Konrad-Zuse-Zentrum für Informationstechnik Berlin

Folkmar A. Bornemann

An Adaptive Multilevel Approach to Parabolic Equations I. General Theory & 1D–Implementation

Preprint SC 90-4 (Mai 1990)

Herausgegeben vom
Konrad-Zuse-Zentrum für Informationstechnik Berlin
Heilbronner Strasse 10
1000 Berlin 31
Verantwortlich: Dr. Klaus André
Umschlagsatz und Druck: Rabe KG Buch– und Offsetdruck Berlin

ISSN 0933-7911

Folkmar A. Bornemann

An Adaptive Multilevel Approach to Parabolic Equations I. General Theory and 1D–Implementation

Abstract

A new adaptive multilevel approach for parabolic PDE's is presented. Full adaptivity of the algorithm is realized by combining multilevel time discretization, better known as extrapolation methods, and multilevel finite element space discretization. In the theoretical part of the paper the existence of asymptotic expansions in terms of time–steps for single–step methods in Hilbert space is established. Finite element approximation then leads to perturbed expansions, whose perturbations, however, can be pushed below a necessary level by means of an adaptive grid control. The theoretical presentation is independent of space dimension. In this part I of the paper details of the algorithm and numerical examples are given for the 1D case only. The numerical results clearly show the significant perspectives opened by the new algorithmic approach.

1 Introduction

A fundamental idea for supporting the development of robust, reliable and efficient software is *adaptivity*. Whereas in the field of ordinary differential equations adaptive techniques are by now standard and much progress due to recent research has been made in the field of *stationary* partial differential equations (cf. [11] and the literature given herein), the area of adaptivity in *time–dependent* partial differential equations, as parabolic equations, is still quite open, see e.g. the survey–article about parabolic Galerkin methods by DUPONT [13].

Nearly all approaches for the numerical solution of parabolic equations *separate* the discretization of time and space both in theory and in computations. One usually develops the theory assuming one discretization (outer discretization) to be carried out first, which leads to a so–called *semi–discrete* problem. After investigating the thus arising type of problem one continues to perform the second discretization (inner discretization), ending up with a *fully discrete* scheme. As long as one uses time-independent uniform or quasi–uniform space grids and fixed time steps, the *sequence* of discretizations (first space then time or vice versa) does not matter. This kind of approach is well analyzed (THOMÉE [25] for Galerkin–methods in space). However, for highly non–uniform grids, possibly varying in time, and adaptive time steps the sequence of discretizations does indeed matter. In addition, the inner discretization can be carried out most easily using adaptive methods, whereas one may run into trouble for the outer discretization.

As an illustration consider the *method of lines.* Discretization in *space first* leads to a block system of ordinary differential equations (ODE's), which can be solved by the available variable–step, variable–order methods very efficiently, which means the *inner* problem is solved accurately and efficiently – however, ignoring the PDE context. But after all one is interested to solve the parabolic problem, so one has to consider errors introduced by the mesh, which one cannot expect to be uniformly small for all time–layers. In the 1D case BIETERMAN/BABUŠKA [5, 6, 7], who use Galerkin method in space, constructed an a–posteriori error estimator for the parabolic problem to overcome this difficulty. At certain times, fixed in advance, they decide whether they have to produce a new mesh (regridding) according to that estimator. MILLER/MILLER [21] optimize the grid in a finite element method while integrating the ODE's – *"moving finite elements"*, thus ending up with a differential–algebraic system. This approach is intimately tied with a fixed number of space nodes – at least, within each outer time step. If a dramatic change of the number of degrees of freedom is required they also have to regrid. Controlling of the complex error propagation introduced by fixing the mesh or the number of nodes over long time layers is difficult and might be nearly impossible in the nonlinear case. Regridding at fixed times may in general be "too late". Adaptivity here would call for a *second* time–step control mechanism (when to regrid) – the first being implemented in the ODE–package.

For this reason the other discretization sequence, *first time then space*, seems to be clearly preferable, and is chosen here. With that sequence it is practicable to perform a *multilevel matching* of the inner and the outer discretization, which involves solution of the inner problem up to an accuracy matched with the accuracy of the outer problem. The top levels consist in a low order single–step discretization in time with *extrapolation in Hilbert space*, which yields variable time steps and variable orders controlled by the problem up to a given accuracy. The occurring elliptic subproblems will be solved by multilevel methods, which produce the adequate individual space–meshes in order to assure an accuracy required by the time discretization.

In *Section 2* we analyze the error–term of a single–step time discretization in Hilbert space in some detail. Since the involved operators have an unbounded spectrum, the known proof techniques need an extension. By virtue of the Dunford–Taylor integral calculus the operator case can be reduced to the case of a single scalar ODE containing some parameter λ varying over the whole spectrum of the operator under consideration. This scalar case has been fully analyzed by LUBICH[18]. Our main result, Theorem 2.7, carefully traces the rôle of inconsistent and non–smooth initial data. We also give an example to show, that our estimates are sharp in a certain sense. This example shows quantitatively that in transient phases the Crank–Nicolson scheme is inferior compared to the implicit Euler.

In *Section 3* we use the just derived asymptotic expansion to establish a semi–discrete time–step control in Hilbert space. Thus the algorithm produces time–steps which really belong to the Hilbert space problem. An adaptive space discretization perturbs the semi–discrete algorithm. As we show, this perturbation can be adaptively pushed below a level indicated by the time–stepping mechanism. The thus derived fully discrete multilevel scheme treats the elliptic solver as a black–box, which has to obey several features.

In *Section 4* the black–box is specified for the 1D case. We show that a certain FEM method has the required features. Challenging numerical examples are included.

2 Single–Step Time Discretization in Hilbert Space

2.1 Preparations and Notation

In this paper *temporally homogeneous* parabolic initial–boundary value problems are studied:

$$\begin{aligned}
&\text{a)} \quad \frac{\partial u(t,x)}{\partial t} + A(x,D)u(t,x) = f(x)\,; \quad x \in \Omega\,, \quad t \in]0,T] \\
&\text{b)} \quad u(t,x) = 0\,; \quad x \in \partial\Omega\,, \quad t \in]0,T] \\
&\text{c)} \quad u(0,x) = u_0(x)\,; \quad x \in \Omega\,.
\end{aligned} \tag{2.1}$$

Here $A(x,D)$ denotes a strongly elliptic operator of second order:

$$A(x,D) = \sum_{0 \le |\rho|,|\sigma| \le 1} (-1)^{|\rho|} D^\rho (a^{\rho\sigma}(x) D^\sigma)\,. \tag{2.2}$$

It is well known, that – provided $\partial\Omega$ and the coefficients $a^{\rho\sigma}$ fulfill certain conditions of smoothness – problem (2.1) possesses a solution

$$u \in C^\infty(]0,T]\,,\ H_0^1(\Omega)) \tag{2.3}$$

continuously depending on

$$f \in H^{-1}(\Omega)\,,\ u_0 \in L^2(\Omega)\,. \tag{2.4}$$

Equation (2.1.a) holds then in the sense of distributions, (2.1.b) in the sense of the trace operator and (2.1.c) as a L^2–limit.
Since we will study the error due to discretization in *time* of problem (2.1) by a *linear* single–step method, it is enough to consider the homogeneous case $f \equiv 0$: Simply subtract the stationary solution $v \in H_0^1(\Omega)$ of $A(x,D)v = f$ and observe, that this commutes with discretization in time.
By A let us denote the following unbounded operator on $L^2(\Omega)$:

$$\begin{aligned}
A:\ D_A := H^2(\Omega) \cap H_0^1(\Omega) \subset L^2(\Omega) &\longrightarrow L^2(\Omega) \\
(Au)(x) := A(x,D)u(x) \quad &\text{for } u \in D_A\,.
\end{aligned} \tag{2.5}$$

We assume that A is positive: There is a $c > 0$, that

$$\Re(Au,u)_{L^2(\Omega)} \ge c(u,u)_{L^2(\Omega)} \quad \text{for } u \in D_A\,. \tag{2.6}$$

Since A is a closed operator, relation (2.6) together with the Lax–Milgram lemma implies that A is maximal accretive and invertible, i.e.

$$0 \in \rho(A)\,. \tag{2.7}$$

Theory of elliptic operators shows that the numerical range $\Theta(A) = \{(Au, u)_{L^2(\Omega)} | u \in D_A\}$ lies in a sector:

$$\Theta(A) \subset \Sigma_\vartheta \text{ for some } \vartheta \in [0, \pi/2[\,. \tag{2.8}$$

Here Σ_ϑ denotes

$$\Sigma_\vartheta := \{z \in \mathbb{C} | |\arg z| \leq \vartheta\} \,.$$

Thus A is m-sectorial with vertex 0 and semi-angle ϑ and $-A$ generates therefore a holomorphic semigroup of contractions.

(For definitions and proofs KATO [16] p.279 f. and 492 f.)

2.2 The Single-Step Methods

Linear single-step schemes for discretization in time define a rational approximation $r(z)$ to the exponential e^{-z} for complex z. For our purposes we restrict ourselves to $A(\vartheta)$*-stable* methods, i.e.

$$|r(z)| \leq 1 \quad \text{for } z \in \Sigma_\vartheta \,. \tag{2.9}$$

The method is called to be of *order* $p \geq 1$, if

$$|r(z)| = e^{-z} + \mathcal{O}(z^{p+1}) \quad \text{for } \Sigma_\vartheta \ni z \to 0 \,. \tag{2.10}$$

Definition 2.1

We distinguish between different types of $A(\vartheta)$-stable methods:

Type (I): $|r(\infty)| < 1$

Type (II): $r(z) = \sigma\,(1 - \gamma/z + \mathcal{O}(1/z^2))$ for $\Sigma_\vartheta \ni z \to \infty$.
Here $|\sigma| = 1$, $\gamma > 0$.

Remark 2.2

$A(\vartheta)$-stable methods of type (I) sometimes are called *strongly* $A(\vartheta)$*-stable*.

Examples of $A(\pi/2)$-stable methods of type (I) are the sub- and subsubdiagonal Padé-approximations (WANNER/HAIRER/NØRSETT [26]) like the implicit Euler scheme, examples of $A(\pi/2)$-stable methods of type (II) are the diagonal Padé-approximations like the implicit trapezoidal rule (Crank-Nicolson scheme).

For later purposes we collect some properties of these methods:

Lemma 2.3

Let $r(z)$ be the stability function of an $A(\vartheta)$–stable single–step method of order p, $0 < \vartheta \leq \pi/2$. Let $0 < \vartheta_0 < \vartheta$.

a) For $0 < \kappa < 1$ there is a $\eta(\kappa) > 0$, such that

$$|r(z)| \leq |e^{-\kappa z}| \quad \text{for } z \in \Sigma_{\vartheta_0}\,, \quad |z| \leq \eta(\kappa)\,. \tag{2.11}$$

b) For methods of type (II) there is for $0 < \kappa < \gamma$ a $\zeta(\kappa) > 0$, such that

$$|r(z)| \leq |e^{-\kappa/z}| \quad \text{for } z \in \Sigma_{\vartheta_0}\,, \quad |z| \geq \zeta(\kappa)\,. \tag{2.12}$$

c) For methods of type (I) there is for $\eta > 0$ a $\rho(\eta) < 1$, such that

$$|r(z)| \leq \rho(\eta) \quad \text{for } z \in \Sigma_{\vartheta_0}\,, \quad |z| \geq \eta\,. \tag{2.13}$$

Proof.

a)

$$\begin{aligned} \left|\frac{r(z)}{e^{-\kappa z}}\right| &= \left|e^{-(1-\kappa)z} + \mathcal{O}(z^{p+1})\right| \\ &= |1-(1-\kappa)z + \mathcal{O}(z^2)| \\ &\leq 1 \text{ for appropriate small } z \in \Sigma_{\vartheta_0} \text{ since } \vartheta_0 < \pi/2\,. \end{aligned}$$

b)

$$\begin{aligned} \left|\frac{r(z)}{e^{-\kappa/z}}\right| &= \left|e^{-(\gamma-\kappa)/z} + \mathcal{O}(1/z^2)\right| \\ &= |1-(\gamma-\kappa)/z + \mathcal{O}(1/z^2)| \\ &\leq 1 \text{ for appropriate big } z \in \Sigma_{\vartheta_0} \text{ since } \vartheta_0 < \pi/2\,, \end{aligned}$$

c) follows easily from the maximum principle for analytical functions. ∎

2.3 *Asymptotic Expansions of Single–Step Methods when Applied to $y' = -\lambda y$ with λ varying in a Sector*

The usual results for asymptotic expansions of the error of single–step methods applied to the scalar

$$y' = -\lambda y\,, \quad y(0) = 1 \tag{2.14}$$

hold for λ varying in a compact set. But since we desire to apply Dunford–Taylor integrals over pathes like $\partial\Sigma_\vartheta$ we are interested in λ to vary over the whole sector Σ_ϑ. This case is studied in the following lemma.

Lemma 2.4
(LUBICH [18], Lemma 6.3) Given an $A(\vartheta)$-stable method of order p with stability function $r(z)$ and $0 < \vartheta_0 < \vartheta \leq \pi/2$. For $z \in \Sigma_{\vartheta_0}$ an asymptotic expansion holds

$$r(z)^n = e^{-nz}\left[1 + P_p(nz)z^p + \ldots + P_N(nz)z^N\right] + R_{N+1}(n,z)\ . \tag{2.15}$$

Here the P_j are polynomials of degree $j-p+1$, $P_j(0) = 0$ and the remainder satisfies

$$|R_{N+1}(n,z)| \leq C|e^{-\kappa nz}z^{N+1}|\ , \tag{2.16}$$

for given $0 < \kappa < 1$, provided that $|z| \leq \eta(\kappa)$ from Lemma 2.3.

Proof. Let $\mathcal{O}(l,m)$ denote a function $g(z,w)$ whose Taylor expansion with respect to z has the form

$$g(z,w) = g_0(w)z^l + \ldots + g_k(w)z^{l+k} + G_{k+1}(z,w)\ , \tag{2.17}$$

where $g_i(w)$ denote polynomials in w of degree $\leq m$.
We proof the following:
Given an iteration–procedure

$$\begin{cases} y_0 &= 1 \\ y_{n+1} &= r(z)y_n + a(z,nz)e^{-nz} \end{cases} \tag{2.18}$$

with the approximation property

$$e^{-z} - r(z) - a(z,w) = \mathcal{O}(j+1, j-p)\ . \tag{2.19}$$

There is an improved iteration–procedure

$$\begin{cases} y_0^* &= 1 \\ y_{n+1}^* &= r(z)y_n^* + a^*(z,nz)e^{-nz} \end{cases} \tag{2.20}$$

with the following update of property (2.19)

$$e^{-z} - r(z) - a^*(z,w) = \mathcal{O}(j+2, j+1-p) \tag{2.21}$$

such that the difference between the two iteration–procedures is given as

$$y_n^* = y_n - e^{-nz}P(nz)z^j\ . \tag{2.22}$$

Here $P(w)$ denotes a polynomial of degree $\leq j-p+1$ with $P(0) = 0$. Furthermore we get the error–estimate

$$|y_n^* - e^{-nz}| \leq C|e^{-\kappa nz}z^{j+1}|\ . \tag{2.23}$$

Equipped with that construction we can start an induction with the single-step scheme as the iteration-procedure, i.e. $a(z,w) \equiv 0$ and $j = p$. Each improved iteration-procedure will give a further term in the asymptotic expansion; we end this at $j = N$. Estimate (2.23) gives the assertion about the remainder term. The approximation property (2.21) is needed to replace (2.19) in the induction.

To end up at this construction, combination of (2.18), (2.20) and (2.22) gives the formal expression

$$a^*(z,w) = a(z,w) + z^j \left[r(z)P(w) - e^{-z}P(z+w)\right] \qquad (2.24)$$

with an unknown polynomial $P(w)$. This will now be determinated by requirement (2.21): From (2.19) we get by interpretation of the $\mathcal{O}$-term

$$e^{-z} - r(z) - a(z,w) = z^{j+1}Q(w) + \mathcal{O}(j+2, j-p)\ , \qquad (2.25)$$

where $Q(w)$ is a polynomial of degree $j-p$. Inserting (2.24) and (2.25) in (2.21) gives

$$\begin{aligned} e^{-z} - r(z) - a^*(z,w) &= z^j[zQ(w) - r(z)P(w) + e^{-z}P(z+w)] \\ &\quad +\mathcal{O}(j+2, j-p) \\ &= z^j[z(Q(w) + P'(w)) + (e^{-z} - r(z))P(w) \\ &\quad +\mathcal{O}(2, j-p)] + \mathcal{O}(j+2, j-p) \end{aligned}$$

which is $\mathcal{O}(j+2, j-p+1)$ if we choose

$$P(w) = -\int_0^w Q(s)ds\ . \qquad (2.26)$$

It remains to prove (2.23): Using Lemma 2.3.a we get for the z under consideration

$$\begin{aligned} |y_n^* - e^{-nz}| &\leq \sum_{k=0}^{n} |r(z)|^{n-k}|e^{-(k+1)z} - r(z)e^{-kz} - a^*(z,kz)e^{-kz}| \\ &\leq \sum_{k=0}^{n} |e^{-\kappa(n-k)z}||e^{-kz}|z^{j+2}Q_k(kz)| \\ &\leq |e^{-\kappa nz}z^{j+2}| \sum_{k=0}^{n} |e^{-(1-\kappa)kz}||Q_k(kz)| \end{aligned}$$

where the $Q_k(\cdot)$ are polynomials of degree $\leq k-p+1$. Now we can estimate as follows

$$\begin{aligned} \sum_{k=0}^{n} |e^{-(1-\kappa)kz}|Q_k(kz)| &\leq C_1 \sum_{k=0}^{n} |e^{-\delta kz}| \\ &\leq \frac{C_1}{1 - e^{-\delta \Re z}} \\ &\leq C_2/\Re z \\ &\leq C_3/|z| \end{aligned}$$

for some $\delta \in]0, 1-\kappa[$.
This gives $|y_n^* - e^{-nz}| \leq C_4 |e^{-\kappa nz} z^{j+1}|$. ∎

Remark 2.5
A totally different proof of Lemma 2.4 for $r(z) = 1/(1+z)$, i.e. the implicit Euler, together with an explicit recurrence relation for the polynomials $P_j(\cdot)$ in that case may be found in [8].

2.4 Asymptotic Expansion for Abstract Cauchy Problems with m–sectorial Operators

Let A be a linear m–sectorial unbounded operator with vertex 0 in a Hilbert space H with semi–angle ϑ_0, i.e. $\Theta(A) \subset \Sigma_{\vartheta_0}$, $0 < \vartheta_0 < \pi/2$. We also assume that $0 \in \rho(A)$.
It follows that for $\alpha \geq 0$ the fractional power A^α can be defined. The corresponding domains of definition

$$H_\alpha := D_{A^\alpha} \tag{2.27}$$

equipped with the norm

$$\|u\|_\alpha := \|A^\alpha u\| \text{ for } u \in H_\alpha = D_{A^\alpha} \tag{2.28}$$

now define a scale of Hilbert spaces H_α (cf. PAZY [22], p. 195 f.). Note that $H_0 = H$ and the embedding $H_\alpha \hookrightarrow H_\beta$ is continuous for $\alpha > \beta$. This scale enables us to emphasize the rôle of inconsistent and non–smooth initial data u_0 of the *abstract Cauchy problem*

$$\begin{cases} u'(t) + Au(t) = 0, \; t \in]0, T] \\ u(0) = u_0 \in H \; . \end{cases} \tag{2.29}$$

If we denote by $\mathcal{U}(t)$ the holomorphic semigroup generated by $-A$, the solution of (2.29) is given as

$$u(t) = \mathcal{U}(t) u_0 \; . \tag{2.30}$$

A single–step method with stability function $r(z)$ generates a discrete semigroup

$$\mathcal{U}_\tau(t) := r(\tau A)^n \; , \tag{2.31}$$

where $\tau = t/n$ denotes the time–step. The result of applying the single–step method up to the time t is given by

$$u_\tau(t) := \mathcal{U}_\tau(t) u_0 \; . \tag{2.32}$$

The idea is now to apply the Dunford–Taylor integral operational calculus: It gives us the possibility to represent the term $\varphi(A)$, where $\varphi(\cdot)$ is a certain scalar function, as an integral with scalar applications of φ only:

$$\varphi(A) = \varphi(\infty)I + \frac{1}{2\pi i}\int_\Gamma \varphi(z)(zI - A)^{-1}dz\,, \tag{2.33}$$

cf. e.g. [12], Section VII.9.

The following lemma essentially estimates such integrals to get an estimate for $\|\varphi(\tau A)\|_{\mathcal{L}(H,H)}$. In addition it contains a neat trick to get $\varphi(z) - \varphi(\infty)$ into play:

Lemma 2.6

(LE ROUX [17])

Let φ be a continuous function on the sector Σ_ϑ, $0 < \vartheta_0 < \vartheta < \pi/2$, which is holomorphic in the interior of Σ_ϑ. If for some constant $R > 0$ and two continuous functions φ_1 and φ_2 from $\mathbb{R}_+$ to $\mathbb{R}_+$ the following estimates hold

$$|\varphi(z)| \le \varphi_1(|z|) \text{ for } z \in \Sigma_\vartheta\,,\quad |z| \le R \tag{2.34}$$

$$|\varphi(z) - \varphi(\infty)| \le \varphi_2(|z|) \text{ for } z \in \Sigma_\vartheta\,,\quad |z| \ge R\,, \tag{2.35}$$

we get a constant C such that

$$\begin{aligned} \|\varphi(\tau A)\|_{\mathcal{L}(H,H)} \;\le\; & \frac{C}{\vartheta - \vartheta_0}\left\{\int_0^R \varphi_1(t)\frac{dr}{r} + \int_R^\infty \varphi_2(r)\frac{dr}{r}\right. \\ & + \left.\left(R + \frac{1}{R}\right)|\varphi(\infty)|\right\} + |\varphi(\infty)| \end{aligned} \tag{2.36}$$

for all $\tau > 0$.

Proof. This lemma is essentially from LE ROUX [17], with the difference that we replaced A in (2.36) by τA. This is possible since A is involved only in two estimates in the proof from [17]:

- $\|A(I + A)^{-1}\|_{\mathcal{L}(H,H)} \le 1$
- $\|(zI - A)^{-1}\|_{\mathcal{L}(H,H)} \le \dfrac{C}{\vartheta - \vartheta_0}\dfrac{1}{|z|}$ for $z \in \Gamma$,

Γ a certain path. But we have

$$\|\tau A(I + \tau A)^{-1}\|_{\mathcal{L}(H,H)} \le 1 \tag{2.37}$$

since τA is also maximal accretive. For $z \in \Gamma$ we estimate

$$\begin{aligned} \|(zI - \tau A)^{-1}\|_{\mathcal{L}(H,H)} &= \|\frac{1}{\tau}(z/\tau I - A)^{-1}\|_{\mathcal{L}(H,H)} \\ &\leq \frac{1}{\tau}\frac{C}{\vartheta - \vartheta_0}\frac{1}{|z|/\tau} = \frac{C}{\vartheta - \vartheta_0}\frac{1}{|z|} \end{aligned} \tag{2.38}$$

since also $z/\tau \in \Gamma$. ∎

Applying this ideas to the error–term

$$\mathcal{U}_\tau(t) - \mathcal{U}(t)$$

yields the main result of this section:

Theorem 2.7
Given an $A(\vartheta)$–stable method of order p with $0 < \vartheta_0 < \vartheta < \pi/2$, there exists an asymptotic expansion

$$\mathcal{U}_\tau(t) - \mathcal{U}(t) = E_p(t)\tau^p + E_{p+1}(t)\tau^{p+1} + \ldots + E_N(t)\tau^N + E_{N+1}(t;\tau)\ . \tag{2.39}$$

For the linear coefficient operators $E_j(t)$ the following estimate holds:

$$\|E_j(t)\|_{\mathcal{L}(H_\alpha,H)} \leq C_\alpha t^{\min(1,\alpha-j)}\ ,\quad \alpha \geq 0\ . \tag{2.40}$$

The remainder operator $E_{N+1}(t;\tau)$ allows for $\tau \leq 1$ and for $\alpha > 0$ the following estimate

$$\|E_{N+1}(t;\tau)\|_{\mathcal{L}(H_\alpha,H)} \leq C_\alpha \left(t^{\min(1,\alpha-(N+1))}\tau^{N+1} + \epsilon_\alpha(t;\tau)\right)\ , \tag{2.41}$$

where the perturbation $\epsilon_\alpha(t;\tau)$ depends on the type of the method:
Methods of type (I):

$$\epsilon_\alpha(t;\tau) \equiv 0 \tag{2.42.I}$$

In this case estimate (2.41) also holds for $\alpha = 0$.
Methods of type (II):

$$\epsilon_\alpha(t;\tau) = \tau^{2\alpha}/t^\alpha\ . \tag{2.42.II}$$

The constant C_α is independent of t and τ.

Proof. We combine Lemma 2.4 and Lemma 2.6. The scalar asymptotic expansion of Lemma 2.4 suggests to set:

$$E_j(t) := P_j(tA)A^j\mathcal{U}(t) \tag{2.43}$$

and

$$E_{N+1}(t;\tau) := \mathcal{U}_\tau(t) - \mathcal{U}(t) - \sum_{j=p}^{N} E_j(t)\tau^j \,. \tag{2.44}$$

Here the $P_j(\cdot)$ denote the polynomials of Lemma 2.4, explicitly given as

$$P_j(z) =: \sum_{k=1}^{j-p+1} \pi_k^j z^k \,. \tag{2.45}$$

For $u_0 \in H_\alpha$ and $\alpha \leq j+1$ we may estimate for the coefficient operators:

$$\begin{aligned}
\|E_j(t)u_0\| &\leq \sum_{k=1}^{j-p+1} |\pi_k^j| t^k \|A^{k+j}\mathcal{U}(t)u_0\| \\
&= \sum_{k=1}^{j-p+1} |\pi_k^j| t^k \|A^{k+j-\alpha}\mathcal{U}(t)A^\alpha u_0\| \\
&\leq C_1 \sum_{k=1}^{j-p+1} |\pi_k^j| t^k \cdot t^{-(k+j-\alpha)} \|A^\alpha u_0\| \\
&\leq C_2 t^{\alpha-j} \|u_0\|_\alpha \,,
\end{aligned}$$

which is (2.40) in that case, whereas in the case $\alpha > j+1$ we get

$$\begin{aligned}
\|E_j(t)u_0\| &\leq C_{j+1} t \|u_0\|_{j+1} \\
&\leq C_\alpha t \|u_0\|_\alpha \,,
\end{aligned}$$

since $H_\alpha \hookrightarrow H_{j+1}$ is continuous. The remainder–term is more difficult to estimate and here Lemma 2.6 gets into play:

First we study the case $0 \leq \alpha < N+1$.

The scalar function under consideration is:

$$\varphi(z) := R_{N+1}(n,z)z^{-\alpha} \,. \tag{2.46}$$

One easily sees

$$\|E_{N+1}(t;\tau)\|_{\mathcal{L}(H_\alpha,H)} \leq \tau^\alpha \|\varphi(\tau A)\|_{\mathcal{L}(H,H)} \,, \tag{2.47}$$

which explains the appearance of the term $z^{-\alpha}$ in (2.46). Putting $\kappa = 1/2$ and $R := \eta(\kappa)$ from Lemma 2.4, we may set as first majorizing function

$$\varphi_1(r) := C_3 e^{-\frac{1}{2} n \cos\vartheta \cdot r} r^{N+1-\alpha} \,, \tag{2.48}$$

since Lemma 2.4 states, that

$$|\varphi(z)| \leq \varphi_1(|z|) \text{ for } z \in \Sigma_\vartheta \text{ and } |z| \leq R\,. \tag{2.49}$$

The integral occurring in the relevant estimate (2.36) is now established as:

$$\begin{aligned} I_1 &:= \int_0^R \varphi_1(r)\frac{dr}{r} \\ &\leq C_3 \int_0^\infty r^{N-\alpha} e^{-\frac{1}{2}n\cos\vartheta\cdot r}\,dr \\ &\leq C_4 n^{\alpha-N-1} \int_0^\infty e^{-\rho}\rho^{N-\alpha}\,d\rho \\ &= C_4\,\Gamma(N-\alpha+1)\frac{n^\alpha}{n^{N+1}}\,, \end{aligned} \tag{2.50}$$

with the usual gamma function Γ.

To define the second majorizing function φ_2 we have to distinguish several cases:

The case $\alpha = 0$ and methods of type (I):

Since for that case in general $\varphi(\infty) \neq 0$ we have to estimate

$$|\varphi(z) - \varphi(\infty)| \leq \sum_{j=p}^{N} |P_j(nz)z^j e^{-nz}| + |e^{-nz}| + |r^n(z) - r^n(\infty)|\,. \tag{2.51}$$

Since the exponential function damps polynomial increase we may estimate for $z \in \Sigma_\vartheta$

$$|P_j(nz)z^j e^{-nz}| \leq C_5 e^{-\frac{1}{2}n\cos\vartheta|z|}\,. \tag{2.52}$$

The second summation–term of (2.51) can be factorized as follows

$$r^n(z) - r^n(\infty) = (r(z) - r(\infty))\sum_{j=0}^{n-1} r^j(z) r^{n-1-j}(\infty)\,. \tag{2.53}$$

Since r is holomorphic in Σ_ϑ we may estimate

$$|r(z) - r(\infty)| \leq C_6/|z| \text{ for } |z| \geq R\,. \tag{2.54}$$

Lemma 2.3 c) together with (2.53) and (2.54) implies:

$$\begin{aligned} |r^n(z) - r^n(\infty)| &\leq C_7/|z|\cdot n\rho^{n-1}\,,\quad 0 \leq \rho = \rho(R) < 1 \\ &\leq C_8/|z|\cdot \rho_1^n \quad \text{for some } \rho < \rho_1 < 1\,. \end{aligned} \tag{2.55}$$

Thus we choose as second majorizing function

$$\varphi_2(r) := C_9\left(e^{-\frac{1}{2}n\cos\vartheta r} + \frac{\rho_1^n}{r}\right)\,. \tag{2.56}$$

The corresponding integral may be estimated as follows:

$$\begin{aligned} I_2 &:= \int_R^\infty \varphi_2(r)\frac{dr}{r} \\ &= C_9\left\{\int_R^\infty \frac{e^{-\frac{1}{2}n\cos\vartheta\cdot r}}{r}dr + \rho_1^n \int_R^\infty \frac{dr}{r^2}\right\} \\ &\leq C_{10}\cdot\frac{1}{n^{N+1}}\,. \end{aligned} \tag{2.57}$$

Finally we observe

$$|\varphi(\infty)| = |r^n(\infty)| \leq \rho^n \leq C_{11}\frac{1}{n^{N+1}} \tag{2.58}$$

where $\rho = \rho(R)$ is taken from Lemma 2.3 c).
Lemma 2.6 now states

$$\|\varphi(\tau A)\|_{\mathcal{L}(H,H)} \leq C_{12}\frac{1}{n^{N+1}} = C_{12}\frac{\tau^{N+1}}{t^{N+1}}\,,$$

that is (2.41) for $\alpha = 0$ and $\epsilon_0 \equiv 0$.
Case $\alpha > 0$:
Here we have $\varphi(\infty) = 0$ and therefore

$$\begin{aligned} |\varphi(z) - \varphi(\infty)| &= |\varphi(z)| \\ &\leq |z|^{-\alpha}\left\{\sum_{j=p}^{N}|P_j(nz)z^j e^{-nz}| + |e^{-nz}| + |r^n(z)|\right\}\,. \end{aligned} \tag{2.59}$$

Lemma 2.3 b) and c) together with (2.52) enforces us to choose as second majorizing function
for methods of type (I)

$$\varphi_2^{(I)} = C_{13}\left(r^{-\alpha}e^{-\frac{1}{2}n\cos\vartheta r} + r^{-\alpha}\rho^n\right)\,, \tag{2.60.I}$$

whereas for methods of type (II)

$$\varphi_2^{(II)} = C_{14}\left(r^{-\alpha}e^{-\frac{1}{2}n\cos\vartheta r} + r^{-\alpha}e^{-\frac{n}{2\cdot\cos\vartheta r}}\right)\,. \tag{2.60.II}$$

The corresponding integrals

$$I_2^{(\cdot)} = \int_R^\infty \varphi_2^{(\cdot)}\frac{dr}{r}\,, \tag{2.61}$$

are estimated as

$$I_2^{(I)} \leq C_{15}\frac{n^\alpha}{n^{N+1}} \tag{2.62.I}$$

and

$$I_2^{(II)} \leq C_{16} \left(\frac{n^\alpha}{n^{N+1}} + n^{-\alpha} \right) . \tag{2.62.II}$$

Lemma 2.6 gives

$$\tau^\alpha \|\varphi(\tau A)\|_{\mathcal{L}(H,H)} \leq C_\alpha \left(\frac{\tau^{N+1}}{t^{N+1}} t^\alpha + \epsilon_\alpha(t;\tau) \right) ,$$

i.e. (2.41).
Now let $\alpha \geq N+1$.
For those α we observe since $\tau \leq 1$:

$$\epsilon_\alpha(t;\tau) \leq \tau^\alpha \leq t^{\min(1,\alpha-(N+1))} \tau^{N+1} , \tag{2.63}$$

that means the first term in (2.41) is dominating.
For $N+1 \leq \alpha \leq N+2$ we estimate as follows

$$\begin{aligned}
\|E_{N+1}(t;\tau)\|_{\mathcal{L}(H_\alpha,H)} &\leq \|E_{N+1}(t)\|_{\mathcal{L}(H_\alpha,H)} \tau^{N+1} \\
&\quad + \|E_{N+2}(t)\|_{\mathcal{L}(H_\alpha,H)} \tau^{N+2} \\
&\quad + \|E_{N+3}(t;\tau)\|_{\mathcal{L}(H_\alpha,H)} \\
&\leq C_\alpha \left(t^{\alpha-(N+1)} \tau^{N+1} + t^{\alpha-(N+2)} \tau^{N+2} + t^{\alpha-(N+3)} \tau^{N+3} + \epsilon_\alpha(t;\tau) \right) \\
&\leq 3C_\alpha \left(t^{\alpha-(N+1)} \tau^{N+1} + \epsilon_\alpha(t;\tau) \right) ,
\end{aligned}$$

that is (2.41).
Finally for $\alpha > N+2$ we get with (2.63)

$$\begin{aligned}
\|E_{N+1}(t;\tau)\|_{\mathcal{L}(H_\alpha,H)} &\leq M_\alpha \|E_{N+1}(t;\tau)\|_{\mathcal{L}(H_{N+2},H)} \\
&\leq 4 M_\alpha C_{N+2} t \tau^{N+1} \\
&\leq 4 M_\alpha C_{N+2} \left(t^{\min(1,\alpha-(N+1))} \tau^{N+1} + \epsilon_\alpha(t;\tau) \right) ,
\end{aligned}$$

where M_α denotes the continuity–constant of the embedding $H_\alpha \hookrightarrow H_{N+2}$. ■

Remarks 2.8

1. A quite different proof of Theorem 2.7 for the implicit Euler runs along the lines of the proof of Theorem 2.13 from BORNEMANN [8], using the recurrence relation for the $P_j(\cdot)$ mentioned in Remark 2.5.

2. The estimate (2.41) has *in the case* $N = p - 1$ several forerunners:

   - For strongly $A(\vartheta)$–stable methods (methods of type (I)) and $\alpha = 0$ it is Theorem 1.2 of LE ROUX [17].
   - For the implicit trapezoidal Rule (Crank–Nicolson–scheme) and $\alpha = 1$, $\alpha = 2$ it is Lemma 3.1 of AUZINGER [1].

3. Theorems like Theorem 2.7 may be used to justify extrapolation in the method of lines. There one approximates A by an A_Δ and solves problem (2.29) for A_Δ and is interested in estimates independent of Δ. Consult AUZINGER[1] about that context.

4. Related questions are estimates for the fully–discrete Galerkin–approximation of parabolic equations. Theorem 2.7 yields in the case $N = p-1$ and $\alpha = 0$ immediately Theorem 2.1 of BAKER/BRAMBLE/THOMÉE [4], who are dealing with selfadjoint A . Cf. also chapter 7 of THOMÉE [25]. Observing that our H_α is identical with their $\dot{H}^{2\alpha}$, Theorem 2.7 yields their Theorem 2.4 and allows a sharpening of their theorem for diagonal Padé–approximations.

5. Also related is Theorem 4.1 of LUBICH [20] on certain approximations of the inverse Laplace transform based on multistep methods. We note that the resolvent map $R(\lambda; -A)$ of $-A$ is the Laplace transform of the semigroup $\mathcal{U}(t)$ generated by $-A$. This fact is reflected by the same approximation term $t^{\alpha-p}\,\tau^p$ in his Theorem 4.1 and our Theorem 2.7 with $N = p-1$, observing that we can put $\mu = 1+\alpha$ in (1.5) of LUBICH [20] since
$$\|R(\lambda; -A)\|_{\mathcal{L}(H_\alpha,H)} \leq M\frac{1}{|\lambda|^{1+\alpha}}$$
   for $\lambda \notin -\Sigma_{\vartheta_0}$.

We close this section by showing that in general the exponents of t and n in (2.41) for *small* t are sharp. This will be done for the *implicit Euler* and the *implicit trapezoidal rule* by means of an

Example 2.9
Consider the heat equations on $\Omega = [0, \pi]$ with homogeneous Dirichlet boundary conditions imposed. This means

$$\begin{aligned} A &= -\frac{d^2}{dx^2}\,, \quad D_A = H^2([0,\pi]) \cap H_0^1([0,\pi]) \\ H &= L^2(\Omega)\,. \end{aligned}$$

We use the following family of functions as initial data

$$\varphi_\vartheta = \sum_{k=1}^{\infty} \frac{1}{k^{1/2+2\vartheta}} \sin(k\cdot) \ . \tag{2.64}$$

As shown in BORNEMANN [8], p. 28, we have

$$\varphi_\vartheta \in H_\alpha \Leftrightarrow \vartheta > \alpha \ . \tag{2.65}$$

I. Implicit Euler applied to (2.29) with $u_0 = \varphi_\vartheta \ , \quad \vartheta > 0$.
Theorem 2.7 shows

$$\|E_1(t;\tau)\varphi_\vartheta\|_{L^2} \le C_\alpha t^\alpha / n \text{ for } \ 0 \le \alpha < \vartheta < 2 \ . \tag{2.66}$$

On the other hand we have Parseval's equality

$$\|E_1(t;\tau)\varphi_\vartheta\|_{L^2}^2 = \frac{\pi}{2} \sum_{k=1}^{\infty} |e_k(t;\tau)|^2 \frac{1}{k^{1+4\vartheta}} \ , \tag{2.67}$$

where

$$e_k(t;\tau) = \left(\frac{1}{1+k^2\tau}\right)^n - e^{-k^2 t} \tag{2.68}$$

denotes the scalar error–term for the eigenvalue k^2. Using the following inequality, which is proven in the appendix,

$$\left(1+\frac{x}{n}\right)^{-n} - e^{-x} \ge \frac{x^2}{2n} e^{-2x} \ \text{ for } x \ge 0 \ , \quad n \in \mathbb{N} \tag{2.69}$$

we derive from (2.67)

$$\begin{aligned} \|E_1(t;\tau)\varphi_\vartheta\|_{L^2}^2 &\ge C_1 \sum_{k=1}^{\infty} \left| \frac{k^4 t^2}{2n} e^{-2k^2 t} \right|^2 \frac{1}{k^{1+4\vartheta}} \\ &= C_2 \sum_{k=1}^{\infty} \frac{k^{7-4\vartheta} t^4}{n^2} e^{-4k^2 t} \ . \end{aligned} \tag{2.70}$$

Now choose $K \in \mathbb{N}$ that

$$\frac{1}{(K+1)^2} < t \le 1/K^2 \ ,$$

which is possible for $0 < t \le 1$.
We thus get from (2.70)

$$\begin{aligned} \|E_1(t;\tau)\varphi_\vartheta\|_{L^2}^2 &\ge C_3 \sum_{k=1}^{K} \frac{k^{7-4\vartheta}}{n^2} \frac{1}{(K+1)^8} \\ &\ge C_4 \frac{K^{8-4\vartheta}}{(K+1)^8} \frac{1}{n^2} \\ &\ge C_5 \frac{1}{K^{4\vartheta}} \frac{1}{n^2} \\ &\ge C_5 \frac{t^{2\vartheta}}{n^2} \ . \end{aligned} \tag{2.71}$$

Altogether we have

$$\|E_1(t;\tau)\varphi_\vartheta\|_{L^2} \geq C_6 \frac{t^\vartheta}{n} \quad \text{for } \; 0 < t \leq 1 \; . \tag{2.72}$$

Comparison with (2.66) shows, that the exponents of t and $1/n$ are sharp for small t.

II. Implicit trapezoidal rule applied to (2.29) with $u_0 = \varphi_\vartheta \; , \;\; \vartheta > 0$.
Theorem 2.7 shows

$$\|E_2(t;\tau)\varphi_\vartheta\|_{L^2} \leq C_\alpha t^\alpha \left(\frac{1}{n^2} + \frac{1}{n^{2\alpha}}\right) \quad \text{for } \; 0 < \alpha < \vartheta < 3 \; . \tag{2.73}$$

On the other hand we have Parseval's equality

$$\|E_2(t;\tau)\varphi_\vartheta\|_{L^2} = \frac{\pi}{2} \sum_{k=1}^{\infty} |e_k(t;\tau)|^2 \frac{1}{k^{1+4\vartheta}} \; , \tag{2.74}$$

where

$$e_k(t;\tau) = \left(\frac{1 - \frac{k^2\tau}{2}}{1 + \frac{k^2\tau}{2}}\right)^n - e^{-k^2 t} \tag{2.75}$$

denotes the scalar error–term for the eigenvalue k^2. We make use of the following two inequalities, for which proofs may be found in the appendix:

$$e^{-x} - \left(\frac{1 - \frac{x}{2n}}{1 + \frac{x}{2n}}\right)^n \geq \frac{\frac{x^3}{12n^2}}{1 + \frac{x^3}{12n^2}} e^{-x} \tag{2.76}$$

for $0 \leq x/2n < 1$ and

$$\left|\frac{(1 - \frac{x}{2n})}{(1 + \frac{x}{2n})}\right|^n \geq e^{-4/3} \quad \text{for } \; 1 \leq n \leq \frac{\sqrt{x}}{2} \; . \tag{2.77}$$

Insertion in (2.74) gives

$$\|E_2(t;\tau)\varphi_\vartheta\|_{L^2}^2 \geq C_1 \left\{ \sum_{k^2 < 2n/t} \left| \frac{\frac{k^6 t^3}{12n^2}}{(1 + \frac{k^6 t^3}{12n^2})} e^{-k^2 t} \right|^2 \frac{1}{k^{1+4\vartheta}} + \sum_{k \geq 2n/\sqrt{t}} \frac{1}{k^{1+4\vartheta}} \right\} . \tag{2.78}$$

Now we choose $K \in \mathbb{N}$ that

$$\frac{1}{(K+1)^2} < t \leq \frac{1}{K^2} \; ,$$

which is possible for $0 < t \leq 1$.

We thus get from (2.78)

$$\|E_2(t;\tau)\varphi_\vartheta\|_{L^2}^2 \geq C_2 \left\{ \sum_{k=1}^{K} \frac{k^{11-4\vartheta} t^6}{n^4} + \frac{1}{\left(\frac{2n}{\sqrt{t}}\right)^{4\vartheta}} \right\} \geq C_3 t^{2\vartheta} \left\{ \frac{1}{n^4} + \frac{1}{n^{4\vartheta}} \right\} . \tag{2.79}$$

Altogether we get

$$\|E_2(t;\tau)\varphi_\vartheta\|_{L^2} \geq C_4 t^\vartheta \left(\frac{1}{n^2} + \frac{1}{n^{2\vartheta}} \right) \text{ for } 0 < t \leq 1 . \tag{2.80}$$

Comparison with (2.73) shows, that the exponents of t and $1/n$ are sharp for small t. ■

2.5 *Appendix: Proofs of inequalities (2.69), (2.76) and (2.77)*

For sake of completeness we include proofs of these inequalities.
Inequality (2.69):
We have

$$\begin{aligned} e^t - \left(1 + \frac{t}{n}\right)^n &\geq \sum_{k=0}^{n} \left(1 - \frac{n!}{(n-k)!n^k}\right) \frac{t^k}{k!} \\ &= \frac{t^2}{2n} + \sum_{k>2}^{n} \left(1 - \frac{n!}{(n-k)!n^k}\right) \frac{t^k}{k!} \\ &\geq \frac{t^2}{2n} \quad \text{for } n \geq 2 , \ \ t \geq 0 , \end{aligned} \tag{2.81}$$

since we have $1 \geq \dfrac{n!}{(n-k)!n^k} = (1 - \dfrac{1}{n}) \ldots (1 - \dfrac{k-1}{n})$.
Inequality (2.81) trivially holds for $n = 1$.
Since $\frac{t^2}{2n} e^{-t} \leq 2e^{-2} < 1$ for $t \geq 0$, $n \geq 1$, we thus get

$$\begin{aligned} (1 + t/n)^{-n} &\geq \frac{1}{e^t - \frac{t^2}{2n}} = \frac{e^{-t}}{1 - \frac{t^2}{2n} e^{-t}} \\ &\geq e^{-t} \left(1 + \tfrac{t^2}{2n} e^{-t}\right) \end{aligned}$$

which is (2.69). ■

Inequality (2.76):

For $0 \le x/2n < 1$ we get

$$\begin{aligned}\left(\frac{1+\frac{x}{2n}}{1-\frac{x}{2n}}\right)^n &= e^{n\log\left(\frac{1+x/2n}{1-x/2n}\right)} \\ &= e^{2n\sum_{k=0}^{\infty}\frac{(x/2n)^{2k+1}}{2k+1}} \\ &\ge e^{2n\left(x/2n+\frac{1}{3}\left(\frac{x}{2n}\right)^3\right)} \\ &= e^{x+\frac{x^3}{12n^2}} \\ &\ge e^x\left(1+\frac{x^3}{12n^2}\right) .\end{aligned} \tag{2.82}$$

Thus we get for those x, n:

$$\left(\frac{1-\frac{x}{2n}}{1+\frac{x}{2n}}\right)^n \le e^{-x}\frac{1}{1+\frac{x^3}{12n^2}} = e^{-x}\left(1-\frac{\frac{x^3}{12n^2}}{1+\frac{x^3}{12n^2}}\right)$$

which is inequality (2.76). ■

Inequality (2.77):

For $4 \le 4n^2 \le x$ we have

$$x/2n \ge 4n^2/2n \ge 2 , \tag{2.83}$$

thus

$$\left|\frac{1-\frac{x}{2n}}{1+\frac{x}{2n}}\right|^n = e^{-n\log\frac{x/2n+1}{x/2n-1}} . \tag{2.84}$$

Comparison of power series shows

$$\log\frac{z+1}{z-1} \le \frac{2}{z}\frac{1}{1-\frac{1}{z^2}} \quad \text{for } z > 1 . \tag{2.85}$$

Observing (2.83), insertion of (2.85) in (2.84) gives

$$\begin{aligned}\left|\frac{1-\frac{x}{2n}}{1+\frac{x}{2n}}\right|^n &\ge e^{-\frac{4n^2}{x}\frac{1}{1-\frac{4n^2}{x^2}}} \\ &\ge e^{-\frac{4n^2}{x}4/3} \\ &\ge e^{-4/3} \quad \text{for those } n, x .\end{aligned}$$

With a different proof technique one can in fact replace $e^{-4/3}$ by the sharp value 1/3. ■

3 The Multilevel Algorithm for Parabolic Equations

3.1 Time–step Control in Hilbert–space

In this section we describe a semi-discrete algorithm for the solution of the parabolic problem (2.1):

We use the *implicit Euler* discretization *in time* and control time–step and order of the method by extrapolation following the ideas of DEUFLHARD [9] for ODE's.

The main purpose of this section will be to show, that the usual results for extrapolation–methods with some modification still hold in $L^2(\Omega)$, instead of some $\mathbb{R}^n$. Also the fully–discrete algorithm has to simulate the time–step and order control of the semi–discrete - in order to obey the requirements of the continuous problem.

The common idea of extrapolation is:

The algorithm suggests an outer time–step $T > 0$ for which

$$\mathcal{U}_{i1} := u_{\tau_i}(T) , \tag{3.1}$$

the implicit Euler discretization with time–step $\tau_i = \frac{T}{n_i}$ as introduced in chapter 2 are computed for a given sequence of increasing n_i:

$$\mathcal{F} = \{n_1, n_2, \ldots\} . \tag{3.2}$$

Since in limit $u(T) = u_{\tau=0}(T)$, we extrapolate the values $(\mathcal{U}_{11}, \ldots, \mathcal{U}_{k1})$ to $\tau = 0$, getting an approximation from which we hope, that it is better than the $\mathcal{U}_{i1}$. This will be made precise now. We compute the interpolation polynomial *with values in* $L^2(\Omega)$

$$p_{jk}(\tau) = e_0 + e_1\tau + \ldots + e_{k-1}\tau^{k-1} , \tag{3.3}$$

$e_0, \ldots, e_{k-1} \in L^2(\Omega)$, such that

$$p_{jk}(\tau_i) = \mathcal{U}_{i1} \text{ for } i = j, j-1, \ldots, j-k+1 . \tag{3.4}$$

This can be done in $L^2(\Omega)$, since the e_j are determinable as linear combinations of the $\mathcal{U}_{i1}$ as we will see later on. Now extrapolation to the limit $\tau \downarrow 0$ consists in using

$$\mathcal{U}_{jk} := p_{jk}(0) = e_0 \quad \in L^2(\Omega) . \tag{3.5}$$

The values $\mathcal{U}_{jk}$ can easily be computed in the *extrapolation table*

$$\begin{array}{lllll}
\mathcal{U}_{11} & & & & \\
\downarrow & \searrow & & & \\
\mathcal{U}_{21} & \rightarrow & \mathcal{U}_{22} & & \\
\downarrow & & \downarrow & \searrow & \\
\vdots & & & & \\
\mathcal{U}_{k1} & \rightarrow & \ldots & \mathcal{U}_{k,k-1} & \rightarrow \quad \mathcal{U}_{kk}
\end{array} \tag{3.6}$$

using the Aitken–Neville algorithm: $j \geq 2$

$$\mathcal{U}_{jk} = \mathcal{U}_{j,k-1} + \frac{\mathcal{U}_{j,k-1} - \mathcal{U}_{j-1,k-1}}{\frac{n_j}{n_{j-k+1}} - 1}\;, \quad k = 2, \ldots, j\;, \tag{3.7}$$

which can be performed in $L^2(\Omega)$.

Now we want to get an idea of the error $\|u(T) - \mathcal{U}_{jk}\|_{L^2(\Omega)}$. This is done by the following

Theorem 3.1

For $u_0 \in H_\alpha\;,\;\; \alpha \geq 0$ we have

$$\epsilon_{jk} := \|u(T) - \mathcal{U}_{jk}\|_{L^2(\Omega)} \leq \gamma_{jk} T^{\min(\alpha,k+1)}\;, \tag{3.8}$$

where asymptotically

$$\gamma_{jk} \doteq [n_{j-k+1} \ldots n_j]^{-1} C_{\text{Toep}} C\;, \tag{3.9}$$

C_{Toep} the Toeplitz–constant associated with $\mathcal{F}$ and C depends on the problem.

Proof. Follows as usual from Theorem 2.7. For instance take the proof of Theorem II. 9.1 in [15]. ∎

Remarks 3.2

1. In our example from the end of Section 2.4 we get for the inconsistent initial data
$$\varphi_{1/4}(x) = \frac{\pi - x}{2}$$
that $\|\mathcal{U}_{jk} - u(T)\|_{L^2(\Omega)} = \mathcal{O}(T^{1/4-\varepsilon})$ for arbitrary small $\varepsilon > 0$.
2. Since one implicit Euler step increases the consistency ($\varphi \in H_\alpha \Rightarrow u_\tau(\tau) \in H_{\alpha+1}$) we have, by Theorem 2.13, *after some basic time-steps*
$$\|\mathcal{U}_{jk} - u(T)\|_{L^2(\Omega)} \leq \gamma_{jk} T^{k+1}\;,$$
which means that we achieved the full and maximal order.

By these remarks and the fact that in general γ_{jk} decreases for increasing k we see, that the *assumption*

$$\epsilon_{j,k+1} \leq \rho \epsilon_{jk} \;, \quad \rho < 1 \tag{3.10}$$

is reasonable. As in DEUFLHARD [9] Section 1.2 we are thus led to the *subdiagonal error criterion*

$$\epsilon_{k+1,k} \doteq \|\mathcal{U}_{k+1,k} - \mathcal{U}_{k+1,k+1}\| =: [\epsilon_{k+1,k}]_{\mathrm{sd}} \tag{3.11}$$

as a reasonable estimator.

Convention 3.3
A single quantity in square brackets denotes a computable estimator for this quantity.

The basic time–step for achieving a prescribed tolerance TOL in line $j+1$ of the extrapolation table is now given as

$$T_{j+1,j} := \left(\frac{\mathrm{TOL}}{[\epsilon_{j+1,j}]_{sd}}\right)^{1/\min([\alpha],k+1)} T \;, \tag{3.12}$$

T the present basic time–step. The estimator $[\alpha]$ will be explained in Section 3.5.

3.2 The Fully Discrete Case: The Multilevel Concept

Now we have to approximate the elliptic problems arising by each implicit Euler step:

$$\begin{aligned} u^1 + \tau A u^1 &= u^0 + \tau f \\ u^1 &\in H_0^1(\Omega) \,. \end{aligned} \tag{3.13}$$

Since we want to use extrapolation in $L^2(\Omega)$ we are interested in global approximations with controllable error. One natural choice in view of irregular boundary geometries are finite element methods.
If we do that, we get instead of (3.6) the perturbed extrapolation table

$$\begin{array}{ccccc} \mathcal{U}_{11} + \delta_{11} & \searrow & & & \\ \downarrow & & & & \\ \vdots & & \ddots & & \\ & & & \searrow & \\ \mathcal{U}_{k1} + \delta_{k1} & \rightarrow & \dots & & \mathcal{U}_{kk} + \delta_{kk} \;, \end{array} \tag{3.14}$$

where the δ_{j1} are produced by the successive solution of the elliptic problems and the δ_{jk} with $k > 1$ are the propagated errors in the table.
Notation:

$$\bar{\mathcal{U}}_{jk} := \mathcal{U}_{jk} + \delta_{jk} \,. \tag{3.15}$$

Since the problem–oriented time–step mechanism (3.12) is connected with the semi–discrete estimator $[\epsilon_{k+1,k}]_{\text{sd}}$ we are naturally forced to achieve two things:

I. A fully discrete estimator $[\epsilon_{k+1,k}]$ with

$$[\epsilon_{k+1,k}]_{\text{sd}} \dot{\leq} [\epsilon_{k+1,k}] \,.$$

II. A control of $\delta_{k+1,k+1}$, so that

$$\bar{\mathcal{U}}_{k+1,k+1} \text{ is a tolerable approximation.}$$

This leads to the following concept: Assuming the existence of estimators $[\delta_{k+1,k}], [\delta_{k+1,k+1}], [\delta_{k+1,k} - \delta_{k+1,k+1}]$, which will be constructed in the next section, we get from the estimate

$$\begin{aligned} [\epsilon_{k+1,k}]_{\text{sd}} &\leq \|\bar{\mathcal{U}}_{k+1,k} - \bar{\mathcal{U}}_{k+1,k+1}\|_{L^2(\Omega)} \\ &+ \|\delta_{k+1,k} - \delta_{k+1,k+1}\|_{L^2(\Omega)} \end{aligned} \tag{3.16}$$

the *fully–discrete estimator*

$$[\epsilon_{k+1,k}] := \|\bar{\mathcal{U}}_{k+1,k} - \bar{\mathcal{U}}_{k+1,k+1}\|_{L^2(\Omega)} + [\delta_{k+1,k} - \delta_{k+1,k+1}] \,, \tag{3.17}$$

a completely computable quantity. It is reasonable to ask for

$$\begin{aligned} &\text{a)} \quad [\delta_{k+1,k}], [\delta_{k+1,k+1}] \leq \text{TOL}/2 \\ &\text{b)} \quad [\epsilon_{k+1,k}] \leq \text{TOL}\,. \end{aligned} \tag{3.18}$$

We also have to replace (3.12) by

$$T_{j+1,j} = \left(\frac{\text{TOL}}{[\epsilon_{j+1,j}]}\right)^{1/\min([\alpha],k+1)} T\,. \tag{3.12'}$$

3.3 *Perturbation of the Extrapolation table*

Here we construct computable $[\delta_{jk}]$. This is done in two steps.
First step: Replay to the δ_{j1}.
Since our extrapolation is linear, we get

$$\delta_{jk} = \sum_{i=j-k+1}^{j} \beta_{jk}^{i} \delta_{i1} \ , \tag{3.19}$$

where the coefficients β_{jk}^{i} only depend on the chosen subdividing sequence $\mathcal{F}$.
Thus we can define

$$[\delta_{jk}] := \sum_{i=j-k+1}^{j} |\beta_{jk}^{i}| [\delta_{i1}] \tag{3.20}$$

and analogously $[\delta_{k+1,k} - \delta_{k+1,k+1}]$. Requirement (3.18.a) can therefore be replaced by

$$[\delta_{j1}] \leq \alpha_j^k \mathrm{TOL} \tag{3.21}$$

where the coefficients α_j^k can be computed once at the beginning, only depending on $\mathcal{F}$. This coefficients can be optimized if the amount of work for the computation of $\mathcal{U}_{j1}$ is known.
Second step: Required errors of the elliptic solver.
For building the extrapolation table up to row k, we have – according to (3.21) – to compute the $\bar{\mathcal{U}}_{j1}$ with error not exceeding α_j^kTOL. This is done by solving j elliptic problems, the implicit Euler steps. The i'th produces its own error $\bar{\Delta}_i$ and the exact problem propagates the previous error Δ_{i-1} by the propagation operator π, thus leading to

$$\Delta_i = \bar{\Delta}_i + \pi \Delta_{i-1} \ . \tag{3.22}$$

The rôle of π, however, can be controlled:

Lemma 3.4
Making the general assumptions of Section 2, we have

$$\|\pi\| \leq 1 \ . \tag{3.23}$$

Proof. This follows easily from the m–accretivity of A. ∎

Assume that we use a reliable elliptic solver. It is started with the required accuracy ϵ and produces solutions together with an estimate $[\bar{\Delta}]$ of the error which was made. By using that we get from (3.22) and (3.23)

$$\|\Delta_j\|_{L^2(\Omega)} \leq j\epsilon \,. \tag{3.24}$$

The requirement (3.21) thus yields

$$\epsilon := \frac{1}{j}\alpha_j^k \mathrm{TOL} \tag{3.25}$$

as accuracy for the elliptic solver in the implicit Euler steps leading to $\mathcal{U}_{j1}$ for an extrapolation table up to row k. This is the *fundamental connection* between the time–control mechanism (extrapolation table) and the space discretization. The elliptic solver has to choose the space mesh relative to the requirement (3.25).

Finally we have

$$[\delta_{j1}] := [\bar{\Delta}_1] + \ldots + [\bar{\Delta}_j] \,. \tag{3.26}$$

The coefficients α_j^k are shown in Table 1 for $\mathcal{F} = \{1, 2, 3, \ldots\}$, the harmonic sequence, optimized with respect to the amount of work formula which will be mentioned in Section 4.

$\alpha_1^2 = 1.67_{10}{-1}$	$\alpha_2^2 = 1.67_{10}{-1}$			
$\alpha_1^3 = 1.18_{10}{-1}$	$\alpha_2^3 = 4.67_{10}{-2}$	$\alpha_3^3 = 5.65_{10}{-2}$		
$\alpha_1^4 = 1.22_{10}{-1}$	$\alpha_2^4 = 2.32_{10}{-2}$	$\alpha_3^4 = 1.35_{10}{-2}$	$\alpha_4^4 = 1.92_{10}{-2}$	
$\alpha_1^5 = 1.62_{10}{-1}$	$\alpha_2^5 = 1.61_{10}{-2}$	$\alpha_3^5 = 5.45_{10}{-3}$	$\alpha_4^5 = 4.02_{10}{-3}$	$\alpha_5^5 = 6.50_{10}{-3}$

Table 1: Coefficients α_j^k up to row $k = 5$

3.4 The order control mechanism

As in DEUFLHARD [9] we control the "order", that means here the row in the extrapolation table, in addition to the time–step. Relation (3.12') supplies us with step–size guesses $T_{j+1,j}$ for convergence of $\mathcal{U}_{j+1,j}$, that means the algorithm expects $\mathcal{U}_{j+1,j}$ to be near to the solution within the given tolerance. As in [9] we define

$$\mathcal{W}_{j+1,j} := \frac{T}{T_{j+1,j}} A_{j+1} \tag{3.27}$$

the *normalized work per unit step*, where A_{j+1} measures the amount of work required to obtain $\mathcal{U}_{j+1,j+1}$. But this will surely depend on the work required by the elliptic solver to solve its problem with accuracy ϵ given in (3.25).

But this ϵ does not only depend on j, the row of the table, but also on k, the final row, to which the table will be build up.

Thus we should replace (3.27) by

$$\mathcal{W}^k_{j+1,j} := \frac{T}{T_{j+1,j}} A^k_{j+1} \,, \tag{3.27'}$$

introducing A^k_{j+1} as the amount of work required to obtain $\mathcal{U}_{j+1,j+1}$ in a table up to $\mathcal{U}_{kk}$. These A^k_{j+1} will depend on the chosen elliptic solver. An example is given in the next chapter for the 1D case. On this basis we can actually determine an *optimal column* index q by

$$\mathcal{W}^{q+1}_{q+1,q} = \min_{j=1,\ldots,k-1} \mathcal{W}^{j+1}_{j+1,j} \,. \tag{3.28}$$

Knowing this q, we certainly use the step–size guess $T_{q+1,q}$ for the next basic time–step and expect convergence in the vicinity of q.
In order to get a reliable code, avoiding pseudo–convergence and related undesirable things, which occur in practice, one has to implement three devices

- convergence monitor
- order window
- a device for possible increase of order greater than q ,

see for instance DEUFLHARD [10].
This can be achieved by comparing the actual behavior in the table with an information–theoretic *standard model* derived in [9]. Here the measure of input–information is the number of Hilbert–space problems which have to be approximated. Thus we get n_j as the quantity of information contained in $\mathcal{U}_{j1}$. Because of the dependency of the final row, we have to change the *order window of* [9]: The error criterion (3.18.b) and the monitoring conditions of [9] are only tested for j in the range of

$$q - 1 \leq j \leq q \,.$$

In all other details the step–size and order control of [9] can be taken without change.

3.5 *The consistency–estimator*

The last missing point for a complete description of the algorithm *without specifying the elliptic solver* remains to be an estimator for α, the consistency of the last approximation $\tilde{u}(t)$:

$$\tilde{u}(t) \in H_\alpha\,, \qquad \alpha \text{ maximal}\,.$$

We start assuming as much consistency we need, that means

$$[\alpha]_{\text{start}} = k_{\max} + 1\,. \tag{3.29}$$

If the estimated $[\alpha]$ is seriously *too large* we will get far too large time–step guesses by (3.12') and therefore a step–size reduction with redoing of the step. Now take the largest possible k, for which with respect to the old time–step T_{old} as well as to the new time–step T_{new} error estimates $[\epsilon_{k+1,k}]_{\text{old}}$ respectively $[\epsilon_{k+1,k}]_{\text{new}}$ are available. By Theorem 3.1 we have

$$\begin{aligned} &\text{a)}\quad [\epsilon_{k+1,k}]_{\text{old}} \doteq C T_{\text{old}}^{\min(\alpha,k+1)} \\ &\text{b)}\quad [\epsilon_{k+1,k}]_{\text{new}} \doteq C T_{\text{new}}^{\min(\alpha,k+1)}\,, \end{aligned} \tag{3.30}$$

that means

$$\min(\alpha, k+1) \approx \frac{\log\left(\frac{[\epsilon_{k+1,k}]_{\text{old}}}{[\epsilon_{k+1,k}]_{\text{new}}}\right)}{\log\left(\frac{T_{\text{old}}}{T_{\text{new}}}\right)} \tag{3.31}$$

leading to the reasonable

$$[\alpha]_{\text{new}} := \min\left([\alpha]_{\text{old}}\,,\ \frac{\log\left(\frac{[\epsilon_{k+1,k}]_{\text{old}}}{[\epsilon_{k+1,k}]_{\text{new}}}\right)}{\log\left(\frac{T_{\text{old}}}{T_{\text{new}}}\right)}\right) \tag{3.32}$$

The log–quotient will be in reasonable behaving cases positive, because of $T_{\text{new}} < T_{\text{old}}$. If not, we do best by trying

$$[\alpha]_{\text{new}} := [\alpha]_{\text{old}}/2\,. \tag{3.33}$$

If we have no step–size reduction and redoing of a step, we have to consider an increase of α since each implicit Euler step increases α by one:

$$[\alpha]_{\text{old}} \longrightarrow \min(k_{\max} + 1\,,\ [\alpha]_{\text{old}} + 1) =: [\alpha]_{\text{new}}\,. \tag{3.34}$$

4 *Algorithmic Details in the 1D Case and Numerical Examples*

In the last section we treated the elliptic solver mainly as a *black box*. In fact we required only two things

1. The elliptic solver is started with a required accuracy ϵ, and gives global solutions together with an error estimate $[\bar{\Delta}]$, which is necessary to realize estimate (3.24).

2. The amount of work A_{j+1}^k as occurring in (3.27') should be computable.

Another feature should also be required:
In order to realize the first requirement, it is reasonable to use an adaptive FEM–method. This will mainly contain the following three modules:

- error–estimator
- linear solver
- refinement–strategy

Since we are dealing with an one–parameter family of elliptic problems

$$u + \tau A u = f \tag{4.1}$$

we have to require:

3. The performance of the error–estimator and linear solver should be independent of τ, especially should work in the vicinity of $\tau = 0$.

4.1 *Time–Step Independent Elliptic Error Estimation and Amount of Work Principle*

Here we restrict ourselves to selfadjoint elliptic operators in one space dimension.
The *difficulty* for constructing τ–independent error estimators lies in the fact, that we get a break–down of $H_0^1(\Omega)$–ellipticity as $\tau \downarrow 0$ for our bilinear form

$$B_\tau(u, v) := (u, v)_{L^2} + \tau a(u, v) \tag{4.2}$$

associated with the elliptic problem (4.1). In this case we get a transition

$$\text{Ritz–projection} \longrightarrow L^2\text{–projection.}$$

So we are led to localize the problem in order to construct an error estimator, i.e. we solve on the subintervals of the mesh the same elliptic problem with imposing the actual FEM–approximation as Dirichlet boundary condition. This should give a reasonable local error estimator. Surely the local problems will not be solved exactly, but it is enough to solve them with higher accuracy using quadratic elements.

The author used in [8] norms, which are extensions of those introduced by BABUŠKA/OSBORNE [2] for the purely elliptic case, in order to get results in the light of the third requirement of the introduction to this section. An example is the following norm, which will occur in the main result about the error estimator.

Definition 4.1

Let $\tau > 0$. We take a subdivision Δ (mesh) of $I := [a, b]$:

$$\begin{aligned} \text{a)} \quad \Delta &:= \{0 = x_0 < x_1 < \ldots < x_n = b\} \\ \text{b)} \quad h_j &:= x_j - x_{j-1}\,;\ I_j :=]x_{j-1}, x_j[\,,\ j = 1, \ldots, n \\ \text{c)} \quad \delta_j &:= (h_j + h_{j+1})/2\,;\ j = 1, \ldots, n-1\,. \end{aligned} \tag{4.3}$$

For $u \in H_0^1(I)$ let

$$\|u\|_{0,\Delta}^2 := \|u\|_0^2 + \sum_{j=1}^{n-1} \delta_j |u(x_j)|^2 \tag{4.4}$$

and define H_Δ^0 to be the completion of $H_0^1(I)$ with respect to this norm.

In the norm $\|\cdot\|_{0,\Delta}$ something like a *discrete L^2–norm* on the mesh Δ is coupled to the $L^2(I)$–norm.

We note that on the family S_Δ the norms $\|\cdot\|_0$ and $\|\cdot\|_{0,\Delta}$ are uniformly equivalent in the sense

$$k_1\|u\|_0 \leq \|u\|_{0,\Delta} \leq k_2\|u\|_0 \text{ for all } u \in S_\Delta\,, \tag{4.5}$$

k_1, k_2 positive constants independent of Δ. This is essentially relation (4.3.c) of [2], but can be shown in our case, linear elements, by direct computation.

Some conclusions which can be drawn from results of [8] are:

1. We have *quasi–optimality* of the FEM–approximation u_Δ with respect to the $\|\cdot\|_{0,\Delta}$ norm *independent of τ*:

$$\|u - u_\Delta\|_{0,\Delta} \leq C \inf_{\varphi \in S_\Delta} \|u - \varphi\|_{0,\Delta}\,, \tag{4.6}$$

   C independent of Δ and τ. Note that for $\tau > 0$ and $f \in L^2$ we have $u \in H_0^1(I) \subset H_\Delta^0$.

2. *Adequate adaptive meshes* can be characterized as follows:

$$\inf_{\varphi \in S_\Delta} \|u - \varphi\|_{0,\Delta} \leq \frac{C}{n^2}\,, \tag{4.7}$$

C fairly independent of u, n the number of degrees of freedom. Note that with I_Δ the interpolation operator and a quasi–uniform mesh Δ we get by (4.6)

$$\begin{aligned} \|u - u_\Delta\|_{0,\Delta} &\leq C\|u - I_\Delta u\|_{0,\Delta} \\ &= C\|u - I_\Delta u\|_0 \\ &\leq \tilde{C}h^2\|u\|_2\,. \end{aligned}$$

3. On those adequate meshes we get

$$\|u - u_\Delta\|_{0,\Delta} \leq \frac{C}{n^2}\,, \tag{4.8}$$

C fairly independent of Δ and τ.

This justifies our

Basic amount of work principle 4.2
Adaptive solution of an elliptic problem from family (4.1) with accuracy ϵ needs

$$n = C/\sqrt{\epsilon} \tag{4.9}$$

degrees of freedom, C *fairly independent* of τ and Δ.

Next we describe the *error estimator*: For $j = 1, \ldots, n$ consider the local elliptic problems

$$\begin{aligned} &\text{a)}\quad w_j + \tau A w_j = f \quad \text{on } I_j \\ &\text{b)}\quad w_j(x_{j-1}) = u_\Delta(x_{j-1})\,,\ \ w_j(x_j) = u_\Delta(x_j)\,. \end{aligned} \tag{4.10}$$

Relation(4.10.b) means that

$$\bar{w}_j := w_j - u_\Delta \in H_0^1(I_j) \tag{4.11}$$

and a weak formulation of (4.10.a) is therefore with

$$e := u - u_\Delta \tag{4.12}$$

the equation

$$B_\tau(\bar{w}_j, v) = B_\tau(e, v) \text{ for every } v \in H_0^1(I_j)\,. \tag{4.13}$$

Let $S_j^Q \subset H_0^1(I_j)$ be the space of quadratic finite elements on the grid $\{x_{j-1}, (x_{j-1}+x_j)/2, x_j\}$, and $\tilde{w}_j \in S_j^Q$ the FEM–approximation of $\bar{w}_j$ on I_j, that means

$$B_\tau(\tilde{w}_j, \varphi) = B_\tau(e, \varphi) = (f, \varphi) - B_\tau(u_\Delta, \varphi) \text{ for all } \varphi \in S_j^Q\,. \tag{4.14}$$

These $\tilde{w}_j$ are *computable.*
Our *computable local error estimator* is now

$$[\eta_j] := \|\tilde{w}_j\|_{0,I_j}\,,\quad j = 1,\ldots,n \tag{4.15}$$

and the global one

$$[\eta] := \left(\sum_{j=1}^{n}[\eta_j]^2\right)^{1/2}\,. \tag{4.16}$$

For relations between $[\eta]$ and $\eta := \|e\|_{0,\Delta}$ we have to introduce a *local* (semi–)norm: $j = 1,\ldots,n$.
For $u \in H_\Delta^0$ set

$$\|u\|_{0,\Delta,I_j}^2 := \|u\|_{0,I_j}^2 + \tfrac{1}{2}h_j(u(x_{j-1})^2 + u(x_j)^2)\,. \tag{4.17}$$

The main result of [8] about the elliptic error estimator is:

Theorem 4.3
The following local and global estimates hold

$$\begin{array}{llllll} \text{a)} & [\eta_j] & \le & K\eta_j & := & K\|u - u_\Delta\|_{0,\Delta,I_j} \\ \text{b)} & [\eta] & \le & K\eta & := & K\|u - u_\Delta\|_{0,\Delta}\,. \end{array} \tag{4.18}$$

Here K denotes a positive constant independent of Δ and τ.

4.2 The Refinement Strategy and the Linear Solver

- *The refinement strategy.*
  Since we are equipped with local error indicators η_j, we are able to build a refinement strategy:

  $$\text{refine } I_j \text{ if } \eta_j > \text{ cut }\,.$$

  The heuristic (4.8) asks for a nearly equidistributed error. In order to achieve that, we determine "cut" following BABUŠKA/RHEINBOLDT

[3]: We use a simple heuristic prediction scheme to forecast what may happen to η_j if I_j is subdivided. *Locally* we may assume

$$\eta_j = c_j h_j^{\lambda_j} \text{ as } h_j \to 0 . \tag{4.19}$$

Suppose I_j was generated by subdividing I_j^{old} with local error η_j^{old} obeying (4.19). The η_j–value after dividing I_j will be thus approximately

$$\eta_j^{\text{new}} = \frac{\eta_j^2}{\eta_j^{\text{old}}} . \tag{4.20}$$

Clearly now, we should refine only those elements I_j which have an η_j–value above the *largest* predicted *new* η–value in the next mesh:

$$\text{cut} := \max_j \eta_j^{\text{new}}$$

- *The linear solver.*
  Since we treat the 1D case here, the stiffness–matrix M is tridiagonal. So linear equations can be solved by direct Gauss-elimination without pivoting in $O(n)$ simple operations.

  However, the global stiffness–matrix M needs not to be assembled: It is enough to know the local stiffness–matrices M^j associated to I_j using a fronting algorithm.

4.3 Realization of Extrapolation

The elliptic–solver produces a first column of the extrapolation table as follows:

$$\begin{gathered} \mathcal{U}_{11} \in S_{\Delta_1} \\ \vdots \\ \mathcal{U}_{k1} \in S_{\Delta_k} . \end{gathered}$$

In order to extrapolate we consider the *common mesh*

$$\Delta = \bigcup_{j=1}^{k} \Delta_j . \tag{4.21}$$

Surely we have $\mathcal{U}_{j1} \in S_\Delta$, which in practice is done by linear interpolation between the nodes since $\mathcal{U}_{j1}$ is linear there. Now we can do the extrapolation in the coefficient vector of the nodal basis for Δ.

In 2D case (4.21) does not work, because the such defined Δ will in general not be a triangulation. So we have to *require* that there exists a triangulation Δ, that

$$\mathcal{U}_{j1} \in S_\Delta , \ j = 1, \ldots, k .$$

This is a requirement on the Δ_j. In this case we will call the Δ_j *compatible.*
In practice this can be achieved as follows:
First we have the triangulation Δ_1 for $\mathcal{U}_{11}$. We set $\Delta^1 := \Delta_1$.
Given Δ^j, $j = 1, \ldots, k-1$ we construct Δ_{j+1} and Δ^{j+1} as follows:
The necessary refinement for Δ_{j+1} is done using the tree for Δ^j, possible *extending* that tree. This extended tree will be Δ^{j+1}, so that Δ^j and Δ_{j+1} are subtrees of Δ^{j+1}. That means, we have

$$S_{\Delta_{j+1}} \subset S_{\Delta^{j+1}}$$

and

$$S_{\Delta^j} \subset S_{\Delta^{j+1}}$$

In the same step we compute the coefficients for $\mathcal{U}_{11}, \ldots, \mathcal{U}_{j+1,1}$ in the nodal basis of $S_{\Delta^{j+1}}$ by linear interpolation of the nodal basis representation of S_{Δ^j}.
As the above Δ we get Δ^k, for which by construction

$$\mathcal{U}_{j1} \in S_{\Delta^k}\ , \quad j = 1, \ldots, k\ .$$

Also we have by construction the $\mathcal{U}_{j1}$ in the nodal basis representation of S_{Δ^k}. In this basis the extrapolation will be performed.
This efficient method should also be used in the 1D case. Numerical experience shows that the method is also preferable for stability reasons.

4.4 *Numerical Examples*

The program KASTIX (*KASkade TIme*–dependent with eXtrapolation) is a realization of the algorithm described in this paper. The 1D elliptic solver is written in analogy to the adaptive multilevel 2D elliptic solver KASKADE [11, 23, 24]. It is written in the language C [8].

All numerical experiments of this paper were made using double precision arithmetic on a SPARC–station1+.

One should note, that there are *no* parameters which had to be fitted to the examples.

The notation used to describe the experimental results has been introduced earlier in this paper or is self–evident, with the exceptions of

$$\text{CPU} = \text{computing time in seconds on a SPARC–station1+},$$

$$[N]_{work} = \left(\sum_{j\geq 1}(\text{no. of nodal–points of time–step j})/\tau_j\right) \Big/ \left(\sum_{j\geq 1} 1/\tau_j\right) ,$$

$$[N]_{mem} = \left(\sum_{j\geq 1}\text{no. of nodal–points of time–step j}\right) /\text{no. of time–steps} ,$$

$$N_{max} = \text{maximum number of nodal–points for a time–layer} .$$

The mean–values $[N]_{work}$ and $[N]_{mem}$ are chosen to equidistribute the number of nodal–points with respect to the same effort of *work* resp. *memory*. These are sensible choices with respect to variable order which has the goal of minimized work. The choice of a time–mean like in [7] would over–weight points of higher–order time–layers.

Example 1. Rise of a traveling wave from a trivial solution. This model problem was created in order to examine how KASTIX handles with sudden events. It consists in a traveling wave which suddenly travels through the space–interval.

The solution of the problem is

$$u(t,x) = -0.5\tanh(40(x - 10t + 6))\,.$$

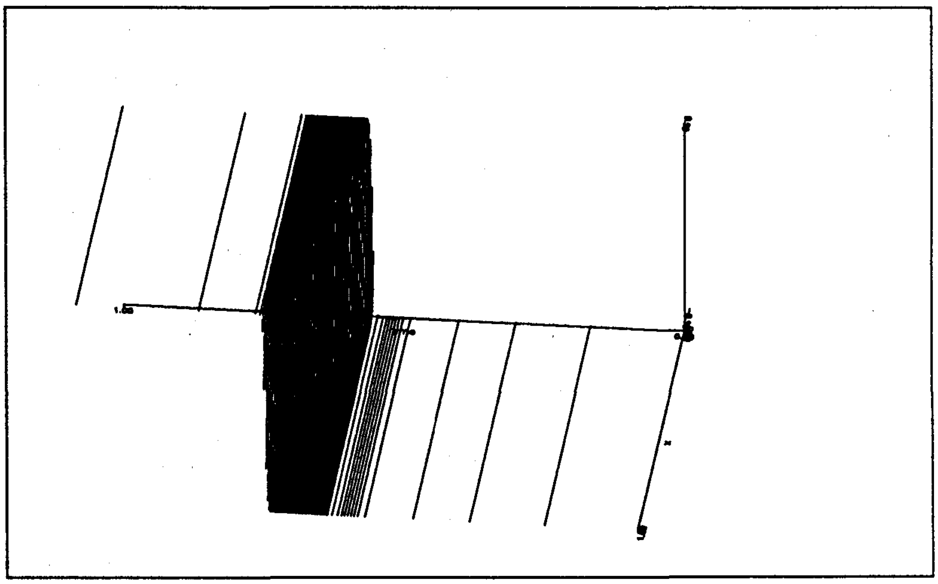

Figure 1: Traveling wave solution (Example 1).

At time $t = 0.59$ the wave enters the interval $I = [0,1]$, travels with speed $v = 10$ through it and leaves it at time $t = 0.71$. The PDE actually solved numerically is the scalar heat equation with a convection term

$$u_t = u_{xx} - 10u_x + f\ ;\ t > 0, x \in [0,1]\,,$$

where f, the initial data at $t = 0$, and the Dirichlet boundary conditions were set that $u(t,x)$ is the solution.

Figure 1 shows the computed solution at the time–layers chosen by the algorithm. The problem was run until time $t = 1$.

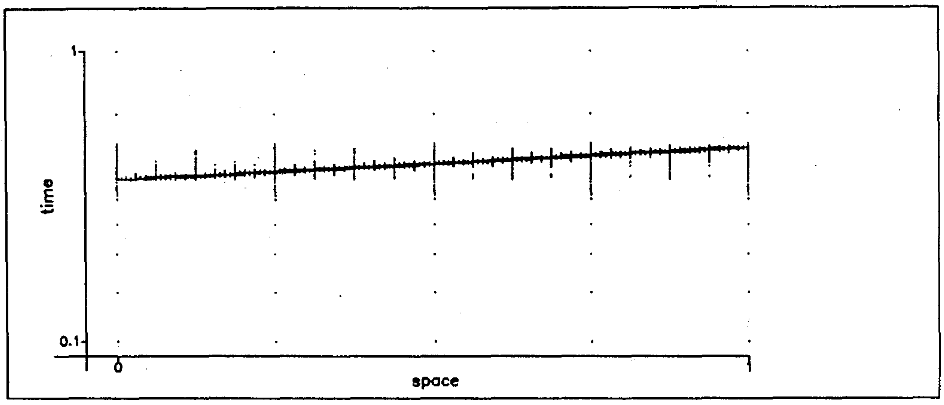

Figure 2: The mesh–development (Example 1).

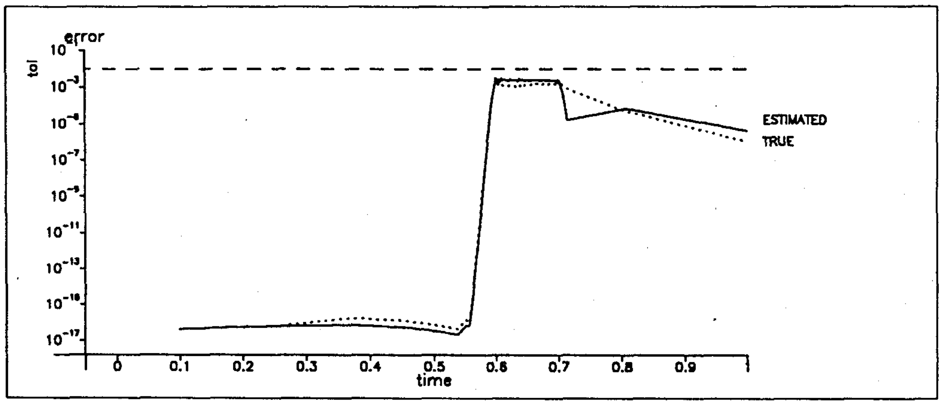

Figure 3: Behavior of the error–estimator (Example 1).

TOL	time–steps	max. order	$[N]_{work}$	$[N]_{mem}$	N_{max}	$L^\infty([0,T], L^2(I))$ norm of true–error	CPU
* 10^{-2}	78	2	35	30	49	$2.73_{10} - 3$	11
10^{-3}	94	3	108	109	315	$2.33_{10} - 4$	71
10^{-4}	134	4	389	430	2335	$4.05_{10} - 5$	720

* run represented in Figs. 1–3

Table 1: KASTIX: performance for variable order (Example 1)

For higher orders we get $[N]_{work} < [N]_{mem}$, which nicely reflects our choice of order to optimize the amount of work.

Example 2. Two counter–traveling waves. This model problem has been proposed by BIETERMAN/BABUŠKA [7] to examine the behavior of the adaptive mesh generation when two waveforms, having different front widths and directions and speed of travel, collide and pass through each other. The solution of this problem is

$$u = u^{(1)} + u^{(2)} ,$$

where

$$\begin{aligned} u^{(1)}(t,x) &= 0.25(1 + \tanh(100(x - 10t))) , \\ u^{(2)}(t,x) &= 0.25(1 + \tanh(80(1 - x - 30t))) ; \end{aligned}$$

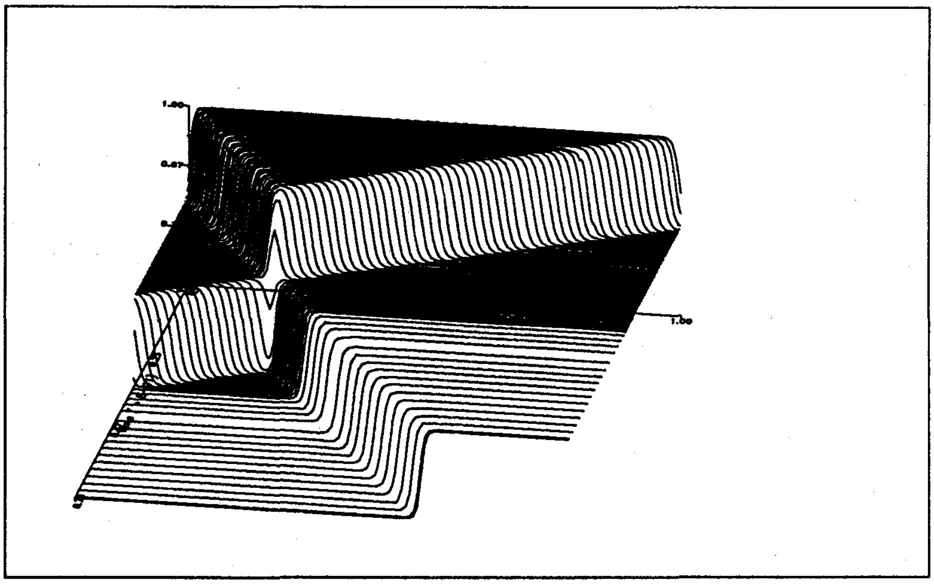

Figure 4: The two counter–traveling waves (Example 2).

$u^{(1)}$ moves towards $x = 1$ at speed $v = 10$ and $u^{(2)}$, whose front width is 25% larger than that of $u^{(1)}$, moves towards $x = 0$ at speed $v = 30$. At time $t = 0.025$ the waves collide, almost extinguishing each other. The PDE actually solved numerically was the scalar heat equation

$$u_t = u_{xx} + f \; ; \; t > 0, x \in [0,1] ,$$

where f, the initial data at $t = 0$, and the Dirichlet boundary conditions were set that $u(t,x)$ is the solution.

Figure 4 shows the computed solution at the time-layers chosen by the algorithm until time $t = 0.07$.

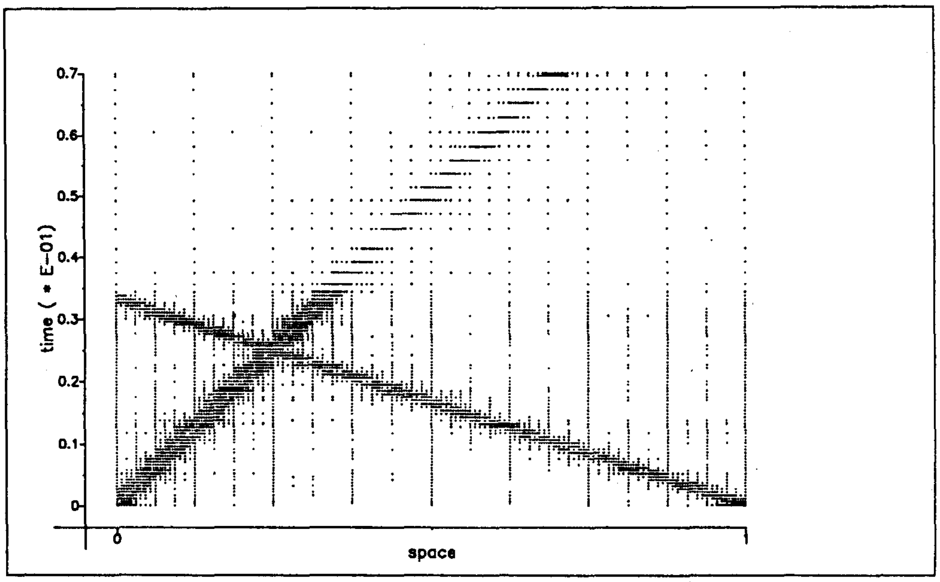

Figure 5: Mesh-development for the two waves (Example 2).

Figure 5 clearly shows that the time-step is chosen automatically with respect to that traveling wave which possesses the larger signal-velocity.

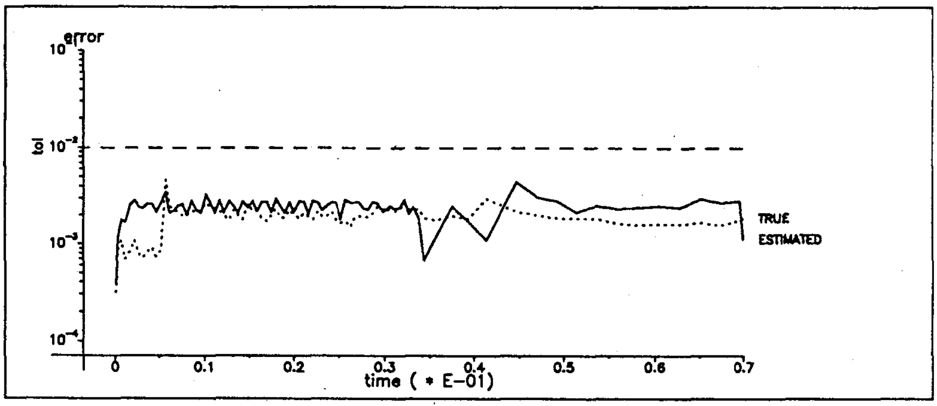

Figure 6: Behavior of the error-estimator (Example 2).

TOL	time–steps	max. order	$[N]_{work}$	$[N]_{mem}$	N_{max}	$L^\infty([0,T],L^2(I))$ norm of true–error	CPU
* 10^{-2}	86	2	53	50	81	$4.65_{10}-3$	32
10^{-3}	98	3	163	203	507	$3.94_{10}-4$	247
$5\cdot 10^{-4}$	130	3	218	261	731	$1.63_{10}-4$	426
$2\cdot 10^{-4}$	175	3	408	488	1256	$6.93_{10}-5$	1169
10^{-4}	87	4	949	1244	3201	$5.14_{10}-5$	2358

* run represented in Figs. 4–6

Table 2: KASTIX: performance for variable order (Example 2)

Once more we get for higher orders $[N]_{work} < [N]_{mem}$, which nicely reflects our choice of order to optimize the amount of work. We also observe that the number of time–steps is nearly constant and drops when the next order in the time discretization is activated, which is a known advantage of extrapolation–methods for ODE's. This occurs here since we have no stationary–phases as in the other examples.

Finally we study the effect which results if we fix the order of the time discretization at order 2, in which case we still can control time–steps adaptively.

TOL	time–steps	$[N]_{work}$	$[N]_{mem}$	N_{max}	$L^\infty([0,T],L^2(I))$ norm of true–error	CPU
10^{-2}	92	54	51	84	$4.64_{10}-3$	36
10^{-3}	414	138	128	247	$8.42_{10}-4$	418
$5\cdot 10^{-4}$	632	207	188	407	$7.14_{10}-4$	953
$2\cdot 10^{-4}$	1114	298	271	712	$3.50_{10}-4$	2490
10^{-4}	> 1084	—	fail	—	$> 2.14_{10}-4$	> 4104

Table 3: KASTIX: performance for fixed maximum order 2 (Example 2)

We observe a drastic increase of time–steps, which leads to more CPU–time and, which is more serious, to more need of storage for the solution–data. We also see that $[N]_{work} > [N]_{mem}$, which had to be expected. Moreover the large number of time–steps yields an error–propagation, which the algorithm is not able to trace. Thus we get global errors exceeding the tolerance TOL, if TOL $< 10^{-3}$. For TOL $= 10^{-4}$ the propagation gets too much influence, so the algorithm even fails. Thus beside other advantages the variable order case appears to be more *reliable.*

TOL	variable order no. of points	fixed order no. of points	factor of save
10^{-2}	4300	4692	1.09
10^{-3}	19894	52992	2.66
$5\cdot 10^{-4}$	33930	118816	3.50
$2\cdot 10^{-4}$	85400	301894	3.54

Table 4: Array storage comparison (Example 2)

Example 3. Point–source. This model problem, which is essentially a 1D version of Example 1 in ERIKSSON/JOHNSON [14], has been proposed by the author in [8] to test the time–stepping procedure. We solve the homogeneous heat equation with the following approximate δ–function as initial data:

$$u_0(x) = 250 \exp(-250x^2) \; .$$

The Dirichlet boundary conditions are chosen to model on $I = [0,2]$ the evolution of u_0 on the whole real axis.

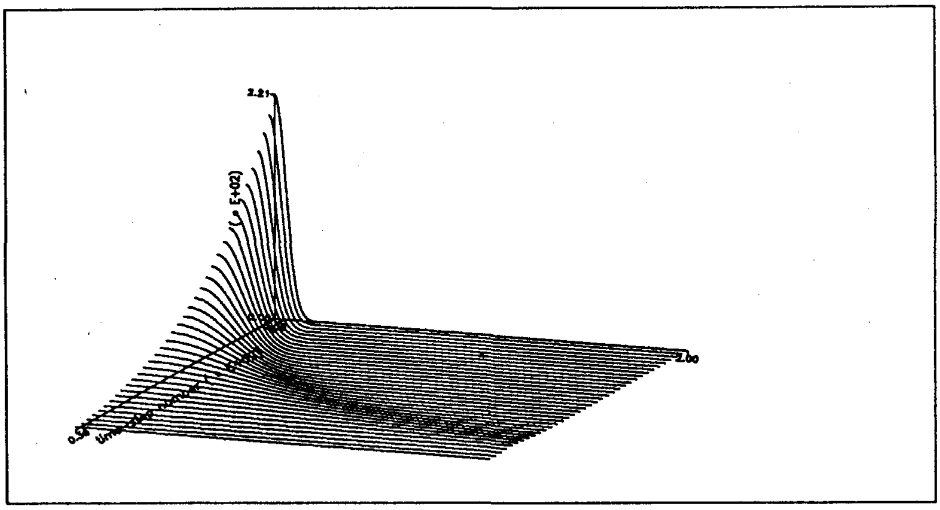

Figure 7: Evolution of point–source, time in log–scale (Example 3).

Because of the exponentially decay of the solution as shown in Figure 7 one expects an increase of the time–step according to a power–law, which really occurs automatically in the performance as shown in Figure 8.

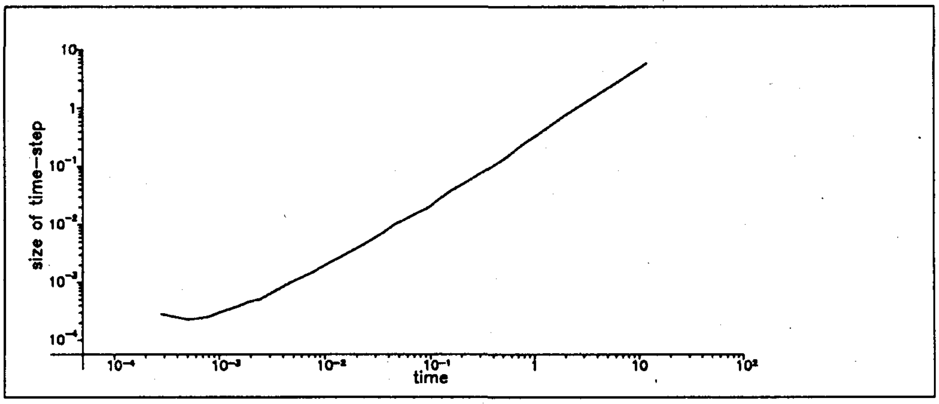

Figure 8: Automatic increase of the time–step (Example 3).

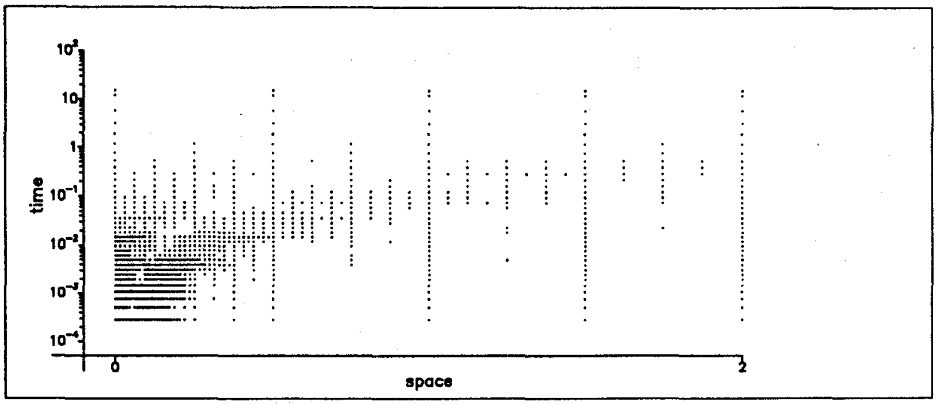

Figure 9: Mesh–development for the point–source (Example 3).

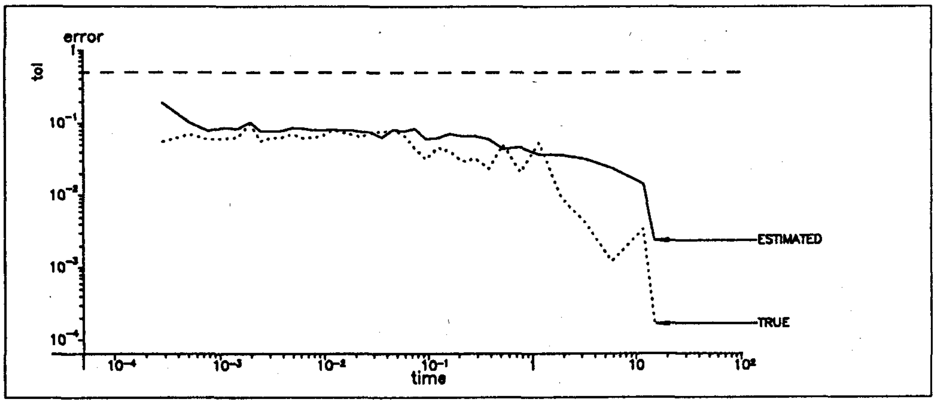

Figure 10: Behavior of the error–estimator (Example 3).

TOL	time–steps	max. order	$[N]_{work}$	$[N]_{mem}$	N_{max}	$L^\infty([0,T],L^2(I))$ norm of true–error	CPU
* $5\cdot10^{-1}$	36	2	59	32	135	$9.82_{10}-2$	3
$5\cdot10^{-2}$	111	2	141	96	804	$2.08_{10}-2$	28
$5\cdot10^{-3}$	186	3	888	311	2601	$1.27_{10}-3$	236

* run represented by Figs. 7–10

Table 5: KASTIX: performance for variable order (Example 3)

In this case the solution runs into the stationary zero–solution. Thus the algorithm chooses very early the low order 2 in time. This causes $[N]_{work} > [N]_{mem}$ here.

Acknowledgements.

My sincere thanks go to C. Lubich for long discussions about asymptotic expansions and his kind allowance to publish his elsewhere nearly unpublished Lemma 6.3 [18], which is Lemma 2.4 in the present paper. I also want to thank P. Deuflhard and M. Wulkow for many helpful discussions. Last but not least my special thanks to Mrs. S. Wacker for her excellent and patient typing of the manuscript.

References

•[1] Auzinger, W.: *On Error Structures and Extrapolation for Stiff Systems with Application in the Method of Lines.* Institut für Angewandte und Numerische Mathematik, Technische Universität Wien, Report Nr. 81/89 (1989).

[2] Babuška, I., Osborne, J.: *Analysis of Finite Element Methods for Second Order Boundary Value Problems using Mesh Dependent Norms.* Num. Math. 34, p. 41–62 (1980).

[3] Babuška, I., Rheinboldt, W.C.: *Error Estimates for Adaptive Finite Element Computations.* SIAM J. Numer. Anal. 15, p. 736–754 (1978).

[4] Baker, G.A., Bramble, J.H., Thomée, V.: *Single Step Galerkin Approximations for Parabolic Problems.* Math. Comp. 31, p. 818–847 (1977).

[5] Bieterman, M., Babuška, I.: *The Finite Element Method for Parabolic Equations. I. A Posteriori Error Estimation.* Num. Math. 40, p. 339–371 (1982).

[6] Bieterman, M., Babuška, I.: *The Finite Element Method for Parabolic Equations. II. A Posteriori Error Estimation and Adaptive Approach.* Num. Math. 40, p. 373–406 (1982).

[7] Bieterman, M., Babuška, I.: *An Adapative Method of Lines with Error Control for Parabolic Equations of the Reaction-Diffusion Type.* J. Comp. Phys. 63, p. 33–66 (1986).

[8] Bornemann, F.A.: *Adaptive multilevel discretization in time and space for parabolic partial differential equations.* Technical Report TR 89-7, ZIB (1989).

[9] Deuflhard, P.: *Order and Stepsize Control in Extrapolation Methods.* Num. Math. 41, p. 399–422 (1983).

[10] Deuflhard, P.: *Recent Progress in Extrapolation Methods for Ordinary Differential Equations.* SIAM Review 27, p. 505–535 (1985).

[11] Deuflhard, P., Leinen, P., Yserentant, H.: *Concepts of an Adaptive Hierarchical Finite Element Code.* IMPACT of Computing in Science and Engineering, 1, p. 3–35 (1989).

[12] Dunford, N., Schwartz, J.T.: *Linear Operators, Part I: General Theory.* Interscience Publishers, Inc., New York (1957).

[13] Dupont, T.F.: *A Short Survey of Parabolic Galerkin Methods.* In: The Mathematical Basis of Finite Element Methods, D.F. Griffiths (ed.), Claredon Press, Oxford (1984).

[14] Eriksson, K., Johnson, C.: *Adaptive Finite Element Method for Parabolic Problems I: A Linear Model Problem.* Preprint 31, Department of Mathematics, University of Göteborg (1988).

[15] Hairer, E., Nørsett, S.P., Wanner, G.: *Solving Ordinary Differential Equations I. Nonstiff Problems.* Springer, Berlin–Heidelberg–New York (1987).

[16] Kato, T.: *Perturbation Theory for Linear Operators.* Second Corrected Printing of the second Edition. Springer, Berlin–Heidelberg–New York (1984).

[17] Le Roux, M.–N.: *Semidiscretization in Time for Parabolic Problems.* Math. Comp. 33 p. 919–931 (1979).

[18] Lubich, C.: *Discretized Operational Calculus Part I: Theory.* Institut für Mathematik und Geometrie, Universität Innsbruck, Institutsnotiz Nr. 4 (1984).

[19] Lubich, C.: *Convolution Quadrature and Discretized Operational Calculus I.* Num. Math. 52, p. 129–145 (1988).

[20] Lubich, C.: *Convolution Quadrature and Discretized Operational Calculus II.* Num. Math. 52, p. 413–425 (1988).

[21] Miller, K., Miller R.N.: *Moving Finite Elements.* I., SIAM J. Numer. Anal. 18, p. 1019–1032 (1981).

[22] Pazy, A.: *Semigroups of Linear Operators and Applications to Partial Differential Equations.* Springer, Berlin–Heidelberg–New York (1983).

[23] Roitzsch, R.: *KASKADE User's Manual.* Technical Report TR89–4, ZIB (1989).

[24] Roitzsch, R.: *KASKADE Programmer's Manual.* Technical Report TR89–5, ZIB (1989).

[25] Thomée, V.: *Galerkin Finite Element Methods for Parabolic Problems.* Lecture Notes in Mathematics, 1054, Springer, Berlin–Heidelberg–New York (1984).

[26] Wanner, G., Hairer, E., Nørsett, S.P., : *Order stars and stability theorems.* BIT 18, p. 475–489 (1978).

Veröffentlichungen des Konrad-Zuse-Zentrum für Informationstechnik Berlin
Preprints Juni 1990

SC 86-1. P. Deuflhard; U. Nowak. *Efficient Numerical Simulation and Identification of Large Chemical Reaction Systems.* (vergriffen) In: Ber. Bunsenges. Phys. Chem., vol. 90, 1986, 940-946
SC 86-2. H. Melenk; W. Neun. *Portable Standard LISP for CRAY X-MP Computers.*

SC 87-1. J. Anderson; W. Galway; R. Kessler; H. Melenk; W. Neun. *The Implementation and Optimization of Portable Standard LISP for the CRAY.*
SC 87-2. Randolph E. Bank; Todd F. Dupont; Harry Yserentant. *The Hierarchical Basis Multigrid Method.* (vergriffen) In: Numerische Mathematik, 52, 1988, 427-458.
SC 87-3. Peter Deuflhard. *Uniqueness Theorems for Stiff ODE Initial Value Problems.*
SC 87-4. Rainer Buhtz. *CGM-Concepts and their Realizations.*
SC 87-5. P. Deuflhard. *A Note on Extrapolation Methods for Second Order ODE Systems.*
SC 87-6. Harry Yserentant. *Preconditioning Indefinite Discretization Matrices.*

SC 88-1. Winfried Neun; Herbert Melenk. *Implementation of the LISP-Arbitrary Precision Arithmetic for a Vector Processor.*
SC 88-2. H. Melenk; H. M. Möller; W. Neun. *On Gröbner Bases Computation on a Supercomputer Using REDUCE.* (vergriffen)
SC 88-3. J. C. Alexander; B. Fiedler. *Global Decoupling of Coupled Symmetric Oscillators.*
SC 88-4. Herbert Melenk; Winfried Neun. *Parallel Polynomial Operations in the Buchberger Algorithm.*
SC 88-5. P. Deuflhard; P. Leinen; H. Yserentant. *Concepts of an Adaptive Hierarchical Finite Element Code.*
SC 88-6. P. Deuflhard; M. Wulkow. *Computational Treatment of Polyreaction Kinetics by Orthogonal Polynomials of a Discrete Variable.* (vergriffen) In: IMPACT, 1, 1989, 269-301.
SC 88-7. H. Melenk; H. M. Möller; W. Neun. *Symbolic Solution of Large Stationary Chemical Kinetics Problems.*
SC 88-8. Ronald H. W. Hoppe; Ralf Kornhuber. *Multi-Grid Solution of Two Coupled Stefan Equations Arising in Induction Heating of Large Steel Slabs.*
SC 88-9. Ralf Kornhuber; Rainer Roitzsch. *Adaptive Finite-Element-Methoden für konvektions-dominierte Randwertprobleme bei partiellen Differentialgleichungen.*
SC 88-10. S -N. Chow; B. Deng; B. Fiedler. *Homoclinic Bifurcation at Resonant Eigenvalues.*

SC 89-1. Hongyuan Zha. *A Numerical Algorithm for Computing the Restricted Singular Value Decomposition of Matrix Triplets.*
SC 89-2. Hongyuan Zha. *Restricted Singular Value Decomposition of Matrix Triplets.*
SC 89-3. Wu Huamo. *On the Possible Accuracy of TVD Schemes.*
SC 89-4. H. Michael Möller. *Multivariate Rational Interpolation: Reconstruction of Rational Functions.*
SC 89-5. Ralf Kornhuber; Rainer Roitzsch. *On Adaptive Grid Refinement in the Presence of Internal or Boundary Layers.*
SC 89-6. Wu Huamo; Yang Shuli. *MmB-A New Class of Accurate High Resolution Schemes for Conservation Laws in Two Dimensions.*
SC 89-7. U. Budde; M. Wulkow. *Computation of Molecular Weight Distributions for Free Radical Polymerization Systems.*
SC 89-8. Gerhard Maierhöfer. *Ein paralleler adaptiver Algorithmus für die numerische Integration.*
SC 89-9. Harry Yserentant. *Two Preconditioners Based on the Multi-Level Splitting of Finite Element Spaces.*
SC 89-10. Ronald H. W. Hoppe. *Numerical Solution of Multicomponent Alloy Solidification by Multi-Grid Techniques.*

SC 90-1. M. Wulkow; P. Deuflhard. *Towards an Efficient Computational Treatment of Heterogeneous Polymer Reactions.*
SC 90-2. Peter Deuflhard. *Global Inexact Newton Methods for Very Large Scale Nonlinear Problems.*
SC 90-3. Karin Gatermann. *Symbolic solution of polynomial equation systems with symmetry.*
SC 90-4. Folkmar A. Bornemann. *An Adaptive Multilevel Approach to Parabolic Equations I. General Theory & 1D-Implementation.*
SC 90-5. Deuflhard; Freund; Walter. *Fast Secant Methods for the Iterative Solution of Large Nonsymmetric Linear System.*
SC 90-6. Daoliu Wang. *On Symplectic Difference Schemes for Hamiltonian Systems.*

Veröffentlichungen des Konrad-Zuse-Zentrum für Informationstechnik Berlin
Technical Reports **Juni 1990**