Summer 2010

A Solution of the Heat Equation with the Discontinuous Galerkin Method Using a Multilivel Calculation Method That Utilizes a Multiresolution Wavelet Basis

Robert Gregory Brown
Old Dominion University

A SOLUTION OF THE HEAT EQUATION WITH THE DISCONTINUOUS GALERKIN METHOD USING A MULTILEVEL CALCULATION METHOD THAT UTILIZES A MULTIRESOLUTION WAVELET BASIS

by

Robert Gregory Brown
M.S. Virginia Commonwealth University 1986

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirement for the Degree of

DOCTOR OF PHILOSOPHY

MATHEMATICS AND STATISTICS

OLD DOMINION UNIVERSITY
August 2010

Approved by:

Richard Noren (Director)

Hideaki Kaneko

Jin Wang

Dimitrie Popescu

ABSTRACT

A SOLUTION OF THE HEAT EQUATION WITH THE DISCONTINUOUS GALERKIN METHOD USING A MULTILEVEL CALCULATION METHOD THAT UTILIZES A MULTIRESOLUTION WAVELET BASIS

Robert Gregory Brown
Old Dominion University, 2010
Director: Dr. Richard Noren

A numerical method to solve the parabolic problem is developed that utilizes the Discontinuous Galerkin Method for space and time discretization. A multilevel method is employed in the space variable. It is shown that use of this process yields the same level of accuracy as the standard Discontinuous Galerkin Method for the heat equation, but with cheaper computational cost. The results are demonstrated using a standard one-dimensional homogeneous heat problem.

ACKNOWLEDGMENTS

It has been a very long and difficult road for me since I decided to undertake this venture, and I have so many people to thank for their inspiration, encouragement, and confidence in me during the many long and difficult days it took to get to this point.

First I must thank my wife Adriana, and my children, Henry and Rebecca, for being so patient and supportive during these years.

To my mother, Constance Brown, I owe so much gratitude for being the firm, kind, and gentle mother that she has always been to me. Her incredible strength and spirit as well as tireless confidence in my efforts have helped me through many difficulties in my life, and particularly the difficulties associated with completing this lengthy and expensive academic endeavor.

I also thank my deceased father, Henry Brown, for showing me that to succeed in life, one must be willing leave their comfort zone far behind, take a risk, work hard and smart into the nights and weekends, and venture into new and unknown places. Further, his ability to keep his sense of humor in hard times taught me much about dealing with the difficulties of graduate studies as well as life in general.

I am especially grateful to my thesis advisor, Richard D. Noren, for all the patient help, guidance, and insight he has provided me through good times and difficult times. His experience and savvy for research mathematics has been inspiration to me, and I do not think I could have found an advisor with more enthusiasm for the subject.

Hideaki Kaneko deserves special praise, for taking an interest in my efforts, pointing the way forward, and always knowing the right move for me to make and when to make it.

Also I extend my gratitude to the other faculty of the Old Dominion University Mathematics Department who granted me the privilege of being a part of this program in the first place, and continuing to believe in me as I suffered, sacrificed, and struggled to stay with it. This includes J. Mark Dorrepaal for providing employment in lean times, and always being so professional and diplomatic; John Tweed and John J. Swetits, who pushed me far beyond my previous academic limits and gave me the motivation to try to be the best I could be; Fang Q. Hu, for his pragmatic lectures on the finite element methods; D. Glenn Lasseigne, for efforts to assist me with Green's functions that went far beyond the call of duty; Gordon Melrose, for enduring two

semesters of me as his tensor calculus student without either of us breaking down; Ruhai Zhou, for teaching me numerical differential equations and methods.

There are several professors from academic times past to whom I must extend my gratitude.

I thank Ronald L. Moore of Randolph-Macon College, for standing by me all these years, and providing endorsement of my efforts.

And I also thank Robert E. Offenbacker, also of Randolph-Macon College, for inspiring me to study higher mathematics in the first place.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Many excellent discrete schemes for the parabolic problem, such as the standard Galerkin method, are derived by first discretizing in the spatial variable using the finite element method, which produces a system of ordinary differential equations with respect to the time variable, and then applying one of the many finite difference time stepping methods to this system, resulting in a fully discrete system and the resulting solution. One characteristic of these schemes is that it can be cumbersome to alter the size of the time steps in the middle of the process; also there may be stability issues as well, depending on the choice of finite difference method used for the time discretization. A method to circumvent these difficulties is to apply the Galerkin finite element method in *both* spatial *and* time variables, and one such scheme that utilizes this strategy is known as the Discontinuous Galerkin Method. This method treats the space and time variables in a similar way, and allows the spatial grid mesh as well as the time steps to be varied as necessary. Such a scheme is advantageous for parabolic problems as it allows small time steps in transients and then larger time steps as the exact solution becomes smoother. This will allow more efficient computation. The approximate solution sought will be a piecewise polynomial in the space variable which will not be required to be continuous at the nodes of the time partition.

The Discontinuous Galerkin Method was introduced in 1981 for ordinary differential equations by M. Delfour, W. Hager, and F. Trochu [8]. Application to partial differential equations appeared in works such as [9] by P. Jamet. Major contributors in the area of parabolic problems are Kenneth Eriksson and Claes Johnson, whose works are too numerous to mention; see [7] as a good example of their work. Eriksson and Johnson are especially noteworthy as one of the error estimates found in [7] served as motivation for the essential error estimates of this thesis. Another major contributor is Vidar Thomee, see [4] and references therein. This source provides not only an extremely comprehensive analysis of the Discontinuous Galerkin Method, but a very complete description of parabolic problem solutions by Galerkin finite element methods. It provided much of the background material for this thesis. Two additional works that also deserve mention are [11] by Beatrice Riviare and [12] by Jan S. Hesthaven and Tim Warburton.

Since the Discontinuous Galerkin Method requires solution of large scale linear systems, a multilevel augmentation method will provide a way to ease the computational cost. This method is based on direct sum decompositions of the range space of the operator and the solution space of the operator equation, along with a matrix splitting scheme. The net effect will be to reduce the task of solving a large linear system to that of solving several linear systems of smaller sizes, thus cutting computation costs, and it is demonstrated in this thesis. The papers [1], [2], and [3] by Zhongying Chen, Bin Wu, Yuesheng Xu, and Charles Micchelli were essential in this area, providing much of the framework for the multilevel method utilized in this thesis.

However, for the multilevel method to function correctly, we need a good, multiresolutional basis, and that is the role of the multiscale orthonormal wavelet bases in Sobolev spaces. Further, these bases will produce sparse matrices in the implementations. Again, the various papers such as [2], and [3] by Zhongying Chen, Bin Wu, and Yuesheng Xu, provide excellent analysis as well as efficient notation for the kind of wavelet bases used in this work.

This thesis provides a numerical scheme for approximating the solution of the parabolic problem using a coarse grid, rather than a fine grid, at a lower computational cost, while at the same time preserving the accuracy of the traditional fine grid, higher cost Discontinuous Galerkin Method. It does this by combining the Discontinuous Galerkin Method with Multilevel Augmentation Method to produce what in effect is an approximate solution to the approximate solution of the problem. We prove the convergence rate of the multilevel Discontinuous Galerkin Method solution is exactly the same as the conventional Discontinuous Galerkin Method solution. We also prove the computational costs are considerably less with this method. Finally, we demonstrate these features with several numerical examples. While these demonstrations are performed using simple one-dimensional problems, the methods introduced in this paper should be able to be generalized in the future to higher dimensions through the use of higher dimensional wavelet bases, and thus become applicable to regions that are thin bodies, such as the wing panels of an airplane, or the hull panels of a spacecraft.

As many actual applications present solutions with weak singularities, special time and spatial discretization schemes are needed to obtain good numerical solutions, and various contributors to this area of study include [6] by Hideaki Kaneko,

Kim S. Bey, and Gene J. W. Hou, [7] by Kenneth Eriksson, and Claes Johnson, and [10] by Dominik Schotzau and Christoph Schwab. The capacity of the Discontinuous Galerkin Method to alter time and space grid resolutions in midstream is quite beneficial here, allowing us to use fine grids during transients, and coarse grids when the solutions have smoothed, altering them as needed from time step to time step. We introduce a new error estimate which is essentially a multilevel version of the time step and grid mesh sizing error estimate detailed by Kaneko, Bey, and Hou in [6]. It shows the accuracy of the error estimate of [6] remains the same when the multilevel method is used to enhance the computational efficiency. As before, we provide numerical demonstrations of these results.

This paper is organized into seven parts, including Chapter I, the introduction. In Chapter II, the parabolic problem and the Discontinuous Galerkin Method are developed, along with a multilevel augmentation method. In conjunction with the multilevel method, a multiscale orthonomal wavelet basis is discussed, and the specific basis used in the implementations is constructed. In Chapter III, these various notions are then blended together as one method, and applied to the basic parabolic problem. New convergence results and error estimates, refined and enhanced from existing multilevel convergence and error results, are then developed in Chapter IV. Further, the main result of Chapter IV, Theorem 4.2, is shown to apply under two different sets of hypotheses. One set, based on the results of [3] by Chen, Wu, and Xu, requires the operator equation to have a uniformly bounded inverse. The other set of hypotheses, introduced in this thesis, allows the norm of the same inverse to go to infinity, which is an intractable situation for the requirements of [3]. Thus, to provide versatility to the method developed here as well as extend the result of [3], we prove both versions of the theorem. Special time and spatial discretizations from [6], designed to treat difficult initial conditions, are described in Chapter V, along with a new error estimate in the form of Theorem 5.5. As before, we show this new result applies under the same two different sets of hypotheses used to prove Theorem 4.2. These concepts are implemented in Chapter VI, where various numerical experiments are outlined and the results tabulated. Finally, some concluding remarks, potential generalizations and possible future projects are discussed in Chapter VII.

CHAPTER II

PRELIMINARIES

II.1 THE PARABOLIC PROBLEM

We consider solving the standard parabolic problem of finding u such that

$$u_t(x,t) - \Delta u(x,t) = f(x,t), \quad x \in \Omega, \quad t > 0, \tag{1}$$

$$u(x,t) = 0, \quad x \in \partial\Omega, \quad t > 0,$$

$$u(x,0) = u_0(x), \quad x \in \Omega,$$

where Ω is a domain in $\mathbf{R}^d$ with smooth boundary $\partial\Omega$, u_t denotes $\partial u/\partial t$, $\Delta := \sum_{j=1}^{d} \partial^2/\partial x_j^2$ is the Laplacian, and the functions f and u_0 are given data. For the spatial discretization of this problem with respect to the space variable $x := (x_1, x_2, \cdots, x_d)$, let Σ be the class of all finite element discretizations (h, T, S) satisfying the following conditions:

1. h is a positive function in $C^1(\bar{\Omega})$ such that $|\nabla h(x)| \leq \lambda$ for all $x \in \bar{\Omega}$ and for some $\lambda > 0$.

2. $T = \{\Omega_K\}$ is a set of closed triangular subdomains of Ω defining a partition of Ω into triangular elements Ω_K of diameter h_K such that

$$c_1 h_K^2 \leq \int_{\Omega_K} dx \tag{2}$$

   for all $\Omega_K \in T$, and associated with the function h through

$$c_2 h_K \leq h(x) \leq h_K \tag{3}$$

   for all $x \in \Omega_K$, $\Omega_K \in T$ where $c_1 > 0, c_2 > 0$ are positive constants.

3. S is the set of all continuous functions on $\overline{\Omega}$ which are polynomials of order r in x for $x \in \Omega_K$ for each $\Omega_K \in T$ and vanish on $\partial\Omega$.

We assume the triangulation is such that the intersection of any two closed triangular elements is either empty, a common face, or a common vertex of the two.

Next, we discuss the Discontinuous Galerkin Method, which will be used for the time discretization of (1).

II.2 THE DISCONTINUOUS GALERKIN METHOD

To introduce the Discontinuous Galerkin Method we utilize much of the discussion in [4]. First, we will write the parabolic problem (1) in its weak form by multiplying both sides of (1) by a function $w \in H_0^1(\Omega)$, that is, the functions w with $\nabla w = \text{grad } w$ in $L_2(\Omega)$ and which vanish on $\partial\Omega$, and integrate over Ω to obtain

$$(u_t, w) - (\Delta u, w) = (f, w), \quad t > 0,$$

where

$$(u, w) := \int_\Omega uw \; dx.$$

Using Green's Formula, given by

$$\int_\Omega (\Delta u) w \; dx = \int_{\partial\Omega} (\nabla u) w \cdot n \; ds - \int_\Omega (\nabla u \cdot \nabla w) \; dx,$$

with $\int_{\partial\Omega} (\nabla u) w \cdot n \; ds = 0$ due to the specified boundary conditions, we obtain the weak form

$$(u_t, w) + (\nabla u, \nabla w) = (f, w) \quad \text{for } w \in H_0^1(\Omega),$$

where

$$(\nabla u, \nabla w) = \int_\Omega \sum_{j=1}^{d} \frac{\partial u}{\partial x_j} \frac{\partial w}{\partial x_j} dx.$$

Next we integrate both sides of the last equation with respect to time t over a fixed interval $[0, t_N]$ to obtain the equation

$$\int_0^{t_N} \{(u_t, w) + (\nabla u, \nabla w)\} \; dt = \int_0^{t_N} (f, w) \; dt.$$

Note that the exact solution of the parabolic problem satisfies this last equation as well. Now, integration of the first term of the last equation by parts gives us

$$\int_0^{t_N} \{-(u, w_t) + (\nabla u, \nabla w)\} \; dt = (u_0, w(0)) + \int_0^{t_N} (f, w) \; dt, \tag{4}$$

where the assumption $w(t_N) = 0$ is made so the term $(u^N, w(t_N))$ in the integration by parts will vanish, per the procedure in [4], due to the eventual decay of $u(x, t)$ as $t \to \infty$. We discretize in time by partitioning the time interval *in a not necessarily uniform fashion* as

$$0 = t_0 < t_1 < t_2 < \cdots < t_N$$

and let

$$I_n := (t_{n-1}, t_n], \quad k_n := t_n - t_{n-1}$$

for $n = 1, \cdots, N$. Further, let $k := \max\{k_1, k_2, \cdots, k_N\}$. For a given positive integer q, we will be looking for an approximate solution to the weak form (4) of the parabolic problem (1) which reduces to a polynomial of degree at most q in t on each subinterval I_n with coefficients in $H_0^1(\Omega)$, or equivalently, a polynomial in the space

$$S_k := \{v : [0,\infty) \rightarrow H_0^1(\Omega); v|_{I_n} \in P_q(I_n), \quad n = 1, \cdots, N\}$$

where

$$P_q(I_n) := \{v(t) = \sum_{j=0}^{q} v_j t^j : v_j \in H_0^1(\Omega), \quad j = 0, \cdots, q\}.$$

Note that these functions are allowed to be discontinuous at the nodal points t_n but will be taken to be continuous from the left there. Further, note that $v(0)$ has to be specified separately for $v \in S_k$ since 0 is not in I_1, and we write S_k^n for the restrictions to I_n of the functions in S_k.

For notational convenience, we write

$$w^n := w(t_n), \quad w^{n,+} := \lim_{t \rightarrow t_n+} w(t), \quad w^{n,-} := \lim_{t \rightarrow t_n-} w(t)$$

for any function w.

Now, replace u in the weak formulation (4) by a function $U \in S_k$ and integrate by parts on each subinterval I_n to obtain for the first term of the left side of (4), with $v^n = v(t_n)$ and $v^N = 0$,

$$\begin{aligned} -\int_0^{t_N} (U, v_t)\, dt &= -\sum_{n=0}^{N} \{(U,v)|_{t_{n-1}}^{t_n} - \int_{I_n} (U_t, v)\, dt\} \\ &= \int_0^{t_N} (U_t, v)\, dt + \sum_{n=0}^{N-1} ([U]_n, v^n) + (U^{0,+}, v^0), \quad v \in S_k. \end{aligned} \tag{5}$$

Here $[U]_n := U^{n,+} - U^n$ denotes the jump of U at t_n and U_t is the piecewise polynomial of degree $n-1$ which agrees with dU/dt on each subinterval I_n. In particular, for the case $q = 0$, we have $U_t \equiv 0$, so the integrand vanishes.

With the first term of the left side of the weak formulation thus modified, the Discontinuous Galerkin Method is defined as follows: Find $U \in S_k$ such that

$$\int_0^{t_N} \{(U_t, v) + (\nabla u, \nabla v)\}\, dt + \sum_{n=1}^{N-1} ([U]^n, v^{n,+}) + (U^{0,+}, v^{0,+}) = (u_0, v^{0,+}) + \int_0^{t_N} (f, v)\, dt \tag{6}$$

for all $v \in S_k$.

Since a function v in S_k is *not required to be continuous* at t_n we may choose its values on the the different time intervals independently, and so by choosing v to

vanish outside the the time interval I_n we reduce (6) to one equation for each time interval I_n, as in [4]. This results in the the following problem: For $n = 1, 2, \cdots, N$, find $U^n \in S_k^n$ such that

$$\int_{I_n} \{(U_t^n, v) + (\nabla U, \nabla v)\}\, dt + ([U]^{n-1}, v^{n-1,+}) = \int_{I_n} (f, v)\, dt \tag{7}$$

for all $v \in S_k^n$, where $U^0 := u_0$ since t_0 is not in I_1. This shows that the discrete solution is independent of the choice of the final nodal point t_N. Further, it can be shown that the exact solution of (1) also satisfies (7).

For the spatial discretization, that is, discretization in the space $H_0^1(\Omega)$, let $M \in \{0, 1, 2, \cdots\}$ and x_m, $m = 0, 1, \cdots, 2^M$ denote the spatial knots. We will use linear splines on Ω, although splines of any order may be employed. At each time step we will approximate $u(x, t_n)$ by

$$U^n(x) = U(x, t_n) = \sum_{i=0}^{2^M} \xi_i^n(t)\phi_i(x), \quad n = 1, 2, \cdots, N.$$

For simplicity of notation, we write $U^n := U^n(x) = U(x, t_n)$. Next, let X_M be the finite dimensional subspace of $H_0^1(\Omega)$ spanned by these splines. Equation (7) may now be stated as follows: For $n = 1, 2, \cdots, N$, given $U^{n-1,-}$, find $U^n \in S_{Mk}^n$ where

$$S_{Mk}^n := \{v : [0, \infty) \to X_M; v|_{I_n} \in P_{Mq}(I_n), \quad n = 1, \cdots, N\}$$

with

$$P_{Mq}(I_n) := \{v(t) = \sum_{j=0}^{q} v_j t^j : v_j \in X_M, \quad j = 0, \cdots, q\}$$

such that

$$\int_{I_n} \{(U_t, v) + (\nabla U, \nabla v)\}\, dt + (U^{n-1,+}, v^{n-1,+}) = \int_{I_n} (f, v)\, dt + (U^{n-1,-}, v^{n-1,+}) \tag{8}$$

for all $v \in P_{Mq}(I_n)$, where $U^{0,-} = u_0$.

For the discretization of the space $H_0^1(\Omega)$ with $\overline{\Omega} = [0, 1]$, denote by $\phi_m(x)$ the spline over $\Omega_m = [x_{m-1}, x_{m+1}]$ for $m = 1, 2, \cdots, 2^M - 1$. Also, denote by $\phi_0(x)$, $\phi_{2^M}(x)$ the splines over $[x_0, x_1]$ and $[x_{2^M-1}, x_{2^M}]$, respectively. Let X_M be the space of these piecewise linear splines on $\Omega = (0, 1)$ with breakpoints $0 = x_0 < x_1 < \cdots < x_{2^M} = 1$ and $h_m := \max_{1 \le m \le 2^M} |x_{m-1} - x_m|$. It follows that X_M is a finite dimensional subspace of $H_0^1(\Omega)$.

From [7] we have the following a priori estimate for the Discontinuous Galerkin Method. Much of the following discussion is paraphrased from [6]. We will utilize this estimate in Chapter IV for the proof of Theorem 4.2.

Theorem 2.1. (Eriksson and Johnson [7]). *Let u be the solution of (1) and U_m that of (8). Assume that $X_{m-1} \subseteq X_m$ for all positive integers m and $k_n \leq \gamma k_{n+1}$ for all n and for some $\gamma > 0$. Then there exists a constant C depending only on c_1 and c_2 from (2) and (3), respectively from above, such that for $q = 0, 1,$ and $N = 1, 2, \cdots,$ we have*

$$\|u - U_m\|_{I_n} \leq C L_N \max_{1 \leq n \leq N} E_{m,qn}(u)$$

where

$$L_N = (\log(\frac{t_N}{k_N}) + 1)^{\frac{1}{2}}$$

and

$$E_{m,qn}(u) = \min_{j \leq q+1} k_n^j \|u_t^{(j)}\|_{I_n} + \|h_{m,n}^2 D^2 u\|_{I_n}$$

with $u_t^{(1)} = u_t$, $u_t^{(2)} = u_{tt}$ and $\|u\|_{I_n} = \max_{t \in I_n} \|u(t)\|_2$.

The term

$$\min_{j \leq q+1} k_n^j \|u_t^{(j)}\|_{I_n}$$

describes the error associated with the time discretization. If $\|u_t^{(j)}\|_{I_n}$ is bounded for each n and $j = 1, 2$, then the Discontinuous Galerkin Method is jth order accurate in time.

The term

$$\|h_{m,n}^2 D^2 u\|_{I_n}$$

describes the spatial discretization error and has second order due to the use of linear splines defining the space X_m. From before, we have $h_m := \max_{1 \leq m \leq 2^M} |x_{m-1} - x_m|$, but since the spatial grid mesh may be varied from time step to time step when the time steps are not uniformly spaced, we use of the double subscript on $h_{m,n}$ to denote this fact.

Next we look at some specific forms of equation (8). We will use the notation

$$[a_{ij}]_{m \times n}$$

to indicate the matrix consisting of m rows and n columns with individual entries a_{ij}, for $1 \leq i \leq m$, $1 \leq j \leq n$.

For the case $q = 0$, where $v(t)$ is piecewise constant in time, with

$$P_{M0}(I_n) := \{v(t) = v_0 : v_0 \in X_M\},$$

we have $\frac{d}{dt}U \equiv 0$ and $U^n = U^{n,-} = U^{n-1,+}$ so (8) reduces to the modified backward Euler Method

$$(U^n, v) + k_n(\nabla U^n, \nabla v) = \int_{I_n} (f, v)dt + (U^{n-1}, v), \quad v \in P_{M0}(I_n), \tag{9}$$

or

$$U^n - k_n \Delta U^n = U^{n-1} + \int_{I_n} f(t)\, dt$$

as in [4], page 183. The $f^n = f(t_n)$ occurring in the standard backward Euler Method, as detailed in [4], page 166, has been replaced by an average of f over the time interval I_n, resulting in the modified version. With $U^n = \sum_{i=0}^{2^M} \xi_i^n \phi_i(x)$ for scalars ξ_i^n, $i = 0, 1, 2, \cdots, 2^M$ and $n = 1, 2, \cdots, N$, equation (9) takes the form

$$\sum_{i=0}^{2^M} \xi_i^n [(\phi_i, \phi_j) + k_n(\nabla\phi_i, \nabla\phi_j)] = \int_{I_n} (f, \phi_j)dt + \sum_{i=0}^{2^M} \xi_i^{n-1}(\phi_i, \phi_j) \tag{10}$$

for $j = 0, 1, 2, \cdots, 2^M$, where ξ_i^{n-1} is known. This system of equations may be written in matrix form as

$$A_M^n \vec{u}_M^n = \vec{f}_M^n$$

where

$$A_M^n := [a_{ij}]_{(2^M+1)\times(2^M+1)},$$

$$a_{ij} := (\phi_i, \phi_j) + k_n(\nabla\phi_i, \nabla\phi_j),$$

$$\vec{u}_M^n := [\xi_i^n]_{(2^M+1)\times 1},$$

$$\vec{f}_M^n := [f_j]_{(2^M+1)\times 1},$$

$$f_j := \int_{I_n} (f, \phi_j)dt + \sum_{i=0}^{2^M} \xi_i^{n-1}(\phi_i, \phi_j).$$

For the case of $q = 1$, where $v(t)$ is piecewise linear in time, with

$$P_{M1}(I_n) := \{v(t) = v_0 + v_1 t : \ v_0, v_1 \in X_M\},$$

we have

$$U|_{I_n} = \overline{\phi}^n(x) + \frac{t - t_{n-1}}{k_n}\overline{\psi}^n(x)$$

on the interval I_n, and obtain the following system from (8):

$$\begin{aligned}
(\overline{\phi}^n, v) + k_n(\nabla\overline{\phi}^n, \nabla v) + (\overline{\psi}^n, v) + \frac{1}{2}k_n(\nabla\overline{\psi}^n, \nabla v) &= (U^{n-1}, v) + \int_{I_n}(f, v)dt, \\
\frac{1}{2}k_n(\nabla\overline{\phi}^n, \nabla w) + \frac{1}{2}(\overline{\psi}^n, w) + \frac{1}{3}k_n(\nabla\overline{\psi}^n, \nabla w) &= \frac{1}{k_n}\int_{I_n}(t - t_n)(f, w)dt,
\end{aligned}$$

for $v, w \in P_{M1}(I_n)$. With

$$\overline{\phi}^n(x) = \sum_{i=0}^{2^M} \xi_i^{\overline{\phi},n}\phi_i(x), \quad \overline{\psi}^n(x) = \sum_{i=0}^{2^M} \xi_i^{\overline{\psi},n}\phi_i(x),$$

for scalars $\xi_i^{\overline{\phi},n}$, $\xi_i^{\overline{\psi},n}$, where $i = 0, 1, 2, \cdots, 2^M$, $n = 1, 2, \cdots, N$, and

$$U^{n-1,+} = \overline{\phi}^n, \quad U^{n-1,-} = \overline{\phi}^{n-1} + \overline{\psi}^{n-1},$$

the last system takes the form

$$\sum_{i=0}^{2^M} \xi_i^{\overline{\phi},n}[(\phi_i, \phi_j) + k_n(\nabla\phi_i, \nabla\phi_j)] + \sum_{i=0}^{2^M} \xi_i^{\overline{\psi},n}[(\phi_i, \phi_j) + \frac{1}{2}k_n(\nabla\phi_i, \nabla\phi_j)]$$

$$= \int_{I_n}(f(t), \phi_j)dt + \sum_{i=0}^{2^M}[\xi_i^{\overline{\phi},n-1} + \xi_i^{\overline{\psi},n-1}](\phi_i, \phi_j), \quad j = 0, 1, \cdots, 2^M,$$

$$\sum_{i=0}^{2^M} \xi_i^{\overline{\phi},n}\frac{1}{2}k_n(\nabla\phi_i, \nabla\phi_j) + \sum_{i=0}^{2^M} \xi_i^{\overline{\psi},n}[\frac{1}{2}(\phi_i, \phi_j) + \frac{1}{3}k_n(\nabla\phi_i, \nabla\phi_j)]$$
$$= \frac{1}{k_n}\int_{I_n}(t - t_{n-1})(f(t), \phi_j)dt, \quad j = 0, 1, \cdots, 2^M.$$

This may be written in matrix form as

$$A_M^n \overrightarrow{\xi}_M^{\overline{\phi},n} + B_M^n \overrightarrow{\xi}_M^{\overline{\psi},n} = f_M^n$$

$$C_m^n \overrightarrow{\xi}_M^{\overline{\phi},n} + D_M^n \overrightarrow{\xi}_M^{\overline{\psi},n} = g_M^n$$

or

$$\begin{bmatrix} A_M^n & B_M^n \\ C_M^n & D_M^n \end{bmatrix} \begin{bmatrix} \overrightarrow{\xi}_M^{\overline{\phi},n} \\ \overrightarrow{\xi}_M^{\overline{\psi},n} \end{bmatrix} = \begin{bmatrix} \overrightarrow{f}_M^n \\ \overrightarrow{g}_M^n \end{bmatrix}$$

where

$$A_M^n := [a_{ij}]_{(2^M+1)\times(2^M+1)}, \quad a_{ij} := (\phi_i, \phi_j) + k_n(\nabla\phi_i, \nabla\phi_j),$$

$$B_M^n := [b_{ij}]_{(2^M+1)\times(2^M+1)}, \quad b_{ij} := (\phi_i, \phi_j) + \frac{1}{2}k_n(\nabla\phi_i, \nabla\phi_j),$$

$$C_M^n := [c_{ij}]_{(2^M+1)\times(2^M+1)}, \quad c_{ij} := \frac{1}{2}k_n(\nabla\phi_i, \nabla\phi_j),$$

$$D_M^n := [d_{ij}]_{(2^M+1)\times(2^M+1)}, \quad d_{ij} := \frac{1}{2}(\phi_i, \phi_j) + \frac{1}{3}k_n(\nabla\phi_i, \nabla\phi_j),$$

$$\overrightarrow{\xi}_M^{\overline{\phi},n} := \left[\xi_i^{\overline{\phi},n}\right]_{(2^M+1)\times 1}, \quad \overrightarrow{\xi}_M^{\overline{\psi},n} := \left[\xi_i^{\overline{\psi},n}\right]_{(2^M+1)\times 1},$$

$$\overrightarrow{f}_M^n := [f_j]_{(2^M+1)\times 1}, \quad f_j := \int_{I_n}(f, \phi_j)dt + \sum_{i=0}^{2^M}\left[\xi_i^{\overline{\phi},n-1} + \xi_i^{\overline{\psi},n-1}\right](\phi_i, \phi_j),$$

$$\overrightarrow{g}_M^n := [g_j]_{(2^M+1)\times 1}, \quad g_j := \frac{1}{k_n}\int_{I_n}(t - t_{n-1})(f, \phi_j)dt.$$

One advantage of the Discontinuous Galerkin Method is that the size of the time steps may be arbitrarily determined with no significant changes to the method, except possibly for a time-dependent change in the spatial mesh. This will be discussed later in Chapter V, when we look at parabolic problems with initial conditions that are incompatible with the prescribed homogeneous boundary conditions.

Next we discuss the multilevel calculation method, which is another essential part of this thesis. Most of the following information is taken from [3].

II.3 THE MULTILEVEL METHOD

II.3.1 Basics

To describe the general setup of the multilevel calculation method, we consider the basic operator equation

$$Au = f \tag{11}$$

where X and Y are Banach spaces, $A : X \to Y$ is a bounded linear operator, $f \in Y$ is assumed, and $u \in X$ is the assumed unique solution that is to be determined. We need two sequences $\{X_m\}$ and $\{Y_m\}$, $m \in M_0 = \{0, 1, 2, \cdots\}$ of *nested*, finite dimensional subspaces of X and Y, respectively, with

$$X_m \subseteq X_{m+1}, \quad m \in M_0, \quad \overline{\bigcup_{m\in M_0} X_m} = X,$$

$$Y_m \subseteq Y_{m+1}, \quad m \in M_0, \quad \overline{\bigcup_{m\in M_0} Y_m} = Y.$$

The nesting of these spaces implies there exists subspaces W_{m+1} of X_{m+1} and Z_{m+1} of Y_{m+1}, respectively, such that

$$X_{m+1} = X_m \oplus W_{m+1}, \quad Y_{m+1} = Y_m \oplus Z_{m+1}, \quad m \in M_0.$$

Further, we need $d_m := \dim(X_m) = \dim(Y_m)$, $m \in \{0, 1, 2, \cdots\}$. Now, assume the equation (11) has the approximate operator equation

$$A_m u_m = f_m \tag{12}$$

where $A_m : X_m \to Y_m$ is an approximate operator of A, $u_m \in X_m$ is the solution to (12) from X_m, and $f_m \in Y_m$ is an approximation of f. We identify the vector $[g_0, g_1]^T \in X_m \otimes W_{m+1}$ with the sum $g_0 + g_1 \in X_m \oplus W_{m+1}$. Likewise, we identify the vector $[g_0, g_1]^T \in Y_m \otimes Z_{m+1}$ with the sum $g_0 + g_1 \in Y_m \oplus Z_{m+1}$. With this notation, we describe the multilevel method for solving the operator equation (12) as a special case of the procedure detailed in [3]. With $m = k + 1$, the last equation takes the form

$$A_{k+1} u_{k+1} = f_{k+1}. \tag{13}$$

We write the solution $u_{k+1} \in X_{k+1}$ to this equation as

$$u_{k+1} = u_{k,0} + v_{k,1} \tag{14}$$

for $u_{k,0} \in X_k$ and $v_{k,1} \in W_{k+1}$. Note that u_{k+1} is identified with $u_k(1) := [u_{k,0}, v_{k,1}]^T$, per the notation of [3]. We refer to the solution of equation (13) as the $(k + 1)$ level solution. The basic idea of the multilevel method is to obtain an approximation of the $(k + 1)$ level solution from the kth level solution in X_k and a correction from W_{k+1}.

Now, define the operators $F_{k,k+1} : W_{k+1} \to Y_k$, $G_{k+1,k} : X_k \to Z_{k+1}$, and $H_{k+1,k+1} : W_{k+1} \to Z_{k+1}$, so the operator A_{k+1} is identified as the matrix of operators

$$A_{k,1} := \begin{bmatrix} A_k & F_{k,k+1} \\ G_{k+1,k} & H_{k+1,k+1} \end{bmatrix}. \tag{15}$$

Equation (13) is now equivalent to

$$A_{k,1} u_k(1) = f_{k+1}. \tag{16}$$

Now we split the operator $A_{k,1}$ into the sum of two operators $B_{k,1} : X_{k+1} \to Y_{k+1}$ and $C_{k,1} : X_{k+1} \to Y_{k+1}$, that is,

$$A_{k,1} = B_{k,1} + C_{k,1} \tag{17}$$

where

$$B_{k,1} := \begin{bmatrix} A_k & F_{k,k+1} \\ 0 & H_{k+1,k+1} \end{bmatrix} \tag{18}$$

and

$$C_{k,1} := \begin{bmatrix} 0 & 0 \\ G_{k+1,k} & 0 \end{bmatrix} \tag{19}$$

such that

$$A_{k,1} = \begin{bmatrix} A_k & F_{k,k+1} \\ G_{k+1,k} & H_{k+1,k+1} \end{bmatrix} = \begin{bmatrix} A_k & F_{k,k+1} \\ 0 & H_{k+1,k+1} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ G_{k+1,k} & 0 \end{bmatrix} = B_{k,1} + C_{k,1}.$$

Thus equation (16) becomes

$$B_{k,1}u_k(1) = f_{k+1} - C_{k,1}u_k(1). \tag{20}$$

Rather than solving equation (20) directly, we use the multilevel method detailed in the algorithm below to approximate the solution of (20). That is, we find an approximation $u_{k,1}$ to the approximate solution u_{k+1}.

Next we describe the algorithm for the multilevel method.

Multilevel Algorithm

Step 1 Solve the equation

$$A_k u_k = f_k \tag{21}$$

exactly, obtaining the kth level solution $u_k \in X_k$.

Step 2 Augment u_k by setting

$$\bar{u}_{k,1} = \begin{bmatrix} u_k \\ 0 \end{bmatrix}$$

and calculate the matrices $F_{k,k+1}$, $G_{k+1,k}$, and $H_{k+1,k+1}$.

Step 3 Solve $u_{k,1} \in X_{k+1}$ where

$$u_{k,1} := \begin{bmatrix} u_{k,0} \\ v_{k,1} \end{bmatrix}$$

from the equation

$$B_{k,1}u_{k,1} = f_{k+1} - C_{k,1}\bar{u}_{k,1}. \tag{22}$$

To be more specific, given the known solution u_k of $A_k u_k = f_k$, we solve the matrix equation

$$\begin{bmatrix} A_k & F_{k,k+1} \\ 0 & H_{k+1,k+1} \end{bmatrix} \begin{bmatrix} u_{k,0} \\ v_{k,1} \end{bmatrix} = \begin{bmatrix} f_k \\ g_k \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ G_{k+1,k} & 0 \end{bmatrix} \begin{bmatrix} u_k \\ 0 \end{bmatrix}$$

for $[u_{k,0}, v_{k,1}]^T$, where $f_{k+1} := [f_k, g_k]^T$. In practical terms, this means we first solve the system

$$H_{k+1,k+1} v_{k,1} = g_k - G_{k+1,k} u_k \tag{23}$$

for $v_{k,1}$, then use this solution to solve the system

$$A_k u_{k,0} = f_k - F_{k,k+1} v_{k,1} \tag{24}$$

for $u_{k,0}$. Then we set $u_{k,1} := u_{k,0} + v_{k,1}$, so $u_{k+1} \approx u_{k,1}$.

The multilevel method is basically a one step predictor-correction method to calculate an approximation to u_{k+1}, using $u_{k,1} = u_{k,0} + v_{k,1}$ as the approximation to u_{k+1}.

II.3.2 Error Estimate of the Multilevel Method

In this section we examine an error estimate for the multilevel method from [3]. Most of this discussion is paraphrased from [3]. For $m = 0, 1, 2, \cdots$, let E_m denote the approximation error in the space X_m for $u \in X$, namely,

$$E_m := \inf\{\|u - v\| : v \in X_m\}.$$

A sequence of nonnegative numbers γ_m, $m = 0, 1, 2, \cdots$, is called a majorization sequence of E_m if $\gamma_m \geq E_m$, $m = 0, 1, 2, \cdots$, and there exists a positive integer M_0 and a positive constant σ such that for $m \geq M_0$, $\frac{\gamma_{m+1}}{\gamma_m} \geq \sigma$.

The following theorem from [3] gives an error estimate for the multilevel method.

Theorem 2.2. (Chen, Wu and Xu [3]). *Suppose*

1. *There exists a positive integer M_0 and a positive constant α such that for $m \geq M_0$,*

$$\|A_m^{-1}\| \leq \alpha^{-1}.$$

2. *The limit*

$$\lim_{m \to \infty} \|C_{m,1}\| = 0$$

holds uniformly for $m = 1, 2, \cdots$.

3. *There exists a positive integer M_0 and a positive constant ρ such that for $m \geq M_0$ and for the solution u_m of equation* (12), *we have*

$$\|u - u_m\| \leq \rho E_m.$$

Let $u \in X$ be the solution of equation (11) and γ_m, $m = 0, 1, 2, \cdots$ a majorization sequence of E_m. Then there exists a positive integer M such that for $m \geq M$,

$$\|u - u_{m,1}\| \leq (\rho + 1)\gamma_{m+1}$$

where $u_{m,1}$ is the solution of (22).

Theorem 2.2 shows that under the assumptions listed, the multilevel solution $u_{m,1}$ approximates the exact solution u at an order comparable to E_{m+1}.

II.3.3 Cost Advantages

The advantage of the multilevel method is the cheaper computational cost incurred in solving several smaller size systems rather than a single system of a larger size. Specifically, to solve the system (13), we must solve a system of size d_{k+1} at an approximate cost of $O(d_{k+1}^3)$. Rather than do this, the multilevel method solves the system (21) of size d_k, obtaining the coarse level solution u_k. Then, using u_k, it solves the system (23) of size $d_{k+1} - d_k$, obtaining $v_{k,1}$. Finally, using $v_{k,1}$, it solves the system (24) of size d_k, obtaining $u_{k,0}$. Then it uses $u_{k,1} := u_{k,0} + v_{k,1}$ as an approximation to the approximate solution u_{k+1}. The cost of solving these systems is approximately $O(d_k^3) + O((d_{k+1} - d_k)^3)$. Even with more systems to solve, the smaller size of the systems will save computational time and effort, especially for high resolution level approximations.

We will discuss the specific savings in more detail in the application section, and then demonstrate these savings with the various numerical experiments in the implementation section.

For this method to work and provide good convergence characteristics, we need bases for X_k, Y_k, Z_k, and W_k, with multiresolutional capability, and for this we employ what we call the *wavelet basis*, which will be described next.

II.4 THE MULTISCALE ORTHONORMAL WAVELET BASIS

II.4.1 Generalities

Most of this section is paraphrased from [2]. Here we assemble the basic facts and structure of multiscale orthonormal wavelet bases for the Sobolev space $H_0^d(0,1)$ of functions u that satisfy the homogeneous boundary conditions

$$u^{(j)}(0) = u^{(j)}(1) = 0, \quad j \in Z_d, \tag{25}$$

where $Z_d := \{0, 1, 2, \cdots, d-1\}$ for a fixed positive integer d.

First, the given boundary conditions enable us to define the inner product as

$$\langle u, v\rangle_d = \int_0^1 u^{(d)}(x) v^{(d)}(x)\, dx, \quad u, v \in H_0^d(0,1)$$

and norm

$$|v|_d := \sqrt{\langle v, v\rangle_d}, \quad v \in H_0^d(0,1)$$

as in [2]. Let $k \geq 2d$ and $\mu > 1$ be fixed positive integers. For $m = 0, 1, 2, \cdots$, denote by X_m the subspace of $H_0^d(0,1)$ whose elements are piecewise polynomials of order k with knots j/μ^m, for $j - 1 \in Z_{\mu^m - 1}$. We have the property of nestedness of the subspaces, that is,

$$X_{m-1} \subset X_m$$

for $m = 1, 2, \cdots$. The dimension of X_m is

$$\dim X_m = (k - d)\mu^m - d.$$

Note that X_0 is the subspace of polynomials of order k satisfying the homogeneous boundary conditions (25), and when $k = 2d$, we have $X_0 = \{0\}$. When $k > 2d$, we have

$$X_0 = \operatorname{span}\{x^{d+j}(1-x)^d : j \in Z_{k-2d}\}.$$

Next, we will look at the orthogonal decomposition of the space X_m in the sense of the inner product $\langle \cdot, \cdot \rangle_d$. For notation, we let $S_1 \oplus S_2$ denote the direct sum of S_1 and S_2 with the property that for any $u \in S_1$, $v \in S_2$, we have $\langle u, v\rangle_d = 0$. Since $X_{m-1} \subset X_m$ for each m, let W_m be the orthogonal complement of X_{m-1} in X_m, that is,

$$X_m = X_{m-1} \oplus W_m.$$

This leads to

$$X_m = X_0 \oplus W_1 \oplus W_2 \oplus W_3 \oplus \cdots \oplus W_m.$$

The dimension of W_m is given by

$$w(m) := \dim\, W_m = \dim\, X_m - \dim\, X_{m-1} = (k-d)(\mu-1)\mu^{m-1}.$$

Once W_1 is determined, the spaces W_m can be constructed in a recursive fashion. To describe the construction, we use the family of affine mappings $\Phi_\mu := \{\phi_e : e \in Z_\mu\}$ with

$$\phi_e(x) := \frac{x+e}{\mu}, \quad e \in Z_\mu.$$

These mappings subdivide $[0,1]$ into the necessary subintervals associated with each space X_m. Associated with these affine mappings we define the family of operators $T_e : L^2[0,1] \to L^2[0,1]$, $e \in Z_\mu$, by

$$T_e v := \mu^{\frac{1}{2}-d} v \circ \phi_e^{-1} \chi_{\varphi_e[0,1]}, \quad e \in Z_\mu.$$

Next is the first of several lemmas from [2] which will help develop the structure of these bases. The first lemma shows the operators $T_e,\ e \in Z_\mu$, to be isometric from $H_0^d(0,1)$ to $H_0^d(0,1)$.

Lemma 2.3. (Chen, Wu, and Xu [2]).

(i) *For* $e \in Z_\mu, T_e$ *maps* $H_0^d(0,1)$ *into* $H_0^d(0,1)$.

(ii) *If* $e, e' \in Z_\mu$, *then for all* $u, v \in H_0^d(0,1)$,

$$\langle T_e u, T_{e'} v\rangle_d = \delta_{e,e'} \langle u, v\rangle_d.$$

Repeated differentiation of $T_e v$ results in the first statement. For the second result, when $e \neq e'$, the intersection of the support of $T_e u$ and that of $T_{e'} v$ has measure zero, so $\langle T_e u, T_{e'} v\rangle_d = 0$. For $e = e'$, using the definition of the operator and the fact that ϕ_e is affine, we have, with a change of variable,

$$\begin{aligned}\langle T_e u, T_e v\rangle_d &= \mu^{1-2d} \int_{\phi_e[0,1]} (u \circ \phi_e^{-1})^{(d)}(x)(v \circ \phi_e^{-1})^{(d)}(x)\, dx \\ &= \int_0^1 u^{(d)}(x) v^{(d)}(x)\, dx = \langle u, v\rangle_d.\end{aligned}$$

The above lemma provides a useful tool which we will utilize later as we set up the stiffness matrix in the application. As it will be necessary to compose the mapping

ϕ_e and the operator T_e for $e \in Z_\mu$ repeatedly, we need the composition mapping which we define next. For $\overrightarrow{e} := (e_0, e_1, \cdots, e_{m-1}) \in Z_\mu^m$, we define the composite map $\phi_{\overrightarrow{e}}$ to be

$$\phi_{\overrightarrow{e}} := \phi_{e_0} \circ \phi_{e_1} \circ \phi_{e_2} \circ \cdots \circ \phi_{e_{m-1}},$$

and the composite operator $T_{\overrightarrow{e}}$ as

$$T_{\overrightarrow{e}} := T_{e_0} \circ T_{e_1} \circ \cdots \circ T_{e_{m-1}}.$$

One can show, using successive compositions of the operators T_{e_i}, $i \in Z_m$ that for $v \in L^2[0,1]$,

$$T_{\overrightarrow{e}} v = \mu^{m(\frac{1}{2}-d)} v \circ \phi_{\overrightarrow{e}}^{-1} \chi_{\phi_{\overrightarrow{e}}[0,1]}.$$

It is the repeated composing of the operator T_e that will produce the required resolution for the problem at hand.

Integration by parts and Hermite interpolatory polynomials result in the next lemma.

Lemma 2.4. (Chen, Wu, and Xu [2]). *If i and j are positive integers with $i \leq j$, $w \in W_1$, $\overrightarrow{e} \in Z_\mu^j$, and $v \in X_i$, then $\langle w, v \circ \phi_{\overrightarrow{e}} \rangle_d = 0$.*

By the definition of the operator $T_{\overrightarrow{e}}$ and the fact that

$$(w \circ \phi_{\overrightarrow{e}}^{-1})^{(d)} = \mu^{di} w^{(d)} \circ \phi_{\overrightarrow{e}}^{-1},$$

we have, using a change of variable, the next lemma.

Lemma 2.5. (Chen, Wu, and Xu [2]). *If i is a positive integer, $\overrightarrow{e} \in Z_\mu^i$, $w \in W_1$, and $v \in X_i$, then $\langle T_{\overrightarrow{e}} w, v \rangle_d = 0$.*

Now, it is absolutely critical that a precise yet simple notational system be used to denote these various wavelets on the various resolution levels, and [2] provides the perfect system for describing the multiscale orthonormal bases for the spaces W_i. To describe this process of recursive construction, we start with the basis $w_{1\,l}$, $l \in Z_r$, where W_1 is given. For $i > 1$ and $\overrightarrow{e} \in Z_\mu^{i-1}$, we set

$$\mu(\overrightarrow{e}) := \mu^{i-2} e_0 + \cdots + \mu e_{i-3} + e_{i-2}.$$

Let $r := w(1) = \dim W_1$. For $i > 1$, $j \in Z_{w(i)}$, there exists the unique factorization

$$j = \mu(\vec{e})r + l, \quad \vec{e} \in Z_\mu^{i-1},\ l \in Z_r,$$

and we define

$$w_{i\ j} := T_{\vec{e}} w_{1\ l}.$$

To better understand this notational system, we will set up notation for the multiscale wavelet basis of the Sobolev space $H_0^2(0,1)$ with $k = 4$, and $\mu = 2$. We have $X_0 = \{0\}$, $w(i) = \dim W_i = 2^i$, $i > 0$, and the orthonormal basis for W_1 will be $\{w_{1\ 0}, w_{1\ 1}\}$ from [2] where

$$w_{1\ 0}(x) = \begin{cases} \frac{1}{2\sqrt{3}} x^2(3-4x) & \text{when } 0 \le x < \frac{1}{2}, \\ \frac{1}{2\sqrt{3}}(1-x)^2(4x-1) & \text{when } \frac{1}{2} \le x \le 1, \end{cases}$$

$$w_{1\ 1}(x) = \begin{cases} \frac{1}{2} x^2(1-2x) & \text{when } 0 \le x < \frac{1}{2}, \\ \frac{1}{2}(1-x)^2(1-2x) & \text{when } \frac{1}{2} \le x \le 1. \end{cases}$$

The specifics of just how such a basis is constructed will be detailed at the end of this section. Our purpose here is the familiarize the reader with the notation system being utilized.

With this set of basis functions, the affine maps

$$\phi_0(x) = \frac{1}{2}x, \quad \phi_1(x) = \frac{1}{2}x + \frac{1}{2},$$

and the operator

$$T_e v := \mu^{\frac{1}{2}-d} v \phi_e^{-1} \chi_{\varphi_e[0,1]}, \quad e = 0, 1,$$

we can recursively construct an orthogonal basis for whatever resolution level m we desire, using the formula

$$X_m = X_0 \oplus W_1 \oplus W_2 \oplus W_3 \oplus \cdots \oplus W_m.$$

Now for some specifics concerning the orthonormal basis $\{w_{10}, w_{11}\}$, where $k = 4$, $\mu = 2$, $r = 2$. To construct the wavelets for W_2, we need to use $i = 2$. We have

$$\vec{e} \in Z_\mu^{i-1} = Z_2^1 = \{e_0 : e_0 \in Z_2^1\} = \{0, 1\},$$

and

$$\mu(\vec{e}) = \mu^{i-2} e_0 = e_0.$$

For notational convenience, let $\overrightarrow{e}_0 = 0$ and $\overrightarrow{e}_1 = 1$ so $\mu(\overrightarrow{e}_0) = 0$ and $\mu(\overrightarrow{e}_1) = 1$. We have

$$j = \mu(\overrightarrow{e}) \cdot 2 + l, \; l \in Z_2 = \{0, 1\},$$

and

$$w_{ij} = T_{\overrightarrow{e}_k} w_{1\ l}.$$

For $\overrightarrow{e}_0 = 0$, $\mu(\overrightarrow{e}_0) = \mu(0) = 0$, $l = 0$, we have $j = 0 \cdot 2 + 0 = 0$ which implies

$$w_{2\ 0} = T_0 w_{1\ 0}.$$

For $\overrightarrow{e}_0 = 0$, $\mu(\overrightarrow{e}_0) = \mu(0) = 0$, $l = 1$, we have $j = 0 \cdot 2 + 1 = 1$ which implies

$$w_{2\ 1} = T_0 w_{1\ 1}.$$

For $\overrightarrow{e}_1 = 1$, $\mu(\overrightarrow{e}_1) = \mu(1) = 1$, $l = 0$, we have $j = 1 \cdot 2 + 0 = 2$ which implies

$$w_{2\ 2} = T_1 w_{1\ 0}.$$

For $\overrightarrow{e}_1 = 1$, $\mu(\overrightarrow{e}_1) = \mu(1) = 1$, $l = 1$, we have $j = 1 \cdot 2 + 1 = 3$ which implies

$$w_{2\ 3} = T_1 w_{1\ 1}.$$

To construct the wavelets for W_3, we need to use $i = 3$. We have

$$\overrightarrow{e} \in Z_\mu^{i-1} = Z_2^2 = \{(e_0, e_1) : e_0, e_1 \in Z_2\} = \{(0,0), (0,1), (1,0), (1,1)\}.$$

Let $\overrightarrow{e}_0 = (0,0)$, $\overrightarrow{e}_1 = (0,1)$, $\overrightarrow{e}_2 = (1,0)$, and $\overrightarrow{e}_3 = (1,1)$. We have

$$\mu(\overrightarrow{e}) = \mu^{i-2} e_0 + \mu^{i-3} e_1 = \mu e_0 + e_1 = 2e_0 + e_1$$

so

$$\mu(\overrightarrow{e}_0) = 2 \cdot 0 + 0 = 0, \quad \mu(\overrightarrow{e}_1) = 2 \cdot 0 + 1 = 1,$$
$$\mu(\overrightarrow{e}_2) = 2 \cdot 1 + 0 = 2, \quad \mu(\overrightarrow{e}_3) = 2 \cdot 1 + 1 = 3.$$

Also, for $j = 0, 1, 2$, we have

$$j = \mu(\overrightarrow{e}_k) \cdot r + l, \; \overrightarrow{e}_k \in Z_2^2, \; l \in Z_2, r = 2, \quad \text{and} \quad w_{ij} = T_{\overrightarrow{e}_k} w_{1\ l}.$$

This leads to the following subscript calculations:

$$0 = 0 \cdot 2 + 0, \quad w_{3\ 0} = T_{\overrightarrow{e}_0} w_{1\ 0},$$

$$1 = 0 \cdot 2 + 1, \quad w_{3\ 1} = T_{\overrightarrow{e}_0} w_{1\ 1},$$

$$2 = 1 \cdot 2 + 0, \quad w_{3\ 2} = T_{\overrightarrow{e}_1} w_{1\ 0},$$

$$3 = 1 \cdot 2 + 1, \quad w_{3\ 3} = T_{\overrightarrow{e}_1} w_{1\ 1},$$

$$4 = 2 \cdot 2 + 0, \quad w_{3\ 4} = T_{\overrightarrow{e}_2} w_{1\ 0},$$

$$5 = 2 \cdot 2 + 1, \quad w_{3\ 5} = T_{\overrightarrow{e}_2} w_{1\ 1},$$

$$6 = 3 \cdot 2 + 0, \quad w_{3\ 6} = T_{\overrightarrow{e}_3} w_{1\ 0},$$

$$7 = 3 \cdot 2 + 1, \quad w_{3\ 7} = T_{\overrightarrow{e}_3} w_{1\ 1}.$$

Next we construct the wavelets for W_4, where $i = 4$. Note that we continue to use $\mu = 2$ and $r = 2$, so this time we have

$$Z_{w(i)} = Z_{w(4)} = Z_4 = \{0, 1, 2, 3\}$$

as well as $Z_\mu = Z_2 = \{0, 1\}$. We use $Z_2^3 = \{(e_0, e_1, e_2) : e_k \in Z_2, k = 0, 1, 2\}$, which when written out, becomes

$$Z_2^3 = \{(0,0,0), (0,0,1), (0,1,0), (0,1,1), (1,0,0), (1,0,1), (1,1,0), (1,1,1)\}.$$

For convenience of notations and subscripts, we define

$$\overrightarrow{e}_0 = (0,0,0), \quad \overrightarrow{e}_1 = (0,0,1), \quad \overrightarrow{e}_2 = (0,1,0), \quad \overrightarrow{e}_3 = (0,1,1),$$

$$\overrightarrow{e}_4 = (1,0,0), \quad \overrightarrow{e}_5 = (1,0,1), \quad \overrightarrow{e}_6 = (1,1,0), \quad \overrightarrow{e}_7 = (1,1,1).$$

Since

$$\mu(\overrightarrow{e}) = \mu^{i-2} e_0 + \mu^{i-3} e_1 + \mu^{i-4} e_2 = \mu^2 e_0 + \mu e_1 + e_2 = 4e_0 + 2e_1 + e_2$$

we have, by the careful choice of subscripts,

$$\mu(\overrightarrow{e}_0) = 0, \quad \mu(\overrightarrow{e}_1) = 1, \quad \mu(\overrightarrow{e}_2) = 2, \quad \mu(\overrightarrow{e}_3) = 3,$$

$$\mu(\overrightarrow{e}_4) = 4, \quad \mu(\overrightarrow{e}_5) = 5, \quad \mu(\overrightarrow{e}_6) = 6, \quad \mu(\overrightarrow{e}_7) = 7.$$

Now, for $j = 0, 1, 2, 3$, $\mu(\overrightarrow{e}_k) = k$, $\ k = 0, 1, \cdots, 7$, and $r = 2$, where

$$j = \mu(\overrightarrow{e}_k) \cdot r + l, \ \overrightarrow{e}_k \in Z_2^3, \ l \in Z_2, r = 2, \quad \text{and} \quad w_{ij} = T_{\overrightarrow{e}_k} w_{1\ l},$$

we have the following subscript calculations:

$$0 = 0 \cdot 2 + 0, \quad w_{4\ 0} = T_{\overrightarrow{e}_0} w_{1\ 0},$$

$$
\begin{aligned}
1 &= 0 \cdot 2 + 1, & w_{4\ 1} &= T_{\vec{e}_0} w_{1\ 1}, \\
2 &= 1 \cdot 2 + 0, & w_{4\ 2} &= T_{\vec{e}_1} w_{1\ 0}, \\
3 &= 1 \cdot 2 + 1, & w_{4\ 3} &= T_{\vec{e}_1} w_{1\ 1}, \\
4 &= 2 \cdot 2 + 0, & w_{4\ 4} &= T_{\vec{e}_2} w_{1\ 0}, \\
5 &= 2 \cdot 2 + 1, & w_{4\ 5} &= T_{\vec{e}_2} w_{1\ 1}, \\
6 &= 3 \cdot 2 + 0, & w_{4\ 6} &= T_{\vec{e}_3} w_{1\ 0}, \\
7 &= 3 \cdot 2 + 1, & w_{4\ 7} &= T_{\vec{e}_3} w_{1\ 1}. \\
8 &= 4 \cdot 2 + 0, & w_{4\ 8} &= T_{\vec{e}_4} w_{1\ 0}, \\
9 &= 4 \cdot 2 + 1, & w_{4\ 9} &= T_{\vec{e}_4} w_{1\ 1}, \\
10 &= 5 \cdot 2 + 0, & w_{4\ 10} &= T_{\vec{e}_5} w_{1\ 0}, \\
11 &= 5 \cdot 2 + 1, & w_{4\ 11} &= T_{\vec{e}_5} w_{1\ 1}, \\
12 &= 6 \cdot 2 + 0, & w_{4\ 12} &= T_{\vec{e}_6} w_{1\ 0}, \\
13 &= 6 \cdot 2 + 1, & w_{4\ 13} &= T_{\vec{e}_6} w_{1\ 1}, \\
14 &= 7 \cdot 2 + 0, & w_{4\ 14} &= T_{\vec{e}_7} w_{1\ 0}, \\
15 &= 7 \cdot 2 + 1, & w_{4\ 15} &= T_{\vec{e}_7} w_{1\ 1}.
\end{aligned}
$$

One then may continue this process, progressively increasing the size of the index i, until the desired resolution level M is reached, and thus obtain the following spaces:

$$
\begin{aligned}
W_1 &= \text{span}\{w_{1\ 0}, w_{1\ 1}\}, \\
W_2 &= \text{span}\{w_{2\ 0}, w_{2\ 1}, w_{2\ 2}, w_{2\ 3}\}, \\
W_3 &= \text{span}\{w_{3\ 0}, w_{3\ 1}, w_{3\ 2}, w_{3\ 3}, w_{3\ 4}, w_{3\ 5}, w_{3\ 6}, w_{3\ 7}\}, \\
W_4 &= \text{span}\{w_{4\ 0}, w_{4\ 1}, w_{4\ 2}, w_{4\ 3}, w_{4\ 4}, w_{4\ 5}, w_{4\ 6}, w_{4\ 7}, \cdots, w_{4\ 15}\}, \\
W_5 &= \text{span}\{w_{5\ 0}, w_{5\ 1}, w_{5\ 2}, \cdots, w_{5\ 31}\}.
\end{aligned}
$$

and so on. In double subscripting system, the first subscript indicates the resolution level of that particular wavelet and the second subscript indicates which particular wavelet on that level. For instance, the double subscript 4 14 denotes the *fifteenth* wavelet on the *fourth* resolution level.

Hence we have

$$\begin{aligned}
X_0 &= \{0\}, \\
X_1 &= X_0 \oplus W_1, \\
X_2 &= X_1 \oplus W_2, \\
X_3 &= X_2 \oplus W_3, \\
X_4 &= X_3 \oplus W_4, \\
X_5 &= X_4 \oplus W_5. \\
\vdots &= \vdots \\
X_m &= X_{m-1} \oplus W_m.
\end{aligned}$$

Note in each case we have dim $X_m = (k-p)\mu^m - p$, where $k = 4$, $p = 2$, and $\mu = 2$.

The following theorem shows that the functions w_{ij} as defined above form an orthonormal basis for the space W_i. The proof is detailed in [2].

Theorem 2.6. (Chen, Wu, and Xu [2]). *Let w_{1j}, $j \in Z_r$, be an orthonormal basis of W_1. Then for any $i > 1$, the functions w_{ij}, $j \in Z_{w(i)}$ form an orthonormal basis for W_i and*

$$H_0^d(0,1) = \overline{X_0 \oplus W_1 \oplus W_2 \oplus \cdots}.$$

We will now give describe an algorithm for the construction of an orthonormal basis for the space W_1. Let Π_k be the space of polynomials of order k on the interval $[0,1]$. We will need the following lemma from [2].

Lemma 2.7 (Chen, Wu, and Xu [2]). *For any $v \in W_1$, v is orthogonal to the space Π_k.*

The following theorem gives an algorithm for the generation of the basis of W_1. Again, the proof is detailed in [2].

Theorem 2.8. (Chen, Wu, and Xu [2]). *A function $v \in W_1$ if and only if*

(i) *v is a piecewise polynomial of degree less than k with knots $\{\frac{j}{\mu} : j-1 \in Z_{\mu-1}\}$,*

(ii) $v^{(i)}$, $i \in Z_d$ *are continuous at the knots* $\{\frac{j}{\mu} : j-1 \in Z_{\mu-1}\}$,

(iii) $v^{(i)}(0) = v^{(i)}(1) = 0, \quad i \in Z_d,$

(iv) *For* $p_j(x) := x^j, \quad \langle v^{(d)}, p_j \rangle_d = 0, \quad j \in Z_{k-d} \backslash Z_d.$

Theorem 2.8 gives us the following method for generating the basis of W_1. By condition (i), any function $v \in W_1$ has a representation of the form

$$v(x) = \sum_{j=0}^{k-1} a_{ij} x^j, \quad t \in [\frac{i+1}{\mu}, \frac{i}{\mu}), \quad i \in Z_\mu.$$

Conditions (ii)-(iv) impose $k + d(\mu - 1)$ restrictions on the coefficients a_{ij}, thus we obtain a homogeneous linear system of equations consisting of $k + d(\mu - 1)$ equations with $k\mu$ unknowns a_{ij}, $i \in Z_\mu$, $j \in Z_k$. The dimension of the solution space is not less than $(k-d)(\mu-1)$. Note that $\dim(W_1) = (k-d)(\mu-1)$, thus the solution space has exact dimension $(k-d)(\mu-1)$. Accordingly, an orthonormal basis of W_1 can be obtained from a solution of the linear system by orthogonalization and normalization.

Next we detail an important example of this process, which will be utilized later in this thesis.

II.4.2 A Linear Spatial Basis

In this section we discuss a linear basis of the space $H_0^1(0,1)$. We choose $k = 2$, and $\mu = 2$. Here the space $X_0 = \{0\}$ because there is no nontrivial linear polynomial which vanishes at both 0 and 1. Further, $\dim W_i = 2^{i-1}$, for $i > 0$. The basis of W_1 is given by

$$w_{1\ 0}(x) = \begin{cases} x & \text{when } 0 \le x < \frac{1}{2}, \\ 1-x & \text{when } \frac{1}{2} \le x \le 1. \end{cases}$$

We will call $w_{1\ 0}$ the *mother wavelet.* Figure 1 shows the plot of this mother wavelet.

The mother wavelet then produces two wavelets $w_{2\ 0}, w_{2\ 1}$ via the operators

$$w_{2\ 0} = T_0 w_{1\ 0}, \quad w_{2\ 1} = T_1 w_{1\ 0},$$

where

$$w_{2\ 0}(x) = \begin{cases} \frac{1}{\sqrt{2}} 2x & \text{when } 0 \le x < \frac{1}{4}, \\ \frac{1}{\sqrt{2}}(1-2x) & \text{when } \frac{1}{4} \le x \le \frac{1}{2}, \\ 0 & \text{otherwise,} \end{cases}$$

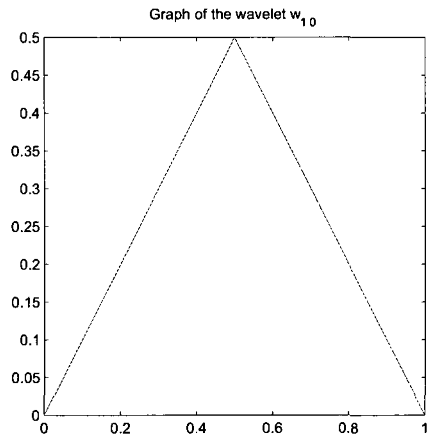

FIG. 1: The Mother Wavelet

$$w_{2\ 1}(x) = \begin{cases} \frac{1}{\sqrt{2}}(2x-1) & \text{when } \frac{1}{2} \leq x < \frac{3}{4}, \\ \frac{1}{\sqrt{2}}(2-2x) & \text{when } \frac{3}{4} \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

The graphs of these two wavelets are shown together in Figure 2. For simplicity, only the nonzero parts of each wavelet are plotted.

Thus we have

$$W_2 := \text{span}\{w_{2\ 0}, w_{2\ 1}\}.$$

Next the mother wavelet produces four wavelets $w_{3\ 0}, w_{3\ 1}, w_{3\ 2}, w_{3\ 3}$ by calculating

$$w_{3\ 0} = T_{(0,0)}w_{1\ 0},\ \ w_{3\ 1} = T_{(0,1)}w_{1\ 0},\ \ w_{3\ 2} = T_{(1,0)}w_{1\ 0},\ \ w_{3\ 3} = T_{(1,1)}w_{1\ 0},$$

where

$$w_{3\ 0}(x) = \begin{cases} \frac{1}{2}4x & \text{when } 0 \leq x < \frac{1}{8}, \\ \frac{1}{2}(1-4x) & \text{when } \frac{1}{8} \leq x \leq \frac{1}{4}, \\ 0 & \text{otherwise,} \end{cases}$$

$$w_{3\ 1}(x) = \begin{cases} \frac{1}{2}(4x-1) & \text{when } \frac{1}{4} \leq x < \frac{3}{8}, \\ \frac{1}{2}(2-4x) & \text{when } \frac{3}{8} \leq x \leq \frac{1}{2}, \\ 0 & \text{otherwise,} \end{cases}$$

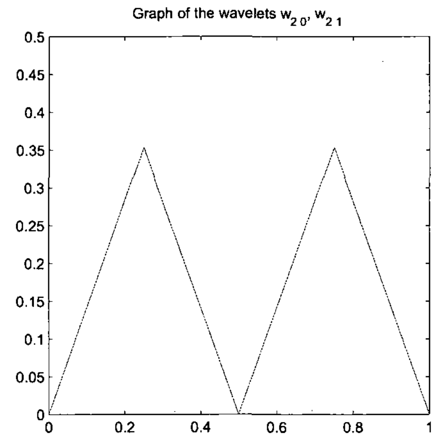

FIG. 2: Two Wavelets

$$w_{3\ 2}(x) = \begin{cases} \frac{1}{2}(4x-2) & \text{when } \frac{1}{2} \le x < \frac{5}{8}, \\ \frac{1}{2}(3-4x) & \text{when } \frac{5}{8} \le x \le \frac{3}{4}, \\ 0 & \text{otherwise,} \end{cases}$$

$$w_{3\ 3}(x) = \begin{cases} \frac{1}{2}(4x-3) & \text{when } \frac{3}{4} \le x < \frac{7}{8}, \\ \frac{1}{2}(4-4x) & \text{when } \frac{7}{8} \le x \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

The graphs of these four wavelets are shown together in Figure 3.

Thus we have

$$W_3 := \operatorname{span}\{w_{3\ 0}, w_{3\ 1}, w_{3\ 2}, w_{3\ 3}\}.$$

Next the mother wavelet produces eight wavelets,

$$w_{4\ 0},\ w_{4\ 1},\ w_{4\ 2},\ w_{4\ 3},\ w_{4\ 4},\ w_{4\ 5},\ w_{4\ 6},\ w_{4\ 7},$$

by calculating

$$w_{4\ 0} = T_{(0,0,0)}w_{1\ 0},\ w_{4\ 1} = T_{(0,0,1)}w_{1\ 0},\ w_{4\ 2} = T_{(0,1,0)}w_{1\ 0},\ w_{4\ 3} = T_{(0,1,1)}w_{1\ 0},$$

$$w_{4\ 4} = T_{(1,0,0)}w_{1\ 0},\ w_{4\ 5} = T_{(1,0,1)}w_{1\ 0},\ w_{4\ 6} = T_{(1,1,0)}w_{1\ 0},\ w_{4\ 7} = T_{(1,1,1)}w_{1\ 0},$$

where

$$w_{4\ 0}(x) = \begin{cases} \frac{1}{2\sqrt{2}}8x & \text{when } 0 \le x < \frac{1}{16}, \\ \frac{1}{2\sqrt{2}}(1-8x) & \text{when } \frac{1}{16} \le x \le \frac{1}{8}, \\ 0 & \text{otherwise,} \end{cases}$$

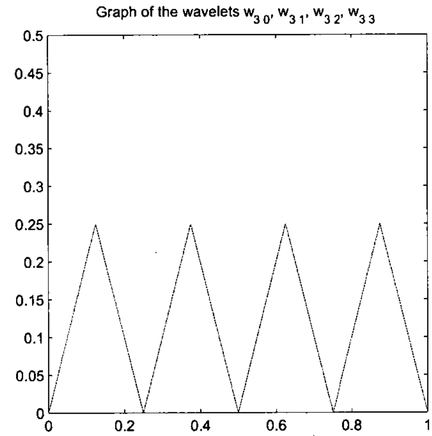

FIG. 3: Four Wavelets

$$w_{4\ 1}(x) = \begin{cases} \frac{1}{2\sqrt{2}}(8x-1) & \text{when } \frac{1}{8} \le x < \frac{3}{16}, \\ \frac{1}{2\sqrt{2}}(2-8x) & \text{when } \frac{3}{16} \le x \le \frac{1}{4}, \\ 0 & \text{otherwise,} \end{cases}$$

$$w_{4\ 2}(x) = \begin{cases} \frac{1}{2\sqrt{2}}(8x-2) & \text{when } \frac{1}{4} \le x < \frac{5}{16}, \\ \frac{1}{2\sqrt{2}}(3-8x) & \text{when } \frac{5}{16} \le x \le \frac{3}{8}, \\ 0 & \text{otherwise,} \end{cases}$$

$$w_{4\ 3}(x) = \begin{cases} \frac{1}{2\sqrt{2}}(8x-3) & \text{when } \frac{3}{8} \le x < \frac{7}{16}, \\ \frac{1}{2\sqrt{2}}(4-8x) & \text{when } \frac{7}{16} \le x \le \frac{1}{2}, \\ 0 & \text{otherwise,} \end{cases}$$

$$w_{4\ 4}(x) = \begin{cases} \frac{1}{2\sqrt{2}}(8x-4) & \text{when } \frac{1}{2} \le x < \frac{9}{16}, \\ \frac{1}{2\sqrt{2}}(5-8x) & \text{when } \frac{9}{16} \le x \le \frac{5}{8}, \\ 0 & \text{otherwise,} \end{cases}$$

$$w_{4\ 5}(x) = \begin{cases} \frac{1}{2\sqrt{2}}(8x-5) & \text{when } \frac{5}{8} \le x < \frac{11}{16}, \\ \frac{1}{2\sqrt{2}}(6-8x) & \text{when } \frac{11}{16} \le x \le \frac{3}{4}, \\ 0 & \text{otherwise,} \end{cases}$$

$$w_{4\ 6}(x) = \begin{cases} \frac{1}{2\sqrt{2}}(8x-6) & \text{when } \frac{3}{4} \le x < \frac{13}{16}, \\ \frac{1}{2\sqrt{2}}(7-8x) & \text{when } \frac{13}{16} \le x \le \frac{7}{8}, \\ 0 & \text{otherwise,} \end{cases}$$

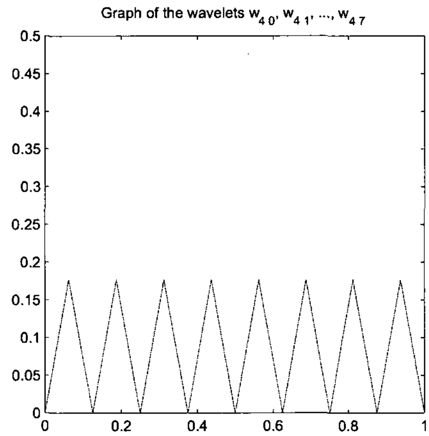

FIG. 4: Eight Wavelets

$$w_{4\ 7}(x) = \begin{cases} \frac{1}{2\sqrt{2}}(8x-7) & \text{when } \frac{7}{8} \le x < \frac{15}{16}, \\ \frac{1}{2\sqrt{2}}(8-8x) & \text{when } \frac{15}{16} \le x \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

The graphs of these eight wavelets are shown in Figure 4.

Thus we have

$$W_4 := \operatorname{span}\{w_{4\ 0}, w_{4\ 1}, w_{4\ 2}, w_{4\ 3}, w_{4\ 4}, w_{4\ 5}, w_{4\ 6}, w_{4\ 7}\}.$$

Next the mother wavelet produces sixteen, wavelets

$$w_{5\ 0}, w_{5\ 1}, w_{5\ 2}, w_{5\ 3}, w_{5\ 4}, w_{5\ 5}, w_{5\ 6}, w_{5\ 7},$$

$$w_{5\ 8}, w_{5\ 9}, w_{5\ 10}, w_{5\ 11}, w_{5\ 12}, w_{5\ 13}, w_{5\ 14}, w_{5\ 15},$$

by calculating

$$w_{5\ 0} = T_{(0,0,0,0)}w_{1\ 0},\ w_{5\ 1} = T_{(0,0,0,1)}w_{1\ 0},\ w_{5\ 2} = T_{(0,0,1,0)}w_{1\ 0},\ w_{5\ 3} = T_{(0,0,1,1)}w_{1\ 0},$$

$$w_{5\ 4} = T_{(0,1,0,0)}w_{1\ 0},\ w_{5\ 5} = T_{(0,1,0,1)}w_{1\ 0},\ w_{5\ 6} = T_{(0,1,1,0)}w_{1\ 0},\ w_{5\ 7} = T_{(0,1,1,1)}w_{1\ 0},$$

$$w_{5\ 8} = T_{(1,0,0,0)}w_{1\ 0},\ w_{5\ 9} = T_{(1,0,0,1)}w_{1\ 0},\ w_{5\ 10} = T_{(1,0,1,0)}w_{1\ 0},\ w_{5\ 11} = T_{(1,0,1,1)}w_{1\ 0},$$

$$w_{5\ 12} = T_{(1,1,0,0)}w_{1\ 0},\ w_{5\ 13} = T_{(1,1,0,1)}w_{1\ 0},\ w_{5\ 14} = T_{(1,1,1,0)}w_{1\ 0},\ w_{5\ 15} = T_{(1,1,1,1)}w_{1\ 0},$$

where

$$w_{5\ 0}(x) = \begin{cases} \frac{1}{4}16x & \text{when } 0 \leq x < \frac{1}{32}, \\ \frac{1}{4}(1-16x) & \text{when } \frac{1}{32} \leq x \leq \frac{1}{16}, \\ 0 & \text{otherwise,} \end{cases}$$

$$w_{5\ 1}(x) = \begin{cases} \frac{1}{4}(16x-1) & \text{when } \frac{1}{16} \leq x < \frac{3}{32}, \\ \frac{1}{4}(2-16x) & \text{when } \frac{3}{32} \leq x \leq \frac{1}{8}, \\ 0 & \text{otherwise,} \end{cases}$$

$$w_{5\ 2}(x) = \begin{cases} \frac{1}{4}(16x-2) & \text{when } \frac{1}{8} \leq x < \frac{5}{32}, \\ \frac{1}{4}(3-16x) & \text{when } \frac{5}{32} \leq x \leq \frac{3}{16}, \\ 0 & \text{otherwise,} \end{cases}$$

$$w_{5\ 3}(x) = \begin{cases} \frac{1}{4}(16x-3) & \text{when } \frac{3}{16} \leq x < \frac{7}{32}, \\ \frac{1}{4}(4-16x) & \text{when } \frac{7}{32} \leq x \leq \frac{1}{4}, \\ 0 & \text{otherwise,} \end{cases}$$

$$w_{5\ 4}(x) = \begin{cases} \frac{1}{4}(16x-4) & \text{when } \frac{1}{4} \leq x < \frac{9}{32}, \\ \frac{1}{4}(5-16x) & \text{when } \frac{9}{32} \leq x \leq \frac{5}{16}, \\ 0 & \text{otherwise,} \end{cases}$$

$$w_{5\ 5}(x) = \begin{cases} \frac{1}{4}(16x-5) & \text{when } \frac{5}{16} \leq x < \frac{11}{32}, \\ \frac{1}{4}(6-16x) & \text{when } \frac{11}{32} \leq x \leq \frac{3}{8}, \\ 0 & \text{otherwise,} \end{cases}$$

$$w_{5\ 6}(x) = \begin{cases} \frac{1}{4}(16x-6) & \text{when } \frac{3}{8} \leq x < \frac{13}{32}, \\ \frac{1}{4}(7-16x) & \text{when } \frac{13}{32} \leq x \leq \frac{7}{16}, \\ 0 & \text{otherwise,} \end{cases}$$

$$w_{5\ 7}(x) = \begin{cases} \frac{1}{4}(16x-7) & \text{when } \frac{7}{16} \leq x < \frac{15}{32}, \\ \frac{1}{4}(8-16x) & \text{when } \frac{15}{32} \leq x \leq \frac{1}{2}, \\ 0 & \text{otherwise,} \end{cases}$$

$$w_{5\ 8}(x) = \begin{cases} \frac{1}{4}(16x-8) & \text{when } \frac{1}{2} \leq x < \frac{17}{32}, \\ \frac{1}{4}(9-16x) & \text{when } \frac{17}{32} \leq x \leq \frac{9}{16}, \\ 0 & \text{otherwise,} \end{cases}$$

$$w_{5\ 9}(x) = \begin{cases} \frac{1}{4}(16x-9) & \text{when } \frac{9}{16} \leq x < \frac{19}{32}, \\ \frac{1}{4}(10-16x) & \text{when } \frac{19}{32} \leq x \leq \frac{5}{8}, \\ 0 & \text{otherwise,} \end{cases}$$

$$w_{5\ 10}(x) = \begin{cases} \frac{1}{4}(16x - 10) & \text{when } \frac{5}{8} \le x < \frac{21}{32}, \\ \frac{1}{4}(11 - 16x) & \text{when } \frac{21}{32} \le x \le \frac{11}{16}, \\ 0 & \text{otherwise,} \end{cases}$$

$$w_{5\ 11}(x) = \begin{cases} \frac{1}{4}(16x - 11) & \text{when } \frac{11}{16} \le x < \frac{23}{32}, \\ \frac{1}{4}(12 - 16x) & \text{when } \frac{23}{32} \le x \le \frac{3}{4}, \\ 0 & \text{otherwise,} \end{cases}$$

$$w_{5\ 12}(x) = \begin{cases} \frac{1}{4}(16x - 12) & \text{when } \frac{3}{4} \le x < \frac{25}{32}, \\ \frac{1}{4}(13 - 16x) & \text{when } \frac{25}{32} \le x \le \frac{13}{16}, \\ 0 & \text{otherwise,} \end{cases}$$

$$w_{5\ 13}(x) = \begin{cases} \frac{1}{4}(16x - 13) & \text{when } \frac{13}{4} \le x < \frac{27}{32}, \\ \frac{1}{4}(14 - 16x) & \text{when } \frac{27}{32} \le x \le \frac{7}{8}, \\ 0 & \text{otherwise,} \end{cases}$$

$$w_{5\ 14}(x) = \begin{cases} \frac{1}{4}(16x - 14) & \text{when } \frac{7}{8} \le x < \frac{29}{32}, \\ \frac{1}{4}(15 - 16x) & \text{when } \frac{29}{32} \le x \le \frac{15}{16}, \\ 0 & \text{otherwise,} \end{cases}$$

$$w_{5\ 15}(x) = \begin{cases} \frac{1}{4}(16x - 15) & \text{when } \frac{15}{16} \le x < \frac{31}{32}, \\ \frac{1}{4}(16 - 16x) & \text{when } \frac{31}{32} \le x \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

The graphs of these sixteen wavelets are shown in Figure 5.

Thus we have

$$W_5 := \operatorname{span}\{w_{5\ 0}, w_{5\ 1}, w_{5\ 2}, w_{5\ 3}, w_{5\ 4}, w_{5\ 5}, w_{5\ 6}, w_{5\ 7},$$

$$w_{5\ 8}, w_{5\ 9}, w_{5\ 10}, w_{5\ 11}, w_{5\ 12}, w_{5\ 13}, w_{5\ 14}, w_{5\ 15}\}.$$

Figure 6 shows the nonzero parts of all 31 of these wavelets plotted together.

One may continue this procedure until the desired resolution level is obtained, and ultimately obtain, for any $m \in M$,

$$X_m = X_0 \oplus W_1 \oplus W_2 \oplus \cdots \oplus W_{m-1} \oplus W_m.$$

By Theorem 2.6, we have

$$H_0^d(0,1) = \overline{X_0 \oplus W_1 \oplus W_2 \oplus \cdots}.$$

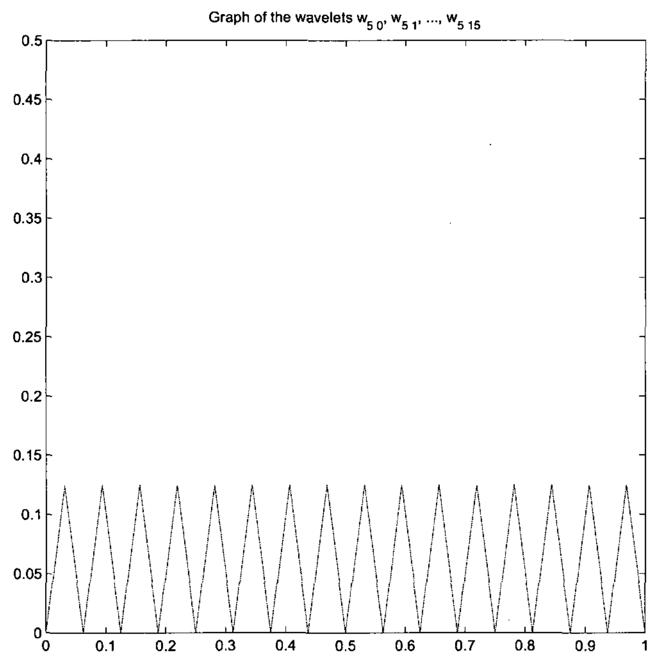

FIG. 5: Sixteen Wavelets

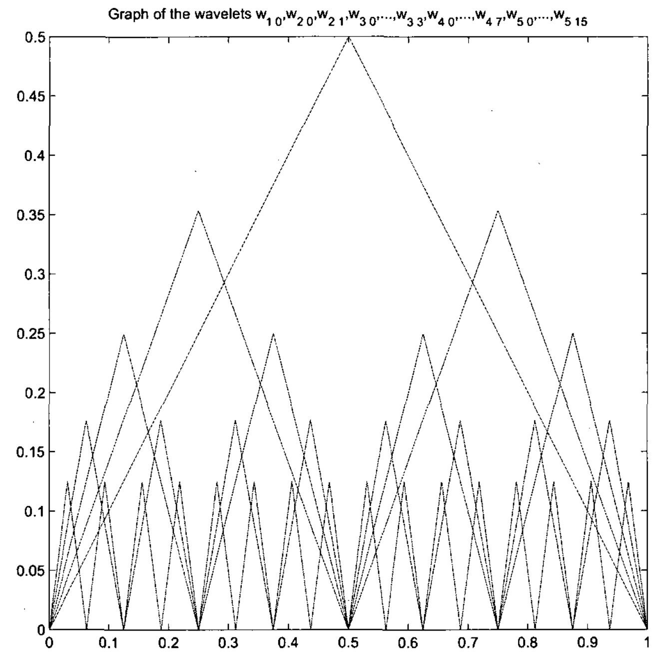

FIG. 6: Thirty-one Wavelets

CHAPTER III

APPLICATION TO THE PROBLEM

III.1 SETUP OF THE CONSTANT IN TIME CASE

In this section we detail the multilevel method and the wavelet basis functions as they apply to the constant in time case of the Discontinuous Galerkin Method. The linear in time case is *considerably* more complicated and involved, and it will be detailed later in the implementation section. For basis functions, we use the linear wavelet basis for $H_0^1(0,1)$ from [2] as constructed in the previous section with $X_0 = \{0\}$ and $W_1 = \text{span}\{w_{1\ 0}\}$. Further, we will use the notations

$$(w_{ij}, w_{i'j'}) := \int_0^1 w_{ij}(x) w_{i'j'}(x)\ dx, \quad a(w_{ij}, w_{i'j'}) := \int_0^1 \frac{d}{dx} w_{ij}(x) \frac{d}{dx} w_{i'j'}(x)\ dx.$$

Recall now that we are solving a system of the form

$$A_M^n \overrightarrow{u}_M^n = \overrightarrow{f}_M$$

where

$$A_M^n := [a_{iji'j'}]_{(2^M-1)\times(2^M-1)},$$

$$a_{iji'j'} := (w_{ij}, w_{i'j'}) + k_n a(w_{ij}, w_{i'j'}),$$

$$\overrightarrow{u}_M^n := [\xi_{i'j'}^n]_{(2^M-1)\times 1},$$

$$\overrightarrow{f}_M := [f_{i'j'}]_{(2^M-1)\times 1},$$

$$f_{i'j'} := \sum_{ij=1\ 0}^{M\ (2^{M-1}-1)} \xi_i^{n-1}(w_{ij}, w_{i'j'}) + \int_{I_n} (f, w_{i'j'}) dt.$$

for

$$ij = 1\ 0, 2\ 0, 2\ 1, 3\ 0, \cdots, 3\ 3, 4\ 0, \cdots, 4\ 7, \cdots, M\ 0, \cdots, M\ (2^{M-1}-1),$$

$$i'j' = 1\ 0, 2\ 0, 2\ 1, 3\ 0, \cdots, 3\ 3, 4\ 0, \cdots, 4\ 7, \cdots, M\ 0, \cdots, M\ (2^{M-1}-1),$$

and

$$X_M = \text{span}\{w_{ij} : ij = 1\,0, 2\,0, 2\,1, 3\,0, \cdots, 3\,3, 4\,0, \cdots, 4\,7, \cdots, M\,0, \cdots, M\,(2^{M-1}-1)\}.$$

For $M = 1$ we have

$$X_1 = X_0 \oplus_1 W_1 = W_1 = \text{span}\{w_{1\ 0}\}.$$

The system will have the form

$$A_1^n \overrightarrow{u}_1^n = \overrightarrow{f}_1$$

where

$$A_1^n = [a_{iji'j'}]_{1\times 1},$$
$$\overrightarrow{u}_1^n := [\xi_{i'j'}^n)]_{1\times 1},$$
$$\overrightarrow{f}_1 := [f_{i'j'}]_{1\times 1},$$

for $ij = 1\ 0$, and $i'j' = 1\ 0$. There is no multilevel decomposition for this resolution level; decomposition begins on the second level.

For $M = 2$ we have

$$X_2 = X_0 \oplus_1 W_1 \oplus_1 W_2 = W_1 \oplus_1 W_2 = \text{span}\{w_{1\ 0}, w_{2\ 0}, w_{2\ 1}\}$$

and the system will have the form

$$A_2^n \overrightarrow{u}_2^n = \overrightarrow{f}_2$$

where

$$A_2^n = [a_{iji'j'}]_{3\times 3},$$
$$\overrightarrow{u}_2^n := [\xi_{i'j'}^n]_{3\times 1},$$
$$\overrightarrow{f}_2 := [f_{i'j'}]_{3\times 1},$$

for $ij = 1\ 0, 2\ 0, 2\ 1$, and $i'j' = 1\ 0, 2\ 0, 2\ 1$.

For $M = 3$ we have

$$X_3 = X_0 \oplus_1 W_1 \oplus_1 W_2 \oplus_1 W_3 = W_1 \oplus_1 W_2 \oplus_1 W_3 = \text{span}\{w_{ij}\}$$

and the system will have the form

$$A_3^n \overrightarrow{u}_3^n = \overrightarrow{f}_3$$

where

$$A_3^n = [a_{iji'j'}]_{7\times 7},$$

$$\overrightarrow{u}_3^n := [\xi^n_{i'j'}]_{7\times 1},$$

$$\overrightarrow{f}_3 := [f_{i'j'}]_{7\times 1},$$

for $ij = 1\ 0, 2\ 0, 2\ 1, 3\ 0, 3\ 1, 3\ 2, 3\ 3$, and $i'j' = 1\ 0, 2\ 0, 2\ 1, 3\ 0, 3\ 1, 3\ 2, 3\ 3$.

For $M = 4$ we have

$$X_4 = X_0 \oplus_1 W_1 \oplus_1 W_2 \oplus_1 W_3 \oplus_1 W_4 = W_1 \oplus_1 W_2 \oplus_1 W_3 \oplus_1 W_4 = \text{span}\{w_{ij}\}$$

and the system will have the form

$$A_4^n \overrightarrow{u}_4^n = \overrightarrow{f}_4$$

where

$$A_4^n = [a_{iji'j'}]_{15\times 15},$$

$$\overrightarrow{u}_4^n := [\xi^n_{i'j'}]_{15\times 1},$$

$$\overrightarrow{f}_4 := [f_{i'j'}]_{15\times 1},$$

for

$$ij = 1\ 0, 2\ 0, 2\ 1, 3\ 0, 3\ 1, 3\ 2, 3\ 3, 4\ 0, 4\ 1, 4\ 2, 4\ 3, 4\ 4, 4\ 5, 4\ 6, 4\ 7,$$

and

$$i'j' = 1\ 0, 2\ 0, 2\ 1, 3\ 0, 3\ 1, 3\ 2, 3\ 3, 4\ 0, 4\ 1, 4\ 2, 4\ 3, 4\ 4, 4\ 5, 4\ 6, 4\ 7.$$

To generalize for the level M, we have

$$X_M = X_0 \oplus_1 W_1 \oplus_1 \cdots \oplus_1 W_M = W_1 \oplus_1 \cdots \oplus_1 W_M = \text{span}\{w_{ij}\}$$

and the system will have the form

$$A_M^n \overrightarrow{u}_M^n = \overrightarrow{f}_M \tag{26}$$

where

$$A_M^n = [a_{iji'j'}]_{(2^M-1)\times(2^M-1)},$$

$$\overrightarrow{u}_M^n := [\xi^n_{i'j'}]_{(2^M-1)\times 1},$$

$$\overrightarrow{f}_M := [\overrightarrow{f}_{i'j'}]_{(2^M-1)\times 1},$$

for

$$ij = 1\ 0, 2\ 0, 2\ 1, 3\ 0, \cdots, 3\ 3, 4\ 0, \cdots, 4\ 7, 5\ 0, \cdots, 5\ 15, \cdots$$

$$\cdots, M\ 0, M\ 1, M\ 2, \cdots, M\ (2^{M-1}-1),$$

and

$$i'j' = 1\ 0, 2\ 0, 2\ 1, 3\ 0, \cdots, 3\ 3, 4\ 0, \cdots, 4\ 7, 5\ 0, \cdots, 5\ 15, \cdots$$
$$\cdots, M\ 0, M\ 1, M\ 2, \cdots, M\ (2^{M-1}-1).$$

To generalize for the level M+1, we write

$$X_{M+1} = X_0 \oplus_1 W_1 \oplus_1 \cdots \oplus_1 W_{M+1} = W_1 \oplus_1 \cdots \oplus_1 W_{M+1} = \text{span}\{w_{ij}\}$$

and the system will have the form

$$A^n_{M+1} \overrightarrow{u}^n_{M+1} = \overrightarrow{f}_{M+1}$$

where

$$A^n_{M+1} = [a_{iji'j'}]_{(2^{M+1}-1)\times(2^{M+1}-1)},$$
$$\overrightarrow{u}^n_{M+1} := [\xi^n_{i'j'}]_{(2^{M+1}-1)\times 1},$$
$$\overrightarrow{f}_{M+1} := [f_{i'j'}]_{(2^{M+1}-1)\times 1},$$

for

$$ij = 1\ 0, 2\ 0, 2\ 1, 3\ 0, \cdots, 3\ 3, 4\ 0, \cdots, 4\ 7, 5\ 0, \cdots, 5\ 15, \cdots$$
$$\cdots, (M+1)\ 0, (M+1)\ 1, (M+1)\ 2, \cdots, (M+1)\ (2^M-1),$$

and

$$i'j' = 1\ 0, 2\ 0, 2\ 1, 3\ 0, \cdots, 3\ 3, 4\ 0, \cdots, 4\ 7, 5\ 0, \cdots, 5\ 15, \cdots$$
$$\cdots, (M+1)\ 0, (M+1)\ 1, (M+1)\ 2, \cdots, (M+1)\ (2^M-1).$$

For the multilevel decomposition on the M+1 level we have

$$X_M = \text{span}\{w_{ij}\}, \quad ij = 1\ 0, \cdots, M\ 0, \cdots, M\ (2^{M-1}-1),$$
$$W_{M+1} = \text{span}\{w_{ij}\}, \quad ij = (M+1)\ 0, \cdots, (M+1)\ (2^M-1),$$

with

$$X_{M+1} = X_M \oplus_1 W_{M+1}.$$

We decompose the matrix A^n_{M+1} as

$$\begin{bmatrix} A^n_M & F^n_{M,M+1} \\ G^n_{M+1,M} & H^n_{M+1,M+1} \end{bmatrix}$$

where the matrices A^n_M, $F^n_{M,M+1}$, $G^n_{M+1,M}$, and $H^n_{M+1,M+1}$ are defined as

$$A^n_M = [a_{iji'j'}]_{(2^M-1)\times(2^M-1)},$$

for $ij = 10,\cdots,M(2^{M-1}-1),\ i'j' = 10,\cdots,M(2^{M-1}-1)$,

$$F^n_{M,M+1} = [a_{iji'j'}]_{(2^M-1)\times(2^M)},$$

for $ij = 10,\cdots,M(2^{M-1}-1),\ i'j' = (M+1)0,\cdots,(M+1)(2^M-1)$,

$$G^n_{M+1,M} = [a_{iji'j'}]_{(2^M)\times(2^M-1)},$$

for $ij = (M+1)0,\cdots,(M+1)(2^M-1),\ i'j' = 10,\cdots,M(2^{M-1}-1)$, and

$$H^n_{M+1,M+1} = [a_{iji'j'}]_{(2^M)\times(2^M)},$$

for $ij = (M+1)0,\cdots,(M+1)(2^M-1),\ i'j' = (M+1)0,\cdots,(M+1)(2^M-1)$.

Next we decompose $\overrightarrow{u}^n_{M+1}$ as

$$\overrightarrow{u}^n_{M+1} = \begin{bmatrix} \overrightarrow{u}^n_{M,0} \\ \overrightarrow{v}^n_{M,1} \end{bmatrix}$$

where

$$\begin{aligned} \overrightarrow{u}^n_{M,0} &:= [\xi^n_{i'j'}]_{(2^M-1)\times 1}, \\ \overrightarrow{v}^n_{M,1} &:= [\eta^n_{i'j'}]_{2^M\times 1}. \end{aligned}$$

Also we have $\overrightarrow{f}_{M+1}$ written as

$$\overrightarrow{f}_{M+1} = \begin{bmatrix} \overrightarrow{f}_M \\ \overrightarrow{g}_M \end{bmatrix}$$

where

$$\begin{aligned} \overrightarrow{f}_M &:= [f_{i'j'}]_{(2^M-1)\times 1}, \\ \overrightarrow{g}_M &:= [g_{i'j'}]_{2^M\times 1}, \end{aligned}$$

with

$$f_{i'j'} := \sum_{ij=10}^{M\ (2^{M-1}-1)} \xi^{n-1}_{ij}(w_{ij}, w_{i'j'}) + \sum_{ij=(M+1)\ 0}^{(M+1)\ (2^M-1)} \eta^{n-1}_{ij}(w_{ij}, w_{i'j'}) + \int_{I_n} (f, w_{i'j'})dt,$$

for $i'j' = 1\ 0, \cdots, M\ (2^{M-1}-1)$, and

$$g_{i'j'} := \sum_{ij=10}^{M\ (2^{M-1}-1)} \xi_{ij}^{n-1}(w_{ij}, w_{i'j'}) + \sum_{ij=(M+1)\ 0}^{(M+1)\ (2^{M}-1)} \eta_{ij}^{n-1}(w_{ij}, w_{i'j'}) + \int_{I_n} (f, w_{i'j'})dt,$$

for $i'j' = (M+1)\ 0, \cdots, (M+1)\ (2^M-1)$.

With these various arrays now defined, the system

$$A_{M+1}^n \overrightarrow{u}_{M+1}^n = \overrightarrow{f}_{M+1}$$

becomes the system

$$\begin{bmatrix} A_M^n & F_{M,M+1}^n \\ G_{M+1,M}^n & H_{M+1,M+1}^n \end{bmatrix} \begin{bmatrix} \overrightarrow{u}_{M,0}^n \\ \overrightarrow{v}_{M,1}^n \end{bmatrix} = \begin{bmatrix} \overrightarrow{f}_M \\ \overrightarrow{g}_M \end{bmatrix}$$

which may be written as

$$\begin{aligned} A_M^n \overrightarrow{u}_{M,0}^n + F_{M,M+1}^n \overrightarrow{v}_{M,1}^n &= \overrightarrow{f}_M, \\ G_{M+1,M}^n \overrightarrow{u}_{M,0}^n + H_{M+1,M+1}^n \overrightarrow{v}_{M,1}^n &= \overrightarrow{g}_M. \end{aligned}$$

To obtain the multilevel approximation $\overrightarrow{u}_{M,1}^n$ of the M+1 level numerical solution $\overrightarrow{u}_{M+1}$, we first solve the coarse grid problem

$$A_M^n \overrightarrow{u}_M^n = \overrightarrow{f}_M$$

obtaining the Mth level solution $\overrightarrow{u}_M^n$.

Next we solve the system

$$H_{M+1,M+1}^n v_{M,1}^n = \overrightarrow{g}_M - G_{M+1,M} \overrightarrow{u}_M^n$$

obtaining

$$\overrightarrow{v}_{M,1}^n = (H_{M+1,M+1}^n)^{-1}(\overrightarrow{g}_M - G_{M+1,M} \overrightarrow{u}_M^n).$$

With $\overrightarrow{v}_{M,1}^n$ now known, we solve the system

$$A_M^n \overrightarrow{u}_{M,0}^n = \overrightarrow{f}_M^n - F_{M,M+1}^n \overrightarrow{v}_{M,1}^n,$$

obtaining

$$\overrightarrow{u}_{M,0}^n = (A_M^n)^{-1}(\overrightarrow{f}_M - F_{M,M+1} \overrightarrow{v}_{M,0}^n).$$

Finally, we set

$$\overrightarrow{u}^n_{M,1} = \begin{bmatrix} \overrightarrow{u}^n_{M,0} \\ \overrightarrow{v}^n_{M,1} \end{bmatrix}$$

Using $\overrightarrow{u}^n_{M,1}$ as an approximation to $\overrightarrow{u}_{M+1}$, that is,

$$\overrightarrow{u}^n_{M+1} \approx \overrightarrow{u}^n_{M,1}.$$

where

$$\overrightarrow{u}^n_{M,1} = \begin{bmatrix} (\xi^n_{i'j'})_{(2^M-1)\times 1} \\ (\eta^n_{i'j'})_{2^M \times 1} \end{bmatrix},$$

we form the linear combination

$$u^n_{M+1} \approx u^n_{M,1} := \sum_{i'j'=10}^{M\ (2^{M-1}-1)} \xi^n_{i'j'} w_{i'j'} + \sum_{i'j'=(M+1)\ 0}^{(M+1)\ (2^M-1)} \eta^n_{i'j'} w_{i'j'}. \tag{27}$$

III.2 COMPUTATIONAL COSTS

III.2.1 Constant in Time Case

Now some specifics as to the advantages of using the multilevel method to approximate the solution of the linear system

$$A^n_{M+1} u^n_{M+1} = f_{M+1}, \tag{28}$$

when the approximating functions are constant in time.

Direct calculation of this system would require us to solve a linear system consisting of $2^{M+1} - 1$ equations and $2^{M+1} - 1$ unknowns resulting of a computational cost of $O((2^{3M+3})$.

Approximating the solution of this system via the *multilevel* method requires solving two systems consisting of $2^M - 1$ equations with $2^M - 1$ unknowns, at a cost of $O(2^{3M})$, and solving one system consisting of $(2^{M+1} - 1) - (2^M - 1)$ equations with $(2^{M+1} - 1) - (2^M - 1)$ unknowns at a cost of $O(2^{3M})$. It will be shown in the implementation section that the multilevel method, despite having more systems to solve, provides a considerable gain in computational efficiencies. This is due to the fact that these systems are of smaller dimensions than the single system of larger dimension used for the computation of the direct method. This gain becomes much more pronounced as the grid resolution M is increased to higher levels.

III.2.2 Linear in Time Case

For the the linear in time case, all of the systems have twice the size of the constant in time case so direct calculation of the system (28) results in a cost of approximately $O(2^{3M+4})$. The cost of using the multilevel method to solve (28) will be approximately $O(2^{3M+1})$ which we will demonstrate to be substantially less in the implementation section.

In the next chapter, we show that the Multilevel Method provides the same degree of accuracy as the standard Discontinuous Galerkin Method. Much of the following discussion, and many of the results are based on information which is taken from [3].

CHAPTER IV

ERROR ANALYSIS AND ESTIMATE

In this section we will look at the the conditions which will allow the multilevel method to be accurate. We will need either the two hypotheses

(I)

$$B_{m,1}^{-1} \text{ exists,}$$

(II)

$$B_{m,1}^{-1} C_{m,1} \text{ is uniformly bounded,}$$

or the two hypotheses

(III) There exists a positive integer M_0 and a positive constant α such that for $m \geq M_0$,

$$\left\| A_m^{-1} \right\| \leq \alpha^{-1},$$

(IV)

$$\lim_{m \to \infty} \| C_{m,1} \| = 0.$$

The reason for each pair of hypotheses is that the main result of this section, Theorem 4.2, can be proven using either hypotheses (I) and (II), or hypotheses (III) and (IV). The advantage to using hypotheses (I) and (II) is that, unlike hypotheses (III) and (IV), the operator A_m^{-1} is not required to be uniformly bounded, and in fact, may even have a norm that approaches infinity. This in fact occurs in Section VI.4. Further, for the version of the proof that utilizes hypotheses (III) and (IV), we need the following lemma from [3], which we state and prove next with additional details provided.

Lemma 4.1. (Chen, Wu, and Xu [3]). *Suppose that hypotheses (III) and (IV) are satisfied. Then there exists a positive integer $M \geq M_0$ such that for $m \geq M$, the*

equation

$$B_{m,1}u_{m,1} = f_{m+1} - C_{m,1}\bar{u}_{m,1}$$

has a unique solution $u_{m,1} \in X_{m+1}$.

Proof. By (17), we have

$$A_{m,1} = B_{m,1} + C_{m,1},$$

so

$$B_{m,1} = A_{m,1} - C_{m,1}.$$

Using hypothesis (III), if $m \geq M_0$, then for $x \in X_{m+1}$,

$$\begin{aligned} \|B_{m,1}x\| &= \|(A_{m,1} - C_{m,1})x\| \\ &\geq \|A_{m,1}x\| - \|C_{m,1}x\| \\ &\geq \alpha\|x\| - \|C_{m,1}\|\|x\| \\ &= (\alpha - \|C_{m,1}\|)\|x\|. \end{aligned}$$

Thus, for $y \in Y_{m+1}$, we have

$$\|y\| \geq (\alpha - \|C_{m,1}\|)\|B_{m,1}^{-1}y\|,$$

so

$$\|B_{m,1}^{-1}y\| \leq \frac{1}{\alpha - \|C_{m,1}\|}\|y\|.$$

Therefore

$$\|B_{m,1}^{-1}\| \leq \frac{1}{\alpha - \|C_{m,1}\|}. \tag{29}$$

But by hypothesis (IV), there exists a positive integer $M > M_0$ such that for $m \geq M$, we have $\| C_{m,1} \| < \alpha/2$. Combining this inequality with the inequality

$$\|B_{m,1}x\| \geq (\alpha - \|C_{m,1}\|)\|x\|$$

from above, we find that for $m \geq M$ we have

$$\|B_{m,1}x\| \geq (\alpha - \frac{\alpha}{2})\|x\| = \frac{\alpha}{2}\|x\|$$

so

$$\frac{\alpha}{2}\|x\| \leq \|B_{m,1}x\|$$

for $x \in X_{m+1}$, which implies

$$\|B_{m,1}^{-1}\| \leq 2\alpha^{-1}.$$

Thus for all $m \geq M$, the equation

$$B_{m,1}u_{m,1} = f_{m+1} - C_{m,1}\bar{u}_{m,1}$$

has a unique solution.

The next theorem, Theorem 4.2, is our main result. It shows that our method provides the same degree of accuracy as the conventional Discontinuous Galerkin Method. Although there are many similarities to Theorem 2.2, there are also several key differences. One of the most apparent is that, unlike Theorem 2.2, our result holds under two somewhat different sets of hypotheses. Further, the error bound of our result is based on the error bound provided by Theorem 2.1 for the Discontinuous Galerkin Method, rather than on the approximation error E_m in the space X_m that is used for Theorem 2.2. Throughout the calculations we use C to denote generic constants whose values change as they appear. Further, as stated, we prove both versions of the theorem, each written in an independent manner that does not rely on the other version for any steps or details.

Theorem 4.2. *Suppose that hypotheses (I) and (II) or (III) and (IV) are satisfied. Let $u \in X$ be the solution of equation (1), and $u_{m,1} \in X_m$ be the solution of equation (22) with $X_m =$ span$\{w_{ij}\}$ where $\{w_{ij}\}$ is the wavelet basis described in Section II.4. Then there exists a positive integer M such that for $m \geq M$ and $n = 1, 2, \cdots, N$, we have*

$$\|u - u_{m,1}\|_{I_n} \leq CL_n \max_{1\leq n\leq N} E_{m+1,qn}(u)$$

where

$$L_N = (\log(\frac{t_N}{k_N}) + 1)^{\frac{1}{2}},$$

$$E_{m,qn}(u) = \min_{j\leq q+1} k_n^j \|u_t^{(j)}\|_{I_N} + \|h_{m,n}^2 D^2 u\|_{I_N}$$

for $u_t^{(1)} = u_t$, $u_t^{(2)} = u_{tt}$, $\|u\|_{I_N} = \max_{t\in I_N} \|u(t)\|_2$, and $h_{m,n} := \max_{1\leq m\leq 2^M} |x_{m-1} - x_m|$ with

$$h_{m+1,n}^2 = \mu^{-2} h_{m,n}^2$$

due to the use of the wavelet basis.

Proof (Version One). Assume hypothesis (I) and (II) are satisfied. We prove this theorem by establishing an estimate on $\|u_{m,1} - U_{m+1}\|_{I_n}$. For this purpose, we subtract (20),

$$B_{m,1}u_m(1) = f_{m+1} - C_{m,1}u_m(1)$$

from (22),

$$B_{m,1}u_{m,1} = f_{m+1} - C_{m,1}\overline{u}_{m,1}$$

to obtain

$$B_{m,1}(u_{m,1} - U_{m+1}) = C_{m,1}(U_{m+1} - \bar{u}_{m,1})$$

where U_{m+1} is identified with $u_m(1) := [u_{m,0}, v_{m,1}]^T$, per the notation of [3], $u_{m,1}$ is the solution of (22), and

$$\bar{u}_{m,1} = \begin{bmatrix} U_m \\ 0 \end{bmatrix},$$

as in Section II.3. Since hypotheses (I) is assumed, we have

$$(u_{m,1} - U_{m+1}) = B_{m,1}^{-1}C_{m,1}(U_{m+1} - \bar{u}_{m,1}).$$

Hypotheses (II) now implies

$$\|u_{m,1} - U_{m+1}\|_{I_n} \;\leq\; C\|U_{m+1} - \bar{u}_{m,1}\|_{I_n}. \tag{30}$$

Since $u_{m,0} := U_m$, we have

$$\|u_{m,0} - U_{m+0}\|_{I_n} = 0.$$

Now using the definition of $\bar{u}_{m,1}$, Theorem 2.1, and the triangle inequality, we obtain

$$\begin{aligned}
\|U_{m+1} - \bar{u}_{m,1}\|_{I_n} &\leq \|U_{m+1} - u\|_{I_n} + \|u - U_{m+0}\|_{I_n} + \|U_{m+0} - u_{m,0}\|_{I_n} \\
&\leq CL_n \max_{1\leq n\leq N} E_{m+1,qn}(u) + CL_n \max_{1\leq n\leq N} E_{m,qn}(u) + 0 \\
&= CL_n \max_{1\leq n\leq N} E_{m+1,qn}(u) + CL_n \max_{1\leq n\leq N} E_{m,qn}(u) \\
&\leq CL_n \max_{1\leq n\leq N} E_{m+1,qn}(u).
\end{aligned} \tag{31}$$

Note that since we are using the wavelet bases for the subspaces X_m and X_{m+1}, we have $\mu^2 h_{m+1,n}^2 = h_{m,n}^2$, which implies

$$L_n \max_{1\leq n\leq N} E_{m,qn}(u) \leq CL_n \max_{1\leq n\leq N} E_{m+1,qn}(u)$$

for some constant C. Substituting the estimate (31) into the right hand side of equation (30), we get

$$\|u_{m,1} - U_{m+1}\|_{I_n} \leq \; CL_n \max_{1\leq n\leq N} E_{m+1,qn}(u).$$

Now using Theorem 2.1 and the last inequality, there exists a positive integer M such that for $m \geq M$, and $n = 1, 2, \cdots, N$, we have

$$\begin{aligned}
\|u - u_{m,1}\|_{I_n} &\leq \; \|u - U_{m+1}\|_{I_n} + \|U_{m+1} - u_{m,1}\|_{I_n} \\
&\leq \; CL_n \max_{1\leq n\leq N} E_{m+1,qn}(u) + CL_n \max_{1\leq n\leq N} E_{m+1,qn}(u) \\
&\leq \; CL_n \max_{1\leq n\leq N} E_{m+1,qn}(u)
\end{aligned}$$

which completes the proof of this version of the theorem .

Proof (Version Two). Assume hypotheses (III) and (IV) are satisfied. This version, like the first, is done by establishing an estimate on $\|u_{m,1} - U_{m+1}\|_{I_n}$. For this purpose, we subtract (20),

$$B_{m,1}u_m(1) = f_{m+1} - C_{m,1}u_m(1)$$

from (22),

$$B_{m,1}u_{m,1} = f_{m+1} - C_{m,1}\overline{u}_{m,1}$$

to obtain

$$B_{m,1}(u_{m,1} - U_{m+1}) = C_{m,1}(U_{m+1} - \bar{u}_{m,1})$$

as we did in the proof of Version One above. Since hypotheses (III) and (IV) are assumed, Lemma 4.1 applies, so

$$(u_{m,1} - U_{m+1}) = B_{m,1}^{-1}C_{m,1}(U_{m+1} - \bar{u}_{m,1}).$$

Thus from the previous equation and the inequality (29),

$$\|B_{m,1}^{-1}\|_{I_n} \leq \frac{1}{\alpha - \|C_{m,1}\|_{I_n}}$$

from the proof of Lemma 4.1, we have

$$\|u_{m,1} - U_{m+1}\|_{I_n} \;\leq\; \frac{\|C_{m,1}\|_{I_n}}{\alpha - \|C_{m,1}\|_{I_n}}\|U_{m+1} - \bar{u}_{m,1}\|_{I_n}. \tag{32}$$

Since $u_{m,0} := U_m$, we have

$$\|u_{m,0} - U_{m+0}\|_{I_n} = 0.$$

Now using the definition of $\bar{u}_{m,1}$, Theorem 2.1, and the triangle inequality, we obtain

$$\begin{aligned}
\|U_{m+1} - \bar{u}_{m,1}\|_{I_n} &\leq \|U_{m+1} - u\|_{I_n} + \|u - U_{m+0}\|_{I_n} + \|U_{m+0} - u_{m,0}\|_{I_n} \\
&\leq CL_n \max_{1\leq n\leq N} E_{m+1,qn}(u) + CL_n \max_{1\leq n\leq N} E_{m,qn}(u) + 0 \\
&= CL_n \max_{1\leq n\leq N} E_{m+1,qn}(u) + CL_n \max_{1\leq n\leq N} E_{m,qn}(u) \\
&\leq CL_n \max_{1\leq n\leq N} E_{m+1,qn}(u). \qquad (33)
\end{aligned}$$

Note that since we are using the wavelet bases for the subspaces X_m and X_{m+1}, we have $\mu^2 h^2_{m+1,n} = h^2_{m,n}$, which implies

$$L_n \max_{1\leq n\leq N} E_{m,qn}(u) \leq CL_n \max_{1\leq n\leq N} E_{m+1,qn}(u)$$

for some constant C. Substituting the estimate (33) into the right hand side of equation (32), we get

$$\|u_{m,1} - U_{m+1}\|_{I_n} \leq \frac{\|C_{m,1}\|_{I_n}}{\alpha - \|C_{m,1}\|_{I_n}} CL_n \max_{1\leq n\leq N} E_{m+1,qn}(u).$$

Now, employing hypothesis (IV), there exists a positive integer M such that for $m \geq M$, we have

$$\|C_{m,1}\| \leq \frac{\alpha}{2}.$$

Thus for $m \geq M$, we have

$$\frac{\|C_{m,1}\|_{I_n}}{\alpha - \|C_{m,1}\|_{I_n}} \leq \frac{\frac{\alpha}{2}}{\alpha - \frac{\alpha}{2}} = 1$$

so

$$\|u_{m,1} - U_{m+1}\|_{I_n} \leq CL_n \max_{1\leq n\leq N} E_{m+1,qn}(u).$$

Now using Theorem 2.1 and the last inequality, there exists a positive integer M such that for $m \geq M$, and $n = 1, 2, \cdots, N$, we have

$$\begin{aligned}
\|u - u_{m,1}\|_{I_n} &\leq \|u - U_{m+1}\|_{I_n} + \|U_{m+1} - u_{m,1}\|_{I_n} \\
&\leq CL_n \max_{1\leq n\leq N} E_{m+1,qn}(u) + CL_n \max_{1\leq n\leq N} E_{m+1,qn}(u) \\
&\leq CL_n \max_{1\leq n\leq N} E_{m+1,qn}(u)
\end{aligned}$$

which completes the proof.

In the next chapter, we examine special time and space discretizations to treat problems with difficult initial conditions.

CHAPTER V

INCOMPATIBLE INITIAL CONDITIONS

V.1 THE DIFFICULTY

In this chapter, most of which is from [6], we consider the case where the initial condition of the parabolic problem is incompatible with the prescribed boundary conditions. For convenience, we follow [6] and use the same notation.

As an example of the above mentioned problem, consider finding u such that

$$u_t(x,t) - u_{xx}(x,t) = 0, \quad x \in \Omega, \quad t > 0,$$

$$u(0,t) = u(1,t) = 0, \quad t > 0,$$

$$u(x,0) = u_0(x), \quad x \in \Omega,$$

where $u_0(x) = 1 - x$ and $\Omega = (0,1)$. Here the actual solution is found, after some work with Fourier Series, to be

$$u(x,t) = \sum_{j=1}^{\infty} u_j^0 e^{-j^2 t} \sin(j\pi x),$$

where the coefficients u_j^0 are given by

$$u_j^0 = 2\int_0^1 (1-\xi)\sin(j\pi\xi)d\xi = \frac{2}{\pi}\left\{\frac{1}{j} - \frac{1}{j^2\pi}\sin(j\pi)\right\} = O\left(\frac{1}{j}\right).$$

In the following discussion, we will let C denote a constant that changes as it appears. For the solution $u(x,t)$ to the above problem, we have

$$\|u_t(t)\|_2^2 = \|u_t(t)\|_{L_2(0,1)}^2 = \sum_{j=1}^{\infty} Cj^2 e^{-2j^2 t} = \sum_{j=1}^{\infty} \frac{d}{dt} Ce^{-2j^2 t} = \frac{d}{dt}\sum_{j=1}^{\infty} Ce^{-2j^2 t},$$

due to the uniform convergence in t of the series $\sum_{k=1}^{\infty} Ce^{-2j^2 t}$. Since $\int_0^{\infty} e^{-\zeta^2} d\zeta < \infty$, using the change of variable $y = j\sqrt{2t}$, we obtain

$$\frac{d}{dt}\sum_{j=1}^{\infty} Ce^{-2j^2 t} = \frac{d}{dt} Ct^{-1/2},$$

which implies

$$\|u_t(t)\|_2 = O(t^{-3/4}).$$

Note that as $t \to 0^+$, we have $\|u_t\|_{I_1} \to \infty$, which will cause difficulty when approximating $u(x,t)$.

A similar situation will also arise when $u_0(x) = \min(x, 1-x)$ in the above problem. Here we have $\|u_t(t)\|_2 = O(t^{-1/4})$.

Both these situations can be treated with the following time/space partitioning scheme from [6].

V.2 THE TIME DISCRETIZATION SCHEME

We will consider the one-dimensional parabolic problem of finding $u(x,t)$ such that

$$u_t(x,t) - u_{xx}(x,t) = 0, \quad x \in \Omega, \quad t > 0, \tag{34}$$

$$u(0,t) = u(1,t) = 0, \quad t > 0,$$

$$u(x,0) = u_0(x), \quad x \in \Omega.$$

where the initial condition $u_0(x)$ is incompatible with the prescribed boundary conditions. As with this previous section, we follow [6]. Assume

$$\|u_t(t)\|_2 = O(t^{-\alpha})$$

where $0 < \alpha < 1$. Let q be a nonnegative integer. We define an *index of singularity* as $Q := \frac{q+1}{1-\alpha}$. For $T > 0$ and a positive integer N, let

$$t_n^* = \left(\frac{n}{N}\right)^Q, \quad n = 0, 1, 2, \cdots, N,$$

and

$$t_n = t_n^* T. \tag{35}$$

As before, we define $I_n = (t_{n-1}, t_n]$ for $n = 1, 2, \cdots, N$ with $k_n = t_n - t_{n-1}$ denoting the length of the time subinterval I_n. Hence, we have

$$k_n = \left[\left(\frac{n}{N}\right)^Q - \left(\frac{n-1}{N}\right)^Q\right] T, \quad n = 1, 2, \cdots, N.$$

By the Mean Value Theorem from calculus,

$$k_n \leq Q \left[\frac{n}{N}\right]^{Q-1} \frac{1}{N} T.$$

The solution of the parabolic problem (34) is then approximated in the time variable t over each time subinterval I_n by a polynomial of degree q. For example, in the case of $q = 1$, let $P_1 w$ be the linear interpolatory projection in time of $w \in H_0^2(\Omega)$ onto S_k where

$$S_k = \{v : [0, \infty) \to H_0^1(\Omega); v|_{I_n} \in P_q(I_n), n = 1, \cdots, N\},$$

$$P_q(I_n) := \{v(t) = \sum_{j=0}^{q} v_i t^i : v_i \in H_0^1(\Omega)\},$$

$$H_0^2(\Omega) = \{v : D^{(j)} v \in L_2(\Omega), \quad j = 0, 1, 2; \quad v \equiv 0 \text{ on } \partial\Omega\}.$$

That is, we have

$$P_1 w(x,t) = \frac{t_n - t}{k_n} w(x, t_{n-1}) + \frac{t - t_{n-1}}{k_n} w(x, t_n)$$

for each $t \in I_n$. Note that P_1, when seen as an operator on $H_0^2(\Omega)$, is bounded with respect to the norm $\|w(t)\|_{\infty, I_n}$ where

$$\|w(t)\|_{\infty, I_n} = \max_{t \in I_n} \|w(t)\|_{L_\infty(\Omega)}.$$

Since Ω is assumed to be a bounded domain, we have P_1 bounded with respect to the norm $\| \cdot \|_{I_n}$ as well. If $w(x,t) = w_0(x)$, that is, constant in time, then we have

$$P_1 w(x,t) = \frac{t_n - t}{k_n} w_0(x) + \frac{t - t_{n-1}}{k_n} w_0(x) = w_0(x).$$

If $w(x,t) = w_0(x) + t w_1(x)$, that is, linear in time, then we have

$$\begin{aligned} P_1 w(x,t) &= \frac{t_n - t}{k_n} w(x, t_{n-1}) + \frac{t - t_{n-1}}{k_n} w(x, t_n) \\ &= \frac{t_n - t}{k_n} [w_0(x) + t_{n-1} w_1(x)] + \frac{t - t_{n-1}}{k_n} [w_0(x) + t_n w_1(x)] \\ &= w_0(x) + t w_1(x). \end{aligned}$$

Thus, P_1 equals the identity on either constant or linear polynomials. If we write the Taylor series in the time variable t about the point t_n to the first or second order, respectively we obtain, for each $n = 1, 2, \cdots, N$,

$$\|u - P_1 u\|_{I_n} \le \int_{I_n} \|u_t(t)\|_2 \, dt,$$

and

$$\|u - P_1 u\|_{I_n} \le C k_n^2 \|u_{tt}\|_{I_n}.$$

For higher values of q, writing the Taylor expansion to the order q, the best possible estimate for the projection $P_q : H_0^2(\Omega) \to S_k$ is given by

$$\|u - P_q u\|_{I_n} \le \min\left\{ \int_{I_n} \|u_t(t)\|_2 \, dt, \ C k_n^{q+1} \left\| \frac{\partial^{q+1}}{\partial t^{q+1}} u \right\|_{I_n} \right\}. \tag{36}$$

Since we are interested in cases where $\|u_t(t)\|_2 = O(t^{-\alpha})$ for $0 < \alpha < 1$, the next lemma will be quite useful.

Lemma 5.1. (Kaneko, Bey, and Hou [6]). *Let $0 < \alpha < 1$, q a nonnegative integer and $T > 0$. Also, assume t_n, $n = 1, 2, \cdots, N$ are defined by (35). Then*

$$\int_{I_n} s^{-\alpha} ds \le C \frac{1}{N^{q+1}},$$

where C is a constant independent of N and for $n > 1$,

$$max_{1 \le n \le N} \int_{I_n} s^{-\alpha} ds \le C \frac{1}{N},$$

where C is a constant independent of N.

Proof. For $n = 1$, we have

$$\int_{I_1} s^{-\alpha} ds = \int_0^{(\frac{1}{N})^Q T} s^{-\alpha} ds = \frac{1}{1-\alpha} \left[\left(\frac{1}{N}\right)^Q T \right]^{1-\alpha} = O\left(\left(\frac{1}{N}\right)^{q+1} \right).$$

For $1 < n \le N$, we have

$$\begin{aligned}
\int_{I_n} s^{-\alpha} ds &\le \int_{I_n} \left[\left(\frac{n-1}{N}\right)^Q T \right]^{-\alpha} ds \text{ as } s^{-\alpha} \text{ is decreasing over } I_n, \\
&= T^{-\alpha} \left(\frac{n-1}{N}\right)^{-\alpha Q} \left[\left(\frac{n}{N}\right)^Q - \left(\frac{n-1}{N}\right)^Q \right] \\
&\le C \left(\frac{n-1}{N}\right)^{-\alpha Q} \left(\frac{n}{N}\right)^{Q-1} \left(\frac{n}{N} - \frac{n-1}{N}\right) \\
&= C \left(\frac{1}{N}\right)^{Q-\alpha Q} \frac{n^{Q-1}}{(n-1)^{\alpha Q}} = C \frac{1}{N^{q+1}} \left(\frac{n}{n-1}\right)^{\alpha Q} n^{Q-1-\alpha Q} \\
&= C \frac{n^q}{N^{q+1}} \le C \frac{N^q}{N^{q+1}} = C \frac{1}{N}.
\end{aligned}$$

Using Lemma 5.1 with $n = 1$ and $q = 1$ will lead to $\int_{I_n} \|u_t(t)\|dt = O(\frac{1}{N^2})$, assuming $\|u_t(t)\|_2 = O(t^{-\alpha})$ for $0 < \alpha < 1$ and $t \in I_1$. Since we assume $\|u_{tt}\|_{I_n}$ is bounded uniformly in $n > 1$, (36) implies

$$\|u - P_1 u\|_{I_n} = O\left(\frac{1}{N^2}\right), \quad n = 1, 2, \cdots, N.$$

The next lemma guarantees the stability of the Discontinuous Galerkin Method by showing $(1 + \log \frac{t_n}{k_n})^{\frac{1}{2}}$ to be uniformly bounded.

Lemma 5.2. (Kaneko, Bey, and Hou [6]). *Assume t_n and k_n are defined by (35). Then, for any positive integer N, we have*

$$\left(1 + \log \frac{t_n}{k_n}\right)^{\frac{1}{2}} \le \sqrt{2}, \quad \textit{for each} \quad n = 0, 1, 2, \cdots, N.$$

Proof. We use the fact that for $0 < x < 1$, we have $\log(1 - x) < -x$ for $x < 1$. Now,

$$\begin{aligned}
\left(1 + \log \frac{t_n}{k_n}\right)^{\frac{1}{2}} &= \left(1 + \log \frac{(\frac{n}{N})^Q}{(\frac{n}{N})^Q - (n - \frac{1}{N})^Q}\right)^{\frac{1}{2}} \\
&= \left(1 + \log \frac{1}{1 - (\frac{n-1}{n})^Q}\right)^{\frac{1}{2}} \\
&= \left(1 - \log\left(1 - \left(\frac{n-1}{n}\right)^Q\right)\right)^{\frac{1}{2}} \\
&\le \left(1 - \log\left(1 - \left(\frac{N-1}{N}\right)^Q\right)\right)^{\frac{1}{2}} \\
&\le \left(1 + \left(\frac{N-1}{N}\right)^Q\right)^{\frac{1}{2}} \\
&\le \sqrt{2}.
\end{aligned}$$

The following theorem is a modification of Theorem 2.1. Minor changes to the proof of this theorem in [7] serve as the proof of this result.

Theorem 5.3. (Kaneko, Bey, and Hou [6]). *Assume there is a constant γ such that the time steps k_n satisfy $k_n \le \gamma k_{n+1}$, $n = 1, 2, \cdots, N - 1$, and let U^n be the solution of (7) approximating u at t_n. Let u^n denote the value of u at t_n. Here u is approximated by a polynomial of degree $q \ge 0$ over each I_n for each $n =$*

$1, 2, \cdots, N-1$. *Then there is a constant C depending only on γ and a constant β, where $\rho_K \geq \beta h_K$ and ρ_K is the diameter of the circle inscribed in K for all $K \in T$, such that for $n = 1, 2, \cdots, N$,*

$$\|u^n - U^n\|_2 \leq C\left(1 + \log\frac{t_n}{k_n}\right)^{\frac{1}{2}} \left\{max_{j\leq n}\|u - P_q u\|_{I_j} + \|h_n^2 D^2 u\|_{I_n}\right\}.$$

The current time and space discretization schemes allow us to use Theorem 5.3 for the following reasons. First, by the construction, the time steps k_n, $n = 1, 2, \cdots, N$ are increasing in size, so the condition $k_n \leq \gamma k_{n+1}$ is satisfied for $\gamma = 1$. Second, for the one-dimensional problem (34), we have $h_K = \rho_K = |\Omega_K|$ so that $\frac{\rho_K}{h_K} = 1$ for all Ω_K. Lemma 5.2 implies the uniformly bounded property of $(1 + \log\frac{t_n}{k_n})^{\frac{1}{2}}$, which results in the stability of the Discontinuous Galerkin Method.

For any nonnegative integer q, by using equation (36) and Lemma 5.1, we have

$$\|u - P_q u\|_{I_n} = O\left(\frac{1}{N^{q+1}}\right),$$

for some $0 < \alpha < 1$, assuming $\|u_t(t)\|_2 = O(t^{-\alpha})$.

By assuming $\|D^2u(t)\|_{I_n}$ is bounded for $n = 1, 2, \cdots, N$ and employing the graded time partitions discussed in Lemma 5.1, Theorem 5.3 can be modified to the following result.

Theorem 5.4. (Kaneko, Bey, and Hou [6]) *For the parabolic problem (1), assume the initial value $u_0(x)$ is defined in such a way that $\|u_t(t)\| = O(t^{-\alpha})$, for $0 < \alpha < 1$. Also, assume $\|D^2u(t)\|_{I_n}$ is bounded for each $n = 1, 2, \cdots, N$, and (0,1) is divided into 2^M subintervals each of equal length. Denote by U_M^n the solution of (7) approximating u at t_n and let u^n denote the value of u at t_n. Let the time discretization $\{t_n\}$ be defined by (35). If q denotes the degree of the approximating polynomials to u in the time variable t, then for each $n = 1, 2, \cdots, N$,*

$$\|u^n - U_M^n\|_2 = O\left(\frac{1}{N^{q+1}} + \frac{1}{2^{2M}}\right).$$

The next theorem is a new result that is similar in many ways to that of Theorem 4.2. It is essentially Theorem 5.4 enhanced with the multilevel calculation method, and shows that the accuracy of Theorem 5.4 is preserved when the multilevel solution

$u^n_{M,1}$ is used to approximate the Discontinuous Galerkin Method solution U^n_{M+1}. The strategy for the proof is similar to that used for proving Theorem 4.2. in that it is proven as two cases, one case using Hypotheses (I) and (II) and the other case, Hypotheses (III) and (IV).

Theorem 5.5. *Suppose hypotheses (I) and (II) or (III) and (IV) from Chapter IV are satisfied. Let $u \in X$ be the solution of equation (1) where $\Omega = (0,1)$, and $u_{m,1} \in X_m$ be the multilevel solution of equation (22) with $X_m = \ span\{w_{ij}\}$ where $\{w_{ij}\}$ is the wavelet basis described in Section II.4. Let the time discretization $\{t_n\}$ be defined by (35), $\|D^2u(t)\|_{I_n}$ bounded for $n = 1, 2, \cdots, N$, and $\|u_t(t)\|_2 = O(t^{-\alpha})$ for $0 < \alpha < 1$. If q denotes the degree of the approximating polynomials to u in the time variable t, then for each $n = 1, 2, \cdots, N$, there exists a positive integer M_0 such that for $M \geq M_0$, with the interval (0,1) subdivided into 2^{M+1} subintervals each of equal length, we have*

$$\|u^n - u^n_{M,1}\|_{I_n} = O\left(\frac{1}{N^{q+1}} + \frac{1}{2^{2(M+1)}}\right)$$

where $\|u\|_{I_n} = \max_{t \in I_n} \|u(t)\|_2$.

Proof (Version One). Assume hypotheses (I) and (II) are satisfied. The proof of this theorem is similar to the proof of Theorem 4.2, employing the same operators and notation. As before, we first establish an estimate on $\|u^n_{M,1} - U^n_{M+1}\|_{I_n}$. For this purpose, we subtract

$$B_{M,1}u_M(1) = f_{M+1} - C_{M,1}u_M(1)$$

from

$$B_{M,1}u_{M,1} = f_{M+1} - C_{M,1}\bar{u}_{M,1}$$

to obtain

$$B_{M,1}(u^n_{M,1} - U^n_{M+1}) = C_{M,1}(U^n_{M+1} - \bar{u}^n_{M,1}).$$

Since hypotheses (I) is assumed, we have

$$(u^n_{M,1} - U^n_{M+1}) = B^{-1}_{M,1}C_{M,1}(U^n_{M+1} - \bar{u}^n_{M,1}).$$

Hypothesis (II) now implies

$$\|u^n_{M,1} - U^n_{M+1}\|_{I_n} \ \leq \ C\|U^n_{M+1} - \bar{u}^n_{M,1}\|_{I_n}. \tag{37}$$

Since $u^n_{M,0} = U^n_M$, we have

$$\|u^n_{M,0} - U^n_{M+0}\|_{I_n} = 0.$$

Using the definition of $\bar{u}^n_{M,1}$ and Theorem 5.4, there exists a positive integer M_0 such that $M \geq M_0$ implies

$$\begin{aligned}
\|U^n_{M+1} - \bar{u}^n_{M,1}\|_{I_n} &\leq \|U^n_{M+1} - u^n\|_{I_n} + \|u^n - U^n_{M+0}\|_{I_n} + \|U^n_{M+0} - u^n_{M,0}\|_{I_n} \\
&= O\left(\frac{1}{N^{q+1}} + \frac{1}{2^{2(M+1)}}\right) + O\left(\frac{1}{N^{q+1}} + \frac{1}{2^{2M}}\right) + 0 \\
&= O\left(\frac{1}{N^{q+1}} + \frac{1}{2^{2(M+1)}}\right).
\end{aligned}$$

where the last equality is justified by the fact that

$$\frac{1}{2^{M+1}} = h_{M+1,n} = \mu^{-1} h_{M,n} = 2^{-1}\frac{1}{2^M}$$

from the use of the wavelet basis. Substituting this estimate into the right hand side of equation (37), there exists a positive integer M_0 such that $M \geq M_0$ implies

$$\|u^n_{M,1} - U^n_{M+1}\|_{I_n} = O\left(\frac{1}{N^{q+1}} + \frac{1}{2^{2(M+1)}}\right).$$

Now, using Theorem 5.4 and the above reasoning, for each $n = 1, 2, \cdots, N$, there exists a positive integer M_0 such that for $M \geq M_0$ with the interval (0,1) subdivided into 2^{M+1} subintervals each of equal length, we have

$$\begin{aligned}
\|u^n - u^n_{M,1}\|_{I_n} &\leq \|u^n - U^n_{M+1}\|_{I_n} + \|U^n_{M+1} - u^n_{M,1}\|_{I_n} \\
&= O\left(\frac{1}{N^{q+1}} + \frac{1}{2^{2(M+1)}}\right)
\end{aligned}$$

which completes the proof of this version of the theorem.

Proof (Version Two). Assume hypotheses (III) and (IV) are satisfied. Again, the proof of this theorem is similar to the proof of Theorem 4.2, employing the same operators and notation. As before, we first establish an estimate on $\|u^n_{M,1} - U^n_{M+1}\|_{I_n}$. For this purpose, we subtract

$$B_{M,1}u_M(1) = f_{M+1} - C_{M,1}u_M(1)$$

from

$$B_{M,1}u_{M,1} = f_{M+1} - C_{M,1}\overline{u}_{M,1}$$

to obtain

$$B_{M,1}(u^n_{M,1} - U^n_{M+1}) = C_{M,1}(U^n_{M+1} - \bar{u}^n_{M,1}).$$

Since hypotheses (III) and (IV) are assumed, Lemma 4.1 applies, so there exists a positive integer M_1 such that $M \geq M_1$ implies

$$(u^n_{M,1} - U^n_{M+1}) = B^{-1}_{M,1}C_{M,1}(U^n_{M+1} - \bar{u}^n_{M,1}).$$

Thus from the previous equation and the inequality

$$\|B^{-1}_{M,1}\|_{I_n} \leq \frac{1}{\alpha - \|C_{M,1}\|_{I_n}}$$

from the proof of Lemma 4.1, there exists a positive integer M_1 such that $M \geq M_1$ implies

$$\|u^n_{M,1} - U^n_{M+1}\|_{I_n} \leq \frac{\|C_{M,1}\|_{I_n}}{\alpha - \|C_{M,1}\|_{I_n}} \|U^n_{M+1} - \bar{u}^n_{M,1}\|_{I_n}. \tag{38}$$

Since $u^n_{M,0} = U^n_M$, we have

$$\|u^n_{M,0} - U^n_{M+0}\|_{I_n} = 0.$$

Using the definition of $\bar{u}^n_{M,1}$ and Theorem 5.4, there exists a positive integer M_1 such that $M \geq M_1$ implies

$$\begin{aligned} \|U^n_{M+1} - \bar{u}^n_{M,1}\|_{I_n} &\leq \|U^n_{M+1} - u^n\|_{I_n} + \|u^n - U^n_{M+0}\|_{I_n} + \|U^n_{M+0} - u^n_{M,0}\|_{I_n} \\ &= O\left(\frac{1}{N^{q+1}} + \frac{1}{2^{2(M+1)}}\right) + O\left(\frac{1}{N^{q+1}} + \frac{1}{2^{2M}}\right) + 0 \\ &= O\left(\frac{1}{N^{q+1}} + \frac{1}{2^{2(M+1)}}\right) \end{aligned}$$

where the last equality is justified by the fact that

$$\frac{1}{2^{M+1}} = h_{M+1,n} = \mu^{-1}h_{M,n} = 2^{-1}\frac{1}{2^M}$$

from the use of the wavelet basis. Substituting this estimate into the right hand side of equation (38), there exists a positive integer M_1 such that $M \geq M_1$ implies

$$\|u^n_{M,1} - U^n_{M+1}\|_{I_n} \leq \frac{\|C_{M,1}\|_{I_n}}{\alpha - \|C_{M,1}\|_{I_n}} O\left(\frac{1}{N^{q+1}} + \frac{1}{2^{2(M+1)}}\right).$$

Employing hypothesis (IV), there exists a positive integer M_2 such that for $M \geq M_2$,

we have

$$\|C_{M,1}\|_{I_n} \leq \frac{\alpha}{2}.$$

Thus for $M \geq M_2$, we have

$$\frac{\|C_{M,1}\|_{I_n}}{\alpha - \|C_{M,1}\|_{I_n}} \leq \frac{\frac{\alpha}{2}}{\alpha - \frac{\alpha}{2}} = 1.$$

so

$$\|u^n_{M,1} - U^n_{M+1}\|_{I_n} = O\left(\frac{1}{N^{q+1}} + \frac{1}{2^{2(M+1)}}\right).$$

Now, using Theorem 5.4 and the above reasoning, for each $n = 1, 2, \cdots, N$, there exists a positive integer $M_0 = \max\{M_1, M_2\}$ such that for $M \geq M_0$ with the interval (0,1) subdivided into 2^{M+1} subintervals each of equal length, we have

$$\begin{aligned} \|u^n - u^n_{M,1}\|_{I_n} &\leq \|u^n - U^n_{M+1}\|_{I_n} + \|U^n_{M+1} - u^n_{M,1}\|_{I_n} \\ &= O\left(\frac{1}{N^{q+1}} + \frac{1}{2^{2(M+1)}}\right) \end{aligned}$$

which completes the proof.

Practical constraints of many applications may force one to abandon the assumption the $\|D^2u(t)\|_{I_n}$ is bounded for each n and t, e.g. the two problems in the last section. In this next section, we examine a space grid partitioning scheme designed to deal with this situation.

V.3 THE SPACE DISCRETIZATION SCHEME

In [6], if $\|u_t(t)\| = O(t^{-\alpha})$, then we select a set of spatial grid points $\{x_m(t_n)\}$ which are dependent on the immediate choice of time step t_n. This is done by selecting a spatial increment h_n based on the size of $\|u_t(t)\|_{\infty,I_n}$. Let

$$h(t) = \max_{1 \leq m \leq 2^M}[x_m(t) - x_{m-1}(t)]$$

for each $t \in (0, T]$. Then $h(t)$ is determined from the condition that

$$(h(t))^2 t^{-\alpha} = O(t^2), \quad \text{as} \quad t \to 0^+.$$

In terms of N, we require that

$$(h(t))^2 t^{-\alpha} = O\left(\frac{1}{N^2}\right).$$

For $t = t_n = (\frac{n}{N})^{\frac{q+1}{1-\alpha}}T$, $n = 1, 2, \cdots, N$, we have , for a constant $C > 0$,

$$(h(t_n))^2 \left[\left(\frac{n}{N}\right)^{\frac{q+1}{1-\alpha}} T \right]^{-\alpha} = C\frac{1}{N^2}.$$

If we solve this last equation for $h(t_n)$, using $C = 1$ for convenience, we obtain

$$h(t_n) = \left[\frac{n^{\alpha Q}}{N^{2+\alpha Q}} T^{\alpha} \right]^{\frac{1}{2}}.$$

With this done, we select a resolution level $M(n)$ such that

$$2^{M(n)} > \frac{L}{h(t_n)} \tag{39}$$

for each time level t_n, $n = 1, 2, \cdots N$ where L denotes the length of the one-dimensional spatial interval and $M(n)$ denotes the fact that the grid resolution M is dependent on the time step t_n. The spatial partition points are defined by

$$x_m(t_n) = \frac{mL}{2^{M(n)}}, \quad m = 1, 2, \cdots, 2^M - 1.$$

Since the space discretization is dependent on the time discretization in this scheme, we have the following theorem which provides an error estimate for the Discontinuous Galerkin Method in both the time and space variables, using the desired number of time steps N.

Theorem 5.6. (Kaneko, Bey, and Hou [6]) *For the parabolic problem (1), assume the initial value $u_0(x)$ is defined in such a way that $\|u_t(t)\| = O(t^{-\alpha})$, for $0 < \alpha < 1$. Denote by U_M^n the solution of (7) approximating u at t_n and let u^n denote the value of u at t_n. Let the time discretization $\{t_n\}$ be defined by (35). Assume (0,1) is divided into $2^{M(n)}$ subintervals each of equal length, where $M(n)$ is defined by (39). If q denotes the degree of the approximating polynomials to u in the time variable t, then for each $n = 1, 2, \cdots, N$,*

$$\|u^n - U_{M(n)}^n\|_2 = O\left(\frac{1}{N^{q+1}} + \frac{1}{N^2}\right).$$

The above scheme allows us to use fewer spatial grid points as $n \to N$, thus lowering the computational effort and expense.

In the next section, we detail an actual one-dimensional implementation of the Discontinuous Galerkin Method, coupled with the Multilevel Augmentation Method described earlier.

CHAPTER VI

IMPLEMENTATION

VI.1 SETUP

For implementation of this method, the linear wavelet basis of the Sobolev space $H_0^1(0,1)$, as constructed in Section II.4.2, is chosen, with the basic notation of Chapter III employed. Note that with this choice of basis, the dimension of each subspace X_M will be $d_M = 2^M - 1$. An arbitrary grid resolution $M+1$ is chosen, an appropriate number of time steps N selected, and the various matrices constructed by augmentation. As before, we have

$$(w_{ij}, w_{i'j'}) := \int_0^1 w_{ij}(x) w_{i'j'}(x)\ dx,$$

and

$$(\frac{d}{dx}w_{ij}, \frac{d}{dx}w_{i'j'}) := \int_0^1 \frac{d}{dx}w_{ij}(x)\frac{d}{dx}w_{i'j'}(x)\ dx = \delta_{iji'j'}$$

where

$$\delta_{iji'j'} := \begin{cases} 1 & \text{when } i = i',\ j = j', \\ 0 & \text{otherwise.} \end{cases}$$

First, the matrix A_1^n is defined as

$$A_1^n = \left[\ (w_{10}, w_{10}) + k_n\ \right]$$

or

$$A_1^n = \left[\ \tfrac{1}{12} + k_n\ \right].$$

From there the matrices

$$F_{1,2}^n = \begin{bmatrix} \frac{\sqrt{2}}{64} & \frac{\sqrt{2}}{64} \end{bmatrix}, \quad G_{2,1}^n = \begin{bmatrix} \frac{\sqrt{2}}{64} \\ \frac{\sqrt{2}}{64} \end{bmatrix}, \quad H_{2,2}^n = \begin{bmatrix} \frac{1}{48} + k_n & 0 \\ 0 & \frac{1}{48} + k_n \end{bmatrix}$$

are constructed, then augmentation of A_1^n gives

$$A_{1,1}^n = \begin{bmatrix} A_1^n & F_{1,2}^n \\ G_{2,1}^n & H_{2,2}^n \end{bmatrix} = \begin{bmatrix} \frac{1}{12}+k_n & \frac{\sqrt{2}}{64} & \frac{\sqrt{2}}{64} \\ \frac{\sqrt{2}}{64} & \frac{1}{48}+k_n & 0 \\ \frac{\sqrt{2}}{64} & 0 & \frac{1}{48}+k_n \end{bmatrix}.$$

This process is then continued, constructing the matrices

$$F_{2,3}^n = \begin{bmatrix} \frac{1}{256} & \frac{3}{256} & \frac{3}{256} & \frac{1}{256} \\ \frac{\sqrt{2}}{256} & \frac{\sqrt{2}}{256} & 0 & 0 \\ 0 & 0 & \frac{\sqrt{2}}{256} & \frac{\sqrt{2}}{256} \end{bmatrix}, \quad G_{3,2}^n = \begin{bmatrix} \frac{1}{256} & \frac{\sqrt{2}}{256} & 0 \\ \frac{3}{256} & \frac{\sqrt{2}}{256} & 0 \\ \frac{3}{256} & 0 & \frac{\sqrt{2}}{256} \\ \frac{1}{256} & 0 & \frac{\sqrt{2}}{256} \end{bmatrix},$$

$$H_{3,3}^n, = \begin{bmatrix} \frac{1}{192}+k_n & 0 & 0 & 0 \\ 0 & \frac{1}{192}+k_n & 0 & 0 \\ 0 & 0 & \frac{1}{192}+k_n & 0 \\ 0 & 0 & 0 & \frac{1}{192}+k_n \end{bmatrix},$$

then augmenting A_2^n to obtain

$$A_{2,1}^n = \begin{bmatrix} A_2^n & F_{2,3}^n \\ G_{3,2}^n & H_{3,3}^n \end{bmatrix}$$

or

$$A_{2,1}^n = \begin{bmatrix} \frac{1}{12}+k_n & \frac{\sqrt{2}}{64} & \frac{\sqrt{2}}{64} & \frac{1}{256} & \frac{3}{256} & \frac{3}{256} & \frac{1}{256} \\ \frac{\sqrt{2}}{64} & \frac{1}{48}+k_n & 0 & \frac{\sqrt{2}}{256} & \frac{\sqrt{2}}{256} & 0 & 0 \\ \frac{\sqrt{2}}{64} & 0 & \frac{1}{48}+k_n & 0 & 0 & \frac{\sqrt{2}}{256} & \frac{\sqrt{2}}{256} \\ \frac{1}{256} & \frac{\sqrt{2}}{256} & 0 & \frac{1}{192}+k_n & 0 & 0 & 0 \\ \frac{3}{256} & \frac{\sqrt{2}}{256} & 0 & 0 & \frac{1}{192}+k_n & 0 & 0 \\ \frac{3}{256} & 0 & \frac{\sqrt{2}}{256} & 0 & 0 & \frac{1}{192}+k_n & 0 \\ \frac{1}{256} & 0 & \frac{\sqrt{2}}{256} & 0 & 0 & 0 & \frac{1}{192}+k_n \end{bmatrix}$$

where $A_2^n := A_{1,1}^n$. One then continues this process until the desired grid resolution level M is obtained. In general, for $m = 2, 3, \cdots, M$, assuming $A_M^n := A_{M-1,1}^n$ has been obtained, one computes the matrices $F_{M,M+1}^n$, $G_{M+1,M}^n$, and $H_{M+1,M+1}^n$, then constructs the matrix

$$A^n_{M,1} = \begin{bmatrix} A^n_M & F^n_{M,M+1} \\ G^n_{M+1,M} & H^n_{M+1,M+1} \end{bmatrix}.$$

As for the actual matrices, the construction of A^n_M has been described above. The matrix $F^n_{M,M+1}$ presents the most computation. Matrix $G^n_{M+1,M}$ is the transpose of $F^n_{M,M+1}$. Matrix $H^n_{M+1,M+1}$ is a simple diagonal matrix, with the main diagonal entries based on the spatial index m and the time step size k_n.

As we did before, we decompose $\overrightarrow{u}^n_{M+1}$ as

$$\overrightarrow{u}^n_{M+1} = \begin{bmatrix} \overrightarrow{u}^n_{M,0} \\ \overrightarrow{v}^n_{M,1} \end{bmatrix}$$

where

$$\begin{aligned} \overrightarrow{u}^n_{M,0} &:= [\xi^n_{i'j'}]_{(2^M-1)\times 1}, \\ \overrightarrow{v}^n_{M,1} &:= [\eta^n_{i'j'}]_{2^M\times 1}. \end{aligned}$$

These will be the scalars we eventually find.

For the constant in time case, we have $\overrightarrow{f}_{M+1}$ written as

$$\overrightarrow{f}_{M+1} = \begin{bmatrix} \overrightarrow{f}_M \\ \overrightarrow{g}_M \end{bmatrix}$$

where

$$\begin{aligned} \overrightarrow{f}_M &:= [f_{i'j'}]_{(2^M-1)\times 1}, \\ \overrightarrow{g}_M &:= [g_{i'j'}]_{2^M\times 1}. \end{aligned}$$

with

$$f_{i'j'} := \sum_{ij=10}^{M\ (2^{M-1}-1)} \xi^{n-1}_{ij}(w_{ij}, w_{i'j'}) + \sum_{ij=(M+1)\ 0}^{(M+1)\ (2^M-1)} \eta^{n-1}_{ij}(w_{ij}, w_{i'j'}) + \int_{I_n} (f, w_{i'j'})dt,$$

for $i'j' = 1\ 0, \cdots, M\ (2^{M-1}-1)$, and

$$g_{i'j'} := \sum_{ij=10}^{M\ (2^{M-1}-1)} \xi^{n-1}_{ij}(w_{ij}, w_{i'j'}) + \sum_{ij=(M+1)\ 0}^{(M+1)\ (2^M-1)} \eta^{n-1}_{ij}(w_{ij}, w_{i'j'}) + \int_{I_n} (f, w_{i'j'})dt,$$

for $i'j' = (M+1)\ 0, \cdots, (M+1)\ (2^M - 1)$.

With these various arrays now defined, the system

$$A^n_{M+1} \overrightarrow{u}^n_{M+1} = \overrightarrow{f}_{M+1}$$

becomes the system

$$\begin{bmatrix} A^n_M & F^n_{M,M+1} \\ G^n_{M+1,M} & H^n_{M+1,M+1} \end{bmatrix} \begin{bmatrix} \overrightarrow{u}^n_{M,0} \\ \overrightarrow{v}^n_{M,1} \end{bmatrix} = \begin{bmatrix} \overrightarrow{f}_M \\ \overrightarrow{g}_M \end{bmatrix}$$

or

$$\begin{aligned} A^n_M \overrightarrow{u}^n_{M,0} + F^n_{M,M+1} \overrightarrow{v}^n_{M,1} &= \overrightarrow{f}_M, \\ G^n_{M+1,M} \overrightarrow{u}^n_{M,0} + H^n_{M+1,M+1} \overrightarrow{v}^n_{M,1} &= \overrightarrow{g}_M. \end{aligned}$$

To obtain the multilevel approximation $u^n_{M,1}$ of the M+1 level numerical solution u_{M+1}, we first solve the coarse grid problem

$$A^n_M \overrightarrow{u}^n_M = \overrightarrow{f}_M$$

obtaining the Mth level solution $\overrightarrow{u}^n_M$.

Next we solve the system

$$H^n_{M+1,M+1} \overrightarrow{v}^n_{M,1} = \overrightarrow{g}_M - G_{M+1,M} \overrightarrow{u}^n_M$$

obtaining

$$\overrightarrow{v}^n_{M,1} = (H^n_{M+1,M+1})^{-1} (\overrightarrow{g}_M - G_{M+1,M} \overrightarrow{u}^n_M).$$

With $\overrightarrow{v}^n_{M,1}$ now known, we solve the system

$$A^n_M \overrightarrow{u}^n_{M,0} = \overrightarrow{f}^n_M - F^n_{M,M+1} \overrightarrow{v}^n_{M,1},$$

obtaining

$$\overrightarrow{u}^n_{M,0} = (A^n_M)^{-1} (\overrightarrow{f}_M - F_{M,M+1} \overrightarrow{v}^n_{M,0}).$$

Finally, we set

$$\overrightarrow{u}^n_{M,1} = \begin{bmatrix} \overrightarrow{u}^n_{M,0} \\ \overrightarrow{v}^n_{M,1} \end{bmatrix} = \begin{bmatrix} (\xi^n_{i'j'})_{(2^M-1)\times 1} \\ (\eta^n_{i'j'})_{2^M \times 1} \end{bmatrix}$$

then form the linear combination

$$u^n_{M+1} \approx u^n_{M,1} := \sum_{i'j'=10}^{M\ (2^{M-1}-1)} \xi^n_{i'j'} w_{i'j'} + \sum_{i'j'=(M+1)\ 0}^{(M+1)\ (2^M-1)} \eta^n_{i'j'} w_{i'j'}. \tag{40}$$

To write the codes for the implementation, some changes of notation are made to simplify the computer codes. To do this, we first define the matrix

$$A_M = [(w_{ij}, w_{i'j'})]_{(2^M-1)\times(2^M-1)},$$

where $(w_{ij}, w_{i'j'}) = \int_0^1 w_{ij} w_{i'j'}\, dx$. Also, we define

$$\begin{aligned} F_M &= [(w_{ij}, w_{i'j'})]_{(2^M-1)\times 2^M} \\ G_M &= [(w_{ij}, w_{i'j'})]_{2^M\times(2^M-1)} \\ H_M &= [(w_{ij}, w_{i'j'})]_{2^M\times 2^M}. \end{aligned}$$

Further, let

$$A_M + k_n I := A^n_M, \quad F_M := F^n_{M,M+1}, \quad G_M := G^n_{M+1,M}, \quad H_M + k_n I := H^n_{M+1,M+1},$$

where the matrices I in $A_M + k_n I$ and I in $H_M + k_n I$ are identity matrices with the same dimensions as A_M and H_M, respectively. The dimension of I is clear from the context. Thus

$$A^n_{M+1} = \begin{bmatrix} A_M + k_n I & F_M \\ G_M & H_M + k_n I \end{bmatrix}.$$

Also, we need

$$\vec{f}_M = \left[\int_{I_n} (f, w_{i'j'}) dt\right]_{(2^M-1)\times 1}$$

Finally, we index the "wavelets" used in the codes with single subscripts rather than double subscripts for simplicity. These notational changes simplify the coding for the implementation.

Next we detail this multilevel procedure as it applies to the implementation of the constant in time case of the Discontinuous Galerkin Method.

VI.2 CONSTANT IN TIME CASE

For the constant in time case, equation (10) on the $M+1$ spatial resolution level becomes

$$(A_{M+1} + k_n I)\overrightarrow{\xi}^n_{M+1} = A_{M+1}\overrightarrow{\xi}^{n-1}_{M+1} + \overrightarrow{f}_{M+1} \tag{41}$$

and the $M+1$ level solution is calculated as

$$\overrightarrow{\xi}^n_{M+1} = (A_{M+1} + k_n I)^{-1}[A_{M+1}\overrightarrow{\xi}^{n-1}_{M+1} + \overrightarrow{f}_{M+1}]$$

where

$$\overrightarrow{\xi}^n_{M+1} = \left[\xi^n_j\right]_{(2^{M+1}-1)\times 1}$$

and

$$\overrightarrow{f}_{M+1} = \left[\int_{I_n}(f, w_j)dt\right]_{(2^{M+1}-1)\times 1}.$$

The numerical solution on the $M+1$ level is written as

$$U^n_{M+1}(x,t) = \sum_{m=1}^{2^{M+1}-1} \xi^n_m w_m(x).$$

For the multilevel method, we replace A_{M+1} with

$$\begin{bmatrix} A_M & F_M \\ G_M & H_M \end{bmatrix}$$

such that

$$A_{M+1} + k_n I = \begin{bmatrix} A_M + k_n I & F_M \\ G_M & H_M + k_n I \end{bmatrix}.$$

Also, we write

$$\overrightarrow{\xi}^n_{M+1} = \begin{bmatrix} \overrightarrow{\xi}^n_M \\ \overrightarrow{\eta}^n_M \end{bmatrix}$$

where

$$\overrightarrow{\xi}^n_M = [\xi^n_m]_{(2^M-1)\times 1}, \quad \overrightarrow{\eta}^n_M = [\eta^n_m]_{(2^M)\times 1}$$

and

$$\overrightarrow{f}_{M+1} = \begin{bmatrix} \overrightarrow{f}_M \\ \overrightarrow{g}_M \end{bmatrix}$$

for

$$\vec{f}_M = \left[\int_{I_n} (f, w_j)dt\right]_{(2^M-1)\times 1}, \quad \vec{g}_M = \left[\int_{I_n} (f, w_j)dt\right]_{2^M\times 1}.$$

The system (41) becomes the system

$$(A_M + k_n I)\vec{\xi}^{\,n}_M + F_M \vec{\eta}^{\,n}_M = A_M \vec{\xi}^{\,n-1}_M + F_M \vec{\eta}^{\,n-1}_M + \vec{f}_M$$

$$G_M \vec{\xi}^{\,n}_M + (H_M + k_n I)\vec{\eta}^{\,n}_M = G_M \vec{\xi}^{\,n-1}_M + H_M \vec{\eta}^{\,n-1}_M + \vec{g}_M$$

which may be written as

$$\vec{\xi}^{\,n}_M = (A_M + k_n I)^{-1}[A_M \vec{\xi}^{\,n-1}_M + F_M \vec{\eta}^{\,n-1}_M + \vec{f}_M - F_M \vec{\eta}^{\,n}_M]$$

$$\vec{\eta}^{\,n}_M = (H_M + k_n I)^{-1}[G_M \vec{\xi}^{\,n-1}_M + H_M \vec{\eta}^{\,n-1}_M + \vec{g}_M - G_M \vec{\xi}^{\,n}_M].$$

The numerical solution is written as

$$U^n_{M+1}(x,t) = \sum_{m=1}^{2^M-1} \xi^n_m w_m(x) + \sum_{m=2^M}^{2^{(M+1)}-1} \eta^n_m w_m(x).$$

To determine the initial values for the scalars, note that on the $M+1$ level, we have

$$U^0 = \sum_{m=1}^{2^M-1} \xi^0_m w_m(x) + \sum_{m=2^M}^{2^{(M+1)}-1} \eta^0_m w_m(x).$$

The basic algorithm for the multilevel method that was used is as follows.

VI.2.1 Multilevel Algorithm-Constant in Time Case

Step 1. Calculate $A_M, F_M, G_M, H_M, \vec{f}_M, \vec{g}_M$, determine the initial values $\vec{\xi}^{\,0}_M$ and $\vec{\eta}^{\,0}_M$.

Step 2. Set $\vec{\xi}^{\,n-1}_M = \vec{\xi}^{\,0}_M, \quad \vec{\eta}^{\,n-1}_M = \vec{\eta}^{\,0}_M$.

Step 3. Main Loop For $n = 1$ to N:

1. Choose k_n.

2. Calculate Mth Level Solution.

$$\overline{\vec{\xi}^{\,n}_M} = (A_M + k_n I)^{-1}[A_M \vec{\xi}^{\,n-1}_M + \vec{f}_M]$$

3. Calculate Multilevel Solution.

$$\overrightarrow{\eta}_M^n = (H_M + k_n I)^{-1}[G_M \overrightarrow{\xi}_M^{n-1} + H_M \overrightarrow{\eta}_M^{n-1} + \overrightarrow{g}_M - G_M \overrightarrow{\xi}_M^n]$$

$$\overrightarrow{\xi}_M^n = (A_M + k_n I)^{-1}[A_M \overrightarrow{\xi}_M^{n-1} + F_M \overrightarrow{\eta}_M^{n-1} + \overrightarrow{f}_M - F_M \overrightarrow{\eta}_M^n]$$

4. Define

$$\overrightarrow{U}_{M+1}^n \approx \overrightarrow{u}_{M,1}^n = \begin{bmatrix} \overrightarrow{\xi}_M^n \\ \overrightarrow{\eta}_M^n \end{bmatrix},$$

$$U_{M+1}^n \approx u_{M,1}^n = \sum_{m=1}^{2^M-1} \xi_m^n w_m(x) + \sum_{m=2^M}^{2^{(M+1)}-1} \eta_m^n w_m(x)$$

5. Update $\overrightarrow{\xi}_M^{n-1}$ and $\overrightarrow{\eta}_M^{n-1}$.

$$\overrightarrow{\xi}_M^{n-1} = \overrightarrow{\xi}_M^n, \quad \overrightarrow{\eta}_M^{n-1} = \overrightarrow{\eta}_M^n$$

6. Update n.

$$n = n + 1$$

End Main Loop

Next we detail the implementation of the multilevel procedure as it applies to the constant in time case of the Discontinuous Galerkin Method.

VI.3 LINEAR IN TIME CASE

The linear in time case is more complex, although the basic plan is similar. The numerical solution on the $M+1$ level is given by

$$\begin{aligned} U_{M+1}^n(x,t) &= \overline{\phi}_n(x) + \frac{t - t_{n-1}}{k_n}\overline{\psi}_n(x) \\ &= \sum_{m=1}^{2^{M+1}-1} \xi_m^{\overline{\phi},n} w_m(x) + \frac{t-t_{n-1}}{k_n}\left[\sum_{m=1}^{2^{M+1}-1} \xi_m^{\overline{\psi},n} w_m(x)\right] \end{aligned}$$

where

$$\overline{\phi}_n(x) := \sum_{m=1}^{2^{M+1}-1} \xi_m^{\overline{\phi},n} w_m(x),$$

$$\overline{\psi}_n(x) := \sum_{m=1}^{2^{M+1}-1} \xi_m^{\overline{\psi},n} w_m(x).$$

The scalars $\xi_m^{\bar{\phi},n}$ and $\xi_m^{\bar{\psi},n}$, $m = 1, 2, \cdots, 2^{M+1} - 1$ are determined from the system

$$(A_{M+1} + k_n I)\overrightarrow{\xi}^{\bar{\phi},n}_{M+1} + (A_{M+1} + \frac{1}{2}k_n I)\overrightarrow{\xi}^{\bar{\psi},n}_{M+1} = f_{M+1} + A_{M+1}(\overrightarrow{\xi}^{\bar{\phi},n-1}_{M+1} + \overrightarrow{\xi}^{\bar{\psi},n-1}_{M+1})$$

$$\frac{1}{2}k_n I\overrightarrow{\xi}^{\bar{\phi},n}_{M+1} + (\frac{1}{2}A_{M+1} + \frac{1}{3}k_n I)\overrightarrow{\xi}^{\bar{\psi},n}_{M+1} = k_n^{-1}\overrightarrow{f}^{n}_{M+1}$$

with

$$\overrightarrow{f}_{M+1} = \left[\int_{I_n} (f(t), w_j)dt\right]_{(2^{M+1}-1)\times 1},$$

$$\overrightarrow{f}^{n}_{M+1} = \left[\int_{I_n} (t - t_{n-1})(f(t), w_j)dt\right]_{(2^{M+1}-1)\times 1},$$

$$\overrightarrow{\xi}^{\bar{\phi},n}_{M+1} = \left[\xi_j^{\bar{\phi},n}\right]_{(2^{M+1}-1)\times 1}, \quad \overrightarrow{\xi}^{\bar{\psi},n}_{M+1} = \left[\xi_j^{\bar{\psi},n}\right]_{(2^{M+1}-1)\times 1}$$

using the same definitions for A_M and I as before. This may be written in matrix form as

$$\begin{bmatrix} A_{M+1} + k_n I & A_{M+1} + \frac{1}{2}k_n I \\ \frac{1}{2}k_n I & \frac{1}{2}A_{M+1} + \frac{1}{3}k_n I \end{bmatrix} \begin{bmatrix} \overrightarrow{\xi}^{\bar{\phi},n}_{M+1} \\ \overrightarrow{\xi}^{\bar{\psi},n}_{M+1} \end{bmatrix} = \begin{bmatrix} \overrightarrow{f}_{M+1} + A_{M+1}(\overrightarrow{\xi}^{\bar{\phi},n-1}_{M+1} + \overrightarrow{\xi}^{\bar{\psi},n-1}_{M+1}) \\ k_n^{-1}\overrightarrow{f}^{n}_{M+1} \end{bmatrix}$$

with the solution of this system being

$$\begin{bmatrix} \overrightarrow{\xi}^{\bar{\phi},n}_{M+1} \\ \overrightarrow{\xi}^{\bar{\psi},n}_{M+1} \end{bmatrix} = \begin{bmatrix} A_{M+1} + k_n I & A_{M+1} + \frac{1}{2}k_n I \\ \frac{1}{2}k_n I & \frac{1}{2}A_{M+1} + \frac{1}{3}k_n I \end{bmatrix}^{-1} \begin{bmatrix} \overrightarrow{f}_{M+1} + A_{M+1}(\overrightarrow{\xi}^{\bar{\phi},n-1}_{M+1} + \overrightarrow{\xi}^{\bar{\psi},n-1}_{M+1}) \\ k_n^{-1}\overrightarrow{f}^{n}_{M+1} \end{bmatrix}$$

where

$$\overrightarrow{u}^{n}_{M+1} = \begin{bmatrix} \overrightarrow{\xi}^{\bar{\phi},n}_{M+1} \\ \overrightarrow{\xi}^{\bar{\psi},n}_{M+1} \end{bmatrix},$$

using

$$\overrightarrow{\xi}^{\overline{\phi},n}_{M+1} = \left[\xi_j^{\overline{\phi},n}\right]_{(2^{M+1}-1)\times 1}, \quad \overrightarrow{\xi}^{\overline{\psi},n}_{M+1} = \left[\xi_j^{\overline{\psi},n}\right]_{(2^{M+1}-1)\times 1}.$$

For the multilevel method, we replace A_{M+1} with

$$\begin{bmatrix} A_M & F_M \\ G_M & H_M \end{bmatrix}$$

such that

$$A_{M+1} + k_n I = \begin{bmatrix} A_M + k_n I & F_M \\ G_M & H_M + k_n I \end{bmatrix},$$

$$A_{M+1} + \frac{1}{2}k_n I = \begin{bmatrix} A_M + \frac{1}{2}k_n I & F_M \\ G_M & H_M + \frac{1}{2}k_n I \end{bmatrix},$$

$$\frac{1}{2}k_n I = \begin{bmatrix} \frac{1}{2}k_n I & 0 \\ 0 & \frac{1}{2}k_n I \end{bmatrix},$$

and

$$\frac{1}{2}A_{M+1} + \frac{1}{3}k_n I = \begin{bmatrix} \frac{1}{2}A_M + \frac{1}{3}k_n I & \frac{1}{2}F_M \\ \frac{1}{2}G_M & \frac{1}{2}H_M + \frac{1}{3}k_n I \end{bmatrix}.$$

We write

$$\overrightarrow{u}^{n}_{M+1} = \begin{bmatrix} \overrightarrow{\xi}^{\overline{\phi},n}_{M+1} \\ \overrightarrow{\xi}^{\overline{\psi},n}_{M+1} \end{bmatrix} = \begin{bmatrix} \overrightarrow{\xi}^{\overline{\phi},n}_{M} \\ \overrightarrow{\eta}^{\overline{\phi},n}_{M} \\ \overrightarrow{\xi}^{\overline{\psi},n}_{M} \\ \overrightarrow{\eta}^{\overline{\psi},n}_{M} \end{bmatrix},$$

where

$$\overrightarrow{\xi}^{\overline{\phi},n}_{M} = \left[\xi_j^{\overline{\phi},n}\right]_{(2^{M}-1)\times 1}, \quad \overrightarrow{\eta}^{\overline{\phi},n}_{M} = \left[\eta_j^{\overline{\phi},n}\right]_{[(2^{M+1}-1)-(2^{M}-1)]\times 1},$$

$$\overrightarrow{\xi}_M^{\overline{\psi},n} = \left[\xi_j^{\overline{\psi},n}\right]_{(2^M-1)\times 1}, \quad \overrightarrow{\eta}_M^{\overline{\psi},n} = \left[\eta_j^{\overline{\psi},n}\right]_{[(2^{M+1}-1)-(2^M-1)]\times 1}.$$

Also, we have

$$\overrightarrow{f}_{M+1} = \begin{bmatrix} \overrightarrow{f}_M \\ \overrightarrow{g}_M \end{bmatrix}, \qquad \overrightarrow{f}^n_{M+1} = \begin{bmatrix} \overrightarrow{f}^n_M \\ \overrightarrow{g}^n_M \end{bmatrix},$$

where

$$\overrightarrow{f}_M = \left[\int_{I_n} (f(t), w_j) dt\right]_{(2^M-1)\times 1}$$

$$\overrightarrow{g}_M = \left[\int_{I_n} (f(t), w_j) dt\right]_{2^M\times 1}$$

$$\overrightarrow{f}^n_M = \left[\int_{I_n} (t - t_{n-1})(f(t), w_j) dt\right]_{(2^M-1)\times 1}.$$

and

$$\overrightarrow{g}^n_M = \left[\int_{I_n} (t - t_{n-1})(f(t), w_j) dt\right]_{2^M\times 1}.$$

After some basic calculation the system becomes a new system consisting of the two matrix equations

$$\begin{bmatrix} A_M + k_n I & A_M + \frac{1}{2}k_n I \\ \frac{1}{2}k_n I & \frac{1}{2}A_M + \frac{1}{3}k_n I \end{bmatrix} \begin{bmatrix} \overrightarrow{\xi}_M^{\overline{\phi},n} \\ \overrightarrow{\xi}_M^{\overline{\psi},n} \end{bmatrix} =$$

$$\begin{bmatrix} \overrightarrow{f}_M + A_M(\overrightarrow{\xi}_M^{\overline{\phi},n-1} + \overrightarrow{\xi}_M^{\overline{\psi},n-1}) + F_M(\overrightarrow{\eta}_M^{\overline{\phi},n-1} + \overrightarrow{\eta}_M^{\overline{\psi},n-1}) - F_M(\overrightarrow{\eta}_M^{\overline{\phi},n} + \overrightarrow{\eta}_M^{\overline{\psi},n}) \\ k_n^{-1}\overrightarrow{f}^n_M - \frac{1}{2}F_M\overrightarrow{\eta}_M^{\overline{\psi},n} \end{bmatrix}.$$

and

$$\begin{bmatrix} H_M + k_n I & H_M + \frac{1}{2}k_n I \\ \frac{1}{2}k_n I & \frac{1}{2}H_M + \frac{1}{3}k_n I \end{bmatrix} \begin{bmatrix} \overrightarrow{\eta}_M^{\overline{\phi},n} \\ \overrightarrow{\eta}_M^{\overline{\psi},n} \end{bmatrix} =$$

$$\begin{bmatrix} \overrightarrow{g}_M + G_M(\overrightarrow{\xi}_M^{\overline{\phi},n-1} + \overrightarrow{\xi}_M^{\overline{\psi},n-1}) + H_M(\overrightarrow{\eta}_M^{\overline{\phi},n-1} + \overrightarrow{\eta}_M^{\overline{\psi},n-1}) - G_M(\overrightarrow{\eta}_M^{\overline{\phi},n} + \overrightarrow{\eta}_M^{\overline{\psi},n}) \\ k_n^{-1}\overrightarrow{g}^n_M - \frac{1}{2}G_M\overrightarrow{\eta}_M^{\overline{\psi},n} \end{bmatrix}$$

for which the solution is given by

$$\begin{bmatrix} \overrightarrow{\xi}_M^{\bar{\phi},n} \\ \overrightarrow{\xi}_M^{\bar{\psi},n} \end{bmatrix} = \begin{bmatrix} A_M + k_n I & A_M + \frac{1}{2}k_n I \\ \frac{1}{2}k_n I & \frac{1}{2}A_M + \frac{1}{3}k_n I \end{bmatrix}^{-1} .$$

$$\begin{bmatrix} \overrightarrow{f}_M + A_M(\overrightarrow{\xi}_M^{\bar{\phi},n-1} + \overrightarrow{\xi}_M^{\bar{\psi},n-1}) + F_M(\overrightarrow{\eta}_M^{\bar{\phi},n-1} + \overrightarrow{\eta}_M^{\bar{\psi},n-1}) - F_M(\overrightarrow{\eta}_M^{\bar{\phi},n} + \overrightarrow{\eta}_M^{\bar{\psi},n}) \\ k_M^{-1}\overrightarrow{f}_M^{n} - \frac{1}{2}F_M \overrightarrow{\eta}_M^{\bar{\psi},n} \end{bmatrix}$$

and

$$\begin{bmatrix} \overrightarrow{\eta}_M^{\bar{\phi},n} \\ \overrightarrow{\eta}_M^{\bar{\psi},n} \end{bmatrix} = \begin{bmatrix} H_M + k_n I & H_M + \frac{1}{2}k_n I \\ \frac{1}{2}k_n I & \frac{1}{2}H_M + \frac{1}{3}k_n I \end{bmatrix}^{-1} .$$

$$\begin{bmatrix} \overrightarrow{g}_M + G_M(\overrightarrow{\xi}_M^{\bar{\phi},n-1} + \overrightarrow{\xi}_M^{\bar{\psi},n-1}) + H_M(\overrightarrow{\eta}_M^{\bar{\phi},n-1} + \overrightarrow{\eta}_M^{\bar{\psi},n-1}) - G_M(\overrightarrow{\xi}_M^{\bar{\phi},n} + \overrightarrow{\xi}_M^{\bar{\psi},n}) \\ k_n^{-1}\overrightarrow{g}_M^{n} - \frac{1}{2}G_M \overrightarrow{\xi}_M^{\bar{\psi},n} \end{bmatrix} .$$

The numerical solution is written as

$$U_{M+1}^n(x,t) = \sum_{m=1}^{2^M-1} \xi_m^{\bar{\phi},n} w_m(x) + \sum_{m=2^M}^{2^{(M+1)}-1} \eta_m^{\bar{\phi},n} w_m(x)$$

$$+\frac{t-t_{n-1}}{k_n}\left[\sum_{m=1}^{2^M-1} \xi_m^{\bar{\psi},n} w_m(x) + \sum_{m=2^M}^{2^{(M+1)}-1} \eta_m^{\bar{\psi},n} w_m(x)\right].$$

The basic algorithm for the multilevel method that was used is as follows.

VI.3.1 Multilevel Algorithm-Linear in Time Case

Step 1 Calculate A_M, F_M, G_M, H_M, determine the initial values $\overrightarrow{\xi}_M^{\bar{\phi},0}$, $\overrightarrow{\xi}_M^{\bar{\psi},0}$ and $\overrightarrow{\eta}_M^{\bar{\phi},0}$, $\overrightarrow{\eta}_M^{\bar{\psi},0}$.

Step 2 Set

$$\overrightarrow{\xi}_M^{\bar{\phi},n-1} = \overrightarrow{\xi}_M^{\bar{\phi},0}, \quad \overrightarrow{\xi}_M^{\bar{\psi},n-1} = \overrightarrow{\xi}_M^{\bar{\psi},0}, \quad \overrightarrow{\eta}_M^{\bar{\phi},n-1} = \overrightarrow{\eta}_M^{\bar{\phi},0}, \quad \overrightarrow{\eta}_M^{\bar{\psi},n-1} = \overrightarrow{\eta}_M^{\bar{\psi},0}.$$

Step 3 Main Loop: For $n = 1$ to N:

1. Calculate $\overrightarrow{f}_M^{n}$, $\overrightarrow{f}_M$, $\overrightarrow{g}_M^{n}$, $\overrightarrow{g}_M$, and choose k_n.

2. Calculate Mth Level Solution.

$$\begin{bmatrix} \overline{\overrightarrow{\xi}^{\bar{\phi},n}_M} \\ \overline{\overrightarrow{\xi}^{\bar{\psi},n}_M} \end{bmatrix} = \begin{bmatrix} A_M + k_n I & A_M + \frac{1}{2}k_n I \\ \frac{1}{2}k_n I & \frac{1}{2}A_M + \frac{1}{3}k_n I \end{bmatrix}^{-1} \begin{bmatrix} \overrightarrow{f}_M + A_M(\overrightarrow{\xi}^{\bar{\phi},n-1}_M + \overrightarrow{\xi}^{\bar{\psi},n-1}_M) \\ k_n^{-1}\overrightarrow{f}^n_M \end{bmatrix}.$$

3. Calculate Multilevel Solution.

$$\begin{bmatrix} \overrightarrow{\eta}^{\bar{\phi},n}_M \\ \overrightarrow{\eta}^{\bar{\psi},n}_M \end{bmatrix} = \begin{bmatrix} H_M + k_n I & H_M + \frac{1}{2}k_n I \\ \frac{1}{2}k_n I & \frac{1}{2}H_M + \frac{1}{3}k_n I \end{bmatrix}^{-1} \cdot$$

$$\begin{bmatrix} \overrightarrow{g}_M + G_M(\overrightarrow{\xi}^{\bar{\phi},n-1}_M + \overrightarrow{\xi}^{\bar{\psi},n-1}_M) + H_M(\overrightarrow{\eta}^{\bar{\phi},n-1}_M + \overrightarrow{\eta}^{\bar{\psi},n-1}_M) - G_M(\overline{\overrightarrow{\xi}^{\bar{\phi},n}_M} + \overline{\overrightarrow{\xi}^{\bar{\psi},n}_M}) \\ k_n^{-1}\overrightarrow{g}^n_M - \frac{1}{2}G_M \overline{\overrightarrow{\xi}^{\bar{\psi},n}_M} \end{bmatrix},$$

$$\begin{bmatrix} \overrightarrow{\xi}^{\bar{\phi},n}_M \\ \overrightarrow{\xi}^{\bar{\psi},n}_M \end{bmatrix} = \begin{bmatrix} A_M + k_n I & A_M + \frac{1}{2}k_n I \\ \frac{1}{2}k_n I & \frac{1}{2}A_M + \frac{1}{3}k_n I \end{bmatrix}^{-1} \cdot$$

$$\begin{bmatrix} \overrightarrow{f}_M + A_M(\overrightarrow{\xi}^{\bar{\phi},n-1}_M + \overrightarrow{\xi}^{\bar{\psi},n-1}_M) + F_M(\overrightarrow{\eta}^{\bar{\phi},n-1}_M + \overrightarrow{\eta}^{\bar{\psi},n-1}_M) - F_M(\overrightarrow{\eta}^{\bar{\phi},n}_M + \overrightarrow{\eta}^{\bar{\psi},n}_M) \\ k_n^{-1}\overrightarrow{f}^n_M - \frac{1}{2}F_M \overrightarrow{\eta}^{\bar{\psi},n}_M \end{bmatrix}.$$

4. Define

$$\overrightarrow{U}^n_{M+1} \approx \overrightarrow{u}^n_{M,1} = \begin{bmatrix} \overrightarrow{\xi}^{\bar{\phi},n}_{M+1} \\ \overrightarrow{\xi}^{\bar{\psi},n}_{M+1} \end{bmatrix} = \begin{bmatrix} \overrightarrow{\xi}^{\bar{\phi},n}_M \\ \overrightarrow{\eta}^{\bar{\phi},n}_M \\ \overrightarrow{\xi}^{\bar{\psi},n}_M \\ \overrightarrow{\eta}^{\bar{\psi},n}_M \end{bmatrix},$$

$$U^n_{M+1} \approx u^n_{M,1} = \sum_{m=1}^{2^M-1} \xi^{\bar{\phi},n}_m w_m(x) + \sum_{m=2^M}^{2^{(M+1)}-1} \eta^{\bar{\phi},n}_m w_m(x)$$

$$+\frac{t-t_{n-1}}{k_n}\left[\sum_{m=1}^{2^M-1} \xi^{\bar{\psi},n}_m w_m(x) + \sum_{m=2^M}^{2^{(M+1)}-1} \eta^{\bar{\psi},n}_m w_m(x)\right].$$

5. Update $\overrightarrow{\xi}^{\bar{\phi},n-1}_M$, $\overrightarrow{\eta}^{\bar{\phi},n-1}_M$, $\overrightarrow{\xi}^{\bar{\psi},n-1}_M$, and $\overrightarrow{\eta}^{\bar{\psi},n-1}_M$.

$$\overrightarrow{\xi}^{\bar{\phi},n-1}_M = \overrightarrow{\xi}^{\bar{\phi},n}_M, \quad \overrightarrow{\eta}^{\bar{\phi},n-1}_M = \overrightarrow{\eta}^{\bar{\phi},n}_M, \quad \overrightarrow{\xi}^{\bar{\psi},n-1}_M = \overrightarrow{\xi}^{\bar{\psi},n}_M, \quad \overrightarrow{\eta}^{\bar{\psi},n-1}_M = \overrightarrow{\eta}^{\bar{\psi},n}_M.$$

6. Update n.

$$n = n + 1$$

End Main Loop

VI.4 ERROR ESTIMATE

Next we show that Theorem 4.2 applies to this scheme. Recall from Chapter IV hypotheses (I) and (II):

(I) $B_{m,1}^{-1}$ exists.

(II) $B_{m,1}^{-1}C_{m,1}$ is uniformly bounded.

We will show hypotheses (I) and (II) are satisfied. Also, we use the linear wavelet basis functions for $H_0^1(0,1)$ developed in Section II.4.2. First, by the definitions of $B_{m,1}$ and $C_{m,1}$ given by (18) and (19), respectively, we have

$$B_{m,1} := \begin{bmatrix} A_m^n & F_{m,m+1}^n \\ 0 & H_{m+1,m+1}^n \end{bmatrix}$$

and

$$C_{m,1} := \begin{bmatrix} 0 & 0 \\ G_{m+1,m}^n & 0 \end{bmatrix}$$

where the matrices A_m^n, $F_{m,m+1}^n$, $G_{m+1,m}^n$, and $H_{m+1,m+1}^n$ are defined as

$$A_m^n = [a_{iji'j'}]_{(2^m-1)\times(2^m-1)},$$

for $ij = 1\ 0, \cdots, m\ (2^{m-1}-1),\ i'j' = 1\ 0, \cdots, m\ (2^{m-1}-1)$,

$$F_{m,m+1}^n = [a_{iji'j'}]_{(2^m-1)\times(2^m)},$$

for $ij = 1\ 0, \cdots, m\ (2^{m-1}-1),\ i'j' = (m+1)\ 0, \cdots, (m+1)\ (2^m-1)$,

$$G_{m+1,m}^n = [a_{iji'j'}]_{(2^m)\times(2^m-1)},$$

for $ij = (m+1)\ 0, \cdots, (m+1)\ (2^m-1),\ i'j' = 1\ 0, \cdots, m\ (2^{m-1}-1)$, and

$$H_{m+1,m+1}^n = [a_{iji'j'}]_{(2^m)\times(2^m)},$$

for $ij = (m+1)\ 0, \cdots, (m+1)\ (2^m-1),\ i'j' = (m+1)\ 0, \cdots, (m+1)\ (2^m-1)$.

To show $B_{m,1}^{-1}$ exists, we need only show

$$\det B_{m,1} \neq 0,$$

since $B_{m,1}$ is a $(2^{m+1}-1)$ by $(2^{m+1}-1)$ matrix. Using a well known property of matrix determinants, we have

$$\det B_{m,1} = \det \begin{bmatrix} A_m^n & F_{m,m+1}^n \\ 0 & H_{m+1,m+1}^n \end{bmatrix} = \det A_m^n \det H_{m+1,m+1}^n.$$

Since the Discontinuous Galerkin Method has a unique solution by the discussion on page 183 of [4], we must have

$$\det A_m^n \neq 0.$$

Using the linear wavelet basis, we have

$$w_{ij}(x) = \begin{cases} (\frac{1}{\sqrt{2}})^{i-1}(2^{i-1}x - j) & \text{when } \frac{j}{2^{i-1}} \leq x < \frac{j}{2^{i-1}} + \frac{1}{2^i}, \\ (\frac{1}{\sqrt{2}})^{i-1}(j+1-2^{i-1}x) & \text{when } \frac{j}{2^{i-1}} + \frac{1}{2^i} \leq x \leq \frac{j+1}{2^{i-1}}, \\ 0 & \text{otherwise.} \end{cases}$$

Thus,

$$\begin{aligned} (w_{ij}, w_{ij}) &= \int_{\frac{j}{2^{i-1}}}^{\frac{j}{2^{i-1}}+\frac{1}{2^i}} [(\frac{1}{\sqrt{2}})^{i-1}(2^{i-1}x - j)]^2\, dx \\ &\quad + \int_{\frac{j}{2^{i-1}}+\frac{1}{2^i}}^{\frac{j+1}{2^{i-1}}} [(\frac{1}{\sqrt{2}})^{i-1}(j+1-2^{i-1}x)]^2\, dx \\ &= \frac{1}{3}\left(\frac{1}{4}\right)^i \end{aligned}$$

so

$$H_{m+1,m+1}^n = \left(\frac{1}{3}\left(\frac{1}{4}\right)^{m+1} + k_n\right) I.$$

This implies

$$\det H_{m+1,m+1}^n \neq 0$$

hence

$$\det B_{m,1} \neq 0.$$

Thus hypothesis (I) is satisfied.

Now we look at the second requirement. Using blockwise inversion on $B_{m,1}$, we have

$$B_{m,1}^{-1} = \begin{bmatrix} A_m^n & F_{m,m+1}^n \\ 0 & H_{m+1,m+1}^n \end{bmatrix}^{-1} = \begin{bmatrix} (A_m^n)^{-1} & -(A_m^n)^{-1}F_{m,m+1}^n(H_{m+1,m+1}^n)^{-1} \\ 0 & (H_{m+1,m+1}^n)^{-1} \end{bmatrix}$$

so

$$
\begin{aligned}
B_{m,1}^{-1} C_{m,1} &= \begin{bmatrix} (A_m^n)^{-1} & -(A_m^n)^{-1} F_{m,m+1}^n (H_{m+1,m+1}^n)^{-1} \\ 0 & (H_{m+1,m+1}^n)^{-1} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ G_{m+1,m}^n & 0 \end{bmatrix} \\
&= \begin{bmatrix} -(A_m^n)^{-1} F_{m,m+1}^n (H_{m+1,m+1}^n)^{-1} G_{m+1,m}^n & 0 \\ (H_{m+1,m+1}^n)^{-1} G_{m+1,m}^n & 0 \end{bmatrix}.
\end{aligned}
$$

Because

$$H_{m+1,m+1}^n = \left(\frac{1}{3}\left(\frac{1}{4}\right)^{m+1} + k_n\right) I,$$

we have

$$(H_{m+1,m+1}^n)^{-1} = \frac{3(4^{m+1})}{1+3(4^{m+1})k_n} I$$

so

$$\|(H_{m+1,m+1}^n)^{-1}\| = \frac{3(4^{m+1})}{1+3(4^{m+1})k_n}.$$

Further, since $\|G_{m+1,m}^n\| = \|(F_{m,m+1}^n)^T\| = \|F_{m,m+1}^n\|$ and the maximum row sum of $F_{m,m+1}^n$ is the first row sum, we have

$$
\begin{aligned}
\|G_{i+1,i}^n\| &= \|F_{i,i+1}^n\| \\
&= \sum_{j=0}^{2^{i-2}-1} (w_{10}, w_{ij}) \\
&= 2 \sum_{j=0}^{2^{i-2}-1} [\int_{\frac{j}{2^{i-1}}}^{\frac{j}{2^{i-1}}+\frac{1}{2^i}} (\frac{1}{\sqrt{2}})^{i-1} (2^{i-1}x - j) x \, dx \\
&\quad + \int_{\frac{j}{2^{i-1}}+\frac{1}{2^i}}^{\frac{j+1}{2^{i-1}}} (\frac{1}{\sqrt{2}})^{i-1} (j+1-2^{i-1}x) x \, dx] \\
&= 2^{\frac{3}{2}-\frac{1}{2}i} \sum_{j=0}^{2^{i-2}-1} \left[(\frac{2^{i-1}}{3} x^3 - \frac{1}{2} j x^2)|_{j2^{1-i}}^{(2j+1)2^{-i}} + (\frac{j+1}{2} x^2 - \frac{2^{i-1}}{3} j x^3)|_{(2j+1)2^{-i}}^{(j+1)2^{1-i}} \right] \\
&= \frac{\sqrt{2}}{16} 2^{-\frac{1}{2}i}.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
\|(H_{m+1,m+1}^n)^{-1} G_{m+1,m}^n\| &\leq \|(H_{m+1,m+1}^n)^{-1}\| \|G_{m+1,m}^n\| \\
&= \left(\frac{3(4^{m+1})}{1+3(4^{m+1})k_n}\right) \left(\frac{\sqrt{2}}{16} 2^{-\frac{1}{2}m}\right) \\
&= \left(\frac{1}{\frac{1}{3}(\frac{1}{4})^{m+1} + k_n}\right) \left(\frac{\sqrt{2}}{16} 2^{-\frac{1}{2}m}\right) \\
&\leq \frac{1}{k_n 2^{\frac{1}{2}m}}.
\end{aligned}
$$

If we let $k_n = 2^{-\frac{1}{2}m}$, we have

$$\|(H^n_{m+1,m+1})^{-1}G^n_{m+1,m}\| \leq 1$$

so the block $(H^n_{m+1,m+1})^{-1}G^n_{m+1,m}$ is uniformly bounded.

Next we look at the uniform bounding of the other block of $B^{-1}_{m,1}C_{m,1}$, namely

$$-(A^n_m)^{-1}F^n_{m,m+1}(H^n_{m+1,m+1})^{-1}G^n_{m+1,m}.$$

We have

$$\begin{aligned}
&\|-(A^n_m)^{-1}F^n_{m,m+1}(H^n_{m+1,m+1})^{-1}G^n_{m+1,m}\| \\
&\leq \|-(A^n_m)^{-1}\|\|F^n_{m,m+1}\|\|(H^n_{m+1,m+1})^{-1}G^n_{m+1,m}\| \\
&\leq \|-(A^n_m)^{-1}\|\|F^n_{m,m+1}\|C
\end{aligned}$$

for some constant C, since $(H^n_{m+1,m+1})^{-1}G^n_{m+1,m}$ was shown to be uniformly bounded. Further, since

$$\|F^n_{m,m+1}\| = \|G^n_{m+1,m}\| = \frac{\sqrt{2}}{16}2^{-\frac{1}{2}m} = \frac{\sqrt{2}}{16}\left(\frac{1}{\sqrt{2}}\right)^m \tag{42}$$

we have $\|F^n_{m,m+1}\| \to 0$ as $m \to \infty$.

We now proceed via induction on m to bound $(A^n_m)^{-1}$. Let

$$k_n = \left(\frac{1}{\sqrt{2}}\right)^{m_0}$$

be fixed for some fixed $m_0 > 0$ and let

$$\epsilon_m = \sqrt{2}^{m-1} > 0.$$

For $m_1 = m_0 + 1$, we have $\epsilon_{m_1} = \sqrt{2}^{m_1-1} = \sqrt{2}^{(m_0+1)-1} = \sqrt{2}^{m_0} > 0$ and for values of m_0 used in practice it can been shown by straightforward calculation that

$$\|(A^n_{m_1})^{-1}\| < k_n^{-1} + \epsilon_{m_1}, \tag{43}$$

so the assertion holds for $m_1 = m_0 + 1$. We do this in our implementation.

Assume, for a fixed $k_n = \left(\frac{1}{\sqrt{2}}\right)^{m_0}$, $m_0 > 0$, that there exists $m > m_0 + 1$ and $\epsilon_m = \sqrt{2}^{m-1} > 0$ such that

$$\|(A^n_m)^{-1}\| \leq k_n^{-1} + \epsilon_m. \tag{44}$$

We will show

$$\|(A^n_{m+1})^{-1}\| \le k_n^{-1} + \epsilon_{m+1}.$$

By the definition of A^n_{m+1}, we have

$$\begin{aligned} A^n_{m+1} &= A_{m+1} + k_n I \\ &= \begin{bmatrix} A_m + k_n I & F^n_{m,m+1} \\ G^n_{m+1,m} & H_{m+1,m+1} + k_n I \end{bmatrix} \\ &= \begin{bmatrix} A_m + k_n I & F^n_{m,m+1} \\ (F^n_{m,m+1})^T & (h_m + k_n) I \end{bmatrix}. \end{aligned}$$

Let

$$S^n_{m+1} := \begin{bmatrix} A_m + k_n I & 0 \\ 0 & (h_m + k_n) I \end{bmatrix}.$$

Using blockwise inversion, we get

$$\begin{aligned} I - (S^n_{m+1})^{-1} A^n_{m+1} &= I - \\ &\quad \begin{bmatrix} (A_m + k_n I)^{-1} & 0 \\ 0 & (h_m + k_n)^{-1} I \end{bmatrix} \begin{bmatrix} A_m + k_n I & F^n_{m,m+1} \\ (F^n_{m,m+1})^T & (h_m + k_n) I \end{bmatrix} \\ &= I - \begin{bmatrix} I & (A_m + k_n I)^{-1} F^n_{m,m+1} \\ (h_m + k_n)^{-1} (F^n_{m,m+1})^T & I \end{bmatrix} \\ &= \begin{bmatrix} 0 & -(A_m + k_n I)^{-1} F^n_{m,m+1} \\ -(h_m + k_n)^{-1} (F^n_{m,m+1})^T & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & -(A^n_m)^{-1} F^n_{m,m+1} \\ -(h_m + k_n)^{-1} (F^n_{m,m+1})^T & 0 \end{bmatrix}. \end{aligned}$$

Then

$$\begin{aligned} \|I - (S^n_{m+1})^{-1} A^n_{m+1}\| &\le \left\| \begin{bmatrix} 0 & -(A^n_m)^{-1} F^n_{m,m+1} \\ -(h_m + k_n)^{-1} (F^n_{m,m+1})^T & 0 \end{bmatrix} \right\| \\ &\le (k_n^{-1} + \epsilon_m) \|F^n_{m,m+1}\| \end{aligned}$$

and

$$\begin{aligned} (A^n_{m+1})^{-1} &= \sum_{k=0}^{\infty} [(S^n_{m+1})^{-1} (S^n_{m+1} - A^n_{m+1})]^k (S^n_{m+1})^{-1} \\ &= \sum_{k=0}^{\infty} [I - (S^n_{m+1})^{-1} A^n_{m+1}]^k (S^n_{m+1})^{-1}. \end{aligned}$$

Thus for $m > m_0 + 1$ and $\epsilon_{m+1} > 0$, we have

$$
\begin{aligned}
\|A^n_{m+1})^{-1}\| &= \left\|\sum_{k=0}^{\infty}[I - (S^n_{m+1})^{-1}A^n_{m+1}]^k (S^n_{m+1})^{-1}\right\| \\
&\leq \left\|\sum_{k=0}^{\infty}[I - (S^n_{m+1})^{-1}A^n_{m+1}]^k\right\| \left\|(S^n_{m+1})^{-1}\right\| \\
&\leq \frac{1}{1 - \|I - (S^n_{m+1})^{-1}A^n_{m+1}\|}\left\|(S^n_{m+1})^{-1}\right\| \\
&\leq \frac{k_n^{-1} + \epsilon_m}{1 - (k_n^{-1} + \epsilon_m)\left\|F^n_{m,m+1}\right\|} \\
&= k_n^{-1} + \left(\frac{k_n^{-1} + \epsilon_m}{1 - (k_n^{-1} + \epsilon_m)\left\|F^n_{m,m+1}\right\|} - k_n^{-1}\right) \\
&= k_n^{-1} + \left(\frac{\sqrt{2}^{m_0} + \sqrt{2}^{m-1}}{1 - (\sqrt{2}^{m_0} + \sqrt{2}^{m-1})\frac{\sqrt{2}}{16}\left(\frac{1}{\sqrt{2}}\right)^m} - \sqrt{2}^{m_0}\right) \\
&= k_n^{-1} + \left(\frac{\sqrt{2}^{m_0} + \sqrt{2}^{m-1}}{1 - (\sqrt{2}^{m_0} + \sqrt{2}^{m-1})\frac{\sqrt{2}^{1-m}}{16}} - \sqrt{2}^{m_0}\right) \\
&= k_n^{-1} + \left(\frac{\sqrt{2}^{m_0} + \sqrt{2}^{m-1}}{1 - \frac{(\sqrt{2})^{m_0+1-m}}{16} - \frac{1}{16}} - \sqrt{2}^{m_0}\right) \\
&= k_n^{-1} + \left(\frac{\sqrt{2}^{m_0} + \sqrt{2}^{m-1}}{\frac{15}{16} - \frac{(\sqrt{2})^{m_0+1-m}}{16}} - \sqrt{2}^{m_0}\right) \\
&= k_n^{-1} + \left(\frac{\sqrt{2}^{m_0} + \sqrt{2}^{m-1}}{\frac{15}{16} - \frac{1}{16}\frac{(\sqrt{2})^{m_0+1}}{\sqrt{2}^m}} - \sqrt{2}^{m_0}\right) \\
&< k_n^{-1} + \left(\frac{\sqrt{2}^{m_0} + \sqrt{2}^{m-1}}{\frac{15}{16} - \frac{1}{16}} - \sqrt{2}^{m_0}\right) \\
&= k_n^{-1} + \frac{8}{7}\sqrt{2}^{m_0} + \frac{8}{7}\sqrt{2}^{m-1} - \sqrt{2}^{m_0} \\
&= k_n^{-1} + \frac{1}{7}\sqrt{2}^{m_0} + \frac{8}{7}\sqrt{2}^{m-1} \\
&< k_n^{-1} + \frac{1}{7}\sqrt{2}^{m} + \frac{8}{7}\sqrt{2}^{m-1} \\
&< k_n^{-1} + \frac{1}{7}\sqrt{2}^{m} + \frac{6}{7}\sqrt{2}^{m} \\
&= k_n^{-1} + \sqrt{2}^{m} \\
&= k_n^{-1} + \epsilon_{m+1}.
\end{aligned}
$$

Thus we now have the norm of the block

$$-(A_m^n)^{-1}F_{m,m+1}^n(H_{m+1,m+1}^n)^{-1}G_{m+1,m}^n$$

bounded as $m \to \infty$. Therefore the matrix

$$B_{m,1}^{-1}C_{m,1}$$

is uniformly bounded for these choices of ϵ and k_n, thus satisfying hypothesis (II).

With these two hypotheses having been satisfied, Theorem 4.2 implies there exists a positive integer M such that for $m \geq M$ we have

$$\|u^n - u_{m,1}^n\| \;\leq\; CL_N \max_{1\leq n\leq N} E_{m+1,qn}(u)$$

where u^n is the exact solution at the time step t_n for $n = 1, 2, \cdots, N$, and

$$u_{m,1}^n := u_{m,0}^n + v_{m,1}^n := \sum_{ij=10}^{(m+1)(2^m-1)} u_{ij}w_{ij}$$

is the multilevel solution, using u_{ij} to denote the scaler entries of the column vector $u_{m,1}$, and the w_{ij} are the wavelet basis functions.

VI.5 NUMERICAL EXPERIMENTS

VI.5.1 A Conventional Example

We use the following one-dimensional heat problem: Find u such that

$$u_t(x,t) - u_{xx}(x,t) = 0, \quad 0 < x < 1, \quad 0 < t < .5,$$

$$u(0,t) = u(1,t) = 0, \quad 0 < t < .5,$$

$$u(x,0) = \sin(\pi x), \quad 0 < x < 1.$$

The standard Discontinuous Galerkin Method solution and the Multilevel Method solution of the above problem were calculated so a comparison of accuracies and efficiencies could be made. Grid resolutions of $M = 2, 3, \ldots, 10$ were chosen, with the number of time steps N chosen as $N = 2^{M-1}$.

As stated in Section VI.5, we need to check that (43) holds. Since $N = 2^{M-1}$ and $t_N = .5$, we have

$$k_n = \frac{t_N}{N} = \frac{\frac{1}{2}}{2^{M-1}} = \frac{1}{2\cdot 2^{M-1}} = \frac{1}{2^M} = \left(\frac{1}{2}\right)^M = \left(\frac{1}{\sqrt{2}}\right)^{2M}$$

thus

$$m_0 = 2M, \quad m_1 = m_0 + 1 = 2M + 1, \quad M = 2, 3, 4, 5, 6, 7, 8, 9, 10.$$

We find by straightforward calculation using Matlab 7, for $M = 2, 3, 4, 5$, that

$$\|(A^n_{m_1})^{-1}\| < k_n^{-1} + \epsilon_{m_1}$$

for $k_n = \left(\frac{1}{\sqrt{2}}\right)^{m_0}$ and $\epsilon_{m_1} = (\sqrt{2})^{m_1 - 1}$, thus insuring the induction base step and the analysis of Section VI.5 apply for $M = 2, 3, 4, 5$. The results of these calculations are shown in Table 1.

TABLE 1: Induction Base Step Norms and Bounds

M	m_0	$\|(A^n_{m_1})^{-1}\|$	$k_n^{-1} + \epsilon_{m_1}$
2	5	4.3321	8.0000
3	7	9.5217	16.0000
4	9	21.3440	32.0000
5	11	48.3825	64.0000

Although computing memory limitations prevented similar calculations for $M = 6, 7, 8, 9, 10$, implementation results were calculated for these resolution levels, and tabulated along with the results for the lower resolution levels. It is believed that the induction base steps are still valid for the higher resolution levels, although there is no direct verification of this.

As stated above, the multiscale linear in space and constant in time basis functions were used, and the results calculated and compiled using Matlab 7. To calculate the error, the actual differences of the exact solution u and the numerical solutions u_{M+1} and $u_{M,1}$ were calculated at each grid point, then $\|u^n - u^n_{M,1}\|$ and $\|u^n - u^n_{M+1}\|$ calculated for each time step n using both the inf-norm and the 2-norm. Finally, for each resolution level, the maximums of each of the inf-norms and 2-norms were selected and tabulated. Table 2 provides the results when the approximating functions are constant in time, while Table 3 provides the results when the approximating functions are linear in time. Further, cpu timings in seconds were taken for both the Discontinuous Galerkin Method and the Multilevel Method loops on each resolution level, and these tabulated as well. In the lower resolutions some cpu elapsed times were too small to be significant, no doubt due to the use of single-precision arithmetic in the Matlab 7.0 software used for the calculations. Further, while successive cpu timings for the same resolution level were not absolutely consistent, they did not vary

by significant amounts. The use of single-precision arithmetic also affected the error values at the higher resolutions of Table 3. It is not believed that these error values are actually zero, rather, these zero values can be attributed again to the limitations of single-precision arithmetic.

Comparison of the norm values at the various resolution levels for both the constant in time and linear in time versions shows comparable values in both the inf-norm and the 2-norm measurements of error. There were very slight differences in the norms at low resolution levels, but at higher resolutions, the norm values for the errors were identical.

TABLE 2: Error and Timing results for DGM and ML Methods (Constant in Time)

M	DG ∞-norm	DG 2-norm	ML ∞-norm	ML 2-norm	DG time	ML time
2	0.1932	0.2733	0.1929	0.2745	0	0
3	0.1533	0.3066	0.1532	0.3069	0	0
4	0.0904	0.2556	0.0904	0.2557	0.0313	0
5	0.0498	0.1990	0.0498	0.1990	0.0313	0
6	0.0266	0.1505	0.0266	0.1505	0.0469	0.0156
7	0.0137	0.1098	0.0137	0.1098	0.4063	0.1250
8	0.0070	0.0789	0.0070	0.0789	4.5313	1.0313
9	0.0035	0.0563	0.0035	0.0563	54.8281	11.5469
10	0.0018	0.0399	0.0018	0.0399	803.4063	144.5625

Figure 7 shows the cpu timings of the resolution levels M of 7, 8, 9, and 10 for the methods when the approximating functions are constant in time. The cpu timings of the lower levels were considered too insignificant to measure so they were not included in the plot. Comparison of the computational times for each method shows substantial saving with the multilevel method, requiring less than half the time to compute while providing the same degree of accuracy as the straight Discontinuous Galerkin Method. Further, the computational costs for each method closely followed the predicted costs, as shown by the plot.

Figure 8 shows the cpu timings of the resolution levels M of 7, 8, 9, and 10 for the methods when the approximating functions are linear in time. Again, the cpu timings of the lower resolution levels, while significantly longer than the constant in time version, were still considered too insignificant to plot. The basic costs of the higher resolution levels were again substantially lower for the multilevel method. As before, the computational costs for each method closely followed the predicted costs.

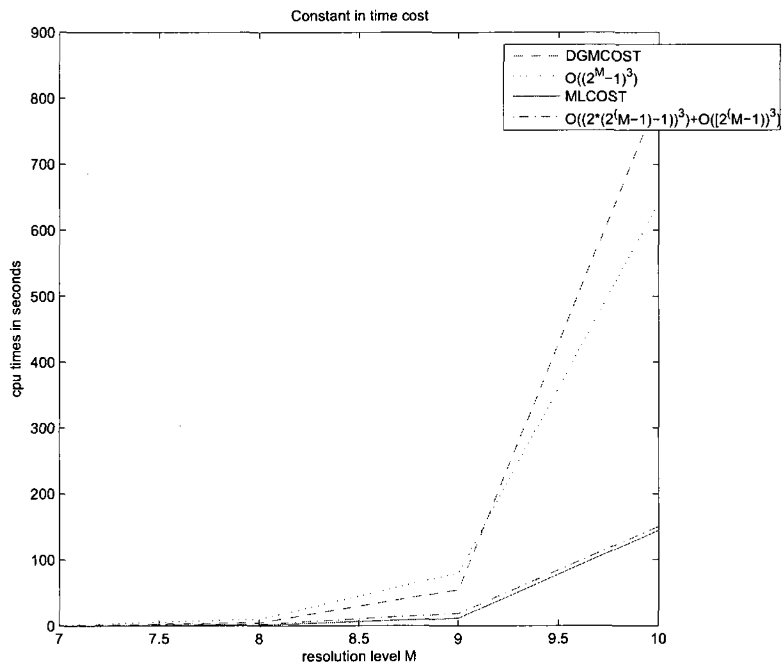

FIG. 7: Constant in Time Computational Cost

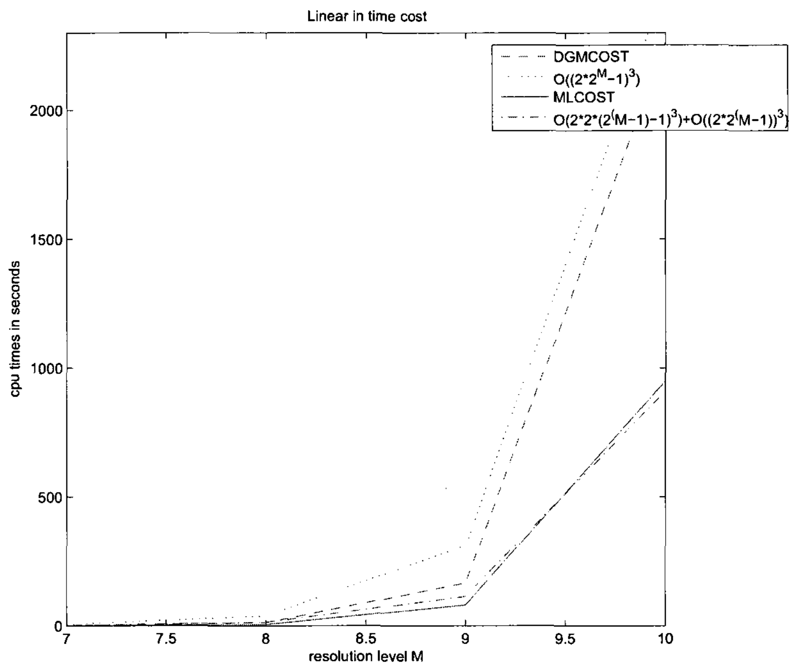

FIG. 8: Linear in Time Computational Cost

TABLE 3: Error and Timing results for DGM and ML Methods (Linear in Time)

M	DG ∞-norm	DG 2-norm	ML ∞-norm	ML 2-norm	DG time	ML time
2	0.0499	0.0706	0.0470	0.0706	0.0156	0.0156
3	0.0125	0.0250	0.0122	0.0245	0.0156	0.0156
4	0.0022	0.0062	0.0022	0.0061	0.0156	0.0156
5	0.0004	0.0017	0.0004	0.0017	0.0156	0.0156
6	0.0001	0.0005	0.0001	0.0005	0.0938	0.0781
7	0.0000	0.0002	0.0000	0.0002	0.9375	0.5938
8	0.0000	0.0001	0.0000	0.0001	12.3281	5.2344
9	0.0000	0.0000	0.0000	0.0000	165.4844	80.1094
10	0.0000	0.0000	0.0000	0.0000	2264.9000	948.9000

In short, these results suggests that the Multilevel Method version of the Discontinuous Galerkin Method provides a cheaper alternative to the traditional straight Discontinuous Galerkin Method, while preserving accuracy of the traditional Discontinuous Galerkin Method.

VI.5.2 An Example with an Incompatible Initial Condition

We use the following one-dimensional heat problem: Find u such that

$$u_t(x,t) - u_{xx}(x,t) = 0, \quad 0 < x < 1, \quad 0 < t < T,$$

$$u(0,t) = u(1,t) = 0, \quad 0 < t < .5,$$

$$u(x,0) = 1 - x, \quad 0 < x < 1.$$

The initial condition is incompatible with the prescribed boundary conditions, so this problem requires the special time and corresponding spatial discretization scheme outlined in Chapter V. Recall that the exact solution is

$$u(x,t) = \sum_{j=1}^{\infty} u_j^0 e^{-j^2 t} \sin(j\pi x),$$

where the coefficients u_j^0 are given by

$$u_j^0 = 2\int_0^1 (1-\xi)\sin(j\pi\xi)d\xi = \frac{2}{\pi}\{\frac{1}{j} - \frac{1}{j^2\pi}\sin(j\pi)\} = O(1/j),$$

and so we have $\|u_t(t)\|_2 = O(t^{-\frac{3}{4}})$. Thus $\alpha = \frac{3}{4}$, and so the index of singularity is $Q = \frac{q+1}{1-\alpha} = 4$ when $q = 0$, that is, the approximating polynomials are constant in

time. The initial spatial resolution level is $M = 9$. The same linear wavelet basis from before was used, as well as the same computing package, Matlab 7.0. Cpu timings were taken in seconds for each time step loop for the two methods. At times, several computations were necessary to obtain meaningful cpu times for lower resolution levels, as these tended to be quite small and were not always detected using single-precision arithmetic. The results of these experiments are shown in Table 4 for the case of $q = 0$. As before, the error of the Multilevel Method matched the error of the Discontinuous Galerkin Method for each specified grid size. Error tended to be greater in early transients and less in late steps, even with the coarser grids used in the late steps, due to the fact that the actual solution becomes smoother as the time steps progress. The size of the error for the first time step was disappointing, no doubt due once again to the single-precision arithmetic.

Although codes were written to calculate the solution for the case of $q = 1$, when the approximating polynomials are linear in time, this was not actually implemented due to the limitations of the available computing equipment, which lacked sufficient memory for spatial grids with resolution levels above $M = 10$. The first time step calculation requires a grid resolution of M=14, far in excess of this limitation.

As before, the multilevel version of the Discontinuous Galerkin Method proved to be more efficient in the higher resolution levels required for the first time steps. Also, at low spatial resolution levels such as those used in the final time steps, there was a less appreciable cost advantage to using the multilevel method.

TABLE 4: Time Step Results for DG and ML Methods (Constant in Time)

n	M	DG 2norm	DG ∞norm	ML 2norm	ML ∞norm	DGtime	MLtime
1	9	0.2311	0.1172	0.2314	0.1176	0.3594	0.1719
2	8	0.4135	0.1053	0.4135	0.1053	0.3438	0.1563
3	7	0.3420	0.0811	0.3417	0.0810	0.0781	0.0469
4	6	0.2614	0.0652	0.2608	0.0651	0.0781	0.0313
5	6	0.1933	0.0543	0.1922	0.0540	0.0313	0.0156
6	5	0.1976	0.0467	0.1975	0.0466	0.0313	0.0156
7	5	0.1568	0.0431	0.1556	0.0428	0.0156	0.0156
8	5	0.1675	0.0421	0.1670	0.0420	0.0156	0.0156
9	4	0.1275	0.0318	0.1272	0.0318	0.0156	0
10	4	0.0466	0.0165	0.0461	0.0163	0.0156	0

CHAPTER VII

CONCLUSIONS AND FUTURE PROJECTS

In this thesis, we have shown that the Discontinuous Galerkin Method can be enhanced with a multilevel calculation method to produce a new method that offers the same level of accuracy as the existing Discontinuous Galerkin Method, but with considerably lower computational costs. Further we have demonstrated that the special time and space discretization schemes of [6] remain valid when enhanced with the multilevel method.

Future projects include a generalization to the cases where the spatial region is taken in R^2 and R^3. Also, an enhancement of the multilevel method that requires us to only solve the linear system corresponding to an initial a coarse level m, then moving from a coarse level $m + k$, where k is any positive integer, to a finer level $m + k + 1$, will be examined. Use of quadratic and cubic wavelet bases will also be considered, along with possible extensions of the method to nonlinear cases.

BIBLIOGRAPHY

[1] Z. Chen, C. A. Micchelli and Y. Xu, *A Multilevel Method for Solving Operator Equations* (Journal of Mathematical Analysis and Applications 262, pp. 688-699, 2001).

[2] Z. Chen, B. Wu and Y. Xu, *Multilevel augmentation methods for differential equations* (Advances in Computational Mathematics 24, pp. 213-238, 2006).

[3] Z. Chen, B. Wu and Y. Xu, *Multilevel Augmentation Methods for Solving Operator Equations,* (Numerical Mathematics 14, pp. 32-55, 2005).

[4] V. Thomee, *Galerkin Finite Element Methods for Parabolic Problems* (Springer-Verlag, Berlin Heidelberg, 1997).

[5] W. Cheney and W. Light, *A Course in Approximation Theory* (Brooks/Cole, Pacific Grove, CA, 2000).

[6] H. Kaneko, K.S. Bey, and G.J.W. Hou, *Discontinuous Galerkin finite element methods for parabolic problems* (Applied Mathematics and Computation 182, pp. 388-402, 2006).

[7] K. Eriksson and C. Johnson, *Adaptive Finite Element Methods for Parabolic Problems I: A Linear Model Problem* (Society for Industrial and Applied Mathematics Journal on Numerical Analysis 28, pp. 43-77, 1991).

[8] M. Delfour, W. Hager, and F. Trochu, *Discontinuous Galerkin Methods for ordinary differential equations*, (Math. Comp., 36, pp. 453-473).

[9] P. Jamet, *Galerkin-type approximations which are discontinuous in time for parabolic problems in a variable domain*, (Society for Industrial and Applied Mathematics Journal of Numerical Analysis, 15, 1978, pp. 912-928).

[10] D. Schotzau and C. Schwab, *Time Discretization of Parabolic Problems by the HP-Version of the Discontinuous Galerkin Finite Element Method*, (Society for Industrial and Applied Mathematics Journal of Numerical Analysis, 38, 2001, pp. 837-875).

[11] B. Riviare, *Discontinuous Galerkin Methods For Solving Elliptic and Parabolic Equations: Theory and Implementation*, (Society for Industrial and Applied Mathematics, 2008).

[12] J. S. Hesthaven and T. Warburton, *Nodal Discontinuous Galerkin Methods Algorithms, Analysis, and Applications*, (Springer Sciences-Business Media, LLC, 2008).

VITA

Robert Gregory Brown
Department of Mathematics and Statistics
Old Dominion University
Norfolk, VA 23529

Education:

M.S. Mathematical Sciences, Virginia Commonwealth University, Richmond, Virginia 1986.

B.S. Mathematics, Randolph-Macon College, Ashland, Virginia 1983.

Professional Experience

Adjunct Instructor, August 2007-present Virginia Commonwealth University, Richmond, Virginia.

Graduate Teaching Assistant, January 2007-May 2007, Old Dominion University, Norfolk, Virginia.

Assistant Professor of Mathematics, August 1987-June 2006, Virginia Union University, Richmond, Virginia.

Adjunct Instructor, August 1986-August 1987, Virginia Commonwealth University, Richmond, Virginia.

Graduate Teaching Assistant, August 1984-May 1986 Virginia Commonwealth University, Richmond, Virginia.

Typeset using LaTeX.