JOHNSON, Joseph Peter, 1945-
EXTRAPOLATED GALERKEN METHODS FOR
PARABOLIC EQUATIONS.

Rice University, Ph.D., 1972
Mathematics

University Microfilms, A XEROX Company, Ann Arbor, Michigan
RICE UNIVERSITY

EXTRAPOLATED GALERKIN METHODS
FOR PARABOLIC EQUATIONS

by

Joseph Peter Johnson

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Thesis Director’s Signature:

Houston, Texas

December, 1971
ACKNOWLEDGEMENTS

I am indebted to Dr. Graeme Fairweather for suggesting this problem and for his advice and encouragement. I also wish to thank Professors Jim Douglas, Jr., Todd Dupont and H. H. Rachford, Jr. for helpful discussions throughout the course of this research. I am especially grateful to Mrs. Wanda Jones, who typed the manuscript accurately and in record time. In addition, I wish to thank the National Aeronautics and Space Administration and Rice University for financial support.
TABLE OF CONTENTS

Section 1  
Introduction and Summary of Results 1

Section 2  
The Galerkin Method and Extrapolation to the Limit 2

Section 3  
Notational Conventions and Definition of a Three-level Galerkin Method 6

Section 4  
Fundamental Lemmas 14

Section 5  
Linear Problems 27

Section 6  
Non-linear Problems 44

Appendix A  
The Existence of y(m) and z(m) in a Special Case 52

Appendix B  
Piecewise Hermite Polynomials 56

Appendix C  
A Special Result for the Case θ = ¼ 58

REFERENCES
DEDICATION

This thesis is dedicated to my parents.
1. Introduction and Summary of Results

In recent years much attention has been focused on the use of Galerkin methods for approximating solutions to parabolic problems [1, 2, 3, 4, 5, 13]. These methods reduce the problem of solving a parabolic partial differential equation to one of solving a system of ordinary differential equations, the latter problem then being approximated by some discrete method. Because the systems of ordinary differential equations are stiff, (see [2]), standard discrete variable methods are unsuitable for their approximate solution. Douglas and Dupont [4] and Wheeler [13] treat several special methods, discrete time Galerkin methods, and obtain error bounds which depend upon both the size of the discrete time step, $\Delta t$, and a term derived from approximation theory. While this latter term can easily be made small, none of the algorithms discussed in [4] and [13] yield time discretization errors smaller than $O((\Delta t)^2)$.

Extrapolation to the limit or Richardson's extrapolation provides a simple method for improving the accuracy of a given discrete variable method for the solution of ordinary differential equations. Although Gragg [8] proves the validity of extrapolation techniques when applied to a wide class of discrete variable methods, his error bounds are stated in terms of the Lipschitz constant of the equation being solved and are inadequate for a discussion of Galerkin methods as the ordinary differential equations arising have large Lipschitz constants by virtue of their stiffness.

This work proves the validity, under certain assumptions of smoothness, of extrapolation techniques as applied to a three level method presented in Dupont, Fairweather and Johnson [5]. Section 2 discusses Galerkin methods and extrapolation to the limit and summarizes certain results obtained by Gragg. Section 3 introduces the notation used throughout this thesis and presents the discrete time method to be considered. Section 4 consists of a series of lemmas which allow one to obtain global error estimates from local error estimates. Section 5 demonstrates the feasibility of repeated extrapolation for linear problems while Section 6 treats non-linear problems and demonstrates that one may extrapolate at least once, the conclusions of both sections requiring certain smoothness hypotheses. These smoothness conditions, the spaces of piecewise Hermite cubics and a three level method not treated elsewhere in the thesis, are briefly discussed in the appendices.
2. The Galerkin Method and Extrapolation to the Limit

The fundamental ideas involved in the Galerkin method for parabolic problems may be exhibited by considering a simple initial-boundary value problem for the heat equation:

(2.1a) \[ u_t - u_{xx} = 0 \text{ on } I \times [0,T] \]

(2.1b) \[ u(0,t) = u(1,t) = 0 \text{ for } t \in [0,T] \]

(2.1c) \[ u(x,0) = u_0(x) \text{ for } x \in I. \]

Here I is the open interval (0,1). Any solution, \( u(x,t) \), of problem (2.1) will be in the space \( H^1_0(I) \) for any fixed time \( t \). (See Section 3 for a definition of \( H^1_0 \).) Therefore an approximate solution, \( U(x,t) \), will be sought which is in \( S^h \), some finite dimensional subspace of \( H^1_0(I) \) for any fixed \( t \). If (2.1a) is multiplied by any element, \( v \), of \( H^1_0(I) \) and the resulting equation is integrated by parts, one obtains

\[
\int_0^1 u_t(x,t)v(x)dx + \int_0^1 u_x(x,t)v_x(x)dx = 0 \quad \forall v \in H^1_0(I).
\]

Analogously, \( U(x,t) \) will be required to satisfy

(2.2) \[
\int_0^1 U_t(x,t)v(x)dx + \int_0^1 U_x(x,t)v_x(x)dx = 0 \quad \forall v \in S^h.
\]

If \( v_1, \ldots, v_q \) is a basis for \( S^h \), then

\[ U(x,t) = \sum_{i=1}^q \alpha_i(t)v_i(t) \]

for some functions \( \alpha_1, \ldots, \alpha_q \) defined on \([0,T]\) and (2.2) becomes

(2.3) \[
\sum_{i=1}^q \alpha_i(t) \int_0^1 v_i(x)v_j(x)dx + \sum_{i=1}^q \alpha_i(t) \int_0^1 v_j^*(x)v_i^*(x)dx = 0 \quad \text{for } 1 \leq j \leq q.
\]

If \( \alpha = [\alpha_1, \ldots, \alpha_q]^T \), \( B = [b_{ij}] \) where

\[ b_{ij} = \int_0^1 v_i(x)v_j(x)dx \]
and \( A = [a_{ij}] \) where
\[
a_{ij} = \int_0^1 v_i(x)v_j(x)\,dx
\]
then \( B \) is non-singular since it is the Grammian of the linearly independent functions \( \{v_1, \ldots, v_q\} \). Thus (2.3) may be rewritten as
\[
(2.4) \quad \alpha' = -B^{-1}A\alpha.
\]
Since \(-B^{-1}A\alpha\) satisfies a Lipschitz condition with respect to \( \alpha \), the ordinary differential equation, (2.4), together with an initial value, \( \alpha(0) \), has a unique solution, \( \alpha(t) \), and thus the corresponding Galerkin approximation of \( u, U(x,t) \), exists uniquely. In practice (2.4) is usually solved by some discrete method. If \( \Delta t \) is the time step used, let \( U(\Delta t) \) denote the resulting approximation to \( U \). \( U(\Delta t) \) is therefore defined on all points \((x,j\Delta t)\) where \( x \in \bar{I} \) and \( j \) is an integer, \( 0 \leq j \leq T/\Delta t \).

Suppose that
\[
(2.5) \quad \|u - (\Delta t)^2y_2 - U(\Delta t)\| = O(A + (\Delta t)^4)
\]
where \( \| \cdot \| \) is some norm or semi-norm, \( y_2 \) is a function independent of \( \Delta t \) and \( A \) is a term depending upon how well \( u(\cdot,t) \) can be approximated by functions in \( S^h \) for arbitrary fixed values of \( t \). Then
\[
\|u - U(\Delta t)\| = \|u - (\Delta t)^2y_2 - U(\Delta t) + (\Delta t)^2y_2\|
\leq \|u - (\Delta t)^2y_2 - U(\Delta t)\| + \|((\Delta t)^2y_2)\|
= O(A + (\Delta t)^4) + O((\Delta t)^2)
= O(A + (\Delta t)^2),
\]
and we may say that \( U(\Delta t) \) approximates \( u \) with that portion of the error due to the time discretization being \( O((\Delta t)^2) \). Richardson [10] suggested a technique which allows one to improve upon the \( O((\Delta t)^2) \) error obtained when \( U(\Delta t) \) is used as the approximating function for \( u \). We define an extrapolate of \( U(\Delta t) \) to be
\[
U^{(1)}(\Delta t) = \frac{1}{3}(4U(\Delta t) - U(\Delta t))
\]
Then
\[
llu-U^{(1)}(\Delta t)ll = \frac{1}{3}(4u-u) - \frac{1}{4}(4U(\Delta t) - U(\Delta t))ll \\
= \frac{1}{3}(4u-u) - \frac{1}{4}(\Delta t^2)\gamma_2 - (\Delta t)^2\gamma_2 - \frac{1}{3}(4U(\Delta t) - U(\Delta t))ll \\
\leq \frac{4}{3}llu-(\Delta t)^2\gamma_2-U(\Delta t)ll + \frac{1}{3}llu-(\Delta t)^2\gamma_2-U(\Delta t)ll \\
= 0(A+(\Delta t)^4).
\]

Thus if the existence of a function \( \gamma_2 \) such that (2.1) holds can be demonstrated, a single extrapolation will suffice to provide an approximate solution to the parabolic problem which is fourth order correct in \( \Delta t \) instead of second order.

Typical choices of \( s^h \) produce an error term, \( A \), which is of the same order as \( (\Delta t)^4 \) or larger. (One such choice, the space of piecewise Hermite cubics, is discussed in appendix B.) Thus \( U^{(1)} \) will usually suffice as an approximation to \( u \). However, if for some \( \gamma_4 \) we have

(2.6)
\[
llu-(\Delta t)^2\gamma_2-(\Delta t)^4\gamma_4-U(\Delta t)ll = 0(A+(\Delta t)^6)
\]

then a short calculation shows that

\[
llu+\frac{3}{4}(\Delta t)^4\gamma_4-U^{(1)}(\Delta t)ll = 0(A+(\Delta t)^6),
\]

and a second extrapolate
\[
U^{(2)}(\Delta t) = \frac{1}{15}(16U^{(1)}(\Delta t) - U^{(1)}(\Delta t))
\]
is seen to produce an error which is \( 0(A+(\Delta t)^6) \). Obviously if there exist additional functions \( \gamma_6, \gamma_8, \ldots \) which account for contributions to the error which are \( 0((\Delta t)^6), 0((\Delta t)^8), \ldots \) respectively, further extrapolation may be carried out, and the \( k \)th extrapolate, defined by
\[
U^{(k)}(\Delta t) = \frac{4^k U^{(k-1)}(\Delta t) - U^{(k-1)}(\Delta t)}{4^k - 1},
\]
will provide an approximation to \( u \) with error \( 0(A+(\Delta t)^{2k+2}) \). Thus, for example, if an
algorithm having an error of $O(A+(\Delta t)^8)$ is desired, one need only use the third extrapolate of the basic second order algorithm.

It has been assumed above that some level of extrapolation would first be chosen and then $\Delta t$ would be chosen sufficiently small so as to obtain the desired accuracy. Gragg [8] has suggested that instead some initial value for $\Delta t$ be chosen and then extrapolation be performed repeatedly until an accurate approximation is obtained. Gragg shows that repeated extrapolation is a stable algorithm for ordinary differential equations provided that the error in an approximate solution can be expressed as a power series in the step size having bounded coefficients. Since this proof is completely independent of the Lipschitz constant of the differential equation in question, it can immediately be extended to apply to Galerkin approximates, the only major change being that one must take into account $A$, the term depending upon the approximation properties of $S^h$.

Heretofore it has been assumed that all extrapolations are based upon a sequence of functions $U(\Delta t), U(\frac{\Delta t}{2}), U(\frac{\Delta t}{4}), \cdots$. An obvious generalization is to use $U(\Delta t_1), U(\Delta t_2), \cdots$ where $\Delta t_1 > \Delta t_2 > \cdots$ and define

$$U(1)(\Delta t_i) = \frac{(\Delta t_i)^2 U(\Delta t_{i+1}) - (\Delta t_{i+1})^2 U(\Delta t_i)}{(\Delta t_i)^2 \Delta t_{i+1}}$$

with higher order extrapolates given by

$$U(k)(\Delta t_i) = \frac{(\Delta t_i)^2 U(k-1)(\Delta t_{i+1}) - (\Delta t_{i+k})^2 U(k-1)(\Delta t_i)}{(\Delta t_i)^2 \Delta t_{i+k}}$$

Gragg [7] discusses repeated extrapolation with these more general sequences of values for $\Delta t$ and shows that repeated extrapolation to the limit remains stable provided that

$$\sup_{n \geq 1} \frac{\Delta t_{n+1}}{\Delta t_n} < 1.$$

We see from the above that once $y_2, y_4, \cdots$ are shown to exist and satisfy inequalities such as (2.1) and (2.2), extrapolation is valid, and if the existence of suitably many uniformly bounded functions, $y_2, y_4, \cdots$, is assumed, repeated extrapolation to the limit is feasible. Thus the remaining sections will deal exclusively with the question of the existence of such functions.
3. Notational Conventions and Definition of a Three-level Galerkin Method

Let $T$ be a positive real number and let $\Omega$ be a bounded open set in $\mathbb{R}^q$ with boundary sufficiently smooth so as to allow integration by parts. (Dendy [1] shows that $\partial \Omega$ being piecewise $C^1$ and satisfying the restricted cone property is adequate.) Let $u$ be a real-valued function defined on $\mathcal{G}$ where $\mathcal{G} = \Omega \times (0, T)$, and suppose that $u$ satisfies

$$
(3.1) \quad \frac{\partial u}{\partial t} - \sum_{i,j=1}^{q} \frac{\partial}{\partial x_i} (a_{ij} \frac{\partial u}{\partial x_j}) - \sum_{i=1}^{q} b_i \frac{\partial u}{\partial x_i} - cu - f = 0 \text{ in } \mathcal{G}.
$$

This equation may be either linear, in which case $a_{ij}, b_i, c$ and $f$ are all defined on $\mathcal{G}$, or quasilinear, in which case the coefficients are defined on $\mathcal{G} \times \mathbb{R}$. In either case the coefficients are assumed to be bounded; the $a_{ij}$ are assumed to have bounded first derivatives; and, if (3.1) is quasilinear, the coefficients are assumed to be uniformly Lipschitz continuous with respect to $u$, that is

$$
l_a(x, t, u_1) - a_{ij}(x, t, u_2) l \leq K |u_1 - u_2| \text{ for all } u_1, u_2 \in \mathbb{R}
$$

where $K$ is some constant independent of $i, j, x$ and $t$ and similar inequalities are valid for $b_i, c$ and $f$. It is assumed that both $u$ and $\frac{\partial u}{\partial x_i}$ are bounded on $\mathcal{G}$ and that $u$ satisfies one of two boundary conditions, either the Dirichlet boundary condition

$$
(3.2) \quad u = 0 \text{ on } \partial \Omega \times (0, T)
$$

or the Neumann boundary condition

$$
(3.3) \quad \sum_{i,j=1}^{q} a_{ij} \frac{\partial u}{\partial x_j} \gamma_i = -g \text{ on } \partial \Omega \times (0, T)
$$

where $g$ is a given function defined on $\partial \Omega \times [0, T]$, the $\gamma_i$ are the components of the outward directed unit normal to $\partial \Omega$ and $\partial \Omega$ is the boundary of $\Omega$. In addition, an initial condition is prescribed for $u$,

$$
u(x, 0) = u_0(x) \quad \forall \ x \in \Omega .$$
Throughout it is assumed that there exists a constant \( c_o > 0 \) such that for all \( \xi \in \mathbb{R}^k \)

\[
\sum_{i,j=1}^{\ell} a_{ij} \xi_i \xi_j \geq c_o \sum_{i=1}^{\ell} \xi_i^2
\]

over the entire domain of the \( a_{ij} \).

Let

\[
L^2(\Omega) = \{ f : \int_{\Omega} f^2 \, dx < \infty \}
\]

\[
H^1(\Omega) = \{ f \in L^2(\Omega) : \frac{\partial f}{\partial x_i} \in L^2(\Omega) \text{ for } 1 \leq i \leq \ell \}
\]

\[
H^1_0(\Omega) = \{ f \in H^1(\Omega) : f = 0 \text{ on } \partial \Omega \},
\]

the corresponding norms being defined by

\[
\| f \|_{L^2(\Omega)} = \left( \int_{\Omega} f^2 \, dx \right)^{\frac{1}{2}} \text{ for } f \in L^2(\Omega)
\]

\[
\| f \|_{H^1(\Omega)} = \left( \int_{\Omega} f^2 \, dx + \sum_{i=1}^{\ell} \left\| \frac{\partial f}{\partial x_i} \right\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}} \text{ for } f \in H^1(\Omega)
\]

\[
\| f \|_{H^1_0(\Omega)} = \left( \sum_{i=1}^{\ell} \left\| \frac{\partial f}{\partial x_i} \right\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}} \text{ for } f \in H^1_0(\Omega).
\]

The inner product in \( L^2(\Omega) \) is written as

\[
\langle u, v \rangle = \int_{\Omega} uv \, dx.
\]

The inner product in \( L^2(\partial \Omega) \) is also of interest and is written

\[
\langle u, v \rangle = \int_{\partial \Omega} uv \, ds.
\]

Of interest also is the bilinear form \( a(w; \cdot, \cdot) \) which, for a given function \( w(x,t) \) defined on \( \overline{\Omega} \), is defined on \( H^1(\Omega) \times H^1(\Omega) \) by
\[
a(w;u,v) = \int_{\Omega} \sum_{i,j=1}^{Q} a_{ij}(x,t,w(x,t)) \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} \, dx \quad \text{where} \quad t \in [0,T].
\]

It follows immediately from (3.4) that

\[
(3.5) \quad a(w;u,u) \geq c_0 \|u\|_{H_0^1}^2
\]

for all \( t \in [0,T] \) and all functions \( w: \overline{\Omega} \to \mathbb{R} \). If the \( a_{ij} \) have domain \( \overline{\Omega} \), then we write \( a(u,v) \) instead of \( a(w;u,v) \).

If (3.1) is multiplied by any function \( v \in H_0^1(\Omega) \) and the resulting equation is integrated over \( \Omega \) one obtains

\[
\langle \frac{\partial u}{\partial t}, v \rangle - \int_{\Omega} \sum_{i,j=1}^{Q} \frac{\partial}{\partial x_j} (a_{ij} \frac{\partial u}{\partial x_j}) v \, dx - \sum_{i=1}^{Q} \langle b_i \frac{\partial u}{\partial x_i}, v \rangle - \langle cu, v \rangle - \langle f, v \rangle = 0.
\]

Provided the \( a_{ij} \) and \( u \) are sufficiently smooth one may apply a generalized Green's theorem (see Dendy [1]) to the second term in this expression to obtain

\[
\langle \frac{\partial u}{\partial t}, v \rangle + a(u;u,v) - \sum_{i=1}^{Q} \langle b_i \frac{\partial u}{\partial x_i}, v \rangle - \langle cu, v \rangle - \langle f, v \rangle = 0 \quad \forall \ v \in H_0^1(\Omega),
\]

the weak form of the Dirichlet problem. If, however, \( v \) is not restricted to lie in \( H_0^1(\Omega) \) and boundary condition (3.3) holds for some function \( g \) defined on \( \partial \Omega \times [0,T] \) then the boundary integral becomes \( (g,v) \) and one obtains

\[
\langle \frac{\partial u}{\partial t}, v \rangle + a(u;u,v) + (g,v) - \sum_{i=1}^{Q} \langle b_i \frac{\partial u}{\partial x_i}, v \rangle - \langle cu, v \rangle - \langle f, v \rangle = 0 \quad \forall \ v \in H^1(\Omega),
\]

the weak form of the Neumann problem.

For the Dirichlet problem one takes \( S^h \) to be a finite dimensional subspace of \( H_0^1(\Omega) \) while for the Neumann problem one takes \( S^h \) to be a finite dimensional subspace of \( H^1(\Omega) \). Then, in analogy with the weak forms above, one seeks a continuous time Galerkin approximation of \( u, U: \overline{\Omega} \to \mathbb{R} \), so that \( U(\cdot, t) \in S^h \quad \forall \ t \in [0,T] \) and
\[
\sum_{i=1}^{q} \langle b_i \frac{\partial U}{\partial x_i}, \nu \rangle - \langle cU, \nu \rangle - \langle f, \nu \rangle = 0 \ \forall \nu \in S^h
\]

in the case of the Dirichlet problem and

\[
\sum_{i=1}^{q} \langle b_i \frac{\partial U}{\partial x_i}, \nu \rangle + \langle g, \nu \rangle - \langle cU, \nu \rangle - \langle f, \nu \rangle = 0 \ \forall \nu \in S^h
\]

in the case of the Neumann problem. If \( U(x, 0) \) is chosen to be an approximation to \( u_0 \), then in either case, by use of techniques similar to those of the previous section, one can show that finding \( U \) is equivalent to solving an initial value problem for a system of ordinary differential equations. As in Section 2, this initial value problem may be shown to possess a unique solution.

Many discrete methods have been proposed for determining approximations to \( U \), [2, 3, 4]. Most of these are two level methods, i.e. they involve approximations to \( U \) at only two time levels. For the solution of the heat equation in one space variable, Descloux [2] uses multilevel methods based on methods of Gear [6] for solving stiff systems of ordinary differential equations. No a priori error estimates have been obtained for Descloux's methods. In [3], Douglas and Dupont describe a three-level method applicable to linear parabolic equations only. Unlike most multilevel methods, this procedure requires a starting procedure of the same accuracy as itself. In this thesis, the three-level method described in [5] will be used. The advantage of this procedure is that it requires the solution of a system of linear algebraic equations at each time step even when the differential equation is nonlinear. Also it requires only a simple starting procedure.

Finally, as is proven in this thesis, under certain conditions the error in the solutions obtained by this three level method may be expressed by inequalities similar to (2.5) and (2.6) and the solutions may therefore be extrapolated.

Let \( t_n = n\Delta t \) where \( \Delta t = T/N \) and \( N \) is an integer. If \( f \) is any function defined on \( \mathcal{G} \), then \( f_n \) is a function of \( x \) with \( f_n(x) = f(x, t_n) \). If \( f \) is a function with domain \( \mathcal{G} \times R \), then \( f(u_n) \) denotes the function of \( x \), \( f(x, t_n, u(x, t_n)) \). Similarly,

\[
a(w_n, u, v) = \int_{\Omega} \sum_{i,j=1}^{q} a_{ij}(x, t_n, u(x, t_n)) \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} dx.
\]
To define the time discrete Galerkin method studied hereafter, a parameter \( \theta > 0 \) is introduced. Then for any function \( f \) defined on \( \mathcal{S} \),

\[
f_{n\theta} = \begin{cases} 
(1-2\theta)f_0 + 2\theta f_1 & \text{if } n = 0 \\
\theta f_{n-1} + (1-2\theta)f_n + \theta f_{n+1} & \text{if } 1 \leq n \leq N-1.
\end{cases}
\]

It should be noted that \( a(u_{n\theta}, v) \) is evaluated with the functions \( a_{ij} \) evaluated at time \( t_n \).

The discrete Galerkin approximations of \( u \) will be a sequence of functions \( U_n \) where \( U_n \in \mathcal{S}^h \). First \( U_0 \) is defined by any convenient method in order to closely approximate \( u_0 \), interpolation in \( \mathcal{S}^h \) and \( L^2 \) projection onto \( \mathcal{S}^h \) being the most commonly used procedures. \( U_1 \) is determined from

\[
(3.6) \quad \langle U_1-U_0, v \rangle + (\Delta t)a(U_0; U_{0\theta}, v) - (\Delta t)(b(U_0) \cdot \nabla U_{0\theta} + c(U_0)U_0 + f(U_0), v) = 0 \quad \forall \ v \in \mathcal{S}^h
\]

in the case of Dirichlet boundary conditions and from

\[
(3.7) \quad \langle U_1-U_0, v \rangle + (\Delta t)a(U_0; U_{0\theta}, v) + (\Delta t)(g_{0\theta}, v)
- (\Delta t)(b(U_0) \cdot \nabla U_{0\theta} + c(U_0)U_0 + f(U_0), v) = 0 \quad \forall \ v \in \mathcal{S}^h
\]

in the case of Neumann boundary conditions. The questions of the existence and uniqueness of \( U_1 \) will be treated later. Then for \( 1 \leq n \leq N-1 \), \( U_{n+1} \) satisfies

\[
(3.8) \quad \langle U_{n+1}-U_{n-1}, v \rangle + (2\Delta t)a(U_n; U_{n\theta}, v)
- (2\Delta t)(b(U_n) \cdot \nabla U_{n\theta} + c(U_n)U_n + f(U_n), v) = 0 \quad \forall \ v \in \mathcal{S}^h
\]

in the case of Dirichlet boundary conditions and

\[
(3.9) \quad \langle U_{n+1}-U_{n-1}, v \rangle + (2\Delta t)a(U_n; U_{n\theta}, v) + (2\Delta t)(g_{n\theta}, v)
- (2\Delta t)(b(U_n) \cdot \nabla U_{n\theta} + c(U_n)U_n + f(U_n), v) = 0 \quad \forall \ v \in \mathcal{S}^h
\]

in the case of Neumann boundary conditions.
To prove the existence of $U_1$ satisfying (3.6) let \( \{v_1, \ldots, v_q\} \) be a basis for $S^h$. Then condition (3.6) is equivalent to

\[
\langle U_1 - U_0, v_i \rangle + (\Delta t)a(U_0; U_0, v_i) - (\Delta t)b(U_0) \cdot \nabla U_0 + c(U_0)U_0 + f(U_0), v_i \rangle = 0 \quad \text{for} \quad 1 \leq i \leq q.
\]

If $U_1$ exists it may be expressed as $U_1 = \sum_{i=1}^{q} \alpha_i v_i$ where $\alpha \in \mathbb{R}^q$. Then the existence and uniqueness of $U_1$ satisfying (3.6) is equivalent to the existence and uniqueness of $\alpha$ satisfying

\[(3.10) \quad G\alpha = \beta\]

where $G$ is a $q \times q$ matrix with

\[(3.11) \quad G_{ij} = \langle v_j, v_i \rangle + (2\theta \Delta t)a(U_0; v_j, v_i) - (2\theta \Delta t)b(U_0) \cdot \nabla v_j, v_i \rangle
\]

and

\[
\begin{align*}
\beta_j &= \langle U_0, v_i \rangle - (1 - 2\theta)(\Delta t)a(U_0; U_0, v_i) \\
&\quad + (\Delta t)(1 - 2\theta)b(U_0) \cdot \nabla U_0 + c(U_0)U_0 + f(U_0), v_i \rangle.
\end{align*}
\]

Equation (3.10) will have a unique solution for all $\beta \in \mathbb{R}^q$ provided that $G\gamma > 0$ for all non-zero $\gamma \in \mathbb{R}^q$ (The dot here denotes the usual inner product in $\mathbb{R}^q$). Given such a $\gamma$, let $w = \sum_{i=1}^{q} \gamma_i v_i$. Then by substitution into (3.1) we see that

\[
G\gamma \cdot \gamma = \langle w, w \rangle + (2\theta \Delta t)a(U_0; w, w) - (2\theta \Delta t)b(U_0) \cdot \nabla w, w \rangle
\]

\[
\geq ||w||_2^2 + (2\theta \Delta t)c_0 ||w||_1^2 - (2\theta \Delta t)b(U_0) \cdot \nabla w, w \rangle.
\]

In the following $C$ will denote a generic positive constant which may be chosen to be as small as desired. Now,
\[(b(U_0) \cdot \nabla w, w) \leq \|b(U_0) \cdot \nabla w\|_L^2 \|w\|_L^2 \leq C \|w\|_{H_0^1} \|w\|_{L^2} \leq c_0 \|w\|_{L^2}^2 + C \|w\|_{L^2}^2,\]

The last expression was obtained using the inequality \(ab \leq \varepsilon a^2 + \frac{1}{4\varepsilon} b^2\), which will be used repeatedly hereafter without comment. Thus

\[G \gamma \cdot \gamma \geq \|w\|_{L^2}^2 + (2\theta \Delta t)c_0 \|w\|_{H_0^1}^2 - (2\theta \Delta t)c_0 \|w\|_{H_0^1}^2 - C \Delta t \|w\|_{L^2}^2 = (1 - C \Delta t) \|w\|_{L^2}^2,\]

which, since \(\gamma\) and therefore \(w\) is not zero, is positive for sufficiently small \(\Delta t\), and the existence and uniqueness of \(U_1\) are both established. It is clear that the same techniques apply not only to (3.7) but also to (3.8) and (3.9).

For a discussion of errors in terms of the \(L^2(\Omega)\) norm some additional concepts are introduced. If \(k < m\) then an \(S^h_{k,m}\) space is any finite dimensional subspace, \(S^h\), of \(H^k(\Omega) \cap H^1_0(\Omega)\) having the property that for all non-negative integers \(j\) and \(l\) with \(l \leq k\) and \(l \leq j \leq m\) and for any \(v \in H^j(\Omega) \cap H^1_0(\Omega)\) there exists a constant \(Q\), independent of \(h\) and \(v\), such that

\[(3.12) \quad \inf_{w \in S^h} \|w - v\|_{H^l} \leq Q h^{j-l} \|v\|_{H^j}.\]

If \(S^h\) is not constrained to lie in \(H^1_0(\Omega)\) and (3.12) holds for arbitrary \(v \in H^j(\Omega)\), then \(S^h\) is said to be an \(S^h_{k,m}\) space.

If \(a(\cdot, \cdot)\) is the bilinear form defined above for linear problems, then \(a^*(\cdot, \cdot)\) is defined by \(a^*(u, v) = a(v, u)\). The form \(a^*\) is said to be zero regular on \(H^1_0(\Omega)\) if given \(f \in L^2(\Omega)\), any solution \(w \in H^1_0(\Omega)\) of

\[a^*(w, v) = (f, v) \quad \forall v \in H^1_0(\Omega)\]

is in \(H^2(\Omega)\).

Of much importance in the derivation of error estimates is the discrete analogue of Gronwall's inequality.

**Lemma (Discrete Gronwall's Inequality):**

Suppose \(\phi\), \(\psi\), and \(\chi\) are nonnegative functions defined on \(\{t_n: 0 \leq n \leq N\}\), the latter function being non-decreasing. If \(C\) is a positive constant and
\[ \phi(t_n) + \psi(t_n) \leq \chi(t_n) + C \Delta t \sum_{i=0}^{n-1} \phi(t_n) \text{ for } 0 \leq n \leq N \]

then

\[ \phi(t_n) + \psi(t_n) \leq \chi(t_n) e^{C t_n} \text{ for } 0 \leq n \leq N. \]

Proof: See Lees [9].
4. Fundamental Lemmas

Suppose that

\[ \xi_n(x) = u_n(x) - (\Delta t)^2 w_2(x,t_n) - \cdots - (\Delta t)^M w_M(x,t_n) \]

and that \( \|\xi_n - u_n\| \), where \( \|\cdot\| \) is some norm, is bounded by \( C(\Delta t)^{M+2} \) plus a term which depends upon the approximation properties of the subspace, \( S^h \), chosen. It then follows from the results of Section 2 that the values of \( u_n \) obtained for successively smaller values of \( \Delta t \) may be extrapolated in the hope of obtaining a better approximation for \( u_n \). In Sections 5 and 6 a function \( \xi_n \) having an expansion similar to that above is introduced. It is further shown that the \( \xi_n \) satisfy equations which are the same as the defining relations for the functions \( u_n \), except that to the right-hand side is added a term of the form \( \langle q_0, v \rangle \) in the case of the equation giving \( U_1 \) and a term of the form \( (\Delta t)\langle q, q \rangle \) in the case of the equation giving \( U_{n+1} \) for \( n \geq 1 \) where in either case \( \|q_n\|_{L^2} = O((\Delta t)^2) \). The following lemmas show that this local error estimate extends to a global error estimate of the form \( O((\Delta t)^{M+2}) \) plus an approximation error dependent upon the choice of \( S^h \).

The first two lemmas deal with the case involving Dirichlet boundary conditions.

**Lemma 4.1:** Let \( \xi_0, \xi_1, \ldots, \xi_N \) be a sequence of elements of \( H^1_0(\Omega) \) such that \( \|\xi_n\|_{L^\infty} \) and \( \|\nabla \xi_n\|_{L^\infty} \) are both bounded and suppose that

\[
(\xi_{n+1} - \xi_{n-1}, v) + (2\Delta t)a(\xi_n; \xi_{n\theta}; v) \]

\[
- (2\Delta t)b(\xi_n; \nabla \xi_{n\theta} + c(\xi_n) \xi_n + f(\xi_n), v) \]

\[
= (\Delta t)\langle q, q \rangle \quad \forall \ v \in H^1_0(\Omega), \ n = 1, \ldots, N-1. \]

Let \( U_0, U_1, \ldots, U_N \) be a sequence of elements of \( S^h \subset H^1_0(\Omega) \) satisfying (3.8). Define \( e_n = \xi_n - U_n \) and, if \( \tilde{\xi}_n \) is an arbitrary element of \( S^h \), define \( \bar{e}_n = \tilde{\xi}_n - \xi_n \). Then if \( \theta > \frac{\pi}{4} \) and \( \Delta t \) is sufficiently small,

\[
\|e_N\|_{L^2}^2 + \beta\Delta t \sum_{n=1}^{N-1} \|e_n\|_{H^1_0}^2 \leq C(\Delta t) \sum_{n=2}^{N-2} \|e_{n+1,\theta} + \bar{e}_{n-1,\theta}\|_{L^2}^2 \]

\[
+ (\Delta t) \sum_{n=1}^{N-1} \left( \|\bar{e}_n\|_{H^1}^2 + \|q_n\|_{L^2}^2 \right) + \max_{1 \leq n \leq N-1} \|\bar{e}_n\|_{L^2}^2 + \|e_0\|_{L^2}^2 + \|\xi_1\|_{L^2}^2 \]
where \( \beta \) is a positive constant independent of \( N \) and \( S^h \).

**Proof:** Subtracting (3.8) from (4.1) and adding \((2\Delta t)a(U_n;\xi_{n\theta},v)\) to both sides of the resulting equation gives, for all \( v \in S^h \),

\[
(4.3) \quad \langle e_{n+1} - e_{n-1}, e_{n\theta} \rangle + (2\Delta t)a(U_n;e_{n\theta},v) = (2\Delta t) [a(U_n;\xi_{n\theta},v) - a(\xi_n;\xi_{n\theta},v)]
\]

\[+ (2\Delta t) \left[ \langle b(\xi_n) \cdot \nabla \xi_{n\theta} - b(U_n) \cdot \nabla U_{n\theta}, v \rangle + \langle c(\xi_n) \xi_n - c(U_n) U_n, v \rangle \right] + \langle f(\xi_n) - f(U_n), v \rangle + (\Delta t) \langle \eta_n, v \rangle \]

\[= (2\Delta t) [a(U_n;\xi_{n\theta},v) - a(\xi_n;\xi_{n\theta},v)]
\]

\[+ (2\Delta t) \left[ \langle b(\xi_n) \cdot \nabla \xi_{n\theta} - b(U_n) \cdot \nabla U_{n\theta}, v \rangle + \langle b(U_n) \cdot \nabla \xi_{n\theta} - b(U_n) \cdot \nabla U_{n\theta}, v \rangle \right]
\]

\[+ (2\Delta t) \left[ \langle c(\xi_n) \xi_n - c(U_n) \xi_n, v \rangle + \langle c(U_n) \xi_n - c(U_n) U_n, v \rangle \right]
\]

\[+ (2\Delta t) \langle f(\xi_n) - f(U_n), v \rangle + (\Delta t) \langle \eta_n, v \rangle \].

Now let \( v = e_{n\theta} + \bar{e}_{n\theta} \). Then

\[
\langle e_{n+1} - e_{n-1}, e_{n\theta} \rangle + (2\Delta t)a(U_n;e_{n\theta},e_{n\theta}) = \langle e_{n-1} - e_{n+1}, \bar{e}_{n\theta} \rangle
\]

\[- (2\Delta t)a(U_n;e_{n\theta},\bar{e}_{n\theta})
\]

\[+ (2\Delta t) [a(U_n;\xi_{n\theta},e_{n\theta}+\bar{e}_{n\theta}) - a(\xi_n;\xi_{n\theta},e_{n\theta}+\bar{e}_{n\theta})] \]

\[+ (2\Delta t) \langle b(\xi_n) - b(U_n) \cdot \nabla \xi_{n\theta} e_{n\theta} + \bar{e}_{n\theta} \rangle
\]

\[+ (2\Delta t) \langle b(U_n) \cdot \nabla \xi_{n\theta} e_{n\theta} + \bar{e}_{n\theta} \rangle + (2\Delta t) \langle c(U_n) e_{n\theta}+\bar{e}_{n\theta} \rangle
\]

\[+ (2\Delta t) \langle c(\xi_n) - c(U_n) \cdot \xi_n e_{n\theta} + \bar{e}_{n\theta} \rangle + (2\Delta t) \langle f(\xi_n) - f(U_n), e_{n\theta}+\bar{e}_{n\theta} \rangle
\]

\[+ (\Delta t) \langle \eta_n, e_{n\theta}+\bar{e}_{n\theta} \rangle \].
\[
\leq \langle e_{n-1} - e_{n+1}, \bar{e}_n \theta \rangle + (C \Delta t) \left[ \| \ell e_n \theta \|_{H_0}^2 \right. \\
+ \left. \| \ell e_n \theta \|_{L^2}^2 \right] \\
+ \langle e_{n-1} - e_{n+1}, \bar{e}_n \theta \rangle + (C \Delta t) \left[ \| \ell e_n \theta \|_{H_0}^2 \right. \\
+ \left. \| \ell e_n \theta \|_{L^2}^2 \right] \\
+ \langle e_{n-1} - e_{n+1}, \bar{e}_n \theta \rangle + (C \Delta t) \left[ \| \ell e_n \theta \|_{H_0}^2 \right. \\
+ \left. \| \ell e_n \theta \|_{L^2}^2 \right] \\
+ \langle e_{n-1} - e_{n+1}, \bar{e}_n \theta \rangle + (C \Delta t) \left[ \| \ell e_n \theta \|_{H_0}^2 \right. \\
+ \left. \| \ell e_n \theta \|_{L^2}^2 \right] \\
+ \langle e_{n-1} - e_{n+1}, \bar{e}_n \theta \rangle + (C \Delta t) \left[ \| \ell e_n \theta \|_{H_0}^2 \right. \\
+ \left. \| \ell e_n \theta \|_{L^2}^2 \right] \\
+ \langle e_{n-1} - e_{n+1}, \bar{e}_n \theta \rangle + (C \Delta t) \left[ \| \ell e_n \theta \|_{H_0}^2 \right. \\
+ \left. \| \ell e_n \theta \|_{L^2}^2 \right] \\
+ \langle e_{n-1} - e_{n+1}, \bar{e}_n \theta \rangle + (C \Delta t) \left[ \| \ell e_n \theta \|_{H_0}^2 \right. \\
+ \left. \| \ell e_n \theta \|_{L^2}^2 \right].
\]

Therefore

\[
\theta (\| \ell e_{n+1} \|_{L^2}^2 - \| \ell e_{n-1} \|_{L^2}^2) + (1 - 2\theta) (\langle e_{n+1}, e_n \rangle - \langle e_n, e_{n-1} \rangle) + (C \Delta t) \beta \| \ell e_n \theta \|_{H_0}^2
\]

\[
\leq \langle e_{n-1} - e_{n+1}, \bar{e}_n \theta \rangle + (C \Delta t) \left[ \| \ell \bar{e}_n \theta \|_{H_1}^2 + \| \ell e_{n+1} \|_{L^2}^2 + \| \ell e_n \|_{L^2}^2 + \| \ell e_{n-1} \|_{L^2}^2 + \| \ell \eta_n \|_{L^2}^2 \right]
\]

where \( \beta \) is some generic positive constant. Now sum both sides from \( n = 1 \) to \( N - 1 \).

\[
\theta (\| \ell e_N \|_{L^2}^2 + \| \ell e_{N-1} \|_{L^2}^2 - \| \ell e_1 \|_{L^2}^2) + \sum_{n=1}^{N-1} \| \ell e_n \theta \|_{H_0}^2
\]

\[
\leq \sum_{n=1}^{N-1} \left[ \langle e_{n-1} - e_{n+1}, \bar{e}_n \theta \rangle + (C \Delta t) \left[ \| \ell \bar{e}_n \theta \|_{H_1}^2 + \| \ell \eta_n \|_{L^2}^2 \right] \right] + (C \Delta t) \sum_{n=0}^{N} \| \ell e_n \|_{L^2}^2.
\]

Now, \((1 - 2\theta) \langle e_{N-1}, e_{N-1} \rangle \geq -\frac{|1 - 2\theta|}{2} (\| \ell e_N \|_{L^2}^2 + \| \ell e_{N-1} \|_{L^2}^2)\).

Also, \[
\sum_{n=1}^{N-1} \langle e_{n-1} - e_{n+1}, \bar{e}_n \theta \rangle = -\langle e_{N-1}, \bar{e}_{N-2}, 0 \rangle - \langle e_N, \bar{e}_{N-1}, 0 \rangle
\]

\[
+ (\Delta t) \sum_{n=2}^{N-2} \frac{\langle e_{n+1} \theta - e_{n-1} \theta \rangle}{\Delta t} + \langle e_{N}, \bar{e}_{1} \theta \rangle + \langle e_{1}, \bar{e}_{2} \theta \rangle.
\]
\[
\begin{align*}
\text{But} \quad \sum_{n=2}^{N-2} \langle \varepsilon_n, \frac{\tilde{e}_{n+1} - \tilde{e}_{n-1}}{\Delta t} \rangle & \leq \sum_{n=2}^{N-2} \left( \|e_n\|_{L_2}^2 + \frac{\|\tilde{e}_{n+1} - \tilde{e}_{n-1}\|_{L_2}^2}{\Delta t} \right); \\
- \langle e_{N-1}, \tilde{e}_{N-2}, \theta \rangle & \leq \varepsilon \|e_{N-1}\|_{L_2}^2 + C \|\tilde{e}_{N-2}, \theta\|_{L_2}^2; \\
- \langle e_{N}, \tilde{e}_{N-1}, \theta \rangle & \leq \varepsilon \|e_{N}\|_{L_2}^2 + C \|\tilde{e}_{N-1}, \theta\|_{L_2}^2; \\
\text{and} \quad \langle e_0, \tilde{e}_0 \rangle + \langle e_1, \tilde{e}_0 \rangle & \leq \|e_0\|_{L_2}^2 + \|e_1\|_{L_2}^2 + \|\tilde{e}_1\|_{L_2}^2 + \|\tilde{e}_2, \theta\|_{L_2}^2.
\end{align*}
\]

Therefore
\[
\begin{align*}
(\theta - \frac{1 - 2\theta}{2}) \|e_N\|_{L_2}^2 + (\theta - \frac{1 - 2\theta}{2}) \|e_{N-1}\|_{L_2}^2 + (\beta \Delta t) \sum_{n=1}^{N-1} \|e_n\|_{H_0}^2 & \\
\leq (C \Delta t) \sum_{n=2}^{N-2} \|\tilde{e}_{n+1} - \tilde{e}_{n-1}\|_{L_2}^2 + (C \Delta t) \sum_{n=0}^{N-1} \|e_n\|_{L_2}^2 \\
+ (C \Delta t) \sum_{n=1}^{N-1} [\|\tilde{e}_n\|_{H_1}^2 + \|e_n\|_{L_2}^2] \\
+ C[\|\tilde{e}_{N-2}, \theta\|_{L_2}^2 + \|\tilde{e}_{N-1}, \theta\|_{L_2}^2 + \|\tilde{e}_1\|_{L_2}^2 + \|\tilde{e}_2, \theta\|_{L_2}^2 + \|e_0\|_{L_2}^2 + \|e_1\|_{L_2}^2].
\end{align*}
\]

If \( \theta > \frac{1}{4} \) then \( \theta - \frac{1 - 2\theta}{2} > 0 \) and for sufficiently small \( \varepsilon \) and \( \Delta t \) one obtains
\[
\begin{align*}
\|e_N\|_{L_2}^2 + (\beta \Delta t) \sum_{n=1}^{N-1} \|e_n\|_{H_0}^2 & \leq C[(\Delta t) \sum_{n=2}^{N-2} \|\tilde{e}_{n+1} - \tilde{e}_{n-1}\|_{L_2}^2 \\
+ (\Delta t) \sum_{n=1}^{N-1} [\|\tilde{e}_n\|_{H_1}^2 + \|e_n\|_{L_2}^2] + \max_{1 \leq n \leq N-1} \|\tilde{e}_n\|_{L_2}^2] \\
+ (C \Delta t) \sum_{n=0}^{N-1} \|e_n\|_{L_2}^2 + C[\|e_0\|_{L_2}^2 + \|e_1\|_{L_2}^2].
\end{align*}
\]
Since no use of the implicit assumption that \( N = T/\Delta t \) has been made, the above inequality holds for all integers between 2 and \( N \) and thus (4.2) follows by an application of the discrete Gronwall's inequality.

For (4.2) to be a useful estimate of \( \| e \|_{L^2}^2 \), estimates of \( \| e_0 \|_{L^2}^2 \) and \( \| e_1 \|_{L^2}^2 \) are required. Clearly the magnitude of \( \| e_0 \|_{L^2}^2 \) will depend upon the ability of elements of \( S^h \) to closely approximate \( \xi \) in the \( L^2 \) norm. The following lemma provides an estimate of \( \| e_1 \|_{L^2}^2 \).

**Lemma 4.2:** Suppose \( \xi_0 \) and \( \xi_1 \) are elements of \( H^1_0(\Omega) \) such that \( \| \xi_n \|_{L^\infty} \) and \( \| \nabla \xi_n \|_{L^\infty} \) are bounded for \( n = 0,1 \) and suppose that

\[
(4.4) \quad \langle \xi_1 - \xi_0, v \rangle + (\Delta t) a(\xi_0, \xi_0, v) - (\Delta t) b(\xi_0) \cdot \nabla \xi_0 + c(\xi_0) \xi_0 + f(\xi_0), v \rangle
= \langle \eta_0, v \rangle \quad \forall \ v \in H^1_0(\Omega)
\]

and suppose \( U_0 \) and \( U_1 \) satisfy (3.6) where \( U_0 \) and \( U_1 \) are elements of \( S^h \subset H^1_0(\Omega) \). Let \( e_n \) and \( \bar{e}_n \) be defined as in Lemma 4.1. Then if \( \theta > 0 \) and \( \Delta t \) is sufficiently small,

\[
(4.5) \quad \| e_1 \|_{L^2}^2 + (\Delta t) \beta \| e_0 \|_{H^1_0}^2 \leq C [ \| e_0 \|_{L^2}^2 + \| \bar{e}_0 \|_{L^2}^2 + \| \eta_0 \|_{L^2}^2 ]
\]

where \( \beta \) is a non-negative constant.

**Proof:** Subtracting (3.6) from (4.4) and adding \( (\Delta t) a(U_0, \xi_0, v) \) to both sides of the resulting equation gives \( \forall \ v \in S^h \)

\[
(4.6) \quad \langle e_1 - e_0, v \rangle + (\Delta t) a(U_0, e_0, v) = (\Delta t) [ a(U_0, \xi_0, v) - a(\xi_0, \xi_0, v) ]
+ (\Delta t) [ b(\xi_0) \cdot \nabla \xi_0 - (b(U_0) \cdot \nabla \xi_0, v) ]
+ (\Delta t) [ c(\xi_0) \xi_0 - c(U_0) \xi_0, v \rangle + (\Delta t) f(\xi_0) - f(U_0), v \rangle + \langle \eta_0, v \rangle
= (\Delta t) [ a(U_0, \xi_0, v) - a(\xi_0, \xi_0, v) ]
+ (\Delta t) [ b(\xi_0) \cdot \nabla \xi_0 - b(U_0) \cdot \nabla \xi_0, v \rangle + (\Delta t) b(U_0) \nabla e_0, v \rangle
+ (\Delta t) [ c(U_0) \xi_0 - c(U_0) \xi_0, v \rangle + (\Delta t) f(\xi_0) - f(U_0), v \rangle + \langle \eta_0, v \rangle.
\]
Now let \( v = e_0 + \bar{e}_0 \in S^h \). Then

\[
\begin{align*}
&\langle e_1 - e_0, 2\theta e_1 + (1-2\theta)e_0 \rangle + (\Delta t) a(U_o^e e_0, e_0) \\
&\quad - (\Delta t) a(U_o^e e_0, \bar{e}_0) + (C\Delta t) [H_1 \| e_0 \|_{L^2} + H_0 \| e_0 + \bar{e}_0 \|_{H^1} + H_0 \| e_0 + \bar{e}_0 \|_{L^2} + H_0 \| e_0 + \bar{e}_0 \|_{L^2}] + H_1 \| e_0 + \bar{e}_0 \|_{L^2} + H_0 \| e_0 + \bar{e}_0 \|_{L^2}.
\end{align*}
\]

Thus

\[
\begin{align*}
&2\theta \| e_1 \|_{L^2}^2 + (1-\theta) \langle e_1, e_0 \rangle - (1-2\theta) \| e_0 \|_{L^2}^2 + (\Delta t) a(U_o^e e_0, \bar{e}_0) \\
&\quad + \| e_0 \|_{L^2}^2 + \| \bar{e}_0 \|_{L^2} + (C\Delta t) [H_1 \| e_0 \|_{H^1} + H_0 \| e_0 + \bar{e}_0 \|_{H^1} + H_0 \| e_0 + \bar{e}_0 \|_{L^2} + H_0 \| e_0 + \bar{e}_0 \|_{L^2}].
\end{align*}
\]

Therefore

\[
\begin{align*}
&(2\theta - \epsilon) \| e_1 \|_{L^2}^2 + (\Delta t) a(U_o^e e_0, \bar{e}_0) \\
&\quad + (\Delta t) [H_1 \| e_0 \|_{H^1} + H_0 \| e_0 + \bar{e}_0 \|_{L^2}] + (C\Delta t) \| e_1 \|_{L^2} + (\epsilon \Delta t) \| e_0 \|_{H^1}.
\end{align*}
\]

Now we can use the fact that \( \Delta t \) is bounded to obtain

\[
\begin{align*}
&(2\theta - \epsilon - C\Delta t) \| e_1 \|_{L^2}^2 + (\Delta t) (a(U_o^e e_0, \bar{e}_0) \\
&\quad + (\Delta t) [H_1 \| e_0 \|_{H^1} + H_0 \| e_0 + \bar{e}_0 \|_{L^2}] + (C\Delta t) \| e_1 \|_{H^1} + (\epsilon \Delta t) \| e_0 \|_{H^1}.
\end{align*}
\]

which, for \( \Delta t \) and \( \epsilon \) sufficiently small, implies (4.5).

The analogous results for the Neumann problem are proved in an almost identical manner. For this reason they are stated as corollaries.

**Corollary 4.1:** Let \( \xi_0, \xi_1, \ldots, \xi_N \) be a sequence of elements of \( H^1(\Omega) \) such that \( \| \xi_n \|_{L^\infty} \) and \( \| \nabla \xi_n \|_{L^\infty} \) are bounded and suppose that
\begin{align}
(4.7) \quad & \langle \xi_{n+1} - \xi_{n-1}, v \rangle + (2\Delta t) a(\xi_{n}, \xi_{n\theta}, v) + (\Delta t) (g_{n\theta}, v) \\
& - (2\Delta t) \langle b(\xi_{n}) \cdot \nabla \xi_{n\theta} + c(\xi_{n}) \xi_{n} + f(\xi_{n}), v \rangle \\
& = (\Delta t) \langle \eta_{n}, v \rangle \quad \forall v \in H^{1}(\Omega), \quad n=1, \ldots, N-1.
\end{align}

Let \( U_{0}, U_{1}, \ldots, U_{N} \) be a sequence of elements of \( S^{h} \subset H^{1}(\Omega) \) satisfying (3.9). Define \( e_{n} \) and \( \bar{e}_{n} \) as in Lemma 4.1. Then if \( \theta > \frac{1}{4} \) and \( \Delta t \) is sufficiently small, (4.2) is again valid.

**Proof:** If one subtracts (3.9) from (4.7) one obtains (4.3). The remainder of the proof of Lemma 4.1 is now valid without change.

**Corollary 4.2:** Suppose \( \xi_{0} \) and \( \xi_{1} \) are elements of \( H^{1}(\Omega) \) such that \( \|\xi_{n}\|_{L^{\infty}} \) and \( \|\nabla \xi_{n}\|_{L^{\infty}} \) are bounded for \( n = 0, 1 \) and

\begin{align}
(4.8) \quad & \langle \xi_{1} - \xi_{0}, v \rangle + (\Delta t) a(\xi_{0}, \xi_{0\theta}, v) + (g_{0\theta}, v) \\
& - (\Delta t) \langle b(\xi_{0}) \cdot \nabla \xi_{0\theta} + c(\xi_{0}) \xi_{0} + f(\xi_{0}), v \rangle = \langle \eta_{0}, v \rangle \quad \forall v \in H^{1}(\Omega)
\end{align}

and suppose \( U_{0} \) and \( U_{1} \) satisfy (3.7) where \( U_{0} \) and \( U_{1} \) are elements of \( S^{h} \subset H^{1}(\Omega) \). Let \( e_{n} \) and \( \bar{e}_{n} \) be defined as in Lemma 4.1. Then if \( \theta > 0 \) and \( \Delta t \) is sufficiently small, (4.5) is again valid.

**Proof:** Subtracting (3.7) from (4.8) gives (4.6). The remainder of the proof of Lemma 4.2 then applies without change.

It should be noted that while Corollaries 4.1 and 4.2 have been proved for the case in which the \( a_{ij}(x, t, u) \) depend upon both \( t \) and \( u \), expressions of the form (4.7) and (4.8) will be shown in Sections 5 and 6 to hold only in the case in which the \( a_{ij} \) are functions of \( x \) alone.

The error bounds given in Lemmas 4.1 and 4.2 give bounds on \( \|e_{n}\|_{L^{2}} \) in terms of \( \|e_{n}\|_{H^{1}} \), \( n=0, \ldots, N \). If \( S^{h} \) is, for example, an \( S_{k,m}^{h} \) space, one can in general expect \( \|e_{n}\|_{H^{1}} \) to be no smaller than \( C h^{m-1} \), even if \( \xi_{n} \) is very smooth. Since proper choice of
$\xi_n \in \mathcal{S}^h$, in this case, gives $\|\tilde{\xi}_n\|^2_{L^2} = O(h^n)$, one would hope to be able to give bounds on $\|\tilde{e}_n\|^2_{L^2}$ in terms of $\|\tilde{e}_n\|^2_{L^2}$, $n=0, \ldots, N$, instead. Wheeler [13] gives such bounds for several time-discrete Galerkin methods and her results can be adapted to the three level methods under consideration here. The following arguments restrict the $a_{ij}$ to be functions of $x$ only; a theory for the more general case has not been obtained.

**Lemma 4.3:** In addition to the hypotheses of Lemma 4.1 assume that the $a_{ij}$ are functions of $x$ alone and that $\frac{\partial b_i(\xi)}{\partial x_j}$ is bounded. Also assume that for each integer $n$ with $0 \leq n \leq N$, $\tilde{\xi}_n$ has the following properties:

\[(4.9) \quad a(\tilde{\xi}_{n+1}, \tilde{\xi}_n, v) = 0 \quad \forall \quad v \in \mathcal{S}^h\]

\[(4.10) \quad \|\tilde{\xi}_{n+1} - \tilde{\xi}_n\|^2_{L^2} = O(h^n) \quad \text{where} \quad h \text{ is a parameter, } 0 < h \leq 1,\]

\[(4.11) \quad \Delta t \sum_{n=1}^{N-1} \left\| \frac{\tilde{\xi}_{n+1} - \tilde{\xi}_n}{\Delta t} - (\tilde{\xi}_{n+1} - \tilde{\xi}_n) \right\|_{L^2} = O(h^n)\]

and $\triangle \tilde{\xi}_n$ and therefore $|\tilde{\xi}_n|$ are both bounded functions of $x$ with bound independent of $h, n$ and $\Delta t$. Then

\[(4.12) \quad \|\tilde{e}_n\|^2_{L^2} \leq C(h^{2r} + \|\tilde{e}_0\|^2_{L^2} + \|\tilde{z}_n\|^2_{L^2} + \Delta t \sum_{n=1}^{N-1} \|\tilde{e}_n\|^2_{L^2})\]

where $z_n = \tilde{\xi}_n - U_n$.

**Proof:** Using (4.9) together with (4.1) gives

\[
\langle \tilde{\xi}_{n+1} - \tilde{\xi}_{n-1}, v \rangle + (2\Delta t)a(\tilde{\xi}_{n+1}, v) - (2\Delta t)b(\tilde{\xi}_n, \nabla \tilde{\xi}_{n+1} + c(\tilde{\xi}_n, \tilde{\xi}_n + f(\xi)_n), v)
\]

\[
= (\Delta t) \langle \eta, v \rangle + \langle (\tilde{\xi}_{n+1} - \tilde{\xi}_{n-1}) - (\tilde{\xi}_{n-1} - \tilde{\xi}_{n-1}), v \rangle.
\]

Then subtracting (3.8) from the above yields
\[(z_{n+1} - z_{n-1}, v) + (2\Delta t)a(z_{n\theta}, v) = (2\Delta t)(b(\xi_n) \cdot \nabla z_{n\theta} - b(U_n) \cdot \nabla U_{n\theta}, v) + (2\Delta t)(c(\xi_n) \cdot \xi_n - c(U_n)U_n + f(\xi_n) - f(U_n), v) + (\Delta t)(\eta_n, v) + (\Delta t)\left(\frac{(\xi_{n+1} - \xi_n) - (\xi_{n-1} - \xi_n)}{\Delta t}\right), v\).

Letting \(v = z_{n\theta}\), we may estimate the terms containing \(c, f\) and \(\eta_n\) as in the proof of Lemma 4.1 to obtain an upper bound for the absolute value of their sum of

\[(C\Delta t)(||e_n||^2_{L^2} + ||z_{n\theta}||^2_{L^2} + ||\eta_n||^2_{L^2}) \leq (C\Delta t)(||z_{n-1}||^2_{L^2} + ||z_n||^2_{L^2} + ||z_{n+1}||^2_{L^2} + ||e_n||^2 + ||\eta_n||^2_{L^2}).\]

Also

\[\leq (\Delta t)\left(\frac||\frac{(\xi_{n+1} - \xi_n) - (\xi_{n-1} - \xi_n)}{\Delta t}||^2_{L^2} + ||z_{n\theta}||^2_{L^2}\right)\]

Now

\[b(\xi_n) \cdot \nabla z_{n\theta} - b(U_n) \cdot \nabla U_{n\theta} = b(\xi_n) \cdot \nabla \xi_{n\theta} - b(\xi_n) \cdot \nabla \tilde{z}_{n\theta}
\]

\[+ b(\xi_n) \cdot \nabla \tilde{z}_{n\theta} - b(U_n) \cdot \nabla \tilde{z}_{n\theta} - b(U_n) \cdot \nabla \tilde{z}_{n\theta}
\]

\[= b(U_n) \cdot \nabla z_{n\theta} + (b(\xi_n) - b(U_n)) \cdot \nabla \tilde{z}_{n\theta} - b(\xi_n) \cdot \nabla (\tilde{z}_{n\theta} - \xi_{n\theta}).\]

Therefore

\[|b(\xi_n) \cdot \nabla z_{n\theta} - b(U_n) \cdot \nabla U_{n\theta}, v|| \leq |b(U_n) \cdot \nabla z_{n\theta}, v|| + |(b(\xi_n) - b(U_n)) \cdot \nabla \tilde{z}_{n\theta}, v|| + |b(\xi_n) \cdot \nabla (\tilde{z}_{n\theta} - \xi_{n\theta}), v||
\]

\[\leq \epsilon ||z_{n\theta}||^2_{H^1_0} + C(||z_{n\theta}||^2_{L^2} + ||e_n||^2_{L^2}) + |b(\xi_n) \cdot \nabla (\tilde{z}_{n\theta} - \xi_{n\theta}), v||.
\]

The last term may be integrated by parts and estimated as follows:
\[
(b(x_n) - \nabla \tilde{x}_n - \tilde{x}_n, \nabla) = -((\tilde{x}_n - x_n) \cdot \nabla b(x_n)) + \frac{1}{4} ((\tilde{x}_n - x_n), (\nabla \times b(x_n)) \times (\nabla \times b(x_n)))
\]

\[
\leq \| (\tilde{x}_n - x_n) \|_{L^2}^2 + \| (\nabla \times b(x_n)) \times (\nabla \times b(x_n)) \|_{L^2}^2
\]

\[
\leq C(h^{2r} + \| x_n \|_{L^2}^2 + \| x_n \|_{L^2}^2 + \| x_n \|_{L^2}^2)
\]

Finally \(2(\Delta t) z_n = (2(\Delta t) c_o \| x_n \|_{L^2}^2 \)

\[
(\tilde{x}_{n+1} - \tilde{x}_n) = \theta (\| x_n+1 \|_{L^2}^2 - \| x_n \|_{L^2}^2) + (1-2\theta) (\| x_n \|_{L^2}^2 - \| x_{n-1} \|_{L^2}^2)
\]

Thus

\[
\theta (\| x_n+1 \|_{L^2}^2 - \| x_n \|_{L^2}^2) + (1-2\theta) (\| x_n \|_{L^2}^2 - \| x_{n-1} \|_{L^2}^2) + (2(\Delta t) c_o \| x_n \|_{L^2}^2
\]

\[
\leq (2(\Delta t) c_o \| x_n \|_{L^2}^2 + C(\Delta t) (h^{2r} + \| x_n \|_{L^2}^2 + \| x_n \|_{L^2}^2 + \| x_{n-1} \|_{L^2}^2) + \| (\tilde{x}_{n+1} - \tilde{x}_n)\|_{L^2}^2
\]

Summing from \(n=1\) to \(N-1\) and the use of (4.11) gives

\[
\theta (\| x_N \|_{L^2}^2 + \| x_{N-1} \|_{L^2}^2) + (1-2\theta) (\| x_N \|_{L^2}^2 + \| x_{N-1} \|_{L^2}^2) \leq C(h^{2r} + \| x_0 \|_{L^2}^2 + \| x_1 \|_{L^2}^2)
\]

\[
+ (C\Delta t) \sum_{n=1}^{N-1} \| x_n \|_{L^2}^2 + (C\Delta t) \sum_{n=0}^{N} \| x_n \|_{L^2}^2
\]

For \(\theta \geq \frac{1}{4}\) we may absorb the \((\| x_N \|_{L^2}^2 + \| x_{N-1} \|_{L^2}^2)\) term in the previous term as in the proof of Lemma 4.1. Thus

\[
\| x_n \|_{L^2}^2 \leq C(h^{2r} + \| x_0 \|_{L^2}^2 + \| x_1 \|_{L^2}^2) + (C\Delta t) \sum_{n=1}^{N-1} \| x_n \|_{L^2}^2 + (C\Delta t) \sum_{n=0}^{N} \| x_n \|_{L^2}^2
\]
As in Lemma 1, we note that no step of the proof has actually required that \( N = T/\Delta t \) and therefore, for \( \Delta t \) sufficiently small, we can use the same reasoning to apply the discrete Gronwall’s inequality and obtain

\[
\|z_N\|_{L^2}^2 \leq C(\|z_0\|_{L^2}^2 + \|z_1\|_{L^2}^2 + h^{2r} + (\Delta t) \sum_{n=1}^{N-1} \|\eta_n\|_{L^2}^2) \leq C(\|e_0\|_{L^2}^2 + \|z_1\|_{L^2}^2 + h^{2r} + (\Delta t) \sum_{n=1}^{N-1} \|\eta_n\|_{L^2}^2).
\]

But \( \|e_N\|_{L^2}^2 \leq 2\|z_N\|_{L^2}^2 + O(h^{2r}) \) so (4.12) follows.

An estimate of \( \|z_1\|_{L^2}^2 \) is now needed.

**Lemma 4.4:** In addition to the hypotheses of Lemma 4.2 assume that the \( a_{ij} \) are functions of \( x \) alone and that \( \frac{\partial (b_j(\xi_n))}{\partial x_i} \) is bounded. Then if \( \nabla \xi_n \) and \( \xi_n \) are bounded for \( n=0,1 \) and \( \tilde{\xi}_0 \) and \( \tilde{\xi}_1 \) satisfy (4.9) and (4.10),

(4.13) \[
\|z_1\|_{L^2}^2 \leq C(h^{2r} + \|e_0\|_{L^2}^2 + \|\eta_0\|_{L^2}^2)
\]

where \( z_n = \tilde{\xi}_n - U_n \).

**Proof:** Substitution of (4.9) into (4.4) gives

\[
\langle \tilde{\xi}_1 - \tilde{\xi}_0, v \rangle + (\Delta t)a(\tilde{\xi}_0, v) - \langle\Delta t\rangle(\nabla \tilde{\xi}_0, v) - \langle\Delta t\rangle(c(\tilde{\xi}_0)\tilde{\xi}_0 + f(\tilde{\xi}_0), v) = \langle \eta_0, v \rangle + \langle \tilde{\xi}_1 - \tilde{\xi}_0 - (\tilde{\xi}_0 - \tilde{\xi}_0), v \rangle \forall v \in S^h.
\]

Subtracting (3.6) from this equation gives

\[
\langle z_1 - z_0, v \rangle + (\Delta t)a(z_0, v) = (\Delta t)(b(\xi_0) \cdot \nabla \tilde{\xi}_0 - b(U_0) \cdot \nabla \tilde{U}_0, v) + (\Delta t)(c(\xi_0)\tilde{\xi}_0 + f(\xi_0) - c(U_0)U_0 - f(U_0), v) + \langle \eta_0, v \rangle
\]

\[
+ \langle (\tilde{\xi}_1 - \tilde{\xi}_0) - (\tilde{\xi}_0 - \tilde{\xi}_0), v \rangle.
\]
Take $v = z_0^\theta$. Then the first two terms on the right hand side become

$$2\theta \|z_1\|_{L^2}^2 - (1 - 2\theta)\|z_0\|_{L^2}^2 + (1 - 4\theta)\langle z_1, z_0 \rangle + (\Delta t) a(z_0^\theta, z_0^\theta)$$

$$\geq (2\theta - \epsilon)\|z_1\|_{L^2}^2 - C\|z_0\|_{L^2}^2 + (\Delta t) c_0 \|z_0\|_{H^1_0}^2.$$  

Terms containing $c$, $f$ and $\eta$ may be bounded in the usual manner by

$$(C \Delta t) (\|x_1\|_{L^2}^2 + \|z_1\|_{L^2}^2) + C(\|x_0\|_{L^2}^2 + \|z_0\|_{L^2}^2 + \|\eta_0\|_{L^2}^2) + \epsilon \|z_1\|_{L^2}^2.$$  

Also

$$\|\tilde{x}_1 - \tilde{x}_0\|_{L^2}^2 \leq (\|x_1 - x_0\|_{L^2}^2 + \|\tilde{x}_0 - \tilde{x}_0\|_{L^2}^2) + \epsilon \|z_1\|_{L^2}^2.$$  

Finally, the term containing $b$ can be bounded by

$$(\epsilon \Delta t) \|z_0\|_{H^1_0}^2 + (C \Delta t) (h^{2r} + \|z_0\|_{L^2}^2 + \|z_1\|_{L^2}^2 + \|x_0\|_{L^2}^2)$$

using techniques similar to those used in the proof of Lemma 4.3. Combining these estimates gives

$$(2\theta - \epsilon)\|z_1\|_{L^2}^2 \leq C(\|z_0\|_{L^2}^2 + \|z_0\|_{L^2}^2 + \|\eta_0\|_{L^2}^2 + h^{2r}) + (C \Delta t)(\|x_1\|_{L^2}^2 + \|z_1\|_{L^2}^2).$$

Since

$$\|x_1\|_{L^2}^2 = 0(\|z_1\|_{L^2}^2 + h^{2r})$$

and

$$\|z_0\|_{L^2}^2 = 0(\|z_0\|_{L^2}^2 + h^{2r})$$

we have

$$(2\theta - \epsilon)\|z_1\|_{L^2}^2 \leq C(\|z_0\|_{L^2}^2 + \|\eta_0\|_{L^2}^2 + h^{2r}) + (C \Delta t)\|z_1\|_{L^2}^2.$$
Therefore if $\epsilon$ and $\Delta t$ are sufficiently small (4.13) follows.

A similar result may be obtained for Neumann boundary conditions. In order to specify a unique $\xi_n$ in the Neumann case, however, (4.9) must be replaced by $a(\tilde{\xi}_n - \xi_n, v) + (\tilde{\xi}_n - \xi_n, v) = 0 \forall v \in S_h$. Also, the integration by parts used to obtain an estimate of the terms containing $b$ is no longer valid since it was assumed that the boundary integral vanishes. Therefore in the Neumann case, $b$ must be assumed to be zero.

One remaining question is that of the existence of functions $\tilde{\xi}_n$ having the properties required by Lemmas 4.3 and 4.4. A function $\tilde{\xi}_n$ satisfying (4.9) may be shown to exist for each $n$ by an argument similar to that of Section 3 showing the existence of $U_n$. Wheeler [13] proves (4.10).

**Theorem (Wheeler):** Let $S^h \subset H^1_0(\Omega)$ be an $S_{k,m}$ space and assume that $a^*(\cdot, \cdot)$ is zero regular on $H^1_0(\Omega)$. If $w \in H^p(\Omega) \cap H^1_0(\Omega)$ for some positive integer $p$ and $\tilde{w}$ is such that $a(w - \tilde{w}, v) = 0$ for all $v \in S^h$, then

$$\|w - \tilde{w}\|_2 \leq Ch^r \|w\|_{H^r(\Omega)}$$

where $r = \min(p,m)$ and $C$ is independent of $w$ and $h$.

Thus if $\xi_n \in H^p(\Omega)$ and $\|\xi_n\|_{H^p(\Omega)}$ is bounded independently of $n$, $h$ and $\Delta t$, then (4.10) is seen to hold for $r = \min(p,m)$. Wheeler [13] shows that if $\Omega$ is a Cartesian product of intervals and $S^h$ is a space of tensor products of piecewise Hermite polynomials then $|\nabla \xi_n|$ and $|\xi_n|$ are bounded under certain mild hypotheses on the smoothness of $\xi_n$ and the mesh spacing of $S^h$. However, a discussion of (4.11) must await the definition of $\xi_n$. 
5. Linear Problems

Although the principles involved in defining the functions $\xi_n$ are identical for both linear and non-linear equations, the notational simplicity of the linear case earns it special treatment. Thus the theorems proved in this section will allow repeated extrapolation whereas those of Section 6 allow only one extrapolation. For notational convenience derivatives are written $D_{x_i} u = \frac{\partial u}{\partial x_i}$.

Define the following linear operators:

$$Pv = \sum_{i,j=1}^{g} D_{x_i}(a_{ij} D_{x_j} v) + b \cdot \nabla v$$

$$L v = P v + cv$$

$$P' v = (4\theta - 1) P v$$

$$L' v = P' v - cv.$$  

In the expression $Pf_{n\theta}$ it will be assumed that the functions $a_{ij}(x,t)$ are evaluated at $t = n\theta$.

Lemmas 5.1 and 5.2 partially define $\{\xi_n\}$ by giving partial differential equations and initial conditions which the coefficients of the powers of $(\Delta t)^2$ must satisfy.

**Lemma 5.1:** Suppose (3.1) is linear and, for even $m$ with $2 \leq m \leq M$ where $M$ is also even, let $y^{(m)}$ and $z^{(m)}$ be functions with domain $\Omega$. Suppose that $D_{x_i} D_{x_j} D_{t}^{M-3} u$, $D_{x_i} D_{x_j} D_{t}^{M-3-m_1} y^{(m)}$ and $D_{x_i} D_{x_j} D_{t}^{M-3-m_2} z^{(m)}$ are all continuous on $\Omega$ for all $1 \leq i, j \leq \ell$. Suppose also that $D_{x_i} a_{ij}$ is continuous on $\Omega$. Assume that

$$D_{t} y^{(m)} - L y^{(m)} = \varphi^{(m)},$$  

$$D_{t} z^{(m)} - L' z^{(m)} = \psi^{(m)},$$  

where
\( \varphi(m) = \frac{D_{t}^{m+1} u - 2\theta(m+1)PD_{T}^{m}u}{(m+1)!} - \sum_{i=2}^{m-2} \frac{D_{t}^{i+1} y^{(m-i)} - 2\theta(i+1)PD_{T}^{i}y^{(m-i)}}{(i+1)!} \)

and

\( \psi(m) = -\sum_{i=2}^{m-2} \frac{D_{t}^{i+1} z^{(m-i)} - 2\theta(i+1)PD_{T}^{i}z^{(m-i)}}{(i+1)!} \)

and let

\( \xi_{n} = u_{n} - \sum_{m=2, \text{m even}}^{M} (y_{n}^{(m)} + (-1)^{n}z_{n}^{(m)}) (\Delta t)^{m} . \)

If \( \eta_{n} \) is defined by

\( \xi_{n+1} - \xi_{n-1} - (2\Delta t)(P \xi_{n} + c\xi_{n} + f_{n}) = (\Delta t)\eta_{n} \)

then \( \eta_{n} = O((\Delta t)^{M+2}) \).

**Proof:** By Taylor’s theorem

\( u_{n+1} = \sum_{i=0}^{M+2} \frac{D_{t}^{i} u_{n}}{i!} (\Delta t)^{i} + O((\Delta t)^{M+3}) \)

and

\( u_{n-1} = \sum_{i=0}^{M+2} \frac{D_{t}^{i} u_{n}}{i!} (-\Delta t)^{i} + O((\Delta t)^{M+3}) \).

As a result
\[ u_{n+1} - u_{n-1} = 2 \sum_{i=1}^{M+1} \frac{D_{i}u_{n}}{i!} (\Delta t)^i + O((\Delta t)^{M+3}) \]

Similarly

\[ y_{n+1}^{(m)} - y_{n-1}^{(m)} = (2\Delta t) \sum_{i=0}^{M-m} \frac{D_{i}y_{n}^{(m)}}{(i+1)!} (\Delta t)^i + O((\Delta t)^{M+3-m}) \]

and

\[ z_{n+1}^{(m)} - z_{n-1}^{(m)} = (2\Delta t) \sum_{i=0}^{M-m} \frac{D_{i}z_{n}^{(m)}}{(i+1)!} (\Delta t)^i + O((\Delta t)^{M+3-m}) \]

Therefore

\[ (-1)^{n+1} z_{n+1}^{(m)} - (-1)^{n-1} z_{n-1}^{(m)} = (-1)^{n-1} (2\Delta t) \sum_{i=0}^{M-m} \frac{D_{i}z_{n}^{(m)}}{(i+1)!} (\Delta t)^i + O((\Delta t)^{M+3-m}) \]

Thus

\[ \xi_{n+1} - \xi_{n-1} = (2\Delta t) \sum_{i=0}^{M} \frac{D_{i}u_{n}}{(i+1)!} (\Delta t)^i \]

\[-(2\Delta t) \sum_{m=2}^{M} \sum_{i=0}^{M-m} \frac{D_{i}y_{n}^{(m)} + (-1)^{n-1} D_{i}z_{n}^{(m)}}{(i+1)!} (\Delta t)^{i+m} + O((\Delta t)^{M+3}) \]
\[(2\Delta t) \sum_{m=0}^{M} \frac{D^m u_n}{(m+1)!} (\Delta t)^m\]

\[-(2\Delta t) \sum_{m=2}^{M} \left[ \sum_{i=0}^{m-2} \frac{D^i u_n}{(i+1)!} \right] (\Delta t)^m + 0((\Delta t)^{M+3})\]

\[= (2\Delta t) D_t u_n + (2\Delta t) \sum_{m=2}^{M} \frac{D^m u_n}{(m+1)!} \]

\[-\sum_{i=0}^{m-2} \frac{D^i u_n}{(i+1)!} (\Delta t)^m + 0((\Delta t)^{M+3})\]

Now

\[u_{n\theta} = u_n + 2\theta \sum_{i=2}^{M} \frac{D_i u_n}{i!} (\Delta t)^i + 0((\Delta t)^{M+2}).\]

Similarly

\[y^{(m)}_{n\theta} = y^{(m)}_n + 2\theta \sum_{i=2}^{M-m} \frac{D_i y^{(m)}_n}{i!} (\Delta t)^i + 0((\Delta t)^{M+2-m})\]

and

\[(-1)^{n+1} \theta z^{(m)}_{n+1} + (-1)^n (1-2\theta) z^{(m)}_n + (-1)^{n-1} \theta z^{(m)}_{n-1}\]

\[= (-1)^{n-1} (\theta z^{(m)}_n + (2\theta-1) z^{(m)}_n + \theta z^{(m)}_{n-1})\]

\[= (-1)^{n-1} (4\theta-1) z^{(m)}_n + (-1)^{n-1} 2\theta \sum_{i=2}^{M-m} \frac{D_i z^{(m)}_n}{i!} (\Delta t)^i + 0((\Delta t)^{M+2-m}).\]
Thus we have

\[ \xi_n^\theta = u_n - \sum_{m=2}^{M} (v_n^{(m)} + (-1)^{n-1}(4\theta - 1)z_n^{(m)})(\Delta t)^m + 2\theta \sum_{m=2}^{M} \frac{D_u u_n}{m!} (\Delta t)^m \]

\[ - \sum_{m=2}^{M} \left[ \sum_{i=2}^{m} \frac{2\theta D_{y_n}^{(m)}}{i!} + (-1)^{n-1} \frac{2\theta D_{z_n}^{(m)}}{i!} (\Delta t)^{i+m} \right] + O((\Delta t)^{M+2}) \]

\[ = u_n - \sum_{m=2}^{M} [v_n^{(m)} - 2\theta \frac{D_{y_n}^{(m)}}{m!} u_n + \sum_{i=2}^{m-2} \frac{2\theta (i+1)D_{y_n}^{(m-i)}}{(i+1)!}] (\Delta t)^m + O((\Delta t)^{M+2}). \]

Therefore

\[ -(2\Delta t)P_{x_n^\theta} = -(2\Delta t)P u_n + (2\Delta t) \sum_{m=2}^{M} [P v_n^{(m)} - 2\theta \frac{PD_{y_n}^{(m)}}{m!} u_n \]

\[ + \sum_{i=2}^{m-2} \frac{2\theta (i+1)PD_{y_n}^{(m-i)}}{(i+1)!}] (\Delta t)^m + O((\Delta t)^{M+3}). \]

Also

\[ -(2\Delta t)c_n^\theta = -(2\Delta t)c u_n + (2\Delta t) \sum_{m=2}^{M} [c v_n^{(m)} (\Delta t)^m - (-1)^{n-1}(2\Delta t) \sum_{m=2}^{M} c z_n^{(m)} (\Delta t)^m]. \]
Thus

$$\Delta t \eta_n = \xi_{n+1} - \xi_{n-1} - (2\Delta t) (P \xi_n \eta + c \xi_n + f_n) = (2\Delta t) (D_t u_n - Pu_n - cu_n - f_n)$$

$$+ (2\Delta t) \sum_{m=2}^{M} \frac{D_t^{m+1} u_n}{(m+1)!} - D_t \gamma_n \sum_{i=2}^{m-2} \frac{D_t^{i+1} \gamma (m-i)}{(i+1)!} + \gamma_n$$

$$- 2\eta \frac{PD_t^{m} u_n}{m!} + \sum_{i=2}^{m-2} \frac{2\eta (i+1) PD_t^{i} \gamma (m-i)}{(i+1)!} + c \gamma_n$$

$$+ (-1)^{n-1} (-D_t \gamma_n) \sum_{i=2}^{m-2} \frac{D_t^{i+1} \gamma (m-i)}{(i+1)!} + \gamma_n$$

$$+ \sum_{i=2}^{m-2} \frac{2\eta (i+1) PD_t^{i} \gamma (m-i)}{(i+1)!} - c \gamma_n (m) \Delta t^m + O(\Delta t^{M+3})$$

and from (3.1), (5.1), (5.2), (5.3) and (5.4) we see that the $O(\Delta t^{M+3})$ term is the only non-zero term on the right hand side, giving the desired result.

**Lemma 5.2:** In addition to the assumptions of Lemma 5.1 assume that

(5.6) \[ y_o^{(m)} = -z_o^{(m)} = \frac{1}{2m!} D_t^{m} u_0 - \frac{\theta PD_t^{m-1} u_0}{(m-1)!} \sum_{i=2}^{m-2} \frac{D_t^{i+1} \gamma (m-i) - D_t^{i} \gamma (m-i)}{2i!} \]

$$+ \sum_{i=2}^{m-2} \frac{\theta PD_t^{i+1} \gamma (m-i) - \theta PD_t^{i} \gamma (m-i)}{(i-1)!}.$$}

If $\eta_0$ is defined by
(5.7) \[ \xi_1 - \xi_0 - (\Delta t)(P' \xi_0 + c\xi_0 + f_0) = \eta_0 \]

then \( \eta_0 = O((\Delta t)^{M+2}) \).

**Proof:** From (5.6) we see that \( \xi_0 = u_0 \).

By the definition of \( \xi_1 \) one has that

\[ \xi_1 = u_1 - \sum_{m=2}^{M} \frac{(y_1^{(m)} - y_1^{(m)}(\Delta t)^m}{m \text{ even}} \]

But

\[ u_1 = \sum_{i=0}^{\frac{M+1}{2}} \frac{D^{l+1}_u}{i!} (\Delta t)^i + O((\Delta t)^{M+2}) \]

\[ y_1^{(m)} = \sum_{i=0}^{\frac{M+1}{2}} \frac{D^{l+1}_u}{i!} (\Delta t)^i + O((\Delta t)^{M+2}) \]

and

\[ z_1^{(m)} = \sum_{i=0}^{\frac{M+1}{2}} \frac{D^{l+1}_u}{i!} (\Delta t)^i + O((\Delta t)^{M+2}) \]

Therefore

\[ \xi_1 - \xi_0 = u_1 - \sum_{m=1}^{M+1} \frac{D^{m+1}_u}{m!} (\Delta t)^m = \sum_{m=2}^{M} \frac{M-m+1}{m \text{ even}} \sum_{i=0}^{\frac{M+1}{2}} \frac{D^{l+1}_u}{i!} (\Delta t)^i - \sum_{m=2}^{M} \frac{M-m+1}{m \text{ even}} \sum_{i=0}^{\frac{M+1}{2}} \frac{D^{l+1}_u}{i!} (\Delta t)^i 
\]

\[ = (\Delta t) \sum_{m=0}^{M} \frac{D^{m+1}_u}{(m+1)!} (\Delta t)^m - \sum_{m=2}^{M} \frac{M-m+1}{m \text{ even}} \sum_{i=0}^{\frac{M+1}{2}} \frac{D^{l+1}_u}{i!} (\Delta t)^i - \sum_{m=2}^{M} \frac{M-m+1}{m \text{ even}} \sum_{i=0}^{\frac{M+1}{2}} \frac{D^{l+1}_u}{i!} (\Delta t)^i \]
\[ + \sum_{m=2}^{M} \frac{D_t^{m} u_0}{m!} (\Delta t)^m - \sum_{m=2}^{M} \sum_{i=0}^{m} \frac{D_i Y_{i0}(m)}{i!} (\Delta t)^i \]

\[ - \sum_{i=0}^{M} \frac{D_i z_{i0}(m)}{i!} (\Delta t)^i ] (\Delta t)^m + 0((\Delta t)^{M+2}) \]

\[ = (\Delta t) \sum_{m=2}^{M} \frac{D_t^{m+1} u_0}{(m+1)!} - \sum_{i=0}^{m-2} \left( \frac{D_t^{i+1} Y_{0}(m-i)}{(i+1)!} - \frac{D_t^{i+1} z_{i0}(m-i)}{(i+1)!} \right) (\Delta t)^m + (\Delta t) D_t u_0 \]

\[ + \sum_{m=2}^{M} \frac{D_t^{m} u_0}{m!} - \sum_{i=0}^{m-2} \left( \frac{D_t^{i} Y_{0}(m-i)}{i!} - \frac{D_t^{i} z_{i0}(m-i)}{i!} \right) (\Delta t)^m + 0((\Delta t)^{M+2}). \]

Also

\[ u_{0\theta} = u_0 + 2\theta \sum_{i=1}^{M} \frac{D_i Y_{i0}}{i!} (\Delta t)^i + 0((\Delta t)^{M+1}), \]

\[ y_{0\theta}^{(m)} = y_{0}^{(m)} + 2\theta \sum_{i=1}^{m} \frac{D_i Y_{i0}(m)}{i!} (\Delta t)^i + 0((\Delta t)^{M-m+1}) \]

and

\[ (1-2\theta)z_{0}^{(m)} - 2\theta z_{1}^{(m)} = -(4\theta-1)z_{0}^{(m)} - 2\theta \sum_{i=1}^{M} \frac{D_i z_{i0}(m)}{i!} (\Delta t)^i + 0((\Delta t)^{M+1-m}). \]

Thus

\[ s_{0\theta} = u_0 + 2\theta \sum_{m=1}^{M} \frac{D_t^{m} u_0}{m!} (\Delta t)^m - \sum_{m=2}^{M} \sum_{i=0}^{m} \frac{D_i Y_{i0}(m)}{i!} (\Delta t)^i \]

\[ - (4\theta-1)z_{0}^{(m)} - 2\theta \sum_{i=1}^{M} \frac{D_i z_{i0}(m)}{i!} (\Delta t)^i ] (\Delta t)^m + 0((\Delta t)^{M+1}) \]
\[
= u_o + \sum_{m=2}^{M} \left[ 2\theta \frac{D_{t}^{m}u_o}{m!} - y_o^{(m)} \right] - 2\theta \sum_{i=2}^{m} \frac{D_{t}^{i}y_o}{i!} (\Delta t)^i + (4\theta - 1)z_o^{(m)} \\
+ 2\theta \sum_{i=2}^{m} \frac{D_{t}^{i}z_o}{i!} (\Delta t)^i - (\Delta t)^i + (4\theta - 1)z_o^{(m)} \\
- \sum_{m=2}^{M} \left[ 2\theta \sum_{i=2}^{m} \frac{D_{t}^{i-1}y_o}{(i-1)!} (\Delta t)^i - 2\theta \sum_{i=2}^{m} \frac{D_{t}^{i-1}z_o}{(i-1)!} (\Delta t)^i + 0((\Delta t)^{M+1}) \right] \\
= u_o + \sum_{m=2}^{M} \left[ 2\theta \frac{D_{t}^{m}u_o}{m!} - 2\theta \sum_{i=2}^{m} \frac{D_{t}^{i}y_o}{i!} + 2\theta \sum_{i=2}^{m} \frac{D_{t}^{i}z_o}{i!} - y_o^{(m)} + (4\theta - 1)z_o^{(m)} \right] (\Delta t)^m \\
+ \sum_{m=2}^{M} \left[ 2\theta \frac{D_{t}^{m-1}u_o}{(m-1)!} - 2\theta \sum_{i=2}^{m} \frac{D_{t}^{i-1}y_o}{(i-1)!} \right] (\Delta t)^{m-1} + 0((\Delta t)^{M+1}) \\
\]

Therefore

\[-(\Delta t)P_{t\theta} = -(\Delta t)P_{u_o}\]

\[-(\Delta t) \sum_{m=2}^{M} \left[ 2\theta \frac{P^{D_{t}^{m}u_o}}{m!} - 2\theta \sum_{i=2}^{m} \frac{P^{D_{t}^{i}y_o}}{i!} \right] \\
+ 2\theta \sum_{i=2}^{m-2} \frac{P^{D_{t}^{i}z_o}}{i!} - P_{y_o}^{(m)} + P_{z_o}^{(m)} (\Delta t)^m \\
- \sum_{m=2}^{M} \left[ 2\theta \frac{P^{D_{t}^{m-1}u_o}}{(m-1)!} - 2\theta \sum_{i=2}^{m-2} \frac{P^{D_{t}^{i-1}y_o}}{(i-1)!} \right] (\Delta t)^m + 0((\Delta t)^{M+2}). \]
Also
\[-(\Delta t)c'\xi_0 = -(\Delta t)cu_0 + (\Delta t)\sum_{m=2}^{M} \frac{c_{y_0}^{(m)} + cz_0^{(m)}}{m \text{ even}}.\]

Therefore
\[\eta_o = \xi_0 - (\Delta t)P_{\xi_0}\theta - (\Delta t)c'\xi_0 - (\Delta t)f_0 = (\Delta t)(Du_0 - Pu_0 - cu_0 - f_0) + (\Delta t)\sum_{m=2}^{M} \frac{PD_{m+1}u_0}{(m+1)!} - \theta \sum_{i=0}^{m-2} \frac{PD_{i+1}y_0}{(i+1)!} - \theta \sum_{i=0}^{m-2} \frac{PD_{i+1}z_0}{(i+1)!} - 2\theta \frac{PD_{m}u_0}{m!}\]
\[+ 2\theta \sum_{i=2}^{m-2} \frac{PD_{i}y_0}{i!} - 2\theta \sum_{i=2}^{m-2} \frac{PD_{i}z_0}{i!}\]
\[+ \sum_{m=2}^{M} \frac{D_{m}u_0}{m!} - \sum_{i=0}^{m-2} \frac{D_{i}y_0}{i!} - \frac{D_{i}z_0}{i!} - 2\theta \frac{PD_{m-1}u_0}{(m-1)!}\]
\[+ \sum_{i=2}^{m-2} (2\theta \frac{PD_{i-1}y_0}{(i-1)!} - 2\theta \frac{PD_{i-1}z_0}{(i-1)!}) (\Delta t)^{m} + O((\Delta t)^{M+2})\]

\[= (\Delta t)\sum_{m=2}^{M} \frac{D_{m+1}u_0}{(m+1)!} - D_{m}y_0 - \sum_{i=2}^{m-2} \frac{D_{i+1}y_0}{(i+1)!} - 2\theta \frac{PD_{m}u_0}{m!}\]
\[+ 2\theta \sum_{i=2}^{m-2} \frac{PD_{i}y_0}{i!} + P_{y_0}(m) + cy_0^{(m)} (\Delta t)^{m}\]
\[+ (\Delta t)\sum_{m=2}^{M} \frac{D_{m}z_0}{m!} + \sum_{i=2}^{m-2} \frac{D_{i+1}z_0}{(i+1)!} - 2\theta \sum_{i=2}^{m-2} \frac{PD_{i}z_0}{i!} - P_{z_0}^{(m)} + cz_0^{(m)} (\Delta t)^{m}\]
the last equality following from (3.1). But from (5.1), (5.2), (5.3), (5.4) and (5.6) the right hand side above is seen to be $O((\Delta t)^{M+2})$, thus proving the desired result.

To complete a specification of $y^{(m)}$ and $z^{(m)}$ it is necessary to give boundary conditions. Given suitable boundary data, one can define $y^{(m)}$ and $z^{(m)}$ recursively for even $m$. That is to say the problem having $y^{(2)}$ and $z^{(2)}$ as a solution is given in terms of $u$ alone, then the problem having $y^{(4)}$ and $z^{(4)}$ as a solution is given in terms of $u$, $y^{(2)}$ and $z^{(2)}$ and so on. In the case that $u$ satisfies Dirichlet boundary conditions, $y^{(m)}$ and $z^{(m)}$ will also be assumed to vanish on $\partial \Omega$ whereas when $u$ satisfies Neumann boundary conditions $y^{(m)}$ and $z^{(m)}$ will be required to satisfy homogeneous Neumann boundary conditions. Theorems 5.1, 5.2 and 5.3 show that these boundary conditions enable us to obtain a bound on $\|e_N\|_{L^2}$.

**Theorem 5.1:** In addition to the conditions of Lemmas 5.1 and 5.2 assume that 
\[
\{u_i, y_i^{(m)}, z_i^{(m)}, 0 \leq i \leq N, 2 \leq m \leq M, m \text{ even}\}
\]
are all elements of $H_0^1(\Omega)$ and that
\[
\{U_0, \ldots, U_N\} \subset S^h \subset H_0^1(\Omega)
\]
is a sequence of functions satisfying (3.6) and (3.8). Then if $e_n$ and $\bar{e}_n$ are defined as in Lemma 4.1 and $\theta > \frac{1}{4}$, one has

\[
(5.8) \quad \|e_N\|^2_{L^2} + (\beta \Delta t) \sum_{n=1}^{N-1} \|e_{n+1} - e_n\|^2_{H_0^1} \leq C((\Delta t)^{M+2}) \sum_{n=2}^{N-2} \|\bar{e}_{n+1} - \bar{e}_n\|^2_{H_0^1} + \|e_n\|_{L^2} + \|e_0\|_{L^2} + O((\Delta t)^{2M+4}).
\]
Proof: Let \( v \in H^1_0(\Omega) \). Then from equation (5.5) one obtains

\[
(5.9) \quad \langle \xi_{n+1} - \xi_{n-1}, v \rangle - (2\Delta t) \langle P_\xi \xi_n, v \rangle - (2\Delta t) \langle P_\xi \xi_n \theta, v \rangle - (2\Delta t) \langle c_0 \xi_n + f_n, v \rangle = (\Delta t) \langle \eta_n, v \rangle
\]

and \( \|\eta_n\| = O((\Delta t)^{M+2}) \). But

\[
\langle P_\xi \xi_n \theta, v \rangle = \int_{\Omega} \sum_{i,j=1}^q \frac{\partial}{\partial x_i} (a_{ij} \frac{\partial \xi_n \theta}{\partial x_j}) v \, dx + \langle b_{n_i} \nabla \xi_n \theta, v \rangle
\]

and the first term may be integrated by parts to obtain

\[
\int_{\partial \Omega} \sum_{i,j=1}^q a_{ij} \frac{\partial \xi_n \theta}{\partial x_j} v \, d\sigma - \int_{\Omega} \sum_{i,j=1}^q a_{ij} \frac{\partial \xi_n \theta}{\partial x_j} \frac{\partial v}{\partial x_i} \, dx.
\]

Since \( v \in H^1_0(\Omega) \) the boundary integral is zero and the integral over \( \Omega \) is \( -a(\xi_n \theta, v) \).

Substitution into (5.9) gives equation (4.1) and thus for \( \Delta t \) sufficiently small Lemma 4.1 gives

\[
\|e_n\|_{L^2}^2 + (\beta \Delta t) \sum_{n=1}^{N-1} \|e_n \theta\|_{H^1_0}^2 \leq C(\Delta t) \sum_{n=2}^{N-2} \|\xi_n \theta - \bar{\xi}_{n-1} \theta\|_{L^2}^2
\]

\[
+ (\Delta t) \sum_{n=1}^{N-1} (\|\xi_n \theta\|_{H^1}^2 + \|e_n \|_{L^2}^2) + \max_{1 \leq n \leq N-1} (\|\xi_n \theta\|_{L^2}^2 + \|e_n \theta\|_{L^2}^2 + \|\eta_n\|_{L^2}^2).
\]

Similarly from equation (5.7) one obtains

\[
(5.10) \quad \langle \xi_0 - \xi_0, v \rangle - (\Delta t) \langle P_\xi \xi_0, v \rangle - (\Delta t) \langle c_0 \xi_0 + f_0, v \rangle = \langle \eta_0, v \rangle
\]

and \( \|\eta_0\|_{L^2} = O((\Delta t)^{M+2}) \). As before \( \langle P_\xi \xi_0, v \rangle = -a(\xi_0 \theta, v) + \langle b_{n_i} \nabla \xi_0 \theta, v \rangle \). Substitution into (5.10) gives equation (4.4) from which Lemma 4.2 implies

\[
\|e_1\|_{L^2}^2 \leq C[\|e_0\|_{L^2}^2 + \|e_0 \theta\|_{H^1}^2 + \|\eta_0\|_{L^2}^2], \text{ this latter being inequality (4.5) with } \beta = 0.
\]
Combination of the two inequalities gives
\[
\|\varepsilon_N\|_{L^2}^2 + (\beta \Delta t) \sum_{n=1}^{N-1} \|\varepsilon_n\|_{H^1}^2 \leq C[(\Delta t) \sum_{n=2}^{N-2} \|\varepsilon_{n+1} - \varepsilon_n\|_{L^2}^2 + (\Delta t) \sum_{n=1}^{N-1} \|\varepsilon_n\|_{H^1}^2 ] + 
\]
\[
+ \max_{1 \leq n \leq N-1} \max_{1 \leq m \leq 0} \|\varepsilon_n\|_{L^2}^2 \|\varepsilon_m\|_{H^1}^2 \leq (N-1)(\Delta t)C(\Delta t)^{2M+4} + C(\Delta t)^{2M+4} = 0(\Delta t)^{2M+4},
\]

and therefore (5.8) is established.

The Neumann problem is handled in a similar fashion except that the boundary integral does not vanish when integration by parts is performed. In order to deal properly with the boundary integral, the coefficients \(a_{ij}\) must be independent of \(t\).

**Theorem 5.2**: In addition to the conditions of Lemmas 5.1 and 5.2 assume the \(a_{ij}\) to be functions of \(x\) alone and that on \(\partial \Omega\)

\[
(5.11) \quad \sum_{i,j=1}^{g} a_{ij} \frac{\partial y(m)}{\partial x_j} \nu_i = \sum_{i,j=1}^{g} a_{ij} \frac{\partial z(m)}{\partial x_j} \nu_i = 0
\]

where \(\nu\) is the outward directed normal to \(\partial \Omega\) while \(u\) satisfies (3.3) on \(\partial \Omega\). Let \(\{U_1, \ldots, U_N\} \subset S^h \subset H^1(\Omega)\) be a sequence of functions satisfying (3.7) and (3.9). Then if \(e_n\) and \(\varepsilon_n\) are defined as in Lemma 4.1 and \(\theta > \frac{1}{4}\), (5.8) is still valid.

**Proof**: If \(v \in H^1(\Omega)\), equation (5.9) remains valid. Upon integrating \(\langle P_n, y, v \rangle\) by parts one obtains
\[
\int_{\partial \Omega} \sum_{i,j=1}^{g} \frac{\partial \xi_{n\theta}}{\partial x_j} \nu_i v d\sigma = a(\xi_{n\theta}, v) + \langle b_n \cdot \nabla \xi_{n\theta}, v \rangle.
\]
Examination of the boundary integral shows that

\[
\int_{\partial\Omega} \sum_{i,j=1}^{q} a_{ij}(x) \frac{\partial n_i}{\partial x_j} \nu_i(x)v(x)\,d\sigma = \int_{\partial\Omega} \sum_{i,j=1}^{q} a_{ij}(x) \frac{\partial u_n}{\partial x_j} \nu_i(x)v(x)\,d\sigma
\]

\[-\sum_{m=2}^{M} \theta \int_{\partial\Omega} \sum_{i,j=1}^{q} a_{ij}(x) \frac{\partial y^{(m)}_{n+1}}{\partial x_j} \nu_i(x)v(x)\,d\sigma
\]

\[+ (1-2\theta) \int_{\partial\Omega} \sum_{i,j=1}^{q} a_{ij}(x) \frac{\partial y^{(m)}_{n}}{\partial x_j} \nu_i(x)v(x)\,d\sigma
\]

\[+ \theta \int_{\partial\Omega} \sum_{i,j=1}^{q} a_{ij}(x) \frac{\partial y^{(m)}_{n-1}}{\partial x_j} \nu_i(x)v(x)\,d\sigma
\]

\[+ (-1)^{n+1} \theta \int_{\partial\Omega} \sum_{i,j=1}^{q} a_{ij}(x) \frac{\partial z^{(m)}_{n+1}}{\partial x_j} \nu_i(x)v(x)\,d\sigma
\]

\[+ (-1)^{n} (1-2\theta) \int_{\partial\Omega} \sum_{i,j=1}^{q} a_{ij}(x) \frac{\partial z^{(m)}_{n}}{\partial x_j} \nu_i(x)v(x)\,d\sigma
\]

\[+ (-1)^{n-1} \theta \int_{\partial\Omega} \sum_{i,j=1}^{q} a_{ij}(x) \frac{\partial z^{(m)}_{n-1}}{\partial x_j} \nu_i(x)v(x)\,d\sigma\] \((\Delta t)^m\).

Since \(a_{ij}\) is independent of \(t\), it follows from (5.11) that

\[
\int_{\partial\Omega} \sum_{i,j=1}^{q} a_{ij}(x) \frac{\partial y^{(m)}_{n+1}}{\partial x_j} \nu_i(x)v(x)\,d\sigma = 0
\]

and the other terms containing \(y^{(m)}\) and \(z^{(m)}\) vanish similarly. This leaves
\begin{align*}
\int_{\partial \Omega} \sum_{i,j=1}^{\varrho} a_{ij}(x) \frac{\partial \xi_{n\theta}}{\partial x_j} \nu_i(x)v(x) d\sigma &= \theta \int_{\partial \Omega} \sum_{i,j=1}^{\varrho} a_{ij}(x) \frac{\partial u_{n+1}}{\partial x_j} \nu_i(x)v(x) d\sigma \\
+ (1-2\theta) \int_{\partial \Omega} \sum_{i,j=1}^{\varrho} a_{ij}(x) \frac{\partial u_n}{\partial x_j} \nu_i(x)v(x) d\sigma + \theta \int_{\partial \Omega} \sum_{i,j=1}^{\varrho} a_{ij}(x) \frac{\partial u_{n-1}}{\partial x_j} \nu_i(x)v(x) d\sigma .
\end{align*}

Since the \(a_{ij}\) are independent of \(t\) it follows from (3.3) that

\[
\sum_{i,j=1}^{\varrho} a_{ij}(x) \frac{\partial u_k}{\partial x_j} \nu_i = -g_k, \quad k=n-1, n, n+1.
\]

Thus the boundary integral containing \(u\) reduces to \((g_{n\theta}, v)\) and

\[
\langle P_{\xi_{n\theta}}, v \rangle = -(g_{n\theta}, v) - a(\xi_{n\theta}, v) + \langle b_n \cdot \nabla \xi_{n\theta}, v \rangle.
\]

Substitution for \(\langle P_{\xi_{n\theta}}, v \rangle\) in (5.9) thus gives equation (4.7) from which for suitably small \(\Delta t\) one obtains, by means of Corollary 4.1, inequality (4.2). Equation (5.10) is also still valid for any \(v \in H^1(\Omega)\) and it is easily seen that

\[
\langle P_{\xi_{n\theta}}, v \rangle = -(g_{n\theta}, v) - a(\xi_{n\theta}, v) + \langle b_o \cdot \nabla \xi_{n\theta}, v \rangle
\]

from which equation (4.8) immediately follows. Thus Corollary 4.2 immediately gives

\[
\|\xi_{n\theta}\|_{L^2}^2 \leq C \left[ \|\xi_{o\theta}\|_{L^2}^2 + \|\xi_{o\theta}\|_{H^1}^2 + \|\xi_{n\theta}\|_{L^2}^2 \right] \text{ and (5.8) follows as in the proof of Theorem 5.1.}
\]

A similar result gives \(L^2\) error estimates for the Dirichlet problem.

**Theorem 5.3:** Under the hypotheses of Lemmas 4.3, 4.4, 5.1 and 5.2

\[
\|\xi_{n\theta}\|_{L^2}^2 \leq C (h^{2r} + \|\xi_{o\theta}\|_{L^2}^2 + (\Delta t)^{2M+4}).
\]

**Proof:** The proof is similar to that of Theorem 5.1. Equation (4.1) is seen to hold, but application of Lemma 4.3 yields
\[ \|e_N\|_2^2 \leq C(h^{2r} + 2\|e_o\|_2^2 + 2|z_1|_2^2 + (\Delta t) \sum_{n=1}^{N-1} |\eta_n|_2^2). \]

Again as in the proof of Theorem 5.1, first (5.10) and then (4.4) can be shown to hold from which Lemma 4.4 gives that

\[ \|z_1\|_2^2 \leq C(h^{2r} + 2\|e_o\|_2^2 + 2|\eta|_2^2). \]

Therefore

\[ \|e_N\|_2^2 \leq C(h^{2r} + 2\|e_o\|_2^2 + 2|\eta|_2^2 + (\Delta t) \sum_{n=1}^{N-1} |\eta_n|_2^2). \]

But the contribution due to the \( \eta_n \)'s is seen to be \( O((\Delta t)^{2M+4}) \) and (5.12) follows.

In Section 4 all of the hypotheses of Lemma 4.3 except (4.11) were shown to be reasonable. Now that the \( \xi_n \) have been defined, (4.11) may be treated. First notice that for \( n \) of a fixed parity we may consider \( \xi_n(x) = \xi(x, t_n) \) where the definition of \( \xi \) depends upon the parity of \( n \). That is, we may take:

\[ \xi = u - \sum_{m=2}^{M} (y^{(m)+z^{(m)}})(\Delta t)^m \text{ if } n \text{ is even} \]

and

\[ \xi = u - \sum_{m=2}^{M} (y^{(m)-z^{(m)}})(\Delta t)^m \text{ if } n \text{ is odd}. \]

In either case under the hypotheses of Lemma 5.1, \( \xi \) will be differentiable with respect to \( t \).

Since \( a(\cdot, \cdot) \) is independent of \( t \) it is readily shown that if \( \tilde{\xi} \) is defined so that

\[ a(\tilde{\xi} - \xi, v) = 0 \quad \forall \ v \in S^h \]

then \( \tilde{\xi} \) also has a derivative with respect to \( t \) and

\[ a\left(\frac{\partial(\tilde{\xi} - \xi)}{\partial t}, v\right) = 0 \quad \forall \ v \in S^h. \]

If \( \frac{\partial \xi}{\partial t} \) is an element of \( H^p(\Omega) \cap H^1_0(\Omega) \) for all \( t \in [0, T] \), if \( S^h \) is an \( S^{h,0}_{k,m} (\Omega) \) space and if \( \|\frac{\partial \xi}{\partial t}\|_{H^p} \) is bounded independently of \( t \) then it follows from
Wheeler [13] that \( \| \frac{\partial (\tilde{\xi} - \xi)}{\partial t} \|_{L^2} \leq C h^r \) where \( r = \min(p, m) \). Now if \( \xi \) is chosen so that \( \xi_{n+1}(x) = \xi(x, t_{n+1}) \) then

\[
(\tilde{\xi}_{n+1} - \xi_{n+1}) - (\tilde{\xi}_{n-1} - \xi_{n-1}) = \int_{t_{n-1}}^{t_{n+1}} \frac{\partial (\tilde{\xi} - \xi)}{\partial t} \, dt.
\]

Therefore

\[
\left\| (\tilde{\xi}_{n+1} - \xi_{n+1}) - (\tilde{\xi}_{n-1} - \xi_{n-1}) \right\|_{L^2}^2 \leq \left\| \int_{t_{n-1}}^{t_{n+1}} \frac{\partial (\tilde{\xi} - \xi)}{\partial t} \, dt \right\|_{L^2}^2 (\Delta t)^{-2}
\]

\[
\leq \left( \int_{t_{n-1}}^{t_{n+1}} \left\| \frac{\partial (\tilde{\xi} - \xi)}{\partial t} \right\|_{L^2}^2 \, dt \right) (\Delta t)^{-2} \leq (2(\Delta t) C h^r (\Delta t)^{-2} = 0(h^2 r)
\]

and (4.11) follows immediately.

As indicated in an earlier section, for extrapolation to be carried out, it is necessary for the coefficients of the powers of \( (\Delta t)^2 \) in the expansion used to define the \( \xi_n \) to be independent of \( n \). The \( \xi_n \) defined in this section do not have these coefficients independent of \( n \) but have coefficients which depend upon the parity of \( n \). However, if attention is restricted to either odd or even \( n \), then the coefficients of the powers of \( (\Delta t)^2 \) are independent of \( n \) and extrapolation is possible.

The results obtained in this section indicate that one may define \( y^{(m)} \) and \( z^{(m)} \) for arbitrarily large \( m \) even if (3.1) is non-linear provided that all functions involved are sufficiently smooth. However, the notation required for such a result is extremely cumbersome and typical choices of \( S^h \) are such that it is usually unnecessary to use algorithms of higher accuracy in time than \( O((\Delta t)^4) \). Thus the following results only define \( y^{(2)} \) and \( z^{(2)} \). For convenience we drop the superscript.

Lemma 6.1: Suppose \( \frac{\partial^4 a_{ij}}{\partial x_i \partial u^3} \), \( \frac{\partial^2 b_i}{\partial u^2} \), \( \frac{\partial^2 c}{\partial u^2} \) and \( \frac{\partial^2 f}{\partial u^2} \) are all continuous on \( \Omega \times \mathbb{R} \) and that \( \frac{\partial^7 u}{\partial x_i \partial x_j \partial t^6} \), \( \frac{\partial^5 y}{\partial x_i \partial y_j \partial t^3} \) and \( \frac{\partial^5 z}{\partial x_i \partial x_j \partial t^3} \) are all continuous on \( \Omega \). Let \( u \) satisfy (3.1) and suppose

\[
\frac{\partial y}{\partial t} - \sum_{i,j=1}^{q} \frac{\partial}{\partial x_i} \left( a_{ij}(u) \frac{\partial y}{\partial x_j} \right) - \sum_{i=1}^{q} \frac{\partial}{\partial x_i} \left( \frac{\partial a_{ij}(u)}{\partial u} \cdot \frac{\partial u}{\partial x_i} \cdot y \right) \\
- \sum_{i=1}^{q} b_i(u) \frac{\partial y}{\partial x_i} - \sum_{i=1}^{q} \frac{\partial b_i(u)}{\partial u} \cdot \frac{\partial u}{\partial x_i} \cdot y - \frac{\partial (c(u) \cdot u + f(u))}{\partial u} = \frac{1}{6} \frac{\partial^3 u}{\partial t^3} - \theta \sum_{i,j=1}^{q} \frac{\partial}{\partial x_i} \left( a_{ij}(u) \frac{\partial^3 u}{\partial x_j \partial t^2} \right) - \theta \sum_{i=1}^{q} b_i(u) \frac{\partial^3 u}{\partial x_i \partial t^2}
\]

and

\[
\frac{\partial z}{\partial t} - (4\theta - 1) \sum_{i,j=1}^{q} \frac{\partial}{\partial x_i} \left( a_{ij}(u) \frac{\partial z}{\partial x_j} \right) + \sum_{i,j=1}^{q} \frac{\partial}{\partial x_i} \left( \frac{\partial a_{ij}(u)}{\partial u} \cdot \frac{\partial u}{\partial x_j} \cdot z \right) \\
- (4\theta - 1) \sum_{i=1}^{q} b_i(u) \frac{\partial z}{\partial x_i} + \sum_{i=1}^{q} \frac{\partial b_i(u)}{\partial u} \cdot \frac{\partial u}{\partial x_i} \cdot z + \frac{\partial (c(u) \cdot u + f(u))}{\partial u} z = 0.
\]

Let \( s^n = u_n - (\Delta t)^2 y_n - (-1)^n (\Delta t)^2 z_n \). If \( \eta_n \) is defined by
\[ (6.3) \quad \xi_{n+1} - \xi_{n-1} - (2\Delta t) \sum_{i,j=1}^{g} \frac{\partial}{\partial x_i} (a_{ij}(\xi_n)) \frac{\partial \xi_n}{\partial x_j} - (2\Delta t) \sum_{i=1}^{g} b_i(\xi_n) \frac{\partial \xi_n}{\partial x_i} \]

\[-(2\Delta t)c(\xi_n) \xi_n - (2\Delta t)f(\xi_n) = (\Delta t)\eta_n \]

then \( \eta_n = O((\Delta t)^4) \).

Proof: From Taylor's theorem we see that

\[ \xi_{n+1} - \xi_{n-1} = (2\Delta t) \frac{\partial u_n}{\partial t} + \frac{(\Delta t)^3}{3} \frac{\partial^3 u_n}{\partial t^3} - 2(\Delta t)^3 \frac{\partial y_n}{\partial t} + 2(-1)^n(\Delta t)^3 \frac{\partial z_n}{\partial t} + O((\Delta t)^5) \]

and

\[ \xi_{n\theta} = u_n + \theta (\Delta t)^2 \frac{\partial^2 u_n}{\partial t^2} - (\Delta t)^2 y_n + (4\theta - 1)(-1)^n(\Delta t)^2 y_n + O((\Delta t)^4) . \]

Now Taylor's theorem may be applied to the \( a_{ij} \)'s to obtain

\[ \sum_{i,j=1}^{g} \frac{\partial}{\partial x_i} (a_{ij}(\xi_n)) \frac{\partial \xi_n}{\partial x_j} \]

\[ = \sum_{i,j=1}^{g} \frac{\partial}{\partial x_i} \left[ (a_{ij}(u_n)) - (\Delta t)^2 \frac{\partial a_{ij}(u_n)}{\partial u} y_n - (-1)^n(\Delta t)^2 \frac{\partial a_{ij}(u_n)}{\partial u} z_n + O((\Delta t)^4) \right] \]

\[ = \sum_{i,j=1}^{g} \frac{\partial}{\partial x_i} (a_{ij}(u_n)) \frac{\partial u_n}{\partial x_j} \]

\[ - (\Delta t)^2 \sum_{i,j=1}^{g} \frac{\partial}{\partial x_i} (a_{ij}(u_n)) \frac{\partial y_n}{\partial x_j} + \frac{\partial a_{ij}(u_n)}{\partial u} \frac{\partial u_n}{\partial x_j} \cdot y_n - \theta a_{ij}(u_n) \frac{\partial^3 u_n}{\partial x_j \partial t^2} \]

\[ + (-1)^n(\Delta t)^2 \sum_{i,j=1}^{g} \frac{\partial}{\partial x_i} ((4\theta - 1)a_{ij}(u_n)) \frac{\partial z_n}{\partial x_j} - \frac{\partial a_{ij}(u_n)}{\partial u} \frac{\partial u_n}{\partial x_j} \cdot z_n + O((\Delta t)^4) . \]
Similarly
\[
\sum_{i=1}^{g} b_i(\xi_n') \frac{\partial \xi_{n\theta}}{\partial x_i}
\]
\[
= \sum_{i=1}^{g} (b_i(u_n') - (\Delta t)^2 \frac{\partial b_i(u_n)}{\partial u} y_n - (-1)^n(\Delta t)^2 \frac{\partial b_i(u_n)}{\partial u} z_n + O((\Delta t)^4))
\]
\[
\left(\frac{\partial u_n}{\partial x_i} + \theta(\Delta t)^2 \frac{\partial^3 u}{\partial x_i \partial t^2} - (\Delta t)^2 \frac{\partial y_n}{\partial x_i} + (-1)^n(\Delta t)^2(4\theta - 1) \frac{\partial z_n}{\partial x_i} + O((\Delta t)^4))\right)
\]
\[
= \sum_{i=1}^{g} b_i(u_n) \frac{\partial u_n}{\partial x_i}
\]
\[
- (\Delta t)^2 \sum_{i=1}^{g} (b_i(u_n) \frac{\partial y_n}{\partial x_i} + \frac{\partial b_i(u_n)}{\partial u} \cdot \frac{\partial u_n}{\partial x_i} y_n - \theta b_i(u_n) \frac{\partial^3 u_n}{\partial x_i \partial t^2})
\]
\[
+ (-1)^n(\Delta t)^2 \sum_{i=1}^{g} ((4\theta - 1)b_i(u_n) \frac{\partial z_n}{\partial x_i} - \frac{\partial b_i(u_n)}{\partial u} \cdot \frac{\partial u_n}{\partial x_i} \cdot z_n) + O((\Delta t)^4).
\]

Finally
\[
c(\xi_n') \xi_n + f(\xi_n) = c(u_n') u_n + f(u_n)
\]
\[
- (\Delta t)^2 \frac{\partial (c(u_n') u_n + f(u_n))}{\partial u} y_n - (-1)^n(\Delta t)^2 \frac{\partial (c(u_n') u_n + f(u_n))}{\partial u} z_n + O((\Delta t)^4).
\]

Therefore
\[
(\Delta t) \eta_n = \xi_{n+1} - \xi_{n-1} - (2\Delta t) \sum_{i,j=1}^{g} \frac{\partial}{\partial x_i} (a_{ij}(\xi_n') \frac{\partial \xi_{n\theta}}{\partial x_j}) - (2\Delta t) \sum_{i=1}^{g} b_i(\xi_n') \frac{\partial \xi_{n\theta}}{\partial x_i}
\]
\[
- (2\Delta t)c(\xi_n') \xi_n - (2\Delta t)f(\xi_n)
\]
\[
= (2\Delta t) \left[ \frac{\partial u_n}{\partial t} - \sum_{i,j=1}^{g} \frac{\partial}{\partial x_i} (a_{ij}(u_n') \frac{\partial u_n}{\partial x_j}) - b(u_n') \cdot \nabla u_n - c(u_n') u_n - f(u_n) \right]
\]
\[-2(\Delta t)^3 \frac{\partial y_n}{\partial t} - \sum_{i,j=1}^{q} \frac{\partial}{\partial x_i} (a_{ij}(u_n) \frac{\partial y_n}{\partial x_j}) - \sum_{i,j=1}^{q} \frac{\partial}{\partial x_i} (\frac{\partial a_{ij}(u_n)}{\partial u} \frac{\partial u_n}{\partial x_j} y_n)\]
\[-\sum_{i=1}^{q} (b_i(u_n) \frac{\partial y_n}{\partial x_i} + \frac{\partial b_i(u_n)}{\partial u} \frac{\partial u_n}{\partial x_i} y_n) - \frac{\partial (c(u_n)u_n + f(u_n))}{\partial u} y_n\]
\[-\frac{1}{6} \frac{\partial^3 u_n}{\partial t^3} + \theta \sum_{i,j=1}^{q} \frac{\partial}{\partial x_i} (a_{ij}(u_n) \frac{\partial^3 u_n}{\partial x_j \partial t^2}) + \theta \sum_{i=1}^{q} b_i(u_n) \frac{\partial^3 u_n}{\partial x_i \partial t^2}\]
\[+2(-1)^n(\Delta t)^3 \left[ \frac{\partial z_n}{\partial t} - (4\theta - 1) \sum_{i,j=1}^{q} \frac{\partial}{\partial x_i} (a_{ij}(u_n) \frac{\partial z_n}{\partial x_j}) + \sum_{i,j=1}^{q} \frac{\partial}{\partial x_i} \left( \frac{\partial a_{ij}(u_n)}{\partial u} \frac{\partial u_n}{\partial x_j} z_n \right) \right.\]
\[\left. - (4\theta - 1) \sum_{i=1}^{q} b_i(u_n) \frac{\partial z_n}{\partial x_i} + \sum_{i=1}^{q} \frac{\partial b_i(u_n)}{\partial u} \frac{\partial u_n}{\partial x_i} z_n + \frac{\partial (c(u_n)u_n + f(u_n))}{\partial u} z_n \right] + O((\Delta t)^5).\]

From (3.1), (6.1) and (6.2) the right hand side of the above equation is seen to be $O((\Delta t)^5)$ and the desired result is established.

**Lemma 6.2:** Suppose that in addition to the hypotheses of Lemma 6.1 we have that

\[(6.4) \quad y_o = -z_o = \frac{1}{4} \frac{\partial^2 u_o}{\partial t^2} - \theta \sum_{i,j=1}^{q} \frac{\partial}{\partial x_i} (a_{ij}(u_o) \frac{\partial^2 u_o}{\partial x_j \partial t}) - \theta \sum_{i=1}^{q} b_i(u_o) \frac{\partial^2 u_o}{\partial x_i \partial t}.\]

If $\eta_o$ is defined by

\[(6.5) \quad \xi - \xi_o - (\Delta t) \sum_{i,j=1}^{q} \frac{\partial}{\partial x_i} (a_{ij}(\xi_o) \frac{\partial \xi_o}{\partial x_j}) - (\Delta t) b(\xi_o) \nabla \xi_o \theta - (\Delta t) c(\xi_o) \xi_o - (\Delta t) f(\xi_o) = \eta_o \]

then $\eta_o = O((\Delta t)^4)$.

**Proof:** Note that $\xi_o = u_o$ since $y_o = -z_o$. But
\[
\xi_1 = u_1 - (\Delta t)^2 y_1 + (\Delta t)^2 z_1 \\
= u_o + (\Delta t) \frac{\partial u_o}{\partial t} + \frac{(\Delta t)^2}{2} \frac{\partial^2 u_o}{\partial t^2} + \frac{(\Delta t)^3}{6} \frac{\partial^3 u_o}{\partial t^3} - (\Delta t)^2 y_o - (\Delta t)^3 \frac{\partial y_o}{\partial t} \\
+ (\Delta t)^2 z_o + O((\Delta t)^4).
\]

Therefore
\[
\xi_1 - \xi_o = (\Delta t) \frac{\partial u_o}{\partial t} + \frac{(\Delta t)^2}{2} \frac{\partial^2 u_o}{\partial t^2} + \frac{(\Delta t)^3}{6} \frac{\partial^3 u_o}{\partial t^3} - 2(\Delta t)^2 y_o - (\Delta t)^3 \frac{\partial y_o}{\partial t} \\
+ (\Delta t)^3 \frac{\partial z_o}{\partial t} + O((\Delta t)^4)
\]

and
\[
\xi_{o\theta} = u_o + 2\theta(\Delta t) \frac{\partial u_o}{\partial t} + \theta(\Delta t)^2 \frac{\partial^2 u_o}{\partial t^2} - 2\theta(\Delta t)^2 y_o + 2\theta(\Delta t)^2 z_o + O((\Delta t)^3)
\]
\[
= u_o + 2\theta(\Delta t) \frac{\partial u_o}{\partial t} + \theta(\Delta t)^2 \frac{\partial^2 u_o}{\partial t^2} - 4\theta(\Delta t)^2 y_o + O((\Delta t)^3).
\]

From this it follows that
\[
\sum_{i,j=1}^{\ell} \frac{\partial}{\partial x_i} - (a_{ij}(\xi_o) \frac{\partial \xi_{o\theta}}{\partial x_j}) = \sum_{i,j=1}^{\ell} \frac{\partial}{\partial x_i} - (a_{ij}(u_o) \frac{\partial u_o}{\partial x_j}) + 2\theta(\Delta t) \sum_{i,j=1}^{\ell} \frac{\partial}{\partial x_i} - (a_{ij}(u_o) \frac{\partial^2 u_o}{\partial x_j \partial t})
\]
\[
+ \theta(\Delta t)^2 \sum_{i,j=1}^{\ell} \frac{\partial}{\partial x_i} - (a_{ij}(u_o) \frac{\partial^3 u_o}{\partial x_j \partial t^2}) - 4\theta(\Delta t)^2 \sum_{i,j=1}^{\ell} \frac{\partial}{\partial x_i} - (a_{ij}(u_o) \frac{\partial y_o}{\partial x_j}) + O((\Delta t)^3)
\]

and
\[
b(\xi_o) \nabla \xi_{o\theta} = b(u_o) \nabla u_o + 2\theta(\Delta t) \sum_{i=1}^{\ell} b_i(u_o) \frac{\partial^2 u_o}{\partial x_i \partial t} + \theta(\Delta t)^2 \sum_{i=1}^{\ell} b_i(u_o) \frac{\partial^3 u_o}{\partial x_i \partial t^2}
\]
\[
- 4\theta(\Delta t)^2 \sum_{i=1}^{\ell} b_i(u_o) \frac{\partial y_o}{\partial x_i} + O((\Delta t)^3).
\]

Finally
\[
c(\xi_o) \xi_o + f(\xi_o) = c(u_o) u_o + f(u_o).
\]
Therefore
\[
\eta_0 = \xi_1 - \xi_0 - (\Delta t) \sum_{i,j=1}^{\frac{g}{2}} \frac{\partial}{\partial x_i} (a_{ij}(\xi_0) \frac{\partial \xi_0}{\partial x_j}) - (\Delta t) \sum_{i=1}^{\frac{g}{2}} b(\xi_0) \cdot \nabla \xi_0 - (\Delta t) c(\xi_0) \xi_0 - (\Delta t) f(\xi_0)
\]
\[
= (\Delta t) \left[ \frac{\partial u_0}{\partial t} - \sum_{i,j=1}^{\frac{g}{2}} \frac{\partial}{\partial x_i} (a_{ij}(u_0) \frac{\partial u_0}{\partial x_j}) - b(u_0) \cdot \nabla u_0 - c(u_0) u_0 - f(u_0) \right]
\]
\[
+ 2(\Delta t)^2 \left[ \frac{1}{4} \frac{\partial^2 u_0}{\partial t^2} \eta_0 - \theta \sum_{i,j=1}^{\frac{g}{2}} \frac{\partial}{\partial x_i} (a_{ij}(u_0) \frac{\partial^2 u_0}{\partial x_j \partial t}) - \theta \sum_{i=1}^{\frac{g}{2}} b_i(u_0) \frac{\partial^2 u_0}{\partial x_i \partial t} \right]
\]
\[
- (\Delta t)^3 \left[ \frac{\partial^2 u_0}{\partial t^2} - \frac{\partial z_0}{\partial t} - \frac{1}{6} \frac{\partial^3 u_0}{\partial t^3} + \theta \sum_{i,j=1}^{\frac{g}{2}} \frac{\partial}{\partial x_i} (a_{ij}(u_0) \frac{\partial^3 u_0}{\partial x_j \partial t^2}) \right]
\]
\[
- 4\theta \sum_{i,j=1}^{\frac{g}{2}} \frac{\partial}{\partial x_i} (a_{ij}(u_0) \frac{\partial y_0}{\partial x_j}) + \theta \sum_{i=1}^{\frac{g}{2}} b_i(u_0) \frac{\partial^3 u_0}{\partial x_i \partial t^2} - 4\theta \sum_{i=1}^{\frac{g}{2}} b_i(u_0) \frac{\partial y_0}{\partial x_i} + O((\Delta t)^4).
\]

The coefficient of $\Delta t$ is zero by equation (3.1) and the coefficient of $(\Delta t)^2$ is zero by equation (6.4). To prove the coefficient of $(\Delta t)^3$ to be zero as well, note that if equation (6.2) is written for $t = 0$ then $-\eta_0$ may be substituted for $z_0$ in every term except $\frac{\partial z_0}{\partial t}$.

This yields
\[
(6.6) \quad \frac{\partial z_0}{\partial t} + (4\theta - 1) \sum_{i,j=1}^{\frac{g}{2}} \frac{\partial}{\partial x_i} (a_{ij}(u_0) \frac{\partial y_0}{\partial x_j}) - \sum_{i,j=1}^{\frac{g}{2}} \frac{\partial}{\partial x_i} \left( \frac{\partial a_{ij}(u_0)}{\partial u} \cdot \frac{\partial u_0}{\partial x_j} \cdot y_0 \right)
\]
\[
+ (4\theta - 1) \sum_{i=1}^{\frac{g}{2}} b_i(u_0) \frac{\partial y_0}{\partial x_i} - \sum_{i=1}^{\frac{g}{2}} \frac{\partial b_i(u_0)}{\partial u} \frac{\partial u_0}{\partial x_i} y_0 - \frac{\partial (c(u_0) u_0 + f(u_0))}{\partial u} y_0 = 0.
\]

Subtracting (6.6) from (6.1) evaluate at $t = 0$ gives
\[
\frac{\partial y_0}{\partial t} - \frac{\partial z_0}{\partial t} - 4\theta \sum_{i,j=1}^{\frac{g}{2}} \frac{\partial}{\partial x_i} (a_{ij}(u_0) \frac{\partial y_0}{\partial x_j}) - 4\theta \sum_{i=1}^{\frac{g}{2}} b_i(u_0) \frac{\partial y_0}{\partial x_i}
\]
\[
= \frac{1}{6} \frac{\partial^3 u_0}{\partial t^3} - \theta \sum_{i,j=1}^{\frac{g}{2}} \frac{\partial}{\partial x_i} (a_{ij}(u_0) \frac{\partial^3 u_0}{\partial x_j \partial t^2}) - \theta \sum_{i=1}^{\frac{g}{2}} b_i(u_0) \frac{\partial^3 u_0}{\partial x_i \partial t^2}.
\]
thus proving the coefficient of \((\Delta t)^3\) to be zero and thereby establishing the lemma.

The remaining work is essentially identical to the theory for linear equations covered in Section 5 and only brief comments will be given in order to clarify the similarities. Notice that the equations for \(y\) and \(z\) are linear and therefore introduce no new difficulties.

**Theorem 6.1:** In addition to the conditions of Lemmas (6.1) and (6.2) assume that \(\{u_i, v_i, w_i, 0 \leq i \leq N\}\) are all elements of \(H^1_0(\Omega)\) and that \(\{U_0, \ldots, U_N\} \subset S^h \subset H^1_0(\Omega)\) is a sequence of functions satisfying (3.6) and (3.8). Then if \(e_n\) and \(\bar{e}_n\) are defined as in Lemma 4.1 and \(\theta > \frac{1}{4}\), (5.8) is valid with \(M = 2\).

**Proof:** Notice that the proof of Theorem 5.2 makes no use of the linearity of (5.9) and (5.10). Non-linear versions of (5.9) and (5.10) follow immediately from (6.3) and (6.5).

The proof of Theorem 5.2 depends heavily upon the \(a_{ij}\) being independent of \(t\). As \(u\) is dependent upon \(t\), the \(a_{ij}\) will vary with \(t\) if they depend upon \(u\). Thus for the proof of Theorem 5.2 to remain valid, the \(a_{ij}\) must be independent of both \(u\) and \(t\).

**Theorem 6.2:** In addition to the conditions of Lemmas 6.1 and 6.2 assume the \(a_{ij}\) to be functions of \(x\) alone and that on \(\partial \Omega\)

\[
\sum_{i,j=1}^{\ell} a_{ij} \frac{\partial v_i}{\partial x_j} = \sum_{i,j=1}^{\ell} a_{ij} \frac{\partial z_i}{\partial x_j} = 0
\]

where \(\nu\) is the outward directed unit normal to \(\partial \Omega\) while \(u\) satisfies (3.3) on \(\partial \Omega\). Let \(\{U_0, \ldots, U_N\} \subset S^h \subset H^1(\Omega)\) be a sequence of functions satisfying (3.7) and (3.9). Then if \(e_n\) and \(\bar{e}_n\) are defined as in Lemma 4.1 and \(\theta > \frac{1}{4}\), (5.8) is valid with \(M = 2\).

**Proof:** Except for the assumption that the \(a_{ij}\) are independent of \(t\) for purposes of evaluating the boundary integrals, the proof of Theorem 5.2 is not dependent upon the linearity of (5.9) and (5.10) and thus serves as a proof of Theorem 6.2.
The $L^2$ error estimate corresponding to Theorem 5.3 is also valid. Notice, however, that nonlinearities cannot occur in the $a_{ij}$.

**Theorem 6.3:** If the hypotheses of Lemmas 4.3, 4.4, 6.1 and 6.2 all hold then

\[(6.7) \quad \|e_N\|_{L^2}^2 \leq C(h^{2r} + (\Delta t)^3 + \|e_0\|_{L^2}^2)\]

**Proof:** As in the proof of Theorem 5.3, both (4.1) and (4.4) are seen to hold from Lemmas 6.1 and 6.2 respectively. But from Lemma 4.3 and 4.4 it is then seen that

\[\|e_N\|_{L^2}^2 \leq C(h^{2r} + \|e_0\|_{L^2}^2 + \|\eta_0\|_{L^2}^2 + (\Delta t) \sum_{n=1}^{N-1} \|\eta_n\|_{L^2}^2)\]

and (6.7) then follows as in the proof of Theorem 5.3.
Appendix A
The Existence of $y^{(m)}$ and $z^{(m)}$ in a Special Case

In Sections 5 and 6 an asymptotic expansion for the error is demonstrated to exist provided that certain associated linear initial-boundary value problems have sufficiently smooth solutions. Although no general theory is known to the author which guarantees this smoothness for any broad class of problems, one can compute the initial conditions and the value of the non-homogeneous term at $t=0$ for these associated problems and use this data to study the smoothness in particular cases.

Study of $y^{(m)}(x,o)$ and $z^{(m)}(x,o)$ requires knowledge of $D_t^i y^{(m-i)}(x,o)$ for $i=m-1$, $m$ and $m+1$ and $D_t^i y^{(m-i)}(x,o)$, $D_t^{i+1} y^{(m-i)}(x,o)$, $D_t^{i-1} z^{(m-i)}(x,o)$, $D_t^i z^{(m-i)}(x,o)$, and $D_t^{i+1} z^{(m-i)}(x,o)$ for all even $i$ with $2 \leq i \leq m-2$. A procedure for determining this data is described below.

In $\mathcal{S}$ the original equation gives $D_t^i u$ as a function of $x$, $t$, $u$, $D_x^i u$, and $D_x^j D_x^j u$ where $i$ and $j$ range between 1 and $l$. If we assume that $u \in C^2(\mathcal{S})$ (as we must if we hope to be able to extrapolate), we may express $D_t^i u(x,o)$ as a function of $x$, $t$, $u_0$, $D_x^i u_0$ and $D_x^j D_x^j u_0$, all of which may be computed from the initial data for $u$. Assuming sufficient smoothness on both $u$ and the coefficients of the equation, we may differentiate the equation and obtain an expression for $D_t^{i+1} u$ in terms of $x$, $t$, $u$, $D_t^i u$, $D_x^i u$, and $D_x^j D_x^j u$. Since at $t=0$ these quantities may be evaluated from the initial data and the function $D_t^i u_0$ computed previously, we can determine $D_t^{i+1} u_0$ provided that $D_t^2 u$ and $D_x^j D_x^j u$ are all continuous in $\mathcal{S}$. Continuing in this manner, we may compute $D_t^i u_0$ for arbitrarily large values of $i$, restricted only by the smoothness of $u$ and of the coefficients of the equation.

Having the derivatives of $u$, one may write down initial conditions for $y^{(2)}$ and $z^{(2)}$. From these initial conditions and $D_t^3 u_0$ expressions for $D_t y^{(2)}_0$ and $D_t z^{(2)}_0$ may be obtained, the expressions being valid under the assumption of adequate smoothness on $y^{(2)}$ and $z^{(2)}$. One may now continue as in the computation of the higher derivatives of $u$, differentiating the equations for $y^{(2)}$ and $z^{(2)}$ to obtain initial values for the higher derivatives of $y^{(2)}$ and $z^{(2)}$, the only change in procedure being that both $D_t^i y^{(2)}_0$ and $D_t^i z^{(2)}_0$ must be computed before either $D_t^{i+1} y^{(2)}_0$ or $D_t^{i+1} z^{(2)}_0$. 
One may now recursively define $y_0^{(m)}$ and $z_0^{(m)}$ in a similar fashion. These manipulations remain valid provided $y^{(m)}$ and $z^{(m)}$ are sufficiently smooth in $\mathcal{A}^\prime$.

For one class of linear problems we not only prove the required smoothness on $y^{(m)}$ and $z^{(m)}$ but actually express these functions in terms of derivatives of $u$. Using the notation of Lemma 3, consider the problem

\[
D_t u - P u = 0
\]
\[
u|_{\partial \Omega} = 0
\]
\[
u(x,0) = u_0(x)
\]

where the coefficients of $P$ are assumed to be time independent. Assume $u \in C^{\infty}(\overline{\Omega} \times [0,\infty))$.

Notice that if $v = D_t^1 u$, then $D_t v - P v = 0$. Also, if

\[
v(x,t) = t^n D_t^1 u(x,t)
\]

then

\[
D_t v - P v = n t^{n-1} D_t^1 u.
\]

If

\[
v(x,t) = D_t^1 u(x,(4\theta-1)t)
\]

then

\[
D_t v - P v = 0
\]

and if

\[
v(x,t) = t^n D_t^1 u(x,(4\theta-1)t)
\]

then

\[
D_t v(x,t) - P v(x,t) = n t^{n-1} D_t^1 u(x,(4\theta-1)t).
\]

Now

\[
\varphi^{(2)} = \left(\frac{1}{6} - \theta\right) D_t^3 u
\]

and

\[
\psi^{(2)} = 0
\]

while

\[
y_0^{(2)} = \left(\frac{1}{4} - \theta\right) D_t^2 u_0
\]
and 
\[ z^{(2)}_0 = -(\frac{1}{4} - \theta) D^2_t u_o. \]

Thus the initial-boundary value problem which \( y^{(2)} \) must solve becomes

\[ D_t y^{(2)} - P_y^{(2)} = (\frac{1}{6} - \theta) D^3_t u \]

\[ y^{(2)}|_{\partial \Omega} = 0 \]

\[ y^{(2)}_o = (\frac{1}{4} - \theta) D^2_t u_o \]

which may be solved by inspection to yield

\[ y^{(2)}(x,t) = t(\frac{1}{6} - \theta) D^3_t u + (\frac{1}{4} - \theta) D^2_t u \]

Similarly we have

\[ z^{(2)}(x,t) = (\theta - \frac{1}{4}) D^2_t u(x,(4\theta - 1)t). \]

Derivation of expressions for \( y^{(4)} \) and \( z^{(4)} \) is more complex, but the techniques are basically those used above.

\[ \psi^{(4)} = \frac{D^5_t u - 10\theta PD^4_t u}{5!} - \frac{D^3_t y^{(2)} - 6\theta PD^2_t y^{(2)}}{6} \]

\[ = (\frac{1}{120} - 10\theta) D^5_t u - \frac{1}{6} D^2_t (\frac{1}{6} - \theta) D^3_t u + t(\frac{1}{6} - \theta) D^4_t u + (\frac{1}{4} - \theta) D^3_t u \]

\[ - 6\theta(t(\frac{1}{6} - \theta) D^4_t u + (\frac{1}{4} - \theta) D^3_t u)) \]

\[ = -(3\theta^2 - \frac{7\theta}{6} + \frac{7}{60})D^5_t u - t(\theta - \frac{1}{6})^2 D^6_t u \]

\[ \psi^{(4)} = \frac{1}{6} D^3_t z^{(2)} + \theta PD^2_t z^{(2)} \]

\[ = \frac{1}{6}(\theta - \frac{1}{4})(4\theta - 1)^3 D^5_t u(x,(4\theta - 1)t) + \theta(\theta - \frac{1}{4})(4\theta - 1)^2 D^5_t u(x,(4\theta - 1)t) \]

\[ = \frac{1}{24}(2\theta + 1)(4\theta - 1)^3 D^5_t u(x,(4\theta - 1)t) \]
\[
\begin{align*}
\gamma_0^{(4)} &= \frac{1}{48} \text{D}^4_t u_0 - \frac{\theta \text{PD}_t^2 u_0}{6} - \frac{1}{4} (\text{D}^2_t \gamma_0^{(2)} - \text{D}^2_t z_0^{(2)}) + \theta \text{PD}_t \gamma_0^{(2)} - \theta \text{PD}_t z_0^{(2)} \\
&= (\frac{1}{48} - \frac{\theta}{6}) \text{D}^4_t u_0 - \frac{1}{4} \text{D}^2_t (t(6-\theta) \text{D}^2_t u + (1-\theta) \text{D}^2_t u - \theta (4\theta - 1) \text{D}^2_t u(x,(4\theta - 1)t)) \bigg|_{t=0} \\
&+ \theta \frac{(1-\theta)}{6} \text{D}^4_t u_0 + (\frac{1}{4} - \theta) \text{D}^4_t u_0 - \theta (\theta - \frac{1}{4}) (4\theta - 1) \text{D}^4_t u_0 \\
&= (\frac{1}{4} - \theta)(\frac{1}{4} + 2\theta) \text{D}^4_t u_0 \\
\zeta_0^{(4)} &= (\theta - \frac{1}{4})(2\theta + \frac{1}{4}) \text{D}^4_t u_0.
\end{align*}
\]

As before, the initial-boundary value problems which \( \gamma^{(4)} \) and \( \zeta^{(4)} \) satisfy are easily solved by inspection to yield

\[
\gamma^{(4)}(x,t) = (\frac{1}{4} - \theta)(\frac{1}{4} + 2\theta) \text{D}^4_t u(x,t) - t(3\theta^2 - \frac{7\theta}{6} + \frac{7}{60}) \text{D}^5_t u(x,t) - \frac{t^2}{2} (\theta - \frac{1}{4}) \text{D}^6_t u(x,t)
\]

and

\[
\zeta^{(4)}(x,t) = (\theta - \frac{1}{4})(2\theta + \frac{1}{4}) \text{D}^4_t u(x,(4\theta - 1)t) + \frac{t}{24} (2\theta + 1)(4\theta - 1)^3 \text{D}^5_t u(x,(4\theta - 1)t).
\]

One may continue in this fashion, obtaining in the differential equation for \( \gamma^{(m)} \) a non-homogeneous part which is a sum of terms like \( t^n \text{D}_t^i u \) where \( i-n = m+1 \). Also, \( \gamma_0^{(m)} \) will be given as a linear combination of terms like \( t^n \text{D}_t^i u_0 \) where \( i-n = m \). The initial-boundary value problem can then be solved to obtain a solution expressed as a linear combination of terms like \( t^n \text{D}_t^i u \) where \( i-n = m \). Similarly, the initial-boundary value problem for \( \zeta^{(m)} \) may be solved to give \( \zeta^{(m)} \) as a linear combination of terms like \( t^n \text{D}_t^i u(x,(4\theta - 1)t) \).
Appendix B

Piecewise Hermite Polynomials

The approximation results obtained in this thesis depend upon the choice of the subspace $S^h$ of $H^1$. One popular choice for $S^h$ is the class of piecewise Hermite polynomials, which we now discuss.

Let $\pi$ be a partition of $\bar{\Omega} = [0,1]$, $0 = x_0 < \cdots < x_n = 1$. Then $H^m(\pi)$ is the set of all functions, $f$, of class $C(m-1)(\bar{\Omega})$ such that $f$ is a polynomial of degree $2m-1$ on each interval $[x_i; x_{i-1}]$. If $\pi$ is a uniform partition with $x_{i+1} - x_i = h$ for each interval then it follows from the work of Schultz [11] that $H^m(\pi)$ is an $S_{m,2m}^0$ space and $H^m(\pi) \cap H_0^1(\Omega)$ is an $S_{m,2m}^1$ space. In fact, Schultz [11] shows that tensor products of $n$ copies of $H^m(\pi)$ is an $S_{m,2m}^0(\Omega^n)$ space while the tensor product of $n$ copies of $H^m(\pi) \cap H_0^1(\Omega)$ is an $S_{m,2m}^1(\Omega^n)$ space.

The most commonly used $H^m(\pi)$ spaces are $H^1(\pi)$, the "chapeau" or "roof-top" functions, and $H^2(\pi)$, the piecewise Hermite cubics. For calculations to be carried out with any choice of $S^h$, a basis for $S^h$ must be chosen, and the calculations are simplified if the basis is chosen so as to make the matrix $G$ defined by (3.11) sparse.

For $H^1(\pi)$, where $\pi$ has a uniform mesh length of $h$, a suitable basis is $v_0, \cdots, v_n$ where

$$v_0(x) = \begin{cases} 1 - \frac{x}{h} & \text{if } x \in [0,h] \\ 0 & \text{otherwise} \end{cases}$$

$$v_n(x) = \begin{cases} 1 - \frac{(1-x)}{h} & \text{if } x \in [1-h,1] \\ 0 & \text{otherwise} \end{cases}$$

and for $1 \leq i \leq n$

$$v_i(x) = \begin{cases} \frac{x-x_{i-1}}{h} & \text{if } x \in [x_{i-1},x_i] \\ \frac{x_{i+1}-x}{h} & \text{if } x \in [x_i,x_{i+1}] \\ 0 & \text{otherwise} \end{cases}$$
With this choice of a basis for \( H^{(1)}(\pi) \), \( G \) is a tridiagonal matrix. If \( v_0 \) and \( v_n \) are eliminated then a basis for \( H^{(1)}(\pi) \cap H^1_0(I) \) is obtained.

For \( H^{(2)}(\pi) \) a suitable basis is the union of two sets of functions, \( v_0, \ldots, v_n \) and \( s_0, \ldots, s_n \) defined as follows:

\[
v_0(x) = 1 - 3\left(\frac{x}{h}\right)^2 + 2\left(\frac{x}{h}\right)^3 \quad \text{if } x \in [0,h]
\]

\[
v_0(x) = s_0(x) = 0 \quad \text{otherwise};
\]

\[
v_n(x) = 1 - 3\left(\frac{1-x}{h}\right)^2 + 2\left(\frac{1-x}{h}\right)^3 \quad \text{if } x \in [1-h,1]
\]

\[
v_n(x) = s_n(x) = 0 \quad \text{otherwise};
\]

and for \( 1 \leq i < n \)

\[
v_i(x) = 1 - 3\left(\frac{x-x_i}{h}\right)^2 - 2\left(\frac{x-x_i}{h}\right)^3 \quad \text{if } x \in [x_{i-1},x_i]
\]

\[
v_i(x) = s_i(x) = \left(\frac{x-x_i}{h}\right)^2 \quad \text{if } x \in [x_i,x_{i+1}]
\]

\[
v_i(x) = s_i(x) = 0 \quad \text{otherwise}.
\]

If \( G \) is defined with this basis ordered as \( v_0, s_0, v_1, s_1, \ldots, v_n, s_n \) then \( G \) is block tridiagonal with 2x2 blocks. If \( v_0 \) and \( v_n \) are deleted then one obtains a basis for \( H^2(\pi) \cap H^1_0(I) \).
Appendix C

A Special Result for the Case $\theta = \frac{1}{4}$

In Sections 5 and 6 extrapolation of the values of $U_n$ has been shown to be possible provided $\theta > \frac{1}{4}$. Results for $\theta \leq \frac{1}{4}$ were not obtained because Lemmas 4.1 and 4.3 do not apply for such values of $\theta$. Indeed, if $\theta < \frac{1}{4}$, the three level method is unstable. If (3.1) is linear with $c=0$ and the $a_{ij}, b_i$ and $f$ are all independent of $t$, then the three level method with $\theta = \frac{1}{4}$ is equivalent to looking at even steps of the three level method with $\theta = \frac{1}{2}$ with the time step halved (the Crank-Nicolson-Galerkin approximation [4]), and in this case extrapolation is possible. It has been conjectured in [4] that a better approximation to $u_n$ is given by $U_{n,\frac{1}{4}}$. We now show that the $U_{n,\frac{1}{4}}$ can be extrapolated even if the $a_{ij}, b_i$ and $f$ depend on $t$.

Lemma: Suppose (3.1) is linear with $c=0$. Let $\xi_0, \ldots, \xi_{N+1}$ be a sequence of elements of $H^1_0(\Omega)$ satisfying (4.1) for $n=1, \ldots, N$ and also satisfying (4.4) with $\theta = \frac{1}{4}$. Let $U_0, \ldots, U_{N+1}$ be a sequence of elements of $S^h \subset H^1_0(\Omega)$ satisfying (3.6) and (3.8) with $\theta = \frac{1}{4}$. Define $e_n$ and $\bar{e}_n$ as in Lemma 4.1. Then for $\Delta t$ sufficiently small,

\[
(C1) \quad \|e_{N,\frac{1}{4}}\|_L^2 \leq C\left(\|e_0\|_L^2 + \|\bar{e}_0, \frac{1}{4}\|_H^1 + \|e_{N,\frac{1}{4}}\|_L^2 + \max_{1 \leq n \leq N} \|e_{n,\frac{1}{4}}\|_L^2 + \|\eta_0\|_L^2\right)
+ (C\Delta t) \sum_{n=1}^{N} \|\eta_n\|_L^2 + \|\bar{e}_{n,\frac{1}{4}}\|_H^1 \right) + (C\Delta t) \sum_{n=2}^{N} \left\|\bar{e}_{n,\frac{1}{4}} - \bar{e}_{n-1,\frac{1}{4}}\right\|_L^2 \frac{\Delta t}{\Delta t} .
\]

Proof: Subtracting (4.1) from (3.8) gives

\[
\langle e_{n+1} - e_{n-1}, v \rangle + (2\Delta t)a(e_{n,\frac{1}{4}}, v) + (2\Delta t)\langle b \cdot \nabla e_{n,\frac{1}{4}}, v \rangle = (\Delta t)\langle \eta_n, v \rangle .
\]

Letting $v = e_{n,\frac{1}{4}} + \bar{e}_{n,\frac{1}{4}}$ gives

\[
\frac{1}{4}\langle (e_{n+1} + e_n) - (e_n + e_{n-1}), (e_{n+1} + e_n) + (e_n + e_{n-1}) \rangle + (2\Delta t)a(e_{n,\frac{1}{4}}, e_{n,\frac{1}{4}})
+ (2\Delta t)\langle b \cdot \nabla e_{n,\frac{1}{4}}, e_{n,\frac{1}{4}} \rangle = (\Delta t)\langle \eta_n, e_{n,\frac{1}{4}} \rangle + (\Delta t)\langle \eta_n, \bar{e}_{n,\frac{1}{4}} \rangle
+ \langle e_{n-1} - e_{n+1}, \bar{e}_{n,\frac{1}{4}} \rangle - (2\Delta t)a(e_{n,\frac{1}{4}}, \bar{e}_{n,\frac{1}{4}}) - (2\Delta t)\langle b \cdot \nabla e_{n,\frac{1}{4}}, \bar{e}_{n,\frac{1}{4}} \rangle .
\]
Therefore
\[ \frac{1}{2}( |l e_{n+1} + e_n |^2_{L^2} - |l e_n + e_{n-1} |^2_{L^2} ) + (2\Delta t) c_0 |l e_n, \gamma |^2_{H^1_0} \leq (2\Delta t) c_0 |l e_n, \gamma |^2_{H^1_0} \]
\[ + (C\Delta t)( |l e_n, \gamma |^2_{L^2} + |l e_n - e_{n+1}, \bar{\gamma} |^2_{L^2} + |l e_{n-1}, \gamma |^2_{L^2} + |l e_{n-1}, \bar{\gamma} |^2_{L^2} ) + \langle e_{n-1} - e_{n+1}, \bar{\gamma} |^2_{L^2} \rangle \]
\[ \leq (2\Delta t) c_0 |l e_n, \gamma |^2_{H^1_0} + (C\Delta t)( |l e_{n+1} + e_n |^2_{L^2} + |l e_{n-1} - e_{n-1}, \bar{\gamma} |^2_{L^2} ) \]
\[ + (C\Delta t)( |l e_n |^2_{L^2} + |l e_{n-1}, \gamma |^2_{L^2} ) + \langle e_{n-1} - e_{n+1}, \bar{\gamma} |^2_{H^1} \rangle . \]

Summing from \( n = 1 \) to \( N - 1 \) gives
\[ |l e_{N+1} + e_N |^2_{L^2} - |l e_1 + e_0 |^2_{L^2} \leq (C\Delta t) \sum_{n=1}^{N} |l e_n - e_{n-1} |^2_{L^2} \]
\[ + (C\Delta t) \sum_{n=1}^{N-1} ( |l e_n |^2_{L^2} + |l e_{n-1} |^2_{L^2} ) + 4 \sum_{n=1}^{N-1} \langle e_{n-1} - e_{n+1}, \bar{\gamma} |^2_{L^2} \rangle . \]

The last term may be summed by parts as follows:
\[ \sum_{n=1}^{N-1} \langle e_{n-1} - e_{n+1}, \bar{\gamma} |^2_{L^2} \rangle = \sum_{n=1}^{N-1} \langle (e_{n-1} + e_n) - (e_n + e_{n+1}), \bar{\gamma} |^2_{L^2} \rangle \]
\[ = \sum_{n=2}^{N-1} \langle e_{n-1} + e_n, \bar{\gamma} |^2_{L^2} - e_{n-1}, \bar{\gamma} |^2_{L^2} \rangle + \langle e_0 + e_1, \bar{\gamma} |^2_{L^2} \rangle - \langle e_N + e_{N-1}, \bar{\gamma} |^2_{L^2} \rangle \]
\[ \leq C( |l e_o |^2_{L^2} + |l e_1 |^2_{L^2} + |l e_{N-1} |^2_{L^2} + |l e_{N-1} |^2_{L^2} ) + (C\Delta t) \sum_{n=2}^{N-1} \left( \frac{ |l e_{n, \gamma} - e_{n-1} |^2_{L^2} }{\Delta t} \right) \]
\[ + 4 |l e_n + e_{n-1} |^2_{L^2} + |l e_{N+1} + e_N |^2_{L^2} . \]

Therefore
\[(1 - \varepsilon - C\Delta t)\|e_{N+1} - e_{N-1}\|_{L^2}^2 \leq (C\Delta t) \sum_{n=1}^{N-1} \|e_n + e_{n-1}\|_{L^2}^2\]

\[+ (C\Delta t) \sum_{n=1}^{N-1} (\|\eta_n\|_{L^2}^2 + \|\tilde{\eta}_n\|_{H^1}^2) + C(\|e_0\|_{L^2}^2 + \|e_1\|_{L^2}^2 + \|\tilde{\eta}_1\|_{H^1}^2)\]

\[+ \max_{1 \leq n \leq N-1} \|\tilde{\eta}_{n,1/4}\|_{L^2}^2 + (4\Delta t) \sum_{n=2}^{N-1} \left\|\frac{\tilde{\eta}_{n,1/4} - \tilde{\eta}_{n-1,1/4}}{\Delta t}\right\|_{L^2}^2\]

and, reasoning as in the proof of Lemma 4.1, it follows from the discrete Gronwall's inequality that

\[\|e_{N+1} - e_{N-1}\|_{L^2}^2 \leq \min\left\{ C(\|e_0\|_{L^2}^2 + \|e_1\|_{L^2}^2 + \|\tilde{\eta}_1\|_{H^1}^2) \right\} \sum_{n=1}^{N-1} (\|\eta_n\|_{L^2}^2 + \|\tilde{\eta}_n\|_{H^1}^2) + (C\Delta t) \sum_{n=2}^{N-1} \left\|\frac{\tilde{\eta}_{n,1/4} - \tilde{\eta}_{n-1,1/4}}{\Delta t}\right\|_{L^2}^2\]

But a review of the proof of Lemma 4.2 reveals that if (3.1) is linear, then the assumptions that \(|\varphi_n|\) and \(i\nabla |\varphi_n|\) are bounded are not needed. As a result, (4.5) holds and

\[\|e_{N+1} - e_{N-1}\|_{L^2}^2 \leq \min\left\{ C(\|e_0\|_{L^2}^2 + \|\tilde{\eta}_0\|_{H^1}^2 + \|\tilde{\eta}_1\|_{H^1}^2 + \|\tilde{\eta}_n\|_{L^2}^2 + \|\eta_0\|_{L^2}^2) \right\} \sum_{n=1}^{N-1} (\|\eta_n\|_{L^2}^2 + \|\tilde{\eta}_n\|_{H^1}^2) + (C\Delta t) \sum_{n=2}^{N-1} \left\|\frac{\tilde{\eta}_{n,1/4} - \tilde{\eta}_{n-1,1/4}}{\Delta t}\right\|_{L^2}^2\]

Similarly,

\[\|e_{N+1} - e_{N-1}\|_{L^2}^2 \leq \min\left\{ C(\|e_0\|_{L^2}^2 + \|\tilde{\eta}_0\|_{H^1}^2 + \|\tilde{\eta}_1\|_{H^1}^2 + \|\tilde{\eta}_n\|_{L^2}^2 + \|\eta_0\|_{L^2}^2) \right\} \sum_{n=1}^{N} (\|\eta_n\|_{L^2}^2 + \|\tilde{\eta}_n\|_{H^1}^2) + (C\Delta t) \sum_{n=2}^{N} \left\|\frac{\tilde{\eta}_{n,1/4} - \tilde{\eta}_{n-1,1/4}}{\Delta t}\right\|_{L^2}^2\]
and since
\[ \| e_{N+1} \|^2 \leq \frac{1}{2} \left( \| e_{N+1} \|^2 + \| e_{N} \|^2 + \| e_{N-1} \|^2 \right), \]
(C1) follows.

If the \( y^{(m)} \) and \( z^{(m)} \) exist and are sufficiently smooth, it is shown on page 31 that \( \xi_{n,\frac{1}{4}} \) has an expansion in powers of \( (\Delta t)^2 \). It then follows from C1 that the \( U_{n,\frac{1}{4}} \) may be extrapolated.

It is worth noting, in fact, that these "smoothed" values do not have an oscillatory contribution arising from the \( y^{(m)} \), which are, in this case, functions of \( x \) only. In [12] Stetter proposes a similar smoothing process for extrapolation methods for the numerical solution of ordinary differential equations.
REFERENCES


