To the Graduate Council:

I am submitting herewith a dissertation written by William J. Layton entitled "The Galerkin Method for First Order Hyperbolic Equations." I recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Mathematics.

Max D. Gunzburger, Major Professor

We have read this dissertation
and recommend its acceptance:

Vassilios Dougalis

Steven M. Serbin

G. Samuel Jordan

Robert Todd Gregory

Accepted for the Graduate Council:

Vice Chancellor
Graduate Studies and Research

THE GALERKIN METHOD FOR FIRST ORDER HYPERBOLIC EQUATIONS

A Dissertation

Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

William J. Layton

August 1980

3045025

To Alice Gordon

ACKNOWLEDGEMENT

I wish to thank my advisor, Dr. Max Gunzburger, and Dr. Vasilios Dougalis for their guidance, encouragement and accessibility. My training in applied math is largely due to these two men. In addition, I would like to thank Mrs. Louise Stokes and Dr. Jerry Scott for teaching me how to do proofs.

I would also like to acknowledge the support given me by friends. In particular I would like to thank Dr. David Dobbs, Rick Gardner, Pam Stafford, Janet Peterson, Chip Eastham, Jim Chesney, Sarah Coode and Alice Gordon, Many others are due thanks.

Special thanks are due to Cindi Blair for the excellent, professional typing job done.

ABSTRACT

The purpose of this thesis is to study and to resolve certain problems arising in the approximate solution, via Galerkin methods, of first order hyperbolic equations and systems of equations. In particular three general questions are considered and particular cases of them answered satisfactorily. The type of equations and questions studied are chosen as models for a theoretical analysis of questions important in applications.

The first question which is addressed is the problem of stability. Just as in finite difference theory, many Galerkin methods, stable and convergent for the Cauchy problem, are unstable for hyperbolic boundary value problems because of improper treatment of the boundary conditions. The usual Galerkin method is seen to be unstable for a linear, constant coefficient, well-posed hyperbolic system on [0, 1]. A method of treating boundary conditions is then proposed which will yield a stable and convergent method for any well-posed, linear hyperbolic system in one dimension. This idea generalizes in a natural way to problems in more than one space dimension. Examples in one and two dimensions are considered.

The approximate solution, by two dissipative Galerkin methods, of the initial-value problem

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x}\ , \quad u(x, 0) = v(x)\ ,$$

is the second problem considered. Convergence estimates in L_2 and

ℓ_2 are derived for smooth solutions for semidiscrete schemes. When v is taken to have a jump discontinuity at zero that discontinuity will propogate along $x + t = 0$. Estimates in L_2 and L_∞ of the pollution effects of the discontinuity are found. These estimates show these effects to decay exponentially in h^{-1} in regions a fixed distance d from the discontinuity and exponentially in d for fixed h. Finally, analogous estimates are shown to hold under the Crank-Nicholson time discretization.

Finally, asymptotic error estimates are derived for an H^{-1} method for an initial-boundary value problem for a hyperbolic equation. This method is optimal in L^2 and computes the approximation in piecewise polynomial spaces.

TABLE OF CONTENTS

CHAPTER PAGE

I. INTRODUCTION AND DISCUSSION 1

II. STABLE GALERKIN METHODS FOR HYPERBOLIC SYSTEMS 11

Introduction 11

Notation 16

The General Hyperbolic System. 17

Galerkin Methods for Hyperbolic Systems. 23

Examples 34

III. ESTIMATES AWAY FROM A DISCONTINUITY. 51

Introduction 51

Notation and Spaces. 53

Factoring the Galerkin Operator. 57

Convergence for Smooth Solutions 63

Discontinuous Initial Data 66

Fully Discrete Estimates 72

IV. ERROR ESTIMATES FOR A VARIABLE COEFFICIENT INITIAL-BOUNDARY VALUE PROBLEM 81

Introduction 81

Notation and Spaces. 82

A Projection Theorem 84

The Semi-Discrete Method 92

LIST OF REFERENCES 98

VITA 105

CHAPTER I

INTRODUCTION AND DISCUSSION

In this dissertation the approximate solution, via Galerkin methods, of first order hyperbolic equations and systems of hyperbolic equations is considered. The question of approximating the solutions of such equations is a problem frequently encountered in many settings. These equations arise in problems in acoustics, fluid flow, wave propagation, meteorology, gas dynamics, movement of glaciers, geometrical optics, neutron transport and many other important areas. Although different physical mechanisms are acting in these phenomena they share the property that mathematically their governing equations are hyperbolic. Thus, such diverse physical phenomena pose a common set of problems in numerical analysis. It is this set of problems that is partially addressed in this dissertation. The goal of this work is a mathematically rigorous analysis of such problems and thus, only rather simple model problems can be considered. Although the problems considered herein are far simpler than those encountered in applications, the difficulties encountered when one attempts to approximate the solution to these simple problems are similar to those encountered in practical problems. Thus, the behavior of Galerkin methods for the approximation of the problems considered here is of interest.

To place the problems and results in this thesis in their proper perspective we shall begin with a historical survey of the literature of finite difference and Galerkin methods for first order

hyperbolic equations and a description of important questions, both solved and unsolved. A good general reference for finite difference theory for hyperbolic and parabolic equations is the set of notes of Kreiss and Oliger [35]. Finite difference theory possesses a large body of literature concerned with the approximate solution of first order hyperbolic equations. This literature contains both rigorous analysis of model problems and experimental studies of difficult real-world problems. The notes [35] and the book [56] contain extensive references to this literature.

One of the first papers studying the finite element method for hyperbolic equations was Swartz and Wendroff [64]. This paper discussed the relationship of semiboundedness of A and stability for the approximation scheme for the equation

$$\frac{\partial u}{\partial t} = Au + f \ . \tag{1.1}$$

This seminal paper contains many ideas developed fully in later work. When $A = A(u)$ and hence (1.1) is nonlinear, it is necessary to solve a nonlinear system at every timestep to calculate u at the new value of t. To avoid this they discussed the idea of extrapolating u in the nonlinearity from old values of t to the new value of t, thus avoiding this costly operation. This idea has recently been analyzed for parabolic equations [19], [10] and the Navier-Stokes equations [8]. Also this paper discusses the use of higher order elements at the boundary to overcome loss of accuracy. Another important idea occurring in [69] is the idea of using nonstandard

variational principles to introduce artificial viscosity for problems with shocks. For example, for the equation

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = 0 \ , \ u(x, 0) = u_0(x)$$

Swartz and Wendroff [64] considered the method of computing $u^h: [0, T] \to S^h$, S^h a space of piecewise linear functions on a grid of meshwidth h, by

$$(\frac{\partial u^h}{\partial t} \ , \ v) + (u^h \frac{\partial u^h}{\partial x} \ , \ v) + \varepsilon(\frac{\partial u^h}{\partial x} \ , \ \frac{\partial v}{\partial x}) = 0 \ ,$$

for all $v \in S^h$. They studied computationally the effect of the size of ε upon the resolution of high gradients. After this paper many authors considered the question of the error in the approximation to the Cauchy problem, both periodic and non-periodic; see for example [20], [22], [67], [68], [69] or [74]. Typically, optimal rates of convergence are shown in spaces of spline functions. These results depend heavily upon Fourier analysis and thus have not extended to problems with non-periodic boundary conditions on bounded domains.

The paper [20] of Dupont presented interesting results. He considered the question of approximating 1-periodic solutions to the initial value problem

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0 \ , \ u = u(x, t) \ , \ -\infty < x < \infty \ , \ t \geq 0 \ ,$$

$$u(x, 0) = u_0(x) \ , \ u_0(x + 1) = u_0(x) \ , \ -\infty < x < \infty \ ,$$

in spaces of 1-periodic piecewise linear functions, 1-periodic Hermite cubic functions and 1-periodic cubic splines. That is, for S^h one of these spaces, he considered the approximation $u^h: [0, T] \to S^h$ defined by

$$\left(\frac{\partial u^h}{\partial t} + \frac{\partial u^h}{\partial x}\right), v) = 0 \text{ , for all } v \in S^h \text{ ,}$$

$$u^h(0) \in S^h \text{ , an approximation to } u_0 \text{ ,}$$

where $(\cdot, \cdot)$ represents the usual inner product on $L^2(0, 1)$. He shows the error, $||u - u^h||$, for piecewise linear functions and cubic splines to be optimal, $O(h^2)$ and $O(h^4)$ respectively. However he also shows the error in the Galerkin approximation in Hermite cubic spaces to be precisely $O(h^3)$ and not $O(h^4)$. This result shows that, in general, one cannot expect optimal rates of convergence for the finite element method for first order hyperbolic equations in piecewise polynomial spaces.

This result of Dupont prompted many people to consider nonstandard Galerkin methods based on other variational principles. Their goal was to formulate methods for which the rates of convergence were optimal, independent of the exact choice of subspaces or the exact form of the error of the interpolant in that space; see, for example [9], [17], [38], [39], [71] or [72]. Some of these methods have been extended to systems, some to problems with nonperiodic boundary conditions, and some to nonlinear problems.

The question of boundary conditions for scalar equations and especially for systems is an important and difficult question. For example, if the analysis of Dupont in [20] is applied to the simple boundary value problem

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0 \ , \ u = u(x, t) \ , \ 0 \le x \le 1 \ , \ 0 \le t \ ,$$

$$u(0, t) = 0 \ , \ u(x, 0) = u_0(x) \ , \ 0 \le x \le 1 \ , \ 0 \le t \ ,$$

using a space of piecewise linear functions vanishing at $x = 0$, the error in the usual Galerkin method is seen to be $O(h^2)$ (optimal) on interior elements but only $O(h^1)$ on the boundary element at $x = 1$. Thus a loss of accuracy is caused by a boundary even in this simple problem. To get around this loss of accuracy nonstandard methods have been proposed for the boundary value problem for the scalar equation [9], [12].

For systems of equations the question of boundary conditions becomes crucial. In 1976 M. Gunzburger considered a simple example of a well posed, constant coefficient, hyperbolic system that was discretized by the usual Galerkin method [26]. The system considered was

$$\frac{\partial u}{\partial t} = A \frac{\partial u}{\partial x} \ , \ u = \begin{pmatrix} u^1 \\ u^{11} \end{pmatrix} \ , \ u^i = u^i(x, t) \ , \ A = \begin{pmatrix} \nu & 1 \\ 1 & \nu \end{pmatrix} \ ,$$

$$u^1(0, t) = 0 = u^1(1, t) \ , \ 0 \le t \ , \tag{1.2}$$

$$u(x, 0) = u_0(x) \ , \ 0 \le x \le 1 \ ,$$

for $0 < \nu < 1$. The usual Galerkin method for this problem consists of a subspace, S^h, of $L^2(0, 1)$ satisfying the boundary conditions (1.2) and an approximation to u, $u^h: [0, T] \to S^h$, computed by

$$\left(\frac{\partial u^h}{\partial t}, v\right) = \left(A \frac{\partial u^h}{\partial x}, v\right), \text{ for all } v \in S^h,$$

$$u^h(0) \in S^h, \text{ an approximation to } u_0. \tag{1.3}$$

He found computationally that the usual Galerkin method (1.3) is unstable! In fact, as $h \to 0$, the norm of the approximate solution, $\|u^h\|$, grows like $\exp(h^{-1})$ although the norm of the true solution is bounded. For scalar equations the effect of boundaries can be a small loss of accuracy; for systems, boundaries can cause gross instabilities. Other work upon systems includes the work of Dupont [12], and Luskin [43] who consider a special system in one dimension. Their system resembles a nonlinear wave equation written as a first-order system. They prove (suboptimal) error estimates in piecewise polynomial spaces. Baker's method [9] can be extended to hyperbolic systems on $[0, 1]$ with a restriction upon the allowable boundary conditions. Lesaint [38], [39], [40] has considered nonstandard methods for hyperbolic systems that make no distinction between space and time variables. He has shown convergence of these methods for a quite wide class of problems using the methods developed by Friedrichs [23] to study the differential equation. Further Dendy [17] and Wahlbin [72], [73] have considered the effect of the introduction of artificial viscosity into the approximation schemes for systems.

These nonstandard methods are stable for the class of problems considered by their authors. However, typically the problems posed have very special boundary conditions. These restrictions of the boundary conditions avoid the question of instabilities caused by the treatment of more general boundary conditions.

A second question not fully resolved is the problem of approximating the solutions of nonlinear hyperbolic equations. Several authors [17], [73], [72] have treated nonlinear equations by the introduction of dissipation and have used energy methods to derive error estimates for smooth solutions to nonlinear equations. Only Dupont [21], [51] and Luskin [43] have considered strictly conservative methods, such as the usual Galerkin method, to approximate the solutions to such problems. Typically these results have needed the nonlinearity to be globally Lipschitz in u for the estimates to hold. Thus the question of the behavior of such methods on nonlinear equations with smooth solutions is yet unresolved.

The question of discontinuous solutions to hyperbolic equations, or shocks, is intimately tied to the question of nonlinear equations. Finite difference theory possesses a large (and not entirely satisfactory) literature upon the problem of discontinuous solutions and solutions with large gradients. In finite element methods the point values of the approximate solution at every mesh point are usually linked in the approximate equations. Thus the equations arising from the finite element method are similar to those arising from implicit finite difference methods and hence the situation in the finite element method is somewhat more complicated and less studied than for explicit finite

difference methods. In problems with discontinuities or shocks one typically seeks to track accurately the shock position and velocity and to approximate well the solution away from the shock front. Ideas employed include using artificial viscosity away from the front to damp spurious oscillations arising from the front, and attempting to track the shock front in order to splice different methods together around the discontinuity.

In this dissertation three problems arising from the approximation of solutions to hyperbolic equations and systems of equations are discussed. The first is the correct treatment of boundary conditions in the approximate equations for first order hyperbolic systems. It is known from finite difference theory that some difference methods, stable and convergent for the Cauchy problem, are unstable for boundary value problems because of improper treatment of the boundary conditions (see [34], [35], [33], or [25]). Analogous instabilities have been observed for semi-discrete and fully discrete Galerkin methods, [26]. In chapter II this question is resolved for Galerkin methods for variable coefficient hyperbolic systems in one and two space dimensions. More precisely, it is shown how relevant boundary forms should be treated in the semi-discrete and fully discrete Galerkin method to ensure stability of that method for boundary value problems in one and two space dimensions. In chapter III the effort of adding dissipation to variational formulations for computing approximations to hyperbolic equations with discontinuous solutions is investigated. This is an important question since hyperbolic equations are nonsmoothing and since an important class of nonlinear hyperbolic equations

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x} f(u) \ , \ u = u(x, t) \ , \ -\infty < x < \infty \ , \quad t \geq 0 \ ,$$

$$u(x, 0) = u_0(x) \ , \ -\infty < x < \infty \ , \tag{1.4}$$

possesses physically interesting discontinuous solutions, even with $u_0 \in C^\infty$. The analysis in chapter III is restricted to a simple model problem

$$\frac{\partial u}{\partial t} = a \frac{\partial u}{\partial x} \ , \ u = u(x, t) \ , \ -\infty < x < \infty \ , \ t \geq 0 \ ,$$

$$u(x, 0) = u_0(x) \ , \ -\infty < x < \infty \ , \tag{1.5}$$

where u_0 is taken to be piecewise smooth with a jump discontinuity at $x = 0$. For small variations in x and t the behavior of Galerkin methods upon the problem (1.5) will model their behavior upon (1.4). The results of chapter III assert that with proper amounts of dissipation added to the variational principle, solutions to (1.5) may be accurately computed.

In chapter IV a method proposed by G. Baker [9] is investigated for the boundary value problem

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} (A(x, t)u) = f(x, t) \ , \ 0 \leq x \leq 1 \ , \ 0 \leq t \leq T \ ,$$

$$u(x, 0) = u_0(x) \ , \quad 0 \leq x \leq 1 \ ,$$

$$u(0, t) = g(t) \ , \ 0 \leq t \leq T \ .$$

The approximate solution is defined using ideas similar to those employed by Rachford and Wheeler [52] for two point boundary value problems. A projection theorem estimates the L^2 and H^{-1} norms of time derivatives of the error in the projection into piecewise polynomial spaces. Since the coefficients of the differential equation depend upon t, the bilinear form involved will depend upon t also. Thus, estimating these time derivatives is the main technical point in the analysis. Next, asymptotic error estimates are derived for a semi-discrete scheme. This scheme gives errors optimal in the L^2-norm when the approximation is computed in piecewise polynomial spaces.

CHAPTER II

STABLE GALERKIN METHODS FOR HYPERBOLIC SYSTEMS

1. Introduction

Until recently the question of stability of projection methods for evolution equations has not been examined. However, in 1976 the usual Galerkin method was first found to be unstable for hyperbolic systems, [26]. The usual Galerkin method has been shown to be stable and convergent (even superconvergent) for the Cauchy problem and the problem with boundary conditions periodic in the space variables [22], [68], [69] and [74]. Gunzburger observed in 1976 that when this method is applied to initial-boundary value problems for hyperbolic systems the usual Galerkin method can be grossly unstable due to the influence of the boundaries.

This situation is understood in finite difference theory. It is known, in finite difference theory, that the treatment of boundary conditions is critical for the stability and convergence of the method. In fact, many difference methods, stable and convergent for the Cauchy problem, are unstable for problems with boundaries precisely because of improper treatment of the boundary conditions, [34]. A large literature exists examining this question for difference methods; see, for example, [35] and the references therein. Therefore, it is not surprising that similar difficulties occur with the Galerkin method.

The question of proper treatment of relevant boundary forms in the Galerkin method to yield a method stable for boundary value problems with non-periodic boundary conditions is discussed and resolved in this

chapter. This work extends work of Gunzburger [26] to cover the general case of the well-posed, linear, hyperbolic system in one space dimension and a large class of hyperbolic systems in two space dimensions.

First the work of Gunzburger [26] shall be discussed. Consider the initial-boundary value problem

$$\begin{aligned} &\frac{\partial u}{\partial t} = A \frac{\partial u}{\partial x} \quad , \; A = \begin{pmatrix} 1/2 & 1 \\ 1 & 1/2 \end{pmatrix} \; , \; u = \begin{pmatrix} u^1 \\ u^1 \end{pmatrix} \; , \\ &u^1(0, t) = u^1(1, t) = 0 \; , \\ &u(x, 0) = u_0(x) \; , \end{aligned} \tag{2.1}$$

where $0 \leq x \leq 1$ and $0 \leq t$. It is known that this problem is well-posed. That is, $u = u(x, t)$ exists uniquely and further, there is a constant C, independent of t, such that

$$||u(t)|| \leq C||u_0|| \; ,$$

where $||\cdot||$ represents the usual norm on $L^2(0, 1)$: for $f: [0, 1] \to \mathbb{R}^2$ define

$$||f||^2 = (f, f) = \int_0^1 |f(x)|^2 \, dx \; , \; |f|^2 = f^T f \; .$$

This equation is discretized by the usual Galerkin method in a space of cubic splines satisfying the linear, homogeneous boundary conditions of the continuous problem (2.1). More precisely, let χ be the characteristic function of $[-1/2, 1/2]$, then define ϕ and ϕ_ℓ, the

B-splines of order 4, by

$$\phi = \chi^{*4} \text{ , and } \phi_\ell(x) = \phi(h^{-1} x - \ell) \text{ .}$$

For more details about these functions see, for example, [58], [59], or [67]. For $0 < h < 1$, $h \cdot N = 1$, let S^h be constructed as follows

$$S^h = \{\phi = \begin{pmatrix} \phi^1 \\ \phi^2 \end{pmatrix} | \phi^1 \in S_1^h \text{ , } \phi^2 \in S_2^h\} \text{ ,}$$

where S_1^h is the subspace spanned by the functions $\{\psi_0, \psi_1, \ldots, \psi_{N-1}, \psi_N\}$ defined by

$$\psi_0 = \phi_0 - 4\,\phi_{-1} \text{ , } \psi_1 = \phi_1 - \phi_{-1} \text{ ,}$$

$$\psi_j = \phi_j \text{ , } 2 \leq j \leq N - 2$$

$$\psi_{N-1} = \phi_{N-1} - \phi_{N+1} \text{ , } \psi_N = \phi_N - 4\,\phi_{N+1} \text{ ,}$$

all restricted to $[0, 1]$. All elements of S_1^h then vanish at $x = 0$ and $x = 1$. S_2^h is the subspace spanned by the functions $\{\phi_{-1}, \phi_0, \ldots, \phi_N, \phi_{N+1}\}$ all restricted to $[0, 1]$. Note that imposing the boundary conditions upon the subspace only changes the construction of the boundary elements.

The semidiscrete Galerkin approximation is then realized as a differentiable map $U: [0, T] \to S^h$ satisfying

$$\left(\frac{\partial U}{\partial t}, v\right) = \left(A \frac{\partial U}{\partial x}, v\right), \text{ for all } v \in S^h,$$

$$U(0) \in S^h, \text{ approximates } u_0 \text{ well.} \tag{2.2}$$

Expanding

$$U(x, t) = \binom{1}{0} \sum_{j=0}^{N} c_j(t)\, \psi_j(x) + \binom{0}{1} \sum_{j=-1}^{N+1} d_j(t)\, \phi_j(x),$$

and substituting U into (2.2) yields a system of ordinary differential equations for the undetermined coefficients of the form

$$M \dot{X}(t) = K X(t), \quad X(0) \text{ given.} \tag{2.3}$$

For more details on M, K, and $X(t)$ see [26] p. 668 formulas (3.10) and (3.11). Associated with the ordinary differential equation (2.3) is the eigenvalue problem

$$\lambda My = Ky. \tag{2.4}$$

In experiments performed in 1976 (see [26] pp. 668, 669) Gunzburger found that:

1. The eigenvalues of K have zero real part.
2. $(2N - 4)$ eigenvalues of (2.4) have zero real part, two have negative real part, and two have positive real part.
3. As $h \to 0$ the real parts of the eigenvalues with positive real parts grow as $1/h$.

Thus, it follows that the norm of the approximate solution will grow like $\exp(c\, h^{-1}\, t)$ as $h \to 0$, where c is a positive constant. Computations performed in 1979 reinforce this conclusion. The ordinary differential equation (2.3) was solved using the Crank-Nicholson method and the norm of the approximate solution computed at $t = 5$. This work is summarized in Table 1. Clearly, as $h \to 0$ the usual Galerkin method becomes unstable. Note that $||v(x, 0)|| = 0.8538 \times 10^{0}$ and $\operatorname{ess\,sup}_x |v(x, 0)| = 0.4000 \times 10^{1}$.

TABLE 1

NORMS OF APPROXIMATE SOLUTION AT $t = 5$
COMPUTED USING USUAL GALERKIN METHOD

$u_0(x) = [x(1 - x)]^2 \qquad v_0(x) = \cos x$

TIMESTEP = 0.010

MESH WIDTH	MAX NORM	L^2 NORM
1/12	0.1071×10^{16}	0.3367×10^{18}
1/24	0.6876×10^{26}	0.5448×10^{25}
1/48	0.1862×10^{34}	0.5991×10^{32}

Thus the usual Galerkin method is unstable and must be modified to treat realistic problems in a stable manner.

This chapter shall be organized as follows. First the notation used shall be defined. In the next section the general hyperbolic system shall be considered. Results concerning the partial differential

equation will be collected here. Next, in section 4, the stability and convergence of a method introduced in that section will be proven. Because of the work in section 3, this method will be fully adequate to treat the general, well-posed, initial-boundary value problem in one space dimension. In the final section concrete examples in one and two space dimensions shall be considered. Those examples shall illustrate the more concrete aspects of the algorithm.

2. Notation

For $\Omega \subset \mathbb{R}^M$ define the L^P-norm of functions $f: \Omega \to \mathbb{R}^n$ by

$$||f||_{L^P} = \left(\int_\Omega |f|^P \, dx\right)^{1/P} , \quad 1 \leq P < \infty ,$$

$$||f||_{L^\infty} = \operatorname*{ess\,sup}_{x \in \Omega} |f(x)| ,$$

where $|\cdot|$ represents the usual Euclidean norm on $\mathbb{R}^n$. The L^2-norm shall be abbreviated as $||f||$. Further, define the usual inner product, $(\cdot, \cdot)$, on $L^2 = L^2(\Omega)$ to be

$$(u, v) = \int_\Omega u^T(x) \, v(x) \, dx .$$

For $-\infty < s < \infty$, the $H^s = H^s(\Omega)$ norm, $||\cdot||_s$, is defined following [1]. In particular, $H^1 = H^1(\Omega)$ consists of those functions, $f: \Omega \to \mathbb{R}^n$, whose weak derivatives of order $|\alpha| \leq 1$ are in L^2 with norm

$$||f||_1 = \left(\sum_{|\alpha| \leq 1} ||\frac{\partial^{|\alpha|}}{\partial x^\alpha} f||^2\right)^{1/2} .$$

Define the spaces $L^P(0, T; H^S(\Omega))$ as the set of $f: [0, T] \to H^S(\Omega)$ with

$$||f||_{L^P(H^S)} = \left(\int_0^T ||f(t)||_S^P \, dt\right)^{1/P} < \infty \ , \ 1 \leq P < \infty \ ,$$

$$||f||_{L^\infty(H^S)} = \operatorname*{ess\,sup}_{0 \leq t \leq T} ||f(t)||_S < \infty \ .$$

Here the function $f(t, x)$ shall frequently be abbreviated by $f(t)$.

C shall denote a generic constant which may have different values in different places.

3. The General Hyperbolic System

In this section the general, initial-boundary value problem for a variable coefficient, symmetric hyperbolic system with constant coefficient, linear, homogeneous boundary conditions is considered. For $u = u(x, t) \in \mathbb{R}^n$, $0 \leq x \leq 1$, $0 \leq t \leq T$, an equation of the form

$$\frac{\partial u}{\partial t} = A(x, t) \frac{\partial u}{\partial x} + B(x, t) u + f(x, t) \tag{2.5}$$

is considered. If A is symmetric and $C^1([0, 1] \times [0, T])$ a change of dependent variables may be made to diagonalize A . Thus, A may be taken to be diagonal:

$$A = \begin{pmatrix} -A^{I} & 0 \\ 0 & A^{II} \end{pmatrix}, \quad A^{I} = \begin{pmatrix} a_1 & & \\ & a_2 & 0 \\ 0 & & \ddots \\ & & & a_r \end{pmatrix}, \quad A^{II} = \begin{pmatrix} a_{r+1} & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}. \tag{2.6}$$

To ensure well-posedness of the boundary value problem for (2.5), boundary conditions must be posed to make each frozen coefficient problem for (2.5) well-posed (see, for example, [32], [23]). In particular assume

$$a_i = a_i(x, t) \geq a_0 > 0 \ , \ 0 \leq x \leq 1 \ , \ 0 \leq t \leq T \ ,$$

and pose boundary conditions of the form

$$u^{I}(0, t) = S^{I} u^{II}(0, t) \ , \ u^{II}(1, t) = S^{II} u^{I}(1, t) \ , \tag{2.7}$$

where $(u^{I}, u^{II})^T$ is the partition of u corresponding to the partition of A in (2.6). S^{I}, S^{II} are given constant $r \times (n - r)$ and $(n - r) \times r$ matrices respectively. Further, the initial condition

$$u(x, 0) = u_0(x) \ , \ 0 \leq x \leq 1 \ , \tag{2.8}$$

will be imposed.

A simple result from linear algebra begins this section.

LEMMA 2.1. Let A be an $n \times n$ matrix, then A may be written $A = P - D$ where P and D are positive definite and D is diagonal.

PROOF: Write $A = P - D$, $P = A + D$, with $D = \mathrm{diag}(d_1, \ldots, d_n) > 0$. Consider now P:

$$x^T Px = x^T Ax + x^T Dx \geq (-\max|\lambda_i(A)| + \min d_i)|x|^2 .$$

Thus $P > 0$ for $d_i > \max|\lambda_i(A)|$.

Next consider the operator

$$L(t)\, u = A(x, t)\, \frac{\partial u}{\partial x}$$

with domain

$$\mathcal{D} = \{u = \begin{pmatrix} u^1 \\ u^{II} \end{pmatrix} | u \text{ is A.C. and satisfies (2.7)}\} .$$

The next proposition concerning the behavior of $L(t)$ on L^2 is fundamental to all following work.

PROPOSITION 2.1. There is an inner product $(\cdot, \cdot)'$ on L^2 with associated norm $||\cdot||'$ equivalent to $||\cdot||$ with respect to which the operator $L(t)$ is semibounded. That is, there is an $\alpha < \infty$ so that

$$(Lu, u)' \leq \alpha ||u||^2 , \quad u \in \mathcal{D} .$$

PROOF: For $G = G(x)$ a positive definite symmetric matrix consider the inner product

$$(u, v)' = (u, Gv) = (Gu, v) .$$

Then, provided G is chosen so that $AG = GA$,

$$(Lu, u)' = \int_0^1 (A \frac{\partial u}{\partial x})^T Gu \, dx =$$

$$= \frac{1}{2}\{u^T AGu \,|_{x=0}^{x=1} + \int_0^1 u^T \frac{\partial}{\partial x} (AG)u \, dx\} .$$

Thus, (2.9) will follow with

$$\alpha = \frac{1}{2} \max_{x,t} |\frac{\partial}{\partial x} (A(x, t) G(x))|$$

provided the boundary form is nonpositive:

$$u^T AGu \,|_{x=0}^{x=1} \leq 0 . \tag{2.10}$$

Applying the boundary conditions (2.7) gives

$$u^T A Gu^1|_0^1 = u^{(1)T} [-A^1(1) G^1(1) + S^{1T} A^{11}(1) G^{11}(1) S^1] u^1(1)$$

$$+ u^1(0)^T [S^{1T} A^1(0) G^1(0) S^1 - A^{11}(0) G^{11}(0)] u^{11}(0) ,$$

where $G = \begin{pmatrix} G^1 & 0 \\ 0 & G^{11} \end{pmatrix}$. Next it is shown that G may be chosen to make

$$S^{1T} A^{11}(1) G^{11}(1) S^1 \leq A^1(1) G^1(1) , \tag{2.11}$$

$$S^{1T} A^{1}(0)\, G^{1}(0)\, S^{1} \leq A^{11}(0)\, G^{11}(0) ,$$

from which (2.10) will follow. Pick G^{1}, G^{11} diagonal matrices with positive entires. Since A and G are now diagonal $AG = GA$. Define G at $x = 0$ by making the diagonal entries of $G^{11}(0)$ sufficiently large for (2.12) to hold. Likewise, by making the entries in $G^{1}(1)$ sufficiently large (2.11) may be found to hold. That this may be accomplished follows from Lemma 2.1. For $0 < x < 1$ define G by linear interpolation. G, thus constructed, will be positive definite, symmetric for $0 \leq x \leq 1$, and will ensure (2.10).

In case A is not a diagonal matrix, existence of the inner product $(\cdot, \cdot)'$ may be shown by diagonalizing A, constructing G for the diagonal system and obtaining $(\cdot, \cdot)'$ by transforming back to the original variables.

The existence of this inner product and associated G matrix will be critical to the analysis that follows, both for the estimates for the continuous equation and the approximation schemes. For example, using this G matrix one may deduce

PROPOSITION 2.2. There are constants C and α such that the true solution, $u(x, t)$, to (2.5), (2.7), (2.8) satisfies

$$||u(t)|| \leq Ce^{\alpha t}||u(0)|| + C \int_0^t e^{\alpha(t-s)} ||f(s)|| \, ds .$$

REMARK. This estimate occurs in [15], [41], [53] or [32], for example, and can be derived by the methods of Kreiss [32].

PROOF: Consider

$$\frac{d}{dt}\,||u||'^2 = (\frac{\partial u}{\partial t}\,,\, u)' + (u,\, \frac{\partial u}{\partial t})'$$

$$= (L(t),\, u)' + (u,\, L(t)\, u)' +$$

$$+ (B(t)u,\, u)' + (u,\, B(t)u)' +$$

$$+ (f(t),\, u)' + (u,\, f(t))' \,.$$

So that, provided B and f are bounded, there follows

$$\frac{d}{dt}\,||u||'^2 \le 2\alpha_1||u(t)||'^2 + 2\alpha_2||u(t)||'^2 +$$

$$+ 2\alpha_3||f(t)||'\;||u(t)||' \,.$$

Whence, with $\alpha = \alpha_1 + \alpha_2$,

$$\frac{d}{dt}\,(e^{-\alpha t}\;||u(t)||') \le \alpha_3||f(t)||' \,,$$

and

$$||u(t)||' \le e^{\alpha t}||u(0)||' + \alpha_3 \int_0^t e^{\alpha(t-s)}||f(s)||'\,ds \,.$$

The proposition now follows from the equivalence of the norm $||\cdot||$ with $||\cdot||'$.

The idea of picking this G matrix in such a way to make boundary forms negative definite occured in a different context in [65].

4. Galerkin Methods for Hyperbolic Systems

In this section a Galerkin method for the boundary value problem for hyperbolic systems is posed. This method sets the approximate problem in a space S^h with the inner product $(\cdot, \cdot)'$, constructed in Proposition 2.1, natural for the study of the continuous equation. The abstract problem considered is the system

$$\frac{\partial u}{\partial t} = L(t)\, u + f(x, t)\ ,\ u(x, t) \in \mathbb{R}^n\ ,\ x \in \Omega\ ,\ t \geq 0\ ,$$
$$u(x, 0) = u_0(x)\ , \tag{2.13}$$

with boundary conditions

$$u(x, t) \in N(x)\ ,\ x \in \partial\, \Omega\ , \tag{2.14}$$

where $N(x)$ is, for each x, a linear space chosen according to [23], [32] to make the problem well posed. Also $L(t)$ is a first order differential operator concerning which the following shall be assumed.

ASSUMPTION A1. There is an inner product $(\cdot, \cdot)'$ on $L^2(\Omega)$ with norm $||\cdot||'$ equivalent to $||\cdot||$ with respect to which the operator $L(t)$ is semibounded; i.e.,

$$(u, L(t)u)' \leq \alpha ||u||'^2 , \ u \in \mathcal{D}(L(t)) , \tag{2.15}$$

where

$$\mathcal{D}(L(t)) = \{u \in L^2 | L(t)u \in L^2 \text{ and } u \text{ satisfies (2.14)}\} .$$

ASSUMPTION A2. There is a constant $C > 0$ such that

$$(L(t)\ u, v)' \leq C||u||_1 \ ||v||' , \ u \in \mathcal{D}(L(t)), \ v \in L^2 .$$

The well-posedness of (2.13), (2.14) shall not be discussed here but the reader is referred to [23], [32], [37], and [61] for this question. An example of such a problem and inner product is the operator

$$L(t) = A(x, t) \frac{\partial u}{\partial x} + B(x, t)u$$

with boundary conditions (2.7) and inner product $(\cdot, \cdot)'$ constructed in Proposition 2.1. Also, assumption A1 yields an estimate upon u of the form

$$||u(t)||' \leq e^{\alpha t}||u_0||' + \int_0^t e^{\alpha(t-s)} \ ||f(s)||' \ ds , \tag{2.16}$$

similar to Proposition 2.2.

For the semidiscrete Galerkin method for (2.13), (2.14) a finite dimensional subspace S^h of $H^1(\Omega) \cap \mathcal{D}(L(t))$ is chosen which satisfies

the approximation property:

For some r and any s, $1 \le s \le r$,

$$\inf_{\chi \in S^h} \{||u - \chi|| + h||u - \chi||_1\} \le Ch^s ||u||_s$$

for all $u \in H^s(\Omega)$ satisfying (2.14). The semidiscrete approximation, U, to u is then sought as a differentiable map $U: [0, T] \to S^h$ satisfying

$$\begin{aligned} &U(0) \in S^h \text{, approximates } u_0 \text{ well ,} \\ &(\frac{\partial U}{\partial t}, v)' = (L(t)U, v)' + (f(t), v)' \text{ , } \forall v \in S^h \text{ .} \end{aligned} \tag{2.15}$$

First the stability of (2.15) shall be discussed.

PROPOSITION 2.3. Let $(\cdot, \cdot)'$ and α be as in assumption A1. Then, the approximate solution, $U(x, t)$, computed by (2.15), satisfies

$$||U(t)||' \le e^{\alpha t}||U(0)||' + \int_0^t e^{\alpha(t-s)}||f(s)||' \, ds \text{ .} \tag{2.16}$$

PROOF: In (2.15) set $v = U(t) \in S^h$, then

$$(\frac{\partial U}{\partial t}, U)' = (L(t)U, U)' + (f(t), U)'$$

$$\frac{1}{2}\frac{d}{dt}||U||'^2 \le \alpha||U||'^2 + ||f(t)||' \, ||U||' \text{ ,}$$

from which (2.16) follows.

THEOREM 2.1. (Convergence) Assume $u, u_t \in H^{s+1}(\Omega)$ for some s, $0 \le s \le r-1$; then there is a constant $C > 0$, independent of h, u_0 and t, such that

$$||u(t) - U(t)|| \le Ce^{\alpha t} ||u(0) - U(0)|| + Ch^s ||u(t)||_{s+1} + \tag{2.17}$$

$$+ C(e^{\alpha t} - 1) h^s \max_{0\le\tau\le t} \{||\frac{\partial u}{\partial t}(\tau)||_{s+1} + ||u(\tau)||_{s+1}\} .$$

PROOF: Let $w(t) = P' u(t)$, where P' is the orthogonal projection into S^h with respect to the inner product $(\cdot, \cdot)'$, and define $\eta(t) = u(t) - w(t)$, $\zeta(t) = U(t) - w(t)$. η then satisfies

$$||\frac{\partial^k}{\partial t^k} \eta(t)||_j \le Ch^{s-j} ||\frac{\partial^k}{\partial t^k} u(t)||_s , \quad j = 0, 1 . \tag{2.18}$$

Next ζ is estimated. For v in S^h, ζ satisfies

$$(\frac{\partial \zeta}{\partial t}, v)' - (L(t)\zeta, v)' = (\frac{\partial \eta}{\partial t}, v)' - (L(t)\eta, v)' .$$

Setting $v = \zeta$ and using assumptions A1 and A2 gives

$$\frac{d}{dt} ||\zeta||'^2 - \alpha||\zeta||'^2 \le ||\frac{\partial \eta}{\partial t}||' \, ||\zeta||' + C||\eta||_1 \, ||\zeta||' .$$

Multiplying through by $e^{-\alpha t}$ and integrating from 0 to t gives

$$||\zeta(t)||' \leq e^{\alpha t}||\zeta(0)|| + \int_0^t e^{\alpha(t-\tau)}\{||\frac{\partial \eta}{\partial t}(\tau)||' + C||\eta||_1\}d\tau \ .$$

The equivalence of $||\cdot||$ and $||\cdot||'$ and (2.18) then yield

$$||u(t) - U(t)|| \leq C\{||\zeta(t)||' + ||\eta(t)||\}$$

$$\leq Ch^{s+1}||u(t)||_{s+1} + e^{\alpha t}||\zeta(0)||' +$$

$$+ \frac{e^{\alpha t} - 1}{\alpha} \max_{0\leq\tau\leq t}\{Ch^{s+1}||\frac{\partial u}{\partial t}||_{s+1} + Ch^s||u||_{s+1}\} \ .$$

Finally $||\zeta(0)||' \leq ||u(0) - U(0)||' + ||\eta(0)||'$. Thus, the Theorem follows.

If a basis, $\{\phi_j\}_{j=1}^J$, is chosen for S^h, and $U(x, t)$ expanded as

$$U(x, t) = \sum_{j=1}^{J} C_j(t)\ \phi_j(x) \ ,$$

then (2.15) becomes the system of ordinary differential equations

$$\sum_{j=1}^{J}\{\frac{d}{dt}C_j(t)\ (\phi_j, \phi_\ell)' - C_j(t)(L(t)\ \phi_j, \phi_\ell)'\} - (f(t), \phi_\ell)' = 0 \ , \tag{2.19}$$

$$\ell = 1,\ldots, J \ .$$

In general, this system cannot be solved exactly and it is necessary to discretize the time variable. For this purpose the common Crank-

Nicholson method for the system (2.19) shall be examined. The case where the operator L is independent of T shall be considered. The analysis can be extended to the case $L = L(t)$.

Choose the timestep $k > 0$, $Nk = T$, and let $U^n(x)$ be the computed approximation to $u(x, nk)$. Specifically $\{U^n\}_{n=0}^N \subset S^h$ is sought, satisfying

$$U^0 \in S^h \text{ , approximates } u_0 \text{ well ,}$$

$$(\frac{U^{n+1} - U^n}{k} , v)' = (L\frac{U^{n+1} + U^n}{2} , v)' + (f^{n+1/2}, v)' . \tag{2.20}$$

In the above, $f^{n+1/2}$ could either be computed as $f((n + 1/2)k)$ or $(f((n + 1)k) + f(nk))/2$. All the following analysis would hold in either case.

The Crank-Nicholson Galerkin approximation shares the stabilty properties of the continuous equation and the semidiscrete approximation. Indeed,

PROPOSITION 2.4. Suppose the solution to (2.13), (2.14) is computed using the Crank-Nicholson Galerkin method (2.20). Then there are constants C , α_k , independent of h, u and t , with $\alpha_k = \alpha(1 - \frac{\alpha k}{2})^{-1} \to \alpha$ as $k \to 0$ such that

$$||U^n||' \leq e^{\alpha_k t} ||U^0||' + Ce^{\alpha_k t} \max_{0 \leq s \leq t} ||f(s)||'$$

holds for $n = 0, 1, \ldots, T/k$, where $t = nk > 0$ and α is the constant occuring in (2.16) in the continuous problem.

PROOF: In (2.20) set $v = U^{n+1} + U^n$; then assumption A1 yields

$$(||U^{n+1}||' - ||U^n||')(||U^{n+1}||' + ||U^n||') \leq$$

$$\leq \frac{k\alpha}{2}(||U^{n+1}||' + ||U^n||')$$

$$+ k||f^{n+1/2}||'(||U^{n+1}||' + ||U^n||') ,$$

or

$$(1 - \tfrac{\alpha k}{2})||U^{n+1}||' - (1 + \tfrac{\alpha k}{2})\ ||U^n||' \leq k||f^{n+1/2}||' \quad . \quad (2.22)$$

This is a difference inequality of the form

$$(1 - \delta)A_{n+1} - (1 + \delta)A_n \leq B_n \ , \ B_n \geq 0 \ , \quad (2.23)$$

which has solution

$$A_n \leq \gamma^n A_0 + C \sum_{i=0}^{n-1} \gamma^{n-1-i} B_i \ , \ \gamma = \frac{1 + \delta}{1 - \delta} \quad . \quad (2.24)$$

But now, $\gamma = 1 + \frac{2\delta}{1 - \delta} \leq 1 + c'\delta \leq e^{c'\delta}$ for $c' = \frac{2}{1 - \varepsilon}$ where $0 < \delta \leq \varepsilon < 1$. Thus, there follows

$$A_n \leq e^{c'n\delta} A_0 + c \sum_{1=0}^{n-1} e^{c'\delta(n-1-i)} B_i \ .$$

Using (2.23), (2.24) in (2.22) yields

$$||U^n||' \leq e^{c'\alpha\, nk/2}\, ||U^0||' + $$

$$+ C\, k \sum_{i=0}^{n-1} e^{c'\alpha\, k/2(n-1-i)}\, ||f^{n+1/2}||' \,.$$

Using $c' = 2(1 - \frac{\alpha k}{2})^{-1}$ and setting $nk = t > 0$ yields (2.21).

THEOREM 2.2. Assume $0 < \frac{\alpha k}{2} < 1$ and that there is an s, $1 \leq s \leq r$, such that

$$\frac{\partial^3 u}{\partial t^3} \in L^\infty(0, T;\ L^\infty(\Omega)) \,,$$

$$u \in L^\infty(0, T;\ H^s(\Omega)) \text{ , and}$$

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

Then there are positive constants C and α_k, where $\alpha_k = \alpha(1 - \frac{\alpha k}{2})^{-1} \to \alpha$ as $k \to 0$, such that

$$\max_{0\leq n\leq N} ||U^n - u(nk)|| \leq C\, e^{\alpha_k T}\, ||U^0 - u(0)|| +$$

$$+ C\, e^{\alpha_k T}\, h^s(||u||_{L^\infty(H^s)} + ||\frac{\partial u}{\partial t}||_{L^1(H^s)}) +$$

$$+ C\, e^{\alpha_k T}\, h^{s-1}\, ||u||_{L^\infty(H^s)} +$$

$$+ C\, e^{\alpha_k T}\, k^2 \max_{\substack{0\leq t\leq T \\ x\in\Omega}} |u_{ttt}| \,.$$

PROOF: Let $u^n(x)$ denote $u(x, nk)$. Then u^n satisfies

$$\left(\frac{u^{n+1} - u^n}{k}, v\right)' - \left(L\frac{u^{n+1} + u^n}{2}, v\right)' = (r^n, v)' + (f^{n+1/2}, v)',$$

where $r^n = \frac{u^{n+1} - u^n}{k} - \frac{\partial}{\partial t}\left(\frac{u^{n+1} + u^n}{2}\right)$. As before, define $w = P'u$, $w^n(x) = w(x, nk)$, $\eta^n = u^n - w^n$, $\zeta^n = U^n - w^n$. ζ^n satisfies the equation

$$\left(\frac{\zeta^{n+1} - \zeta^n}{k}, v\right)' - \left(L\left(\frac{\zeta^{n+1} + \zeta^n}{2}\right), v\right)' =$$

$$= \left(\frac{\eta^{n+1} - \eta^n}{k}, v\right)' - \left(L\left(\frac{\eta^{n+1} + \eta^n}{2}\right), v\right)' + (r^n, v)'.$$

Setting $v = \zeta^{n+1} + \zeta^n$ in the above and using assumptions A1 and A2 then yields

$$\left(||\zeta^{n+1}||' - ||\zeta^n||'\right)\left(||\zeta^{n+1}||' + ||\zeta^n||'\right) - \frac{\alpha k}{2}||\zeta^{n+1} + \zeta^n||'^2 \leq$$

$$\leq ||\eta^{n+1} - \eta^n||' \, ||\zeta^{n+1} + \zeta^n||' +$$

$$+ C\,k||\eta^{n+1} + \eta^n||_1 \, ||\zeta^{n+1} + \zeta^n||' +$$

$$+ k||r^n||' \, ||\zeta^{n+1} + \zeta^n||'.$$

The triangle inequality yields

$$\left(1 - \frac{\alpha k}{2}\right) ||\zeta^{n+1}||' - \left(1 + \frac{\alpha k}{2}\right)||\zeta^n||' \leq$$

$$\leq ||\eta^{n+1} - \eta^n||' + C\,k||\eta^{n+1} + \eta^n||_1 + k||r^n||' \quad .$$

This is a difference inequality of the form (2.23). Solving this in the manner that (2.23) was solved yields

$$||\zeta^n||' \leq e^{(nk)\alpha_k}\,||\zeta^0||' + \\ + C \sum_{i=0}^{n-1} e^{k\alpha_k(n-i-1)} \{||\eta^{i+1} - \eta^i||' + C\,k||\eta^{i+1} + \eta^i||_1\} + \\ + C\,k \sum_{i=0}^{n-1} e^{k\alpha_k(n-i-1)}\,||r^n||' \quad . \tag{2.26}$$

Each term in (2.26) will be considered separately. First, for $nk = t > 0$

$$e^{nk\,\alpha_k}\,||\zeta^0||' \leq e^{\alpha_k t} \{||u(0) - U(0)||' + ||\eta(0)||'\} \\ \leq e^{\alpha_k t} \{||u(0) - U(0)||' + C\,h^s||u||_s\} \quad . \tag{2.27}$$

Also,

$$\eta^{n+1} - \eta^n = \int_{nk}^{(n+1)k} \frac{d\eta}{dt}\,dt \quad ,$$

so that

$$||\eta^{n+1} - \eta^n||' \leq \int_{nk}^{(n+1)k} ||\frac{d\eta}{dt}||'\,dt \quad .$$

Hence,

$$\sum_{i=0}^{n-1} e^{k\alpha_k(n-i-1)} ||\eta^{i+1} - \eta^i||' \le e^{\alpha_k t} \int_0^t ||\frac{d\eta}{dt}||' \, dt$$

$$\le C \, e^{\alpha_k t} \, h^s \int_0^t ||\frac{\partial u}{\partial t}||_s \, dt \, . \tag{2.28}$$

Furthermore,

$$C \, k \sum_{i=0}^{n-1} e^{k\alpha_k(n-i-1)} ||\eta^{i+1} - \eta^i||_1 \le C \max_{0\le\tau\le t} ||\eta(\tau)||_1 \, k \sum_{i=0}^{n} e^{k\alpha_k(i-1)}$$

$$\le C \, e^{\alpha_k t} \max_{0\le\tau\le t} ||\eta(\tau)||_1 \tag{2.29}$$

$$\le C \, e^{\alpha_k t} \, h^s \max_{0\le\tau\le t} ||u(\tau)||_s \, .$$

Also, a Taylor series argument yields

$$|r^n| \le C \, k^2 \max|u_{ttt}|$$

and thus,

$$k \sum_{i=0}^{n-1} ||r^i||' \, e^{k\alpha_k(n-i-1)} \le C \, e^{nk\,\alpha_k} \, k^2 \max|u_{ttt}| \, . \tag{2.30}$$

(2.26) - (2.30) together give

$$\max_{0\le n\le N} ||\zeta^n||' \le e^{\alpha_k T} \, ||u(0) - U^0||' \, +$$

$$+ C\, e^{\alpha_k T} h^s \left(||u||_s + \int_0^T ||\tfrac{\partial u}{\partial t}||_s \, dt \right) +$$

$$+ C\, e^{\alpha_k T} h^{s-1} \max_{0 \le t \le T} ||u(s)||_s +$$

$$+ C\, e^{\alpha_k T} k^2 \max |u_{ttt}| \, .$$

Since $||u(nk) - U^n||' \le ||\eta^n||' + ||\zeta^n||'$ and $||\cdot||'$ is equivalent to $||\cdot||$, (2.25) follows.

4. Examples

In Proposition 2.1 the existence of a G matrix and associated inner product $(\cdot, \cdot)'$ is shown for the symmetric, hyperbolic system on $(0, 1)$. For the methods (2.15) and (2.20) to be useful this G matrix must be constructable in an easy way from the data of the problem. Here certain specific hyperbolic systems are considered and the construction of the relevant G matrix is discussed. Example 1 is the example discussed in the introduction, i.e. the example for which the usual Galerkin method was first found to be unstable. Example 2 illustrates the fact that each boundary must be treated in a different way, in general, and contains the key idea necessary to extend the work of Gunzburger [26] and Gunzburger and Plemmons [27] to general, well-posed systems in one dimension, and certain systems in two space dimensions. Example 3 is the half space problem. The relationship between the ideas herein and ideas in finite difference theory [28], [31], [33] are discussed heuristically. Example 4 is the shallow water equation with boundary conditions periodic in one variable. For any such problem, with boundary conditions periodic in one direction, the theory

is fully adequate. Example 5 is a problem in two space dimensions. The methods described here extend in a natural way to this and many other problems in two space dimensions, but it is not known if they can be extended to the general well-posed problem on the unit square. One reason for this is that the question of which boundary conditions are in fact well-posed is far from settled [36], [47], [48], [57].

Example 1. Here the example discussed in 1976 by Gunzburger [26] shall be considered:

$$\begin{cases} \dfrac{\partial u}{\partial t} = A \dfrac{\partial u}{\partial x}\ , \ u = \begin{pmatrix} u^{1} \\ u^{11} \end{pmatrix}\ , \ A = \begin{pmatrix} 1/2 & 1 \\ 1 & 1/2 \end{pmatrix}\ , \\ \\ u^{1}(0,\ t) = u^{1}(1,\ t) = 0\ , \\ \\ u(x,\ 0) = u_{0}(x)\ . \end{cases} \tag{2.27}$$

Since here A is not diagonal, the matrix G is constructed by first diagonalizing A, applying the analysis of proposition 2.1 then transforming back to the original variables. In practice, this means that when A is nondiagonal the matrix G must be taken to be non-diagonal. As pointed out before the introduction of the matrix

$$G = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

is essential to the treatment of this problem. Indeed, when u satisfies

the boundary conditions in (2.3i)

$$(A \frac{\partial u}{\partial x} , Gu) = 0 ,$$

so that $||\dot{u}||'$ is conserved:

$$||u(t)||' = ||u_0||' , t \geq 0 .$$

The space S^h is taken to be the space of cubic spline functions on a uniform mesh constructed in the introduction. That is,

$$S^h = \{\phi = \begin{pmatrix} \phi^1 \\ \phi 2 \end{pmatrix} | \phi^1 \in S_1^h , \phi^2 \in S_2^h\} ,$$

where S_1^h is the space spanned by the functions $\{\psi_0, \psi_1, \ldots, \psi_{N-1}, \psi_N\}$ and S_2^h is the space spanned by the functions $\{\phi_{-1}, \phi_0, \ldots, \phi_N, \phi_{N+1}\}$, all restricted to $[0, 1]$, as constructed in the introduction to chapter 2.

The fully-discrete approximation is then computed as a sequence $\{U^n\}_{n=0}^J$, where $0 < k < 1$, $Jk = T > 0$, and $U^n \approx u(t = nk)$ satisfies

$$(\frac{U^{n+1} - U^n}{k} , Gv) = (A \frac{\partial}{\partial x} (\frac{U^{n+1} + U^n}{2}) , Gv) \qquad (2.32)$$

for all $v \in S^h$ and $n = 0, 1, \ldots, J - 1$.

Expanding

$$U^n(x) = \binom{1}{0} \sum_{j=0}^{N} c_j^n \psi_j(x) + \binom{0}{1} \sum_{i=-1}^{N+1} d_i^n \phi_i(x)$$

and substituting into (2.32) yields the linear system for the unknowns c_j^{n+1}, d_i^{n+1} :

$$\begin{aligned} &\sum_{j=0}^{N} c_j^{n+1}\left[\left(\binom{1}{0} \psi_j, Gv\right) - \frac{k}{2}\left(A\binom{1}{0} \psi_j', Gv\right)\right] + \\ &+ \sum_{i=-1}^{N+1} d_i^{n+1}\left[\left(\binom{0}{1} \phi_i, Gv\right) - \frac{k}{2}\left(A\binom{0}{1} \phi_j', Gv\right)\right] = \\ &= \sum_{j=0}^{N} c_j^{n}\left[\left(\binom{1}{0} \psi_j, Gv\right) + \frac{k}{2}\left(A\binom{1}{0} \psi_j', Gv\right)\right] + \\ &+ \sum_{i=-1}^{N+1} d_i^{n}\left[\left(\binom{0}{1} \phi_i, Gv\right) + \frac{k}{2}\left(A\binom{0}{1} \phi_j', Gv\right)\right], \end{aligned} \tag{2.33}$$

for all $v \in S^h$ and $n = 0, 1, \ldots, J - 1$.

This system may be written in the form

$$(M - \frac{k}{2} K)\, X^{n+1} = (M + \frac{k}{2} K)\, X^n \,, \quad n = 0, 1, \ldots, J - 1 \,. \tag{2.34}$$

Computational experiments have been performed on (2.34). Computations indicate that $||U^n||'$ is indeed conserved by this method and that this method is stable:

$$||U^n||' = ||U^0||',$$

$$||U^n|| \leq C||U^0|| \ , \quad C = \text{constant} \ .$$

This is a sharp contrast with the work outlined in the introduction to Chapter II in which the usual method is seen to be grossly unstable. This work is summarized in Tables 2 and 3.

TABLE 2

NORMS OF APPROXIMATE SOLUTION AT $t = 5$
COMPUTED USING G-MATRIX METHOD

$u_0(x) = [x(1 - x)]^2 \qquad v_0(x) = \cos x$

TIMESTEP = 0.010

MESH WIDTH	MAX-NORM	$\|\|U\|\|$	$\|\|U\|\|'$
1/12	0.3270×10^1	0.7785×10^0	0.8992×10^0
1/24	0.3331×10^1	0.7786×10^0	0.8992×10^0
1/48	0.3363×10^1	0.7787×10^0	0.8992×10^0

TABLE 3

EVOLUTION OF NORMS OF APPROXIMATE SOLUTION
USING G-MATRIX METHOD

$u_0(x) = [x(1-x)]^2$		$v_0(x) = \cos x$	
TIMESTEP = 0.010		MESHWIDTH = 1/10	
TIME	MAX-NORM	L^2-NORM	WEIGHTED L^2-NORM
0	0.4000×10^{1}	0.8538×10^{0}	0.8992×10^{0}
1	0.3306×10^{1}	0.8161×10^{0}	0.8992×10^{0}
2	0.4041×10^{1}	0.8917×10^{0}	0.8992×10^{0}
3	0.3053×10^{1}	0.7880×10^{0}	0.8992×10^{0}
4	0.4099×10^{0}	0.8759×10^{0}	0.8992×10^{0}
5	0.3328×10^{1}	0.7789×10^{0}	0.8992×10^{0}

Example 2: A Boundary Value Problem. Consider the initial, boundary-value problem

$$\begin{cases} \frac{\partial u}{\partial t} = A \frac{\partial u}{\partial x}, \; u = \begin{pmatrix} u^{I} \\ u^{II} \end{pmatrix}, \; A = \begin{pmatrix} -1 & 0 \\ 0 & 10 \end{pmatrix}, \\ u^{I}(0, t) = 5\, u^{II}(0, t) \\ u^{II}(1, t) = u^{I}(1, t) \\ u(x, 0) = u_0(x) \end{cases} \tag{2.35}$$

Set $G = \begin{pmatrix} g_1(x) & 0 \\ 0 & g_2(x) \end{pmatrix}$ and, for u satisfying the boundary conditions in (2.35), consider the boundary form

$$(A \frac{\partial u}{\partial x}, Gu) = -\frac{1}{2} \int_0^1 u^T (A \frac{\partial}{\partial x} G)\, u\, dx +$$

$$+ u^{112} [-1\, g_1(1) + 10\, g_2(1)]$$

$$+ u^{12}(0)[+ 25\, g_1(0) - 10\, g_2(0)] .$$

In this case, the necessary conditions are

$$g_2(1) \leq \frac{1}{10} g_1(1) ,$$

$$g_1(0) \leq \frac{10}{25} g_2(0) ,$$

and thus, G cannot be taken to be a constant matrix as in example 1. However, taking $g_1(s) = 8x + 2$, $g_2(x) = -4x + 5$ is sufficient. With this choice

$$(A \frac{\partial u}{\partial x}, Gu) \leq 40 ||u||^2 ,$$

and the analysis of section 2 remains valid. In this case, the necessity of treating two boundary forms requires the use of variable G. This simple problem is an example of one for which the analysis of Gunzburger [26] and Gunzburger and Plemmons [27] does not apply.

If Example 1 is transformed into characteristic variables, then, as in Example 2, two boundary forms are present. Both forms may be simultaneously made to be negative semidefinite with G chosen to be a constant matrix. This is not the case here and G must be allowed to vary.

Example 3: The Half Space Problem. In this section considered is the half space problem for symmetric, hyperbolic systems of the form

$$\frac{\partial u}{\partial t} = A \frac{\partial u}{\partial x} - \sum_{j=2}^{d} B_j \frac{\partial u}{\partial y_j} + f(x, t)$$

for $t \geq 0$, $x \geq 0$, $-\infty < y_j < \infty$. A, B_j are assumed to be constant $m \times m$ symmetric matrices. Without loss of generality, A may be taken to be diagonal

$$A = \begin{pmatrix} -A^{I} & 0 \\ 0 & A^{II} \end{pmatrix}, \quad A^{I} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_r \end{pmatrix}, \quad A^{II} = \begin{pmatrix} \lambda_{r+1} & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{pmatrix}, \quad \lambda_i > 0 .$$

The boundary and initial conditions

$$u = \begin{pmatrix} u^{I} \\ u^{II} \end{pmatrix}, \quad u^{I}(0, t) = S\, u^{II}(0, t) \tag{2.37}$$

$$u(x, 0) = u_0(x) \tag{2.38}$$

are imposed, where S is a fixed $r \times (m - r)$ matrix. A sufficient condition for the existence of a unique classical solution to (2.36), (2.37), (2.38) is that for real η and real $\tau > 0$ there are no

$v = (v_1, \ldots, v_r)^T$ (apart from 0) eigenvectors corresponding to eigenvalues ξ with $\mathrm{Re}\,\xi < 0$ of

$$A^{-1}[\tau + i \sum_j \eta_j B_j]\tilde{v} = \xi \tilde{v}\,, \quad \tilde{v} = \begin{pmatrix} Sv \\ v \end{pmatrix},$$

(see, for example, Rauch [54] or Strang [61]). There is an extensive literature in finite difference theory concerned with stable difference approximations to the above problem, e.g. see [30], [33], [34], [35], [28], and [25].

For the Galerkin method a subspace $S^h \subset H^1$ is chosen satisfying the linear homogeneous boundary conditions (2.37). That is, $v \in S^h$ means

(i) $v: (0, \infty) \times \prod_{i=2}^{d} (-\infty, \infty) \to \mathbb{R}^m$,

(ii) $v \in H^1$

(iii) $v = \begin{pmatrix} v^1 \\ v^{11} \end{pmatrix}$ satisfies (2.37).

The continuous-in-time Galerkin approximation is then sought as a differentiable map $U: [0, T] \to S^h$ satisfying for all $v \in S^h$

$$\begin{cases} \left(\frac{\partial U}{\partial t}, Gv\right) = \left(A \frac{\partial U}{\partial x}, Gv\right) - \left(\sum_{j=2}^{d} B_j \frac{\partial U}{\partial y_j}, Gv\right) + (f, Gv)\,, \\ G = G^T > 0\,, \\ U(x, 0) \in S^h, \text{ an approximation to } u_0\,. \end{cases} \tag{2.39}$$

G is chosen so that (2.39) yields a stable and convergent method. To see the correct choice of G, consider, for $u \in S^h$ vanishing at ∞,

$$(A \frac{\partial u}{\partial x}, Gu) = \int_0^\infty \frac{\partial u}{\partial x} (GA)\, u\, dx$$

$$= \frac{1}{2} u^T GA\, u \Big|_0^\infty$$

$$= u^{II}(0)^T (S^T G^I A^I S - G^{II} A^{II})\, u^{II}(0) ,$$

where $G = \begin{pmatrix} G & 0 \\ 0 & G^{II} \end{pmatrix}$ is partitioned in the manner of A and u. Thus for the half space problem to be treated stably a constant matrix G is chosen so that G is positive definite, symmetric and the boundary form

$$S^T G^I A^I S - G^{II} A^{II} \leq 0$$

is negative semidefinite. That this may be done follows, for example, from Lemma 2.1. Thus, it is seen that a constant G matrix is sufficient to treat a problem with one boundary.

The role of G in the boundary value problem on $[0, 1]$ is now more easily understood. A different treatment of each boundary form is necessary to make each half plane problem stable. The two half plane problems are then patched together by interpolating G between the two boundaries, e.g. between $G(0)$ and $G(1)$, using a method that preserves other relevant properties of G. An analogous approach

is used to analyze the boundary value problem in finite difference theory, for example [33], [28], [30], or [31].

Example 4: The Shallow Water Equations. Consider the linearized shallow water equations on $[0, 1] \times [0, 1]$

$$\begin{cases} \dfrac{\partial w}{\partial t} = A \dfrac{\partial w}{\partial x} + B \dfrac{\partial w}{\partial y} + Cw \ , \ w = w(x, y, t), \\ \\ w(t = 0) = w_0 \ , \end{cases} \tag{2.40}$$

where

$$w = \begin{pmatrix} u \\ v \\ \phi \end{pmatrix} , \ A = \begin{pmatrix} U & 0 & C \\ 0 & U & 0 \\ C & 0 & U \end{pmatrix} , \ B = \begin{pmatrix} V & 0 & 0 \\ 0 & V & C \\ 0 & C & V \end{pmatrix} , \ C = \begin{pmatrix} 0 & f & 0 \\ -f & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} ,$$

for $U^2 + V^2 < C^2$ in $0 \leq x, y \leq 1$ and $t \geq 0$. In many problems in geophysics boundary conditions periodic in x are important:

$$\begin{pmatrix} u \\ v \\ \phi \end{pmatrix} (x = 0) = \begin{pmatrix} u \\ v \\ \phi \end{pmatrix} (x = 1) \ . \tag{2.42}$$

The effect of these boundary conditions is to reduce the problem to one of one space dimension. To pose boundary conditions along $y = 0$ and $y = 1$ the matrix B is diagonalized and boundary conditions are posed according to the results in [19] , [28], or [33] . Note that

$$T^*BT = \begin{pmatrix} V & 0 & 0 \\ 0 & \frac{V+C}{2} & 0 \\ 0 & 0 & \frac{V-C}{2} \end{pmatrix} , \quad T = \frac{1}{2}\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & 1 \end{pmatrix} .$$

Since $U^2 + V^2 < C^2$ boundary conditions may be imposed upon the first two components of $z = T^*w$ at $y = 1$ and on the third component of $z = T^*w$ at $y = 0$.

Thus, at $y = 0$ impose

$$\left(\frac{\phi - v}{2}\right)(y = 0) = S^1 \begin{pmatrix} u \\ \frac{\phi + v}{2} \end{pmatrix} (y = 0) , \tag{2.43}$$

and at $y = 1$

$$\begin{pmatrix} u \\ \frac{\phi + v}{2} \end{pmatrix} (y = 1) = S^{11}\left(\frac{\phi - v}{2}\right)(y = 0) , \tag{2.44}$$

where S^1, S^{11} are 1×2 and 2×1 matrices respectively. These are the general, linear, homogeneous, constant coefficient, local boundary conditions leading to a well posed problem for (2.40). For simplicity the boundary conditions are taken to be homogeneous.

As before, a subspace S^h of $H^1([0, 1] \times [0 \times 1])$ is constructed satisfying the periodic boundary conditions in x (2.42), and the reflection conditions in y (2.43), (2.44). The semidiscrete approximation is computed as a map $w^h: [0, T] \to S^h$ satisfying

$$\begin{cases} {}^{h}(0) \in S^h \quad \text{approximates} \quad w_0 \quad \text{well} \\ \left(\frac{\partial w^h}{\partial t}, G\chi\right) = \left(A\frac{\partial w^h}{\partial}, G\chi\right) + \left(B\frac{\partial w^h}{\partial y}, G\chi\right) + \left(Cw^h, G\chi\right) \quad (2.45) \\ \text{for all} \quad \chi = \begin{pmatrix} \chi_1 \\ \chi_2 \\ \chi_3 \end{pmatrix} \in S^h . \end{cases}$$

Here $G = G(y)$ is a positive definite, symmetric matrix function of y chosen to make

$$\left(A\frac{\partial w}{\partial x}, Gw\right) + \left(B\frac{\partial w}{\partial y}, Gw\right) + (Cw, Gw) \leq \alpha ||w||^2 \qquad (2.46)$$

for all $w \in S^h$ and some $\alpha < \infty$. Because G is a function of y alone, the boundary conditions are periodic in x, and the matrix C is skew symmetric the left hand side of (2.46) reduces to

$$\left(A\frac{\partial w}{\partial x}, Gw\right) + \left(B\frac{\partial w}{\partial y}, Gw\right) + (Cw, Gw) = \left(B\frac{\partial w}{\partial y}, Gw\right) .$$

It is in this sense that the boundary conditions, 1-periodic in x, reduce the choice of G to a one dimensional problem. G is then chosen as in the discussion in proposition 2.1 and the examples.

Example 5: A Two-Dimensional Problem. Problems in two space dimensions occur frequently in practice as do problems in domains with corners. For example, when computing the approximate solution to any exterior problem artificial boundaries must be imposed at some distance from the body. If those boundaries are to be noncharacteristic

everywhere then necessarily the boundaries must have corners. It is also known that many problems in domains with corners are ill-posed while each half plane problem is well-posed. For further information on the state of the theory of the differential equation see [36], [47], [48], or [57].

Here considered will be the differential equation

$$\frac{\partial u}{\partial t} = \begin{pmatrix} -1 & 0 \\ 0 & +1 \end{pmatrix} \frac{\partial u}{\partial x} + \begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix} \frac{\partial u}{\partial y} \quad , \quad u = \begin{pmatrix} u^{I} \\ u^{II} \end{pmatrix}$$

$$u = u(x, y, t) \quad , \quad 0 \leq x \leq 1 \quad , \quad 0 \leq t \leq T \quad .$$

Boundary conditions are imposed to make the problem in the x-variables well-posed for frozen y and the problem in the y-variables well-posed for frozen x, i.e.

$$\begin{cases} u^{I}(0, y, t) = S^{I} u^{II}(0, y, t) \\ \\ u^{II}(1, y, t) = S^{II} u^{I}(1, y, t) \end{cases} \tag{2.48}$$

$$\begin{cases} u^{II}(x, 0, t) = T^{I} u^{I}(x, 0, t) \\ \\ u^{I}(x, 1, t) = T^{II} u^{II}(x, 1, t) \quad , \end{cases} \tag{2.49}$$

together with the initial condition

$$u(x, y, 0) = u_0(x, y) \ . \tag{2.50}$$

The corners $(0, 0)$ and $(1, 1)$ are treated both as inflow and outflow points by the boundary conditions (2.48), (2.49). Thus, a compatibility condition is necessary at these corners for (2.47) - (2.50) to be well-posed.

PROPOSITION 2.5. The initial, boundary-value problem (2.47) - (2.50) is well posed provided

$$(S^{11} T^{11})^2 \leq 1 \ , \ (S^1 T^1)^2 \leq 1 \ . \tag{2.51}$$

PROOF: This result will follow from the analysis of the discrete problem to follow.

It will be shown that, provided the compatibility conditions (2.51) are satisfied at $(0, 0)$ and $(1, 1)$, there is a G matrix, $G = G^T > 0$, with respect to which

$$(A \frac{\partial u}{\partial x} + B \frac{\partial u}{\partial y} \ , Gu) \leq \alpha ||u||^2 \ , \ \alpha < \infty \ ,$$

for all u satisfying (2.48) and (2.49). Set

$$G = \begin{pmatrix} g_1(x, y) & 0 \\ 0 & g_2(x, y) \end{pmatrix} \ , \ g_i(x, y) > 0 \ ,$$

and consider the boundary form. From integration by parts there follows

$$(A \frac{\partial u}{\partial x} + B \frac{\partial u}{\partial y}, Gu) \leq \alpha ||u||^2 + B ,$$

where B is the boundary form:

$$B = \int_0^1 u^1(1, y)^2[-g_1(1, y) + g_2(1, y)(S^{11})^2] +$$

$$+ u^{11}(0, y)^2 [(S^1)^2 g_1(0, y) - g_2(0, y)] dy +$$

$$+ \int_0^1 u^{11}(x, 1)^2[(T^{11})^2 g_1(x, 1) - g_2(x, 1)] +$$

$$+ u^1(x, 0)^2 [-g_1(x, 0) + (T^1)^2 g_2(x, 0)]dx .$$

Thus, the four conditions

$$\begin{cases} -g_1(1, y) + g_2(1, y)(S^{11})^2 \leq 0 \\ (S^1)^2 g_1(0, y) - g_2(0, y) \leq 0 \\ (T^{11})^2 g_1(x, 1) - g_2(x, 1) \leq 0 \\ -g_1(x, 0) + (T^1)^2 g_2(x, 0) \leq 0 \end{cases} \tag{2.52}$$

are sufficient to ensure that (2.52) holds. These four conditions must be fulfilled with no singularities at the corners. That is, at each of $(0, 0)$, $(1, 1)$, $(0, 1)$, $(1, 0)$ two of the conditions (2.52) must simultaneously hold. Indeed, the compatibility condition (2.51) is sufficient for (2.52) to hold for some $g_1, g_2 > 0$ at all four corners.

For example, suppose that

$$S^{I} = 4 \ , \ S^{II} = \frac{1}{3} \ , \ T^{I} = \frac{1}{4} \ , \ T^{II} = 3 \ ,$$

so that $(S^{I}T^{I})^2 = 1$, $(S^{II}\,T^{II})^2 = 1$. In this case the functions

$$g_1(x, y) = -\frac{1}{2}\,xy + 9x - \frac{1}{2}\,y + 1$$

$$g_2(x, y) = -15\,\frac{15}{16}\,xy + 8\,\frac{15}{16}\,x + 7\,\frac{15}{16}\,y + \frac{15}{16}$$

fulfill (2.52) and the other restrictions upon G .

The existence of a G, continuously differentiable, positive definite and symmetric, making the boundary forms B negative semidefinite follows from the condition (2.51) by defining G at $(0, 0)$, $(0, 1)$, $(1, 0)$ and $(1, 1)$ to satisfy (2.52) and between the corners by interpolation, linear in x for fixed y and linear in y for fixed x .

CHAPTER III

ESTIMATES AWAY FROM A DISCONTINUITY

1. Introduction

In this chapter the question of approximating, by Galerkin methods, discontinuous solutions of the initial-value problem

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x}\,, \quad x \in \mathbb{R}\,, \ 0 \le t \le T\,, \tag{3.1}$$

$$u(x, 0) = v(x) \tag{3.2}$$

is considered. Here v is taken to be an $L_2(-\infty, \infty)$ function, smooth, except for a jump discontinuity at $x = 0$. This discontinuity will propagate along the characteristic $x + t = 0$ of the true solution $u(x, t) = v(x + t)$ of (3.1), (3.2).

In the usual, continuous-in-time Galerkin method for (3.1), (3.2) a subspace S_h of $H^1(-\infty, \infty)$ is chosen and the approximate solution $u^h(x, t)$ is computed as a differentiable map $u^h: [0, T] \to S_h$ by

$$(\frac{\partial u^h}{\partial t}\,, \phi) = (D\,u^h, \phi)\,, \text{ for } \phi \in S_h\,, \ D = \frac{\partial}{\partial x}\,, \tag{3.3}$$

$$u^h(0) \in S_h\,, \text{ an approximation to } v\,, \tag{3.4}$$

where $(\cdot, \cdot)$ represents the usual inner product on $L_2(-\infty, \infty)$.

The usual Galerkin method is not appropriate for the solution of (3.1), (3.2) with $v(x)$ as described for two reasons:

(i) The rate of convergence of the method (3.3), (3.4) depends upon the global smoothness of the true solution u, which might in fact be piecewise C^∞ but globally only L_2.

(ii) If ϕ is chosen to be $u^h(t)$ in (3.3) then u^h is seen to satisfy

$$||u^h(t)||_{L_2} = ||u^h(0)||_{L_2} .$$

Methods satisfying this condition are called strictly conservative methods. Experience with such methods indicates that the approximate solution will be hopelessly polluted by the discontinuity in u even far away from that discontinuity.

For these reasons two dissipative Galerkin methods are considered for the problem (3.1), (3.2). Estimates of the error outside a neighborhood of the discontinuity are derived for each method. These estimates imply, for example, that the error for (3.1), (3.2) has two parts, the first being the error of the method applied to smooth solutions of (3.1). The second component is the pollution effect of the discontinuity. This part is shown to decay exponentially in h^{-1} for regions a fixed distance d from the shock front and exponentially in d for a fixed h.

The plan of this chapter is as follows. In section 2 the function spaces needed and the notation used are defined and the two dissipative

Galerkin methods considered are formulated. In section 3 using methods developed by Thomée [67] and Thomée and Wendroff [69] a finite difference operator is associated with the Galerkin approximation. The properties of this operator are analyzed using the symbol of the operator in transform space. In section 4 convergence for smooth solutions is proven using results from finite difference and spline interpolation theory. In section 5 estimates outside a neighborhood of a discontinuity of u are found in the L_∞ and L_2 norms. A typical time discretization is finally considered in section 6 to indicate that the estimates in section 5 remain valid for fully discrete schemes.

2. Notation and spaces.

For $p = 2$ and ∞ consider the L_p spaces of equivalence classes of real-valued Lebesque measurable functions $f: (-\infty, \infty) \to (-\infty, \infty)$ with

$$||f||_{L_2}^2 = (f, f) = \int_{-\infty}^{\infty} |f|^2 \, dx < \infty ,$$

$$||f||_{L_\infty} = \operatorname{ess\,sup}_x |f(x)| < \infty .$$

For $0 \leq s < \infty$ define the Sobolev spaces $H^s = H^s(-\infty, \infty)$ as the set of those $u \in L_2(-\infty, \infty)$ with

$$\int_{-\infty}^{\infty} (1 + |y|^2)^s \, |\hat{u}(y)|^2 \, dy < \infty ,$$

where $\hat{u}$, $\check{u}$ shall denote the Fourier and inverse Fourier transforms

of u, respectively,

$$\hat{u}(y) = (Fu)(y) = \int_{-\infty}^{\infty} u(x)\, e^{-ixy}\, dx ,$$

$$\check{u}(x) = (F^{-1}u)(x) = (2\pi)^{-1} \int_{-\infty}^{\infty} u(y)\, e^{+ixy}\, dy .$$

For s a nonnegative integer H^s may be identified with the set of functions $u \in L_2$ having s generalized derivatives in L_2, with norm

$$||u||_{H^s} = \Big(\sum_{j=0}^{s} ||D^j u||_{L_2}^2 \Big)^{1/2} , \quad D = \frac{\partial}{\partial x} .$$

Spaces of doubly infinite sequences shall also be used. For $c = (c_j)_{j=-\infty}^{\infty}$ define the ℓ_2 norm

$$||c||_{\ell_2}^2 = \sum_{j=-\infty}^{\infty} |c_j|^2 .$$

For $h > 0$ and a function $v \in L_2(-\infty, \infty)$ write $v \in \ell_{2,h}$ provided

$$||v||_{\ell_{2,h}}^2 = h \sum_{j=-\infty}^{\infty} |v(jh)|^2 < \infty .$$

If v^2 is Riemann integrable note that, as $h \to 0$, $||v||_{\ell_{2,h}} \to ||v||_{L_2}$, but in general the finiteness of $||v||_{L_2}$ does not imply the finiteness of $||v||_{\ell_{2,h}}$.

Let $\mathcal{S}$ denote

$$\mathcal{S} = \{c = (c_j): c_j = O(|j|^{-q}) \text{ as } |j| \to \infty \text{ for all } q > 0\} .$$

Then for $c \in \mathcal{S}$ define the discrete Fourier transform of c to be

$$\tilde{c}(\theta) = (\mathcal{F} c)(\theta) = \sum_{j=-\infty}^{\infty} c_j e^{-ij\theta} .$$

Next define following e.g. Schoenberg [58] and Thomée [67] the subspaces in which our approximate problem is posed. Let χ be the characteristic function of $[-1/2, 1/2]$. As usual, define ϕ and ϕ_ℓ, the B-splines of order $\mu \geq 2$, as

$$\phi = \chi^{*\mu} , \quad \phi_\ell(x) = \phi(h^{-1} x - \ell) .$$

The approximating subspace S^h is taken to be the space of splines of at most power growth

$$S_h = \{\sum_\ell c_\ell \phi_\ell : c_\ell = O(|\ell|^q) \text{ as } |\ell| \to \infty \text{ for some } q\} .$$

The elements of S_h are $C^{\mu-2}$ functions which reduce to polynomials of degree $\mu - 1$ on each interval $[jh, (j + 1)h]$ for μ even and $[(j - 1/2)h , (j + 1/2)h]$ for μ odd .

Two nonstandard Galerkin methods will be considered for the problem (3.1), (3.2). Both methods are dissipative. In method 1 the introduction of the term $(-1)^{\gamma+1} h^\delta (D^\gamma u^h, D^\gamma \phi_\ell)$ is analogous to

the explicit introduction of an artificial viscosity term to the equations of inviscid fluid flow; whereas the dissipation present in method 2 is implicit in the variational formulation.

The continuous in time Galerkin approximation is computed as a differentiable map $u^h: [0, T] \to S_h$ satisfying

Method 1. For some $\delta > 0$ and integer γ, $0 \leq \gamma \leq \mu - 1$,

$$(\frac{\partial u^h}{\partial t}, \phi_\ell) - (Du^h, \phi_\ell) + (-1)^{\gamma+1} h^\delta (D^\gamma u^h, D^\gamma \phi_\ell) = 0, \quad \ell \in \mathbb{Z}, \tag{3.5}$$

$$u^h(0) = I_h v \in S_h. \tag{3.6}$$

Method 2.

$$(\frac{\partial u^h}{\partial t} - Du^h, \phi_\ell - h D \phi_\ell) = 0, \quad \ell \in \mathbb{Z}, \tag{3.7}$$

$$u^h(0) = I_h v \in S_h. \tag{3.8}$$

By I_h in (3.6) and (3.8) we denote the spline interpolation operator, that is, $I_h v$ is that element of S_h satisfying

$$(I_h v)(jh) = v(jh), \quad j \in \mathbb{Z}.$$

Method 2 is similar to a method of Dendy [17] and was first analyzed by Wahlbin [72], and [73] who proved convergence for smooth solutions for some linear and nonlinear hyperbolic equations.

3. Factoring the Galerkin Operator

In this section techniques developed by Thomée [67] and Thomée and Wendroff [69] are used to show that one may attach to the solution operator for either Galerkin method, $G_h(t)$, a finite difference operator, $F_k(t)$, so that $G_h(t) = I_h F_k(t)$, with I_h the spline interpolation operator. Results from finite difference theory and the theory of spline interpolation are then used to derive estimates on the error in either method. The rest of this paper depends heavily on the methods and results of Thomee in [67] and [68]. Therefore the notation of these papers shall be followed as closely as possible.

For $\mu \geq 2$ and integer σ , $0 \leq \sigma \leq 2\mu - 2$, define the trigonometric polynomials

$$g_{\mu,\sigma}(\theta) = h^{\sigma-1}(-i)^{\sigma-2\nu} \sum_{\ell=-\infty}^{\infty} (D^{\sigma-\nu} \phi_0, D^{\nu} \phi_\ell)\, e^{-i\ell\theta} \qquad (3.9)$$

with $\nu = [\sigma / 2]$. Note that the factor $h^{\sigma-1}$ makes $g_{\mu,\sigma}$ independent of h .

Consider now method 1. Expanding $u^h(x, t)$,

$$u^h(x, t) = \sum_{j=-\infty}^{\infty} c_j(t)\, \phi_j(x) ,$$

(3.5) is then the equivalent to the system of ordinary differential equations

$$\sum_{j=-\infty}^{\infty} \{\frac{d}{dt} c_j(t)(\phi_j, \phi_\ell) - c_j(t)(D\phi_j, \phi_\ell)$$

(3.10)

$$+ c_j(t)\ (-i)^{\gamma+1}\ h^{\delta}(D^{\gamma}\phi_j, D^{\gamma}\phi_\ell) = 0\ ,\ \ell \in \mathbb{Z}\ .$$

Let $\tilde{c}(\theta, t)$ be the discrete Fourier transform of $c(t) = (c_j(t))$; (3.10) then yields

$$h\, g_{\mu,0}(\theta)\, \frac{d}{dt}\, \tilde{c}(\theta, t) - i\, g_{\mu,1}(\theta)\, \tilde{c}(\theta, t)$$

$$+ h^{-2\gamma+\delta+1}\, g_{\mu,2\gamma}(\theta)\, \tilde{c}(\theta, t) = 0\ .$$

This ordinary differential equation has solution

$$\tilde{c}(\theta, t) = \exp\left(\frac{t}{h}\left[i\, \frac{g_{\mu,1}(\theta)}{g_{\mu,1}(\theta)} - h^{-2\gamma+\delta+1}\, \frac{g_{\mu,2\gamma}(\theta)}{g_{\mu,0}(\theta)}\right]\right) \tilde{c}(\theta, 0) \qquad (3.11)$$

Expanding the exponential in a Fourier series yields

$$\exp\left(\frac{t}{h}\left[i\, \frac{g_{\mu,1}(\theta)}{g_{\mu,0}(\theta)} - h^{-2\gamma+\delta+1}\, \frac{g_{\mu,2\gamma}(\theta)}{g_{\mu,0}(\theta)}\right]\right) = \sum_{\ell=-\infty}^{\infty} f_\ell(\frac{t}{h}, h) e^{-i\theta\ell}\ . \qquad (3.12)$$

If $u^h(0)$ is taken as indicated, i.e. as the interpolant of v at $h\mathbb{Z}$, then (3.11), (3.12) means that the Galerkin solution coincides at the meshpoints with a finite difference approximation to (3.1), (3.2)

$$u^h(jh, t) = \sum_{\ell=-\infty}^{\infty} f_\ell(\frac{t}{h}\ , h)\ v(jh - \ell h)\ ,\ j \in \mathbb{Z}\ .$$

Fix $\lambda < 0$ and consider for $k = \lambda h$ t of the form $t = nk$. The above discussion shows that there is a finite difference operator

$$(F_k v)(x) = \sum_{\ell=-\infty}^{\infty} f_\ell(\lambda, h)\, v(x - \ell h) \tag{3.13}$$

such that, with $t = nk$,

$$u^h = I_h (F_k)^n v . \tag{3.14}$$

Furthermore F_k has symbol

$$a(\theta) = \exp\left(\lambda \left(i \frac{g_{\mu,1}(\theta)}{g_{\mu,0}(\theta)} - h^{-2\gamma+\delta+1} \frac{g_{\mu,2\gamma}(\theta)}{g_{\mu,0}(\theta)} \right) \right) . \tag{3.15}$$

Before we analyze F_k we shall state a result concerning trigonometric polynomials that will be used later. The reader is referred to Thomée [67] for proofs and more details.

LEMMA 3.1. For integer σ, $0 \le \sigma \le 2\mu - 2$, there follows

$$\begin{aligned} g_{\mu,0}(\theta) &= \sum_{\ell=-\infty}^{\infty} (\theta + 2\pi\ell)^\sigma \hat{\chi}(\theta + 2\pi\ell)^{2\mu} \\ &= \theta^\sigma \hat{\chi}(\theta)^{2\mu} + R_{\mu,\sigma}(\theta) , \end{aligned} \tag{3.16}$$

where, as $|\theta| \to 0$,

$$R_{\mu,\sigma}(\theta) = \begin{cases} O(\theta^{2\mu}) & \text{, for } \sigma \text{ even,} \\ O(\theta^{2\mu+1}) & \text{, for } \sigma \text{ odd .} \end{cases} \tag{3.17}$$

For σ even, $g_{\mu,\sigma}(\theta) > 0$ for $0 < |\theta| \leq \pi$. ■

Recall, [12], that the finite difference operator F_k

$$(F_k\, v)(x) = \sum_{j=-\infty}^{\infty} f_j\, v(x - jh)$$

is dissipative if its symbol

$$a(\theta) = \sum_{j=-\infty}^{\infty} f_j\, e^{-ij\theta}$$

may be written as

$$a(\theta) = \exp(i\lambda\theta + \psi(\theta)) \tag{3.18}$$

where $\operatorname{Re}\psi(\theta) < 0$ for $0 < |\theta| \leq \pi$. Further, if F_k is a constant coefficient difference operator, F_k is accurate of order r provided $\psi(\theta) = O(\theta^{r+1})$ as $\theta \to 0$, and dissipative of order s provided

$$\operatorname{Re}\psi(\theta) \leq -\Gamma\, |\theta|^s\,, \quad |\theta| \leq \pi\,, \tag{3.19}$$

where Γ is some positive constant.

LEMMA 3.2. Consider the finite difference operator, F_k, attached to method 1 and given by (3.13). Its symbol, $a(\theta)$, may be written as

$$a(\theta) = \exp(i\lambda\theta + \psi(\theta, h)) \tag{3.20}$$

where as, $\theta \to 0$,

$$\psi(\theta, h) = O(\theta^{2\mu+1}) - h^{-2\gamma+\delta+1}\ \lambda\theta^{2\gamma} - h^{-2\gamma+\delta+1}\ O(\theta^{2\mu}) .$$

PROOF: (3.20) may be written with

$$\psi(\theta, h) = -i\ \lambda\theta + i\lambda\ \frac{g_{\mu,1}(\theta)}{g_{\mu,0}(\theta)} - \lambda h^{-2\gamma+\delta+1}\ \frac{g_{\mu,2\gamma}(\theta)}{g_{\mu,0}(\theta)} . \quad (3.21)$$

By Lemma 3.1 we have that as $\theta \to 0$ $i\lambda\ \frac{g_{\mu,1}(\theta)}{g_{\mu,0}(\theta)} = i\lambda\theta + O(\theta^{2\mu+1})$,

and $\lambda h^{-2\gamma+\delta+1}\ \frac{g_{\mu,2\gamma}(\theta)}{g_{\mu,0}(\theta)} = \lambda h^{-2\gamma+\delta+1}\ \theta^{2\gamma} + \lambda h^{-2\gamma+\delta+1}\ O(\theta^{2\mu})$. The lemma now follows from these relations.

LEMMA 3.3. F_k as given in (3.13) is a dissipative operator. In particular there is a positive constant Γ , independent of h and θ , so that

$$\text{Re}\ \psi(\theta, h) \leq -h^{-2\gamma+\delta+1}\ \Gamma\ |\theta|^{2\gamma} , \ |\theta| \leq \pi . \quad (3.22)$$

PROOF: (3.21) implies

$$\text{Re}\ \psi(\theta, h) = -\lambda\ h^{-2\gamma+\delta+1}\ \frac{g_{\mu,2\gamma}(\theta)}{g_{\mu,0}(\theta)} .$$

Since, by Lemma 3.1, $g_{\mu,2\gamma}$ and $g_{\mu,0}$ are positive for $0 < |\theta| \leq \pi$, it follows that

$$\text{Re}\ \psi(\theta, h) < 0 , \ 0 < |\theta| \leq \pi ,$$

and F_k is dissipative. (3.22) follows from Lemma 3.1 by virtue of the fact that

$$\text{Re } \psi(\theta, h) = -\lambda h^{-2\gamma+\delta+1} \frac{\theta^{2\gamma} \hat{\chi}(\theta)^{2\mu} + R_{\mu,2\gamma}(\theta)}{\hat{\chi}(\theta)^{2\mu} + R_{\mu,0}(\theta)}$$

$$= -\lambda h^{-2\gamma+\delta+1} (O(\theta^{2\gamma}) + O(\theta^{2\mu})) .$$

A corresponding analysis may be carried out for method 2. This was done in Wahlbin [73] in which the finite difference operator associated with method 2 is found. For method 2 F_k has symbol

$$a(\theta) = \exp\left(\lambda \frac{i\, g_{\mu,1}(\theta) - g_{\mu,2}(\theta)}{g_{\mu,0}(\theta) + i\, g_{\mu,1}(\theta)}\right) ,$$

$$= \exp(\lambda P(\theta)) .$$

Wahlbin shows that

$$\text{Re } P(\theta) \leq -c\, \theta^{2\mu} , \quad 0 \leq |\theta| \leq \pi ,$$

and

$$\text{Im } P(\theta) = \theta(1 + O(\theta^{2\mu})) \quad \text{as} \quad \theta \to 0 ,$$

and thus F_k is dissipative of order 2μ and accurate of order $2\mu - 1$.

4. Convergence for Smooth Solutions

The basic error estimates for smooth solutions of (3.1), (3.2) using method 2 are contained in the works [72] and [73]. These assert L_2-optimality of the error; for example

THEOREM 3.1. Suppose the approximate solution to the initial value problem (3.1), (3.2) is computed using method 2. Then there is a positive constant C, such that for $v \in H^{\mu}(-\infty, \infty)$ we have

$$\max_{0\le t\le T} ||u(t) - u^h(t)||_{L_2} \le C\, h^{\mu} \max_{0\le t\le T} ||u(t)||_{H^{\mu}} .$$

PROOF: See Wahlbin [72] or [73]. ∎

Consider now method 1. It will be shown that the error in $\ell_{2,h}$ is $O(h^{\delta})$ for $1/2 < \delta \le 2\mu$ and in L_2 is $O(h^{\delta})$ for $1/2 < \delta \le \mu$. Following Thomée [68] define

$$P_h(\theta) = h^{-1}\left(i\,\frac{g_{\mu,1}(h\theta)}{g_{\mu,0}(h\theta)} - h^{-2\gamma+\delta+1}\,\frac{g_{\mu,2\gamma}(h\theta)}{g_{\mu,0}(h\theta)}\right) ,$$

so that

$$u^h(\ell h, t) = F^{-1}(\exp(t\,P_h(\theta))\,\hat{v}(\theta))\,(\ell h) , \quad \ell \in \mathbb{Z} .$$

LEMMA 3.3. There is a positive constant C, such that for $|h\theta| \le \pi$ we have

$$|P_h(\theta) - i\theta| \leq C\, h^{\delta}|\theta|^{2\gamma} + C\, h^{2\mu}|\theta|^{2\mu+1} . \tag{3.23}$$

PROOF: By Lemma 3.1 we have that, as $\alpha \to 0$,

$$h(P_h(h^{-1}\alpha)) = i\alpha + O(\alpha^{2\mu+1}) - h^{-2\gamma+\delta+1}O(\alpha^{2\gamma}).$$

Hence, for $|\alpha| \leq \pi$,

$$|h\, P_h(h^{-1}\alpha) - i\alpha| \leq C(h^{-2\gamma+\delta+1}\,|\alpha|^{2\gamma} + |\alpha|^{2\mu+1}) .$$

(3.23) now follows by letting $\alpha = h\theta$ for $|h\theta| \leq \pi$. ∎

This shall be used to estimate the solution operators in the transform space.

LEMMA 3.4. For $|h\theta| \leq \pi$, there follows

$$\exp(t\, P_h(\theta)) - \exp(t\, i\theta) =$$

$$= t(P_h(\theta) - i\theta) \int_0^1 \exp(st\, P_h(\theta) + (1 - s)\, t\, i\theta)ds , \tag{3.25}$$

and thus

$$|\exp(t\, P_h(\theta)) - \exp(t\, i\theta)| \leq$$

$$\leq t|P_h(\theta) - i\theta| \int_0^1 \exp(\mathrm{Re}(s\, t\, P_h(\theta)))\, ds .$$

Integrating the above exactly gives

$$|\exp(t\, P_h(\theta)) - \exp(t\, i\theta)| \leq$$

$$\leq |P_h(\theta) - i\theta| \sum_{k=0}^{\infty} \frac{1}{(k+1)!} \left(-h^{-2\gamma+\delta} \frac{g_{\mu,2\gamma}(h\theta)}{g_{\mu,0}(h\theta)}\right)^k$$

$$\leq C\, |P_h(\theta) - i\theta| \,,$$

since $g_{\mu,\sigma}(\theta) \geq 0$ for σ even. (3.24) now follows from Lemma 3.4. ■

THEOREM 3.2. There is a positive constant C, independent of h, v, and t, such that for $1/2 < \delta \leq 2\mu$ and for $v \in H^k(-\infty, \infty)$, $k = \max\{2\gamma\,,\ \delta + 1\}$, we have

$$||u^h(t) - u(t)||_{\ell_{2,h}} \leq C\, h^{\delta} ||v||_{H^k} \,.$$

PROOF: For $s > 1/2$ we may estimate $||w||_{\ell_{2,h}}$ as follows (see Lemma 4.4 in [68])

$$||w||^2_{\ell_{2,h}} \leq C\{ \int_{-\pi/h}^{+\pi/h} |\hat{w}(\theta)|^2\, d\theta + h^{2s} \int_{-\infty}^{+\infty} |\theta|^{2s} |\hat{w}(\theta)|^s d\theta \} \,.$$

Setting $w = u^h - u$ yields

$$||u^h(t) - u(t)||^2_{\ell_{2,h}} \leq$$

$$\leq C\{ \int_{-\pi/h}^{+\pi/h} |\exp(t\, P_h(\theta)) - \exp(t\, i\theta)|^2 |\hat{v}(\theta)|^2\, d\theta +$$

$$+ h^{2s}\, ||v||^2_{H^s} \} \,.$$

Lemma 3.5 implies that for $|h\theta| \leq \pi$ and $\delta \leq 2\mu$

$$|\exp(t\, P_h(\theta)) - \exp(t\, i\theta)| \leq C(h^{\delta}|\theta|^{2\gamma} + h^{\delta}|\theta|^{\delta+1})$$

$$\leq C\, h^{\delta}|\theta|^{k},\ k = \max\{2\gamma,\ \delta + 1\}\ ,$$

from which the result follows. ∎

THEOREM 3.3. There is a $C > 0$, independent of h, v, and t, such that for $v \in H^k$, $k = \max\{2\gamma, \delta\}$, we have

$$||u^h(t) - u(t)||_{L^2} \leq C\, h^{\min\{\delta,\ \mu\}}||v||_{H^k}\ .$$

PROOF: This follows since spline interpolation is $O(h^{\mu})$. ∎

5. Discontinuous Initial Data

In this section the problem outlined in the introduction is considered. The initial data is taken to be an L_2 function vanishing for positive x with a jump discontinuity at zero and smooth on $(-\infty, 0)$. The case of a function smooth on $(-\infty, 0)$ and $(0, \infty)$ with a jump at zero may be reduced to a function vanishing for positive x by substracting a smooth function agreeing with the initial data on $(0, \infty)$. The analogous work on this question in finite difference theory was carried out in [60], [2], [13] and [29]. These estimates are shown here to hold in the finite element case as well. This depends upon some estimates on spline interpolation.

To be more specific, the true solution of (3.1), (3.2) for such a $v(x)$ vanishes for $x > -t$. Therefore $u^h(x, t)$ is estimated in

terms of the distance d of x parallel to the x-axis to the shock front $x + t = 0$. If χ_y denotes the characteristic function of (y, ∞), for method 1 and 2 estimated are

$$||\chi_{d-t}\, u^h(t)||_{L_2}, \text{ and } ||\chi_{d-t}\, u^h(t)||_{L_\infty}.$$

The basic result is that the effect of the discontinuity decays exponentially in d for fixed h and exponentially in h^{-1} for fixed d.

Recall that method 2 is accurate to order $2\mu - 2$ and dissipative of order 2μ. For method 1 choose $\delta = 2\gamma - 1$. This choice has two consequences. With this choice both dissipative Galerkin methods generate a constant coefficient difference operator. Both methods then satisfy the Kreiss condition for stability in L_2, [56], and the Brenner and Thomée [12] condition for stability in L_∞. Because of this a good portion of the analysis that follows is independent of the method under consideration. The two theorems proven in this section are:

THEOREM 3.4. Suppose the solution of (3.1), (3.2) is approximated by Method 1 or Method 2 with μ even. Suppose also $v(x) \in \mathcal{S}_{2,h} \cap L_2 \cap L_\infty$ and that v vanishes for positive x. Then there are positive constants C, c, and α (C and α depend upon μ) such that

$$||\chi_{d-t}\, u^h(t)||_{L_2} \le C||v||_{L_\infty}\, h^{1/2} \exp(-\alpha\, h^{-1}\, d) + \tag{3.26}$$

$$+ \exp(-c\, n \min\{d_t^K\,,\, d_t\}\, ||v||_{\ell_{2,h}}$$

with $t = nk > 0$, and $d_t = dt^{-1}$. Also

$$K = \begin{cases} \dfrac{2\mu}{2\mu - 1} & \text{, for method 2 ,} \\[2ex] \dfrac{2\gamma}{2\gamma - 1} & \text{, for method 1 .} \end{cases}$$

THEOREM 3.5. Suppose the hypotheses of Theorem 3.4 hold. Furthermore assume $|x|^{-a}\, v(x) \in L_\infty$ for some $a \ge 0$. Then

$$||\chi_{d-t}\, u^h(t)||_{L_\infty} \le C||v||_{L_\infty} \exp(-\alpha\, h^{-1}\, d) + \tag{5.3}$$

$$+ C\, h^{a/K} \exp(-cn \min\{d_t^K\,,\, d_t\})\; ||\, |x|^{-a}\, v||_{L_\infty}\,,$$

holds for $t \le T$, with d_t, C, c, α, and K as in Theorem 5.1.

PROOF of the Theorems: For $d > 0$ and $p = 2$ and ∞ there follows

$$||\chi_{d-t}\, I_h\, F_k(t)v||_{L_p(-\infty,\, \infty)} \le$$

$$\le ||\chi_{\frac{d}{2} - t}\, (I_h\, F_k(t)\, v) - I_h(\chi_{\frac{d}{2} - t}\, F_k(t)v)||_{L_p(d-t,\, \infty)} \tag{3.29}$$

$$+ ||I_h(\chi_{d-t}\, F_k(t)v)||_{L_p(d-t,\, \infty)}\,.$$

Since spline interpolation is a bounded linear operator between L_2 and ℓ_2 and between L_∞ and L_∞ [55], [59] we have

$$||I_h(\chi_{d-t}\, F_k(t)v)||_{L_2} \le C||\chi_{d-t}\, F_k(t)v||_{\ell_{2,h}} ,$$

$$||I_h(\chi_{d-t}\, F_k(t)v||_{L_\infty} \le C||\chi_{d-t}\, F_k(t)v||_{L_\infty} . \tag{3.30}$$

Finite difference theory shall be used to estimate the above two quantities. The remaining term in (3.29) shall be estimated using the theory of spline interpolation.

In [13] estimates on $||\chi_{d-t}\, F_k(t)v||_{L_p}$ are derived for $p = 2$ and ∞ using inequalities for the symbol $a(\theta)$ of F_k in a neighborhood of the real axis. If $F_k(t)$ is considered as a Toeplitz matrix acting between $\ell_{2,h}$ and $\ell_{2,h}$ instead of L_2 and L_2 those estimates on the symbol of F_k may be used to derive ℓ_2 estimates on $\chi_{d-t}\, F_k(t)v$. Specifically using the notation of Theorems 3.4 and 3.5,

$$||\chi_{d-t}\, F_k(t)\, v||_{\ell_{2,h}} \le \exp(-cn \min\{d_t^K , d_t\})||v||_{\ell_{2,h}} . \tag{3.31}$$

The corresponding L_∞ estimate in [13] reads

$$||\chi_{d-t}\, F_k(t)v||_{L_\infty} \le Ch^{a/K} \exp(-cn \min\{d_t^K, d_t\})||\, |x|^{-a}v||_{L_\infty} \tag{3.32}$$

To estimate the first part of (3.29) consider the cardinal spline interpolation problem:

For μ even $L_h(x) \in S_h$ is sought satisfying

$$L_h(jh) = \begin{cases} 1 \ , \ \text{if} \ \ j = 0 \ , \\ \\ 0 \ , \ \text{if} \ \ j \neq 0 \ , \end{cases} \tag{3.33}$$

$$L_h(x) \in L_\infty(-\infty, \infty) \ . \tag{3.34}$$

The following basic result is known about this problem [58], [59].

THEOREM 3.6. For μ even there is a unique L_h in S_h satisfying (3.33), (3.34). Furthermore, there are positive constants A and α depending upon μ such that

$$|L_h(x)| < A \exp(-\alpha \, h^{-1}|x|) \ , \quad x \in \mathbb{R} \ .$$

Also, for $c = (c_j) \in \ell_\infty$, the unique function $S_h(x)$ in S_h satisfying

$$S_h(jh) = c_j \ , \ S_h \in L_\infty(-\infty, \infty)$$

is given by

$$S_h(x) = \sum_{j=-\infty}^{\infty} c_j \, L_h(x - jh) \ ,$$

the series converging uniformly and absolutely on each compact subset of $\mathbb{R}^1$. ∎

Given a function, $f(x) \in L_\infty$, expand $I_h f$ and $I_h(\chi_{\frac{d}{2}-t} f)$ using Theorem 3.6 :

$$(I_h f)(x) = \sum_{j=-\infty}^{\infty} f(jh)\, L_h(x - jh)$$

$$I_h(\chi_{\frac{d}{2}-t} f)(x) = \sum_{j=J}^{\infty} f(jh)\, L_h(x - jh)$$

where $J = \min\{j \in \mathbb{Z} \;\; jh > \frac{d}{2} - t\}$. Thus

$$|(I_h f)(x) - I_h(\chi_{\frac{d}{2}-t} f)(x)| = |\sum_{j=-\infty}^{J-1} f(jh)\, L_h(x - jh)|$$

$$\leq A\, ||f||_{L_\infty} \sum_{j=-\infty}^{J-1} \exp(-\alpha|xh^{-1} - j|) \quad .$$

And now for $x > d - t$ there follows

$$|(I_h f)(x) - I_h(\chi_{\frac{d}{2}-t} f)(x)| \leq$$

$$\leq C||f||_{L_\infty} \exp(-\alpha\, x\, h^{-1}) \sum_{j=-\infty}^{J-1} \exp(\alpha\, j) \; .$$

Summing the geometric series and using the definition of J gives that for $x > d - t$

$$|(I_h f)(x) - I_h(\chi_{\frac{d}{2}-t} f)(x)| \leq \tag{3.35}$$

$$\leq C||f||_{L_\infty} \exp(-\alpha\, h^{-1}(x - (\tfrac{d}{2} - t))) \; .$$

Setting $f = F_k(t)v$, integrating the square of the above inequality between $d - t$ and ∞ yields

$$\|\chi_{\frac{d}{2} - t} I_h(F_k(t)v) - I_h(\chi_{\frac{d}{2} - t} F_k(t)v)\|^2_{L_2(d-t, \infty)} \leq$$

$$\leq C\|F_k(t)v\|^2_{L_\infty} \int_{d-t}^{\infty} \exp(-2\alpha h^{-1}(x - (\tfrac{d}{2} - t))dx .$$

Since F_k is stable in L_∞

$$\|F_k(t)v\|_{L_\infty} \leq C\|v\|_{L_\infty} , \tag{3.36}$$

this yields that

$$\|\chi_{\frac{d}{2} - t} I_h(F_k(t)v) - I_h(\chi_{\frac{d}{2} - t} F_k(t)v)\|_{L^2(d-t, \infty)} \leq \tag{3.37}$$

$$\leq C h^{1/2} \|v\|_{L_\infty} \exp(-\tfrac{\alpha}{2} h^{-1} d) .$$

Theorem 3.4 now follows from (3.29), (3.30), (3.31), and (3.37).

Theorem 3.5 follows from (3.29), (3.30), (3.32), (3.36) and the pointwise estimate (3.35). ∎

6. Fully Discrete Estimates

Next turn to fully discrete schemes. The aim of this section is not to be exhaustive, but to show that the estimates of section 5 hold true under appropriate time discretizations. Therefore the usual Crank-Nicholson time discretization for method 1 is considered as an example.

Recall that method 1 yields an infinite system of ordinary differential equations

$$\sum_{j=-\infty}^{\infty} \{\frac{d}{dt} c_j(t)(\phi_j, \phi_\ell) - c_j(t)(D\,\phi_j, \phi_\ell) \tag{3.38}$$

$$+ h^{\delta}(-1)^{\gamma+1}\,(D^{\gamma}\,\phi_j, D^{\gamma}\,\phi_\ell)\} = 0\ ,\ \ell \in \mathbb{Z}\ .$$

Define the vector $c(t) = (c_j(t))$ and the Toeplitz matrices

$$M_{ij} = h^{-1}(\phi_i, \phi_j)\ ,\ S_{ij} = (D\,\phi_i, \phi_j)\ ,\ \text{and}\ \mathcal{D}_{ij} = h^{2\gamma-1}(D^{\gamma}\phi_i, D^{\gamma}\phi_j)$$

for $i, j \in \mathbb{Z}$. The system (3.38) may then be written as

$$h\,M\,\frac{d}{dt}\,c(t) - S\,c(t) + (-1)^{\gamma+1}\,h^{-2\gamma+\delta+1}\,\mathcal{D}\,c(t) = 0\ , \tag{6.2}$$

and thus has solution

$$c(t) = \exp(-\frac{t}{h}\,M^{-1}((-1)^{\gamma+1}\,h^{-2\gamma+\delta+1}\,\mathcal{D} - S))\,c(0)\ .$$

The Crank-Nicholson method consists in replacing the exponential above by a rational approximation to the exponential, $r(z)$, given by

$$r(z) = \frac{p(z)}{q(z)}\ ,\ p(z) = 1 - \frac{1}{2}\,z\ ,\ q(z) = 1 + \frac{1}{2}\,z\ ,$$

so that, for $\mathrm{Re}\ z \leq 0$,

$$r(z) = \exp(-z) + O(z^3) \ , \text{ as } \ z \to 0 \ .$$

The approximate solution is then given by

$$c(t) \approx r(-\frac{t}{h} M^{-1}((-1)^{\gamma+1} h^{-2\gamma+\delta+1} \mathcal{D} - S))\ c(0) \ . \qquad (3.40)$$

Let c^n be the approximation to $c(t)$ for $t = nk \geq 0$. (3.40) may then be written

$$[M - \frac{1}{2}\frac{k}{h}((-1)^{\gamma+1} h^{-2\gamma+\delta+1} \mathcal{D} - S)]\ c^{n+1}$$

$$= [M + \frac{1}{2}\frac{k}{h}((-1)^{\gamma+1} h^{-2\gamma+\delta+1} \mathcal{D} - S)]\ c^n \ .$$

Define the bilinear forms

$$P_k(u, v) = (u, v) - \frac{1}{2} k[(-1)^{\gamma+1} h^{\delta}(D^{\gamma}u, D^{\gamma}v) - (Du, v)] \ ,$$

$$Q_k(u, v) = (u, v) + \frac{1}{2} k[(-1)^{\gamma+1} h^{\delta}(D u, D v) - (Du, v)] \ .$$

The discrete in time method may alternately be posed as follows. $u_n^h \in S_h$ is sought approximating $u(\cdot, nk)$ for $k = \lambda h$, and determined recursively by

$$u_0^h = I_h\, v$$

$$Q_k(u_{n+1}^h\,,\ \phi_\ell) = P_k(u_n^h\,,\ \phi_\ell)\ ,\ \ell \in \mathbb{Z}\ ,\ n = 1,\ldots,\ T/k\ . \tag{3.41}$$

Setting, as before,

$$u_n^h(x) = \sum_{j=-\infty}^{\infty} c_j^n\, \phi_j(x)\ ,\quad c^n = (c_j^n)\ ,$$

(3.41) becomes

$$\sum_{j=-\infty}^{\infty} c_j^{n+1}\, Q_k(\phi_j,\ \phi_\ell) = \sum_{j=-}^{\infty} c_j^n\, P_k(\phi_j,\ \phi_\ell)\ ,\ \ell \in \mathbb{Z}\ .$$

Put $\tilde{c}^n(\theta) = (\mathcal{J}\, c^n)(\theta)$; then the transform of the above equation is

$$\tilde{c}^{n+1}(\theta)\, q_k(\theta) = \tilde{c}^n(\theta)\, p_k(\theta)\ ,$$

where p_k, q_k are the symbols of P_k and Q_k respectively

$$p_k(\theta) = \sum_{\ell=-\infty}^{\infty} P_k(\phi_0,\ \phi_\ell)\ e^{-i\ell\theta}\ ,$$

$$q_k(\theta) = \sum_{\ell=-\infty}^{\infty} Q_k(\phi_0,\ \phi_\ell)\ e^{-i\ell\theta}\ .$$

Expanding

$$\rho(\theta) = \frac{P_k(\theta)}{q_k(\theta)} = \sum_{j=-\infty}^{\infty} f_j \, e^{-ij\theta} \quad ,$$

and defining the finite difference operator,

$$(F_k \, v)(x) = \sum_{j=-\infty}^{\infty} f_j \, v(x - jh), \tag{3.42}$$

there follows

LEMMA 3.6. Fix $\lambda > 0$ and, for $k = \lambda h$, consider t of the form $t = nk$. For such t, k the dissipative Crank-Nicholson Galerkin solution, determined by equation (3.41), may be written as

$$u_n^h(x) = (I_h \, F_k^n)(x) \, , \tag{3.43}$$

where F_k is the finite difference operator given by (3.42) with symbol

$$\rho(\theta) = \frac{P_k(\theta)}{q_k(\theta)} = \frac{g_{\mu,0}(\theta) + \frac{1}{2}\lambda \, (h^{-2\gamma+\delta+1} \, g_{\mu,2\gamma}(\theta) + i \, g_{\mu,1}(\theta))}{g_{\mu,0}(\theta) - \frac{1}{2}\lambda \, (h^{-2\gamma+\delta+1} \, g_{\mu,2\gamma}(\theta) + i \, g_{\mu,1}(\theta))} \tag{3.44}$$

PROOF: The representation (3.43) follows as did (3.14) in section 3. (3.44) follows from the definition of p_k, q_k, and $g_{\mu,\sigma}$. ■

LEMMA 3.7. The finite difference operator F_k, described in Lemma 3.6, with $\delta = 2\gamma - 1$, is a constant coefficient finite difference operator accurate of order $r = \min\{2, \delta\}$.

PROOF: By Lemma 3.1, as $|\theta| \to 0$,

$$\frac{g_{\mu,1}(\theta)}{g_{\mu,0}(\theta)} = i\theta + O(\theta^{2\mu+1}) \ , \text{ and}$$

$$\frac{g_{\mu,2\gamma}(\theta)}{g_{\mu,0}(\theta)} = \theta^{2\gamma} + O(\theta^{2\mu}) \ .$$

Thus, as $|\theta| \to 0$,

$$\rho(\theta) = \frac{(1 + \frac{1}{2}\lambda\, i\theta) + \frac{1}{2}\lambda\, \theta^{2\gamma} + O(\theta^{2\mu})}{(1 - \frac{1}{2}\lambda\, i\theta) - \frac{1}{2}\lambda\, \theta^{2\gamma} + O(\theta^{2\mu})}$$

$$= \frac{P(\lambda i\theta)}{q(\lambda i\theta)} + O(\theta^{2\gamma}) + O(\theta^{2\mu})$$

$$= \exp(\lambda i\theta) + O(\theta^{r+1})$$

$$= \exp(\lambda i\theta + O(\theta^{r+1})) \ ,$$

since $2\gamma = \delta + 1$ and $r = \min\{2, \delta\}$. ■

LEMMA 3.8. F_k , as defined in (3.42), is a dissipative operator. If $\delta = 2\gamma - 1$ then there is a positive constant Γ , independent of h and θ , such that

$$|\rho(\theta)| \leq \exp(-\Gamma\, |\theta|^{2\gamma}) \ , \quad 0 \leq |\theta| \leq \pi \ ,$$

and thus F_k is dissipative of order 2γ .

PROOF: Consider $\rho(\theta)$

$$\rho(\theta) = \frac{1 - \frac{1}{2}\lambda(h^{-2\gamma+\delta+1}\frac{g_{\mu,2\gamma}(\theta)}{g_{\mu,0}(\theta)} + i\frac{g_{\mu,1}(\theta)}{g_{\mu,0}(\theta)})}{1 + \frac{1}{2}\lambda(h^{-2\gamma+\delta+1}\frac{g_{\mu,2\gamma}(\theta)}{g_{\mu,0}(\theta)} + i\frac{g_{\mu,1}(\theta)}{g_{\mu,0}(\theta)})} .$$

To show $|\rho(\theta)| < 1$ for $0 < |\theta| \leq \pi$ it is sufficient to prove

$$\mathrm{Re}\left\{h^{-2\gamma+\delta+1}\frac{g_{\mu,2\gamma}(\theta)}{g_{\mu,0}(\theta)} + i\frac{g_{\mu,1}(\theta)}{g_{\mu,0}(\theta)}\right\} > 0 \text{ , for } 0 < |\theta| \leq \pi .$$

But this follows from Lemma 3.1. Thus F_k is dissipative.

When $\delta = 2\gamma - 1$ set

$$z = z(\theta) = \lambda(g_{\mu,2\gamma}(\theta) + i\, g_{\mu,1}(\theta))/\, g_{\mu,0}(\theta) \quad ,$$

so that $\rho(\theta) = r(z)$. As before, $\mathrm{Re}(z(\theta)) = O(\theta^{2\gamma})$ as $\theta \to 0$. Since F_k is dissipative ($|\rho(\theta)| < 1$, $\theta \neq 0$) it is sufficient to prove (3.45) in a neighborhood of $\theta = 0$ (hence a neighborhood at $z = 0$) . This will be accomplished when the following claim is shown.

Claim: There is a constant Γ so that, for $|z|$ sufficiently small and $\mathrm{Re}\ z \geq 0$, we have:

$$|r(z)| \leq \exp(-\Gamma\ \mathrm{Re}\ z) . \tag{3.46}$$

To see (3.46) let $z = x + iy$ and consider

$$|r(z)|^2 = r(z)\,\overline{r(z)} = 1 - \frac{8x}{4 + 4x + |z|^2} \quad .$$

But for $|z| < 1$ and $\mathrm{Re}\, z \geq 0$ $\quad 4 + 4x + |z|^2 < 9$. Thus,

$$|r(z)|^2 \leq 1 - \frac{8}{9} x \leq \exp(-\frac{8}{9} x) \quad .$$

Hence, $|r(z)| \leq \exp(-\frac{4}{9} x)$, which is (3.46) . ∎

For the rest of this section set $\gamma = 2\delta - 1$. The discrete-in-time analog to Theorem 3.3 is then

THEOREM 3.7. With $\delta = 2\gamma - 1$, there is a positive constant C such that for $\frac{1}{2} < \delta \leq 2$ and for $t = nk$, $k = \lambda h$,

$$||u_n^h - u(t)||_{L_2} \leq C\, h^{\delta} ||v||_{H^{\delta+1}} \quad . \qquad \blacksquare$$

An argument similar to the one used to establish Theorems 3.4 and 3.5 then yields

THEOREM 3.8. Suppose the solution of (3.1), (3.2) is approximated by the dissipative Crank-Nicholson Galerkin method, as described by equation (3.41), with μ even and $\delta = 2\gamma - 1$. Suppose, also, $v(x) \in L_2 \cap \ell_{2,h} \cap L_\infty$ and that v vanishes for positive x . Then there are positive constants C, c, and α , depending upon μ , such that

$$||\chi_{d-t}\, u_n^h||_{L_2} \leq C||v||_{L_\infty} h^{1/2} \exp(-\alpha\, h^{-1}\, d) +$$

$$+ \exp(-c\, n \min\{d_t^K , d_t\})\ ||v||_{\ell_{2,h}}$$

with $t = nk > 0$ and $d_t = dt^{-1}$. Also $K = 2\gamma / \min\{2, \delta\}$. ■

THEOREM 3.9. In addition to the hypotheses of Theorem 3.8, assume also $|x|^{-a} v(x) \in L_\infty$ for some $a \geq 0$. Then

$$||\chi_{d-t}\, u_n^h||_{L_\infty} \leq C||v||_{L_\infty} \exp(-\alpha\, h^{-1}\, d) +$$

$$+ C\, h^{a/K} \exp(-cn \min\{d_t^K\, ,\, d_t\})||\, |x|^{-a}\, v||_{L_\infty}$$

holds for $t \leq T$, with d_t , C, c, α, and K as in Theorem 3.8. ■

CHAPTER IV

ERROR ESTIMATES FOR A VARIABLE COEFFICIENT INITIAL-BOUNDARY VALUE PROBLEM

1. Introduction

In many physical situations there arise nonlinear hyperbolic equations of the form

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(A(u, x, t)u) = f(x, t) . \tag{4.1}$$

The problem of approximating the solutions to such equations is thus an important question. Little is known about Galerkin methods for such equations. In [72] Wahlbin considers a dissipative method for approximating periodic solutions of equations of the type (4.1). Further Dupont [21], [20] has considered the usual Galerkin method for some specific systems of equations of the type (4.1) with boundary conditions. He derives suboptimal error estimates in L^2 for approximations in piecewise polynomial spaces. Also Luskin [45] has studied both standard and nonstandard Galerkin methods for systems of the form (4.1) in piecewise polynomial spaces. In this section a method proposed by G. Baker [9] for the boundary value problem for (4.1) when $A \equiv 1$ is studied in the case $A = A(x, t)$ thus extending his analysis to the case of time dependent coefficients. This scheme resembles the H^{-1} methods proposed by Rachford and Wheeler [52] for two point boundary value problems and analyzed for parabolic equations by Diaz [18]. For extensions of this work to collocation methods see for example Archer [3], and Archer and Diaz [4] .

The continuous-in-time method yields an optimal error in L^2 of $O(h^r)$.

Considered therefore will be the problem

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(A(x, t)u) = f(x, t) \ , \ 0 \leq x \leq 1 \ , \ 0 \leq t \leq T, \quad (4.2)$$

$$u(x, 0) = u_0(x) \ , \ 0 \leq x \leq 1 \ , \quad (4.3)$$

$$u(0, t) = g(t) \ , \ 0 \leq t \leq T \ , \quad (4.4)$$

$$A(x, t) \geq a > 0 \ , \quad (4.5)$$

where (4.5) is assumed to make the boundary condition (4.4) well-posed. The functions A, f, g, and u_0 are given and assumed to be sufficiently smooth for a unique and sufficiently smooth solution to (4.2), (4.3), (4.4) and (4.5) to exist.

2. Notation and Spaces

The notation of Baker's paper [9], will be followed as closely as possible. With $\Omega = (0, 1)$ the Sobolev spaces $H^s = H^s(\Omega)$ are defined following [1], [5] or [6] and are equipped with their standard norm denoted $||\cdot||_s$. For $v: [0, T] \to H^s(\Omega)$ define the norms

$$||v||_{L^p(0, T; H^s)} = \left(\int_0^T ||v(\cdot, t)||_s^p \, dt\right)^{1/p} \ , \ 1 \leq p < \infty \ ,$$

$$||v||_{L^\infty(0, T; H^s)} = \operatorname*{ess\,sup}_{0 \leq t \leq T} ||v(\cdot, t)||_s \ .$$

$L^p(0, T; H^s)$, the space of v with $||v||_{L^p(0, T; H^s)}$ finite, will be abbreviated $L^p(H^s)$. Also the $L^2 = H^0$-norm, $||\cdot||_0$, shall be abbreviated $||\cdot||$ (no subscript). For $u, v \in L^2(0, 1)$ define the inner product

$$(u, v) = \int_0^1 u(x)\, v(x)\, dx .$$

Further, denote by $\mathring{H}(\Omega)$ the subspace of $H^1(\Omega)$ given by $\{v \in H^1(\Omega) | v(1) = 0\}$.

Next define, following Baker op. cit., the subspaces in which the approximate problem is posed. For Δ a partition of $(0, 1)$

$$\Delta: 0 = x_0 < x_1 < \ldots < x_N = 1 , N \geq 1 ,$$

set $h = \max(x_i - x_{i-1})$. For integer $r \geq 0$ and $D \subset \Omega$ let $P_r(D)$ denote the set of polynomials of degree $\leq r$ on D. Define the subspaces

$$G_\Delta^r(\Omega) = \{v | v \in P_{r-1}(x_{i-1}, x_i) , i = 1, \ldots, N\} , \text{ and}$$

$$\mathring{H}_\Delta^r(\Omega) = \{v \in C(\Omega) | v \in P_r(x_{i-1}, x_i) , i = 1, \ldots, N, v(1) = 0\} .$$

Note that $G_\Delta^r(\Omega)$ is a subspace of $L^2(\Omega)$ and $\mathring{H}_\Delta^r(\Omega)$ is a subspace of $\mathring{H}(\Omega)$ and that (see [9], p. 997)

$$\text{Dim } G_{\Delta}^{r}(\Omega) = \text{Dim } \overset{\circ}{H}{}_{\Delta}^{r}(\Omega) = N \cdot r \ .$$

In what follows the dependence in the notation upon Δ and Ω will be supressed. The proofs of the following two lemmata concerning G^r and $\overset{\circ}{H}{}^r$ may be found in ([9], pp. 997 - 998) .

LEMMA 4.1. Let $\{\alpha_1, \ldots, \alpha_d\}$ and $\{\beta_1, \ldots, \beta_d\}$ be bases for G^r and $\overset{\circ}{H}{}^r$ respectively, $d = N \cdot r$. Then the matrix $G = ((\alpha_j, \beta_i))$, $i, j = 1, \ldots, d$, is nonsingular.

LEMMA 4.2. There exists a constant C , $0 < C < \infty$, such that for all $u \in H^s(\Omega)$

$$\inf_{\chi \in G^r} ||u - \chi|| \leq C h^s ||u||_s \ , \ 0 \leq s \leq r \ ,$$

and for all $u \in H^s(\Omega) \cap \overset{\circ}{H}(\Omega)$

$$\inf_{\chi \in \overset{\circ}{H}{}^r} ||u - \chi||_1 \leq C h^{s-1} ||u||_s \ , \ 1 \leq s \leq r + 1 \ .$$

C is independent of Δ , h, and u but depends upon r .

3. A Projection Theorem

The following projection theorem is similar to many that may be found in the literature, see for example [6], [7], or [46] .

For $u \in L^2$ and $v \in H^1$ define the bilinear form

$$b(u, v) = -\int_0^1 A(x, t)\, u(x)\, \frac{\partial v}{\partial x}\, dx \ . \tag{4.6}$$

Then the following lemma, analogous to Baker's (L. 2.3), follows easily:

LEMMA 4.3. Assume $0 < a \le |A(x, t)| \le M < \infty$, $0 \le x \le 1$, $0 \le t \le T$, then $b(u, v)$ is continuous on $L^2 \times H^1$. That is

$$|b(u, v)| \le M||u|| \; ||v||_1 , \; u \in L^2 , \; v \in H^1 . \tag{4.7}$$

Also, there is a positive constant C such that, for every $u \in G^r$,

$$\sup_{v \in \mathring{H}^r;\ ||v||_1 \le 1} |b(u, v)| \ge C||u|| . \tag{4.8}$$

PROOF. (4.7) is trivial. For (4.8) let $u \; (\neq 0) \in G^r$ be given and put

$$v(x) = -\int_1^x u(s)\, ds , \; 0 \le x \le 1 .$$

Then $v \in \mathring{H}^r$ and by the Cauchy-Schwarz inequality

$$||v||_1 \le \sqrt{2}\, ||u|| . \tag{4.9}$$

Then, using (4.5),

$$|b(u, v / ||v||_1)| = ||v||_1^{-1} \left| \int_0^1 A(x, t)\, u(x)\, u(x)dx \right|$$

$$\ge a||v||_1^{-1} \; ||u||^2 \ge (a/\sqrt{2}) \; ||u|| ,$$

which implies (4.8).

The next projection theorem is critical to the work that follows. Abbreviate $u(\cdot, t) = u(t)$, $A(\cdot, t) = A(t)$.

THEOREM 4.1. Assume $A(x, t) \geq a > 0$ and that $|\frac{\partial^k}{\partial t^k} A(x, t)| \leq M < \infty$, $|\frac{\partial A}{\partial x}(x, t)| \leq M < \infty$ for $0 \leq x \leq 1$, $0 \leq t \leq T$, and $k = 0, 1, 2$. Given $u = u(t) \in L^2(0, 1)$ there exists a unique $\tilde{u}: [0, T] \to G^r$ such that

$$b(\tilde{u} - u, v) = 0 , \quad \forall v \in \mathring{H}^r , \ 0 \leq t \leq T .$$

Further, for $0 \leq s \leq r$, $j = 0, 1$, and $u \in L^\infty(0, T; H^s)$ there follows

$$\left|\left|\frac{\partial^k}{\partial t^k}(u - \tilde{u})\right|\right|_{L^\infty(H^{-j})} \leq C\, h^{s+j} \sum_{\ell=0}^{k} \left|\left|\frac{\partial^\ell}{\partial t^\ell} u\right|\right|_{L^\infty(H^s)} . \quad (4.10)$$

PROOF: Existence and uniqueness of u and (4.10) for $k = 0$, $j = 0$ and $j = +1$ follow from lemmas 4.2, 4.3, Baker's Theorem 3.1 p. 1000 of [9] with the identification $G = L^2$, $H = \mathring{H}(\Omega)$, $\mathcal{G} = G^r$ $\mathcal{H} = \mathring{H}^r$ and the proof of Baker's Theorem 3.1, (3.8).

For (4.10) with $k = 1, 2$, $j = 0, 1$, let $w(t)$ be the L^2-projection of $u(t)$ onto G^r . That is, $w(t)$ is given by

$$(w(t) - u(t), v) = 0 , \quad \forall v \in G^r ,$$

and satisfies

$$\left\| \frac{\partial^k}{\partial t^k} (u - w) \right\|_{L^\infty(H^j)} \leq C\, h^{s+j} \left\| \frac{\partial^k}{\partial t^k} u \right\|_{L^\infty(H^s)} , \tag{4.11}$$

for $0 \leq s \leq r$, $k = 0, 1, 2, \ldots$, and $j = 0, 1$. Define $\eta = u - w$, $\phi = \tilde{u} - w$. ϕ then satisfies

$$\int_0^1 A(x, t)\, \phi\, \frac{\partial v}{\partial x}\, dx = \int_0^1 A(x, t)\, \eta\, \frac{\partial v}{\partial x}\, dx , \tag{4.12}$$

$$\begin{aligned} \int_0^1 A(x, t)\, \frac{\partial \phi}{\partial t}\, \frac{\partial v}{\partial x}\, dx &= - \int_0^1 \frac{\partial A}{\partial t}\, \phi\, \frac{\partial v}{\partial x} + \\ &+ \int_0^1 \frac{\partial A}{\partial t}\, \eta\, \frac{\partial v}{\partial x}\, dx \int_0^1 A\, \frac{\partial \eta}{\partial t}\, \frac{\partial v}{\partial x}\, dx , \end{aligned} \tag{4.13}$$

$$\begin{aligned} \int_0^1 A(x, t)\, \frac{\partial^2 \phi}{\partial t^2}\, \frac{\partial v}{\partial x}\, dx &= -2 \int_0^1 \frac{\partial A}{\partial t}\, \frac{\partial \phi}{\partial t}\, \frac{\partial v}{\partial x}\, dx \\ &- \int_0^1 \frac{\partial^2 A}{\partial t^2}\, \phi \frac{\partial v}{\partial x}\, dx + \int_0^1 \frac{\partial^2 A}{\partial t^2}\, \eta \frac{\partial v}{\partial x}\, dx + \\ &+ 2 \int_0^1 \frac{\partial A}{\partial t}\, \frac{\partial \eta}{\partial t}\, \frac{\partial v}{\partial x}\, dx + \int_0^1 A\, \frac{\partial^2 \eta}{\partial t^2}\, \frac{\partial v}{\partial x}\, dx , \end{aligned} \tag{4.14}$$

for all $v \in \mathring{H}^r$. To see (4.10) when $k = 1$ $j = 0$ set $v = -\int_1^x \frac{\partial \phi}{\partial t}(s)ds$ in (4.13). The bounds upon A and its time derivatives gives then

$$\begin{aligned} a \left\| \frac{\partial \phi}{\partial t}(t) \right\|^2 &\leq M \|\phi(t)\| \left\| \frac{\partial \phi}{\partial t}(t) \right\| + \\ &+ M \|\eta(t)\| \left\| \frac{\partial \phi}{\partial t}(t) \right\| + M \left\| \frac{\partial \eta}{\partial t}(t) \right\| \left\| \frac{\partial \phi}{\partial t}(t) \right\| . \end{aligned}$$

Thus from (4.11) and (4.10) with $k = 0$, $j = 0$

$$\left|\left|\frac{\partial\phi}{\partial t}\right|\right|_{L^\infty(L^2)} \leq C\left(\left|\left|\phi\right|\right|_{L^\infty(L^2)} + \right.$$

$$\left. + \left|\left|\eta\right|\right|_{L^\infty(L^2)} + \left|\left|\frac{\partial\eta}{\partial t}\right|\right|_{L^\infty(L^2)}\right)$$

$$\left|\left|\frac{\partial\phi}{\partial t}\right|\right|_{L^\infty(L^2)} \leq C\, h^s\left(\left|\left|\frac{\partial u}{\partial t}\right|\right|_{L^\infty(H^s)} + \left|\left|u\right|\right|_{L^\infty(H^s)}\right) . \qquad (4.15)$$

(4.10) with $k = 1$, $j = 0$ now follows by the above (4.15), (4.11) and the triangle inequality.

For (4.10) when $k = 2$, $j = 0$ set $v = -\int_1^x \frac{\partial^2\phi}{\partial t^2}(s)\, ds$ in (4.14). Once again, the bounds upon A, $\frac{\partial A}{\partial t}$ and $\frac{\partial^2 A}{\partial t^2}$, and the Cauchy-Schwarz inequality yield

$$a\left|\left|\frac{\partial^2\phi}{\partial t^2}\right|\right|^2 \leq 2M\left|\left|\frac{\partial\phi}{\partial t}\right|\right| \left|\left|\frac{\partial^2\phi}{\partial t^2}\right|\right|$$

$$+ M\left|\left|\phi\right|\right| \left|\left|\frac{\partial^2\phi}{\partial t^2}\right|\right| + M\left|\left|\eta\right|\right| \left|\left|\frac{\partial^2\phi}{\partial t^2}\right|\right|$$

$$+ 2M\left|\left|\frac{\partial\eta}{\partial t}\right|\right| \left|\left|\frac{\partial^2\phi}{\partial t^2}\right|\right| + M\left|\left|\frac{\partial^2\eta}{\partial t^2}\right|\right| \left|\left|\frac{\partial^2\phi}{\partial t^2}\right|\right| .$$

Optimality of $\left|\left|\frac{\partial\phi}{\partial t}\right|\right|$ and $\left|\left|\phi\right|\right|$ has previously been established. Optimality of $\left|\left|\eta\right|\right|$, $\left|\left|\frac{\partial\eta}{\partial t}\right|\right|$, and $\left|\left|\frac{\partial^2\eta}{\partial t^2}\right|\right|$ is known. Thus (4.11), (4.10) with $k = 1, 2$, $j = 0$, and the above yields

$$\left|\left|\frac{\partial^2 \phi}{\partial t^2}\right|\right|_{L^\infty(L^2)} \leq C\, h^s \left(\sum_{\ell=0}^{2} \left|\left|\frac{\partial^k u}{\partial t^k}\right|\right|_{L^\infty(H^s)} \right) .$$

(4.10) with $k = 2$, $j = 0$ follows from this inequality and (4.11) with $k = 2$, $j = 0$ and the triangle inequality:

$$\left|\left|\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 \tilde{u}}{\partial t^2}\right|\right| \leq \left|\left|\frac{\partial^2 \phi}{\partial t^2}\right|\right| + \left|\left|\frac{\partial^2 \eta}{\partial t^2}\right|\right| .$$

For the estimate when $j = +1$, let $w \in C^\infty$ be given and define

$$\psi(x) = -\int_1^x w(s)\, ds \ , \ 0 \leq x \leq 1 .$$

Then $\psi \in C^\infty(\Omega) \cap \mathring{H}(\Omega)$ and

$$||\psi||_2 \leq \sqrt{2}\, ||w||_1 \ .$$

Let $\tilde{\psi} \in \mathring{H}^r$ be such that

$$||\psi - \tilde{\psi}||_1 \leq C\, h||\psi||_2 \leq \sqrt{2}\, C\, h||w||_1 \ .$$

Let $e(t) = u(t) - \tilde{u}(t)$. By Liebnitz' formula $\frac{\partial^k e}{\partial t^k}$ then satisfies

$$\int_0^1 A(t) \frac{\partial^k e}{\partial t^k} \frac{\partial v}{\partial x}\, dx = -\sum_{n=1}^{k} \binom{k}{n} \int_0^1 \frac{\partial^n A}{\partial t^n} \frac{\partial^{(k-n)} e}{\partial t^{(k-n)}} \frac{\partial v}{\partial x}\, dx$$

for all $v \in \mathring{H}^r$. Thus

$$|(A(t)\frac{\partial^k e}{\partial t^k}, w)| = |(A(t)\frac{\partial^k e}{\partial t^k}, \frac{\partial \psi}{\partial x})| .$$

Since $\psi \in \mathring{H}^r$ (4.16) may be applied to the above to yield

$$|(A(t)\frac{\partial^k e}{\partial t^k}, w)| \leq |(A(t)\frac{\partial^k e}{\partial t^k}, \frac{\partial \psi}{\partial x} - \frac{\partial \tilde{\psi}}{\partial x})| +$$

$$+ \sum_{n=1}^{k} \binom{k}{n} |(\frac{\partial^n A}{\partial t^n} \frac{\partial^{(k-1)} e}{\partial t^{k-n}}, \frac{\partial \tilde{\psi}}{\partial x})|$$

$$\leq C||\frac{\partial^n e}{\partial t^k}|| \; ||\psi - \tilde{\psi}||_1 +$$

$$+ C \sum_{n=1}^{k} |(\frac{\partial^n A}{\partial t^n} \frac{\partial^{(k-n)} e}{\partial t^{(k-n)}}, \frac{\partial}{\partial x}(\tilde{\psi} - \psi) + \frac{\partial \psi}{\partial x})|$$

$$\leq C||\frac{\partial^k e}{\partial t^k}|| \; ||\psi - \tilde{\psi}||_1 +$$

$$+ C \sum_{n=1}^{k} (||\frac{\partial^{(k-n)} e}{\partial t^{(k-n)}}|| \; ||\psi - \tilde{\psi}||_1 + ||\frac{\partial^{(k-n)} e}{\partial t^{(k-n)}}||_{-1} \; ||\frac{\partial \psi}{\partial x}||_1) .$$

Since $||\psi - \tilde{\psi}||_1 \leq C\, h||w||_1$ and $||\psi||_2 \leq C||w||_1$ there follows

$$|(A(t)\frac{\partial^k e}{\partial t^k}, w)| \leq C\, h \sum_{n=0}^{k} ||\frac{\partial^{(k-n)} e}{\partial t^{(k-n)}}|| \; ||w||_1 +$$

$$+ C \sum_{n=1}^{k} ||\frac{\partial^{(k-n)} e}{\partial t^{(k-n)}}||_{-1} \; ||w||_{+1} .$$

Now since $a \leq A(x, t) \leq M$ and $|\frac{\partial}{\partial x} A(x, t)| \leq M$ it follows that $||w||_1 \leq C||A(t)\, w||_1$. Thus

$$|(\frac{\partial^k e}{\partial t^k} , A(t)\ w)| \le C\ h \sum_{n=0}^{k} ||\frac{\partial^{(k-n)} e}{\partial t^{(k-n)}}||\ \ ||A(t)\ w||_1 +$$

$$+ C \sum_{n=1}^{k} ||\frac{\partial^{(k-n)} e}{\partial t^{(k-n)}}||_{-1}\ ||A(t)\ w||_1 .$$

Dividing by $||A(t)\ w||_1$ and taking sup over $w \in C^\infty$ gives that

$$||\frac{\partial^k e}{\partial t^k}||_{-1} \le C\ h \sum_{n=0}^{k} ||\frac{\partial^{(k-n)} e}{\partial t^{(k-n)}}|| + C \sum_{n=1}^{k} ||\frac{\partial^{(k-n)} e}{\partial t^{(k-n)}}||_{-1} \quad (4.17)$$

Thus, by induction, the theorem will follow for all k if it can be shown for $k = 1$. But from (4.17) it is clear that the theorem holds for $k = 1$ provided $||e||_{-1}$ is of optimal order. To see this, consider

$$|(A(t)\ e, w)| = |(A(t)\ e\ , \frac{\partial \psi}{\partial x})|$$

$$= |A(t)\ e\ , \frac{\partial \psi}{\partial x} - \frac{\partial \tilde{\psi}}{\partial x})|$$

$$\le C||e||\ \ ||\psi - \tilde{\psi}||_{+1}$$

$$\le C||e||\ \ h||w||_1$$

Thus, because $||w||_1 \le C||A(t)\ w||_1$,

$$\sup_{w \in C^\infty} \frac{|(e, A(t)\, w)|}{||A(t)\, w||_{+1}} \leq C\, h||e|| \quad ,$$

and the theorem follows.

4. The Semi-Discrete Method

The continuous-in-time Galerkin approximation is defined as a differentiable map $U\colon [0, T] \to G^r$ satisfying

$$(U(0), v) = (u_0, v) \quad , \quad \forall v \in G^r \tag{4.18}$$

$$(\frac{\partial U}{\partial t} , v) + b(U, v) = (f(t), v) + \\ + A(0, t)\, g(t)\, v(0) \quad , \forall v \in \mathring{H}^r \quad . \tag{4.19}$$

The following theorem then asserts the optimal order of convergence of the continuous-in-time Galerkin approximation in L^2 .

THEOREM 4.2. Let u be the solution to the continuous problem (4.2), (4.3), (4.4). Then the Galerkin approximation, $U(t)$, defined by (4.18), (4.19) exists and is unique for $t \geq 0$.

Furthermore, provided $A(x, t) \geq a > 0$, $|\frac{\partial A}{\partial x}(x, t)| \leq M < \infty$, $|\frac{\partial^k}{\partial t^k} A(x, t)| \leq M < \infty$, $k = 0, 1, 2$, $0 \leq t \leq T$, $0 \leq x \leq 1$ and

$$u \in L^\infty(0, T; H^s(\Omega)) \quad ,$$

$$\frac{\partial u}{\partial t} , \frac{\partial^2 u}{\partial t^2} \in L^\infty(0, T; H^{s-1}(\Omega))$$

there follows, for $0 \leq s \leq r$,

$$\|u - U\|_{L^\infty(H^\circ)} \le C\|u(0) - U(0)\| + \tag{4.20}$$

$$+ C\, h^s \{\|u\|_{L^\infty(H^s)} + \|\frac{\partial u}{\partial t}\|_{L^\infty(H^{s-1})} + \|\frac{\partial^2 u}{\partial t^2}\|_{L^\infty(H^{s-1})}\}$$

PROOF OF THEOREM 4.2. Expanding $U(t) = \sum_{k=1}^{d} \lambda_k(t)\, \alpha_k$ yields the initial-value problem for the system of ordinary differential equations

$$G \frac{d\lambda}{dt} + B\, \lambda = F(t) , \quad 0 < t \le T , \tag{4.21}$$

$$A\, \lambda(0) = u_0$$

where λ, F and u are the vectors

$$\lambda(t) = (\lambda_1(t), \ldots, \lambda_d(t))^T$$

$$F(t) = ((f(\cdot, t), \beta_1), \ldots, (f(\cdot, t), \beta_d))^T +$$

$$+ (A(0, t)\, g(t)\, \beta_1(0), \ldots, A(0, t)\, g(t)\, \beta_d(0))^T$$

$$u_0 = ((u_0, \alpha_1), \ldots, (u_0, \alpha_d))^T ,$$

and G, B, A are the $d \times d$ matrices

$$G_{ij} = (\alpha_j, \beta_i) , \quad B_{ij} = b(\alpha_j, \beta_i) , \quad A_{ij} = (\alpha_j, \alpha_i) .$$

Since A is nonsingular, and by lemma 4.1 G is nonsingular this system has a unique solution and $U(t)$ exists and is unique.

To derive (4.20), let $\tilde{u}$ be the projection of u on G^r with respect to b defined in Theorem 4.1, that is, let

$$b(u - \tilde{u}, v) = 0 , \quad \forall v \in \mathring{H}^r ,$$

and define $\eta = u - \tilde{u}$, $\phi = U - \tilde{u}$, and $e = u - U$. ϕ estimated in terms of η, $\frac{\partial \eta}{\partial t}$, and $\frac{\partial^2 \eta}{\partial t^2}$. ϕ satisfies the equation

$$(\frac{\partial \phi}{\partial t} , v) + b(\phi, v) = (\frac{\partial \eta}{\partial t} , v) , \quad v \in \mathring{H}^r . \tag{4.22}$$

Define

$$\forall v = \hat{v}(x, t) = -\int_1^x \frac{\partial \phi}{\partial t} (s, t)ds , \psi(x, t) = -\int_1^x \phi(s, t)ds$$

so that $\frac{\partial \psi}{\partial t} = \hat{v}$ and $\frac{\partial \hat{v}}{\partial x} = - \frac{\partial \phi}{\partial t}$. Setting $v = \hat{v}$ in (4.22) yields

$$-(\frac{\partial \hat{v}}{\partial x} , \hat{v}) + b(\phi, \hat{v}) = (\frac{\partial \eta}{\partial t} , \frac{\partial \psi}{\partial t}) .$$

Since $-(\frac{\partial \hat{v}}{\partial x} , \hat{v}) = \frac{1}{2} \hat{v}(0, t)^2$ this becomes

$$b(\phi, v) \leq (\frac{\partial \eta}{\partial t} , \frac{\partial \psi}{\partial t}) .$$

Consider first the left-hand side of (4.23).

$$b(\phi, v) = \int_0^1 A(x, t) \phi(x, t) \frac{\partial \phi}{\partial t} (x, t)dx$$

$$\geq \frac{\partial}{\partial t} \int_0^1 \frac{1}{2} A \, \phi^2 \, dx - \frac{M}{a} \int_0^1 \frac{1}{2} A \, \phi^2 \, dx \ .$$

Multiplying (4.23) by $\exp(-Mt/a)$ and using the above yields

$$\frac{\partial}{\partial t} \left(e^{-\frac{M}{a} t} \int_0^1 \frac{1}{2} A \, \phi^2 \, dx\right) \leq e^{-\frac{M}{a} t} \left(\frac{\partial \eta}{\partial t} , \frac{\partial \psi}{\partial t}\right) \ .$$

Hence,

$$e^{-\frac{M}{a} t} \int_0^1 \frac{1}{2} A(t) \, \phi(t)^2 \, dx \leq \int_0^1 \frac{1}{2} A(0) \, \phi(0)^2 \, dx + \int_0^t e^{-\frac{M}{a} s} \left(\frac{\partial \eta}{\partial t}(s), \frac{\partial \psi}{\partial t}(s)\right) ds \ . \tag{4.24}$$

But now

$$\left(\frac{\partial \eta}{\partial t} , \frac{\partial \psi}{\partial t}\right) = \frac{\partial}{\partial t} \left(\frac{\partial \eta}{\partial t} , \psi\right) - \left(\frac{\partial^2 \eta}{\partial t^2} , \psi\right) ,$$

and hence,

$$\int_0^t e^{-\frac{M}{a} s} \left(\frac{\partial \eta}{\partial t} , \frac{\partial \psi}{\partial t}\right) ds = \int_0^t e^{-\frac{M}{a} s} \frac{\partial}{\partial t} \left(\frac{\partial \eta}{\partial t} , \psi\right) ds - \int_0^t e^{-\frac{M}{a} s} \left(\frac{\partial^2 \eta}{\partial t^2} , \psi\right) ds$$

$$= e^{-\frac{M}{a} t} \left(\frac{\partial \eta}{\partial t}(t), \psi(t)\right) - \left(\frac{\partial \eta}{\partial t}(0) , \psi(0)\right)$$

$$- \int_0^t (-\frac{M}{a})\, e^{-\frac{M}{a}s} (\frac{\partial^2 \eta}{\partial t^2}, \psi) ds$$

$$- \int_0^t e^{-\frac{M}{a}s} (\frac{\partial^2 \eta}{\partial t^2}, \psi) ds \quad .$$

Thus, the Cauchy-Schwarz-Young inequality yields

$$\left| \int_0^t e^{-\frac{M}{a}s} (\frac{\partial \eta}{\partial t}, \frac{\partial \psi}{\partial t}) ds \right| \leq C\, \varepsilon^{-1} \left|\left| \frac{\partial \eta}{\partial t} \right|\right|_{L^\infty(H^{-1})} + \tag{4.25}$$

$$+ C\, \varepsilon^{-1} \left|\left| \frac{\partial^2 \eta}{\partial t^2} \right|\right|_{L^\infty(H^{-1})} + C\, \varepsilon \left|\left| \psi \right|\right|_{L^\infty(H^1)} \quad ,$$

where C depends upon M, a, and T but not u, h, η or ψ .

Using (4.25) and the bounds on A and $\frac{\partial^k A}{\partial t^k}$ in (4.24) yields

$$C_1 ||\phi(t)||^2 \leq C||\phi(0)||^2 + C\, \varepsilon \left|\left| \psi \right|\right|_{L^\infty(H^1)} \quad ,$$

where, for example, $C_1 = \frac{1}{2} a \exp(-M\, T/a)$ is sufficient. Taking sup over t , $0 \leq t \leq T$, in the above and dividing by C_1 gives

$$||\phi||^2_{L^\infty(H^0)} \leq C||\phi(0)||^2 +$$

$$C\, \varepsilon^{-1} \left|\left| \frac{\partial \eta}{\partial t} \right|\right|^2_{L^\infty(H^{-1})} + C\, \varepsilon^{-1} \left|\left| \frac{\partial^2 \eta}{\partial t^2} \right|\right|^2_{L^\infty(H^{-1})} +$$

$$+ C\, \varepsilon\, ||\phi||^2_{L^\infty(H^0)} \quad ,$$

where $||\psi||_1 \leq C||\phi||$ was used. Picking $\varepsilon = C/2$ and using Theorem 4.1 gives that

$$||\phi||^2_{L^\infty(H^0)} \leq C||\phi(0)||^2 +$$

$$C\, h^{2s}(||\frac{\partial u}{\partial t}||^2_{L^\infty(H^{s-1})} + ||\frac{\partial^2 \eta}{\partial t^2}||^2_{L^\infty(H^{s-1})})$$

holds for $0 \leq s \leq r$.

The triangle inequality

$$||u - U|| \leq ||\phi|| + ||\eta|| ,$$

$$||\phi(0)|| = ||U(0) - u(0) + u(0) - \tilde{u}(0)|| ,$$

then yield

$$||u - U||_{L^\infty(H^0)} \leq C||u(0) - U(0)|| +$$

$$+ C\, h^s(||u||_{L^\infty(H^s)} + ||\frac{\partial u}{\partial t}||_{L^\infty(H^{s-1})} + ||\frac{\partial^2 \eta}{\partial t^2}||_{L^\infty(H^{s-1})}$$

which is (4.20). Note that the definition of $U(0)$ is arbitrary up to being an optimal in L^2 approximation to u_0 in G^r.

LIST OF REFERENCES

LIST OF REFERENCES

1. Adams, R.A., Sobolev Spaces, New York: Academic Press, 1975.

2. Apelkrans, M.Y.T., "On Difference Schemes for Hyperbolic Equations with Discontinuous Initial Values," Math. Comp., 22(1968), pp. 525 - 539.

3. Archer, D., "Collocation in C^0 spaces for First Order Hyperbolic Equations I: Optimal Order Global Estimates for Quasilinear Problems," SIAM J. Numer. Anal., 15(1978), pp. 271 - 281.

4. Archer, D. and Diaz, J., "A Collocation-Galerkin Method for a First Order Hyperbolic Equation with Space and Time Dependent Coefficients," (1979), to appear in SIAM J. Numer. Anal. .

5. Aubin, J.P., Applied Functional Analysis, New York: Wiley-Interscience, 1979.

6. Aubin, J.P., Approximation of Elliptic Boundary Value Problems, New York: Wiley-Interscience, 1972.

7. Babuska, I. and Aziz, A.K., "Survey Lectures on the Mathematical Foundations of the Finite Element Method," The Mathematical Foundations of the Finite Method with Applications to Partial Differential Equations, New York: Academic Press, (1972).

8. Baker, G.A., "Galerkin Approximations for the Navier-Stokes Equations," preprint, 1976.

9. Baker, G.A., "A Finite Element Method for First Order Hyperbolic Equations," Math. Comp., 29(1975), pp. 995 - 1006.

10. Bramble, J.H. and Sammon, P.H., "Efficient Higher Order Single Step Methods for Parabolic Problems, Part I," M.R.C. T.S.R. 1958, May, (1979).

11. Bramble, J.H. and Thomee, V., "Discrete Time Galerkin Methods for a Parabolic Boundary Value Problem," Ann. Mat. Pura. Appl., 101(1974), pp. 115 - 152.

12. Brenner, Ph. and Thomee, V., "Stability and Convergence Rates in L^p for Certain Difference Schemes," Math. Scand., 27(1970), pp. 5 - 23.

13. Brenner, Ph. and Thomee, V., "Estimates near Discontinuities for some Difference Schemes," Math. Scand., 28(1971), pp. 329 - 340.

14. Brenner, Ph., Thomee, V. and Wahlbin, L., Besov Spaces and Applications to Difference Methods for Initial Value Problems, Lecture Notes in Mathematics vol. 434, New York: Springer-Verlag, 1975.

15. Butzer, P.L. and Berens, H., Semigroups of Operators and Approximation, Berlin, Heidelberg, New York: Springer-Verlag, 1967.

16. Calderon, A., Spitzer, F. and Widom, H., "Inversion of Toeplitz Matrices," Illinois J. Math, 13(1959), pp. 490 - 498.

17. Dendy, J.E., "Two Methods of Galerkin Type Achieving Optimum L^2-Accuracy for First Order Hyperbolics," SIAM J. Numer. Anal., 11(1974), pp. 637 - 653.

18. Diaz, J.C., "Collocation H^{-1}-Galerkin Method for Parabolic Problems with Time Dependent Coefficients," SIAM J. Numer. Anal., 16(1979), pp. 911 - 922.

19. Douglas, J. Jr. and Dupont, T., "The Effect of Interpolating the Coefficients in Nonlinear Parabolic Galerkin Procedure," Math. Comp., 29(1975), pp. 360 - 389.

20. Dupont, T., "Galerkin Methods for First Order Hyperbolics: An Example," SIAM J. Numer. Anal., 10(1973), pp. 890 - 899.

21. Dupont, T., "Galerkin Methods for Modeling Gas Pipelines," Constructive and Computational Methods for Differential and Integral Equations, D.L. Colton and R.P. Gilbert (Editors), Lecture Notes in Math., vol. 430, New York: Springer-Verlag, 1974, pp. 112 - 130.

22. Fix, G. and Nassif, N., "On Finite Element Approximations to Time Dependent Problems," Numer. Math., 19(1972), pp. 127 - 135.

23. Friedrichs, K.O., "Symmetric Hyperbolic Linear Differential Equations," Comm. Pure Appl. Math., 17(1954), pp. 345 - 392.

24. Friedrichs, K.O. and Lax, P.D., "On Symmetrizable Differential Operators," Proc. Symp. in Pure Math., vol. 10, Amer. Math. Soc., 1967, pp. 128 - 137.

25. Gary, J., "On Boundary Conditions for Hyperbolic Difference Schemes," J. Comp. Phys., 26(1978), pp. 339 - 351.

26. Gunzburger, M., "On the Stability of Galerkin Methods for Initial-Boundary Value Problems for Hyperbolic Systems." Math. Comp., 31(1977), pp. 661 - 675.

27. Gunzburger, M. and Plemmons, R., "Energy Conserving Norms for the Solution of Hyperbolic Systems of Partial Differential Equations," Math. Comp., 33(1979), pp. 1 - 10.

28. Gustavson, B., Kreiss, H.O. and Sundstrom, A., "Stability Theory of Difference Approximations for Mixed Initial Boundary Value Problems, II," Math. Comp., 26(1972), pp. 649 - 686.

29. Hedstrom, G.W., "The Rate of Convergence of Some Difference Schemes," SIAM J. Numer. Anal., 5(1968), pp. 363 - 406.

30. Katsanis, Th., "Numerical Solution of Symmetric Positive Differential Equations," Math. Comp., 22(1968), pp. 763 - 783.

31. Kreiss, H.O., "Uber implizite Differenzmethoden fur partielle Differentialgleichungen," Numer. Math., 5(1963), pp. 24 - 47.

32. Kreiss, H.O., "Initial Boundary Value Problems for Hyperbolic Systems," Comm. Pure. Appl. Math., 23(1970), pp. 277 - 298.

33. Kreiss, H.O., "Stability Theory for Difference Approximations of Mixed Initial Boundary Value Problems, I," Math. Comp., 22(1968), pp. 1 - 12.

34. Kreiss, H.O. and Lundquist, E., "On Difference Approximations with the Wrong Boundary Values," Math. Comp., 22(1968), pp. 1 - 12.

35. Kreiss, H.O. and Oliger, J., Methods for the Approximate Solution of Time Dependent Problems, GARP Publication Series No. 10, 1973.

36. Kupka, I.A.K. and Osher, S., "On the Wave Equation in a Multidimensional Corner," Comm. Pure. Appl. Math., 24(1971), pp. 381 - 393.

37. Lax, P.D. and Phillips, R.S., "Local Boundary Conditions for Dissipative Symmetric Linear Differential Operators," Comm. Pure. Appl. Math., 13(1960), pp. 427 - 455.

38. Lesaint, P., "Finite Element Methods for Symmetric Hyperbolic Systems," Numer. Math., 21(1973/74), pp. 244 - 255.

39. Lesaint, P., "Finite Element Methods for the Transport Equation," R.A.I.R.O. Serie Math., August(1974), pp. 67 - 94.

40. Lesaint, P., "Continuous and Discontinuous Finite Element Methods for Solving the Transport Equation," The Mathematics of Finite Elements and Applications II, J.R. Whiteman (Editor), New York: Academic Press, 1976, pp. 151 - 161.

41. Lions, J.L., Equations Differentielles Operationelles, New York: Springer, 1961.

42. Lions, J.L. and Magenes, E., Non-Homogeneous Boundary Value Problems and Applications, vol. I, New York: Springer-Verlag, 1972.

43. Luskin, M., "A Finite Element Method for First Order Hyperbolic Systems," submitted to Math. Comp., (1979).

44. Mock. M.S., "Projection Methods with Different Test and Trial Spaces," Math. Comp., 30(1976), pp. 400 - 416.

45. Mock, M.S. and Lax, P.D., "The Computation of Discontinuous Solutions of Linear Hyperbolic Equations," Comm. Pure. Appl. Math., 31(1978), pp. 423 - 430.

46. Nitsche, J., "Ein Kriterium fur die Quasi-Optimalitat des Ritzschen Verfahrens," Numer. Math., 11(1968), pp. 346 - 348.

47. Osher, S., "Initial Boundary Value Problems for Hyperbolic Equations in Regions with Corners, II," Trans. Amer. Math. Soc., 198(1973), pp. 141 - 164.

48. Osher, S., "An Ill-Posed Problem for a Strictly Hyperbolic Equation in Two Unknowns near a Corner," Bull. A.M.S., 80(1974), pp. 705 - 708.

49. Parter, S.V., "On the Roles of "Stability" and "Convergence" in Semidiscrete Projection Methods for Initial-Value Problems," Math. Comp., 34(1980), pp. 127 - 154.

50. Peetre, J. and Thomee, V., "On the rate of Convergence for Discrete Initial Value Problems," Math. Scand. 21(1967), pp. 159 - 176.

51. Rachford, H.H., Jr. and Dupont, T., "A Fast, High Accuracy Model for Transient Flow in Gas Pipeline Systems by Variational Methods," Soc. Petroleum Eng. Journ., 4005, Fall, (1972).

52. Rachford, H.H., Jr. and Wheeler, M.F., "H^{-1} Galerkin Procedure for the Two Point Boundary Value Porblem," Mathematical Aspects of the Finite Element Method, C. DeBoor (Editor), New York: Academic Press, 1974.

53. Rauch, J., "Energy and Resolvent Inequalities for Hyperbolic Mixed Problems," J. Differential Eqns., 11(1972), pp. 528 - 540.

54. Rauch, J., "L^2 is a Continuable Initial Condition for Kreiss' Mixed Problems," Comm. Pure. Appl. Math., 15(1972), pp. 265 - 285.

55. Richards, R., "Uniform Spline Interpolation Operators in L^2," M.R.C. T.S.R. #1305, November, 1972.

56. Richtmeyer, R.D. and Morton, K.W., Difference Methods for Initial Value Problems, Second Edition, New York: Interscience, 1967.

57. Sarason, L. and Smoller, J., "Geometrical Optics and the Corner Problem," Arch. Rat. Mech. and Anal., 56(1975), pp. 34 - 69.

58. Schoenberg, I.J., Cardinal Spline Interpolation, Regional Conference Series in Applied Math., no. 12, Philadelphia: SIAM, 1973.

59. Schoenberg, I.J., "Cardinal Interpolation and Spline Functions II," J. App. Thy., 6(1972), pp. 404 = 420.

60. Serdyukova, S.I., "Oscillations Which Occur in the Numerical Computation of the Discontinuous Solutions of Differentiable Equations," Zh. vychisl. Mat. Fiz., 11(1971), pp. 411 - 424.

61. Strang, G., "Hyperbolic Initial-Boundary Value Problems in Two Unknowns," J. Diff. Equs., 16(1969), pp. 161 - 171.

62. Strikwerda, J., "Initial Boundary Value Problems for the Method of Lines," I.C.A.S.E. Report #78-16, 1978.

63. Swartz, B. and Wendroff, B., "Generalized Finite Difference Schemes," Math. Comp., 20(1969), pp. 37 - 50.

64. Swartz, B. and Wendroff, B., "The Relative Efficiency of Finite Difference and Finite Element Methods I. Hyperbolic Problems and Splines," SIAM J. Numer. Anal., 11(1974), pp. 979 - 993.

65. Thomee, V., "Estimates of the Friedrichs-Lewy Type for Mixed Problems in the Theory of Linear Hyperbolic Differential Equations in Two Independent Variables," Math. Scand., 5(1957), pp. 93 - 113.

66. Thomee, V., "On Maximum-Norm Stable Difference Operators," Proceedings of Symposium on Partial Differential Equations, New York: Academic Press, (1966), pp. 125 - 151.

67. Thomee, V., "Spline Approximation and Difference Schemes for the Heat Equation," The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, A.K. Aziz (Editor), New York: Academic Press, (1972), pp. 715 - 746.

68. Thomee, V., "Convergence Estimates for Semidiscrete Galerkin Methods for Initial Value Problems," Numerische, insbesondere approximations theoretische Behandlung von Funktionalgleichungen, Lecture Notes in Math. vol. 333, Berlin: Springer, 1973.

69. Thomee, V. and Wendroff, B., "Convergence Estimates for Galerkin Methods for Variable Coefficient Initial Value Problems," SIAM J. Numer. Anal., 11(1974), pp. 1059 - 1068.

70. Varga, R.S., Functional Analysis and Approximation Theory in Numerical Analysis, Regional Conference Series in Applied Math, no. 3, Philadelphia: SIAM, 1971.

71. Wahlbin, L., "A Modified Galerkin Procedure with Hermite Cubics For Hyperbolic Problems," Math. Comp., 29(1975), pp. 978 - 984.

72. Wahlbin, L., "A Dissipative Galerkin Method Applied to some Quasi-linear Hyperbolic Equations," R.A.I.R.O. Serie Math., 8(1974), pp. 109 - 117.

73. Wahlbin, L., "A Dissipative Galerkin Method for the Numerical Solution of First Order Hyperbolic Equations," Mathematical Aspects of Finite Elements in Partial Differential Equations, C. DeBoor (Editor), New York: Academic Press, 1974.

74. Wendroff, B., "On Finite Elements for Equations of Evolution," Tech. Rep. LA-DC-72-1220, Los Alamos, N.M., 1972.

VITA

William J. Layton was born in Atlanta, Georgia on August 21, 1956. He attended elementary schools in that city. Later he moved to Milledgeville, Georgia where he graduated from Baldwin High School in June 1974. He attended George College in Milledgeville where he received a Bachelor of Science degree in Mathematics.

Mr. Layton enrolled in graduate school at The University of Tennessee, Knoxville in the spring of 1976 where he was awarded the Dryzer fellowship. Later he obtained a teaching assistantship. He completed the thesis requirement for the Doctor of Philosophy degree with a major in Mathematics in August 1980.