Determination of the Effective Hydraulic Conductivity for Heterogeneous Porous Media Using a Numerical Spectral Approach

1. Method

BRUCE B. DYKAAR AND PETER K. KITANIDIS

Department of Civil Engineering, Stanford University, Stanford, California

A numerical spectral method is described for solving the equations governing the effective hydraulic conductivity of a stationary heterogeneous porous medium. The numerical method is efficient in the sense that the solution can be achieved in order $N^2 \ln (N^2)$ operations and order $N$ storage, where $N^2$ is the total number of sampling points used in an $n$-dimensional flow domain. The numerical method is extensively tested and shown to be quite robust. In the test cases, very accurate results are achieved even for an extremely heterogeneous porous medium. The developed numerical algorithm is quite general and can be used to compute effective conductivities for cases where analytic results are unavailable. In the companion paper (Dykaar and Kitanidis, this issue) the method is used to find the effective conductivity of (1) multidimensional isotropic stationary, lognormally distributed conductivity and (2) core samples of a shale and sandstone formation.

1. Introduction

Natural porous media exhibit significant spatial variability in most attributes of hydrogeologic interest. For instance, it is quite typical for hydraulic conductivity to vary orders of magnitude over short distances. Data from the Mt. Simon aquifer in Illinois showed a four order of magnitude variation in conductivity [Gelhar, 1986], while even for the relatively homogeneous Borden aquifer the conductivity varied by a factor of ten over a distance of only 1 m in the vertical direction [Sidlecky, 1986]. This severe and ubiquitous spatial heterogeneity makes modeling groundwater flow and solute transport very challenging.

The equation describing fluid flow through a porous medium at the Darcy scale is

$$\nabla \cdot [K(x) \nabla \phi(x, t)] = S \frac{\partial \phi(x, t)}{\partial t}$$

(1)

where $x = (x_1, x_2, x_3)$ is the spatial coordinate, $K(x)[L/T]$ is the hydraulic conductivity tensor, $S [L^2/T]$ is specific storativity, and $\phi(L)$ is the hydraulic head. In general, analytic solutions to (1) are not available when the hydraulic conductivity is an arbitrary spatial function of $x$. Numerical methods provide the only general solution procedure, but they have their difficulties. When the spatial variability of $K(x)$ is large, equation (1) may be difficult to solve for two reasons. First, regardless of the numerical scheme used, large variations in conductivity lead to solving an ill-conditioned matrix equation. We will show that this is the case for the numerical spectral method used in this study. Second, the large number of points it would take to represent the function in detail in multidimensions would quickly overwhelm even a supercomputer. An additional consideration is that it is impractical, and probably impossible, to know the function $K(x)$ precisely at the Darcy scale, since this would require a very fine measurement grid with sample locations of the order of only centimeters apart.

It would be desirable then, to replace (1) with another equation which would be valid at some scale larger than the Darcy scale, where the variations in $K(x)$ are lumped into a constant macroscopic or effective value $K'$. For a porous medium of volume $V$ the effective hydraulic conductivity is usually defined, in an analogous way to Darcy’s law, as the constant of proportionality between the mean head gradient and mean specific discharge:

$$q(x) = -K' \nabla \langle \phi \rangle$$

(2)

where the angle brackets $\langle \cdot \rangle$ denote mean value, $q[L/T]$ is the specific discharge vector, and $K'[L/T]$ is the effective conductivity tensor. The heterogeneous medium is thereby replaced with an “equivalent” homogeneous one of the same size, and (1) is replaced by an equation of the same form:

$$\nabla \cdot [K' \nabla \phi(x, t)] = S \frac{\partial \phi(x, t)}{\partial t}$$

(3)

valid at scales of the size of $V$.

For the upscaling procedure to be applicable, two disparate scales of variability in $K(x)$ are required: the highly variable Darcy or local scale and the larger, more slowly varying, macroscale [Dagan, 1986; Gelhar, 1986]. If local-scale fluctuations are of the order of length $\ell$, macroscale variations are of the order of length $L$, and the averaging volume $V$ is of the order of length $L$, then the condition

$$\ell \ll L \ll D$$

(4)

is required for effective parameters to exist and be well defined. This observation points out the inherent difficulty with numerical schemes for determining $K'$: fine sampling is needed at the small scale $\ell$ to represent adequately the function $K(x)$, yet averaging is done over the much larger scale $L$. This will in most cases require an extremely large number of data points to sample the entire averaging volume. For numerical upscaling methods, this ultimately leads to solving a partial equation with no more solutions than direct methods for solution, such as LU decomposition.
become impractical due to computer storage and computational time limitations.

This work contributes to the study of effective parameters by providing an efficient and accurate numerical scheme for calculating the effective hydraulic conductivity tensor based on the results of Kitanidis [1990]. Following the Taylor-Aris method of moments, as generalized by Brenner [1980, 1982], Kitanidis derived the governing system of equations for the effective hydraulic conductivity of a periodic porous medium. The key assumptions made in the method of moments are (1) the conductivity function is periodic or, in the limit as \( L \to \infty \), stationary, (2) the boundary conditions are periodic, and (3) the flow is gradually varying; specifically, the zeroth, first, and second central moments of the head field computed over the averaging volume \( V \) reach a steady rate of change. Compared with other upscaling methods, these three assumptions are probably the least restrictive.

The relevance of the periodic assumptions has provoked some debate and confusion in the hydrogeology community and merits further discussion at this point. The moment method assumes that the averaging volume \( V \) encompasses one period of an idealized periodic porous medium. Of course, no natural porous medium is truly periodic, but it can be approximated as such by neglecting variability at scales larger than the size of \( V \). Assuming periodicity then places an upper limit on the largest-scale components used in the mathematical representation of the conductivity field. For a realistic nonperiodic porous medium, which has variability at all scales, the computed effective conductivity will, in general, be a function of the size of the averaging volume used. Viewing \( V \) as an independent variable, the moment method allows exploration of the importance of these large-scale components on the value of \( K' \). The restrictiveness of assuming periodic boundary conditions is less intuitive and is addressed in the companion paper through several examples and by comparison with other methods which make different assumptions.

2. Moment Method Approach

In this section the final results of Kitanidis [1990] are summarized, and the stage is set for the numerical solution. Finding the effective conductivity via the moment method is a two-step procedure. First, \( n \) uncoupled partial differential equations are solved for the unknown auxiliary functions \( g^i(x) \), \( i = 1, \cdots, n \), where \( n \) is the dimension of the flow domain. The partial differential equations to solve are linear elliptic equations with variable coefficients given by

\[
\nabla \cdot [K(x) \nabla g^i(x)] = \nabla \cdot \kappa_i(x) \quad i = 1, \cdots, n \tag{5}
\]

where \( \kappa_i \) is the \( i \)th column of \( K \). The domain of interest is a parallelepiped defined in, for example, three dimensions by the volume \( V = L_1 L_2 L_3 \) with sides of length \( L_1 \), \( L_2 \), and \( L_3 \). Equation (5) is subject to periodic boundary conditions, which means \( g^i \) and all its derivatives are periodic at \( x_j = 0, L_j \). The second step uses the functions \( g^i \) in an integral equation to determine the effective conductivity tensor:

\[
K_{ij}^e = -\frac{1}{2V} \int_V (\kappa_i \cdot \nabla g^j + \kappa_j \cdot \nabla g^i) \, dx + \bar{K}_{ij} \tag{6}
\]

where \( \bar{K}_{ij} \) is the spatial average of \( K_{ij} \):

\[
\bar{K}_{ij} = \frac{1}{V} \int_V K_{ij}(x) \, dx \tag{7}
\]

While the method is general, the situation investigated in this paper is for locally isotropic conductivity. This means that, at the local scale, the porous medium exhibits no change in conductivity with respect to direction. This assumption does not preclude the possibility of anisotropy at the large scale; local-scale heterogeneity displaying direction-dependent variability will produce anisotropic effective parameters in the upscaling process [Gelhar and Axness, 1983]. The conductivity tensor for a locally isotropic heterogeneous medium is

\[
K = \begin{bmatrix} K(x) & 0 & 0 \\ 0 & K(x) & 0 \\ 0 & 0 & K(x) \end{bmatrix} \tag{8}
\]

Equations (5) and (6) become, with the locally isotropic assumption, respectively,

\[
\sum_{p=1}^{n} \frac{\partial}{\partial x_p} \left( K \frac{\partial g^i}{\partial x_p} \right) = \frac{\partial K}{\partial x_i} \quad i = 1, \cdots, n \tag{9}
\]

\[
K_{ij}^e = \frac{1}{2V} \int_V K \left( \frac{\partial g^j}{\partial x_i} + \frac{\partial g^i}{\partial x_j} \right) \, dx + \bar{K} \delta_{ij} \tag{10}
\]

where \( \delta_{ij} \) is the Kronecker delta, \( \delta_{ij} = 1 \) for \( i = j \), and \( \delta_{ij} = 0 \) otherwise. In the following sections a numerical spectral scheme to solve equations (9) and (10) is described.

3. Numerical Solution

3.1. Spectral Discretization

Spectral methods are well suited to the solution of (9) for several reasons. First, the periodic boundary conditions are easily handled. Second, spectral representations of rapidly fluctuating functions, such as \( K \) at the local scale, are known to be much more accurate than, say, a finite difference discretization for the same number of nodes [Canuto et al., 1988]. Or equivalently, a spectral representation requires far fewer sampling points for a given degree of accuracy than the finite difference representation. As alluded to earlier, the upscaling problem is inherently computationally intensive, and the enhanced accuracy of spectral methods can result in substantial savings in computational resources, particularly in higher dimensions.

Spectral methods are a class of discretization schemes within the more general category of the method of weighted residuals. Particular schemes within the method of weighted residuals are distinguished by the particular basis function \( \beta_k(x) \) used for representation:

\[
K(x) = \sum_k K_k \beta_k(x) \tag{11}
\]

where \( K_k \) are the expansion coefficients, and the weight function \( w(x) \) used in the weighted residual condition:

\[
\int_V \nabla y(x) \cdot \nabla w(x) \, dx = 0 \tag{12}
\]
where $V$ is the domain of interest and $\Lambda(\cdot)$ is a differential operator. In this case the differential operator is

$$
\Lambda(\cdot) = \nabla \cdot [K \nabla (\cdot)] - \nabla \cdot \mathbf f,
$$

(13)

The choice of the weight function distinguishes between the three most common methods: Galerkin, tau, and collocation. Selection of the basis function is partially dictated by the boundary conditions. For periodic boundary conditions the complex exponentials (trigonometric polynomials) are ideally suited for use as the basis functions because they automatically satisfy the boundary conditions. If the same functions are also used as weight functions, then the scheme is known as the Fourier Galerkin method. The Fourier Galerkin method is used to solve equation (9) and is described next.

For $\beta_k = \exp (i2\pi x \cdot f)$, where $i = (-1)^{1/2}$, $f = (k_1/L_1, k_2/L_2, k_3/L_3)$ is the frequency vector, $k_i$ is the integer wave number spanning $-N_i/2 \leq k_i \leq N_i/2$, $N_i$ is the even number of sampling points in direction $i$, and $L_i$ is the periodic length in direction $i$, (11) becomes the usual truncated complex Fourier series representation of the function $K(x)$. For the weight function $w(x) = \exp (-i2\pi x \cdot f)$ the weighted residual condition (12) is a Fourier transform of $\Lambda(g')$. The Fourier Galerkin method then amounts to using Fourier series representations and requiring that the Fourier transform of the differential operator vanish.

Applying this methodology, first expand $K$ and $g'$ in their respective truncated Fourier series:

$$
K(x) \approx \sum_{|k| \leq k'} \hat K_k e^{i2\pi x \cdot f},
$$

(14)

$$
g'(x) \approx \sum_{|k| \leq k'} \hat g'_k e^{i2\pi x \cdot f},
$$

(15)

where the hat symbol signifies the usual Fourier series coefficient and $k'$ is the cutoff wave number vector. An important practical consideration is that the value of $k' = (N_1/2, N_2/2, N_3/2)$ must be large enough for the accurate representation of the function $K(x)$ or $g'(x)$. The advantage of spectral methods is that, in many cases of practical interest, to achieve a certain degree of accuracy, $N_i$ is relatively small compared to the number of collocation points required in other discretization schemes. For an infinitely differentiable function with periodic derivatives, the $k$th coefficient of the Fourier series expansion decays faster than any inverse power of $k$ [Canuto et al., 1988]. This type of decay is called exponential convergence and is also known in the literature as "spectral accuracy." A Fourier series truncated after a few terms into this rapid decay represents a very accurate approximation to the function. Hence with spectral representations the order of accuracy increases dramatically with the number of discretization points once exponential convergence is achieved. This is in marked contrast to finite elements, for example, where the order of accuracy is usually very low (typically $2$ or $3$) and is set a priori with the particular formulation used. However, this characteristic convergence of spectral methods occurs only after there are enough coefficients to capture the essential structure of the represented function.

Applying a Fourier transform to (9) using (14) and (15) yields a system of linear algebraic equations in the unknown Fourier series coefficients $\hat g'_k$ given in the three-dimensional case by

$$
17 \pi \sum_{p=1}^{N_1} \sum_{l_1=-N_1/2}^{N_1/2} \sum_{l_2=-N_2/2}^{N_2/2} \sum_{l_3=-N_3/2}^{N_3/2} \frac{m_p l_p}{L_p} \hat K_{m_p l_p} \hat g'_k \frac{1}{2L_i} \hat K_{m_i} = \frac{m_i}{2L_i} \hat K_{m_i}
$$

$$
m_1 = -N_1/2, \cdots, N_1/2
$$

$$
m_2 = -N_2/2, \cdots, N_2/2
$$

$$
m_3 = -N_3/2, \cdots, N_3/2
$$

(16)

where $m_p, m_i$, and $l_p$ are wave number components of the three-dimensional wave number vectors $m = (m_1, m_2, m_3)$ and $l = (l_1, l_2, l_3)$. The details of the derivation are given in Appendix A. In matrix notation, (16) is written

$$
A \hat g' = \mathbf b'
$$

(17)

where $\hat g'$ is the vector of unknown Fourier coefficients with the $m$th element given by

$$
(\hat g'_m) = \hat g'_m
$$

(18)

the $m$th element of the vector $\mathbf b'$ is

$$
(b'_m) = \frac{m_i}{2L_i} \hat K_{m_i}
$$

(19)

and the $(m, l)$th element of $A$ is

$$
(A)_{m, l} = 17 \pi \sum_{p=1}^{n} \frac{m_p l_p}{L_p} \hat K_{m_p l_p}
$$

(20)

Methods for solving (17) are investigated in the following section.

Consider now the problem of numerically evaluating (10). This is conveniently accomplished using the Fourier series representations for $K$ and $g'$, equations (14) and (15), respectively. The volume integration can be performed with the assistance of the orthogonality condition for the complex exponentials given in Appendix A by $(\Lambda)_{k+}$, yielding

$$
K'_{ij} = K \delta_{ij} - \pi \sum_{k=0}^{k'} \frac{k_i}{L_i} \hat K_{k} \hat g'_k + \frac{k_j}{L_j} \hat K_{k} \hat g'_k
$$

(21)

Since the conductivity field is real, the Fourier coefficients are related by $K_{k} = K_{k}^*$, where the asterisk denotes complex conjugate. Using this fact, the above equation becomes

$$
K'^{c}_{ij} = K \delta_{ij} - \pi \sum_{k=0}^{k'} \frac{k_i}{L_i} \hat K_{k} \hat g'_k + \frac{k_j}{L_j} \hat K_{k} \hat g'_k
$$

(22)

Equations (16) and (22) are the discretized versions of (9) and (10). However, these equations are not yet in a form suitable for general numerical implementation, and the next section includes a discussion on the necessary modifications. For notational simplicity, in the remainder of this paper we will use the same number of sampling points in each direction $N_i = N$ and take $l_i = 1$ so the frequency and wave number are equivalent, $f = k$. 
3.2. Numerical Implementation

A difficulty with the numeric implementation of (16) and (22) is that the Fourier coefficients of the conductivity field given by

$$\tilde{K}_k = \frac{1}{N} \sum_{i=0}^{N-1} K(x_i) e^{-i 2 \pi k x/L}$$

(23)

are required. In general, these coefficients cannot be found in closed form for an arbitrary function $K(x)$. However, the Fourier coefficients can be approximated by the discrete Fourier coefficients which, for example, in one dimension are given by

$$\tilde{K}_k = \frac{1}{N} \sum_{i=0}^{N-1} K(x_i) e^{-i 2 \pi k x/L} \quad k = -N/2, \ldots, N/2 - 1$$

(24)

computed on the equally spaced sampling points

$$x_i = \frac{i}{N} L \quad i = 0, 1, \ldots, N - 1$$

(25)

The tilde is used to denote a discrete Fourier coefficient. Equation (24) is known as the discrete Fourier transform. From the discrete Fourier coefficients the function values at the sampling points can be recovered using the inverse discrete Fourier transform

$$K(x_i) = \frac{1}{N} \sum_{k=-N/2}^{N/2-1} \tilde{K}_k e^{i 2 \pi k x/L}$$

(26)

$$i = 0, 1, \ldots, N - 1$$

Using an algorithm called the fast Fourier transform (FFT) [Press et al., 1988], the discrete Fourier transform and its inverse can be efficiently evaluated compared with order $N^{2n}$ operations for the direct computation of (24) or (26). The fastest and most popular FFT programs are base 2 algorithms which require $N$ to be a power of two and have an operation count of order $N^{n}$ ln ($N^{n}$). Base 2 FFT programs were used in this research.

For periodic functions such as the spatial function $K$ given by (14), a relation between the continuous and discrete Fourier coefficients can be derived. Substituting the continuous Fourier series formula (14) into the discrete equation (24) and using the orthogonality relation

$$\frac{1}{N} \sum_{i=0}^{N-1} e^{i 2 \pi k x/L} = \begin{cases} 1 & k = 0, \pm 1, \pm 2, \ldots \\ 0 & \text{otherwise} \end{cases}$$

(27)

yields an expression for the discrete Fourier coefficients in terms of the continuous coefficients which is, in one dimension,

$$\tilde{K}_k = \tilde{K}_k + \sum_{m=0}^{\infty} \tilde{K}_{k+N_m} \quad k = \frac{N}{2}, \ldots, \frac{N}{2} - 1$$

(28)

Equation (28) shows that the $k$th frequency of $\tilde{K}_k$ depends not only on the $k$th frequency of $K_k$ but also on all the modes $K_{k+N_m}$. The frequencies $k + N_m$ for $m = -\infty, \ldots, \infty$ are superimposed on, or "alias" the frequency $k$. At some number of sampling points, say $N_s$, the magnitude of the aliasing modes $K_{k+N_m}$ is small compared with $K_k$ over the range of frequencies used, say $k = -N_k/2, \ldots, N_k/2$. Then the discrete Fourier coefficients are good approximations for the continuous coefficients, $\tilde{K}_k = K_k$, for these frequencies. The required value of $N_s$ depends on the particular function $K(x)$; the richer $K$ is in high-frequency energy, the greater $N_s$ will have to be. Using the discrete Fourier series coefficients $\tilde{K}_k$, equations (14)-(22) remain essentially the same except that the Fourier coefficients $K_k$ and $\tilde{K}_k$ are replaced with the discrete Fourier coefficients $\tilde{K}_k$ and $\tilde{G}_k$.

Now we are ready to describe the method used to solve (16) for the unknown discrete Fourier coefficients $\tilde{G}_k$. For reference, we note the potential of using a direct matrix solver such as LU decomposition [Strang, 1988]. The matrix $A$ defined by (20) is a full matrix, so LU decomposition for $N$ discretization points in each dimension $n$ requires order $N^{3n}$ operations and order $N^{2n}$ storage. For values considered in this paper, such as $N = 80$ and $n = 3$, the operation count and storage requirement is clearly way beyond the capability of even a supercomputer.

Iterative methods offer a viable alternative to direct methods. Since the differential operator $A(x)$ is self-adjoint and $K$ is a symmetric positive definite tensor, the Fourier Galerkin discretization given by (16) is a Hermitean positive definite system of linear equations (the complex analogy to symmetric positive definite). Among the several iterative methods available for symmetric positive definite systems [Hageman and Young, 1981] the conjugate gradient method was selected. Following Golub and Van Loan [1989], the conjugate gradient algorithm for solving $Ax = b$ is given in Table 1, where $q$ is the iteration index and for $x$ the value of the solution vector at iteration $q$ then $r = b - Ax$ is the residual, $\rho_{q-1}$ is the solution component of the largest eigenvalue of $x + \rho_{q-1} x$ and $\beta$ is the direction vector. The loop beginning at line cg2 continues

<table>
<thead>
<tr>
<th>Line</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>cg1</td>
<td>Initialize $q = 0; x = 0; r = b; r_n = 1$</td>
</tr>
<tr>
<td>cg2</td>
<td>$q = q + 1$</td>
</tr>
<tr>
<td>cg3</td>
<td>If $q = 1$</td>
</tr>
<tr>
<td>cg4</td>
<td>$p = r$</td>
</tr>
<tr>
<td>cg5</td>
<td>Else</td>
</tr>
<tr>
<td>cg6</td>
<td>$\beta = |r|_2 = 1$</td>
</tr>
<tr>
<td>cg7</td>
<td>$p = r + \beta p$</td>
</tr>
<tr>
<td>cg8</td>
<td>$w = Ap$</td>
</tr>
<tr>
<td>cg9</td>
<td>$a = r^T w$</td>
</tr>
<tr>
<td>cg10</td>
<td>$x = x + \alpha p$</td>
</tr>
<tr>
<td>cg11</td>
<td>$r = r - \alpha w$</td>
</tr>
<tr>
<td>cg12</td>
<td>$r_n = |r|_2$</td>
</tr>
<tr>
<td>cg13</td>
<td>End</td>
</tr>
<tr>
<td>cg14</td>
<td>End</td>
</tr>
</tbody>
</table>

Table 1. Conjugate Gradient Algorithm for Solving $Ax = b$
executing until either the normalized residual is smaller than some specified error criterion \( \epsilon \) or the maximum allowed iterations \( q_{\text{max}} \) is reached. The quantities \( a \) and \( \beta \) are scalar and, in our case, are always real valued since \( A \) is Hermitian. The superscript \( H \) stands for conjugate transpose, and the norm of a vector is defined as \( ||b|| = (b^H b)^{1/2} \).

An efficient implementation of the conjugate gradient method, such as the one above, requires just one matrix-vector multiplication per iteration, as in line cg10. This single operation is the bottleneck in terms of operation count and storage, both of order \( N^{2n} \). Although a considerable improvement over direct methods, it is still not good enough to tackle the upsampling problem. For our case there is a much faster way available to perform the matrix-vector multiplication based on the fast Fourier transform.

Consider performing the matrix-vector multiplication \( w = Ap \), where \( A \) is given by (20). As an example, let \( n = 2 \). Then,

\[
  w_m = \pi m_1 \sum_{l=0}^{n} \hat{K}_{m-l} \hat{X}_l^{(1)} + \pi m_2 \sum_{l=0}^{n} \hat{K}_{m-l} \hat{X}_l^{(2)}
\]

where we have defined

\[
  \hat{X}_l^{(1)} = l_1 p_1
\]

\[
  \hat{X}_l^{(2)} = l_2 p_1
\]

The summations \( \sum \hat{K}_{m-l} \hat{X}_l^{(1)} \) and \( \sum \hat{K}_{m-l} \hat{X}_l^{(2)} \) are known as convolution sums. A convolution sum can be performed in order \( N^n \) ln \( (N^n) \) operations by employing the convolution theorem of the Fourier transform and using the FFT to perform the transform. The basic idea is to inverse Fourier transform \( \hat{K}_m \) and \( \hat{X}_l^{(i)} \) in order to get \( K(x) \) and \( X'(x) \), respectively. Multiply the functions in the physical domain \( Z'(x) = K(x) X'(x) \), then transform \( Z'(x) \) to get \( Z_m \). Equation (29) becomes

\[
  w_m = \pi n_1 m_2 Z_m^{(1)} + \pi n_2 m_2 Z_m^{(2)} + E_m
\]

where \( E_m \) is an error term. Canuto et al. [1988, p. 83] observe that this technique "was the single most important development which made spectral Galerkin methods practical for large-scale computations." With the matrix-vector multiplication of the conjugate gradient scheme performed via FFT convolutions, the operation count per iteration is of the order \( N^n \) log \( (N^n) \), and the storage requirement is reduced to order \( N^n \), since only the vector \( \hat{K}_m \) needs to be stored and not the entire matrix \( A \). Now the problem is within the reach of current computational resources.

The error term in (32) arises because the FFT convolution procedure performs a circular convolution and so some frequency components may become corrupted because of "overlap" [Brigham, 1988]. The error term can be made exactly zero if the two series to be convolved are sufficiently lengthened by zero padding, or alternatively, the corrupted frequencies are simply not used. The standard procedure for computing error-free convolution sums [Canuto et al., 1988] is to truncate both series at the same length, say \( N \), and then to extend the length of both series to their padded length of \( N_p \) points by padding the high frequencies with zeros, where \( 3N/2 - 1 \leq N_p \). This ensures that the frequency components of \( Z_k \) for \( |k| \leq N/2 \) are not corrupted by overlap. The cost of the new method of the FFT convolution is of order \( N_p^n / (N_p^n) \), and storage is of order \( N_p^n \). Also, recall that \( N_p \) must be a power of 2 for use with a base 2 FFT program. This procedure is precisely equivalent to the matrix-vector multiplication \( w = Ap \), with the \( (m, m) \)th element of the matrix \( A \) given by

\[
  (A')_{m,m} = \pi n_1 \sum_{p=1}^{n} m_p l_p \hat{K}_{m-p} \quad |m_p - l_p| \leq N/2
\]

(33)

\[
  (A')_{m,1} = 0 \quad N/2 < |m_p - l_p| \leq N
\]

where, as before, \( m_1, m_2, m_3, l_1, l_2, l_3 = N/2, \ldots, N/2 \). Compare this with the full matrix defined in (20).

In the course of the numerical experiments conducted it was noticed that for a porous medium that was "too heterogeneous", or rich in high-frequency energy, the conjugate gradient method would not converge using the standard FFT convolution procedure described above. For example, for lognormally distributed media generated from a Gaussian covariance function this behavior occurred for roughly \( \sigma^2 > 2 \). Some investigating was done to determine the cause of this behavior. The eigenvalues of a Hermitian positive definite matrix are all real and greater than zero. For the conductivity fields of interest in this paper, lognormally and bimodally distributed, several one-dimensional examples were created, and the eigenvalues of the matrices defined by (20) and (33) were computed. It was found that beyond some very moderate level, measured either by the condition number of the matrix or the variance of \( K(x) \), \( A' \) would cease to be positive definite even though \( A \) was positive definite for all the cases tried.

The conjugate gradient method requires that the matrix used be positive definite, so the standard method for performing the FFT convolution is inadequate for our purposes. To make the matrix positive definite, a modified version of the FFT convolution was devised which includes more frequency components of \( K \) than \( g' \). If the highest frequency of \( g' \) is \( N_p/2 \), then \( K \) should include all frequencies up to \( N_p/2 \), where \( N_p \approx N_p \). The amount of zero padding required to ensure that the FFT convolution of two different length series does not corrupt the frequencies \( N_p/2 \) is

\[
  N_p \geq \frac{N_p}{2} + N_p + 1
\]

(34)

where \( N_p \) is the padded series length, \( N_p \) is the length of the \( g' \) series, and \( N_p \) is the length of the \( K \) series. We have found that for \( N_p \approx 2N_p \), the conjugate gradient program converged very smoothly for all cases tried. Using the empirical relation \( 2N_p = N_p \), then, for a given padded series length, the longest possible length for the unknown coefficients is \( N_p = (N_p - 1)/2 \). The authors are unaware of any other published work which points out the difficulties of the standard FFT convolution procedure when used with the conjugate gradient method and provides a solution.

For some functions, most notably discontinuous ones, the Fourier series coefficients decay very slowly. It might be desirable to attenuate the high-frequency components of \( \hat{K}_m \) by applying a low-pass filter. In the physical domain this operation is equivalent to smoothing \( K(x) \) by convolving (the moving average) it with the Fourier transform of the filter. This would tend to round the corners and fill in the sharp peaks and valleys of \( K(x) \). In this paper and the companion one [Dykaar...
and Kitanidis, this issue) the impact of filtering on $K'$ and on the performance of the conjugate gradient method is explored.

Filtering is implemented by multiplying the discrete Fourier coefficient by a weighting factor $\sigma_k$. For example, the filtered version of the conductivity becomes

$$K_f(x) = \sum_{|k| = k'} \sigma_k K_k e^{i2\pi x \cdot k}$$  \hspace{1cm} (35)

One of the low-pass filters used in this study was the raised cosine filter, which in one dimension is given by

$$\sigma_k^{\text{cos}} = \frac{1}{2} \left[ 1 + \cos \left( \frac{2\pi k}{N} \right) \right] \hspace{1cm} k = \frac{N}{2}, \cdots, \frac{N}{2}$$  \hspace{1cm} (36)

It is enlightening to view the truncated Fourier series as the infinite Fourier series filtered by the rectangular filter given by

$$\sigma_k^{\text{rect}} = \begin{cases} 1 & k = \frac{N}{2}, \cdots, \frac{N}{2} \\ 0 & \text{otherwise} \end{cases}$$  \hspace{1cm} (37)

Summarizing, the numerical implementation of the Fourier Galerkin discretization is subject to three possible sources of error: (1) aliasing, (2) series truncation or low-pass filtering, and (3) numerical round off. Aliasing errors are caused by discrete sampling $K(x)$ at an insufficient rate $N_x/L_x$. For a fixed periodic length $L_x$, aliasing errors are controlled by $N_x$. Low-pass filtering causes errors by attenuating or completely extinguishing important Fourier series coefficients and is controlled by the choice of $N_x$, $N_y$, and filter weights $\sigma_k$. Numerical round off can potentially be cause for concern in severely heterogeneous media. The effect of spatial heterogeneity on the condition number of $A$ is explored in section 6.

While we are on the subject of numerical errors and the idea of filtering, it is worth noting that all numerical schemes, such as finite differences or finite elements, are subject to the same error sources for the same reasons. In addition, the finite difference or finite element operator itself acts as a filter to varying degrees, depending on the particular differential operator. In Appendix B this is illustrated by way of a model spatially variable one-dimensional function. This appendix compares the performance of second-order finite differences with spectral methods for the simplest differential operator, $A(\cdot) = d(\cdot)/dx$. Despite its simplicity, this example does provide insight into how finite differences will tend to behave when representing highly fluctuating functions. The results show that second-order finite differences will attenuate the actual frequency component $k$ by an amount given by the low-pass filter weighting factor:

$$\sigma_k^{\text{fd}} = \frac{N}{2\pi k} \sin \left( \frac{2\pi k}{N} \right)$$  \hspace{1cm} (38)

where $N$ is the number of sampling points in a domain of length one. Figure 1 is a plot of $\sigma_k^{\text{fd}}$: note the high-frequency attenuation. No distortion occurs with spectral differentiation of the model function. This partially explains why finite differences require many more sampling points than spectral methods to represent accurately rapidly fluctuating functions. Another reason is that aliasing errors can be isolated

![Fig. 1. Second-order finite difference weighting factor $\sigma_k^{\text{fd}}$ for differentiation of the model function given in Appendix B. This plot shows how the finite difference operator acts like a low-pass filter. Note the attenuation of the high frequencies.](image)

4. Generation of a Stationary Medium

Many researchers have investigated the behavior of flow and transport by solving the relevant equations in a numerically simulated conductivity field. The methods used to generate each realization of the simulated conductivity field are designed to preserve certain statistics, usually the mean and covariance function. A widely used method in the hydrology literature is known as the turning bands method [Mantoglou and Wilson, 1982] which was employed by Ahabou et al. [1989] to generate the three-dimensional conductivity field used in their numerical experiment. In the companion paper an extensive numerical study is made on the effective conductivity of stationary lognormally distributed conductivity fields. In this section it will be shown how a realization of such a random field can be efficiently produced from a covariance function [Newland, 1984]. For the case of a finite domain the covariance function is defined

$$R(\xi) = \frac{1}{V} \int_V Y(x) Y(x + \xi) \, dx - m_Y^2$$  \hspace{1cm} (39)

where $Y(x) = \ln K(x)$ and $m_Y$ is the mean of $Y$. Expanding $Y$ in a Fourier series,

$$Y(x) = \sum_{k} \hat{Y}_k e^{i2\pi k \cdot x}$$  \hspace{1cm} (40)

and replacing $Y$ with this series expansion in the definition for $R$ we have

$$R(\xi) = \sum_{k} \sum_{m} \hat{Y}_k \hat{Y}_m e^{i2\pi m \cdot \xi} \frac{1}{V} \int_V e^{i2\pi (k - m) \cdot x} \, dx - m_Y^2$$

Using the orthogonality property (A7),
\[ R(\xi) = \sum_{k \neq 0} \hat{Y}_{-k} \hat{Y}_k e^{i2\pi k \cdot \xi} \quad (42) \]

Since \( Y(x) \) is a real function, \( \hat{Y}_{-k} = \hat{Y}_k^* \):

\[ R(\xi) = \sum_{k \neq 0} S_k e^{i2\pi k \cdot \xi} \quad (43) \]

where

\[ S_k = \hat{Y}_k \hat{Y}_k^* \quad (44) \]

is called the power spectrum of \( \hat{Y}_k \). Equation (43) is a Fourier series with expansion coefficients given by the power spectrum. The coefficients are given by the usual formula for Fourier series:

\[ S_k = \frac{1}{V} \int_V R(\xi) e^{-i2\pi k \cdot \xi} \, d\xi \quad (45) \]

The above derivation suggests a way to obtain a realization of \( K(x) \) from \( R(\xi) \). First, Fourier transform \( R(\xi) \) to get \( S_k \), as indicated by (45). From (44), the magnitude of \( Y_k \) is

\[ |\hat{Y}_k| = (S_k)^{1/2} \]

Any complex number, such as \( \hat{Y}_k \), can be decomposed into a magnitude and a phase, \( \hat{Y}_k = r_k e^{i\theta_k} \), where \( r_k \) is the magnitude and \( \theta_k \) is the phase. Since the covariance function only specifies the magnitude of \( \hat{Y}_k \) and gives no phase information, the trick is to add a random phase such that \( \hat{Y}_{-k} = \hat{Y}_k^* \), which ensures that the realization \( Y \) is real valued. To summarize, \( \hat{Y}_k = (S_k)^{1/2} e^{i\theta_k} \), where \( \theta_k \) is a random phase with the stipulation that \( \theta_k = -\theta_{-k} \).

The realization of \( Y \) is then found by evaluating (40). The conductivity field is then simply \( K = \exp(Y) \). The mean of \( Y \) is an arbitrary constant and for the work in this paper it is set to zero.

To implement this algorithm numerically we need to evaluate (40) and (45) numerically. Equation (45) is a Fourier transform and can be evaluated at the collocation points given by

\[ x_i = \frac{i}{N_x} \quad i = 1, \ldots, N_x - 1 \quad (46) \]

using the FFT. The continuous Fourier series coefficients \( S_k \) and \( \hat{Y}_k \) are then replaced by the discrete Fourier series coefficients \( S_k \) and \( \hat{Y}_k \). Equation (40) at the collocation points is an inverse discrete Fourier and can be evaluated using the inverse FFT. In this way, hydraulic conductivity realizations can be generated in order \( N_x^2 \ln(N_x^2) \) operations. Also, since this algorithm is entirely numerical, it can handle any legitimate covariance function.

The power spectra of the covariance functions used in this paper can be found analytically, so in the interest of computational speed, the conductivity realizations were actually generated from the power spectrum, thus requiring one less FFT. The power spectrum of the Gaussian covariance function given by (24) in the companion paper is

\[ S_k = \sigma^2 e^{i(\pi/2) \sigma^2 \xi^2} \quad (47) \]

where \( k^2 = \sum_{j=1}^n k_j^2 \). The power spectrum of the exponential covariance function given by (23) in the companion paper is

\[ S_k = \frac{2\pi e^{i\sigma^2}}{(1 + 4\pi^2 \sigma^2 k^2)^{3/2}} \quad n = 2 \quad (49) \]

It is worth clarifying one point about this method. A realization is generated from a periodic covariance function. This results in correlation between two points that are decreasing for separation distances up to \( \xi = \frac{1}{2} \) and then increasing until the points are perfectly correlated at unit separation \( \xi = 1 \). One could avoid this problem, if deemed necessary, by only using a section of the random field for which no two points are separated by a distance greater than \( \frac{1}{2} \). A typical realization of the entire field is shown in Figure 2.

5. Test Cases and Comparison with Finite Difference Solution

The spectral method was tested on several cases for which an analytic solution is available. In one dimension, for example, the effective conductivity is always the harmonic mean. For comparison, it is interesting to see how well second-order finite differences and block-centered finite differences using the harmonic mean perform. The second-order finite difference discretization of (9) in one dimension is

\[ (-K_{j+1} + 4K_j + K_{j-1})g_{j-1} - 8K_j g_j + (K_{j+1} + 4K_j) \]

\[ -K_{j-1}g_{j+1} = 2\delta x(K_{j+1} - K_{j-1}) \quad (50) \]

at the sampling point locations given in (25) and \( \delta x = 1/N_K \). The block-centered finite difference discretization of (9) in one dimension is

\[ K_{j-1/2}g_{j-1} - (K_{j-1/2} + K_{j+1/2})g_j + K_{j+1/2}g_{j+1} = \delta x(K_{j+1/2} - K_{j-1/2}) \quad (51) \]
where the harmonic means of the adjacent nodal conductivities are used for the midnodal conductivities:

$$k_{j+1/2} = \frac{2k_j k_{j+1}}{k_j + k_{j+1}}$$  \hspace{1cm} (52)

The boundary conditions and the conductivity are periodic, so \( g_0 = g_{N_x} \) and \( K_0 = K_{N_x} \) in (50) and (51). The above finite difference discretization both yield systems of linear equations \( A_{g,x} = b_{g,x} \). Equation (10) was integrated using the rectangular rule of integration [Press et al., 1988], giving

$$K^e = K - \frac{1}{2} \sum_{j=0}^{N_x-1} K_j (g_j + 1 - g_{j-1})$$  \hspace{1cm} (53)

In one dimension a major simplification can be made in the governing equation (9) by noting that we need only solve for the gradient of \( g \) for use in (10). This simplification was not made for either the spectral or finite difference method since no analogous simplification exists for the multidimensional case. The basis of comparison between the spectral and finite difference schemes is the number of nodes \( g_j \) used to sample the function \( g(x) \), because this is the number which determines the size of the respective matrices \( A \) and \( A_{g,x} \) and hence the eventual cost in three-dimensional problems. In one dimension the number of unknowns is relatively small, so the spectral and finite difference systems of equations are both easily solved with standard direct methods.

The test case is the conductivity given by the function

$$K(x) = \exp \{ \sin (2\pi f_1 x) + 2 \cos (2\pi f_2 x)$$

$$+ \sin (2\pi f_3 x + 1/3) \}$$  \hspace{1cm} (54)

where the frequencies used are \( f_1 = 2, f_2 = 8, \) and \( f_3 = 32 \). A plot of \( K(x) \) is shown in Figure 3. (For the spectral equation the test function was sampled with \( N_x = 1024 \) points.) Figure 4 shows the effective conductivity computed using the three methods as a function of the number of sampling points of \( g(x) \). It is evident from this example that the spectral discretization requires far fewer sampling points to obtain an answer of a given accuracy. If we compare the number of sampling points needed to get within 5% of the harmonic mean, the block-centered finite difference solution needs a factor of 3.7 times more points than the spectral solution while second-order finite differences require a factor of 8 times more points. Also note the relatively rapid convergence of the spectral method to the analytic solution compared with the extended tail on the finite difference solution. This illustrates the benefits of spectral accuracy compared with the second-order accuracy of both finite difference methods.

The final test case for the spectral method is the perfectly stratified medium shown in Figure 5. In this case the effective conductivity is the arithmetic mean in the direction parallel to the bedding and the harmonic mean in the direction perpendicular to the bedding. In the companion paper we will be interested in finding the effective conductivity of two-dimensional bimodal conductivity distributions composed of shale and sandstone, so it is useful to see how the method behaves in a simple case where the answer is known. Numerically, a bimodal conductivity distribution is a particularly challenging case because it is a discontinuous function. Representing a discontinuous function using a continuous one, such as the trigonometric polynomials in the Fourier Galerkin method, leads to certain difficulties. For the Fourier method, the Gibbs phenomenon describes the classic oscillatory behavior of the truncated Fourier series

---

The figure on the right shows a comparison between the spectral method, second-order finite differences, and block-centered finite differences using the harmonic mean for the test function shown in Figure 3.
near a point of discontinuity. The Gibbs phenomenon influences the behavior of the Fourier series over the entire periodic domain, not only near the point of discontinuity. Unlike the case of an infinitely differentiable function with periodic derivatives, where spectral accuracy is achieved, it can be shown for the discontinuous case [Canuto et al., 1988] that the kth coefficient of the Fourier series decays like k⁻¹.

This case was solved as a two-dimensional problem using the numerical spectral algorithm described. The conductivity of the sandstone was normalized to $K_{ss} = 1$, and the conductivity of the shale $K_{sh}$ was varied to obtain three different conductivity contrasts of $K_{ss}/K_{sh} = 10$, 50, and 100. The medium was sampled on a grid with 512 by 512 points, and in the frequency domain we used $N_K = 120$ and $N_y = 60$. The algorithm easily found the solution, in the direction parallel to the layering in one iteration, to be the arithmetic mean to almost machine precision. In the direction perpendicular to the layering the percent errors were 2.5, 5, and 6% for contrasts of 10, 50, and 100, respectively.

Since the Gibbs phenomenon is associated with the slow decay of the Fourier coefficients of a discontinuous function, it was thought that low-pass filtering the conductivity to attenuate the high-frequency coefficients might help the conjugate gradient solution. Filtering $K(x)$ with the raised cosine filter given by (36) was found not to help by either enhancing the rate of convergence or improving the accuracy. In fact, to get the same level of accuracy with filtering, more frequency components had to be included.

6. DISCUSSION

One of the numerical difficulties in solving (16) for an extremely heterogeneous medium is that the condition number of the matrix A can become large. Using direct solution methods, the computer can lose log [cond (A)] decimal places to round off error [Strang, 1988]. In addition, with iterative solution techniques such as conjugate gradients, a large condition number results in a slower rate of convergence to the solution [Golub and Van Loan, 1989]. The condition number of A can be reduced, and hence the rate of convergence increased, by using preconditioning techniques [Wong et al., 1986; Meijerink and Van Der Vorst, 1981]. Although preconditioners could probably enhance the numerical method described in this paper, they were not used in this study because of the difficulty in their implementation in three dimensions. It is interesting to look at the condition number for our system of equations. For a Hermitian matrix the eigenvalues are real and the condition number is defined as

$$\text{cond} (A) = \frac{e_{\text{max}}}{e_{\text{min}}} \quad (55)$$

where $e_{\text{min}}$ is the smallest eigenvalue and $e_{\text{max}}$ is the largest eigenvalue. The behavior of the condition number was studied, in one dimension, for the two scenarios investigated in the companion paper, a bimodal and lognormally distributed conductivity. Figure 6 shows how the condition number increases as a function of the conductivity contrast, where the bimodal conductivity distribution is given by Figure 5. Figure 7 shows how the condition number as a function of $\sigma^2$ for a random conductivity field generated using a Gaussian covariance function with $L/\ell = 40$. The condition number increases proportional to about $(\sigma^2)^{1/3}$. The plot also shows that the condition number is a function of the number of frequencies, $N_K$, used. For the Fourier Galerkin method the condition number increases in proportion to $N_y^2$ [Wong et al., 1986]. It is evident that computations are best performed in double precision to assure that enough significant decimal places remain when the spatial variations in $K(x)$ are large.

Recall that the governing equation (5) is a nonlinear system due to the nonconstant coefficients (not to be confused with linear partial differential equations). Unlike a linear system which can only modify the amplitude and phase of the input spectrum (such as the small perturbation method), a nonlinear system can, in addition, create new frequencies. The convolution sums of (16) show that each frequency of $\tilde{g}_{i}$ depends on all the frequency components of the conductivity field $\tilde{K}_{i}$. To get an idea of the behavior of our particular nonlinear system, (16) was solved directly for a simple one-dimensional case. The variability of the con-

Fig. 6. Condition number of matrix A as a function of the conductivity contrast $K_{ss}/K_{sh}$ for the bimodal distribution shown in Figure 5. The matrix A has $N_K = 160$ and $N_y = 80$.

Fig. 7. Condition number of matrix A as a function of $\sigma^2$ for $N_K = 160$, $N_y = 80$ and $N_K = 80$, $N_y = 40$. The one-dimensional matrix A was generated using a Gaussian covariance function with $L/\ell = 40$. 
conductivity field was composed of just two different frequencies:

\[ K(x) = 3 + \sin(2\pi f_1 x) + \sin(2\pi f_2 x) \]  

(56)

where the frequencies are \( f_1 = 13 \) and \( f_2 = 15 \). Figure 8 shows the power spectrum of \( g(x) \). Not only are the original frequencies present but so are all the possible harmonics: \( m - f_1, f_2, f_1 \pm f_2, 2f_1, 2f_2, 2f_1 \pm 2f_2, \) etc. It is particularly interesting to note that the two “high” frequencies \( f_1 \) and \( f_2 \) yielded a very potent “low” frequency component at \( f_2 - f_1 = 2 \), with more than six times the energy than is at either of the original frequencies. Different frequencies can interact to shift energy up and down the frequency spectrum in significant amounts. From a physical point of view this indicates that the interaction between different scales of variability can potentially be very significant.

7. Conclusion

A numerical spectral technique known as the Fourier Galerkin method was applied to the problem of determining the effective hydraulic conductivity in a rigid, saturated porous medium. The equations for the effective conductivity under the assumptions that \( K(x) \) is periodic and the flow is gradually varying have been derived previously by Kitidis [1990]. The numerical method is efficient in that the solution can be achieved in order \( N_f \ln(N_f) \) operations and order \( N_f \) storage, where \( N_f \) is the total number of sampling points used in a flow domain of dimension \( n \). The efficiency of the method stems from decomposing the matrix-vector multiplication in the conjugate gradient algorithm into \( n \) convolution sums which can be performed efficiently using FFTs. However, this well-documented practice [Canuto et al., 1988] was found not to work with severely heterogeneous media, so a modified version of the standard FFT convolution was devised.

The spectral method was tested for several cases for which an analytic solution was available and found to give accurate results even when the spatial variations in \( K(x) \) were quite large. For these test cases the spectral results were also compared to a second-order finite difference solution. The finite difference method required about four times more nodes than the spectral method to give the same accuracy. It was also shown that the condition number of the spectral coefficient matrix increases with increasing spatial variability. In one dimension the condition number was proportional to \( (\sigma^2)_{ij}^{1/2} \) for a conductivity field generated from a Gaussian covariance function. This implies that for large \( \sigma^2 \), numerical solutions should be performed in double precision.

Appendix A

Rewriting (9) by expanding the derivative, we have

\[ \sum_{\rho=1}^{n} \left[ \frac{\partial K}{\partial x_p} \frac{\partial g^i}{\partial x_p} + K \frac{\partial^2 g^i}{\partial x_p^2} \right] \frac{\partial K}{\partial x_i} = 0 \quad i = 1, \ldots, n \]

(A1)

For notational simplicity, let \( L_1 = L_2 - L_3 - 1 \) so \( f = k \). It is straightforward at the end to remove this restriction by dividing the wave number component by its corresponding periodic length. Applying a Fourier transform to (A1) by substituting equations (14) and (15), we obtain the corresponding frequency domain equation:

\[ \sum_{\rho=1}^{n} \left[ \sum_{k=0}^{n} 12\pi k_p \hat{K}_k e^{i2\pi k x \cdot k} \sum_{l=0}^{n} 12\pi l_p \hat{g}^i e^{i2\pi k x \cdot l} \right] + \sum_{k=0}^{n} \hat{K}_k e^{i2\pi x \cdot k} \sum_{l=0}^{n} (12\pi l_p)^2 \hat{g}^i e^{i2\pi x \cdot l} \]  

\[ - \sum_{k=0}^{n} 12\pi k_p \hat{K}_k e^{i2\pi x \cdot k} = R(x) \]

(A2)

where it is implied in (A2) that summations over indices representing wave numbers cover the range of values \( k = -N_f/2, \ldots, N_f/2 \) and \( i = 1, \ldots, n \). Since (14) and (15) are truncated series representations, the governing equation is not exactly satisfied except in the limit as \( N_f \to \infty \). The residual is defined to be

\[ R(x) = -\sum_{\rho=1}^{n} \left[ \frac{\partial K}{\partial x_p} \frac{\partial g^i}{\partial x_p} + K \frac{\partial^2 g^i}{\partial x_p^2} \right] \frac{\partial K}{\partial x_i} \]

(A3)

Factoring out the \( k = 0 \) term in the second summation over \( k \) gives

\[ -4\pi^2 \sum_{\rho=1}^{n} \left[ \sum_{k=0}^{n} k_p \hat{K}_k e^{i2\pi k x \cdot (k+1)} \right] \]  

\[ + \sum_{k=0}^{n} \sum_{l=0}^{n} l_p^2 \hat{K}_k e^{i2\pi x \cdot l} e^{i2\pi x \cdot (k+1)} + \hat{K}_0 \sum_{l=0}^{n} l_p^2 \hat{g}^i e^{i2\pi x \cdot l} \]  

\[ - \sum_{k=0}^{n} 12\pi k_p \hat{K}_k e^{i2\pi x \cdot k} = R(x) \]

(A4)

Then, combining terms,

\[ -4\pi^2 \sum_{\rho=1}^{n} \left[ \sum_{k=0}^{n} (k+1)^2 \hat{K}_k e^{i2\pi x \cdot (k+1)} \right] \]  

\[ + \sum_{k=0}^{n} \sum_{l=0}^{n} (k+1)^2 \hat{K}_k e^{i2\pi x \cdot (k+1)} \]  

\[ - \sum_{k=0}^{n} 12\pi k_p \hat{K}_k e^{i2\pi x \cdot k} = R(x) \]

(A4)
\[ + \hat{K}_0 \sum_{l=0}^{\infty} \hat{f}_l e^{i2\pi x \cdot k} \left\{ \sum_{k=0}^{\infty} 12\pi k \hat{K}_4 e^{i2\pi x \cdot k} = R(x) \right\} \]

The weighted residual condition states that we want to minimize the residual in the sense that the weighted residual tends to zero over the domain of interest. In the Galerkin scheme the weight functions are the complex conjugates of the basis functions. The method of weighted residuals condition is then

\[ \int_V R(x)e^{-i2\pi x \cdot m} \, dx = 0 \]  

(A6)

where, again, \( m = -N/2, \cdots, N/2 \) and \( i = 1, \cdots, n \).

This integral can be evaluated using the orthogonality condition of the trigonometric polynomials:

\[ \int_V e^{i2\pi x \cdot k} e^{-i2\pi x \cdot m} \, dx = \delta_{km} \]  

(A7)

Using (A5) in (A3), replacing that into (A6), and employing the orthogonality condition, we have

\[ 0 = 12\pi m_i \hat{K}_m + 4\pi^2 \sum_{j=1}^{n} \left\{ \hat{K}_0 m_j^2 \delta_{jm} + 4\pi^2 \sum_{\substack{l=0 \atop m \neq 1}}^{n} \frac{m_j}{L_j} K_m - \hat{f}_l \right\} \]  

(A8)

Noting that the last two terms can be combined and reintroducing the periodic lengths \( L_i \), we obtain the desired result:

\[ \frac{1}{2} \frac{1}{L_i} K_m = \pi \sum_{j=1}^{n} \frac{m_j}{L_j} K_m - \hat{f}_l \]  

(A9)

**APPENDIX B**

This appendix compares the accuracy with which spectral methods and second-order finite differences can differentiate. The model problem for a rapidly fluctuating spatial function is, in one dimension, the sinusoid:

\[ f(x) = e^{i2\pi x \cdot k_0}, \quad 0 \leq x \leq 1 \]  

(B1)

The domain is discretized with \( N \) sample points \( x_i = i/N \) for \( i = 0, 1, \cdots, N - 1 \), so \( f(x_i) = \exp(i2\pi ki/N) \). The derivative of \( f \) is first found using second-order finite differences:

\[ \frac{df}{dx} \bigg|_{x_i} = \frac{1}{2\Delta x} [f(x_{i+1}) - f(x_{i-1})] \]  

(B2)

\[ = \frac{1}{2\Delta x} [e^{i2\pi k/N} - e^{-i2\pi k/N} - e^{i2\pi k/N} e^{-i2\pi k/N}] \]  

(B3)

\[ = \frac{1}{2\Delta x} [e^{i2\pi k/N} - e^{-i2\pi k/N}] e^{i2\pi k/N} \]  

(B4)

\[ = \frac{1}{2\Delta x} [e^{i\pi k/N} - e^{-i\pi k/N}] f(x_i) \]  

(B5)

where \( \Delta x = 1/N \). Using the identity \( \sin(x) = \frac{1}{2i} [\exp(i2\theta) - \exp(-i2\theta)] \), we have

\[ \frac{df}{dx} \bigg|_{x_i} = \frac{1}{\sin(2\pi k/N)} f(x_i) \]  

(B6)

For any value of the wave number between \( k = -N/2, \cdots, N/2 \) the spectral method is an exact differentiator:

\[ \frac{df}{dx} \bigg|_{x_i} = 12\pi k e^{i2\pi x \cdot k} = 12\pi k f(x_i) \]  

(B8)

The finite difference derivative includes a frequency-dependent weight factor \( \sigma^\text{Fd}_k \) defined by dividing (B7) by (B8), yielding

\[ \sigma^\text{Fd}_k = \frac{\pi}{2\Delta x} \sin \left( \frac{\pi k}{N} \right) \quad k = -\frac{N}{2}, \cdots, \frac{N}{2} \]  

(B9)

**Acknowledgments.** Funding for this study was provided by the Office of Research and Development, U.S. Environmental Protection Agency, under agreement R-815738-01 through the Western Region Hazardous Substance Research Center. Funding for this work has also been provided by the National Science Foundation under grant EET 87-16585. The content of this study does not necessarily represent the views of these agencies. The authors thank Thomas Van Lent of Stanford University for valuable assistance.

**REFERENCES**


B. B. Dykaar and P. K. Kitanidis, Department of Civil Engineering, Stanford University, Stanford, CA 94305.

(Received February 18, 1991, revised November 15, 1991; accepted December 9, 1991.)