M2AN. Mathematical modelling and numerical analysis - Modélisation mathématique et analyse numérique

Kenneth Eriksson
Claes Johnson
Vidar Thomée

Time discretization of parabolic problems by the discontinuous Galerkin method

M2AN. Mathematical modelling and numerical analysis - Modélisation mathématique et analyse numérique, tome 19, nº 4 (1985), p. 611-643

<http://www.numdam.org/item?id=M2AN_1985__19_4_611_0>

MATHEMATICAL MODELLING AND NUMERICAL ANALYSIS
MODÉLISATION MATHÉMATIQUE ET ANALYSE NUMÉRIQUE

(vol. 19, n° 4, 1985, p. 611 à 643)

TIME DISCRETIZATION OF PARABOLIC PROBLEMS BY THE DISCONTINUOUS GALERKIN METHOD

by Kenneth ERIKSSON, Claes JOHNSON, Vidar THOMÉE (*)

Abstract. — *This paper analyzes the discontinuous Galerkin method for the time discretization of parabolic type problems in a general setting. Error estimates are derived at the nodal points as well as uniformly in time both for smooth and non-smooth initial data. These estimates are then combined with known estimates for semi-discrete in space Galerkin approximations of parabolic problems to yield error estimates for complete discretizations of such problems.*

Résumé. — *Dans ce travail nous analysons dans un contexte général, la méthode de Galerkin discontinue pour la discrétisation temporelle de problèmes de type parabolique. Des estimations d'erreur sont établies aux points nodaux, uniformément en temps, que la donnée initiale soit régulière ou non. Ces estimations sont ensuite combinées avec les estimations connues pour les approximations de Galerkin semi-discrètes en espace de problèmes paraboliques et cela conduit à des estimations d'erreur pour les discrétisations complètes de tels problèmes.*

Key words : Parabolic problem, Time discretization, Discontinuous Galerkin method, Error estimates.

1. INTRODUCTION

In this paper we shall analyze the discontinuous Galerkin method for the discretization in time t of the parabolic type problem

$$\begin{aligned} y_t + Ay &= f \quad \text{for} \quad t \geqslant 0\,, \\ y(0) &= y_0\,, \end{aligned} \tag{1.1}$$

where y is a function of t with values in a Hilbert space H, y_t denotes the derivative of y, A is a self-adjoint, positive definite, linear operator on H (independent of t), and y_0 and $f = f(t)$ are given data. We shall assume that H is real and separable, and that A is densely defined on H and has a compact inverse A^{-1}.

Manuscrit reçu le 10 septembre 1984.
(*) Department of mathematics, Chalmers university of technology and the university of Göteborg, S-412 96 Göteborg, Sweden.
AMS (1980) Mathematics subject classification. 65M60, 65M15.

As an example of a problem of the given form we shall consider the initial boundary-value problem

$$\begin{aligned} u_t - \Delta u &= f \quad \text{in} \quad \Omega, t \geqslant 0 \\ u &= 0 \quad \text{on} \quad \partial\Omega, t \geqslant 0 \\ u &= v \quad \text{for} \quad t = 0, \end{aligned} \tag{1.2}$$

where Ω is a bounded domain in R^d with smooth boundary $\partial\Omega$, $H = L_2(\Omega)$, and A is the operator $-\Delta$ defined for functions which are twice differentiable in $L_2(\Omega)$ and vanish on $\partial\Omega$. Ultimately, we are interested in a complete discretization of this problem and shall therefore be concerned mainly with the corresponding semi-discrete problem which is also of the form (1.1). In fact, if we discretize (1.2) with respect to the space variables using the standard Galerkin finite element method we are left with the problem to find $u_h(t) \in S_h$ such that

$$\begin{aligned} (u_{h,t}, \chi) + (\nabla u_h, \nabla \chi) &= (f, \chi) \quad \text{for} \quad \chi \in S_h, \quad t \geqslant 0, \\ u_h(0) &= v_h, \end{aligned} \tag{1.3}$$

where S_h is a finite dimensional subspace of $H_0^1(\Omega)$ consisting of piecewise polynomial functions on a partition of Ω into elements of diameter at most h, and where $v_h \in S_h$ is an appropriate approximation of v. Defining $\Delta_h : S_h \to S_h$ by

$$-(\Delta_h \psi, \chi) = (\nabla \psi, \nabla \chi) \quad \text{for} \quad \chi \in S_h,$$

we may write (1.3) as

$$\begin{aligned} u_{h,t} - \Delta_h u_h &= P_0 f \quad \text{for} \quad t \geqslant 0, \\ u_h(0) &= v_h, \end{aligned}$$

where P_0 denotes the L_2-projection onto S_h, and thus as (1.1) with $A = -\Delta_h$ and $H = S_h$, considered with the inner product and norm of $L_2(\Omega)$.

The method we shall analyze for the discretization of (1.1) is the following: Let $\{t_n\}_0^\infty$ be real numbers with $t_0 = 0$ and $t_n < t_{n+1}$ for $n \geqslant 0$, set $I_n = (t_n, t_{n+1}]$, and denote by $\mathscr{S}_n$ the set of polynomials of degree less than q on I_n with values in H. Given y_0 and $f = f(t)$ in H, set $Y_0 = y_0$ and let $Y|_{I_n} \in \mathscr{S}_n$ be recursively determined for $n \geqslant 0$ by

$$\int_{I_n} (Y_t + AY, V)\, ds + (Y_n^+, V_n^+) = (Y_n, V_n^+) + \int_{I_n} (f, V)\, ds \quad \text{for} \quad V \in \mathscr{S}_n, \tag{1.4}$$

where $(.,.)$ is the inner product in H, Y_n^+ is the limit of $Y(s)$ at t_n from above, and $Y_n = Y(t_n)$. Note that no continuity of $Y(s)$ is required at the nodes t_n. As we shall see below, (1.4) admits a unique solution so that the discrete problem is well posed.

In order to describe our results we define a family of norms associated with the operator A, namely

$$\| v \|_s = \| v \|_{s,A} = \left(\sum_{j=1}^{\infty} \lambda_j^{2s} (v, \varphi_j)^2 \right)^{1/2},$$

where $\{ \lambda_j \}_1^\infty$ and $\{ \varphi_n \}_1^\infty$ are the positive eigenvalues and corresponding orthonormal eigenvectors of A, respectively. We note, in particular, that $\| . \|_0 = \| . \| = (.,.)^{1/2}$ is the norm in H. Also, let $k_n = t_{n+1} - t_n$ be the length of the interval I_n and set $k = \max_n k_n$.

We begin our analysis by considering the case of the homogeneous equation. Our first result (Theorem 1) states that for " smooth " initial data y_0 we have

$$\| Y(t) - y(t) \| \leqslant C k^q \| y_0 \|_q \quad \text{for} \quad t \geqslant 0,$$

where C only depends on q. We then derive a more precise estimate for the error at the nodal points, namely (Theorem 2)

$$\| Y_N - y(t_N) \| \leqslant C k^{2q-1} \| y_0 \|_{2q-1} \| y_0 \| \quad \text{for} \quad t_N > 0,$$

where, again, C only depends on q. For the case of " non-smooth " initial data we show (Theorem 3) that at a fixed positive time t the rate of convergence is of the same order as above for arbitrary y_0 in H or, more precisely, that

$$\| Y(t) - y(t) \| \leqslant C k^q t^{-q} \| y_0 \| \quad \text{for} \quad t > 0,$$

and, at the nodal points,

$$\| Y_N - y(t_N) \| \leqslant C k^{2q-1} t^{-(2q-1)} \| y_0 \| \quad \text{for} \quad t_N > 0,$$

where the constants C now also depend on an upper bound for the ratio k_n/k_{n-1}.

For the case of the inhomogeneous equation we show first (Theorem 4) that

$$\| Y(t) - y(t) \| \leqslant C k^q \left(\| y^{(q)}(0) \| + \| f^{(q-1)}(0) \| + \int_0^{t_n} \| f^{(q)} \| \, ds \right),$$

$$\text{for} \quad 0 \leqslant t \leqslant t_n,$$

again with C only depending on q. At the nodal points we find now that

$$\| Y_N - y(t_N) \| \leqslant Ck^{2q-1}\left(\| y^{(q)}(0) \|_{q-1} + \int_0^{t_N} \| f^{(q)} \|_{q-1} \, ds \right) \quad \text{for} \quad t_N \geqslant 0 .$$

As an alternative to this estimate when $q \geqslant 2$, we derive an error estimate of order $0\left(k^{q+1} \log \frac{1}{k}\right)$ at the nodal points where only the norm in H enters in the error bound, namely (Theorem 5)

$$\| Y_N - y(t_N) \| \leqslant Ck^{q+1} \log \frac{1}{k}\left(\| y^{(q+1)}(0) \| + \| f^{(q)}(0) \| + \int_0^{t_N} \| f^{(q+1)} \| \, ds \right) \quad \text{for} \quad t_N \geqslant 0 ,$$

where C now depends on q and on $\max k_{n-1}/k_n$.

Finally, we shall consider in some detail an application of our time discretization scheme (1.4) to the semi-discrete problem (1.3), thus combining the discontinuous Galerkin method in time with the standard Galerkin finite element method in the space variables into a completely discrete scheme for the parabolic problem (1.2). Assuming that the order of accuracy in the discretization in the space variables is $0(h^r)$ we shall then be able to derive various estimates of order $0(k^q + h^r)$ and, at the nodal points t_n, $0(k^{2q-1} + h^r)$ or $0\left(k^{q+1} \log \frac{1}{k} + h^r\right)$ for the fully discrete scheme. For example, in the case $q = 2$ we shall conclude that for a suitable choice of discrete initial data v_h we have for the completely discrete solution U that

$$\| U_N - u(t_N) \| \leqslant Ck^3 \log \frac{1}{k}\left(\| u^{(3)}(0) \| + \| f^{(2)}(0) \| + \int_0^{t_N} \| f^{(3)} \| \, ds \right) + C(u)\, h^r ,$$

where $\| \cdot \|$ is the $L_2(\Omega)$-norm. In these applications it is essential that our previous error estimates hold with constants independent of the specific Hilbert space used.

The virtues of the discontinuous Galerkin method are, in particular, the following : High order almost optimal error estimates can be proved under general hypothesis such as, for instance, variable time steps and variable coefficients. This is accomplished basically by exploiting the fact that the discontinuous Galerkin method admits a variational formulation which makes duality arguments applicable, together with the very good stability properties of the method. Previous error estimates for parabolic problems have generally been restricted to particular methods such as the backward

Euler and Crank-Nicolson methods, under general assumptions on time steps and coefficients, or higher order methods such as subdiagonal Padé methods under assumptions of constant time steps and coefficients, using spectral representations.

The almost optimal error estimates for the discontinuous Galerkin method may be taken as a basis for rational methods for automatic time step control. Further, the discontinuous Galerkin method can be naturally extended to nonlinear parabolic problems. A first step towards an error analysis in this case is taken in Johnson [4] where the discontinuous Galerkin method for stiff ordinary differential equations is analyzed and automatic time step control is discussed.

In the present note we treat for simplicity only the case of constant coefficients; the case of variable coefficients and extensions to nonlinear parabolic problems will be considered in subsequent work.

The discontinuous Galerkin method (1.4) was first analyzed for linear non-stiff ordinary differential equations by Delfour, Hager, and Trochu [2] who proved nodal convergence of order $0(k^{2q-1})$, and for linear parabolic problems by Jamet [3] who proved $0(k^q)$-results. For $q = 1$ (piecewise constants), the scheme (1.4) is the same as the variant of the backward Euler method analyzed earlier by Luskin and Rannacher [5] by techniques similar to those used here. For the homogeneous equation the discontinuous Galerkin method coincides at the nodal points with the subdiagonal Padé scheme of order $(q, q-1)$. For an analysis of such schemes and further references to finite difference methods for (1.1) we refer to Baker, Bramble and Thomée [1].

2. PRELIMINARIES

For non-negative s, let $\dot{H}^s$ be the linear space of all v in H for which the norm $\| v \|_s$ introduced above is finite. Clearly $\dot{H}^s$ forms a Hilbert space with the inner product

$$(v, w)_s = \sum_{j=1}^{\infty} \lambda_j^{2s}(v, \varphi_j)(w, \varphi_j).$$

For negative s, let $\dot{H}^s$ be the dual space of $\dot{H}^{-s}$. Using the notation $(.,.)$ also for the pairing of $\dot{H}^s$ and $\dot{H}^{-s}$ we then have

$$\| v \|_s = \sup_{\substack{w \in \dot{H}^{-s} \\ w \neq 0}} \frac{(v, w)}{\| w \|_{-s}} \quad \text{for} \quad v \in \dot{H}^s,$$

and it is easy to see that $\| v \|_s$ can be represented as in Section 1 in terms of the eigenvalues of A also for $s < 0$. For $s < 0$ the operator A is defined on $\dot{H}^{s+1}$ by duality in the usual manner.

We shall frequently use the following fact :

PROPOSITION : *The operator A is an isometry from $\dot{H}^{s+1}$ onto $\dot{H}^{s}$.*

Proof : Since each v can be represented as $v = \sum_1^\infty (v, \varphi_j)\, \varphi_j$ we have

$$Av = \sum_{j=1}^{\infty} \lambda_j (v, \varphi_j)\, \varphi_j ,$$

and hence

$$\| Av \|_s^2 = \sum_{j=1}^{\infty} \lambda_j^{2s} \lambda_j^2 (v, \varphi_j)^2 = \| v \|_{s+1}^2 ,$$

which is the proposition.

In our analysis we shall consider, in particular, the homogeneous problem

$$y_t + Ay = 0 \quad \text{for} \quad t \geqslant 0, \; y(0) = y_0 . \tag{2.1}$$

We shall then need the following estimates of y in terms of the initial data y_0. Here and below, C denotes a positive constant independent of the particular functions involved and also, which is essential for our later applications, of the specific Hilbert space under consideration.

LEMMA 1 : *Let y be the solution of* (2.1), *let m be nonnegative, and let j be a non-negative integer. Then*

$$t^{2m} \| y^{(j)}(t) \|_l^2 + \int_0^t s^{2m} \| y^{(j)} \|_{l+1/2}^2 \, ds \leqslant C \| y_0 \|_{j+l-m}^2 \quad \textit{for} \quad t > 0 ,$$

where $y^{(j)}$ denotes the jth derivative of y, and $C = C(m)$.

Proof : It suffices to show that for any v in H such that $v_t + Av = 0$ for $t \geqslant 0$ we have

$$t^{2m} \| v(t) \|^2 + \int_0^t s^{2m} \| v \|_{1/2}^2 \, ds \leqslant C \| v(0) \|_{-m}^2 \quad \text{for} \quad t > 0 ,$$

since the general case then follows, in view of the proposition, by taking $v = A^l y^{(j)}$ and using the fact that then $v(0) = (-1)^j A^{l+j} y_0$.

In order to prove the above estimate we note first that the solution of (2.1) can be represented as

$$v(t) = \sum_{j=1}^{\infty} e^{-\lambda_j t} (v(0), \varphi_j)\, \varphi_j \quad \text{for} \quad t \geqslant 0 ,$$

whence

$$\| v(t) \|^2 = \sum_{j=1}^{\infty} e^{-2\lambda_j t}(v(0), \varphi_j)^2 ,$$

and thus

$$t^{2m} \| v(t) \|^2 \leqslant \max_{j \geqslant 1} \sup_{t \geqslant 0} \{ t^{2m} \lambda_j^{2m} e^{-2\lambda_j t} \} \| v(0) \|_{-m}^2 = C(m) \| v(0) \|_{-m}^2 .$$

Similarly

$$\int_0^t s^{2m} \| v(s) \|_{1/2}^2 \, ds \leqslant$$

$$\leqslant \max_{j \geqslant 1} \left\{ \lambda_j^{2m+1} \int_0^t s^{2m} e^{-2\lambda_j s} \, ds \right\} \| v(0) \|_{-m}^2 = C'(m) \| v(0) \|_{-m}^2 .$$

Together these estimates thus complete the proof of the lemma.

Our next result concerns the inhomogeneous equation.

LEMMA 2 : *Let y be the solution of* (1.1) *and let j be a nonnegative integer. Then*

$$\| y^{(j)}(t) \|_l + \left(\int_0^t \| y^{(j)} \|_{l+1/2}^2 \, ds \right)^{1/2} \leqslant C\left(\| y^{(j)}(0) \|_l + \int_0^t \| f^{(j)} \|_l \, ds \right),$$

where C is a constant.

Proof : As before, it suffices to prove the estimate for $j = l = 0$. By (1.1) we have

$$(y_t, y) + (Ay, y) = (f, y) \qquad \text{for} \qquad t \geqslant 0 ,$$

or, equivalently,

$$\frac{1}{2} \frac{d}{dt} \| y \|^2 + \| y \|_{1/2}^2 = (f, y) \qquad \text{for} \qquad t \geqslant 0 .$$

Integrating this identity we obtain by obvious estimates

$$\frac{1}{2} \| y(t') \|^2 + \int_0^{t'} \| y \|_{1/2}^2 \, ds \leqslant \frac{1}{2} \| y(0) \|^2 + \int_0^{t'} \| f \| \, \| y \| \, ds \leqslant$$

$$\leqslant \sup_{0 \leqslant s \leqslant t} \| y(s) \| \left\{ \frac{1}{2} \| y(0) \| + \int_0^t \| f \| \, ds \right\} \quad \text{for} \quad 0 \leqslant t' \leqslant t .$$

With t' such that

$$\| y(t') \| = \sup_{0 \leqslant s \leqslant t} \| y(s) \| ,$$

we first conclude that

$$\sup_{0 \leqslant s \leqslant t} \| y(s) \| \leqslant \| y(0) \| + 2 \int_0^t \| f \| \, ds ,$$

and hence, with $t' = t$, that

$$\frac{1}{2} \| y(t) \|^2 + \int_0^t \| y \|_{1/2}^2 \, ds \leqslant \frac{1}{2} \| y(0) \|^2 + \sup_{0 \leqslant s \leqslant t} \| y(s) \| \int_0^t \| f \| \, ds \leqslant$$

$$\leqslant \| y(0) \|^2 + \frac{5}{2} \left(\int_0^t \| f \| \, ds \right)^2 ,$$

from which the assertion of the lemma then follows.

Remark : If instead we estimate (f, v) above as

$$(f, v) \leqslant \frac{1}{2} \| f \|_{-1/2}^2 + \frac{1}{2} \| v \|_{1/2}^2 ,$$

we find that

$$\| y^{(j)}(t) \|_l^2 + \int_0^t \| y^{(j)} \|_{l+1/2}^2 \, ds \leqslant \| y^{(j)}(0) \|_l^2 + \int_0^t \| f^{(j)} \|_{l-1/2}^2 \, ds ,$$

which is an alternative and sometimes more useful estimate for y.

We have not yet verified that the discrete problem is well-posed, that is, that (1.4) admits a unique solution $Y \in \mathcal{S}_n$ for $n \geqslant 0$. In order to do so, we fix an orthonormal basis $\{ p_j \}_1^q$ for the polynomials of degree less than q on $I_n = (t_n, t_{n+1}]$ and, identifying each $V = \sum_1^q v_j p_j \in \mathcal{S}_n$ with the element $\hat{v} = (v_j)$ in the product Hilbert space $\hat{H} = [H]^q$, write (1.4) equivalently as

$$[\hat{M}\hat{y}, \hat{v}] = [\hat{f}, \hat{v}] \qquad \text{for} \qquad \hat{v} \in \hat{H} ,$$

or

$$\hat{M}\hat{y} = \hat{f},$$

where $[.,.]$ is the inner product in $\hat{H}$, and the matrix $\hat{M} = (m_{ij})$ and $\hat{f} = (f_i)$ are given by

$$m_{ij} = \int_{I_n} p_{j,t} \, p_i \, ds I + \int_{I_n} p_j \, p_i \, ds A + p_j(t_n) \, p_i(t_n) \, I , \quad (I = \text{identity on } H)$$

and

$$f_i = p_i(t_n)\, Y_n + \int_{I_n} f p_i \, ds ,$$

respectively. We shall verify that (1.4) is well-posed by showing that $\hat{M}$, as an operator on the Hilbert space $\hat{H}$, is both one-to-one and onto. Since A is positive, we have for all $\hat{v}$ in the domain of $\hat{M}$ that

$$[\hat{M}\hat{v}, \hat{v}] = \frac{1}{2} \| V_{n+1} \|^2 + \int_{I_n} (AV, V)\, ds + \frac{1}{2} \| V_n \|^2$$

$$\geqslant c \int_{I_n} \| V \|^2 \, ds = c \, |[\, \hat{v} \,]|^2 ,$$

where c is a positive constant, $|[.]|$ denotes the norm in H, and the last step follows by the orthonormality of $\{ p_j \}_1^q$. This shows that $\hat{M}$ is one-to-one, but also that the adjoint $\hat{M}^*$ of $\hat{M}$ is one-to-one. Hence we conclude that the range of $\hat{M}$ is dense in $\hat{H}$, and to complete the argument we just note that the range of $\hat{M}$ is also closed. This follows easily from the two facts that $\hat{M}$ is closed, since A is closed, and the inverse of $\hat{M}$ continuous, which is obvious by above. We have thus shown that if Y_n and $f = f(t)$ belongs to H and $L_1(I_n, H)$, respectively, then (1.4) admits a unique solution $Y \in \mathcal{S}_n$ such that $Y(t) \in \dot{H}^1$ on I_n. By analogous arguments it follows that if $Y_n \in \dot{H}^s$ and $f \in L_1(I_n, \dot{H}^s)$, then $Y(t) \in \dot{H}^{s+1}$ on I_n.

Setting $Y = \sum_0^{q-1} \tilde{Y}_j \, k_n^{-j}(t - t_n)^j$ on I_n we have the following system for the determination of the coefficients $\tilde{Y}_j \in H$, namely, with $\delta_{i,j}$ the Kronecker delta,

$$\int_{I_n} \left(\sum_{j=1}^{q-1} j \tilde{Y}_j k_n^{-j} (s - t_n)^{j-1} + \sum_{j=0}^{q-1} A \tilde{Y}_j k_n^{-j} (s - t_n)^j \right) k_n^{-l} (s - t_n)^l \, ds + \tilde{Y}_0 \, \delta_{l,0} =$$

$$= Y_n \, \delta_{l,0} + \int_{I_n} f(s) \, k_n^{-l} (s - t_n)^l \, ds \quad \text{for} \quad l = 0, 1, ..., q - 1 ,$$

or, after evaluating the integrals on the left,

$$\sum_{j=1}^{q-1} \frac{j}{j + l} \tilde{Y}_j + \sum_{j=0}^{q-1} \frac{1}{j + l + 1} k_n \, A \tilde{Y}_j + \tilde{Y}_0 \, \delta_{l,0} =$$

$$= Y_n \, \delta_{l,0} + k_n^{-l} \int_{I_n} f(s) \, (s - t_n)^l \, ds , \qquad l = 0, 1, ..., q - 1 . \quad (2.2)$$

In particular, for $q = 1$ (piecewise constants in t) we obtain for the determination of $Y(t) = \tilde{Y}_0 = Y_{n+1}$,

$$(I + k_n A) Y_{n+1} = Y_n + \int_{I_n} f \, ds ,$$

which is a version of the backward Euler method (*cf.* [5]) and reduces to the standard such method if the integral is evaluated by the quadrature rule $k_n f(t_{n+1})$. For $q = 2$ (piecewise linears in t) we find

$$\begin{aligned} (I + k_n A) \tilde{Y}_0 + \left(I + \frac{1}{2} k_n A\right) \tilde{Y}_1 &= Y_n + \int_{I_n} f \, ds , \\ \frac{1}{2} k_n A \tilde{Y}_0 + \left(\frac{1}{2} I + \frac{1}{3} k_n A\right) \tilde{Y}_1 &= k_n^{-1} \int_{I_n} f(s) (s - t_n) \, ds , \end{aligned}$$

from which $\tilde{Y}_0$, $\tilde{Y}_1$ and $Y(t)$ are easily determined. Note, in particular, that for the homogeneous equation we then have by a simple calculation

$$\begin{aligned} Y_{n+1} = \tilde{Y}_0 + \tilde{Y}_1 = \left(I + \frac{2}{3} k_n A + \frac{1}{6} k_n^2 A^2\right)^{-1} \left(I - \frac{1}{3} k_n A\right) Y_n = \\ = r_{2,1}(- k_n A) Y_n . \end{aligned}$$

Here we recognize $r_{2,1}(\lambda)$ as the subdiagonal Padé approximant of e^{λ} of order (2, 1).

In general, solving the system (2.2) we find $\tilde{Y}_j$ as a linear combination of rational functions of $k_n A$ acting on the right hand side of (2.2), and, in particular, for the homogeneous equation we have

$$Y_{n+1} = \sum_{j=0}^{q-1} \tilde{Y}_j = r(k_n A) Y_n ,$$

where, by Cramer's rule, $r(\lambda)$ has numerator and denominator of degree at most $q - 1$ and q, respectively.

We conclude this section by introducing some notation which will be useful in our subsequent analysis. Assuming that we are interested in the solution of (1.1) on the interval $(0, t^*)$ and, in particular, at $t = t^*$, we shall use partitions with $t_N = t^*$ and introduce the global bilinear form

$$B(V, W) = \sum_{n=0}^{N-1} \int_{I_n} (V_t + AV, W) \, ds + \sum_{n=1}^{N-1} (V_n^+ - V_n, W_n^+) + (V_0^+, W_0^+) .$$

If we let $\mathcal{S}$ denote the discrete space of piecewise polynomial functions V

such that $V|_{I_n} \in \mathcal{S}_n$ for $n = 0, 1, \ldots, N-1$, we can now write the discrete equations compactly as

$$B(Y, V) = (y_0, V_0^+) + \int_0^{t_N} (f, V)\, ds \quad \text{for} \quad V \in \mathcal{S}\,.$$

Since clearly the solution of the continuous problem (1.1) satisfies

$$B(y, V) = (y_0, V_0^+) + \int_0^{t_N} (f, V)\, ds \quad \text{for} \quad V \in \mathcal{S}\,,$$

we have for the error $Y - y$ that

$$B(Y - y, V) = 0 \quad \text{for} \quad V \in \mathcal{S}\,. \tag{2.3}$$

By integration by parts we see that the bilinear form may also be represented as

$$B(V, W) = \sum_{n=0}^{N-1} \int_{I_n} (V, -W_t + AW)\, ds + \sum_{n=1}^{N-1} (V_n, W_n - W_n^+) + (V_N, W_N)\,. \tag{2.4}$$

In our analysis we shall also consider the " backward " homogeneous problem

$$-z_t + Az = 0 \quad \text{for} \quad t \leqslant t_N\,, \quad z(t_N) = \varphi\,. \tag{2.5}$$

It is clear from the latter representation of $B(.,.)$ that the associated discrete problem, analogous to (1.4), is to find $Z \in \mathcal{S}$ such that

$$B(V, Z) = (V_N, \varphi) \quad \text{for} \quad V \in \mathcal{S}\,,$$

and that results obtained for the forward problem will have counterparts for the backward problem. In particular, the latter problem is well-posed.

3. ENERGY ESTIMATES

Our error analysis below will rely on the following lemma which contains the technical energy estimates needed.

LEMMA 3 : *Let $\rho = \rho(t) : I_n \to H$ be a given function such that ρ_t and $A\rho$*

are in $L_1(I_n, H)$ *and such that* $\rho_{n+1} = \rho(t_{n+1}) = 0$, *let* $\theta_n \in H$, *and assume that* $\theta \in \mathcal{S}_n$ *satisfies*

$$\int_{I_n} (\theta_t + A\theta, V)\, ds + (\theta_n^+, V_n^+) = (\theta_n, V_n^+) + \int_{I_n} (\rho, V_t - AV)\, ds \quad \text{for} \quad V \in \mathcal{S}_n . \tag{3.1}$$

Then for any real l *we have the following estimates, with constants only depending on* q,

$$\| \theta_{n+1} \|_l^2 + \int_{I_n} \| \theta \|_{l+1/2}^2\, ds \leqslant \| \theta_n \|_l^2 + C \int_{I_n} (\| \rho_t \|_{l-1/2}^2 + \| \rho \|_{l+1/2}^2)\, ds , \tag{3.2}$$

$$\| \theta_{n+1} \|_l^2 + \int_{I_n} \| \theta \|_{l+1/2}^2\, ds \leqslant \| \theta_n \|_l^2 + + C \int_{I_n} (k_n \| \rho_t \|_l^2 + k_n \| \rho \|_{l+1}^2 + k_n^{-1} \| \rho \|_l^2)\, ds , \tag{3.3}$$

$$\int_{I_n} \| \theta_t \|_{l-1/2}^2\, ds \leqslant C \int_{I_n} (\| \theta \|_{l+1/2}^2 + \| \rho_t \|_{l-1/2}^2 + \| \rho \|_{l+1/2}^2)\, ds , \tag{3.4}$$

$$\left(\int_{I_n} \| \theta_t \|_l\, ds \right)^2 \leqslant k_n \int_{I_n} \| \theta_t \|_l^2\, ds \leqslant \leqslant C \int_{I_n} (\| \theta \|_{l+1/2}^2 + k_n \| \rho_t \|_l^2 + k_n \| \rho \|_{l+1}^2)\, ds . \tag{3.5}$$

Proof : Again it suffices to consider the case $l = 0$. Setting $V = \theta$ in (3.1) we have

$$\frac{1}{2}(\| \theta_{n+1} \|^2 - \| \theta_n^+ \|^2) + \int_{I_n} \| \theta \|_{1/2}^2\, ds + \| \theta_n^+ \|^2 =$$

$$= (\theta_n, \theta_n^+) + \int_{I_n} (A^{1/2} \rho, A^{-1/2}(\theta_t - A\theta))\, ds$$

$$\leqslant \frac{1}{2} \| \theta_n \|^2 + \frac{1}{2} \| \theta_n^+ \|^2 + \varepsilon \int_{I_n} (\| \theta_t \|_{-1/2}^2 + \| \theta \|_{1/2}^2)\, ds + C_\varepsilon \int_{I_n} \| \rho \|_{1/2}^2\, ds , \tag{3.6}$$

whence, for ε sufficiently small,

$$\| \theta_{n+1} \|^2 + \frac{3}{2} \int_{I_n} \| \theta \|_{1/2}^2 \, ds \leqslant \| \theta_n \|^2 + 2 \varepsilon \int_{I_n} \| \theta_t \|_{-1/2}^2 \, ds + C_\varepsilon \int_{I_n} \| \rho \|_{1/2}^2 \, ds ,$$

from which (3.2) follows once (3.4) is demonstrated. In order to show (3.3) we estimate the integrand in the second line of (3.6) instead as

$$| (\rho, \theta_t - A\theta) | \leqslant \frac{1}{4} \| \theta \|_{1/2}^2 + \| \rho \|_{1/2}^2 + \varepsilon k_n \| \theta_t \|^2 + C_\varepsilon k_n^{-1} \| \rho \|^2 ,$$

to obtain

$$\| \theta_{n+1} \|^2 + \frac{3}{2} \int_{I_n} \| \theta \|_{1/2}^2 \, ds \leqslant \| \theta_n \|^2 + 2 \varepsilon k_n \int_{I_n} \| \theta_t \|^2 \, ds + \\ + C_\varepsilon \int_{I_n} (\| \rho \|_{1/2}^2 + k_n^{-1} \| \rho \|^2) \, ds ,$$

from which (3.3) follows once (3.5) is shown, if we take into account also the obvious inequality

$$\| \rho \|_{1/2}^2 \leqslant \| \rho \| \; \| \rho \|_1 \leqslant \frac{1}{2} k_n^{-1} \| \rho \|^2 + \frac{1}{2} k_n \| \rho \|_1^2 .$$

For the proof of (3.4) we set $V = (t - t_n) A^{-1} \theta_t$ in (3.1) and obtain now, after an integration by parts in the last term, since $\rho_{n+1} = V_n^+ = 0$,

$$\int_{I_n} (s - t_n) \| \theta_t \|_{-1/2}^2 \, ds = - \int_{I_n} (s - t_n) (A\theta, A^{-1} \theta_t) \, ds - \\ - \int_{I_n} (s - t_n) (\rho_t + A\rho, A^{-1} \theta_t) \, ds ,$$

and hence

$$\int_{I_n} (s - t_n) \| \theta_t \|_{-1/2}^2 \, ds \leqslant \frac{1}{2} \int_{I_n} (s - t_n) \| \theta_t \|_{-1/2}^2 \, ds + \int_{I_n} (s - t_n) \| \theta \|_{1/2}^2 \, ds + \\ + 2 \int_{I_n} (s - t_n) (\| \rho_t \|_{-1/2}^2 + \| \rho \|_{1/2}^2) \, ds .$$

Using also the inverse inequality

$$k_n \int_{I_n} \| \theta_t \|_{-1/2}^2 \, ds \leqslant C \int_{I_n} (s - t_n) \| \theta_t \|_{-1/2}^2 \, ds , \tag{3.7}$$

which is valid since $\| \theta_t \|_{-1/2}^2$ is a polynomial of degree at most $2q - 4$, we conclude the proof of (3.4) and hence of (3.2).

Setting instead $V = (t - t_n)\, \theta_t$ in (3.1) we obtain

$$\int_{I_n} (s - t_n) \| \theta_t \|^2 \, ds = - \int_{I_n} (s - t_n) \frac{1}{2} \frac{d}{ds} \| \theta \|_{1/2}^2 \, ds - $$
$$- \int_{I_n} (s - t_n)(\rho_t + A\rho, \theta_t)\, ds \leqslant \frac{1}{2} \int_{I_n} \| \theta \|_{1/2}^2 \, ds$$
$$+ \frac{1}{2} \int_{I_n} (s - t_n) \| \theta_t \|^2 \, ds + \int_{I_n} (s - t_n)(\| \rho_t \|^2 + \| \rho \|_1^2)\, ds,$$

from which (3.5) follows if we again use (3.7). This completes the proof.

We shall begin our application of the energy lemma by showing a stability result for our discrete method for the homogeneous equation, which may be written

$$\int_{I_n} (Y_t + AY, V)\, ds + (Y_n^+, V_n^+) = (Y_n, V_n^+) \quad \text{for} \quad V \in \mathcal{S}_n, \quad n \geqslant 0,$$
$$Y_0 = y_0 . \tag{3.8}$$

LEMMA 4 : *The solution of* (3.8) *satisfies*

$$\| Y(t) \| \leqslant C \| y_0 \| \quad \textit{for} \quad t > 0 .$$

Proof : We have at once by Lemma 3, with $\theta = Y$, $\rho = 0$, and $l = 0$, that

$$\| Y_{n+1} \| \leqslant \| Y_n \| \quad \text{for} \quad n \geqslant 0 ,$$

so that at the nodal points

$$\| Y_n \| \leqslant \| y_0 \| \quad \text{for} \quad n \geqslant 0 .$$

For $t \in I_n$ we have

$$\| Y(t) \| \leqslant \| Y_{n+1} \| + \int_{I_n} \| Y_t \| \, ds ,$$

and thus it only remains to show the desired bound for the latter term. But Lemma 3 again yields

$$\left(\int_{I_n} \| Y_t \| \, ds \right)^2 \leqslant \int_{I_n} \| Y \|_{1/2}^2 \, ds \leqslant C \| Y_n \|^2 \leqslant C \| y_0 \|^2 ,$$

which completes the proof.

Before we begin the error analysis of our method we state a lemma which describes the approximation properties of interpolating polynomials in $\mathcal{S}_n$.

LEMMA 5 : *Let* $\tilde{Y} \in \mathcal{S}_n$ *be the interpolant of the solution* y *of* (2.1) (*or* (1.1)) *defined on* I_n *by* $\tilde{Y}_{n+1} = \tilde{Y}(t_{n+1}) = y(t_{n+1})$, *and, if* $q > 1$,

$$\tilde{Y}\left(t_{n+1} - \frac{m}{q-1}k_n\right) = y\left(t_{n+1} - \frac{m}{q-1}k_n\right) \quad \textit{for} \quad m = 1, \dots, q-1 .$$

Then there is a constant C depending only on q such that for $\rho = \tilde{Y} - y$ *we have*

$$\sup_{I_n} \| \rho \|_l \leqslant C k_n^{j-1} \int_{I_n} \| y^{(j)} \|_l \, ds \quad \textit{for} \quad j = 1, \dots, q ,$$

and, if $q > 1$,

$$\sup_{I_n} \| \rho_t \|_l \leqslant C k_n^{j-2} \int_{I_n} \| y^{(j)} \|_l \, ds \quad \textit{for} \quad j = 2, \dots, q .$$

Proof : These results follow easily with the aid of the Lagrange interpolation formula, after transformation to the unit interval.

4. ERROR ESTIMATES

Below we shall derive various estimates for the error $e = Y - y$. For this purpose we write the error, with $\tilde{Y}$ the interpolant defined in Lemma 5, as

$$e = Y - y = (Y - \tilde{Y}) + (\tilde{Y} - y) = \theta + \rho ,$$

so that ρ satisfies the estimates of Lemma 5, and θ belongs to $\mathcal{S}_n$ for $n \geqslant 0$. Including, for the purpose of later use, the case of the non-homogeneous equation, we find by our definitions that for $V \in \mathcal{S}_n$,

$$\int_{I_n} (\theta_t + A\theta, V)\, ds + (\theta_n^+, V_n^+) = \int_{I_n} (Y_t + AY, V)\, ds + (Y_n^+, V_n^+) -$$

$$- \int_{I_n} (\tilde{Y}_t + A\tilde{Y}, V)\, ds - (\tilde{Y}_n^+, V_n^+)$$

$$= (Y_n, V_n^+) + \int_{I_n} (f, V)\, ds - \int_{I_n} (\tilde{Y}_t + A\tilde{Y}, V)\, ds - (\tilde{Y}_n^+, V_n^+)$$

$$= (Y_n, V_n^+) + \int_{I_n} (y_t + Ay, V)\, ds - \int_{I_n} (\tilde{Y}_t + A\tilde{Y}, V)\, ds - (\tilde{Y}_n^+, V_n^+)$$

$$= (\theta_n, V_n^+) - (\rho_n^+, V_n^+) - \int_{I_n} (\rho_t + A\rho, V)\, ds = (\theta_n, V_n^+) + \int_{I_n} (\rho, V_t - AV)\, ds ,$$

where we have used the fact that $\rho_n = \rho_{n+1} = 0$ by the definition of $\tilde{Y}$. This equation is of the form used in Lemma 3.

We are now ready for our first error estimate in which we consider the homogeneous equation.

THEOREM 1 : *Let y and Y be the solutions of* (2.1) *and* (3.8), *respectively. Then*

$$\| Y(t) - y(t) \| \leqslant Ck^l \| y_0 \|_l \quad \text{for} \quad t \geqslant 0, 0 \leqslant l \leqslant q,$$

where C is a constant only depending on q.

Proof : By Lemmas 1 and 4 the result holds for $l = 0$. We shall show that it holds for $l = q$ which then yields the result for general l by interpolation. In fact, writing $y_0 = y_{0,k} + (y_0 - y_{0,k})$ where $y_{0,k} = \sum_{k\lambda_j \leqslant 1} (y_0, \varphi_j)\, \varphi_j$ we have easily,

$$k^q \| y_{0,k} \|_q = \left(\sum_{k\lambda_j \leqslant 1} (k\lambda_j)^{2q} (y_0, \varphi_j)^2 \right)^{1/2} \leqslant \left(\sum_j (k\lambda_j)^{2l} (y_0, \varphi_j)^2 \right)^{1/2} = k^l \| y_0 \|_l$$

and

$$\| y_0 - y_{0,k} \| = \left(\sum_{k\lambda_j > 1} (y_0, \varphi_j)^2 \right)^{1/2} \leqslant \left(\sum_j (k\lambda_j)^{2l} (y_0, \varphi_j)^2 \right)^{1/2} = k^l \| y_0 \|_l ,$$

and hence, using the two extremal cases, that

$$\| Y(t) - y(t) \| \leqslant Ck^q \| y_{0,k} \|_q + C \| y_0 - y_{0,k} \| \leqslant Ck^l \| y_0 \|_l .$$

With $\tilde{Y}$ and $\rho = \tilde{Y} - y$ as above we have by Lemma 5 for $t \in I_n$,

$$\| \rho(t) \| \leqslant Ck^{q-1} \int_{I_n} \| y^{(q)} \| \, ds \leqslant Ck^q \sup_{I_n} \| y^{(q)} \| \leqslant Ck^q \| y_0 \|_q .$$

It remains to estimate $\theta = Y - \tilde{Y}$ and we begin by doing so at the nodal points. Let thus t_N be a nodal point, let $\varphi \in H$ be arbitrary, and let $Z \in \mathscr{S}$ solve the backward homogeneous problem with $Z_N^+ = \varphi$, so that

$$B(V, Z) = (V_N, \varphi) \quad \text{for} \quad V \in \mathscr{S} .$$

Setting $V = \theta$ we obtain, using (2.3) and (2.4), that

$$(\theta_N, \varphi) = B(\theta, Z) = - B(\rho, Z) = \sum_{n=0}^{N-1} \int_{I_n} (\rho, Z_{t_N} - AZ) \, ds \leqslant$$

$$\leqslant \left(\int_0^{t_N} \| \rho \|_{1/2}^2 \, ds \right)^{1/2} \left(\sum_{n=0}^{N-1} \int_{I_n} (\| Z_t \|_{-1/2}^2 + \| Z \|_{1/2}^2) \, ds \right)^{1/2} . \quad (4.1)$$

By Lemma 3 applied to the backward problem we find easily,

$$\sum_{n=0}^{N-1} \int_{I_n} (\| Z_t \|_{-1/2}^2 + \| Z \|_{1/2}^2)\, ds \leqslant C \| Z_N^+ \|^2 = C \| \varphi \|^2 ,$$

and by Lemmas 5 and 1,

$$\int_0^{t_N} \| \rho \|_{1/2}^2\, ds \leqslant C k^{2q} \int_0^{t_N} \| y^{(q)} \|_{1/2}^2\, ds \leqslant C k^{2q} \| y_0 \|_q^2 ,$$

so that by (4.1),

$$(\theta_N, \varphi) \leqslant C k^q \| y_0 \|_q \| \varphi \| ,$$

and thus

$$\| \theta_N \| \leqslant C k^q \| y_0 \|_q .$$

This also completes the proof if $q = 1$ since θ is then piecewise constant. For $q \geqslant 2$ we have for a general point $t \in I_n$ that

$$\| \theta(t) \| \leqslant \| \theta_{n+1} \| + \int_{I_n} \| \theta_t \|\, ds ,$$

and it remains now to bound the last term. By Lemma 3 we have

$$\left(\int_{I_n} \| \theta_t \|\, ds \right)^2 \leqslant C \int_{I_n} (\| \theta \|_{1/2}^2 + k_n \| \rho_t \|^2 + k_n \| \rho \|_1^2)\, ds$$

$$\leqslant C \| \theta_n \|^2 + C \int_{I_n} (k_n \| \rho_t \|^2 + k_n \| \rho \|_1^2 + k_n^{-1} \| \rho \|^2)\, ds ,$$

and using the already proven estimate for θ_n and Lemmas 5 and 1 we conclude

$$\int_{I_n} \| \theta_t \|\, ds \leqslant C k^q \left\{ \| y_0 \|_q + \sup_{I_n} (\| y^{(q)} \| + \| y^{(q-1)} \|_1) \right\} \leqslant C k^q \| y_0 \|_q ,$$

which completes the proof.

We shall now show that superconvergence occurs at the nodal points and that the order of convergence there is $0(k^{2q-1})$.

THEOREM 2 : *Let $q > 1$, and let y and Y be the solutions of* (2.1) *and* (3.8), *respectively. Then at each nodal point t_N we have*

$$\| Y_N - y(t_N) \| \leqslant C k^l \| y_0 \|_l \quad \textit{for} \quad 0 \leqslant l \leqslant 2q - 1 ,$$

where $C = C(q)$.

Proof : As in the proof of Theorem 1 it suffices to consider the extremal case $l = 2q - 1$, since the general case then follows by the stability and interpolation. As above, let $\varphi \in H$ be arbitrary and let z be the solution of the backward homogeneous problem with $z(t_N) = \varphi$, and let Z be the solution of the corresponding discrete problem. Then with $\tilde{Y}$ as in Lemma 5 and $\rho = \tilde{Y} - y$, $e = Y - y$ and $\eta = Z - z$ we have

$$(e_N, \varphi) = B(e, z) = B(e, z - Z) = B(\tilde{Y} - y, z - Z)$$
$$= \sum_{n=0}^{N-1} \int_{I_n} (\rho, \eta_t - A\eta)\, ds \leqslant C\left(\int_0^{t_N} \| \rho \|_{q-1/2}^2\, ds\right)^{1/2} \times$$
$$\times \left(\sum_{n=0}^{N-1} \int_{I_n} (\| \eta_t \|_{-q+1/2}^2 + \| \eta \|_{-q+3/2}^2)\, ds\right)^{1/2}.$$

Here, by Lemma 5,

$$\int_0^{t_N} \| \rho \|_{q-1/2}^2\, ds \leqslant Ck^{2q} \int_0^{t_N} \| y^{(q)} \|_{q-1/2}^2\, ds \leqslant Ck^{2q} \| y_0 \|_{2q-1}^2 .$$

We shall show now that

$$\sum_{n=0}^{N-1} \int_{I_n} (\| \eta_t \|_{-q+1/2}^2 + \| \eta \|_{-q+3/2}^2)\, ds \leqslant Ck^{2q-2} \| \varphi \|^2 , \qquad (4.2)$$

which would yield

$$(e_N, \varphi) \leqslant Ck^{2q-1} \| y_0 \|_{2q-1} \| \varphi \| ,$$

and thus complete the proof.

In order to show (4.2) we adopt the notation of the corresponding forward problem, so that we want to show now the equivalent assertion

$$\sum_{n=0}^{N-1} \int_{I_n} (\| e_t \|_{-q+1/2}^2 + \| e \|_{-q+3/2}^2)\, ds \leqslant Ck^{2q-2} \| y_0 \|^2 .$$

We have first by Lemma 5,

$$\sum_{n=0}^{N-1} \int_{I_n} (\| \rho_t \|_{-q+1/2}^2 + \| \rho \|_{-q+3/2}^2)\, ds \leqslant$$
$$\leqslant Ck^{2q-2} \int_0^{t_N} (\| y^{(q)} \|_{-q+1/2}^2 + \| y^{(q-1)} \|_{-q+3/2}^2)\, ds$$
$$\leqslant Ck^{2q-2} \int_0^{t_N} \| y \|_{1/2}^2\, ds \leqslant Ck^{2q-2} \| y_0 \|^2 .$$

It remains to show the corresponding estimate for $\theta = Y - \tilde{Y}$. We have at once by summation of (3.2) and then (3.4) of Lemma 3 that, since $\theta_0 = 0$,

$$\sum_{n=0}^{N-1} \int_{I_n} (\| \theta_t \|_{-q+1/2}^2 + \| \theta \|_{-q+3/2}^2) \, ds \leqslant$$

$$\leqslant C \sum_{n=0}^{N-1} \int_{I_n} (\| \rho_t \|_{-q+1/2}^2 + \| \rho \|_{-q+3/2}^2) \, ds \leqslant Ck^{2q-2} \| y_0 \|^2 ,$$

which completes the proof.

Recall that the step from Y_n to Y_{n+1} for the homogeneous problem corresponds to applying the operator $r(k_n A)$ where r is a rational function with numerator and denominator of degree $q - 1$ and q, respectively. Since the only such function of accuracy of order $2q - 1$ is the corresponding Padé approximant we conclude from Theorem 2 that, with a uniform partition in t, the present method reduces at the nodal points to the $(q, q - 1)$-subdiagonal Padé scheme.

We shall now show that non-smooth data estimates are valid for the present situation, that is, for fixed positive time t we have the same order of convergence as above for arbitrary initial data in H.

THEOREM 3 : *Assume that there is a constant γ, independent of the partition, such that $k_n \leqslant \gamma k_{n-1}$ for $n > 0$. Then for y and Y the solutions of* (2.1) *and* (3.8) *we have*

$$\| Y(t) - y(t) \| \leqslant Ck^q \, t^{-q} \| y_0 \| \quad \textit{for} \quad t > 0 \, , \tag{4.3}$$

and, at the nodal points,

$$\| Y_N - y(t_N) \| \leqslant Ck^{2q-1} \, t_N^{-(2q-1)} \| y_0 \| \quad \textit{for} \quad t_N > 0 \, , \tag{4.4}$$

where the constants C only depend on q and γ.

Proof : We shall first show that

$$\| Y(t) - y(t) \| \leqslant Ck^{1/2} \, t^{-1/2} \| y_0 \| \quad \text{for} \quad t > 0 \, , \tag{4.5}$$

and then use an iteration argument to complete the proof.

As usual we write the error as $Y - y = (Y - \tilde{Y}) + (\tilde{Y} - y) = \theta + \rho$ where now $\tilde{Y}$ is the piecewise constant interpolant with $\tilde{Y}_n = y(t_n)$. In order to show (4.5) at the nodal points we need only estimate θ_n. By Lemmas 3 and 5 (with $q = 1$) we have

$$\| \theta_{n+1} \|^2 \leqslant \| \theta_n \|^2 + C \int_{I_n} (k_n \| \rho_t \|^2 + k_n \| \rho \|_1^2 + k_n^{-1} \| \rho \|^2) \, ds$$

$$\leqslant \| \theta_n \|^2 + C \int_{I_n} (k_n \| y_t \|^2 + k_n^3 \| y_t \|_1^2) \, ds ,$$

and, after multiplication by $t_{n+1} = t_n + k_n$,

$$t_{n+1} \| \theta_{n+1} \|^2 \leqslant t_n \| \theta_n \|^2 + k_n \| \theta_n \|^2 + C t_{n+1} \int_{I_n} (k_n \| y \|_1^2 + k_n^3 \| y \|_2^2) \, ds .$$

Hence, for $n \geqslant 1$, using $k_n \leqslant \gamma k_{n-1}$ and an inverse estimate for the second term on the right,

$$t_{n+1} \| \theta_{n+1} \|^2 \leqslant t_n \| \theta_n \|^2 + C \int_{I_{n-1}} \| \theta \|^2 \, ds + C k_n \int_{I_n} (s \| y \|_1^2 + s^3 \| y \|_2^2) \, ds ,$$

and after summation, since $\theta_1 = Y_1 - y(t_1)$,

$$t_N \| \theta_N \|^2 \leqslant k_1 \| \theta_1 \|^2 + C \int_0^{t_N} \| \theta \|^2 \, ds + C k \int_0^{t_N} (s \| y \|_1^2 + s^3 \| y \|_2^2) \, ds$$

$$\leqslant C k \| y_0 \|^2 + C \int_0^{t_N} \| \theta \|^2 \, ds .$$

In order to estimate the latter integral we find by Lemma 3 that

$$\int_0^{t_N} \| \theta \|^2 \, ds \leqslant C \sum_{n=0}^{N-1} \int_{I_n} (k_n \| \rho_t \|_{-1/2}^2 + k_n \| \rho \|_{1/2}^2 + k_n^{-1} \| \rho \|_{-1/2}^2) \, ds .$$

Here

$$\int_0^{t_1} k_0 \| \rho \|_{1/2}^2 \, ds \leqslant 2 k_0 \int_0^{t_1} \| y \|_{1/2}^2 \, ds + 2 k_0^2 \| y(t_1) \|_{1/2}^2 \leqslant C k \| y_0 \|^2 ,$$

and, similarly to above,

$$\sum_{n=1}^{N-1} \int_{I_n} (k_n \| \rho_t \|_{-1/2}^2 + k_n^{-1} \| \rho \|_{-1/2}^2) \, ds + \sum_{n=1}^{N-1} \int_{I_n} k_n \| \rho \|_{1/2}^2 \, ds \leqslant$$

$$\leqslant C k \int_0^{t_N} \| y_t \|_{-1/2}^2 \, ds + C k \int_0^{t_N} s^2 \| y_t \|_{1/2}^2 \, ds \leqslant C k \| y_0 \|^2 ,$$

so that altogether

$$t_N \| \theta_N \|^2 \leqslant Ck \|y_0\|^2 ,$$

which completes the proof of (4.5) at $t = t_N$.

In order to prove (4.5) for general t we first observe that the result follows at once on I_0 by stability since $kt^{-1} \geqslant 1$ there, and note then that for $t \in I_n$ with $n \geqslant 1$ we have since $k_{n-1} t_n^{-1} \leqslant 1$ and $t_n^{-1} \leqslant Ct^{-1}$ that

$$\| \rho(t) \| \leqslant Ck_n \sup_{I_n} \| y_t \| \leqslant Ck_{n-1} t_n^{-1} \| y_0 \| \leqslant Ck^{1/2} t^{-1/2} \| y_0 \| ,$$

and, by obvious similar estimates (with $\theta_t = 0$ if $q = 1$),

$$\begin{aligned} \| \theta(t) \| &\leqslant \| \theta_{n+1} \| + \int_{I_n} \| \theta_t \| \, ds \leqslant \\ &\leqslant C \{ \| \theta_{n+1} \| + \| \theta_n \| + \sup_{I_n} (k_n \| \rho_t \| + k_n \| \rho \|_1 + \| \rho \|) \} \\ &\leqslant Ck^{1/2} t^{-1/2} \| y_0 \| . \end{aligned}$$

This completes the proof of (4.5).

We now turn to the iteration argument for the proof of (4.3). By stability we may restrict the consideration to $t \geqslant ck$ for c an arbitrary fixed positive constant. Let $S(t)$ be the solution operator of the continuous homogeneous problem, and let $S_k(t, t_i)$ be the corresponding discrete solution operator starting at t_i. Set $E_k(t, t_i) = S(t - t_i) - S_k(t, t_i)$ and let t_M be a nodal point such that $| t_M - t/2 | \leqslant k/2$. We have the identity

$$E_k(t, 0) = E_k(t, t_M) S(t_M) + S(t - t_M) E_k(t_M, 0) - E_k(t, t_M) E_k(t_M, 0) .$$

By Theorem 1 we have

$$\| E_k(t, t_i) \varphi \| \leqslant Ck^q \| \varphi \|_q ,$$

and also, since A commutes with $E_k(t, t_i)$,

$$\| E_k(t, t_i) \varphi \|_{-q} \leqslant Ck^q \| \varphi \| .$$

Hence, using Lemma 1, we have, in particular,

$$\| E_k(t, t_M) S(t_M) y_0 \| \leqslant Ck^q \| S(t_M) y_0 \|_q \leqslant Ck^q t_M^{-q} \| y_0 \| \leqslant Ck^q t^{-q} \| y_0 \|$$

and

$$\| S(t - t_M) E_k(t_M, 0) y_0 \| \leqslant C(t - t_M)^{-q} \| E_k(t_M, 0) y_0 \|_{-q} \leqslant Ct^{-q} k^q \| y_0 \| .$$

Finally, by (4.5),

$$\| E_k(t, t_M) E_k(t_M, 0) y_0 \| \leqslant Ck^{1/2}(t - t_M)^{-1/2} \| E_k(t_M, 0) y_0 \| \leqslant$$
$$\leqslant Ck^{1/2} t^{-1/2} \| E_k(t_M, 0) y_0 \| ,$$

and thus altogether,

$$\| E_k(t, 0) y_0 \| \leqslant Ck^q t^{-q} \| y_0 \| + Ck^{1/2} t^{-1/2} \| E_k(t_M, 0) y_0 \| .$$

The first assertion (4.3) of the theorem now follows in the obvious way by repeated application. The proof of (4.4) is analogous.

We shall now turn to the non-homogeneous equation and prove first the following estimates.

THEOREM 4 : *Let y be the solution of* (1.1) *and Y the corresponding approximate solution defined by* (1.4). *Then*

$$\| Y(t) - y(t) \| \leqslant Ck^q \left\{ \| y^{(q)}(0) \| + \| f^{(q-1)}(0) \| + \int_0^{t_{n+1}} \| f^{(q)} \| \, ds \right\}$$
$$\text{for} \quad t \leqslant t_{n+1} ,$$

and, at the nodal points,

$$\| Y_N - y(t_N) \| \leqslant Ck^{2q-1} \left\{ \| y^{(q)}(0) \|_{q-1} + \int_0^{t_N} \| f^{(q)} \|_{q-1} \, ds \right\},$$

with constants only depending on q.

Proof : We write as usual $Y - y = (Y - \tilde{Y}) + (\tilde{Y} - y) = \theta + \rho$ where, by Lemma 5,

$$\| \rho(t) \| \leqslant Ck^q \sup_{I_n} \| y^{(q)} \| \quad \text{for} \quad t \in I_n .$$

Exactly as in the proof of Theorem 1 we find that

$$\| \theta_n \|^2 \leqslant \int_0^{t_n} \| \rho \|_{1/2}^2 \, ds \leqslant Ck^{2q} \int_0^{t_n} \| y^{(q)} \|_{1/2}^2 \, ds ,$$

and for $t \in I_n$, that

$$\| \theta(t) \|^2 \leqslant \left(\| \theta_{n+1} \| + \int_{I_n} \| \theta_t \| \, ds \right)^2$$
$$\leqslant C \left\{ \| \theta_{n+1} \|^2 + \| \theta_n \|^2 + \right.$$
$$\left. + \int_{I_n} (k_n \| \rho_t \|^2 + k_n \| \rho \|_1^2 + k_n^{-1} \| \rho \|^2) \, ds \right\}$$
$$\leqslant C k^{2q} \left\{ \int_0^{t_{n+1}} \| y^{(q)} \|_{1/2}^2 \, ds + \sup_{I_n} (\| y^{(q)} \|^2 + \| y^{(q-1)} \|_1^2) \right\}.$$

Since, using the differential equation,

$$\| y^{(q-1)} \|_1 = \| A y^{(q-1)} \| \leqslant \| y^{(q)} \| + \| f^{(q-1)} \| ,$$

the first estimate of Theorem 4 now follows by Lemma 2 and the obvious fact that

$$\sup_{I_n} \| f^{(q-1)} \| \leqslant \| f^{(q-1)}(0) \| + \int_0^{t_{n+1}} \| f^{(q)} \| \, ds .$$

For the error estimate at the nodal points we recall form the proof of Theorem 2 that

$$\| e_N \|^2 \leqslant C k^{2q-2} \int_0^{t_N} \| \rho \|_{q-1/2}^2 \, ds \leqslant C k^{4q-2} \int_0^{t_N} \| y^{(q)} \|_{q-1/2}^2 \, ds ,$$

by which, again, the conclusion follows by Lemma 2. This completes the proof.

In the second estimate of Theorem 4 we require, in particular, that $f^{(q)}$ is in the domain of definition of A^{q-1} on $(0, t_N)$. In order to apply this result to the case when $A = -\Delta$ with zero boundary conditions imposed, as in our introductory example, we thus have to require not only that f has a certain degree of smoothness but also that $\Delta^j f$ vanishes on the boundary for $0 \leqslant j \leqslant q - 2$ and t positive. In the following theorem we avoid such artificial boundary conditions on f and yet obtain a higher rate of convergence at the nodal points than can be naturally expected. In particular, for $q = 2$ (piecewise linears) we obtain (essentially) the same rate of convergence as in Theorem 4.

THEOREM 5 : *Let y be the solution of* (1.1) *and Y that of* (1.4) *with* $q \geqslant 2$

and $k_{n-1} \leqslant \gamma k_n$ *for* $n \geqslant 1$ *with* γ *independent of the partition. Then*

$$\| Y_N - y(t_N) \| \leqslant C k^{q+1} \log \frac{1}{k} \left\{ \| y^{q+1}(0) \| + \| f^{(q)}(0) \| + \int_0^{t_N} \| f^{(q+1)} \| \, ds \right\},$$

where C *only depends on* q *and* γ.

Proof : With $\varphi \in H$, z the solution of the backward homogeneous equation (2.5) with $z(t_N) = \varphi$, Z the corresponding discrete solution, $\tilde{Y}$ the interpolant of y as in Lemma 5, $e = Y - y$, $\rho = \tilde{Y} - y$ and $\eta = Z - z$ we have (*cf.* the proof of Theorem 2)

$$\begin{aligned}(e_N, \varphi) = \sum_{n=0}^{N-1} \int_{I_n} (\rho, \eta_t - A\eta) \, ds \leqslant \\ &\leqslant \sup_{s \leqslant t_N} \| \rho(s) \|_1 \sum_{n=0}^{N-1} \int_{I_n} (\| \eta_t \|_{-1} + \| \eta \|) \, ds \\ &\leqslant C k^q \sup_{s \leqslant t_N} \| y^{(q)}(s) \|_1 \sum_{n=0}^{N-1} \int_{I_n} (\| \eta_t \|_{-1} + \| \eta \|) \, ds \, .\end{aligned}$$

We shall show below that

$$\sum_{n=0}^{N-1} \int_{I_n} (\| \eta_t \|_{-1} + \| \eta \|) \, ds \leqslant C k \log \frac{1}{k} \| \varphi \| \, , \tag{4.6}$$

from which we conclude that

$$\| e_N \| \leqslant C k^{q+1} \log \frac{1}{k} \sup_{s \leqslant t_N} \| y^{(q)}(s) \|_1 \, . \tag{4.7}$$

Here

$$\| y^{(q)} \|_1 = \| A y^{(q)} \| \leqslant \| y^{(q+1)} \| + \| f^{(q)} \| \, , \tag{4.8}$$

and hence the desired result follows easily in view of Lemma 2.

It remains to demonstrate (4.6), or, equivalently, the corresponding estimate for the homogeneous forward equation. For this we shall show, with y and Y now denoting the solutions of the forward homogeneous problems with initial data y_0 that

$$\int_{I_n} (\| e_t \|_{-1} + \| e \|) \, ds \leqslant C \min (k_n, k k_n t_n^{-1}) \| y_0 \| \quad \text{for} \quad n \geqslant 0 \, . \tag{4.9}$$

Assuming this for a moment, and, using also that $k_n \leqslant \gamma k_{n-1}$ and choosing t_m

such that $k \leqslant t_m \leqslant 2k$, we have

$$\sum_{n=0}^{N-1} \int_{I_n} (\| e_t \|_{-1} + \| e \|) \, ds \leqslant C\left(\sum_{n=0}^{m} k_n + k \sum_{n=m+1}^{N-1} \frac{k_{n-1}}{t_n} \right) \| y_0 \| \leqslant$$

$$\leqslant C\left(t_{m+1} + k \log \frac{t_N}{t_m} \right) \| y_0 \| \leqslant Ck \log \frac{1}{k} \| y_0 \| ,$$

which shows (4.6).

In order to show (4.9) we first note that by Lemma 5 we have for $n \geqslant 1$.

$$\sup_{I_n} (\| \rho_t \|_{-1} + \| \rho \|) \leqslant Ck_n \sup_{I_n} (\| y_{tt} \|_{-1} + \| y_t \|) \leqslant Ck_n t_n^{-1} \| y_0 \| ,$$

and by Theorem 3,

$$\sup_{I_n} \| \theta \| \leqslant \sup_{I_n} (\| e \| + \| \rho \|) \leqslant Ckt_n^{-1} \| y_0 \| ,$$

so that

$$\int_{I_n} (\| \theta \| + \| \rho_t \|_{-1} + \| \rho \|) \, ds \leqslant Ckk_n t_n^{-1} \| y_0 \| .$$

Also, by Lemma 3,

$$\int_{I_n} \| \theta_t \|_{-1} \, ds \leqslant Ck_n^{1/2} \left(\int_{I_n} \| \theta_t \|_{-1}^2 \, ds \right)^{1/2}$$

$$\leqslant Ck_n^{1/2} \left(\int_{I_n} (\| \theta \|^2 + \| \rho_t \|_{-1}^2 + \| \rho \|^2) \, ds \right)^{1/2}$$

$$\leqslant Ck_n \sup_{I_n} (\| \theta \| + \| \rho_t \|_{-1} + \| \rho \|) \leqslant Ckk_n t_n^{-1} \| y_0 \| ,$$

so that altogether

$$\int_{I_n} (\| e_t \|_{-1} + \| e \|) \, ds \leqslant Ckk_n t_n^{-1} \| y_0 \| .$$

By stability we have

$$\int_{I_n} \| e \| \, ds \leqslant Ck_n \sup_{I_n} (\| Y \| + \| y \|) \leqslant Ck_n \| y_0 \| .$$

Also

$$\int_{I_n} \| y_t \|_{-1} \, ds = \int_{I_n} \| y \| \, ds \leqslant Ck_n \| y_0 \| ,$$

and, by Lemma 3,

$$\int_{I_n} \| Y_t \|_{-1} \, ds = Ck_n^{1/2} \left(\int_{I_n} \| Y \|^2 \, ds \right)^{1/2} \leqslant Ck_n \| y_0 \| ,$$

so that

$$\int_{I_n} \| e_t \|_{-1} \leqslant Ck_n \| y_0 \| .$$

Together these estimates prove (4.9) and thus complete the proof of the theorem.

Remark : By (4.7) and (4.8) the estimate of Theorem 5 for $q \geqslant 2$ may alternatively be formulated as

$$\| Y_N - y(t_N) \| \leqslant Ck^{q+1} \log \frac{1}{k} \sup_{s \leqslant t_N} (\| y^{(q+1)}(s) \| + \| f^{(q)}(s) \|) \quad \text{for} \quad t_N \geqslant 0 .$$

Similarly we have for $q = 1$ that

$$\| Y_N - y(t_N) \| \leqslant Ck \log \frac{1}{k} \sup_{s \leqslant t_N} \| y_t(s) \| . \tag{4.10}$$

This estimate follows easily from the representation

$$(e_N, \varphi) = - \sum_{n=0}^{N-1} \int_{I_n} (\rho, AZ) \, ds ,$$

using Lemma 5 and the stability estimate

$$\| AZ(t) \| = \| AZ_{n+1} \| \leqslant \frac{C \| \varphi \|}{t_N - t_n} \quad \text{for} \quad t \in I_n , \quad n < N ,$$

which may be proved by our above methods. Note that the latter estimate is a discrete analogue of the estimate

$$\| z_t(t) \| = \| Az(t) \| \leqslant \frac{C \| \varphi \|}{t_N - t} , \quad \text{for} \quad t < t_N ,$$

for the backward homogeneous continuous problem (2.5). In Johnson [4] an estimate of the form (4.10) is suggested as a basis for rational methods for automatic time step control.

5. APPLICATIONS

In this section we shall apply our discontinuous Galerkin method to the parabolic problem (1.2). In doing so we shall combine our time discretization procedure (1.4) with the standard Galerkin semi-discretization of (1.2) with respect to the space variables to obtain a completely discrete scheme for this problem and give the corresponding error estimates.

Considered as an operator on the Hilbert space $L_2(\Omega)$ with domain of definition $H^2(\Omega) \cap H_0^1(\Omega)$, the operator $-\Delta$ is selfadjoint and positive definite and admits a compact inverse $T = (-\Delta)^{-1}$. We may thus apply our time discretization procedure (1.4) to (1.2) with $H = L_2(\Omega)$ and $A = -\Delta$. It is then natural to adapt the notation to the fact that Δ is of second order in the space variables and, with $\| \cdot \|$ and $(.,.)$ the norm and inner product in $L_2(\Omega)$, and $\{ \lambda_j \}_1^\infty$ and $\{ \varphi_j \}_1^\infty$ the eigenvalues and corresponding orthonormal eigenfunctions of $-\Delta$ with zero boundary values, set

$$\| v \|_s = \left(\sum_{j=1}^{\infty} \lambda_j^s (v, \varphi_j)^2 \right)^{1/2} = \| (-\Delta)^{s/2} v \| ,$$

and correspondingly, for $s \geqslant 0$,

$$\dot{H}^s(\Omega) = \{ v \in L_2(\Omega) : \| v \|_s < \infty \} .$$

For s a non-negative integer $\dot{H}^s(\Omega)$ consists of the functions $v \in H^s(\Omega)$ for which $\Delta^j v = 0$ on $\partial\Omega$ for $j < s/2$ and the norm $\| \cdot \|_s$ is equivalent to the usual Sobolev norm $\| \cdot \|_{H^s(\Omega)}$ on $\dot{H}^s$ (*cf.* e.g. [6] p. 34). For the homogeneous equation, i.e. if $f = 0$, we now have by Theorems 1 and 2, with $\tilde{u}$ the discrete solution, that

$$\| \tilde{u}(t) - u(t) \| \leqslant Ck^q \| v \|_{2q} \quad \text{for} \quad t \geqslant 0 , \quad v \in \dot{H}^{2q}(\Omega) ,$$

and, at the nodal points,

$$\| \tilde{u}_N - u(t_N) \| \leqslant Ck^{2q-1} \| v \|_{4q-2} \quad \text{for} \quad t_N \geqslant 0 , \quad v \in \dot{H}^{4q-2}(\Omega) .$$

On each time interval I_n, however, the discrete solution $\tilde{u}$ is now determined by a system of partial differential equations, which then in practice has to be discretized in the space variables. In our discussion below we shall assume instead that we have first discretized (1.2) with respect to the space variables, and then apply our method of time discretization to obtain a fully discrete scheme for (1.2).

More precisely, let $\{ S_h \}_{0 < h \leqslant 1}$ be a family of finite dimensional spaces contained in $H_0^1(\Omega)$ such that, for some integer $r \geqslant 2$,

$$\inf_{\chi \in S_h} \| \psi - \chi \|_{H^1(\Omega)} \leqslant Ch^{s-1} \| \psi \|_{H^s(\Omega)}$$
$$\text{for} \quad \psi \in H^s(\Omega) \cap H_0^1(\Omega) , \quad 1 \leqslant s \leqslant r ,$$

and consider the semi-discrete problem (1.3) or, equivalently, as described in Section 1,

$$u_{h,t} - \Delta_h u_h = P_0 f \quad \text{for} \quad t \geqslant 0 \tag{5.1}$$
$$u_h(0) = v_h .$$

For this problem we quote the error estimates (*cf.* e.g. [6])

$$\| u_h(t) - u(t) \| \leqslant \| v_h - v \| + Ch^r \left\{ \| v \|_{H^r(\Omega)} + \int_0^t \| u_t \|_{H^r(\Omega)} ds \right\}, \tag{5.2}$$

and, in the case of the homogeneous equation,

$$\| u_h(t) - u(t) \| \leqslant \| v_h - P_0 v \| + Ct^{-j/2} h^r \| v \|_{r-j} \quad \text{for} \quad 0 \leqslant j \leqslant r , \tag{5.3}$$

valid under the appropriate regularity and compatibility assumptions. We now apply the discontinuous Galerkin method (1.4) to (5.1) with $A = -\Delta_h$ and $H = S_h$, considered with the inner product and norm of $L_2(\Omega)$, and thus seek a piecewise polynomial U in time, of degree less than q and with coefficients in S_h, determined by

$$\int_{I_n} (U_t - \Delta_h U, V) \, ds + (U_n^+, V_n^+) = (U_n, V_n^+) + \int_{I_n} (P_0 f, V) \, ds$$
$$\text{for} \quad V \in \mathcal{S}_n , \quad n \geqslant 0 ,$$
$$U_0 = v_h .$$

The theorems of Section 4 now give estimates for $U - u_h$ in terms of k and the data v_h and $P_0 f$ of the semi-discrete problem, with constants independent of h, and with the data measured in norms associated with the operator Δ_h and denoted below by

$$\| \psi \|_{s,h} = \left(\sum_{j=1}^{N(h)} \Lambda_j^s (\psi, \phi_j)^2 \right)^{1/2} .$$

where $\{ \Lambda_j \}_1^{N(h)}$ and $\{ \phi_j \}_1^{N(h)}$ are the eigenvalues and corresponding orthonormal eigenfunctions of $- \Delta_h$, respectively. For example, for $f = 0$ we have by Theorem 1 that

$$\| U(t) - u_h(t) \| \leqslant Ck^l \| v_h \|_{2l,h} = Ck^l \| (- \Delta_h)^l v_h \| \quad \text{for} \quad t \geqslant 0 , \quad 0 \leqslant l \leqslant q , \tag{5.5}$$

and, by Theorem 2, at the nodal points,

$$\| U_N - u_h(t_N) \| \leqslant Ck^l \| (- \Delta_h)^l v_h \| \quad \text{for} \quad t_N \geqslant 0 , \quad 0 \leqslant l \leqslant 2q - 1 . \tag{5.6}$$

Our purpose now, however, is to derive error bounds for $U - u$ in terms of k, h, and the data v and f of the original problem. For the particular choice $v_h = (- \Delta_h)^{-q} (- \Delta)^q v$, for instance, we find at once by (5.5), for the homogeneous equation, that

$$\| U(t) - u_h(t) \| \leqslant Ck^q \| (- \Delta)^q v \| = Ck^q \| v \|_{2q} \quad \text{for} \quad t \geqslant 0 , \quad v \in \dot{H}^{2q}(\Omega) ,$$

and, using also (5.3), it is possible to show that

$$\| U(t) - u(t) \| = 0(h^r + k^q) \quad \text{for} \quad t \geqslant 0 , \quad v \in \dot{H}^{\max(r, 2q)}(\Omega) .$$

We shall now prove such an error estimate for more natural choices of discrete initial data. For our analysis we introduce the discrete solution operator $T_h = (- \Delta_h)^{-1} P_0 : L_2(\Omega) \to S_h$ of the associated elliptic problem, and recall that the standard error estimate for this problem may be expressed as

$$\| (T_h - T) f \| \leqslant Ch^r \| f \|_{H^{r-2}(\Omega)} . \tag{5.7}$$

We then have the following.

THEOREM 6 : *Let u be the solution of* (1.2) *with* $f = 0$, *and let U be the associated completely discrete solution determined by* (5.4) *with v_h chosen so that*

$$\| v_h - v \| \leqslant Ch^r \| v \|_r .$$

Then

$$\| U(t) - u(t) \| \leqslant C \{ k^q \| v \|_{2q} + h^r \| v \|_r \} \quad \textit{for} \quad t \geqslant 0 , \quad v \in \dot{H}^{\max(2q, r)}(\Omega) ,$$

and, at the modal points,

$$\| U_N - u(t_N) \| \leqslant C \{ k^{2q-1} \| v \|_{4q-2} + h^r \| v \|_r \} \quad \textit{for} \quad t_N \geqslant 0 , \quad v \in \dot{H}^{\max(4q-2, r)}(\Omega) .$$

Proof : By the stability of the discrete solution we may assume that $v_h = P_0 v$. In fact, if $\tilde{U}$ denotes the discrete solution corresponding to this choice of initial approximation, we have by Lemma 4,

$$\| U(t) - \tilde{U}(t) \| \leqslant C \| v_h - P_0 v \| \leqslant C(\| v_h - v \| + \| v - P_0 v \|) \leqslant Ch^r \| v \|_r .$$

Given the time discretization parameter k we set $v_k = \sum_{k\lambda_j \leqslant 1} (v, \varphi_j) \cdot \varphi_j$, and find easily that

$$\| v - v_k \| \leqslant k^q \| v \|_{2q} ,$$

$$\| v_k \|_{2q} \leqslant \| v \|_{2q} ,$$

and

$$\| v_k \|_{r+2j} \leqslant k^{-j} \| v \|_r \quad \text{for} \quad 0 \leqslant j \leqslant q - 1 .$$

Let $E(t)$ be the error operator defined by $E(t)\, v = U(t) - u(t)$ where U and u are the solutions of (5.4) and (1.2) with $f = 0$ and $v_h = P_0 v$. By the stability of $E(t)$ we have at once that

$$\| E(t)(v - v_k) \| \leqslant C \| v - v_k \| \leqslant Ck^q \| v \|_{2q} \quad \text{for} \quad t \geqslant 0 .$$

In order to bound $E(t)\, v_k$ we shall use the identity

$$v_k = \sum_{j=0}^{q-1} T_h^j (T - T_h)(-\Delta)^{j+1} v_k + T_h^q (-\Delta)^q v_k .$$

Using (5.5), (5.7), and the properties of v_k, we then have

$$\begin{aligned} \| E(t)\, v_k \| &\leqslant C \sum_{j=0}^{q-1} k^j \| (T - T_h)(-\Delta)^{j+1} v_k \| + Ck^q \| (-\Delta)^q v_k \| \\ &\leqslant Ch^r \sum_{j=0}^{q-1} k^j \| v_k \|_{r+2j} + Ck^q \| v_k \|_{2q} \\ &\leqslant C \{ h^r \| v \|_r + k^q \| v \|_{2q} \} \quad \text{for} \quad t \geqslant 0 . \end{aligned}$$

Together our estimates prove the first part of the theorem. The error bound at the nodal points follows similarly, using (5.6) instead of (5.5).

We shall now consider the case of non-smooth initial data.

THEOREM 7 : *Let u be the solution of (1.2) and U that of (5.4) with $f = 0$ and $v_h = P_0 v$. Then*

$$\| U(t) - u(t) \| \leqslant C \{ t^{-q} k^q + t^{-r/2} h^r \} \| v \| \quad \textit{for} \quad t > 0 , \quad v \in L_2(\Omega) ,$$

and, at the nodal points,

$$\| U_N - u(t_N) \| \leqslant C \{ t_N^{-(2q-1)} k^{2q-1} + t_N^{-r/2} h^r \} \| v \| \quad \textit{for} \quad t_N > 0 , \quad v \in L_2(\Omega) .$$

Proof : The proof follows at once by (5.3) (with $j = r$) and Theorem 3, since clearly $\| P_0 v \| \leqslant \| v \|$.

Finally, we shall briefly consider the case of the non-homogeneous equation. We first prove a uniform error bound in t assuming that the solution is appropriately smooth for $t \geqslant 0$.

THEOREM 8 : *Let u be the solution of* (1.2) *and U that of* (5.4) *with v_h chosen as the " quasi-projection " defined by*

$$v_h = P_0 v + \sum_{j=0}^{q-1} (- T_h)^j P_0(P_1 - I) u^{(j)}(0) .$$

where $P_1 = T_h(- \Delta)$. Then

$$\| U(t) - u(t) \| \leqslant C k^q \left\{ \| u^{(q)}(0) \| + \| f^{(q-1)}(0) \| + \int_0^{t_N} \| f^{(q)} \| \, ds \right\} +$$

$$+ C h^r \left\{ \| v \|_{H^r(\Omega)} + \sum_{j=1}^{q-1} \| u^{(j)}(0) \|_{H^{\max(r-2j,2)}(\Omega)} + \int_0^t \| u_t \|_{H^r(\Omega)} \, ds \right\}$$

for $t \leqslant t_N$.

Proof : The proof follows from Theorem 4 and (5.2), since with our choice of initial data v_h we have (*cf.* [6] pp. 85-86).

$$u_h^{(q)}(0) = P_0 u^{(q)}(0) .$$

and since

$$\| v_h - v \| \leqslant C h^r \left\{ \| v \|_{H^r(\Omega)} + \sum_{j=1}^{q-1} \| u^{(j)}(0) \|_{H^{\max(r-2j,2)}(\Omega)} \right\} . \qquad (5.8)$$

For the proof of (5.8) we note first that

$$\| v_h - v \| \leqslant \| P_0 v - v \| + \| P_0(P_1 - I) v \| +$$

$$+ \sum_{j=1}^{q-1} \| T_h^j P_0(P_1 - I) u^{(j)}(0) \| .$$

Here, by (5.7),

$$\| P_0 v - v \| + \| P_0(P_1 - I) v \| \leqslant 2 \| (P_1 - I) v \| =$$

$$= 2 \| (T_h - T)(- \Delta) v \| \leqslant C h^r \| v \|_{H^r(\Omega)} .$$

To complete the proof we note first that $T_h P_0 = T_h$ so that

$$\| T_h^j P_0(P_1 - I) u^{(j)}(0) \| = \| (P_1 - I) u^{(j)}(0) \|_{-2j,h} \quad \text{for} \quad j > 0,$$

and then, applying the negative norm estimate (*cf.* e.g. [6], Lemmas 2 and 3 of Chapter 6)

$$\| (P_1 - I) \varphi \|_{-p,h} \leqslant Ch^{p+q+2} \| \varphi \|_{H^{q+2}(\Omega)} \quad \text{for} \quad 0 \leqslant p, \quad q \leqslant r - 2,$$

we conclude that

$$\| (P_1 - I) u^{(j)}(0) \|_{-2j,h} \leqslant Ch^r \| u^{(j)}(0) \|_{H^{\max(r-2j,2)}(\Omega)} \quad \text{for} \quad j > 0.$$

Remark : Note that the functions $u^{(j)}(0)$ needed for the determination of the quasi-projection v_h in the theorem may be computed from the data v and f of the original problem by means of the differential equation.

In order to apply the second estimate of Theorem 4 to obtain a more precise error estimate at the nodal points, we would need

$$\int_0^{t_N} \| P_0 f^{(q)} \|_{q-1,h} \, ds$$

to be bounded, independently of h, which is not the case in general unless $\Delta^j f^{(q)}$ vanishes on $\partial\Omega$ for t positive and $0 \leqslant j \leqslant q - 2$. Alternatively, we may apply Theorem 5 for which no such boundary conditions for f are needed. For $q \geqslant 2$ and under the appropriate assumptions we then obtain an error bound of order $0\left(k^{q+1} \log \frac{1}{k} + h^r\right)$ at the nodal points. To illustrate this in more precise terms we consider the case $q = 2$ (piecewise linears in t) for which Theorem 5 yields essentially the same high order rate of convergence as does Theorem 4 for compatible data f.

THEOREM 9 : *Assume that $k_{n-1} \leqslant \gamma k_n$ for $n \geqslant 1$ with γ independent of the partition. Let u be the solution of* (1.2) *and U that of* (5.4), *with $q = 2$ and*

$$v_h = T_h f(0) + T_h^2 \Delta f(0) + T_h^2 \Delta^2 v .$$

Then

$$\| U_N - u(t_N) \| \leqslant Ck^3 \log \frac{1}{k} \left\{ \| u^{(3)}(0) \| + \| f^{(2)}(0) \| + \int_0^{t_N} \| f^{(3)} \| \, ds \right\} +$$

$$+ Ch^r \left\{ \| v \|_{H^r(\Omega)} + \| u_t(0) \|_{H^{\max(r-2,2)}(\Omega)} + \int_0^{t_N} \| u_t \|_{H^r(\Omega)} \, ds \right\} \quad \text{for} \quad t_N \geqslant 0 .$$

Proof : We observe that since $u_t(0) = \Delta v + f(0)$ our choice of discrete initial data is precisely the quasi-projection defined in Theorem 8, with $q = 2$. The result therefore follows from Theorem 5 and (5.2) by the same arguments as above.

REFERENCES

[1] G. A. BAKER, J. H. BRAMBLE and V. THOMÉE, *Single step Galerkin approximations for parabolic problems*. Math. comp. 31, 818-847 (1977).

[2] M. C. DELFOUR, W. W. HAGER and F. TROCHU, *Discontinuous Galerkin methods for ordinary differential equations*. Math. Comp. 36, 455-473 (1981).

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

[4] C. JOHNSON, *On error estimates for numerical methods for stiff o.d.e's*. Preprint, Department of Mathematics, University of Michigan, 1984.

[5] M. LUSKIN and R. RANNACHER, *On the smoothing property of the Galerkin method for parabolic equations*. SIAM J. Numer. Anal. 19, 93-113 (1981).

[6] V. THOMÉE, *Galerkin Methods for Parabolic Problems*, Springer Lecture Notes in Mathematics, No. 1054, 1984.