The numerical performance of wavelets for PDEs: the multi-scale finite element

Mark Christon, David Roach

The numerical performance of wavelets for PDEs: the multi-scale finite element

M. A. Christon
Computational Physics R&D Department,
Sandia National Laboratories,
M/S 0819, P.O. Box 5800 Albuquerque,
New Mexico 87185-0819, USA

D. W. Roach
Mathematics Department, University of Georgia, USA

Sandia is a multiprogram laboratory operated by Sandia Corporation for the United States Department of Energy under Contract DE-AC04-94AL85000.

Abstract The research summarized in this paper is part of a multi-year effort focused on evaluating the viability of wavelet bases for the solution of partial differential equations. The primary objective for this work has been to establish a foundation for hierarchical/wavelet simulation methods based upon numerical performance, computational efficiency, and the ability to exploit the hierarchical adaptive nature of wavelets. This work has demonstrated that hierarchical bases can be effective for problems with a dominant elliptic character. However, the strict enforcement of orthogonality in the usual L^2 sense is less desirable than orthogonality in the energy norm. This conclusion has led to the development of a multi-scale linear finite element based on a hierarchical change-of-basis. This work considers the numerical and computational performance of the hierarchical Schauder basis in a Galerkin context. A unique row-column lumping procedure is developed with multi-scale solution strategies for 1-D and 2-D elliptic partial differential equations.

1 Introduction

Wavelets are a relatively new mathematical tool that dissect data, functions, and differential operators into components with an associated resolution that is matched to the scale of each component. The name "wavelet" or "ondelette" was coined in the early 1980s by French researchers Morlet, Arens, Fourgeua, Giard and Grossman [9, 8, 5]. However, functions with the attributes of wavelets have been known for almost 100 years. Meyer [7] points out that there are seven primary origins for wavelets that date from around 1930 with the Haar wavelet dating back to 1909. However, the literature from this era does not explicitly use the term "wavelet".

In recent years, the application of wavelet bases to the solution of partial differential equations (PDEs) has evolved to the point where there are a number of competing formulations that include, but are not limited to wavelet-Galerkin, wavelet-collocation, and reproducing kernel methods. Despite the growing number of wavelet-based formulations and solution algorithms for PDEs, the field is still very new, and many technical issues remain. That is to say, there is promise in the approach, but there is a clear need for fundamental research that characterizes the numerical and computational performance of wavelets for the solution of partial differential equations.

A complete historical perspective of wavelet based approaches to solving PDEs is beyond the scope of this paper. However, a survey of the literature on wavelet based methods for solving PDEs may be found in Christon et al. [2]. Over the past 8 years, there has been a great deal of attention paid to the wavelet-Galerkin method because of the generality that the Galerkin method provides, and the ease with which alternative bases may be implemented and tested. In addition, the trend in recent years has been away from traditional Daubachies wavelets with L^2 orthogonality and towards wavelets that are constructed with weaker orthogonality properties yielding biorthogonal and semi-orthogonal wavelets (see for example Dahmen and Micchelli [4]). This approach has led to the idea that wavelets may be most useful for solving PDEs if they are used to simply "complete" a subspace with the concomitant relationship between grid scales. A related approach that combines a spectral method with biorthogonal wavelet bases adapted to non-square domains and higher spatial dimensions may be found in Canuto et al. [1].

In this work, orthogonality with respect to the energy norm is used with a nodal change-of-basis to develop multi-scale finite elements that make use of a hierarchical Schauder basis. In 1-D, the uniform stability of the Schauder basis in $\mathscr{H}^1$ yields uniformly bounded condition numbers, i.e., independent of refinement level (grid resolution), for elliptic operators in a Galerkin form. This property was previously described by Zienkiewicz [12] as the "diagonality" feature for one dimensional hierarchical shape functions. However, the relationship between uniform stability, a nodal change-of-basis and the wavelet transform was not considered. In the ensuing discussion, the theoretical issues surrounding the hierarchical Schauder basis are presented, and the details of a uniformly stable basis are made concrete. The use of the wavelet transform in conjunction with the assembly of the mass and stiffness operators for a hierarchical basis is presented in the framework of 1-D and 2-D multi-scale finite elements. In addition, a comparison between the hierarchical Schauder basis and the linear finite element

basis is presented. The application of an ad-hoc row-column lumping procedure for the hierarchical basis is outlined, and its effectiveness for solving elliptic boundary value problems in one and two-dimensions is demonstrated.

2 The hierarchical Schauder basis

In this section, the formulation issues associated with the hierarchical Schauder basis are outlined. For this development, the boundary value problem under consideration is

$$-\varepsilon\nabla^2 u + \gamma u = f \text{ on } \Omega \ , \tag{1}$$

with homogeneous Dirichlet boundary conditions

$$u = 0 \text{ on } \Gamma \ . \tag{2}$$

The weak form of the problem, after introducing the finite dimensional subspace, $\mathscr{V}_k$, is

$$\int_\Omega \varepsilon\nabla u_k \cdot \nabla\mathbf{\Phi}_k \, d\Gamma + \int_\Omega \gamma u_k \mathbf{\Phi}_k \, d\Gamma = \int_\Gamma f\mathbf{\Phi}_k \, d\Gamma \ . \tag{3}$$

Here, $u_k \in \mathscr{H}_0^1$, and $\mathbf{\Phi}_k$ is a finite dimensional basis for the subspace $\mathscr{V}_k$ at scale k. In matrix form, this may be written as

$$[\mathbf{M}_k + \mathbf{K}_k]\{u_k\} = \{F_k\} \tag{4}$$

where $\mathbf{M}_k$ is the mass matrix, and $\mathbf{K}_k$ is the stiffness matrix associated with scale k. Alternatively, in a bilinear form, $a(\mathbf{\Phi}_k, \mathbf{\Phi}_k)u_k = F(\mathbf{\Phi}_k)$.

Now, consider an alternative basis for $\mathscr{V}_k$,

$$\mathbf{\Psi}_k = \mathbf{W}_k^{\mathrm{T}}\mathbf{\Phi}_k \ , \tag{5}$$

where $\mathbf{W}_k$ is a nonsingular $N \times N$ matrix. From this transformation, $u_k = \Delta u^{\mathrm{T}}\mathbf{\Psi}$ may also be found by solving

$$a(\mathbf{\Psi}_k, \mathbf{\Psi}_k)\Delta u = F(\mathbf{\Psi}_k) \tag{6}$$

for the multi-scale representation of the field, Δu. The linear system, $\mathbf{A}_k^{\mathbf{\Psi}}$, that derives from the change of basis may also be viewed as the linear system obtained by preconditioning $\mathbf{A}_k^{\mathbf{\Phi}} = [\mathbf{M}_k + \mathbf{K}_k]$ with $\mathbf{W}_k$, i.e.,

$$\mathbf{A}_k^{\mathbf{\Psi}} = \mathbf{W}_k^{\mathrm{T}}\mathbf{A}_k^{\mathbf{\Phi}}\mathbf{W}_k \ . \tag{7}$$

Note that the superscript $\mathbf{\Psi}$ ($\mathbf{\Phi}$) indicates that $\mathbf{\Psi}_k$ ($\mathbf{\Phi}_k$) is used as both the test and trial function in obtaining the weak form.

The matrix $\mathbf{W}_k$ is a wavelet transform and provides the mechanism for decomposing a field into multiple scales. In order to understand the role of the wavelet transform, consider that

$$\mathscr{V}_0 \subset \mathscr{V}_1 \subset \cdots \subset \mathscr{V}_k \subset \cdots \tag{8}$$

is a one-sided sequence of nested finite-dimensional subspaces of $\mathscr{H}$, a Hilbert space, such that $\overline{\bigcup \mathscr{V}_k} = \mathscr{H}$. Now, define the "coarse-grid" wavelet subspace, $\mathscr{W}_0 = \mathscr{V}_0$ and, for $k \geq 1$, choose $\mathscr{W}_k$ so that completion of subspace $\mathscr{V}_{k-1}$ may be achieved via a direct sum,

$$\mathscr{V}_k = \mathscr{V}_{k-1} \oplus \mathscr{W}_k \ . \tag{9}$$

Here $\oplus$ denotes a direct sum, but not an orthogonal direct sum.

Beginning with Eq. (5), the wavelet transform may be written recursively in terms of a two-scale transform as

$$\mathbf{W}_k = T_k \begin{bmatrix} T_{k-1} & 0 \\ 0 & I_{k-1} \end{bmatrix} \cdots \begin{bmatrix} T_1 & 0 \\ 0 & I_1 \end{bmatrix} \tag{10}$$

where $\mathbf{I}_k$ is the $N_{\mathbf{\Psi}} \times N_{\mathbf{\Psi}}$ identity matrix with $N_{\mathbf{\Psi}} = \dim(\mathscr{W}_k)$. Here, T_k is a two-scale transform such that

$$\begin{Bmatrix} \mathbf{\Phi}_{k-1} \\ \mathbf{\Psi}_k \end{Bmatrix} = T_k^{\mathrm{T}}\mathbf{\Phi}_k \ . \tag{11}$$

As an aside, one very important aspect of this formulation is that it does not require the decomposition matrix $\mathbf{W}_k^{-1}$. This is significant because it permits the relaxation of strict orthogonality in the selection of the wavelet bases.

One choice for $\mathscr{W}_k$ that satisfies Eq. (9) is the hierarchical Schauder basis [3, 10–13]. This basis consists of the elements of the linear finite element basis at multiple scales where the relationship between $\mathbf{\Psi}_k$ and $\mathbf{\Phi}_k$ is simply

$$\psi_k^j = \phi_k^{2j-1}, \quad j = 1, \ldots, \mathrm{Nnp}_k \ , \tag{12}$$

where $\mathbf{\Phi}_k$ is the linear finite element basis. Here the superscript j indicates the node number, k indicates the scale, and Nnp_k is the number of node points at scale k. The Schauder basis is illustrated in Fig. 1.

With the relationship between basis elements in the Schauder basis given by Eq. (12), the construction of the two-scale transformation may be written in terms of the grid spacing. In order to define the wavelet transform, let $\mathbf{H}_k$ be the $(\mathrm{Nnp}_k - 1) \times (\mathrm{Nnp}_{k-1} - 1)$ matrix $\mathbf{H}_k = (h_k^{j,j'})_{j',j}$ and let $\mathbf{G}_k$ be the $(\mathrm{Nnp}_k - 1) \times (\mathrm{Nnp}_{k-1})$ matrix

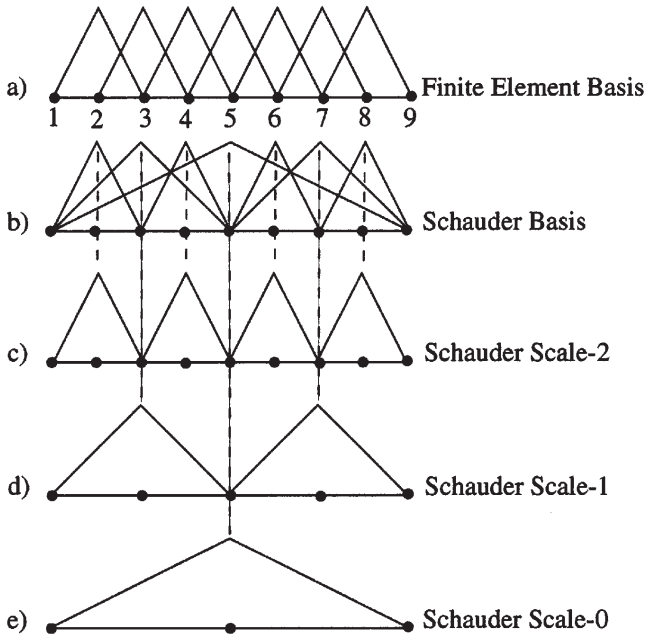


Fig. 1. The hierarchical Schauder basis for scale $k = 2$ with 8 elements and 9 nodes showing **a** the equivalent finite element basis, **b** the composite basis, **c** the basis functions for Scale-2, **d** the basis functions for Scale-1, and **e** the basis functions for Scale-0

$\mathbf{G}_k = (g_k^{j,j'})_{j',j}$. Then, the two-scale transformation for the Schauder basis is given by

$$T_k = [\mathbf{H}_k|\mathbf{G}_k] \ . \tag{13}$$

The function values $h_k^{j,j'}$ and $g_k^{j,j'}$ for ϕ_{k-1}^j and ψ_k^j respectively at the node points x_k^j are given by

$$h_k^{j,j'} = \begin{cases} \frac{\Delta_k^{j'}}{\Delta_{k-1}^j}, & j' = 2j-1 \\ 1, & j' = 2j \\ \frac{\Delta_k^{j'+1}}{\Delta_{k-1}^j}, & j' = 2j+1 \\ 0, & \text{otherwise} \end{cases} \tag{14}$$

$$g_k^{j,j'} = \delta_{j',2j-1} \ . \tag{15}$$

In one dimension, the element length from x_k^{j-1} to x_k^j is defined as $\Delta_k^j = x_k^j - x_k^{j-1}$.

The action of $h_k^{j,j'}$ applied at scale k is to average, while $g_k^{j,j'}$ acts to inject at scale k. This can be seen in Fig. 1 where the relationship between basis elements is given by

$$\phi_{k-1}^j = \sum_{j'} h_k^{j,j'} \phi_k^{j'} \tag{16}$$

$$\psi_k^j = \sum_{j'} g_k^{j,j'} \phi_k^{j'} \ . \tag{17}$$

With the Schauder basis defined, attention is turned to the question of stability. To begin the discussion on stability, several definitions are required.

Definition 1 *Given a basis $\mathbf{\Phi} = (\phi^1, \phi^2, \phi^3, \ldots)^{\mathrm{T}}$ for a subspace, $\mathscr{V}$, of a Hilbert Space, $\mathscr{H}$, with norm $\|\cdot\|_{\mathscr{H}} = \sqrt{(\cdot,\cdot)_{\mathscr{H}}}$, define the associated discretized matrix as*

$$\mathbf{A}^{\mathbf{\Phi}} = a(\mathbf{\Phi}, \mathbf{\Phi})_{\mathscr{H}} \ . \tag{18}$$

Definition 2 *Suppose the vector $\mathbf{\Phi}$ is a normalized basis for a Hilbert space, $\mathscr{H}$. Then $\mathbf{\Phi}$ is said to be stable in $\mathscr{H}$ if there exists constants $\alpha, \beta > 0$ such that*

$$\alpha \|u\|_{\ell^2}^2 \leq \|u^{\mathrm{T}}\mathbf{\Phi}\|_{\mathscr{H}}^2 \leq \beta \|u\|_{\ell^2}^2 \ , \tag{19}$$

where $\|u\|_{\ell^2}^2 = \sum_i u_i^2$.

This definition of stability is equivalent to the definition of a Riesz basis with respect to the $\|\cdot\|_{\mathscr{H}}$ norm.

Now consider the following set of nested subspaces of $\mathscr{H}$,

$$\mathscr{V}_0 \subset \mathscr{V}_1 \subset \cdots \subset \mathscr{V}_k \cdots \subset \mathscr{H} \ , \tag{20}$$

such that $\mathbf{\Phi}_k$ is a basis for $\mathscr{V}_k$.

Definition 3 *Given a nested sequence of subspaces of a Hilbert space, $\mathscr{H}$, and their respective bases $\{\mathbf{\Phi}_k\}_{k\in\mathbb{Z}^+}$ (normalized in $\mathscr{H}$), then $\{\mathbf{\Phi}_k\}_{k\in\mathbb{Z}^+}$ is uniformly stable in $\mathscr{H}$ if there exists constants $\alpha, \beta > 0$ independent of k such that*

$$\alpha \|u_k\|_{\ell^2}^2 \leq \|u_k^{\mathrm{T}}\mathbf{\Phi}_k\|_{\mathscr{H}}^2 \leq \beta \|u_k\|_{\ell^2}^2 \quad k = 0, 1, \ldots \tag{21}$$

With these definitions, the relationship between the stability of a basis and the condition number of an operator discretized in the basis is outlined. In specific, it will be demonstrated that a uniformly stable set of bases yields uniformly bounded condition numbers for the discrete operators independent of the refinement level.

Lemma 1 *Let $\mathbf{\Phi}_k$ be a basis for a subspace, $\mathscr{V}_k$, of a Hilbert space $\mathscr{H}$ such that $\mathbf{A}_k^{\mathbf{\Phi}}$ is positive definite. Let $\alpha = \|\mathbf{A}_k^{\mathbf{\Phi}^{-1}}\|_2^{-1}$ and $\beta = \|\mathbf{A}_k^{\mathbf{\Phi}}\|_2$. Then*

$$\alpha \|u_k\|_{\ell_2}^2 \leq \|u_k^{\mathrm{T}}\mathbf{\Phi}_k\|_{\mathscr{H}}^2 \leq \beta \|u_k\|_{\ell^2}^2 \ , \tag{22}$$

and

$$\mathrm{cond}(\mathbf{A}_k^{\mathbf{\Phi}}) = \frac{\beta}{\alpha} \ . \tag{23}$$

Proof The condition number of a symmetric matrix is given by $\mathrm{cond}(\mathbf{A}_k^{\mathbf{\Phi}}) = \|\mathbf{A}_k^{\mathbf{\Phi}}\|_2 \|[\mathbf{A}_k^{\mathbf{\Phi}}]^{-1}\|_2$.

For the stability condition, note that

$$\frac{\|u_k^{\mathrm{T}}\mathbf{\Phi}_k\|_{\mathscr{H}}^2}{\|u_k\|_{\ell^2}^2} = \frac{\left(u_k^{\mathrm{T}}\mathbf{\Phi}_k, u_k^{\mathrm{T}}\mathbf{\Phi}_k\right)_{\mathscr{H}}}{u_k^{\mathrm{T}}u_k} = \frac{u_k^{\mathrm{T}}\mathbf{A}_k^{\mathbf{\Phi}}u_k}{u_k^{\mathrm{T}}u_k} \ . \tag{24}$$

Moreover, for symmetric matrices

$$\beta = \|\mathbf{A}_k^{\mathbf{\Phi}}\|_2 = \rho(\mathbf{A}_k^{\mathbf{\Phi}}) = \max_{u_k} \frac{u_k^{\mathrm{T}}\mathbf{A}_k^{\mathbf{\Phi}}u_k}{u_k^{\mathrm{T}}u_k} \tag{25}$$

which gives us the upper bound. For the lower bound, observe that for symmetric positive definite matrices

$$\alpha = \|\mathbf{A}_k^{\mathbf{\Phi}^{-1}}\|_2^{-1} = \left(\max_{u_k} \frac{u_k^{\mathrm{T}}\mathbf{A}_k^{\mathbf{\Phi}^{-1}}u_k}{u_k^{\mathrm{T}}u_k}\right)^{-1} = \min_{u_k} \frac{u_k^{\mathrm{T}}\mathbf{A}_k^{\mathbf{\Phi}}u_k}{u_k^{\mathrm{T}}u_k} \ . \tag{26}$$

The proof of Lemma 1 states that the best stability bounds are provided by the smallest and largest eigenvalues of the discretized matrix – a well known result.

Now consider the stability of a linear finite element basis, shown in Fig. 1a for $k = 2$, where $\mathscr{H} = L^2(0,2)$ and

$$\mathscr{V}_k = \mathrm{span}\{\phi_k^j, j = 1, \ldots, 2^{k+1} - 1\}, \ k = 0, 1, \ldots, \tag{27}$$

$$\mathbf{\Phi}_k = \left(\phi_k^1, \phi_k^2, \ldots, \phi_k^{2^{k+1}-1}\right)^{\mathrm{T}} \ . \tag{28}$$

Lemma 2 *The sequence of bases $\mathbf{\Phi}_k$ for the linear finite element basis is uniformly stable in $L^2(0,2)$.*

Proof Let the discretized matrix $\mathbf{A}_k^{\mathbf{\Phi}}$ be a diagonally scaled finite element mass matrix. In 1-D this is a symmetric tridiagonal matrix consisting of 1 on the diagonal and 1/4 on the non-zero off diagonals. Note that for the ensuing proof, $\mathbf{\Phi}_k$ has been normalized in a manner that is equivalent to diagonal preconditioning applied to the finite element mass matrix. Using Gershgorin's theorem, the eigenvalues for $\mathbf{A}_k^{\mathbf{\Phi}}$ lie in the interval $[1/2, 3/2]$ giving

$$\tfrac{1}{2}\|u_k\|_{\ell^2}^2 \leq \|u_k^{\mathrm{T}}\mathbf{\Phi}_k\|_{L^2}^2 \leq \tfrac{3}{2}\|u_k\|_{\ell^2}^2 \ , \tag{29}$$

and $\mathrm{cond}(\mathbf{A}_k^{\mathbf{\Phi}}) \leq 3$. Therefore, the sequence of bases, $\mathbf{\Phi}_k$, is uniformly stable in L^2.

The stability associated with the finite element basis in L^2 is consistent with the empirical observation that the consistent mass matrix is well behaved in terms of its

condition number. In practice, this is reflected in the ability to easily solve mass matrix dominated problems with simple iterative techniques.

Theorem 1 (*Poincaré-Friedrich*) *For bounded domains, the $\mathscr{H}^m$ norm and semi-norm are equivalent in the sense that there exists constants $\alpha, \beta > 0$ such that*

$$\alpha|u|_{\mathscr{H}^m} \leq \|u\|_{\mathscr{H}^m} \leq \beta|u|_{\mathscr{H}^m} \quad \text{for } u \in \mathscr{H}_0^m \ . \tag{30}$$

Recall that the norm and semi-norm for the Sobolev space $\mathscr{H}^1$ for one space dimension are

$$\|u\|_{\mathscr{H}^1} = \|u'\|_{L^2} + \|u\|_{L^2}$$
$$|u|_{\mathscr{H}^1} = \|u'\|_{L^2} \ .$$

Lemma 3 *After re-normalization with respect to the $\mathscr{H}^1$ semi-norm, the sequence of bases $\mathbf{\Phi}_k$ is not uniformly stable in $\mathscr{H}_0^1(0,2)$.*

Proof Using the equivalence of the norm and semi-norm, a sequence of vectors u_k is constructed for which uniform bounds are not possible. If

$$u_k = (\overbrace{1,1,1,\ldots,1,1}^{2^{k+1}-1}) \quad \text{for } k = 0,1,2,\ldots, \tag{31}$$

then

$$\|u_k\|_{\ell^2}^2 = 2^{k+1} - 1$$

and

$$|u_k^{\mathrm{T}}\mathbf{\Phi}_k|_{\mathscr{H}^1}^2 = 1 \ \text{ for } k = 0,1,2,\ldots, \tag{32}$$

and stability for the basis can not be demonstrated. Therefore, using the equivalence of the $\mathscr{H}^1$ norm and semi-norm, there are no uniform bounds for the sequence of bases $\mathbf{\Phi}_k$ in $\mathscr{H}_0^1$.

Next, the hierarchical Schauder Basis is considered where $\mathbf{\Phi}_0$ corresponds to the linear finite element basis described above. Let

$$\mathscr{W}_k = \text{span}\left\{\phi_k^j, \ j = 0,2,4,\ldots,2^{k+1}-2\right\} \tag{33}$$

$$\mathscr{V}_k = \mathscr{V}_0 \bigoplus_{j=1}^{k} \mathscr{W}_j \tag{34}$$

$$\mathbf{\Psi}_k = \left(\phi_0^0, \phi_1^0, \phi_1^2, \ldots, \phi_k^0, \phi_k^2, \ldots, \phi_k^{2^{k+1}-2}\right)^{\mathrm{T}}. \tag{35}$$

The hierarchical basis for $k = 2$ is shown in Fig. 1.

Lemma 4 *After re-normalization with respect to the $\mathscr{H}^1$ semi-norm, the sequence of bases $\mathbf{\Psi}_k$ for the Schauder basis is uniformly stable in $\mathscr{H}_0^1(0,2)$.*

Proof The $\mathscr{H}^1$ semi-norm of $\mathbf{\Psi}_k$ is equivalent to the L^2 norm of the orthogonal Haar wavelet basis. The Haar basis is an orthogonal basis for $L^2[0,2]$ and thus $\mathbf{A}_k^{\mathbf{\Psi}}$ is the identity with respect to the semi-norm for $k = 0,1,2,\ldots,\infty$.

Lemma 5 *After re-normalization with respect to the L^2 norm, the sequence of bases $\mathbf{\Psi}_k$ for the Schauder basis is not uniformly stable in $L^2(0,2)$.*

Proof Define the sequence u_k which gives the hat function centered at one with support $[1-2^k, 1+2^k]$ as

$$u_k = (1, -1/2, -1/2, 0, -1/4, -1/4, \\ 0,0,0,0,-1/8,-1/8,0,0,0,\ldots) \ . \tag{36}$$

Thus,

$$\|u_k\|_{\ell^2}^2 \geq 1 \ , \tag{37}$$

and

$$\|u_k^{\mathrm{T}}\mathbf{\Psi}_k\|_{L^2}^2 = \frac{2}{3}\frac{1}{2^k} \ . \tag{38}$$

Because $\|u_k^{\mathrm{T}}\mathbf{\Psi}_k\|_{L^2}^2 \to 0$ as $k \to \infty$ the stability condition in Eq. (3) can not be met.

3 The multi-scale finite element

This section outlines the 1-D and 2-D multi-scale finite elements. The multi-scale elements are based upon the hierarchical Schauder basis with a change-of-basis incorporated at the element-level in order to make use of the well-known finite element assembly procedure [6].

Before embarking on a description of the 1-D multi-scale finite element, a brief review and interpretation of the multi-scale transformation is presented. Recall from Eq. (5) that $\mathbf{\Psi}_k = \mathbf{W}_k^{\mathrm{T}}\mathbf{\Phi}_k$ where $\mathbf{W}_k$ is the multi-scale transformation matrix. In order to make this transformation concrete, consider the following example. Beginning with a 1-D grid consisting of 5 grid points and 4 linear finite elements, the nodal basis will be decomposed into a coarse-grid consisting of two elements and the associated "psuedo-wavelets". This decomposition is shown schematically in Fig. 2.

Remark 1 *The term pseudo-wavelet is used here to indicate that the elements of $\mathscr{W}_k$ in the hierarchical Schauder basis do not possess the property that their zeroth moment is zero. However, the elements of $\mathscr{W}_k$ used to complete the subspace $\mathscr{V}_{k-1}$ are uniformly stable in $\mathscr{H}^1$. Thus, the term pseudo-wavelet seems appropriate.*

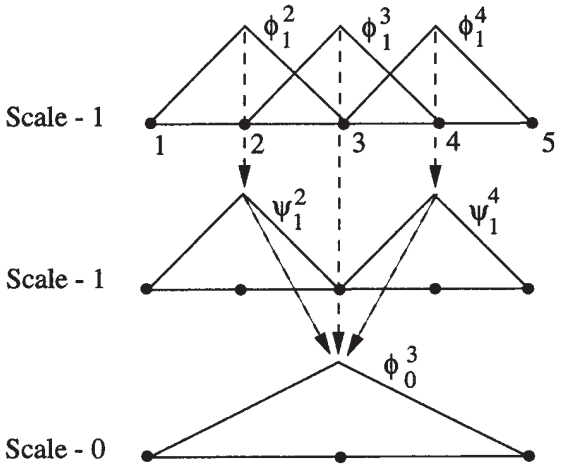


Fig. 2. One-dimensional two-scale decomposition of a finite element nodal basis at scale-1

In this example, the wavelet transform, $\mathbf{\Psi}_1 = W_1^{\mathrm{T}}\mathbf{\Phi}_1$, is

$$\begin{Bmatrix} \phi_0^3 \\ \psi_1^2 \\ \psi_1^4 \end{Bmatrix} = \begin{bmatrix} 1/2 & 1 & 1/2 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} \phi_1^2 \\ \phi_1^3 \\ \phi_1^4 \end{Bmatrix} \ , \tag{39}$$

where the subscript indicates the scale, and the superscript indicates the node number. From this, it is clear that the wavelet transform performs an averaging procedure to obtain the coarse-grid basis elements, $\mathbf{\Phi}_0$, and an injection to obtain the coarse-grid pseudo-wavelets, $\mathbf{\Psi}_1$.

The decomposition of nodal variables, u, may be accomplished with the wavelet transform, but its inverse is required to obtain the coarse-grid coefficients, i.e., $\Delta u = \mathbf{W}^{-1}u$, where Δu is the multi-scale component of the field. Relying on the inverse wavelet transform is impractical because the orthogonality constraint between the wavelets and scaling functions has been relaxed in the hierarchical Schauder basis. However, incorporating the hierarchical Schauder basis at the element level yields solution algorithms that compute the multi-scale solution directly and rely only on the reconstruction algorithm, i.e., $u = \mathbf{W}\Delta u$. Thus, given a multi-scale representation of the field that corresponds to the multi-scale basis elements in Fig. 2, the reconstruction algorithm is simply

$$\begin{Bmatrix} u_1^2 \\ u_1^3 \\ u_1^4 \end{Bmatrix} = \begin{bmatrix} 1/2 & 1 & 0 \\ 1 & 0 & 0 \\ 1/2 & 0 & 1 \end{bmatrix} \begin{Bmatrix} u_0^3 \\ \Delta u_1^2 \\ \Delta u_1^4 \end{Bmatrix} \ . \tag{40}$$

As an aside, the wavelet transform is comprised of two components and both may be viewed in terms of a discrete convolution. Using the nomenclature introduced above for the wavelet transform, $T_1 = [\mathbf{H}_1|\mathbf{G}_1]$, where

$$\mathbf{H}_1 = \begin{Bmatrix} 1/2 \\ 1 \\ 1/2 \end{Bmatrix}, \quad \mathbf{G}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \ . \tag{41}$$

In this simple example of a two-scale decomposition, $\mathbf{W}_1 = T_1$, but in general, the wavelet transform is computed recursively according to Eq. (10).

The 1-D multi-scale element

The description of the multi-scale element begins with the linear finite element for which the shape functions are

$$N_i = \tfrac{1}{2}(1 - \xi_i\xi) \ . \tag{42}$$

Here, ξ is the natural coordinate, ξ_i is the nodal value of the natural coordinate, $i = 1, 2$, and $-1 \le \xi \le 1$.

The concept of scale is introduced at the element level by injecting degrees-of-freedom (DOF) that are supported by the pseudo-wavelets of the hierarchical Schauder basis. The 1-D multi-scale element is shown in Fig. 3 where a single "internal" degree-of-freedom located at $\xi = 0$ in the element is introduced at Scale-1, two DOF are introduced at Scale-2, and four DOF at Scale-3.

At Scale-1, the pseudo-wavelet is

$$\psi_1(\xi) = \begin{cases} 1 + \xi & \text{if } -1 \le \xi \le 0 \\ 1 - \xi & \text{if } \ \ 0 \le \xi \le 1 \end{cases} \ . \tag{43}$$

More generally, the psuedo-wavelets for the multi-scale DOF may be written in terms of the translates and dilates of $\psi(\xi)$ as

$$\psi_k(\xi) = \psi(\tilde{\xi}) \ , \tag{44}$$

where

$$\tilde{\xi} = 2^{k-1}(1 + \xi) - 2j - 1 \ , \tag{45}$$

and

$$\begin{aligned} 2^{2-k}j - 1 \le \ \xi \ \le 2^{2-k}j + 2^{1-k} - 1 \\ 2^{2-k}j + 2^{1-k} - 1 \le \ \xi \ \le 2^{2-k}j + 2^{2-k} - 1 \ . \end{aligned} \tag{46}$$

Here k indicates the scale, and j indicates the translates in the element parametric space ($-1 \le \xi \le 1$).

The derivatives of the shape functions yield constant functions that are orthogonal to the derivatives of the pseudo-wavelets at all scales. The derivatives of the pseudo-wavelets yield Haar wavelets as illustrated in Fig. 3, and at any given scale they are orthogonal with the derivatives at all other scales in the 1-D multi-scale element.

The reconstruction algorithm may be viewed as an element-by-element procedure that relies only on the multi-scale information in each element. The reconstruction is shown schematically in Fig. 3 where the DOF located at $\xi = -3/4$ is computed as a linear combination of the detail, Δu, at scales $0-3$. Here, a canonical node-numbering scheme is used where the node numbers are n_1, n_2 at Scale-0, n_3 at Scale-1, n_4, n_5 at Scale-2, etc. as shown in Fig. 3.

With this numbering scheme, the reconstruction for the DOF located at n_6 may be written as

$$u_6 = N_1u_0^1 + N_2u_0^2 + \psi_1^3\Delta u_1^3 + \psi_2^4\Delta u_2^4 + \psi_3^6\Delta u_3^6 \ , \tag{47}$$

where each of the basis elements is evaluated at $\xi = -3/4$. Thus, the reconstruction algorithm begins with the interpolant of the coarse-grid solution and injects refinements, or detail, up to the desired scale. The reconstruction may be written more generally as

$$u_k^{2+2^{k-1}+j} = \sum_{i=1}^{\mathrm{Nnpe}} N_i u_0^i + \sum_{k=1}^{\mathrm{Nscale}} \psi_k^{2+2^{k-1}+j}\Delta u_k^{2+2^{k-1}+j} \tag{48}$$

where the basis elements (N_i, and ψ_k) are evaluated at locations corresponding to the DOF location in the parent element.

The 1-D multi-scale operators

The computation of the stiffness for the model problem, Eq. (1), is a straight-forward procedure that begins with the coarse-grid stiffness.

At the element level, i.e., Scale-0, this is simply

$$K_0^e = \frac{\varepsilon}{h}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \ , \tag{49}$$

where h is the node-spacing for the coarse-grid. Making use of orthogonality, the diagonal stiffness entries for the scale DOF, i.e., for $k > 0$, are

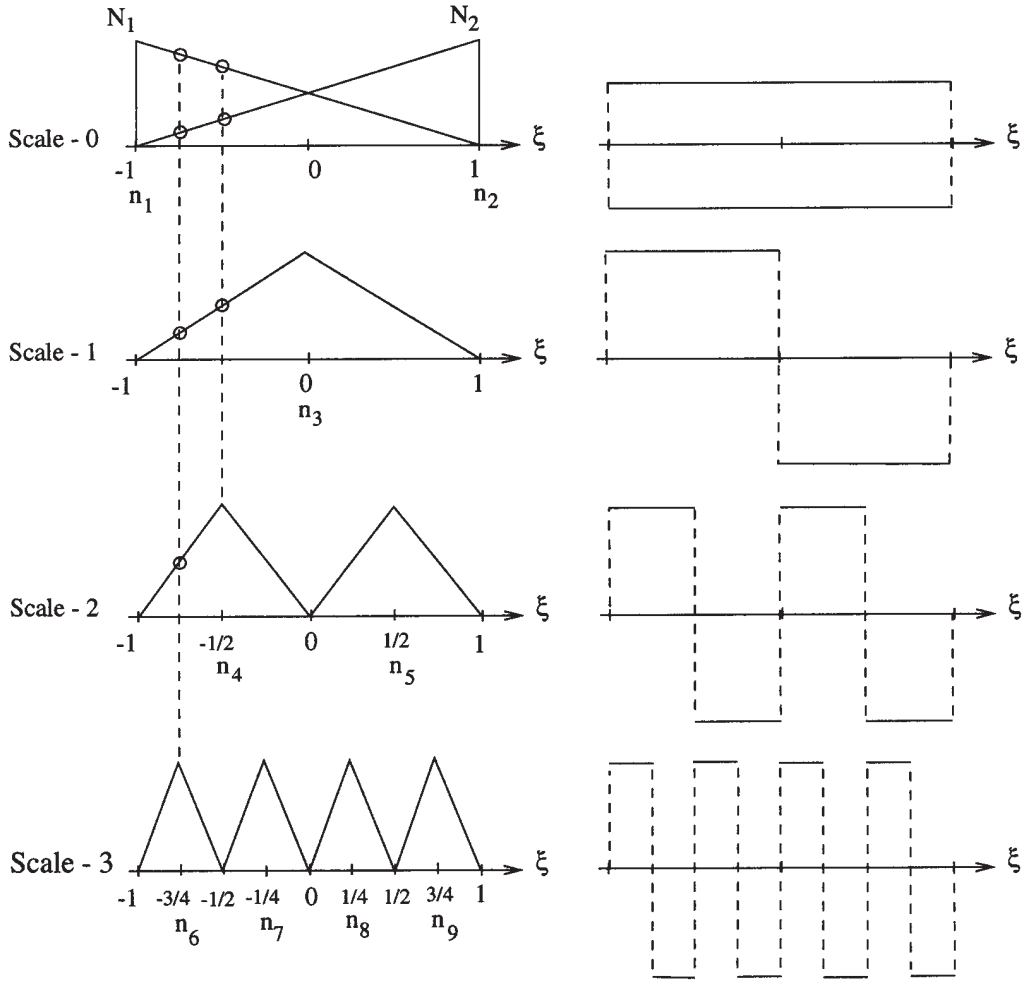


Fig. 3. Basis elements and their piecewise derivatives for the one-dimensional multi-scale element with three refinement scales. The derivatives have been scaled by $1/k$ for scales $k = 1, 2, 3$

$$K_k^{2+2^{k-1}+j} = \frac{2^{k+1}\varepsilon}{h} \quad . \tag{50}$$

The resulting multi-scale element-level stiffness matrix is

$$K^e = \frac{\varepsilon}{h}\begin{bmatrix} 1 & -1 & & & & & \\ -1 & 1 & & & & & \\ & & 4 & & & & \\ & & & 8 & & & \\ & & & & 8 & & \\ & & & & & \ddots & \\ & & & & & & 2^{k+1} \end{bmatrix} . \tag{51}$$

With this form of the element stiffness, only the coarse-scale terms contribute to the element assembly procedure since all internal DOF simply require a point evaluation and do not rely on information outside the element. Here, a uniform mesh refinement was applied to inject the scale DOF, but this is not a necessity for the multi-scale element.

The mass matrix computation is somewhat more involved because it requires the calculation of inner-products that involve the basis elements across scale. The computation of the multi-scale mass matrix consists of the following

$$\begin{aligned} M_0^e &= \int_{\Omega^e} \gamma N_i N_j \, \mathrm{d}\Omega, \quad M_{01} = \int_{\Omega^e} \gamma N_i \psi_1 \mathrm{d}\Omega, \\ M_{10} &= \int_{\Omega^e} \gamma \psi_1 N_j \, \mathrm{d}\Omega, \quad M_{11} = \int_{\Omega^e} \gamma \psi_1 \psi_1 \mathrm{d}\Omega \\ M_{12} &= \int_{\Omega^e} \gamma \psi_1 \psi_2 \, \mathrm{d}\Omega, \ldots \end{aligned} \tag{52}$$

This series of element-level integrals leads to the finger-diagonal matrix structure shown in Fig. 4 with the inset coarse-grid element mass matrix.

A row-column lumping procedure is shown schematically in Fig. 4 for the multi-scale DOF at n_7 in the element. Note that the traditional row-sum lumping for the coarse-grid mass matrix can be used for the entries corresponding to the coarse-grid, but this type of mass lumping can not be used for the multi-scale DOF in the element. This concept is discussed further in the comparison between the multi-scale and linear finite elements.

A multi-scale algorithm

As an example of a multi-scale algorithm, consider the model problem, Eq. (2), with $\varepsilon = 1$, $\gamma = 0$, $f(x) = 1$, and essential boundary conditions, $u(0) = 0$, $u(1) = 0$. In

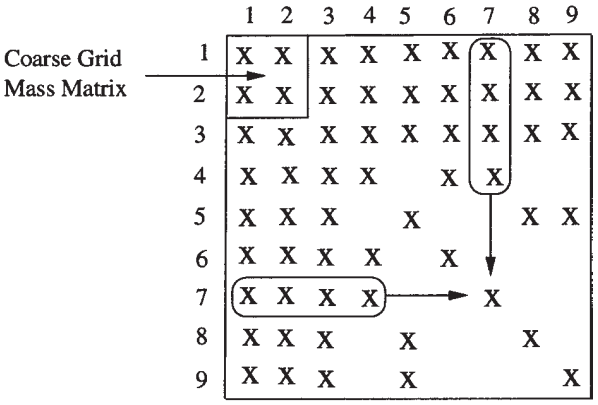


Fig. 4. Finger diagonal structure of multi-scale element mass matrix

this example, three scale solutions are computed. Scale-0 corresponds to the coarse-grid solution using two elements. Scale-1 corresponds to the injection of one multi-scale DOF per element, while Scale-2 corresponds to the injection of two multi-scale DOF per element. This may be seen in Fig. 5 where the multi-scale DOF are shown relative to the elements (e_1, e_2) of the one-dimensional grid. At the finest resolution, the multi-scale solution corresponds to the solution associated with 8 linear finite elements. The multi-scale algorithm proceeds as follows.

Algorithm 1 *Multi-scale solution algorithm*

1. *Form the coarse-grid operator* K_0*, and solve the coarse-grid problem,* $K_0 U_0 = F_0$.
2. *For each element, inject one scale DOF and solve for the wavelet coefficient,* Δu.

$$\{\Delta u_k\} = [\mathbf{K}_k]^{-1}\{F_k\} \tag{53}$$

3. *Compute the termination measure for the scale DOF injection. One possibility for the termination measure relies on stopping when the scale DOF are small relative to the overall solution.* (*Other stopping criteria have not been investigated.*)

$$\varepsilon = \frac{\Delta u_k^{2+2^{k-1}+j}}{\|u_k\|} \tag{54}$$

4. *Repeat 1–3 until* ε *is smaller than some user-specified criteria.*
5. *Perform the element-by-element multi-scale reconstruction using Eq.* (48).

For the example problem, the multi-scale solution for $k = 1$ corresponds to $\Delta u_1^4 = \Delta u_1^5 = 1/32$, and for $k = 2$, the $\Delta u_2^6 = \Delta u_2^9 = 1/128$, and $\Delta u_2^7 = \Delta u_2^8 = 1/128$. After the reconstruction procedure, the scale DOF yield solution values that interpolate the exact solution – a result that is expected for linear problems. Similar results have been obtained for problems with non-linear source terms, and for problems with inhomogeneous essential and natural boundary conditions.

The algorithm presented for the 1-D multi-scale element possesses the property that all scale injection relies only on element-local data and does not require a re-solve of the coarse-grid problem to improve the solution. For problems that involve a significant mass-matrix contribution, the correction procedure outlined in the following sections is required when the row-column lumped mass is used.

The 2-D multi-scale element

Attention is now turned to the 2-D multi-scale element. As in the 1-D case, the 2-D bilinear element provides the element-level components of the global basis functions. To begin, Fig. 6 shows a four-patch of bilinear finite elements with the injected multi-scale DOF corresponding to $k = 1$. The configuration of the multi-scale DOF in the parent element is shown in Fig. 7. Like the shape functions, the components of the pseudo-wavelets at the element level take on a value of 1 at the DOF location, and they are zero at all other node locations.

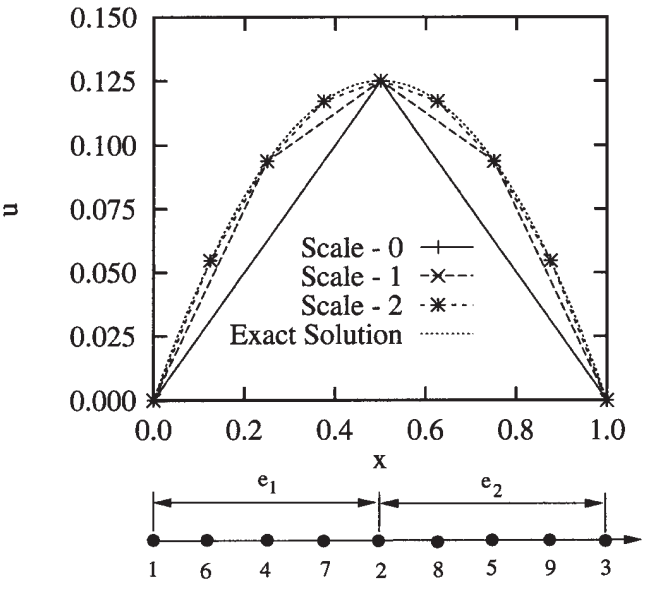


Fig. 5. Exact solution and scale solutions for $k = 0, 1, 2$ (Node numbers 4−8 correspond to the insertion of multi-scale DOF)

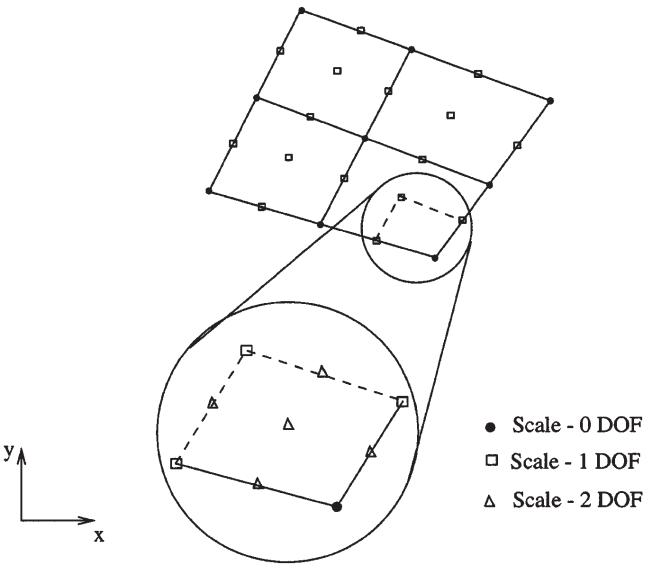


Fig. 6. Four-patch of bilinear elements with multi-scale DOF

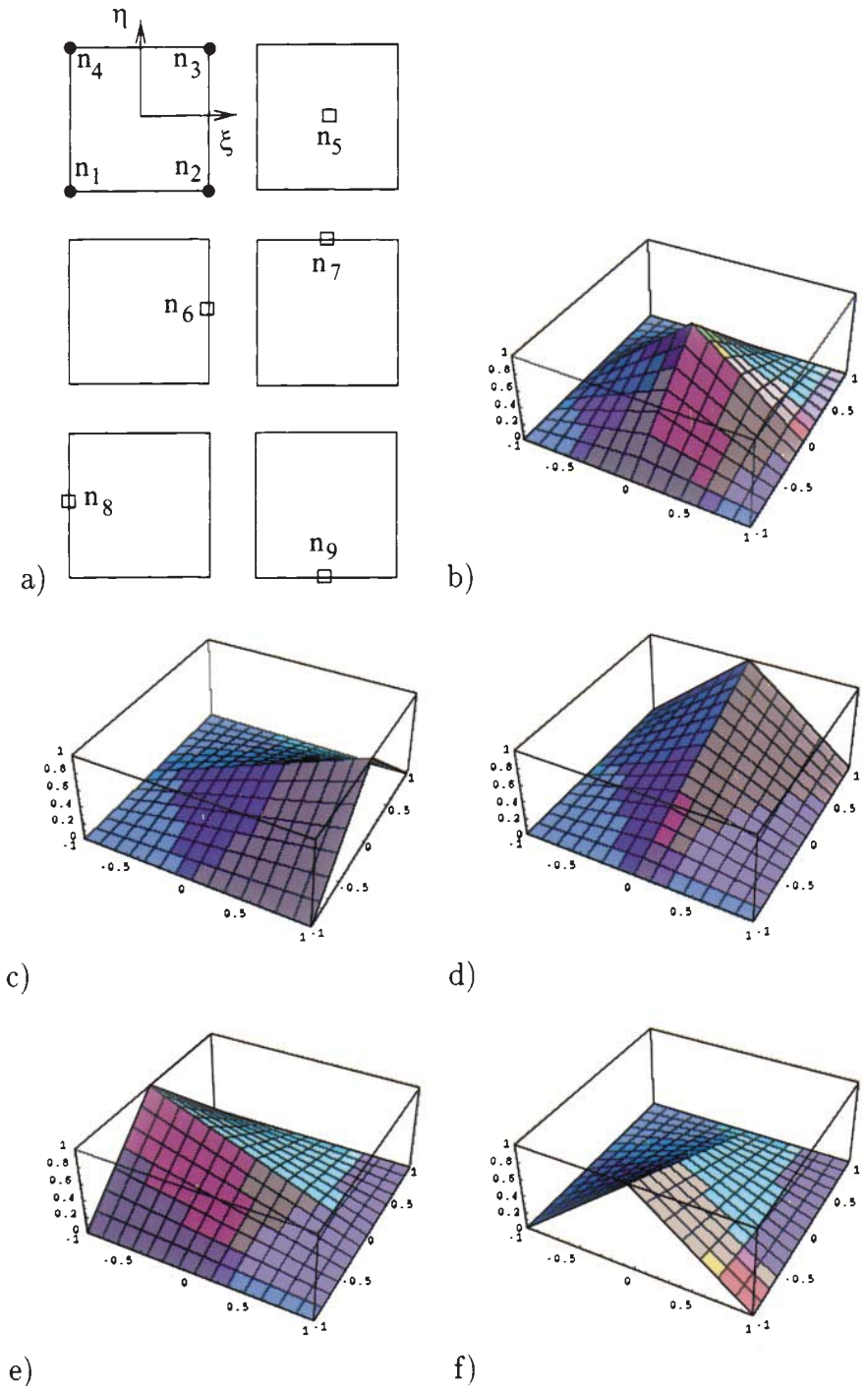


Fig. 7a–f. Parent element and multi-scale DOF for $k = 1$. **a** reference element and basis elements for multi-scale DOF with $k = 1$, **b** ψ_1^5, **c** ψ_1^6, **d** ψ_1^7, **e** ψ_1^8, and **f** ψ_1^9

The shape functions for the 2-D bilinear element are

$$N_i = \tfrac{1}{4}(1 + \xi_i\xi)(1 + \eta_i\eta) \ , \tag{55}$$

where $i = 1, 2, 3, 4$, and $-1 \le \xi, \eta \le 1$. At the first scale, $k = 1$, the psuedo-wavelets are

$$\begin{aligned} &\psi_1^5 = (1 - |\xi|)(1 - |\eta|), \ \psi_1^6 = \tfrac{1}{2}(1 + \xi)(1 - |\eta|), \\ &\psi_1^7 = \tfrac{1}{2}(1 - |\xi|)(1 + \eta), \ \ \psi_1^8 = \tfrac{1}{2}(1 - \xi)(1 - |\eta|), \\ &\psi_1^9 = \tfrac{1}{2}(1 - |\xi|)(1 - \eta) \ . \end{aligned} \tag{56}$$

In a more general way, the pseudo-wavelets for the multi-scale DOF may be written in terms of the translates and dilates of the basis functions at scale $k = 1$. The pseudo-wavelets in two-dimensions are

$$\psi_k^m(\xi, \eta) = \psi_1^m(\tilde{\xi}, \tilde{\eta}) \ , \tag{57}$$

where

$$\begin{aligned} \tilde{\xi} &= 2^{k-1}(1 + \xi) - 2j - 1 \\ \tilde{\eta} &= 2^{k-1}(1 + \eta) - 2j - 1 \ , \end{aligned} \tag{58}$$

and

$$
\begin{aligned}
&0 \le j \le k-1 \\
&2^{2-k}j - 1 \le \tilde{\xi} \le 2^{2-k}j + 2^{2-k} - 1 \\
&2^{2-k}j - 1 \le \tilde{\eta} \le 2^{2-k}j + 2^{2-k} - 1 \ .
\end{aligned} \tag{59}
$$

Here k indicates the scale, j indicates the translates in the element parametric space ($-1 \le \xi, \eta \le 1$), and $m = 5, 6, 7, 8, 9$ for the element-local numbering of the pseudo-wavelets. With the basis elements defined this way, the use of recursion at the element level as indicated in Fig. 6 can be used to automate the computation of the mass and stiffness operators with a given scale of resolution, Nscale.

There are several key points regarding the 2-D multi-scale element. First, the orthogonality of the derivatives of the pseudo-wavelets is not preserved in two dimensions – even on an orthogonal grid. The finger-diagonal matrices that arise from this discretization can lead to excessive storage costs if the matrices are used without threshholding, row-column lumping, or element-by-element procedures. Surprisingly, the row-column lumping can be applied to both the multi-scale mass and stiffness operators, albeit only for the rows and columns corresponding to the multi-scale DOF. The 2-D multi-scale element is compatible with many h-adaptivity strategies being implemented in finite element codes today, however, a detailed discussion of this aspect is beyond the scope of this paper. The ability to use this element as a change-of-basis preconditioner is just beginning to be explored. Finally, the numerical performance of the 1-D and 2-D multi-scale element in terms of dispersive behavior, rate of convergence, etc. is identical to the linear (bilinear) element since any multi-scale solution can be cast in terms of the reconstructed solution in the finite element basis at the finest grid scale.

4 Multi-scale vs. linear elements

In this section, a comparison of the storage, matrix conditioning, and computational complexity for the hierarchical Schauder and linear finite element bases is presented in the context of the model elliptic problem developed in Sect. 2. The approximations inherent in using a lumped multi-scale mass matrix are discussed, and an efficient row-column lumping procedure with a lumped-mass correction algorithm is presented to illustrate that only the largest wavelet coefficients need be corrected in the approximate row-column lumped-mass solution. The row-column lumping procedure is applied to both the multi-scale mass and stiffness matrices in 2-D, and the lumped-mass correction algorithm is extended to account for the approximations introduced by the row-column lumped mass and stiffness matrices. Although a uniform discretization is considered in the ensuing discussion, it should be noted that in 1-D the hierarchical Schauder basis retains all of its properties for a non-uniform grid. However, in two-dimensions, orthogonality in the $\mathscr{H}^1$ sense is lost even for uniform grids.

1-D comparison

The non-zero fill pattern for the mass and stiffness matrices are shown in Fig. 8 for both the hierarchical Schauder and the linear finite element bases. The formulae for the number of non-zeros associated with the mass and stiffness operators for both the hierarchical Schauder and finite element bases are shown in Table 1. Because the hierarchical Schauder basis diagonalizes the stiffness,

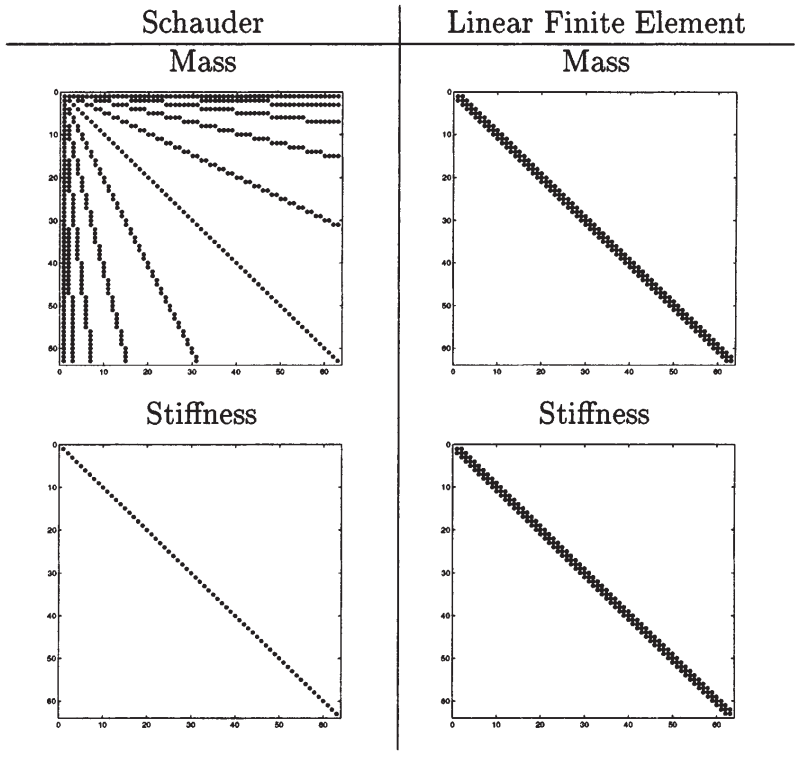


Fig. 8. The non-zero entries in the mass and stiffness matrices for the hierarchical Schauder and linear finite element bases for a mesh with 64 nodes

Table 1. Formulae for the number of non-zeros in the 1-D mass and stiffness matrices where Ndof is the number of degrees-of-freedom, Ndof $= 2^{k+1} - 1$, and k indicates the scale

Basis – matrix	No. of non zeros
Linear finite element – Mass matrix ($\mathbf{M}_k$)	3Ndof–2
Linear finite element – Stiffness matrix ($\mathbf{K}_k$)	3Ndof–2
Schauder – Mass matrix ($\mathbf{M}_k$)	(2k–1) (Ndof + 1) + 3
Schauder – Stiffness matrix ($\mathbf{K}_k$)	Ndof

only Ndof (number of DOF) words of storage are required for the diagonal, although in practice a coarse-grid finite element matrix is required for the multi-scale element. The consistent mass matrix, although sparse, requires increased storage relative to the tri-diagonal finite element mass matrix due to the finger-diagonal structure. In comparison to the linear finite element, the storage for the multi-scale mass matrix scales with both the number of DOF, Ndof, and the scale, k.

As demonstrated in Sect. 2, the finite element mass and the hierarchical Schauder stiffness matrices are both well conditioned. In fact, the hierarchical Schauder stiffness is ideal in 1-D since the nodal change-of-basis results in a diagonal stiffness operator when the full wavelet transformation is applied as a preconditioner. In contrast, the finite element stiffness and the Schauder mass matrices are both poorly conditioned. To be more precise, the finite element basis is uniformly stable in L^2, but not in $\mathscr{H}^1$. However, the hierarchical Schauder basis is uniformly stable in $\mathscr{H}^1$, but not in L^2.

In order to illustrate the differences in stability between the finite element and the hierarchical Schauder bases, consider the condition numbers associated with the mass and stiffness matrices as shown in Table 2. Here, multiple levels of mesh refinement are considered with $k = 1$ corresponding to a mesh with 3 nodes (2 elements). The growth of the condition numbers for $\mathbf{K}_k$ is seen to be proportional to $O(h^{-2})$ for the linear finite element bases, while the condition numbers for the mass matrix are bounded asymptotically at 3. The condition number associated with the combined finite element mass and stiffness ($\mathbf{M}_k + \mathbf{K}_k$) is dominated by the stiffness since $\varepsilon = \gamma = 1$ in this case. In contrast, the mass matrix for the hierarchical Schauder basis yields a condition number that grows approximately as $O(h^{-3/2})$ while the condition number for the stiffness is uniformly bounded at 1. The combined mass and stiffness operator ($\mathbf{M}_k + \mathbf{K}_k$) for the Schauder basis is also uniformly bounded indicating the dominance of the stiffness.

Although the multi-scale mass matrix associated with the hierarchical Schauder basis is not as well conditioned as the finite element mass matrix, its primary drawback is the storage associated with the finger-diagonal fill pattern. However, the entries in the finger diagonals decrease in amplitude with increasing scale suggesting that the relative importance of the matrix entries become negligible with increasing scale. This effect is reflected in the integrals for the multi-scale mass matrix in Eq. (52). Therefore, either a mass lumping procedure or threshholding based on the relative size of the entries can be used to limit the matrix fill-in. In order to test this idea, an ad-hoc procedure for mass lumping is considered. In a finite element setting, mass lumping refers to the row-sum lumping procedure used to obtain a diagonal mass matrix. However, in the context of a multi-scale basis, the physical interpretation of mass lumping is not simple, and the idea is perhaps less well founded for the multi-scale element than for the linear finite element.

Regardless of these issues, experimentation with mass lumping strategies has suggested that a viable lumping procedure for the multi-scale mass matrix is to sum the values to the left and above the diagonal entry to the diagonal as shown in Fig. 4. The use of the row-column lumping procedure in Algorithm 1 permits the point-wise evaluation of the multi-scale DOF for problems where the mass matrix is required, but the subsequent reconstruction procedure can only yield an approximate solution. However, inspection of the resulting approximate solution, Δu, obtained with the lumped multi-scale mass indicates that a small correction to the largest wavelet coefficients can dramatically improve the solution quality. In other words, the row-column lumping procedure cannot be used over the entire range of scales without degrading the solution accuracy, i.e., the influence of mass lumping is "felt" more at the coarser grid scales. A lumped-mass correction algorithm has been developed to account for the approximate nature of the solution and proceeds as follows.

Algorithm 2 *Lumped-mass correction algorithm*

1. *Row-column sum lump the multi-scale mass matrix resulting in a diagonal matrix,* $\tilde{\mathbf{A}} = \mathbf{M}_k^L + \mathbf{K}_k$. *Here, the superscript L indicates a lumped operator. For the multi-scale element, only those entries in the mass matrix associated with the multi-scale DOF are lumped.*

Table 2. Condition numbers for the 1-D mass and stiffness matrices after diagonal preconditioning for $\varepsilon = 1$, $\gamma = 1$. (Ndof $= 2^{k+1} - 1$)

Mesh parameters		Linear finite element			Schauder		
Scale (k)	Ndof	$\mathbf{M}_k$	$\mathbf{K}_k$	$\mathbf{M}_k + \mathbf{K}_k$	$\mathbf{M}_k$	$\mathbf{K}_k$	$\mathbf{M}_k + \mathbf{K}_k$
1	3	2.09	6	4	7	1	1.23
2	7	2.71	25	18	18	1	1.27
3	15	2.92	103	74	49	1	1.29
4	31	2.98	414	295	121	1	1.29
5	63	3.00	1659	1181	288	1	1.29
6	127	3.00	6640	4724	667	1	1.29
7	255	3.00	26560	18900	1517	1	1.29

2. *Calculate the approximate wavelet coefficients, $\Delta\tilde{u}$, by dividing the right-hand-side by the diagonal entries of* $\tilde{\mathbf{A}}$.
3. *Permute the components of $\Delta\tilde{u}$ so that the N largest coefficients are at the top. Let $\Delta\tilde{u}^N$ be the largest N lumped-approximate wavelet coefficients and $\Delta\tilde{u}^R$ be the remaining coefficients such that*

$$\mathrm{perm}(\Delta\tilde{u}) = \left\{ \begin{array}{c} \Delta\tilde{u}^N \\ \Delta\tilde{u}^R \end{array} \right\} \quad . \tag{60}$$

4. *Permute the rows and columns of* $\mathbf{A}$ *to match the permutation of $\Delta\tilde{u}$ using a consistent multi-scale mass only for the $\Delta\tilde{u}^N$ unknowns. Here, the permuted matrix, $\mathbf{A} = \mathbf{M}_k + \mathbf{K}_k$ is formed with the consistent mass matrix, $\mathbf{M}_k$.*
5. *Solve the smaller system using the consistent mass matrix for Δu^N, i.e., the coarse-grid correction,*

$$A^N \Delta u^N = f^N - A^R \Delta\tilde{u}^R \quad . \tag{61}$$

Here, A^N is the operator for the coarsest grid, and A^R corresponds to the rows of $\mathbf{A}$ *associated with $\Delta\tilde{u}^R$ but formed with the consistent mass matrix. An efficient means of computing the matrix-vector product, $A^R\Delta\tilde{u}^R$ is through the use of the wavelet transform in Eq.*(7). *Multiplication of Δu by the matrix* $\mathbf{W}$ *is equivalent to the application of the wavelet transform, i.e., the reconstruction algorithm, which can be implemented in an* $O(\mathrm{Nnp})$ *operation algorithm. It should be noted that multiplication by $\mathbf{W}^{\mathrm{T}}$ is not equivalent to the inverse wavelet transform, i.e., $\mathbf{W}^{\mathrm{T}}\mathbf{W} \neq I$. However, the properties of $\mathbf{W}^{\mathrm{T}}$ permit implementation of a matrix-vector product using convolution techniques similar to the reconstruction algorithm.*

In practice, the DOF associated with the coarse-grid tend to be the largest and are the ones that need to be corrected. Thus, the permutation step can be replaced with a solve for the coarse-grid DOF followed by the injection of the multi-scale DOF. This algorithmic approach is exactly what the multi-scale finite element does since the coarse-grid solution is computed with only the injection of scale relying on the hierarchical Schauder basis. Figure 9 illustrates the effectiveness of the lumped-mass correction algorithm for the model problem with $\varepsilon = 1$, $\gamma = 1$, and $f = -x^8(x-2)^3$. In Fig. 9a, the exact solution is shown with the approximate solution computed with the lumped multi-scale mass matrix. In Fig. 9b, the approximate solution has been improved by correcting only the three largest wavelet coefficients.

2-D comparison

In two dimensions, the behavior of the hierarchical Schauder basis changes rather significantly relative to the 1-D basis. The most apparent changes are the additional fill-in for the multi-scale mass matrix and the structure of the stiffness matrix which has a finger diagonal fill pattern rather than being diagonal. This is shown in Fig. 10 and is due to the multi-scale interaction of the tensor product basis in 2-D. The storage for the finger diagonal mass matrix grows as $4k^2\mathrm{Ndof}$ while the storage for the stiffness is proportional to $4k\mathrm{Ndof}$ where $\mathrm{Ndof} = (2^{k+1}-1)^2$. In comparison, the storage for the finite element mass and stiffness matrices scales linearly in the number of DOF as $\sim$9Ndof. As in the 1-D case, the amplitude of the entries in the finger-diagonals of the multi-scale mass matrix decrease with increasing scale.

In addition to the change in the fill pattern, the $O(1)$ condition number for the 1-D Schauder stiffness becomes $O(\log(h^{-2}))$ in 2-D. Again, the condition number for the finite element stiffness grows as $O(h^{-2})$ regardless of the dimensionality. Table 3 illustrates how the condition number varies with mesh resolution for the model problem in Eq. (1). Here, the L^2 stability of the finite element basis is reflected in the uniformly bounded condition number for the mass matrix and the $O(h^{-2})$ growth of the condition number for the stiffness. In contrast, the condition number for the multi-scale mass matrix grows faster than $O(h^{-2})$ although firm estimates have not been made. For the stiffness, the condition number for $\mathbf{K}_k$ grows as $\log(h^{-2})$. Although the condition number for the multi-scale mass matrix degrades rapidly, the condition number for $\mathbf{M}_k + \mathbf{K}_k$ is dominated by the stiffness and is much better behaved than the finite element basis as shown in Table 3.

Following the 1-D mass lumping, the concept of row-column lumping is applied in 2-D. However, here the lumping procedure is applied to both the multi-scale mass and the stiffness matrices. Although the row-sum lumping procedure is not valid for the finite element stiffness, the row-column lumping is possible for the multi-scale

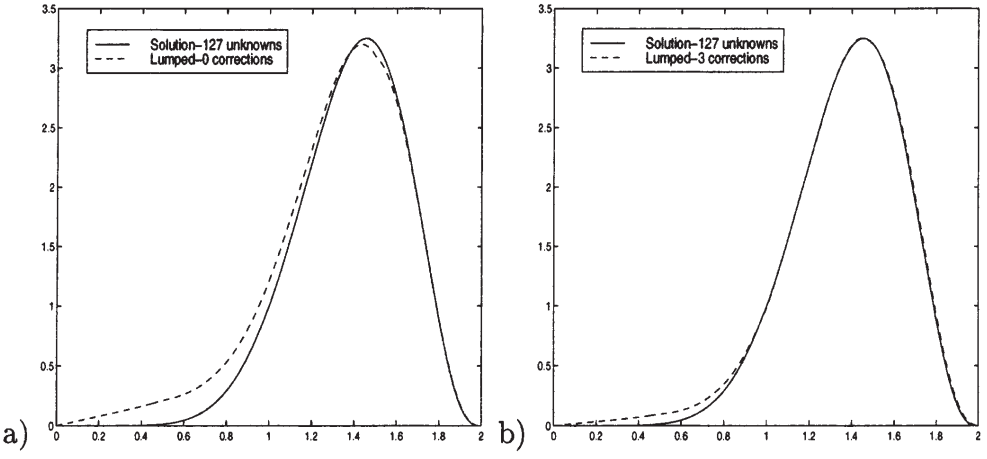


Fig. 9a, b. Solutions, u, on $0 \leq x \leq 2$ obtained using the hierarchical Schauder basis with mass-lumping and **a** no corrections and **b** 3 wavelet coefficient corrections

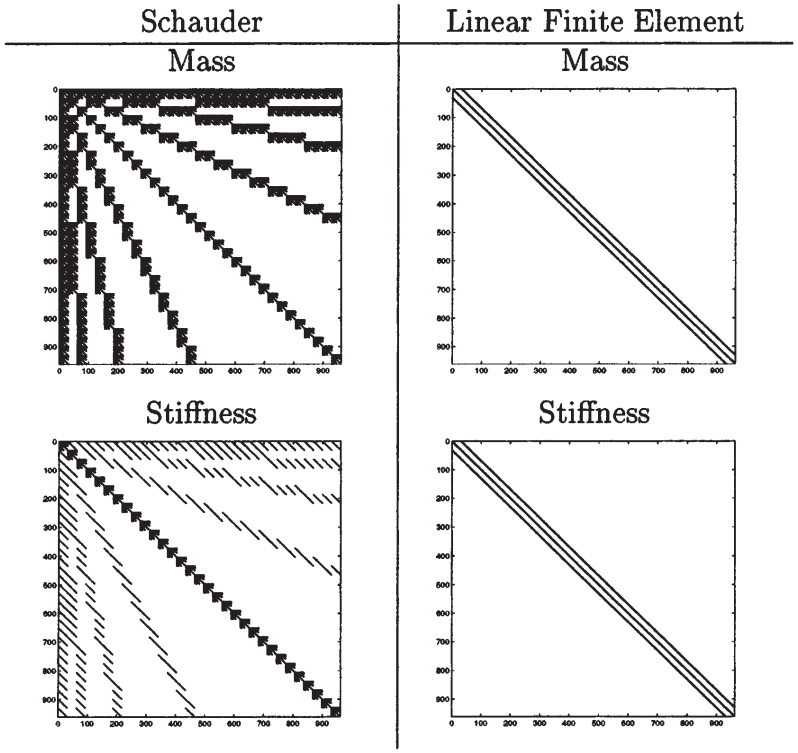


Fig. 10. The non-zero entries in the mass and stiffness matrices associated with the hierarchical Schauder and linear finite element bases in 2-D for a 32×32 mesh

Table 3. Condition numbers for the 2-D mass and stiffness matrices after diagonal scaling for $\varepsilon = 1$, $\gamma = 1$. (Ndof = $(2^{k+1}-1)^2$)

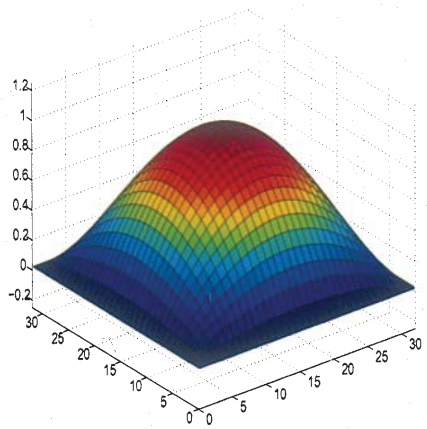
Mesh parameters		Linear finite element			Schauder basis		
Scale (k)	Ndof	$\mathbf{M}_k$	$\mathbf{K}_k$	$\mathbf{M}_k + \mathbf{K}_k$	$\mathbf{M}_k$	$\mathbf{K}_k$	$\mathbf{M}_k + \mathbf{K}_k$
1	9	4	3	3	49	7	8
2	49	7	13	11	334	18	19
3	225	9	52	43	2396	49	51
4	576	9	207	172	14621	121	123
5	1225	9	830	690	82832	288	290

representation of the stiffness – an artifact of the nodal change-of-basis. In other words, the multi-scale representation of the stiffness does not retain the "row-sum to zero" property of the nodal stiffness matrix. In order to illustrate this property, consider the application of Algorithm 2 in two-dimensions to the model problem with $\varepsilon = \gamma = 1$ and $f = xy(2-x)(2-y)$. However, in this case, the stiffness is replaced by the "lumped" stiffness, i.e., $\tilde{\mathbf{A}} = \mathbf{M}_k^L + \mathbf{K}_k^L$.

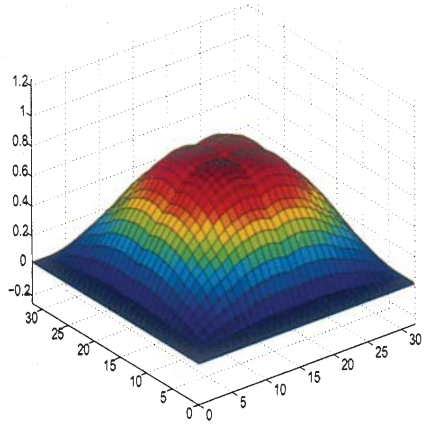
Figure 11b and c show the lumped-approximate solution with no wavelet correction and the associated error relative to the exact solution. The approximate solution was obtained by performing one vector divide and a multi-scale reconstruction resulting in a very reasonable initial solution. In Fig. 11c, the peak in the error is associated with the largest approximate wavelet coefficient. Figure 11d illustrates the lumped-approximation after the largest wavelet coefficient has been corrected following Algorithm 2. Nearly an order of magnitude reduction in the error has been achieved by correcting only the largest wavelet coefficient. In Fig. 11e, the peaks in the error are associated with the next 8 largest wavelet coefficients. Although the row-column lumping procedure yields an approximate solution, it is thought that this approach may generate a good initial solution for an iterative procedure or alternatively provide a multi-scale preconditioner.

Although the hierarchical Schauder basis is promising, the question of computational complexity remains. In 2-D, the change in fill pattern and matrix conditioning make computational complexity a more important issue in the comparison of the Schauder and linear finite element bases. In order to evaluate the computational complexity associated with the Schauder basis, a brief comparison of the cost of solving the model problem using a conjugate gradient method was performed. In this comparison, complete fill-in of the multi-scale mass and stiffness was accounted for ignoring the advantages of row-column lumping.

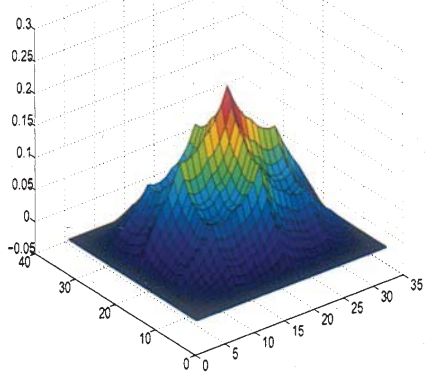
In order to perform the comparison, both the number of non-zero entries in the matrix, and the number of iterations required to solve the problem are required. Table 4 shows the number of non-zero entries and associated iteration count for the Schauder and linear finite element bases for the model problem with $0 \leq \varepsilon \leq 1000$ and $\gamma = 1$. For problems where the mass matrix dominates, $\varepsilon = 0$, the uniform stability of the finite element basis in L^2



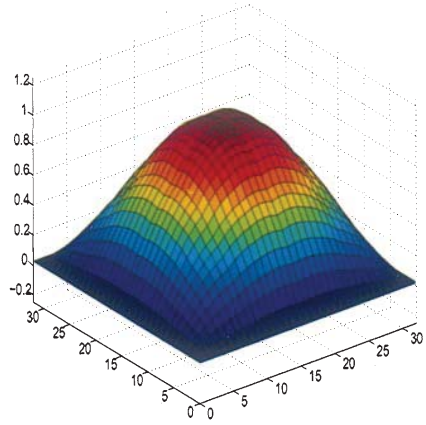
a) Exact Solution: u^e



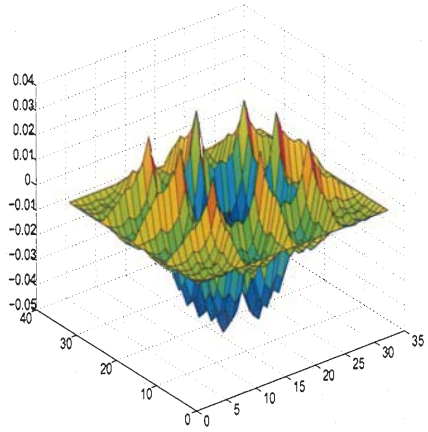
b) Lumped Approximation (u^h) without Correction



c) Error: $u^e - u^h$



d) Lumped Approximation (u^h) with 1 Correction



e) Error: $u^e - u^h$

Fig. 11a–e. The exact and approximate solutions using the lumping procedure with and without wavelet correction for a 32×32 2-D mesh (u^e is the interpolant of the exact solution, and u^h is the discrete solution)

results in a computational cost that is bounded independent of the mesh resolution. In contrast, the hierarchical Schauder basis results in a lower-order computational cost for problems where the stiffness dominates due to the multi-scale preconditioning of the stiffness – despite the additional fill-in associated with the finger-diagonal matrices. This result is further emphasized by the data in Table 5 for a "purely" elliptic problem with no mass matrix contribution.

In summary, Table 6 provides order of magnitude estimates for the computational complexity for the hierarchical Schauder and linear finite element bases. Here, the number of floating point operations required to solve the model problem grows as $9\text{Ndof}\sqrt{\text{Ndof}}$ for the finite

Table 4. Number of non-zeros and iteration count for the model problem using the **a** Schauder and **b** linear finite element bases in 2-D. The iteration count is based on using Jacobi preconditioned conjugate gradient with $0 \leq \varepsilon \leq 1000$, $\gamma = 1$

Mesh parameters		No. of non-zeros	Iteration count				
Scale (k)	Ndof		$\varepsilon = 0$	$\varepsilon = 0.1$	$\varepsilon = 1$	$\varepsilon = 10$	$\varepsilon = 1000$
a Schauder basis							
1	9	49	3	3	3	3	3
2	49	729	11	10	10	9	9
3	225	6889	50	27	25	22	22
4	961	51529	187	48	42	37	37
5	3969	335241	488	73	61	52	51
6	16129	1990921	979	107	87	70	59
b Linear finite element basis							
1	9	49	3	3	3	3	3
2	49	361	6	6	7	7	7
3	225	1849	7	13	15	16	16
4	961	8281	6	27	31	31	31
5	3969	34969	5	54	62	63	63
6	16129	143641	3	110	125	126	126

Table 5. Storage requirements, matrix fill, and iteration count using Jacobi preconditioned conjugate for the Schauder basis with the "purely" elliptic model problem with $\varepsilon = 1$, $\gamma = 0$

Mesh parameters		No. of non-zeros	Iteration count
Scale (k)	Ndof		
1	9	33	3
2	49	329	9
3	225	2265	22
4	961	13113	37
5	3969	68985	51
6	16129	342265	59

Table 6. Computational complexity for the linear finite element and the hierarchical Schauder basis in 2-D where Ndof = $(2^{k+1}-1)^2$, and k indicates the scale

Basis	No. of non-zeros	Iteration count	Floating point operations
Linear finite element $(\mathbf{M}_k + \mathbf{K}_k)$	9Ndof	$\sqrt{\text{Ndof}}$	$9\text{Ndof}\sqrt{\text{Ndof}}$
Schauder basis $(\mathbf{M}_k + \mathbf{K}_k)$	$4k^2$Ndof	$12k$	$48k^3$Ndof
Schauder basis $(\mathbf{K}_k)$	$4k$Ndof	$12k$	$48k^2$Ndof

element basis. In comparison, the number of floating point operations required to solve $\mathbf{M}_k + \mathbf{K}_k$ using the Schauder basis scales as $48k^3$Ndof and as $48k^2$Ndof for the purely elliptic problem. From this data, the Schauder basis does eventually have a lower computational cost, but the benefit is only apparent for relatively large problems. For the "purely" elliptic operator, $\mathbf{K}_k$, the hierarchical Schauder basis wins, i.e., has lower computational cost, for mesh resolution exceeding Ndof $= 10^5$. However, for $\mathbf{M}_k + \mathbf{K}_k$, the Schauder basis does not win until the mesh resolution exceeds Ndof $\geq 10^8$.

5 Summary and conclusions

In the search for a wavelet basis with good numerical and computational performance for multi-scale simulations, the following "shopping list" of characteristics was developed as the goal for the ideal wavelet basis. The attributes include: (1) Low order, e.g., linear, for computational efficiency, (2) Consistent reproduction of polynomials, e.g., reproduce $\{1, x, y, xy\}$ in two-dimensions, (3) Nodal, i.e., possesses the Kronecker delta property, (4) Hierarchical: $\mathscr{V}_1 = \mathscr{V}_0 \oplus \mathscr{W}_1$, (5) Element based – compatible with isoparametric finite elements, (6) Analytic expressions for the basis elements ϕ and ψ, (7) Easy treatment of boundary conditions, (8) Good numerical performance, e.g., dispersion characteristics, truncation error, etc., (9) Appropriate for both Eulerian and Lagrangian computations, (10) Computationally efficient decomposition and reconstruction of fields, and (11) Extensible to multiple spatial dimensions. Based on these characteristics and the results of this exploratory effort the following conclusions are drawn.

The hierarchical Schauder basis, and the multi-scale elements, are prototypical of what the ideal multi-scale basis should be although this basis does not satisfy all 11 characteristics of an ideal basis. Unfortunately, the storage and computational cost associated with the finger-diagonal matrices from this type of basis is a significant penalty. However, the use of ad-hoc lumping procedures ameliorates this problem and offers the potential for the development of fast, simple preconditioners. Currently, the real value of the multi-scale elements lies in the application to elliptic problems. A preconditioner based on the multi-scale element would be of great value in applications where a dominant elliptic component is present, e.g., in time-dependent incompressible flow and quasi-static electro-magnetics.

As demonstrated in the discussion of the hierarchical Schauder basis, it is difficult to construct a wavelet basis that is stable in both L^2 and in $\mathscr{H}^1$, i.e., for all possible combinations of mass and stiffness operators. The design of wavelet bases that are customized for a specific partial differential equation remains an open topic of active research, and the use of wavelet bases for the solution of partial differential equations remains a research topic that is currently centered in the mathematics community. However, hierarchical solution procedures that use wavelets tailored to the physical problem appear to be the most viable candidates for using wavelet bases.

References

[1] **Canuto C, Tobaco A, Urban K** (1998) The wavelet element method part ii: Realization and additional features in 2d and 3d, Tech. Rep. 98-17, Institute for Computer Applications in Science, NASA Langley Research Center, Hampton, VA April 1998

[2] **Christon MA, Baty RS, Burns SP, Roach DW, Trucano TG, Voth TE, Weatherby JR, Womble D** (1998) An investigation of wavelet bases for grid-based multi-scale simulations – Final Report, Tech. Rep. SAND 98-2456, Sandia National Laboratories, Albuquerque, New Mexico, September 1998

[3] **Dahmen W** (1997) Wavelet and multiscale methods for operator equations. Acta Numerica 55–228

[4] **Dahman W, Micchelli CA** (1995) Biorthogonal wavelet expansions, Tech. Rep. 114, Institute for Geometry and Applied Mathematics – RWTH Aachen, Aachen, Germany, May 1995

[5] **Grossmann A, Morlet J** (1984) Decomposition of hardy functions into square integrable wavelets of constant shape. SIAM Journal on Mathematical Analysis 15:723–736

[6] **Hughes TJR** (1987) The Finite Element Method, Prentice-Hall, Inc., Englewood Cliffs, New Jersey

[7] **Meyer Y** (1993) Wavelets Algorithms and Applications, SIAM, Philadelphia, Pennsylvania

[8] **Morlet J** (1983) NATO ASI Series, Issues in Acoustic Signal/Image Processing and Recognition, vol. 1, Springer, Berlin, 1983, ch. Sampling theory and wave propagation, pp. 233–261

[9] **Morlet J, Arens G, Fourgeau I, Giard D** (1982) Wave propagation and sampling theory, Geophys. 47:203–236

[10] **Yserentant H** (1986) On the multi-level splitting of finite element spaces. Numerishce Mathematik 49:379–412

[11] (1990) Two preconditioners based on the multi-level splitting of finite element spaces. Numerische Mathematik 58:163–184

[12] **Zienkiewicz OC** (1994) The Finite Element Method, McGraw-Hill, England, 4th ed., 1994, ch. 7, pp. 110–149

[13] **Zienkiewicz OC, Kelley DW, Gago J, Babuska I** (1982) The mathematics of finite elements and applications IV, Academic Press, London, England, 1982, ch. Hierarchical finite element approaches, error estimates and adaptive refinement