SIAM J. NUMER. ANAL.
Vol. 34, No. 1, pp. 389–401, February 1997
© 1997 Society for Industrial and Applied Mathematics
016

ON THE STABILITY OF THE DISCONTINUOUS GALERKIN METHOD FOR THE HEAT EQUATION*

CH. G. MAKRIDAKIS† AND I. BABUŠKA‡

Abstract. This paper analyzes stability properties of a class of discontinuous Galerkin methods for the heat equation. It is shown that the finite element projection associated with these methods is stable with respect to a mesh-dependent norm—a discrete analogue of the space–time L^2-norm. Optimal order-regularity error bounds in $L^2([0,T];L^2(\Omega))$ are derived.

Key words. finite elements, discontinuous Galerkin

AMS subject classifications. Primary, 65M60, 65M12; Secondary, 65M15

PII. S0036142994261658

1. Introduction.

In this paper we study the stability properties of the discontinuous Galerkin method with respect to certain mesh-dependent norms for the model problem

$$\begin{aligned} &u_t(x,t)-\Delta u(x,t)=f(x,t) \quad \text{in} \quad \Omega\times[0,T],\\ &u(x,t)=0 \quad \text{on} \quad \partial\Omega\times[0,T],\\ &u(x,0)=u^0(x) \quad \text{in} \quad \Omega, \end{aligned} \tag{1.1}$$

where Ω is a bounded polygonal domain in $\mathbb{R}^N$, $N=1,2$ with boundary $\partial\Omega$. The discontinuous Galerkin method has been introduced for first-order hyperbolic problems in [15] and for parabolic problems in [14]. The method proposed by Jamet in [14] allows completely unstructured space–time finite element meshes, but in this (general) form there are no optimal order error estimates available. Under some assumptions on the structure of the mesh the method has been further developed, including issues such as optimal order error estimates, a posteriori estimates, and adaptivity, in a series of important papers by Eriksson, Johnson, and Thomée [13], and by Eriksson and Johnson [9], [10], [11], [12]. For other works on space–time finite element methods for parabolic problems cf., e.g., [1], [3], [6], [8].

We shall introduce some notation and describe the method in the form that we consider here. Let $0=t_0<t_1<\cdots<t_J=T$ be a partition of $[0,T]$,

$$I_n=(t_{n-1},t_n], \quad k_n=t_n-t_{n-1}\,.$$

We associate a partition $\mathcal{T}_{hn}$ of Ω and a finite element space S_{hn} to each interval I_n :

$$S_{hn}=\{\chi\in H_0^1(\Omega):\chi|_K\in P_p(K),K\in\mathcal{T}_{hn}\},$$

where $P_p(S)$ is the space of polynomials of degree p. In what follows we shall denote by K and e the elements and the "edges," respectively, of the partition $\mathcal{T}_{hn}$ and by

*Received by the editors January 17, 1994; accepted for publication (in revised form) April 26, 1995.

http://www.siam.org/journals/sinum/34-1/26165.html

†Department of Mathematics, University of Crete, 714 09 Heraklion, Crete, Greece and Department of Mathematics, University of Tennessee, Knoxville, TN 37966 (makr@math.uch.gr).

‡Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742. Current address: Texas Institute for Computational and Applied Mathematics, University of Texas at Austin, TAY 2.400, Austin, TX 78712.

Γ_{hn} the collection of the interior edges of $\mathcal{T}_{hn}$. Also, h_K stands for the diameter of the element K, and $h_n = \max_{K\in\mathcal{T}_{hn}} h_K$. We assume that the family of partitions $\mathcal{T}_{hn}$ satisfies the minimal angle condition (uniformly with respect to h and n). Following standard notation, we let $H^k(S)$ be the Sobolev space of order k, and we denote its norm by $\|\cdot\|_{k,S}$ ($\|\cdot\|_k$ if $S=\Omega$.) Also, $(\cdot,\cdot)$ shall denote the inner product, and $\|\cdot\|$ the corresponding norm on $L^2(\Omega)$.

Now let $V_{hk} = V_{hk}(q)$, q a positive integer, be the space of piecewise polynomial functions $\varphi : \Omega\times[0,T]\to\mathbb{R}$ of the form

$$\varphi|_{\Omega\times I_n} = \sum_{j=0}^{q} t^j\chi_j(x)\,, \quad \chi_j\in S_{hn}\,.$$

Hence, the functions of V_{hk} are, for each $t\in I_n$, elements of S_{hn} and for each $x\in\Omega$ piecewise polynomial functions of degree q with possible discontinuities at the nodes t_n, $n=0,\dots,J-1$. Let also $V_{hk}^n = \{\varphi|_{\Omega\times I_n} \;:\; \varphi\in V_{hk}\}$.

The finite element method is defined as follows: for $U^0 = u^0$ find $U\in V_{hk}$ such that

$$B(U,\varphi) = F(\varphi) \qquad \forall\varphi\in V_{hk}, \tag{1.2}$$

where

$$\begin{aligned} B(U,\varphi) = &\sum_{n=1}^{J}\int_{I_n}\{(U_t,\varphi)+(\nabla U,\nabla\varphi)\}dt \\ &+\sum_{n=1}^{J-1}(U^{n+}-U^n,\varphi^{n+})+(U^{0+},\varphi^{0+}), \end{aligned} \tag{1.3}$$

$$F(\varphi) = \int_0^T (f,\varphi)dt + (U^0,\varphi^{0+}), \tag{1.4}$$

where $v^n = v(t_n)$, $v^{n+} = \lim_{s\to 0^+} v(t_n+s)$.

This method belongs to the family of methods introduced in [14] and allows considerable flexibility in the construction of the meshes (we allow variable time step k_n and variable with n spaces S_{hn}). For the more general case considered in [14] it has been proved that (1.2) has a unique solution and that the finite element approximation U converges in $L^2([0,T],H^1(\Omega))$ to the solution u of (1.1). For $p=1$ and $q=0$ or 1, Eriksson and Johnson [9], [10] prove optimal a priori error estimates, a posteriori estimates, and propose an adaptive method for the selection of the mesh. Here we consider the case $p>1$, $q>1$ as well, approach the stability of the method in a different way, and, under certain assumptions on the spaces S_{hn}, cf. (1.8) below, we show that the finite element projection is stable with respect to a mesh-dependent norm (discrete analogue of the $L^2([0,T],L^2(\Omega))$). As a consequence of this, we obtain a "symmetric" bound for the error of the form

$$\|u-U\|_{0,hk}\le C\inf_{\varphi\in V_{hk}}\|u-\varphi\|_{0,hk},$$

and optimal order-regularity error bounds in $L^2([0,T],L^2(\Omega))$. (For other symmetric estimates for parabolic problems cf. Dupont [8] and Bank and Santos [6].) The analysis is carried out by proving a discrete inf-sup (Babuška–Brezzi) condition for the

bilinear form $B(\cdot,\cdot)$ with an appropriate choice of a pair of mesh-dependent norms—this technique was introduced for the error analysis of finite element approximations to elliptic problems in [4], [5]. Following the standard way of proving conditions of this type [2] it turns out that the stability estimate rests (in a completely natural way) on the stability of a discrete dual problem associated with (1.2). The importance of the stability analysis of the discrete dual problem for the discontinuous Galerkin method was previously realized in [13], [9], [10], where the error analysis was based on the stability of a discrete dual problem similar to the one considered here.

The space–time mesh-dependent norms. Let u be the exact solution of (1.1) and $\varphi \in V_{hk}$. Integrating (1.3) by parts, in both space and time variables we have

$$\begin{aligned} B(u,\varphi) = \sum_{n=1}^{J} \Bigg\{ \int_{I_n} \Bigg\{ \sum_{K\in\mathcal{T}_{hn}} - \int_K u(\varphi_t + \Delta\varphi) \\ + \sum_{e\in\Gamma_{hn}} \int_e u \left[\frac{\partial\varphi}{\partial\nu}\right] dS \Bigg\} dt \Bigg\} \\ + \sum_{n=1}^{J-1} (u^n, \varphi^n - \varphi^{n+}) + (u^J, \varphi^J), \end{aligned} \tag{1.5}$$

where if $e = \partial K^1 \cap \partial K^2$, $[\frac{\partial\varphi}{\partial\nu}] = \frac{\partial\varphi}{\partial\nu^{[1]}} + \frac{\partial\varphi}{\partial\nu^{[2]}}$, and $\nu^{[j]}$ is the unit normal to e, exterior to K^j; i.e., since $\nu^{[1]} = -\nu^{[2]}$, $[\frac{\partial\varphi}{\partial\nu}]$ denotes the jump of $\frac{\partial\varphi}{\partial\nu}$ across e. If $u \in X = [\bigoplus_{n=1}^{J} C^1(I_n; L^2(\Omega))] \cap L^2([0,T]; H^1(\Omega))$ the above integral terms involving u are well defined.

For $u \in X$ we define the mesh-dependent norm $\|\cdot\|_{0,hk}$ by

$$\begin{aligned} \|u\|_{0,hk}^2 = \sum_{n=1}^{J} \Bigg\{ \int_{I_n} \left(\|u\|^2 + \sum_{e\in\Gamma_{hn}} h_e \int_e |u|^2 dS \right) dt \\ + k_n \|u^n\|^2 \Bigg\}, \end{aligned} \tag{1.6}$$

where h_e is the length of e. (If $N = 1$, $h_e = (h_K + h_{K'})/2$, where e is the common point of K and K'.) If u is the exact solution of (1.1) and U is the finite element approximation defined by (1.2), then U is also characterized by

$$B(U,\varphi) = B(u,\varphi) \quad \forall\varphi \in V_{hk}.$$

If we define the finite element projection $\Pi_1 : X \to V_{hk}$ by

$$B(\Pi_1 u, \varphi) = B(u,\varphi) \quad \forall\varphi \in V_{hk}, \tag{1.7}$$

then our goal is to show that Π_1 is stable with respect to the $\|\cdot\|_{0,hk}$ norm. For this, we introduce first the mesh-dependent norm in V_{hk}:

$$\begin{aligned} \|\varphi\|_{2,hk}^2 = \sum_{n=1}^{J} \Bigg\{ \int_{I_n} \Bigg(\sum_{K\in\mathcal{T}_{hn}} \int_K (|\varphi_t|^2 + |\Delta\varphi|^2 + |\nabla\varphi|^2) dx \\ + \sum_{e\in\Gamma_{hn}} h_e^{-1} \int_e \left[\frac{\partial\varphi}{\partial\nu}\right]^2 dS \Bigg) dt \Bigg\} \\ + \sum_{n=1}^{J-1} k_n^{-1} \|\varphi^n - \varphi^{n+}\|^2 + k_J^{-1} \|\varphi^J\|^2, \quad \varphi \in V_{hk}. \end{aligned}$$

We shall also need the stationary analogues of these norms. Following [5], for $v \in S_{hn}$, let

$$\|v\|_{0,h}^2 = \|v\|^2 + \sum_{e \in \Gamma_{hn}} h_e \int_e |v|^2 dS$$

and

$$\|v\|_{2,h}^2 = \sum_{K \in \mathcal{T}_{hn}} \int_K |\Delta v|^2 dx + \sum_{e \in \Gamma_{hn}} h_e^{-1} \int_e \left[\frac{\partial v}{\partial \nu}\right]^2 dS.$$

Our results are based on the assumption that the finite element spaces S_{hn} satisfy the following Babuška–Brezzi (inf-sup) condition with respect to these norms:

$$\sup_{\psi \in S_{hn}} \frac{|(\nabla v, \nabla \psi)|}{\|\psi\|_{2,h}} \geq \gamma_0 \|v\|_{0,h} \quad \forall v \in S_{hn}. \tag{1.8}$$

For $N = 1$, (1.8) is proved in [4]. For $N = 2$, when Ω is convex it is proved in [5] that (1.8) holds with a constant γ_0 independent of h for quasi-uniform meshes. And, also, Lemma 8.1 of Eriksson and Johnson [9] implies (1.8) in the case of piecewise linear elements under a weak local quasi-uniformity assumption on the mesh. Whether (1.8) is valid or not on nonconvex polygonal domains or without any restriction hypothesis on the mesh even for the piecewise linear case is an open problem.

We shall also assume that Ω is such that the elliptic regularity estimate

$$|v|_{2,\Omega} \leq C\|\Delta v\| \qquad \forall v \in H_0^1(\Omega) \cap H^2(\Omega) \tag{1.9}$$

holds.

Now we are ready to discuss our basic results. In section 2 we study the stability of this method and prove the following result.

THEOREM 1.1. *Assume that the spaces S_{hn} satisfy the inf-sup condition* (1.8) *with a constant γ_0 independent of h and n, and that* (1.9) *holds. Let $k_n \leq C_1 k_{n+1}$ for some constant $C_1 > 0$, and $h_n^2 \leq \beta k_n$, for $\beta > 0$ sufficiently small, cf. Lemma* 2.1 *below, for all n. Then there is a positive constant $\gamma = \gamma(C_1, \beta, \gamma_0)$ such that*

$$\sup_{\psi \in V_{hk}} \frac{|B(v, \psi)|}{\|\psi\|_{2,hk}} \geq \gamma \|v\|_{0,hk} \quad \forall v \in V_{hk}. \tag{1.10}$$

A basic tool for the proof of Theorem 1.1 is a stability estimate for a dual problem, cf. (2.3), (2.4) below, the proof of which is carried out along the lines of the proof of a similar estimate in [10].

If now $u \in X$, a direct consequence of (1.5) and of the definition of mesh-dependent norms is the relation

$$|B(u, \varphi)| \leq \|u\|_{0,hk} \|\varphi\|_{2,hk} \quad \forall \varphi \in V_{hk}, u \in X. \tag{1.11}$$

Combining (1.10) and (1.11) we obtain [2]

$$\|\Pi_1 u\|_{0,hk} \leq \frac{1}{\gamma} \|u\|_{0,hk} \quad \forall u \in X, \tag{1.12}$$

and, hence, we get the approximability result

$$\|u - \Pi_1 u\|_{0,hk} \leq \left(1 + \frac{1}{\gamma}\right) \inf_{\varphi \in V_{hk}} \|u - \varphi\|_{0,hk}. \tag{1.13}$$

So we have the following a priori estimates whose proof is given in section 2.

THEOREM 1.2. *Let u be the solution of* (1.1) *and U the finite element approximation defined by* (1.3), (1.4). *Under the assumptions of Theorem* 1.1 *we have*

$$\|u-U\|_{0,hk} \le \left(1+\frac{1}{\gamma}\right) \inf_{\varphi\in V_{hk}} \|u-\varphi\|_{0,hk}, \tag{1.14}$$

and also for a positive constant C

$$\|u-U\|_{L^2(\Omega\times[0,T])} \le C(\|k^{q+1}u^{(q+1)}\|_k + \|h^{p+1}D^{(p+1)}u\|_h), \tag{1.15}$$

where $u^{(q+1)}$ is the $q+1$ order derivative of u with respect to t and

$$\|k^{q+1}u^{(q+1)}\|_k^2 = \sum_{n=1}^{J} k_n^{2(q+1)} \int_{I_n} \|u^{(q+1)}\|^2 dt\,,$$

$$\|h^{p+1}D^{(p+1)}u\|_h^2 = \sum_{n=1}^{J} \int_{I_n} \left(\sum_{K\in\mathcal{T}_{hn}} h_K^{2(p+1)} \|u\|_{p+1,K}^2 \right) dt.$$

In this theorem, (1.14) is an optimal approximability result with respect to the norm $\|\cdot\|_{0,hk}$ and (1.15) is an *optimal order-regularity* result with respect to the global space–time L^2-norm.

We note here that for $q = 0, 1$ Eriksson and Johnson [9], [10] have proved an estimate for the error $u - U$ in $L^\infty([0,T];L^2(\Omega))$ which is of order $q+1$ globally in time. They also prove a superconvergence result of order $2q+1$ at the discrete time levels t_n . These results hold provided $S_{hn} \subset S_{h(n-1)}$ or that a weak mesh condition $h_n^2 \le ck_n$, c small enough, holds.

Remark. Analogous results to (1.14) and (1.15) above can also be derived for other mesh-dependent norms. For instance, if we consider the discrete analogue of the $L^\infty([0,T];L^2(\Omega))$-norm we can verify that an estimate of the form (1.14) holds for this norm but with a constant that grows logarithmically as $k_n \to 0$.

2. Error analysis. In this section we shall analyze the stability properties of the method (1.2)–(1.4) and shall prove Theorems 1.1 and 1.2. The proof of the inf-sup condition of Theorem 1.1 is given according to the standard argument: in order to get (1.10), let v be a given element of V_{hk}; then, it suffices to find $\varphi \in V_{hk}$ and two positive constants β_0, β_1 independent of k, h such that

$$B(v,\varphi) \ge \beta_0 \|v\|_{0,hk}^2, \qquad \beta_0 > 0, \tag{2.1}$$

and

$$\|\varphi\|_{2,hk} \le \beta_1 \|v\|_{0,hk}, \qquad \beta_1 > 0. \tag{2.2}$$

Hence, for a given $v \in V_{hk}$, let φ be the unique element of V_{hk} satisfying

$$\int_{I_J} \{-(\psi,\varphi_t) + (\nabla\psi,\nabla\varphi)\}dt + (\psi^J,\varphi^J) = \int_{I_J} (\psi,v)dt \tag{2.3}$$

and

$$\int_{I_n} \{-(\psi,\varphi_t) + (\nabla\psi,\nabla\varphi)\}dt + (\psi^n,\varphi^n-\varphi^{n+}) = \int_{I_n} (\psi,v)dt, \tag{2.4}$$
$$n = 1,\dots,J-1 \,\forall\psi \in V_{hk}.$$

Taking now $\psi = v$ in (2.3)–(2.4) we have

$$\begin{aligned} B(v,\varphi) &= \sum_{n=1}^{J} \int_{I_n} \{ -(v,\varphi_t) + (\nabla v, \nabla \varphi)\} dt \\ &\quad + \sum_{n=1}^{J-1} (v^n, \varphi^n - \varphi^{n+}) + (v^J, \varphi^J) \\ &= \sum_{n=1}^{J} \int_{I_n} \|v\|^2. \end{aligned}$$

Since v is a polynomial of degree q with respect to t in each I_n, we have for all $\psi \in V_{hk}$

$$\|\psi^n\|^2 \le \sup_{s\in I_n} \|\psi(s)\|^2 \le \alpha_1 k_n^{-1} \int_{I_n} \|\psi(s)\|^2 ds\,, \quad \alpha_1 > 0. \tag{2.5}$$

In what follows we shall frequently use (2.5).

Also, one can easily verify (e.g., by doing step-by-step the proof of Lemma 1 of [5] without the assumption that the mesh is quasi uniform) that for all $\chi \in S_{hn}$ we have

$$\|\chi\| \le \|\chi\|_{0,h} \le c_1 \|\chi\|. \tag{2.6}$$

Combining (2.5) and (2.6) we get

$$\sum_{n=1}^{J} \int_{I_n} \|v\|^2 \ge \beta_0 \|v\|_{0,hk}^2, \quad \beta_0 > 0, \tag{2.7}$$

and, hence, (2.1) holds. In order to complete the proof of Theorem 1.1 it remains to prove (2.2). For this we introduce the discrete operator $\Delta_h^n : S_{hn} \to S_{hn}$ defined for $n = 1, \dots, J$ by

$$(\Delta_h^n \psi, \chi) = -(\nabla \psi, \nabla \chi) \quad \forall \chi \in S_{hn}. \tag{2.8a}$$

If $R_n : H^1 \to S_{hn}$ is the elliptic projection defined by

$$(\nabla R_n w, \nabla \chi) = (\nabla w, \nabla \chi) \quad \forall \chi \in S_{hn}, \tag{2.8b}$$

then, assuming (1.9), we have the well-known estimate $\|w - R_n w\| \le c_o h_n \|\nabla(w - R_n w)\|$. Using this estimate one can show that, cf. [10, proof of Lemma 2.2],

$$\|w - R_n w\| \le C_0 h_n (h_n + h_{n+1}) \|\Delta_h^{n+1} w\|, \quad w \in S_{h(n+1)}, \tag{2.9}$$

where $C_0 = 2c_o^2$. We are using this estimate in the proof of Lemma 2.1 below. The proof of this result follows the lines of the proof of Lemma 2.2 of [10].

LEMMA 2.1. *Let φ be the solution of* (2.3)–(2.4). *Assume that*

$$k_n \le C_1 k_{n+1} \tag{i}$$

for some positive constant C_1. If

$$h_n^2 \le \beta k_n, \quad \beta > 0 \tag{ii}$$

where β is a small enough constant such that $C_0\beta\alpha_1(3+2C_1) \le \frac{1}{4}$, then the following estimate holds: there is a positive constant C independent of k_n and h_n, $n = 1, 2, \ldots, J$ such that

$$\sum_{n=1}^{J}\left(\int_{I_n}\|\varphi_t\|^2dt + \int_{I_n}\|\Delta_h^n\varphi\|^2dt\right) + \frac{1}{k_J}\|\varphi^J\|^2 + \sum_{n=1}^{J-1}\frac{1}{k_n}\|\varphi^n - \varphi^{n+}\|^2 \le C\sum_{n=1}^{J}\int_{I_n}\|v\|^2dt. \tag{2.10}$$

Proof. Using (2.8) the equation (2.3) becomes

$$\int_{I_J} -(\psi, \varphi_t + \Delta_h^J\varphi)dt + (\psi^J, \varphi^J) = \int_{I_J}(\psi, v)dt.$$

Now let $\psi = -\Delta_h^J\varphi \in V_{hk}^J$:

$$\int_{I_J} -(\nabla\varphi, \nabla\varphi_t)dt + \int_{I_J}\|\Delta_h^J\varphi\|^2dt + (\nabla\varphi^J, \nabla\varphi^J) = \int_{I_J}(\Delta_h^J\varphi, v)dt$$

or

$$\frac{1}{2}\|\nabla\varphi^J\|^2 + \frac{1}{2}\|\nabla\varphi^{J-1+}\|^2 + \int_{I_J}\|\Delta_h^J\varphi\|^2dt \le \int_{I_J}\|v\|^2dt + \frac{1}{4}\int_{I_J}\|\Delta_h^J\varphi\|^2dt. \tag{2.11}$$

For $\psi = -\Delta_h^n\varphi$ in (2.4) we have

$$-\frac{1}{2}\|\nabla\varphi^n\|^2 + \frac{1}{2}\|\nabla\varphi^{n-1+}\|^2 + \int_{I_n}\|\Delta_h^n\varphi\|^2dt - (\Delta_h^n\varphi^n, \varphi^n - \varphi^{n+}) \le \int_{I_n}|(\Delta_h^n\varphi, v)|dt.$$

Now, following [10],

$$\begin{aligned}
-\frac{1}{2}\|\nabla\varphi^n\|^2 &-(\Delta_h^n\varphi^n, \varphi^n - \varphi^{n+}) \\
&= \frac{1}{2}\|\nabla\varphi^n\|^2 - (\nabla\varphi^n, \nabla R_n\varphi^{n+}) - (\Delta_h^n\varphi^n, R_n\varphi^{n+} - \varphi^{n+}) \\
&= \frac{1}{2}\|\nabla(\varphi^n - R_n\varphi^{n+})\|^2 - \frac{1}{2}\|\nabla R_n\varphi^{n+}\|^2 - (\Delta_h^n\varphi^n, R_n\varphi^{n+} - \varphi^{n+}) \\
&\ge \frac{1}{2}\|\nabla(\varphi^n - R_n\varphi^{n+})\|^2 - \frac{1}{2}\|\nabla\varphi^{n+}\|^2 - (\Delta_h^n\varphi^n, R_n\varphi^{n+} - \varphi^{n+}).
\end{aligned}$$

Therefore,

$$-\frac{1}{2}\|\nabla\varphi^{n+}\|^2 + \frac{1}{2}\|\nabla\varphi^{n-1+}\|^2 + \int_{I_n}\|\Delta_h^n\varphi\|^2dt + \frac{1}{2}\|\nabla(\varphi^n - R_n\varphi^{n+})\|^2 \le \int_{I_n}|(\Delta_h^n\varphi, v)|dt + |(\Delta_h^n\varphi^n, R_n\varphi^{n+} - \varphi^{n+})|.$$

Using the hypotheses (i) and (ii), the relation (2.9), and the inverse estimate (2.5) we have

$$
\begin{aligned}
-\frac{1}{2}\|\nabla\varphi^{n+}\|^2+\frac{1}{2}\|\nabla\varphi^{n-1+}\|^2+\int_{I_n}\|\Delta_h^n\varphi\|^2dt&\\
&\le\int_{I_n}\|v\|^2+\frac{1}{4}\int_{I_n}\|\Delta_h^n\varphi\|^2+C_0h_n^2\|\Delta_h^n\varphi^n\|^2\\
&\quad+C_0(h_n+h_{n+1})^2\|\Delta_h^{n+1}\varphi^{n+}\|^2\\
&\le\int_{I_n}\|v\|^2+\frac{1}{4}\int_{I_n}\|\Delta_h^n\varphi\|^2+C_0k_n\|\Delta_h^n\varphi^n\|^2\\
&\quad+2C_0\beta(k_n+k_{n+1})\|\Delta_h^{n+1}\varphi^{n+}\|^2\\
&\le\int_{I_n}\|v\|^2+\left(\frac{1}{4}+C_0\beta\alpha_1\right)\int_{I_n}\|\Delta_h^n\varphi\|^2dt\\
&\quad+2C_0\beta\alpha_1(1+C_1)\int_{I_{n+1}}\|\Delta_h^{n+1}\varphi\|^2dt.
\end{aligned}
$$

(2.12)

Summing (2.11) and (2.12), $n=1,\dots,J-1$, we have

$$
\begin{aligned}
\frac{1}{2}\|\nabla\varphi^J\|^2+\frac{1}{2}\|\nabla\varphi^{0+}\|^2+\sum_{n=1}^{J}\int_{I_n}\|\Delta_h^n\varphi\|^2dt&\\
&\le\sum_{n=1}^{J}\int_{I_n}\|v\|^2+\left[\frac{1}{4}+C_0\beta\alpha_1(3+2C_1)\right]\sum_{n=1}^{J}\int_{I_n}\|\Delta_h^n\varphi\|^2dt.
\end{aligned}
$$

Therefore, for

$$
(*)\qquad C_0\beta\alpha_1(3+2C_1)\le\frac{1}{4},
$$

we have

$$
\|\nabla\varphi^J\|^2+\|\nabla\varphi^{0+}\|^2+\sum_{n=1}^{J}\int_{I_n}\|\Delta_h^n\varphi\|^2dt\le 2\sum_{n=1}^{J}\int_{I_n}\|v\|^2. \tag{2.13}
$$

Our next step will be to estimate $\int_{I_n}\|\varphi_t\|^2$. For this let $\psi=-(t_J-t)\varphi_t$ in (2.3):

$$
\begin{aligned}
\int_{I_J}(t_J-t)\|\varphi_t\|^2dt&\le\int_{I_J}(t_J-t)|(\varphi_t,v)|dt+\int_{I_J}(t_J-t)|(\varphi_t,\Delta_h^J\varphi)|dt\\
&\le k_J\int_{I_J}\|\varphi_t\|\|v\|dt+k_J\int_{I_J}\|\varphi_t\|\|\Delta_h^J\varphi\|dt.
\end{aligned}
$$

Using the inverse estimate

$$
k_n\int_{I_n}\omega^2dt\le C_3\int_{I_n}(t_n-t)\omega^2dt\quad\forall\omega\in P_{q-1}(I_n), \tag{2.14}
$$

we have

$$
\begin{aligned}
k_J\int_{I_J}\|\varphi_t\|^2dt\le C_3^2k_J\int_{I_J}\|v\|^2dt+C_3^2k_J\int_{I_J}\|\Delta_h^J\varphi\|^2dt&\\
+\frac{1}{2}k_J\int_{I_J}\|\varphi_t\|^2dt,&
\end{aligned}
$$

or

$$\int_{I_J} \|\varphi_t\|^2 dt \le 2C_3^2 \left(\int_{I_J} \|v\|^2 dt + \int_{I_J} \|\Delta_h^J \varphi\|^2 dt \right). \tag{2.15}$$

Choosing $\psi = -(t_n - t)\varphi_t$ in (2.4), we show that

$$\int_{I_n} \|\varphi_t\|^2 dt \le 2C_3^2 \left(\int_{I_n} \|v\|^2 dt + \int_{I_n} \|\Delta_h^n \varphi\|^2 dt \right), \quad n = 1, 2, \ldots, J-1, \tag{2.16}$$

exactly as we did for (2.15). Summing (2.15) and (2.16), $n = 1, 2, \ldots, J-1$, and using (2.13) we conclude

$$\begin{aligned} \sum_{n=1}^{J} \int_{I_n} \|\varphi_t\|^2 dt &\le 2C_3^2 \sum_{n=1}^{J} \left(\int_{I_n} \|v\|^2 dt + \int_{I_n} \|\Delta_h^n \varphi\|^2 dt \right) \\ &\le 6C_3^2 \sum_{n=1}^{J} \int_{I_n} \|v\|^2 dt. \end{aligned} \tag{2.17}$$

The last step is the estimation of the terms $\frac{1}{k_n}\|\varphi^n - \varphi^{n+}\|^2$. Let P_n be the L^2 projection onto S_{hn}. The function $P_n(\varphi^n - \varphi^{n+})$ can be considered as an (constant in t) element of V_{hk}^n. For $\psi = P_n(\varphi^n - \varphi^{n+})$ in (2.4) we have

$$(P_n(\varphi^n - \varphi^{n+}), \varphi^n - \varphi^{n+}) = \int_{I_n} (P_n(\varphi^n - \varphi^{n+}), v) dt + \int_{I_n} (P_n(\varphi^n - \varphi^{n+}), \varphi_t - \Delta_h^n \varphi) dt.$$

Now using the fact that $\|P_n\chi - \chi\| \le \|R_n\chi - \chi\|$ we obtain

$$\begin{aligned} \|\varphi^n - \varphi^{n+}\|^2 &= ((\varphi^n - \varphi^{n+}) - P_n(\varphi^n - \varphi^{n+}), \varphi^n - \varphi^{n+}) + (P_n(\varphi^n - \varphi^{n+}), \varphi^n - \varphi^{n+}) \\ &= (P_n\varphi^{n+} - \varphi^{n+}, \varphi^n - \varphi^{n+}) \\ &\quad + \int_{I_n} (P_n(\varphi^n - \varphi^{n+}), v) dt + \int_{I_n} (P_n(\varphi^n - \varphi^{n+}), \varphi_t - \Delta_h^n \varphi) dt \\ &\le \|P_n\varphi^{n+} - \varphi^{n+}\| \|\varphi^n - \varphi^{n+}\| \\ &\quad + \int_{I_n} \left(\frac{1}{4k_n} \|P_n(\varphi^n - \varphi^{n+})\|^2 + k_n \|v\|^2 \right) dt \\ &\quad + \int_{I_n} \left(\frac{1}{4k_n} \|P_n(\varphi^n - \varphi^{n+})\|^2 + 2k_n (\|\varphi_t\|^2 + \|\Delta_h^n \varphi\|^2) \right) dt \\ &\le \|R_n\varphi^{n+} - \varphi^{n+}\|^2 + \frac{3}{4} \|\varphi^n - \varphi^{n+}\|^2 \\ &\quad + k_n \int_{I_n} \left(\|v\|^2 + 2(\|\varphi_t\|^2 + \|\Delta_h^n \varphi\|^2) \right) dt. \end{aligned}$$

As before, (i), (ii), (2.5), (2.9), and $(*)$ give

$$\begin{aligned}\|R_n\varphi^{n+} - \varphi^{n+}\|^2 &\le 2C_0^2(h_n^2 + h_{n+1}^2)h_n^2\|\Delta_h^{n+1}\varphi^{n+}\|^2\\ &\le 2C_0^2\beta(k_n + k_{n+1})k_n\|\Delta_h^{n+1}\varphi^{n+}\|^2\\ &\le 2C_0^2\beta(1+C_1)k_nk_{n+1}\|\Delta_h^{n+1}\varphi^{n+}\|^2\\ &\le 2C_0^2\beta(1+C_1)k_n\alpha_1\int_{I_{n+1}}\|\Delta_h^{n+1}\varphi\|^2dt\\ &\le \frac{1}{2}C_0k_n\int_{I_{n+1}}\|\Delta_h^{n+1}\varphi\|^2dt.\end{aligned}$$

Hence, we obtain for $n = 1, \dots J-1$

$$\begin{aligned}\frac{1}{4}\|\varphi^n - \varphi^{n+}\|^2 \le k_n\int_{I_n}\left(\|v\|^2 + 2(\|\varphi_t\|^2 + \|\Delta_h^n\varphi\|^2)\right)dt\\ +\frac{1}{2}C_0k_n\int_{I_{n+1}}\|\Delta_h^{n+1}\varphi\|^2dt\,.\end{aligned} \tag{2.18}$$

Using an analogous estimate for $\|\varphi^J\|^2$, summing with respect to n, and using (2.13), (2.17) we conclude

$$\frac{1}{k_J}\|\varphi^J\|^2 + \sum_{n=1}^{J-1}\frac{1}{k_n}\|\varphi^n - \varphi^{n+}\|^2 \le C_*\sum_{n=1}^{J}\int_{I_n}\|v\|^2dt, \tag{2.19}$$

where $C_* = (3+12C_3^2+C_0/2)$. The relations (2.13), (2.17), and (2.19) give the desired result. □

We are now ready to prove Theorem 1.1.

Proof of Theorem 1.1. In order to complete the proof of the theorem we need to show (2.2). In view of (2.10) and of the definition of $\|\cdot\|_{2,hk}$ norm it remains to prove that

$$\sum_{n=1}^{J}\int_{I_n}\|\varphi\|_{2,h}^2dt \le C\sum_{n=1}^{J}\int_{I_n}\|v\|^2dt.$$

The inf-sup condition (1.8) is equivalent to [5]

$$\sup_{\psi\in S_{hn}}\frac{|(\nabla w, \nabla\psi)|}{\|\psi\|_{0,h}} \ge \gamma_0\|w\|_{2,h} \quad \forall v \in S_{hn}.$$

Using this relation, (2.6), and the definition of Δ_h^n, we obtain

$$\begin{aligned}\gamma_0\int_{I_n}\|\varphi(t)\|_{2,h}^2dt &\le \int_{I_n}\left(\sup_{\psi\in S_{hn}}\frac{|(\nabla\varphi(t),\nabla\psi)|}{\|\psi\|_{0,h}}\right)^2dt\\ &\le \int_{I_n}\left(\sup_{\psi\in S_{hn}}\frac{|(\Delta_h^n\varphi(t),\psi)|}{\|\psi\|}\right)^2dt\\ &= \int_{I_n}\|\Delta_h^n\varphi(t)\|^2dt.\end{aligned}$$

Therefore, (2.13) implies

$$\sum_{n=1}^{J}\int_{I_n}\|\varphi\|_{2,h}^2 dt \le \frac{2}{\gamma_0}\sum_{n=1}^{J}\int_{I_n}\|v\|^2 dt, \tag{2.20}$$

which completes the proof of the theorem. □

We now prove Theorem 1.2.

Proof of Theorem 1.2. Theorem 1.1 and the relation (1.11) give [2]

$$\|\Pi_1 u\|_{0,hk} \le \frac{1}{\gamma}\|u\|_{0,hk} \quad \forall u \in X,$$

and, therefore, (1.14) follows.

For the proof of (1.15) we will show that

$$\inf_{\varphi\in V_{hk}}\|u-\varphi\|_{0,hk} \le C(\|k^{q+1}u^{(q+1)}\|_k + \|h^{p+1}D^{(p+1)}u\|_h). \tag{2.21}$$

Since φ is discontinuous in time it suffices to consider

$$\inf_{\varphi\in V_{hk}^n}\int_{I_n}\left(\|u-\varphi\|^2 + \sum_{e\in\Gamma_{hn}} h_e\int_e |u-\varphi|^2 dS\right)dt + k_n\|u^n-\varphi^n\|^2 = \inf_{\varphi\in V_{hk}^n}\int_{I_n}\|u-\varphi\|_{0,h}^2 dt + k_n\|u^n-\varphi^n\|^2.$$

Let $I_h v \in S_{hn}$ be the interpolant of v from the finite element space S_{hn}, and φ any element of V_{hk}^n. Since $I_h u(t) - \varphi(t) \in S_{hn}$ for all t in I_n we have using (2.6)

$$\begin{aligned}\int_{I_n}\|u-\varphi\|_{0,h}^2 dt + k_n\|u^n-\varphi^n\|^2 &\le \left(\int_{I_n}\|I_h u-\varphi\|_{0,h}^2 dt + k_n\|u^n-\varphi^n\|^2\right) \\ &\quad + \int_{I_n}\|u-I_h u\|_{0,h}^2 dt \\ &\le \left(C\int_{I_n}\|I_h u-\varphi\|^2 dt + k_n\|u^n-\varphi^n\|^2\right) \\ &\quad + \int_{I_n}\|u-I_h u\|_{0,h}^2 dt.\end{aligned} \tag{2.22}$$

Now let Π_n be the $L^2(I_n)$ projection onto $P_q(I_n)$, and $\widetilde{I}^n$ be an interpolant operator of smooth functions from $P_q(I_n)$, such that

$$\widetilde{I}^n y(t_n) = y(t_n), \tag{2.23}$$

with the approximation property

$$\|\widetilde{I}^n y - y\|_{L^2(I_n)} \le C k_n^{q+1}\|y^{(q+1)}\|_{L^2(I_n)}. \tag{2.24}$$

An operator with this property is constructed in [13]. Now note that $\widetilde{I}^n u - \varphi$ is a polynomial of degree q with respect to the t variable; therefore, using (2.23) and the

inverse inequality (2.5), we obtain

$$
\begin{aligned}
\int_{I_n} \|u-\varphi\|_{0,h}^2 dt + k_n \|u^n-\varphi^n\|^2 &\le \left(C\int_{I_n} \|I_h u-\varphi\|^2 dt + k_n\|u^n-\varphi^n\|^2 \right) \\
&\quad + 2\int_{I_n} \|u-I_h u\|_{0,h}^2 dt \\
&\le \left(C\int_{I_n} \|I_h u-\varphi\|^2 dt + k_n\|\widetilde{I}^n u^n-\varphi^n\|^2 \right) \\
&\quad + 2\int_{I_n} \|u-I_h u\|_{0,h}^2 dt \\
&\le C\int_{I_n} (\|I_h u-\varphi\|^2 + \|\widetilde{I}^n u-\varphi\|^2) dt \\
&\quad + 2\int_{I_n} \|u-I_h u\|_{0,h}^2 dt.
\end{aligned} \tag{2.25}
$$

Using well-known properties of the finite element interpolant [7] we have

$$
\int_{I_n} \|u-I_h u\|_{0,h}^2 dt \le C\int_{I_n} \left(\sum_{K\in\mathcal{T}_{hn}} h_K^{2(p+1)} \|u\|_{p+1,K}^2 \right) dt. \tag{2.26}
$$

It remains to estimate the first two terms of the right-hand side of (2.25). For this let $\varphi = \Pi_n I_h u$. Using the stability properties of the L^2 projection we obtain

$$
\begin{aligned}
\int_{I_n} \|I_h u-\Pi_n I_h u\|^2 dt &\le 3\int_{I_n} \|I_h u-u\|^2 dt + 3\int_{I_n} \|u-\Pi_n u\|^2 dt \\
&\quad + 3\int_{I_n} \|\Pi_n(u-I_h u)\|^2 dt \\
&\le 6\int_{I_n} \|I_h u-u\|^2 dt + 3\int_{I_n} \|u-\widetilde{I}^n u\|^2 dt.
\end{aligned} \tag{2.27}
$$

Also,

$$
\begin{aligned}
\int_{I_n} \|\widetilde{I}^n u-\Pi_n I_h u\|^2 dt &\le 2\int_{I_n} \|\widetilde{I}^n u-u\|^2 dt + 2\int_{I_n} \|u-\Pi_n I_h u\|^2 dt \\
&\le 2\int_{I_n} \|\widetilde{I}^n u-u\|^2 dt + 2\int_{I_n} \|u-\Pi_n u\|^2 dt \\
&\quad + 2\int_{I_n} \|\Pi_n(u-I_h u)\|^2 dt \\
&\le 4\int_{I_n} \|\widetilde{I}^n u-u\|^2 dt + 2\int_{I_n} \|u-I_h u\|^2 dt.
\end{aligned} \tag{2.28}
$$

Finally, using (2.24) and (2.26) we complete the proof of the theorem. □

Acknowledgment. The first author would like to thank Professor John E. Osborn for several helpful discussions on mesh-dependent norms.

REFERENCES

[1] A. K. Aziz and P. Monk *Continuous finite elements in space and time for the heat equation,* Math. Comp., 52 (1989), pp. 255–274.

[2] I. Babuška and A. K. Aziz, *Survey lecture on the mathematical foundations of the finite element method*, in The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, A. K. Aziz, ed., Academic Press, New York, 1972, pp. 5–359.

[3] I. Babuška and T. Janik, *The $h-p$ version of the finite element method for parabolic equations, Part* I*: The p-version in time*, Numer. Methods Partial Differential Equations, 5 (1989), pp. 363–399; *Part* II: *The $h-p$ version in time*, Numer. Methods Partial Differential Equations, 6 (1990), pp. 343–369.

[4] I. Babuška and J. E. Osborn, *Analysis of finite element methods for second order boundary value problems using mesh dependent norms*, Numer. Math., 34 (1980), pp. 41–62.

[5] I. Babuška, J. E. Osborn, and J. Pitkäranta, *Analysis of mixed methods using mesh dependent norms*, Math. Comp., 35 (1980), pp. 1039–1062.

[6] R. E. Bank and R. F. Santos, *Analysis of some moving space-time finite element methods*, SIAM J. Numer. Anal., 30 (1993), pp. 1–18.

[7] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North–Holland, Amsterdam, 1978.

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

[9] K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems* I: *A linear model problem*, SIAM J. Numer. Anal., 28 (1991), pp. 43–77.

[10] K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems* II: *Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$*, SIAM J. Numer. Anal., 32 (1995), pp. 706–740.

[11] K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems* IV: *A nonlinear problem*, SIAM J. Numer. Anal., 32 (1995), pp. 1729–1749.

[12] K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems* V: *Long-time integration*, SIAM J. Numer. Anal., 32 (1995), pp. 1750–1762.

[13] K. Eriksson, C. Johnson, and V. Thomée, *Time discretization of parabolic problems by the discontinuous Galerkin method*, RAIRO Anal. Numer., 19 (1985), pp. 611–643.

[14] 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.

[15] P. Lesaint and P. A. Raviart, *On a finite element method for solving the neutron transport equation*, in Mathematical Aspects of Finite Elements in Partial Differential Equations, C. deBoor, ed., Academic Press, New York, 1974, pp. 89–123.