

GALERKIN TIME-STEPPING METHODS FOR NONLINEAR PARABOLIC EQUATIONS

GEORGIOS AKRIVIS¹ AND CHARALAMBOS MAKRIDAKIS^{2,†}

Abstract. We consider discontinuous as well as continuous Galerkin methods for the time discretization of a class of nonlinear parabolic equations. We show existence and local uniqueness and derive optimal order optimal regularity *a priori* error estimates. We establish the results in an abstract Hilbert space setting and apply them to a quasilinear parabolic equation.

Mathematics Subject Classification. 65M15, 65M50.

Received: December 5, 2002.

Dedicated to professor Michel Crouzeix on the occasion of his 60th birthday, September 30, 2004.

1. INTRODUCTION

The interest for time Galerkin and corresponding space-time finite element methods has been linked during the last decade to the development of adaptivity of mesh selection for evolution PDE's. Certain issues as, *e.g.*, *a posteriori* estimates, estimates of optimal order and regularity, fully discrete schemes with mesh modification, etc., have been extensively considered in the framework of Continuous and Discontinuous Galerkin methods, *cf.*, *e.g.*, [2, 7–16]. This is probably partly due to the fact that in principle space-time Galerkin methods provide freedom for (almost) arbitrary selection of the space time mesh, [11], and partly due merely to the fact that, as in the elliptic case, the properties of variational methods can be studied in an easier, more systematic and clearer way than properties of their pointwise counterparts, *i.e.*, finite difference methods. Still many issues related to the above problems have to be investigated, mainly for nonlinear evolution PDE's. The purpose of this paper is to provide a rather comprehensive *a priori* analysis of a class of variational in time methods, and in particular of the Discontinuous and Continuous Galerkin methods, for nonlinear parabolic equations. It turns out that the limitations of the approach of [7, 8] that was further developed (although from a different perspective) for nonlinear problems in [2] can be overcome by adopting a direct approach based on energy type variational arguments.

Keywords and phrases. Nonlinear parabolic equations, local Lipschitz condition, continuous and discontinuous Galerkin methods, *a priori* error analysis, monotone operators.

¹ Computer Science Department, University of Ioannina, 451 10 Ioannina, Greece. e-mail: akravis@cs.uoi.gr

² Department of Applied Mathematics, University of Crete, 71409 Heraklion-Crete, Greece,

and

Institute of Applied and Computational Mathematics, FORTH, 71110 Heraklion-Crete, Greece. e-mail: makr@tem.uoc.gr

[†] Partially supported by E.U. RT Network Hyke HPRN-CT-2002-00282.

We will discretize initial value problems of the form: find $u : [0, T] \rightarrow D(A)$ satisfying

$$\begin{cases} u'(t) + F(t, u(t)) = 0, & 0 < t < T, \\ u(0) = u^0, \end{cases} \tag{1.1}$$

with $F(t, v) = Av - B(t, v)$, A a positive definite, selfadjoint, linear operator on a Hilbert space $(H, (\cdot, \cdot))$ with domain $D(A)$ dense in H , $B(t, \cdot) : D(A) \rightarrow H, t \in [0, T]$, a (possibly) nonlinear operator, and $u^0 \in H$. Subsequently, we will precisely describe the properties of B and therefore also of F . Essentially the assumption that F admits the above splitting is made for purely technical reasons, in fact our framework, see below, covers a wide class of nonlinear parabolic equations.

The time Galerkin methods. Let $0 = t^0 < t^1 < \dots < t^N = T$ be a partition of $[0, T]$, $I_n := (t^n, t^{n+1}]$, $k_n := t^{n+1} - t^n$, and $q \in \mathbb{N}$. We shall analyze the discretization of (1.1) both by the discontinuous and the continuous Galerkin methods.

To formulate the *discontinuous Galerkin* method, we let \mathcal{V}_q^d be the space of functions that are piecewise polynomials of degree at most $q - 1$ in time in each I_n , with coefficients in $D(A^{1/2})$, i.e.,

$$\mathcal{V}_q^d := \left\{ \varphi : [0, T] \rightarrow D(A^{1/2}) / \varphi|_{I_n}(t) = \sum_{j=0}^{q-1} v_j t^j \right\}.$$

The elements of \mathcal{V}_q^d are allowed to be discontinuous at the nodes t^n , but are taken to be continuous to the left there. The time discrete discontinuous Galerkin approximation U to u is defined as follows: we seek $U \in \mathcal{V}_q^d$ such that

$$\int_{I_n} [(U', v) + (F(t, U), v)] dt + (U^{n+} - U^n, v^{n+}) = 0 \quad \forall v \in \mathcal{V}_q(I_n), \tag{1.2}$$

for $n = 0, \dots, N - 1$. Here $U(0) = u(0)$, $v^n := v(t^n)$, $v^{n+} := \lim_{t \downarrow t^n} v(t)$ and $\mathcal{V}_q(I_n) := \{\varphi|_{I_n} : \varphi \in \mathcal{V}_q^d\}$.

To formulate the *continuous Galerkin* method, we let the space \mathcal{V}_q^c consist of continuous functions that are piecewise polynomials of degree at most $q - 1$ in time in each I_n , with coefficients in $D(A^{1/2})$, i.e.,

$$\mathcal{V}_q^c := \left\{ \varphi \in C([0, T]; D(A^{1/2})) : \varphi|_{I_n}(t) = \sum_{j=0}^{q-1} v_j t^j \right\}.$$

The time discrete continuous Galerkin approximation U to u is defined as follows: we seek $U \in \mathcal{V}_q^c$ such that $U(0) = u(0)$ and

$$\int_{I_n} [(U', v) + (F(t, U), v)] dt = 0 \quad \forall v \in \mathcal{V}_{q-1}(I_n), \tag{1.3}$$

for $n = 0, \dots, N - 1$.

We emphasize that both the discontinuous and continuous Galerkin methods are independent of the particular splitting of F in the form $F(t, v) = Av - B(t, v)$; in applications F is given and the splitting is only used for the analysis of the methods.

The problem framework. We let $V := D(A^{1/2})$ and denote the norms in H and V by $|\cdot|$ and $\|\cdot\|$, $\|v\| := |A^{1/2}v|$, respectively. We assume that $\|\cdot\|$ dominates $|\cdot|$ in V . We identify H with its dual, and let V' be the dual of V , $V \subset H \subset V'$. We still denote by (\cdot, \cdot) the duality pairing between V' and V , and by $\|\cdot\|_*$ the dual norm on V' , $\|v\|_* := |A^{-1/2}v|$.

We assume that $B(t, \cdot)$ can be extended to an operator from V into V' . A natural condition for (1.1) to be locally of parabolic type is the local one-sided Lipschitz condition on $B(t, \cdot)$,

$$(B(t, v) - B(t, w), v - w) \leq \bar{\lambda}\|v - w\|^2 + \bar{\mu}|v - w|^2 \quad \forall v, w \in T_u \tag{1.4}$$

in a tube $T_u, T_u := \{v \in V : \min_t \|u(t) - v\| \leq 1\}$, around the solution u , uniformly in t , with a constant $\bar{\lambda}$ less than one. It is easily seen that (1.4) can be written in the form of a Gårding-type inequality,

$$(F(t, v) - F(t, w), v - w) \geq (1 - \bar{\lambda})\|v - w\|^2 - \bar{\mu}|v - w|^2 \quad \forall v, w \in T_u. \tag{1.4'}$$

In this paper, we shall assume the following, stronger, local Lipschitz condition

$$\|B(t, v) - B(t, w)\|_* \leq \lambda\|v - w\| + \mu|v - w| \quad \forall v, w \in T_u \tag{1.5}$$

with a constant λ less than one and a constant μ . The tube T_u is defined here in terms of the norm of V for concreteness. The analysis may be modified to yield error estimates under conditions analogous to (1.5) for v and w belonging to tubes defined in terms of other norms, not necessarily the same for both arguments. It is actually more natural to have two tubes, because in the error analysis one of the arguments in (1.5) will be a time interpolant of the exact solution (or a time interpolant of an elliptic projection of the exact solution in the fully discrete case) for which estimates in stronger norms might be available, and the other argument will be the approximate solution; it is advantageous to define the second tube in terms of weak norms since this allows the derivation of error estimates under mild mesh conditions, cf. [1, 3, 4], and Section 6.

In particular the above framework covers the following class of quasilinear equations: let $\Omega \subset \mathbb{R}^d, d = 1, 2, 3$, be a bounded interval or a bounded domain with smooth boundary $\partial\Omega$. For $T > 0$ we seek a real-valued function u , defined on $\bar{\Omega} \times [0, T]$, satisfying

$$\begin{cases} u_t = \operatorname{div}(c(x, t, u)\nabla u + g(x, t, u)) + f(x, t, u) & \text{in } \Omega \times [0, T], \\ u = 0 & \text{on } \partial\Omega \times [0, T], \\ u(\cdot, 0) = u^0 & \text{in } \Omega, \end{cases}$$

with $c : \bar{\Omega} \rightarrow (0, \infty), f : \bar{\Omega} \times [0, T] \times \mathbb{R} \rightarrow \mathbb{R}, g : \bar{\Omega} \times [0, T] \times \mathbb{R} \rightarrow \mathbb{R}^d$, and $u^0 : \bar{\Omega} \rightarrow \mathbb{R}$ given smooth functions.

Let $\mathcal{U} := [-1 + \min_{x,t} u, 1 + \max_{x,t} u]$, $c_* > 0$ and c^* be such that

$$c_* \leq c(x, t, y) \leq c^* \quad \forall x \in \bar{\Omega}, t \in [0, T], y \in \mathcal{U}.$$

We set

$$\begin{aligned} a &:= \frac{c_* + c^*}{2}, \quad b(x, t, y) := c(x, t, y) - a, \\ A &:= -a\Delta, \quad B(t, v) := \operatorname{div}(b(\cdot, t, v)\nabla v) + \operatorname{div}g(\cdot, t, y) + f(\cdot, t, y). \end{aligned}$$

Then, $V = H_0^1 = H_0^1(\Omega)$ and the norm $\|\cdot\|$ in $V, \|v\| = \sqrt{a}|\nabla v|$, is equivalent to the H^1 -norm. Let

$$\lambda := \sup\{|b(x, t, y)|/a : x \in \bar{\Omega}, t \in [0, T], y \in \mathcal{U}\};$$

it is shown in Section 6 that $\lambda = 1 - \frac{c_*}{a} < 1$ and that (1.5) is satisfied, in appropriately defined tubes.

Condition (1.5), with appropriately small λ , is used in [4], see also [3] for a similar but more stringent condition, for the analysis of implicit-explicit multistep schemes for (1.1), and in [1] for the analysis of more general linearly implicit methods. In these papers, A and B are discretized in different ways and their knowledge is crucial. In contrast, in this paper both A and B are discretized in the same way; thus for the methods only F matters while for the analysis solely the existence of A and B suffices.

Description of the results. In this paper we present a comprehensive *a priori* analysis of the Discontinuous and Continuous Galerkin methods for the above class of nonlinear parabolic equations. Our approach is related to the one of [12, 13] in the sense that we still use the properties of Radau and Gauss–Legendre quadrature rules that are naturally associated to Discontinuous and Continuous Galerkin methods. On the other hand the

proofs in this paper are based on entirely variational arguments and in particular on novel stability lemmata, cf. Lemma 2.1, Corollary 2.1 and Lemma 5.1. These lemmata allow us to gain the necessary control in $L^2(I_n; H)$ that is missing in order for the energy method to be successfully applied. The lack of control in $L^2(I_n; H)$ is also the reason that the proof of [18] for the linear case is not easily extendable to nonlinear equations. (Note the interesting relation between the test functions we choose in Lem. 2.1, Cor. 2.1 and Lem. 5.1.)

First, we show existence and local uniqueness of the Galerkin approximations. We then derive optimal order *a priori* error estimates. Note that the required regularity of the exact solution is minimal and corresponds to similar estimates in [18]. The analysis is extended also to the case of fully discrete schemes, *i.e.*, the combination of Galerkin time stepping methods with discretization in space. For simplicity we do not consider here schemes combined with mesh modification in space, but our results can be extended to this case by appropriately adopting ideas from [14] to our case. We derive our results in an abstract Hilbert space setting and apply them to a concrete example, namely to a quasilinear parabolic equation.

We consider the continuous and discontinuous Galerkin methods as *base* schemes for the discretization of nonlinear parabolic equations. In many cases, to obtain implementable methods, further discretization of the base schemes, such as linearization and approximation of integrals by quadrature rules, may be required. Some additional difficulties arise in the analysis of the resulting methods; they are not addressed in this paper.

The paper is organized as follows: In Section 2 we show existence and uniqueness of the discontinuous Galerkin approximations for a modified equation with globally Lipschitz continuous nonlinearity. Our proofs are variational and simplify those in [12]. These results combined with the error estimates yield existence and local uniqueness of the discontinuous Galerkin approximations for problem (1.1). Section 3 is devoted to the *a priori* error analysis for the discontinuous Galerkin method. In Section 4 we consider fully discrete schemes, *i.e.*, we combine the discontinuous Galerkin time stepping with space discrete schemes. The continuous Galerkin method is analyzed in Section 5: we prove existence and local uniqueness of continuous Galerkin approximations and derive optimal order error estimates. The results are presented in a more condensed way, avoiding details for arguments already used in the analysis of the discontinuous Galerkin method. In addition we do not include the analysis for fully discrete continuous Galerkin approximations since our results can be extended to fully discrete schemes in a similar fashion as in the discontinuous Galerkin case presented in Section 4. In Section 6 we briefly discuss the application of the abstract results to a quasilinear parabolic equation.

2. THE DG CASE: EXISTENCE AND UNIQUENESS

In this section we show existence and uniqueness of the discontinuous Galerkin approximations for a modified equation. This serves as an intermediate step and will be used in the sequel to establish existence and local uniqueness of the discontinuous Galerkin approximations for our original equation.

We assume that $B(t, \cdot)$ can be modified to yield an operator $\bar{B}(t, \cdot) : V \rightarrow V'$ coinciding with $B(t, \cdot)$ in the tube T_u , $\bar{B}(t, v) = B(t, v)$ for all $t \in [0, T]$ and all $v \in T_u$, and satisfying the global Lipschitz condition, cf. (1.5),

$$\|\bar{B}(t, v) - \bar{B}(t, w)\|_* \leq \lambda \|v - w\| + \mu |v - w| \quad \forall v, w \in V. \tag{2.1}$$

The discontinuous Galerkin method for the modified equation

$$\begin{cases} u'(t) + Au(t) = \bar{B}(t, u(t)), & 0 < t < T, \\ u(0) = u^0, \end{cases} \tag{2.2}$$

is to seek $U \in \mathcal{V}_q^d$ satisfying

$$\int_{I_n} [(U', v) + (AU, v)] dt + (U^{n+} - U^n, v^{n+}) = \int_{I_n} (\bar{B}(t, U), v) dt, \tag{2.3}$$

for all $v \in \mathcal{V}_q(I_n)$, for $n = 0, \dots, N - 1$. Here $U(0) = u(0)$. It is easily seen that the solution u of (1.1) is also a solution of (2.2); further, (2.1) yields easily uniqueness of (smooth) solutions of (2.2).

In this section we show existence and uniqueness of the solution of the scheme (2.3). Later on we will see that $U \in T_u$ and will easily conclude existence and local uniqueness of the solution of scheme (1.2). Existence and uniqueness of discontinuous Galerkin approximations for the nonlinear Schrödinger equation were established in [12]. Our approach is motivated by the one in [12], simplifies it and relies on the following auxiliary result.

Lemma 2.1. *Let $0 < \tau_1 < \dots < \tau_q = 1$ and w_1, \dots, w_q be the abscissae and the weights of the Radau quadrature rule in $[0, 1]$, $p \in \mathbb{P}_{q-1}$ and $\tilde{p} \in \mathbb{P}_{q-1}$ be the interpolant of φ , $\varphi(t) := p(t)/t$, at the Radau points. Then*

$$\int_0^1 p' \tilde{p} dt + p(0)\tilde{p}(0) = \frac{1}{2} \left[|p(1)|^2 + \sum_{i=1}^q w_i |\tau_i^{-1} p(\tau_i)|^2 \right]; \tag{2.4}$$

in particular,

$$\int_0^1 p' \tilde{p} dt + p(0)\tilde{p}(0) \geq \frac{1}{2} \left[|p(1)|^2 + \int_0^1 |p|^2 dt \right]. \tag{2.5}$$

Proof. Let $v \in \mathbb{P}_{q-2}$ be given by $v(t) := \frac{p(t)-p(0)}{t}$, i.e., be such that $p(t) = p(0) + tv(t)$; then, obviously, $\varphi(t) = v(t) + p(0)\frac{1}{t}$. Therefore, polynomial \tilde{p} can be written in the form $\tilde{p} = v + p(0)\Lambda$ with $\Lambda \in \mathbb{P}_{q-1}$ the interpolant of $\frac{1}{t}$ at the Radau points, i.e., $\Lambda(\tau_i) = \frac{1}{\tau_i}$, $i = 1, \dots, q$. To start with, we first note that it is easily seen that

$$\int_0^1 p' \tilde{p} dt = \frac{1}{2} \int_0^1 v^2 dt + \frac{1}{2} |v(1)|^2 + p(0) \left[\int_0^1 tv'(t)\Lambda(t) dt + \int_0^1 v(t)\Lambda(t) dt \right]. \tag{2.6}$$

Now, for $s \in \mathbb{P}_{q-1}$, using the exactness of the Radau quadrature formula, we have

$$\int_0^1 ts'(t)\Lambda(t) dt = \int_0^1 s'(t) dt = s(1) - s(0);$$

in particular

$$\int_0^1 tv'(t)\Lambda(t) dt = v(1) - v(0) \tag{2.7}$$

and

$$\int_0^1 t\Lambda'(t)\Lambda(t) dt = 1 - \Lambda(0). \tag{2.8}$$

Since also $v\Lambda$ is integrated exactly by the Radau quadrature formula, using (2.7) in (2.6), we get

$$\int_0^1 p' \tilde{p} dt = \frac{1}{2} \int_0^1 v^2 dt + \frac{1}{2} |v(1)|^2 + p(0) \left[v(1) - v(0) + \sum_{i=1}^q w_i v(\tau_i)\tau_i^{-1} \right]. \tag{2.9}$$

Since,

$$\begin{aligned} p(0)\tilde{p}(0) &= p(0)v(0) + \Lambda(0)p(0)^2, \\ v(1) &= p(1) - p(0), \end{aligned}$$

and v^2 is integrated exactly by the Radau quadrature formula, relation (2.9) can be written as

$$\int_0^1 p' \tilde{p} dt + p(0)\tilde{p}(0) = \frac{1}{2} |p(1)|^2 + \left(\Lambda(0) - \frac{1}{2} \right) p(0)^2 + \frac{1}{2} \sum_{i=1}^q w_i [v(\tau_i)^2 + 2v(\tau_i)\tau_i^{-1} p(0)]. \tag{2.10}$$

Now, in view of (2.8),

$$\Lambda(0) = 1 - \frac{1}{2} \int_0^1 t(\Lambda^2)'(t) dt = \frac{1}{2} + \frac{1}{2} \int_0^1 (\Lambda(t))^2 dt,$$

i.e.,

$$\Lambda(0) - \frac{1}{2} = \frac{1}{2} \sum_{i=1}^q w_i \tau_i^{-2}. \tag{2.11}$$

Using (2.11) in (2.10) and then the fact that

$$(p(0))^2 \tau_i^{-2} + (v(\tau_i))^2 + 2p(0)v(\tau_i)\tau_i^{-1} = |\tau_i^{-1}p(\tau_i)|^2,$$

we obtain (2.4). Further,

$$\int_0^1 p' \tilde{p} dt + p(0)\tilde{p}(0) \geq \frac{1}{2} \left[|p(1)|^2 + \sum_{i=1}^q w_i |p(\tau_i)|^2 \right] = \frac{1}{2} \left[|p(1)|^2 + \int_0^1 |p(t)|^2 dt \right]. \quad \square$$

A second proof to Lemma 2.1 is given in the Appendix. A change of variables shows that the results of Lemma 2.1 transform to the interval $[t^n, t^{n+1}]$ as follows:

Corollary 2.1. *Let $p \in \mathbb{P}_{q-1}$ be a polynomial in $[t^n, t^{n+1}]$, $\varphi(t) := k_n p(t)/(t - t^n)$, and $\tilde{p} \in \mathbb{P}_{q-1}$ be the interpolant of φ at the Radau points $t^{n,i}, i = 1, \dots, q$, shifted to $[t^n, t^{n+1}]$, $t^{n,i} := t^n + k_n \tau_i$. Then*

$$\int_{t^n}^{t^{n+1}} p' \tilde{p} dt + p(t^n)\tilde{p}(t^n) = \frac{1}{2} \left[|p(t^{n+1})|^2 + \sum_{i=1}^q w_i |\tau_i^{-1} p(t^{n,i})|^2 \right] \tag{2.12}$$

and

$$\int_{t^n}^{t^{n+1}} p' \tilde{p} dt + p(t^n)\tilde{p}(t^n) \geq \frac{1}{2} \left[|p(t^{n+1})|^2 + \frac{1}{k_n} \int_{t^n}^{t^{n+1}} |p|^2 dt \right]. \tag{2.13}$$

With $\|\cdot\|$ denoting either one of the norms $|\cdot|$, $\|\cdot\|$ or $\|\cdot\|_*$, we introduce in $\mathcal{V}_q(I_n)$ the norms $|\cdot|_n$, $\|\cdot\|_n$ and $\|\cdot\|_{*n}$ by

$$\|v\|_n := \left[k_n \sum_{i=1}^q w_i \tau_i^{-1} \|v(t^{n,i})\|^2 \right]^{1/2}.$$

Let $U, V \in \mathcal{V}_q(I_n)$, $W := U - V$, and $\tilde{W} \in \mathcal{V}_q(I_n)$ be the interpolant of $k_n W(t)/(t - t^n)$ at the Radau points $t^{n,i}, i = 1, \dots, q$. Then, we have, cf. (2.12),

$$\int_{I_n} (W', \tilde{W}) dt + (W^{n+}, \tilde{W}^{n+}) \geq \frac{1}{2} \left[|W^{n+1}|^2 + \frac{1}{k_n} |W|_n^2 \right]. \tag{2.14i}$$

Further, since the Radau quadrature rule integrates polynomials of degree at most $2q - 2$ exactly,

$$\int_{I_n} (AW, \tilde{W}) dt = k_n \sum_{i=1}^q w_i (AW(t^{n,i}), \tilde{W}(t^{n,i})) = k_n \sum_{i=1}^q w_i \tau_i^{-1} (AW(t^{n,i}), W(t^{n,i})),$$

i.e.,

$$\int_{I_n} (AW, \tilde{W}) dt = \|W\|_n^2. \tag{2.14ii}$$

Moreover, using first (2.1) we get

$$(\bar{B}(t, U) - \bar{B}(t, V), \tilde{W}) \leq (\lambda \|W\| + \mu |W|) \|\tilde{W}\|$$

and using then the fact that $(\lambda\|W\| + \mu|W|)\|\tilde{W}\|$ is integrated exactly by the Radau quadrature rule we obtain

$$\begin{aligned} \int_{I_n} (\bar{B}(t, U) - \bar{B}(t, V), \tilde{W}) dt &\leq k_n \sum_{i=1}^q w_i \left[\lambda\|W(t^{n,i})\| + \mu|W(t^{n,i})| \right] \tau_i^{-1} \|W(t^{n,i})\| \\ &= \lambda\|W\|_n^2 + \mu k_n \sum_{i=1}^q w_i \tau_i^{-1} |W(t^{n,i})| \|W(t^{n,i})\|; \end{aligned}$$

therefore,

$$\int_{I_n} (\bar{B}(t, U) - \bar{B}(t, V), \tilde{W}) dt \leq (\lambda + \varepsilon)\|W\|_n^2 + \frac{\mu^2}{4\varepsilon}\|W\|_n^2 \tag{2.14iii}$$

for any positive ε .

The discontinuous Galerkin approximate solution $U \in \mathcal{V}_q^d$ is defined in I_n by its value U^n at t^n , which has been determined from the conditions in the preceding time interval I_{n-1} , and by (2.3).

We introduce in $\mathcal{V}_q(I_n)$ the inner product $\langle \cdot, \cdot \rangle$ by

$$\langle v, w \rangle := \int_{I_n} (v, \tilde{w}) dt$$

with $\tilde{w} \in \mathcal{V}_q(I_n)$ denoting the interpolant of $k_n w(t)/(t - t^n)$ at the Radau points $t^{n,i}, i = 1, \dots, q$. It is readily seen that $\langle \cdot, \cdot \rangle$ can also be written in the form

$$\langle v, w \rangle = k_n \sum_{i=1}^q w_i \tau_i^{-1} (v(t^{n,i}), w(t^{n,i})).$$

We now define a map $\mathcal{G} : \mathcal{V}_q(I_n) \rightarrow \mathcal{V}_q(I_n)$ by

$$\langle \mathcal{G}(v), w \rangle = \int_{I_n} [(v', \tilde{w}) + (Av, \tilde{w}) - (\bar{B}(t, v), \tilde{w})] dt + (v^{n+} - U^n, \tilde{w}^{n+}) \tag{2.15}$$

for all $w \in \mathcal{V}_q(I_n)$. It is easily seen that \mathcal{G} is well defined. We establish existence and uniqueness of the discontinuous Galerkin approximation in I_n , *i.e.*, existence and uniqueness of $U \in \mathcal{V}_q(I_n)$ such that $\mathcal{G}(U) = 0$, by showing that \mathcal{G} is Lipschitz continuous and strongly monotone.

2.1. Strong monotonicity of \mathcal{G}

Let $v, w \in \mathcal{V}_q(I_n)$, $\vartheta := v - w$, and, as usual, $\tilde{\vartheta} \in \mathcal{V}_q(I_n)$ be the interpolant of $k_n \vartheta(t)/(t - t^n)$ at the Radau points $t^{n,i}, i = 1, \dots, q$. Then

$$\langle \mathcal{G}(v) - \mathcal{G}(w), v - w \rangle = \langle \vartheta', \vartheta \rangle + \langle A\vartheta, \vartheta \rangle - \langle \bar{B}(t, v) - \bar{B}(t, w), \vartheta \rangle + (\vartheta^{n+}, \tilde{\vartheta}^{n+})$$

and, in view of (2.14), we obtain

$$\begin{aligned} \langle \mathcal{G}(v) - \mathcal{G}(w), v - w \rangle &\geq \frac{1}{2} \left(|\vartheta^{n+1}|^2 + \frac{1}{k_n} |\vartheta|_n^2 \right) + (1 - \lambda - \varepsilon)\|\vartheta\|_n^2 - \frac{\mu^2}{4\varepsilon} |\vartheta|_n^2 \\ &\geq \frac{1}{k_n} \left(1 - \frac{\mu^2}{4\varepsilon} k_n \right) |\vartheta|_n^2 + (1 - \lambda - \varepsilon)\|\vartheta\|_n^2. \end{aligned}$$

Therefore, for $\varepsilon = (1 - \lambda)/2$ and $k_n < 2(1 - \lambda)/\mu^2$, we have the following strong monotonicity property of \mathcal{G}

$$\langle \mathcal{G}(v) - \mathcal{G}(w), v - w \rangle \geq \frac{1 - \lambda}{2} \|v - w\|_n^2 \quad \forall v, w \in \mathcal{V}_q(I_n). \tag{2.16}$$

In particular, (2.16) yields immediately uniqueness of the discontinuous Galerkin approximation, for sufficiently small $k, k := \max_n k_n$.

2.2. Lipschitz continuity of \mathcal{G}

Let $v \in \mathcal{V}_q(I_n)$ and $\tilde{v} \in \mathcal{V}_q(I_n)$ be the interpolant of $k_n v(t)/(t - t^n)$ at the Radau points $t^{n,i}, i = 1, \dots, q$. Since $|\tilde{v}|^2$ is integrated exactly by the Radau quadrature rule we have

$$\int_{I_n} |\tilde{v}|^2 dt = k_n \sum_{i=1}^q w_i \tau_i^{-2} |v(t^{n,i})|^2$$

and thus easily

$$\int_{I_n} |\tilde{v}|^2 dt \leq \tau_1^{-1} |v|_n^2. \tag{2.17i}$$

Similarly,

$$\int_{I_n} \|\tilde{v}\|^2 dt \leq \tau_1^{-1} \|v\|_n^2. \tag{2.17ii}$$

Further, we have

$$\tilde{v}^{n+} = \sum_{i=1}^q w_i \tau_i^{-1} v(t^{n,i}) \prod_{\substack{j=1 \\ j \neq i}}^q \frac{-\tau_j}{\tau_i - \tau_j},$$

and hence

$$|\tilde{v}^{n+}|^2 \leq \frac{c_1}{k_n} |v|_n^2. \tag{2.18i}$$

In the sequel, we will also use the following well-known inverse inequalities

$$|v^{n+}|^2 \leq \frac{c_2}{k_n} |v|_n^2 \tag{2.18ii}$$

and

$$|v'|_n \leq \frac{c_3}{k_n} |v|_n; \tag{2.18iii}$$

the latter can also be written in the form

$$\left(\int_{I_n} |v'(t)|^2 dt \right)^{1/2} \leq \frac{\bar{c}_3}{k_n} \left(\int_{I_n} |v(t)|^2 dt \right)^{1/2}. \tag{2.18'iii}$$

Let now $v, w, \omega \in \mathcal{V}_q(I_n)$ and let $\tilde{\omega} \in \mathcal{V}_q(I_n)$ be defined in the usual way. Then, with $\vartheta := v - w$,

$$\langle \mathcal{G}(v) - \mathcal{G}(w), \omega \rangle = \langle \vartheta', \omega \rangle + (\vartheta^{n+}, \tilde{\omega}^{n+}) + \int_{I_n} (A\vartheta, \tilde{\omega}) dt - \int_{I_n} (\bar{B}(t, v) - \bar{B}(t, w), \tilde{\omega}) dt$$

and, in view of (2.17), (2.18) and (2.1) we easily obtain

$$\langle \mathcal{G}(v) - \mathcal{G}(w), \omega \rangle \leq C k_n^{-1} |v - w|_n |\omega|_n + C (\|v - w\|_n + |v - w|_n) \|\omega\|_n. \tag{2.19}$$

For fixed, sufficiently small k_n , we conclude from (2.16) and (2.19) that

$$\langle \mathcal{G}(v) - \mathcal{G}(w), v - w \rangle \geq c \|v - w\|_n^2 \quad \forall v, w \in \mathcal{V}_q(I_n) \tag{2.16'}$$

and

$$\|\mathcal{G}(v) - \mathcal{G}(w)\|_{\star n} \leq L \|v - w\|_n \quad \forall v, w \in \mathcal{V}_q(I_n). \tag{2.19'}$$

In analogy to $\mathcal{V}_q(I_n)$, we define $\mathcal{H}_q(I_n)$ by

$$\mathcal{H}_q(I_n) := \left\{ \varphi : I_n \rightarrow H / \varphi|_{I_n}(t) = \sum_{j=0}^{q-1} v_j t^j, \quad v_j \in H \right\}.$$

Obviously, $(\mathcal{H}_q(I_n), \langle \cdot, \cdot \rangle)$ is a Hilbert space. It easily follows from (2.16') and (2.19') that

$$\bar{\mathcal{G}} : \mathcal{H}_q(I_n) \rightarrow \mathcal{H}_q(I_n), \quad \bar{\mathcal{G}}(v) = A^{-1/2} \mathcal{G}(A^{-1/2} v),$$

is strongly monotone and Lipschitz continuous in $(\mathcal{H}_q(I_n), \langle \cdot, \cdot \rangle)$,

$$\langle \bar{\mathcal{G}}(v) - \bar{\mathcal{G}}(w), v - w \rangle \geq c|v - w|_n^2 \quad \forall v, w \in \mathcal{H}_q(I_n)$$

and

$$|\bar{\mathcal{G}}(v) - \bar{\mathcal{G}}(w)|_n \leq L|v - w|_n \quad \forall v, w \in \mathcal{H}_q(I_n).$$

Then, it readily follows that $\mathcal{F}, \mathcal{F}(v) := v - \frac{c}{L^2} \bar{\mathcal{G}}(v)$, is a contraction in $(\mathcal{H}_q(I_n), |\cdot|_n)$; this fact is known as Zarantonello's fixed point theorem. If $v \in \mathcal{H}_q(I_n)$ is the unique fixed point of \mathcal{F} , then $A^{-1/2} v \in \mathcal{V}_q(I_n)$ is the unique solution of $\mathcal{G}(v) = 0$.

3. THE DG CASE: ERROR ESTIMATES

In this section we establish optimal order estimates for $u - U$, U being the solution of (2.3), under the assumption that $\bar{B}(t, \cdot) : V \rightarrow V'$ coincides with $B(t, \cdot)$ in the tube T_u , $\bar{B}(t, v) = B(t, v)$ for all $t \in [0, T]$ and all $v \in T_u$, and satisfies the global Lipschitz condition (2.1). After having established the error estimate we can show that for sufficiently small time steps the solution of (2.3) satisfies also (1.2), and, thus, we will have error estimates for the original equation.

Let $W \in \mathcal{V}_q^d$ be defined by $W^0 = u^0$ and

$$\begin{aligned} W^{n+1} &= u^{n+1}, \\ \int_{I_n} (u - W, v) dt &= 0 \quad \forall v \in \mathcal{V}_{q-1}(I_n), \end{aligned} \tag{3.1}$$

$n = 0, \dots, N - 1$, cf. [18], p. 185. The function W will play an important role in the error analysis in the sequel. It is well known that W is well defined by (3.1) and satisfies the error estimates

$$\|(u - W)(t)\|^2 \leq Ck_n^{2q-1} \int_{I_n} \|u^{(q)}(s)\|^2 ds, \quad t \in I_n, \tag{3.2}$$

with $\|\cdot\|$ standing for either one of the norms $|\cdot|$, $\|\cdot\|$ and $\|\cdot\|_*$, cf. [18]. From (3.1) we easily obtain, for $v \in \mathcal{V}_q(I_n)$,

$$\int_{I_n} (u' - W', v) dt = - \int_{I_n} (u - W, v') dt - (u^n - W^{n+}, v^{n+}),$$

i.e.,

$$\int_{I_n} (W', v) dt + (W^{n+} - W^n, v^{n+}) = \int_{I_n} (u', v) dt \quad \forall v \in \mathcal{V}_q(I_n). \tag{3.3}$$

With $\rho := u - W$ and $\Theta := W - U$, we split the error $u - U$ in the form

$$u - U = \rho + \Theta.$$

Since ρ has been estimated in (3.2), our main goal in this section will be the estimation of Θ . We will achieve this in two stages: first we will show consistency of the numerical scheme for W and subsequently we will show stability.

3.1. Consistency

Let $R \in \mathcal{V}_q(I_n)$ denote the consistency error of the discontinuous Galerkin method for W ,

$$\int_{I_n} (R(t), v) dt = \int_{I_n} (W' + AW - \bar{B}(t, W), v) dt + (W^{n+} - W^n, v^{n+}) \quad (3.4)$$

for all $v \in \mathcal{V}_q(I_n)$. Then, in view of (3.3) and (1.1),

$$\int_{I_n} (R(t), v) dt = - \int_{I_n} (A\rho, v) dt + \int_{I_n} (\bar{B}(t, u) - \bar{B}(t, W), v) dt. \quad (3.5)$$

Letting $v := A^{-1}R$ in (3.5) and using (2.1), we have

$$\int_{I_n} \|R(t)\|_{\star}^2 dt \leq - \int_{I_n} (A^{1/2}\rho, A^{-1/2}R) dt + \int_{I_n} [\lambda\|\rho\| + \mu|\rho|] \|R\|_{\star} dt,$$

and thus easily

$$\int_{I_n} \|R(t)\|_{\star}^2 dt \leq 2 \int_{I_n} \|\rho(t)\|^2 dt + 2 \int_{I_n} [\lambda|\rho(t)| + \mu|\rho(t)|]^2 dt. \quad (3.6)$$

From (3.2) and (3.6) we obtain the consistency estimate

$$\int_{I_n} \|R(t)\|_{\star}^2 dt \leq Ck_n^{2q} \int_{I_n} \|u^{(q)}(t)\|^2 dt. \quad (3.7)$$

3.2. Stability

From (2.3) and (3.4) we obtain an error equation for Θ ,

$$\int_{I_n} (\Theta', v) dt + (\Theta^{n+}, v^{n+}) + \int_{I_n} (A\Theta, v) dt = (\Theta^n, v^{n+}) + \int_{I_n} (\bar{B}(t, W) - \bar{B}(t, U), v) dt + \int_{I_n} (R(t), v) dt \quad (3.8)$$

for all $v \in \mathcal{V}_q(I_n)$. First, letting $v := 2\Theta$ in (3.8) we obtain

$$|\Theta^{n+1}|^2 + |\Theta^{n+}|^2 - 2(\Theta^n, \Theta^{n+}) + 2 \int_{I_n} \|\Theta\|^2 dt = 2 \int_{I_n} (\bar{B}(t, W) - \bar{B}(t, U), \Theta) dt + 2 \int_{I_n} (R(t), \Theta) dt. \quad (3.9)$$

Now,

$$|\Theta^{n+}|^2 - 2(\Theta^n, \Theta^{n+}) \geq -|\Theta^n|^2 \quad (3.10i)$$

and

$$2 \int_{I_n} (R(t), \Theta) dt \leq \frac{1}{\varepsilon} \int_{I_n} \|R(t)\|_{\star}^2 dt + \varepsilon \int_{I_n} \|\Theta\|^2 dt; \quad (3.10ii)$$

further, in view of (2.1),

$$2 \int_{I_n} (\bar{B}(t, W) - \bar{B}(t, U), \Theta) dt \leq (2\lambda + \varepsilon) \int_{I_n} \|\Theta\|^2 dt + \frac{\mu^2}{\varepsilon} \int_{I_n} |\Theta|^2 dt. \quad (3.10iii)$$

Using (3.10) in (3.9), we get

$$|\Theta^{n+1}|^2 + \alpha \int_{I_n} \|\Theta(t)\|^2 dt \leq |\Theta^n|^2 + \frac{\mu^2}{\varepsilon} \int_{I_n} |\Theta(t)|^2 dt + \frac{1}{\varepsilon} \int_{I_n} \|R(t)\|_*^2 dt,$$

with $\alpha := 2(1 - \lambda - \varepsilon)$, for any positive ε . Thus, choosing $\varepsilon < 1 - \lambda$, we have

$$|\Theta^{n+1}|^2 + c \int_{I_n} \|\Theta(t)\|^2 dt \leq |\Theta^n|^2 + \bar{c} \int_{I_n} |\Theta(t)|^2 dt + \bar{c} \int_{I_n} \|R(t)\|_*^2 dt. \tag{3.11}$$

The standard energy proof can not be completed since we do not have control of the term $\int_{I_n} |\Theta(t)|^2 dt$. To overcome this barrier we will make use of the stability Lemma 2.1 and the corresponding Corollary 2.1. To this end, letting in (3.8) $v := \tilde{\Theta}$, the interpolant of $k_n \Theta(t)/(t - t^n)$ at the Radau points $t^{n,i}, i = 1, \dots, q$, we obtain, cf. (2.14) and the derivation of (2.16),

$$\frac{1}{2} \left[|\Theta^{n+1}|^2 + \frac{1}{k_n} |\Theta_n^2| \right] + (1 - \lambda - \varepsilon) \|\Theta\|_n^2 \leq \frac{\mu^2}{4\varepsilon} |\Theta_n^2| + (\Theta^n, \tilde{\Theta}^{n+}) + \int_{I_n} (R(t), \tilde{\Theta}) dt, \tag{3.12}$$

for any positive ε . Now,

$$(R(t), \tilde{\Theta}) \leq \frac{1}{4\tau_1\varepsilon} \|R(t)\|_*^2 + \varepsilon\tau_1 \|\tilde{\Theta}\|^2,$$

and thus

$$\int_{I_n} (R(t), \tilde{\Theta}) dt \leq \frac{1}{4\tau_1\varepsilon} \int_{I_n} \|R(t)\|_*^2 dt + \varepsilon \|\Theta\|_n^2, \tag{3.13i}$$

cf. (2.17ii); moreover

$$(\Theta^n, \tilde{\Theta}^{n+}) \leq c_1 |\Theta^n|^2 + \frac{1}{4k_n} |\Theta_n^2|, \tag{3.13ii}$$

cf. (2.18i). Using (3.13) in (3.12) we get

$$|\Theta^{n+1}|^2 + \frac{1}{2} \left(\frac{1}{k_n} - \frac{\mu^2}{\varepsilon} \right) |\Theta_n^2| + \alpha \|\Theta\|_n^2 \leq 2c_1 |\Theta^n|^2 + \frac{1}{2\tau_1\varepsilon} \int_{I_n} \|R(t)\|_*^2 dt,$$

with $\alpha = 2(1 - \lambda - 2\varepsilon)$, for any positive ε . Hence, in view of the equivalence of the norms $\|\cdot\|_n$ and $\left(\int_{I_n} \|\cdot\|^2 dt\right)^{1/2}$ in $\mathcal{V}_q(I_n)$ with constants independent of I_n , with $\|\cdot\|$ standing for either one of the norms $|\cdot|$ and $\|\cdot\|$, we have, for $\varepsilon < (1 - \lambda)/2$ and k_n sufficiently small,

$$|\Theta^{n+1}|^2 + \frac{c}{k_n} \int_{I_n} |\Theta(t)|^2 dt + c \int_{I_n} \|\Theta(t)\|^2 dt \leq c_1 |\Theta^n|^2 + c_1 \int_{I_n} \|R(t)\|_*^2 dt. \tag{3.14}$$

In particular,

$$\int_{I_n} |\Theta(t)|^2 dt \leq Ck_n |\Theta^n|^2 + Ck_n \int_{I_n} \|R(t)\|_*^2 dt. \tag{3.15}$$

Using (3.15) in (3.11), we obtain

$$|\Theta^{n+1}|^2 + c \int_{I_n} \|\Theta(t)\|^2 dt \leq (1 + Ck_n) |\Theta^n|^2 + C \int_{I_n} \|R(t)\|_*^2 dt$$

and this yields easily

$$|\Theta^n|^2 + \int_0^{t^n} \|\Theta(t)\|^2 dt \leq ce^{Ct^n} \left[|\Theta^0|^2 + \int_0^{t^n} \|R(t)\|_*^2 dt \right]. \tag{3.16}$$

3.3. Convergence

3.3.1. Error estimation at the nodes

In view of (3.7) and the fact that $\Theta(0) = 0$, from (3.16) we obtain

$$|\Theta(t^n)|^2 \leq c e^{Ct^n} \sum_{j=0}^{n-1} k_j^{2q} \int_{I_j} \|u^{(q)}(t)\|^2 dt. \quad (3.17)$$

Now, since ρ vanishes at the nodes of the partition, see (3.1), (3.17) yields immediately the desired estimate at the nodes

$$\max_{0 \leq n \leq N} |(u - U)(t^n)|^2 \leq c e^{CT} \sum_{j=0}^{N-1} k_j^{2q} \int_{I_j} \|u^{(q)}(t)\|^2 dt. \quad (3.18)$$

3.3.2. Error analysis in $L^\infty(H)$

Combining the inverse inequality

$$|\Theta|_{L^\infty(I_n; H)}^2 \leq c k_n^{-1} |\Theta|_{L^2(I_n; H)}^2 \quad (3.19)$$

with (3.15) and using (3.7) and (3.17), we easily conclude

$$|\Theta|_{L^\infty(I_n; H)}^2 \leq c e^{Ct^{n+1}} \sum_{j=0}^n k_j^{2q} \int_{I_j} \|u^{(q)}(t)\|^2 dt. \quad (3.20)$$

Finally, (3.2) and (3.20) yield the desired uniform in time error estimate

$$\max_{0 \leq t \leq T} |(u - U)(t)|^2 \leq c \max_{0 \leq n \leq N-1} \left(k_n^q \max_{t \in I_n} |u^{(q)}(t)| \right)^2 + c e^{CT} \sum_{j=0}^{N-1} k_j^{2q} \int_{I_j} \|u^{(q)}(t)\|^2 dt. \quad (3.21)$$

3.3.3. Error analysis in $L^2(V)$

In view of (3.7) and the fact that $\Theta(0) = 0$, from (3.16) we obtain

$$\int_0^T \|\Theta(t)\|^2 dt \leq C \sum_{j=0}^{N-1} k_j^{2q} \int_{I_j} \|u^{(q)}(t)\|^2 dt. \quad (3.22)$$

Finally, (3.2) and (3.22) yield the desired error estimate

$$\int_0^T \|(u - U)(t)\|^2 dt \leq C \sum_{j=0}^{N-1} k_j^{2q} \int_{I_j} \|u^{(q)}(t)\|^2 dt. \quad (3.23)$$

3.3.4. Error analysis for the original equation

Our analysis up to this point concerns the discontinuous Galerkin approximation U , see (2.3), to the modified equation (2.2). Here, we give our main result in this section, namely error estimates for the discontinuous Galerkin approximation to the original equation (1.1).

Theorem 3.1. *Let the solution u of (1.1) be sufficiently smooth and $U \in \mathcal{V}_q^d$ be a discontinuous Galerkin approximation to (1.1), i.e., a solution of (1.2). Let $\underline{k} := \min_{0 \leq n \leq N-1} k_n$. Then, under the mesh condition*

“ $k^{2q}\underline{k}^{-1}$ sufficiently small” we have the error estimate

$$\max_{0 \leq t \leq T} |(u - U)(t)|^2 \leq C \max_{0 \leq n \leq N-1} \left(k_n^q \max_{t \in I_n} |u^{(q)}(t)| \right)^2 + C \sum_{j=0}^{N-1} k_j^{2q} \int_{I_j} \|u^{(q)}(t)\|^2 dt. \tag{3.24}$$

Proof. Let for the time being U be the discontinuous Galerkin approximation to the modified equation (2.2). First, from (3.2) we conclude, for k sufficiently small,

$$\max_{0 \leq t \leq T} \|(u - W)(t)\| \leq \frac{1}{2}. \tag{3.25}$$

Further, from (3.22) and the analogous to (3.19) inverse inequality, we get

$$\max_{0 \leq t \leq T} \|(W - U)(t)\|^2 \leq Ck^{2q}\underline{k}^{-1}, \tag{3.26}$$

and thus, under our mesh condition,

$$\max_{0 \leq t \leq T} \|(W - U)(t)\| \leq \frac{1}{2}. \tag{3.27}$$

It immediately follows from (3.25) and (3.27) that $U \in T_u$. Therefore, the discontinuous Galerkin approximation U to the modified equation (2.2) is also a (locally unique) discontinuous Galerkin approximation to the original equation (1.1), and (3.24) follows from (3.21). \square

Remark 3.1. The constants in this and previous sections as well as conditions like “ k sufficiently small” do not directly depend on the particular choice of the operators A and B ; they only depend on λ, μ , the discretization scheme and on various norms of the solution u . This fact will play a crucial role in the analysis of fully discrete schemes in the next section.

4. THE DG CASE: FULLY DISCRETE SCHEMES

In this section we consider fully discrete schemes; we combine the discontinuous Galerkin time stepping with space discrete schemes. We establish optimal order error estimates.

For the space discretization we use a family $\mathcal{V}_h, 0 < h < 1$, of finite dimensional subspaces of V . For simplicity, we will use the same finite dimensional space \mathcal{V}_h throughout the interval $[0, T]$; the analysis can be modified to take into account possible changes of this space. In this section the following discrete operators will play an essential role: define $P_o : V' \rightarrow \mathcal{V}_h, A_h : V \rightarrow \mathcal{V}_h$ and $B_h(t, \cdot) : V \rightarrow \mathcal{V}_h$ by

$$\begin{aligned} (P_o v, \chi) &= (v, \chi) & \forall \chi \in \mathcal{V}_h \\ (A_h \varphi, \chi) &= (A\varphi, \chi) & \forall \chi \in \mathcal{V}_h \\ (B_h(t, \varphi), \chi) &= (B(t, \varphi), \chi) & \forall \chi \in \mathcal{V}_h. \end{aligned}$$

The space discrete problem corresponding to (1.1) is to seek a function $u_h : [0, T] \rightarrow \mathcal{V}_h$ satisfying

$$\begin{cases} u_h'(t) + A_h u_h(t) = B_h(t, u_h(t)), & 0 < t < T, \\ u_h(0) = u_h^0, \end{cases} \tag{4.1}$$

with $u_h^0 \in \mathcal{V}_h$ a given approximation to u^0 .

To construct a fully discrete scheme, we discretize (4.1) in time by the discontinuous Galerkin method. With the notation of the previous sections and

$$\mathcal{V}_{qh}^d := \left\{ \varphi : [0, T] \rightarrow \mathcal{V}_h / \varphi|_{I_n}(t) = \sum_{j=0}^{q-1} w_j t^j \right\},$$

the fully discrete approximation $U_h \in \mathcal{V}_{qh}^d$ to u is defined by

$$\int_{I_n} [(U_h', v) + (F_h(t, U_h), v)] dt + (U_h^{n+} - U_h^n, v^{n+}) = 0 \quad \forall v \in \mathcal{V}_{qh}(I_n), \tag{4.2}$$

for $n = 0, \dots, N - 1$, with $F_h(t, v) := A_h v - B_h(t, v)$ and $U_h(0) = u_h^0$.

Let $B(t, \cdot) : V \rightarrow V'$ be differentiable, and assume that the linear operator $M(t)$, $M(t) := A - B'(t, u(t)) + \sigma I$, is uniformly positive definite, for an appropriate constant σ . Following [4], we introduce the ‘elliptic’ projection $R_h(t) : V \rightarrow \mathcal{V}_h, t \in [0, T]$, by

$$P_o M(t) R_h(t) v = P_o M(t) v. \tag{4.3}$$

We will show consistency of the space discrete scheme for $W_h, W_h(t) := R_h(t)u(t)$; to this end we shall use approximation properties of the elliptic projection operator $R_h(t)$. We assume that $R_h(t)$ satisfies the estimates

$$|u(t) - W_h(t)| + h^{d/2} \|u(t) - W_h(t)\| \leq Ch^r, \tag{4.4}$$

and

$$\left| \frac{d}{dt} [u(t) - W_h(t)] \right| \leq Ch^r, \tag{4.5}$$

with two integers r and d , $2 \leq d \leq r$. Note here that d corresponds to the order of the operator A , e.g., if A is a second order elliptic operator, as in Section 6, then $d = 2$. We further assume that

$$\int_0^T \left\| \frac{d^q}{dt^q} W_h(t) \right\|^2 dt \leq C. \tag{4.6}$$

For consistency purposes, we assume for the nonlinear part the estimate

$$\|B(t, u(t)) - B(t, W_h(t)) - B'(t, u(t))(u(t) - W_h(t))\|_* \leq Ch^r. \tag{4.7}$$

Let $E_h(t) \in \mathcal{V}_h$ denote the consistency error of the space discrete equation (4.1) for W_h ,

$$E_h(t) := W_h'(t) + A_h W_h(t) - B_h(t, W_h(t)), \quad 0 \leq t \leq T. \tag{4.8}$$

From the definition of W_h we easily conclude

$$(A_h W_h(t), \chi) = (Au(t) - [B'(t, u(t)) - \sigma I](u(t) - W_h(t)), \chi) \quad \forall \chi \in \mathcal{V}_h.$$

Therefore, using (1.1),

$$E_h(t) = W_h'(t) - P_o u'(t) + \sigma [P_o u(t) - W_h(t)] + P_o [B(t, u(t)) - B(t, W_h(t)) - B'(t, u(t))(u(t) - W_h(t))], \tag{4.3}$$

and, in view of (4.4), (4.5) and (4.7), we easily obtain the following optimal order estimate for the consistency error E_h ,

$$\max_{0 \leq t \leq T} \|E_h(t)\|_* \leq Ch^r. \tag{4.9}$$

The main result in this paper concerning the discontinuous Galerkin method is given in the following theorem:

Theorem 4.1. *Let the solution u of (1.1) be sufficiently smooth. Assume we are given an initial approximation $U_h^0 = u_h^0 \in \mathcal{V}_h$ to u^0 such that*

$$|u^0 - u_h^0| \leq Ch^r. \tag{4.10}$$

Then, for k and h sufficiently small, there exists a locally unique solution $U_h \in \mathcal{V}_{qh}^d$ to (4.2); further, there exists a constant C , independent of k and h , such that, for $k^{2q}\underline{k}^{-1}$ and $h^{2r}\underline{k}^{-1}$ sufficiently small, we have the error estimate

$$\max_{0 \leq t \leq T} |(u - U_h)(t)|^2 \leq C \left[h^{2r} + \sum_{j=0}^{N-1} k_j^{2q+1} \right]. \tag{4.11}$$

Proof. Existence and local uniqueness follow easily from our analysis in Section 2. To prove the error estimate, let first $\rho := u - W_h$; according to (4.4) we have

$$\max_{0 \leq t \leq T} |\rho(t)| \leq Ch^r. \tag{4.12}$$

Further, according to (4.4),

$$\max_{0 \leq t \leq T} \|\rho(t)\| \leq Ch^{r-\frac{d}{2}} \tag{4.13}$$

and thus, for h sufficiently small, $W_h(t) \in T_u$, $t \in [0, T]$. Now, let $\bar{B}_h(t, \cdot) : V \rightarrow \mathcal{V}_h$ be defined by

$$(\bar{B}_h(t, \varphi), \chi) = (\bar{B}(t, \varphi), \chi) \quad \forall \chi \in \mathcal{V}_h,$$

and $\bar{W}_h \in \mathcal{V}_{qh}^d$ be such that

$$\int_{I_n} [(\bar{W}'_h, v) + (\bar{F}_h(t, \bar{W}_h), v)] dt + (\bar{W}_h^{n+} - \bar{W}_h^n, v^{n+}) = 0 \quad \forall v \in \mathcal{V}_{qh}(I_n), \tag{4.14}$$

for $n = 0, \dots, N - 1$, with $\hat{B}_h(t, v) := \bar{B}_h(t, v) + E_h(t)$, $\bar{F}_h(t, v) := A_h v - \hat{B}_h(t, v)$ and $\bar{W}_h(0) = W_h^0$, i.e., \bar{W}_h is a modified fully discrete discontinuous Galerkin approximation corresponding to equation (4.8). Then, according to (3.21), and in view of (3.7) and (4.6),

$$\max_{0 \leq t \leq T} |(W_h - \bar{W}_h)(t)|^2 \leq C \sum_{j=0}^{N-1} k_j^{2q+1}. \tag{4.15}$$

Further,

$$\max_{0 \leq t \leq T} \|(W_h - \bar{W}_h)(t)\|^2 \leq Ck^{2q}\underline{k}^{-1}, \tag{4.16}$$

cf. (3.26). In view of (4.12) and (4.15), it remains to estimate $\Theta := \bar{W}_h - U_h$. We now temporarily change the meaning of U_h and let it denote the modified discontinuous Galerkin approximation,

$$\int_{I_n} [(U'_h, v) + (A_h U_h, v)] dt + (U_h^{n+} - U_h^n, v^{n+}) = \int_{I_n} (\bar{B}_h(t, U_h), v) dt \tag{4.17}$$

for all $v \in \mathcal{V}_{qh}(I_n)$, for $n = 0, \dots, N - 1$, with $U_h(0) = u_h^0$; subsequently, after having established the desired error estimate we will show that U_h is also a solution to (4.2). From (4.14) and (4.17), we obtain

$$|\Theta^n|^2 + \int_0^{t^n} \|\Theta(t)\|^2 dt \leq c e^{Ct^n} \left[|\Theta^0|^2 + \int_0^{t^n} \|E_h(t)\|_*^2 dt \right], \tag{4.18}$$

see (3.16). Now, $\Theta^0 = (W_h^0 - u^0) + (u^0 - u_h^0)$; thus, in view of (4.12) and (4.10), $|\Theta^0| \leq Ch^r$, and, using also (4.9), (4.18) yields

$$|\Theta^n|^2 + \int_0^{t^n} \|\Theta(t)\|^2 dt \leq Ch^{2r}. \tag{4.19}$$

Further,

$$\int_{I_n} |\Theta(t)|^2 dt \leq Ck_n |\Theta^n|^2 + Ck_n \int_{I_n} \|E_h(t)\|_*^2 dt,$$

see (3.15), and thus, in view of (3.19), (4.19) and (4.9),

$$\max_{0 \leq t \leq T} |\Theta(t)| \leq Ch^r. \tag{4.20}$$

Moreover, from (4.19) we obtain

$$\max_{0 \leq t \leq T} \|\Theta(t)\|^2 \leq Ch^{2r} \underline{k}^{-1}. \tag{4.21}$$

Now, it easily follows from (4.13), (4.16) and (4.21) that, for $k^{2q} \underline{k}^{-1}$ and $h^{2r} \underline{k}^{-1}$ sufficiently small, $U_h(t) \in T_u, t \in [0, T]$, and thus that U_h is a discontinuous Galerkin approximation to the original equation, *i.e.*, it satisfies (4.2). Finally, (4.12), (4.15) and (4.20) yield the error estimate (4.11) and the proof is complete. \square

5. THE CONTINUOUS GALERKIN METHOD

In this section we analyze the continuous Galerkin method for problem (1.1). We show existence and local uniqueness of the continuous Galerkin approximations and establish optimal order error estimates. Let us emphasize that in this section no error estimates in $L^2(V)$ are derived, and thus the tube T_u is defined in terms of the norm of H , $T_u := \{v \in V : \min_t |u(t) - v| \leq 1\}$, see Remark 5.1.

5.1. Existence and uniqueness

We begin by showing existence and uniqueness of the continuous Galerkin approximations for a modified equation. As before, this serves as an intermediate step and will be used in the sequel to establish existence and local uniqueness of the continuous Galerkin approximations for our original equation.

Existence and uniqueness of continuous Galerkin approximations for the nonlinear Schrödinger equation were established in [13]. The approach in [13] is based on properties of Gauss–Legendre quadrature formulae and interpolation. Our approach is direct and simplifies the proofs of [13].

As in the analysis of the discontinuous Galerkin method, we assume that $B(t, \cdot)$ can be modified to yield an operator $\bar{B}(t, \cdot) : V \rightarrow V'$ coinciding with $B(t, \cdot)$ in the tube T_u , $\bar{B}(t, v) = B(t, v)$ for all $t \in [0, T]$ and all $v \in T_u$, and satisfying the global Lipschitz condition, *cf.* (2.1),

$$\|\bar{B}(t, v) - \bar{B}(t, w)\|_* \leq \lambda \|v - w\| + \mu |v - w| \quad \forall v, w \in V. \tag{5.1}$$

The continuous Galerkin method for the modified equation

$$\begin{cases} u'(t) + Au(t) = \bar{B}(t, u(t)), & 0 < t < T, \\ u(0) = u^0, \end{cases} \tag{5.2}$$

is to seek $U \in \mathcal{V}_q$ satisfying

$$\int_{I_n} [(U', v) + (AU, v)] dt = \int_{I_n} (\bar{B}(t, U), v) dt \quad \forall v \in \mathcal{V}_{q-1}(I_n) \tag{5.3}$$

for $n = 0, \dots, N - 1$. Here $U(0) = u(0)$. It is easily seen that the solution u of (1.1) is also a solution of (5.2); further, (5.1) yields easily uniqueness of (smooth) solutions of (5.2).

Next, we shall show existence and uniqueness of the solution of the scheme (5.3). Later on, we will see that $U \in T_u$ and will easily conclude existence and local uniqueness of the solution of scheme (1.3).

The continuous Galerkin approximate solution $U \in \mathcal{V}_q^c$ is defined in I_n by its value $U(t^n)$ at t^n (which has been determined from the conditions in the preceding time interval I_{n-1}) and by (5.3). Now, since $U(t^n)$ is considered given, U can be written in I_n in the form

$$U(t) = U(t^n) + (t - t^n)W(t), \quad t \in I_n, \quad \text{with } W \in \mathcal{V}_{q-1}(I_n). \tag{5.4}$$

We now consider $W \in \mathcal{V}_{q-1}(I_n)$ our unknown and use (5.4) to rewrite (5.3) as

$$\begin{aligned} \int_{I_n} [(W, v) + (t - t^n)(W', v) + (AU(t^n), v) + (t - t^n)(AW, v)] dt \\ = \int_{I_n} (\bar{B}(t, U(t^n) + (t - t^n)W), v) dt \quad \forall v \in \mathcal{V}_{q-1}(I_n). \end{aligned} \tag{5.5}$$

Let P_{q-2} denote the L^2 orthogonal projection operator onto $\mathcal{V}_{q-1}(I_n)$. It is then easily seen that (5.5) can be written in the form

$$\mathcal{G}(W) = 0 \tag{5.6}$$

with $\mathcal{G} : \mathcal{V}_{q-1}(I_n) \rightarrow \mathcal{V}_{q-1}(I_n)$,

$$\mathcal{G}(v) := v + (t - t^n)v' + AU(t^n) + P_{q-2}((t - t^n)Av) - P_{q-2}\bar{B}(t, U(t^n) + (t - t^n)v). \tag{5.7}$$

To establish existence and uniqueness of the continuous Galerkin approximation in I_n , *i.e.*, existence and uniqueness of $W \in \mathcal{V}_{q-1}(I_n)$ at which \mathcal{G} vanishes, we show that \mathcal{G} is Lipschitz continuous and strongly monotone.

With $\|\cdot\|$ denoting either one of the norms $|\cdot|$, $\|\cdot\|$ or $\|\cdot\|_*$, for the analysis of the continuous Galerkin method, we introduce in $\mathcal{V}_q(I_n)$ the norms $|\cdot|_n$, $\|\cdot\|_n$ and $\|\cdot\|_{*n}$ by

$$\|v\|_n := \left(\int_{I_n} \|v(t)\|^2 dt \right)^{1/2}.$$

Obviously, $|\cdot|_n$ is induced by the inner product $\langle \cdot, \cdot \rangle$,

$$\langle v, w \rangle = \int_{I_n} (v, w) dt.$$

In the sequel, we will make use of the inverse inequality

$$k_n \|v\|_n^2 \leq c \int_{I_n} (t - t^n) \|v\|^2 dt, \tag{5.8}$$

cf. [18].

5.1.1. *Strong monotonicity of \mathcal{G}*

Let $v, w \in \mathcal{V}_{q-1}(I_n)$ and $\vartheta := v - w$. Now,

$$\begin{aligned} \int_{I_n} (t - t^n)(\vartheta', \vartheta) dt &= \frac{1}{2} \int_{I_n} (t - t^n) \frac{d}{dt} |\vartheta|^2 dt \\ &= \frac{1}{2} [(t - t^n)|\vartheta(t)|^2]_{t=t^n}^{t=t^{n+1}} - \frac{1}{2} \int_{I_n} |\vartheta|^2 dt, \end{aligned}$$

i.e.,

$$\langle (t - t^n)\vartheta', \vartheta \rangle = \frac{1}{2} k_n |\vartheta(t^{n+1})|^2 - \frac{1}{2} |\vartheta|_n^2. \quad (5.9)$$

Further, it immediately follows from (5.1) that

$$(\bar{B}(t, U(t^n)) + (t - t^n)v - \bar{B}(t, U(t^n)) + (t - t^n)w, \vartheta) \leq (t - t^n) \left[(\lambda + \varepsilon) \|\vartheta\|^2 + \frac{\mu^2}{4\varepsilon} |\vartheta|^2 \right], \quad (5.10)$$

for any positive ε . In view of (5.9) and (5.10), we obtain

$$\langle \mathcal{G}(v) - \mathcal{G}(w), v - w \rangle \geq \frac{1}{2} k_n |\vartheta(t^{n+1})|^2 + \frac{1}{2} |\vartheta|_n^2 + \int_{I_n} (t - t^n) \left[(1 - \lambda - \varepsilon) \|\vartheta\|^2 - \frac{\mu^2}{4\varepsilon} |\vartheta|^2 \right] dt.$$

Using here (5.8) and letting $c' := 1/c$, we have

$$\langle \mathcal{G}(v) - \mathcal{G}(w), v - w \rangle \geq (1 - \lambda - \varepsilon) c' k_n \|\vartheta\|_n^2 + \frac{1}{2} \left(1 - \frac{\mu^2}{4\varepsilon} k_n \right) |\vartheta|_n^2.$$

Therefore, for $\varepsilon = (1 - \lambda)/2$ and $k_n \leq (1 - \lambda)/\mu^2$, we have the following strong monotonicity property of \mathcal{G}

$$\langle \mathcal{G}(v) - \mathcal{G}(w), v - w \rangle \geq \frac{1 - \lambda}{2} c' k_n \|v - w\|_n^2 \quad \forall v, w \in \mathcal{V}_{q-1}(I_n). \quad (5.11)$$

In particular, (5.11) yields immediately uniqueness of the continuous Galerkin approximation, for sufficiently small $k, k = \max_n k_n$.

5.1.2. Lipschitz continuity of \mathcal{G}

Let $v, w, \omega \in \mathcal{V}_{q-1}(I_n)$ and $\vartheta := v - w$. Using (5.1) we have

$$\langle \mathcal{G}(v) - \mathcal{G}(w), \omega \rangle \leq |\vartheta|_n |\omega|_n + k_n |\vartheta'|_n |\omega|_n + k_n \|\vartheta\|_n \|\omega\|_n + k_n (\lambda \|\vartheta\|_n + \mu |\vartheta|_n) \|\omega\|_n$$

and thus, in view also of the inverse inequality

$$k_n |v'|_n \leq c |v|_n \quad \forall v \in \mathcal{V}_{q-1}(I_n) \quad (5.12)$$

we easily conclude, for fixed k_n ,

$$\langle \mathcal{G}(v) - \mathcal{G}(w), \omega \rangle \leq C \|v - w\|_n \|\omega\|_n,$$

i.e.,

$$\|\mathcal{G}(v) - \mathcal{G}(w)\|_{\star n} \leq C \|v - w\|_n \quad \forall v, w \in \mathcal{V}_{q-1}(I_n). \quad (5.13)$$

Existence and uniqueness of the continuous Galerkin approximation, for sufficiently small $k := \max_n k_n$, follows now easily from (5.11) and (5.13), see the corresponding result for the discontinuous Galerkin approximation.

5.2. Error estimates

We shall establish optimal order estimates for $u - U$, U being the solution of (5.3), under the assumption that $\bar{B}(t, \cdot) : D(A) \rightarrow H$ coincides with $B(t, \cdot)$ in the tube T_u , $\bar{B}(t, v) = B(t, v)$ for all $t \in [0, T]$ and all $v \in T_u$, and satisfies the global Lipschitz condition (5.1). After having established the error estimate we can show that for sufficiently small time steps the solution of (5.3) satisfies also (1.3), and, thus, we will have error estimates for the original equation.

Let $W \in \mathcal{V}_q^c$ be defined by

$$\begin{aligned} W(0) &= u(0), \\ \int_0^T (u' - W', v') dt &= 0, \quad \forall v \in \mathcal{V}_q^c, \end{aligned} \tag{5.14}$$

cf. [5]. The function W will play an important role in the error analysis. With

$$v(t) := \begin{cases} (u - W)(t^n)t, & 0 \leq t \leq t^n, \\ (u - W)(t^n)t^n, & t^n < t \leq T, \end{cases} \tag{5.15}$$

relation (5.14) yields $W(t^n) = u(t^n)$, i.e., W has the following interpolation property

$$W(t^n) = u(t^n), \quad n = 0, \dots, N. \tag{5.16}$$

It is then easily seen that the restriction of W in I_n could have been alternatively defined by

$$\begin{aligned} W(t^n) &= u(t^n), \quad W(t^{n+1}) = u(t^{n+1}), \\ \int_{I_n} (u' - W', v) dt &= 0 \quad \forall v \in \mathcal{V}_{q-1}(I_n). \end{aligned} \tag{5.17}$$

Existence, uniqueness and the error estimate

$$\|(u - W)(t)\|^2 \leq Ck_n^{2q-1} \int_{I_n} \|u^{(q)}(s)\|^2 ds, \quad t \in I_n, \tag{5.18}$$

with $\|\cdot\|$ standing for either one of the norms $|\cdot|$, $\|\cdot\|$ or $\|\cdot\|_\star$, can be established by standard arguments, cf. (3.1) and (3.2).

With $\rho := u - W$ and $\Theta := W - U$, we split the error $u - U$ in the form

$$u - U = \rho + \Theta.$$

Since ρ has been estimated in (5.18), our main goal in this section will be the estimation of Θ . We will achieve this in two stages: first we will show consistency of the numerical scheme for the interpolant W and subsequently we will show stability.

5.3. Consistency

Let $R \in \mathcal{V}_{q-1}(I_n)$ denote the consistency error of the continuous Galerkin method (5.3) for W ,

$$R = W' + P_{q-2}(AW - \bar{B}(\cdot, W)). \tag{5.19}$$

Let $v \in \mathcal{V}_{q-1}(I_n)$. Then,

$$\int_{I_n} (R, v) dt = \int_{I_n} (W' + AW - \bar{B}(t, W), v) dt,$$

and thus, in view of (1.1) and (5.17),

$$\int_{I_n} (R, v) dt = - \int_{I_n} (A\rho, v) dt + \int_{I_n} (\bar{B}(t, u) - \bar{B}(t, W), v) dt. \tag{5.20}$$

Letting $v := A^{-1}R$ in (5.20) and using (5.1) we have

$$\int_{I_n} \|R(t)\|_\star^2 dt \leq - \int_{I_n} (A^{1/2}\rho, A^{-1/2}R) dt + \int_{I_n} [\lambda\|\rho\| + \mu|\rho|] \|R(t)\|_\star dt,$$

and thus easily

$$\int_{I_n} \|R(t)\|_*^2 dt \leq 2 \int_{I_n} \|\rho(t)\|^2 dt + 2 \int_{I_n} [\lambda \|\rho(t)\| + \mu |\rho(t)|]^2 dt. \tag{5.21}$$

From (5.18) and (5.21) we obtain the consistency estimate

$$\int_{I_n} \|R(t)\|_*^2 dt \leq C k_n^{2q} \int_{I_n} \|u^{(q)}(s)\|^2 ds. \tag{5.22}$$

5.4. Stability

From (5.3) and (5.19) we obtain an error equation for Θ ,

$$\int_{I_n} (\Theta', v) dt + \int_{I_n} (A\Theta, v) dt = \int_{I_n} (\bar{B}(t, W) - \bar{B}(t, U), v) dt + \int_{I_n} (R(t), v) dt \tag{5.23}$$

for all $v \in \mathcal{V}_{q-1}(I_n)$. The stability analysis relies on the following auxiliary result.

Lemma 5.1. *Let $0 < \tau_1 < \dots < \tau_{q-1} < 1$ be the abscissae of the Gauss–Legendre quadrature formula in $[0, 1]$, and $t^{n,i}, t^{n,i} := t^n + k_n \tau_i$, $i = 1, \dots, q - 1$, denote the corresponding nodes shifted to I_n . Let $\Theta \in \mathbb{P}_{q-1}(I_n)$, and $\hat{\Theta}, \tilde{\Theta} \in \mathbb{P}_{q-2}(I_n)$ be the interpolants of Θ and φ , $\varphi(t) := k_n \Theta(t)/(t - t^n)$, at $t^{n,i}$, $i = 1, \dots, q - 1$, respectively. Then*

$$\int_{I_n} \Theta' \hat{\Theta} dt = \int_{I_n} \Theta' \Theta dt \tag{5.24}$$

and

$$k_n \int_{I_n} \Theta' \tilde{\Theta} dt \geq \frac{1}{5} \int_{I_n} |\Theta|^2 dt - ck_n (|\Theta(t^{n+1})|^2 + |\Theta(t^n)|^2), \tag{5.25}$$

with a positive constant c .

Proof. Let w_1, \dots, w_{q-1} denote the weights of the Gauss–Legendre quadrature formula in $[0, 1]$ with $q - 1$ nodes. It is well known that the corresponding Gauss–Legendre quadrature formula Q ,

$$Q(v) = k_n [w_1 v(t^{n,1}) + \dots + w_{q-1} v(t^{n,q-1})],$$

integrates in I_n polynomials of degree at most $2q - 3$ exactly. Since $\Theta(\tau_i) = \hat{\Theta}(\tau_i)$, obviously $Q(\Theta' \Theta) = Q(\Theta' \hat{\Theta})$, and (5.24) follows. To show (5.25), we first write Θ in the form

$$\Theta(t) = \Theta(t^n) + (t - t^n)Z(t),$$

with $Z \in \mathbb{P}_{q-2}(I_n)$,

$$Z(t) = \frac{1}{t - t^n} [\Theta(t) - \Theta(t^n)].$$

Then

$$\varphi(t) = k_n Z(t) + \Theta(t^n) \frac{k_n}{t - t^n}.$$

Therefore, with $\Lambda \in \mathbb{P}_{q-2}(I_n)$ denoting the interpolant of $k_n/(t - t^n)$ at $t^{n,i}$, $i = 1, \dots, q - 1$, i.e.,

$$\Lambda(t^{n,i}) = \frac{1}{\tau_i}, \quad i = 1, \dots, q - 1,$$

$\tilde{\Theta}$ may be written in the form $\tilde{\Theta} = k_n Z + \Theta(t^n) \Lambda$. Now, cf. (5.9),

$$\int_{I_n} \Theta'(t) Z(t) dt = \int_{I_n} [Z(t) + (t - t^n) Z'(t)] Z(t) dt = \frac{1}{2} k_n |Z(t^{n+1})|^2 + \frac{1}{2} \int_{I_n} |Z|^2 dt,$$

i. e.,

$$k_n \int_{I_n} \Theta'(t)Z(t) dt = \frac{1}{2}|\Theta(t^{n+1}) - \Theta(t^n)|^2 + \frac{1}{2}k_n \int_{I_n} |Z|^2 dt. \tag{5.26}$$

Further, from the definition of Z we easily get

$$\int_{I_n} |\Theta|^2 dt \leq 2k_n|\Theta(t^n)|^2 + 2k_n^2 \int_{I_n} |Z|^2 dt. \tag{5.27}$$

Relations (5.26) and (5.27) yield immediately

$$k_n^2 \int_{I_n} \Theta'(t)Z(t) dt \geq \frac{1}{2}k_n [|\Theta(t^{n+1}) - \Theta(t^n)|^2 - |\Theta(t^n)|^2] + \frac{1}{4} \int_{I_n} |\Theta|^2 dt. \tag{5.28}$$

Next, we shall estimate $\int_{I_n} \Theta' \Lambda dt$. We rewrite this term in the form

$$\int_{I_n} \Theta' \Lambda dt = - \int_{I_n} \Theta \Lambda' dt + \Theta(t^{n+1})\Lambda(t^{n+1}) - \Theta(t^n)\Lambda(t^n),$$

and use the fact that Λ and $k_n \Lambda'$ are uniformly bounded to obtain

$$k_n \left| \int_{I_n} \Theta' \Theta(t^n) \Lambda dt \right| \leq \varepsilon \int_{I_n} |\Theta|^2 dt + k_n [c|\Theta(t^{n+1})|^2 + c_\varepsilon|\Theta(t^n)|^2] \tag{5.29}$$

for any positive ε . Choosing $\varepsilon \leq 1/20$ in (5.29), from (5.28) and (5.29) we easily conclude that (5.25) is valid and the proof is complete. \square

Letting $v = \hat{\Theta} \in \mathcal{V}_{q-1}(I_n)$ in (5.23) be the interpolant of Θ at $t^{n,i}, i = 1, \dots, q-1$, and using (5.24), we have

$$\int_{I_n} (\Theta', \Theta) dt + \int_{I_n} (A\Theta, \hat{\Theta}) dt = \int_{I_n} (\bar{B}(t, W) - \bar{B}(t, U), \hat{\Theta}) dt + \int_{I_n} (R(t), \hat{\Theta}) dt. \tag{5.30}$$

In view of the exactness of the Gaussian quadrature rule, the left-hand side can be written in the form

$$\int_{I_n} (\Theta', \Theta) dt + \int_{I_n} (A\Theta, \hat{\Theta}) dt = \frac{1}{2} [|\Theta^{n+1}|^2 - |\Theta^n|^2] + k_n \sum_{i=1}^{q-1} w_i \|\Theta(t^{n,i})\|^2. \tag{5.31}$$

To estimate the first term on the right-hand side of (5.30), we first note that using (5.1) we have

$$\begin{aligned} (\bar{B}(t, W) - \bar{B}(t, U), \hat{\Theta}) &\leq \|\bar{B}(t, W) - \bar{B}(t, U)\|_* \|\hat{\Theta}\| \leq \lambda \|\Theta\| \|\hat{\Theta}\| + \mu |\Theta| \|\hat{\Theta}\| \\ &\leq \left[\lambda \|\Theta\| \|\hat{\Theta}\| + \frac{1}{2} \mu \varepsilon \|\hat{\Theta}\|^2 \right] + \frac{1}{2\varepsilon} \mu |\Theta|^2; \end{aligned}$$

since the first term on the right-hand side is integrated exactly by the Gaussian quadrature formula, we obtain

$$\int_{I_n} (\bar{B}(t, W) - \bar{B}(t, U), \hat{\Theta}) dt \leq \left(\lambda + \frac{1}{2} \mu \varepsilon \right) k_n \sum_{i=1}^{q-1} w_i \|\Theta(t^{n,i})\|^2 + \frac{1}{2\varepsilon} \mu \int_{I_n} |\Theta|^2 dt. \tag{5.32}$$

Further, $\|\hat{\Theta}\|^2$ is integrated exactly by the Gaussian quadrature formula, and the second term on the right-hand side of (5.30) can be estimated in the form

$$\int_{I_n} (R(t), \hat{\Theta}) dt \leq \frac{1}{2\varepsilon} \int_{I_n} \|R(t)\|_*^2 dt + \frac{1}{2} \varepsilon k_n \sum_{i=1}^q w_i \|\Theta(t^{n,i})\|^2. \tag{5.33}$$

Using (5.31), (5.32) and (5.33) in (5.30), we obtain

$$|\Theta^{n+1}|^2 + \alpha k_n \sum_{i=1}^{q-1} w_i \|\Theta(t^{n,i})\|^2 \leq |\Theta^n|^2 + \mu_\varepsilon \int_{I_n} |\Theta|^2 dt + \frac{1}{\varepsilon} \int_{I_n} \|R(t)\|_*^2 dt \tag{5.34}$$

with $\alpha := 2(1 - \lambda) - (\mu + 1)\varepsilon$ and $\mu_\varepsilon := \frac{1}{\varepsilon}\mu$; we assume in the sequel that ε is sufficiently small such that α be positive.

Next, we would like to estimate $|\Theta|_{L^2(I_n;H)} = \left(\int_{I_n} |\Theta|^2 dt\right)^{1/2}$. Letting $v := \tilde{\Theta} \in \mathcal{V}_{q-1}(I_n)$ in (5.23) be the interpolant of $k_n \frac{1}{t-t^n} \Theta(t)$ at $t^{n,i}$, $i = 1, \dots, q$, and using (5.25), we have

$$\frac{1}{5} \int_{I_n} |\Theta|^2 dt + k_n \int_{I_n} (A\Theta, \tilde{\Theta}) dt \leq ck_n (|\Theta^{n+1}|^2 + |\Theta^n|^2) + k_n \int_{I_n} (\bar{B}(t, W) - \bar{B}(t, U), \tilde{\Theta}) dt + k_n \int_{I_n} (R(t), \tilde{\Theta}) dt. \tag{5.35}$$

Now, in view of the exactness of the Gaussian quadrature formula,

$$\int_{I_n} (A\Theta, \tilde{\Theta}) dt = k_n \sum_{i=1}^{q-1} \frac{w_i}{\tau_i} \|\Theta(t^{n,i})\|^2. \tag{5.36}$$

Further,

$$\int_{I_n} (\bar{B}(t, W) - \bar{B}(t, U), \tilde{\Theta}) dt \leq \lambda k_n \sum_{i=1}^{q-1} \frac{w_i}{\tau_i} \|\Theta(t^{n,i})\|^2 + \mu \int_{I_n} |\Theta| \|\tilde{\Theta}\| dt. \tag{5.37}$$

From (5.35), (5.36) and (5.37), we obtain

$$\begin{aligned} \frac{1}{5} \int_{I_n} |\Theta|^2 dt + k_n^2(1 - \lambda) \sum_{i=1}^{q-1} \frac{w_i}{\tau_i} \|\Theta(t^{n,i})\|^2 \\ \leq ck_n (|\Theta^{n+1}|^2 + |\Theta^n|^2) + \mu k_n \int_{I_n} |\Theta| \|\tilde{\Theta}\| dt + k_n \int_{I_n} \|R(t)\|_* \|\tilde{\Theta}\| dt. \end{aligned} \tag{5.38}$$

Now, $\|\tilde{\Theta}\|^2$ is integrated exactly by the Gaussian quadrature formula, and we have

$$\int_{I_n} \|\tilde{\Theta}\|^2 dt = k_n \sum_{i=1}^{q-1} \frac{w_i}{\tau_i^2} \|\Theta(t^{n,i})\|^2;$$

hence

$$\int_{I_n} \|\tilde{\Theta}\|^2 dt \leq \frac{1}{\tau_1} k_n \sum_{i=1}^{q-1} \frac{w_i}{\tau_i} \|\Theta(t^{n,i})\|^2. \tag{5.39}$$

Further,

$$\mu |\Theta| \|\tilde{\Theta}\| \leq \frac{\mu^2}{2\varepsilon\tau_1} |\Theta|^2 + \frac{\tau_1\varepsilon}{2} \|\tilde{\Theta}\|^2,$$

and, in view of (5.39), we obtain

$$\mu \int_{I_n} |\Theta| \|\tilde{\Theta}\| dt \leq \frac{\mu^2}{2\varepsilon\tau_1} \int_{I_n} |\Theta|^2 dt + \frac{\varepsilon}{2} k_n \sum_{i=1}^{q-1} \frac{w_i}{\tau_i} \|\Theta(t^{n,i})\|^2. \tag{5.40}$$

Similarly,

$$\int_{I_n} \|R(t)\|_* \|\tilde{\Theta}\| dt \leq \frac{1}{2\varepsilon\tau_1} \int_{I_n} \|R(t)\|_*^2 dt + \frac{\varepsilon}{2} k_n \sum_{i=1}^{q-1} \frac{w_i}{\tau_i} \|\Theta(t^{n,i})\|^2. \tag{5.41}$$

From (5.38), (5.40) and (5.41) we get

$$\begin{aligned} \left(1 - \frac{5\mu^2}{2\varepsilon\tau_1} k_n\right) \int_{I_n} |\Theta|^2 dt + 5k_n^2(1 - \lambda - \varepsilon) \sum_{i=1}^{q-1} \frac{w_i}{\tau_i} \|\Theta(t^{n,i})\|^2 \\ \leq ck_n (|\Theta^{n+1}|^2 + |\Theta^n|^2) + \frac{5}{2\varepsilon\tau_1} k_n \int_{I_n} \|R(t)\|_*^2 dt; \end{aligned} \tag{5.42}$$

therefore, for $\varepsilon < 1 - \lambda$ and k_n sufficiently small,

$$\int_{I_n} |\Theta|^2 dt \leq ck_n (|\Theta^{n+1}|^2 + |\Theta^n|^2) + ck_n \int_{I_n} \|R(t)\|_*^2 dt. \tag{5.43}$$

From (5.34) and (5.43), we easily obtain, for k_n sufficiently small,

$$|\Theta^{n+1}|^2 + k_n \sum_{i=1}^{q-1} w_i \|\Theta(t^{n,i})\|^2 \leq (1 + Ck_n) |\Theta^n|^2 + C \int_{I_n} \|R(t)\|_*^2 dt \tag{5.44}$$

and this yields easily

$$|\Theta^n|^2 + \sum_{\ell=0}^{n-1} k_\ell \sum_{i=1}^{q-1} w_i \|\Theta(t^{\ell,i})\|^2 \leq ce^{Ct^n} \left[|\Theta^0|^2 + \int_0^{t^n} \|R(t)\|_*^2 dt \right], \tag{5.45}$$

cf. (3.16).

5.5. Convergence

5.5.1. Error estimation at the nodes

In view of (5.22) and the fact that $\Theta(0) = 0$, from (5.45) we obtain

$$|\Theta(t^n)|^2 \leq ce^{Ct^n} \sum_{j=0}^{n-1} k_j^{2q} \int_{I_j} \|u^{(q)}(t)\|^2 dt. \tag{5.46}$$

Now, since ρ vanishes at the nodes of the partition, see (5.16), (5.46) yields immediately the desired estimate at the nodes

$$\max_{0 \leq n \leq N} |(u - U)(t^n)|^2 \leq ce^{CT} \sum_{j=0}^{N-1} k_j^{2q} \int_{I_j} \|u^{(q)}(t)\|^2 dt. \tag{5.47}$$

5.5.2. Error analysis in $L^\infty(H)$

Combining the inverse inequality

$$|\Theta|_{L^\infty(I_n; H)}^2 \leq ck_n^{-1} |\Theta|_{L^2(I_n; H)}^2,$$

cf. (3.19), with (5.43) and using (5.22) and (5.46), we easily conclude

$$|\Theta|_{L^\infty(I_n; H)}^2 \leq c e^{Ct^{n+1}} \sum_{j=0}^n k_j^{2q} \int_{I_j} \|u^{(q)}(t)\|^2 dt. \tag{5.48}$$

Finally, (5.18) and (5.48) yield the desired uniform in time error estimate

$$\max_{0 \leq t \leq T} |(u - U)(t)|^2 \leq C \max_{0 \leq n \leq N-1} \left(k_n^q \max_{t \in I_n} |u^{(q)}(t)| \right)^2 + C e^{cT} \sum_{j=0}^{N-1} k_j^{2q} \int_{I_j} \|u^{(q)}(t)\|^2 dt. \tag{5.49}$$

Up to this point in this section U is considered a continuous Galerkin approximation for the modified equation (5.2). However, it immediately follows from (5.49) that, for sufficiently small k , we have $U \in T_u$; therefore U is also a continuous Galerkin approximation for the original equation (1.1).

Remark 5.1. In contrast to the discontinuous Galerkin method, error estimates in $L^2(V)$ for the continuous Galerkin method are not directly obtained. The essential reason for this difference between the two methods might be the fact that the continuous Galerkin method has less advantageous smoothing properties than the discontinuous Galerkin method, cf. [18]. Concerning our approach, estimates in $L^2(V)$ do not follow at once for the continuous Galerkin approximations, since the expression

$$\left(k_n \sum_{i=1}^{q-1} w_i \|v(t^{n,i})\|^2 \right)^{1/2} \tag{5.50}$$

cf., e.g., (5.34), is a seminorm in $\mathcal{V}_q(I_n)$ rather than a norm. An expression of the form (5.50) is a norm in $\mathcal{V}_q(I_n)$ equivalent to

$$\left(\int_{I_n} \|v(t)\|^2 dt \right)^{1/2}$$

if the sum contains at least q terms. In concrete applications one might derive estimates in $\|\cdot\|$ at an additional point in the interval I_n and then combine this with our results to obtain estimates in $L^2(V)$. The lack of estimates in $L^2(V)$ is the reason for the definition of the tube T_u in terms of the norm of H in this section. As noted in the introduction, in applications the choice of the appropriate tube depends on the concrete problem. Thus in the analysis of fully discrete schemes the inclusion of the approximate solution to the appropriate tube is verified using the estimates obtained, certain inverse inequalities, and appropriate mesh conditions, cf. Section 6 for example. For the continuous Galerkin method in the fully discrete case the mesh conditions needed can be relaxed by obtaining extra control of the error in the norm of V at an additional point in the interval I_n (different from $t^{n,i}, i = 1, \dots, q - 1$). The implementation of this task depends on the particular application.

6. APPLICATION TO A QUASILINEAR EQUATION

In this section we shall briefly discuss the application of our abstract results to a class of quasilinear equations: Let $\Omega \subset \mathbb{R}^\nu, \nu = 1, 2, 3$, be a bounded domain with smooth boundary $\partial\Omega$. For $T > 0$ we seek a real-valued function u , defined on $\bar{\Omega} \times [0, T]$, satisfying

$$\begin{cases} u_t = \operatorname{div}(c(x, t, u)\nabla u + g(x, t, u)) + f(x, t, u) & \text{in } \Omega \times [0, T], \\ u = 0 & \text{on } \partial\Omega \times [0, T], \\ u(\cdot, 0) = u^0 & \text{in } \Omega, \end{cases} \tag{6.1}$$

with $c : \bar{\Omega} \rightarrow (0, \infty)$, $f : \bar{\Omega} \times [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$, $g : \bar{\Omega} \times [0, T] \times \mathbb{R} \rightarrow \mathbb{R}^\nu$, and $u^0 : \bar{\Omega} \rightarrow \mathbb{R}$ given smooth functions. We are interested in approximating smooth solutions of this problem, and assume therefore that the data are smooth and compatible such that (6.1) gives rise to a sufficiently regular solution.

For the discretization of (6.1) by implicit-explicit finite element multistep methods we refer to [4]; other applications are included in [3].

Let $H^s = H^s(\Omega)$ be the usual Sobolev spaces of order s , and $\|\cdot\|_{H^s}$ be the norm of H^s . The inner product in $H := L^2(\Omega)$ is denoted by (\cdot, \cdot) , and the induced norm by $|\cdot|$; the norm of $L^s(\Omega)$, $1 \leq s \leq \infty$, is denoted by $\|\cdot\|_{L^s}$.

Let $\mathcal{U} := [-1 + \min_{x,t} u, 1 + \max_{x,t} u]$, and

$$\begin{aligned} \tilde{T}_u &:= \{v \in V \cap L^\infty : \min_t \|u(t) - v\|_{L^\infty} \leq 1\}, \\ \hat{T}_u &:= \{v \in V \cap W_\infty^1 : \min_t \|u(t) - v\|_{W_\infty^1} \leq 1\}. \end{aligned}$$

Let $c_\star > 0$ and c^\star be such that

$$c_\star \leq c(x, t, y) \leq c^\star \quad \forall x \in \bar{\Omega}, t \in [0, T], y \in \mathcal{U}.$$

We set

$$a := \frac{c_\star + c^\star}{2}, \quad b(x, t, y) := c(x, t, y) - a,$$

$$A := -a\Delta, \quad B(t, v) := \operatorname{div}(b(\cdot, t, v)\nabla v) + \operatorname{div}g(\cdot, t, y) + f(\cdot, t, y).$$

Then, obviously, $V = H_0^1 = H_0^1(\Omega)$ and the norm $\|\cdot\|$ in V , $\|v\| = \sqrt{a}|\nabla v|$, is equivalent to the H^1 -norm.

Let now

$$\lambda := \sup\{|b(x, t, y)|/a : x \in \Omega, t \in [0, T], y \in \mathcal{U}\};$$

it is then easily seen that $\lambda = 1 - \frac{c_\star}{a} < 1$.

For $v, w, \varphi \in V$,

$$\begin{aligned} (B(t, v) - B(t, w), \varphi) &= -(b(\cdot, t, w)\nabla(v - w), \nabla\varphi) - ([b(\cdot, t, v) - b(\cdot, t, w)]\nabla v, \nabla\varphi) \\ &\quad - (g(\cdot, t, v) - g(\cdot, t, w), \nabla\varphi) + (f(\cdot, t, v) - f(\cdot, t, w), \varphi), \end{aligned}$$

and we easily see that

$$\|B(t, v) - B(t, w)\|_\star \leq \lambda\|v - w\| + \mu|v - w| \quad \text{for all } v \in \hat{T}_u, w \in \tilde{T}_u; \tag{6.2}$$

thus, a stability condition of the form (1.5) is satisfied for $v \in \hat{T}_u$ and $w \in \tilde{T}_u$.

Further,

$$B'(t, v)w = \operatorname{div}(b(\cdot, t, v)\nabla w) + \operatorname{div}(\partial_3 b(\cdot, t, v)w\nabla v) + \operatorname{div}(\partial_3 g(\cdot, t, v)w) + \partial_3 f(\cdot, t, v)w,$$

and, therefore, $A - B'(t, u(t)) + \sigma I$ is, for an appropriate constant σ , uniformly positive definite in H_0^1 .

Let V_h be the subspace of V defined on a regular finite element partition \mathcal{T}_h of Ω , and consisting of piecewise polynomial functions of degree at most $r - 1$, $r \geq 2$. Let h_K denote the diameter of an element $K \in \mathcal{T}_h$, and $h := \max_{K \in \mathcal{T}_h} h_K$. We define the elliptic projection operator $R_h(t)$, $R_h(t) : V \rightarrow V_h$, $t \in [0, T]$, by

$$\begin{aligned} ([a(\cdot) + b(\cdot, t, u(\cdot, t))]\nabla(v - R_h(t)v), \nabla\chi) &+ ([\partial_3 b(\cdot, t, u(\cdot, t))]\nabla u(\cdot, t) + \partial_3 g(\cdot, t, u(\cdot, t)))(v - R_h(t)v), \nabla\chi \\ &- ([\partial_3 f(\cdot, t, u(\cdot, t)) - \sigma](v - R_h(t)v), \chi) = 0 \quad \forall \chi \in V_h. \end{aligned}$$

It is well known from the error analysis for elliptic equations that

$$|v - R_h(t)v| + h\|v - R_h(t)v\| \leq Ch^r\|v\|_{H^r}, \quad v \in H^r \cap H_0^1, \tag{6.3}$$

i.e., the estimate (4.4) is satisfied with $d = 2$. Further,

$$\left| \frac{d}{dt} [u(\cdot, t) - R_h(t)u(\cdot, t)] \right| \leq Ch^r \tag{6.4}$$

and

$$\left| \frac{d^q}{dt^q} R_h(t)v \right| + h \left\| \frac{d^q}{dt^q} R_h(t)v \right\| \leq Ch^r \|v\|_{H^r}, \quad v \in H^r \cap H_0^1, \tag{6.5}$$

cf., *e.g.*, [6]; thus (4.5) and (4.6) are valid. We further assume, *cf.* [17], that

$$\sup_t \|u(\cdot, t) - R_h(t)u(\cdot, t)\|_{W_\infty^1} \leq \frac{1}{2}. \tag{6.6}$$

Next, we will verify (4.7). We have

$$\begin{aligned} B(t, u(t)) - B(t, R_h(t)u(t)) - B'(t, u(t))(R_h(t)u(t) - u(t)) = \\ - \int_0^1 \tau B''(t, R_h(t)u(t) - \tau[R_h(t)u(t) - u(t)]) d\tau [R_h(t)u(t) - u(t)]^2 \end{aligned} \tag{6.7i}$$

and

$$\begin{aligned} B''(t, v)w^2 = \operatorname{div}(\partial_3^2 b(\cdot, t, v)w^2 \nabla v) + 2\operatorname{div}(\partial_3 b(\cdot, t, v)w \nabla w) \\ + \operatorname{div}(\partial_3^2 g(\cdot, t, v)w^2) + \partial_3^2 f(\cdot, t, v)w^2. \end{aligned} \tag{6.7ii}$$

It easily follows from (6.7) and (6.3), in view of (6.6), that

$$\|B(t, u(t)) - B(t, R_h(t)u(t)) - B'(t, u(t))(u(t) - R_h(t)u(t))\|_{H^{-1}} \leq Ch^r, \tag{6.8}$$

i.e., (4.7) is satisfied.

We further assume we are given an initial approximation $u_h^0 \in V_h$ to u^0 such that

$$|u^0 - u_h^0| \leq ch^r. \tag{6.9}$$

The discontinuous Galerkin scheme. We define $U_h \in \mathcal{V}_{qh}^d, U_h(0) = u_h^0$, recursively by the fully discrete discontinuous Galerkin scheme

$$\begin{aligned} \int_{I_n} [(U_h', v) + (c(\cdot, t, U_h)\nabla U_h + g(\cdot, t, U_h), \nabla v) - (f(\cdot, t, U_h)\nabla U_h, v)] dt \\ + (U_h^{n+} - U_h^n, v^{n+}) = 0 \quad \forall v \in \mathcal{V}_{qh}(I_n), \end{aligned} \tag{6.10}$$

for $n = 0, \dots, N - 1$. Then, Theorem 4.1 yields, in view of (6.6), for sufficiently small k and provided that the approximate solutions $U_h(t), t \in [0, T]$, are in \tilde{T}_u , the error estimate

$$\max_{0 \leq t \leq T} |(u - U_h)(t)|^2 \leq C \left[h^{2r} + \sum_{j=0}^{N-1} k_j^{2q+1} \right]. \tag{6.11}$$

To ensure that $U_h(t) \in \tilde{T}_u, t \in [0, T]$, we define $\underline{h} := \min_{K \in \mathcal{T}_h} h_K$ and will distinguish three cases: $\nu = 1, \nu = 2$ and $\nu = 3$.

(i). $\nu = 1$. First, since the H^1 -norm dominates the L^∞ -norm in one space dimension, we have

$$\max_{0 \leq t \leq T} \|\Theta(t)\|_{L^\infty} \leq C \max_{0 \leq t \leq T} \|\Theta(t)\|,$$

and thus, according to (4.21),

$$\max_{0 \leq t \leq T} \|\Theta(t)\|_{L^\infty}^2 \leq Ch^{2r} \underline{k}^{-1}.$$

Therefore, for $\underline{k}^{-1}k^{2q}$ and $\underline{k}^{-1}h^{2r}$ sufficiently small, in view of (6.6), $U_h(t) \in \tilde{T}_u, t \in [0, T]$. We easily conclude that the convergence result holds.

(ii). $\nu = 2$. First, we note that

$$\|\chi\|_{L^\infty} \leq C|\log(\underline{h})|^{1/2}\|\chi\|_{H^1} \quad \forall \chi \in V_h,$$

(cf. [18], p. 68). It is then easily seen that the convergence result holds, if k and h are chosen such that $|\log(\underline{h})|k^{2q}\underline{k}^{-1}$ and $|\log(\underline{h})|h^{2r}\underline{k}^{-1}$ are sufficiently small.

(iii). $\nu = 3$. In this case,

$$\|\chi\|_{L^\infty} \leq C\underline{h}^{-1/2}\|\chi\|_{H^1} \quad \forall \chi \in V_h,$$

and the result (6.11) holds, provided that $k^{2q}\underline{h}^{-1}\underline{k}^{-1}$ and $h^{2r}\underline{k}^{-1}\underline{h}^{-1}$ are sufficiently small.

The continuous Galerkin scheme. The fully discrete continuous Galerkin scheme is defined by seeking $U_h \in \mathcal{V}_{qh}^c, U_h(0) = u_h^0$, recursively by

$$\int_{I_n} [(U'_h, v) + (c(\cdot, t, U_h)\nabla U_h + g(\cdot, t, U_h), \nabla v) - (f(\cdot, t, U_h)\nabla U_h, v)] dt = 0 \quad \forall v \in \mathcal{V}_{(q-1)h}(I_n), \quad (6.12)$$

for $n = 0, \dots, N - 1$. Then, the fully discrete version of the analysis in Section 5 yields, for sufficiently small k and provided that the approximate solutions $U_h(t), t \in [0, T]$, are in \tilde{T}_u , the error estimate

$$\max_{0 \leq t \leq T} |(u - U_h)(t)|^2 \leq C \left[h^{2r} + \sum_{j=0}^{N-1} k_j^{2q+1} \right]. \quad (6.13)$$

To ensure that $U_h(t) \in \tilde{T}_u, t \in [0, T]$, as before we distinguish the cases: $\nu = 1, \nu = 2$ and $\nu = 3$. It is to be noted that in the continuous Galerkin case we do not gain at once control of the norm $\int_0^T \|\Theta(t)\|^2 dt$, cf. Remark 5.1. Thus alternatively we shall use the inverse inequality

$$\|\chi\|_{L^\infty} \leq C\underline{h}^{-\nu/2}|\chi| \quad \forall \chi \in V_h,$$

concluding that the result (6.13) holds, provided that $k^{2q}\underline{h}^{-\nu}$ and $h^{2r}\underline{h}^{-\nu}$ are sufficiently small. The derivation of estimates for $\int_0^T \|\Theta(t)\|^2 dt$ in the fully discrete case that will lead to milder mesh conditions, is of course a feasible task but we will not insist on it in the present work.

A. SECOND PROOF OF LEMMA 2.1

Since $p\tilde{p}'$ is integrated exactly by the Radau quadrature rule, we have

$$\int_0^1 p\tilde{p}' dt = \sum_{i=1}^q w_i p(\tau_i) \tilde{p}'(\tau_i). \quad (A.1)$$

Obviously, with $\ell_1, \dots, \ell_q \in \mathbb{P}_{q-1}$ denoting the Lagrange polynomials associated with τ_1, \dots, τ_q ,

$$\tilde{p} = \sum_{j=1}^q \varphi(\tau_j) \ell_j,$$

and (A.1) yields

$$\int_0^1 p\tilde{p}' dt = \sum_{i,j=1}^q w_i \ell_j'(\tau_i) \tau_i \varphi(\tau_i) \varphi(\tau_j).$$

Therefore,

$$\int_0^1 p\tilde{p}' dt = \sum_{i,j=1}^q w_i \ell_j'(\tau_i) \ell_i(\tau_i) \tau_i \varphi(\tau_i) \varphi(\tau_j),$$

i.e.,

$$\int_0^1 p\tilde{p}' dt = \sum_{i=1}^q w_i \ell_i'(\tau_i) \ell_i(\tau_i) \tau_i |\varphi(\tau_i)|^2 + \sum_{1 \leq i < j \leq q} [w_i \ell_j'(\tau_i) \ell_i(\tau_i) \tau_i + w_j \ell_i'(\tau_j) \ell_j(\tau_j) \tau_j] \varphi(\tau_i) \varphi(\tau_j).$$

Now, for $1 \leq i < j \leq q$,

$$w_i \ell_j'(\tau_i) \ell_i(\tau_i) \tau_i + w_j \ell_i'(\tau_j) \ell_j(\tau_j) \tau_j = \int_0^1 [\ell_j'(t) \ell_i(t) + \ell_j(t) \ell_i'(t)] t dt = \ell_i(1) \ell_j(1) - \int_0^1 \ell_i(t) \ell_j(t) dt = 0,$$

and the second sum vanishes. Similarly,

$$w_i \ell_i'(\tau_i) \ell_i(\tau_i) \tau_i = \frac{1}{2} [\ell_i(1) - w_i],$$

and thus

$$\int_0^1 p\tilde{p}' dt = \frac{1}{2} \left[|\varphi(1)|^2 - \sum_{i=1}^q w_i |\varphi(\tau_i)|^2 \right],$$

and (2.4) follows. Further,

$$\int_0^1 p' \tilde{p} dt + p(0) \tilde{p}(0) \geq \frac{1}{2} \left[|p(1)|^2 + \sum_{i=1}^q w_i |p(\tau_i)|^2 \right] = \frac{1}{2} \left[|p(1)|^2 + \int_0^1 |p(t)|^2 dt \right]. \quad \square$$

Acknowledgements. We would like to thank an anonymous referee and Y.-S. Smyrlis for useful suggestions.

REFERENCES

- [1] G. Akrivis and M. Crouzeix, Linearly implicit methods for nonlinear parabolic equations. *Math. Comp.* **73** (2004) 613–635.
- [2] G. Akrivis and C. Makridakis, *Convergence of a time discrete Galerkin method for semilinear parabolic equations*. Preprint (2002).
- [3] G. Akrivis, M. Crouzeix and C. Makridakis, Implicit-explicit multistep finite element methods for nonlinear parabolic problems. *Math. Comp.* **67** (1998) 457–477.
- [4] G. Akrivis, M. Crouzeix and C. Makridakis, Implicit-explicit multistep methods for quasilinear parabolic equations. *Numer. Math.* **82** (1999) 521–541.
- [5] A.K. Aziz and P. Monk, Continuous finite elements in space and time for the heat equation. *Math. Comp.* **52** (1989) 255–274.
- [6] J.H. Bramble and P.H. Sammon, Efficient higher order single step methods for parabolic problems: Part I, *Math. Comp.* **35** (1980) 655–677.
- [7] K. Eriksson and C. Johnson, Adaptive finite element methods for parabolic problems. I. A linear model problem. *SIAM J. Numer. Anal.* **28** (1991) 43–77.

- [8] K. Eriksson and C. Johnson, Adaptive finite element methods for parabolic problems. IV. Nonlinear problems. *SIAM J. Numer. Anal.* **32** (1995) 1729–1749.
- [9] K. Eriksson, C. Johnson and S. Larsson, Adaptive finite element methods for parabolic problems. VI. Analytic semigroups. *SIAM J. Numer. Anal.* **35** (1998) 1315–1325.
- [10] D. Estep and S. Larsson, The discontinuous Galerkin method for semilinear parabolic problems. *RAIRO Modél. Math. Anal. Numér.* **27** (1993) 35–54.
- [11] P. Jamet, Galerkin-type approximations which are discontinuous in time for parabolic equations in a variable domain. *SIAM J. Numer. Anal.* **15** (1978) 912–928.
- [12] O. Karakashian and C. Makridakis, A space-time finite element method for the nonlinear Schrödinger equation: the discontinuous Galerkin method. *Math. Comp.* **67** (1998) 479–499.
- [13] O. Karakashian and C. Makridakis, A space-time finite element method for the nonlinear Schrödinger equation: the continuous Galerkin method, *SIAM J. Numer. Anal.* **36** (1999) 1779–1807.
- [14] O. Karakashian and C. Makridakis, Convergence of a continuous Galerkin method with mesh modification for nonlinear wave equations. *Math. Comp.* (to appear).
- [15] C. Makridakis and I. Babuška, On the stability of the discontinuous Galerkin method for the heat equation. *SIAM J. Numer. Anal.* **34** (1997) 389–401.
- [16] C. Makridakis and R.H. Nochetto, *A posteriori error estimates for a class of dissipative schemes for nonlinear evolution equations*. Preprint (2002).
- [17] A.H. Schatz and L.B. Wahlbin, Interior maximum-norm estimates for finite element methods: Part II. *Math. Comp.* **64** (1995) 907–928.
- [18] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*. Springer-Verlag, Berlin (1997).