 1 DAEs, see [8, 9].

2.2 Second order problems

Let us discuss how this approach can be extended to the case of direct collocation approximations $\tilde{u} \in C^1[a, b]$ for a second order two-point boundary value problem, for a nonlinear ODE with leading part $\mathcal{L}u = u''$,

$$u''(x) = F(x, u(x), u'(x)), \quad x \in (a, b)$$

For simplicity of presentation, we assume an equidistant mesh $\{x_j\}$ with meshwidth h; but this is in no way essential. Lu is the standard second order difference operator. On an interval $[x_{j-1}, x_{j+1}] = [x_j - h, x_j + h]$, the identity

$$Lu(x_j) = (\mathcal{K}u'')(x_j), \quad \text{with} \quad (\mathcal{K}f)(x) := \int_{-1}^{1} K(\xi)\, f(x + \xi h)\, d\xi, \quad K(\xi) = 1 - |\xi|, \tag{11}$$

is valid for each $u \in C^2[a, b]$. This extends to the case where a jump in u'' occurs at $x = x_j$ (this is the case for $u = \tilde{u}$ at the end of a collocation subinterval). This leads to a reformulation, in the spirit of (5), of the given ODE in the form of an 'exact finite difference scheme',

$$Lu(x_j) = (\mathcal{K}f)(x_j), \quad f(x) = F(x, u(x), u'(x))$$

where x_j runs over the complete collocation mesh including the endpoints of the subintervals. In this way, the error estimation procedure is well-defined along the lines of Section 1.1. Sufficiently accurate quadrature approximations for the occurring integrals of the type (11) are readily constructible, and minor details like approximation of the first derivative in the basic difference approximation, or incorporating boundary conditions containing first derivatives are easy to fix. A complete analysis of this approach can be found in [14], see also [11]. Sharp estimates for $\Delta F(\tilde{u}) - \Delta F(u_*)$ are derived making use of available convergence results for the error $\tilde{e} = \tilde{u} - u_*$ and its first and second derivatives. It turns out that for the second order case the resulting estimate is always of a very high quality, namely with deviation $\|\hat{\sigma}\| = \|\hat{\varepsilon} - \tilde{e}\| = \mathcal{O}(h^{m+2})$ for $\|\tilde{e}\| = \mathcal{O}(h^m)$.

This approach is related to the iterated defect correction method from [10], based on finite-difference approximations for two-point boundary value problems.

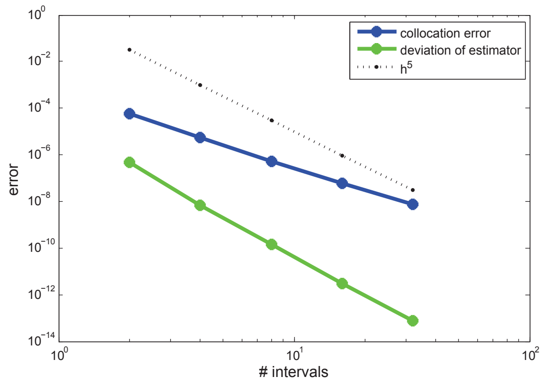

Figure 1: Error and deviation of estimate for Example (12)

2.3 Numerical example

Let us illustrate the method for the nonlinear boundary value problem

$$u''(x) = 1 - (u'(x))^2, \quad u(0) = u(1) = 0, \tag{12}$$

with known analytic solution. $\tilde{u}$ is computed as a piecewise polynomial collocating approximation of degree $m = 3$ using irregularly spaced collocation nodes. On coherent refinement of the collocation subintervals we obtain the error history shown as log-log-plot in Figure 1. The observed orders are $m = 3$ for $||\tilde{e}||_\infty$ and $m + 2 = 5$ for $||\hat{\sigma}||_\infty = ||\hat{\varepsilon} - \tilde{e}||_\infty$.

3 Exponential splitting for evolution equations

We consider a linear evolution equation $u' = Hu$ for an unknown function $u = u(t)$, which we write in the form

$$\partial_t u(t) = Hu(t), \qquad u(0) = u_0, \tag{13}$$

with exact solution

$$u_*(t) = \Phi_*(t; u_0) = \mathrm{e}^{tH} u_0.$$

Here, $\Phi_*(t; \cdot)$ denotes the exact flow of the given problem, i.e., $\Phi_*(t; v) = \mathrm{e}^{tH} v$.

We are aiming at error estimates for exponential splitting approximations to $u_*(t)$, based on the operator splitting

$$H = A + B.$$

We assume that H, A, B generate strongly continuous semigroups of operators in a Banach space $\mathcal{B}$ bounded in the induced norm $\| \cdot \| = \| \cdot \|_{\mathcal{B} \leftarrow \mathcal{B}}$. Without essentially restricting of the generality of our arguments, we assume that these operators generate semigroups of contractions. All explicit error bounds given below refer to this particular case.

The purpose of this section is to exemplify the idea how local error estimates for such approximations are designed, and to discuss their asymptotic behavior. In the terminology of Section 1.1, we have $\tilde{u} = u_{bas}$ where $\tilde{u}$ is a splitting approximation (16) or (44) to u_*. The precise analysis of these estimators including explicit bounds in terms of problem data is given forthcoming papers, see e.g. [2], with special focus on linear and nonlinear evolutionary Schrödinger equations. Special regularity properties of the initial data u_0 are not assumed, therefore this approach can be applied to an arbitrary integration step, with the goal of mesh adaptivity in time t (possibly also for adaptivity in space for an underlying spatial discretization).

Here we consider the linear case only; however, we use a denotation for the flows [4] which is more natural in view of generalization to the nonlinear problems. With the operator $\mathcal{L}$ acting on C^1-vector fields $X(t;v)$,

$$(\mathcal{L}X)(t;v) := \partial_t X(t;v) - AX(t;v) - BX(t;v), \tag{14}$$

the exact flow satisfies

$$(\mathcal{L}\Phi_*)(t;v) = 0, \quad \Phi_*(0;v) = v. \tag{15}$$

3.1 Lie-Trotter splitting

One step of the Lie-Trotter approximation to $u_*(t)$ over an interval $[0,h]$ is defined as

$$\tilde{u}(h) := \mathrm{e}^{hB}\mathrm{e}^{hA}u_0 \approx \mathrm{e}^{hH}u_0 = u_*(h). \tag{16}$$

We assume that both e^{hA} and e^{hB} are easier/cheaper to evaluate than e^{hH}. With the individual flows

$$\Phi_A(t;v) := \mathrm{e}^{tA}v, \quad \Phi_B(t;v) := \mathrm{e}^{tB}v, \tag{17}$$

(16) reads

$$\tilde{u}(h) := \Phi_B(h;\Phi_A(h;u_0)) \approx \Phi_*(h;u_0) = u_*(h), \tag{18}$$

with splitting error

$$\tilde{e}(h) := \tilde{u}(h) - u_*(h). \tag{19}$$

3.1.1 Splitting flow and its defect

The main idea in our approach to error estimation is to argue in terms of flows. We consider the approximate 'splitting flow' from (18),

$$\tilde{\Phi}(t;v) := \Phi_B(t;\Phi_A(t;v)) = \mathrm{e}^{tB}\mathrm{e}^{tA}v$$

as a continuous approximation for the exact flow $\Phi_*(t;v)$. It satifies is a modified version of (15),

$$\begin{aligned}\partial_t\,\tilde{\Phi}(t;v) &= \mathrm{e}^{tB}\,\mathrm{e}^{tA}Av + B\,\mathrm{e}^{tB}\,\mathrm{e}^{tA}v\\ &= \Phi_B(t;\Phi_A(t;Av)) + B\,\Phi_B(t;\Phi_A(t;v))\\ &= \tilde{\Phi}(t;Av) + B\,\tilde{\Phi}(t;v).\end{aligned}$$

I.e., with the operator L acting on C^1-vector fields $X(t;v)$,

$$(LX)(t;v) := \partial_t X(t;v) - X(t;Av) - BX(t;v), \tag{20}$$

the splitting flow $\tilde{\Phi}(t;v)$ satisfies

$$(L\tilde{\Phi})(t;v) = 0, \quad \tilde{\Phi}(0;v) = v. \tag{21}$$

This takes the form of a *Sylvester equation* for $\tilde{\Phi}(t;v)$, with (20),(21) approximating (14),(15). In the spirit of Section 1.1, we now rewrite the original problem (15) in the equivalent form (cf. (4)),

$$(L\Phi_*)(t;v) = G(\Phi_*)(t;v), \quad G(\Phi_*)(t;v) = A\,\Phi_*(t;v) - \Phi_*(t;Av), \tag{22}$$

and define the defect (rather: *defect flow*), according to (5), of $\tilde{\Phi}(t;v)$ with respect to (22),

$$\tilde{\mathcal{D}}(t;v) := (L\tilde{\Phi})(t;v) - G(\tilde{\Phi})(t;v) = -(A\,\tilde{\Phi}(t;v) - \tilde{\Phi}(t;Av)). \tag{23}$$

[4] Here and in the sequel, v denotes a generic initial argument for the flows considered.

Remark 3.1 *Here, the defect has been defined following the general formalism from Section 1.1. In the present context we can also directly write*

$$\tilde{\mathcal{D}}(t;v) = (\mathcal{L}\tilde{\Phi})(t;v) = \partial_t \tilde{\Phi}(t;v) - A\,\tilde{\Phi}(t;v) - B\,\tilde{\Phi}(t;v),$$

the residual of $\tilde{\Phi}(t;v)$ *with respect to* (20),(21). *This is a computable approximation for* $-\mathcal{T}(t;v)$, *where*

$$\begin{aligned} \mathcal{T}(t;v) &= (L\Phi_*)(t;v) = \partial_t \Phi_*(t;v) - \tilde{\Phi}(t;Av) - B\,\tilde{\Phi}(t;v) \\ &= (A\,\Phi_*(t;v) - \Phi_*(t;Av)) \end{aligned}$$

is the truncation error, the residual of $\Phi_*(t;v)$ *with respect to* (14),(15).

In the linear case considered here, $\Phi_*(t;v)$ and $\tilde{\Phi}(t;v)$ are linear operators, and $\tilde{\mathcal{D}}(t;v)$,$\mathcal{T}(t;v)$ can be expressed as commutators of linear operators,

$$\tilde{\mathcal{D}}(t;v) = -(A\,\tilde{\Phi}(t;v) - \tilde{\Phi}(t;Av)) = -[A,\tilde{\Phi}](t;v), \tag{24}$$

$$\mathcal{T}(t;v) = (A\,\Phi_*(t;v) - \Phi_*(t;Av)) = [A,\Phi_*](t;v). \tag{25}$$

3.1.2 Inhomogeneous flow equations. Representation of the splitting error

For the *error flow*

$$\tilde{E}(t;v) := \tilde{\Phi}(t;v) - \Phi_*(t;v),$$

with $\tilde{E}(0;v) = 0$, the identities

$$(\mathcal{L}\tilde{E})(t;v) = \tilde{\mathcal{D}}(t;v), \tag{26}$$

$$(L\tilde{E})(t;v) = -\mathcal{T}(t;v), \tag{27}$$

are valid, which take the form of inhomogeneous evolution equations of original and Sylvester type. These are our basic relations for the purpose of estimating the local error of the actual approximation,

$$\tilde{e}(h) := \tilde{u}(h) - u_*(h) = \tilde{E}(h;u_0).$$

Proceeding from (26) or (27), we now derive formulate two alternative representations for the splitting error $\tilde{e}(h)$ by means of the following variation-of-constants identities.

- Relation (26) is of the type [5]

$$(\mathcal{L}X)(t;v) = R(t;v), \quad X(0;v) = X_0(v), \tag{28}$$

with $\mathcal{L}$ from (15), a given vector field $R(t;\,\cdot\,)$, and a given initial state described by a linear vector field $X_0(\cdot)$. The solution of (28) is given by the variation-of-constants formula

$$X(t;v) = \Phi_*(t;X_0(v)) + \int_0^t \Phi_*(t-s;R(s;v))\,ds, \tag{29}$$

or equivalently,

$$X(t;v) = \mathrm{e}^{tH} X_0(v) + \int_0^t \mathrm{e}^{(t-s)H} R(s;v)\,ds. \tag{30}$$

[5] In (28) as well as (31), inhomogeneous initial states are also admitted, leading to the general variation-of-constants identities (29),(32).

- Relation (27) is of inhomogeneous Sylvester type,

$$(LX)(t;v) = R(t;v), \quad X(0;v) = X_0(v), \tag{31}$$

with L from (20), a given vector field $R(t;\cdot)$, and a given initial state described by a linear vector field $X_0(\cdot)$. With (17), the solution of (28) is given by the modified variation-of-constants formula

$$X(t;v) = \Phi_B(t; X_0(\Phi_A(t;v)) + \int_0^t \Phi_B(t-s; R(s;\Phi_A(t-s)))\,ds, \tag{32}$$

or equivalently,

$$X(t;v) = \mathrm{e}^{tB} X_0(\mathrm{e}^{tA} v) + \int_0^t \mathrm{e}^{(t-s)B} R(s; \mathrm{e}^{(t-s)A} v)\,ds. \tag{33}$$

From (24) or (25) and with (26),(30) or (27),(33), respectively, we obtain two alternative representations for the splitting error $\tilde{E}(h;v)$:

$$\tilde{E}(h;v) = -\int_0^h \mathrm{e}^{(h-s)H} \tilde{\mathcal{D}}(s;v)\,ds = \int_0^h \mathrm{e}^{(h-s)B}\, \mathcal{T}(s; \mathrm{e}^{(h-s)A} v)\,ds. \tag{34}$$

For a given initial value u_0, this gives a representation for the error (19) in the form

$$\tilde{e}(h) = \tilde{E}(h;u_0).$$

3.1.3 *Estimates for defect and splitting error*

Clearly, the defect $\tilde{\mathcal{D}}(t;\cdot)$, the truncation error $\mathcal{T}(t;\cdot)$ and the splitting error $\tilde{E}(t;\cdot)$ vanish identically if A commutes with B, i.e., if $[A,B] = AB - BA = 0$. It is also well-known (cf. e.g. [13]) that Lie-Trotter splitting is first order accurate, i.e., $\|\tilde{e}(h)\| = \mathcal{O}(h^2)$, where the error constant depends on $\|[A,B]\|$. The explicit dependence of $\tilde{e}(h)$ on $[A,B]$ is also known; however, it useful to note how this can be concluded in our context. The idea is first to derive an evolution equation for the defect [6] $\tilde{\mathcal{D}}(t;\cdot)$:

Lemma 3.1 *The defect* (23) *is the solution of the inhomogeneous Sylvester equation (see* (20),(21)*)*

$$(L\tilde{\mathcal{D}})(t;v) = -[A,B]\,\tilde{\Phi}(t;v), \quad \tilde{\mathcal{D}}(0;v) = 0. \tag{35}$$

Proof: $\tilde{\mathcal{D}}(0;v) = 0$ is obvious. Furthermore, from (23),(21) we have

$$\partial_t \tilde{\Phi}(t;v) = \tilde{\Phi}(t;Av) + B\,\tilde{\Phi}(t;v), \qquad \partial_t \tilde{\Phi}(t;Av) = \tilde{\Phi}(t;A^2 v) + B\,\tilde{\Phi}(t;Av),$$

and differentiation of $\tilde{\mathcal{D}}(t;v)$ gives

$$\begin{aligned}
\partial_t \tilde{\mathcal{D}}(t;v) &= -A\,\partial_t \tilde{\Phi}(t;Av) + \partial_t \tilde{\Phi}(t;Av) = \\
&= -A(\tilde{\Phi}(t;Av) + B\,\tilde{\Phi}(t;v)) + (\tilde{\Phi}(t;A^2 v) + B\,\tilde{\Phi}(t;Av)) \\
&= -(A\,\tilde{\Phi}(t;Av) - \tilde{\Phi}(t;A^2 v)) - (\underline{BA\,\tilde{\Phi}(t;v)} - B\,\tilde{\Phi}(t;Av)) + (\underline{BA\,\tilde{\Phi}(t;v)} - AB\,\tilde{\Phi}(t;v)) \\
&= -(A\,\tilde{\Phi}(t;Av) - \tilde{\Phi}(t;A^2 v)) - B(A\,\tilde{\Phi}(t;v)) - \tilde{\Phi}(t;Av)) - (AB - BA)\tilde{\Phi}(t;v) \\
&= \tilde{\mathcal{D}}(t;Av) + B\,\tilde{\mathcal{D}}(t;v) - [A,B]\,\tilde{\Phi}(t;v),
\end{aligned}$$

which is equivalent to (35) (see (20)). □

[6] A similar relation can be derived for the truncation error $\mathcal{T}(t;\cdot)$, see [2].

Together with (32),(33), Lemma 3.1 yields the representation

$$\begin{aligned}\tilde{\mathcal{D}}(t;v) &= -\int_0^t \Phi_B(t-s;[A,B]\,\tilde{\Phi}(s;\Phi_A(t-s;v)))\,ds\\ &= -\int_0^t \Phi_B(t-s;[A,B]\,\Phi_B(s;\underbrace{\Phi_A(s;\Phi_A(t-s;v))}_{=\Phi_A(t;v)}))\,ds\\ &= -\int_0^t \mathrm{e}^{(t-s)B}[A,B]\,e^{sB}\mathrm{e}^{tA}v\,ds.\end{aligned}$$

Consequently, the estimate

$$\|\tilde{\mathcal{D}}(t;v)\| \le t\,\|[A,B]\| = \mathcal{O}_1(t) \tag{36}$$

is valid, where the notation $\mathcal{O}_1(\,\cdot\,)$ indicates that the error constant depends on the first order commutator $\|[A,B]\|$. Together with (26) this implies the first order a priori error estimate for a step of Lie-Trotter splitting:

Proposition 3.1 *The error of the Lie-Trotter splitting applied to* (13) *satisfies*

$$\tilde{E}(t;v) = -\int_0^t \mathrm{e}^{(t-s)H}\,\tilde{\mathcal{D}}(s;v)\,ds,$$

hence

$$\|\tilde{e}(h)\| = \|\tilde{E}(h;u_0)\| \le \frac{h^2}{2}\,\|[A,B]\| = \mathcal{O}_1(h^2).$$

Proof: This is obtained via the variation-of-constants identity (30) as a direct consequence of (26) and (36). □

3.1.4 A posteriori error estimates

Relations (26),(27) suggest to approximate the error flow $\tilde{E}(t;v)$ by the solution $\tilde{\mathcal{E}}(t;v)$ of

$$(L\tilde{\mathcal{E}})(t;v) = \tilde{\mathcal{D}}(t;v), \tag{37}$$

i.e., the solution of the inhomogeneous Sylvester equation

$$\partial_t\,\tilde{\mathcal{E}}(t;v) = \tilde{\mathcal{E}}(t;Av) + B\,\tilde{\mathcal{E}}(t;v) + \tilde{\mathcal{D}}(t;v), \qquad \tilde{\mathcal{E}}(0;v) = 0. \tag{38}$$

This means that, in (27), $-\mathcal{T}(t;v)$ is approximated by $\tilde{\mathcal{D}}(t;v)$. The solution $\tilde{\mathcal{E}}(t;v)$ of (38) would result in a defect-based a posteriori estimate of the form (see (32),(33))

$$\tilde{\mathcal{E}}(h;v) = \int_0^h \Phi_B(h-s;\tilde{\mathcal{D}}(s;\Phi_A(h-s)))\,ds = \int_0^h \mathrm{e}^{(h-s)B}\tilde{\mathcal{D}}(s;\mathrm{e}^{(h-s)A}v)\,ds, \tag{39}$$

or equivalently, with representation (24) for $\tilde{\mathcal{D}}(t;\,\cdot\,)$,

$$\tilde{\mathcal{E}}(h;v) = -\int_0^h \mathrm{e}^{(h-s)B}[A,\Phi_*](h;\mathrm{e}^{(h-s)A}v)\,ds. \tag{40}$$

The integral in (40) is sort of an a posteriori approximation to (34), but, of course, it is not directly computable and is to be approximated by quadrature. Ignoring the effect of quadrature for the moment, we first consider the question whether $\tilde{\mathcal{E}}(t;v)$ is an asymptotically correct estimate for $\tilde{E}(t;v)$, i.e., whether its deviation

$$\Sigma(h;v) := \tilde{\mathcal{E}}(h;v) - \tilde{E}(h;v) \tag{41}$$

is of higher order than $\tilde{E}(h;v)$. To this end, we consider the difference of the Sylvester equations (37) and (27), resulting in

$$(L\Sigma)(t;v) = \tilde{\mathcal{D}}(t;v) + \mathcal{T}(t;v), \quad \Sigma(0;v) = 0. \tag{42}$$

From (24),(25) we see that the inhomogeneity in (42) is of the form

$$\tilde{\mathcal{D}}(t;v) + \mathcal{T}(t;v) = -[A, \tilde{\Phi} - \Phi_*](t;v) = -[A, \tilde{E}](t;v),$$

with $\|\tilde{E}(t;v)\| = \mathcal{O}_1(t^2)$ due to Proposition 3.1. This shows $\|\Sigma(h;v)\| = \mathcal{O}(h^3)$ asymptotically for $h \to 0$; but for a rigorous estimate, the commutator expression $-[A, \tilde{E}](t;v)$ has to be studied in more detail. This can be done by means of an extension of the arguments and techniques from the proof of Lemma 3.1, and from this one can conclude the desired asymptotical correctness property:

Lemma 3.2 *The deviation* (41) *satisfies*

$$\|\Sigma(h;v)\| = \mathcal{O}_2(h^3),$$

where the notation $\mathcal{O}_2(\,\cdot\,)$ *indicates that the error constant depends on commutators of* A *and* B *up to second order.*

Proof: See [2]. □

To obtain a *computable* error estimate, we now approximate the integral in (40) by the trapezoidal rule over the interval $[0,h]$ (note that $\tilde{\mathcal{D}}(0;v) = 0$):

$$\tilde{\mathcal{E}}(h;v) \approx -\frac{h}{2}\,[A, \tilde{\Phi}](h;v) = \frac{h}{2}\,\tilde{\mathcal{D}}(h;v) =: \hat{\mathcal{E}}(h;v).$$

For a particular initial value u_0 this results in the estimate

$$\tilde{e}(h) \approx \hat{\varepsilon}(h) := \hat{\mathcal{E}}(h;u_0) = -\frac{h}{2}\,[A, \tilde{\Phi}](h;u_0) = \frac{h}{2}\big(\tilde{\Phi}(h;Au_0) - A\,\tilde{u}(h)\big), \tag{43}$$

which involve on further evaluation of $\tilde{\Phi}(h;\,\cdot\,)$. To conclude that this gives an asymptotically correct a posteriori estimate for $\tilde{e}(h)$, the quadrature error $\hat{\mathcal{E}}(h;v) - \tilde{\mathcal{E}}(h;v)$ remains to be studied. Here is the outcome of this analysis:

Lemma 3.3 *The deviation involving quadrature,*

$$\hat{\Sigma}(h;v) := \hat{\mathcal{E}}(h;v) - \tilde{E}(h;v) = \Sigma(h;v) + \hat{\mathcal{E}}(h;v) - \tilde{\mathcal{E}}(h;v)$$

satisfies an analogous estimate as $\Sigma(h;v)$*,*

$$\|\hat{\Sigma}(h;v)\| = \mathcal{O}_2(h^3).$$

Proof: See [2]. □

Consequently, $\hat{\varepsilon}(h)$ from (43) provides an asymptotically estimate for $\tilde{e}(h)$ since Lemma 3.3 implies

$$\|\hat{\varepsilon}(h) - \tilde{e}(h)\| = \mathcal{O}_2(h^3).$$

3.2 Strang splitting

We will be brief. The ideas and techniques are the same as for the Lie-Trotter case but, naturally, the analysis is technically more involved, see [2].

The symmetric Strang splitting scheme (a composition of two alternate Lie-Trotter steps)

$$\tilde{u}(h) := \mathrm{e}^{\frac{h}{2}B}\,\mathrm{e}^{hA}\,\mathrm{e}^{\frac{h}{2}B}u_0 \approx \mathrm{e}^{hH}u_0 = u_*(h) \tag{44}$$

is second order accurate. With notation as in Section 3.1, we again consider the splitting flow

$$\tilde{\Phi}(t;v) := \mathrm{e}^{\frac{t}{2}B}\,\mathrm{e}^{tA}\,\mathrm{e}^{\frac{t}{2}B}v,$$

and its defect

$$\tilde{\mathcal{D}}(t;v) = (\mathcal{L}\tilde{\Phi})(t;v) = \partial_t\tilde{\Phi}(t;v) - A\,\tilde{\Phi}(t;v) - B\,\tilde{\Phi}(t;v). \tag{45}$$

A simple Taylor argument shows

$$\tilde{\mathcal{D}}(0;v) = 0, \quad \partial_t\tilde{\mathcal{D}}(0;v) = 0, \tag{46}$$

hence $\|\tilde{\mathcal{D}}(t;v)\| = \mathcal{O}(t^2)$ asymptotically for $t \to 0$. In [2], the theory developed in Section 3.1 is extended to the Strang scheme (and more general triple-splitting schemes). It is shown that, analogously as in Section 3.1.4, the error flow $\tilde{E}(t;v)$ can be approximated by the solution of a (generalized) Sylvester equation with the defect (45) on its right-hand side. Furthermore, the integral representation of its solution can be approximated by quadrature with sufficient accuracy. For efficient evaluation, property (46) suggests to use the third order Hermite quadrature formula

$$h\Big(\frac{2}{3}\,f(0) + \frac{h}{6}\,f'(0) + \frac{1}{3}\,f(h)\Big) = \int_0^h f(s)\,ds + \mathcal{O}(h^4)$$

for this purpose. Eventually, the estimator

$$\hat{\varepsilon}(h) = \hat{\mathcal{E}}(h;u_0) = \frac{h}{3}\,\tilde{\mathcal{D}}(h;u_0) \tag{47}$$

turns out to be an asymptotically correct local error estimate for one step of Strang splitting. Similarly as in (43), evaluation of $\tilde{\mathcal{D}}(h;u_0)$ is based on quantities already available plus some further 'splitting arithmetic'.

Remark 3.2 *Naturally, the Strang scheme is also a good error estimator for the Lie-Trotter scheme, but the computational effort is somewhat higher compared to (43). On the other hand, questions worth investigating are whether the Lie-Trotter scheme may serve as a cheap 'auxiliary scheme', in the spirit of Section 1.1, for estimating the error for Strang or higher order splitting methods, or whether iteration in the sense of Iterated Defect Correction [15] is a competitive approach.*

3.3 Numerical example

The following simple test problem serves as a preliminary example to illustrate the performance of our estimators.

$$\partial_t U(x,t) = \partial_{xx}U(x,t) + a(x)\,U(x,t), \quad a(x) = x(1-x)\,\mathrm{e}^x, \tag{48}$$

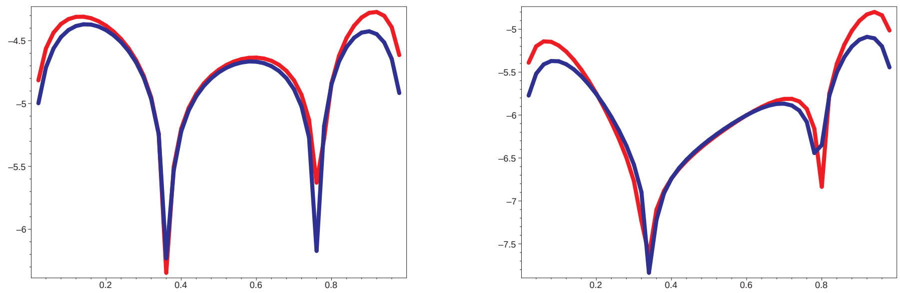

Figure 2: Error and its estimate for Example (48). Left: Lie-Trotter; right: Strang.

with $x \in [0,1]$ and subject to homogeneous Dirichlet boundary conditions. Discretizing (48) in space by standard finite differences using an equidistant spatial mesh over 50 subintervals yields a stiff initial value problem of the form

$$\partial_t u(t) = Au(t) + Bu(t), \tag{49}$$

where the diagonal matrix A represents pointwise evaluation of $a(\cdot)$, and B corresponds the finite difference approximation of $\partial_{xx}U$. We start from the initial function $U(x,0) = \frac{x(1-x)}{1+(x-\frac{1}{2})^2}$. In Figure 2 we show the errors $\tilde{\varepsilon}(h)$ (blue) and estimates $\hat{\varepsilon}(h)$ (red) after one splitting step with $h = 0.01$. The errors have been computed with respect to an accurate numerical solution $u_*(t)$ of (49), and common logarithms of absolute values are shown. The slightly irregular behavior near the boundary is to be attributed to the lack of smoothness in the exact solution $U(x,t)$. An appropriate form of spatial adaptivity is required in practice, based on the information available from the error estimate.

References

[1] W. Auzinger, H. Hofstätter, W. Kreuzer, and E. Weinmüller, Modified defect correction algorithms for ODEs. Part I: General theory. *Numer. Algorithms* **36**, 135–156 (2001).

[2] W. Auzinger, O. Koch, and M. Thalhammer, Defect-based local error estimation of splitting methods for nonlinear evolutionary Schrödinger equations. Part I. The linear case. In preparation.

[3] W. Auzinger, O. Koch, and E. Weinmüller, Efficient collocation schemes for singular boundary value problems. *Numer. Algorithms* **85**, 5–25 (2002).

[4] W. Auzinger, O. Koch, and E. Weinmüller, Analysis of a new error estimate for collocation methods applied to singular boundary value problems. *SIAM J. Numer. Anal.* **42**, 2366–2386 (2005).

[5] W. Auzinger, O. Koch, and E. Weinmüller, Efficient mesh selection for collocation methods applied to singular BVPs. *J. Comp. Appl. Math* **180/1**, 213–227 (2005).

[6] W. Auzinger, G. Kneisl, O. Koch, and E. Weinmüller, SBVP 1.0 – A MATLAB solver for singular boundary value problems. Technical Report / ANUM Preprint No.2/02, Vienna University of Technology (2002). See also:

http://www.mathworks.com/matlabcentral/fileexchange/1464-sbvp-1-0-package

[7] W. Auzinger, O. Koch, D. Praetorius, and E. Weinmüller, New a posteriori error estimates for singular boundary value problems. *Numer. Algorithms* **40**, 79–100 (2005).

[8] W. Auzinger, H. Lehner, and E. Weinmüller, Defect-based a-posteriori error estimation for differential-algebraic equations. *PAMM* **7**, 1023101–1023102 (2007).

[9] W. Auzinger, H. Lehner, and E. Weinmüller, An efficient asymptotically correct error estimator for collocation solutions to singular index-1 DAEs. To appear in *BIT Numerical Mathematics*.

[10] J.C. Butcher, J.R. Cash, G. Moore, and R.D. Russell, Defect correction for two-point boundary value problems on nonequidistant meshes. *Math. Comp.* **64/210**, 629–648 (1995).

[11] G. Kitzler, A posteriori Fehlerschätzer für Zweipunkt-Randwertprobleme mittels Defektkorrektur, Diploma Thesis, Vienna University of Technology (2010).

[12] O. Koch, Asymptotically correct error estimation for collocation methods applied to singular boundary value problems. *Numer. Math.* **101**, 143–164 (2005).

[13] C. Lubich, On splitting methods for Schrödinger-Poisson and cubic nonlinear Schrödinger equations. *Math. Comp.* **77**, 2141–2153 (2008).

[14] A. Saboor Bagherzadeh, Defect-based error estimation for higher order differential equations. PhD Thesis, Vienna University of Technology (2011). In preparation.

[15] H.J. Stetter, The defect correction principle and discretization methods. *Numer. Math.* **29**, 425–443 (1978).

[16] P.E. Zadunaisky, On the estimation of errors propagated in the numerical integration of ODEs. *Numer. Math.* **27**, 21–39 (1976).