SIAM J. NUMER. ANAL.
Vol. 28, No. 1, pp. 43-77, February 1991

003

ADAPTIVE FINITE ELEMENT METHODS FOR PARABOLIC PROBLEMS I: A LINEAR MODEL PROBLEM*

KENNETH ERIKSSON† AND CLAES JOHNSON†

Abstract. This paper is the first part in a series of papers on adaptive finite element methods for parabolic problems. In this paper, an adaptive algorithm is presented and analyzed for choosing the space and time discretization in a finite element method for a linear parabolic problem. The finite element method uses a space discretization with meshsize variable in space and time and a third-order accurate time discretization with timesteps variable in time. The algorithm is proven to be (i) reliable in the sense that the L_2-error in space is guaranteed to be below a given tolerance for all timesteps and (ii) efficient in the sense that the approximation error is for most timesteps not essentially below the given tolerance. The adaptive algorithm is based on an a posteriori error estimate which proves (i), and sharp a priori error estimates are used to prove (ii). Analogous results are given for the corresponding stationary (elliptic) problem. In the following papers in this series extensions are made, e.g., to timesteps variable also in space and to nonlinear problems.

Key words. adaptive finite element procedures, a priori error estimates, a posteriori error estimates, automatic error control, discontinuous Galerkin method, elliptic problems, parabolic problems

AMS(MOS) subject classifications. 65N15, 65N30

Introduction and outline of main results. This is the first part in a series of papers containing also [E1], [EJ3], and [EJ4] on adaptive finite element methods for parabolic problems. The present work is related to our earlier work on adaptive methods for parabolic or stiff initial value problems initiated in [J] and continued in [EJ2], [JNT], and [L], but contains essential new features in the design and analysis of the adaptive algorithms. We also include in this paper results on an adaptive method for a corresponding stationary (elliptic) model problem.

The problem of constructing adaptive finite element methods is of great practical importance and has recently been considered by several researchers. For pioneering work with mathematical aspects we refer to Babuska et al. ([BB1], [BB2], [BB3], [BM], see also [B] and [EW]) and for recent work in the engineering field to Löhner, Morgan, and Zienkiewicz [LMZ] and Oden et al. [ODSD]. An adaptive algorithm may be considered to be a computational procedure for constructing a finite element discretization for a given problem (not requiring any a priori information on the exact solution) such that, ideally, the error of the corresponding approximate solution is within a given tolerance in a given norm and such that the number of degrees of freedom is minimal. Successful adaptive finite element methods may be expected to lead to substantial savings in computational work for a given accuracy, and quantitative error control is of obvious interest in applications. Adaptive codes are now entering into applications, and adaptivity may be expected to become a standard feature of finite element software in the future.

As is well known, the smoothness of solutions of parabolic problems in general vary considerably in space and time with, for instance, initial transients where highly oscillatory components of the solution are rapidly decaying. Efficient computational methods for parabolic problems therefore require space and timesteps which are variable, ideally in both space and time, with small steps in transients and larger steps as the exact solution becomes smoother.

* Received by the editors May 21, 1986; accepted for publication (in revised form) January 30, 1990. This work was supported by the Swedish Board for Technical Development (STU).

† Department of Mathematics, Chalmers University of Technology, 41296, Göteborg, Sweden.

In this paper we construct an adaptive algorithm for choosing the space and timesteps in a finite element method for a linear parabolic model problem and prove the following results:

(a) The algorithm is *reliable* in the sense that the L_2-error in space is guaranteed to be within a given tolerance for all time.

(b) The algorithm is *efficient* in the sense that the computational mesh generated by the algorithm is not overly refined for the given accuracy (more precisely we prove here, up to a modification of the tolerance by a constant factor, that an optimal mesh is accepted by the algorithm and that the actual L_2-error in space on a mesh produced by the algorithm is almost never below the tolerance).

Corresponding results are given for the stationary problem.

The adaptive algorithms are based on sharp a posteriori error estimates estimating the error in terms of quantities depending on the computed solution and given data, which proves (a). The efficiency (b) is related to the sharpness of the a posteriori error estimates and is demonstrated using a priori error estimates, where the error is estimated in terms of the exact solution. To prove efficiency we need a priori estimates with minimal regularity requirements on the exact solution and allowing the space and timesteps to be variable in both space and time. Since such a priori estimates are not available in the literature, a considerable part of the work will be devoted to this problem. The a posteriori and a priori error estimates are both proved using duality techniques involving a continuous and a discrete dual problem, respectively, and rely on the strong stability properties of the dual problems. The fact that a continuous dual problem is used to prove the a posteriori estimate makes this estimate easier to prove than the a priori estimate contrary to the standard opinion, that a posteriori estimates are difficult to obtain. Our technique for proving the a posteriori estimates appears to be new, although as indicated, it is simple and natural. For a related but somewhat different approach to a posteriori error estimates and adaptivity, where the a posteriori estimates are not obtained directly but via a priori estimates, see [JNT] and [L]. For an a posteriori estimate of a different form than those used in this paper for a backward Euler semidiscretization in time of a parabolic problem, we refer to Lippold [Li].

The finite element method underlying the adaptive algorithm in the time-dependent case is the discontinuous Galerkin method which is based on a space-time discretization. We consider here the case of piecewise linear (continuous) approximation in space and piecewise polynomial (discontinuous) approximation in time of degree $q=0$ or 1, giving a timestepping method which is second-order accurate in space and accurate of order $2q+1$ in time. The advantages of using the discontinuous Galerkin method in the present context are as follows: (i) The method has excellent stability properties which makes it suitable for parabolic problems and in particular allows sharp a posteriori and a priori error analysis, which is used in the design and analysis of the adaptive algorithm. (ii) The method naturally allows space and timesteps which are variable in both space and time.

In this paper we shall consider a linear parabolic model problem with constant coefficients and we shall assume that the timesteps are constant in space. The case of timesteps variable in space will be considered in [EJ3]. In [EJ4] we plan to consider parabolic problems with variable coefficients and certain nonlinear parabolic problems and also a class of stiff systems of ordinary differential equations.

An outline of this paper is as follows: In § 1 we introduce the discretization methods. In § 2 we state the basic a priori and a posteriori error estimates for these methods. In § 3 we formulate the adaptive algorithms and state our results on their reliability and efficiency. In § 4 we give some preliminary results and in § 5 we prove the a posteriori error estimates for the elliptic problem. In §§ 6 and 7 we prove the a priori and a posteriori error estimates for the parabolic problem. In § 8 we prove the results on the reliability and efficiency of the methods. Finally, in § 9 we give some numerical results illustrating the performance of the algorithms.

To be explicit, let us emphasize some of the unique features of the present paper together with [E1] and [EJ3] all concerned with linear model problems:

(i) sharp a posteriori error estimates,

(ii) a priori estimates with minimal requirements on the regularity of the exact solution and which allow the space and timesteps to be variable in both space and time,

(iii) new adaptive algorithms which prove to be reliable and efficient.

Concerning (ii), note that earlier a priori error estimates for parabolic problems require excessive regularity of the exact solution and generally assume the space mesh to be quasi uniform in space and constant in time (except in the early work by Dupont [Du] where nonoptimal a priori estimates were derived with the space mesh variable in time, see also [DF]).

The techniques used to establish (i)-(iii) are general in nature and not restricted to the model problems considered. Extensions to more general problems will be given in [EJ4].

1. The discretization methods. For simplicity we shall restrict our considerations to the standard model problems of elliptic and parabolic type, namely, to find u such that

$$\begin{aligned} \Delta u(x) &= f(x), \qquad x \in \Omega, \\ u(x) &= 0, \qquad x \in \partial\Omega, \end{aligned} \tag{1.1}$$

and

$$\begin{aligned} u_t(x,t) - \Delta u(x,t) &= f(x,t), \qquad x \in \Omega, \quad t \in \mathbb{R}^+, \\ u(x,t) &= 0, \qquad x \in \partial\Omega, \quad t \in \mathbb{R}^+, \\ u(x,0) &= u_0(x), \qquad x \in \Omega, \end{aligned} \tag{1.2}$$

respectively. Here Ω is a bounded polygonal domain in $\mathbb{R}^2$ with boundary $\partial\Omega$, $\mathbb{R}^+ = (0, \infty)$, $\Delta = (\partial^2 u/\partial x_1^2) + (\partial^2 u/\partial x_2^2)$, $u_t = \partial u/\partial t$, and the functions f and u_0 are given data.

For the discretization of these problems with respect to the space variable $x = (x_1, x_2)$, let Σ be the class of all finite element discretizations $\mathcal{S} = (h, T, S)$ defined as follows:

(A.1) h is a positive function in $C^1(\bar{\Omega})$ such that

$$|\nabla h(x)| \leqq \mu \quad \forall x \in \bar{\Omega}. \tag{1.3a}$$

(A.2) $T = \{K\}$ is a set of triangular subdomains of Ω defining a partition of Ω into triangles K of diameter h_K such that

$$c_1 h_K^2 \leqq \int_K dx \quad \forall K \in T, \tag{1.3b}$$

and associated with the function h through

$$c_2 h_K \leqq h(x) \leqq h_K \quad \forall x \in K, \forall K \in T, \tag{1.3c}$$

where c_1, c_2, and μ are given positive constants.

(A.3) S is the set of all continuous functions on $\bar{\Omega}$ which are linear in x on each $K \in T$ and vanish on $\partial\Omega$.

As usual we require the triangulation to be regular in the sense that the intersection of any two (closed) triangles in T is either empty or a common edge or a common vertex of the two.

With a triangulation T as above we associate the set $E=\{\tau\}$ consisting of those line segments in $\mathbb{R}^2$ which appear as an edge of some $K\in T$. We further denote by E_i those τ in E which are interior to Ω (not part of $\partial\Omega$) and note that by (1.3b)

$$h_K=\max_{\tau\in\partial K} h_\tau \leqq C \min_{\tau\in\partial K} h_\tau,$$

where h_τ is the length of τ.

The stationary problem (1.1) may now be approximated in the usual way: Let $\mathcal{S}_h=(h, T_h, S_h)\in\Sigma$ and find $u_h\in S_h$ such that

$$(\nabla u_h, \nabla v)=(f, v) \quad \forall v\in S_h, \tag{1.4}$$

where $(\cdot,\cdot)$ denotes the usual inner product in $[L_2(\Omega)]^d$, $d=1, 2$.

By introducing the L_2 projection operator $P_h: L_2(\Omega)\to S_h$

$$(P_h w, v)=(w, v) \quad \forall v\in S_h,$$

and the discrete Laplacian $\Delta_h: H^1(\Omega)\to S_h$ defined by

$$(\Delta_h w, v)=-(\nabla w, \nabla v) \quad \forall v\in S_h,$$

we may write (1.4) equivalently as

$$-\Delta_h u_h=P_h f,$$

which has a more obvious resemblance to (1.1).

Let us now turn to the time dependent problem (1.2). For a full discretization of this problem with the discontinuous Galerkin method we consider partitions $0=t_0<t_1<\cdots<t_n<\cdots$ of $\mathbb{R}^+$ into subintervals $I_n=(t_{n-1}, t_n)$ of length $k_n=t_n-t_{n-1}$, and associate with each such time interval a space discretization $\mathcal{S}_n=(h_n, T_n, S_n)\in\Sigma$. For q a nonnegative integer we define

$$V_{qn}=\left\{v: v=\sum_{j=0}^{q} t^j\varphi_j,\ \varphi_j\in S_n\right\},$$

and discretize (1.2) as follows: Find U such that for $n=1,2,\cdots$, $U|_{\Omega\times I_n}\in V_{qn}$ and

$$\int_{I_n}\{(U_t, v)+(\nabla U, \nabla v)\}\,dt+([U]_{n-1}, v_{n-1}^+)=\int_{I_n}(f, v)\,dt \quad \forall v\in V_{qn}, \tag{1.5}$$

where

$$[w]_n=w_n^+-w_n^-,\ w_n^{+(-)}=\lim_{s\to 0^{+(-)}} w(t_n+s), \qquad U_0^-=u_0.$$

For $q=0$ equation (1.5) reduces to the following variant of the Euler backward scheme where $U_n=U|_{t\in I_n}$ and $\Delta_n=\Delta_{h_n}$:

$$(U_n-P_nU_{n-1})/k_n-\Delta_nU_n=\int_{I_n} P_nf\,dt/k_n. \tag{1.6}$$

For $q=1$ we may write $U|_{I_n}$ as $\phi_n+(t-t_{n-1})\psi_n/k_n$ for some ϕ_n and ψ_n in S_n and (1.5) reduces to the system of equations (cf. [EJT])

$$\begin{aligned}
&\psi_n-k_n\Delta_n\phi_n-\frac{k_n}{2}\Delta_n\psi_n+\phi_n=P_nU_{n-1}^-+\int_{I_n}P_nf\,dt,\\
&\frac{1}{2}\psi_n-\frac{k_n}{2}\Delta_n\phi_n-\frac{k_n}{3}\Delta_n\psi_n=\int_{I_n}(t-t_{n-1})P_nf\,dt/k_n.
\end{aligned} \tag{1.7}$$

The time discretization method is in this case basically the same as the subdiagonal Padé scheme of order (2, 1) and is third-order accurate in U_n^- at the nodal points t_n.

Remark. Note that in the discretization (1.5) the space and timesteps may vary in time and that the space discretization may be variable also in space, whereas the timesteps k_n are kept constant in space. Clearly, optimal mesh design requires the timesteps to be variable also in space. Now, it is easy to extend the method (1.5) to admit timesteps which are variable in space simply by defining

$$V_{qn}=\left\{v\colon v(x,t)=\sum_i v_i(t)\chi_i(x)\right\},$$

where $\{\chi_i\}$ is a basis for S_n and the coefficients v_i now are piecewise polynomial of degree q in t, without continuity requirements, on partitions of I_n which may vary with i. The discrete functions may now be discontinuous also inside the "slab" $\Omega\times I_n$. The discontinuous Galerkin method again takes the form (1.5) with the difference that the term $([U]_{n-1}, v_{n-1}^+)$ is replaced by a sum over all jumps of U in $\Omega\times[t_{n-1}, t_n)$ and further the discontinuities of U_t are discarded in the integral involving U_t. Adaptive methods for the discontinuous Galerkin method in this generality will be considered in [EJ3].

2. A priori and a posteriori error estimates. In this section we state a priori and a posteriori error estimates for the discretization methods (1.4) and (1.5).

For the stationary problem we have the following a priori estimate.

THEOREM 2.1. *Let $f\in L_2(\Omega)$ and let u and u_h be the solutions of* (1.1) *and* (1.4), *respectively. Then there exist constants C depending only on the constants c_1 and c_2 in* (1.3), *such that the following holds: If μ is sufficiently small, then*

$$\|D^1(u-u_h)\|\leqq C\|hD^2u\|, \tag{2.1a}$$

and if in addition Ω is convex, then

$$\|u-u_h\|\equiv\|D^0(u-u_h)\|\leqq C\|h^2D^2u\|, \tag{2.1b}$$

where $\|\cdot\|$ is the $L_2(\Omega)$ norm and $D^m u=(\sum_{|\alpha|=m}|D^\alpha u|^2)^{1/2}$.

Remark 2.0. Note the way in which the local meshsize $h(x)$ enters in these error estimates showing that large second derivatives of u may be compensated for by a (locally) small meshsize so as to control the quantity $\|h^mD^2u\|$, $m=1, 2$, bounding the error. This indicates the possiblity of adaptively choosing the meshsize to control the error if D^2u may be computationally estimated, an idea which was explored in [EJ1]. In this note, however, we will follow a related but different adaptive strategy directly based on a posteriori error estimates.

Remark 2.1. The error bounds in (2.1a, b) can in fact be replaced by

$$C\left\|\min_{1\leqq m\leqq 2} h^{m-j}D^m u\right\|,$$

with $j=1$ and $j=0$, respectively. These bounds are of interest, e.g., when $u\notin H^2(\Omega)$ and in cases when ∇u has rapid oscillations.

Remark 2.2. Note further that the estimates (2.1a, b) are optimal in the sense that there exists a constant c such that "for most u" (e.g., if $D^\alpha u$, $|\alpha|=2$, is roughly constant on each element),

$$\inf_{v\in S_h}\|D^{2-m}(u-v)\|\geqq c\|h^mD^2u\|, \qquad m=1,2$$

which indicates that error control based on (2.1) should be efficient.

The error estimate (2.1a) is classical, whereas the estimate (2.1b) in the present generality can be found in [E1]. For quasi-uniform partitions (corresponding to taking h constant) the estimate (2.1b) is well known. Let us further remark that the estimates (2.1a, b) may also be derived for a (convex or nonconvex) domain Ω with smooth boundary, in the case (2.1b) with the constant C depending on Ω.

To state the a posteriori estimate underlying the adaptive algorithm for the stationary problem we need some notation. With each $\tau \in E_i$ associate a vector n_τ of length one normal to τ and define for $v \in S_h$

$$\left[\frac{\partial v}{\partial n_\tau}\right] = \lim_{s\to 0^+} (\nabla v(x+sn_\tau) - \nabla v(x-sn_\tau)) \cdot n_\tau, \qquad x \in \tau,$$

that is, $[\partial v/\partial n_\tau]$ is the jump across τ in the normal component of ∇v. We now introduce, with h_τ the length of τ, the discrete norm

$$D_{h,m}(v) = \left(\sum_{\tau \in E_i} h_\tau^{2m} \left|\left[\frac{\partial v}{\partial n_\tau}\right]\right|^2\right)^{1/2}, \qquad v \in S_h,$$

which should be viewed as a discrete counterpart of $\|h^m D^2 v\|$.

We may then state the following a posteriori error estimates for the stationary problem, which are proved in § 5 below.

THEOREM 2.2. *There are constants α_i and β_i only depending on the constants c_1 and c_2 such that if $f \in L_2(\Omega)$ and u and u_h are the solutions of* (1.1) *and* (1.4), *respectively, and μ is sufficiently small, then*

$$\|D^1(u-u_h)\| \leqq \alpha_1 \|hf\| + \beta_1 D_{h,1}(u_h), \tag{2.2a}$$

and, if in addition Ω is convex,

$$\|u-u_h\| \leqq \alpha_2 \|h^2 f\| + \beta_2 D_{h,2}(u_h). \tag{2.2b}$$

Remark 2.3. Note that by defining a piecewise constant quantity $D_h^2 u_h$ by

$$D_h^2 u_h\big|_K = \max_{K' \in N(K)} \frac{|\nabla u_h(P') - \nabla u_h(P)|}{|P'-P|},$$

where $N(K)$ is the set of triangles K' in T_h sharing one edge with K, and P' and P denote the centers of gravity of the triangles K' and K, we have for some constant C only depending on c_1 and c_2,

$$\frac{1}{C} D_{h,m}(u_h) \leqq \|h^m D_h^2 u_h\| \leqq C D_{h,m}(u_h), \qquad m = 1, 2,$$

and thus we can rewrite (2.2a, b) as

$$\|D^{2-m}(u-u_h)\| \leqq C(\|h^m f\| + \|h^m D_h^2 u_h\|), \qquad m = 1, 2, \tag{2.3}$$

which shows a strong similarity with the optimal a priori estimates (2.1a, b).

Remark 2.4. Note that without further analysis the amount of information in the a posteriori estimates (2.2a, b) is not obvious. Clearly, if for a given $\mathcal{S}_h \in \Sigma$ we compute u_h using (1.4), then we may bound the error using (2.2a, b) by evaluating the right-hand sides of these estimates. If these quantities turn out to be sufficiently small, then we may be satisfied and quit. However, without further analysis it is conceivable that the right-hand sides of (2.2a, b) would always be large and then these estimates would be useless. In fact, a posteriori error estimates of the form (2.2a, b) may be derived also for unstable methods and in such cases the right-hand side quantities could be large regardless of the meshsize. In our case we shall prove that in fact (2.2a, b) are sharp, and thus may be useful in practice, by comparing them with the optimal a priori estimates (2.1a, b).

Remark 2.5. If $f \in H^2(\Omega)$, then the f-terms in (2.2a, b) or (2.3) may be replaced by $\alpha_3 \| h^{4-m} D^2 f\|$.

Let us next state optimal a priori estimates for the time dependent problems. These estimates are improved variants of those of [EJT] allowing the space mesh to be variable in space and time and with minimal regularity requirements on the exact solution. For simplicity we assume that Ω is convex. The estimates may be extended to general domains with smooth boundary with the constants C depending on Ω.

THEOREM 2.3. *Let u be the solution of* (1.2) *and U that of* (1.5), *suppose μ is small enough and assume that for each n one of the following two assumptions hold*:

$$S_n \subset S_{n-1}, \tag{2.4a}$$

$$\bar{h}_n^2 \leqq \gamma k_n, \tag{2.4b}$$

where $\bar{h}_n = \max_{x \in \bar{\Omega}} h_n(x)$ and γ is sufficiently small and that for all n, $k_n \leqq C k_{n+1}$. Then there exist constants C only depending on c_1 and c_2 (if Ω is convex) such that for $q = 0$, 1, and $N = 1, 2, \cdots$,

$$\|u - U\|_{I_N} \leqq C L_N \max_{1 \leqq n \leqq N} E_{qn}(u), \tag{2.5a}$$

and for $q = 1$, $N = 1, 2, \cdots$,

$$\|u(t_N) - U_N^-\| \leqq C L_N \max_{1 \leqq n \leqq N} E_{2n}(u), \tag{2.5b}$$

where $L_N = \frac{1}{4}(\log(t_N/k_N) + 1)^{1/2}$,

$$E_{qn}(u) = \min_{j \leqq q+1} k_n^j \|u_t^{(j)}\|_{I_n} + \|h_n^2 D^2 u\|_{I_n}, \qquad q = 0, 1, 2$$

with $u_t^{(1)} = u_t$, $u_t^{(2)} = u_{tt}$, $u_t^{(3)} = \Delta u_{tt}$ and $\|w\|_{I_n} = \max_{t \in I_n} \|w(t)\|$.

Remark 2.6. Note that (2.5) states that the discontinuous Galerkin method (1.5) is of order $q+1$ globally in time and of order $2q+1$ at the discrete time levels t_n for $q = 0$, 1. Further, the estimates (2.5) are optimal in the sense that for some positive constant c

$$\inf_{v \in V_{qn}} \|u - v\|_{I_n} \geqq c E_{qn}(u), \qquad q = 0, 1, 2, \tag{2.6}$$

if here, in the definition of $E_{qn}(u)$, we put $u_t^{(3)} = u_{ttt}$ and restrict the variation of $u_t^{(3)}$ and $D^\alpha u$ for $|\alpha| = 2$ as in Remark 2.2. Note that for the "super approximation" result (2.5b) it is relevant to compare with approximation in V_{2n}.

Remark 2.7. With quasi-uniform space meshes with $h_n(x) \sim \bar{h}_n$ we expect to have $\bar{h}_n^2 \sim k_n$ for $q = 0$ and $\bar{h}_n^2 \ll k_n$ if $q = 1$, since the discontinuous Galerkin method is of second order in space and of order $2q+1$ in time, $q = 0$, 1. Thus, in particular for $q = 1$ the condition (2.4b) does not appear to be restrictive and in fact allows a considerable variation of $h_n(x)$. In certain extreme situations, however, e.g., with initial data u_0 highly concentrated in space, (2.4b) may impose a restriction on the mesh. It is possible that (2.4b) may be weakened to a condition of the form $\bar{h}_n^2 \leqq \gamma K_n$, where $K_n = t_{n^*} - t_{n-1}$, and $S_m = S_n$ for $m = n, n+1, \cdots, n^*$.

The proof of Theorem 2.3 in the case (2.4a) is given in § 6 below, while the case (2.4b) is considered in [E1].

We finally state a posteriori estimates for the time dependent problem. These estimates are proved in § 7 below. Again, we assume that Ω is convex, but generalizations to smooth nonconvex domains are possible (cf. Remark 2.10).

THEOREM 2.4. *Let u be the solution of* (1.2) *and U that of* (1.5), *suppose Ω is convex and μ sufficiently small. Then for $N \geqq 1$, we have for $q=0$*

$$\|u(t_N)-U_N^-\| \leqq \max_{1\leqq n\leqq N} \mathscr{E}_{0n}(U), \tag{2.7a}$$

and for $q=1$

$$\|u(t_N)-U_N^-\| \leqq \max_{1\leqq n\leqq N} \mathscr{E}_{2n}(U), \tag{2.7b}$$

where

$$\mathscr{E}_{0n}(U)=C_1\|h_n^2 f\|_{I_n}+C_2\int_{I_n}\|f\|\,dt+C_3 D_{n,2}(U_n)+C_4\|[U_{n-1}]\|$$
$$+C_5\|h_n^2[U_{n-1}]/k_n\|^*,$$
$$\mathscr{E}_{2n}(U)=C_6\|h_n^2 f\|_{I_n}+C_7 k_n^2\int_{I_n}\|f_{tt}\|\,dt+C_8\max_{t\in I_n} D_{n,2}(U(t))$$
$$+\min\,(C_9\|[U_{n-1}]\|,\,C_{10}k_n\|\Delta_n P_n[U_{n-1}]\|)+C_{11}\|h_n^2[U_{n-1}]/k_n\|^*,$$

where $D_n^2=D_{h_n}^2$, and a star indicates that the corresponding term is present only if $S_n \not\supseteq S_{n-1}$. Furthermore, the C_i are constants given by

$$C_1=\alpha_2 L,\quad C_2=L+2,\quad C_3=\beta_2(L+2),\quad C_4=L+1,$$
$$C_5=\alpha_2(L+\exp(-1)),\quad C_6=\alpha_2(L+2),\quad C_7=\gamma_3(\gamma_1 L+\gamma_0+1),$$
$$C_8=2\beta_2(L+1),\quad C_9=C_4,\quad C_{10}=\gamma_2 L+\gamma_1,\quad C_{11}=C_5,$$
$$L=\max_N L_N,$$

where α_2 and β_2 are certain constants depending on c_1 and c_2 related to approximation by functions in S_h (*see Lemma* 4.3), *and the γ_i are absolute constants related to one-dimensional approximation by linear functions* (*see* (7.14) *and* (7.15)).

Remark 2.8. The term $C_i\|h_n^2 f\|_{I_n}$ in $\mathscr{E}_{0n}$ and $\mathscr{E}_{2n}$ may be replaced by $\bar{C}_i\|h_n^4 D^2 f\|_{I_n}$, $i=1,6$, the term $C_2\int_{I_n}\|f\|\,dt$ in $\mathscr{E}_{0n}$ by $\bar{C}_2 k_n\int_{I_n}\|f_t\|\,dt$ and $C_7 k_n^2\int_{I_n}\|f_{tt}\|\,dt$ in $\mathscr{E}_{2n}$ by $\bar{C}_7 k_n^3\int_{I_n}\|\Delta f_{tt}\|\,dt$ with modified constants $\bar{C}_i$.

Remark 2.9. The comments of Remark 2.4 are also relevant for the a posteriori estimate (2.7). By comparison with the optimal a priori estimate (2.5) we shall prove below that (2.7) is sharp and thus may be used as a basis for an efficient adaptive algorithm.

Remark 2.10. In the general case with the boundary of Ω smooth (some of) the constants C_i should be replaced by constants $\hat{C}_i=C_s C_i$, where C_s is a stability constant depending on Ω defined by (cf. Lemma 4.1)

$$C_s=\sup_{\substack{v\in H_0^1(\Omega)\cap H^2(\Omega)\\ v\neq 0}}\frac{\|D^2 v\|}{\|\Delta v\|}.$$

The approximation constants α_2 and β_2 (depending on c_1 and c_2) and the absolute constants γ_i entering in the C_i, may be estimated once and for all (values of these constants used in our numerical compuatations are given in § 9 below), while the stability constant C_s in general depends on Ω. It is possible that a relevant value of C_s may be found by computing the quotient $\|D_h^2 v\|/\|\Delta_h v\|$ for some properly chosen $v \in S_h$. The a posteriori estimates may be generalized also to problems with variable coefficients or nonlinear problems (see [EJ4]). In this case the C_i should be replaced by $\hat{C}_i = C_s(u)C_i$, where $C_s(u)$ is a "stability constant" depending on Ω and the coefficients, and also "mildly" on u. It is likely that such constants may be estimated through the adaptive procedure; see [EJ2] and [EJ4].

Remark 2.11. We can prove direct analogues of Theorem 2.1–2.4 replacing the $L_2(\Omega)$-norm by the $L_p(\Omega)$-norm, $1 \leqq p \leqq \infty$ (see [E1]).

Remark 2.12. Defining the H^{-1}-norm of the residual $r(u_h, f)$ of the discrete solution u_h of (1.4) by

$$\|(r(u_h, f)\|_{H^{-1}} \equiv \sup_{\substack{\|\nabla \varphi\|=1 \\ \varphi \in H_0^1(\Omega)}} [(f, \varphi) - (\nabla u_h, \nabla \varphi)],$$

we have, writing $e = u - u_h$ with u the solution of (1.1),

$$\|\nabla e\|^2 = (\nabla(u - u_h), \nabla e) = (f, e) - (\nabla u_h, \nabla e) \leqq \|r(u_h, f)\|_{H^{-1}} \|\nabla e\|,$$

which gives the following "abstract" a posteriori estimate

$$\|\nabla e\| \leqq \|r(u_h, f)\|_{H^{-1}}. \tag{2.8}$$

The estimate (2.2a) may be viewed as "concretization" of (2,8) where the right-hand side of (2.8) is estimated following the proof of Theorem 2.2.

3. The adaptive algorithms. Reliability and efficiency. Let us first consider the elliptic problem (1.1) with solution u and the corresponding discrete problem (1.4) with solution $u_h \in S_h$ associated with the discretization $\mathcal{S}_h = (h, T_h, S_h) \in \Sigma$.

Let now $\delta > 0$ be a given *tolerance* and suppose we want to determine an approximation u_h at minimal cost so that for $m = 1$ or $m = 2$,

$$\|D^{2-m}(u - u_h)\| \leqq \delta. \tag{3.1}$$

Starting from the a priori estimate (2.1) this problem may be reformulated as follows: Find $\mathcal{S}_h \in \Sigma$ such that

(3.2a) $\quad \|h^m D^2 u\| \leqq \dfrac{\delta}{C}$,

(3.2b) $\quad \mathcal{S}_h$ is *optimal* in the sense that the number of degrees of freedom of S_h is minimal.

To find u_h and $\mathcal{S}_h$ satisfying (3.1) and (3.2) it is natural to consider the following *adaptive method* based on the a posteriori estimate (2.3): Find $\mathcal{S}_h \in \Sigma$ so that the corresponding $u_h \in S_h$ satisfies for $m = 1$ or 2,

(3.3a) $\quad \alpha_m \|h^m f\| + \beta_m D_{h,m}(u_h) \leqq \delta$,

(3.3b) $\quad$ the number of degrees of freedom of S_h is (nearly) minimal.

Since u_h depends on $\mathcal{S}_h$ this is a complex nonlinear minimization problem. The existence of a $\mathcal{S}_h$ with corresponding u_h satisfying (3.3a) follows from Theorem 3.1 below. To solve the problem (3.3) approximately we consider the following *adaptive algorithm*: With T_h^0 a given arbitrary initial triangulation determine successively mesh functions

$h_j(x)$ with corresponding triangulations T_h^j, finite-dimensional spaces S_h^j, and approximate solutions $u_h^j \in S_h^j$ such that for $j = 1, 2, \cdots, J$, h_j is maximal under the condition (1.3a) and

$$\alpha_m \|h_j^m f\|_{L_2(K)} + \beta_m \left(\frac{1}{2} \sum_{\tau \in \partial K} h_{j,\tau}^{2m} \left| \left[\frac{\partial u_h^{j-1}}{\partial n_\tau} \right] \right|^2 \right)^{1/2} \leqq \frac{\theta \delta}{\sqrt{N_{j-1}}}, \qquad K \in T_h^{j-1}, \tag{3.4}$$

where $h_{j,\tau} = \max_{x \in \tau} h_j(x)$, $\tau \in \partial K$ is a side of $K \in T_h^{j-1}$, and N_j is the number of degrees of freedom of S_h^j. The number of "trials" J is the smallest integer such that (3.3a) holds with h and u_h replaced by h_J and u_h^J, and θ is a parameter ($\theta \sim 1$) chosen so that (3.3a) is satisfied with near equality and J is small (cf. [E2], where in a similar situation it is proved that θ may be chosen so that always $J \leqq 2$). Note that (3.4) attempts to make all element contributions equal, which would seem to correspond to the most efficient distribution of the degrees of freedom and thus would imply (3.3b). Note also that if f is smooth, then the f-term in (3.3a) and (3.4) may be modified according to Remark 2.5.

By the a posteriori estimate (2.2) it follows that (3.3a) guarantees the error control (3.1a) and thus the adaptive method (3.3) is reliable. Concerning the efficiency of the algorithm (3.3) we have the following result.

THEOREM 3.1. *Under the assumptions of Theorem* 2.1 *there is a constant C such that for* $m = 1, 2$,

$$\alpha_m \|h^m f\| + \beta_m D_{h,m}(u_h) \leqq C \|h^m D^2 u\|. \tag{3.5}$$

Proof. Let $u_i \in S_h$ be the standard Lagrangian interpolant of u interpolating at the nodes of T_h and note that by standard interpolation theory there is a constant C depending only on c_1 and c_2, such that for $m = 1, 2$,

$$\|h^m D_h^2 u_i\| + \|D^{2-m}(u - u_i)\| \leqq C \|h^m D^2 u\|. \tag{3.6}$$

Recalling Remark 2.3 we now have by an inverse estimate and Theorem 2.1 for $m = 1, 2$,

$$\begin{aligned} D_{h,m}(u_h) &\leqq C \|h^m D_h^2 u_h\| \leqq C(\|h^m D_h^2 (u_h - u_i)\| + \|h^m D_h^2 u_i\|) \\ &\leqq C \|D^{2-m}(u_h - u_i)\| + C \|h^m D^2 u\| \\ &\leqq C \|D^{2-m}(u - u_h)\| + C \|D^{2-m}(u - u_i)\| + C \|h^m D^2 u\| \leqq C \|h^m D^2 u\|, \end{aligned}$$

which proves the desired result. $\square$

Remark. Note that the assumption that μ is sufficiently small may be enforced explicitly in the algorithm by somewhat restricting the mesh variation.

We now discuss the efficiency of the adaptive method (3.4). We then start out noting that by (3.5) it follows that (3.3a) may be realized by taking h small enough. Thus, the method is *operative* in the sense that for any δ there is a mesh satisfying (3.3a) (cf. Remark 2.4). Next we note that if $C\|h^m D^2 u\| \leqq \delta$, then $\alpha_m \|h^m f\| + \beta_m D_{h,m}(u_h) \leqq \delta$. This means, that for an optimal mesh satisfying (3.2) the criterion (3.3a) would be satisfied (up to a constant). Thus, an optimal mesh would be accepted by the algorithm in the sense that (3.3a) would be satisfied. This is an indication of efficiency of the algorithm. Next, we note that if we apply the algorithm so that (3.3a) holds, then by Theorem 3.1 (cf. also Remarks 2.1 and 2.2) for the actual error we would have $\|D^{2-m}(u - u_h)\| \geqq \delta / C$. This also indicates efficiency in the sense that the algorithm cannot generate a mesh for which the actual error is significantly smaller than the tolerance. However, it is not obvious from Theorem 3.1 that the mesh $\mathcal{S}_h$ actually generated by the adaptive algorithm will be optimal in the sense (3.3b). To strictly prove this would require an estimate from below of $|h^m(x)D^2u(x)|$ *locally* for $x \in \Omega$ and not just globally as stated in Theorem 3.1. For a local analysis proving the efficiency of (a variant of (3.3)) in a more precise sense, we refer to [E2] and [EJ5], (cf. also [EJ1]).

We summarize our results on the adaptive method (3.3) for the stationary problem as follows:

THEOREM 3.2. *The adaptive method* (3.3) *is*

(a) **operative** *in the sense that for any* $\delta > 0$ *there is a* $\mathcal{S}_h \in \Sigma$ *with corresponding* $u_h \in S_h$ *satisfying* (3.3a),

(b) **reliable** *in the sense that if* (3.3*a*) *holds, then the error control* (3.1) *is guaranteed,*

(c) **efficient** *in the sense that* (*up to a modification of the tolerance by a constant factor*) *an optimal mesh is accepted by the algorithm and that the actual error on any mesh produced by the algorithm is not below the tolerance.*

We now turn our attention to the time dependent problem (1.2) and the corresponding discrete problem (1.5) with solution U based on a space-time discretization $\mathbf{S}_h = \{(h_n, T_n, S_n, k_n)\}_{n \geqq 1}$ where $(h_n, T_n, S_n) \in \Sigma$, $n = 1, 2, \cdots$. For a given tolerance δ, consider the problem of finding $\mathbf{S}_h$ such that

(3.7a) $\quad \|u(t_n) - U_n^-\| \leqq \delta, \quad n = 1, 2, \cdots,$

(3.7b) $\quad \mathbf{S}_h$ is optimal, in the sense that the number of degrees of freedom is minimal.

To solve this problem we are, in analogy with the elliptic problem, led to the following *adaptive method* based on the a posteriori estimate (2.7a, b): Choose $\mathbf{S}_h$ such that for $n = 1, 2, \cdots,$

(3.8a) $\quad \mathscr{E}_{0n}(U) \leqq \delta \quad \text{if } q = 0,$

(3.8b) $\quad \mathscr{E}_{2n}(U) \leqq \delta \quad \text{if } q = 1,$

(3.8c) $\quad$ the number of degrees of freedom of $\mathbf{S}_h$ is minimal,

As a practical implementation of this method, we consider the following *adaptive algorithm* for choosing $\mathbf{S}_h$: for each $n = 1, 2, \cdots$, with T_n^0 a given initial space mesh and k_{n0} an initial timestep, determine meshes T_n^j with N_j elements of size $h_{nj}(x)$, and timesteps k_{nj} and corresponding approximate solutions U^j defined on I_{nj}, such that for $j = 0, 1, \cdots, \hat{n} - 1$,

(3.9a)
$$C_6 \max_{t \in I_{nj}} \|h_{n,j+1}^2 f\|_{L_2(K)} + C_8 \max_{t \in I_{nj}} \left(\frac{1}{2} \sum_{\tau \in \partial K} h_{n,j+1,\tau}^4 \left| \left[\frac{\partial U^j}{\partial n_\tau} \right] \right|^2 \right)^{1/2} + C_{11} \|h_{n,j+1}^2 [U^j]_{n-1}/k_{nj}\|_{L_2(K)}^* = \frac{\theta\delta}{2\sqrt{N_j}} \quad \forall K \in T_n^j,$$

(3.9b)
$$k_{n,j+1}(C_2\|f\|_{I_{nj}} + C_4\|[U^j]_{n-1}/k_{nj}\|) = \frac{\delta}{2} \quad \text{if } q = 0,$$

(3.9c)
$$k_{n,j+1}^3(C_7\|f_{tt}\|I_{nj} + \min(C_{10}\|\Delta_n P_n[U^j]_{n-1}/k_{nj}^2\|, C_9\|[U^j]_{n-1}/k_{nj}^3\|)) = \frac{\delta}{2} \quad \text{if } q = 1,$$

where $I_{nj} = (t_{n-1}, t_{n-1} + k_{nj})$, $[U^j]_{n-1} = U_{n-1}^{j+} - U_{n-1}^-$, and define $T_n = T_n^{\hat{n}}$, $k_n = k_{n\hat{n}}$, and $h_n = h_{n\hat{n}}$ (the constants in (3.9a) should be modified in the obvious way if $q = 0$). As above, for each n the number of "trials" $\hat{n}$ is the smallest integer such that for $j = \hat{n}$, (3.8a or b) holds with U replaced by $U^{\hat{n}}$, and $\theta \sim 1$ is a parameter chosen so that $\hat{n}$ is small and (3.8a, b) is satisfied with near equality with U replaced by $U^{\hat{n}}$. With smooth f the method (3.8) may be modified in the obvious way using Remark 2.8. In the applications reported on below, we chose $T_n^0 = T_{n-1}$ for $n = 2, 3, \cdots$, and $\theta = 1$ and we then found that usually $\hat{n} = 1$.

By the a posteriori estimate (2.7a, b) it follows that the adaptive method (3.8) is reliable in the sense that if (3.8a, b) holds, then the error control (3.7a) is guaranteed. Concerning the efficiency of (3.8) we shall prove certain results giving bounds for the a posteriori quantities $\mathscr{E}_{qn}(U)$, $q = 0, 2$, in terms of the a priori quantities $E_{qn}(u)$, $q = 0$, 2. We then do not include the f-terms in $\mathscr{E}_{qn}(U)$, since these terms may be controlled independently. We shall further use a slight modification $\hat{E}_{2n}$ of E_{2n} defined as follows:

$$\hat{E}_{2n}(u) = E_{2n}(u) + k_n^2 \|u_t^{(2)}\|_{I_n} \quad \text{if } S_{n-1} \neq S_n,$$

and $\hat{E}_{2n}(u) = E_{2n}(u)$ otherwise, corresponding to allowing the accuracy to drop to second order on a timestep immediately following a change of the space mesh. We also assume that there is a constant C such that $(1/C)k_{n-1} \leqq k_n \leqq Ck_{n-1}$, $n = 2, 3, \cdots$. We now first state Theorem 3.3.

THEOREM 3.3. *Under the assumptions of Theorem* 2.3, *and provided* $\bar{h}_n^2 \leqq Ck_n$ *whenever* $S_{n-1} \not\subseteq S_n$, *there exists a constant* C *such that for* $q = 0$

$$\mathscr{E}_{0N}(U) \leqq CL^2 \max_{n \leqq N} E_{0n}(u), \tag{3.10a}$$

and for $q = 1$

$$\mathscr{E}_{2N}(U) \leqq CL^2 \max_{n \leqq N} \hat{E}_{2n}(u). \tag{3.10b}$$

These results state that if $\mathbf{S}_h = \{(h_n, T_n, S_n, k_n)\}$ is optimal for control of $\|u(t_n) - U_n^-\|$ on the tolerance level δ so that $CE_{0n}(u) \leqq \delta$ or $C\hat{E}_{2n}(u) \leqq \delta$, $n = 1, 2, \cdots$, (cf. Remark 2.6), then (up to a modification of the tolerance by a constant factor) $\mathbf{S}_h$ will be accepted by (3.8) in the sense that (3.8a or b) will be satisfied. This indicates that (3.8) is efficient as a method of order 2 in space and order $2q+1$ in time (if the space mesh is not changed too often).

We finally give a result stating that for a discretization $\mathbf{S}_h$ generated by the method (3.8), the actual truncation error $E_{qn}(u)$ will be bounded below by $c\delta/L$ if we maximize over a few n. This result may be viewed as a variant of (3.10) which is localized in time. To prove an efficiency result localized also in space requires a more careful analysis which we do not attempt here (cf. the elliptic case discussed above).

THEOREM 3.4. *Let* $\mathbf{S}_h$ *be a discretization such that the corresponding* U *satisfies* (3.8a *or* b). *Suppose that* $S_n = S_{N^*}$ *for* $N^* \leqq n \leqq N$ *where* $k_N/(t_{N-1} - t_{N^*})$ *is sufficiently small, and suppose further that if* $q = 1$ *then* (3.8) *is implemented so that with* λ *sufficiently large*

$$\frac{\delta}{CL} = \max_{I_n} D_{n,2}(U) = \lambda k_n \|\Delta_n[U_{n-1}]\|, \qquad N^* \leqq n \leqq N. \tag{3.11}$$

Then there is a positive constant c *such that*

$$\max_{N^* \leqq n \leqq N} E_{0n}(u) \geqq \frac{c\delta}{L}, \qquad q = 0, \tag{3.12a}$$

$$\max_{N^* \leqq n \leqq N} E_{2n}(u) \geqq \frac{c\delta}{L}, \qquad q = 1. \tag{3.12b}$$

Remark. The assumption in Theorem 3.4 requiring the space mesh S_n to be constant over a few steps is probably not necessary.

We now summarize our results on the adaptive method (3.8) for the time dependent problem as follows:

THEOREM 3.5. *The adaptive method* (3.8) *is*

(a) **operative** *in the sense that* (3.8a, b) *may be realized,*

(b) **reliable** *in the sense that* (3.8a, b) *guarantees the desired error control* (3.7a),

(c) **efficient** *in the sense that* (*up to a modification of the tolerance by a constant*) *an optimal mesh is accepted by* (3.8a, b) *and on a mesh generated by* (3.8), *the actual* L_2*-error in space is most of the time not below the tolerance.*

Remark 3.1. Recalling Remark 2.11 we may in the obvious way generalize the adaptive methods (3.3) and (3.8) to control the quantities $\|u-U\|_{L_p(\Omega)}$ and $\max_{n\leqq N}\max_{I_n}\|u-U\|_{L_p(\Omega)}$, respectively, $1\leqq p\leqq\infty$.

Remark 3.2. Letting S_h in (1.4) be a finite element space of piecewise polynomials of degree $r-1$ with $r>2$, the a posteriori estimate (2.2) takes the form

$$\|D^{2-m}(u-u_h)\|\leqq C(\|h^{m+r-2}D^{r-2}f\|+D_{h,m}(u_h)+\|h^2(\Delta u_h-P_h\Delta u_h)\|),\qquad m=1,2,$$

where Δu_h is computed elementwise only so that the jumps of $\partial u_h/\partial n_\tau$ are not included. The adaptive method (3.3) and Theorems 3.1 and 3.2 may be generalized in the obvious way to the present situation. Generalizations to more general finite element methods such as mixed methods for elliptic problems are also possible.

4. Preliminaries. We first recall the following regularity (or stability) result.

LEMMA 4.1. *If* Ω *is a convex polygonal domain in* $\mathbb{R}^2$, *then*

$$\|D^2\varphi\|\leqq\|\Delta\varphi\|\quad\forall\varphi\in H_0^1(\Omega)\cap H^2(\Omega).\tag{4.1}$$

For the time dependent problem we shall use the following stability result.

LEMMA 4.2. *Let* u *be the solution of* (1.2) *for* $f=0$. *Then for* $t>0$

$$\|u(t)\|\leqq\|u_0\|,\tag{4.2a}$$

$$\|\Delta u(t)\|\leqq\exp(-1)t^{-1}\|u_0\|,\tag{4.2b}$$

$$\int_0^t\|\nabla u(s)\|^2\,ds\leqq\frac{1}{2}\|u_0\|^2,\tag{4.2c}$$

$$\int_0^t s\|\Delta u(s)\|^2\,ds\leqq\frac{1}{4}\|u_0\|^2,\tag{4.2d}$$

$$\int_0^t s\|u_t(s)\|^2\,ds\leqq\frac{1}{4}\|u_0\|^2.\tag{4.2e}$$

In our analysis below we shall also need the following approximation results for the L_2-projection P_h.

LEMMA 4.3. *Given* c_1 *and* c_2 *there exist positive constants* α_i *and* β_i *such that if* μ *is sufficiently small and* $(h, T_h, S_h)\in\Sigma$, *then for* $m=1,2$,

$$|(f,(I-P_h)v)|\leqq\alpha_m\|h^mf\|\,\|D^mv\|\quad\forall f\in L_2(\Omega),\quad\forall w\in H_0^1(\Omega)\cap H^m(\Omega),\tag{4.3}$$

$$|(\nabla w,\nabla(I-P_h)v)|\leqq\beta_mD_{h,m}(w)\|D^mv\|\quad\forall w\in S_h,\quad\forall v\in H_0^1(\Omega)\cap H^m(\Omega).\tag{4.4}$$

Proof. To prove (4.4), we integrate by parts and use Cauchy's inequality to get for any $\rho\in H_0^1(\Omega)$

$$\begin{aligned}|(\nabla w,\nabla\rho)|&=\left|\sum_{\tau\in E_i}\left[\frac{\partial w}{\partial n_\tau}\right]\int_\tau\rho\,ds\right|\\&\leqq D_{h,m}(w)\left(\sum_{\tau\in E_i}h_\tau^{-2m}\left(\int_\tau\rho\,ds\right)^2\right)^{1/2},\qquad m=1,2.\end{aligned}$$

From the scaled trace inequality

$$\int_\tau |\rho|\,ds \leqq C\left(\int_K |\nabla\rho|\,dx + h_K^{-1}\int_K |\rho|\,dx\right), \qquad \tau\in\partial K,$$

we obtain

$$\left(\sum_{\tau\in E_i} h_\tau^{-2m}\left(\int_\tau \rho\,ds\right)^2\right)^{1/2} \leqq C\left(\sum_{K\in T_h} h^{2-2m}\int_K |\nabla\rho|^2\,dx + h_K^{-2m}\int_K \rho^2\,dx\right)^{1/2}$$
$$\leqq C(\|h^{1-m}\nabla\rho\| + \|h^{-m}\rho\|), \qquad m=1,2.$$

We now put $\rho=(I-P_h)v$ and recall (e.g., from [E1]) that

$$\|h^{1-m}\nabla\rho\| + \|h^{-m}\rho\| \leqq C(\|h^{1-m}\nabla(v-\chi)\| + \|h^{-m}(v-\chi)\|),$$
$$\forall v\in H_0^1(\Omega), \quad \forall\chi\in S_h, \quad m=1,2.$$

Taking χ to be the usual Lagrangian interpolant v_i of v we obtain from above

$$|(\nabla w,\nabla\rho)| \leqq CD_{h,2}(w)(\|h^{-1}\nabla(v-v_i)\| + \|h^{-2}(v-v_i)\|) \leqq \beta_2 D_{h,2}(w)\|D^2v\|$$
$$\forall v\in H_0^1(\Omega)\cap H^2(\Omega),$$

for v only in $H_0^1(\Omega)$ the usual interpolant v_i does not necessarily exist, since the functions in $H_0^1(\Omega)$ do not have point values. However, we may then replace v_i by an interpolant $\tilde v_i$ of v defined from local average values of v around each nodal point (e.g., as in [E3]) to obtain

$$|(\nabla w,\nabla\rho)| \leqq CD_{h,1}(w)(\|\nabla(v-\tilde v_i)\| + \|h^{-1}(v-\tilde v_i)\|) \leqq \beta_1 D_{h,1}(w)\|Dv\|$$
$$\forall v\in H_0^1(\Omega).$$

Finally, to prove (4.3) we note that by Cauchy's inequality

$$|(f,(I-P_h)v)| \leqq \|h^m f\|\,\|h^{-m}(I-P_h)v\|,$$

and the desired result now follows from the weighted norm estimates for the L_2-projection P_h given in, e.g., [E1]. This completes the proof of Lemma 4.3. □

5. Proof of Theorem 2.2. Equation (1.1) may be written in weak form as

$$(\nabla u,\nabla v) = (f,v) \quad \forall v\in H_0^1(\Omega).$$

Subtracting $(\nabla u_h,\nabla v)$ from each side of this equation and using (1.4) we get

$$(\nabla(u-u_h),\nabla v) = (f,v) - (\nabla u_h,\nabla v)$$
$$= (f,(I-P_h)v) - (\nabla u_h,\nabla(I-P_h)v) \quad \forall v\in H_0^1(\Omega).$$

By Lemma 4.3, we have for $v\in H_0^1(\Omega)$

$$|(f,(I-P_h)v)| \leqq \alpha_1\|hf\|\,\|\nabla v\|,$$
$$|(\nabla u_h,\nabla(I-P_h)v)| \leqq \beta_1 D_{h,1}(u_h)\|\nabla v\|,$$

from which (2.2a) now follows at once by taking $v=u-u_h$. For the proof of (2.2b) we have similarly for $v\in H_0^1(\Omega)\cap H^2(\Omega)$,

$$|(f,(I-P_h)v)| \leqq \alpha_2\|h^2 f\|\,\|D^2v\|,$$
$$|(\nabla u_h,\nabla(I-P_h)v)| \leqq \beta_2 D_{h,2}(u_h)\|D^2v\|.$$

We now take v to be the solution of the dual problem

$$-\Delta v = u - u_h \quad \text{in } \Omega, \qquad v = 0 \quad \text{on } \partial\Omega,$$

and apply Lemma 4.1 to obtain

$$\begin{aligned} \|u-u_h\|^2 &= (u-u_h, -\Delta v) = (\nabla(u-u_h), \nabla v) \\ &= (f, (I-P_h)v) - (\nabla u_h, \nabla(I-P_h)v) \\ &\leqq (\alpha_2 \|h^2 f\| + \beta_2 D_{h,2}(u_h)) \|D^2 v\| \\ &\leqq (\alpha_2 \|h^2 f\| + \beta_2 D_{h,2}(u_h)) \|u-u_h\|, \end{aligned}$$

which proves (2.2b). This completes the proof of Theorem 2.2. □

6. Proof of Theorem 2.3. In this section we give the proof of the a priori estimates (2.5) in the case (2.4a). The case (2.4b) is considered in [E1]. The proof is naturally divided into the following steps indicating the overall structure of the argument:

(a) An error representation formula using duality.
(b) A strong stability estimate for the discrete dual problem.
(c) Choice of interpolant and end of proof of (2.5a).
(d) Choice of interpolant and end of proof of (2.5b).

6.1. An error representation formula using duality. Given a discrete time level $t_N > 0$ we write the set of equations (1.5) determining the discrete solution $U \in V = \{v: v|_{\Omega\times I_n} \in V_{qn}, n = 1, \cdots, N\}$ up to time t_N in compact form as

$$B(U, v) = (u_0, v_0^+) + \int_0^{t_N} (f, v)\, dt \quad \forall v \in V,$$

where

$$B(w, v) \equiv \sum_{n=1}^{N} \int_{I_n} \{(w_t, v) + (\nabla w, \nabla v)\}\, dt + (w_0^+, v_0^+) + \sum_{n=1}^{N-1} ([w]_n, v_n^+).$$

Clearly, in view of (1.2), we then have for the error $e \equiv u - U$

$$B(u-U, v) = 0 \quad \forall v \in V. \tag{6.1}$$

Given e_N^- let $Z \in V$ be defined by

$$B(v, Z) = (v_N^-, e_N^-) \quad \forall v \in V. \tag{6.2}$$

Note that Z is the discrete solution corresponding to our above method of discretization of the associated dual (or "backward") problem

$$\begin{aligned} -z_t - \Delta z &= 0 \quad \text{in } \Omega, t < T, \\ z(T) &= e_N^- \quad \text{in } \Omega, \end{aligned}$$

with "initial" data given at time $T = t_N$. This follows from the fact that the bilinear form $B(\cdot, \cdot)$, after time integration by parts, can also be written as

$$B(w, v) = \sum_{n=1}^{N} \int_{I_n} \{(w, -v_t) + (\nabla w, \nabla v)\}\, dt + \sum_{n=1}^{N-1} (w_n^-, -[v]_n) + (w_N^-, v_N^-). \tag{6.3}$$

In view of (6.2) and (6.1) we have for any $v \in V$, with $u_N \equiv u(t_N)$,

$$\begin{aligned} \|e_N^-\|^2 &= (u_N - v_N^-, e_N^-) + (v_N^- - U_N^-, e_N^-) \\ &= (u_N - v_N^-, e_N^-) + B(v-U, Z) \\ &= (u_N - v_N^-, e_N^-) + B(v-u, Z). \end{aligned} \tag{6.4}$$

Taking here $v \in V$ to be a suitable interpolant of u, we thus have a representation of the error at time T in terms of the truncation error $u - v$ and the discrete solution Z of the associated backward problem.

We will now derive a stability estimate for Z in terms of the initial data e_N^-. For convenience we establish this result by proving a corresponding estimate for the discrete forward problem (1.5) with $f = 0$. This estimate is analogous to the stability estimate (4.2) for the continuous forward problem (1.2).

6.2. The strong stability estimate.

LEMMA 6.1. *Assume that $S_{n-1} \subseteq S_n$, and that $\sigma k_n \leqq t_n$ for $n > 1$ where $\sigma > 1$ is a constant, and let U be the solution of* (1.5) *with $f \equiv 0$. Then for $N = 1, 2, \cdots$, we have*

$$\|U_N^-\|^2 + 2\int_0^{t_N} \|\nabla U\|^2\, dt + \sum_{n=0}^{N-1} \|[U]_n\|^2 = \|u_0\|^2, \tag{6.5}$$

$$\sum_{n=1}^{N} t_n \int_{I_n} \{\|U_t\|^2 + \|\Delta_n U\|^2\}\, dt + \sum_{n=2}^{N} t_n \|[U]_{n-1}\|^2 / k_n \leqq C\|u_0\|^2, \tag{6.6}$$

and

$$\sum_{n=1}^{N} \int_{I_n} \{\|U_t\| + \|\Delta_n U\|\}\, dt + \sum_{n=1}^{N-1} \|[U]_n\| \leqq C\left(\log \frac{T}{k_1} + 1\right)^{1/2} \|u_0\|, \tag{6.7}$$

where C only depends on σ and q.

Proof. In order to prove the basic stability estimate (6.5) we put $v = U$ in (1.5) to obtain

$$\int_{I_n} \{(U_t, U) + (\nabla U, \nabla U)\}\, dt + ([U]_{n-1}, U_{n-1}^+) = 0. \tag{6.8}$$

Here

$$\int_{I_n} (U_t, U)\, dt = \frac{1}{2}\|U_n^-\|^2 - \frac{1}{2}\|U_{n-1}^+\|^2,$$

and

$$([U]_{n-1}, U_{n-1}^+) = \tfrac{1}{2}\|U_{n-1}^+\|^2 + \tfrac{1}{2}\|[U]_{n-1}\|^2 - \tfrac{1}{2}\|U_{n-1}^-\|^2,$$

so that (6.5) follows at once by summation of (6.8) over n.

For the proof of (6.6) we put $v = -\Delta_n U$ in (1.5), to obtain

$$\frac{1}{2}\|\nabla U_n^-\|^2 + \int_{I_n} \|\Delta_n U\|^2\, dt + \frac{1}{2}\|\nabla U_{n-1}^+\|^2 = (U_{n-1}^-, -\Delta_n U_{n-1}^+), \tag{6.9}$$

where we have used the definition of Δ_n and evaluated the integral over I_n of the term involving U_t.

For $n = 1$ we deduce from (6.9) that for any $\varepsilon > 0$

$$\int_{I_n} \|\Delta_1 U\|^2\, dt \leqq \|u_0\|\, \|\Delta_1 U_0^+\| \leqq \frac{1}{4\varepsilon k_1}\|u_0\|^2 + \varepsilon k_1 \|\Delta_1 U_0^+\|^2.$$

Here $\|\Delta_1 U\|^2$ is a polynomial in t of degree at most $2q$ on I_1 so that

$$k_1 \|\Delta_1 U_0^+\|^2 \leqq C(q) \int_{I_n} \|\Delta_1 U\|^2\, dt.$$

Taking ε suitably small we thus conclude from above that

$$t_1 \int_{I_1} |\Delta_1 U\|^2\, dt \leqq C\|u_0\|^2. \tag{6.10}$$

For $n>1$ we write (6.9) in the form

$$\frac{1}{2}\|\nabla U_n^-\|^2+\int_{I_n}\|\Delta_n U\|^2\,dt+\frac{1}{2}\|[\nabla U]_{n-1}\|^2-\frac{1}{2}\|\nabla U_{n-1}^-\|^2=0,$$

where we have used our assumption $S_{n-1}\subseteq S_n$. Multiplying this identity by $2t_n$ we obtain

$$\begin{aligned}t_n\|\nabla U_n^-\|^2+2t_n\int_{I_n}\|\Delta_n U\|^2\,dt+t_n\|[\nabla U]_{n-1}\|^2-t_{n-1}\|\nabla U_{n-1}^-\|^2\\ =k_n\|\nabla U_{n-1}^-\|^2\leqq k_n(\|[\nabla U]_{n-1}\|+\|\nabla U_{n-1}^+\|)^2\\ \leqq \sigma k_n\|[\nabla U]_{n-1}\|^2+\frac{\sigma k_n}{\sigma-1}\|\nabla U_{n-1}^+\|^2\quad\text{for } \sigma>1.\end{aligned}$$

Hence, using our assumption $\sigma k_n\leqq t_n$, we conclude that

$$\begin{aligned}t_n\|\nabla U_n^-\|^2+2t_n\int_{I_n}\|\Delta_n U\|^2\,dt-t_{n-1}\|\nabla U_{n-1}^-\|^2&\leqq\frac{\sigma k_n}{\sigma-1}\|\nabla U_{n-1}^+\|^2\\ &\leqq C\int_{I_n}\|\nabla U\|^2\,dt,\end{aligned}\tag{6.11}$$

where in the last step we have again used an inverse estimate based on the fact that $\|\nabla U\|^2$ is a polynomial in t on I_n. By summation of (6.11) over n we find that

$$\begin{aligned}t_N\|\nabla U_N^-\|^2+2\sum_{n=2}^N t_n\int_{I_n}\|\Delta_n U\|^2\,dt&\leqq t_1\|\nabla U_1^-\|^2+C\int_{t_1}^{t_N}\|\nabla U\|^2\,dt\\ &\leqq C\int_0^{t_N}\|\nabla U\|^2\,dt\leqq C\|u_0\|^2,\end{aligned}\tag{6.12}$$

where in the last step we used (6.5).

In view of (6.10) and (6.12) it now suffices for the proof of (6.6) to show that

$$\int_{I_n}\|U_t\|^2\,dt\leqq C\int_{I_n}\|\Delta_n U\|^2\,dt\quad\text{for } n\geqq 1,\tag{6.13}$$

and

$$\|[U]_{n-1}\|^2/k_n\leqq C\int_{I_n}\|\Delta_n U\|^2\,dt\quad\text{for } n>1.\tag{6.14}$$

In order to do so we first put $v=(t-t_{n-1})U_t$ in (1.5), which yields

$$\begin{aligned}\int_{I_n}(t-t_{n-1})\|U_t\|^2\,dt&=\int_{I_n}(t-t_{n-1})(\Delta_n U,U_t)\,dt\\ &\leqq\left(\int_{I_n}(t-t_{n-1})\|\Delta_n U\|^2\,dt\right)^{1/2}\left(\int_{I_n}(t-t_{n-1})\|U_t\|^2\,dt\right)^{1/2}.\end{aligned}$$

Using another inverse estimate we deduce that

$$\begin{aligned}\int_{I_n}\|U_t\|^2\,dt&\leqq Ck_n^{-1}\int_{I_n}(t-t_{n-1})\|U_t\|^2\,dt\\ &\leqq Ck_n^{-1}\int_{I_n}(t-t_{n-1})\|\Delta_n U\|^2\,dt\leqq C\int_{I_n}\|\Delta_n U\|^2\,dt,\end{aligned}$$

which shows (6.13).

For the proof of (6.14) we put $v=[U]_{n-1}$ in (1.5), for which again we need our assumption $S_{n-1}\subseteq S_n$, and find that

$$\|[U]_{n-1}\|^2=-\int_{I_n}(U_t-\Delta_n U,[U]_{n-1})\,dt$$
$$\leqq k_n\int_{I_n}\{\|U_t\|^2+\|\Delta_n U\|^2\}\,dt+\frac{1}{2}\|[U]_{n-1}\|^2,$$

from which (6.14) follows if we take (6.13) into account. Finally (6.7) follows from (6.6) by applying Cauchy's inequality. This completes the proof of Lemma 6.1. □

We now apply the above stability lemma to the solution Z of the backward problem (6.2) under the assumption $S_n\subseteq S_{n-1}$ for $1<n\leqq N$ and $\sigma k_n\leqq t_N-t_{n-1}$ for $1\leqq n<N$, where $\sigma>1$. Note here that the latter condition is in fact somewhat weaker than the condition $k_n\leqq Ck_{n+1}$ used in the theorem. We then conclude from (6.5) and (6.7) that there is a constant C_* only depending on σ and q such that

$$\max_{t\in(0,t_N]}\|Z(t)\|+\sum_{n=1}^{N}\int_{I_n}\{\|Z_t\|+\|\Delta_n Z\|\}\,dt+\sum_{n=1}^{N-1}\|[Z]_n\|$$
$$\leqq C_*\left(\log\frac{T}{k_N}+1\right)^{1/2}\|e(T)\|. \tag{6.15}$$

6.3. Proof of (2.5a). We will first show that the desired error bound holds at the nodal points $t=t_N>0$. By Theorem 2.1 we have for the elliptic projection $\pi_n: H_0^1(\Omega)\to S_n$ defined by

$$(\nabla(\varphi-\pi_n\varphi),\nabla\psi)=0\quad\forall\psi\in S_n, \tag{6.16}$$

that

$$\|\varphi-\pi_n\varphi\|\leqq C\|h_n^2D^2\varphi\|\quad\forall\varphi\in H^2(\Omega)\cap H_0^1(\Omega). \tag{6.17}$$

We now take $v=\tilde u$ in the error representation (6.4), where $\tilde u|_{I_n}=\tilde I_n\pi_n u$ is the L_2 projection in time onto V_n of $\pi_n u$; that is, for $q=1$

$$\tilde u|_{I_n}=\tilde I_n\pi_n u$$
$$=k_n^{-1}\int_{I_n}\pi_n u\,ds+12k_n^{-3}(t-t_{n-1}-k_n/2)\int_{I_n}(s-t_{n-1}-k_n/2)\pi_n u\,ds, \tag{6.18}$$

and for $q=0$ the second term on the right-hand side of (6.18) is not present. With this choice of v (6.4) reduces to

$$\|e_N^-\|^2=(u_N-\tilde u_N^-,e_N^-)+\sum_{n=1}^{N}\int_{I_n}(u-\pi_n u,Z_t)\,dt$$
$$+\sum_{n=1}^{N-1}(u_n-\tilde u_n^-,[Z]_n)-(u_N-\tilde u_N^-,Z_N^-), \tag{6.19}$$

where we have used (6.16) and the fact that

$$\int_{I_n}(\pi_n u-\tilde u,v)\,dt=0\quad\forall v\in V_n,$$

and thus, in particular, for $v=Z_t$ and $v=\Delta_n Z$.

Since the L_2-projection operator $\tilde{I}_n$ is bounded with respect to the norm $\|\cdot\|_{I_n}$, we have

$$\|u-\tilde{u}\|_{I_n} \leqq \|u-\tilde{I}_n u\|_{I_n} + \|\tilde{I}_n(u-\pi_n u)\|_{I_n}$$

$$\leqq C\left(\min_{j\leqq q+1} k_n^j \|u_t^{(j)}\|_{I_n} + \|h_n^2 D^2 u\|_{I_n}\right), \tag{6.20}$$

where the bound for $u-\tilde{I}_n u$ follows at once from the Taylor expansion

$$u(t) = u(t_n) + \int_{t_n}^{t} u_t(s)\,ds$$

$$= u(t_n) + (t-t_n)u_t(t_n) + \int_{t_n}^{t} (t-s)u_{tt}(s)\,ds,$$

since $\tilde{I}_n$ equals the identity on the polynomial part of $u(t)$ of degree q.

From (6.19) we thus obtain

$$\|e_N^-\|^2 \leqq C \max_{1\leqq n\leqq N}\left(\min_{j\leqq q+1} k_n^j \|u_t^{(j)}\|_{I_n} + \|h_n^2 D^2 u\|_{I_n}\right)$$

$$\cdot\left(\|e_N^-\| + \sum_{n=1}^{N}\int_{I_n} \|Z_t\|\,dt + \sum_{n=1}^{N-1} \|[Z]_n\| + \|Z_N^-\|\right),$$

and conclude in view of (6.15) that

$$\|e_N^-\| \leqq C_*\left(\log\frac{t_N}{k_N}+1\right)^{1/2} \max_{1\leqq n\leqq N} C\left(\min_{j\leqq q+1} k_n^j \|u_t^{(j)}\|_{I_n} + \|h_n^2 D^2 u\|_{I_n}\right). \tag{6.21}$$

We shall now show by a local error analysis that (6.21) can be extended to hold uniformly in t on I_N. For this purpose we write the error $e=u-U$ as $e=\eta+\theta$ with $\eta=u-\tilde{u}$ and $\theta=\tilde{u}-U\in V$, where $\tilde{u}$ is defined as above. In view of (6.20) it then suffices to find an appropriate estimate for $\|\theta\|_{I_n}$. For $q=0$ we have

$$\|\theta\|_{I_n} = \|\theta_n^-\| \leqq \|e_n^-\| + \|\eta_n^-\|,$$

so that the desired estimate follows at once from (6.21) and (6.20). It thus suffices to consider the case $q=1$. If again we write $e=\eta+\theta$ and rearrange terms and integrate by parts in the local error equation

$$\int_{I_n} \{(e_t, v) + (\nabla e, \nabla v)\}\,dt + ([e]_{n-1}, v_{n-1}^+) = 0 \quad \forall v\in V_n,$$

we obtain the following equation for θ:

$$\int_{I_n} \{(\theta_t, v) + (\nabla\theta, \nabla v)\}\,dt + (\theta_{n-1}^+, v_{n-1}^+)$$

$$= (e_{n-1}^-, v_{n-1}^+) + \int_{I_n} (\eta, v_t)\,dt - \int_{I_n} (\nabla\eta, \nabla v)\,dt - (\eta_n^-, v_n^-) \tag{6.22}$$

$$= (e_{n-1}^-, v_{n-1}^+) + \int_{I_n} (u-\pi_n u, v_t)\,dt - (\eta_n^-, v_n^-) \quad \forall v\in V_n,$$

where we have again used (6.16) and the properties of $\tilde{u}$.

We choose $v=\theta$ in (6.22) to find that

$$\begin{aligned}\frac{1}{2}\|\theta_n^-\|^2+\int_{I_n}\|\nabla\theta\|^2\,dt+\frac{1}{2}\|\theta_{n-1}^+\|^2&=(e_{n-1}^-,\theta_{n-1}^+)+\int_{I_n}(u-\pi_n u,\theta_t)\,dt-(\eta_n^-,\theta_n^-)\\ &\leqq\|e_{n-1}^-\|^2+\frac{1}{4}\|\theta_{n-1}^+\|^2+Ch_n^4\|D^2u\|_{I_n}^2\\ &\quad+\frac{1}{8}(\|\theta_{n-1}^+\|^2+\|\theta_n^-\|^2)+\|\eta_n^-\|^2+\frac{1}{4}\|\theta_n^-\|^2,\end{aligned}\tag{6.23}$$

where we have used the fact that

$$\int_{I_n}\|\theta_t\|\,dt=\|\theta_n^- -\theta_{n-1}^+\|\leqq\|\theta_n^-\|+\|\theta_{n-1}^+\|,\tag{6.24}$$

since θ is piecewise linear in t. From (6.24) and (6.23) we deduce that

$$\begin{aligned}\|\theta\|_{I_n}&\leqq\|\theta_n^-\|+\int_{I_t}\|\theta_t\|\,dt\\ &\leqq C(\|e_{n-1}^-\|+\|\eta_n^-\|+\|h_n^2D^2u\|_{I_n}).\end{aligned}\tag{6.25}$$

Together our estimates (6.25), (6.20) and (6.21) now prove (2.5a) also for $q=1$. □

6.4. Proof of (2.5b). Again we start from the error representation formula (6.4) but this time we choose $v=\tilde{\tilde{u}}$ where $\tilde{\tilde{u}}$ is defined so that $(\tilde{\tilde{u}})_n^-=\pi_n u_n$ and so that $\pi_n u-\tilde{\tilde{u}}$ has mean value zero over I_n (recall that now $q=1$), i.e., we take

$$\tilde{\tilde{u}}|_{I_n}:=\tilde{\tilde{I}}_n\pi_n u:=\pi_n u_n+(t-t_n)\frac{2}{k_n^2}\int_{I_n}\pi_n(u-u_n)\,ds.\tag{6.26}$$

With this particular choice of v (6.4) reduces to

$$\begin{aligned}\|e_N^-\|^2&=(u_N-\pi_N u_N,e_N^-)+\sum_{n=1}^N\int_{I_n}(u-\pi_n u,Z_t)\,dt+\sum_{n=1}^N\int_{I_n}(\nabla(\pi_n u-\tilde{\tilde{u}}),\nabla Z)\,dt\\ &\quad+\sum_{n=1}^{N-1}(u_n-\pi_n u_n,[Z]_n)-(u_N-\pi_N u_N,Z_N^-),\end{aligned}\tag{6.27}$$

where in the first sum we have used the fact that $\pi_n u-\tilde{\tilde{u}}$, having mean value zero over I_n, is orthogonal to Z_t being constant in t on I_n, whereas in the second sum we have again used (6.16).

For the terms in the second sum on the right in (6.27) we have

$$(\nabla(\pi_n u-\tilde{\tilde{u}}),\nabla Z)=(\nabla(\pi_n u-\tilde{\tilde{u}}),\nabla Z_n^-)+(\nabla(\pi_n u-\tilde{\tilde{u}}),(t-t_n)\nabla Z_t),$$

so that by our choice of $\tilde{\tilde{u}}$,

$$\begin{aligned}\left|\int_{I_n}(\nabla(\pi_n u-\tilde{\tilde{u}}),\nabla Z)\,dt\right|&=\left|\int_{I_n}(\nabla(\pi_n u-\tilde{\tilde{u}}),(t-t_n)\nabla Z_t)\,dt\right|\\ &\leqq k_n\|\Delta_n(\pi_n u-\tilde{\tilde{u}})\|_{I_n}\int_{I_n}\|Z_t\|\,dt.\end{aligned}\tag{6.28}$$

Using Taylor expansions, we easily find that

$$\|\Delta_n(\pi_n u-\tilde{\tilde{u}})\|_{I_n}\leqq Ck_n^2\|\Delta_n\pi_n u_t^{(2)}\|_{I_n}.\tag{6.29}$$

Finally, we note that for any $\varphi \in H^2(\Omega) \cap H_0^1(\Omega)$ we have

$$(-\Delta_n \pi_n \varphi, \psi) = (\nabla \pi_n \varphi, \nabla \psi) = (\nabla \varphi, \nabla \psi) = (-\Delta \varphi, \psi) \quad \forall \psi \in S_n,$$

from which we deduce at once by taking $\psi = -\Delta_n \pi_n \varphi$ that

$$\|\Delta_n \pi_n \varphi\| \leqq \|\Delta \varphi\| \quad \forall \varphi \in H^2(\Omega) \cap H_0^1(\Omega). \tag{6.30}$$

By another twist of the argument we can replace (6.28) by

$$\begin{aligned} \left| \int_{I_n} (\nabla(\pi_n u - \tilde{\tilde{u}}), \nabla Z)\, dt \right| &= \left| \int_{I_n} (\pi_n u - \tilde{\tilde{u}}, -\Delta_n Z)\, dt \right| \\ &\leqq (\|u - \pi_n u\|_{I_n} + \|u - \tilde{\tilde{I}}_n u\|_{I_n} \\ &\quad + \|\tilde{\tilde{I}}(u - \pi_n u)\|_{I_n}) \int_{I_n} \|\Delta_n Z\|\, dt \\ &\leqq C\left(\|h_n^2 D^2 u\|_{I_n} + \min_{j \leqq q+1} k_n^j \|u_t^{(j)}\|_{I_n} \right) \int_{I_n} \|\Delta_n Z\|\, dt. \end{aligned} \tag{6.31}$$

If we now put everything together and use (6.26) through (6.31) together with (6.17) and (6.15) we find (with $u_t^{(3)} = \Delta u_{tt}$)

$$\begin{aligned} \|e_N^-\|^2 &\leqq C \max_{1 \leqq n \leqq N} \left(\|h_n^2 D^2 u\|_{I_n} + \min_{j \leqq q+2} k_n^j \|u_t^{(j)}\|_{I_n} \right) \\ &\quad \cdot \left(\|e_N^-\| + \sum_{n=1}^{N} \int_{I_n} \|Z_t\|\, dt + \sum_{n=1}^{N-1} \|[Z]_n\| + \|Z_N^-\| \right) \\ &\leqq \|e_N^-\| C_* \left(\log \frac{T}{k_N} + 1 \right)^{1/2} \max_{1 \leqq n \leqq N} C\left(\min_{j \leqq q+2} k_n^j \|u_t^{(j)}\|_{I_n} + \|h_n^2 D^2 u\|_{I_n} \right), \end{aligned}$$

from which the desired estimate (2.5b) follows. This completes the proof of Theorem 2.3. □

7. Proof of Theorem 2.4. In order to get a representation of the error $e_N^- = u_N - U_N^-$ we introduce the dual problem

$$\begin{aligned} &z_t(x, t) + \Delta z(x, t) = 0, \qquad x \in \Omega, \quad t < t_N, \\ &z(x, t) = 0, \qquad x \in \partial\Omega, \quad t < t_N, \\ &z(x, t_N) = e_N^-(x), \qquad x \in \Omega, \end{aligned} \tag{7.1}$$

and use (6.3) to conclude that

$$\|e_N^-\|^2 = B(e, z) = B(u, z) - B(U, z).$$

Since

$$B(u, z) = (u_0, z_0^+) + \int_0^{t_N} (f, z)\, dt,$$

and by subtracting a zero quantity obtained from (1.5), we get the following error representation in terms of U, z and data

$$
\begin{aligned}
\|e_N^-\|^2 &= B(U, v-z)+(u_0, z_0^+ - v_0^+)+\int_0^{t_N} (f, z-v)\,dt \\
&= \sum_{n=1}^{N} \int_{I_n} \{(U_t, v-z)+(\nabla U, \nabla(v-z))\}\,dt \\
&\quad + \sum_{n=0}^{N-1} ([U]_n, v_n^+ - z_n^+)+\int_0^{t_N} (f, z-v)\,dt \\
&= \mathrm{I}+\mathrm{II}+\mathrm{III} \quad \forall v \in V,
\end{aligned} \tag{7.2}
$$

where $U_0^- = u_0$.

We now first consider the case $q=0$. We then choose $v \in V$ in (7.2) to be the L_2 projection of z, i.e., $v|_{\Omega \times I_n} \in V_n$ is the $L_2(\Omega \times I_n)$ projection of z for all n, so that in particular

$$
\int_{I_n} (U_t, v-z)\,dt = 0, \qquad 1 \leqq n \leqq N.
$$

Moreover,

$$
\int_{I_n} (\nabla U, \nabla(v - P_n z))\,dt = \int_{I_n} (-\Delta_n U, v - P_n z)\,dt = 0, \qquad 1 \leqq n \leqq N,
$$

and consequently the sum I reduces to

$$
\mathrm{I} = \sum_{n=1}^{N} \int_{I_n} (\nabla U, \nabla(P_n - I)z)\,dt = \sum_{n=1}^{N} \left(\nabla U_n, \nabla(P_n - I)\int_{I_n} z\,dt\right), \tag{7.3}
$$

where $P_n = P_{h_n}$ is the L_2-projection onto S_n and $U_n = U|_{I_n}$. Using (4.4) and (4.1) and the fact that

$$
\Delta \int_{I_n} z\,dt = \int_{I_n} \Delta z\,dt = \int_{I_n} z_t\,dt = z(t_n) - z(t_{n-1}),
$$

we find that

$$
\begin{aligned}
|\mathrm{I}| &\leqq \beta_2 \sum_{n=1}^{N} D_{n,2}(U_n) \left\| \Delta \int_{I_n} z\,dt \right\| \\
&\leqq \beta_2 \max_{1 \leqq n \leqq N} D_{n,2}(U_n)\left(\int_0^{t_{N-1}} \|z_t\|\,dt + 2\|z\|_{I_N}\right).
\end{aligned} \tag{7.4}
$$

Further, we have by (4.3) and Lemma 4.1

$$
|([U_{n-1}], (P_n - I)z_{n-1}^+)| \leqq \alpha_2 \|h_n^2[U_{n-1}]\|^* \|\Delta z_{n-1}^+\|, \tag{7.5}
$$

noting that the left-hand side is zero if $S_{n-1} \subset S_n$. By obvious stability and approximation properties of the $L_2(I_n)$ projections onto the set of constant functions on I_n, we also have

$$
\|v - P_n z\|_{I_n} \leqq \|P_n z\|_{I_n} \leqq \|z\|_{I_n} \tag{7.6}
$$

and

$$
\|v - P_n z\|_{I_n} \leqq \int_{I_n} \|P_n z_t\|\,dt \leqq \int_{I_n} \|z_t\|\,dt. \tag{7.7}
$$

We thus conclude that

$$\begin{aligned}|\mathrm{II}| &\leqq \alpha_2 \max_{1\leqq n\leqq N} \|h_n^2[U_{n-1}]/k_n\|^* \sum_{n=1}^{N} k_n\|\Delta z_{n-1}^+\| \\ &\quad + \max_{1\leqq n\leqq N} \|[U_{n-1}]\|\left(\int_0^{t_{N-1}} \|z_t\|\,dt + \|z\|_{I_N}\right).\end{aligned} \tag{7.8}$$

It remains to estimate the term III. Using (7.6) and (7.7) we find

$$\begin{aligned}\left|\sum_{n=1}^{N}\int_{I_n} (f, P_n z - v)\,dt\right| &\leqq \sum_{n=1}^{N}\int_{I_n} \|f\|\,dt\|P_n z - v\|_{I_n} \\ &\leqq \max_{1\leqq n\leqq N}\int_{I_n}\|f\|\,dt\left(\int_0^{t_{N-1}}\|z_t\|\,dt + \|z\|_{I_N}\right).\end{aligned} \tag{7.9}$$

We also have

$$\left|\int_{I_N} (f, (I-P_N)z)\,dt\right| \leqq \int_{I_N} \|f(t)\|\,dt\|z\|_{I_N}, \tag{7.10}$$

and by (4.1)

$$\begin{aligned}\left|\sum_{n=1}^{N-1}\int_{I_n} (f, (I-P_n)z)\,dt\right| &\leqq \alpha_2 \sum_{n=1}^{N-1}\int_{I_n} \|h_n^2 f\|\,\|\Delta z\|\,dt \\ &\leqq \alpha_2 \max_{1\leqq n\leqq N-1} \|h_n^2 f\|_{I_n}\int_0^{t_{N-1}} \|\Delta z\|\,dt.\end{aligned} \tag{7.11}$$

We now note that by Lemma 4.2 and Cauchy's inequality we have

$$\begin{aligned}\sum_{n=1}^{N-1} k_n\|w_{n-1}^+\| &\leqq \int_0^{t_{N-1}} \|w\|\,dt \\ &\leqq \left(\int_0^{t_{N-1}} (t_N - t)^{-1}\,dt\right)^{1/2}\left(\int_0^{t_N} (t_N - t)\|w\|^2\,dt\right)^{1/2} \\ &\leqq \frac{1}{4}\left(\log\frac{t_N}{k_N}\right)^{1/2}\|e_N^-\| = L_N\|e_N^-\|, \qquad w = z_t \text{ or } w = \Delta z.\end{aligned} \tag{7.12a}$$

$$k_N\|\Delta z_{N-1}^+\| \leqq \exp(-1)\|e_N^-\|. \tag{7.12b}$$

Together, our estimates (7.4) and (7.8)-(7.12) now prove the desired a posteriori error estimate (2.7a) for $q=0$.

Let us now turn to the case $q=1$ and the proof of (2.7b). Again we have to estimate the terms I, II, and III in the error representation (7.2). In order to estimate I, let $\zeta(x,t) = \int_{t_n}^{t} z(x,s)\,ds$, $t \in I_n$. Using integration by parts we obtain

$$\begin{aligned}\int_{I_n} (\nabla U, \nabla(P_n - I)z)\,dt &= \left(\nabla U_{n-1}^+, \nabla(P_n - I)\int_{I_n} z\,dt\right) \\ &\quad - \left(\nabla(U_n^- - U_{n-1}^+), \nabla(P_n - 1)\int_{I_n}\frac{\zeta\,dt}{k_n}\right).\end{aligned}$$

From (4.4) and (4.1) we thus get

$$\left|\int_{I_n} (\nabla U, \nabla(P_n - I)z)\,dt\right| \leqq \beta_2 \max_{t\in I_n} D_{n,2}(U(t))\left(\left\|\Delta\int_{I_n} z\,dt\right\| + \left\|\Delta\int_{I_n}\frac{\zeta\,dt}{k_n}\right\|\right),$$

where, as above,

$$\Delta \int_{I_n} z\, dt = z(t_n) - z(t_{n-1}) = \int_{I_n} z_t\, dt,$$

$$\Delta \int_{I_n} \zeta\, dt = \int_{I_n} \Delta\zeta\, dt = \int_{I_n} \int_{t_n}^{t} \Delta z\, ds\, dt = \int_{I_n} (z(t) - z(t_n))\, dt.$$

We thus end up with the estimate

$$|\mathrm{I}| \leqq 2\beta_2 \max_{1\leqq n\leqq N} \max_{t\in I_n} D_{n,2}(U(t)) \left(\int_0^{t_{N-1}} \|z_t\|\, dt + \|z\|_{I_N} \right). \tag{7.13}$$

In order to estimate II we shall again use (7.5) but replace the use of (7.6) and (7.7) as follows: we have that

$$\begin{aligned}([U]_{n-1}, v_{n-1}^+ - P_n z_{n-1}^+) &= (U_{n-1}^+ - P_n U_{n-1}^-, v_{n-1}^+ - P_n z_{n-1}^+)\\ &= (\Delta_n (U_{n-1}^+ - P_n U_{n-1}^-), \Delta_n^{-1}(v_{n-1}^+ - P_n z_{n-1}^+)),\end{aligned}$$

where Δ_n^{-1} is the inverse of Δ_n onto S_n. By the approximation properties of v we have

$$\|\Delta_n^{-1}(v - P_n z)\|_{I_n} \leqq \gamma_i k_n^{i-1} \int_{I_n} \|\Delta_n^{-1} P_n z_t^{(i)}\|\, dt, \qquad i = 1, 2, \tag{7.14a}$$

$$\|\Delta_n^{-1}(v - P_n z)\|_{I_n} \leqq \gamma_0 \|\Delta_n^{-1} P_n z\|_{I_n}\, dt, \tag{7.14b}$$

where the γ_i are absolute constants. But here $P_n z_{tt} = P_n \Delta z_t = \Delta_n z_t$ and $P_n z_t = \Delta_n z$, so that using also (7.5), we get

$$\begin{aligned}|\mathrm{II}| \leqq& \sum_{n=1}^{N} (\|\Delta_n (U_{n-1}^+ - P_n U_{n-1}^-)\|\, \|\Delta_n^{-1}(v - P_n z)\|_{I_n}\\ &+ \alpha_2 \|h_n^2 [U_{n-1}]\|^* \|\Delta_n z_{n-1}^+\|)\\ \leqq& \max_{1\leqq n\leqq N} k_n \|\Delta_n (U_{n-1}^+ - P_n U_{n-1}^-)\| \left(\gamma_2 \int_0^{t_{N-1}} \|z_t\|\, dt + \gamma_1 \|z\|_{I_N} \right)\\ &+ \alpha_2 \left(\max_{1\leqq n\leqq N} \|h_n^2 [U_{n-1}]/k_n\|^* \sum_{n=1}^{N} k_n \|\Delta z_{n-1}^+\| \right).\end{aligned}$$

It now remains to estimate III. By our choice of v we have

$$\int_{I_n} (f, P_n z - v)\, dt = \int_{I_n} (f - \bar{f}, P_n z - v)\, dt,$$

where $\bar{f}$ is, say, the linear (in time) interpolant of f which equals f at t_n and t_{n-1}. (Note that $\bar{f}$ need not be projected onto S_n since both $P_n z$ and v belong to S_n.) This gives

$$\begin{aligned}\left| \sum_{n=1}^{N} \int_{I_n} (f, P_n z - v)\, dt \right| &\leqq \sum_{n=1}^{N} \int_{I_n} \|f - \bar{f}\|\, dt \|v - P_n z\|_{I_n}\\ &\leqq \gamma_3 \sum_{n=1}^{N} k_n^2 \int_{I_n} \|f_{tt}\|\, dt \|P_n z - v\|_{I_n} \qquad (7.15)\\ &\leqq \gamma_3 \max_{1\leqq n\leqq N} k_n^2 \int_{I_n} \|f_{tt}\|\, dt \left(\gamma_1 \int_0^{t_{N-1}} \|z_t\|\, dt + \gamma_0 \|z\|_{I_N} \right).\end{aligned}$$

Further, using (4.1) we get

$$\left|\sum_{n=1}^{N-1}\int_{I_n}(f,(I-P_n)z)\,dt\right|\leqq \alpha_2\max_{1\leqq n\leqq N-1}|h_n^2 f|_{I_n}\int_0^{t_{N-1}}\|\Delta z\|\,dt.$$

We finally need a separate estimate for the Nth subinterval I_N. We have

$$\left|\int_{I_N}(f-\bar f,(I-P_N)z)\,dt\right|\leqq \gamma_3 k_N^2\int_{I_N}\|f_{tt}\|\,dt\|z\|_{I_N},$$

and, with ζ defined as above,

$$\begin{aligned}\left|\int_{I_N}(\bar f,(I-P_N)z)\,dt\right| &\leqq \left|\left(f(t_{N-1}),(I-P_N)\int_{I_N}z\,dt\right)\right|\\ &\quad+\left|\left(f(t_N)-f(t_{N-1}),(I-P_N)\int_{I_N}\frac{\zeta\,dt}{k_N}\right)\right|\\ &\leqq \alpha_2\|h_N^2 f\|_{I_N}\left(\left\|\Delta\int_{I_N}z\,dt\right\|+\left\|\Delta\int_{I_N}\zeta\,dt\right\|\Big/k_N\right)\\ &\leqq 2\alpha_2\|h_N^2 f\|_{I_N}\|z\|_{I_N}.\end{aligned}$$

Together our estimates starting from (7.12) prove (2.7b) and the proof of Theorem 2.4 is thus complete. □

8. Proof of Theorem 3.3. In this section we prove Theorem 3.3 and Theorem 3.4 indicating that the algorithm (3.8) is efficient (and thus in particular operative). We shall use the following auxiliary results, the proofs of which are given below. A proof of Lemma 8.1 in the case of a quasi-uniform mesh is given in [BOP].

LEMMA 8.1. *If Ω is convex then there exists a constant C depending only on c_1 and c_2 such that if μ is small enough and $(h, T_h, S_h)\in\Sigma$, then*

$$\|D_h^2 v\|\leqq C\|\Delta_h v\|\quad \forall v\in S_h.$$

LEMMA 8.2. *Let $q=1$, let $S_n=S_{N^*}$ for $N^*\leqq n\leqq N$, and suppose for some constant γ that $k_n\leqq\gamma k_{n+1}$. Then, there exists a constant C such that with $\theta|_{I_N}=U|_{I_N}-\tilde I_N\pi_N u$,*

$$k_N\|\Delta_N\theta_N^-\|\leqq C\left[\max_{N^*\leqq n\leqq N}E_{2n}(u)+\frac{k_N}{t_N-t_{N^*-1}}\|e_{N^*-1}^-\|\right], \tag{8.1a}$$

where $\tilde I_N$ is the L_2-projection defined by (6.18) *and π_N the elliptic projection onto S_N defined by* (6.16). *Further,*

$$k_N\|\Delta_N\theta_N\|_{I_N}\leqq C(E_{2N}(u)+\|e_{N-1}^-\|+\|e_N^-\|). \tag{8.1b}$$

Proof of Theorem 3.3. We first consider the case $q=0$. To bound the main space discretization term in $\mathscr{E}_{0N}(U)$ in terms of the a priori quantities we argue as follows: By an inverse estimate we have, letting $u_i(\cdot,t)\in S_N$ be the usual Lagrangian interpolant of $u(\cdot,t)$,

$$\begin{aligned}\frac{1}{C}D_{N,2}(U_N) &\leqq \|h_N^2 D_N^2 U\|_{I_N}\leqq\|h_N^2 D_N^2(U-u_i)\|_{I_N}+\|h_N^2 D_N^2 D_N^2 u_i\|_{I_N}\\ &\leqq \|U-u_i\|_{I_N}+C\|h_N^2 D^2 u\|_{I_N}\leqq\|u-U\|_{I_N}+\|u-u_i\|_{I_N}+C\|h_N^2 D^2 u\|_{I_N}\\ &\leqq CL\max_{n\leqq N}E_{0n}(u),\end{aligned}$$

where Theorem 2.3 was used to bound $\|u-U\|_{I_N}$. For the remaining part

$\|h_N^2[U]_{N-1}/k_N\|^*$ of the space discretization error, we have since $\bar{h}_N^2 \leqq Ck_N$,

$$\|h_N^2[U]_{N-1}/k_N\| \leqq C\|[U]_{N-1}\| \leqq C\|[U-u]_{N-1}\|$$
$$\leqq C(\|e\|_{I_N} + \|e\|_{I_{N-1}})$$
$$\leqq CL \max_{n \leqq N} E_{0n}(u),$$

which also proves the desired bound for the time discretization term $\|[U]_{N-1}\|$. This completes the proof of Theorem 3.3 in the case $q=0$ since the f-terms are not included in the modified $\mathscr{E}_{q0}(U)$.

We now consider the case $q=1$. By Lemma 8.1 and (8.1b) we have using the assumption $\bar{h}_N^2 \leqq Ck_N$ and with θ as in the proof of Lemma 8.2:

$$\|h_N^2 D_N^2 U\|_{I_N} = \|h_N^2 D_N^2 \theta\|_{I_N} + \|h_N^2 D_N^2 \tilde{I}_N \pi_N u\|_{I_N}$$
$$\leqq C(k_N\|\Delta_N \theta\|_{I_N} + \|h_N^2 D_N^2 \pi_N u\|_{I_N})$$
$$\leqq C(k_N\|\Delta_N \theta\|_{I_N} + \|h_N^2 D^2 u\|_{I_N})$$
$$\leqq C(E_{2N}(u) + \|e_{N-1}^-\| + \|e_N^-\|)$$
$$\leqq CL \max_{n \leqq N} E_{2n}(u),$$

which proves the desired estimate for the main part of the space discretization error. The remaining part $\|h_N^2[U]_{N-1}/k_N\|^*$ only occurs if $S_N \not\supseteq S_{N-1}$ and is estimated as follows:

$$\|h_N^2[U]_{N-1}/k_N\| \leqq C\|[U]_{N-1}\| \leqq C\|[U-u]_{N-1}\|$$
$$\leqq C\|e_{N-1}^-\| + C\|e_{N-1}^+\|$$
$$\leqq CL \max_{n \leqq N-1} E_{2n}(u) + CL\hat{E}_{2N}(u),$$

where we used (6.25) to estimate $\|e_{N-1}^+\|$ in terms of $\|e_{N-1}^-\|$ and the interpolation error on I_N adding a hat on $E_{2N}(u)$.

Next, we turn to the time discretization term. If $S_N = S_{N-1}$, then we may consider this term to be given by $k_N\|\Delta_N[U]_{N-1}\|$ and estimate it as follows using (8.1b) and Theorem 2.3

$$k_N\|\Delta_N[U]_{N-1}\| \leqq k_N\|\Delta_N[\theta]_{N-1}\| + k_N\|\Delta_N[\tilde{I}_N \pi_N u]_{N-1}\|$$
$$\leqq k_N(\|\Delta_N \theta\|_{I_N} + \|\Delta_{N-1}\theta\|_{I_{N-1}}) + C \max_{I_N \cup I_{N-1}} k_N^3 \|\Delta u_t^{(2)}\|$$
$$\leqq CL \max_{n \leqq N} E_{2n}(u).$$

Finally, if $S_N \not\supseteq S_{N-1}$ then we may consider the time discretization term to be given by $\|[U]_{N-1}\|$, which is estimated as above. This completes the proof of Theorem 3.3. $\square$

Proof of Theorem 3.4. We have, recalling (3.11) and using Lemmas 8.1 and 8.2,

$$\frac{\delta}{CL} = \|h_N^2 D_N^2 U\|_{I_N} \leqq \|h_N^2 D_N^2 \theta\|_{I_N} + \|h_N^2 D_N^2 \tilde{I}_N \pi_N u\|_{I_N}$$
$$\leqq Ck_N\|\Delta_N \theta\|_{I_N} + C\|h_N^2 D^2 u\|_{I_N}$$
$$\leqq Ck_N(\|\Delta_N \theta_N^-\| + \|\Delta_N[\theta]_{N-1}\| + \|\Delta_N \theta_{N-1}^-\|) + C\|h_N^2 D^2 u\|_{I_N}$$
$$\leqq C\left(\max_{N^* \leqq n \leqq N} E_{2n}(u) + \frac{k_N}{t_{N-1} - t_N^*}\|e_{N^*-1}^-\|\right) + Ck_N\|\Delta_N[U]_{N-1}\|$$
$$+ Ck_N\|\Delta_N[\tilde{I}_N \pi_N u]_{N-1}\|$$

$$\leqq C \max_{N^* \leqq n \leqq N} E_{2n}(u) + \frac{Ck_N}{t_{N-1} - t_{N^*}} \delta + \frac{C\delta}{\lambda CL},$$

where we used the algorithmic control of the term $k_N \|\Delta_N [U]_{N-1}\|$ and the fact that by the a posteriori error estimate (2.7) and (3.8) we have that $\|e^-_{N^*-1}\| \leqq \delta$. This proves (3.12) if λ is large enough and $k_N/(t_{N-1} - t_{N^*})$ is small enough. □

Proof of Lemma 8.2. We consider the case $q = 1$ with the case $q = 0$ being similar but easier. We introduce the discrete dual problem: Find $Z|_{I_n} \in V_n$ for $n = N, N-1, \cdots, N^*$, such that for $n = N, N-1, \cdots, N^*$,

$$-\int_{I_n} (v, Z_t)\, dt - (v^-_n, [Z]_n) + \int_{I_n} (\nabla v, \nabla Z)\, dt = 0 \quad \forall v \in V_n, \tag{8.2}$$

where $Z^+_N = \Delta^2_N \theta^-_N$. Arguing as in the proof of Lemma 6.1 and choosing successively $v = t^4_n \Delta^2_n Z$, $t^3_n \Delta_n Z$, $t^2_n Z$, $t_n \Delta^{-1}_n Z$ and $\Delta^{-2}_n Z$ on I_n, we get

$$\|Z_t\|_{I_n} + \|\Delta_n Z\|_{I_n} \leqq \frac{C}{(t_N - t_{n-1})^2} \|\Delta^{-1}_N Z^+_N\| \leqq \frac{C}{(t_N - t_{n-1})^2} \|\Delta_N \theta^-_N\|, \quad n = N^*, \cdots, N, \tag{8.3a}$$

$$\|[Z]_{n-1}/k_{n-1}\| + \|[\Delta^{-1}_n Z]_{n-1}/k^2_n\| \leqq \frac{C}{(t_N - t_{n-1})^2} \|\Delta_N \theta^-_N\|, \quad n = N^* + 1, \cdots, N, \tag{8.3b}$$

$$k_N \|Z^-_N\| + \|\Delta^{-1}_N Z^-_N\| \leqq C \|\Delta_N \theta^-_N\|, \tag{8.3c}$$

$$\|Z^+_n\| \leqq \frac{C}{(t_N - t_n)} \|\Delta_N \theta^-_N\|, \qquad n = N^* - 1, \cdots, N-1. \tag{8.3d}$$

Note that (8.3b) is obtained from (8.2) by choosing $v \in S_n$ with zero mean value over I_n and $v^-_n = [Z]_n$ or $v^-_n = \Delta^{-2}_n [Z]_n$.

Choosing now $v = U^-_N - \tilde{I}_N \pi_N u(t_N)$ in (8.2) and using also (1.2) and (1.5) we find writing $u - \tilde{I}_N \pi_N u = \eta$ on I_N, and using the definition of $\tilde{I}_n$

$$\begin{aligned}
\|\Delta_N \theta^-_N\|^2 &= -\sum_{n=N^*}^{N} \int_{I_n} \{(\eta, Z_t) + (\nabla \eta, \nabla Z)\}\, dt - \sum_{n=N^*+1}^{N-1} (\eta^-_n, [Z]_n) \\
&\quad - (\eta^-_N, Z^-_N) - (e^-_{N^*-1}, Z^+_{N^*-1}) \\
&= \sum_{n=N^*}^{N} \int_{I_n} (u - \pi_N u, Z_t)\, dt - \sum_{n=N^*+1}^{N-1} ((u - \pi_n u)^-_n, [Z]_n) \\
&\quad + (\Delta_n (\pi_n u - \tilde{I}_n \pi_n u)^-_n, [\Delta^{-1}_n Z])) - ((u - \pi_N u)^-_N, Z^-_N) \\
&\quad - (\Delta_N (\pi_N u - \tilde{I}_N \pi_N u)^-_N, \Delta^{-1}_N Z^-_N).
\end{aligned}$$

Now using Theorem 2.1 to estimate $u - \pi_n u$ and standard estimates for the L_2-projection $\tilde{I}_n$ together with (8.3), we obtain (8.1a).

To prove (8.1b) we shall estimate $k_N \|\Delta_N \bar{\theta}_N\|$, where $\bar{\theta}_N$ is the time average of θ over I_N. To this end we note that for all $v \in S_N$

$$\begin{aligned}
&\int_{I_N} \{(\theta_t, v) + (\nabla \theta, \nabla v)\}\, dt + ([\theta]_{N-1}, v^+_{N-1}) \\
&\quad = \int_{I_N} \{(\eta_t, v) + (\nabla \eta, \nabla v)\}\, dt + ([\eta]_{N-1}, v^+_{N-1})
\end{aligned}$$

so that

$$\int_{I_N} \{(\theta, -v_t) - (\Delta_N \theta, v)\}\, dt + (\theta_N^-, v_N^-) - (\theta_{N-1}^-, v_{N-1}^+)$$

$$= \int_{I_N} \{(\eta, -v_t) + (\nabla \eta, \nabla v)\}\, dt + (\eta_N^-, v_N^-) - (\eta_{N-1}^-, v_{N-1}^+).$$

Choosing now $v = k_N \Delta_N \bar{\theta}_N$ we get since $e = \eta - \theta$,

$$k_N \|\Delta_N \bar{\theta}_N\| \leqq C(\|e_{N-1}^-\| + \|e_N^-\|).$$

Together with (8.1a) with $N^* = N-1$ this proves (8.1b) and the proof of the lemma is complete. □

Proof of Lemma 8.1. Let w be the function which is linear in x on each $K \in T_h$, continuous at the midpoints x_τ of the interfaces $\tau \in E_i$, and determined by

$$w(x_\tau) = \begin{cases} [\partial v/\partial n_\tau]/h_\tau, & \tau \in E_i \\ 0, & \tau \in E \setminus E_i, \end{cases}$$

so that with $(f, g)_h = \sum_{K \in T_h} \int_K f \cdot g\, dx$,

$$c\|D_h^2 v\|^2 \leqq \sum_{\tau \in E_i} \left[\frac{\partial v}{\partial n_\tau}\right]^2 = \sum_{\tau \in E_i} \left[\frac{\partial v}{\partial n_\tau}\right] h_\tau w(x_\tau) = \sum_{\tau \in E_i} \int_\tau \left[\frac{\partial v}{\partial n_\tau}\right] w\, ds$$

$$= -(\nabla v, \nabla w)_h.$$

Defining $Rw \in S_h$ by

$$(\nabla \chi, \nabla Rw) = (\nabla \chi, \nabla w)_h \quad \forall \chi \in S_h,$$

we get

$$(\nabla v, \nabla w)_h = (\nabla v, \nabla Rw) = (-\Delta_h v, Rw),$$

and hence conclude that

$$\sum_{\tau \in E_i} \left[\frac{\partial v}{\partial n_\tau}\right]^2 \leqq \|\Delta_h v\|\, \|Rw\|.$$

Since the function w admits the inverse estimates

$$\|h \nabla w\|_h + \|w\| \leqq C\left(\sum_{\tau \in E_i} h_\tau^2 w(x_\tau)^2\right)^{1/2},$$

where $\|\cdot\|_h = \sqrt{(\cdot, \cdot)_h}$, it remains to show that

$$\|Rw\| \leqq C(\|h \nabla w\|_h + \|w\|). \tag{8.4}$$

To this end let φ be the solution of the adjoint problem $-\Delta \varphi = Rw$ in Ω, $\varphi = 0$ on $\partial\Omega$, so that

$$\|Rw\|^2 = (Rw, -\Delta\varphi) = (\nabla Rw, \nabla \varphi)$$

$$= (\nabla(Rw - w), \nabla \varphi)_h + (\nabla w, \nabla \varphi)_h = \mathrm{I} + \mathrm{II}.$$

Here, for the usual Lagrangian interpolant $\varphi_i \in S_h$ of φ,

$$|\mathrm{I}| = |(\nabla(Rw - w), \nabla(\varphi - \varphi_i))|$$

$$\leqq \|h \nabla(Rw - w)\|_h \|h^{-1} \nabla(\varphi - \varphi_i)\|$$

$$\leqq C(\|h \nabla Rw\| + \|h \nabla w\|)\|Rw\|,$$

since by the approximation properties of φ_i and elliptic regularity, we have

$$\|h^{-1}\nabla(\varphi-\varphi_i)\| \leqq C\|D^2\varphi\| \leqq C\|\Delta\varphi\|.$$

Furthermore, with $[w]_\tau(x)=\lim_{s\to 0^+}(w(x+sn_\tau)-w(x-sn_\tau))$, $x\in\tau$, we have

$$\mathrm{II} = -\sum_{\tau\in E_i}\int_\tau [w]_\tau \frac{\partial\varphi}{\partial n_\tau}\,ds + (w, -\Delta\varphi) = \mathrm{III}+\mathrm{IV},$$

where

$$|\mathrm{IV}| \leqq \|w\|\,\|Rw\|.$$

For the sum III we have, since $[w]_\tau$ has mean value zero on τ,

$$\begin{aligned}|\mathrm{III}| &= \left|\sum_{\tau\in E_i}\int_\tau [w]_\tau\left(\frac{\partial\varphi}{\partial n_\tau}-c_\tau\right)ds\right|\\ &\leqq \left(\sum_{\tau\in E_i} h_\tau^2\|[w]\|^2_{L_\infty(\tau)}\right)^{1/2}\left(\sum_{\tau\in E_i} h_\tau^{-1}\int_\tau\left[\frac{\partial\varphi}{\partial n_\tau}-c_\tau\right]^2 ds\right)^{1/2},\\ &\leqq C\|w\|\left(\sum_{\tau\in E_i} h_\tau^{-1}\int_\tau\left(\frac{\partial\varphi}{\partial n_\tau}-c_\tau\right)^2 ds\right)^{1/2},\end{aligned}$$

where c_τ are arbitrary constants.

From the scaled trace inequality

$$\int_\tau \psi^2\,ds \leqq C\left(h_K\int_K|\nabla\psi|^2\,dx + h_K^{-1}\int_K\psi^2\,dx\right),\qquad \tau\in\partial K,$$

and the fact that we may choose the constants c_τ in such a way that

$$\int_K\left(\frac{\partial\varphi}{\partial n_\tau}-c_\tau\right)^2 dx \leqq Ch_K^2\int_K (D^2\varphi)^2\,dx,$$

we get

$$\begin{aligned}\sum_{\tau\in E_i} h_\tau^{-1}\int_\tau\left(\frac{\partial\varphi}{\partial n_\tau}-c_\tau\right)^2 ds &\leqq C\sum_{K\in T_h}\int_K (D^2\varphi)^2\,dx\\ &= C\|D^2\varphi\|^2 \leqq C\|\Delta\varphi\|^2 = C\|Rw\|^2.\end{aligned}$$

Summarizing what we have obtained so far we find that

$$\|Rw\| \leqq C(\|h\nabla Rw\| + \|h\nabla w\|_h + \|w\|).$$

As in the proof of Lemma 1 in [E1], it can be proved that for μ sufficiently small

$$\|h\nabla Rw\| \leqq C(\|h\nabla w\|_h + \mu\|Rw\|),$$

which together with the above estimate for $\|Rw\|$ proves (8.4) if μ small. This completes the proof. □

9. Numerical results. In this section we report on some numerical experiments with adaptive algorithms (3.4) with $m=2$ and (3.9) with $q=1$ for the elliptic problem (1.1) and the linear parabolic problem (1.2), respectively. In the practical implementation of (3.4) and (3.9) we have used two different mesh generators which for a given mesh function $h(x)$ both produce a mesh T_h satisfying (1.3b, c). The first mesh generator is based on successive refinements, where certain triangles (fathers) are subdivided into four similar triangles (sons), or by the reverse process where four sons

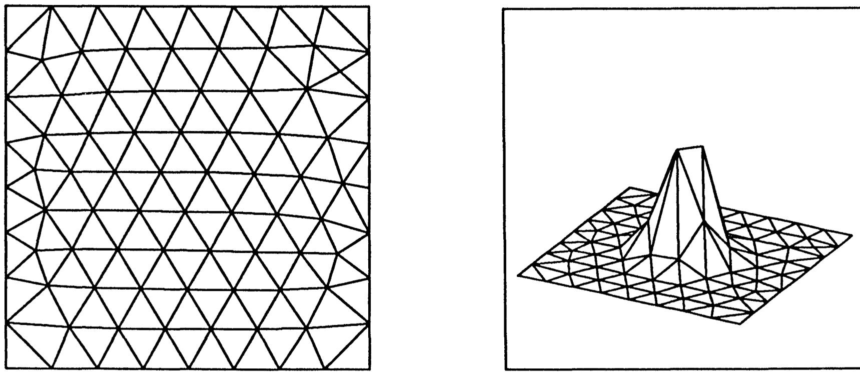

FIG. 9.1a. *Initial mesh,* 144 *elements,* L_2*-error: exact* 0.144, *estimated* 0.329.

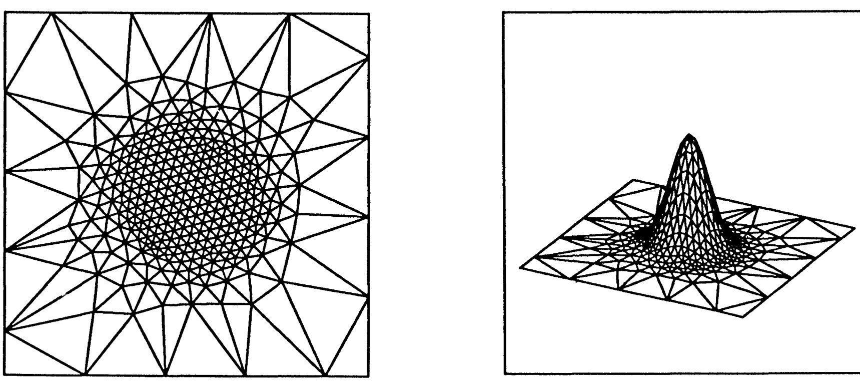

FIG. 9.1b. *First refinement,* 634 *elements,* L_2*-error: exact* 0.0152, *estimated* 0.0229.

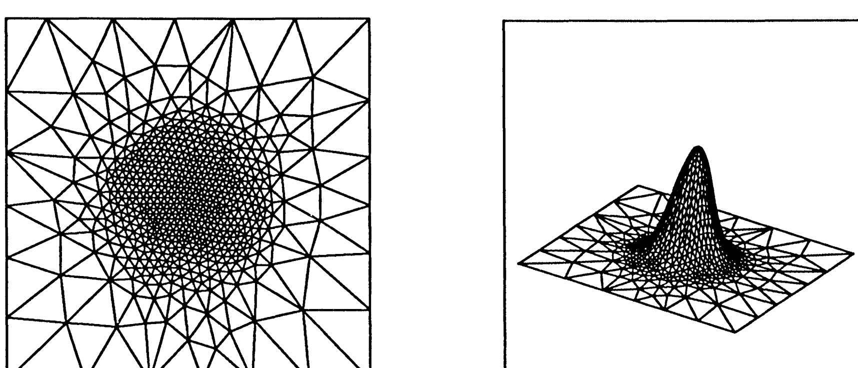

FIG. 9.1c. *Second refinement,* 1140 *elements,* L_2*-error: exact* 0.0116, *estimated* 0.0119.

are replaced by their father. The meshes obtained in this way, starting from a given coarse triangulation, are completed into triangulations by introducing special transition triangles subdivided into two triangles connecting zones with different meshsize. This mesh generator was used in the experiment with the algorithm (3.9) for parabolic problems. In this case $T_n^0 = T_{n-1}$, $n = 1, 2, 3, \cdots$, where T_0 is a given coarse triangulation, and the successive meshes $T_n^j, j \geqq 1$, are obtained through the indicated refinement or derefinement process.

The second mesh generator, used in the numerical experiment with (3.4) presented in Example 9.1, is a so-called "front generator," where elements of a given size are dynamically created at a "front" separating the part of the region already triangulated and the remaining part. The initial front is usually taken to be the boundary of the region. Furthermore, various smoothing procedures may be applied to post process a mesh given by the front method to obtain meshes with "smoother" variation (cf. [H]).

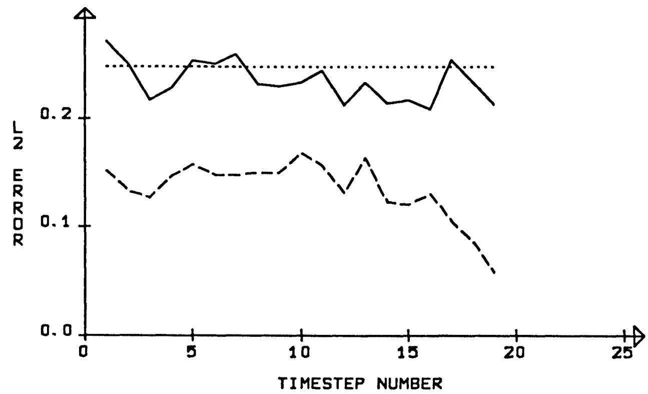

FIG. 9.2a.

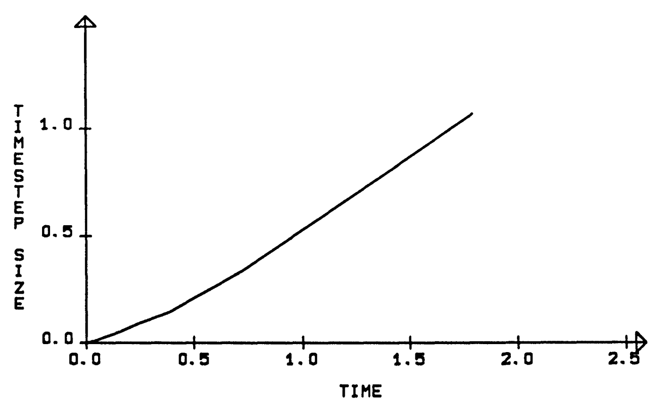

FIG. 9.2b.

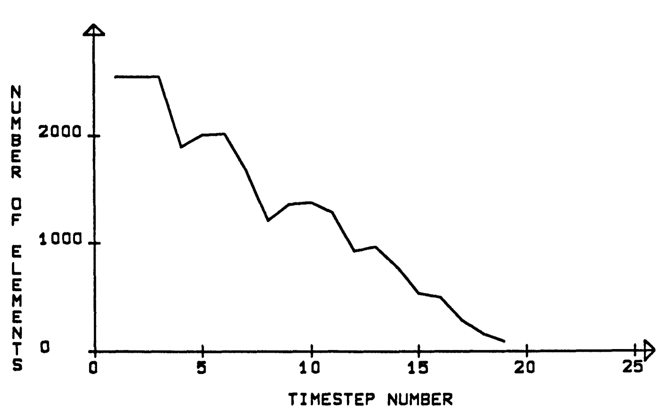

FIG. 9.2c.

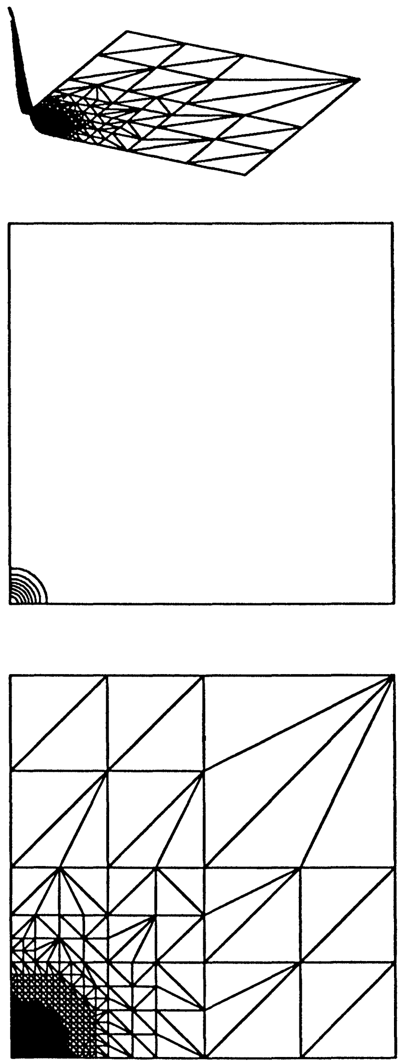

FIG. 9.2d. *Timestep No.* 1.

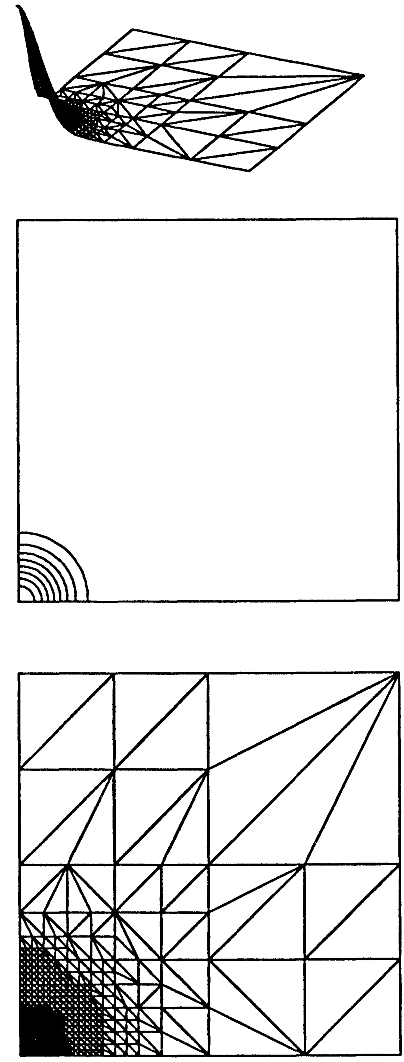

FIG. 9.2e. *Timestep No.* 5.

The constants α_i, β_i, and γ_i which appear in (3.4) and implicitly in (3.9) are essentially pure interpolation error constants, because of the local character of the L_2-projection. For efficiency reasons we would like to find the minimal values of these constants. For example, in our numerical experiments we have taken $\alpha_2 = 0.15$ and $\beta_2 = 0.3$. To give a detailed account for this choice would carry too far. Just to indicate the type of argument used, consider the reference triangular element K with vertices in $(0, 0)$, $(1, 0)$ and $(0, 1)$ and the function $z = x(1-x) + y(1-y)$ with the interpolant $z_i = 0$. We find in this case that the constant C in the interpolation error estimate

$$\max_K |z - z_i| \leqq C \max_K |D^2 z|$$

is $C = 1/2^{5/2}$. Similar calculations using several "generic" functions z and the relevant norms have led us to the given (presumably close minimal) values of α_1 and β_2.

For full reliability, the constants C_i in (3.9) should be determined from the constants α_2, β_2, and γ_2 as in Theorem 2.4. In the numerical experiments the constants

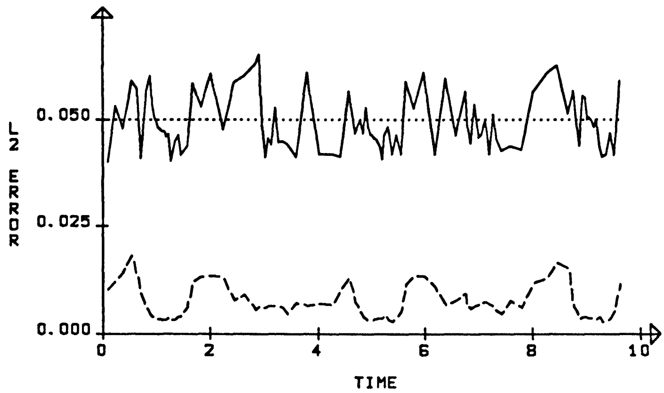

FIG. 9.3a.

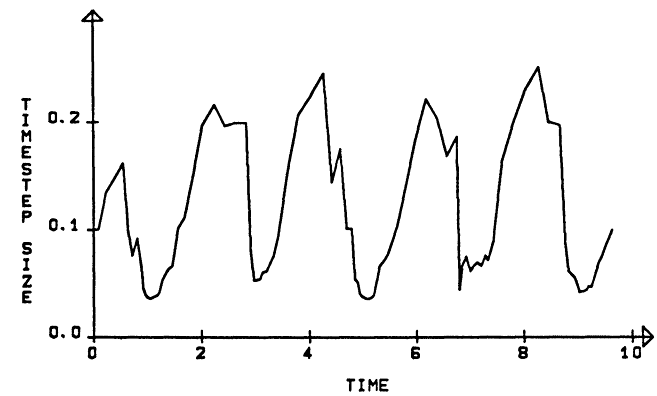

FIG. 9.3b.

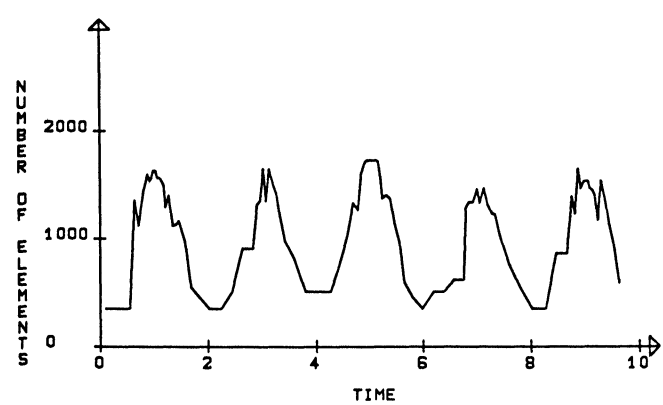

FIG. 9.3c.

C_i in (3.9) have been chosen somewhat smaller since the estimates (2.7) are upper bounds and as such are not quite sharp. Below we have thus used $C_6 = 0.15$, $C_7 = \frac{1}{36}$, $C_8 = 0.3$, $C_9 = 2$, $C_{10} = \frac{1}{6}$, and $C_{11} = 0.2$.

Example 9.1. In this case we let in (1.1) $\Omega = (-1, 1) \times (-1, 1)$, $f(x) = -\Delta \exp(-|x|^2/\varepsilon)$, $\varepsilon = 0.05$, and $\delta = 0.015$. In Fig. 9.1a–c we give the sequence of meshes and surface plots of the corresponding approximate solutions produced by the algorithm (3.4) and the successive quantities $\alpha_2\|h_j f\| + \beta_2 D_{h_j,2}(u_h^j)$, $j = 0, 1, 2$. We note the close correspondence of the estimated and true L_2-error.

Example 9.2. In this case $\Omega = (0, 2) \times (0, 2)$, $f = 0$, and u_0 is the following "approximate δ-function" at $x = 0$:

$$u_0(x) = \frac{1}{\varepsilon} \exp(-|x|^2/\varepsilon),$$

with $\varepsilon = 1/250$, and $\delta = 0.25$. The L_2-error $\|e_n^-\|$, timesteps k_n, number of elements N_n of T_n generated by (3.9) as functions of n or time are given in Fig. 9.2a–c. We note that $\|e_n^-\|$ is kept approximately on the level $\delta/2$ throughout the computation. The theoretically correct timestep sequence is given by $k_n \approx C\delta^{1/3} t_n^{7/6}$, which follows from the fact that for the exact solution u we have $\|u_t^{(3)}(t)\| \approx Ct^{-7/2}$. Thus, k_n should depend almost linearly on t_n which is in accordance with Fig. 9.2b. The theoretically correct number of elements $N(t)$ at time t (from pure approximation theory) is of order $\delta^{-1} t^{-1/2}$ for $t > C\delta^{1/3}\varepsilon^{7/6}$. Figure 9.2d,e shows surface plots of the computed solution, level curves, and the meshes at timesteps 1 and 5, respectively.

Example 9.3. In this case $\Omega = (0, 2) \times (0, 2)$, $u_0 = (x_1(1 - x_1) + x_2(1 - x_2))/4$, and

$$f(x, t) = \left(\sin \frac{\pi t}{2}\right)^5 \exp\left(-\frac{|x - \bar{x}|^2}{\varepsilon}\right) + 1,$$

with $\varepsilon = 0.04$, $\bar{x} = (1, 1)$, and $\delta = 0.05$. The exact solution is here periodic in time with a variable degree of smoothness requiring variable space and timesteps. In Fig. 9.3a–c we give the L_2-error $\|e_n^-\|$, timesteps k_n, and number of elements N_n of T_n as functions of time. We see that the error is kept roughly on the level $\delta/3$ over many periods of the time dependent factor $(\sin \pi t/2)^5$. (The error distribution was similar up to time $t \simeq 120$ corresponding to thirty periods.)

REFERENCES

[B] R. BANK, *Analysis of a local a posteriori error estimate for elliptic equations*, in Accuracy Estimates and Adaptive Refinements in Finite Element Computations, I. Babuska, O. C. Zienkiewicz, J. Gago, and E. R. de A. Oliveira, eds., John Wiley, New York, 1986.

[BB1] M. BIETERMAN AND I. BABUSKA, *The finite element method for parabolic equations*, I. *A posteriori error estimation*, Numer. Math., 40 (1982), pp. 339–371.

[BB2] ———, *The finite element method for parabolic equations*, II. *A posteriori error estimation and adaptive approach*, Numer. Math., 40 (1982), pp. 373–406.

[BB3] ———, *An adaptive method of lines with error control for parabolic equations of the reaction-diffusive type*, J. Comput. Phys., 63 (1986), pp. 33–66.

[BM] I. BABUSKA AND A. MILLER, *A posteriori error estimates and adaptive techniques for the finite element method*, Univ. of Maryland, Institute for Physical Science and Technology, Tech. note BN-968, College Park, MD, 1981.

[BOP] I. BABUSKA, J. OSBORN, AND J. PITKÄRANTA, *Analysis of mixed methods using mesh dependent norms*, Math. Comp., 35 (1980), pp. 1039–1062.

[DF] S. F. DAVIS AND J. E. FLAHERTY, *An adaptive finite element method for initial-boundary value problems for partial differential equations*, SIAM J. Sci. Statist. Comput., 3 (1982), pp. 6–27.

[Du] T. DUPONT, *Mesh modification for evolution equations*, Math. Comp., 39 (1982), pp. 85–107.

[E1] K. ERIKSSON, *Adaptive finite element methods for parabolic problems* II: *A priori error estimates in $L_\infty(L_2)$ and $L_\infty(L_\infty)$*, Tech. report, Department of Mathematics, Chalmers University of Technology, Göteborg, Sweden, 1988.

[E2] ———, *Adaptive finite element methods based on optimal error estimates for linear elliptic problems*, Tech. report, Dept. of Mathematics, Chalmers University of Technology, Göteburg, Sweden, 1987.

[E3] K. ERIKSSON, *Improved accuracy by adapted mesh-refinement in the finite element method*, Math. Comp., 44 (1985), pp. 321-343.

[E4] ———, *Error estimates for the $H_0^1(\Omega)$ and $L_2(\Omega)$ projections onto finite element spaces under weak mesh regularity assumptions*, Dept. of Mathematics, Chalmers University of Technology, Göteborg, Sweden, 1988.

[EJ1] K. ERIKSSON AND C. JOHNSON, *An adaptive finite element method for linear elliptic problems*, Math. Comp., 50 (1988), pp. 361-383.

[EJ2] ———, *Error estimates and automatic time step control for nonlinear parabolic problems*, I, SIAM J. Numer. Anal., 24 (1987), p. 12-23.

等到找 [EJ3] ———, *Adaptive finite element methods for parabolic problems* III: *Time steps variable in space*, to appear.

[EJ4] ———, *Adaptive finite element methods for parabolic problems* IV: *A nonlinear problem*, to appear.

[EJ5] ———, *Adaptive finite element element methods for elliptic problems based on a posteriori error estimates*, to appear.

[EJT] K. ERIKSSON, C. JOHNSON, AND V. THOMÉE, *Time discretization of parabolic problems by the Discontinuous Galerkin method*, RAIRO Math. Anal., 19 (1985), pp. 611-643.

[EW] D. EWING, *Adaptive mesh refinements in large-scale fluid flow simulation*, in Accuracy Estimates and Adaptive Refinements in Finite Element Computations, I. Babuska, O. C. Zienkiewicz, J. Gago, and E. R. de A. Oliveira, eds., John Wiley, New York, 1986.

[H] P. HANSBO, *Adaptivity and streamline diffusion procedures in the finite element method*, Ph.D. thesis, Chalmers University of Technology, Göteborg, Sweden, 1989.

[J] C. JOHNSON, *Error estimates and adaptive time step control for a class of one step methods for stiff ordinary differential equations*, SIAM J. Numer. Anal., 25 (1988), pp. 908-926.

[Ja] P. JAMET, *Galerkin-type approximations which are discontinuous in time for parabolic equations in a variable domain*, SIAM J. Numer. Anal., 15 (1978), pp. 912-928.

[JNT] C. JOHNSON, Y.-Y. NIE, AND V. THOMÉE, *An a posteriori error estimate and adaptive timestep control for a backward Euler discretization of a parabolic problem*, SIAM J. Numer. Anal., 27 (1990), pp. 277-291.

[L] J. LENNBLAD, *An adaptive finite element method for a linear parabolic problem*, Tech. report, Dept. of Mathematics, Chalmers University of Technology, Göteborg, Sweden, 1988.

[Li] G. LIPPOLD, *Error estimates and step-size control for the approximate solution of a first order evolution equation*, preprint, Akademie der Wissenschaffen der Karl-Weierstrass-Institut für Matematik, Berlin, 1988.

[LMZ] R. LÖHNER, K. MORGAN, AND O. C. ZIENKIEWICZ, *Adaptive grid refinement for the compressible Euler equations*, in Accuracy Estimates and Adaptive Refinements in Finite Element Computations, I. Babuska, O. C. Zienkiewicz, J. Gago, and E. R. de A. Oliveira, eds., John Wiley, New York, 1986.

[ODSD] J. T. ODEN, L. DEMKOWICZ, T. STROUBOULIS, AND P. DEVLOO, *Adaptive methods for problems in solid and fluid mechanics*, in Accuracy Estimates and Adaptive Refinements in Finite Element Computations, I. Babuska, O. C. Zienkiewicz, J. Gago, and E. R. de A. Oliveira, eds., John Wiley, New York, 1986.