MATHEMATICS OF COMPUTATION
VOLUME 37, NUMBER 156
OCTOBER 1981

Some Convergence Estimates for Semidiscrete Type Schemes for Time-Dependent Nonselfadjoint Parabolic Equations

By Huang Mingyou and Vidar Thomée

Abstract. L_2-norm error estimates are shown for semidiscrete (continuous in time) Galerkin finite element type approximations to solutions of general time-dependent nonselfadjoint second order parabolic equations under Dirichlet boundary conditions. The semidiscrete solutions are defined in terms of given methods for the corresponding elliptic problem such as the standard Galerkin method in which the boundary conditions are satisfied exactly but also methods for which this is not necessary. The results are proved by energy arguments and include estimates for the homogeneous equation with both smooth and nonsmooth initial data.

1. Introduction. In this paper we shall discuss semidiscrete Galerkin finite element type approximations of the general second order parabolic initial boundary value problem ($u_t = \partial u/\partial t$)

$$\begin{aligned} u_t + Au &= f \quad \text{in } \Omega \times I, \\ u &= 0 \quad \text{on } \partial\Omega \times I, \\ u &= v \quad \text{in } \Omega \text{ for } t = 0, \end{aligned} \tag{1.1}$$

where Ω is a bounded domain in R^n with sufficiently smooth boundary $\partial\Omega$, $I = (0, t^0]$ a finite interval in time, f a function of x and t, and A the uniformly elliptic differential operator

$$Au = A(x, t)u = -\sum_{j,k=1}^{n} \frac{\partial}{\partial x_j}\left(a_{jk}(x, t)\frac{\partial u}{\partial x_k}\right) + \sum_{j=1}^{n} b_j(x, t)\frac{\partial u}{\partial x_j} + c(x, t)u,$$

with a_{jk}, b_j, and c smooth functions in $\bar{\Omega} \times \bar{I}$, and $a_{jk} = a_{kj}$. Our purpose here is to consider the modifications necessary in order to carry over the L_2 error estimates of Bramble, Schatz, Thomée, and Wahlbin [4] for the homogeneous equation with A selfadjoint, positive, and time-independent to the general situation stated.

The semidiscrete approximations in [4] were defined by the equation

$$T_h u_{h,t} + u_h = 0 \quad \text{for } t \in I,$$

where, with h a small positive parameter, $\{T_h\}$ denotes a family of approximations of $T = A^{-1}$ with range in finite element spaces $\{S_h\}$ such that

(i) T_h is selfadjoint, positive semidefinite on $L_2(\Omega)$, and positive definite on S_h;

There is an integer $r \geqslant 2$ such that

(ii) $\|(T_h - T)f\| \leqslant Ch^s\|f\|_{s-2}$ for $2 \leqslant s \leqslant r$.

Received December 4, 1980.
1980 *Mathematics Subject Classification*. Primary 65N15, 65N30.

Here and below $\|\cdot\|_k$ is the norm in $H^k(\Omega)$ and $\|\cdot\| = \|\cdot\|_0$ that in $L_2(\Omega)$. The approach of [4] was extended to nonhomogeneous equations in Thomée [11] and to time-dependent A in Sammon [7].

We introduce the bilinear form corresponding to the elliptic operator A,

$$A(v, w) = A(t; v, w) = \int_\Omega \left(\sum_{j,k=1}^{n} a_{jk} \frac{\partial v}{\partial x_k} \frac{\partial w}{\partial x_j} + \sum_{j=1}^{n} b_j \frac{\partial v}{\partial x_j} w + cvw \right) dx,$$

and note that if $\kappa > \sup(\frac{1}{2} \sum_j \partial b_j/\partial x_j - c)$ we have

$$A(v, v) + \kappa\|v\|^2 \geq c_1\|v\|_1^2 \quad \text{for } v \in H_0^1(\Omega) \text{ with } c_1 > 0;$$

κ will be fixed in this way in the rest of this paper. We shall now associate with the study of our parabolic equation the elliptic problem

$$A_\kappa u \equiv (A + \kappa)u = f \quad \text{in } \Omega, \qquad u = 0 \quad \text{on } \partial\Omega, \tag{1.2}$$

or in weak form,

$$A_\kappa(u, \varphi) \equiv A(u, \varphi) + \kappa(u, \varphi) = (f, \varphi) \quad \forall \varphi \in H_0^1(\Omega).$$

We denote by $T = T(t)$: $L_2(\Omega) \to H^2(\Omega) \cap H_0^1(\Omega)$ (or $H^{-1}(\Omega) \to H_0^1(\Omega)$) the solution operator of this problem, so that

$$A_\kappa(T(t)f, \varphi) = (f, \varphi) \quad \forall \varphi \in H_0^1(\Omega).$$

Introducing in (1.1) $\tilde{u} = e^{-\kappa t}u$ as a new dependent variable, we have

$$\tilde{u}_t + A_\kappa \tilde{u} = \tilde{f} \equiv e^{-\kappa t} f \quad \text{for } t \in I, \ \tilde{u}(0) = v,$$

which may now also be written

$$T\tilde{u}_t + \tilde{u} = T\tilde{f} \quad \text{for } t \in I, \ \tilde{u}(0) = v. \tag{1.3}$$

Following [4] we define a semidiscrete (discrete with respect to x) approximation of (1.1) in terms of an approximate method to solve the associated elliptic problem (1.2): Let $\{S_h\}$ be a family of finite-dimensional subspaces of $L_2(\Omega)$ and $T_h = T_h(t)$: $L_2(\Omega) \to S_h$ an approximation of T with properties to be stated below. Consider then as an approximate solution of (1.1) a function u_h: $\bar{I} \to S_h$ such that $u_h = e^{\kappa t}\tilde{u}_h$ where

$$T_h \tilde{u}_{h,t} + \tilde{u}_h = T_h \tilde{f} \quad \text{for } t \in I, \ \tilde{u}_h(0) = v_h, \tag{1.4}$$

with v_h an approximation of v in S_h. For example, assuming $S_h \subset H_0^1(\Omega)$, we may take for T_h the solution operator associated with the standard Galerkin method as defined by

$$A_\kappa(T_h f, \chi) = (f, \chi) \quad \forall \chi \in S_h.$$

The problem (1.4) then reduces to the standard weak formulation of the parabolic equation

$$(u_{h,t}, \chi) + A(u_h, \chi) = (f, \chi) \quad \forall \chi \in S_h.$$

We shall now describe the conditions which will be placed upon $\{T_h\}$ for (1.4) to define a good approximation of the solution of (1.1) and present corresponding convergence estimates. The latter will result from considering the following equation satisfied by the error $e = \tilde{u}_h - \tilde{u} = e^{-\kappa t}(u_h - u)$, namely

$$T_h e_t + e = \rho \equiv (T_h - T)A_\kappa \tilde{u} \quad \text{for } t \in I,$$

which is obtained by subtracting (1.3) from (1.4). Our hypotheses below, which generalize the basic assumptions (i) and (ii) of [4] to the nonselfadjoint time-dependent situation, will be the following:

(ia) $(f, T_h f) \geq 0$ for $f \in L_2(\Omega)$ and $(\chi, T_h \chi) > 0$ for $0 \neq \chi \in S_h$;

(ib) $$|(T_h f, g) - (f, T_h g)| \leq C(f, T_h f)^{1/2} \|T_h g\|.$$

There is an integer $r \geq 2$ such that

$(\mathrm{ii})_k$ $$\|(T_h^{(j)} - T^{(j)})f\| \leq Ch^s \|f\|_{s-2} \quad \text{for } 2 \leq s \leq r, 0 \leq j \leq k.$$

Here and below we use $B^{(j)}$ to denote $(d/dt)^j B$, with $B' = B^{(1)}$, for B a function or operator depending on t, and similarly for a bilinear form. Note that (ia) implies that (1.4) is uniquely solvable for $t \geq 0$ since T_h^{-1} then exists on S_h. The assumption (ib) bounds the degree of nonselfadjointness of T_h; for the standard Galerkin method it is a simple consequence of the fact that

$$|A_\kappa(v, w) - A_\kappa(w, v)| \leq C\|v\|_1 \cdot \|w\|.$$

The present condition $(\mathrm{ii})_0$ was also used in [4].

In Section 3 below we will discuss the validity of our hypotheses for some different choices of elliptic approximations $\{T_h\}$, namely the standard Galerkin method, the Langrange multiplier method of Babuška [1] which restricts S_h by requiring its elements to be orthogonal to a separate finite-dimensional space on the boundary, the method studied by Berger, Scott, and Strang [2] and Scott [9] in which the elements of S_h vanish at certain carefully chosen boundary points, and finally the method of Nitsche [6] in which certain boundary terms are included in the bilinear form defining T_h.

Section 2 below contains the derivation of our L_2-norm error estimates. We begin by proving that in the case of the nonhomogeneous equation with sufficiently smooth solution we have under assumptions (ia) and $(\mathrm{ii})_1$ (Theorem 1)

$$\|u_h(t) - u(t)\| \leq Ch^r \left\{ \|v\|_r + \int_0^t \|u_t\|_r \, d\tau \right\} \quad \text{for } t \in \bar{I},$$

if v_h is a suitably chosen approximation of v. We then direct our attention to the homogeneous equation. Assuming now (ia), (ib), and $(\mathrm{ii})_1$ we first prove an estimate for the case of smooth initial data, compatible with the differential equation, again with v_h suitably chosen, namely (Theorem 2)

$$\|u_h(t) - u(t)\| \leq Ch^r \|v\|_r \quad \text{for } t \in \bar{I},$$

and then an estimate for nonsmooth data, now with v_h the L_2-projection of v onto S_h (Theorems 3 and 4),

$$\|u_h(t) - u(t)\| \leq Ch^r t^{-r/2} \|v\| \quad \text{for } t \in I.$$

For $r > 2$ the proof of this latter estimate requires the analogue of $(\mathrm{ii})_1$ to be valid for the adjoints of T and T_h, so that

$(\mathrm{ii})_1^*$ $$\|(T_h^{*(j)} - T^{*(j)})f\| \leq Ch^s \|f\|_{s-2} \quad \text{for } 2 \leq s \leq r, j = 0, 1.$$

For the standard Galerkin method we find easily

$$(\chi, f) = A_\kappa(\chi, T_h^* f) \quad \forall \chi \in S_h,$$

so that T_h^* is the standard Galerkin approximation of $T^* = (A_\kappa^*)^{-1}$, and $(\text{ii})_1^*$ is proved in the same way as $(\text{ii})_1$. We end Section 2 by showing that similarly for the error in the time derivatives, assuming now (ia), (ib), $(\text{ii})_{l+1}$, and $(\text{ii})_1^*$,

$$\|D_t^l u_h(t) - D_t^l u(t)\| \leqslant Ch^r t^{-r/2-l}\|v\| \quad \text{for } t \in I.$$

The proofs of the nonsmooth data estimates in [4] and preceding work depended on spectral representation, whereas, as was the case in [7], our present proofs are by energy arguments. The basic step is Lemma 3 below which generalizes the inequality (2.8) of [4], used there to prove the smooth data result.

We end this introductory section by recalling some regularity properties of the solution of the homogeneous parabolic equation which will be explicitly used in our error analysis below (cf. [10] and [7]). Letting u be a solution of

$$(1.9) \qquad u_t + Au = 0 \quad \text{in } \Omega \times I, \qquad u = 0 \quad \text{on } \partial\Omega \times I,$$

we define the solution operator $E(t, s)$ for $0 \leqslant s \leqslant t \in I$ by $u(t) = E(t, s)u(s)$. This operator has the properties $E(s, s) = I$ and

$$(1.10) \qquad E(t, y)E(y, s) = E(t, s) \quad \text{for } 0 \leqslant s \leqslant y \leqslant t \in I.$$

The solution $u(t) = E(t, 0)v$ may be defined even for $v \in L_2(\Omega)$; it is smooth for $t > 0$ and for any $k \geqslant 0$,

$$(1.11) \qquad \|E(t, 0)v\|_k \leqslant Ct^{-k/2}\|v\| \quad \text{for } t \in I.$$

More generally, the solution operator satisfies

$$(1.12) \qquad \|E(t, s)v\|_k \leqslant C(t-s)^{-k/2}\|v\| \quad \text{for } 0 \leqslant s < t \in I.$$

For the solution to be smooth on $\bar{I}$, it is necessary to demand, in addition to smoothness of v, that this function be compatible with the differential equation and the boundary condition at $\partial\Omega$ for $t = 0$, so that the time derivatives of the solution for $t = 0$, as formally computed from the differential equation, vanish at $\partial\Omega$. More precisely, set

$$v_0 = v, \qquad v_{l+1} = -\sum_{j=0}^{l} \binom{l}{j} A^{(l-j)}(0) v_j.$$

We say that v is compatible of order k with (1.9), where k is a positive integer, if $v_j \in H^{k-2j}(\Omega) \cap H_0^1(\Omega)$ for $j < k/2$. When this holds we have

$$(1.13) \qquad \|E(t, 0)v\|_{k+j} \leqslant Ct^{-j/2}\|v\|_k \quad \text{for } j \geqslant 0,\ t \in I.$$

With v satisfying the corresponding compatibility conditions of order k with (1.9) at $t = s$ (which is always the case for $v = E(s, 0)w$ if $s > 0$, $w \in L_2(\Omega)$) we have similarly

$$\|E(t, s)v\|_{k+j} \leqslant C(t-s)^{-j/2}\|v\|_k \quad \text{for } j \geqslant 0,\ 0 \leqslant s < t \in I.$$

During the completion of this paper we learned about related independent work by Luskin and Rannacher [5], and Sammon [8]. In [8], L_2 error estimates similar to ours were derived under inverse hypotheses on $\{S_h\}$, and in [5] nonsmooth data results of the type described above were obtained without inverse assumptions in the case of the standard Galerkin method, using a parabolic duality argument.

2. The Error Estimates. We shall start the L_2-norm error analysis by proving our result for the nonhomogeneous equation. Its proof will be based on the following lemma.

LEMMA 1. *Assume that* (ia) *holds and that*

$$T_h w_t + w = g \quad \text{for } t \in I.$$

Then

$$\|w(t)\| \leqslant \|w(0)\| + C\left\{\|g(0)\| + \int_0^t \|g_t\| d\tau\right\} \quad \text{for } t \in \bar{I}.$$

Proof. Multiplying the equation for w by w_t, we have

$$(T_h w_t, w_t) + \frac{1}{2}\frac{d}{dt}\|w\|^2 = (g, w_t) = \frac{d}{dt}(g, w) - (g_t, w),$$

so that using (ia), after integration,

$$\|w(t)\|^2 \leqslant \|w(0)\|^2 + 2\|g(t)\|\,\|w(t)\| + 2\|g(0)\|\,\|w(0)\| + 2\int_0^t \|g_t\|\,\|w\|\,d\tau$$

$$\leqslant \sup_{\tau < t} \|w(\tau)\| \left\{\|w(0)\| + C \sup_{\tau < t} \|g(\tau)\| + C\int_0^t \|g_t\|\,d\tau\right\}.$$

This estimate is valid with t replaced by any $t_1 \leqslant t$ and in particular for t_1 such that $\|w(t_1)\| = \sup_{\tau \leqslant t} \|w(\tau)\|$. Hence

$$\|w(t_1)\|^2 \leqslant \sup_{\tau < t} \|w(\tau)\| \cdot \left\{\|w(0)\| + C \sup_{\tau < t} \|g(\tau)\| + C\int_0^t \|g_t\|\,d\tau\right\}.$$

After cancelling a common factor $\|w(t_1)\|$ and noting that

$$\|g(\tau)\| \leqslant \|g(0)\| + \int_0^t \|g_t\|\,d\tau \quad \text{for } \tau \leqslant t,$$

we get

$$\|w(t)\| \leqslant \|w(t_1)\| \leqslant \|w(0)\| + C\left\{\|g(0)\| + \int_0^t \|g_t\|\,d\tau\right\},$$

which proves the lemma.

The following is now our error estimate for the nonhomogeneous equation:

THEOREM 1. *Assume that* (ia) *and* (ii)$_1$ *hold and that* v_h *is chosen so that* $\|v_h - v\| \leqslant Ch^r\|v\|_r$. *Then, for* $t \in I$,

$$\|u_h(t) - u(t)\| \leqslant Ch^r\left\{\|v\|_r + \int_0^t \|u_t\|_r\,d\tau\right\}.$$

Proof. We know from Section 1 that $e(t) = e^{-\kappa t}(u_h(t) - u(t))$ satisfies

$$T_h e_t + e = \rho \equiv (T_h - T)A_\kappa \tilde{u}.$$

Since, by (ii)$_1$,

$$\|\rho(0)\| = \|(T_h - T)A_\kappa v\| \leqslant Ch^r\|A_\kappa v\|_{r-2} \leqslant Ch^r\|v\|_r,$$

and, with A' the operator obtained from A (or A_κ) by time differentiation,

$$\|\rho_t\| \leqslant \|(T_h' - T')A_\kappa \tilde{u}\| + \|(T_h - T)(A'\tilde{u} + A_\kappa \tilde{u}_t)\| \leqslant Ch^r(\|u\|_r + \|u_t\|_r),$$

with

$$\|u(t)\|_r \leqslant \|v\|_r + \int_0^t \|u_t\|_r\,d\tau,$$

the desired result follows at once by Lemma 1.

Note that in Theorem 1 we have considerable freedom in our choice of v_h. For instance, we could choose v_h to be the elliptic projection of v defined by $v_h = T_h A_\kappa v$, since

$$\|v_h - v\| = \|(T_h - T)A_\kappa v\| \leq Ch^r\|A_\kappa v\|_{r-2} \leq Ch^r\|v\|_r,$$

but we could also take $v_h = P_h v$, where P_h denotes the orthogonal projection onto S_h with respect to the L_2 inner product since then

$$\|v_h - v\| = \inf_{\chi \in S_h} \|\chi - v\| \leq \|T_h A_\kappa v - v\| \leq Ch^r\|v\|_r.$$

We now turn to the homogeneous equation. Our problem is thus

$$\begin{aligned} u_t + Au &= 0 \quad \text{in } \Omega \times I, \\ u &= 0 \quad \text{on } \partial\Omega \times I, \\ u(\cdot, 0) &= v \quad \text{in } \Omega, \end{aligned} \tag{2.1}$$

and, with our previous notation, the error equation now takes the form

$$T_h e_t + e = \rho \equiv -(T_h - T)\tilde{u}_t.$$

We shall first prove some bounds for the time derivatives of T_h.

LEMMA 2. *Assume that* (ii)$_k$ *holds with* $r = 2, k \geq 1$. *Then*

$$|(T_h' f, f)| \leq C\big((T_h f, f) + h^2\|f\|^2\big),$$

and

$$\|T_h^{(j)} f\| \leq C\big(\|T_h f\| + h^2\|f\|\big) \quad \textit{for } j \leq k.$$

Proof. We shall show the continuous counterparts of these estimates, namely

$$|(T'f, f)| \leq C(Tf, f), \tag{2.2}$$

and

$$\|T^{(j)} f\| \leq C\|Tf\| \quad \text{for } j \leq k. \tag{2.3}$$

Once this is done, the desired results follow easily by (ii)$_k$, as for example for the first inequality,

$$\begin{aligned} |(T_h' f, f)| &= |(T'f, f) + ((T_h' - T')f, f)| \leq C\big((Tf, f) + h^2\|f\|^2\big) \\ &\leq C\big((T_h f, f) + h^2\|f\|^2\big). \end{aligned}$$

For the continuous inequalities we differentiate the equation

$$A_\kappa(Tf, \varphi) = (f, \varphi) \quad \forall \varphi \in H_0^1(\Omega),$$

to obtain

$$A_\kappa(T'f, \varphi) + A'(Tf, \varphi) = 0 \quad \forall \varphi \in H_0^1(\Omega).$$

Hence, in particular,

$$\begin{aligned} |(T'f, f)| &= |A_\kappa(T'f, T^*f)| = |A'(Tf, T^*f)| \leq C\|Tf\|_1 \cdot \|T^*f\|_1 \\ &\leq C(A_\kappa(Tf, Tf))^{1/2}(A_\kappa(T^*f, T^*f))^{1/2} \\ &= C(f, Tf)^{1/2}(T^*f, f)^{1/2} = C(Tf, f), \end{aligned}$$

which is (2.2). To show (2.3) we use induction over j. Noting that it is trivially valid for $j = 0$, we assume that it holds for $j \leqslant l-1$ with $l > 0$. Since for any $\varphi \in L_2(\Omega)$,

$$(T^{(l)}f, \varphi) = A_\kappa(T^{(l)}f, T^*\varphi) = -\sum_{i=1}^{l} \binom{l}{i} A^{(i)}(T^{(l-i)}f, T^*\varphi)$$
$$= -\sum_{i=1}^{l} \binom{l}{i} (T^{(l-i)}f, A^{(i)*}T^*\varphi),$$

we have

$$|(T^{(l)}f, \varphi)| \leqslant C \sum_{i=1}^{l} |(T^{(l-i)}f, A^{(i)*}T^*\varphi)| \leqslant C\|Tf\|\cdot\|T^*\varphi\|_2 \leqslant C\|Tf\|\cdot\|\varphi\|,$$

which shows (2.3) for $j = l$. The proof of the lemma is now complete.

Our basic lemma for the subsequent error analysis in both the smooth data and the nonsmooth data cases is the following.

LEMMA 3. *Assume that* (ia), (ib), *and* (ii)$_1$ (*with* $r = 2$) *hold and that*

$$T_h w_t + w = g \quad \text{for } t \in I, \qquad T_h w(0) = 0.$$

Then, for sufficiently small h,

$$\int_0^t \|w\|^2\, d\tau \leqslant C \int_0^t \|g\|^2\, d\tau. \tag{2.4}$$

Further, for ε *positive there is a* C_ε *such that, for* $t \in I$,

$$\|w(t)\|^2 \leqslant \frac{\varepsilon^2}{t} \int_0^t \tau^2 \|g_t\|^2\, d\tau + C_\varepsilon \left\{ \|g(t)\|^2 + \frac{1}{t}\int_0^t \|g\|^2\, d\tau \right\} \tag{2.5}$$

and

$$\|w(t)\| \leqslant \varepsilon \sup_{\tau < t} (\tau\|g_t(\tau)\|) + C_\varepsilon \sup_{\tau < t} \|g(\tau)\|. \tag{2.6}$$

Proof. We multiply the equation for w by w to obtain

$$\frac{d}{dt}(T_h w, w) + 2\|w\|^2 = 2(g, w) + (T_h' w, w) + [(T_h w, w_t) - (T_h w_t, w)].$$

For the second term on the right we have by Lemma 2, for small h,

$$|(T_h' w, w)| \leqslant C(T_h w, w) + \tfrac{1}{4}\|w\|^2,$$

and for the last term, using (ib) and the equation for w,

$$|(T_h w, w_t) - (T_h w_t, w)| \leqslant C(T_h w, w)^{1/2} \|T_h w_t\|$$
$$\leqslant C(T_h w, w)^{1/2}(\|g\| + \|w\|) \leqslant C(T_h w, w) + C\|g\|^2 + \tfrac{1}{4}\|w\|^2.$$

Hence

$$\frac{d}{dt}(T_h w, w) + \|w\|^2 \leqslant C(\|g\|^2 + (T_h w, w)).$$

Using Gronwall's lemma and the fact that $T_h w(0) = 0$, this yields (2.4). In order to show (2.5) we multiply the equation for w by $2tw_t$ and obtain after some manipulation

$$2t(T_h w_t, w_t) + \frac{d}{dt}(t\|w\|^2) = 2\frac{d}{dt}(t(g, w)) - 2(g, w) - 2t(g_t, w) + \|w\|^2,$$

and hence, after integration and obvious estimates,

$$t\|w(t)\|^2 \leqslant \varepsilon^2 \int_0^t \tau^2 \|g_t\|^2 \, d\tau + C_\varepsilon \left\{ t\|g(t)\|^2 + \int_0^t \left(\|g\|^2 + \|w\|^2\right) d\tau \right\}.$$

This, together with (2.4), shows (2.5) which obviously implies (2.6) and thus completes the proof of the lemma.

We are now ready for our error estimates for the homogeneous equation and begin by considering the case of a smooth solution:

THEOREM 2. *Assume that* (ia), (ib), *and* $(\text{ii})_1$ *hold, that* $v \in H^r(\Omega)$ *is compatible of order r with* (2.1), *and that* $\|v_h - v\| \leqslant Ch^r \|v\|_r$. *Then*

$$\|u_h(t) - u(t)\| \leqslant Ch^r \|v\|_r \quad \text{for } t \in \bar{I}.$$

Proof. We first note that it suffices to consider the case $v_h = P_h v$. In fact, let $\tilde{u}_h$ and $\tilde{\tilde{u}}_h$ denote the solutions corresponding to the initial values v_h and $P_h v$, respectively. Then obviously

$$T_h\left(\tilde{u}_h - \tilde{\tilde{u}}_h\right)_t + \left(\tilde{u}_h - \tilde{\tilde{u}}_h\right) = 0,$$

and hence, by Lemma 1, for any $t \geqslant 0$,

$$\|\tilde{u}_h(t) - \tilde{\tilde{u}}_h(t)\| \leqslant \|v_h - P_h v\| \leqslant \|v_h - v\| + \|v - P_h v\| \leqslant Ch^r \|v\|_r.$$

We also note that if $v_h = P_h v$, then $T_h e(0) = 0$. For since $e(0) = P_h v - v$ is orthogonal to S_h, we have by (ib), for any $\chi \in S_h$,

$$|(T_h e(0), \chi)| = |(T_h e(0), \chi) - (e(0), T_h \chi)| \leqslant C(T_h e(0), e(0))^{1/2} \|T_h \chi\| = 0.$$

In order to complete the proof of the theorem we may thus assume $v_h = P_h v$ and apply Lemma 3. By assumption $(\text{ii})_1$ and the regularity estimate (1.13) we have

$$\|\rho(\tau)\| = \|(T_h - T)\tilde{u}_t(\tau)\| \leqslant Ch^r \|\tilde{u}_t(\tau)\|_{r-2} \leqslant Ch^r \|\tilde{u}(\tau)\|_r \leqslant Ch^r \|v\|_r,$$

and

$$\begin{aligned} \|\rho_t(\tau)\| &\leqslant \|(T_h' - T')\tilde{u}_t(\tau)\| + \|(T_h - T)\tilde{u}_{tt}(\tau)\| \\ &\leqslant Ch^r\left(\|\tilde{u}_t(\tau)\|_{r-2} + \|\tilde{u}_{tt}(\tau)\|_{r-2}\right) \leqslant Ch^r \|\tilde{u}(\tau)\|_{r+2} \leqslant Ch^r \tau^{-1} \|v\|_r, \end{aligned}$$

and hence, by Lemma 3, $\|e(t)\| \leqslant Ch^r \|v\|_r$, which completes the proof of the theorem.

We shall now proceed to prove our error estimate for nonsmooth initial data. The choice of discrete initial approximation will then be restricted to the L_2-projection $P_h v$. We start by considering the case $r = 2$ separately.

THEOREM 3. *Assume that* (ia), (ib), *and* $(\text{ii})_1$ (*with* $r = 2$) *hold, that* $v \in L_2(\Omega)$, *and* $v_h = P_h v$. *Then*

$$\|u_h(t) - u(t)\| \leqslant Ch^2 t^{-1} \|v\| \quad \text{for } t \in I.$$

Proof. We shall prove

$$\|e(t)\| \leqslant Ct^{-1} \sup_{\tau \leqslant t} \left\{\tau^2 \|\rho_t\| + \tau\|\rho\| + \|R\| + h^2 \|e\|\right\}, \tag{2.6}$$

where $R(t) = \int_0^t \rho \, d\tau$. Here, similarly as above, we have

$$\tau\|\rho(\tau)\| = \tau\|(T_h - T)\tilde{u}_t(\tau)\| \leqslant Ch^2 \tau \|\tilde{u}_t(\tau)\| < Ch^2 \|v\|,$$

and

$$\tau^2\|\rho_t(\tau)\| \leqslant \tau^2\|(T_h' - T')\tilde{u}_t\| + \tau^2\|(T_h - T)\tilde{u}_{tt}\|$$
$$\leqslant Ch^2\tau^2(\|\tilde{u}_t(\tau)\| + \|\tilde{u}_{tt}(\tau)\|) \leqslant Ch^2\|v\|.$$

Since

$$R(t) = \int_0^t (T_h - T)\tilde{u}_t\, d\tau = \left[(T_h - T)\tilde{u}(\tau)\right]_0^t - \int_0^t (T_h' - T')\tilde{u}\, d\tau,$$

we also have

$$\|R(\tau)\| \leqslant Ch^2 \sup_{\sigma < \tau} \|\tilde{u}(\sigma)\| \leqslant Ch^2\|v\|,$$

and the stability of the solution operators gives at once

$$\|e(\tau)\| \leqslant \|\tilde{u}_h(\tau)\| + \|\tilde{u}(\tau)\| \leqslant 2\|v\|.$$

Together these estimates show $\|e(t)\| \leqslant Ct^{-1}h^2\|v\|$, which proves the desired result.

In order to show (2.6) we set $w = te$. We shall demonstrate

(2.7) $$\|w(t)\| \leqslant C \sup_{\tau < t} \left\{\tau^2\|\rho_t\| + \tau\|\rho\| + \|T_h e\|\right\},$$

and thereafter

(2.8) $$\|T_h e(t)\| \leqslant C \sup_{\tau < t} \left\{\tau\|\rho\| + \|R\| + h^2\|e\|\right\}.$$

Together these estimates imply (2.6).

We have from multiplication of the error equation by t that

$$T_h w_t + w = \tilde{\rho} \equiv t\rho + T_h e.$$

Note that by the error equation and Lemma 2, since h is bounded,

$$\|\tilde{\rho}_t\| = \|t\rho_t + \rho + T_h e_t + T_h' e\| \leqslant C(t\|\rho_t\| + \|\rho\| + \|e\| + \|T_h e\|).$$

Hence, by Lemma 3, with suitable choice of ε, for $t \in \bar{I}$,

$$\|w(t)\| \leqslant \varepsilon \sup_{\tau < t}(\tau\|\tilde{\rho}_t\|) + C_\varepsilon \sup_{\tau < t}\|\tilde{\rho}\|$$
$$\leqslant \frac{1}{2} \sup_{\tau < t} \|w(\tau)\| + C \sup_{\tau < t} \left\{\tau^2\|\rho_t\| + \tau\|\rho\| + \|T_h e\|\right\},$$

which easily gives (2.7).

For (2.8) we integrate the error equation to obtain, for $\tilde{e} = \int_0^t e\, d\tau$,

$$T_h \tilde{e}_t + \tilde{e} = R + \int_0^t T_h' e\, d\tau.$$

Lemma 3 yields, with Lemma 2,

$$\|\tilde{e}(t)\| \leqslant \varepsilon \sup_{\tau < t}(\tau\|R'\| + \tau\|T_h' e\|) + C_\varepsilon \sup_{\tau < t}\left(\|R\| + \left\|\int_0^\tau T_h' e\, d\sigma\right\|\right)$$
$$\leqslant \varepsilon \sup_{\tau < t}\left(\tau\|\rho\| + C\tau\|T_h e\| + C\tau h^2\|e\|\right) + C_\varepsilon \sup_{\tau < t} \|R\|$$
$$+ C_\varepsilon \int_0^t \left(\|T_h e\| + h^2\|e\|\right) d\sigma,$$

and observing that $\tilde{e}_t = e$ this shows

$$\|T_h e(t)\| \leqslant \|\tilde{e}\| + \|R\| + \left\|\int_0^t T_h' e\, d\tau\right\|$$
$$\leqslant \varepsilon C \sup_{\tau < t}(\tau\|T_h e\|) + C_\varepsilon \sup_{\tau < t}(\tau\|\rho\| + \|R\| + th^2\|e\|) + C_\varepsilon \int_0^t \|T_h e\|\, d\tau.$$

Choosing ε such that $\varepsilon C t^0 = \frac{1}{2}$, this gives, for $t \in \bar{I} = [0, t^0]$,

$$\|T_h e(t)\| \leqslant C \sup_{\tau < t} (\tau \|\rho\| + \|R\| + h^2 \|e\|) + C \int_0^t \|T_h e\| \, d\tau,$$

which implies (2.8) by Gronwall's lemma, thus completing the proof of the lemma.

We have finally the following result for general r.

THEOREM 4. *Assume that* (ia), (ib), $(\mathrm{ii})_1$ *and* $(\mathrm{ii})_1^*$ *hold, that* $v \in L_2(\Omega)$, *and* $v_h = P_h v$. *Then*

$$\|u_h(t) - u(t)\| \leqslant C h^r t^{-r/2} \|v\| \quad \text{for } t \in I.$$

Proof. Recall from the introduction the properties of the solution operator $E(t, s)$ of the homogeneous equation. It is easy to see that the adjoint $E(t, s)^*$ of $E(t, s)$ is the solution operator $\tilde{E}(s, t)$ of the backward problem for the equation

$$w_t - A^* w = 0, \tag{2.9}$$

where A^* is the formal adjoint of A, so that, for a solution w of this equation, $w(s) = \tilde{E}(s, t) w(t)$ for $s \leqslant t$. In particular, then, we have analogously to the corresponding property (1.12) of $E(t, s)$,

$$\|E(t, s)^* v\|_r = \|\tilde{E}(s, t) v\|_r \leqslant C(t - s)^{-r/2} \|v\| \quad \text{for } 0 \leqslant s < t \in I.$$

We introduce similarly the solution operator $E_h(t, s)$ of the semidiscrete counterpart of (2.1) and the error operator for this problem,

$$F_h(t, s) = E_h(t, s) P_h - E(t, s).$$

With this notation the result of Theorem 3 can be stated as

$$\|F_h(t, 0) v\| \leqslant C h^2 t^{-1} \|v\| \quad \text{for } t \in I,$$

and it is clear that then also

$$\|F_h(t, s) v\| \leqslant C h^2 (t - s)^{-1} \|v\| \quad \text{for } 0 \leqslant s < t \in I. \tag{2.10}$$

Similarly, Theorem 2 can be expressed

$$\|F_h(t, s) v\| \leqslant C h^r \|v\|_r \quad \text{for } 0 \leqslant s \leqslant t \in \bar{I}, \tag{2.11}$$

for v compatible with the differential equation (2.1) at $t = s$, and since $F_h(t, s)^*$ is the error operator for the semidiscrete analogue of the backward equation (2.9), we also have, for v compatible with Eq. (2.9) at $s = t$,

$$\|F_h(t, s)^* v\| \leqslant C h^r \|v\|_r.$$

After these preparations we are now ready to complete the proof of the theorem for $r > 2$. Since

$$\|F_h(t, 0) v\| \leqslant \|E_h(t, 0) P_h v\| + \|E(t, 0) v\| \leqslant 2 e^{\kappa t} \|v\| \leqslant C \|v\|,$$

it suffices to consider the case $h^2 t \leqslant 1$. Using (1.10) and its discrete analogue, we have by a simple computation

$$F_h(t, 0) = F_h(t, t/2) E(t/2, 0) + E(t, t/2) F_h(t/2, 0) + F_h(t, t/2) F_h(t/2, 0).$$

By (2.11) and (1.11) we have for the first term, since $E(t/2, 0) v$ is compatible with (2.1) at time $t/2$,

$$\|F_h(t, t/2) E(t/2, 0) v\| \leqslant C h^r \|E(t/2, 0) v\|_r \leqslant C h^r t^{-r/2} \|v\|.$$

For the second term we conclude, since an operator and its adjoint have the same norm, and since $E(t, t/2)^*v$ is compatible with (2.9) at time $t/2$,

$$\|E(t, t/2)F_h(t/2, 0)v\| \leqslant Ch^r t^{-r/2}\|v\|.$$

Using (2.10) for the third term, we have thus

$$\|F_h(t, 0)v\| \leqslant Ch^r t^{-r/2}\|v\| + Ch^2 t^{-1}\|F_h(t/2, 0)v\|.$$

By iteration this yields, for $2s \geqslant r$,

$$\|F_h(t, 0)v\| \leqslant Ch^r t^{-r/2}\|v\| + C(h^2 t^{-1})^s\|F_h(t\cdot 2^{-s}, 0)v\| \leqslant Ch^r t^{-r/2}\|v\|.$$

This completes the proof of the theorem.

We complete this section by a nonsmooth data error estimate for an arbitrary time derivative.

THEOREM 5. *Assume that* (ia), (ib), $(\mathrm{ii})_{l+1}$, *and* $(\mathrm{ii})_1^*$ *hold, that* $v \in L_2(\Omega)$, *and that* $v_h = P_h v$. *Then, for each* $l \geqslant 0$,

$$\|u_h^{(l)}(t) - u^{(l)}(t)\| \leqslant Ch^r t^{-r/2-l}\|v\| \quad \text{for } t \in I.$$

Proof. By differentiation of the error equation we have

$$T_h e_t^{(l)} + e^{(l)} = \rho^{(l)} - \sum_{i=0}^{l-1}\binom{l}{i}T_h^{(l-i)}e^{(i+1)}.$$

Setting $w_{(l)}(t) = t^{l+r/2}e^{(l)}(t)$ and $\gamma_{(l)}(t) = t^{l+r/2}\rho^{(l)}(t)$, we shall prove

$$\|w_{(l)}(t)\| \leqslant C\sup_{\tau\leqslant t}\left\{\sum_{i=0}^{l+1}\|\gamma_{(i)}(\tau)\| + \sum_{i=0}^{l-1}\|w_{(i)}(\tau)\|\right\}. \tag{2.12}$$

If this has been done, the conclusion of the theorem follows easily by induction since we know from Theorem 4 that

$$\|w_{(0)}(\tau)\| = \tau^{r/2}\|e(\tau)\| \leqslant Ch^r\|v\|,$$

and since, by $(\mathrm{ii})_{l+1}$, for each $i \leqslant l+1$,

$$\|\gamma_{(i)}(\tau)\| = \tau^{i+r/2}\|\rho^{(i)}(\tau)\| = \tau^{i+r/2}\left\|\sum_{j=0}^{i}\binom{i}{j}(T_h^{(j)} - T^{(j)})\tilde{u}_t^{(i-j)}\right\|$$

$$\leqslant Ch^r\tau^{i+r/2}\sum_{j=0}^{i}\|\tilde{u}_t^{(i-j)}(\tau)\|_{r-2} \leqslant Ch^r\tau^{i+r/2}\|\tilde{u}(\tau)\|_{2i+r} \leqslant Ch^r\|v\|.$$

To show (2.12) we observe that $w_{(l)}$ satisfies

$$\begin{aligned} T_h w_{(l),t} + w_{(l)} = \sigma_{(l)}(t) &\equiv \gamma_{(l)}(t) + \left(l + \frac{r}{2}\right)\gamma_{(l-1)}(t) - lT_h' w_{(l)} \\ &- \left(l + \frac{r}{2}\right)w_{(l-1)} - \sum_{i=0}^{l-2}\binom{l}{i}t^{l-1-i}T_h^{(l-i)}w_{(i+1)} \\ &- \left(l + \frac{r}{2}\right)\sum_{i=0}^{l-2}\binom{l-1}{i}t^{l-2-i}T_h^{(l-1-i)}w_{(i+1)}. \end{aligned}$$

Noticing that by Lemma 2, for any $\varepsilon > 0$ and h small,

$$\|T_h' w_{(l)}\| \leqslant C\|T_h w_{(l)}\| + \varepsilon\|w_{(l)}\| \leqslant C(\|\sigma_{(l-1)}\| + \|w_{(l-1)}\|) + \varepsilon\|w_{(l)}\|,$$

and using the fact that the $T_h^{(i)}$ are bounded we find easily, for $t \in \bar{I}$,

$$\|\sigma_{(l)}\| \leqslant C\left(\sum_{i=0}^{l} \|\gamma_{(i)}\| + \sum_{i=0}^{l-1} \|w_{(i)}\|\right) + \varepsilon\|w_{(l)}\|. \tag{2.13}$$

Since

$$w_{(l),t}(\tau) = \tau^{-1}w_{(l+1)}(\tau) + \left(l + \frac{r}{2}\right)\tau^{-1}w_{(l)},$$

and analogously for $\gamma_{(l),t}$, we find similarly

$$\tau\|\sigma_{(l),t}(\tau)\| \leqslant C\sum_{i=0}^{l+1} (\|\gamma_{(i)}(\tau)\| + \|w_{(i)}(\tau)\|). \tag{2.14}$$

We now apply our basic Lemma 3. First, by (2.5) we have

$$\|w_{(l)}(t)\|^2 \leqslant C\left\{\|\sigma_{(l)}(t)\|^2 + \frac{1}{t}\int_0^t (\|\sigma_{(l)}\|^2 + \tau^2\|\sigma_{(l),t}\|^2)\, d\tau\right\},$$

and hence, using (2.13) with ε sufficiently small and (2.14),

$$\begin{aligned} \|w_{(l)}(t)\|^2 \leqslant C\Bigg\{ &\sum_{i=0}^{l} \|\gamma_{(i)}(t)\|^2 + \sum_{i=0}^{l-1} \|w_{(i)}(t)\|^2 \\ &+ \frac{1}{t}\int_0^t \left(\sum_{i=0}^{l+1} \|\gamma_{(i)}\|^2 + \sum_{i=0}^{l+1} \|w_{(i)}\|^2\right) d\tau\Bigg\}. \end{aligned} \tag{2.15}$$

Similarly, we have by (2.4), for each $i \geqslant 0$,

$$\int_0^t \|w_{(i)}\|^2\, d\tau \leqslant C\int_0^t \left(\sum_{j=0}^{i} \|\gamma_{(j)}\|^2 + \sum_{j=0}^{i-1} \|w_{(j)}\|^2\right) d\tau.$$

Using this with $i = l + 1$ and l in (2.15) yields

$$\begin{aligned} \|w_{(l)}(t)\|^2 \leqslant C\Bigg\{ &\sum_{i=0}^{l} \|\gamma_{(i)}(t)\|^2 + \sum_{i=0}^{l-1} \|w_{(i)}(t)\|^2 \\ &+ \frac{1}{t}\int_0^t \left(\sum_{i=0}^{l+1} \|\gamma_{(i)}\|^2 + \sum_{i=0}^{l-1} \|w_{(i)}\|^2\right) d\tau\Bigg\}, \end{aligned}$$

which implies the desired inequality (2.12). The proof of the theorem is now complete.

3. Applications. In this final section we shall consider several different choices of approximate solution operators T_h of the elliptic problem and show the validity of our hypotheses for these. We shall assume here for simplicity, without restricting the generality, that $\kappa = 0$ so that $A_\kappa(\cdot, \cdot) = A(\cdot, \cdot)$. The analysis below has several points in common with that in [7].

I. *The Standard Galerkin Method.* Recall that, for this method, $\{T_h\}$ is defined from a family $\{S_h\}$ with $S_h \subset H_0^1(\Omega)$ such that

$$\begin{aligned} &\inf_{\chi\in S_h} \{\|v - \chi\| + h\|v - \chi\|_1\} \\ &\qquad \leqslant Ch^s\|v\|_s \quad \text{for } v \in H^s(\Omega) \cap H_0^1(\Omega),\ 1 \leqslant s \leqslant r, \end{aligned} \tag{3.1}$$

by T_h: $L_2(\Omega) \to S_h$ and

$$A(T_hf, \chi) = (f, \chi) \quad \forall\chi \in S_h.$$

We have at once

$$(f, T_h f) = A(T_h f, T_h f) \geqslant c\|T_h f\|_1^2 \geqslant 0,$$

which shows the first inequality of (ia) and also that equality holds only if $T_h f = 0$. Assume that then $f = \chi \in S_h$. Using the definition of T_h once more, we have

$$\|\chi\|^2 = A(T_h \chi, \chi) = 0,$$

so that $\chi = 0$ which shows the second inequality of (ia).

By our definitions we have with $v_h = T_h f$, $w_h = T_h g$,

$$\begin{aligned}(T_h f, g) - (f, T_h g) &= A(T_h g, T_h f) - A(T_h f, T_h g)\\ &= \int_\Omega \sum_j b_j\left(\frac{\partial w_h}{\partial x_j} v_h - \frac{\partial v_h}{\partial x_j} w_h\right) dx\\ &= -\int_\Omega \sum_j \left(2b_j \frac{\partial v_h}{\partial x_j} w_h + \frac{\partial b_j}{\partial x_j} v_h w_h\right) dx\\ &\leqslant C\|v_h\|_1\|w_h\| = C\|T_h f\|_1\|T_h g\| \leqslant C(f, T_h f)^{1/2}\|T_h g\|,\end{aligned} \tag{3.2}$$

which shows (ib).

We now turn to the conditions $(\text{ii})_k$. We shall show that, for each $j \geqslant 0$,

$$\|(T_h^{(j)} - T^{(j)})f\|_q \leqslant Ch^{s-q}\|f\|_{s-2} \quad \text{for } q = 0, 1, \quad 2 \leqslant s \leqslant r. \tag{3.3}$$

For $j = 0$ these are the well-known standard Galerkin elliptic error estimates and it remains to consider the successive time derivatives. We begin by proving the H^1-result ($q = 1$) by induction over j and assume thus that this has been accomplished for $j < l$. Letting $e = T_h f - Tf$, we obtain by differentiating the equation $A(e, \chi) = 0$ that

$$\sum_{j=0}^{l} \binom{l}{j} A^{(l-j)}(e^{(j)}, \chi) = 0 \quad \forall \chi \in S_h, \tag{3.4}$$

and hence

$$\begin{aligned}\|e^{(l)}\|_1^2 &\leqslant CA(e^{(l)}, e^{(l)})\\ &= C\Bigg\{A(e^{(l)}, e^{(l)} - \chi)\\ &\qquad + \sum_{j=0}^{l-1} \binom{l}{j}\left[A^{(l-j)}(e^{(j)}, e^{(l)} - \chi) - A^{(l-j)}(e^{(j)}, e^{(l)})\right]\Bigg\}\\ &\leqslant \frac{1}{2}\|e^{(l)}\|_1^2 + C\sum_{j=0}^{l-1} \|e^{(j)}\|_1^2 + C\|e^{(l)} - \chi\|_1^2.\end{aligned} \tag{3.5}$$

Since $\chi \in S_h$ is arbitrary, we obtain, using the induction assumption and the approximability assumption (3.1),

$$\|e^{(l)}\|_1 \leqslant C\sum_{j=0}^{l-1} \|e^{(j)}\|_1^2 + C \inf_{\chi \in S_h} \|T^{(l)}f - \chi\|_1 \leqslant Ch^{s-1}\|f\|_{s-2},$$

which is the desired result. In the last step we have applied the estimate

$$\|T^{(l)}f\|_s \leqslant C\|f\|_{s-2} \quad \text{for } s \geqslant 2, l \geqslant 0,$$

which is easily proved by induction over l using the fact that $w = T^{(l)}f$ solves the elliptic problem

$$Aw = -\sum_{j=0}^{l-1} A^{(l-j)}T^{(j)}f \quad \text{in } \Omega, \qquad w = 0 \quad \text{on } \partial\Omega.$$

We now consider the estimate (3.3) for $q = 0$. Let φ be arbitrary in $L_2(\Omega)$ and assume that the estimate has been shown for $j < l$. Setting $z = T^*\varphi \in H^2(\Omega) \cap H_0^1(\Omega)$, we have, for any $\chi \in S_h$,

$$\begin{aligned}(e^{(l)}, \varphi) &= A(e^{(l)}, z)\\ &= \sum_{j=0}^{l} \binom{l}{j} A^{(l-j)}(e^{(j)}, z - \chi) - \sum_{j=0}^{l-1} \binom{l}{j} (e^{(j)}, A^{*(l-j)}z),\end{aligned} \tag{3.6}$$

and hence

$$\begin{aligned}|(e^{(l)}, \varphi)| &\leqslant C\sum_{j=0}^{l} \|e^{(j)}\|_1 \inf_{\chi \in S_h} \|z - \chi\|_1 + C\sum_{j=0}^{l-1} \|e^{(j)}\| \, \|z\|_2\\ &\leqslant Ch^s\|f\|_{s-2}\|z\|_2 \leqslant Ch^s\|f\|_{s-2}\|\varphi\|,\end{aligned} \tag{3.7}$$

thus completing the proof of (3.3).

Recall that T^* and T_h^* are simply the operators corresponding to the adjoint A^* of A, so that $(\text{ii})_1^*$ follows as $(\text{ii})_1$.

II. *Babuška's Method.* This method can be formulated as follows ([1], [3]). Let $\{\mathbb{S}_h\}$ be a family of subspaces of $H^1(\Omega)$, and $\{\mathbb{S}_h'\}$ a family of subspaces of $H^1(\partial\Omega)$ enjoying the approximation properties

$$\inf_{\chi \in \mathbb{S}_h} \{\|w - \chi\| + h\|w - \chi\|_1\} \leqslant Ch^s\|w\|_s \quad \text{for } 1 \leqslant s \leqslant r, \tag{3.8}$$

and

$$\begin{aligned}&\inf_{\chi' \in \mathbb{S}_h'} \{h^{-1/2}\|w' - \chi'\|_{H^{-1/2}(\partial\Omega)} + h^{1/2}\|w' - \chi'\|_{H^{1/2}(\partial\Omega)}\}\\ &\qquad \leqslant Ch^s\|w'\|_{H^s(\partial\Omega)} \quad \text{for } \tfrac{1}{2} \leqslant s \leqslant r - \tfrac{3}{2},\end{aligned} \tag{3.9}$$

and the inverse properties

$$\|\chi\|_1 \leqslant Ch^{-1}\|\chi\| \quad \text{for } \chi \in \mathbb{S}_h,$$

and

$$\|\chi'\|_{H^1(\partial\Omega)} \leqslant Ch^{-1}\|\chi'\|_{L_2(\partial\Omega)} \quad \text{for } \chi' \in \mathbb{S}_h'.$$

Then with δ a sufficiently small positive number we define a family $\{S_h\}$ by

$$S_h = \{\chi \in \mathbb{S}_{\delta h};\ \langle \chi, \chi' \rangle = 0\ \forall \chi' \in \mathbb{S}_h'\},$$

where $\langle \cdot\,, \cdot \rangle$ is the inner product in $L_2(\partial\Omega)$, and a family $\{T_h\}$ with T_h: $L_2(\Omega) \to S_h$ by

$$A(T_hf, \chi) = (f, \chi) \qquad \forall \chi \in S_h. \tag{3.10}$$

The discussion of our hypotheses will follow that for the standard Galerkin method. Condition (ia) is satisfied in the same way as before. Since $S_h \not\subset H_0^1(\Omega)$, we obtain an additional term in the integration by parts in (3.2) in the proof of (ib), so that we need to show now also

$$\left|\int_{\partial\Omega} b_j n_j v_h w_h \, ds\right| \leqslant C\|v_h\|_1 \|w_h\|, \tag{3.11}$$

where $v_h = T_h f$, $w_h = T_h g$, and n_j are the components of the exterior normal at $\partial\Omega$. By the properties of $\mathbb{S}_h$, $\mathbb{S}_h'$ and the definition of S_h, and using also the trace inequality $\|v\|_{H^{1/2}(\partial\Omega)} \leqslant C\|v\|_1$, we have, for a suitably chosen $\chi_j' \in \mathbb{S}_h'$,

$$\begin{aligned}\left|\int_{\partial\Omega} b_j n_j v_h w_h \, ds\right| &= \left|\int_{\partial\Omega} (b_j n_j v_h - \chi_j') w_h \, ds\right| \\ &\leqslant C\|b_j n_j v_h - \chi'\|_{H^{-1/2}(\partial\Omega)} \|w_h\|_{H^{1/2}(\partial\Omega)} \leqslant Ch\|b_j n_j v_h\|_{H^{1/2}(\partial\Omega)} \|w_h\|_{H^{1/2}(\partial\Omega)} \\ &\leqslant Ch\|b_j n_j v_h\|_1 \|w_h\|_1 \leqslant C\|v_h\|_1 \|w_h\|.\end{aligned}$$

For $(\text{ii})_k$ it is again known that (3.3) holds for $j = 0$. In particular, we shall use below that, for $z|_{\partial\Omega} = 0$,

$$\inf_{\chi \in S_h} \|z - \chi\|_1 \leqslant \|(T - T_h)Az\|_1 \leqslant Ch^{s-1}\|z\|_s \quad \text{for } 2 \leqslant s \leqslant r. \tag{3.12}$$

We begin the proof of (3.3) for $j > 0$ by showing the H^1 estimate by induction over j. This time, with $e = T_h f - Tf$, (3.4) is not valid since the elements of S_h do not vanish on $\partial\Omega$. By differentiation of (3.10) we still have

$$\sum_{j=0}^{l} \binom{l}{j} A^{(l-j)}(T_h^{(j)} f, \chi) = 0 \quad \forall \chi \in S_h, \tag{3.13}$$

and, since

$$\sum_{j=0}^{l} \binom{l}{j} A^{(l-j)} T^{(j)} f = 0,$$

we obtain now

$$\sum_{j=0}^{l} \binom{l}{j} A^{(l-j)}(e^{(j)}, \chi) = -\sum_{j=0}^{l} \binom{l}{j} \left\langle \frac{\partial}{\partial \nu^{(l-j)}} T^{(j)} f, \chi \right\rangle, \tag{3.14}$$

where $\partial/\partial\nu^{(l-j)} = \sum_{j,k} n_j a_{jk}^{(l-j)} \partial/\partial x_k$. Therefore, we have to add to the right in the estimate (3.5) the term

$$C \sum_{j=0}^{l} \left| \left\langle \frac{\partial}{\partial \nu^{(l-j)}} T^{(j)} f, \chi \right\rangle \right|. \tag{3.15}$$

Here by the definition of S_h we have, for the appropriate $\chi_j' \in \mathbb{S}_h'$,

$$\begin{aligned}\left| \left\langle \frac{\partial}{\partial \nu^{(l-j)}} T^{(j)} f, \chi \right\rangle \right| &= \left| \left\langle \frac{\partial}{\partial \nu^{(l-j)}} T^{(j)} f - \chi_j', \chi \right\rangle \right| \\ &\leqslant \left\| \frac{\partial}{\partial \nu^{(l-j)}} T^{(j)} f - \chi_j' \right\|_{H^{-1/2}(\partial\Omega)} \|\chi\|_{H^{1/2}(\partial\Omega)} \\ &\leqslant Ch^{s-1} \left\| \frac{\partial}{\partial \nu^{(l-j)}} T^{(j)} f \right\|_{H^{s-3/2}(\partial\Omega)} \|\chi\|_1 \\ &\leqslant Ch^{s-1}\|f\|_{s-2}\left(\|e^{(l)}\|_1 + \|e^{(l)} - \chi\|_1\right).\end{aligned} \tag{3.16}$$

We may therefore conclude as from (3.5), using the induction assumption, (3.16), and (3.12), that

$$\|e^{(l)}\|_1 \leqslant Ch^{s-1}\|f\|_{s-2} + C \inf_{\chi \in S_h} \|T^{(l)}f - \chi\|_1 \leqslant Ch^{s-1}\|f\|_{s-2}. \tag{3.17}$$

We now turn to the case $q = 0$. Similarly to above, since (3.4) is not valid, we have instead of (3.6) that, for any $\chi \in S_h$, and with $z = T^*\varphi$,

$$\begin{aligned}(e^{(l)}, \varphi) = (e^{(l)}, A^*z) &= \left(\sum_{j=0}^{l} - \sum_{j=0}^{l-1}\right)\binom{l}{j}(e^{(j)}, A^{*(l-j)}z)\\ &= \sum_{j=0}^{l}\binom{l}{j}\big[\big((e^{(j)}, A^{*(l-j)}z) - A^{(l-j)}(e^{(j)}, z)\big)\\ &\qquad + A^{(l-j)}(e^{(j)}, z - \chi) + A^{(l-j)}(e^{(j)}, \chi)\big]\\ &\quad - \sum_{j=0}^{l-1}\binom{l}{j}(e^{(j)}, A^{*(l-j)}z).\end{aligned} \tag{3.18}$$

Here, since $e^{(j)} = T_h^{(j)}f$ on $\partial\Omega$, we have with suitably chosen $\chi_j' \in \mathbb{S}_h'$, for $j = 0, \ldots, l$,

$$\begin{aligned}\left|(e^{(j)}, A^{*(l-j)}z) - A^{(l-j)}(e^{(j)}, z)\right| &= \left|\left\langle e^{(j)}, \frac{\partial z}{\partial \nu^{(l-j)}} - \chi_j'\right\rangle\right|\\ &\leqslant C\|e^{(j)}\|_{H^{1/2}(\partial\Omega)}\left\|\frac{\partial z}{\partial \nu^{(l-j)}} - \chi'\right\|_{H^{-1/2}(\partial\Omega)} \leqslant Ch\|e^{(j)}\|_1\left\|\frac{\partial z}{\partial \nu^{(l-j)}}\right\|_{H^{1/2}(\partial\Omega)}\\ &\leqslant Ch^s\|f\|_{s-2}\|z\|_2 \leqslant Ch^s\|f\|_{s-2}\|\varphi\|.\end{aligned}$$

Further, with χ suitable in S_h, for $j = 0, \ldots, l$,

$$\left|A^{(l-j)}(e^{(j)}, z - \chi)\right| \leqslant C\|e^{(j)}\|_1\|z - \chi\|_1 \leqslant Ch^s\|f\|_{s-2}\|\varphi\|, \tag{3.19}$$

and, with the same χ and $\chi_j' \in \mathbb{S}_h'$ suitable, from (3.9),

$$\begin{aligned}\left|\sum_{j=0}^{l}\binom{l}{j}A^{(l-j)}(e^{(j)}, \chi)\right| &= \left|\sum_{j=0}^{l}\binom{l}{j}\left\langle \frac{\partial}{\partial \nu^{(l-j)}}T^{(j)}f - \chi_j', z - \chi\right\rangle\right|\\ &\leqslant Ch^{s-1}\left\|\frac{\partial}{\partial \nu^{(l-j)}}T^{(j)}f\right\|_{H^{s-3/2}(\partial\Omega)}\|z - \chi\|_1 \leqslant Ch^s\|f\|_{s-2}\|\varphi\|.\end{aligned} \tag{3.20}$$

Finally, by the induction assumption, for $j = 0, \ldots, l - 1$,

$$\left|(e^{(j)}, A^{*(l-j)}z)\right| \leqslant C\|e^{(j)}\|\,\|z\|_2 \leqslant Ch^2\|f\|_{s-2}\|\varphi\|. \tag{3.21}$$

Together these estimates show

$$\left|(e^{(j)}, \varphi)\right| \leqslant Ch^s\|f\|_{s-2}\|\varphi\|,$$

which completes the proof by induction of (3.7) and thus of (3.3) in the present situation.

As for the standard Galerkin method, we find at once that T^* and T_h^* satisfy

$$A(\varphi, T^*f) = (\varphi, f) \quad \forall \varphi \in H_0^1(\Omega),$$

and

$$A(\chi, T_h^*f) = (\chi, f) \quad \forall \chi \in S_h,$$

and hence that the above proof of $(\mathrm{ii})_1$ is easily modified to yield $(\mathrm{ii})_1^*$.

III. *The Method of Interpolated Boundary Conditions.* For this method applicable to a plane region Ω (cf. [9], [2]), S_h consists of continuous piecewise polynomials of degree at most $r - 1$ (with $r \geqslant 3$) on a nondegenerate triangulation of Ω which vanish at the boundary vertices and certain carefully chosen boundary edge nodes, and T_h: $L_2(\Omega) \to S_h$ is defined as before by (3.10).

It is shown in [9] that, for $\chi \in S_h$,

$$\|\chi\|_{H^{-s}(\partial\Omega)} \leqslant Ch^{s+3/2}\|\chi\|_1 \quad \text{for } -1 \leqslant s \leqslant r-2, \tag{3.22}$$

and

$$\|\chi\|_{L_\infty(\partial\Omega)} \leqslant Ch^{-1}\|\chi\|. \tag{3.23}$$

Letting $u_I \in S_h$ denote the interpolant of $u \in H_0^1(\Omega)$ as defined in [9], we also quote

$$\|u - u_I\|_1 \leqslant Ch^{s-1}\|u\|_s \quad \text{for } 2 \leqslant s \leqslant r, \tag{3.24}$$

and extract, from the proof of Lemma 6 of [9],

$$\|u_I\|_{H^{-q}(\partial\Omega)} = \|u - u_I\|_{H^{-q}(\partial\Omega)} \leqslant Ch^{s+q}\|u\|_s \quad \text{for } 0 \leqslant q \leqslant r-2, 2 \leqslant s \leqslant r. \tag{3.25}$$

Our condition (ia) is satisfied for this method as before, and, as for Babuška's method, (ib) will follow from (3.11); we have now by (3.22) and (3.23),

$$\begin{aligned}\left|\int_{\partial\Omega} b_j n_j v_h w_h \, ds\right| &\leqslant C\|v_h\|_{L_2(\partial\Omega)}\|w_h\|_{L_2(\partial\Omega)} \\ &\leqslant Ch^{3/2}\|v_h\|_1\|w_h\|_{L_\infty} \leqslant C\|v_h\|_1\|w_h\|.\end{aligned}$$

For $(\text{ii})_k$ it remains again to show (3.3) for $j > 0$ since the case $j = 0$ is treated in [9]. Similarly to Babuška's method, the proof of (3.3) for $q = 1$ requires an estimate for the sum in (3.15). This time we have by (3.22), with $2 \leqslant s \leqslant r$,

$$\begin{aligned}\left|\left\langle \frac{\partial}{\partial\nu^{(l-j)}} T^{(j)}f, \chi\right\rangle\right| &\leqslant \left\|\frac{\partial}{\partial\nu^{(l-j)}} T^{(j)}f\right\|_{H^{s-3/2}(\partial\Omega)} \|\chi\|_{H^{-(s-3/2)}(\partial\Omega)} \\ &\leqslant C\|T^{(j)}f\|_s\|\chi\|_{H^{-(s-2)}(\partial\Omega)} \leqslant Ch^{s-1/2}\|f\|_{s-2}\|\chi\|_1 \\ &\leqslant Ch^{s-1}\|f\|_{s-2}(\|e^{(l)}\|_1 + \|e^{(l)} - \chi\|_1),\end{aligned}$$

and from this the proof may be concluded as before. For the proof of (3.3) for $q = 0$ we again employ the representation (3.18), now with $\chi = z_I$. Then (3.19) and (3.21) are still valid. We have by (3.22), (3.24), (3.25), and the result for $q = 1$, for $j \leqslant l$,

$$\begin{aligned}\|e^{(j)}\|_{L_2(\partial\Omega)} &= \|T_h^{(j)}f\|_{L_2(\partial\Omega)} \leqslant \|T_h^{(j)}f - (T^{(j)}f)_I\|_{L_2(\partial\Omega)} + \|(T^{(j)}f)_I\|_{L_2(\partial\Omega)} \\ &\leqslant Ch^{3/2}\|T_h^{(j)}f - (T^{(j)}f)_I\|_1 + Ch^s\|T^{(j)}f\|_s \\ &\leqslant Ch^{3/2}(\|T_h^{(j)}f - T^{(j)}f\|_1 + \|T^{(j)}f - (T^{(j)}f)_I\|_1) + Ch^s\|f\|_{s-2} \\ &\leqslant Ch^s\|f\|_{s-2},\end{aligned}$$

and hence

$$\begin{aligned}\left|\left\langle e^{(j)}, \frac{\partial z}{\partial\nu^{(l-j)}}\right\rangle\right| &\leqslant \|e^{(j)}\|_{L_2(\partial\Omega)}\left\|\frac{\partial z}{\partial\nu^{(l-j)}}\right\|_{L_2(\partial\Omega)} \\ &\leqslant Ch^s\|f\|_{s-2}\|z\|_2 \leqslant Ch^s\|f\|_{s-2}\|\varphi\|,\end{aligned}$$

so that

$$\left|\sum_{j=0}^{l}\binom{l}{j}\left[(e^{(j)}, A^{*(l-j)}z) - A^{(l-j)}(e^{(j)}, z)\right]\right|$$
$$= \left|\sum_{j=0}^{l}\binom{l}{j}\left\langle e^{(j)}, \frac{\partial z}{\partial \nu^{(l-j)}}\right\rangle\right| \leqslant Ch^s\|f\|_{s-2}\|\varphi\|.$$

Finally, by (3.25) we have

$$\left|\sum_{j=0}^{l}\binom{l}{j}A^{(l-j)}(e^{(j)}, z_l)\right| = \left|\sum_{j=0}^{l}\binom{l}{j}\left\langle \frac{\partial}{\partial \nu^{(l-j)}} T^{(j)}f, z_l\right\rangle\right|$$
$$\leqslant C\|z_l\|_{H^{-(s-3/2)}(\partial\Omega)}\sum_{j=0}^{l}\left\|\frac{\partial}{\partial \nu^{(l-j)}} T^{(j)}f\right\|_{H^{s-3/2}(\partial\Omega)}$$
$$\leqslant Ch^{s+1/2}\|f\|_{s-2}\|z\|_2 \leqslant Ch^s\|f\|_{s-2}\|\varphi\|,$$

which shows (3.20). This completes the proof of (3.3) as above.

The proof of $(\text{ii})_1^*$ is again analogous to that of $(\text{ii})_1$.

IV. *Nitsche's Method.* This method (cf. [6], [3]) uses the bilinear form

$$B_h(v, w) = A(v, w) - \left\langle \frac{\partial v}{\partial \nu}, w\right\rangle - \left\langle v, \frac{\partial w}{\partial \nu} + \sum_j b_j n_j w\right\rangle + \beta h^{-1}\langle v, w\rangle,$$

with β a positive constant. Introducing the norm

$$|||v||| = \|\nabla v\| + h^{1/2}\|\nabla v\|_{L_2(\partial\Omega)} + h^{-1/2}\|v\|_{L_2(\partial\Omega)},$$

we assume that $\{S_h\}$ is a family of finite-dimensional spaces such that $|||\cdot|||$ is well defined on S_h, that S_h satisfies the approximation property

$$(3.26)\qquad \inf_{\chi\in S_h}\{\|v-\chi\| + h|||v-\chi|||\} \leqslant Ch^s\|v\|_s \quad \text{for } v \in H^s(\Omega),\ 2 \leqslant s \leqslant r,$$

and that the inverse estimates

$$(3.27)\qquad \|\chi\|_{L_2(\partial\Omega)} \leqslant Ch^{-1/2}\|\chi\| \quad \forall \chi \in S_h,$$

and

$$\|\nabla\chi\|_{L_2(\partial\Omega)} \leqslant Ch^{-1/2}\|\chi\|_1 \quad \forall \chi \in S_h,$$

hold. For the derivatives with respect to time, we have

$$|B_h^{(j)}(v, w)| \leqslant C|||v|||\cdot|||w||| \quad \text{for } j \geqslant 0,$$

and with the appropriate β, fixed in this manner below,

$$B_h(\chi, \chi) \geqslant C|||\chi|||^2 \quad \forall \chi \in S_h.$$

Noting that by Green's formula

$$(f, \chi) = (ATf, \chi) = B_h(Tf, \chi) \quad \text{for } f \in L_2(\Omega),\ \chi \in S_h,$$

we now define T_h: $L_2(\Omega) \to S_h$ by

$$B_h(T_hf, \chi) = (f, \chi) \quad \forall \chi \in S_h,$$

and quote the error estimate

$$(3.28)\qquad \|(T_h - T)f\| + h|||(T_h - T)f||| \leqslant Ch^s\|f\|_{s-2} \quad \text{for } 2 \leqslant s \leqslant r.$$

We now turn to the hypotheses of Section 1. We have here

$$(f, T_h f) = B_h(T_h f, T_h f) \geqslant C|||T_h f|||^2 \geqslant 0$$

with equality only for $T_h f = 0$. If then $f \in S_h$, we conclude

$$\|f\|^2 = B_h(T_h f, f) = 0,$$

which shows that (ia) is satisfied. With $v_h = T_h f$, $w_h = T_h g$ we have

$$(T_h f, g) - (f, T_h g) = B_h(T_h g, T_h f) - B_h(T_h f, T_h g)$$

$$= A(w_h, v_h) - A(v_h, w_h) = -\int_\Omega \sum_j 2b_j \frac{\partial v_h}{\partial x_j} w_h \, dx + \int_{\partial\Omega} \sum_j b_j n_j v_h w_h \, ds. \tag{3.29}$$

Here, using the definition of $|||\cdot|||$ and the inverse estimate (3.27), the last term is bounded by

$$C\|v_h\|_{L_2(\partial\Omega)}\|w_h\|_{L_2(\partial\Omega)} \leqslant C|||v_h||| \cdot \|w_h\| \leqslant C(f, T_h f)^{1/2}\|T_h g\|.$$

Since the first term on the right in (3.29) is similarly bounded, (ib) follows at once.

We now turn to $(\text{ii})_k$. Instead of (3.3) with $q = 1$ we shall show by induction over j the stronger result

$$|||e^{(j)}||| \leqslant Ch^{s-1}\|f\|_{s-2} \quad \text{for } 2 \leqslant s \leqslant r. \tag{3.30}$$

This is known for $j = 0$ by (3.28) and we shall assume now that it holds for $j < l$. We have here

$$\sum_{j=0}^{l} \binom{l}{j} B_h^{(l-j)}(e^{(j)}, \chi) = 0 \quad \forall \chi \in S_h,$$

and hence, for any $\chi \in S_h$,

$$|||\chi|||^2 \leqslant CB_h(\chi, \chi) = C\left\{ B_h(\chi - e^{(l)}, \chi) - \sum_{j=0}^{l-1} \binom{l}{j} B_h^{(l-j)}(e^{(j)}, \chi) \right\}$$

$$\leqslant C\left(|||e^{(l)} - \chi||| + \sum_{j=0}^{l-1} |||e^{(j)}||| \right) |||\chi|||.$$

This yields

$$|||e^{(l)}||| \leqslant |||e^{(l)} - \chi||| + |||\chi||| \leqslant C\left(|||e^{(l)} - \chi||| + \sum_{j=0}^{l-1} |||e^{(j)}||| \right),$$

and thus, by the induction assumption and (3.26),

$$|||e^{(l)}||| \leqslant C\left\{ \inf_{\chi \in S_h} |||T^{(l)}f - \chi||| + h^{s-1}\|f\|_{s-2} \right\} \leqslant Ch^{s-1}\|f\|_{s-2},$$

which shows (3.30). For the induction step in the proof of (3.3) for $q = 0$ we write, with $z = T^*\varphi$,

$$(e^{(l)}, \varphi) = B_h(e^{(l)}, z) = \sum_{j=0}^{l} \binom{l}{j} B_h^{(l-j)}(e^{(j)}, z - \chi) - \sum_{j=0}^{l-1} \binom{l}{j} B_h^{(l-j)}(e^{(j)}, z).$$

Here

$$B_h^{(l-j)}(e^{(j)}, z) = (e^{(j)}, A^{(l-j)*}z),$$

and we conclude, using the induction assumption,

$$|(e^{(l)}, \varphi)| \leqslant C\left\{ \sum_{j=0}^{l} |||e^{(j)}||| \inf_{\chi \in S_h} |||z - \chi||| + \sum_{j=0}^{l-1} \|e^{(j)}\| \, \|z\|_2 \right\}$$
$$\leqslant Ch^s \|f\|_{s-2} \|z\|_2 \leqslant Ch^s \|f\|_{s-2} \|\varphi\|,$$

which completes the proof.

The property (ii)$_1^*$ follows again similarly since

$$B_h(\chi, T_h^* f) = (\chi, f) \quad \forall \chi \in S_h.$$

Department of Mathematics
Jilin University, Changchun
The People's Republic of China

Department of Mathematics
Chalmers University of Technology
S-41296 Göteborg, Sweden

1. I. Babuška, "The finite element method with Lagrangian multipliers," *Numer. Math.*, v. 20, 1973, pp. 179–192.

2. A. Berger, R. Scott & G. Strang, *Approximate Boundary Conditions in the Finite Element Method*, Symposia Mathematica X, Academic Press, New York, 1972, pp. 259–313.

3. J. H. Bramble & J. E. Osborn, "Rate of convergence estimates for nonselfadjoint eigenvalue approximations," *Math. Comp.*, v. 27, 1973, pp. 525–549.

4. J. H. Bramble, A. H. Schatz, V. Thomée & L. B. Wahlbin, "Some convergence estimates for semidiscrete Galerkin type approximations for parabolic equations," *SIAM J. Numer. Anal.*, v. 14, 1977, pp. 218–241.

5. M. Luskin & R. Rannacher, "On the smoothing property of the Galerkin method for parabolic equations." (Preprint.)

6. J. H. Nitsche, "Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen die keinen Randbedingungen unterworfen sind," *Abh. Sem. Univ. Hamburg*, v. 36, 1971, pp. 9–15.

7. P. H. Sammon, *Approximations for Parabolic Equations With Time Dependent Coefficients*, Thesis, Cornell Univ., Ithaca, N. Y., 1978.

8. P. H. Sammon, *Convergence Estimates for Semidiscrete Parabolic Equation Approximations*, MRC Technical Summary Report No. 2053, Univ. of Wisconsin, 1980.

9. R. Scott, "Interpolated boundary conditions in the finite element method," *SIAM J. Numer. Anal.*, v. 12, 1975, pp. 404–427.

10. P. E. Sobolevskiĭ, "Equations of parabolic type in a Banach space," *Trudy Moskov. Mat. Obšč.*, v. 10, 1961, pp. 297–350; English transl., *Amer. Math. Soc. Transl.* (2), v. 49, 1966, pp. 1–62.

→ V. Thomée, "Negative norm estimates and superconvergence in Galerkin methods for parabolic problems," *Math. Comp.*, v. 34, 1980, pp. 93–113.