Radio Science, Volume 29, Number 4, Pages 965–977, July–August 1994

Utilization of wavelet concepts in finite elements for an efficient solution of Maxwell's equations

Tapan K. Sarkar and Raviraj S. Adve
Department of Electrical Engineering, Syracuse University, Syracuse, New York

Luis Emilio García-Castillo and Magdalena Salazar-Palma
Signals, Systems, and Radiocommunication Department, Escucla Tecnica Superior de Ingenieros Telecomunicación, Universidad Politécnica de Madrid, Ciudad Universitaria, Madrid, Spain

Abstract. The principles of dilation and shift are two important properties that are attributed to wavelets. It is shown that inclusion of such properties in the choice of a basis in Galerkin's method can lead to a slow growth of the condition number of the system matrix obtained from the discretization of the differential form of Maxwell's equations. It is shown that for one-dimensional problems the system matrix can be diagonalized. For two-dimensional problems, however, the system matrix can be made mostly diagonal. This paper illustrates the application of the new type of "dilated" basis for a Galerkin's method (or equivalent, for example, finite element method) for the efficient solution of waveguide problems. Typical numerical results are presented to illustrate the concepts.

Introduction

Differential forms of Maxwell's equations are generally solved utilizing the finite difference and the finite element method. These techniques transform the operator equation to a matrix equation and then a sparse matrix solver is used to solve the problem. However, one of the problems with these techniques is that as the dimension of the problem increases, the size of the matrix equation increases, and typically the condition number of the system matrix grows as $\theta(1/h^2)$ (where $\theta(1/h^2)$ denotes "of the order of $1/h^2$," where h is the discretization step). This is in contrast to the electric field integral equation utilized in the method of moments where the growth of the condition number of the system matrix is $\theta(1/h)$ and for the magnetic field integral equation the growth of the condition number can be independent of h. The above holds as long as the integral equations have a unique solution (that is, the problem is not solved at a frequency corresponding to an internal resonance of the closed structure) [*Peterson*, 1987].

The objective of this paper is to demonstrate that if the principle of dilation is introduced into the choice of basis functions in a finite element method then most of the system matrix can be made diagonal. If that is the case then the growth of condition number can be checked by proper scaling.

In section 2 the concept of wavelets are introduced and with it the principle of dilation. It is shown in section 3 that for the one-dimensional Laplace's equation the system matrix can be made exactly diagonal. This concept is then extended to two-dimensional problems, and in section 4 we demonstrate how such a basis can be chosen. Typical numerical results are presented in section 5.

Wavelets: A Cursory Preview

Wavelets have been studied extensively over the last two decades by both mathematicians and engineers resulting in some excellent documentation [*Daubechies*, 1992; *Chui*, 1992; *Vaidyanathan*, 1993]

Paper number 94RS00975.
0048-6604/94/94RS-00975$08.00

explaining the various mathematical subtleties and their properties. Two of the items (the integral wavelet transform and the wavelet series and their interconnections) have provided some interesting results. The properties of the wavelets are summarized next.

The integral wavelet transform of a square integrable function f(i.e., $f \in \mathfrak{L}^2$) is defined as

$$W_f[a,b] = \langle f;\ \psi_{a,b} \rangle \tag{1}$$

where $\langle \bullet;\ \bullet \rangle$ denotes the usual Hilbert inner product and where the doubly indexed family of wavelets $\psi_{a,b}(x)$ are generated from the basic wavelet ψ (often called the mother wavelet) by dilation and translation, that is,

$$\psi_{a,b}(x) = \frac{\psi(x-b)}{\sqrt{|a|}} \tag{2}$$

where a,b are real variables with $a \neq 0$. Substitution of (2) into (1) results in

$$W_f[a,b] = \int_{-\infty}^{\infty} dx\ f(x)\ \frac{\bar{\psi}[\frac{x-b}{a}]}{\sqrt{|a|}} \tag{3}$$

where the overbar denotes the complex conjugate.

The inverse wavelet transform recovers the function f from the values of $W_f[a,b]$. This is achieved by

$$f(x) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{da\,db}{a^2}\ W_f[a,b]\ \psi_{a,b}(x) \tag{4a}$$

where

$$C_\psi = 2\pi \int_{-\infty}^{\infty} d\xi\ \frac{|\hat{\psi}(\xi)|^2}{|\xi|} < \infty \tag{4b}$$

and $\hat{\psi}(\xi)$ is the Fourier transform of $\psi(x)$.

The convergence of the integral in (3) is defined in a weak sense [*Daubechies*, 1992], that is, taking the inner product of both sides of (4) with any function $g(x) \in \mathfrak{L}^2$ and commuting the inner product with the integral over a,b, in the right-hand side leads to the true formula,

$$\lim_{\substack{A_1 \to 0 \\ A_2, B \to \infty}} \left\| f - \frac{1}{C_\psi} \iint_{\substack{A_1 \le |a| \le A_2 \\ |b| \le B}} \frac{da\,db}{a^2}\ W_f[a,b]\ \psi_{a,b}(x) \right\| = 0 \tag{5}$$

where the double bar denotes the norm. Since for any absolutely integrable function ψ, $\hat{\psi}(\xi)$ (the Fourier transform of $\psi(x)$) is continuous, (5) can only be satisfied if

$$\hat{\psi}(0) = 0 \tag{6a}$$

or equivalently

$$\int_{-\infty}^{\infty} \psi(x)\ dx = 0 \tag{6b}$$

that is, the wavelet ψ does not have any "dc" value. The paradox now is that the wavelets have integral zero, so how can any superposition of them approximate f, which has a nonzero integral. The reason is that (4) holds in a $\mathfrak{L}^2$ sense and not in $\mathfrak{L}^1$ sense.

Next, we choose the parameters a,b as

$$b = k/2^j \tag{7a}$$

$$a = 1/2^j \tag{7b}$$

so that the integral wavelet transform is defined as

$$W_f[j,k] = \int_{-\infty}^{\infty} dx\ f(x)\ 2^{-j/2}\ \psi[2^{-j}x - k] \tag{8}$$

The shift integers k are chosen in such a way that $\psi[2^{-j}x - k]$ covers the whole x axis. The wavelet transform thus separates the "object" into different components in its transform domain and studies each component with a resolution matched to its scale.

The wavelet series amounts to expanding the function f in terms of wavelets $\psi_{j,k}(x)$, so that

$$f(x) = \sum_{j,k=-\infty}^{\infty} C_{j,k}\ \psi_{j,k}(x) \tag{9}$$

If we further assume that the wavelets $\psi_{j,k}(x)$ are orthogonal, then

$$C_{j,k} = \langle f; \psi_{j,k} \rangle \tag{10}$$

By comparing (8) and (10) it is apparent that the $(j,k)^{th}$ wavelet coefficient of f is given by the integral wavelet transform of f if the same orthogonal wavelets are used in both the integral wavelet transform and in the wavelet series. The problem now at hand is: are there any numerically stable algorithms to compute the wavelet coefficients $C_{j,k}$ in (10)? Specifically, in real life, f is not a given function but is sampled. Computing the integrals of $\langle f; \psi_{j,k} \rangle$ then requires some quadrature formula. For the smallest value of j, often referred to as the scale parameter, that is, most negative j, this will not involve many samples of f and one can do the computation quickly. For large scales, however, one faces large integrals, which might considerably slow down the computation of the wavelet transform of any given function. Especially for on-line implementations, one should avoid having to compute these long integrals. One way out is the technique used in multirate/multiresolution analysis, by introducing an auxiliary function $\phi(x)$, so that

$$\psi(x) = \sum_{m=-\infty}^{\infty} d_m \, \phi(x-m) \tag{11}$$

$$\phi(x) = \sum_{m=-\infty}^{\infty} c_m \, \phi(2x-m) \tag{12}$$

where in each case only finitely many coefficients c_m and d_m are different from zero. Here ϕ does not have integral zero but ψ does, and ϕ is normalized such that

$$\int_{-\infty}^{\infty} \phi(x)dx = 1 \tag{13}$$

and we define $\phi_{j,k}$ even though ϕ is not a wavelet, that is,

$$\phi_{j,k} = 2^{-j/2} \, \phi(2^{-j}x - k) \tag{14}$$

Since $\phi(x)$ satisfies an dilation equation in (12), it is called the scaling function. Hence

$$\langle f; \psi_{j,k} \rangle = \sum_{m=-\infty}^{\infty} \langle f; \phi_{j,k+m} \rangle \tag{15}$$

So the problem of finding the wavelet coefficient reduces to that of computing $< f ; \phi_{j,k} >$. Note

$$\langle f; \phi_{j,k} \rangle = \sum_{m=-\infty}^{\infty} c_m \langle f; \phi_{j-l;\, 2k+m} \rangle \tag{16}$$

in which case $< f; \phi_{j,k} >$ can be computed recursively starting from the smallest scale (most negative j) to the largest scale. Under certain conditions the advantage of this procedure is that it is numerically robust (because, even though the wavelet coefficients $C_{j,k}$ in (10) are computed with low precision, say with a couple of bits) one can still reproduce f with comparatively much higher precision [*Daubechies*, 1992, p. 98]. However, from practical considerations the limits in the sum can never be infinite but have to be truncated to a finite sum. Also, the wavelets have no dc value, so they cannot provide a good approximation for functions with nonzero mean. Both these situations can be avoided in the hybrid representation, where the function $f(x)$ is approximated by both the scaling function $\phi(x)$ and wavelets $\psi(x)$. The scaling functions provide the dc value, as they have nonzero integrals as per (13). The coefficients of the functions in the hybrid representation can also be computed efficiently by utilizing the discrete wavelet transform and terminating the infinite summation after a finite number of terms. The number of terms chosen depends on the number of samples of data provided. The details are available by *Chui* [1992], *Daubechies* [1992], and *Vaidyanathan* [1993].

The added advantage of the hybrid representation is that for continuous functions $f(t)$, they provide uniform convergence. However, like the Fourier techniques, the hybrid representation does display the Gibbs phenomenon at a discontinuity of the function $f(t)$. The Gibbs phenomenon occurs because it has been assumed that the wavelets chosen for the approximation problem are continuous. However, that need not be the case. One could in principle, and in practice, choose wavelets that are discontinuous: for example, the Haar wavelets (or equivalently, in engineering notation, the Walsh functions). For the case of a discontinuous wavelet basis the Gibbs phenomenon cannot occur! So whether the Gibbs

phenomenon manifests itself in the hybrid representation of discontinuous functions depends on the choice of the wavelet basis.

Another salient feature of the wavelets is that it provides localization of the result in both the original and in the transform domain. It is claimed that in the time-frequency characterization of a signal, a wavelet approximation provides better localization properties than the Fourier techniques simultaneously in time and in frequency. If one delves deeper into the subject, one observes that a function cannot simultaneously be limited in both frequency and time. Also, the degree of resolution achievable in both time and frequency is limited by the Heisenberg principle of uncertainty, that is,

$$\Delta t \, \Delta f \geq 1/2 \tag{17}$$

Here Δt and Δf are the resolutions achievable in time and frequency, respectively. Both the Fourier techniques and the wavelets are dictated by the principle of uncertainty. The only difference, between Fourier techniques and the wavelets is that, for the wavelets, the approximation is made of the function as a sum of functions which have nonoverlapping octaves bandwidth. For example, in a wavelet representation we take the spectrum of $f(t)$ and separate the spectrum into octaves of widths $\Delta\omega_j$ that is, the frequency band ω has been divided into $[2\pi$ to $2^{j+1}\pi]$ for all values of j, and now we define wavelets in each frequency bin $\Delta\omega_j$ and approximate $f(t)$ by it. If we choose

$$\phi(t) = \frac{\sin \pi t}{\pi t} \tag{18}$$

$$\psi(t) = 2\phi(2t) - \phi(t) \tag{19}$$

and define

$$\psi_{j,k}(t) = 2^{j/2} \, \psi(2^j t - k) \tag{20}$$

then the wavelet expansion of $f(t)$ with respect to ψ is

$$f(t) = \Sigma f_j(t) = \sum_{j,k} C_{j,k} \, \psi_{j,k}(t) \tag{21}$$

The functions $\psi_{j,k}(t)$ are orthonormal because their bandwidths are nonoverlapping: namely for a fixed j, $f_j(\omega)$ has bandwidth $\Delta\omega_j$ which is $[2^j\pi, 2^{j+1}\pi]$. So the wavelet expansion of a function is complete in the sense that it makes an approximation by orthogonal functions that have nonoverlapping bandwidth.

There are also some similarities between the wavelets and the Fourier series. Both of them use the principle of dilation. In the next section the principle of dilation used in the choice of basis functions will be discussed. Note that in the approximation of a function by a Fourier series results in decomposing a function into a sum of orthogonal functions (e^{jnt}) just as in (21).

$$f(t) \sim \sum_{n=-\infty}^{\infty} C_n \, e^{jnt} \tag{22a}$$

where

$$C_n = \frac{1}{2\pi} \int_0^{2\pi} f(t) \, e^{-jnt} \, dt \tag{22b}$$

Observe that the orthogonal functions into which $f(t)$ is decomposed in (22) is generated by integer dilations of a single function e^{jt}. In contrast, for the Fourier transform, the spectrum is decomposed into noninteger dilations of the function e^{jt}.

In the next section we show how a proper choice of basis functions can change the structure of the system matrix.

Solution of One-Dimensional Problems Utilizing the Wavelet Basis

In the solution of operator equations, particularly differential equations, the above concepts of dilation and shift in the choice of the hybrid basis functions (a combination of scaling functions and wavelets) could provide some computational advantages. As an example, consider the solution of the one-dimensional differential form of Maxwell's equations, that is,

$$\nabla^2 u(x) = F(x) \quad a \leq x \leq b \tag{23}$$

where ∇^2 is the Laplacian operator $= d^2/dx^2$ in one-dimension and u is the unknown to be solved for the given excitation F. The boundary conditions are left undefined at this point because it can be either Dirichlet (homogeneous, that is, $u(a$ or $b) = 0$, or inhomogeneous, $u(a) = A$ and $u(b) = B$) or Neumann type (homogeneous, that is, du/dx (at $x = a$ or $b) = 0$

or inhomogeneous $du/dx\ (x = a) = C$ and $du/dx\ (x = b) = D$).

The development is independent of the nature of the boundary conditions. However, the boundary conditions are needed for the complete solution of the problem.

Galerkin's method is now used to solve (23) which gives the fundamental equations of the finite element method. Hence consider the weighting function $v(x)$, which multiplies both sides of (23) and the product is integrated by parts from a to b to yield

$$-\int_a^b \frac{du}{dx}\frac{dv}{dx}\,dx + \frac{du(x)}{dx}\Big|_{x=a} v(a) - \frac{du(x)}{dx}\Big|_{x=b} v(b) = \int_a^b v(x)\,F(x)\,dx \tag{24}$$

Next it is assumed that the unknown $u(x)$ can be represented by a complete set of basis functions $\phi_i(x)$, which has first-order differentability. Then

$$u(x) \sim u_N(x) = \sum_{i=1}^{N} a_i\phi_i(x) + a_{01}\phi_{01}(x) + a_{02}\phi_{02}(x) \tag{25}$$

where a_i and a_{oj} are the unknowns to be solved for. Basically, the functions $\phi_i(x)$ satisfy the homogeneous boundary conditions and ϕ_{oj} take care of the inhomogeneous Dirichlet conditions.

In a Galerkin's procedure the weighting functions $v(x)$ are of the form

$$v(x) = \phi_j(x)\ ;\ \phi_{01}(x)\ ;\ \phi_{02}(x) \tag{26}$$

Substitution of (25) and (26) into (24) results in a system of equations which can be written in the following matrix form:

$$\begin{bmatrix} \langle\phi_1';\phi_1'\rangle & \langle\phi_1';\phi_n'\rangle & \langle\phi_1';\phi_{01}'\rangle & \langle\phi_1';\phi_{02}'\rangle \\ \langle\phi_2';;\phi_1'\rangle & \langle\phi_2';\phi_n'\rangle & \vdots & \\ \vdots & & \vdots & \vdots \\ \langle\phi_n';\phi_1'\rangle & \langle\phi_n';\phi_n'\rangle & \langle\phi_n';\phi_{01}'\rangle & \langle\phi_n';\phi_{02}'\rangle \\ \hline \langle\phi_{01}';\phi_1'\rangle & \langle\phi_{01}';\phi_n'\rangle & \langle\phi_{01}';\phi_{01}'\rangle & \langle\phi_{01}';\phi_{02}'\rangle \\ \langle\phi_{02}';\phi_1'\rangle & \langle\phi_{02}';\phi_n'\rangle & \langle\phi_{02}';\phi_{01}'\rangle & \langle\phi_{02}';\phi_{02}'\rangle \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \\ \hline a_{01} \\ a_{02} \end{bmatrix} = - \begin{bmatrix} \langle F;\phi_1'\rangle \\ \langle F;\phi_2'\rangle \\ \vdots \\ \langle F;\phi_n'\rangle \\ \hline \langle F;\phi_{01}'\rangle \\ \langle F;\phi_{02}'\rangle \end{bmatrix} - \begin{bmatrix} 0 \\ \hline - \dfrac{du}{\partial x}\Big|_{x=a} \\ \hline \dfrac{du}{\partial x}\Big|_{x=b} \end{bmatrix} \tag{27a}$$

or equivalently

$$\bar{\bar{Z}}A = Y \tag{27b}$$

where the superscript $'$ denotes the first derivative of the function and $\langle c,d\rangle$ denotes the classical Hilbert inner product, that is,

$$\langle c,d\rangle = \int_a^{\ell} c(x)\,\bar{d}(x)\,dx \tag{28}$$

Here the overbar denotes complex conjugate.

The solution of (27) then provides the unknowns a_i and a_{01} and a_{02}. The crux of the problem therefore lies in the solution of large matrix equations. The stability of the solution of large systems is dictated by the condition number of the matrix and by the number of effective bits t with which the solution is carried out on the computer.

Specifically, in the solution of (27) if $\Delta\bar{\bar{Z}}$ is the Y is ΔY, then the error in the solution ΔA is bounded by [*Sarkar et. al*, 1981]

$$\frac{\|\Delta A\|}{\|A\|} \le \frac{cond(\bar{\bar{Z}})}{1 - \sqrt{N}\,cond(\bar{\bar{Z}})2^{-t}} \left[\frac{\|\Delta Y\|}{\|Y\|} + \frac{\|\Delta\bar{\bar{Z}}\|}{\|\bar{\bar{Z}}\|}\right] \tag{29}$$

where N is the dimension of the matrix $\bar{\bar{Z}}$ and the norm is defined as the Euclidian norm. It is therefore clear that the choice of the basis functions, which determines the condition number of the matrix $\bar{\bar{Z}}$, has a tremendous influence on the efficiency and accuracy on the solution of (27).

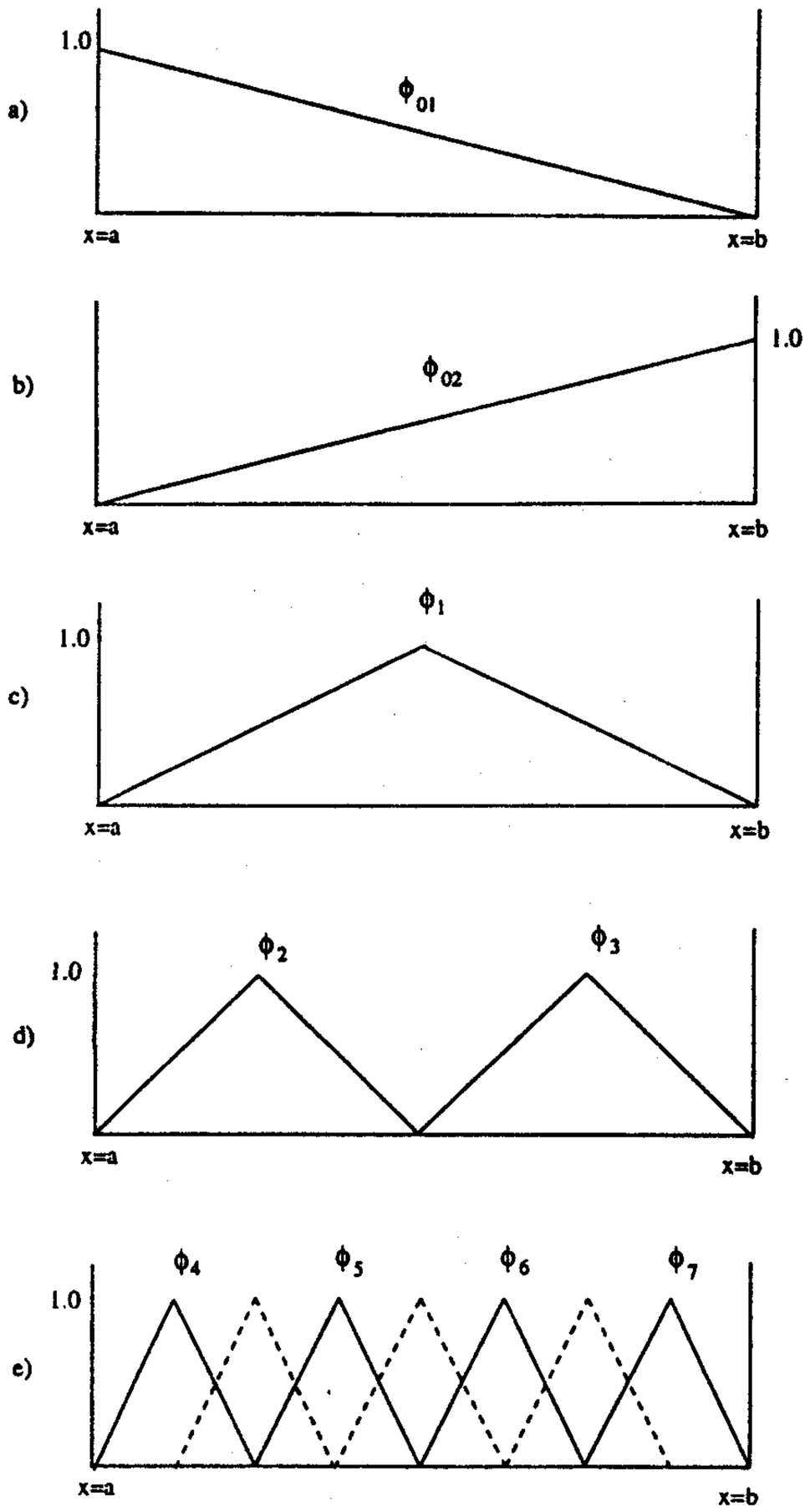

Figure 1. The basis functions: (a)ϕ_{01}; (b)ϕ_{02}; (c)ϕ_1; (d)ϕ_2; ϕ_3; and (e) subsectional basis seven piecewise triangle functions.

The problem with the finite element method lies in the solution of large matrix equation. Also, as the number of basis functions increases, the condition number of the matrix also increases. An increase in the condition number of the matrix creates various types of solution problems. For example, the condition number directly dictates the solution procedure, as a highly ill-conditioned matrix prohibits application of a direct matrix solver like Gaussian elimination [*Golub and VanLoan*, 1989] and a more sophisticated technique like singular value decomposition may have to be introduced. There are various ways to eliminate the increase of the condition number as the dimension of the matrix increases. *Mikhlin* [1971] and *Krasnoselskii* [1972] choose the basis functions such that the growth of the condition number is controlled.

A good way to choose the basis functions is shown in Figure 1. It is interesting to point out that these basis functions are similar to the classical triangular functions used by *Harrington* [1967] in the method of moments. However, unlike the method of moments, these basis functions are not the subdomain basis functions. In the classical subdomain basis functions the choice would be the seven piecewise triangle functions as shown in Figure 1e. The seven basis functions would consist of the four solid line triangular functions and in addition the three dotted line triangular functions shown in the same figure.

In the new basis, which we call the hybrid wavelet basis, we have the seven basis functions shown in Figures 1a-1e marked by ϕ_{01}, ϕ_{02}, and $\phi_{1\text{-}7}$. The difference is that instead of the three dotted triangular basis functions we have three different nested basis functions ϕ_1, ϕ_2, and ϕ_3, which in the finite element literature are called hierarchical basis functions. The functions ϕ_{01} and ϕ_{02} are there to treat arbitrary boundary conditions. The basis functions shown in Figures 1c-1e are termed the "wavelet" basis as they are the dilated and shifted version of the same function [R. A. H. Lorentz and W. R. Madych, Wavelets and generalized box splines, Unpublished manuscript, 1994]. These basis functions are derived from the Battle-Lamarie type of wavelets. Here the basis are chosen as the ϕ functions and not the ψ functions. The natural question that arises is, what is the advantage of this type of the wavelet basis over the conventional subsectional basis functions? The disadvantage of the wavelet basis is clear, for example, for ϕ_1, ϕ_2, and ϕ_3 more calculations need to be carried out over the domain of interest as opposed to the three dotted triangular basis functions shown for the classical subsectional basis. However, as a final solution, both the subsectional and the wavelet type basis provide the same information content about the approximation.

In spite of the additional computation the reason for the choice of the wavelet basis is that as the dimension of the problem increases, the condition number of the solution matrix does not go up as fast for the wavelet basis. This has been rigorously shown by *Jaffard* [1992]. There is another computational advantage which we will describe later.

In summary, the distinct feature of this present approach is a different choice of basis functions for the unknown. This choice is different from the classical subsection basis. For the approach presented in this paper the "granularity" of the basis function is different for different orders of approximation, some entire domain and some subdomain, so that the span of the basis function is complete. We now proceed with the solution procedure to describe the other salient features. Under the new basis functions we have

$$\langle \phi_i' ; \phi_j' \rangle = 0 \qquad i \neq j \tag{30}$$

$$\langle \phi_i' ; \phi_{01}' \rangle = 0 \tag{31}$$

$$\langle \phi_i' ; \phi_{02}' \rangle = 0 \tag{32}$$

Because of the proper choice of the wavelet basis, we have (30)-(32) and therefore (27) reduces to

$$\begin{bmatrix} \langle \phi_1' ; \phi_1' \rangle & & & 0 & & 0 \\ & \ddots & & & & \\ 0 & & \langle \phi_n' ; \phi_n' \rangle & & & \\ \hline & & & 0 & \langle \phi_{01} ; \phi_{01}' \rangle & \langle \phi_{01}' ; \phi_{02}' \rangle \\ & & & & \langle \phi_{02} ; \phi_{01}' \rangle & \langle \phi_{02} ; \phi_{01}' \rangle \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \\ \hline a_{01} \\ a_{02} \end{bmatrix} = - \begin{bmatrix} \langle F ; \phi_1 \rangle \\ \langle F ; \phi_2 \rangle \\ \vdots \\ \langle F ; \phi_n \rangle \\ \hline \langle F ; \phi_{01} \rangle \\ \langle F ; \phi_{02} \rangle \end{bmatrix} + \begin{bmatrix} 0 \\ \hline - \dfrac{\partial u}{\partial x} \Big|_{x=a} \\ \dfrac{\partial u}{\partial x} \Big|_{x=b} \end{bmatrix} \tag{33}$$

In the solution of (33) the boundary conditions of the problem are implicitly provided. For example, if the boundary conditions are purely Dirichlet type then a_{01} are known and for the Neumann condition the right-hand sides are known and a_{0i} values are to be solved for. For mixed boundary conditions a combination of the above are required and a_{0j} need to be solved for the jth boundary which has a specified Neumann condition. So, from (33) we have

$$a_i = - \frac{\langle F ; \phi_i \rangle}{\langle \phi_i' ; \phi_i' \rangle} \tag{34}$$

$$\begin{bmatrix} \langle \phi_{01}' ; \phi_{01} \rangle & \langle \phi_{01}' ; \phi_{02} \rangle \\ \langle \phi_{02}' ; \phi_{01}' \rangle & \langle \phi_{02}' ; \phi_{01} \rangle \end{bmatrix} \begin{bmatrix} a_{01} \\ a_{02} \end{bmatrix} = - \begin{bmatrix} \langle F ; \phi_{01} \rangle \\ \langle F ; \phi_{02} \rangle \end{bmatrix} + \begin{bmatrix} - \dfrac{du}{dx} \Big|_{x=a} \\ - \dfrac{\partial u}{\partial x} \Big|_{x=b} \end{bmatrix} \tag{35}$$

where two of the four parameters a_{01}; a_{02}; $\partial u/\partial x$ $(x=a)$ and $\partial u/\partial x$ $(x=b)$ have been fixed by the boundary conditions (Dirichlet, Neumann, or mixed) of the problem.

The application of the wavelet basis is now clear. For one-dimensional problems the system matrix can be made almost diagonal, and hence its solution is trivial.

In summary, if the dilation and shift principles of wavelets are utilized in the choice of basis functions such that their first derivatives are orthogonal to each other then the system matrix can be diagonalized corresponding to the unknowns. So the choice of a wavelet type basis makes the system matrix almost diagonal, simplifying the computational complexity.

We now extend this procedure to the solution of two-dimensional problems and apply it to some waveguide problems.

Solution of $\nabla^2 u + k^2 u = F$ for Two-Dimensional Problems Utilizing the Wavelet Basis

Consider the solution of

$$\nabla^2 u(x,y) + k^2 u(x,y) = F(x,y) \tag{36}$$

restricted to the region R defined by the contour Γ. In this paper we have focused our attention only to

rectangular regions and have assumed that any arbitrary shaped region $\Re$ can be made of rectangular regions R of the type:

$$R: 0 \le x \le a \quad 0 \le y \le b \tag{37}$$

To apply Galerkin's method, one integrates (36) in the domain R to obtain

$$\int_R v(x,y)\ \nabla^2 u(x,y)\ dxdy + k^2 \int_R u(x,y)\, v(x,y)\, dxdy = \int_R F(x,y)\, v(x,y)\ dxdy \tag{38}$$

which, after integration by parts, yields

$$-\int_R (\nabla u)(\nabla v)dr + \int_\Gamma v \frac{du}{dn} d\Gamma + k^2 \int_R uvdr = \int_R Fvdr \tag{39}$$

Generation of a wavelet type basis in two dimensions can be done by utilizing the multiresolution analysis and essentially following the one-dimensional construction. Another method consists in obtaining the wavelets using the one-dimensional reconstruction.

In this paper we follow *Jaffard* [1992] in the development of a two-dimensional wavelet type basis, where it is shown that three wavelets (ψ^1, ψ^2, and ψ^3) are required in two dimensions, and they are generated from

$$\psi^1(x,y) = \psi(x)\ \ \phi(y) \tag{40a}$$

$$\psi^2(x,y) = \phi(x)\ \ \psi(y) \tag{40b}$$

$$\psi^3(x,y) = \psi(x)\ \ \psi(y) \tag{40c}$$

where the significance of the ϕ and ψ functions are explained in section 2.

Let the basis chosen for this case, be

$$u(x,y) \sim \hat{u} = \sum_{i=1}^{M}\sum_{j=1}^{N} A_{ij}\phi_{ij}(x,y) + \sum_{i=1}^{4}\sum_{j=1}^{K} B_{ij}N_{ij}(x,y) + \sum_{i=1}^{4} C_i T_i(x,y) \tag{41}$$

and A_{ij}, B_{ij}, and C_i are the unknown constants to be solved for. In this expansion, $\phi_{ij}(x,y)$ are zero on the rectangular boundary as before, and they are explicitly chosen as

$$\phi_{ij}(x,y) = \eta_i(x)\,\eta_j(y) = \sin\left(\frac{\pi i x}{a}\right)\sin\left(\frac{\pi j y}{b}\right) \tag{42}$$

These functions are not only orthogonal to themselves, but their partial derivatives are also orthogonal; that is,

$$\langle \phi_{ij}; \phi_{pq} \rangle = 0 \quad for \quad i \neq p; j \neq q \tag{43}$$

$$\langle \nabla\phi_{ij}; \nabla\phi_{pq} \rangle = 0 \quad for \quad i \neq p; j \neq q \tag{44}$$

$$\int_\Gamma \phi_{ij}\ \frac{\partial}{\partial n}\ \phi_{pq}\ d\Gamma = 0 \tag{45}$$

where the inner product in the two-dimensional rectangular region is defined as

$$\langle c; d \rangle \equiv \int_0^a dx \int_0^b dy\ c(x,y)\ \bar{d}(x,y) \tag{46}$$

In addition we need four edge basis functions N_{ij} where, for example, N_{ij} is zero everywhere on the boundary (i.e., on all edges) except on edge E1 (refer to Figure 2). The two-dimensional basis has the representation

$$N_{1j}(x,y) = G(y)\,\eta_j(x) \tag{47a}$$

$$N_{2j}(x,y) = H(x)\,\eta_j(y) \tag{47b}$$

$$N_{3j}(x,y) = H(y)\,\eta_j(x) \tag{47c}$$

$$N_{4j}(x,y) = G(x)\,\eta_j(y) \tag{47d}$$

where the polynomials $G(x)$, $G(y)$, $H(x)$ and $H(y)$, are chosen based on the differentiability conditions. In this case,

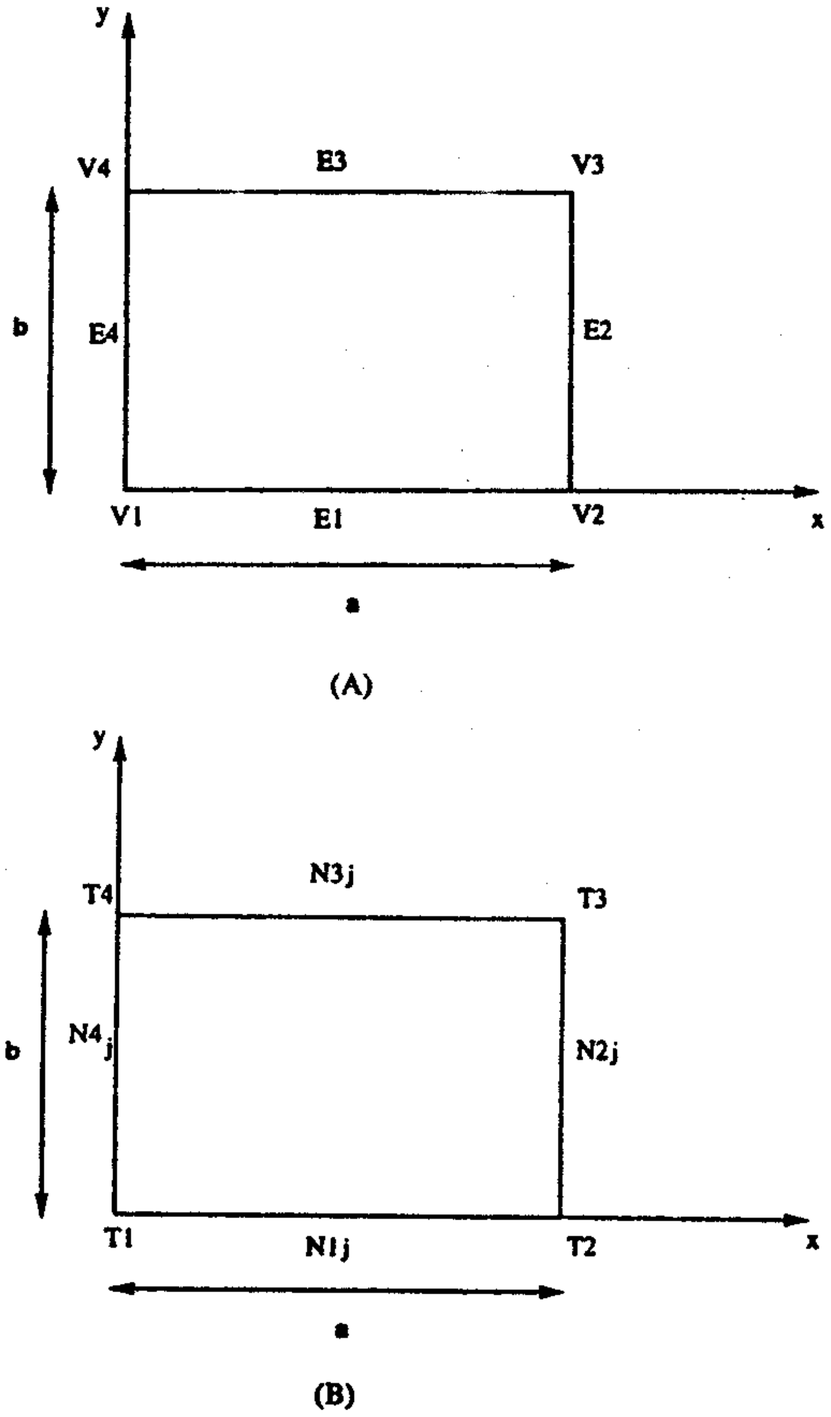

Figure 2. Geometry for the two-dimensional basis functions: (a) edges and nodes; (b) basis functions corresponding to the edges and nodes.

$$G(x) = 1 - \frac{x}{a} \tag{48a}$$

$$G(y) = 1 - \frac{y}{b} \tag{48b}$$

$$H(x) = \frac{x}{a} \tag{48c}$$

$$H(y) = \frac{y}{b} \tag{48d}$$

Therefore N_{2j} only participates in providing the match corresponding to edge E2 only.

In an analogous fashion, one can illustrate that the wavelet basis T_i in (41) provides the matching conditions needed for the four vertices V_1, V_2 V_3 and V_4 as shown in Figure 2. Specifically, the wavelet basis associated with the four vertices can be written as

$$T_1(x,y) = G(x)\ G(y) \tag{49a}$$

$$T_2(x,y) = H(x)\ G(y) \tag{49b}$$

$$T_3(x,y) = H(x)\ H(y) \tag{49c}$$

$$T_4(x,y) = G(x)\ H(y) \tag{49d}$$

Substitution of (42), (47), (48), and (49) in (41) and utilizing ϕ_{ij}; N_{ij}, and T_i functions, as the weighting function results into a system matrix

$$[P][A] + k^2[Q][A] = [V_F] + [V_B] \tag{50}$$

where the system matrix $[P]$ and $[Q]$ has the form

$$[P];[Q] = \begin{bmatrix} [D]_{n\times n} & [G]^{*}_{n\times m} \\ [G]_{m\times n} & [B]_{m\times m} \end{bmatrix} \tag{51}$$

where $[D]$ is a diagonal matrix and $[G]$ and $[B]$ are sparse matrices. The system matrices again would be mostly diagonal. What percentage of the matrix is diagonal depends on how many rectangular regions the original region has been divided into.

Case A: Dirichlet

If the boundary condition over Γ is purely Dirichlet then the maximum dimension of the system matrix will be $L(N^2+4N+4)$ where the original domain has been subdivided into L regions and the highest order of approximation M, N and K in (41) are assumed to be the same, all N, that is, they are considered to be the same in all L regions for comparison purposes.

Because of the choice of the wavelet type basis, out of the maximum dimension of $L(N^2+4N+4)$, the rank of the diagonal submatrix $[D]$ in (51) will be LN^2. This clearly demonstrates that as the number of unknowns N increase majority of the system matrix becomes diagonal. This is because the row size increase of $[P]$ is dominated by the term LN^2 and so is the row size of the diagonal matrix $[D]$. The

rectangular submatrix [G] has the row dimension as N^* (the number of internal edges + number of internal corners) and the column dimension is LN^2. The square matrix [B] has a row and column dimension of N^* (the number of internal edges + number of internal corners). Hence the size of B goes up as essentially $\theta[(L+1)N]$. Therefore the computational complexity goes up as $\theta[\{(L+1)N\}^3]$ when the number of unknowns go up by LN^2. This amounts to a significant decrease in reduction of computational complexity.

Case B: Neumann

For this case the diagonal submatrix is of the same size as that for the previous case of Dirichlet boundary conditions. However, now the coefficients of all the matching functions are unknowns. Hence the size of the system matrix [P] is LN^2+N^* {(number of edges) + (number of corners)}. The size of the diagonal matrix [D] is the same as before, that is, LN^2.

Case C: Mixed

It is easy to extrapolate the results to a mixed Dirichlet and Neumann conditions. The important point is that due to the choice of the wavelet basis, the major portion of the system matrix [P] and [Q] are diagonal.

Application To Some Waveguide Problems

As an example, consider the solution of cutoff frequencies of the transverse electric (TE) and transverse magnetic (TM) modes of various conducting waveguides structures. So, in this case the objective is to solve an eigenvalue problem, as F=0.

Therefore we are solving for

$$\nabla^2 u + k_c^2 u = 0 \tag{52}$$

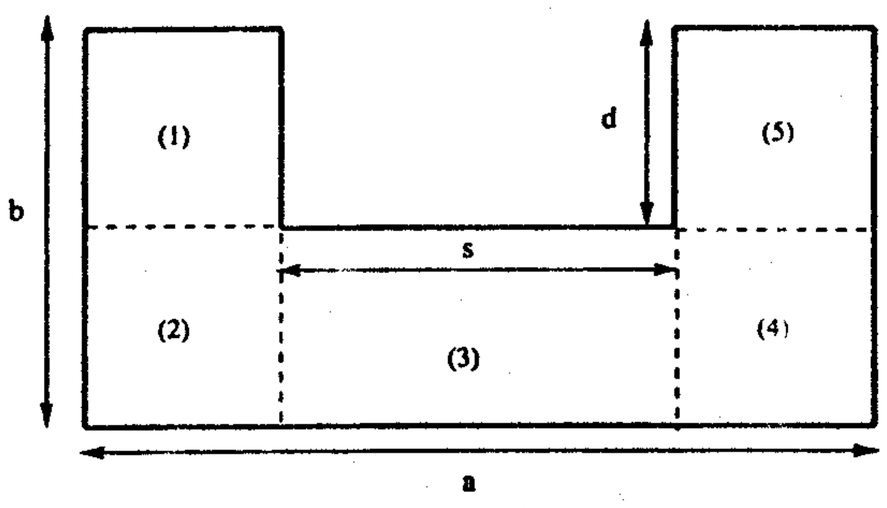

Figure 3. Single ridge waveguide (a = 1.0 m; b = 0.5 m; d = 0.25 m; s = 0.5 m).

Table 1. Cut-off Wave Numbers of the Transverse. Magnetic (TM) Modes of a Single Ridge Waveguide

	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5
N=2	12.20	12.5	14.02	15.62	16.67
N=4	12.16	12.45	14.02	15.6	16.66
N=6	12.15	12.44	14.01	15.6	16.66
N=8	12.14	12.43	14.01	15.59	16.65
N=10	12.05	12.43	14.01	15.59	16.65
Sarkar et. al. [1989]	12.05	12.32	13.86	15.34	16.28
Swaminathan et. al. [1990]	12.04	12.29	14.00	15.99	...

and the system matrix equation is

$$[P][A] + k_c^2[Q][A] = 0 \tag{53}$$

The objective is to solve this generalized eigenvalue problem for the eigenvalue k_c^2 and the eigenvectors [A]. Since [P] and [Q] are sparse mostly diagonal matrices, the computational complexity has been greatly reduced and the conjugate gradient method [*Chen et. al.*, 1986] can be used efficiently to find a few of the generalized eigenvalues k_c^2 and the eigenvector associated with it.

For the first example, consider the solution of the TM modes of a single ridge waveguide of dimensions shown in Figure 3. The single ridge waveguide was divided into five rectangular regions as shown by the dotted lines. Table 1 provides the cutoff wave numbers of the TM modes of the single ridge waveguide. Table 2 shows the percentage of the system

Table 2. Percentage of the Matrix That is Diagonal for the TM Modes.

	Matrix Size	Row Size of the Diagonal Block	Percent Diagonal
N=2	28	20	71.43
N=4	96	80	83.33
N=6	204	180	88.24
N=8	312	320	90.91
N=10	540	500	92.59

Table 3. Cut-off Wave Numbers of the Transverse. Electric (TE) Modes of a Single Ridge Waveguide

	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5
N=2	2.26	4.90	6.48	7.53	9.85
N=4	2.25	4.87	6.46	7.52	9.83
N=6	2.25	4.87	6.46	7.52	9.83
N=8	2.25	4.87	6.46	7.72	9.83
N=10	2.25	4.86	6.46	7.52	9.83
Sarkar et.al. [1989]	2.23	4.78	6.40	7.48	9.71
Swaminathan et. al. [1990]	2.25	4.94	6.52	7.56	...

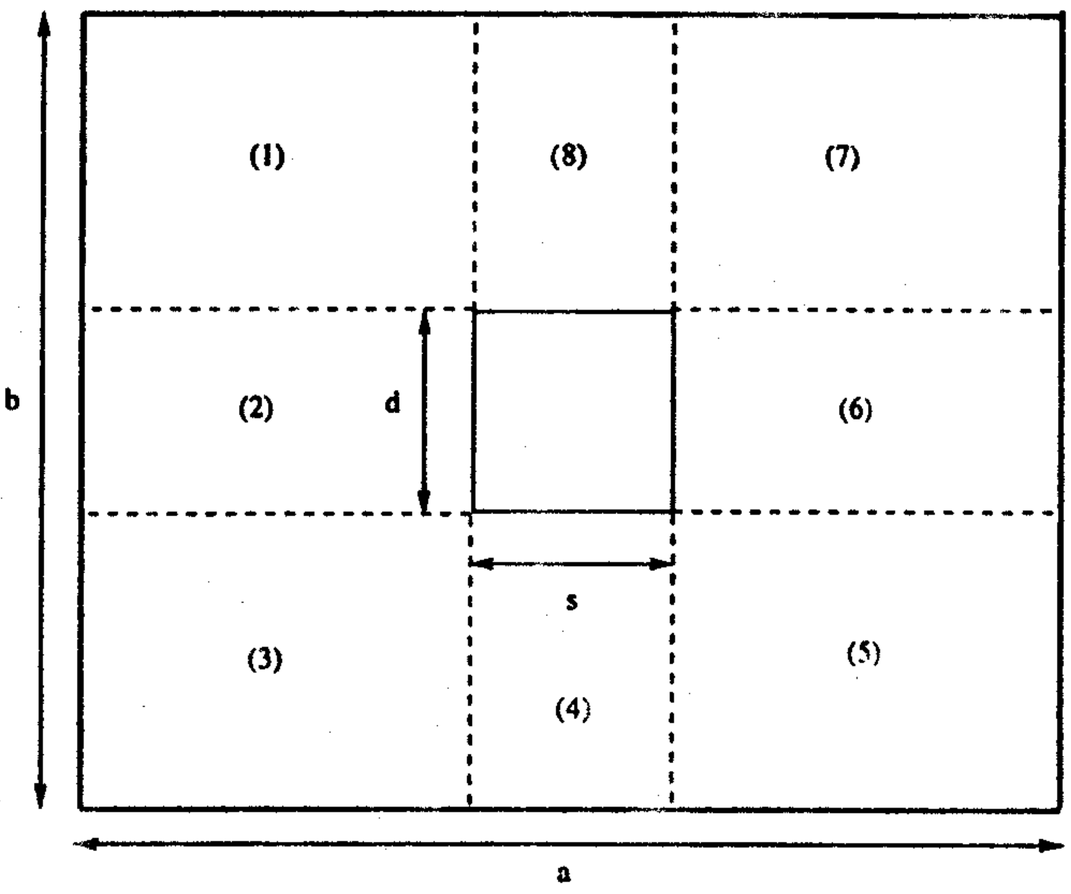

Figure 4. Coaxial rectangular waveguide (a = 1.25 m; b = 1.0 m; d = 0.25 m; s = 0.25 m).

matrix that is diagonal. Table 3 provides the cutoff wave numbers of the TE modes of the same single ridge waveguide, and Table 4 shows the percentage of the system matrix that is diagonal.

Comparison of these results with data published in the literature utilizing a finite difference technique due to *Sarkar et. al.* [1989] and a surface integral equation due to *Swaminathan et. al.* [1990] show that the new approach is quite accurate and provides fast convergence.

As a second example consider the coaxial rectangular waveguide shown in Figure 4. The waveguide has been divided into eight regions. Table 5 provides the cutoff wave numbers of the TM modes this waveguide. Table 6 shows the percentage of the system matrix that is diagonal as the order of the approximation increases. Table 7 provides the cutoff wave numbers for the TE modes, and Table 8 shows the percentage of the system matrix that is diagonal as the order of the system increases. Again, good agreement has been obtained with other published results. The three dots indicate that the results are not available.

Table 4. Percentage of the Matrix That is Diagonal for the TE Modes.

	Matrix Size	Row Size of the Diagonal Block	Percent Diagonal
N=2	64	20	31.25
N=4	156	80	51.28
N=6	288	180	62.50
N=8	460	320	69.57
N=10	672	500	74.40

Table 5. Cut-off Wave Numbers of the TM Modes of a Coaxial Rectangular Waveguide.

	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5
N=2	6.95	6.96	8.68	8.72	10.97
N=4	6.94	6.95	8.66	8.70	10.94
N=6	6.94	6.95	8.66	8.70	10.92
N=8	6.94	6.95	8.66	8.70	10.92
N=10	6.94	6.95	8.66	8.70	10.92
Sarkar et.al. [1989]	6.91	6.96	8.5	8.51	10.57

Table 6. Percentage of the Matrix That is Diagonal for the TM Modes.

	Matrix Size	Row Size of the Diagonal Block	Percent Diagonal
N=2	48	32	66.67
N=4	160	128	80.00
N=6	336	228	85.71
N=8	576	572	88.89

Table 7. Cut-off Wave Numbers of the TE Modes of a Coaxial Rectangular Waveguide.

	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5
N=2	1.90	2.84	3.91	5.18	5.78
N=4	1.89	2.84	3.91	5.16	5.77
N=6	1.89	2.84	3.91	5.16	5.77
N=8	1.89	2.84	3.91	5.16	5.76
Sarkar et. al. [1989]	1.85	2.81	3.89	5.05	5.68

Table 8. Percentage of the Matrix That is Diagonal for the TE Modes.

	Matrix Size	Row Size of the Diagonal Block	Percent Diagonal
N=2	96	32	38.33
N=4	240	128	53.33
N=6	448	288	64.29
N=8	720	512	71.11

Conclusion

The principal of dilation, extensively used by the wavelet concept, can be introduced into finite element techniques for efficient choice of basis functions. With the new basis functions the large finite element method system matrices can be made mostly diagonal and the computational complexity can be significantly reduced. This approach can easily be extended to three-dimensional problems or to waveguides containing inhomogeneous materials.

References

Alpert, B. K., Wavelet and other bases for fast numerical linear algebra, in *Wavelets: A Tutorial in Theory and Applications*, edited by in C. K. Chui, pp. 181-216, Academic Press, San Diego, Calif., 1992.

Chen, H., T. K. Sarkar, S. A. Dianat, and J. D. Brulé, Adaptive spectral estimation by the conjugate gradient method, *IEEE Trans., Acoust. Speech, Signal Processing*, *34*, 272-284, 1986.

Chui, C. K., An introduction to wavelets", Academic, San Diego, Calif., 1992.

Daubechies, I., Ten lectures on wavelets, SIAM, CBMS -NSF Regional Conference Series in Applied Mathematics, *61*, 1992.

Flanagan, J. L., *Speech Analysis, Synthesis and Percep tion*, Springer-Verlag, New York, 1972.

Golub, G., and C. F. VanLoan, *Matrix computations*, Johns Hopkins University Press, Baltimore, MD, 1989.

Harrington, R. F., *Field Computation by Moment Methods*, Macmillan, New York, 1967.

Jaffard, S., Wavelets and analysis of partial differential equations, in *Probabilistic and Stochastic Methods in Analysis With Applications*, edited by J. S. Byrnes et. al., pp. 3-13, Kluwer Academic, Norwell, Mass, 1992.

Jaffard, S., and P. Laurencot, Orthonormal wavelets, analysis of operators, and applications to numerical analysis, in *Wavelets: A Tutorial in Theory and Applications*, edited by C. Chui, Academic, San Diego, Calif, 1992.

Krasnoseliskii, M. A., *Approximate Solution of Operator Equations*, Noordhoff, Leiden Netherlands, 1972.

Mikhlin, S. G., *The Numerical Performance of Variational Methods*, Noordhoff, Leiden Netherlands, 1971.

Nelson, G. A., L. L. Pfeiffer, and R. C. Wood, High speed octave band digital filtering", *IEEE Trans. on Audio and Electroacoust*, *20*, 8-65, 1972.

Peterson, A. F., Eigenvalue projection theory for linear operator equations of electromagnetics, *UILV-ENG-87-2252*, Coord. Sci. Lab., Univ. of Ill., 1987.

Sarkar, T. K. (Ed.), *Application of conjugate gradient method to electromagnetics and signal analysis*, Prog. Electromagn. Res., *5*, Elsevier, New York, 1991.

Sarkar, T. K., K. R. Siarkiewicz, and R. F. Stratton, Survey of numerical methods for solution of large systems of linear equations for electromagnetic field problems, *IEEE Trans. Antennas & Propag.*, *29*, 847-856, 1981.

Sarkar, T. K., K. Athar, E. Arvas, M. Manela, and R. Lade, Computation of the propagation characteristics of TE and TM modes in arbitrary shaped hollow guides utilizing the conjugate gradient method, J. Electromagn. Waves Appl., *3*, 143-165, 1989.

Schafer, R. W., L. R. Rabiner, and O. Herrmann, FIR digital filter banks for speech analysis, *Bell Syst. Tech. J.*, *54*, 531-544, 1975.

Swaminathan, M., E. Arvas, T. K. Sarkar, and A. R. Djordevic, Computation of cut-off wavenumbers of TE and TM modes in waveguides of arbitrary cross sections using a surface integral formulation, *IEEE Trans. Microwave Theory Tech.*, *38*, 154-159, 1990.

Vaidyanathan, P. P., *Multirate systems and filter banks*, Prentice-Hall, Englewood Cliffs, N.J., 1993.

T. K. Sarkar and R. S. Adve, Department of Electrical and Computer Engineering, Syracuse University, New York, Syracuse, NY 13244. (email:tksarkar@rodan.syr.edu)

Luis Emilio García-Castillo and Magdalena Salazar-Palma, Signals, Systems and Radiocommunication Department Escucla Tecnica Superior de Ingenieros Telecomunicación, Universidad Politécnica de Madrid, Ciudad Universitaria, Madrid.

(Received August 11, 1993; revised February 7, 1994; accepted March 10, 1994.)