Numer. Math. 20, 213–237 (1973)

Galerkin Methods for Parabolic Equations with Nonlinear Boundary Conditions

Jim Douglas, Jr., and Todd Dupont*

University of Chicago

Received April 18, 1972

Abstract. A variety of Galerkin methods are studied for the parabolic equation $u_t = \nabla \cdot \big(a(x)\nabla u\big)$, $x \in \Omega \subset \mathbb{R}^n$, $t \in (0, T]$, subject to the nonlinear boundary condition $u_\nu = g(x, t, u)$, $x \in \partial\Omega$, $t \in (0, T]$ and the usual initial condition. Optimal order error estimates are derived both in $L^2(\Omega)$ and $H^1(\Omega)$ norms for all methods treated, including several that produce linear computational procedures.

1. Introduction

We shall study the numerical solution of the parabolic problem

$$\begin{aligned}
&\text{(a)} \quad \frac{\partial u}{\partial t} = \nabla \cdot (a \nabla u), && x \in \Omega,\ 0 < t \leqq T,\\
&\text{(b)} \quad u(x, 0) = f(x), && x \in \Omega,\ t = 0, \qquad (1.1)\\
&\text{(c)} \quad a\frac{\partial u}{\partial \nu} = g(x, t, u), && x \in \partial\Omega,\ 0 < t \leqq T,
\end{aligned}$$

where Ω is a bounded domain in $\mathbb{R}^n$ and $\partial/\partial\nu$ denotes the exterior normal derivative, by a number of Galerkin methods. The primary object of this paper is the treatment of the nonlinear boundary condition (1.1c); consequently, we shall isolate its effect by taking a very simple differential equation. We shall assume that

$$a = a(x), \quad x \in \Omega, \quad 0 \leqq t \leqq T, \tag{1.2}$$

throughout this paper. The more general nonlinear problem in which (1.1a) is replaced by

$$\frac{\partial}{\partial t} c(x, t, u) = \sum_i \frac{\partial}{\partial x_i} a_i(x, t, u, \nabla u) + b(x, t, u, \nabla u) \tag{1.1a$'$}$$

will be treated in another paper by H. H. Rachford, Jr., and the present authors. There are sufficient complications in the simpler problem to justify its separate presentation.

Let (only real functions will arise)

$$\langle v, w\rangle = \int_\Omega v w\, dx, \qquad \langle v, w\rangle_{\partial\Omega} = \int_{\partial\Omega} v w\, d\sigma. \tag{1.3}$$

* The authors were partially supported by The National Science Foundation during the preparation of this paper.

Let $H^1(\Omega)=\{v\in L^2(\Omega)\,|\,\partial v/\partial x_i\in L^2(\Omega),\ i=1,\dots,n\}$ and norm it in the usual way [7]:

$$\|v\|^2_{H^1(\Omega)}=\|v\|^2_{L^2(\Omega)}+\sum_{i=1}^n\left\|\frac{\partial v}{\partial x_i}\right\|^2_{L^2(\Omega)}=\|u\|^2_{L^2(\Omega)}+\|\nabla v\|^2_{L^2(\Omega)}. \tag{1.4}$$

Let $L^p(0,T;X)$ denote those vector-valued maps of $[0,T]$ into X such that

$$\|v\|^p_{L^p(0,T;X)}=\int_0^T\|v(t)\|^p_X\,dt<\infty,\qquad 1\leqq p<\infty, \tag{1.5}$$

and $L^\infty(0,T;X)$ those maps such that

$$\|v\|_{L^\infty(0,T;X)}=\sup_{0<t<T}\|v(t)\|_X<\infty. \tag{1.6}$$

Denote by

$$\|v\|_{H^1_0(\Omega)}=\|\nabla v\|_{L^2(\Omega)},\qquad v\in H^1(\Omega), \tag{1.7}$$

the seminorm on $H^1(\Omega)$ that is a norm, equivalent to the H^1-norm, on the subspace $H^1_0(\Omega)$ of $H^1(\Omega)$ obtained by closing $C_0^\infty(\Omega)$ in $H^1(\Omega)$. Adopt the definitions of Lions-Magenes [7] for $H^s(\Omega)$ and $H^s(\partial\Omega)$, s real, whenever these constructions can be carried out. Since we shall not use very large or small values of s, it is not necessary that Ω have a C^∞ boundary. Extend the inner product notations of (1.3) to represent the duality between H^s and $(H^s)'$ for both Ω and $\partial\Omega$. If $\Omega=(a,b)\subset\mathbb{R}$, $H^s(\partial\Omega)\cong L^\infty(\partial\Omega)$ for all s.

Some properties of Ω and the coefficient $a(x)$ will be critical in our convergence analyses below. Certain trace theorems will be vital both in the proofs and, surprisingly, in the actual choice of the computationally more efficient discrete time Galerkin methods. The two trace inequalities that we shall consider are the following:

There exists a constant C_T such that, for $0<\varepsilon\leqq 1$

$$\|v\|_{L^2(\partial\Omega)}\leqq C_T[\varepsilon\|v\|_{H^1_0(\Omega)}+\varepsilon^{-1}\|v\|_{L^2(\Omega)}],\qquad v\in H^1(\Omega). \tag{T1}$$

There exists a constant $C_{T,s}$ such that

$$\|v\|_{H^s(\partial\Omega)}\leqq C_{T,s}\|v\|_{H^{s+\frac12}(\Omega)},\qquad v\in H^{s+\frac12}(\Omega),\qquad 0<s\leqq\tfrac32,\qquad s\neq1. \tag{T2}$$

Both (T1) and (T2) are standard for domains with smooth boundaries [7], but it is also clear that they hold under less restrictive conditions. In particular, they hold for Ω a rectangular parallelopiped or an interval. Only (T1) will be required for the H^1 estimates.

Elliptic regularity plays an important role in obtaining optimal L^2 estimates. Consider the Neumann problem

$$\begin{aligned}\nabla\cdot\big(a(x)\nabla v\big)-\lambda v&=\varphi,\qquad x\in\Omega,\\ a\frac{\partial v}{\partial\nu}&=\psi,\qquad x\in\partial\Omega,\end{aligned} \tag{1.8}$$

where

$$0<m\leqq a(x)\leqq M<\infty \tag{1.9}$$

and λ is a sufficiently large positive constant. We shall at various times need the following a priori inequalities:

For $\psi = 0$, there exists a constant C such that

$$\|v\|_{H^2(\Omega)} \leqq C \|\varphi\|_{L^2(\Omega)}, \qquad \varphi \in L^2(\Omega). \tag{R1}$$

For $\varphi = 0$, there exists a constant C such that

$$\|v\|_{H^s(\Omega)} \leqq C \|\psi\|_{H^{s-\frac{3}{2}}(\partial\Omega)}, \qquad \psi \in H^{s-\frac{3}{2}}(\partial\Omega), \qquad s = 0, 1, 2. \tag{R2}$$

These results are well known for smooth $a(x)$ and smooth $\partial\Omega$; see Lions-Magenes [7]. Note that we must be able to define $H^s(\partial\Omega)$, $-\frac{3}{2} \leqq s \leqq \frac{3}{2}$, in order for (R2) to make sense. For $n = 1$ and $a(x)$ smooth, (R1) and (R2) are elementary. Elliptic regularity is not needed for the H^1 estimates.

Galerkin methods are based on approximability of functions in the solution space, $H^1(\Omega)$ in our case, by functions in conveniently chosen subspaces. Usually these subspaces are selected in some systematic fashion depending on some parameter, such as the spacing of nodes associated with the elements of a particular basis for a subspace. Bramble and Schatz [4] formulated a useful definition that isolated extractly the properties that we shall need in order to derive optimal L^2 estimates. Let $h > 0$ be a parameter and let $\mathcal{M}_h$ be a (finite-dimensional) subspace of $H^1(\Omega)$ associated with h. Then the family $\{\mathcal{M}_h\}$ is said to be an S_h^r family if the following inequalities hold.

For $0 \leqq s \leqq 1$ and $s \leqq q \leqq r$, there exists a constant C such that

$$\inf_{\chi \in \mathcal{M}_h} \|v - \chi\|_{H^s(\Omega)} \leqq C \|v\|_{H^q(\Omega)} h^{q-s}, \qquad v \in H^q(\Omega). \tag{S_h^r}$$

In fact, the above is a special case of their definition, since we are interested only in $s \in [0, 1]$. Essentially all of the standardly used subspaces satisfy this requirement for some choice of $r \geqq 2$. The Hermite spaces, the smooth spline spaces, and the spaces based on triangles all are S_h^r spaces, at least if some modest regularity is practiced in the choice of nodes [2, 3, 4, 5, 6].

In many discussions of Galerkin or other projection methods for approximating solutions of partial differential equations, a so-called "inverse hypothesis" has been used. Typical of these is the assumption that

$$\|v\|_{H^1(\Omega)} \leqq C h^{-1} \|v\|_{L^2(\Omega)}, \qquad v \in \mathcal{M}_h. \tag{$*$}$$

Babuška [1] has shown that for many domains Ω such that $\partial\Omega$ contains a smooth piece having non-zero curvature it is impossible to construct an S_h^r family with the subspace being a tensor product of spaces of piecewise polynomials having local bases such that ($*$) is satisfied. It is important to note that we do *not* assume ($*$) or any other such embedding result.

Let us turn to the formulation of the Galerkin methods that we shall consider for approximating the solution of (1.1). A weak form of (1.1) is as follows. Find a function $u(x, t)$ such that

$$\begin{gathered} u \in L^2(0, T; H^1(\Omega)), \qquad \frac{\partial u}{\partial t} \in L^2(0, T; H^1(\Omega)'), \\ \left\langle \frac{\partial u}{\partial t}, z \right\rangle + \langle a \nabla u, \nabla z \rangle - \langle g(u), z \rangle_{\partial\Omega} = 0, \qquad z \in H^1(\Omega), \qquad 0 < t \leqq T, \\ u(\cdot, 0) = f \in L^2(\Omega), \end{gathered} \tag{1.10}$$

where the explicit dependence of g on x and t has not been written for notational convenience. It is this weak form that we propose to approximate in the numerical procedures. Let

$$\mathcal{M} = \operatorname{Span}[v_1, v_2, \ldots, v_N] \subset H^1(\Omega)$$

and define the *continuous-time-Galerkin* method as follows. Find a function $U(x, t)$ such that

$$\begin{aligned} &U(\cdot, t) \in \mathcal{M}, \quad 0 \leqq t \leqq T, \\ &U(\cdot, 0) - f \quad \text{"small"}, \\ &\left\langle \frac{\partial U}{\partial t}, v \right\rangle + \langle a \nabla U, \nabla v \rangle - \langle g(U), v \rangle_{\partial\Omega} = 0, \quad v \in \mathcal{M}, \quad 0 < t \leqq T. \end{aligned} \tag{1.11}$$

The specification of U at time $t = 0$ is obviously vague; the usual ways to obtain $U(\cdot, 0)$ are interpolation of f into $\mathcal{M}$, L^2-projection of f into $\mathcal{M}$, and H^1-projection. We shall not be specific at the moment, since the various estimates that we shall derive below require different approximations of the initial condition.

We shall derive both H^1 and L^2 estimates for $u - U$ in the sections to follow. Under various hypotheses we shall see that

$$\|u - U\|_{L^2(0, T; H^1(\Omega))} = O(h^{r-1})$$

and

$$\|u - U\|_{L^\infty(0, T; L^2(\Omega))} = O(h^r)$$

if $\mathcal{M} = \mathcal{M}_h$ comes from an S_h^r family, u is sufficiently smooth, and $U(x, 0)$ is chosen reasonably.

The Eq. (1.11) is a system of ordinary differential equations, nonlinear if g is nonlinear in u, for the coefficients of the basis elements $v_i(x)$. As such, it is not usually solvable directly in any form that is numerically usable; hence, we shall introduce several methods of differencing in the time variable to produce solvable algebraic problems at each time level. Let $v^n = v(t^n) = v(n\Delta t)$ and set

$$v^{n+\frac12} = \tfrac12 (v^n + v^{n+1}). \tag{1.12}$$

Note that $v^{n+\frac12} \neq v(t^{n+\frac12})$ in general. The simplest differencing leads to the *Crank-Nicolson-Galerkin* procedure given by

$$\langle d_t U^n, v \rangle + \langle a \nabla U^{n+\frac12}, \nabla v \rangle - \langle g(t^{n+\frac12}, U^{n+\frac12}), v \rangle_{\partial\Omega} = 0, \quad v \in \mathcal{M}, \tag{1.13}$$

where

$$d_t U^n = (U^{n+1} - U^n)/\Delta t. \tag{1.14}$$

Note that the time is actually inserted at $(n + \frac12)\Delta t$, but the average of U^n and U^{n+1} is used in the evaluation of g. This equation is centered on time at $t^{n+\frac12}$, and it is easy to see that it is locally second order correct in the time step. We shall show that, under sufficient hypotheses,

$$\left(\sum_n \|u^{n+\frac12} - U^{n+\frac12}\|^2_{H^1(\Omega)} \Delta t \right)^{\frac12} = O\left(h^{r-1} + (\Delta t)^2\right)$$

and

$$\max_n \|u^n - U^n\|_{L^2(\Omega)} = O\left(h^r + (\Delta t)^2\right).$$

Note that the H^1 estimate concerns the *averaged* solution. It indicates that the numerical solution is in general better at the times $t^{n+\frac12}$ than it is at the time t^n.

This to a large extent explains the famous Crank-Nicolson "bounce" that engineers have observed in printing out answers at the time levels rather than at average time levels. (Forget the fact that they usually employ finite-difference methods; analogous results can be derived in the finite-difference case.) It is also very important to take account of the nature of the H^1 estimate when modifying the Crank-Nicolson procedure to give algebraically more efficient procedures.

The algebraic system generated by the Crank-Nicolson method at each time step is nonlinear if g is nonlinear in u. Thus, it seems attractive to modify the evaluation of the boundary term so as to obtain a linear algebraic system at each time step. The simplest way which maintains local second order accuracy to do is to extrapolate using U^n and U^{n-1} to approximate $U^{n+\frac{1}{2}}$, since

$$u^{n+\frac{1}{2}} = \tfrac{3}{2} u^n - \tfrac{1}{2} u^{n-1} + O\big((\Delta t)^2\big).$$

However, it is clear that if this extrapolation is chosen, there will be terms such as $\|u^n - U^n\|_{L^2(\partial\Omega)}$ arising in the error analysis. The trace inequality (T1) is sharp in the sence that it is false if the $H_0^1(\Omega)$ term is omitted; hence, there is no place to hide the $L^2(\partial\Omega)$ terms at the time levels using the positive definite terms related to Ω. This indicates very strongly that any linearization of the boundary term should be done using only the values $U^{n-\frac{1}{2}}$, $U^{n-\frac{3}{2}}$, etc. We shall write down three such schemes out of the large set of possibilities and shall analyze them as typical.

The extrapolation of the boundary values mentioned above can be changed to $u^{n+\frac{1}{2}} = 2u^{n-\frac{1}{2}} - u^{n-\frac{3}{2}} + O\big((\Delta t)^2\big)$, where simply for notational convenience we restrict ourselves to constant time steps. This leads to the *Extrapolated-Crank-Nicolson* procedure

$$\langle d_t U^n, v\rangle + \langle a \nabla U^{n+\frac{1}{2}}, \nabla v\rangle - \langle g(t^{n+\frac{1}{2}}, 2U^{n-\frac{1}{2}} - U^{n-\frac{3}{2}}), v\rangle_{\partial\Omega} = 0, \quad v \in \mathcal{M}. \tag{1.15}$$

The second method is based on linearizing $g(u)$ about $u^{n-\frac{1}{2}}$:

$$g(u^{n+\frac{1}{2}}) = g(u^{n-\frac{1}{2}}) + \frac{\partial g}{\partial u}(u^{n-\frac{1}{2}})(u^{n+\frac{1}{2}} - u^{n-\frac{1}{2}}) + O\big((\Delta t)^2\big).$$

Thus, we can define the *Linearized-about-$u^{n-\frac{1}{2}}$-Crank-Nicolson* procedure as the following:

$$\begin{aligned}&\langle d_t U^n, v\rangle + \langle a \nabla U^{n+\frac{1}{2}}, \nabla v\rangle \\ &\qquad - \langle g(t^{n+\frac{1}{2}}, U^{n-\frac{1}{2}}) + \frac{\partial g}{\partial u}(t^{n+\frac{1}{2}}, U^{n-\frac{1}{2}})(U^{n+\frac{1}{2}} - U^{n-\frac{1}{2}}), v\rangle_{\partial\Omega} \\ &\quad = 0, \ v \in \mathcal{M}.\end{aligned} \tag{1.16}$$

Finally, $g(u)$ can be linearized about the extrapolation:

$$g(u^{n+\frac{1}{2}}) = g(2u^{n-\frac{1}{2}} - u^{n-\frac{3}{2}}) + \frac{\partial g}{\partial u}(2u^{n-\frac{1}{2}} - u^{n-\frac{3}{2}})(u^{n+\frac{1}{2}} - 2u^{n-\frac{1}{2}} + u^{n-\frac{3}{2}}) + O\big((\Delta t)^4\big).$$

Thus, a *Linearized-Extrapolated-Crank-Nicolson* equation results:

$$\begin{aligned}&\langle d_t U^n, v\rangle + \langle a \nabla U^{n+\frac{1}{2}}, \nabla v\rangle \\ &\qquad - \Big\langle g(t^{n+\frac{1}{2}}, E^{n+\frac{1}{2}}) + \frac{\partial g}{\partial u}(t^{n+\frac{1}{2}}, E^{n+\frac{1}{2}})(U^{n+\frac{1}{2}} - E^{n+\frac{1}{2}}), v\Big\rangle_{\partial\Omega} = 0, \\ &\qquad\qquad v \in \mathcal{M},\end{aligned} \tag{1.17}$$

where

$$E^{n+\frac{1}{2}} = 2U^{n-\frac{1}{2}} - U^{n-\frac{3}{2}}. \tag{1.18}$$

All three methods produce linear algebraic problems at each time level. Note that each requires a separate start-up procedure, since (1.15) and (1.17) are defined only for $n \geqq 2$ and (1.16) for $n \geqq 1$. We shall indicate in the proofs that care must be exercised in establishing that these initial values can be obtained in a reasonable way so that the inherent accuracy can be preserved. A method will be suggested to accomplish the desired initialization, and then it will be shown that the same error estimates can be demonstrated as for the Crank-Nicolson scheme.

Throughout the paper the function g will be assumed to have a bounded derivative with respect to u:

$$\left|\frac{\partial g}{\partial u}(x, t, u)\right| \leqq K, \quad x \in \partial\Omega, \quad 0 \leqq t \leqq T, \quad -\infty < u < \infty. \tag{1.19}$$

Frequently, it will be assumed much smoother. However, we do *not* make an assumption on the sign of $\partial g/\partial u$; i.e., we do not need to assume the stability of the associated steady-state problem.

2. A Nonlinear H^1 Projection of the Solution

It is very convenient to make a preliminary study of a particular nonlinear projection of the solution u of (1.10) into the subspace $\mathcal{M}$ of $H^1(\Omega)$ in which the solution of (1.11) lies. Let the function $W(\cdot, t) \in \mathcal{M}$ be defined by the relation

$$\begin{gathered}\langle a\nabla(u-W), \nabla v\rangle + \lambda\langle u-W, v\rangle - \langle g(u)-g(W), v\rangle_{\partial\Omega} = 0, \quad v \in \mathcal{M}, \\ 0 \leqq t \leqq T,\end{gathered} \tag{2.1}$$

where λ is a positive constant to be fixed sufficiently large that existence and uniqueness of W are assured. We shall set

$$\eta = u - W \tag{2.2}$$

and shall derive estimates in both $H^1(\Omega)$ and $L^2(\Omega)$ for η and $\partial\eta/\partial t$. These estimates will be absolutely fundamental in our analysis of the convergence of the Galerkin approximations to u. Our method of analysis is related to earlier ones by Wheeler [10] and Fix and Strang [6] in the U and W are compared.

The $H^1(\Omega)$ estimates are much easier to come by than the $L^2(\Omega)$ ones; thus, it is reasonable to start with them. Rewrite Eq. (2.1) in the form

$$\langle a\nabla\eta, \nabla v\rangle + \lambda\langle\eta, v\rangle - \langle g(u)-g(W), v\rangle_{\partial\Omega} = 0, \quad v \in \mathcal{M},$$

and take $v = \tilde{u} - W = (\tilde{u} - u) + \eta = \delta u + \eta$, $\tilde{u} \in \mathcal{M}$. Then (1.9) and (1.19) imply that

$$\begin{aligned} m\|\eta\|^2_{H_0^1(\Omega)} + \lambda\|\eta\|^2_{L^2(\Omega)} &\leqq -\langle a\nabla\eta, \nabla\delta u\rangle - \lambda\langle n, \delta u\rangle + \langle g(u)-g(W), \eta+\delta u\rangle_{\partial\Omega} \\ &\leqq \frac{m}{4}\|\eta\|^2_{H_0^1(\Omega)} + \frac{\lambda}{2}\|\eta\|^2_{L^2(\Omega)} + C\|\delta u\|^2_{H^1(\Omega)} \\ &\quad + K\left[\frac{3}{2}\|\eta\|^2_{L^2(\partial\Omega)} + \frac{1}{2}\|\delta u\|^2_{L^2(\partial\Omega)}\right]. \end{aligned}$$

Assume that the trace inequality (T1) holds for Ω. Then,

$$\frac{3}{2} K\|\eta\|^2_{L^2(\partial\Omega)} \leqq 3KC_T^2[\varepsilon\|\eta\|^2_{H_0^1(\Omega)} + \varepsilon^{-1}\|\eta\|^2_{L^2(\Omega)}] = \frac{m}{4}\|\eta\|^2_{H_0^1(\Omega)} + 36K^2C_T^4 m^{-1}\|\eta\|^2_{L^2(\Omega)}$$

for the choice $\varepsilon = m/(12KC_T^2)$. Hence, if

$$\lambda \geqq m + 72K_T^2C_T^4 m^{-1}, \tag{2.3}$$

it follows that

$$\|\eta\|_{H^1(\Omega)} \leqq C[\|\delta u\|_{H^1(\Omega)} + \|\delta u\|_{L^2(\partial\Omega)}].$$

Since $\tilde{u}$ was an arbitrary element of $\mathscr{M}$.

$$\begin{aligned}\|\eta\|_{H^1(\Omega)} &\leqq C \inf_{\chi\in\mathscr{M}} [\|u-\chi\|_{H^1(\Omega)} + \|u-\chi\|_{L^2(\partial\Omega)}] \\ &\leqq C' \inf_{\chi\in\mathscr{M}} \|u-\chi\|_{H^1(\Omega)}, \quad 0\leqq t\leqq T.\end{aligned} \tag{2.4}$$

In particular, (2.3) is sufficient to insure existence and uniqueness of W, given the hypotheses (1.9) and (1. 19).

The estimate (2.4) required nothing on the subspace $\mathscr{M}$ and is essentially the best possible $H^1(\Omega)$ estimate, since only C' is subject to improvement. If $\mathscr{M} = \mathscr{M}_h$ is an S_h^r space, then, for $0\leqq t\leqq T$,

$$\|\eta\|_{H^1(\Omega)} \leqq C\|u\|_{H^k(\Omega)} h^{k-1}, \quad 1\leqq k\leqq r. \tag{2.5}$$

Eq. (2.1) can be differentiated with respect to time to give

$$\begin{aligned}&\left\langle a\nabla\frac{\partial\eta}{\partial t}, \nabla v\right\rangle + \lambda\left\langle\frac{\partial\eta}{\partial t}, v\right\rangle - \left\langle\frac{\partial g}{\partial u}(W)\frac{\partial\eta}{\partial t}, v\right\rangle_{\partial\Omega} \\ &\quad = \left\langle\left\{\frac{\partial g}{\partial t}(u) - \frac{\partial g}{\partial t}(W)\right\} + \frac{\partial u}{\partial t}\left\{\frac{\partial g}{\partial u}(u) - \frac{\partial g}{\partial u}(W)\right\}, v\right\rangle_{\partial\Omega}, \quad v\in\mathscr{M}.\end{aligned} \tag{2.6}$$

Assume that $\partial u/\partial t \in L^\infty(0, T; L^\infty(\partial\Omega))$ and that

$$\left|\frac{\partial^2 g}{\partial t\,\partial u}\right|, \quad \left|\frac{\partial^2 g}{\partial u^2}\right| \leqq K_1 < \infty. \tag{2.7}$$

Then, essentially the same argument as above shows that

$$\left\|\frac{\partial\eta}{\partial t}\right\|_{H^1(\Omega)} \leqq C' \inf_{\chi\in\mathscr{M}} \left\|\frac{\partial u}{\partial t} - \chi\right\|_{H^1(\Omega)} + C''\|\eta\|_{L^2(\partial\Omega)}, \quad 0<t\leqq T, \tag{2.8}$$

provided that (2.3) holds. Again, if $\mathscr{M} = \mathscr{M}_h$ is an S_h^r space,

$$\left\|\frac{\partial\eta}{\partial t}\right\|_{H^1(\Omega)} \leqq C\left[\|u\|_{H^k(\Omega)} + \left\|\frac{\partial u}{\partial t}\right\|_{H^k(\Omega)}\right] h^{k-1}, \quad 1\leqq k\leqq r. \tag{2.9}$$

The above results can be summarized in the following theorem.

Theorem 2.1. Let $\eta = u - W$ be defined by (2.1) and assume that (2.3) is valid. Assume that Ω is such that the trace inequality (T1) holds and assume that the function g has its derivatives $\partial^2 g/\partial t\,\partial u$ and $\partial^2 g/\partial u^2$ bounded. Then, if $\mathscr{M} = \mathscr{M}_h$ is taken from an S_h^r family of subspaces of $H^1(\Omega)$,

$$\|\eta\|_{L^2(0,T;H^1(\Omega))} \leqq C\|u\|_{L^2(0,T;H^k(\Omega))} h^{k-1}, \quad 1\leqq k\leqq r,$$

and

$$\left\| \frac{\partial \eta}{\partial t} \right\|_{L^2(0,T;H^1(\Omega))} \leqq C' \left(\|u\|_{L^2(0,T;H^k(\Omega))} + \left\| \frac{\partial u}{\partial t} \right\|_{L^2(0,T;H^k(\Omega))} \right) h^{k-1}, \qquad 1 \leqq k \leqq r,$$

where C and C' are independent of h but C' depends on a bound for $\partial u/\partial t$ on $\partial\Omega \times [0, T]$.

The bounding of η and $\partial\eta/\partial t$ in $L^2(\Omega)$ is a more delicate problem. These bounds will be obtained using a line of argument introduced by Nitsche [8, 9], although our versions vary considerably from his.

Let $\alpha = \alpha(t) \in H^1(\Omega)$ satisfy

$$\langle a\nabla\alpha, \nabla z\rangle + \lambda\langle\alpha, z\rangle - \langle G\alpha, z\rangle_{\partial\Omega} = \langle\eta, z\rangle, \qquad z \in H^1(\Omega), \tag{2.10}$$

where λ agrees with its previous value and the function G is defined by

$$G(x,t) = \int_0^1 \frac{\partial g}{\partial u}\left(x, t, \theta u(x,t) + (1-\theta) W(x,t)\right) d\theta. \tag{2.11}$$

Note first that $|G(x,t)| \leqq K$, the same bound as for $\partial g/\partial u$. The value of λ assures the existence and uniqueness of α. Now, the choice $z = \eta$ in (2.10) and the definition of λ lead to the following:

$$\begin{aligned}
\|\eta\|^2_{L^2(\Omega)} &= \langle a\nabla\alpha, \nabla\eta\rangle + \lambda\langle\alpha, \eta\rangle - \langle G\alpha, \eta\rangle_{\partial\Omega} \\
&= \langle a\nabla\eta, \nabla(\alpha-\chi)\rangle + \lambda\langle\eta, \alpha-\chi\rangle - \langle G\eta, \alpha\rangle + \langle g(u) - g(W), \chi\rangle_{\partial\Omega} \\
&= \langle a\nabla\eta, \nabla(\alpha-\chi)\rangle + \lambda\langle\eta, \alpha-\chi\rangle - \langle g(u) - g(W), \alpha-\chi\rangle_{\partial\Omega} \\
&\leqq C\|\eta\|_{H^1(\Omega)} \|\alpha-\chi\|_{H^1(\Omega)}, \qquad \chi \in \mathcal{M}.
\end{aligned}$$

Assume from now on in this section that $\mathcal{M} = \mathcal{M}_h$ is drawn from an S_h^r family. Then

$$\begin{aligned}
\|\eta\|^2_{L^2(\Omega)} &\leqq C\|\eta\|_{H^1(\Omega)} \inf_{\chi\in\mathcal{M}_h} \|\alpha-\chi\|_{H^1(\Omega)} \\
&\leqq C\|\eta\|_{H^1(\Omega)} \|\alpha\|_{H^2(\Omega)} h,
\end{aligned} \tag{2.12}$$

by (S_h^r). If the elliptic regularity hypotheses (R1) and (R2) hold for (2.10), then

$$\|\alpha\|_{H^2(\Omega)} \leqq C\left[\|\eta\|_{L^2(\Omega)} + \|G\alpha\|_{H^{\frac{1}{2}}(\partial\Omega)}\right]. \tag{2.13}$$

It is convenient at this point to consider the following lemma that allows us to consider pointwise multiplication on $H^{\frac{1}{2}}(\partial\Omega)$. Also, assume $\dim(\Omega) \leqq 3$.

Lemma 2.2. Let $\dim(\Omega) \leqq 3$ and assume that (T2) holds. Let $F \in H^{\frac{3}{2}+\varepsilon}(\Omega)$, some $\varepsilon > 0$, and $G \in H^{\frac{1}{2}}(\partial\Omega)$. Then $FG \in H^{\frac{1}{2}}(\partial\Omega)$ and

$$\|FG\|_{H^{\frac{1}{2}}(\partial\Omega)} \leqq C_\varepsilon \|F\|_{H^{\frac{3}{2}+\varepsilon}(\Omega)} \|G\|_{H^{\frac{1}{2}}(\partial\Omega)}.$$

Proof. Since $\dim(\Omega) \leqq 3$,

$$\|F\|_{L^\infty(\Omega)} \leqq C\|F\|_{H^{\frac{3}{2}+\varepsilon}(\Omega)}.$$

Hence, if $G \in L^2(\partial\Omega)$,

$$\|FG\|_{L^2(\partial\Omega)} \leqq C\|F\|_{H^{\frac{3}{2}+\varepsilon}(\Omega)} \|G\|_{L^2(\partial\Omega)}.$$

If $G \in H^1(\partial\Omega)$, then $\nabla(FG) = F\nabla G + G\nabla F$ and

$$\|F\nabla G\|_{L^2(\partial\Omega)} \leqq C\|F\|_{H^{\frac{3}{2}+\varepsilon}(\Omega)}\|G\|_{H^1(\partial\Omega)}.$$

Since $F \in H^{1+\varepsilon}(\partial\Omega)$ and $\dim(\partial\Omega) \leqq 2$, $\nabla F \in L^{2+\delta}(\partial\Omega)$, $\delta = \delta(\varepsilon) > 0$, and $G \in L^p(\partial\Omega)$, any $p < \infty$. Hence,

$$\|G\nabla F\|_{L^2(\partial\Omega)} \leqq C\|F\|_{H^{\frac{3}{2}+\varepsilon}(\Omega)}\|G\|_{H^1(\partial\Omega)}.$$

Thus, it follows from standard interpolation theorems [7] that $FG \in H^{\frac{1}{2}}(\partial\Omega)$ if $G \in H^{\frac{1}{2}}(\partial\Omega)$ and that

$$\|FG\|_{H^{\frac{1}{2}}(\partial\Omega)} \leqq C(\varepsilon)\|F\|_{H^{\frac{3}{2}+\varepsilon}(\Omega)}\|G\|_{H^{\frac{1}{2}}(\partial\Omega)}.$$

Note that it is a trivial consequence of the compactness of the injection of $H^2(\Omega)$ into $H^{\frac{3}{2}+\varepsilon}(\Omega)$, $\varepsilon < \frac{1}{2}$, that

$$\|FG\|_{H^{\frac{1}{2}}(\partial\Omega)} \leqq \delta\|F\|_{H^2(\Omega)} + C(\varepsilon, \delta, \|G\|_{H^{\frac{1}{2}}(\partial\Omega)})\|F\|_{L^2(\Omega)}.$$

Let us apply the above inequality to $G\alpha$, where G is given by (2.11). Assume that g can be extended to the closure of Ω. Then it is sufficient to show that $G \in H^1(\Omega)$, boundedly for $t \in [0, T]$. First, $G \in L^2(\Omega)$, since it is bounded. It is an easy calculation to see that

$$\frac{\partial G}{\partial x_i}(x,t) = \int_0^1 \left[\frac{\partial^2 g}{\partial x_i \partial u}(x,t,\mathscr{D}) + \left\{\theta\frac{\partial u}{\partial x_i}(x,t) + (1-\theta)\frac{\partial W}{\partial x_i}(x,t)\right\}\frac{\partial^2 g}{\partial u^2}(x,t,\mathscr{D})\right] d\theta,$$

where $\mathscr{D} = \theta u(x,t) + (1-\theta)W(x,t)$. Hence, if $g \in C^2(\bar{\Omega} \times [0,T] \times \mathbb{R})$ with each of the derivatives appearing above being bounded, then (2.4), taken with $\chi = 0$, implies that

$$\|G\|_{L^\infty(0,T;H^{\frac{1}{2}}(\partial\Omega))} \leqq C(\|u\|_{L^\infty(0,T;H^1(\Omega))} + 1). \tag{2.14}$$

It follows from (2.13), (2.14), and Lemma 2.2 that

$$\|\alpha\|_{H^2(\Omega)} \leqq C\|\eta\|_{L^2(\Omega)}.$$

Thus, it follows from (2.12) that

$$\|\eta\|_{L^2(\Omega)} \leqq C\|\eta\|_{H^1(\Omega)}h \leqq C(1 + \|u\|_{H^1(\Omega)})\|u\|_{H^k(\Omega)}h^k, \qquad 1 \leqq k \leqq r. \tag{2.15}$$

Now consider an $L^2(\Omega)$ bound for $\partial\eta/\partial t$. For $t \in (0, T]$, let $\varphi \in H^1(\Omega)$ be the solution of

$$\begin{aligned} &\langle a\nabla\varphi, \nabla z\rangle + \lambda\langle\varphi, z\rangle - \left\langle\frac{\partial g}{\partial u}(W)\varphi, z\right\rangle_{\partial\Omega} \\ &\quad = \left\langle\left\{\frac{\partial g}{\partial t}(u) - \frac{\partial g}{\partial t}(W)\right\} + \frac{\partial u}{\partial t}\left\{\frac{\partial g}{\partial u}(u) - \frac{\partial g}{\partial u}(W)\right\}, z\right\rangle_{\partial\Omega}, \quad z \in H^1(\Omega). \end{aligned} \tag{2.16}$$

Thus, φ is the weak solution of

$$-\nabla\cdot(a\nabla\varphi) + \lambda\varphi = 0, \qquad x \in \Omega,$$

$$a\frac{\partial\varphi}{\partial\nu} - b\varphi = \gamma, \qquad x \in \partial\Omega,$$

where

$$b = \frac{\partial g}{\partial u}(W) \in L^\infty\left(0, T; H^{\frac{1}{2}}(\partial\Omega)\right),$$

$$\gamma = \frac{\partial g}{\partial t}(u) - \frac{\partial g}{\partial t}(W) + \frac{\partial u}{\partial t}\left[\frac{\partial g}{\partial u}(u) - \frac{\partial g}{\partial u}(W)\right].$$

Note that the function is actually in $H^2(\Omega)$ for smooth u. We can consider the adjoint problem

$$-\nabla \cdot (a \nabla \psi) + \lambda \psi = \varphi, \qquad x \in \Omega,$$

$$a\frac{\partial \psi}{\partial \nu} - b\psi = 0, \qquad x \in \partial\Omega.$$

Then,

$$\begin{aligned}
\|\varphi\|^2_{L^2(\Omega)} &= \langle \varphi, -\nabla \cdot (a\nabla\psi) + \lambda\psi \rangle \\
&= \langle -\nabla \cdot (a \nabla \varphi) + \lambda \varphi, \psi \rangle + \left\langle a\frac{\partial \varphi}{\partial \nu}, \psi \right\rangle_{\partial\Omega} - \left\langle \varphi, a\frac{\partial \psi}{\partial \nu} \right\rangle_{\partial\Omega} \\
&= \langle \gamma, \psi \rangle_{\partial\Omega} \\
&\leq \|\gamma\|_{H^{-\frac{3}{2}}(\partial\Omega)} \|\psi\|_{H^{\frac{3}{2}}(\partial\Omega)} \\
&\leq C\|\gamma\|_{H^{-\frac{3}{2}}(\partial\Omega)} \|\psi\|_{H^2(\Omega)}.
\end{aligned}$$

If the regularity hypotheses (R1) and (R2) hold, a short calculation shows that

$$\|\psi\|_{H^2(\Omega)} \leq C \|\varphi\|_{L^2(\Omega)};$$

hence,

$$\|\varphi\|_{L^2(\Omega)} \leq C\|\gamma\|_{H^{-\frac{3}{2}}(\Omega)}. \tag{2.17}$$

Now,

$$\frac{\partial g}{\partial t}(u) - \frac{\partial g}{\partial t}(W) = \eta(x,t) \int_0^1 \frac{\partial^2 g}{\partial t\,\partial u}\left(x, t, W(x,t) + \theta\,[u(x,t) - W(x,t)]\right) d\theta.$$

Since the integrand is boundedly in $H^{\frac{1}{2}}(\partial\Omega)$ for $\theta \in [0, 1]$ and $t \in (0, T]$ if $g \in C^3$ with bounded derivatives through order three, it follows that

$$\frac{\partial g}{\partial t}(u) - \frac{\partial g}{\partial t}(W) = G_1 \eta, \qquad G_1 \in L^\infty\left(0, T; H^{\frac{1}{2}}(\partial\Omega)\right).$$

In fact,

$$\|G_1\|_{L^\infty(0,T; H^{\frac{1}{2}}(\partial\Omega))} \leq C\left[1 + \|u\|_{L^\infty(0,T;H^1(\Omega))}\right].$$

Lemma 2.2 (with $\varepsilon = \frac{1}{2}$) shows that multiplication by G_1 is a continuous map of $H^{\frac{3}{2}}(\partial\Omega)$ into $H^{\frac{1}{2}}(\partial\Omega)$; by transposition it is also a continuous map of $H^{-\frac{1}{2}}(\partial\Omega)$ into $H^{-\frac{3}{2}}(\partial\Omega)$. Thus,

$$\left\|\frac{\partial g}{\partial t}(u) - \frac{\partial g}{\partial t}(W)\right\|_{H^{-\frac{3}{2}}(\partial\Omega)} \leq C\|\eta\|_{H^{-\frac{1}{2}}(\partial\Omega)}, \tag{2.18}$$

with the constant depending on $\|u\|_{H^1(\Omega)}$.

Similarly,

$$\frac{\partial u}{\partial t}\left[\frac{\partial g}{\partial u}(u) - \frac{\partial g}{\partial u}(W)\right] = \eta \frac{\partial u}{\partial t} \int_0^1 \frac{\partial^2 g}{\partial u^2}(W + \theta\,[u - W])\, d\theta = G_2 \eta.$$

If $\partial u/\partial t \in H^{\frac{3}{2}}(\partial\Omega)$, then $G_2 \in H^{\frac{1}{2}}(\partial\Omega)$. It is clear that for smooth $a(x)$, $u \in H^4(\Omega)$ implies $\partial u/\partial t \in H^{\frac{3}{2}}(\partial\Omega)$. Hence,

$$\left\| \frac{\partial u}{\partial t} \left[\frac{\partial g}{\partial u}(u) - \frac{\partial g}{\partial u}(W) \right] \right\|_{H^{-\frac{3}{2}}(\partial\Omega)} \leq C \|\eta\|_{H^{-\frac{1}{2}}(\partial\Omega)}, \tag{2.19}$$

where

$$C = C\left(\left\| \frac{\partial u}{\partial t} \right\|_{H^{\frac{3}{2}}(\partial\Omega)}, \|u\|_{H^1(\Omega)} \right).$$

Note that $H^{\frac{3}{2}}(\partial\Omega)$ can be replaced by $H^{1+\varepsilon}(\partial\Omega)$ for $\varepsilon > 0$; however, $u \in H^4(\Omega)$ will be needed anyway for the optimal rate of convergence to occur when the subspace contains at least the smooth cubic splines.

Clearly, we need to estimate η in the $H^{-\frac{1}{2}}(\partial\Omega)$-norm. This can be done using a different variant of the Nitsche lemma. Let $\beta = \beta(t) \in H^1(\Omega)$ be the solution of

$$\langle a \nabla\beta, \nabla z \rangle + \lambda \langle \beta, z \rangle - \langle G\beta, z \rangle_{\partial\Omega} = \langle \delta, z \rangle_{\partial\Omega}, \quad z \in H^1(\Omega), \tag{2.20}$$

where $\delta \in H^{\frac{1}{2}}(\partial\Omega)$ is such that

$$\begin{aligned} \|\delta\|_{H^{\frac{1}{2}}(\partial\Omega)} &= \|\eta\|_{H^{-\frac{1}{2}}(\partial\Omega)}, \\ \langle \delta, \eta \rangle_{\partial\Omega} &= \|\eta\|^2_{H^{-\frac{1}{2}}(\partial\Omega)}. \end{aligned} \tag{2.21}$$

The existence of such a δ is a simple consequence of the Hahn-Banach theorem. Then, choosing $z = \eta$ and using (2.1) leads to

$$\begin{aligned} \|\eta\|^2_{H^{-\frac{1}{2}}(\partial\Omega)} &= \langle a \nabla\beta, \nabla\eta \rangle + \lambda \langle \beta, \eta \rangle - \langle G\beta, \eta \rangle_{\partial\Omega} \\ &= \langle a \nabla\eta, \nabla(\beta - \chi) \rangle + \lambda \langle \eta, \beta - \chi \rangle - \langle g(u) - g(W), \beta - \chi \rangle_{\partial\Omega} \\ &\leq C \|\eta\|_{H^1(\Omega)} \|\beta - \chi\|_{H^1(\Omega)}, \quad \chi \in \mathcal{M}. \end{aligned}$$

Hence,

$$\begin{aligned} \|\eta\|^2_{H^{-\frac{1}{2}}(\partial\Omega)} &\leq C \|\eta\|_{H^1(\Omega)} \|\beta\|_{H^2(\Omega)} h \\ &\leq C \|\eta\|_{H^1(\Omega)} \|\delta\|_{H^{\frac{1}{2}}(\partial\Omega)} h \\ &= C \|\eta\|_{H^1(\Omega)} \|\eta\|_{H^{-\frac{1}{2}}(\partial\Omega)} h, \end{aligned}$$

using (S_h^r) and (R2). Thus,

$$\|\eta\|_{H^{-\frac{1}{2}}(\partial\Omega)} \leq C \|u\|_{H^k(\Omega)} h^k, \quad 1 \leq k \leq r. \tag{2.22}$$

This result can be combined with (2.17), (2.18), and (2.19) to show that

$$\|\varphi\|_{L^2(\Omega)} \leq C \|u\|_{H^k(\Omega)} h^k, \quad 1 \leq k \leq r, \tag{2.23}$$

with

$$C = C\left(\left\| \frac{\partial u}{\partial t} \right\|_{H^{\frac{3}{2}}(\partial\Omega)}, \|u\|_{H^1(\Omega)} \right). \tag{2.24}$$

Let

$$\psi = \varphi - \frac{\partial\eta}{\partial t} = \varphi - \frac{\partial u}{\partial t} + \frac{\partial W}{\partial t}. \tag{2.25}$$

Then

$$\langle a, \nabla\psi, \nabla v \rangle + \lambda \langle \psi, v \rangle - \left\langle \frac{\partial g}{\partial u}(W)\psi, v \right\rangle_{\partial\Omega} = 0, \quad v \in \mathcal{M}.$$

Hence, the choice $v=\psi+(\chi-\varphi+\partial u/\partial t)\in\mathcal{M}$ implies that

$$\begin{aligned}\|\psi\|_{H^1(\Omega)} &\leqq C\inf_{\chi\in\mathcal{M}}\left\|\chi-\varphi+\frac{\partial u}{\partial t}\right\|_{H^1(\Omega)}\\ &\leqq C\left[\|\varphi\|_{H^1(\Omega)}+\inf_{\chi\in\mathcal{M}}\left\|\chi+\frac{\partial u}{\partial t}\right\|_{H^1(\Omega)}\right]\\ &\leqq C\left[\|\eta\|_{H^1(\Omega)}+\left\|\frac{\partial u}{\partial t}\right\|_{H^k(\Omega)}h^{k-1}\right]\\ &\leqq C\left[\|u\|_{H^k(\Omega)}+\left\|\frac{\partial u}{\partial t}\right\|_{H^k(\Omega)}\right]h^{k-1},\qquad 1\leqq k\leqq r,\end{aligned}$$

where C is of the same form as in (2.24). The same argument that was used to lift the $H^1(\Omega)$ estimate of η to an $L^2(\Omega)$ estimate can be used again with the only change being the redefinition of G to be $\partial g/\partial u(W)$ to obtain the inequality

$$\|\psi\|_{L^2(\Omega)}\leqq C\left[\|u\|_{H^k(\Omega)}+\left\|\frac{\partial u}{\partial t}\right\|_{H^k(\Omega)}\right]h^k,\qquad 1\leqq k\leqq r. \tag{2.26}$$

The estimates (2.23) and (2.26) combine to give

$$\left\|\frac{\partial\eta}{\partial t}\right\|_{L^2(\Omega)}\leqq C\left[\|u\|_{H^k(\Omega)}+\left\|\frac{\partial u}{\partial t}\right\|_{H^k(\Omega)}\right]h^k,\qquad 1\leqq k\leqq r, \tag{2.27}$$

where C is of the form in (2.24).

We collect the L^2 estimates in the following way.

Theorem 2.3. Let $\dim(\Omega)\leqq 3$ and assume that the trace inequalities (T1) and (T2) and the regularity inequalities (R1) and (R2) hold. Let $\mathcal{M}=\mathcal{M}_h$ be taken from an S_h^r family. Let the solution u of (1.1) be such that the norms in (2.24) are finite. Then, a constant C of the form (2.24) exists so that

$$\|\eta\|_{L^2(\Omega)}+\left\|\frac{\partial\eta}{\partial t}\right\|_{L^2(\Omega)}\leqq C\left[\|u\|_{H^k(\Omega)}+\left\|\frac{\partial u}{\partial t}\right\|_{H^k(\Omega)}\right]h^k,\qquad 1\leqq k\leqq r.$$

3. H^1 Estimates for the Continuous-Time-Galerkin Method

The most natural estimate that can be derived for the error

$$\zeta=u-U \tag{3.1}$$

between the solution u of (1.10) and the solution U of (1.11) is a bound in $L^2(0,T;H^1(\Omega))$, since this bound can be obtained without limiting the subspace $\mathcal{M}$ of $H^1(\Omega)$ in any way. In the case that $\mathcal{M}$ satisfies the Bramble-Schatz condition (S_h^r) then the bound will be of optimal order in the parameter h. The estimate will be derived making use of the projection $W(\cdot,t)\in\mathcal{M}$ defined by (2.1).

Let

$$\xi=W-U,\qquad \eta=u-W. \tag{3.2}$$

Then, it is easy to see from (2.1) that

$$\begin{aligned}&\left\langle\frac{\partial W}{\partial t},v\right\rangle+\langle a\nabla W,\nabla v\rangle-\langle g(W),v\rangle_{\partial\Omega}\\ &\quad=\left\langle\frac{\partial u}{\partial t},v\right\rangle+\langle a\nabla u,\nabla v\rangle-\langle g(u),v\rangle_{\partial\Omega}\\ &\qquad-\left\langle\frac{\partial u}{\partial t},v\right\rangle+\lambda\langle\eta,v\rangle\\ &\quad=-\left\langle\frac{\partial\eta}{\partial t},v\right\rangle+\lambda\langle\eta,v\rangle,\qquad v\in\mathcal{M}.\end{aligned}$$

Eq. (1.11) can be subtracted from (3.3) to obtain the relation

$$\left\langle \frac{\partial \xi}{\partial t}, v \right\rangle + \langle a \nabla \xi, \nabla v \rangle - \langle g(W) - g(U), v \rangle_{\partial\Omega} = -\left\langle \frac{\partial \eta}{\partial t}, v \right\rangle + \lambda \langle \eta, v \rangle, \quad v \in \mathcal{M}. \tag{3.4}$$

Since $\xi(\cdot, t) \in \mathcal{M}$, it can be used as the test function in (3.4); then (1.9) and (1.19) imply that

$$\frac{1}{2}\frac{d}{dt}\|\xi\|^2_{L^2(\Omega)} + m\|\xi\|^2_{H_0^1(\Omega)} - K\|\xi\|^2_{L^2(\partial\Omega)} \leqq \frac{1}{2}\left\|\frac{\partial \eta}{\partial t}\right\|^2_{L^2(\Omega)} + \frac{1}{2}\lambda\|\eta\|^2_{L^2(\Omega)} + \frac{1}{2}(\lambda+1)\|\xi\|^2_{L^2(\Omega)}. \tag{3.5}$$

Assume that the domain Ω is such that the trace inequality (T1) holds. Then,

$$K\|\xi\|^2_{L^2(\partial\Omega)} \leqq \tfrac{1}{2} m\|\xi\|^2_{H_0^1(\Omega)} + C_1\|\xi\|^2_{L^2(\Omega)}$$

and

$$\frac{d}{dt}\|\xi\|^2_{L^2(\Omega)} + m\|\xi\|^2_{H_0^1(\Omega)} \leqq C\left[\|\xi\|^2_{L^2(\Omega)} + \left\|\frac{\partial \eta}{\partial t}\right\|^2_{L^2(\Omega)} + \|\eta\|^2_{L^2(\Omega)}\right]. \tag{3.6}$$

Thus, for $0 < t \leqq T$,

$$\|\xi(t)\|^2_{L^2(\Omega)} + m\|\xi\|^2_{L^2(0,t;H_0^1(\Omega))} \leqq C\left[\|\xi\|^2_{L^2(0,t;L^2(\Omega))} + \|\xi(0)\|^2_{L^2(\Omega)} + \left\|\frac{\partial \eta}{\partial t}\right\|^2_{L^2(0,t;L^2(\Omega))} + \|\eta\|^2_{L^2(0,t;L^2(\Omega))}\right].$$

It follows easily from the Gronwall lemma that

$$\|\xi\|^2_{L^\infty(0,T;L^2(\Omega))} + \|\xi\|^2_{L^2(0,T;H^1(\Omega))} \leqq C\left[\|\xi(0)\|^2_{L^2(\Omega)} + \left\|\frac{\partial \eta}{\partial t}\right\|^2_{L^2(0,T;L^2(\Omega))} + \|\eta\|^2_{L^2(0,T;L^2(\Omega))}\right]. \tag{3.7}$$

Hence,

$$\|\zeta\|_{L^\infty(0,T;L^2(\Omega))} + \|\zeta\|_{L^2(0,T;H^1(\Omega))} \leqq C\left[\|\zeta(0)\|_{L^2(\Omega)} + \|\eta\|_{L^2(0,T;H^1(\Omega))} + \left\|\frac{\partial \eta}{\partial t}\right\|_{L^2(0,T;L^2(\Omega))}\right], \tag{3.8}$$

where we used the triangle inequality, (3.7), and the fact that

$$\|\eta\|_{L^\infty(0,T;L^2(\Omega))} \leqq C\left[\|\eta\|_{L^2(0,T;L^2(\Omega))} + \left\|\frac{\partial \eta}{\partial t}\right\|_{L^2(0,T;L^2(\Omega))}\right].$$

Assume that g is such that (2.7) holds. Then it follows from (2.4) and (2.8) that

$$\|\zeta\|_{L^\infty(0,T;L^2(\Omega))} + \|\zeta\|_{L^2(0,T;H^1(\Omega))} \leqq C\left[\|\zeta(0)\|_{L^2(\Omega)} + \left(\int_0^T \left(\inf_{\chi_1(t)\in\mathcal{M}} \|u-\chi_1\|^2_{H^1(\Omega)}\right)^2 dt\right)^{\frac{1}{2}} + \left(\int_0^T \left(\inf_{\chi_2(t)\in\mathcal{M}} \left\|\frac{\partial u}{\partial t} - \chi_2\right\|_{L^2(\Omega)}\right)^2 dt\right)^{\frac{1}{2}}\right], \tag{3.9}$$

$$C = C\left(\left\|\frac{\partial u}{\partial t}\right\|_{L^\infty(0,T;L^\infty(\partial\Omega))}\right).$$

Let $\mathcal{M} = \mathcal{M}_h$ be chosen from an S_h^r family and retain the other hypotheses from above. Assume that $U(0)$ is selected so that

$$\|\zeta(0)\|_{L^2(\Omega)} \leqq C\|f\|_{H^k(\Omega)} h^k; \tag{3.10}$$

this can be done by using the projection of f into $\mathcal{M}_h$ with respect to either $L^2(\Omega)$ or $H^1(\Omega)$. Usually (3.10) can be obtained by interpolating f into $\mathcal{M}_h$; this is almost always the easiest way to initialize the Galerkin problem. It then follows from Theorem 2.1, (3.8), and (3.10) that

$$\begin{gathered}\|\zeta\|_{L^\infty(0,T;L^2(\Omega))} + \|\zeta\|_{L^2(0,T;H^1(\Omega))} \leqq C h^{k-1}, \qquad 1 \leqq k \leqq r, \\ C \leqq C_1 \left(1 + \left\|\frac{\partial u}{\partial t}\right\|_{L^\infty(0,T;L^\infty(\partial\Omega))}\right)\left(\|u\|_{L^2(0,T;H^k(\Omega))} + \left\|\frac{\partial u}{\partial t}\right\|_{L^2(0,T;H^k(\Omega))}\right).\end{gathered} \tag{3.11}$$

Theorem 3.1. Let the trace inequality (T1) hold and let $\partial^2 g/\partial t \partial u$ and $\partial^2 g/\partial u^2$ be bounded. Let $\mathcal{M}_h$ be selected from an S_h^r family. Let $f \in H^k(\Omega)$, $1 \leqq k \leqq r$, and assume that $U(0)$ satisfies (3.10). Then there exists a constant C_1 such that the error $\zeta = u - U$ can be estimated by the inequality (3.11).

Note that (3.11) is optimal in the exponent on the parameter h. It should also be noted that the bound (3.11) implies that

$$\|\zeta\|_{L^2(0,T;H^{\frac{1}{2}}(\partial\Omega))} \leqq C h^{k-1}$$

as well, if $H^{\frac{1}{2}}(\partial\Omega)$ can be defined.

4. L^2 Estimates for the Continuous-Time-Galerkin Method

The principal result of this paper concerning the continuous-time-Galerkin method is embodied in the following theorem.

Theorem 4.1. Let $\dim(\Omega) \leqq 3$ and assume that Ω, $a(x)$, and $g(x, t, u)$ are such that the trace inequalities (T1) and (T2) hold and the elliptic regularity inequalities hold for λ sufficiently large. Let $g \in C^3(\Omega \times [0, T] \times \mathbb{R})$ have bounded derivatives of order less than or equal to three. Let $\mathcal{M}_h$ be selected from an S_h^r family, where $r \geqq 2$. Suppose that for some k, $1 \leqq k \leqq r$,

$$\begin{gathered} u \in L^\infty(0, T; H^k(\Omega)), \\ \frac{\partial u}{\partial t} \in L^2(0, T; H^k(\Omega)) \cap L^\infty(0, T; H^{\frac{1}{2}}(\partial\Omega)). \end{gathered} \tag{4.1}$$

Then there exists a function $C(u)$, with the dependence on u expressible in terms of the norms of the three spaces in (4.1), such that (u being the solution of (1.10) and U that of (1.11))

$$\|u - U\|_{L^\infty(0,T;L^2(\Omega))} \leqq C(u) h^k, \tag{4.2}$$

provided that $U(0)$ is chosen so that

$$\|f - U(0)\|_{L^2(\Omega)} \leqq C\|f\|_{H^k(\Omega)} h^k. \tag{4.3}$$

It should be noted that there is an immediate corollary of Theorems 3.1 and 4.1 and standard interpolation theory [7].

Corollary 4.2. Assume the hypotheses of Theorem 4.1. Then

$$\|u-U\|_{L^2(0,T;H^s(\Omega))} \leqq C(u)\,h^{k-s}, \qquad 0\leqq s\leqq 1, \qquad 1\leqq k\leqq r, \tag{4.4}$$

with $C(u)$ of the same form as in Theorem 4.1.

The proof of Theorem 4.1 is simple, given the preliminaries that have been covered. If follows trivially from (3.7) that

$$\|\zeta\|_{L^\infty(0,T;L^2(\Omega))} \leqq C\left[\|\xi(0)\|_{L^2(\Omega)} + \|\eta\|_{L^\infty(0,T;L^2(\Omega))} + \left\|\frac{\partial\eta}{\partial t}\right\|_{L^2(0,T;L^2(\Omega))}\right]. \tag{4.5}$$

Now, (4.3) and (2.15) imply that

$$\|\xi(0)\|_{L^2(\Omega)} \leqq C h^k,$$

and (4.1), (2.15), and (2.27) imply that

$$\|\eta\|_{L^\infty(0,T;L^2(\Omega))} + \left\|\frac{\partial\eta}{\partial t}\right\|_{L^2(0,T;L^2(\Omega))} \leqq C h^k$$

with the proper form of the constant being clear from the inequalities of Section 2. Thus, (4.2) is proved.

It should be particularly noted that the only condition imposed on the subspace $\mathscr{M}_h$ was that it come from an S_h^r family, and it was shown that the optimal rate of convergence resulted. This is not always the case; for some hyperbolic problems it is known that something besides the Bramble-Schatz condition is needed for optimal convergence rates to occur.

5. The Crank-Nicolson-Galerkin Method

Both $H^1(\Omega)$ and $L^2(\Omega)$ estimates will be constructed for the error $\zeta = u - U$ between the solutions of (1.10) and U of (1.13), the Crank-Nicolson-Galerkin method. Let W, ξ, and η retain their earlier definitions in terms of u and U. Also, let

$$\begin{aligned}
&\|v\|^2_{L^2_{\Delta t}(0,T;X)} = \sum_{0\leqq t^n\leqq T} \|v^n\|_X^2\,\Delta t, \qquad t^n = n\,\Delta t, \qquad v^n = v(t^n),\\
&\|v\|_{L^\infty_{\Delta t}(0,T;X)} = \max_{0\leqq t^n\leqq T} \|v^n\|_X,\\
&\|v\|^2_{\hat{L}^2_{\Delta t}(0,T;X)} = \sum_{0<t^{n+\frac12}<T} \|v^{n+\frac12}\|_X^2\,\Delta t, \qquad v^{n+\frac12} = \tfrac12(v^n + v^{n+1}).
\end{aligned} \tag{5.1}$$

Again we shall estimate ξ in terms of η and $d_t\eta$ and make strong use of the estimates of section 2 on η and $\partial\eta/\partial t$. First, it follows from (2.1) that

$$\begin{aligned}
&\langle a\nabla W^{n+\frac12}, \nabla v\rangle + \lambda\langle W^{n+\frac12}, v\rangle - \langle g(t,W)^{n+\frac12}, v\rangle_{\partial\Omega}\\
&\quad = \langle a\nabla u^{n+\frac12}, \nabla v\rangle + \lambda\langle u^{n+\frac12}, v\rangle - \langle g(t,u)^{n+\frac12}, v\rangle_{\partial\Omega}, \qquad v\in\mathscr{M}.
\end{aligned}$$

It is clear from (1.10) that

$$\left\langle \frac{\partial u^{n+\frac12}}{\partial t}, z\right\rangle + \langle a\nabla u^{n+\frac12}, \nabla z\rangle - \langle g(t,u)^{n+\frac12}, z\rangle_{\partial\Omega} = 0, \qquad z\in H^1(\Omega).$$

Hence, a simple calculation shows that

$$\begin{aligned}\langle d_t W^n, v\rangle + \langle a\nabla W^{n+\frac12}, \nabla v\rangle - \langle g(t^{n+\frac12}, W^{n+\frac12}), v\rangle_{\partial\Omega} \\ = -\langle d_t\eta^n, v\rangle + \lambda\langle \eta^{n+\frac12}, v\rangle - \langle g(t^{n+\frac12}, W^{n+\frac12}) - g(t, W)^{n+\frac12}, v\rangle_{\partial\Omega} \\ + \left\langle d_t u^n - \frac{\partial u^{n+\frac12}}{\partial t}, v\right\rangle, \quad v\in\mathcal{M}.\end{aligned} \tag{5.2}$$

Substraction of (1.13) from (5.2) gives the equation for the evolution of ξ:

$$\begin{aligned}\langle d_t \xi^n, v\rangle + \langle a\nabla \xi^{n+\frac12}, \nabla v\rangle - \langle g(t^{n+\frac12}, W^{n+\frac12}) - g(t^{n+\frac12}, U^{n+\frac12}), v\rangle_{\partial\Omega} \\ = -\langle d_t\eta^n, v\rangle + \lambda\langle \eta^{n+\frac12}, v\rangle - \langle g(t^{n+\frac12}, W^{n+\frac12}) - g(t, W)^{n+\frac12}, v\rangle_{\partial\Omega} \\ + \left\langle d_t u^n - \frac{\partial u^{n+\frac12}}{\partial t}, v\right\rangle, \quad v\in\mathcal{M}.\end{aligned} \tag{5.3}$$

Recall that

$$\left(d_t u^n - \frac{\partial u^{n+\frac12}}{\partial t}\right)(x) = -\frac{1}{2\Delta t}\int_{t^n}^{t^{n+1}} (t-t^n)(t^{n+1}-t)\frac{\partial^3 u}{\partial t^3}(x,t)\,dt;$$

thus, we can estimate this truncation term by

$$\left\| d_t u^n - \frac{\partial u^{n+\frac12}}{\partial t}\right\|^2_{L^2(\Omega)} \leqq \frac{(\Delta t)^3}{120}\left\|\frac{\partial^3 u}{\partial t^3}\right\|^2_{L^2(t^n, t^{n+1}; L^2(\Omega))}.$$

Also,

$$\|d_t\eta^n\|^2_{L^2(\Omega)} \leqq (\Delta t)^{-1}\left\|\frac{\partial\eta}{\partial t}\right\|^2_{L^2(t^n, t^{n+1}; L^2(\Omega))}.$$

Now, employ $v=\xi^{n+\frac12}$ as the test function in (5.3):

$$\begin{aligned}(2\Delta t)^{-1}(\|\xi^{n+1}\|^2_{L^2(\Omega)} - \|\xi^n\|^2_{L^2(\Omega)}) + m\|\xi^{n+\frac12}\|^2_{H_0^1(\Omega)} - K\|\xi^{n+\frac12}\|^2_{L^2(\partial\Omega)} \\ \leqq (2\Delta t)^{-1}\left\|\frac{\partial\eta}{\partial t}\right\|^2_{L^2(t^n, t^{n+1}; L^2(\Omega))} + \frac12\lambda\|\eta^{n+\frac12}\|^2_{L^2(\Omega)} \\ + \frac12\|g(t^{n+\frac12}, W^{n+\frac12}) - g(t, W)^{n+\frac12}\|^2_{L^2(\partial\Omega)} \\ + \frac{(\Delta t)^3}{240}\left\|\frac{\partial^3 u}{\partial t^3}\right\|^2_{L^2(t^n, t^{n+1}; L^2(\Omega))} + C\|\xi^{n+\frac12}\|^2_{L^2(\Omega)} + \frac12\|\xi^{n+\frac12}\|^2_{L^2(\partial\Omega)}\end{aligned} \tag{5.4}$$

After having used the trace inequality (T1), sum on n for $n=0,\ldots,m$, and then use the discrete form of the Gronwall inequality. It follows that

$$\begin{aligned}\|\xi\|^2_{L^\infty_{\Delta t}(0,T;L^2(\Omega))} + \|\xi\|^2_{L^2_{\Delta t}(0,T;H_0^1(\Omega))} \\ \leqq C\Big[\|\xi^0\|^2_{L^2(\Omega)} + \|\eta\|^2_{L^2_{\Delta t}(0,T;L^2(\Omega))} \\ + \left\|\frac{\partial\eta}{\partial t}\right\|^2_{L^2(0,T;L^2(\Omega))} + (\Delta t)^4\left\|\frac{\partial^3 u}{\partial t^3}\right\|^2_{L^2(0,T;L^2(\Omega))} \\ + \|g(t^{n+\frac12}, W^{n+\frac12}) - g(t, W)^{n+\frac12}\|^2_{L^2_{\Delta t}(0,T;L^2(\partial\Omega))}\Big].\end{aligned} \tag{5.5}$$

Note that some terms on the right involve a discrete L^2 norm and some the ordinary L^2 norm in time. The only term that requires a new treatment is the last one.

We need to show that

$$\|g(t^{n+\frac12}, W^{n+\frac12}) - g(t, W)^{n+\frac12}\|_{L^2(\Omega)} = O((\Delta t)^2)$$

in order to obtain global second order accuracy in Δt. Let

$$\partial_t^2 v^{n+\frac{1}{2}} = 4\left(v^{n+1} - 2v(t^{n+\frac{1}{2}}) + v^n\right)(\Delta t)^{-2}$$

denote the second time difference associated with the step $\Delta t/2$. Then, recall that

$$\partial_t^2 v^{n+\frac{1}{2}} = \frac{4}{(\Delta t)^2} \int_{t^n}^{t^{n+1}} \left(\frac{\Delta t}{2} - |t^{n+\frac{1}{2}} - s|\right) \frac{d^2 v}{dt^2}(s)\, ds.$$

Let $\widetilde{W}$ be the piecewise-linear interpolation (in time) of $\{W^n\}$; then

$$\begin{aligned} g(t^{n+\frac{1}{2}}, W^{n+\frac{1}{2}}) - g(t, W)^{n+\frac{1}{2}} &= -\frac{1}{8}(\Delta t)^2 \partial_t^2 g(t, \widetilde{W})^{n+\frac{1}{2}} \\ &= -\frac{1}{2} \int_{t^n}^{t^{n+1}} \left(\frac{\Delta t}{2} - |t^{n+\frac{1}{2}} - s|\right) \frac{d^2}{dt^2} g(t, \widetilde{W})\Big|_{t=s} ds. \end{aligned} \tag{5.6}$$

Now,

$$\frac{d^2}{dt^2} g(t, \widetilde{W}) = \frac{\partial^2 g}{\partial t^2}(t, W) + 2\frac{\partial^2 g}{\partial t\, \partial u}(t, W)\, d_t W^n + \frac{\partial^2 g}{\partial u^2}(t, W)(d_t W^n)^2 + \frac{\partial g}{\partial u}(t, W)\frac{\partial^2 \widetilde{W}}{\partial t^2},$$
$$t^n < t < t^{n+1},$$

and we have three integrals to consider, since the last term vanishes. If we retain the assumption that g has bounded second derivatives, then it is clear that

$$\begin{aligned} &\|g(t^{n+\frac{1}{2}}, W^{n+\frac{1}{2}}) - g(t, W)^{n+\frac{1}{2}}\|_{L^2(\partial\Omega)}^2 \\ &\quad \le C(\Delta t)^3 \left[\Delta t + \left\|\frac{\partial W}{\partial t}\right\|_{L^2(t^n, t^{n+1}; L^2(\partial\Omega))}^2 + \left\|\left(\frac{\partial W}{\partial t}\right)^2\right\|_{L^2(t^n, t^{n+1}; L^2(\partial\Omega))}^2\right], \end{aligned}$$

and

$$\begin{aligned} &\|g(t^{n+\frac{1}{2}}, W^{n+\frac{1}{2}}) - g(t, W)^{n+\frac{1}{2}}\|_{L^2_{\Delta t}(0, T; L^2(\partial\Omega))}^2 \\ &\quad \le C(\Delta t)^4 \left[1 + \left\|\frac{\partial W}{\partial t}\right\|_{L^2(0, T; L^2(\partial\Omega))}^2 + \left\|\left(\frac{\partial W}{\partial t}\right)^2\right\|_{L^2(0, T; L^2(\partial\Omega))}^2\right]. \end{aligned} \tag{5.7}$$

Now, it follows from Section 2 that

$$\begin{aligned} \left\|\frac{\partial W}{\partial t}\right\|_{L^2(0, T; L^2(\partial\Omega))}^2 &\le C\left\|\frac{\partial W}{\partial t}\right\|_{L^2(0, T; H^1(\Omega))}^2 \\ &\le 2C\left(\left\|\frac{\partial u}{\partial t}\right\|_{L^2(0, T; H^1(\Omega))}^2 + \left\|\frac{\partial \eta}{\partial t}\right\|_{L^2(0, T; H^1(\Omega))}^2\right) \\ &\le C_1\left(\left\|\frac{\partial u}{\partial t}\right\|_{L^2(0, T; H^1(\Omega))}^2 + \|u\|_{L^2(0, T; H^1(\Omega))}^2\right), \end{aligned}$$

where C_1 involves the $L^\infty(0, T; L^\infty(\partial\Omega))$ norm of $\partial u/\partial t$. For convenience in this argument, let us only consider the case when we can expect to get L^2 estimates; i.e., let (T1), (T2), (R1), and (R2) hold and assume dim $(\Omega) \le 3$. Let $\mathscr{M} = \mathscr{M}_h$ be chosen from an S_h^r family. Note that, since $\dim(\partial\Omega) \le 2$, $H^{\frac{1}{2}}(\partial\Omega) \subset L^4(\partial\Omega)$ with continuous injection. Thus, it is also the case that

$$\left\|\left(\frac{\partial W}{\partial t}\right)^2\right\|_{L^2(0, T; L^2(\partial\Omega))}^2 \le C_1\left(\left\|\frac{\partial u}{\partial t}\right\|_{L^4(0, T; H^1(\Omega))} + \|u\|_{L^4(0, T; H^1(\Omega))}\right)^4$$

with C_1 having the same form as above. The last few inequalities show that

$$\|g(t^{n+\frac12}, W^{n+\frac12}) - g(t, W)^{n+\frac12}\|^2_{L^2_{\Delta t}(0,T;L^2(\partial\Omega))} \leqq C_2(\Delta t)^4, \tag{5.8}$$

where

$$C_2 = C\left(1 + \left\|\frac{\partial u}{\partial t}\right\|^2_{L^\infty(0,T;L^\infty(\partial\Omega))}\right)\left(\|u\|^4_{L^4(0,T;H^1(\Omega))} + \left\|\frac{\partial u}{\partial t}\right\|^4_{L^4(0,T;H^1(\Omega))} + 1\right). \tag{5.9}$$

Note that we can replace the $L^2_{\Delta t}(0, T; L^2(\Omega))$-norm on η by the $L^2(0, T; L^2(\Omega))$-norm in (5.5), since

$$\|\eta\|_{L^2_{\Delta t}(0,T;L^2(\Omega))} \leqq C\|\eta\|_{L^\infty(0,T;L^2(\Omega))} \leqq C\left[\|\eta\|_{L^2(0,T;L^2(\Omega))} + \left\|\frac{\partial\eta}{\partial t}\right\|_{L^2(0,T;L^2(\Omega))}\right].$$

These results can be combined to give the final estimate for ξ:

$$\begin{aligned}\|\xi\|^2_{L^\infty_{\Delta t}(0,T;L^2(\Omega))} &+ \|\xi\|^2_{L^2_{\Delta t}(0,T;H^1_0(\Omega))} \\ &\leqq C\left[\|\xi(0)\|^2_{L^2(\Omega)} + \|\eta\|^2_{L^2(0,T;L^2(\Omega))} + \left\|\frac{\partial\eta}{\partial t}\right\|^2_{L^2(0,T;L^2(\Omega))}\right] + C_3(\Delta t)^4,\end{aligned} \tag{5.10}$$

where

$$C_3 = C_2 + C\left\|\frac{\partial^3 u}{\partial t^3}\right\|^2_{L^2(0,T;L^2(\Omega))}. \tag{5.11}$$

It is now an immediate consequence of (5.10) and the bound on η and $\partial\eta/\partial t$ derived in section 2 that the following theorem is valid.

Theorem 5.1. Let $\dim(\Omega) \leqq 3$ and assume that (T1), (T2), (R1), and (R2) are valid. Let $g(x, t, u)$ have continuous and bounded derivatives of order order less than or equal to three. Let $\mathcal{M}_h$ be selected from an S^r_h family for some $r \geqq 2$. Let the solution u of (1.10) satisfy the following constraints for some k such that $1 \leqq k \leqq r$:

$$\begin{gathered} u \in L^\infty(0, T; H^k(\Omega)), \\ \frac{\partial u}{\partial t} \in L^2(0, T; H^k(\Omega)) \cap L^\infty(0, T; H^{\frac12}(\partial\Omega)), \\ \frac{\partial^3 u}{\partial t^3} \in L^2(0, T; L^2(\Omega)). \end{gathered} \tag{5.12}$$

Let U be the solution of the Crank-Nicolson-Galerkin equation (1.13). Then, there exist constants C and $C(s)$, $0 \leqq s \leqq 1$, depending on the norms of u, $\partial u/\partial t$, and $\partial^3 u/\partial t^3$ in the spaces listed above such that

$$\|u - U\|_{L^\infty_{\Delta t}(0,T;L^2(\Omega))} \leqq C(h^k + (\Delta t)^2) \tag{5.13}$$

and

$$\|u - U\|_{L^2_{\Delta t}(0,T;H^s(\Omega))} \leqq C(s)(h^{k-s} + (\Delta t)^2), \quad 0 \leqq s \leqq 1, \tag{5.14}$$

provided that $\|f - U(0)\|_{L^2(\Omega)} \leqq C\|f\|_{H^k(\Omega)} h^k$.

We shall now consider the three modifications of the Crank-Nicolson-Galerkin equations that were introduced in Section 1. Fortunately, their analyses can be reduced in each case to that of the Crank-Nicolson equation plus the consideration of a perturbation term.

6. The Extrapolated-Crank-Nicolson-Galerkin Procedure

Let U now denote the solution of the Extrapolated-Crank-Nicolson-Galerkin Eq. (1.15) for $t=t^n$, $n\geqq 2$. Assume for the moment that U^0, U^1, and U^2 are known; we shall specify rules for finding these initial values later in this section. Also, assume all of the hypotheses of Theorem 5.1.

Modify the evolution equation (5.2) for W to read

$$\begin{aligned}\langle d_t W^n, v\rangle + \langle a\nabla W^{n+\frac12}, \nabla v\rangle - \langle g(t^{n+\frac12}, 2W^{n-\frac12} - W^{n-\frac32}), v\rangle_{\partial\Omega} \\ = -\langle d_t\eta^n, v\rangle + \lambda\langle\eta^{n+\frac12}, v\rangle - \langle g(t^{n+\frac12}, W^{n+\frac12}) - g(t, W)^{n+\frac12}, v\rangle_{\partial\Omega} \\ + \left\langle d_t u^n - \frac{\partial u^{n+\frac12}}{\partial t}, v\right\rangle + \langle g(t^{n+\frac12}, W^{n+\frac12}) - g(t^{n+\frac12}, 2W^{n-\frac12} - W^{n-\frac32}), v\rangle_{\partial\Omega}\end{aligned} \tag{6.1}$$

for $v\in\mathcal{M}_h$. Subtract (1.15) from (6.1) and then use $\xi^{n+\frac12}$ as the test function v as before:

$$\begin{aligned}(2\Delta t)^{-1}(\|\xi^{n+1}\|^2_{L^2(\Omega)} - \|\xi^n\|^2_{L^2(\Omega)}) + m\|\xi^{n+\frac12}\|^2_{H_0^1} - K\|2\xi^{n-\frac12} - \xi^{n-\frac32}\|^2_{L^2(\partial\Omega)} \\ \leqq (2\Delta t)^{-1}\left\|\frac{\partial\eta}{\partial t}\right\|^2_{L^2(t^n, t^{n+1}; L^2(\Omega))} + \frac12\lambda\|\eta^{n+\frac12}\|^2_{L^2(\Omega)} \\ + \frac12\|g(t^{n+\frac12}, W^{n+\frac12}) - g(t, W)^{n+\frac12}\|^2_{L^2(\partial\Omega)} + \frac{(\Delta t)^3}{240}\left\|\frac{\partial^3 u}{\partial t^3}\right\|^2_{L^2(t^n, t^{n+1}; L^2(\Omega))} \\ + \frac12\|g(t^{n+\frac12}, W^{n+\frac12}) - g(t^{n+\frac12}, 2W^{n-\frac12} - W^{n-\frac32})\|^2_{L^2(\partial\Omega)} \\ + C\|\xi^{n+\frac12}\|^2_{L^2(\Omega)} + \|\xi^{n+\frac12}\|^2_{L^2(\partial\Omega)}, \qquad n\geqq 2.\end{aligned} \tag{6.2}$$

The boundary term on the left-hand side of (6.2) can be handled by the trace theorem (T1) without any subtlety, it then follows by essentially the same argument as in Section 5 that

$$\begin{aligned}\|\xi\|^2_{L^\infty_{\Delta t}(2\Delta t, T; L^2(\Omega))} + \|\xi\|^2_{L^2_{\Delta t}(2\Delta t, T; H_0^1(\Omega))} \\ \leqq C\Big[\|\xi^2\|^2_{L^2(\Omega)} + \|\eta\|^2_{L^2_{\Delta t}(2\Delta t, T; L^2(\Omega))} + \left\|\frac{\partial\eta}{\partial t}\right\|^2_{L^2(2\Delta t, T; L^2(\Omega))} \\ + (\Delta t)^4\left\|\frac{\partial^3 u}{\partial t^3}\right\|^2_{L^2(2\Delta t, T; L^2(\Omega))} \\ + \|g(t^{n+\frac12}, W^{n+\frac12}) - g(t, W)^{n+\frac12}\|^2_{L^2_{\Delta t}(2\Delta t, T; L^2(\partial\Omega))} \\ + \|g(t^{n+\frac12}, W^{n+\frac12}) - g(t^{n+\frac12}, 2W^{n-\frac12} - W^{n-\frac32})\|^2_{L^2(2\Delta t, T; L^2(\partial\Omega))} \\ + (\|\xi^{\frac12}\|^2_{H_0^1(\Omega)} + \|\xi^{\frac32}\|^2_{H_0^1(\Omega)})\Delta t\Big].\end{aligned} \tag{6.3}$$

The three initial terms are different from the one in (5.5), and there is a new boundary term. However,

$$\begin{aligned}\|g(t^{n+\frac12}, W^{n+\frac12}) - g(t^{n+\frac12}, 2W^{n-\frac12} - W^{n-\frac32})\|^2_{L^2(\partial\Omega)} \\ \leqq K^2\|W^{n+\frac12} - 2W^{n-\frac12} + W^{n-\frac32}\|^2_{L^2(\partial\Omega)} \\ \leqq \tfrac12 K^2\|W^{n+1} - 2W^n + W^{n-1}\|^2_{L^2(\partial\Omega)} + \tfrac12 K^2\|W^n - 2W^{n-1} + W^{n-2}\|^2_{L^2(\partial\Omega)}.\end{aligned}$$

It follows simply that

$$\begin{aligned}\|g(t^{n+\frac12}, W^{n+\frac12}) - g(t^{n+\frac12}, 2W^{n-\frac12} - W^{n-\frac32})\|^2_{L^2_{\Delta t}(2\Delta t, T; L^2(\partial\Omega))} \\ \leqq C(\Delta t)^4\left\|\frac{\partial^2 W}{\partial t^2}\right\|^2_{L^2(0, T; L^2(\partial\Omega))}.\end{aligned}$$

The bounding of $\partial^2 W/\partial t^2$ remains, and this will be facilitated by producing an estimate for $\partial^2\eta/\partial t^2$. For our purposes it is sufficient that this term be bounded; it is not necessary that it tend to zero with h. Eq. (2.1) can be differentiated twice with respect to time to give

$$\left\langle a\nabla\frac{\partial^2\eta}{\partial t^2}, \nabla v\right\rangle + \lambda\left\langle\frac{\partial^2\eta}{\partial t^2}, v\right\rangle - \left\langle\frac{d^2}{dt^2}[g(t,u) - g(t,W)], v\right\rangle_{\partial\Omega} = 0, \qquad v\in\mathcal{M}_h.$$

Thus,

$$\left\langle a\nabla\frac{\partial^2\eta}{\partial t^2}, \nabla v\right\rangle + \lambda\left\langle\frac{\partial^2\eta}{\partial t^2}, v\right\rangle - \left\langle\frac{\partial g}{\partial u}(W)\frac{\partial^2\eta}{\partial t^2}, v\right\rangle_{\partial\Omega} = \langle R, v\rangle_{\partial\Omega}, \qquad v\in\mathcal{M}_h,$$

where

$$\begin{aligned} R = &\left(\frac{\partial g}{\partial u}(u) - \frac{\partial g}{\partial u}(W)\right)\frac{\partial^2 u}{\partial t^2} + \frac{\partial^2 g}{\partial t^2}(u) - \frac{\partial^2 g}{\partial t^2}(W) \\ &+ 2\frac{\partial^2 g}{\partial t\,\partial u}(u)\frac{\partial u}{\partial t} - 2\frac{\partial^2 g}{\partial t\,\partial u}(W)\frac{\partial W}{\partial t} \\ &+ \frac{\partial^2 g}{\partial u^2}(u)\left(\frac{\partial u}{\partial t}\right)^2 - \frac{\partial^2 g}{\partial u^2}(W)\left(\frac{\partial W}{\partial t}\right)^2. \end{aligned}$$

Let $v = \partial^2\eta/\partial t^2 - \partial^2 u/\partial t^2 \in \mathcal{M}_h$. Then

$$\begin{aligned} \left\|\frac{\partial^2\eta}{\partial t^2}\right\|^2_{H^1(\Omega)} &\le C\left[\left\|\frac{\partial^2 u}{\partial t^2}\right\|^2_{H^1(\Omega)} + \|R\|^2_{L^2(\partial\Omega)}\right] \\ &\le C_1\left[\left\|\frac{\partial^2 u}{\partial t^2}\right\|^2_{H^1(\Omega)} + \|u\|^4_{H^1(\Omega)} + \left\|\frac{\partial u}{\partial t}\right\|^4_{H^1(\Omega)}\right. \\ &\qquad\left. + \left\|\frac{\partial^2 u}{\partial t^2}\right\|^2_{H^1(\Omega)} + 1\right]. \end{aligned}$$

with C_1 being of exactly the same form as above. We have used Lemma 2.2 and the estimates immediately preceding for $\partial W/\partial t$ and $(\partial W/\partial t)^2$. Thus,

$$\left\|\frac{\partial^2\eta}{\partial t^2}\right\|_{L^2(\partial\Omega)} \le C\left\|\frac{\partial^2\eta}{\partial t^2}\right\|_{H^1(\Omega)} \le C_1\left[\|u\|^2_{H^1(\Omega)} + \left\|\frac{\partial u}{\partial t}\right\|^2_{H^1(\Omega)} + \left\|\frac{\partial^2 u}{\partial t^2}\right\|_{H^1(\Omega)} + 1\right].$$

Thus, it is clear from the above estimate and (6.3) that

$$\begin{aligned} &\|\xi\|_{L^\infty_{\Delta t}(2\Delta t, T; L^2(\Omega))} + \|\xi\|^2_{L^2_{\Delta t}(2\Delta t, T; H^{\frac{1}{2}}_0(\Omega))} \\ &\qquad \le C'[h^{2k} + (\Delta t)^4] + C[\|\xi^2\|^2_{L^2(\Omega)} + \Delta t(\|\xi^{\frac{1}{2}}\|^2_{H^{\frac{1}{2}}_0(\Omega)} + \|\xi^{\frac{3}{2}}\|^2_{H^{\frac{1}{2}}_0(\Omega)})], \end{aligned} \tag{6.4}$$

with C' depending on precisely the set of norms of u, $\partial u/\partial t$, and $\partial^3 u/\partial t^3$ indicated in (5.12) and the $L^2(0, T; H^1(\Omega))$-norm of $\partial^2 u/\partial t^2$. Assume

$$\frac{\partial^2 u}{\partial t^2} \in L^2(0, T; H^1(\Omega)). \tag{6.5}$$

Let us consider the problem of specifying the three initial functions U^0, U^1, and U^2 that are needed to start up (1.15). Conceptually, the easiest method would be to use either an $L^2(\Omega)$ or a $H^1(\Omega)$ projection of f into $\mathcal{M}_h$ so that

$$\|\xi^0\|_{L^2(\Omega)} \le C\|f\|_{H^k(\Omega)} h^k, \qquad 1 \le k \le r,$$

and then to follow with two steps using the Crank-Nicolson Eq. (1.13), since (5.10) applied with $T=2\Delta t$ states that

$$\|\xi^2\|^2_{L^2(\Omega)} + (\|\xi^{\frac{1}{2}}\|^2_{H^1_0(\Omega)} + \|\xi^{\frac{3}{2}}\|^2_{H^1_0(\Omega)})\,\Delta t \leqq C^1\,[h^{2k} + (\Delta t)^4]. \tag{6.6}$$

To do this would involve solving nonlinear algebraic systems for U^1 and U^2; however, these systems can be iterated by successive substitution using almost exactly the same computer code as would be used for taking a time step by (1.15). Thus, no great loss of efficiency would occur if Crank-Nicolson were used for two steps. (This iteration converges for Δt sufficiently small; the constraint on Δt depends only on m, K, and C_T, but not on $\mathcal{M}$.)

Alternately, a predictor-corrector procedure based on (1.13) could be used for two steps to find U^1 and U^2. This amounts, of course, to taking two iterations in the successive substitution discussed above and accepting the second iterate as the solution. It can be shown that (6.6) is valid for U^1 and U^2 determined by the predictor-corrector procedure by a modest complication of the Crank-Nicolson argument, but this will be left to the reader. (In fact, the predictor-corrector procedure could be used as the basic computing method to give an algebraically linear procedure. The error estimates would be of exactly the same form as those for the three algebraically linear procedures we analyze here. Since the computing requirements would be twice those for the other three, we have not presented the predictor-corrector method in this paper.)

Given U^0, U^1, and U^2 by either of the methods mentioned above, we then have

$$\|\xi\|_{L^\infty_{\Delta t}(0,T;L^2(\Omega))} + \|\xi\|_{\tilde{L}^2_{\Delta t}(0,T;H^1_0(\Omega))} \leqq C'\,[h^k + (\Delta t)^2].$$

Consequently, the following theorem has been proved.

Theorem 6.1. Let U be the solution of (1.15) starting from values U^0, U^1, and U^2 satisfying (6.6). Assume (6.5) and the remaining hypotheses of Theorem 5.1. Then, there exists constants C and $C(s)$, $0\leqq s\leqq 1$, depending on the norms in (5.12) and (6.5) and the constant C^1 of (6.6) such that the error bounds (5.13) and (5.14) are valid.

It should be mentioned that the extrapolation

$$2g(t^{n+\frac{1}{2}}, U^{n-\frac{1}{2}}) - g(t^{n+\frac{1}{2}}, U^{n-\frac{3}{2}})$$

could just as easily have been employed as the one chosen. The results are the same and the proof is almost unaltered. Slightly less arithmetic is required per time step if U is first extrapolated and then g evaluated rather than vice versa if the flow rate g depends explicitly on the time, as it usually does in practice. This is why we chose our form of the extrapolation over the other.

7. Linearization About $u^{n-\frac{1}{2}}$

For $t=t^n$, $n\geqq 1$, let U be the solution of the equation (1.16) resulting from the linearization of g about the most recently known average of the solution at two successive time levels; i.e., about $U^{n-\frac{1}{2}}$. As in the last section, assume that the necessary initial values, namely U^0 and U^1, have been specified and assume the hypotheses of Theorem 6.1. Write the evolution equation for W in the form

$$\begin{aligned}
&\langle d_t W^n, v\rangle + \langle a \nabla W^{n+\frac12}, \nabla v\rangle \\
&\quad - \left\langle g(t^{n+\frac12}, W^{n-\frac12}) + \frac{\partial g}{\partial u}(t^{n+\frac12}, W^{n-\frac12})(W^{n+\frac12} - W^{n-\frac12}), v\right\rangle_{\partial\Omega} \\
&= -\langle d_t \eta^n, v\rangle + \lambda\langle \eta^{n+\frac12}, v\rangle - \langle g(t^{n+\frac12}, W^{n+\frac12}) - g(t, W)^{n+\frac12}, v\rangle_{\partial\Omega} \qquad (7.1)\\
&\quad + \left\langle d_t u^n - \frac{\partial u^{n+\frac12}}{\partial t}, v\right\rangle \\
&\quad + \left\langle g(t^{n+\frac12}, W^{n+\frac12}) - g(t^{n+\frac12}, W^{n-\frac12}) - \frac{\partial g}{\partial u}(t^{n+\frac12}, W^{n-\frac12})(W^{n+\frac12} - W^{n-\frac12}), v\right\rangle_{\partial\Omega}
\end{aligned}$$

for $v \in \mathcal{M}_h$. Only the first and last boundary terms in (7.1) are different from terms that we have already analyzed. We shall omit writing the time, since it is $t^{n+\frac12}$ in all cases. Then

$$\begin{aligned}
&\Big|\left\langle g(W^{n-\frac12}) + \frac{\partial g}{\partial u}(W^{n-\frac12})(W^{n+\frac12} - W^{n-\frac12}), \xi^{n+\frac12}\right\rangle_{\partial\Omega} \\
&\qquad - \left\langle g(U^{n-\frac12}) + \frac{\partial g}{\partial u}(U^{n-\frac12})(U^{n+\frac12} - U^{n-\frac12}), \xi^{n+\frac12}\right\rangle_{\partial\Omega}\Big| \\
&\quad \le K\|\xi^{n-\frac12}\|_{L^2(\partial\Omega)}\|\xi^{n+\frac12}\|_{L^2(\partial\Omega)} + \left|\left\langle \frac{\partial g}{\partial u}(U^{n-\frac12})(\xi^{n+\frac12} - \xi^{n-\frac12}), \xi^{n+\frac12}\right\rangle_{\partial\Omega}\right| \\
&\qquad + \left|\left\langle \left\{\frac{\partial g}{\partial u}(W^{n-\frac12}) - \frac{\partial g}{\partial u}(U^{n-\frac12})\right\}(W^{n+\frac12} - W^{n-\frac12}), \xi^{n+\frac12}\right\rangle_{\partial\Omega}\right| \\
&\quad \le \left(\frac12 K + \frac32 K_1\right)\|\xi^{n+\frac12}\|^2_{L^2(\partial\Omega)} + \left(\frac12 K + \frac12 K_1\right)\|\xi^{n-\frac12}\|^2_{L^2(\partial\Omega)} \qquad (7.2)\\
&\qquad + K_1\|W^{n+\frac12} - W^{n-\frac12}\|_{L^4(\partial\Omega)}\|\xi^{n-\frac12}\|_{L^{\frac83}(\partial\Omega)}\|\xi^{n+\frac12}\|_{L^{\frac83}(\partial\Omega)} \\
&\quad \le \left(\frac12 K + \frac32 K_1\right)\|\xi^{n+\frac12}\|^2_{L^2(\partial\Omega)} + \left(\frac12 K + \frac12 K_1\right)\|\xi^{n-\frac12}\|^2_{L^2(\partial\Omega)} \\
&\qquad + (\Delta t)^{\frac12}\left\|\frac{\partial W}{\partial t}\right\|_{L^2(t^{n-1}, t^{n+1}; L^4(\partial\Omega))}\big[(\|\xi^{n+\frac12}\|^2_{H^1(\Omega)} + \|\xi^{n-\frac12}\|^2_{H^1(\Omega)})\,\delta \\
&\qquad + C_\delta(\|\xi^{n+\frac12}\|^2_{L^2(\Omega)} + \|\xi^{n-\frac12}\|^2_{L^2(\Omega)})\big],
\end{aligned}$$

since the injection of $H^{\frac34}(\Omega)$ into $L^{\frac83}(\partial\Omega)$ is continuous, the injection of $H^1(\Omega)$ into $H^{\frac34}(\Omega)$ is compact, and

$$\begin{aligned}
\|W^{n+\frac12} - W^{n-\frac12}\|_{L^4(\partial\Omega)} &= \frac12\left\|\int_{t^{n-1}}^{t^{n+1}} \frac{\partial W}{\partial t}\,dt\right\|_{L^4(\partial\Omega)} \\
&\le 2^{-\frac12}(\Delta t)^{\frac12}\left\|\frac{\partial W}{\partial t}\right\|_{L^2(t^{n-1}, t^{n+1}; L^4(\partial\Omega))} \\
&\le C(\Delta t)^{\frac12}\left(\|u\|_{L^2(0,T;H^1(\Omega))} + \left\|\frac{\partial u}{\partial t}\right\|_{L^2(0,T;H^1(\Omega))}\right).
\end{aligned}$$

In particular, the coefficient of $\|\xi^{n+\frac12}\|^2_{H^1(\Omega)} + \|\xi^{n-\frac12}\|^2_{H^1(\Omega)}$ can be taken smaller than $m/8$.

Next, we need to see that the last term of (7.1) is $O\big((\Delta t)^4 + \|v\|^2_{L^2(\partial\Omega)}\big)$. This is a consequence of the following estimate:

$$\left\| g(W^{n+\frac{1}{2}}) - g(W^{n-\frac{1}{2}}) - \frac{\partial g}{\partial u}(W^{n-\frac{1}{2}})(W^{n+\frac{1}{2}} - W^{n-\frac{1}{2}})\right\|^2_{L^2(\partial\Omega)}$$
$$\leqq \frac{1}{4} K_1^2 \|(W^{n+\frac{1}{2}} - W^{n-\frac{1}{2}})^2\|^2_{L^2(\partial\Omega)}$$
$$\leqq \frac{1}{2} K_1 (\Delta t)^3 \left\|\left(\frac{\partial W}{\partial t}\right)^2\right\|^2_{L^2(t^{n-1}, t^{n+1}; L^2(\partial\Omega))}.$$

Hence,

$$\frac{1}{2} K_1 \sum_{n=1}^{T/\Delta t - 1} \left\|\left(\frac{\partial W}{\partial t}\right)^2\right\|^2_{L^2(t^{n-1}, t^{n+1}; L^2(\partial\Omega))} (\Delta t)^4 \leqq K_1 (\Delta t)^4 \left\|\left(\frac{\partial W}{\partial t}\right)^2\right\|^2_{L^2(0, T; L^2(\partial\Omega))}$$
$$\leqq C(\Delta t)^4 \left\|\frac{\partial W}{\partial t}\right\|^4_{L^4(0, T; H^1(\Omega))}.$$

It then follows that

$$\|\xi\|^2_{L^\infty_{\Delta t}(\Delta t, T; L^2(\Omega))} + \|\xi\|_{L^2_{\Delta t}(\Delta t, T; H_0^1(\Omega))}$$
$$\leqq C'[h^{2k} + (\Delta t)^4] + C[\|\xi^1\|^2_{L^2(\Omega)} + \Delta t \|\xi^{\frac{1}{2}}\|^2_{H_0^1(\Omega)}], \tag{7.3}$$

with C' again depending on the norms of $u, \ldots, \partial^3 u/\partial t^3$ listed in Theorem 6.1.

The evaluation of U^1 can again be accomplished by either solving the Crank-Nicolson equation for one step or by using a predictor-corrector form of Crank-Nicolson for one step. In either case,

$$\|\xi^1\|^2_{L^2(\Omega)} + \Delta t \|\xi^{\frac{1}{2}}\|^2_{H_0^1(\Omega)} \leqq C[h^{2k} + (\Delta t)^4]. \tag{7.4}$$

We have proved the following theorem.

Theorem 7.1. Let U be the solution of (1.16) starting from initial values U^0 and U^1 such that (7.4) holds. Then, the conclusions of Theorem 6.1 follow from its remaining hypotheses.

8. Extrapolation and Linearization

For $t = t^n$, $n \geqq 2$, let U denote the solution of (1.17), the result of linearizing about the projected value $2U^{n-\frac{1}{2}} - U^{n-\frac{3}{2}}$ in the evaluation of g at $t = t^{n+\frac{1}{2}}$. Assume that U^1 and U^2 are computed in such a way that

$$\|\xi^2\|^2_{L^2(\Omega)} + \Delta t (\|\xi^{\frac{1}{2}}\|^2_{H_0^1(\Omega)} + \|\xi^{\frac{3}{2}}\|^2_{H_0^1(\Omega)}) \leqq C[h^{2k} + (\Delta t)^4]. \tag{8.1}$$

Write the evolution equation for W in the form (g and $\partial g/\partial u$ being evaluated for $t = t^{n+\frac{1}{2}}$ everywhere but in $g(W)^{n+\frac{1}{2}}$)

$$\begin{aligned}
&\langle d_t W^n, v\rangle + \langle a \nabla W^{n+\frac{1}{2}}, \nabla v\rangle \\
&\quad - \left\langle g(2W^{n-\frac{1}{2}} - W^{n-\frac{3}{2}}) + \frac{\partial g}{\partial u}(2W^{n-\frac{1}{2}} - W^{n-\frac{3}{2}})(W^{n+\frac{1}{2}} - 2W^{n-\frac{1}{2}} + W^{n-\frac{3}{2}}), v\right\rangle_{\partial\Omega} \\
&= -\langle d_t \eta^n, v\rangle + \lambda\langle \eta^{n+\frac{1}{2}}, v\rangle - \langle g(W^{n+\frac{1}{2}}) - g(W)^{n+\frac{1}{2}}, v\rangle_{\partial\Omega} + \left\langle d_t u^n - \frac{\partial u}{\partial t}^{n+\frac{1}{2}}, v\right\rangle \\
&\quad + \left\langle g(W^{n+\frac{1}{2}}) - g(2W^{n-\frac{1}{2}} - W^{n-\frac{3}{2}}) - \frac{\partial g}{\partial u}(2W^{n-\frac{1}{2}} - W^{n-\frac{3}{2}}) \right. \\
&\quad \left. \cdot (W^{n+\frac{1}{2}} - 2W^{n-\frac{1}{2}} + W^{n-\frac{3}{2}}), v\right\rangle_{\partial\Omega}
\end{aligned} \tag{8.2}$$

for $v\in\mathcal{M}_h$. Again it is necessary to treat only the first and last boundary terms. The first of these leads to the following three terms:

$$\begin{aligned}
&|\langle g(2W^{n-\frac12}-W^{n-\frac32})-g(2U^{n-\frac12}-U^{n-\frac32}),\ \xi^{n+\frac12}\rangle_{\partial\Omega}|\\
&\qquad\le K\|\xi^{n+\frac12}\|_{L^2(\partial\Omega)}\|2\xi^{n-\frac12}-\xi^{n-\frac32}\|_{L^2(\partial\Omega)},\\
&\left|\left\langle\frac{\partial g}{\partial u}(2U^{n-\frac12}-U^{n-\frac32})(\xi^{n+\frac12}-2\xi^{n-\frac12}+\xi^{n-\frac32}),\ \xi^{n+\frac12}\right\rangle_{\partial\Omega}\right|\\
&\qquad\le K(\|\xi^{n+\frac12}\|_{L^2(\partial\Omega)}+2\|\xi^{n-\frac12}\|_{L^2(\partial\Omega)}+\|\xi^{n-\frac32}\|_{L^2(\partial\Omega)})\|\xi^{n+\frac12}\|_{L^2(\partial\Omega)},\\
&\left|\left\langle\left(\frac{\partial g}{\partial u}(2W^{n-\frac12}-W^{n-\frac32})-\frac{\partial g}{\partial u}(2U^{n-\frac12}-U^{n-\frac32})\right)(W^{n+\frac12}-2W^{n-\frac12}+W^{n-\frac32}),\ \xi^{n+\frac12}\right\rangle_{\partial\Omega}\right|\\
&\qquad\le 2K\|W^{n+\frac12}-2W^{n-\frac12}+W^{n-\frac32}\|_{L^2(\partial\Omega)}\|\xi^{n+\frac12}\|_{L^2(\partial\Omega)}\\
&\qquad\le\frac12\|\xi^{n+\frac12}\|^2_{L^2(\partial\Omega)}+C(\Delta t)^3\left\|\frac{\partial^2 W}{\partial t^2}\right\|^2_{L^2(t^{n-2},t^{n+1};L^2(\partial\Omega))}.
\end{aligned}$$

Notice that no delicacy is required; errors that are fourth order in Δt for sufficiently smooth u are easily shown to be second order. Since the basic Crank-Nicolson process is only second order, there is no gain attached to being more painstaking. Similarly, the last term can be treated a bit cavalierly:

$$\begin{aligned}
&\Big\|g(W^{n+\frac12})-g(2W^{n-\frac12}-W^{n-\frac32})\\
&\qquad-\frac{\partial g}{\partial u}(2W^{n-\frac12}-W^{n-\frac32})(W^{n+\frac12}-2W^{n-\frac12}+W^{n-\frac32})\Big\|^2_{L^2(\partial\Omega)}\\
&\le 2\|g(W^{n+\frac12})-g(2W^{n-\frac12}-W^{n-\frac32})\|^2_{L^2(\partial\Omega)}+2K^2\|W^{n+\frac12}-2W^{n-\frac12}+W^{n-\frac32}\|^2_{L^2(\partial\Omega)}\\
&\le C'(\Delta t)^4,
\end{aligned}$$

where C' depends on the same collection of norms of $u,\dots,\partial^3u/\partial t^3$ as in Theorem 6.1.

The usual argument again shows that

$$\begin{aligned}
&\|\xi\|^2_{L^\infty_{\Delta t}(2\Delta t,T;L^2(\Omega))}+\|\xi\|^2_{L^2_{\Delta t}(2\Delta t,T;H^1_0(\Omega))}\\
&\qquad\le C'[h^{2k}+(\Delta t)^4]+C[\|\xi^2\|^2_{L^2(\Omega)}+\Delta t(\|\xi^{\frac12}\|^2_{H^1_0(\Omega)}+\|\xi^{\frac32}\|^2_{H^1_0(\Omega)})],
\end{aligned}$$

and the usual theorem results.

Theorem 8.1. If U is the solution of (1.17) and U^0, U^1, and U^2 are chosen so that (8.1) holds, then the remaining hypotheses of Theorem 6.1 imply its conclusion.

Computationally, (1.17) is almost exactly equivalent to the extrapolation alone or linearization alone. Since it is a somewhat closer approximation to the Crank-Nicolson equation, the authors would tend to recommend it as the first choice for practical implementation. Obviously, there is nothing in the formal arguments given above to make us insistent on this recommendation.

References

1. Babuška, I.: A remark to the finite element method. Commentationes Mathematicae Universitatis Carolinae **12**, 367–376 (1971).
2. Birkhoff, G., Schultz, M. H., Varga, R. S.: Piecewise Hermite interpolation in one and two variables with applications to partial differential equations. Numer. Math. **11**, 232–256 (1968).

3. Bramble, J. H., Hilbert, S. R.: Bounds for a class of linear functionals with applications to Hermite interpolation. Numer. Math. **16**, 362–369 (1971).
4. Bramble, J. H., Schatz, A. H.: Rayleigh-Ritz-Galerkin methods for Dirichlet's problem using subspaces without boundary conditions. Comm. Pure and Appl. Math. **23**, 653–675 (1970).
5. Bramble, J. H., Zlámal, M.: Triangular elements in the finite element method. Math. Comp. **24**, 809–820 (1970).
6. Fix, G., Strang, G.: An analysis of the finite element method. Prentice-Hall, to appear.
7. Lions, J. L., Magenes, E.: Problèmes aux limites non homogènes et applications. Paris: Dunod 1968.
8. Nitsche, J.: Ein Kriterium für die Quasi-Optimalität des Ritzschen Verfahrens. Numer. Math. **11**, 346–348 (1968).
9. Nitsche, J.: Lineare Spline-Funktionen und die Methoden von Ritz für elliptische Randwertprobleme. Archive for Rational Mechanics and Analysis **36**, 348–355 (1970).
10. Wheeler, M. F.: Thesis, Rice University, 1971.

Jim Douglas, Jr.
Todd Dupont
Department of Mathematics
University of Chicago
Chicago, Illinois 60637/U.S.A.