

Adaptive multilevel discretization in time and space for parabolic partial differential equations

Diploma thesis of Folkmar A. Bornemann

at

Fachbereich Mathematik, Freie Universität Berlin

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 0933-789X

For all who supported my way to mathematics

Folkmar A. Bornemann

Adaptive multilevel discretization
in time and space for parabolic
partial differential equations

Abstract

The present paper develops an adaptive multilevel approach for parabolic PDE's – as a first step, for one linear scalar equation. Full adaptivity of the algorithm is conceptually realized by simultaneous multilevel discretization in both time and space. Thus the approach combines multilevel time discretization, better known as extrapolation methods, and multilevel finite element space discretization such as the hierarchical basis method. The algorithmic approach is theoretically backed by careful application of fundamental results from semigroup theory. These results help to establish the existence of asymptotic expansions (in terms of time-steps) in Hilbert space. 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 arising space grids are not required to satisfy any quasi-uniformity assumption. Even though the theoretical presentation is independent of space dimension details of the algorithm and numerical examples are given for the 1-D case only. For the 1-D elliptic solver, which is used, an error estimator is established, which works uniformly well for a family of elliptic problems. The numerical results clearly show the significant perspectives opened by the new algorithmic approach.

Contents

1	Introduction	1
2	Theory of time-discretization	3
2.1	The parabolic problem and notation from semigroup theory	3
2.2	The exponential formula	5
2.3	Application to the parabolic problem	9
2.4	Asymptotic expansions I: The selfadjoint case	14
2.5	Asymptotic expansions II: General C_0 -semigroups	30
3	Presentation of the algorithm	34
3.1	The semi-discrete case and the time-step control mechanism	34
3.2	The fully discrete case: the multilevel concept	37
3.3	Perturbation of the extrapolation-table	38
3.4	The order control mechanism	41
3.5	The consistency-estimator $[\gamma_k]$	43
4	Algorithmical details in the 1-D case	45
4.1	Required features of an elliptic solver	45
4.2	Theory for τ -independent elliptic error estimation and amount of work principle	45
4.3	The refinement strategy and the linear solver	53
4.4	Realization of extrapolation	55
4.5	Some remarks about 2-D difficulties	56
5	Realization of the algorithm	57
5.1	The program KASTIX, a short description	57
5.2	Numerical examples	59
	Acknowledgements	72
	References	73

1. Introduction

*Because I know that time is always time
And place is always and only place
And what is actual is actual only for one time ...
from T.S. Eliot: Ash Wednesday*

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 ([14] 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 [15].

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. For this context the method is well analyzed (THOMÉE [36] for Galerkin-methods in space). However, for highly non-uniform grids, possibly varying in time, and adaptive time steps the sequence of discretizations does 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, that 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 1-D case BIETERMAN and BABUŠKA [6, 7] (Galerkin method in space) constructed an a-posteriori error estimator for the parabolic problem to overcome this difficulty. At fixed time-points they produce a new mesh (regridding), possibly doing the last outer time step again by solving an ODE-system. MILLER and MILLER [26] optimize the grid in a finite element method while integrating the ODE's - "*moving finite elements*", thus end-

ing 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 one has 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, what the above difficulties strongly advise – 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 one 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. These levels require the solution of several elliptic problems up to an *individual* accuracy computed by the above levels. These elliptic problems will be solved again by multilevel methods, which produce the adequate individual space–meshes.

This matching seems to be quite natural and this work will show, how it has to be done in detail – on a thorough theoretical basis.

The necessary theory is developed in Chapter 2. Firstly existence theory in the framework of semigroups yields the implicit Euler scheme, secondly asymptotic expansions demanded for extrapolation in the appropriate Hilbert–space are derived.

In Chapter 3 we will explain the multilevel principle used and how the table of extrapolation controls the adaptive meshing in space. In contrast to the foregoing two chapters, which do not depend on the space dimension, we will be concerned in Chapter 4 with special questions arising while implementing the ideas in the 1–D case. These are mainly questions in connection with the elliptic solver. In Chapter 5 the 1–D program KASTIX (KAScade TIme–dependent with eXtrapolation) will be introduced, and some numerical examples are given.

In the opinion of the author the advocated approach seems to be promising and encouraging. The present restriction to scalar parabolic equations is understood as a first step.

2. Theory of time–discretization

In this chapter we shall develop the theory necessary to justify the semi-discrete part of the algorithm explained in chapter 3. Existence of solutions for the parabolic problem is shown using semigroups, which enlightens the distinct rôle of the time variable as already indicated in the introduction. The proof we have chosen directly leads to a discretization in time for which *asymptotic expansions* of the global error are derived in order to get a sound theoretical basis for the procedure of repeated extrapolation.

2.1 The parabolic problem and notation from semigroup theory

Throughout this chapter the following notation and assumptions are valid.

Let $\Omega \subset \mathbb{R}^n$ be a bounded domain and $A(x, D)$ a strongly elliptic differential operator of order $2m$ in divergence form

$$A(x, D)u = \sum_{0 \leq |\rho|, |\sigma| \leq m} (-1)^{|\rho|} D^\rho (a^{\rho\sigma}(x) D^\sigma u) \quad (2.1)$$

in the usual multiindex notation. It is assumed that

$$a^{\rho\sigma} \in C^{|\rho|}(\bar{\Omega}) , \quad (2.2)$$

making the rewriting

$$A(x, D)u = \sum_{0 \leq |\alpha| \leq 2m} a_\alpha(x) D^\alpha u \quad (2.3)$$

possible.

For ease of representation the attention is restricted to the case

$$\partial\Omega \in C^{2m} . \quad (2.4)$$

For the most important case of non-smooth domains i.e. corner domains versions of the theorems on elliptic operators stated later on can be found in DAUGE [9].

Consider the autonomous parabolic *initial-boundary value problem*

$$\begin{aligned}
\text{a)} \quad & \frac{\partial u(t, x)}{\partial t} + A(x, D)u(t, x) = f(x) \\
& \text{for every } x \in \Omega, \quad t \in]0, T] \\
\text{b)} \quad & u(t, x) = \frac{\partial u(t, x)}{\partial n} = \dots = \frac{\partial^{m-1} u(t, x)}{\partial n^{m-1}} = 0 \\
& \text{for every } x \in \partial\Omega, \quad t \in]0, T] \\
\text{c)} \quad & u(0, x) = \varphi(x) \\
& \text{for every } x \in \Omega.
\end{aligned} \tag{2.5}$$

The function spaces for f, φ , a possible solution u and the *sense of solution* as well will be specified later on. $\frac{\partial}{\partial n}$ denotes derivation in direction of the outer normal on $\partial\Omega$.

Remark. The choice of *homogeneous Dirichlet boundary conditions* (2.5.b) was made to avoid clumsy notation but is in no way essential for the following, whilst time-dependent A and f would ask for a non-trivial extension of the theory.

Let X be a space of functions in space variable x , which obey the Dirichlet conditions (2.5.b). The homogeneous problem ($f \equiv 0$) reads then

$$\begin{aligned}
\text{a)} \quad & \frac{d}{dt}u(t) = -Au, \quad t \in]0, T] \\
\text{b)} \quad & u(0) = \varphi.
\end{aligned} \tag{2.5'}$$

This is an *abstract Cauchy-problem*. In analogy to linear ordinary differential equations one expects a “flow” giving the solution formal as

$$u(t) = e^{-At}\varphi, \tag{2.6}$$

indicating the expected properties. But the difficulty arises here that regarding the differential operator $A(x, D)$ as an endomorphism in X leads to unboundedness because of the loss of regularity. The powerful concept of semigroups overcomes that difficulty.

For the following we replace A by $-A$.

A one parameter family $T(t)$ (taking the rôle of e^{At}), $0 \leq t < \infty$, of bounded linear operators from Banach space X into X is called a *semigroup*, if

$$\begin{aligned}
\text{a)} \quad & T(0) = I, \text{ the identity operator on } X \\
\text{b)} \quad & T(t+s) = T(t)T(s) \text{ for every } t, s \geq 0 \\
& \text{(the semigroup property)}
\end{aligned} \tag{2.7}$$

These are exactly the desired properties of the flow (2.6) for *positive time-direction*.

To get a connection to a differential equation like (2.5'.a) we define

$$\begin{aligned} \text{a)} \quad D_A &:= \left\{ x \in X \left| \lim_{t \downarrow 0} \frac{T(t)x - x}{t} \text{ exists} \right. \right\} \\ \text{b)} \quad Ax &:= \lim_{t \downarrow 0} \frac{T(t)x - x}{t} = \left. \frac{d^+ T(t)x}{dt} \right|_{t=0} \\ &\text{for } x \in D_A . \end{aligned} \tag{2.8}$$

A is called the *infinitesimal generator* of the semigroup $T(t)$ with domain D_A .

For a meaningful interpretation of initial values like (2.5'.b) we need the property

$$\text{c)} \quad \lim_{t \downarrow 0} T(t)x = x \quad \text{for every } x \in X . \tag{2.7}$$

Such semigroups are called *strongly continuous* or simply C_0 -semigroups.

For such semigroups is in case of $\varphi \in D_A$

$$u(t) := T(t)\varphi$$

the unique solution of the abstract Cauchy problem (2.5'), $-A$ being replaced by A ([35], Theorems 3.2.1 and 3.2.2). For general $\varphi \in X$ the still existing $u(t) := T(t)\varphi$ can be considered as a generalized solution. Therefore the question arises, which operators A constitute a C_0 -semigroup and how it will be constructed from A , subject of the next section.

2.2 The exponential formula

In this section we state the classical HILLE–YOSIDA theorem, characterizing the infinitesimal generators of C_0 -semigroups of *contractions*, that means

$$\text{d)} \quad \|T(t)\| \leq 1 \quad \text{for every } t > 0 . \tag{2.7}$$

We present the main idea of the original proof given by HILLE [21] for mainly two reasons:

- The construction of the semigroup from its infinitesimal generator made there directly leads to the time-discretization we shall discuss.
- The convergence proof is done by means of an *integral-representation* (formula 2.15), which will be the starting point for the derivation of an asymptotic expansion in section 2.5 – by a formulation given in corollary 2.3.

Theorem 2.1 (HILLE, YOSIDA independently 1948). A linear, possibly unbounded operator A in a B -space X is the infinitesimal generator of a C_0 -semigroup of contractions $T(t)$, $t \geq 0$ iff

a) A is closed and D_A dense in X .

b) The resolvent set $\rho(A)$ of A contains \mathbb{R}^+ and for every $\lambda > 0$ holds

$$\|R(\lambda; A)\| \leq 1/\lambda. \quad (2.9)$$

Here $R(\lambda; A) := (\lambda I - A)^{-1}$ is the resolvent of A .

Proof. (The construction of $T(t)$ in the sufficiency part). The idea is to imitate the exponential formula of EULER

$$e^{\lambda t} = \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda t}{n}\right)^{-n}, \quad (2.10)$$

that means trying to get

$$T(t)\varphi = \lim_{n \rightarrow \infty} (I - tn^{-1}A)^{-n}\varphi. \quad (2.10')$$

The difficulty is now to show the existence of the limit in the right hand side of (2.10').

Because of assumption b) the expression

$$(I - \lambda^{-1}A)^{-1} = \lambda R(\lambda; A) \quad (2.11)$$

exists for $\lambda > 0$ and by (2.9)

$$\|(I - \lambda^{-1}A)^{-1}\| \leq 1. \quad (2.12)$$

Hence, if $\varphi \in D_A$,

$$\begin{aligned} \|(I - \lambda^{-1}A)^{-1}\varphi - \varphi\| &= \|(I - \lambda^{-1}A)^{-1}(\varphi - (I - \lambda^{-1}A)\varphi)\| \\ &= \|(I - \lambda^{-1}A)^{-1}\lambda^{-1}A\varphi\| \\ &\leq \lambda^{-1}\|A\varphi\| \longrightarrow 0 \end{aligned}$$

as $\lambda \longrightarrow \infty$. Since D_A is dense in X and (2.12) holds,

$$(I - \lambda^{-1}A)^{-1}\varphi \longrightarrow \varphi \quad (2.13)$$

strongly for all $\varphi \in X$ as $\lambda \longrightarrow \infty$. Now put $u_n(t) := (I - tn^{-1}A)^{-n}\varphi$ for a $\varphi \in X$. Aiming to show that $\{u_n(t)\}_n$ is fundamental in X for every $t > 0$, we are interested in a representation

$$\begin{aligned} u_n(t) - u_m(t) &= (I - tn^{-1}A)^{-n}\varphi - (I - tm^{-1}A)^{-m}\varphi \\ &= \int_0^t \frac{d}{ds} \left\{ \left(I - \frac{t-s}{m}A\right)^{-m} \left(I - \frac{s}{n}A\right)^{-n} \varphi \right\} ds. \end{aligned} \quad (2.14)$$

Unfortunately it is not at all clear that this integral exists, because differentiation with respect to ε in

$$(1 - \varepsilon A)^{-1} \varphi = \frac{1}{\varepsilon} R\left(\frac{1}{\varepsilon}; A\right) \varphi$$

at $\varepsilon = 0$ causes trouble.

Therefore the approximation

$$\begin{aligned} \Delta_\varepsilon &:= \int_\varepsilon^{t-\varepsilon} \frac{d}{ds} \left\{ \left(I - \frac{t-s}{m} A \right)^{-m} \left(I - \frac{s}{n} A \right)^{-n} \varphi \right\} ds \\ &= \left(I - \frac{\varepsilon}{m} A \right)^{-m} \left(I - \frac{t-\varepsilon}{n} A \right)^{-n} \varphi \\ &\quad - \left(I - \frac{t-\varepsilon}{m} A \right)^{-m} \left(I - \frac{\varepsilon}{n} A \right)^{-n} \varphi \end{aligned}$$

is considered for $0 < \varepsilon < t$. By (2.13) we obtain

$$\Delta_\varepsilon \longrightarrow u_n(t) - u_m(t)$$

strongly as $\varepsilon \longrightarrow 0$. Hence, (2.14) exists at least as an improper Riemann integral in X .

Performing the differentiation elementary calculations give

$$\begin{aligned} u_n(t) - u_m(t) &= \int_0^t \left\{ -A \left(I - \frac{t-s}{m} A \right)^{-m-1} \left(I - \frac{s}{n} A \right)^{-n} \varphi \right. \\ &\quad \left. + \left(I - \frac{t-s}{m} A \right)^{-m} A \left(I - \frac{s}{n} A \right)^{-n-1} \varphi \right\} ds \end{aligned}$$

thus yielding, for $\varphi \in D_{A^2}$ the necessary commutations being allowed,

$$\begin{aligned} u_n(t) - u_m(t) &= \int_0^t \left(I - \frac{t-s}{m} A \right)^{-m-1} \left(I - \frac{s}{n} A \right)^{-n-1} \\ &\quad \left\{ -A \left(I - \frac{s}{n} A \right) \varphi + \left(I - \frac{t-s}{m} A \right) A \varphi \right\} ds \\ &= \int_0^t \left(I - \frac{t-s}{m} A \right)^{-m-1} \left(I - \frac{s}{n} A \right)^{-n-1} \left(\frac{s}{n} - \frac{t-s}{m} \right) A^2 \varphi ds. \end{aligned} \tag{2.15}$$

This implies together with (2.12)

$$\begin{aligned} \|u_n(t) - u_m(t)\| &\leq \int_0^t \left| \frac{s}{n} - \frac{t-s}{m} \right| \|A^2 \varphi\| ds \\ &\leq \int_0^t \left(\frac{s}{n} + \frac{t-s}{m} \right) \|A^2 \varphi\| ds \\ &= \frac{t^2}{2} \left(\frac{1}{n} + \frac{1}{m} \right) \|A^2 \varphi\|. \end{aligned}$$

Hence, for $\varphi \in D_{A^2}$ the limit $\lim_{n \rightarrow \infty} u_n(t)$ exists uniformly on compact intervals $t \in [0, T]$. A further density argument together with (2.12) leads to the uniform convergence on compact intervals on $u_n(t)$ for every $\varphi \in X$. Then it can be shown, that this limit has the desired properties of a C_0 -semigroup of contractions with generator A . ■

Remarks.

- For the whole proof cf. HILLE [21] or TANABE [35], page 57 f.
- In many textbooks the construction of the semigroups is done by means of the YOSIDA-Approximations

$$A_\lambda := A(I - \lambda^{-1}A)^{-1}$$

which are *bounded* linear operators with $A_\lambda \varphi \rightarrow A\varphi$ for $\varphi \in D_A$ as $\lambda \rightarrow \infty$. The validity of the exponential formula is shown then using the fact, that the resolvent is the Laplace transform of the semigroup

$$R(\lambda; A)\varphi = \int_0^\infty e^{-\lambda s} T(s)\varphi ds \text{ for every } \varphi \in X. \quad (2.16)$$

This approach is not useful for section 2.5.

Because of its importance for the following, let us state the construction of the proof as

Corollary 2.2. (The exponential formula)

Let A be the infinitesimal generator of the C_0 -semigroup of contractions $T(t)$. Then the following holds for every $\varphi \in X$:

$$T(t)\varphi = \lim_{n \rightarrow \infty} (I - tn^{-1}A)^{-n}\varphi.$$

■

Interpretation

Consider again the abstract Cauchy problem

$$\begin{aligned} \frac{d}{dt}u(t) &= Au(t) \\ u(0) &= \varphi \end{aligned}$$

Using the *implicit Euler scheme* to discretize in time leads to

$$\frac{u_n\left(\frac{jt}{n}\right) - u_n\left(\frac{(j-1)t}{n}\right)}{t/n} = Au_n\left(\frac{jt}{n}\right) \quad (2.17)$$

$$u_n(0) = \varphi \quad , \quad j = 1, \dots, n \quad ,$$

that means

$$u_n(t) = (I - tn^{-1}A)^{-n}\varphi \quad .$$

Corollary 2.2 asserts that the implicit Euler converges for all $\varphi \in X$, and for $\varphi \in D_A$ to the unique solution of the Cauchy problem (see the text at the end of section 2.1). This is again an argument to consider $T(t)\varphi$ as a generalized solution in case $\varphi \in X \setminus D_A$.

For later purposes (section 2.5) we state the following:

Corollary 2.3. *Let $u_\tau(\cdot)$ denote the implicit Euler approximation (2.17) with timestep $\tau = t/n$. For $\varphi \in D_{A^2}$ holds*

$$u_\tau(t) - u(t) = \int_0^t \frac{s}{n} A^2 (I - \frac{s}{n} A)^{-n-1} T(t-s) \varphi ds \quad (2.18)$$

where $u(t) = T(t)\varphi$.

Proof. Take formula (2.15) and let $m \rightarrow \infty$. Keeping in mind the uniform convergence on compact intervals of the exponential formula and the relation before (2.13) we get the desired result after rearranging. ■

2.3 Application to the parabolic problem

Now the results of the previous section are applied to the parabolic problem (2.5) stated in section 2.1. For this *linear* problem the Hilbert space $L^2(\Omega)$ turns out to be the suitable function space, being considered in the following. Surely the elliptic operator $A(x, D)$ defines an operator A in $L^2(\Omega)$ with domain

$$D_A := H^{2m}(\Omega) \cap H_0^m(\Omega) \quad (2.19)$$

as follows

$$A : D_A \subset L^2(\Omega) \longrightarrow L^2(\Omega)$$

$$(Au)(x) := A(x, D)u(x) \text{ for } u \in D_A \quad . \quad (2.20)$$

Here the spaces $H^k(\Omega)$, $H_0^k(\Omega)$ denote the usual isotropic Sobolev spaces with norm denoted by $\|\cdot\|_k$. The intersection with $H_0^m(\Omega)$ in the definition of D_A

is taken in view of (2.5.b), the homogeneous Dirichlet boundary condition represented here in the sense of the trace operator. Furthermore let $a(\cdot, \cdot)$ denote the bilinear form on $H_0^m(\Omega) \times H_0^m(\Omega)$ associated with $A(x, D)$

$$a(u, v) := \sum_{0 \leq |\rho|, |\sigma| \leq m} \int_{\Omega} a^{\rho\sigma}(x) D^{\sigma} u(x) D^{\rho} v(x) dx . \quad (2.21)$$

For the following $a(\cdot, \cdot)$ is assumed to be $H_0^m(\Omega)$ -elliptic, that means

$$a(u, u) \geq c \|u\|_m^2 \quad (2.22)$$

for every $u \in H_0^m(\Omega)$, c being a positive constant.

For the *homogeneous* problem (2.5), $f \equiv 0$, this is no loss of generality. In fact, generally *Gårding's inequality* ($H_0^m(\Omega)$ -coercivity of $a(\cdot, \cdot)$) holds for strongly elliptic operators

$$a(u, u) \geq c \|u\|_m^2 - \kappa \|u\|_0^2 \quad (2.23)$$

for every $u \in H_0^m(\Omega)$, c, κ positive constants. Now introducing the elliptic operator

$$A_{\kappa}(x, D)u := A(x, D)u + \kappa u \quad (2.24)$$

and transforming the possible solution u of (2.5) to

$$\tilde{u}(t, x) = e^{-\kappa t} u(t, x) \quad (2.25)$$

we get the equivalent homogeneous parabolic problem

$$\begin{aligned} \text{a)} \quad & \tilde{u}_t + A_{\kappa} \tilde{u} = 0 \\ \text{b)} \quad & \tilde{u} \text{ obeys homogeneous Dirichlet conditions} \\ \text{c)} \quad & \tilde{u}(0, \cdot) = \varphi . \end{aligned} \quad (2.26)$$

Since the bilinear form $a_{\kappa}(\cdot, \cdot)$ associated with $A_{\kappa}(x, D)$ is obviously $H_0^m(\Omega)$ -elliptic, the attention can be restricted to that case: transformation of the results to the general case are by now an easy task.

The connection to section 2.2 is established by

Lemma 2.4. *— A as defined in (2.19) and (2.20) is the infinitesimal generator of a C_0 -semigroup of contractions $T(t)$ in $L^2(\Omega)$.*

Proof. According to theorem 2.1 the following three parts have to be shown:

I. D_A is dense in $L^2(\Omega)$.

This is a well known fact from the theory of Sobolev spaces (cf. ADAMS [1]).

II. A is closed.

Let $\{u_i\}_{i=1}^{\infty} \subset D_A$ be a sequence with

$$\begin{aligned} u_i &\longrightarrow u && \text{in } L^2(\Omega) \\ Au_i &\longrightarrow \psi && \text{in } L^2(\Omega) \end{aligned}$$

From elliptic regularity theory the following a priori inequality is known

$$\|u_i - u_j\|_{2m} \leq C (\|A(u_i - u_j)\|_0 + \|u_i - u_j\|_0)$$

(cf. [17] theorem 18.1: here the assumptions from the beginning of section 2.1 are essential.) Hence, $\{u_i\}_i$ is fundamental in $H^{2m}(\Omega)$ and therefore

$$\|u_i - v\|_{2m} \longrightarrow 0 \text{ as } i \longrightarrow \infty$$

for some $v \in H^{2m}(\Omega)$. Since $H^{2m}(\Omega) \hookrightarrow L^2(\Omega)$ is a continuous embedding we get $u = v$. Furthermore $u \in H_0^m(\Omega)$ since $H_0^m(\Omega)$ is the kernel of a trace operator in $H^{2m}(\Omega)$. Thus $u \in D_A$. A , considered as a mapping from $H^{2m}(\Omega)$ to $L^2(\Omega)$, is continuous implying now $Au = \psi$, together with $u \in D_A$ the desired result.

III. $\|R(\lambda; -A)\| \leq 1/\lambda$ for $\lambda > 0$.

Let $\lambda > 0$ and $\psi \in L^2(\Omega)$ be given. Due to the $H_0^m(\Omega)$ -ellipticity of $a(\cdot, \cdot)$ the elliptic problem

$$(\lambda I + A)u = \psi$$

is readily seen by the Lax–Milgram lemma to have a unique *weak* solution $u \in H_0^m(\Omega)$, which is then by elliptic regularity also a strong solution in $H^{2m}(\Omega)$ (cf. [17] theorem 17.2). Thus the resolvent $R(\lambda; -A)$ exists. Furthermore

$$\begin{aligned} \langle \psi, u \rangle_{L^2(\Omega)} &= \langle (\lambda I + A)u, u \rangle_{L^2(\Omega)} \\ &= a(u, u) + \lambda \|u\|_0^2 \\ &\geq \lambda \|u\|_0^2. \end{aligned}$$

The Cauchy–Schwarz inequality yields

$$\lambda \|u\|_0^2 \leq \|u\|_0 \|\psi\|_0, \text{ that means}$$

$$\|u\|_0 \leq 1/\lambda \|\psi\|_0$$

$$\text{implying } \|R(\lambda; -A)\| \leq 1/\lambda$$

■

Remark. In the literature this lemma is usually proven by the Lumer-Phillips theorem using the dissipativity of $-A$. The proof given here is included mainly to impose the rôle of regularity results for the elliptic resolvent equation. Again we refer to DAUGE [9] for non-smooth domains Ω .

Before the central existence and uniqueness result is stated, we shortly introduce the concept of a *holomorphic* C_0 -semigroup of contractions: A C_0 -semigroup of contractions $T(t)$ is called holomorphic if it admits a holomorphic extension $T(t)$ to some sector

$$\Delta_\varphi := \{z | z \neq 0 \wedge |\arg z| < \varphi\}.$$

For us the most important properties of such semigroups are:

- i) for each $t > 0$ $A^k T(t)$ is a bounded operator in X with

$$\|A^k T(t)\| \leq \frac{C}{t^k}, \quad (2.27)$$

C a constant *not* depending on t .

- ii)

$$\frac{d}{dt} T(t)\varphi = AT(t)\varphi \text{ for every } \varphi \in X, \quad t > 0. \quad (2.28)$$

(cf. [40] chapter IX.10)

In extension of lemma 2.4 we get

Lemma 2.5. $-A$ as defined in (2.19) and (2.20) generates a holomorphic C_0 -semigroup of contractions $T(t)$ in $L^2(\Omega)$. This gives a unique $C^\infty([0, T[, D_A)$ solution for the homogeneous problem by (2.28) for all $\varphi \in L^2(\Omega)$.

Proof. A nice proof can be found in chapter 7 of [28]. ■

The *central existence and uniqueness result* is

Theorem 2.6. For $f \in H^{-m}(\Omega) = (H_0^m(\Omega))^*$ and $\varphi \in L^2(\Omega)$ problem (2.5) has a unique solution

$$u \in C^\infty([0, \infty[, H_0^m(\Omega)) \quad (2.29)$$

depending continuously on f and φ .

Differentiation in space has to be understood in the sense of distributions on Ω . The Dirichlet boundary condition is understood to hold in the sense of the trace-operator.

Proof. Following Duhamel's principle the solution is tried to be found in the decomposed form

$$u = w + v , \quad (2.30)$$

where w is solution of the homogeneous equation and v does not depend on the time variable t .

The *stationary problem*

$$Av = f \quad (2.31)$$

has a unique solution

$$v \in H_0^m(\Omega)$$

continuously depending on f . This follows from the $H_0^m(\Omega)$ -ellipticity of $a(\cdot, \cdot)$, cf. section 7.3 in chapter 2 of [24].

Now the *time-dependent homogeneous problem*

$$\begin{aligned} \text{a) } \frac{d}{dt}w + Aw &= 0 \\ \text{b) } w(0) &= \varphi - v \end{aligned} \quad (2.32)$$

is considered. Since $\varphi - v \in L^2(\Omega)$ lemma 2.5 gives a unique solution $w \in C^\infty([0, \infty[, D_A)$ as $w(t) = T(t)(\varphi - v)$, surely depending continuously on φ, v . Defining now u as in (2.30) constitutes the unique solution u with the desired properties, formally given as

$$u(t) = T(t)(\varphi - A^{-1}f) + A^{-1}f . \quad (2.33)$$

■

Remark. The initial value φ is taken in the sense

$$\|u(t) - \varphi\|_{L^2(\Omega)} \longrightarrow 0 \text{ as } t \downarrow 0 .$$

Proof. $T(t)$ is a C_0 -semigroup. ■

Remark. Usually initial data $\varphi \in L^2(\Omega)$ of a parabolic initial boundary value problem, which do not obey the boundary conditions imposed for the solution, are called *inconsistent*. Also a certain amount of smoothness is normally required, otherwise the data are called *nonsmooth*. In our context consistency and smoothness of the initial data $\varphi \in L^2(\Omega)$, shortly *consistency* in future, can be characterized by the largest number m of

$$\varphi \in D_{A^m} .$$

In fact m can even be chosen to be fractional.

Now the convergence of the implicit Euler discretization in time is analysed as suggested by the interpretation of corollary 2.2.

Lemma 2.7. *Notation as in the proof of theorem 2.6. Let $u_\tau(t) \in L^2(\Omega)$ denote the implicit Euler discretization in time of problem (2.5), $w_\tau(t) \in L^2(\Omega)$ the corresponding of problem (2.32). Here $\tau = t/n$ is the time-step. The following results hold:*

$$u_\tau(t) = w_\tau(t) + v \quad (2.34)$$

$$u_\tau(t) - u(t) = w_\tau(t) - w(t) \quad (2.35)$$

$$\|u_\tau(t) - u(t)\|_{L^2(\Omega)} \longrightarrow 0 \text{ as } \tau \rightarrow 0. \quad (2.36)$$

Proof. (2.34) has to be shown, the rest follows from (2.30) and corollary 2.2. The proof of (2.34) is done by induction ($t \rightarrow t + \tau$). Equation (2.32.b) ensures the result for $t = 0$. Furthermore, we have by writing the implicit Euler steps for (2.5) and (2.32):

$$\begin{aligned} \text{a)} \quad & u_\tau(t + \tau) + \tau A u_\tau(t + \tau) = u_\tau(t) + \tau f \\ \text{b)} \quad & w_\tau(t + \tau) + \tau A w_\tau(t + \tau) = w_\tau(t) \end{aligned} \quad (2.37)$$

Taking the difference leads, together with (2.31) and (2.34), to

$$\{u_\tau(t + \tau) - w_\tau(t + \tau)\} + \tau A \{u_\tau(t + \tau) - w_\tau(t + \tau)\} = v + \tau A v. \quad (2.38)$$

Keeping in mind, that the elliptic problem $\bar{v} + \tau A \bar{v} = g$, $g \in H^{-m}(\Omega)$ has a *unique* solution $\bar{v} \in H_0^m(\Omega)$, we get

$$u_\tau(t + \tau) - w_\tau(t + \tau) = v \quad (2.34')$$

■

Remark. The implicit Euler yields to *elliptic problems*. In the literature sometimes the use of the implicit Euler time-discretization as a semi-discretization for parabolic problems is called *Rothe's method* (ROTHE 1930 [32]). Cf. also NEČAS [27], KAČUR [22],[23], REKTORYS [29]. They analyse the convergence directly and apply it to several concrete problems.

2.4 Asymptotic expansions I: The selfadjoint case

In this and the next section the global error term (2.36) of the implicit Euler discretization in $L^2(\Omega)$ is analysed in detail. For extrapolation one is

interested in an *asymptotic expansion*

$$\begin{aligned} \text{a)} \quad & u_\tau(t) - u(t) = e_1(t)\tau + \dots + e_N(t)\tau^N + E_{N+1}(t; \tau)\tau^{N+1} \\ \text{b)} \quad & E_{N+1}(t; \tau) = \mathcal{O}(1) \text{ as } \tau \downarrow 0. \end{aligned} \tag{2.39}$$

Unfortunately the known proof-techniques for asymptotic expansions of one-step methods for ODE's like GRAGG [18], STETTER [34], HAIRER/LUBICH [19] (cf. also [33] and [20] p. 211–214) are *not applicable* here: Due to the *unboundedness* of A , Lipschitz-estimates, Taylor expansions etc. are not possible.

Formula (2.35) of lemma 2.7 allows to restrict the attention to the homogeneous case. For that case two methods will be presented:

1. *Expansion into eigenfunctions of A* (this section). The main features of this approach are
 - inconsistent and non-smooth initial data are possible
 - A must be selfadjoint.
2. *Semigroup approach using integral (2.18)* (next section). Here the main features are
 - N in (2.39) depends on the consistency of the initial data
 - A may be an arbitrary infinitesimal generator of a C_0 -semigroup in a B -space X .

Now the first method is presented. In addition to the assumptions on the operator $A(x, D)$ made at the beginning of sections 2.1 and 2.3 $A(x, D)$ is *assumed to be selfadjoint* and $0 \in \rho(A)$ throughout this section.

Due to the compact embeddings between Sobolev spaces of different indices the Riesz–Schauder theory for compact operators is applicable, leading to:

Lemma 2.8. *There exists an orthonormal basis $\{\chi_n\}_{n=1}^\infty \subset D_A$ of $L^2(\Omega)$ consisting of eigenfunctions of A with corresponding eigenvalues $\{\lambda_n\}_{n=1}^\infty$, $0 < \lambda_1 \leq \lambda_2 \leq \dots$ and $\lambda_n \rightarrow \infty$.*

Proof. See e.g. WLOKA [39]. ■

The starting point of the analysis is

Lemma 2.9. Denote the expansion of φ by

$$\varphi = \sum_{n=1}^{\infty} \hat{\varphi}(n) \chi_n \text{ in } L^2(\Omega) . \quad (2.40)$$

I. The solution of the homogeneous problem (2.5) has the expansion

$$u(t) = \sum_{n=1}^{\infty} \hat{u}(n, t) \chi_n \text{ in } L^2(\Omega) , \quad (2.41)$$

where $\hat{u}(n, t)$ is the solution of the scalar ordinary differential equation

$$\begin{aligned} a) \quad y' &= -\lambda_n y \\ b) \quad y(0) &= \hat{\varphi}(n) \end{aligned} \quad (2.42)$$

namely

$$\hat{u}(n, t) = e^{-\lambda_n t} \hat{\varphi}(n) . \quad (2.43)$$

II. The implicit Euler discretization u_τ has the expansion

$$u_\tau(t) = \sum_{n=1}^{\infty} \hat{u}_\tau(n, t) \chi_n \text{ in } L^2(\Omega) , \quad (2.44)$$

where $\hat{u}_\tau(n, t)$ denotes the implicit Euler discretization of the scalar problem (2.42) for time step τ .

Proof. I. Taking (2.41) as an ansatz, we get

$$\frac{d}{dt} u(t) = \sum_{n=1}^{\infty} \frac{d}{dt} \hat{u}(n, t) \chi_n ,$$

whenever $\hat{u}(n, t)$ is a differentiable real function and

$$\sum_{n=1}^{\infty} \left| \frac{d}{dt} \hat{u}(n, t) \right|^2 < \infty . \quad (2.45)$$

Since A is closed (see part II in the proof of lemma 2.4) we can commute as follows

$$\begin{aligned} Au(t) &= \sum_{n=1}^{\infty} \hat{u}(n, t) A \chi_n \\ &= \sum_{n=1}^{\infty} \hat{u}(n, t) \lambda_n \chi_n \end{aligned} \quad (2.46)$$

and have

$$u(t) \in D_A \quad (2.47)$$

as soon as

$$\sum_{n=1}^{\infty} |\hat{u}(n, t) \lambda_n|^2 < \infty . \quad (2.48)$$

Assuming (2.45) and (2.48) to hold,

$$\frac{d}{dt} u(t) + Au(t) = 0$$

implies

$$\sum_{n=1}^{\infty} \left(\frac{d}{dt} \hat{u}(n, t) + \lambda_n \hat{u}(n, t) \right) \chi_n = 0 .$$

Since $\{\chi_n\}_n$ is an ON-Basis, we have

$$\frac{d}{dt} \hat{u}(n, t) = -\lambda_n \hat{u}(n, t) \text{ for all } n \geq 1 . \quad (2.49)$$

That means, $\hat{u}(n, t)$ necessarily obeys (2.42.a). Requiring $u(t) \rightarrow \varphi$ in $L^2(\Omega)$ as $t \downarrow 0$ leads obviously to

$$\hat{u}(n, 0) = \hat{\varphi}(n) .$$

Thus $\hat{u}(n, \cdot)$ must be the solution of the initial value problem (2.42), that is

$$\hat{u}(n, t) = e^{-\lambda_n t} \hat{\varphi}(n) . \quad (2.43)$$

In order to show that these are the coefficients of the (by the argument unique) solution, we have to show for (2.43) the validity of (2.45) and (2.48). Because of (2.49) they are the same condition.

Since for fixed $t > 0$

$$|\lambda e^{-\lambda t}| \leq \frac{1}{t} e^{-1} \text{ for every } \lambda \geq 0 , \quad (2.50)$$

we get

$$\sum_{n=1}^{\infty} |\hat{u}(n, t) \lambda_n|^2 \leq \left(\frac{1}{t} e^{-1} \right)^2 \sum_{n=1}^{\infty} |\hat{\varphi}(n)|^2 < \infty . \quad (2.51)$$

(Note that Parseval's equality implies now

$$\|Au(t)\|_{L^2(\Omega)} \leq \frac{1}{t} e^{-1} \|\varphi\|_{L^2(\Omega)} ,$$

that is (2.27).)

II. Taking (2.44) as an ansatz, it is shown by induction, that

$\hat{u}_\tau(n, t)$ is the implicit Euler discretization of (2.42). The case $t = 0$ is trivial.

$$u_\tau(t + \tau) + \tau Au_\tau(t + \tau) = u_\tau(t)$$

implies by the closedness of A

$$\sum_{n=1}^{\infty} \left(\hat{u}_{\tau}(n, t + \tau) + \tau \lambda_n \hat{u}_{\tau}(n, t + \tau) \right) \chi_n = \sum_{n=1}^{\infty} \hat{u}_{\tau}(n, t) \chi_n .$$

whenever

$$\sum_{n=1}^{\infty} |\lambda_n \hat{u}_{\tau}(n, t + \tau)|^2 < \infty . \quad (2.52)$$

Thus we have necessarily

$$\hat{u}_{\tau}(n, t + \tau) = \frac{1}{1 + \tau \lambda_n} \hat{u}_{\tau}(n, t)$$

an implicit Euler step for (2.42).

Expression (2.52) is shown using

$$\left| \frac{\lambda}{1 + \tau \lambda} \right| \leq 1/\tau \quad (2.53)$$

and the square summability of $\{\hat{u}_{\tau}(n, t)\}_n$, making thus the induction step perfect. \blacksquare

Remark. En passant the proof of lemma 2.9 has given an independent proof of theorem 2.6 in the *homogeneous selfadjoint case*. For the inhomogeneous case and *this method* see e.g. TRIEBEL [37] §44.

Lemma 2.9 gives the global error of the implicit Euler as

$$u_{\tau}(t) - u(t) = \sum_{n=1}^{\infty} [\hat{u}_{\tau}(n, t) - \hat{u}(n, t)] \chi_n . \quad (2.54)$$

We are therefore led to analyse the implicit Euler discretization of Dahlquist's old scalar test equation

$$\begin{aligned} \text{a) } y' &= -\lambda y, \quad \lambda \geq 0 \\ \text{b) } y(0) &= y_0 \end{aligned} \quad (2.55)$$

in dependence of λ , not for a fixed λ as in the theory of stiff ODE's.

This is done by the following essential lemma (the proof is a little bit monstrous) which is the key-lemma of this paper.

Lemma 2.10.

I. The global error of the implicit Euler discretization for the scalar test equation (2.55) has the asymptotic expansion

$$\begin{aligned} y_\tau(t) - y(t) = & \lambda P_1(t\lambda) e^{-\lambda t} y_0 \tau \\ & + \lambda^2 P_2(t\lambda) e^{-\lambda t} y_0 \tau^2 + \dots \\ & + \lambda^N P_N(t\lambda) e^{-\lambda t} y_0 \tau^N + R_{N+1}(t, \lambda, \tau) y_0 \tau^{N+1}, \end{aligned} \quad (2.56)$$

where $P_k(\cdot)$ are k^{th} order polynomials with $P_k(0) = 0$ and for $R_{N+1}(t, \lambda, \tau)$, given $t, \tau_0 > 0$, holds

$$R_{N+1}(t, \lambda, \tau) \text{ uniformly bounded in } \lambda \in]0, \infty[, \tau \in [0, \tau_0]. \quad (2.57)$$

II. The following recurrence formula holds for the polynomials $P_k(\cdot)$

$$\begin{aligned} a) \quad & P_0 \equiv 1 \\ b) \quad & P'_k(x) + \frac{k P_k(x)}{x} = \sum_{j=0}^{k-1} (-1)^{k-j-1} P_j(x), \quad k \geq 1 \\ c) \quad & P_k(0) = 0, \quad k \geq 1 \end{aligned} \quad (2.58)$$

(Note that $\frac{k P_k(x)}{x}$ is no singular term because of c)).

Proof.

Part I. a) We have

$$\begin{aligned} y_\tau(t) &= (1 + \tau\lambda)^{-t/\tau} y_0 \\ &= \exp \left\{ -\frac{t}{\tau} \log(1 + \tau\lambda) \right\} y_0 \end{aligned} \quad (2.59)$$

Inserting the power-series for $\log(1 + \cdot)$, converging for $0 < \tau\lambda < 1$, yields

$$\begin{aligned} y_\tau(t) &= \exp \left\{ -t\lambda + t\lambda \frac{\tau\lambda}{2} - t\lambda \frac{\tau^2\lambda^2}{3} + \dots \right\} y_0 \\ &= \exp \left\{ t\lambda \left(\frac{\tau\lambda}{2} - \frac{\tau^2\lambda^2}{3} + \dots \right) \right\} y(t) \end{aligned}$$

Since $\exp(\cdot)$ is an entire function, we get

$$\begin{aligned} y_\tau(t) &= \left\{ 1 + t\lambda \left(\frac{\tau\lambda}{2} - \dots \right) + \frac{1}{2!} (t\lambda)^2 \left(\frac{\tau\lambda}{2} - \dots \right)^2 + \dots \right\} y(t) \\ &= (1 + P_1(t\lambda)\lambda\tau + P_2(t\lambda)\lambda^2\tau^2 + P_3(t\lambda)\lambda^3\tau^3 + \dots) y(t), \end{aligned} \quad (2.60)$$

then $P_k(\cdot)$ being polynomials of order k with $P_k(0) = 0$, thus (2.58.c).

b) The power-series (2.60) gives us the coefficients of the Taylor expansion with respect to τ for arbitrary $\tau\lambda > 0$:

$$y_\tau(t) = y(t) + P_1(t\lambda)\lambda y(t)\tau + \dots + P_N(t\lambda)\lambda^N y(t)\tau^N + R_{N+1}(t, \lambda, \tau)y_0\tau^{N+1}, \quad (2.61)$$

where $R_{N+1}(t, \lambda, \tau)$ is continuous in all entries, that is (2.56).

c) To finish part I we have to show the assertion (2.57) of uniform boundedness. This will be rather lengthy and is accomplished in several steps:

Step 1: The Lagrangian form of the remainder $R_N(t, \lambda, \tau)$ yields

$$R_N(t, \lambda, \tau)y_0 = \left(\frac{\partial}{\partial\tau}\right)^N y_\tau(t) \Big|_{\tau=\vartheta} \text{ for some } \vartheta \in]0, \tau[. \quad (2.62)$$

Now we can represent the right hand side in the form

$$\left(\frac{\partial}{\partial\tau}\right)^N y_\tau(t) = \sum_{j=1}^N \frac{t^j P_j^{(N)}(1 + \lambda\tau, \log(1 + \lambda\tau))}{\tau^{N+j}(1 + \lambda\tau)^N} y_\tau(t), \quad (2.63)$$

where the $P_j^{(N)}(x, y)$ are polynomials of degree N in x and degree j in y . This can readily be seen by induction, which will be done in the rest of this step. The case $N = 1$ is given by (2.64) below.

$$\begin{aligned} \left(\frac{\partial}{\partial\tau}\right)^{N+1} y_\tau(t) &= \frac{\partial}{\partial\tau} \left[\sum_{j=1}^N \frac{t^j P_j^{(N)}(\cdot, \cdot)}{\tau^{N+j}(1 + \tau\lambda)^N} y_\tau(t) \right] \\ &+ \left[\sum_{j=1}^N \frac{t^j P_j^{(N)}(\cdot, \cdot)}{\tau^{N+j}(1 + \tau\lambda)^N} \right] \frac{\partial}{\partial\tau} y_\tau(t). \end{aligned}$$

By (2.59) we get

$$\frac{\partial}{\partial\tau} y_\tau(t) = \frac{t[(1 + \tau\lambda) \log(1 + \tau\lambda) - \tau\lambda]}{\tau^2(1 + \tau\lambda)} y_\tau(t). \quad (2.64)$$

Thus

$$\begin{aligned} \left(\frac{\partial}{\partial \tau}\right)^{N+1} y_\tau(t) = & \left\{ -(N+j) \sum_{j=1}^N \frac{t^j P_j^{(N)}(\cdot, \cdot)(1+\tau\lambda)}{\tau^{N+1+j}(1+\tau\lambda)^{N+1}} \right. \\ & + \sum_{j=1}^N \frac{t^j \left(\frac{d}{dx} P_j^{(N)}\right)(\cdot, \cdot) \lambda \cdot \tau(1+\tau\lambda)}{\tau^{N+j}(1+\tau\lambda)^N \tau(1+\tau\lambda)} \\ & + \sum_{j=1}^N \frac{t^j \left(\frac{d}{dy} P_j^{(N)}\right)(\cdot, \cdot) \frac{\lambda}{1+\tau\lambda} \tau(1+\tau\lambda)}{\tau^{N+j}(1+\tau\lambda)^N \tau(1+\tau\lambda)} \\ & + (-N) \sum_{j=1}^N \frac{t^j P_j^{(N)}(\cdot, \cdot) \cdot \lambda \cdot \tau}{\tau^{N+j}(1+\lambda\tau)^{N+1}\tau} \\ & \left. + \sum_{j=1}^N \frac{t^j P_j^{(N)}(\cdot, \cdot)}{\tau^{N+j}(1+\lambda\tau)^N} \frac{t [\log(1+\tau\lambda)(1+\tau\lambda) - \tau\lambda]}{\tau^2(1+\tau\lambda)} \right\} y_\tau(t). \end{aligned}$$

Rearranging of terms leads to

$$\begin{aligned} \left(\frac{\partial}{\partial \tau}\right)^{N+1} y_\tau(t) = & \left\{ \sum_{j=1}^N t^j \tau^{-(N+1+j)} \left[-(N+j) P_j^{(N)}(\cdot, \cdot)(1+\tau\lambda) \right] \right. \\ & + \sum_{j=1}^N t^j \tau^{-(N+1+j)} \left[\left(\frac{d}{dx} P_j^{(N)}\right)(\cdot, \cdot)(1+\tau\lambda)(1+\tau\lambda-1) \right] \\ & + \sum_{j=1}^N t^j \tau^{-(N+1+j)} \left[\left(\frac{d}{dy} P_j^{(N)}\right)(\cdot, \cdot)(1+\tau\lambda-1) \right] \\ & + \sum_{j=1}^N t^j \tau^{-(N+1+j)} \left[-N P_j^{(N)}(\cdot, \cdot)(1+\tau\lambda-1) \right] \\ & + \sum_{j=2}^{N+1} t^j \tau^{-(N+1+j)} \left[P_{j-1}^{(N)}(\cdot, \cdot)((1+\tau\lambda) \log(1+\tau\lambda) \right. \\ & \left. \left. - (1+\tau\lambda) + 1) \right] \right\} \frac{y_\tau(t)}{(1+\tau\lambda)^{N+1}} \end{aligned}$$

The sum of the terms in square brackets constitutes just the desired polynomials $P_j^{(N+1)}((1+\tau\lambda), \log(1+\tau\lambda))$ of order $N+1$ in $(1+\tau\lambda)$ and order j in $\log(1+\tau\lambda)$.

Step 2: Aim is to use (2.63) for estimating $R_N(t, \lambda, \tau)$ in the vicinity of $\tau = 0$. To this end we first observe, that by (2.61) for $t > 0$, $\lambda > 0$, using Taylor's theorem

$$R_N(t, \lambda, \tau) y_0 \longrightarrow P_N(t\lambda) \lambda^N y(t) \text{ as } \tau \downarrow 0. \quad (2.65)$$

Comparison with (2.62) and (2.63) yields

$$\frac{P_j^{(N)}(1+\lambda\tau, \log(1+\tau\lambda))}{\tau^{N+j}(1+\tau\lambda)^N} y_\tau(t) \longrightarrow \pi_j^N \lambda^{N+j} y(t) \text{ as } \tau \downarrow 0. \quad (2.66)$$

Here π_j^N denotes the j^{th} coefficient of $P_N(\cdot)$.

Since $y_\tau(t) \rightarrow y(t) \neq 0$ as $\tau \downarrow 0$, we have

$$\frac{P_j^{(N)}(1 + \lambda\tau, \log(1 + \lambda\tau))}{\tau^{N+j}(1 + \tau\lambda)^N} \rightarrow \pi_j^N \lambda^{N+j}$$

as $\tau \downarrow 0$, or for every $\lambda > 0$,

$$\frac{P_j^{(N)}(1 + \lambda\tau, \log(1 + \lambda\tau))}{(\tau\lambda)^{N+j}(1 + \tau\lambda)^N} \rightarrow \pi_j^N \text{ as } \tau \downarrow 0. \quad (2.67)$$

Furthermore we have, because of the orders of $P_j^{(N)}(\cdot)$:

$$\begin{aligned} \frac{P_j^{(N)}(1 + \xi, \log(1 + \xi))}{\xi^{N+j}(1 + \xi)^N} &= \mathcal{O}\left(\frac{(1 + \xi)^N \log^j(1 + \xi)}{(1 + \xi)^N \xi^j} \frac{1}{\xi^N}\right) \\ &= \mathcal{O}\left(\frac{1}{\xi^N}\right) \text{ as } \xi \rightarrow \infty. \end{aligned} \quad (2.68)$$

If we put $\tau = \xi$, $\lambda = 1$ in (2.67) we thus get constants $C_j^N > 0$ with

$$\left| \frac{P_j^{(N)}(1 + \xi, \log(1 + \xi))}{\xi^{N+j}(1 + \xi)^N} \right| \leq C_j^N, \quad \forall \xi > 0. \quad (2.69)$$

Inserting $\xi = \tau\lambda$ yields

$$\left| \frac{P_j^{(N)}(1 + \tau\lambda, \log(1 + \tau\lambda))}{\tau^{N+j}(1 + \tau\lambda)^N} \right| \leq C_j^N \lambda^{N+j}, \quad \forall \tau > 0, \lambda > 0. \quad (2.70)$$

Thus by (2.63) for $t, \tau, \lambda > 0$:

$$\left| \left(\frac{\partial}{\partial \tau} \right)^N y_\tau(t) \right| \leq \left\{ \sum_{j=1}^N C_j^N t^j \lambda^{N+j} \right\} |y_\tau(t)|. \quad (2.71)$$

Now is for fixed $\lambda, t > 0$

$$|y_\tau(t)| \text{ isotone in } \tau. \quad (2.72)$$

In fact (2.64) yields

$$\frac{\partial}{\partial \tau} (1 + \tau\lambda)^{-t/\tau} = \frac{t}{\tau^2} \left[\log(1 + \tau\lambda) - \frac{\tau\lambda}{1 + \tau\lambda} \right] (1 + \tau\lambda)^{-t/\tau}$$

This is positive, since the term in square brackets is easily seen to be positive.

Thus (2.71) leads, if we restrict $\tau \in [0, \bar{\tau}_0]$ for some $\bar{\tau}_0$ specified later on, to

$$\left| \left(\frac{\partial}{\partial \tau} \right)^N y_\tau(t) \right| \leq \sum_{j=1}^N C_j^N t^j \lambda^{N+j} \frac{1}{(1 + \bar{\tau}_0 \lambda)^{t/\bar{\tau}_0}} |y_0|. \quad (2.73)$$

Now we set

$$\bar{\tau}_0 := t/2N > 0 \quad (2.74)$$

and get

$$\left| \left(\frac{\partial}{\partial \tau} \right)^N y_\tau(t) \right| \leq \tilde{C} \frac{\lambda^{2N}}{(1 + \bar{\tau}_0 \lambda)^{2N}} |y_0| \leq C |y_0|,$$

for all $\lambda > 0$, where \tilde{C} and C depend only on $t > 0$ and N . Therefore we have proven (2.57) for the vicinity $\tau \in [0, \bar{\tau}_0]$.

Step 3: We now show (2.57) for τ bounded away from zero: $\tau \in [\bar{\tau}_0, \tau_0]$, $\bar{\tau}_0$ defined in (2.74). Formula (2.61) yields together with (2.72)

$$\begin{aligned} |R_N(t, \lambda, \tau) y_0| &\leq \frac{1}{\tau^N} \left\{ |y_\tau(t)| + |y(t)| + \sum_{j=1}^{N-1} \lambda^j |P_j(t\lambda)| e^{-\lambda t} |y_0| \tau^j \right\} \\ &\leq \frac{1}{\bar{\tau}_0^N} \left\{ \frac{1}{(1 + \tau_0 \lambda)^{t/\tau_0}} |y_0| + e^{-\lambda t} |y_0| \right. \\ &\quad \left. + \sum_{j=1}^{N-1} \lambda^j |P_j(t\lambda)| e^{-\lambda t} |y_0| \tau_0^j \right\} \\ &\leq C \text{ for } \lambda \in]0, \infty[\end{aligned} \quad (2.75)$$

since all terms in braces are vanishing as $\lambda \rightarrow \infty$. C depends only on t and N . This finishes the proof of part I.

Part II. The recurrency formula bases on (2.60) and the possibility of two different representations for $y_\tau(t + \tau)$. Let $0 < \tau\lambda < 1$ and $P_0 \equiv 1$ (that is (2.58.a)).

Representation 1. “One further Euler step”. Formula (2.60) yields

$$\begin{aligned} y_\tau(t + \tau) &= \frac{1}{1 + \tau\lambda} y_\tau(t) \\ &= \sum_{j=0}^{\infty} (-\tau\lambda)^j \sum_{k=0}^{\infty} P_k(t\lambda) \lambda^k \tau^k y(t) \\ &= \sum_{k=0}^{\infty} \left\{ \sum_{j=0}^k (-1)^j P_{k-j}(t\lambda) \right\} \lambda^k \tau^k y(t) \end{aligned} \quad (2.76)$$

Representation 2. “Differentiation with respect to λ ”. We have

$$\begin{aligned}\frac{d}{d\lambda}y_\tau(t) &= -\frac{t}{\tau}\tau(1+\tau\lambda)^{-(t/\tau+1)} \\ &= -ty_\tau(t+\tau).\end{aligned}\tag{2.77}$$

Therefore

$$\begin{aligned}y_\tau(t+\tau) &= -\frac{1}{t}\frac{d}{d\lambda}y_\tau(t) \\ &= -\frac{1}{t}\frac{d}{d\lambda}\sum_{k=0}^{\infty}P_k(t\lambda)\lambda^k\tau^ky(t) \\ &= -\frac{1}{t}\sum_{k=0}^{\infty}\left\{P'_k(t\lambda)t\lambda^k\tau^k + P_k(t\lambda)k\lambda^{k-1}\tau^k\right. \\ &\quad \left.- P_k(t\lambda)t\lambda^k\tau^k\right\}y(t) \\ &= \sum_{k=0}^{\infty}\left\{P_k(t\lambda) - k\frac{P_k(t\lambda)}{t\lambda} - P'_k(t\lambda)\right\}\lambda^k\tau^ky(t)\end{aligned}\tag{2.78}$$

Since $P_k(0) = 0$ for $k \geq 1$ was already proven in part I, the middle term is not singular for $t\lambda = 0$. (For $k = 0$ it does not exist at all!)

Comparison of (2.76) and (2.78) yields

$$\sum_{j=0}^k(-1)^jP_{k-j}(t\lambda) = P_k(t\lambda) - k\frac{P_k(t\lambda)}{t\lambda} - P'_k(t\lambda)$$

which is (2.58.b) for $x = t\lambda$. ■

Remark. The recurrence formula (2.58) is needed in the next section. Since we are also interested in t -dependence of (2.56) and (2.57) we prove

Corolary 2.11. *Notation as in lemma 2.10. For given $t > 0$, $0 \leq \tau_0 \leq t$ (note that only $\tau \leq t$ need to be considered since for the implicit Euler $\tau = t/n$ for some $n \in \mathbb{N}$) we have*

$$\begin{aligned}a) \quad |\lambda^k P_k(t\lambda)e^{-\lambda t}| &\leq \frac{C}{t^k}, \quad 1 \leq k \leq N \\ b) \quad |R_{N+1}(t, \lambda, \tau)| &\leq \frac{C}{t^{N+1}}, \quad \forall \tau \in [0, \tau_0]\end{aligned}\tag{2.79}$$

uniformly in $\lambda > 0$; C depends only on N .

Proof. a) The function $\lambda^m e^{-\lambda t}$ take for given $t > 0$ its maximum in $\lambda \in [0, \infty]$ at

$$\lambda_{\max} = \frac{m}{t} .$$

Thus

$$\lambda^m e^{-\lambda t} \leq m^m \frac{e^{-m}}{t^m} ,$$

which yields

$$\begin{aligned} |\lambda^k P_k(t\lambda) e^{-\lambda t}| &\leq \sum_{j=1}^k |\pi_j^k| t^j \lambda^{k+j} e^{-\lambda t} \\ &\leq \sum_{j=1}^k |\pi_j^k| \frac{e^{-(k+j)t}}{t^{k+j}} (k+j)^{k+j} \\ &\leq C/t^k , \end{aligned}$$

that means (2.79.a).

For b) we have to work a little harder. Starting point is formula (2.71) together with (2.62). Inequality (2.71) suggests in analogy with part a) to search the maximum in λ for fixed $t, \tau > 0$ of the term:

$$\lambda^m \frac{1}{(1 + \tau\lambda)^{t/\tau}} .$$

Obviously this makes sense only for $\tau \in [0, t/m]$. Differentiation with respect to λ yields

$$\frac{\partial}{\partial \lambda} \left[\lambda^m \frac{1}{(1 + \tau\lambda)^{t/\tau}} \right] = \lambda^{m-1} \left[m - \frac{\lambda t}{1 + \tau\lambda} \right] y_\tau(t) , \quad (2.80)$$

thus giving

$$\lambda_{\max} = \frac{m}{t - m\tau} . \quad (2.81)$$

Inserting gives for $t > 0, \tau \in [0, t/m[$

$$\lambda^m \frac{1}{(1 + \tau\lambda)^{t/\tau}} \leq \frac{m^m}{(t - m\tau)^m} \frac{1}{(1 + \frac{\tau m}{t - m\tau})^{t/\tau}} , \quad (2.82)$$

for every $\lambda \geq 0$. This gives by rearranging

$$t^m \lambda^m \frac{1}{(1 + \tau\lambda)^{t/\tau}} \leq m^m (1 + m \frac{\tau}{t - m\tau})^{-\frac{t-m\tau}{\tau}} . \quad (2.83)$$

If we set

$$\eta = \frac{\tau}{t - m\tau} \quad (2.84)$$

and observe

$$\begin{aligned} \text{a)} \quad & (1 + m\eta)^{-1/\eta} \longrightarrow e^{-m} \text{ as } \eta \rightarrow 0 \\ \text{b)} \quad & (1 + m\eta)^{-1/\eta} \longrightarrow 1 \text{ as } \eta \rightarrow \infty, \end{aligned} \quad (2.85)$$

we get

$$t^m \lambda^m \frac{1}{(1 + \tau \lambda)^{t/\tau}} \leq C \quad (2.86)$$

for every $t > 0$, $\lambda \geq 0$, $\tau \in [0, t/m]$.

Thus (2.71) and (2.62) give for $t > 0$, $\tau \in [0, \bar{\tau}_0]$, $\bar{\tau}_0$ as in (2.74)

$$|R_{N+1}(t, \lambda, \tau)| \leq \frac{C}{t^{N+1}} \text{ uniformly in } \lambda,$$

C depending only on N . For the remaining $\tau \in [\bar{\tau}_0, \tau_0]$ formula (2.75) yields, keeping $\tau_0 \leq t$ in mind,

$$|R_{N+1}(t, \lambda, \tau)y_0| \leq \frac{1}{\bar{\tau}_0^{N+1}} \left\{ \frac{1}{(1 + \tau_0 \lambda)^{t/\tau_0}} |y_0| + e^{-\lambda t} |y_0| + \sum_{j=1}^N \lambda^j |P_j(t\lambda)| e^{-\lambda t} |y_0| t^j \right\}.$$

By (2.79.a) the term in braces is bounded independently of $t, \lambda > 0$. Thus with (2.74)

$$|R_{N+1}(t, \lambda, \tau)| \leq \frac{C}{t^{N+1}}$$

for $t > 0$, $\tau \in [\bar{\tau}_0, \tau_0]$ uniformly in $\lambda > 0$, C depending only on N ; completing the proof of (2.79.b). \blacksquare

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

Theorem 2.12. *Notation as before. For $t > 0$, $\tau_0 \leq t$ there exists an asymptotic expansion in $L^2(\Omega)$:*

$$u_\tau(t) - u(t) = e_1(t)\tau + \dots + e_N(t)\tau^N + E_{N+1}(t; \tau)\tau^{N+1}. \quad (2.87)$$

For $\tau \in [0, \tau_0]$ we have

$$\begin{aligned} \text{a)} \quad & \|e_k(t)\|_{L^2(\Omega)} \leq \frac{C}{t^k} \|\varphi\|_{L^2(\Omega)}, \quad 1 \leq k \leq N \\ \text{b)} \quad & \|E_{N+1}(t; \tau)\|_{L^2(\Omega)} \leq \frac{C}{t^{N+1}} \|\varphi\|_{L^2(\Omega)} \end{aligned} \quad (2.88)$$

C depends only on N . The functions $e_k(t)$ are explicitly given by

$$e_k(t) = A^k P_k(tA)u(t); \quad (2.89)$$

here the $P_k(\cdot)$ are the polynomials as defined in (2.58).

Proof. Using (2.54) and lemma 2.10 we get

$$\begin{aligned}
u_\tau(t) - u(t) &= \sum_{n=1}^{\infty} \left[\lambda_n P_1(t\lambda_n) e^{-\lambda_n t} \hat{\varphi}(n) \tau + \dots \right. \\
&\quad \left. + \lambda_n^N P_N(t\lambda_n) e^{-\lambda_n t} \hat{\varphi}(n) \tau^N \right. \\
&\quad \left. + R_{N+1}(t, \lambda_n, \tau) \hat{\varphi}(n) \tau^{N+1} \right] \chi_n \\
&= \sum_{n=1}^{\infty} A P_1(tA) \hat{u}(n, t) \chi_n \tau + \dots + \sum_{n=1}^{\infty} A^N P_N(tA) \hat{u}(n, t) \chi_n \tau^N \\
&\quad + \sum_{n=1}^{\infty} R_{N+1}(t, \lambda_n, \tau) \hat{\varphi}(n) \chi_n \tau^{N+1} .
\end{aligned}$$

The closedness of A and corollary 2.11 give the assertion. ■

We close this section with some remarks on the estimates (2.88).

Remarks.

1. Formula (2.88.a) and even more can be proved using theory of holomorphic C_0 -semigroups and (2.89) (keeping lemma 2.5 in mind). Let $\varphi \in D_{A^m}$ for some $m \in \mathbb{N}_0$. Then with C only depending on N :

$$\begin{aligned}
\text{a) } \|e_k(t)\|_{L^2(\Omega)} &\leq \frac{C}{t^{k-m}} \|A^m \varphi\|_{L^2(\Omega)} \quad \text{for } k \geq m \\
\text{b) } \|e_k(t)\|_{L^2(\Omega)} &\leq C t \|A^{k+1} \varphi\|_{L^2(\Omega)} \quad \text{for } k < m
\end{aligned} \tag{2.90}$$

Proof.

$$\text{a) } e_k(t) = A^{k-m} \sum_{j=1}^k \pi_j^k t^j A^j T(t) A^m \varphi, \text{ thus with (2.27)}$$

$$\begin{aligned}
\|e_k(t)\|_{L^2(\Omega)} &\leq \sum_{j=1}^k |\pi_j^k| t^j \|A^{k-m+j} T(t)\| \|A^m \varphi\|_{L^2(\Omega)} \\
&\leq \sum_{j=1}^k |\pi_j^k| \tilde{C} t^{j-(k-m+j)} \|A^m \varphi\|_{L^2(\Omega)} \\
&\leq C/t^{k-m} \|A^m \varphi\|_{L^2(\Omega)}
\end{aligned}$$

$$\text{b) } e_k(t) = \sum_{j=1}^k \pi_j^k t^j A^{j-1} T(t) A^{k+1} \varphi,$$

thus with (2.27)

$$\begin{aligned}
\|e_k(t)\| &\leq \sum_{j=1}^k |\pi_j^k| \tilde{C} t^{j-(j-1)} \|A^{k+1} \varphi\|_{L^2(\Omega)} \\
&\leq C t \|A^{k+1} \varphi\|_{L^2(\Omega)}.
\end{aligned}$$

■

2. A formula analogous to (2.90) holds, whenever $\varphi \in D_{A^\alpha}$ with broken α . Compare Theorem 6.13 of chapter 2 in [28]. In our context fractional powers of A are intimately connected with Sobolev spaces of broken index, see e.g. [24].
3. The inequalities (2.88.a) respectively (2.90) are the best, that means *in general* the exponent of t can not be relaxed. We show this by the following

Example. Consider the heat-equation on $\Omega = [0, \pi]$ with homogeneous Dirichlet boundary conditions imposed. That means

$$A = -\frac{d^2}{dx^2} . \quad (2.91)$$

Here we have

$$\begin{aligned} \text{a)} \quad \lambda_n &= n^2 \\ \text{b)} \quad \chi_n &= \sqrt{\frac{2}{\pi}} \sin(n \cdot) . \end{aligned} \quad (2.92)$$

Now we consider the family

$$\varphi_\vartheta = \sum_{n=1}^{\infty} \frac{1}{n^{1/2+2\vartheta}} \sin(n \cdot) \quad (2.93)$$

of initial data, which is in $L^2[0, \pi]$ for $\vartheta > 0$. E.g. we have

$$\varphi_{1/4}(x) = \frac{\pi - x}{2} . \quad (2.94)$$

Since

$$P_1(x) = \frac{x}{2} , \quad (2.95)$$

we get

$$e_1(t) = \sum_{n=1}^{\infty} \frac{n^4}{2} t e^{-n^2 t} \frac{1}{n^{1/2+2\vartheta}} \sin(n \cdot) , \quad (2.96)$$

or by Parseval's equality

$$\|e_1(t)\|_{L^2(\Omega)}^2 = \frac{\pi}{2} \sum_{n=1}^{\infty} \frac{n^8}{4} t^2 e^{-2n^2 t} \frac{1}{n^{1+4\vartheta}} . \quad (2.97)$$

Choosing $t = \frac{1}{N^2}$ for an integer N , we estimate

$$\begin{aligned} \|e_1(t)\|_{L^2(\Omega)}^2 &\geq C_1 \sum_{n=1}^N \frac{n^{7-4\vartheta}}{N^4} \\ &\geq C_2 \frac{N^{8-4\vartheta}}{N^4} = C_2 t^{-(2-2\vartheta)} \end{aligned}$$

Thus we get the lower bound

$$\|e_1(t)\|_{L^2(\Omega)} \geq \frac{C_3}{t^{1-\vartheta}}. \quad (2.98)$$

On the other hand for $\psi = \sum_{n=1}^{\infty} \hat{\psi}_n \sin(n\cdot) \in L^2[0, \pi]$ and $0 < \alpha < 1$ we have

$$\begin{aligned} A^{-\alpha} \psi &= \frac{\sin \pi \alpha}{\pi} \int_0^{\infty} \sigma^{-\alpha} \left(\sigma I - \frac{d^2}{dx^2} \right)^{-1} \psi d\sigma \\ &= \sum_{n=1}^{\infty} \frac{\sin \pi \alpha}{\pi} \hat{\psi}_n \int_0^{\infty} \sigma^{-\alpha} \left(\sigma I - \frac{d^2}{dx^2} \right)^{-1} \sin(n\cdot) d\sigma \\ &= \sum_{n=1}^{\infty} \frac{\sin \pi \alpha}{\pi} \hat{\psi}_n \int_0^{\infty} \frac{\sigma^{-\alpha} d\sigma}{\sigma + n^2} \sin(n\cdot) \\ &= \sum_{n=1}^{\infty} \frac{\hat{\psi}_n}{n^{2\alpha}} \sin(n\cdot). \end{aligned} \quad (2.99)$$

Consider $\varphi = \sum_{n=1}^{\infty} \hat{\varphi}_n \sin(n\cdot) \in L^2[0, \pi]$. We deduce from (2.99), since $D_{A^\alpha} = \text{range}(A^{-\alpha})$ (cf. theorem 6.8 in chapter 2 of [28])

$$\varphi \in D_{A^\alpha} \Leftrightarrow \{\hat{\varphi}_n n^{2\alpha}\}_n \in l^2. \quad (2.100)$$

Therefore

$$\varphi_\vartheta \in D_{A^\alpha} \text{ for } \alpha < \vartheta. \quad (2.101)$$

Inequality (2.90.a) gives (cf. remark 2!) now the upper bound

$$\|e_1(t)\|_{L^2(\Omega)} \leq \frac{C_4}{t^{1-\alpha}} \text{ for } \alpha < \vartheta. \quad (2.102)$$

The connection of (2.102) and (2.98) shows the assertion. \blacksquare

4. We look once again closer to $e_1(t)$: (2.95) shows

$$e_1(t) = \frac{t}{2} A^2 u(t). \quad (2.103)$$

Formally this looks like

$$e_1(t) = \frac{t}{2} u_{tt}(t), \text{ which would be}$$

the *formal solution* of the inhomogeneous parabolic problem

$$\begin{aligned} \text{a) } e_t + Ae &= \frac{1}{2} u_{tt} \\ \text{b) } e(0) &= 0, \end{aligned} \quad (2.104)$$

the problem we get, if we *formally apply* the technique of successive elimination of error terms due to HAIRER/LUBICH [19]. (Assuming an expansion of the local error). But since in general $\|e_1(t)\|_{L^2(\Omega)} \not\rightarrow 0$ as $t \downarrow 0$, (2.104) has not even a mild solution in the sense of PAZY [28].

5. Inequality (2.88.b) suggests that a result analogous to (2.90) holds also for the remainder term $E_{N+1}(t; \tau)$. We *conjecture* strongly that this is the case. In any case the question arises, which consistency has to be imposed to get a result like (2.90.b) for the remainder term. An answer, even though not the conjectured one, will be given in the next section in a more general context.

2.5 Asymptotic expansions II: General C_0 -semigroups

Here we tie up to the 5th remark at the end of the last section and the question asked therein:

Under which conditions we get rid of the singularity at $t = 0$ in (2.88.b)?

We are able to answer this in the context of general C_0 -semigroups $T(t)$ of contractions in a B -space X and generator $-A$; at the expense of higher consistency.

Theorem 2.13. *Given $T, \tau_0 > 0$, notation as before. If*

$$\varphi \in D_{A^{2N+2}} \quad (2.105)$$

we have the asymptotic expansion in X :

$$u_\tau(t) - u(t) = e_1(t)\tau + \dots + e_N(t)\tau^N + E_{N+1}(t; \tau)\tau^{N+1}. \quad (2.106)$$

The functions $e_k(t)$ are explicitly given by (2.89) as

$$e_k(t) = A^k P_k(tA)u(t), \quad 1 \leq k \leq N;$$

here the $P_k(\cdot)$ are the polynomials from (2.58). Furthermore we have

$$\begin{aligned} a) \quad & \|e_k(t)\| \leq Ct, \quad 1 \leq k \leq N \\ b) \quad & \|E_{N+1}(t; \tau)\| \leq Ct \end{aligned} \quad (2.107)$$

for every $t \in [0, T]$, $\tau \in [0, \tau_0]$. Here C depends only on T, τ_0, N and φ .

Proof. The proof bases on the second part of the key-lemma 2.10 and on corollary 2.3.

First we *define* the $e_k(\cdot)$ by formula (2.89) and observe, that they are well-defined because of (2.105) and that (2.107.a) holds by (2.58.c). We now show by induction, that the remainder term $E_{N+1}(t; \tau)$ fulfills (2.107.b).

Step 1. $N = 0$. We have $\varphi \in D_{A^2}$. By corollary 2.3, formula (2.18) and (2.12) and (2.7.d) we estimate

$$\begin{aligned}\|E_1(t; \tau)\tau\| &\leq \int_0^t \frac{s}{n} \|A^2\varphi\| ds \\ &= \frac{t^2}{2} \frac{1}{n} \|A^2\varphi\| \\ &= \frac{1}{2} t \tau \|A^2\varphi\| .\end{aligned}$$

Therefore

$$\|E_1(t; \tau)\| \leq \frac{1}{2} t \|A^2\varphi\| , \quad (2.108)$$

that is (2.107.b) for $N = 0$.

Step 2. Assume the assertion to be true for $E_N(\cdot, \cdot)$ for all possible initial data $\psi \in D_{A^{2N}}$. Now we have

$$E_{N+1}(t; \tau)\tau^{N+1} = u_\tau(t) - u(t) - e_1(t)\tau - \dots - e_N(t)\tau^N . \quad (2.109)$$

The difference $u_\tau(t) - u(t)$ we represent with the integral (2.18) of corollary 2.3. The corresponding representation of the other terms is done by

$$e_k(t)\tau^k = \int_0^t \frac{d}{ds} \left\{ A^k P_k(sA) \frac{s^k}{n^k} \right\} T(t)\varphi ds , \quad 1 \leq k \leq N . \quad (2.110)$$

(remember $\tau = t/n$) This exists because of (2.105). Thus we get for (2.109), performing the differentiation in (2.110)

$$\begin{aligned}E_{N+1}(t; \tau)\tau^{N+1} &= \int_0^t \frac{s}{n} A^2 \left\{ \left(I + \frac{s}{n} A \right)^{-(n+1)} \right. \\ &\quad \left. - \sum_{k=1}^N \frac{A^{k-2} s^{k-2}}{n^{k-1}} \left[k P_k(sA) + s A P'_k(sA) \right] T(s) \right\} T(t-s)\varphi ds .\end{aligned} \quad (2.111)$$

(The term behind the Σ -sign has to be interpreted for $k = 1$ in such a way, that no negative powers of A and s occur, this is possible because of $P_k(0) = 0$ for $k \geq 1$).

The term in square brackets is just of the form, that we can apply the

recurrency formula (2.58.b) of the key-lemma:

$$\begin{aligned}
E_{N+1}(t; \tau) \tau^{N+1} &= \int_0^t \frac{s}{n} A^2 \left\{ \left(I + \frac{s}{n} A \right)^{-(n+1)} \right. \\
&\quad \left. - \sum_{k=1}^N \frac{A^{k-1} s^{k-1}}{n^{k-1}} \sum_{j=0}^{k-1} (-1)^{k-j-1} P_j(sA) T(s) \right\} T(t-s) \varphi ds \\
&= \int_0^t \frac{s}{n} T(t-s) \left(I + \frac{s}{n} A \right)^{-1} \left\{ \left(I + \frac{s}{n} A \right)^{-n} \right. \\
&\quad \left. - \sum_{k=1}^N \frac{A^{k-1} s^{k-1}}{n^{k-1}} \sum_{j=0}^{k-1} (-1)^{k-j-1} P_j(sA) T(s) \right. \\
&\quad \left. - \sum_{k=1}^N \frac{A^k s^k}{n^k} \sum_{j=0}^{k-1} (-1)^{k-j-1} P_j(sA) T(s) \right\} A^2 \varphi ds \\
&= \int_0^t \frac{s}{n} T(t-s) \left(I + \frac{s}{n} A \right)^{-1} \left\{ \left(I + \frac{s}{n} A \right)^{-n} \right. \\
&\quad \left. - \sum_{j=0}^{N-1} P_j(sA) \frac{A^j s^j}{n^j} \left[\sum_{k=j+1}^N \frac{A^{k-j-1} s^{k-j-1}}{n^{k-j-1}} (-1)^{k-j-1} \right. \right. \\
&\quad \left. \left. - \sum_{k=j+1}^N \frac{A^{k-j} s^{k-j}}{n^{k-j}} (-1)^{k-j} \right] T(s) \right\} A^2 \varphi ds
\end{aligned}$$

Observing the telescope-sum in square brackets, we get:

$$\begin{aligned}
E_{N+1}(t; \tau) \tau^{N+1} &= \int_0^t \frac{s}{n} T(t-s) \left(I + \frac{s}{n} A \right)^{-1} \left\{ \left(I + \frac{s}{n} A \right)^{-n} \right. \\
&\quad \left. - \sum_{j=0}^{N-1} P_j(sA) \frac{A^j s^j}{n^j} T(s) \right. \\
&\quad \left. + \sum_{j=0}^{N-1} (-1)^{N-j} P_j(sA) \frac{A^N s^N}{n^N} T(s) \right\} A^2 \varphi ds. \tag{2.112}
\end{aligned}$$

If we denote by $\tilde{u}_\sigma(s)$, $\tilde{u}(s)$, $\tilde{e}_j(s)$, $\tilde{E}_N(s; \sigma)$ the terms corresponding to $u_\tau(t)$, $u(t)$, $e_j(t)$, $E_N(t; \tau)$ respectively for initial data $A^2 \varphi$ and time-step $\sigma = s/n$, we can rewrite (2.112) by (2.89)

$$\begin{aligned}
E_{N+1}(t; \tau) \tau^{N+1} &= \int_0^t \frac{s}{n} T(t-s) \left(I + \frac{s}{n} A \right)^{-1} \left\{ \tilde{u}_\sigma(s) - \tilde{u}(s) \right. \\
&\quad \left. - \tilde{e}_1(s) \sigma - \dots - \tilde{e}_{N-1}(s) \sigma^{N-1} \right. \\
&\quad \left. + \sum_{j=0}^{N-1} (-1)^{N-j} P_j(sA) \frac{A^N s^N}{n^N} T(s) A^2 \varphi \right\} ds \\
&= \int_0^t \frac{s}{n} T(t-s) \left(I + \frac{s}{n} A \right)^{-1} \left\{ \tilde{E}_N(s; \sigma) \sigma^N \right. \\
&\quad \left. + \sum_{j=0}^{N-1} (-1)^{N-j} P_j(sA) A^N \sigma^N T(s) A^2 \varphi \right\} ds. \tag{2.113}
\end{aligned}$$

By the assumption of the beginning of this step and (2.12), (2.7.d) we can estimate

$$\begin{aligned}
\|E_{N+1}(t; \tau) \tau^{N+1}\| &\leq C_1 \int_0^t \frac{s}{n} \sigma^N ds \\
&= C_1 \int_0^t \frac{s^{N+1}}{n^{N+1}} ds \\
&= C_2 \frac{t^{N+2}}{n^{N+1}} \\
&= C_2 t \tau^{N+1} .
\end{aligned} \tag{2.114}$$

Thus

$$\|E_{N+1}(t; \tau)\| \leq C_2 t , \tag{2.115}$$

that is (2.107.b) for $N + 1$. ■

Remarks. 1. $T(t)$ being a C_0 -semigroup of *contractions* is not essential here. In general

$$\|T(t)\| \leq M e^{\beta t} \tag{2.116}$$

holds for some $M > 0$, $\beta \in \mathbb{R}$ (cf. [35]). Instead of (2.9) now holds (theorem 3.13 of [35])

$$\|R(\lambda; -A)^n\| \leq \frac{M}{(\lambda - \beta)^n} \text{ for } \lambda > \beta . \tag{2.117}$$

That means

$$\|(I + \tau A)^{-n}\| \leq \frac{M}{(1 - \tau \beta)^n} \text{ for } \tau < 1/\beta , \text{ if } \beta > 0 . \tag{2.118}$$

Now the proof of theorem 2.13 is exactly the same by using (2.116) and (2.118) instead of (2.7.d) and (2.12), keeping in mind (2.72) for the right hand side of (2.118), *but* we get the additional *time-step bound*

$$\tau < 1/\beta , \text{ in case } \beta > 0 . \tag{2.119}$$

2. The proof of theorem 2.13 does not render theorem 2.12 and it's proof superfluous: From theorem 2.12 and the succeeding remarks we learn the rôle of inconsistent data.

3. Presentation of the algorithm

3.1 The semi-discrete case and the time-step control mechanism

In this section we describe a semi-discrete algorithm for the solution of the parabolic problem (2.5):

We use the implicit Euler discretization *in time*, assuming the thus arising elliptic problems are solved exactly, and control time-step and order of the method by extrapolation following the ideas of DEUFLHARD [11] for ODE's.

The main purpose of this section will be to show, that the usual results 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: For a fixed basic time-step $T > 0$

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

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

$$\mathcal{F} = \{n_1, n_2, \dots\} . \quad (3.2)$$

Since in limit $u(T) = u_{\tau=0}(T)$, we extrapolate the values $(\mathcal{U}_{11}, \dots, \mathcal{U}_{k1})$ 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 $L^2(\Omega)$*

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

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

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

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

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

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

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

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

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

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. *Assume that the following asymptotic expansion holds in $L^2(\Omega)$:*

$$u_\tau(T) = u(T) + e_1(T)\tau + \dots + e_{k-1}(T)\tau^{k-1} + E_k(T; \tau)\tau^k \tag{3.8}$$

with $e_j(T)$, $E_k(T; \tau) \in L^2(\Omega)$ for $T > 0$, $\tau \in [0, \tau_0]$. Furthermore we assume

$$\|E_k(T; \tau)\| \leq CT^{-\gamma_k} \text{ uniformly in } \tau \in [0, \tau_0]. \tag{3.9}$$

Conditions for (3.8) and (3.9) to hold have been derived in sections 2.4 and 2.5, showing

$$\gamma_k \in [-1, k]. \tag{3.10}$$

This yields

$$\varepsilon_{jk} := \|u(T) - \mathcal{U}_{jk}\|_{L^2(\Omega)} \leq \gamma_{jk} T^{k-\gamma_k} \tag{3.11}$$

Proof. We follow [20], proof of theorem 9.1. We consider the Vandermonde matrix

$$A_{jk} := \begin{bmatrix} 1 & \frac{1}{n_j} & \dots & \frac{1}{n_j^{k-1}} \\ \vdots & \vdots & & \vdots \\ 1 & \frac{1}{n_{j-k+1}} & \dots & \frac{1}{n_{j-k+1}^{k-1}} \end{bmatrix}, \tag{3.12}$$

so we can write (3.3) and (3.8) in new form

$$\begin{bmatrix} \mathcal{U}_{j1} \\ \vdots \\ \mathcal{U}_{j-k+1,1} \end{bmatrix} = A_{jk} \begin{bmatrix} \mathcal{U}_{jk} \\ e_1 T \\ \vdots \\ e_{k-1} T^{k-1} \end{bmatrix} \text{ in } [L^2(\Omega)]^k. \tag{3.3'}$$

$$\begin{bmatrix} \mathcal{U}_{j1} \\ \vdots \\ \vdots \\ \mathcal{U}_{j-k+1,1} \end{bmatrix} = A_{jk} \begin{bmatrix} u(T) \\ e_1(T)T \\ \vdots \\ e_{k-1}(T)T^{k-1} \end{bmatrix} + \begin{bmatrix} E_k(T; \tau_j) \tau_j^k \\ \vdots \\ \vdots \\ E_k(T; \tau_{j-k+1}) \tau_{j-k+1}^k \end{bmatrix} \quad (3.8')$$

Since A is a *scalar matrix* (3.3') shows the already stated computability of the e_j 's in $L^2(\Omega)$. Subtracting of (3.8') from (3.3') yields

$$A_{jk} \begin{bmatrix} \mathcal{U}_{jk} - u(T) \\ \vdots \\ (e_{k-1} - e_{k-1}(T))T^{k-1} \end{bmatrix} = \begin{bmatrix} E_k(T; \tau_j) \tau_j^k \\ \vdots \\ E_k(T; \tau_{j-k+1}) \tau_{j-k+1}^k \end{bmatrix}. \quad (3.13)$$

Therefore, since A_{jk}^{-1} exists:

$$\|\mathcal{U}_{jk} - u(T)\|_{L^2(\Omega)} \leq \|A_{jk}^{-1}\|_{\infty} \max_{j \geq i \geq j-k+1} \|E_k(T; \tau_i) \tau_i^k\|_{L^2(\Omega)}. \quad (3.14)$$

Together with (3.9) this yields

$$\|\mathcal{U}_{jk} - u(T)\|_{L^2(\Omega)} \leq \gamma_{jk} T^{k-\gamma_k}. \quad (3.15)$$

■

Remarks.

1. Because of (3.14) the analysis of the γ_{jk} can be done as usual, yielding asymptotically

$$\gamma_{jk} \doteq [n_{j-k+1} \dots n_j]^{-1} C(\alpha) \cdot C, \quad (3.16)$$

where $C(\alpha)$ is the Toeplitz-constant associated with \mathcal{F} and C depends on the problem.

2. In our example from the end of section 2.4 we get for the inconsistent initial data

$$\varphi_{1/4} = \frac{\pi - x}{2}$$

that $\gamma_k = k - 1/4$. Therefore we have order 1/4 in the whole table.

3. Since one implicit Euler step increases the consistency ($\varphi \in D_{A^m} \Rightarrow u_\tau(\tau) \in D_{A^{m+1}}$) we have, by theorem 2.13, *after some basic time-steps* $\gamma_k = -1$, that means

$$\|\mathcal{U}_{jk} - u(T)\|_{L^2(\Omega)} \leq \gamma_{jk} T^{k+1}$$

the full and maximal order.

By these remarks and the fact that in general

γ_{jk} will be small for j, k large

we see, that the *assumption*

$$\varepsilon_{j,k+1} \ll \varepsilon_{jk} \quad (3.17)$$

is reasonable. As in DEUFLHARD [11] section 1.2 we are thus led to the *subdiagonal error criterion*

$$\varepsilon_{k+1,k} \doteq \|\mathcal{U}_{k+1,k} - \mathcal{U}_{k+1,k+1}\| =: [\varepsilon_{k+1,k}]_{\text{sd}} \quad (3.18)$$

as a reasonable estimator.

A single quantity in square brackets shall denote by now and in future a computable estimator for this quantity.

The basic time-step for achieving a prescribed tolerance TOL in line $j + 1$ of the extrapolation table will be given by

$$T_{j+1,j} := \left(\frac{\text{TOL}}{[\varepsilon_{j+1,j}]_{\text{sd}}} \right)^{1/(j-[\gamma_j])} T, \quad (3.19)$$

T the present basic time-step. The estimator $[\gamma_j]$ 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:

$$\begin{aligned} u^1 + \tau A u^1 &= u^0 + \tau f \\ u^1 &\in H_0^m(\Omega). \end{aligned} \quad (3.20)$$

Since we want to use extrapolation in $L^2(\Omega)$ we are interested in global approximations with controllable error. One natural possibility in view of irregular boundary geometries are finite element methods.

If we do that, we get instead of (3.6) the perturbed extrapolation table

$$\begin{array}{ccc} \mathcal{U}_{11} + \delta_{11} & \searrow & \\ \downarrow & & \ddots \\ \vdots & & \searrow \\ \mathcal{U}_{k1} + \delta_{k1} & \rightarrow \dots & \mathcal{U}_{kk} + \delta_{kk}, \end{array} \quad (3.21)$$

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:

$$\bar{\mathcal{U}}_{jk} := \mathcal{U}_{jk} + \delta_{jk} . \quad (3.22)$$

Since the problem-oriented time-step mechanism (3.19) is connected with the semi-discrete estimator $[\varepsilon_{q+1,q}]_{\text{sd}}$ we are naturally led to achieve two things:

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

$$[\varepsilon_{q+1,q}]_{\text{sd}} \leq [\varepsilon_{q+1,q}] .$$

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

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

This leads to the following concept: Assuming estimators $[\delta_{k+1,k}]$, $[\delta_{k+1,k+1}]$ (they will be constructed in the next section) we get from

$$\begin{aligned} [\varepsilon_{k+1,k}]_{\text{sd}} &\leq \|\bar{\mathcal{U}}_{k+1,k} - \bar{\mathcal{U}}_{k+1,k+1}\|_{L^2(\Omega)} \\ &+ \|\delta_{k+1,k}\|_{L^2(\Omega)} + \|\delta_{k+1,k+1}\|_{L^2(\Omega)} \end{aligned} \quad (3.23)$$

the *fully-discrete estimator*

$$[\varepsilon_{k+1,k}] := \|\bar{\mathcal{U}}_{k+1,k} - \bar{\mathcal{U}}_{k+1,k+1}\|_{L^2(\Omega)} + [\delta_{k+1,k}] + [\delta_{k+1,k+1}] , \quad (3.24)$$

a completely computable quantity. The error of the available approximation $\bar{\mathcal{U}}_{k+1,k+1}$ is now estimated by

$$\|\bar{\mathcal{U}}_{k+1,k+1} - u(T)\|_{L^2(\Omega)} \leq [\delta_{k+1,k+1}] + [\varepsilon_{k+1,k}] , \quad (3.25)$$

so it is reasonable to ask for

$$\begin{aligned} \text{a) } [\delta_{k+1,k}], [\delta_{k+1,k+1}] &\leq \text{TOL}/4 \\ \text{b) } [\varepsilon_{k+1,k}] &\leq \text{TOL} \end{aligned} \quad (3.26)$$

in connection with the replacement of (3.19) by

$$T_{j+1,j} = \left(\frac{\text{TOL}}{[\varepsilon_{j+1,j}]} \right)^{1/(j-[\gamma_j])} T . \quad (3.19')$$

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

$$\delta_{jk} = \sum_{i=j-k+1}^j \beta_{jk}^i \delta_{i1} , \quad (3.27)$$

where the coefficients β_{jk}^i only depend on the chosen subdividing sequence \mathcal{F} .

Thus we can define

$$[\delta_{jk}] := \sum_{i=j-k+1}^j |\beta_{jk}^i| [\delta_{i1}] \quad (3.28)$$

and requirement (3.26.a) can be replaced by

$$[\delta_{j1}] \leq \min_{1 \leq q \leq k, 1 \leq i=q, q+1 \leq k} \left\{ \frac{\text{TOL}}{4|\beta_{iq}^j|q} \right\} =: \alpha_j^k \text{TOL} . \quad (3.29)$$

The coefficients α_j^k can be computed once at the beginning, they also depend only on \mathcal{F} .

Second step: Required errors of the elliptic solver.

For building the extrapolation table up to row k , we have to compute the $\bar{\mathcal{U}}_{j1}$ with error not exceeding $\alpha_j^k \text{TOL}$ (see (3.29)). 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

$$\Delta_i = \bar{\Delta}_i + \pi \Delta_{i-1} . \quad (3.30)$$

The rôle of π can be controlled:

Lemma 3.2. *Making the general assumptions of chapter 2, we have*

$$\|\pi\| \leq 1 . \quad (3.31)$$

Proof. Instead of solving one implicit Euler step

$$(I + \tau A)u^1 = u^0 + \tau f \quad (3.32)$$

we have an error Δ additional to u^0 producing an error $\pi\Delta$ additional to u^1

$$(I + \tau A)(u^1 + \pi\Delta) = u^0 + \Delta + \tau f . \quad (3.33)$$

The difference leads

$$\begin{aligned}\pi &= (I + \tau A)^{-1} \\ &= \frac{1}{\tau} R\left(\frac{1}{\tau}; -A\right) .\end{aligned}\tag{3.34}$$

By the Hille–Yosida theorem 2.1 and lemma 2.4 we therefore know the result:

$$\|\pi\| = \left\| \frac{1}{\tau} R\left(\frac{1}{\tau}; -A\right) \right\| \leq \frac{1/\tau}{1/\tau} = 1 .$$

■

If we use a reliable elliptic solver, which produces solutions with required accuracy ϵ delivering an estimation $[\bar{\Delta}]$, we have for constant requirement from (3.30) and (3.31)

$$\|\Delta_j\|_{L^2(\Omega)} \leq j\epsilon .\tag{3.35}$$

The requirement (3.29) thus yields

$$\epsilon := \frac{1}{j} \alpha_j^k \text{TOL}\tag{3.36}$$

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 (3.36).

Finally we have

$$[\delta_{j1}] := [\bar{\Delta}_1] + \dots + [\bar{\Delta}_j] .\tag{3.37}$$

The coefficients α_j^k/j are shown in figure 1 for $\mathcal{F} = \{1, 2, 3, \dots\}$, the harmonic sequence.

The coefficients α_j^k up to row 5				
$\alpha_1^1 = 1/4$				
$\alpha_1^2 = 1/8$	$\alpha_2^2 = 1/16$			
$\alpha_1^3 = 1/8$	$\alpha_2^3 = 1/48$	$\alpha_3^3 = 1/54$		
$\alpha_1^4 = 1/8$	$\alpha_2^4 = 1/64$	$\alpha_3^4 = 1/216$	$\alpha_4^4 = 3/512$	
$\alpha_1^5 = 1/8$	$\alpha_2^5 = 1/64$	$\alpha_3^5 = 1/405$	$\alpha_4^5 = 3/2560$	$\alpha_5^5 = 6/3125$

Table 3.1 The coefficients α_j^k until row 5

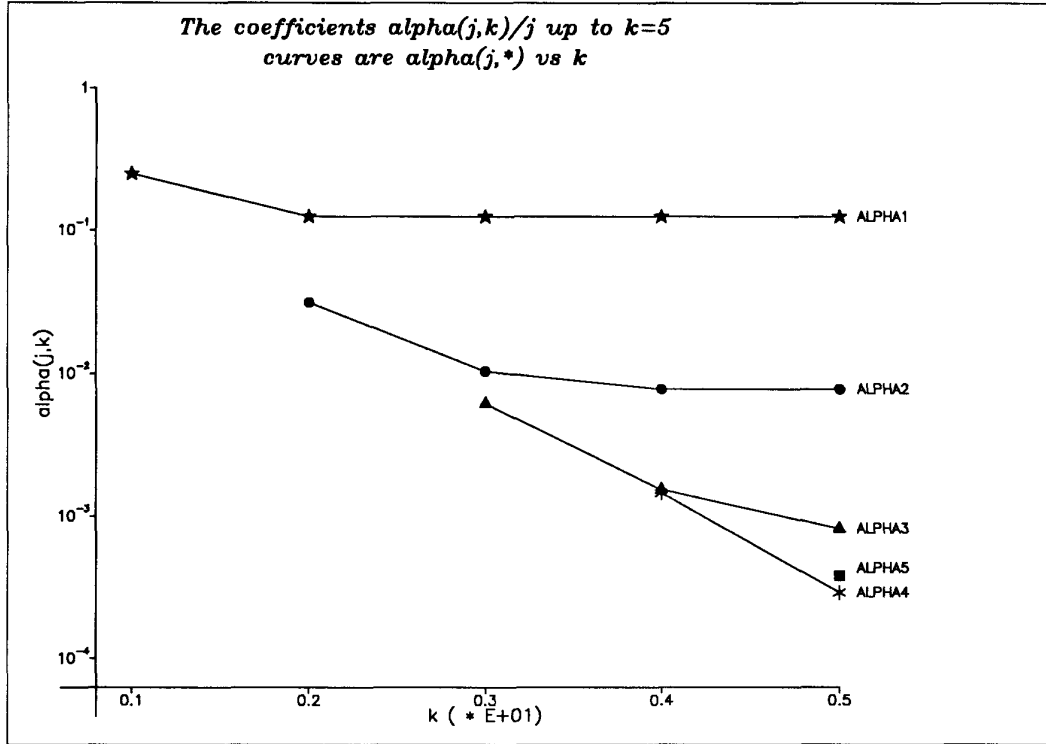


Figure 3.1 The coefficient α_j^k/j

3.4 The order control mechanism

As in DEUFLHARD [11] we control the “order”, that means here the row in the extrapolation table, in addition to the time-step. Relation (3.19') supplies us with step-size guesses $T_{j+1,j}$ for convergence of $\mathcal{U}_{j+1,j}$. As in [11] we define

$$\mathcal{W}_{j+1,j} := \frac{T}{T_{j+1,j}} A_{j+1} \quad (3.38)$$

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 (3.36).

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 (3.38) by

$$\mathcal{W}_{j+1,j}^k := \frac{T}{T_{j+1,j}} A_{j+1}^k, \quad (3.38')$$

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

$$\mathcal{W}_{q+1,q}^k = \min_{j=1,\dots,k-1} \mathcal{W}_{j+1,j}^k. \quad (3.39)$$

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 things

- convergence monitor
- order window
- device for possible increase of order greater than k .

(See DEUFLHARD [12])

This can be achieved by comparing the actual behavior in the table with an information-theoretic *standard model* derived in [11]. Since it depends on the amount of work A_{j+1}^k and the exponents of estimate (3.11) we have to perform several small changes. For sake of completeness we give all changed quantities and the new convergence monitor etc. – without giving any reasons, since one can use exactly the arguments given in [11].

Instead of the $\alpha_{ik}^{(q)}$ ([11] 3.7) we get

$$\alpha_{ij}^{(q,k)} := \text{TOL} \left\{ \frac{A_i^k - A_{i-j+1}^k}{A_{q+1}^k - A_1^k + 1} \right\} \quad (3.40)$$

and new

$$\alpha(j, q, k) := \left\{ \frac{\alpha_{j+1,j}^{(q,k)}}{\text{TOL}} \right\}^{1/(j - \lfloor \gamma_j \rfloor)} \quad \text{for } j \leq q. \quad (3.41)$$

1. Possible increase of order:

Situation: $q = k - 1 < k_{\max} - 1$.

Increase “order”, if

$$A_{q+1}^{q+1} \alpha(q, q+1, q+2) > A_{q+2}^{q+2} \quad (3.42)$$

and take the time-step

$$T_{\text{new}} := T_{q+1,q} \cdot \alpha(q, q+1, q+2) \quad (3.43)$$

2. Convergence monitor:

Let q be the expected optimal order. If for some previous entry of the table (specified in 3)

$$\frac{T}{T_{j+1,j}} > \begin{cases} \alpha(j, q+1, k), & \text{if } q < k-1 \\ \alpha(j, q+1, k+1), & \text{if } q = k-1 \end{cases} \quad (3.44)$$

then redo the step with

$$T_{\text{red}} := T_{j+1,j} \alpha(j, q, k) \cdot \sigma \quad (3.45)$$

where σ denotes a safety-factor $0 < \sigma < 1$.

3. Order-window:

Both the error criterion (3.26.b) and the monitoring condition (3.44) are only tested for j in the range

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

Moreover the next optimal “order” index is restricted to the condition

$$q_{\text{new}} \leq q+1$$

3.5 The consistency-estimator $[\gamma_k]$

The last missing point for a complete description of the algorithm *without specifying the elliptic solver* remains to be an estimator for $[\gamma_k]$. Relation (2.90) makes the assumption

$$\gamma_j = \max(-1, j - \gamma) \quad (3.46)$$

for some γ reflecting the consistency of the last approximation $\tilde{u}(t)$ plausible:

$$\tilde{u}(t) \in D_{A^\gamma}, \quad \gamma \text{ maximal}.$$

So we need an estimator $[\gamma]$ for γ .

We start assuming as much consistency we need, that means

$$[\gamma]_{\text{start}} = k_{\text{max}} + 1. \quad (3.47)$$

If the estimated $[\gamma_j]$ are seriously *too less* ($[\gamma]$ seriously too large) we will get far too large time-step guesses by (3.19') 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 $[\varepsilon_{k+1,k}]_{\text{old}}$ respectively $[\varepsilon_{k+1,k}]_{\text{new}}$ are available. By theorem 3.1 we have

$$\begin{aligned} \text{a)} \quad & [\varepsilon_{k+1,k}]_{\text{old}} \doteq CT_{\text{old}}^{k-\gamma_k} \\ \text{b)} \quad & [\varepsilon_{k+1,k}]_{\text{new}} \doteq CT_{\text{new}}^{k-\gamma_k}, \end{aligned} \tag{3.48}$$

that means

$$\gamma_k \approx k - \frac{\log\left(\frac{[\varepsilon_{k+1,k}]_{\text{old}}}{[\varepsilon_{k+1,k}]_{\text{new}}}\right)}{\log\left(\frac{T_{\text{old}}}{T_{\text{new}}}\right)} \tag{3.49}$$

leading to the reasonable

$$[\gamma]_{\text{new}} := \min\left([\gamma]_{\text{old}}, \frac{\log\left(\frac{[\varepsilon_{k+1,k}]_{\text{old}}}{[\varepsilon_{k+1,k}]_{\text{new}}}\right)}{\log\left(\frac{T_{\text{old}}}{T_{\text{new}}}\right)}\right) \tag{3.50}$$

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

$$[\gamma]_{\text{new}} := [\gamma]_{\text{old}}/2. \tag{3.51}$$

Now we set

$$[\gamma_j]_{\text{new}} := \max(-1, j - [\gamma]_{\text{new}}). \tag{3.52}$$

If we have no step-size reduction and redoing of a step, we have to consider an increase of γ (each implicit Euler step increases γ by one):

$$[\gamma] \longrightarrow \min(k_{\text{max}} + 1, [\gamma] + 1) =: [\gamma]_{\text{new}}, \tag{3.53}$$

and define again the $[\gamma_j]$ by (3.52).

4. Algorithmical details in the 1-D case

4.1 Required features of an elliptic solver

In the last chapter 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 $[\bar{\Delta}]$ (see the text before estimate (3.35))
2. The amount of work A_{j+1}^k as occurring in (3.38') 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

$$u + \tau Au = f \quad (4.1)$$

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 $\tau = 0$.

4.2 Theory for τ -independent elliptic error estimation and amount of work principle

For the rest of this chapter we restrict our attention to second order elliptic operators ($m = 1$) in *one space dimension*, using linear finite elements.

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

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

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

Ritz-projection $\longrightarrow L^2$ -projection.

Therefore residual-methods for error-estimation, which strongly use ellipticity bounds, are *ruled out*.

Here we use localization, 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. This will give a τ -independent error-estimator from below as we will see later on.

The rest of this section is devoted to make this idea precise and to show the relevant inequality. This will be rather long and technical.

The procedure uses norms which are extensions of those introduced by BABUŠKA/OSBORNE [4] for the purely elliptic case. Extensions, because we also have to consider the effect of the parameter τ .

Since we do not want to excess the technical effort here, we restrict the attention to the *heat equation* on $\Omega = I := [0, 1]$, thus considering the bilinear form

$$B_\tau(u, v) := \int_0^1 u(x)v(x)dx + \tau \int_0^1 u'(x)v'(x)dx. \quad (4.3)$$

Changes for general Sturm-Liouville operators A instead of $-d^2/dx^2$ will be indicated in the remarks at the end of this section.

Let $\tau > 0$. We take a subdivision Δ (mesh) of I :

$$\begin{aligned} \text{a) } \Delta &:= \{0 = x_0 < x_1 < \dots < x_n = 1\} \\ \text{b) } h_j &:= x_j - x_{j-1}; \quad I_j :=]x_{j-1}, x_j[, \quad j = 1, \dots, n \\ \text{c) } \delta_j &:= (h_j + h_{j+1})/2; \quad j = 1, \dots, n-1. \end{aligned} \quad (4.4)$$

Now we introduce mesh and parameter τ dependent norms:

For $u \in H_0^1(I)$ let (remember Sobolev-lemma!)

$$\|u\|_{0,\Delta}^2 := \|u\|_0^2 + \sum_{j=1}^{n-1} \delta_j |u(x_j)|^2 \quad (4.5)$$

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.

Further we define

$$\begin{aligned} \text{a) } H_{\Delta,\tau}^2 &:= \{u \in H_0^1(I) \mid u|_{I_j} \in H^2(I_j), \quad j = 1, \dots, n\} \\ \text{b) } \|u\|_{2,\Delta,\tau}^2 &:= \|u\|_0^2 + \tau^2 \left(\sum_{j=1}^n |u|_{2,I_j}^2 + \sum_{j=1}^{n-1} |\mathcal{J}u'(x_j)|^2 \delta_j^{-1} \right) \end{aligned} \quad (4.6)$$

where

$$\mathcal{J}u'(x_j) := u'(x_j^+) - u'(x_j^-) \quad (4.7)$$

denotes the jump of the first derivative at x_j (again Sobolev-lemma is used!) and $|\cdot|_{2,I_j}$ the seminorm $|u|_{2,I_j}^2 = \int_{I_j} |u''|^2 dx$.

$S_\Delta \subset H_0^1(I)$ shall denote the space of linear C_0 -finite elements on Δ . Note that $S_\Delta \subset H_\Delta^0$ and $S_\Delta \subset H_{\Delta,\tau}^2$.

Partial integration shows for $u \in H_0^1(I)$ and $v \in H_{\Delta,\tau}^2$ that

$$B_\tau(u, v) = \int_0^1 uv - \tau \sum_{j=1}^n \int_{I_j} uv'' - \tau \sum_{j=1}^{n-1} u(x_j) \mathcal{J}v'(x_j), \quad (4.8)$$

so we can extend $B_\tau(\cdot, \cdot)$ as a bilinear form on $H_\Delta^0 \times H_{\Delta,\tau}^2$, noting the following

Lemma 4.1. *For $u \in H_\Delta^0$ and $v \in H_{\Delta,\tau}^2$ we have*

$$|B_\tau(u, v)| \leq C_1 \|u\|_{0,\Delta} \|v\|_{2,\Delta,\tau} \quad (4.9)$$

with C_1 independent of Δ and τ .

Proof. Repeated application of Cauchy-Schwarz inequality to integrals and sums yields

$$\begin{aligned} |B_\tau(u, v)| &\leq \int_0^1 |uv| + \tau \sum_{j=1}^n \int_{I_j} |uv''| + \tau \sum_{j=1}^{n-1} |u(x_j)| |\mathcal{J}v'(x_j)| \\ &\leq \|u\|_0 \|v\|_0 + \sum_{j=1}^n \|u\|_{0,I_j} \tau \|v\|_{2,I_j} \\ &\quad + \sum_{j=1}^{n-1} |u(x_j)| \delta_j^{+1/2} \tau |\mathcal{J}v'(x_j)| \delta_j^{-1/2} \\ &\leq \left\{ \|u\|_0^2 + \sum_{j=1}^n \|u\|_{0,I_j}^2 + \sum_{j=1}^{n-1} |u(x_j)|^2 \delta_j \right\}^{1/2} \\ &\quad \left\{ \|v\|_0^2 + \tau^2 \left(\sum_{j=1}^n |v|_{2,I_j}^2 + \sum_{j=1}^{n-1} |\mathcal{J}v'(x_j)|^2 \delta_j^{-1} \right) \right\}^{1/2} \\ &\leq \sqrt{2} \|u\|_{0,\Delta} \|v\|_{2,\Delta,\tau}. \end{aligned}$$

■

Lemma 4.2. *There exists a constant $C_2 > 0$, independent of Δ and τ , such that*

$$\sup_{v \in S_\Delta, v \neq 0} \frac{|B_\tau(u, v)|}{\|v\|_{2,\Delta,\tau}} \geq C_2 \|u\|_{0,\Delta} \quad (4.10)$$

for all $u \in S_\Delta$.

Proof. For a given $u \in S_\Delta$ let $v \in S_\Delta$ be the solution of

$$B_\tau(\varphi, v) = (u, \varphi)_{L^2} \text{ for all } \varphi \in S_\Delta . \quad (4.11)$$

Therefore

$$B_\tau(u, v) = \|u\|_0^2 . \quad (4.12)$$

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

$$k_1 \|u\|_0 \leq \|u\|_{0,\Delta} \leq k_2 \|u\|_0 \text{ for all } u \in S_\Delta , \quad (4.13)$$

k_1, k_2 positive constants independent of Δ . This is essentially relation (4.3.c) of [4], but can be shown in our case, linear elements, by direct computation. Thus we get from (4.12)

$$|B_\tau(u, v)| \geq C \|u\|_{0,\Delta}^2 , \quad (4.14)$$

$C > 0$ independent of Δ and τ .

Next we have to estimate $\|u\|_{2,\Delta,\tau}$. For that purpose we look closer at (4.11):

$$\begin{aligned} \int_0^1 u \varphi &= \int_0^1 (\tau v' \varphi' + v \varphi) \\ &= \sum_{j=1}^n \int_{I_j} v \varphi - \tau \sum_{j=1}^{n-1} \mathcal{J} v'(x_j) \varphi(x_j) , \quad \varphi \in S_\Delta , \end{aligned} \quad (4.15)$$

since $v'' \equiv 0$ on I_j .

Let $\{\varphi_k\} \subset S_\Delta$ be the nodal basis

$$\varphi_k(x_j) = \delta_{kj} \quad k = 1, \dots, n-1 . \quad (4.16)$$

Equation (4.15) gives by inserting $\varphi = \varphi_k$:

$$\mathcal{J} v'(x_k) = \frac{1}{\tau} \left(\int_0^1 v \varphi_k - \int_0^1 u \varphi_k \right) . \quad (4.17)$$

Now we have

$$\begin{aligned} \left| \int_0^1 (v - u) \varphi_k \right| &\leq \int_{x_{k-1}}^{x_{k+1}} |v - u| |\varphi_k| \\ &\leq \left\{ \int_{x_{k-1}}^{x_{k+1}} |v - u|^2 \right\}^{1/2} \left\{ \int_{x_{k-1}}^{x_{k+1}} \varphi_k^2 \right\}^{1/2} \\ &\leq \|v - u\|_{0,I_k \cup I_{k+1}} \sqrt{2\delta_k} , \end{aligned}$$

therefore

$$\begin{aligned} \tau^2 \sum_{j=1}^{n-1} |\mathcal{J} v'(x_j)|^2 \delta_j^{-1} &\leq 2 \sum_{j=1}^{n-1} \|v - u\|_{0,I_j \cup I_{j+1}}^2 \\ &\leq 4 \|v - u\|_0^2 . \end{aligned} \quad (4.18)$$

Hence we get, remembering $v'' = 0$ on each I_j ,

$$\begin{aligned}\|v\|_{2,\Delta,\tau}^2 &= \|v\|_0^2 + \tau^2 \sum_{j=1}^{n-1} |\mathcal{J}v'(x_j)|^2 \delta_j^{-1} \\ &\leq \|v\|_0^2 + 4\|v - u\|_0^2 \\ &\leq 9\|v\|_0^2 + 8\|u\|_0^2.\end{aligned}\tag{4.19}$$

But inserting $\varphi = v$ into (4.11) yields

$$\|v\|_0^2 \leq B_\tau(v, v) = (u, v)_{L^2} \leq \|u\|_0 \|v\|_0,$$

thus

$$\|v\|_0 \leq \|u\|_0, \tag{4.20}$$

and

$$\|v\|_{2,\Delta,\tau}^2 \leq 17\|u\|_0^2 \leq 17\|u\|_{0,\Delta}^2. \tag{4.21}$$

This and relation (4.14) give the assertion. \blacksquare

Before we consider the error estimation, we shall discuss some *consequences* of these lemmas:

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

$$\|u - u_\Delta\|_{0,\Delta} \leq C \inf_{\varphi \in S_\Delta} \|u - \varphi\|_{0,\Delta}, \tag{4.22}$$

C independent of Δ and τ . Note that for $\tau > 0$ and $f \in L^2$ we have $u \in H_0^1(I) \subset H_\Delta^0$. The relation (4.22) follows from general results given in BABUŠKA/AZIZ [2], using Lemma 4.1 and 4.2.

2. Now *adequate adaptive meshes* can be characterized as follows:

$$\inf_{\varphi \in S_\Delta} \|u - \varphi\|_{0,\Delta} \leq \frac{C}{n^2}, \tag{4.23}$$

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 (4.22)

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

As *heuristic arguments* for requirement (4.23) can serve section 6.b and 6.c of [4] and theorem 5.2 of [3], as well as numerical experience.

3. On those adequate meshes we get

$$\|u - u_\Delta\|_{0,\Delta} \leq \frac{C}{n^2}, \quad (4.24)$$

C fairly independent of Δ and τ .

This justifies our *basic amount of work principle*:

Adaptive solution of an elliptic problem from family (4.1) with accuracy ϵ needs

$$n = C/\sqrt{\epsilon} \quad (4.25)$$

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

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

$$\begin{aligned} \text{a) } w_j + \tau A w_j &= f \quad \text{on } I_j \\ \text{b) } w_j(x_{j-1}) &= u_\Delta(x_{j-1}), \quad w_j(x_j) = u_\Delta(x_j). \end{aligned} \quad (4.26)$$

Relation(4.26.b) means that

$$\bar{w}_j := w_j - u_\Delta \in H_0^1(I_j) \quad (4.27)$$

and a weak formulation of (4.26.a) is therefore with

$$e := u - u_\Delta \quad (4.28)$$

the equation

$$B_\tau(\bar{w}_j, v) = B_\tau(e, v) \text{ for every } v \in H_0^1(I_j). \quad (4.29)$$

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

$$B_\tau(\tilde{w}_j, \varphi) = B_\tau(e, \varphi) = (f, \varphi) - B_\tau(u_\Delta, \varphi) \text{ for all } \varphi \in S_j^Q. \quad (4.30)$$

These \tilde{w}_j are *computable*.

Our *computable local error estimator* is now

$$[\eta_j] := \|\tilde{w}_j\|_{0,I_j}, \quad j = 1, \dots, n \quad (4.31)$$

and the global one

$$[\eta] := \left(\sum_{j=1}^n [\eta_j]^2 \right)^{1/2}. \quad (4.32)$$

To show relations between $[\eta]$ and $\eta := \|e\|_{0,\Delta}$ we have to introduce *local* (semi-)norms: $j = 1, \dots, n$.

For $u \in H_\Delta^0$ set

$$\|u\|_{0,\Delta,I_j}^2 := \|u\|_{0,I_j}^2 + \frac{1}{2}h_j(u(x_{j-1})^2 + u(x_j)^2), \quad (4.33)$$

and for $v \in H^2(I_j)$

$$\|v\|_{2,I_j,\tau}^2 := \|v\|_{0,I_j}^2 + \tau^2 \left(|v|_{2,I_j}^2 + 2h_j^{-1}(|v'(x_{j-1})|^2 + |v'(x_j)|^2) \right). \quad (4.34)$$

Before we state our main result, we localize lemma 4.1 and 4.2.

Lemma 4.3. *For $u \in H_\Delta^0$ and $v \in H^2(I_j)$, extended to I by zero, we have*

$$|B_\tau(u, v)| \leq C_1 \|u\|_{0,\Delta,I_j} \|v\|_{2,I_j,\tau} \quad (4.35)$$

with the τ and Δ -independent constant $C_1 = \sqrt{2}$ of lemma 4.1.

Proof. The same as for lemma 4.1. ■

Lemma 4.4. *There exists a constant $C_3 > 0$ independent of Δ and τ , such that*

$$\sup_{v \in S_j^Q, v \neq 0} \frac{|B_\tau(u, v)|}{\|v\|_{2,I_j,\tau}} \geq C_3 \|u\|_{0,I_j}$$

for all $u \in S_j^Q$.

Proof. First we observe that $v \in S_j^Q$ can be written as follows

$$\begin{aligned} \text{a) } v(x) &= \alpha(v) \cdot \frac{4}{h_j^2} (x - x_{j-1})(x_j - x) \\ \text{b) } \alpha(v) &:= v \left(\frac{x_{j-1} + x_j}{2} \right). \end{aligned} \quad (4.36)$$

Inserting and direct computation gives for $v \in S_j^Q$

$$\begin{aligned} \text{a) } |v|_{2,I_j}^2 &= 2h_j^{-1} (|v'(x_{j-1})|^2 + |v'(x_j)|^2) = 64\alpha(v)^2/h_j^3, \\ \text{b) } |v|_{1,I_j}^2 &= \frac{16}{3}\alpha(v)^2/h_j \\ \text{c) } \|v\|_{0,I_j}^2 &= \frac{8}{15}\alpha(v)^2 h_j. \end{aligned} \quad (4.37)$$

Given $u \in S_j^Q$ we solve the problem: $v \in S_j^Q$

$$B_\tau(v, \varphi) = (u, \varphi)_{L^2(I_j)} \text{ for all } \varphi \in S_j^Q, \quad (4.38)$$

that means

$$\frac{8}{15}\alpha(v)h_j + \tau/h_j \frac{16}{3}\alpha(v) = \frac{8}{15}h_j\alpha(u) , \quad (4.39)$$

or

$$\alpha(v) = \frac{\alpha(u)}{1 + 10\tau/h_j^2} . \quad (4.40)$$

Thus we get

$$\begin{aligned} \|v\|_{2,I_j,\tau}^2 &= \left(\frac{8}{15}h_j + 128\tau^2/h_j^3 \right) \frac{\alpha(u)^2}{(1+10\tau/h_j^2)^2} \\ &= \frac{(h_j^2 + 240\tau^2/h_j^2)}{(h_j + 10\tau/h_j)^2} \|u\|_{0,I_j}^2 \end{aligned} \quad (4.41)$$

Set

$$\chi(s, t) := \frac{(s^2 + 240t^2)}{(s + 10t)^2} . \quad (4.42)$$

For fixed $t > 0$ we have

$$\begin{aligned} \text{a) } \lim_{s \rightarrow \infty} \chi(s, t) &= 1 \\ \text{b) } \chi(0, t) &= \frac{12}{5} \\ &\text{and one local extremum by } s_{\text{ex}} = 24t: \\ \text{c) } \chi(s_{\text{ex}}, t) &= \frac{12}{17} . \end{aligned} \quad (4.43)$$

Therefore

$$0 \leq \chi(s, t) \leq \frac{12}{5} \text{ for every } s \geq 0 , \ t \geq 0 \quad (4.44)$$

and thus

$$\|v\|_{2,I_j,\tau}^2 \leq \frac{12}{5} \|u\|_{0,I_j}^2 . \quad (4.45)$$

Inserting $\varphi = u$ into (4.38) gives

$$B_\tau(u, v) = \|u\|_{0,I_j}^2 , \quad (4.46)$$

therefore

$$\frac{|B_\tau(u, v)|}{\|v\|_{2,I_j,\tau}} \geq \frac{1}{6} \sqrt{15} \|u\|_{0,I_j} , \quad (4.47)$$

that implies our assertion. ■

Now we get our main result:

Theorem 4.5 *The following local and global estimates hold*

$$\begin{aligned} \text{a) } [\eta_j] &\leq K\eta_j := K\|u - u_\Delta\|_{0,\Delta,I_j} \\ \text{b) } [\eta] &\leq K\eta := K\|u - u_\Delta\|_{0,\Delta} , \end{aligned} \quad (4.48)$$

K a positive constant independent of Δ and τ .

Proof. By (4.30) we get using lemma 4.3 and 4.4:

$$\begin{aligned}
[\eta_j] = \|\tilde{w}_j\|_{0,I_j} &\leq \frac{1}{C_3} \sup_{v \in S_j^Q, v \neq 0} \frac{|B_\tau(\tilde{w}_j, v)|}{\|v\|_{2,I_j,\tau}} \\
&= \frac{1}{C_3} \sup_{v \in S_j^Q, v \neq 0} \frac{|B_\tau(e, v)|}{\|v\|_{2,I_j,\tau}} \\
&\leq \frac{C_1}{C_3} \|e\|_{0,\Delta,I_j} =: K\eta_j .
\end{aligned}$$

That gives (4.48.a). Relation (4.48.b) follows by taking the definition of $[\eta]$ and observing that

$$\|v\|_{0,\Delta}^2 = \sum_{j=1}^n \|v\|_{0,\Delta,I_j}^2 \text{ for all } v \in H_\Delta^0 . \quad (4.49)$$

■

Remarks.

1. In the case of our model-problem, the heat-equation, we have

$$K = \frac{2}{5}\sqrt{30} \lesssim 2.2 .$$

This is seen by relation (4.35) and (4.47) and the constants given therein.

2. All results hold (except the values of C_1 , C_2 , C_3 and K) for general Sturm–Liouville operators A . This can be seen with techniques used in [4] applied to our proof for the model-problem. Even L^p -versions are possible, see [4].

4.3 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:

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

The heuristic (4.23) asks for a nearly equidistributed error. In order to achieve that, we determine “cut” following BABUŠKA/RHEINBOLDT [5]: We use a simple heuristic prediction scheme to forecast what may happen to η_j if I_j is subdivided. *Locally* we may assume

$$\eta_j = c_j h_j^{\lambda_j} \text{ as } h_j \rightarrow 0 . \quad (4.50)$$

Suppose I_j was generated by subdividing I_j^{old} with local error η_j^{old} obeying (4.50). The η_j -value after dividing I_j will be thus approximately

$$\eta_j^{\text{new}} = \frac{\eta_j^2}{\eta_j^{\text{old}}} . \quad (4.51)$$

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 1-D 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 the following algorithm:

$$Mu = f \text{ on } \Delta \quad (4.52)$$

can be solved as follows

I. LR-Decomposition:

$$\begin{aligned} \text{a)} \quad r_2 &= M_{11}^2 + M_{22}^1 \\ \text{b)} \quad j &= 2, \dots, n-1 : \\ l_j &= M_{12}^j / r_j \\ r_{j+1} &= M_{11}^{j+1} + M_{22}^j - l_j M_{21}^j \end{aligned} \quad (4.53)$$

delivering to vectors

$$\begin{aligned} r &= (r_2, \dots, r_n) \\ l &= (l_2, \dots, l_{n-1}) \end{aligned}$$

which have to be stored.

II. Forward Substitution:

$$\begin{aligned} \text{a)} \quad w(x_1) &= f(x_1) \\ \text{b)} \quad j &= 2, \dots, n-1 : \\ w(x_j) &= f(x_j) - l_j w(x_{j-1}) \end{aligned} \quad (4.54)$$

giving a vector $w = (w(x_1), \dots, w(x_{n-1}))$.

III. Backward Substitution:

- a) $u(x_{n-1}) = w(x_{n-1})/r_n$
- b) $j = n - 2, \dots, 1 :$
 $u(x_j) = (w(x_j) - M_{21}^{j+1}w(x_{j+1}))/r_{j+1}$

giving the solution vector

$$u = (u(x_1), \dots, u(x_{n-1})) .$$

The storage of u may overwrite w .

This finishes our description for the elliptic solver.

4.4 Realization of extrapolation

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

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

In order to extrapolate we consider the *common mesh*

$$\Delta = \bigcup_{j=1}^k \Delta_j . \quad (4.55)$$

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

In 2-D case (4.55) 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

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

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 , $j = 1, \dots, k - 1$ 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 $\mathcal{U}_{11}, \dots, \mathcal{U}_{j+1,1}$ 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

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

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 1-D case. It yields the same Δ as (4.55).

4.5 Some remarks about 2-D difficulties

In principle we could try to use our τ -extrapolation algorithm for 2-D problems as well, the theory of chapter 2 and the description of chapter 3 do not depend on the dimension. But the use in connection with the *temporary version* of the multi-level elliptic solver KASKADE is for two reasons not possible:

1. The implemented error-estimator would not be τ -independent
2. The iterative linear solver would not work τ -independent

Motivation for 2:

τ “small”

the stiffness-matrix to B_τ behaves essentially like the mass-matrix: proper preconditioning would be the use of a *scaled nodal basis*, WATHEN [38].

τ “big”

the stiffness-matrix to B_τ behaves essentially like the stiffness matrix with respect to A : proper preconditioning is the use of *hierarchical basis* YSERENTANT [41], as used in KASKADE.

The preconditioning of B_τ for all τ is a still open question. For *quasi-uniform grids* exists a suggestion by YSERENTANT [42].

5. Realization of the algorithm

5.1 The program KASTIX, a short description

The program KASTIX (KASKade TIme dependent with eXtrapolation) is a realization of the algorithm described in chapters 2 – 4 for the 1-D case. The central 1-D elliptic solver, which plays an essential rôle as we know, is written analogous to the 2-D elliptic solver KASKADE, in its temporary version at the ZIB written by ROITZSCH [30] [31]. This analogy consists in the use of the same structuring, hiding of relevant parts and parts of the data-structures. This should simplify a future extension to the 2-D case using KASKADE as the elliptic solver.

Essentially new parts in the elliptic solver are the simultaneous handling of a family of grids, the error-estimator, the refinement strategy and the linear solver; parts which are subject to change in KASKADE when it will be used sometime for the 2-D case, see the problems mentioned at the end of chapter 4.

The temporary experimental batch version of KASTIX was written by the author in the language C (the source has the size of 64K) and was developed and tested on a SUN3 workstation.

The following modules reveal the main structure of KASTIX:

- Grid management (files: delete.c, gridbasic.c, gridinter.c, find.c, read-grid.c, refine.c)
- Assembling (assemble.c, problem.c)
- Solve (estimate.c, solve.c)
- Extrapolation (euler.c, extrapol.c, startconst.c)
- System (sunutil.c, msg.c)
- Graphics (writegraz.c)
- Main (kastix.c)

Grid management:

Routines for creating and deleting grids, refinement of elements, dynamic storage control, evaluating grid functions, interpolation on finer grids, combining of several grids to a common one.

Assembling:

Routines for the assembling of *local* stiffness- and mass-matrices, the problem management.

Solve:

Routines for the τ -independent error-estimation of the elliptic subproblem, marking of elements for refinement, the Gauss-elimination process of section 4.3.

Extrapolation:

Routines for computing the constants from section 3.4, performing implicit Euler-steps, extrapolation, time and order control.

Systems:

Memory routines, clock and message handling.

Graphics:

Interface to the ZIB-Graphic GRAZIL, which was used to make the plots of section 5.2.

Main:

Starting KASTIX with user-specified data.

Because even the source without comments would take too much place (70 pages), we regret to refer the reader to forthcoming manuals of later versions.

Remarks.

- In the assembling module we use two point Gauss-quadrature for the L^2 -product of a function with a linear form function and three point Gauss-quadrature for the product with a quadratic form function.
- In the current version the **problem** may have the following form

$$\begin{aligned} \text{a) } & u_t(t, x) - p(x)u_{xx}(t, x) - r(x)u_x(t, x) - q(x)u(t, x) = f(t, x) \\ & x \in [a, b], t \geq 0 \\ \text{b) } & u(0, x) = \varphi(x), \quad x \in [a, b] \\ \text{c) } & u(t, a) = g_1(t) \\ & u(t, b) = g_2(t) \end{aligned}$$

The convection term r should not “dominate”.

5.2 Numerical examples

Remark. In the following examples grids for time-sections will be shown. However, it should be noted, that these nodes are not a single grid to represent a stationary solution, but are the union of the grids in the extrapolation-table. So they also contain information about the time-development.

Example 1 This is the easiest example of simple heat conduction.

Problem:

$$\begin{aligned} \text{a)} \quad u_t(t, x) - u_{xx}(t, x) &= \begin{cases} 0.0 & \text{for } 0 \leq x \leq 1 \\ 3.0 & \text{for } 1 < x \leq 2 \end{cases} \\ \text{b)} \quad \varphi := u(0, x) &= \begin{cases} x & \text{for } 0 \leq x \leq 1 \\ 2 - x & \text{for } 1 \leq x \leq 2 \end{cases} \\ \text{c)} \quad u(t, 0) = u(t, 2) &= 0 \end{aligned}$$

The initial data φ are here consistent, that is $\varphi \in H_0^1([0, 2])$. But due to the Sobolev-lemma, $\varphi \notin H^2([0, 2])$; so we have $\varphi \notin D_A$, i.e. φ is non-smooth. But Fourier analysis shows

$$\varphi \in D_{A^\alpha} \text{ for all } \alpha < 3/4 .$$

We have computed the solution until $t = 5.0$ with a prescribed tolerance $\text{TOL} = 10^{-1}$.

step	time	L2-error	order	nodes
0	0.00e+00	0.00e+00	5	3
1	5.00e-05	1.42e-07	2	5
2	1.05e-01	3.17e-02	2	63
3	5.69e-01	1.65e-02	2	41
4	3.42e+00	6.87e-03	2	35
5	5.00e+00	7.13e-03	2	35

Table 5.1 Performance of KASTIX in example 1

The right hand side has been chosen so, that the stationary solution is linear for $0 \leq x \leq 1$ and quadratic for $1 \leq x \leq 2$. Figure 1 nicely reflects in the picture for $t = 5$ the quality of the error-estimation: only one interior point in the linear part, equidistant points in the quadratic part. A SUN4 workstation needed 1.67 seconds.

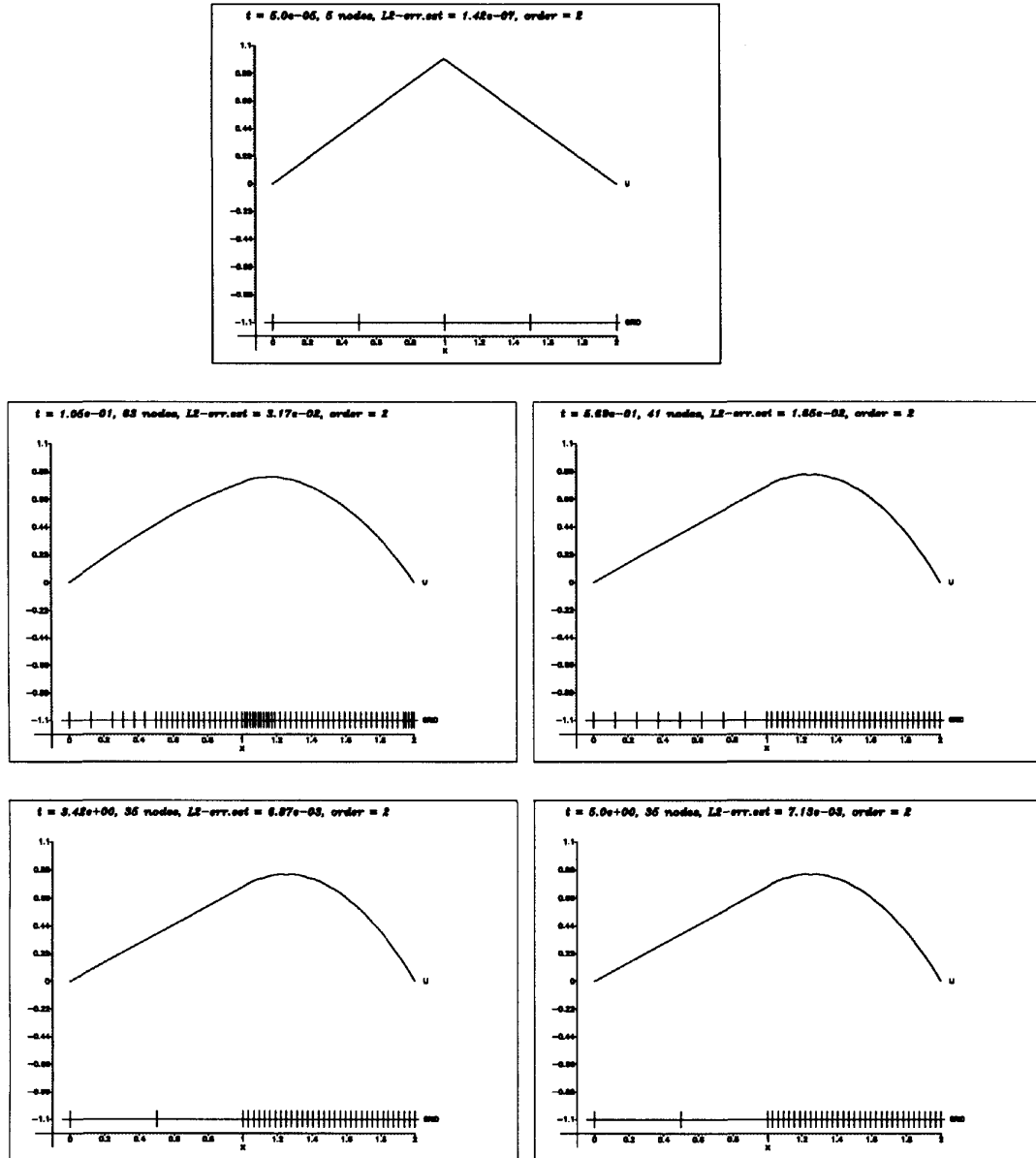


Figure 5.1 Time development for example 1

Example 2 This is an example for heat conduction with inconsistent initial data.

Problem: As in example 1, but

$$b') \quad \varphi := u(0, x) = \begin{cases} -1 & \text{for } 0 \leq x < 1 \\ +1 & \text{for } 1 < x \leq 2 \end{cases}$$

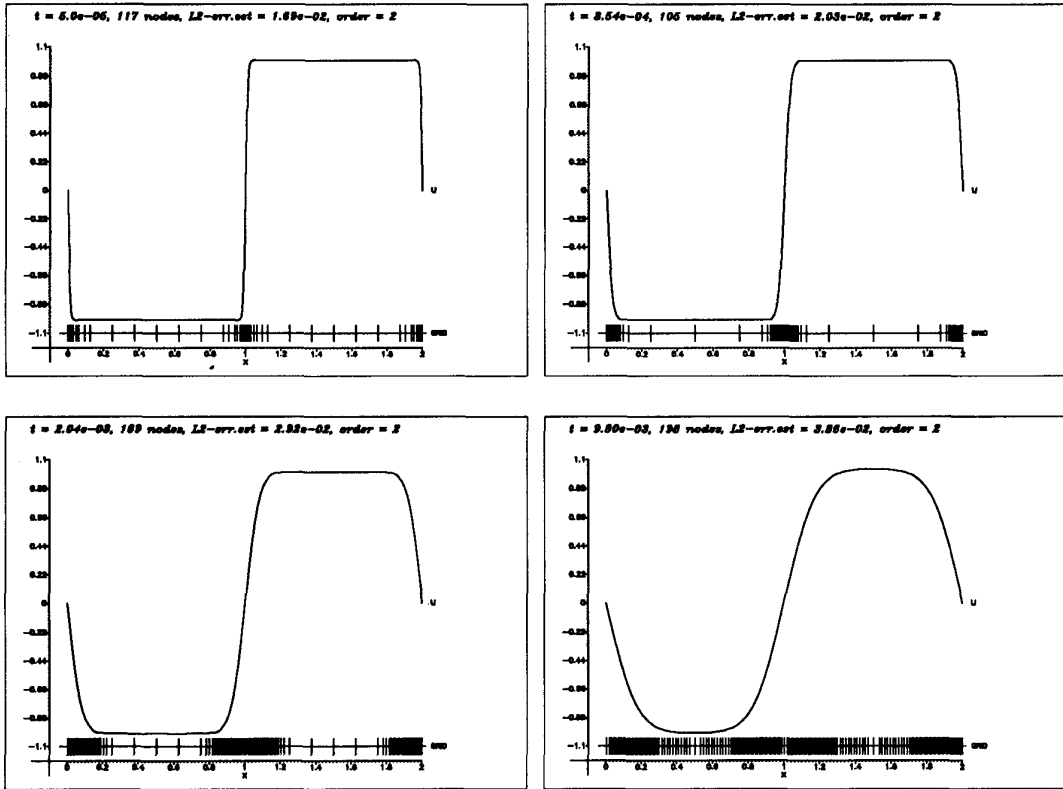
Here we have even $\varphi \notin H_0^1([0, 2])$, but again by Fourier analysis we get

$$\varphi \in D_{A^\alpha} \quad \text{for all } \alpha < 1/4.$$

Again the solution is computed until $t = 5.0$ with a prescribed tolerance $\text{TOL} = 10^{-1}$.

step	time	L2-error	order	nodes
0	0.00e+00	0.00e+00	5	3
1	5.00e-05	1.69e-02	2	117
2	3.54e-04	2.03e-02	2	105
3	2.04e-03	2.92e-02	2	169
4	9.80e-03	3.86e-02	2	198
5	4.10e-02	4.02e-02	2	128
6	1.64e-01	9.34e-02	2	123
7	4.81e-01	6.08e-02	2	48
8	1.49e+00	4.58e-02	2	44
9	5.00e+00	1.00e-02	2	35

Table 5.2 Performance of KASTIX in example 2



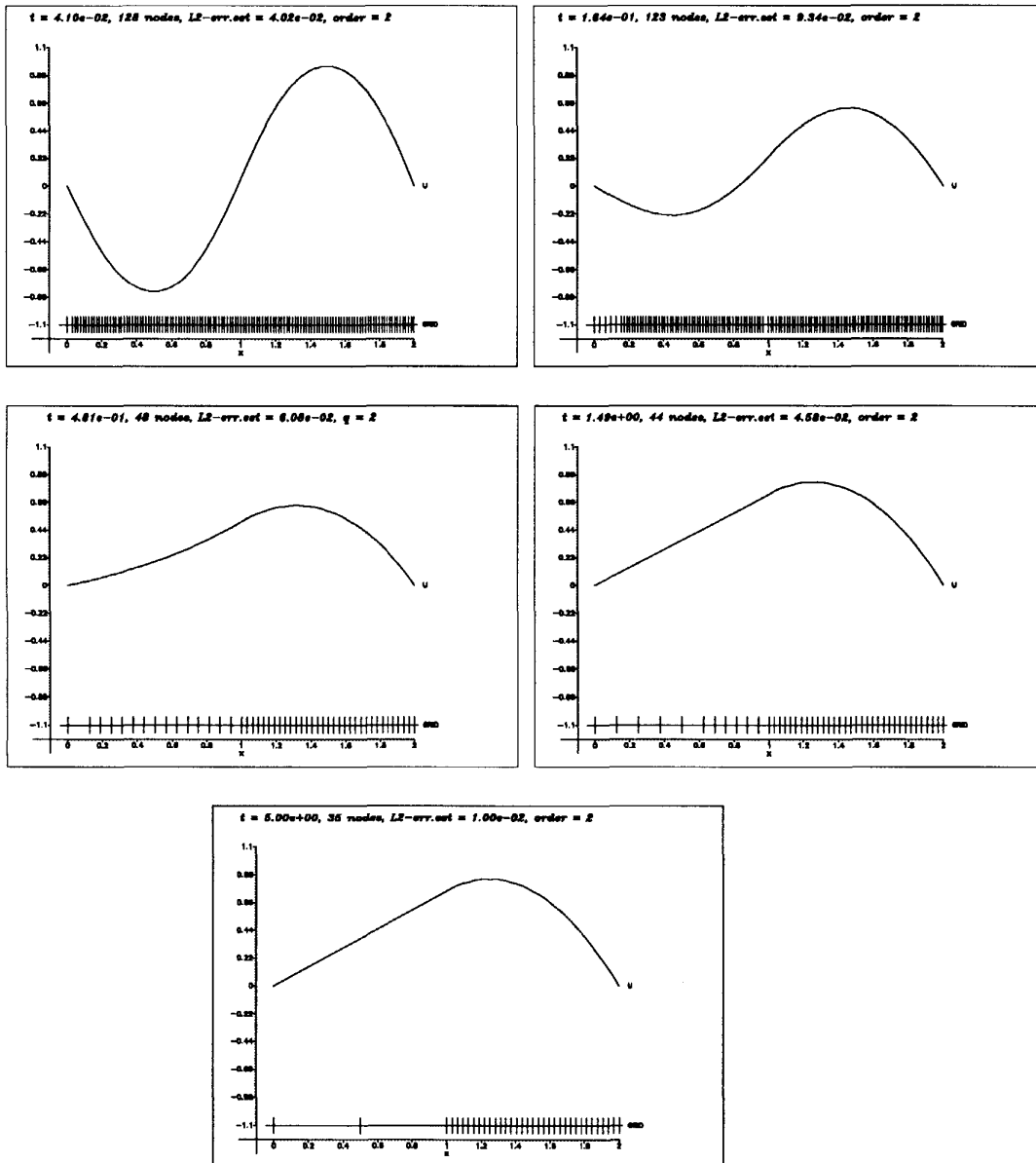


Figure 5.2 Time development for example 2

A SUN4 workstation needed 15.74 seconds.

Example 3 This is a 1-D version of example 9.2 from ERIKSSON/JOHNSON [16]. In this case we solve on $[0, 2]$ a homogeneous heat equation with the following “approximate δ -function” at $t = 0$ as φ

$$\varphi := u(0, x) = \frac{1}{\varepsilon} \exp(-x^2/\varepsilon)$$

with $\varepsilon = 1/250$ and prescribed tolerance $\text{TOL} = 0.5$. The exact solution is

$$u(t, x) = \frac{1}{\sqrt{\varepsilon}\sqrt{4t + \varepsilon}} \exp\left(-x^2/(4t + \varepsilon)\right) ,$$

which is the evolution of the Gauss-kernel.

This example is interesting because of the development from a sharp peak to the zero-solution, which is challenging to the time-step control mechanism. Also we have the opportunity to test the error estimation.

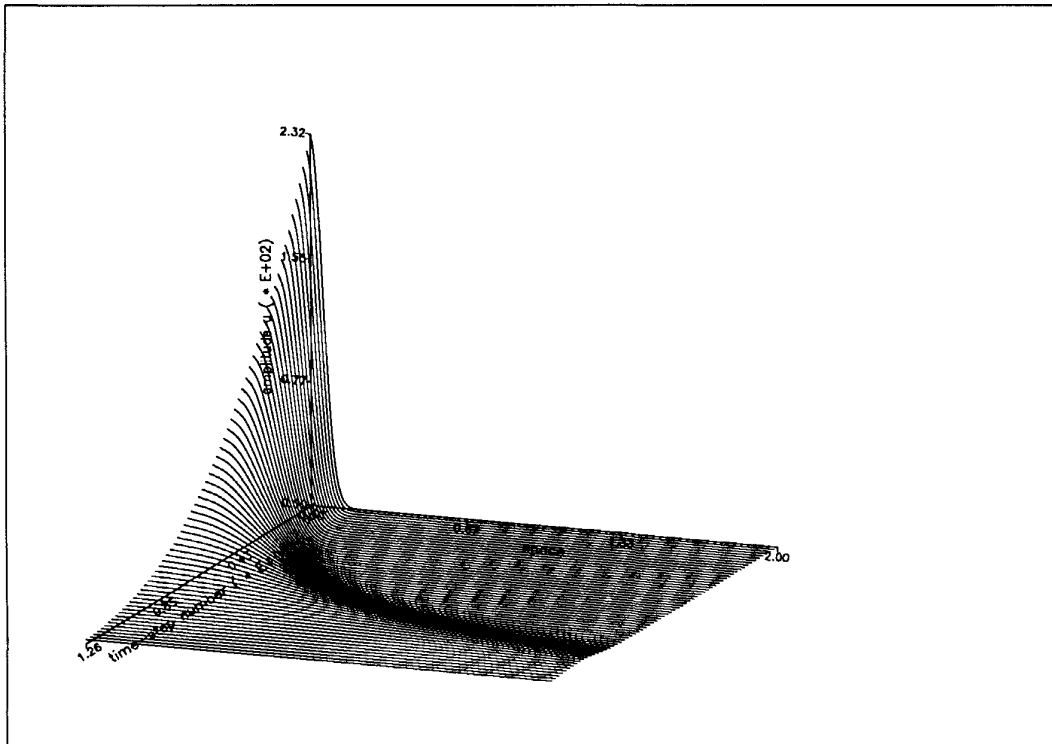


Figure 5.3 Surface plot of the computed solution, example 3

Here the time axis is chosen in logarithmic scale and is pointing towards the reader.

This representation nicely illustrates the fact that the time-step control mechanism reflects reasonable changes of the solution.

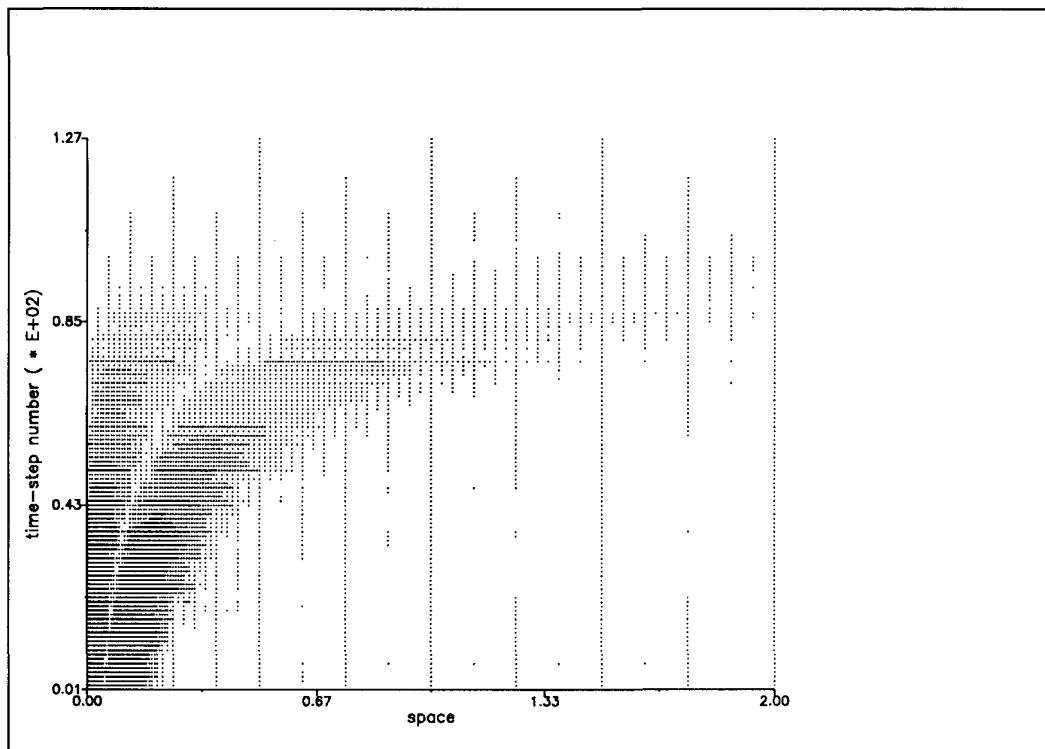


Figure 5.4 Evolution of nodes, example 3

Here the time axis is again in logarithmic scale.

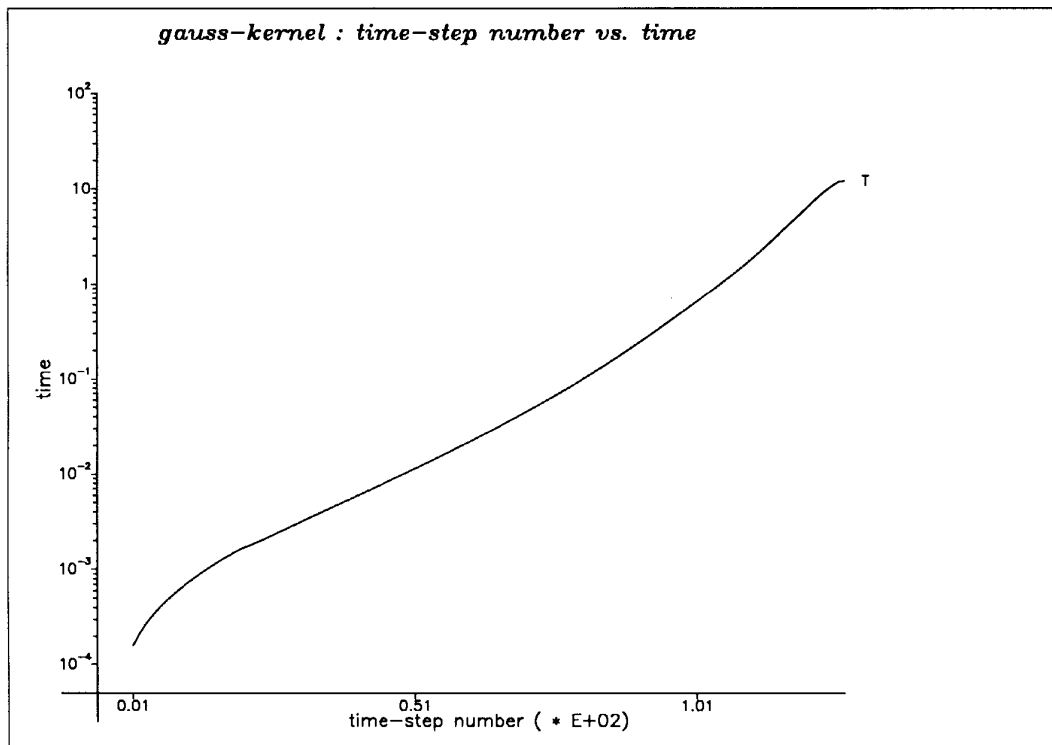


Figure 5.5 Time-step evolution in example 3

Here we can see the behavior of the time-step control.

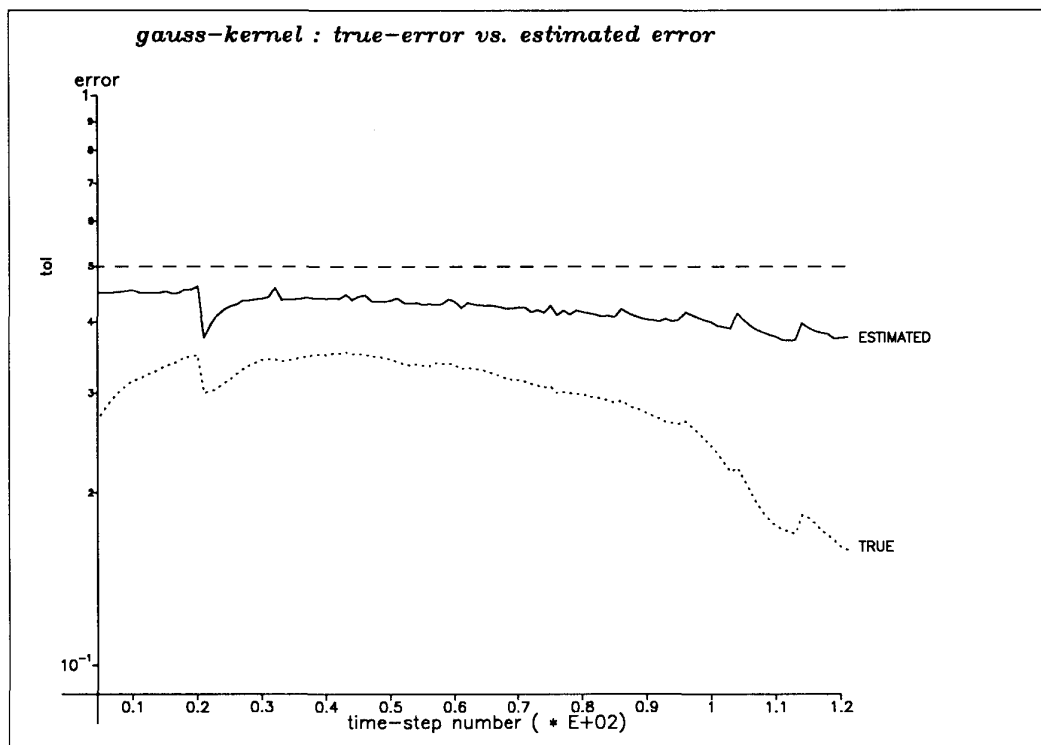


Figure 5.6 True error versus estimated error, example 3

Here we see the reliability of the error estimation.

A SUN4 workstation needed 132.69 seconds to compute the solution till $t = 12.0$.

Example 4 This is example 2 of BIETERMAN/BABUŠKA [7]. In this case we solve a heat equation with a small convection term with righthand-side, Dirichlet conditions and initial data chosen so, that the solution $u(t, x)$ is a “travelling wave”.

On $J = [0, 1]$ we solve for $0 \leq t \leq 0.9$ and $\text{TOL} = 2 \cdot 10^{-2}$

- a) $u_t - u_{xx} - u_x = -c^2 \tanh[c(x - t)] (1 - \tanh^2[c(x - t)])$
- b) $u(0, x) = (1.1 - \tanh[cx])/2$
- c) $u(t, 0) = (1.1 - \tanh[-ct])/2$
 $u(t, 1) = (1.1 - \tanh[c(1 - t)])/2 .$

The exact solution is

$$u(t, x) = (1.1 - \tanh[c(x - t)])/2 .$$

This example is interesting, because there should occur a “moving” of the condensing of nodes along the jump of the wave.

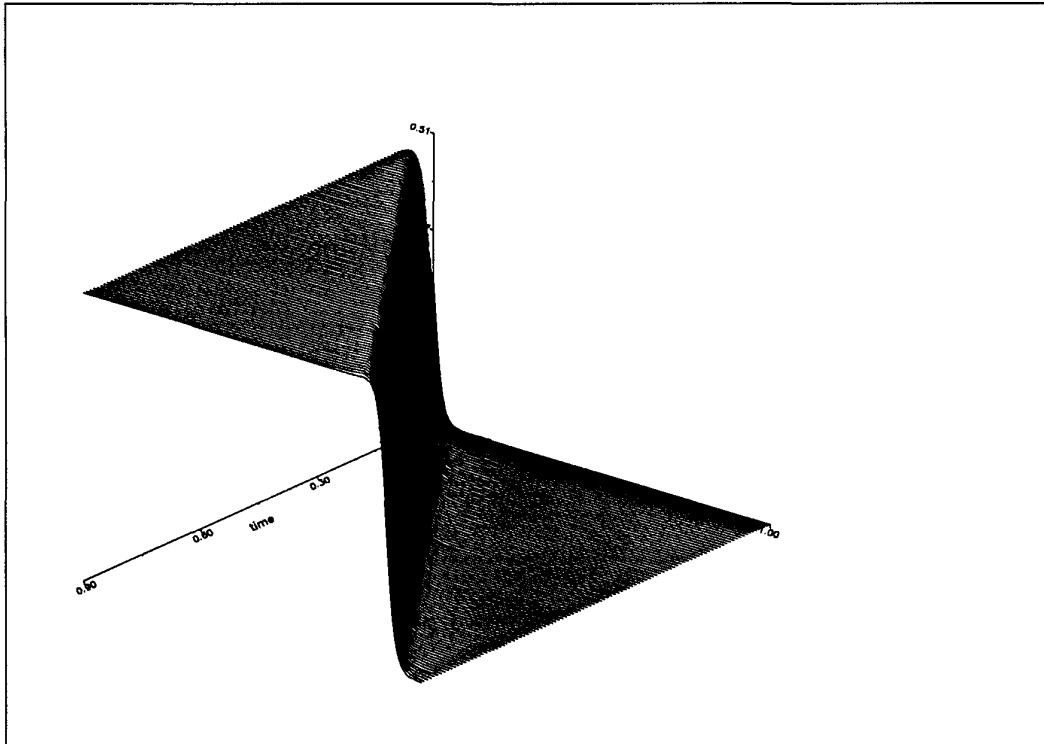


Figure 5.7 Surface plot of the travelling wave, example 4

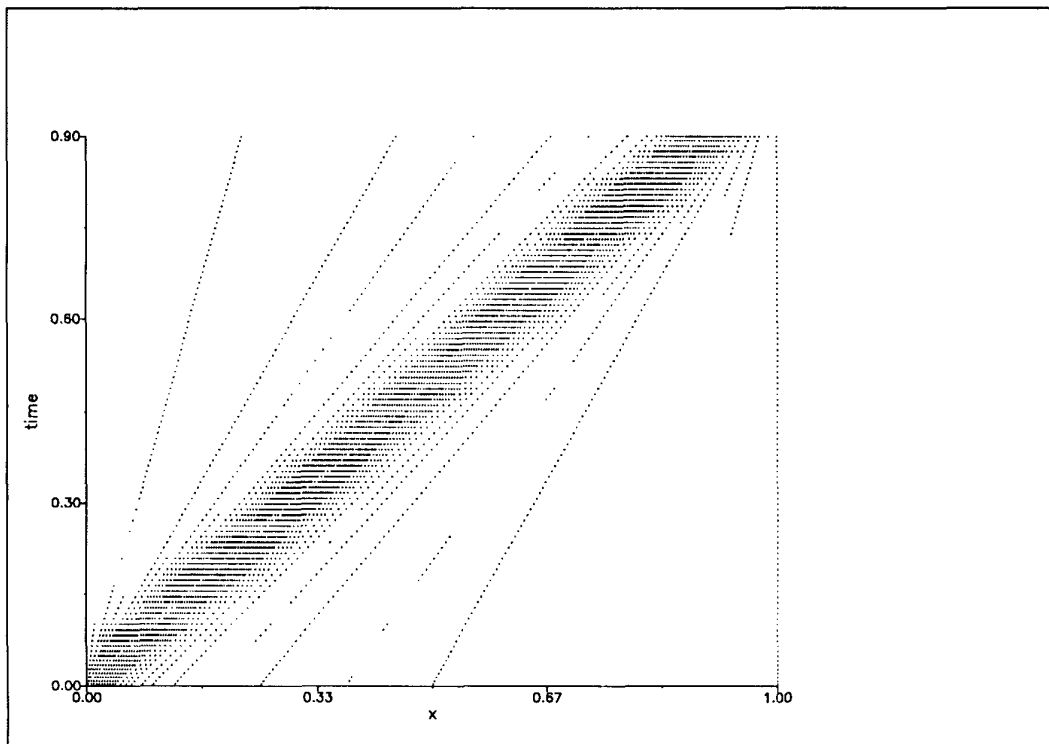


Figure 5.8 “Moving” of nodes, example 4

Remember that our algorithm realizes “moving nodes” without explicit moving of nodes.

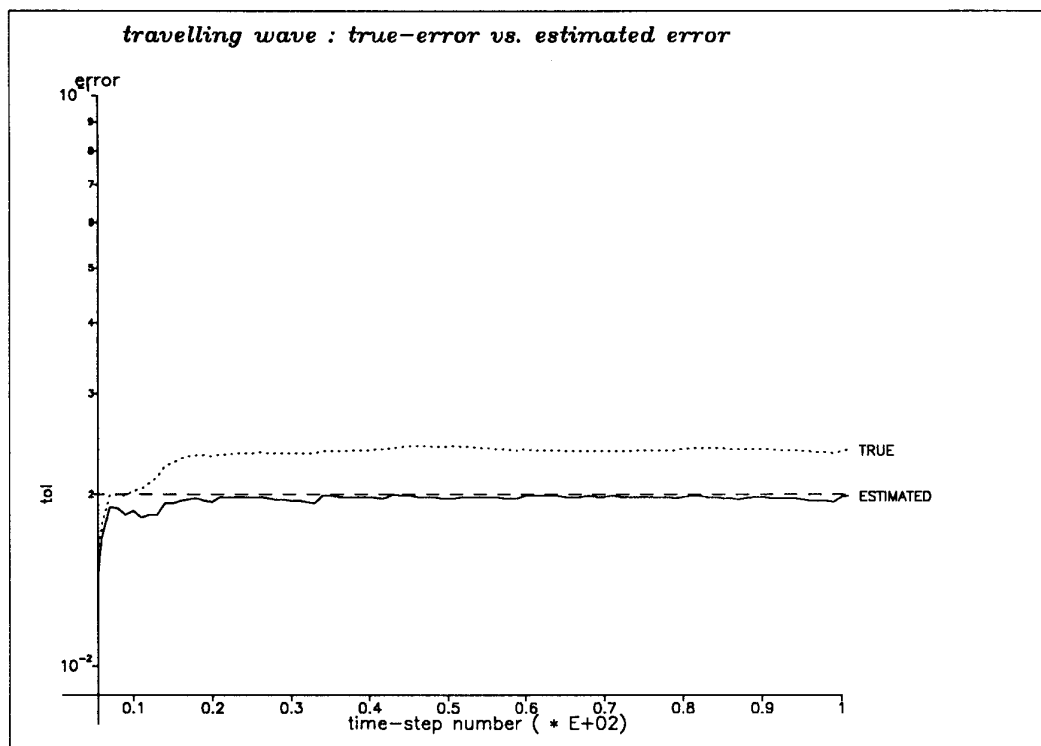


Figure 5.9 Behavior of error-estimation, example 4

A SUN4 workstation needed 94.10 seconds to compute the solution till $t = 0.9$.

Acknowledgements

First I would like to express my gratitude to my advisor, Prof. P. Deuffhard, for several reasons: He has given me the most interesting and deliberately open task for my diploma thesis to find an appropriate matching of time and space discretization for parabolic equations. Due to him I had the possibility of doing this work as a student worker at the Konrad-Zuse-Center (ZIB) in Berlin. At this most pulsating research place I had the opportunity to get contact to several researchers. Last but not least I thank for his most personal education in numerical mathematics and a lot of discussions.

I also thank Michael Wulkow for long and intensive discussions about matching problems. He was the first, who encouraged me to follow my final idea. He also read the manuscript and made several suggestions to clarify the representation.

For discussions I thank Prof. H. Wu, Dr. R. Hoppe, Dr. B. Fiedler and Dr. R. Kornhuber. Dr. Hoppe encouraged me to use semi-group theory and Dr. Fiedler drew my attention to reference [23].

I want to thank R. Roitzsch for his pretty version of KASKADE: From this source I learned the language C and in analogy I could write most parts of my 1-D elliptic solver. This has been invaluable for me, otherwise I would not have finished this work in such a short time.

Special thanks for Mrs. S. Wacker's careful typing of my manuscript.

References

- [1] Adams, R.A.: *Sobolev Spaces*. Academic Press, New York (1975).
- [2] Babuška, I., Aziz, A.K.: *Survey Lectures on the Mathematical Foundations of the Finite Element Method*. The Mathematical Foundation of the Finite Element Method, A.K. Aziz (ed.), Academic Press, New York, p. 3–362 (1972).
- [3] Babuška, I., Kellogg, R.B., Pitkäranta, J.: *Direct and Inverse Estimates for Finite Elements with Mesh Refinements*. Num. Math. 15, p. 447–471 (1979).
- [4] 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).
- [5] Babuška, I., Rheinboldt, W.C.: *Error Estimates for Adaptive Finite Element Computations*. SIAM J. Numer. Anal. 15, p. 736–754 (1978).
- [6] Bieterman, M., Babuška, I.: *The Finite Element Method for Parabolic Equations. I. A Posteriori Error Estimation*. Num. Math. 40, p. 339–371 (1982).
- [7] 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).
- [8] Ciarlet, P.G.: *The Finite Element Method for Elliptic Problems*. North Holland, Amsterdam (1978).
- [9] Dauge, M.: *Elliptic Boundary Value Problems on Corner Domains*. Lecture Notes in Mathematics, 1341, Springer, Berlin–Heidelberg–New York (1988).
- [10] Dautray, R., Lions, J.L.: *Mathematical Analysis and Numerical Methods for Science and Technology*. Vol. 2, Functional and Variational Methods, Springer, Berlin–Heidelberg–New York (1988).
- [11] Deuffhard, P.: *Order and Stepsize Control in Extrapolation Methods*. Num. Math. 41, p. 399–422 (1983).
- [12] Deuffhard, P.: *Recent Progress in Extrapolation Methods for Ordinary Differential Equations*. SIAM Review 27, p. 505–535 (1985).

- [13] Deuffhard, P.: *Numerik von Anfangswertmethoden für gewöhnliche Differentialgleichungen*. Technical Report TR89-2, ZIB (1989).
- [14] Deuffhard, 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).
- [15] Dupont, T.F.: *A Short Survey of Parabolic Galerkin Methods*. In: *The Mathematical Basis of Finite Element Methods*, D.F. Griffiths (ed.), Clarendon Press, Oxford (1984).
- [16] 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).
- [17] Friedman, A.: *Partial Differential Equations*. Robert E. Krieger Publishing Company, Huntington, New York (1976).
- [18] Gragg, W.B.: *Repeated Extrapolation to the Limit in the Numerical Solution of Ordinary Differential Equations*. SIAM J. Numer. Anal. 2, p. 384–403 (1965).
- [19] Hairer, E., Lubich, C.: *Asymptotic Expansions of the Global Error of Fixed-Stepsize Methods*. Num. Math. 45, p. 345–360 (1984).
- [20] Hairer, E., Nørsett, S.P., Wanner, G.: *Solving Ordinary Differential Equations I. Nonstiff Problems*. Springer, Berlin–Heidelberg–New York (1987).
- [21] Hille, E.: *Functional Analysis and Semi-Groups*. Amer. Math. Soc. Colloq. Publ., 31, New York (1948).
- [22] Kačur, J.: *Application of Rothe's Method to Nonlinear Evolution Equations*. Mat. Čas. 25, p. 63–81 (1975).
- [23] Kačur, J.: *Method of Rothe in Evolution Equations*. Teubner, Leipzig (1985).
- [24] Lions, J.L., Magenes, E.: *Non-Homogeneous Boundary Value Problems and Applications*. Springer, Berlin–Heidelberg–New York (1972).
- [25] McBride, A.C.: *Semigroups of Linear Operators: An Introduction*. Pitman Research Notes in Mathematics Series 156, Longman, Harlow (1987).
- [26] Miller, K., Miller R.N.: *Moving Finite Elements*. I., SIAM J. Numer. Anal. 18, p. 1019–1032 (1981).

- [27] Nečas, J.: *Application of Rothe's Method to Abstract Parabolic Equations*. Czech. Math. J. 24, p. 496–500 (1974).
- [28] Pazy, A.: *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer, Berlin–Heidelberg–New York (1983).
- [29] Rektorys, K.: *The Method of Discretization in Time and Partial Differential Equations*. D. Reidel Publishing Company, Dordrecht (1982).
- [30] Roitzsch, R.: *KASKADE User's Manual*. Technical Report TR89–4, ZIB (1989).
- [31] Roitzsch, R.: *KASKADE Programmer's Manual*. Technical Report TR89–5, ZIB (1989).
- [32] Rothe, E.: *Zweidimensionale parabolische Randwertaufgabe als Grenzfall eindimensionaler Randwertaufgaben*. Math. Annalen 102, p. 650–670 (1930).
- [33] Stetter, H.J.: *Asymptotic Expansions for the Error of Discretization Algorithms for Non-Linear Functional Equations*. Num. Math. 7, p. 18–31 (1965).
- [34] Stetter, H.J.: *Analysis of Discretization Methods for Ordinary Differential Equations*. Springer, Berlin–Heidelberg–New York (1973).
- [35] Tanabe, H.: *Equations of Evolution*. Pitman, London (1979).
- [36] Thomée, V.: *Galerkin Finite Element Methods for Parabolic Problems*. Lecture Notes in Mathematics, 1054, Springer, Berlin–Heidelberg–New York (1984).
- [37] Triebel, H.: *Höhere Analysis*. Verlag Harri Deutsch, Thun-Frankfurt (Main) (1980).
- [38] Wathen, A.J.: *Realistic Eigenvalue Bounds for the Galerkin Mass Matrix*. IMA J. Numer. Anal. 7, p. 449–457 (1987).
- [39] Wloka, J.: *Partielle Differentialgleichungen*. Teubner, Stuttgart (1982).
- [40] Yosida, K.: *Functional Analysis*. Springer, Berlin–Heidelberg–New York (1978).
- [41] Yserentant, H.: *On the Multilevel Splitting of Finite Element Spaces*. Num. Math. 49, p. 379–412 (1986).
- [42] Yserentant, H.: *Hierarchical Bases in the Numerical Solution of Parabolic Problems*. Large Scale Scientific Computing, Deuffhard, P., Engquist, B. (eds.), Birkhäuser, Boston, p. 22–37 (1987).

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)

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.*