The *hp*-Multigrid Method Applied to *hp*-Adaptive Refinement of Triangular Grids

William F. Mitchell*
Mathematical and Computational Sciences Division
National Institute of Standards and Technology
Gaithersburg, MD 20899-8910

Abstract

Recently the *hp* version of the finite element method, in which adaptivity occurs in both the size, h, of the elements and in the order, p, of the approximating piecewise polynomials, has received increasing attention. It is desirable to combine this optimal order discretization method with an optimal order algebraic solution method, such as multigrid. An intriguing notion is to use the values of p as the levels of a multilevel method. In this paper we present such a method, known as *hp*-multigrid, for high order finite elements and *hp*-adaptive grids. We present a survey of the development of *p*-multigrid and *hp*-multigrid, define an *hp*-multigrid algorithm based on the *p*-hierarchical basis for the p levels and *h*-hierarchical basis for an *h*-multigrid solution of the $p = 1$ "coarse grid" equations, and present numerical convergence results using *hp*-adaptive grids. The numerical results suggest the method has a convergence rate of 1/2 for Poisson's equation.

Keywords: elliptic partial differential equations, finite elements, hp adaptive refinement, multigrid, p-multigrid

1 Introduction

The numerical solution of partial differential equations (PDEs), is the most compute-intensive part of a wide range of scientific and engineering applications. So the development and application of faster and more accurate methods for solving PDEs is a very important field that has received much attention in the past fifty years. For simplicity, we consider the elliptic PDE

$$-\nabla \cdot (A\nabla u) = f \quad \text{in } \Omega \tag{1}$$

$$u = g \quad \text{on } \partial\Omega \tag{2}$$

 Preprint. Appeared in *Numerical Linear Algebra with Applications*, **17** (2010) 211–228.

where Ω is a bounded, connected, open region in $\mathbb{R}^2$, $f \in L^2(\Omega)$, and A is a piecewise constant positive definite symmetric matrix, although the algorithm described here can be applied to more general second order elliptic PDEs and boundary conditions. Determining the best grid and approximation space on which to efficiently compute the solution with the finite element method (FEM) is a central concern in this regard. Unfortunately, it is rarely possible to determine an optimal algorithm in advance. Thus, developing self-adaptive techniques which lead to optimal resource allocations is critical for future progress in many fields.

Self-adaptive methods have been studied for nearly 30 years now. Most of the work has focused on h-adaptive methods. But recently the research community has begun to focus more attention on hp-adaptive methods. In these methods, one not only locally adapts the size of the mesh, h, but also the degree of the polynomials, p. The attraction of hp-adaptivity is that, for problems with piecewise analytic data and with a properly chosen grid, the H^1 norm of the error approaches zero at an exponential rate in the number of degrees of freedom [17], as opposed to a polynomial rate for h-adaptivity with fixed p. In two dimensions the error behaves as $||e|| = O(e^{-c\sqrt[3]{N}})$, where N is the number of degrees of freedom, i.e. the dimension of the linear system that results from discretization, and c is a constant independent of N. Much of the theoretical work showing the advantages of hp methods was done in the 1980's by Babuška and coworkers (see [4] for an overview) but it wasn't until the 1990's that practical implementation began to be studied (see [39] for a survey).

A critical piece of an optimal-time hp-adaptive algorithm is an optimal $O(N)$ method for solving the linear system of equations. For the traditional h finite element method, such methods are obtained by using a multigrid method [8, 20] or by the preconditioned conjugate gradient method where the preconditioner is a Schwarz-based domain decomposition consisting of local block solves and a coarse grid solve [13]. Several researchers have examined the use of these Schwarz-based methods for high order finite element methods. The direct application of algebraic multigrid methods has also been considered. References [26] and [30] contain extensive bibliographies on these approaches. In this paper, we examine the use of a different type of multilevel method in which the order of the polynomials are used as the multi"grids".

Following [43], we refer to this as the *hp-multigrid method*. The hp in the name refers to the combination of p-multigrid (using lower order for "coarse grids") with h-multigrid (using larger elements for coarse grids) on the $p = 1$ level. On the surface, the method in this paper seems similar to the one in [43]. However, the method in [43] is developed for the discontinuous Galerkin method for conservation laws with non-adaptive unstructured meshes using an agglomeration h-multigrid method on the $p = 1$ level, while in this paper we develop the method for the continuous hp-adaptive finite element method for second order elliptic PDEs using an h-hierarchical multigrid method on the $p = 1$ level.

The remainder of the paper is organized as follows. In Section 2 we give

a history of the *hp*-multigrid method. Section 3 gives the definition of *h*-hierarchical and *p*-hierarchical basis functions. In Section 4 we define an *hp*-multigrid method. Computational complexity is considered in Section 5. Section 6 presents numerical results to examine the rate of convergence of the *hp*-multigrid method with 2D elliptic PDEs in both uniform and *hp*-adaptive grid settings. To the author's knowledge, this is the first presentation of numerically determined convergence rates for *hp*-multigrid using *hp*-adaptive grids with continuous finite elements.

2 History

The early development of *hp*-multigrid occurred in two contexts: the *p*-version of the finite element method, and the spectral element method.

The use of p to define the levels of a multilevel method was first suggested by Craig and Zienkiewicz [10] in 1985 for the *p*-version of the finite element method, where it was simply referred to as "multigrid based on the hierarchical basis". The presentation was abstract enough to define a multigrid method using either the *h*-hierarchical basis for the *h*-version of the finite element method, or the *p*-hierarchical basis for the *p*-version of the finite element method. They did not combine the two, but rather used an exact solver on the coarsest grid. The sequence of p's for the *p*-version is implied to be p, $p-1$, $p-2$, ..., 2, 1, 2, ..., $p-1$, p, which we will refer to as a V-cycle with an *arithmetic p sequence*. The presentation was in the context of 2D rectangular elements and elasticity equations.

Bussino *et al.* [9, 16] followed up with a parallel implementation of a "multilevel method for *p*-version FEM". They replaced the exact solver at $p = 1$ with a conjugate gradient solver, and presented numerical experiments to demonstrate the convergence rate and the speed up obtained on a parallel computer using a 2D Poisson equation.

Babuška *et al.* [2] also examined what they called a "multi-*p* iterative procedure" (as opposed to multi-grid) for *p*-version FEM. Their work was with Laplace's equation in 2D using rectangular elements, and used Successive Over-relaxation (SOR) for the relaxation part of the multi-*p* algorithm. They presented computationally-determined convergence rates for several different types of cycles. In addition to the full arithmetic p sequence, they considered using just even p, using just odd p, and the sequence p, $p/2$, $p/4$, ..., 2, 1, 2, 4, ..., $p/2$, p (which we will refer to as a V-cycle with a *geometric p sequence*). They also considered V-cycles, downward cycles, and upward cycles.

Field and Pressburger [15] presented what they called an "*h*-*p*-multigrid" method for "*hp*-FEM" for 3D structural analysis using tetrahedral elements. However, they only used quadratic elements and the solution method was an iteration between the quadratic and linear elements with SOR relaxation, so it is not truly an *hp*-adaptive FEM as we know it, and it is at best a two-level *p*-multigrid method.

In 1995, Hu and Katz [28] presented the first theoretical results for the

"multi-p V-cycle of Babuška *et al.*" in the context of 2D rectangular elements and second order elliptic PDEs. This work used an arithmetic p sequence and an exact solver at $p = 1$. They proved that the method converges, but did not prove that the convergence rate is independent of p. In Hu, Guo and Katz [27] they started calling it the "algebraic multi-p method", and also a "multi-p preconditioner" by using it as a preconditioner for conjugate gradients. As a preconditioner, the smoother is changed to a symmetric Gauss-Seidel iteration. In the context of elastostatic examples in 2D and 3D, they proved convergence with condition number $O(1+\log(p)^2)$, and gave numerical evidence of p-independent convergence rates in both 2D and 3D. Further work by this team is found in [18, 19].

Sun *et al.* [49] extended the method of Babuška *et al.* to a "p-multilevel-ILU preconditioned conjugate gradients" solver for Maxwell's equation. Although presented as a multiplicative Schwarz algorithm, it is equivalent to a V-cycle with the arithmetic p sequence.

The first paper from the spectral element community was by Rønquist and Patera [46] in 1987, where they call it "spectral element multigrid". This method uses the geometric p sequence and an exact solver on $p = 1$. They presented the method for Poisson's equation in 1D, and numerically demonstrated p-independent convergence. A year later, a paper by Maday and Muñoz [33] gave the convergence theory for the spectral element multigrid method. They proved that the convergence rate is independent of p in 1D, but in 2D it is $O(1 - c/p)$ for some constant c. Other presentations of work by Rønquist, Patera, Maday and Muñoz can be found in [34, 35, 41]

In 2001, Helenbrook [21] applied the spectral multigrid method to a 2D incompressible Navier-Stokes equation using triangular elements. For the first time, a geometric multigrid method was used for the solution at $p = 1$. Two years later, Helenbrook, Mavriplis and Atkins [25] changed the name of the method to "p-multigrid", but still used a geometric p sequence. This paper considered both a (non-adaptive) hp-FEM discretization of the Laplace equation in 2D, and Streamline-Upwind Petrov-Galerkin (SUPG) and Discontinuous Galerkin discretizations of the convection equation. In 2006, Nastase and Mavriplis [43] coined the term "hp-multigrid" for their method using an algebraic p sequence and h-multigrid at $p = 1$. The paper addressed the Discontinuous Galerkin method for inviscid compressible flows in 3D.

Starting in 2005, many more researchers began using multilevel methods based on reducing p, with most of them using the name "p-multigrid". Most of this work is either with spectral element methods or with the Discontinuous Galerkin method for fluid flow problems. At this point it becomes difficult to be exhaustive in the citations, but a sampling of the work can be found in [7, 11, 14, 23, 22, 24, 29, 31, 32, 36, 42, 45, 51].

The h-multigrid method has a much more extensive history, beginning with the seminal paper of Brandt [8]. In fact, the MGNet Bibliography [12] contains about 3600 entries. In this paper we mainly consider the h-hierarchical multigrid method developed in the late 1980's [6, 37].

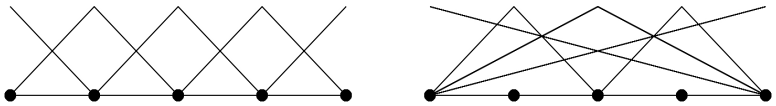

Figure 1: The nodal basis (left) and h-hierarchical basis (right) for 1D linear finite elements.

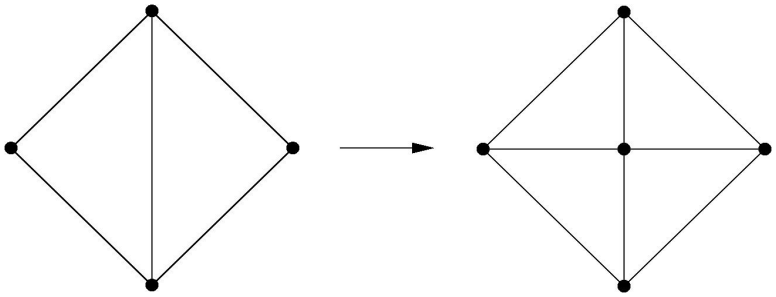

Figure 2: Bisection of a pair of triangles.

3 Hierarchical Bases

The simplicity of the p-multigrid method depends on using a p-hierarchical basis for the finite element space. Although other forms of geometric or algebraic h-multigrid methods can be used for the $p = 1$ level, the h-multigrid method in Section 4 will be similarly defined in terms of the h-hierarchical basis [5, 6, 37, 44, 53]. In this section, the relevant properties of hierarchical bases are given.

The defining characteristic of hierarchical bases is that they are created as a hierarchy. Figure 1 illustrates this using linear bases in 1D. The left side of the figure shows the usual "hat function" nodal basis for a grid with four intervals. The right side shows the h-hierarchical basis for the same finite element space. One begins with a grid consisting of a single interval and defines the nodal basis on that grid, i.e. the two basis functions that have the value 1.0 at one end point and 0.0 at the other. The grid is refined by adding a node at the middle of the interval, and a new basis function, which would be a nodal basis function on this grid, is added, but the existing basis functions are not changed. Finally, the grid is refined again, and two more basis functions are added, leaving the existing basis functions unchanged.

The definition of the linear h-hierarchical basis for bisected triangles is the same. With each refinement of a pair of triangles as in Figure 2, a new vertex is added with a corresponding piecewise linear basis function whose support is the four triangles just created.

In practice, the h-hierarchical basis is not used directly because the stiffness matrix is too dense, due to the large supports of the coarse level bases [53].

Instead, the h-hierarchical basis is used implicitly in the h-multigrid method of Section 4 through a sequence of 2-level transformations of vectors and the stiffness matrix. For any $l \geq 2$, consider a 2-level h-hierarchical basis where the lower level is the nodal basis of the coarse grid consisting of h-refinement up to level $l-1$, and the fine grid consists of h-refinement up to level l. Any function in the finite element space of the fine grid has an expansion in both the nodal basis and this hierarchical basis,

$$f = \sum_{i=1}^{N} \alpha_i^{(N)} \phi_i^{(N)} = \sum_{i=1}^{N} \alpha_i^{(H)} \phi_i^{(H)}$$

where $\phi_i^{(N)}$ are the nodal bases and $\phi_i^{(H)}$ are the hierarchical bases. Conversion between the nodal and hierarchical bases is a linear process, so the coefficients in the expansion are related by $\alpha^{(N)} = S_l \alpha^{(H)}$ and $\alpha^{(H)} = S_l^{-1} \alpha^{(N)}$ where S_l has the form

$$S_l = \begin{bmatrix} I & 0 \\ s_l & I \end{bmatrix}$$

with the lower block corresponding to the lower level (coarse grid) part of the hierarchical basis. For bisected triangles, s_l has two nonzeros per row. Note that S_l^{-1} has the same form as S_l but with s_l negated, so basis conversion in either direction can be performed with a small number of operations. The relationship between the nodal and 2-level h-hierarchical stiffness matrices is given by $A_l^{(H)} = S_l^T A_l^{(N)} S_l$. By the equivalence of $A_l^{(N)} x_l^{(N)} = b_l^{(N)}$ and

$$(S_l^T A_l^{(N)} S_l)(S_l^{-1} x_l^{(N)}) = S_l^T b_l^{(N)}$$

we have $x_l^{(H)} = S_l^{-1} x_l^{(N)}$ and $b_l^{(H)} = S_l^T b_l^{(N)}$.

Similarly, the p-hierarchical bases are defined by beginning with nodal linear basis on the fine grid, adding bases of exact degree 2, then bases of exact degree 3, etc. The exact definition of the basis functions is not important for this paper, but we mention that the numerical results in Section 6 were obtained using the p-hierarchical basis of Szabo and Babuška [50].

For triangular elements, the p-hierarchical basis functions can be divided into three groups:

1. Vertex bases. There is one piecewise linear basis function associated with each vertex. These are the usual linear nodal basis functions.

2. Edge bases. An edge of degree p has $p-1$ associated basis functions, one each of exact degree $2, 3, ..., p$. They are 0.0 at the end points of the edge. The support is the union of the two triangles that share that edge (one triangle if the edge is on the boundary), and they are a polynomial over each of the triangles.

3. Element bases, sometimes called face bases or bubble functions. An element of degree p has $q-2$ element basis functions of exact degree q, $q = 3, 4, ..., p$, for a total of $(p-1)(p-2)/2$ element basis functions. The support is the triangle, and they are 0.0 on the boundary of the triangle.

4 The *hp*-Multigrid Method

In this section we define one form of the *hp*-multigrid algorithm. Obviously many other variants are possible by using other smoothers, coarse grid ($p = 1$) solvers, cycle strategies (e.g. W-cycle), etc., for example the method in [43]. Here we use a Gauss-Seidel smoother, h-hierarchical multigrid on $p = 1$, and V-cycle.

The first observation is that the element bases can be eliminated by static condensation [47, 52]. Using subscript e to denote parts associated with elements, and v to denote parts associated with vertices and edges, the linear system can be partitioned as

$$\begin{bmatrix} A_e & A_{ev} \\ A_{ev}^T & A_v \end{bmatrix} \begin{bmatrix} x_e \\ x_v \end{bmatrix} = \begin{bmatrix} b_e \\ b_v \end{bmatrix}.$$

Because the support of the element bases is a single triangle, A_e is block diagonal with the size of the blocks being the number of element bases in the corresponding triangles, so it is easy to multiply by A_e^{-1} using a direct solver. Performing block Gauss elimination gives the condensed system

$$(A_v - A_{ev}^T A_e^{-1} A_{ev}) x_v = b_v - A_{ev}^T A_e^{-1} b_e.$$

The Schur complement $A_v - A_{ev}^T A_e^{-1} A_{ev}$ has the same nonzero structure as A_v. After the condensed system is solved for x_v, the coefficients for the element bases are computed by solving $A_e x_e = b_e - A_{ev} x_v$.

For the remainder of this section it is to be understood that the matrices and vectors are from the condensed system given above.

The next observation is that, because of the p-hierarchical nature of the basis, the matrix for the coarse p-level with degree $p-1$ is contained within the matrix for the fine p-level with degree p. If A_p denotes the stiffness matrix with all bases up to degree p, and $A_{\overline{p}}$ denotes the stiffness matrix with all bases of exact degree p, then

$$A_p = \begin{bmatrix} A_{\overline{p}} & A_{\overline{p},p-1} \\ A_{\overline{p},p-1}^T & A_{p-1} \end{bmatrix}.$$

Using similar notation for the vectors, the coarse level equations are simply given by

$$A_{p-1} x_{p-1} = b_{p-1} - A_{\overline{p},p-1}^T x_{\overline{p}}.$$

Note that the restriction operator simply amounts to removing the section of the coefficient vector corresponding to bases of exact degree p, and the prolongation operator simply amounts to reinstating them.

The Gauss-Seidel relaxations at any level p are limited to the region near elements of degree at least p. In Section 5 we show that, for uniform grids, the number of operations for one V-cycle is proportional to the number of nonzero entries in the stiffness matrix, M, divided by p. But if the p refinement is very local, these operation counts will not hold and the total number of operations to solve the linear system will be larger than $O(M/p)$. The same suboptimality was

procedure hp-multigrid
repeat
  p-multigrid($A_v - A_{ev}^T A_e^{-1} A_{ev}$,$x_v$,$b_v - A_{ev}^T A_e^{-1} b_e$,P)
until converged
$x_e = A_e^{-1}(b_e - A_{ev} x_v)$
end procedure

Figure 3: Algorithm for *hp*-multigrid, including static condensation. P is the maximum degree of any element in the grid.

observed for h-multigrid with highly localized h refinement in [37]. The solution given there is to use "local black" relaxation. The term black comes from the red-black coloring of the vertices with the black vertices being the coarse grid vertices. The term local refers to performing Gauss-Seidel relaxation only at red vertices and black vertices that are immediate neighbors of red vertices. This recovers the optimal order number of operations and does not significantly effect the convergence rate in the numerical experiments of [37]. We propose the same approach for p-multigrid: on level p, perform relaxations only on equations coming from bases of degree p and equations that are directly connected to equations coming from bases of degree p.

A similar observation about nestedness of the matrices holds for the h-hierarchical basis for the $p = 1$ equations, i.e., the linear system for level l with the 2-level hierarchical basis has the form

$$\begin{bmatrix} A_l^{(H)} & A_{l,l-1}^{(H)} \\ A_{l,l-1}^{(H)^T} & A_{l-1}^{(N)} \end{bmatrix} \begin{bmatrix} x_l^{(H)} \\ x_{l-1}^{(N)} \end{bmatrix} = \begin{bmatrix} b_l^{(H)} \\ b_{l-1}^{(N)} \end{bmatrix}.$$

This leads to an h-multigrid algorithm that is in essence the same as the p-multigrid algorithm. The algorithm given here first appeared in [37], although similar, but suboptimal, algorithms are given in [6] and [10]. The difference between the h-multigrid algorithm and the p-multigrid algorithm is that we begin an h-level with the nodal basis, and have to convert to the 2-level hierarchical basis to get the coarse grid equations.

Let P be the maximum degree and L be the maximum number of h refinement levels. The *hp*-multigrid algorithm is given in Figures 3, 4 and 5.

5 Computational Complexity

In this section we consider the number of operations required by one V-cycle of the *hp*-multigrid method in the case of a uniform grid. In particular, we show that the number of operations is $O(M/p)$ where M is the number of nonzero entries in the full (noncondensed) stiffness matrix.

Let V be the number of vertices in the grid, and let every triangle have degree p. The number of edges is $E \approx 3V$ and the number of triangles is $T \approx 2V$. Since there is one basis associated with each vertex, $p-1$ bases associated with each

procedure p-multigrid(A_p,x_p,b_p,p)
if p==1 **then**
  h-multigrid(A_1,x_1,b_1,L)
else
  perform Gauss-Seidel relaxation on $A_p x_p = b_p$
  p-multigrid(A_{p-1},x_{p-1},$b_{p-1} - A^T_{\overline{p},p-1} x_{\overline{p}}$,$p-1$)
  perform Gauss-Seidel relaxation on $A_p x_p = b_p$
end if
end procedure

Figure 4: Algorithm for p-multigrid part of hp-multigrid. L is the maximum h-refinement level of any element in the grid.

procedure h-multigrid($A_l^{(N)}$,$x_l^{(N)}$,$b_l^{(N)}$,l)
if l==1 **then**
  solve $A_1^{(N)} x_1^{(N)} = b_1^{(N)}$ with a direct solver
else
  perform Gauss-Seidel relaxation on $A_l^{(N)} x_l^{(N)} = b_l^{(N)}$
  $A_l^{(H)} = S_l^T A_l^{(N)} S_l$
  $x_l^{(H)} = S_l^{-1} x_l^{(N)}$
  $b_l^{(H)} = S_l^T b_l^{(N)}$
  h-multigrid($A_{l-1}^{(N)}$,$x_{l-1}^{(N)}$,$b_{l-1}^{(N)} - A_{\overline{l},l-1}^{(H)^T} x_{\overline{l}}^{(H)}$,$l-1$)
  $x_l^{(N)} = S_l \begin{bmatrix} x_{\overline{l}}^{(H)} \\ x_{l-1}^{(N)} \end{bmatrix}$
  perform Gauss-Seidel relaxation on $A_l^{(N)} x_l^{(N)} = b_l^{(N)}$
end if
end procedure

Figure 5: Algorithm for h-multigrid part of hp-multigrid.

edge, and $(p-2)(p-1)/2$ bases associated with each triangle, the number of degrees of freedom, and number of rows in the full matrix, is

$$N \approx V + 3V(p-1) + 2V(p-2)(p-1)/2 = p^2 V.$$

The number of rows in the condensed matrix is $N_c \approx 3pV - 2V = O(pV)$. In the full matrix there are $O(p^2)$ nonzero entries in each row, giving $M = O(p^2 N) = O(p^4 V)$. In the condensed matrix there are $O(p)$ nonzeros in each row, so the number of nonzeros in the condensed matrix is $M_c = O(pN_c) = O(p^2 V)$.

The number of operations required for a Gauss-Seidel sweep at p-level i , $i = 2, 3, ..., p$, is $O(i^2 V)$. The computation of $A^T_{\bar{i},i-1} x_{\bar{i}}$ to get the coarse grid equations for level $i-1$ requires only $O(iV)$. Also, the h-multigrid on p-level 1 uses $O(V)$ operations, so the total operation count for one V-cycle is

$$Operations = O(\sum_{i=1}^{p} i^2 V) = O(p^3 V) = O(M/p).$$

If the convergence rate of the multigrid cycles is bounded away from 1.0 independent of h and p, then when used in a full multigrid scheme where a fixed number of iterations on each intermediate grid suffices to reduce the algebraic error below the discretization error, the number of operations used in the multigrid cycles to solve the final linear system is $O(M/p)$. In addition, it can be shown that the number of operations for static condensation, which occurs once on each intermediate grid, is $O(p^2 M)$. Thus the total number of operations for solution is $O(p^2 M)$.

Note that the number of operations required to generate the linear system is necessarily at least $O(M)$ unless the matrix entries are simply constants that are known in advance.

6 Experimental Convergence Rates

We examine the convergence rate of the hp-multigrid method defined in Section 4 numerically using newest vertex bisection of triangles [37] and four examples. In the first example we use a simple Poisson problem with different values of uniform h and uniform p to demonstrate that the convergence rate appears to be bounded away from 1.0 independent of both h and p. In the other examples we use hp-adaptive grids with two Poisson problems that are commonly used as test cases for adaptive grids, and a problem with discontinuous coefficients.

By convergence rate, we mean the contraction factor of the L_2 norm of the residual. Specifically, let $Ax = b$ be the discrete system, and let $x^{(i)}$ be the approximate solution after i V-cycles of the hp-multigrid iteration. The relative residual after iteration i is defined as

$$r^{(i)} = ||b - Ax^{(i)}||/||b||$$

and the contraction factor of iteration i is defined as

$$\rho^{(i)} = r^{(i)}/r^{(i-1)}$$

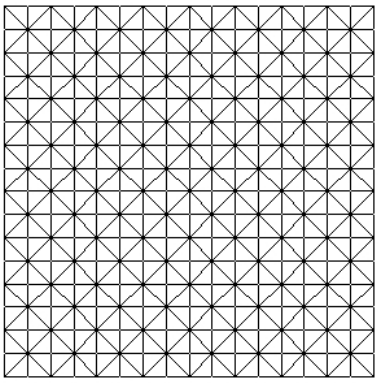

Figure 6: Grid with h=1/16 for the first test problem.

h	p=1	2	3	4	8	16
1/16	.06	.47	.42	.45	.43	.40
1/32	.09	.48	.41	.45	.41	.40
1/64	.09	.48	.41	.45	.41	.39
1/128	.09	.49	.41	.45	.41	.39
1/256	.09	.49	.40	.45	.41	*
1/512	.09	.49	.40	.45	.41	*
1/1024	.09	.49	.40	.45	*	*

Table 1: Convergence rates for uniform grids. * indicates insufficient memory.

for $i \geq 1$. Starting with $x^{(0)} = 0$, we iterate until $r^{(i)} < 10^{-9}$ and report the largest $\rho^{(i)}$ as the contraction factor.

These computations were performed using the program PHAML Version 1.5.1 [38] on one core of a dedicated 8 dual-core AMD Opteron Processor 8218 shared memory system with 64 Gbytes RAM operating under the 64-bit CentOS 5.2 distribution of Linux. [1] PHAML was compiled with Intel Fortran Version 11.0.

For the first problem we use Poisson's equation on the unit square with Dirichlet boundary conditions and the right hand side chosen so that the solution is $2^{40}x^{10}(1-x)^{10}y^{10}(1-y)^{10}$. The power 10 was chosen so that the degree of the polynomial is larger than any p to be considered, and 2^{40} is a normalization factor so that the maximum value of the solution is 1.0. We computed the convergence rate for grids with h ranging from 1/16 to 1/1024 and $p = 1, 2, 3, 4, 8$ and 16. The grid for $h = 1/16$ is shown in Figure 6. The convergence rates

[1] The mention of specific products, trademarks, or brand names is for purposes of identification only. Such mention is not to be interpreted in any way as an endorsement or certification of such products or brands by the National Institute of Standards and Technology. All trademarks mentioned herein belong to their respective owners.

p=2	4	5	6	7	8	12	16
.49	.45	.43	.42	.42	.41	.40	.39

Table 2: Additional convergence rates for Example 1 indicating monotonic improvement as p increases.

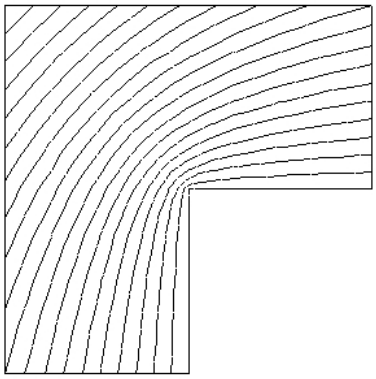

Figure 7: Solution of the L-shaped domain problem.

are given in Table 1. We observe that the convergence rate is bounded by 1/2 throughout the table. The asymptotic convergence rate is quickly achieved as h is decreased. Note that the column labeled "$p = 1$" reflects h-multigrid for linear elements, and we observe the usual rate of convergence of approximately 1/10, while the other columns show a slower convergence rate for p-multigrid. We also observer that if $p = 1$ and $p = 3$ are not considered, the rate of convergence actually improves as p is increased. This is further illustrated in Table 2 which contains convergence rates for additional values of p, using sufficiently small h to have reached the asymptotic limit.

The second example is the L-shaped domain problem used frequently in the illustration of adaptive grid techniques. It is Laplace's equation on the domain $(-1,1)\times(-1,1)\setminus(0,1)\times(-1,0)$. The solution has a singularity at the reentrant corner at the origin. It is given in polar coordinates by $r^{2/3}\sin(2\theta/3)$. Dirichlet boundary conditions are set accordingly. A contour plot of the solution is shown in Figure 7.

Adaptive refinement is performed by refining elements with large error indicators until the number of equations in the discrete system is doubled, at which point the hp-multigrid solution is performed. We used a classical element residual error indicator [1] in which the error indicator for triangle T is given by the energy norm of the solution of

$$Le = f - L\hat{u} \text{ in } T$$

$$e_n = [\hat{u}_n] \text{ on } \partial T$$

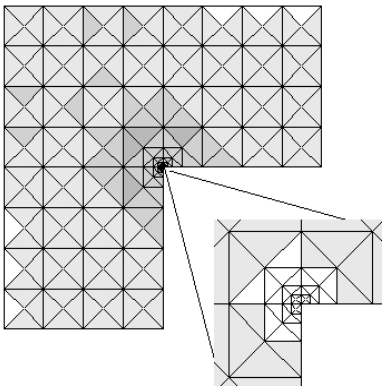

Figure 8: An *hp*-adaptive grid for the L-shaped domain problem.

N	$max\ ref$	$min\ p$	$max\ p$	rate
226	9	1	2	.44
453	15	1	2	.47
904	25	1	3	.46
1809	35	1	5	.42
3616	45	1	6	.44
7233	53	1	8	.42

Table 3: Convergence rates for the L-shaped domain example.

where L is the differential operator, f is the right hand side of the differential equation, $\hat{u}$ is the approximate solution, subscript n denotes the outward normal derivative, and $[\hat{u}_n]$ is the jump in the normal derivative of the approximate solution on the boundary of T. The local Neumann problem is solved using bases one degree higher than the current degree of T. The *hp* strategy, i.e. selection of whether to refine an element by h (subdivide by bisection) or by p (increase the degree by one), for this problem is to refine by h if the element touches the reentrant corner, and by p otherwise. An example *hp*-refined grid is illustrated in Figure 8. Here the gray scale indicates the degree of the element with white being degree 1 and darker colors indicating higher degree. The inset shows the detail of the grid near the reentrant corner.

Table 3 contains the convergence rates of *hp*-multigrid on each of the grids in the sequence obtained by *hp*-adaptive refinement. Here N is the number of equations in the discrete system, $max\ ref$ is the maximum number of times any element was refined by h (minimum $h \propto 2^{-(max\ ref)/2}$), $min\ p$ is the minimum element degree, and $max\ p$ is the maximum element degree. We again find that the convergence rates are bounded by 1/2.

The third example is another Poisson equation commonly used in the adap-

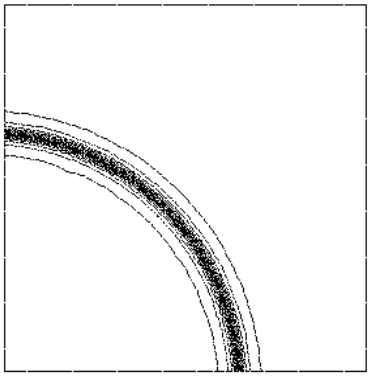

Figure 9: Solution of the wave front problem.

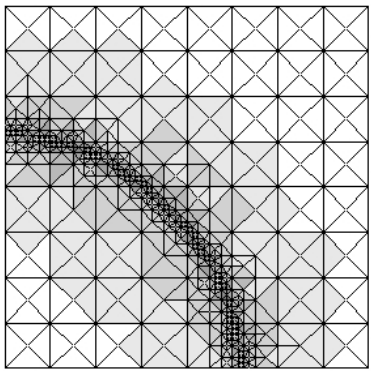

Figure 10: An *hp*-adaptive grid for the wave front problem.

tive grid refinement literature. The right hand side and Dirichlet boundary conditions on the unit square are chosen so that the solution is

$$u(x, y) = \tan^{-1}(\alpha\sqrt{(x - x_c)^2 + (y - y_c)^2} - r_0)$$

which contains a circular wavefront with position determined by (x_c, y_c) and r_0 and sharpness determined by α. It also has a mild singularity at (x_c, y_c). We choose $(x_c, y_c) = (-.05, -.05)$ so that the singularity does not have an effect in this example. For the other parameters, we use $r_0 = .7$ and $\alpha = 100$. A contour plot of the solution is shown in Figure 9.

We use the same approach to adaptive refinement as in Example 2, except for the *hp* strategy. For this example we use a strategy based on that of Süli *et al.* [48], which is called PRIOR2P in PHAML. For an element of degree p, an error estimate, η_{p-1}, for the degree $p-1$ part of the solution is given by the

N	$max\ ref$	$min\ p$	$max\ p$	local black rate	full black rate
291	1	1	3	.46	.20
580	3	1	4	.39	.23
1174	5	1	5	.41	.18
2320	7	1	5	.42	.27
4640	8	1	5	.45	.34
9281	9	2	7	.46	.18
18561	11	2	7	.44	.24
37181	13	2	8	.41	.24
74325	14	3	9	.43	.24
148480	15	3	10	.44	.26
297138	17	3	10	.42	.29
593921	18	3	11	.40	.30
1187894	19	3	12	.40	.29
2375722	22	3	13	.40	.27
4751622	22	3	14	.41	.25

Table 4: Convergence rates for the wave front example.

norm of the part of the solution from the p-hierarchical bases of exact degree p. Similarly η_{p-2} estimates the error of the degree $p-2$ part of the solution using the bases of exact degree $p-1$. Using the *a priori* error estimate [3]

$$||e||_{H^1} \leq C\frac{h^\mu}{p^{(m-1)}}||u||_{H^m}$$

where $\mu = \min(p, m-1)$, C is a constant independent of h and p, u is the true solution, and H^1 and H^m are the usual Sobolev spaces, take the ratio of the error estimates for $p-1$ and $p-2$ to estimate the smoothness

$$m \approx 1 - \frac{\log(\eta_{p-1}/\eta_{p-2})}{\log((p-1)/(p-2))}.$$

Then use h refinement if $m < p+1$ and p refinement otherwise. Since this formula requires $p \geq 3$, use p refinement if $p < 3$. An example of an hp-adaptive grid for this problem is shown in Figure 10. The convergence rates are given in the column labeled "local black rate" in Table 4. Once again, all convergence rates are less than 1/2.

The second and third examples suggest that the local black relaxation does not harm the convergence rate of the hp-multigrid method, compared to the convergence rate obtained with uniform grids. In the final column of Table 4 we give the convergence rates for the wave front problem if relaxation is performed at all the black points instead of just those that are neighbors of red points. This results in a substantial reduction in the convergence rate, typically to about 1/4,

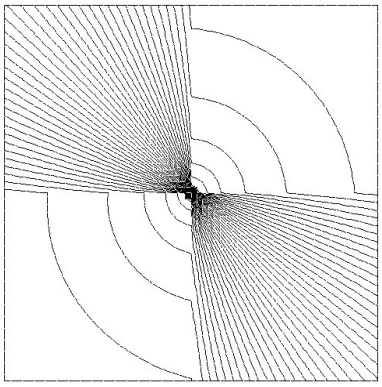

Figure 11: Solution of the Kellogg problem.

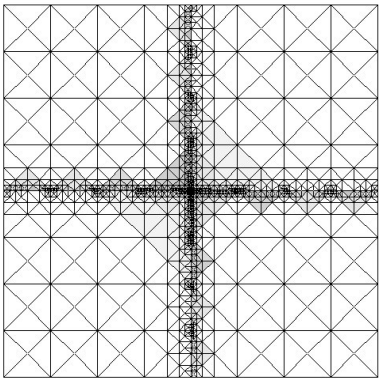

Figure 12: An *hp*-adaptive grid for the Kellogg problem.

which implies the full black relaxation could converge in half as many V-cycles. However, unless p is fairly uniform, a V-cycle with full black relaxation will require considerably more than twice as many operations as a V-cycle with local black relaxation, so the local black relaxation is still advised in general.

As a final example we consider a problem with discontinuous coefficients, the Kellogg problem as given in [40]. In Equation 1, $\Omega = (-1,1) \times (-1,1)$, $A = I$ in the first and third quadrants, and $A = RI$ in the second and fourth quadrants, where R will be defined shortly, and $f = 0$. The solution is given in

N	$max\ ref$	$min\ p$	$max\ p$	1 h V-cycle	20 h V-cycles
309	4	1	5	.33	.33
647	7	1	6	.45	.34
1179	11	1	7	.58	.36
2528	15	1	7	.70	.38
4653	21	1	7	.78	.39
9553	29	1	8	.85	.41
22412	39	1	9	.89	.46
39419	49	1	9	.92	.50
81033	53	1	9	.93	.43

Table 5: Convergence rates for the Kellogg example.

polar coordinates by $u(r,\theta) = r^\gamma \mu(\theta)$ where

$$\mu(\theta) = \begin{cases} \cos((\pi/2-\sigma)\gamma)\cdot\cos((\theta-\pi/2+\rho)\gamma) & \text{if } 0 \le \theta \le \pi/2 \\ \cos(\rho\gamma)\cdot\cos((\theta-\pi+\sigma)\gamma) & \text{if } \pi/2 \le \theta \le \pi \\ \cos(\sigma\gamma)\cdot\cos((\theta-\pi+\rho)\gamma) & \text{if } \pi \le \theta \le 3\pi/2 \\ \cos((\pi/2-\rho)\gamma)\cdot\cos((\theta-3\pi/2-\sigma)\gamma) & \text{if } 3\pi/2 \le \theta \le 2\pi \end{cases}$$

with $\gamma = 0.1$, $R = 161.4476387975881$, $\rho = \pi/4$ and $\sigma = -14.92256510455152$. The solution is shown in Figure 11 and an example grid in Figure 12. We use the same adaptive strategy as in the wave front example.

The convergence rates are given in Table 5. Under the column labeled "1 h V-cycle" it is seen that the convergence rate deteriorates as the mesh is refined. The column labeled "20 h V-cycles" gives the convergence rate of an hp-multigrid V-cycle if we use 20 V-cycles for the h-multigrid on p-level 1. Evidently, the deterioration of the convergence rate is due to the h-multigrid, and the p-multigrid converges with a rate that is independent of the mesh if the p-level 1 equations are solved accurately enough. This suggests that, for this problem, it may be better to use an h-multigrid that is tuned for discontinuous coefficients, or perhaps an algebraic h-multigrid, for the p-level 1 equations.

7 Conclusion

In this paper we presented the hp-multigrid method for high order and hp-adaptive finite elements using triangular elements for 2D elliptic partial differential equations. Extensions to any finite element space with a p-hierarchical basis are obvious. The crux of the method is to use the polynomial degree p as the multilevels of a V-cycle, and use a standard h-multigrid method to solve the equations at $p = 1$. Although the roots of the method go back nearly 25 years, only a handful of researchers considered it until a few years ago.

Most of the research on this and related methods has been in the context of the p version of the finite element method, spectral element methods, and Discontinuous Galerkin methods. In this paper we examined it in the context of hp-adaptive continuous finite elements. The number of operations required for one V-cycle with uniform grids is $O(M/p)$ where M is the number of nonzero entries in the stiffness matrix. Numerical results suggest the convergence rate for Poisson's equation and uniform grids is bounded by $1/2$ independent of both h and p. Together these imply that the number of operations for the multigrid cycles in a full multigrid method is $O(M/p)$. The number of operations for static condensation is $O(p^2M)$. Thus, for uniform grids and Poisson's equation, the method is optimal in the sense that the number of operations to solve the linear system is of the same order as the number of operations required to generate the linear system up to a factor of p^2.

For hp-adaptive grids, the algorithm is modified by using local black relaxation to maintain the $O(M/p)$ operation count. Numerical results with hp-adaptive grids using two Poisson examples that are commonly used in the adaptive refinement community also show convergence rates bounded by $1/2$ independent of both h and p. Numerical results with hp-adaptive grids for an elliptic PDE with discontinuous coefficients also show convergence rates bounded by $1/2$ provided the p-level 1 equations are solved to sufficient accuracy.

References

[1] M. Ainsworth and J. T. Oden, *A Posteriori Error Estimation in Finite Element Analysis*, John Wiley and Sons, New York, 2000.

[2] I. Babuška, M. Griebel, and J. Pitkaranta, *The problem of selecting the shape functions for a p-type finite element*, Internat. J. Numer. Methods Engrg., 28 (1989), pp. 1891–1908.

[3] I. Babuška and M. Suri, *The optimal convergence rate of the p-version of the finite element method*, SIAM J. Numer. Anal., 24 (1987), pp. 750–776.

[4] ——, *The p- and h-p versions of the finite element method, an overview*, Comput. Methods Appl. Mech. Engrg., 80 (1990), pp. 5–26.

[5] R. E. Bank, *Hierarchical bases and the finite element method*, Acta Numerica, 5 (1996), pp. 1–43.

[6] R. E. Bank, T. F. Dupont, and H. Yserentant, *The hierarchical basis multigrid method*, Numer. Math., 52 (1988), pp. 427–458.

[7] F. Bassi, A. Ghidoni, S. Rebay, and P. Tesini, *High-order accurate p-multigrid discontinuous Galerkin solution of the Euler equations*, Internat. J. Numer. Methods Fluids, (2008), p. online.

[8] A. Brandt, *Multi-level adaptive solutions to boundary value problems*, Math. Comp., 31 (1977), pp. 333–390.

[9] G. Brussino, R. Herbin, Z. Christidis, and V. Sonnad, *Parallel multilevel finite element method with hierarchical basis functions*, in Proceedings of the Third SIAM Conference on Parallel Processing for Scientific Computing, G. H. Rodrigue, ed., Philadelphia, 1999, SIAM, pp. 146–150.

[10] A. W. Craig and O. C. Zienkiewicz, *A multigrid algorithm using a hierarchical finite element basis*, in Multigrid Methods for Integral and Differential Equations, D. J. Paddon and Holstein, eds., Oxford, 1985, Clarendon Press, pp. 310–312.

[11] V. Dolean, R. Pasquetti, and F. Rapetti, *p-multigrid for Fekete spectral element method*, in Domain Decomposition Methods in Science and Engineering XVII, vol. 60 of Lecture Notes in Computational Science and Engineering, Springer Berline Heidelberg, 2008, pp. 485–492.

[12] C. C. Douglas and M. B. Douglas, *MGNet Bibliography*, Yale University, Department of Computer Science, New Haven, CT, 1991-2002. http://www.mgnet.org/bib/mgnet.bib.

[13] M. Dryja, B. Smith, and O. Widlund, *Schwarz analysis of iterative substructuring algorithms for elliptic problems in three dimensions*, SIAM J. Numer. Anal., 31 (1994), pp. 1662–1694.

[14] K. J. Fidkowski, T. A. Oliver, J. Lu, and D. L. Darmofal, *p-multigrid solution of high-order discontinuous Galerkin discretizations of the compressible Navier-Stokes equations*, J. Comput. Phys., 207 (2005), pp. 92–113.

[15] D. A. Field and Y. Pressburger, *An h-p-multigrid method for finite element analysis*, Internat. J. Numer. Methods Engrg., 36 (1993), pp. 893–908.

[16] S. Foresti, G. Brussino, S. Hassanzadeh, and V. Sonnad, *Multilevel solution method for the p-version of the finite elements*, Comput. Phys. Comm., 53 (1989), pp. 349–355.

[17] B. Guo and I. Babuška, *The h-p version of the finite element method. Part 1: The basic approximation results*, Comput. Mech., 1 (1986), pp. 21–41.

[18] X. Z. Guo and I. N. Katz, *Performance enhancement of the multi-p preconditioner*, Comput. Math. Appl., 36 (1998), pp. 1–8.

[19] ——, *A parallel multi-p method*, Comput. Math. Appl., 39 (2000), pp. 115–123.

[20] W. Hackbusch, *Multigrid Methods and Applications*, Vol. 4 of Computational Mathematics, Springer-Verlag, Berlin, 1985.

[21] B. T. Helenbrook, *A two-fluid spectral element method*, Comput. Methods Appl. Mech. Engrg., 191 (2001), pp. 273–294.

[22] B. T. Helenbrook and H. L. Atkins, *Coupling p-multigrid to geometric multigrid for discontinuous Galerkin formulations of the Poisson equation*, to appear.

[23] ——, *Application of p-multigrid to discontinuous Galerkin formulations of the Poisson equation*, 44 (2005), pp. 566–575.

[24] ——, *Solving discontinuous Galerkin formulations of Poisson's equation using geometric and p multigrid*, 46 (2008), pp. 894–902.

[25] B. T. Helenbrook, D. Mavriplis, and H. L. Atkins, *Analysis of 'p'-multigrid for continuous and discontinuous finite element discretizations*, AIAA Paper 2003-3989, (2003).

[26] J. J. Heys, T. A. Manteuffel, S. F. McCormick, and L. N. Olson, *Algebraic multigrid (AMG) for higher-order finite elements*, J. Comput. Phys., 204 (2005), pp. 520–532.

[27] N. Hu, X.-Z. Guo, and I. Katz, *Multi-p preconditioners*, SIAM J. Sci. Comput., 18 (1997), pp. 1676–1697.

[28] N. Hu and I. Katz, *Multi-p methods: iterative algorithms for the p-version of finite element analysis*, J. Sci. Comput., 16 (1995), pp. 1308–1332.

[29] C. Liang, R. Kannan, and Z. J. Wang, *A p-multigrid spectral difference method with explicit and implicit smoothers on unstructured triangular grids*, Computers and Fluids, 38 (2009), pp. 254–265.

[30] J. W. Lottes and P. F. Fischer, *Hybrid Schwarz-multigrid methods for the spectral element method*, J. Sci. Comput., 24 (2004), pp. 45–78.

[31] H. Luo, J. D. Baum, and R. Löhner, *A p-multigrid discontinuous Galerkin method for the Euler equations on unstructured grids*, J. Comput. Phys., 211 (2006), pp. 767–783.

[32] ——, *A fast, p-multigrid discontinuous Galerkin method for compressible flows at all speeds*, 46 (2008), pp. 635–652.

[33] Y. Maday and R. Muñoz, *Spectral element multigrid; numerical analysis*, J. Sci. Comput., 3 (1988), pp. 323–354.

[34] ——, *Numerical analysis of a multigrid method for spectral approximations*, in Proc. of the 11th Int. Conf. on Numerical Methods in Fluid Dynamics, M. Hussaini, D. Dwoyer, and R. Voigt, eds., vol. 323 of Lecture Notes in Physics, New York, 1989, Springer, pp. 389–394.

[35] Y. Maday, R. Muñoz, A. T. Patera, and E. M. Rønquist, *Spectral element multigrid methods*, in Proc. of the IMACS Int. Symposium on Iterative Methods in Linear Algebra, P. de Groen and R. Beauwens, eds., Amsterdam, 1992, Elsevier, pp. 191–201.

[36] B. S. Mascarenhas, B. T. Helenbrook, and H. L. Atkins, *Application of the p-multigrid algorithm to discontinuous Galerkin formulations of the compressible Euler equation*, Collection of Technical Papers - 18th AIAA Computational Fluid Dynamics Conference 2, (2007), pp. 1550–1579.

[37] W. F. Mitchell, *Unified multilevel adaptive finite element methods for elliptic problems*, PhD thesis, University of Illinois, Champaign-Urbana, 1988.

[38] ——, *PHAML user's guide*, Tech. Report NISTIR 7374, National Institute of Standards and Technology, Gaithersburg, MD, 2006. Software available at `http://math.nist.gov/phaml`.

[39] W. F. Mitchell and M. A. McClain, *A survey of hp-adaptive strategies for elliptic partial differential equations*, submitted.

[40] P. Morin, R. H. Nochetto, and K. G. Siebert, *Data oscillation and convergence of adaptive FEM*, SIAM J. Numer. Anal., 38 (2000), pp. 466–488.

[41] R. Muñoz, *Theoretical analysis of some spectral multigrid methods*, Comput. Methods Appl. Mech. Engrg., 80 (1990), pp. 287–294.

[42] C. R. Nastase and D. J. Mavriplis, *Discontinuous Galerkin methods using an hp-multigrid solver for inviscid compressible flows on three-dimensional unstructured meshes*, AIAA Paper 2006-0107, (2006).

[43] ——, *High-order discontinuous Galerkin methods using an hp-multigrid approach*, J. Comput. Phys., 213 (2006), pp. 330–357.

[44] M. E. G. Ong, *Hierarchical basis preconditioners in three dimensions*, SIAM J. Sci. Comput., 18 (1997), pp. 479–498.

[45] R. Pasquetti and F. Rapetti, *p-multigrid method for Fekete-Gauss spectral element approximations of elliptic problems*, Communications in Computational Physics, 5 (2009), pp. 667–682.

[46] E. Rønquist and A. Patera, *Spectral element multigrid. I. Formulation and numerical results*, J. Sci. Comput., 2 (1987), pp. 389–406.

[47] P. Šolín, K. Segeth, and I. Doležel, *Higher-Order Finite Element Methods*, Chapman & Hall/CRC, New York, 2004.

[48] E. Süli, P. Houston, and C. Schwab, *hp-finite element methods for hyperbolic problems*, in The Mathematics of Finite Elements and Applications X. MAFELAP, J. Whiteman, ed., Elsevier, 2000, pp. 143–162.

[49] D. K. Sun, J. F. Lee, and Z. Cendes, *Construction of nearly orthogonal Nedelec bases for rapid convergence with multilevel preconditioned solvers*, SIAM J. Sci. Comput., 23 (2001), pp. 1053–1076.

[50] B. Szabo and I. Babuška, *Finite Element Analysis*, John Wiley and Sons, New York, 1991.

[51] K. van den Abeele, T. Broeckhoven, and C. Lacor, *Dispersion and dissipation properties of the 1D spectral volume method and application to a p-multigrid method*, J. Comput. Phys., 224 (2007), pp. 616–636.

[52] E. L. Wilson, *The static condensation algorithm*, Internat. J. Numer. Methods Engrg., 8 (1974), pp. 198–203.

[53] H. Yserentant, *On the multi-level splitting of finite element spaces*, Numer. Math., 49 (1986), pp. 379–412.