Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Computational Physics 205 (2005) 706–718
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

A Fourier–Wachspress method for solving Helmholtz's equation in three-dimensional layered domains

Christopher R. Anderson [a,*], Thomas C. Cecil [b]

[a] UCLA Mathematics Department, University of California, P.O. Box 951555, Los Angeles, CA 90095-1555, USA
[b] ICES, University of Texas at Austin, 1 University Station, Austin, TX 78712, USA

Received 16 June 2004; received in revised form 10 November 2004; accepted 20 November 2004

Abstract

In this paper, we present a fast direct method for solving Poisson's or Helmholtz's equation in three-dimensional layered domains. The method combines a Fourier method for two dimensions and a variant of Wachspress's method in the third dimension. The resulting scheme is capable of efficiently creating solutions that are highly accurate even when the coefficients defining the layered structure are discontinuous or extreme mesh refinement is used.

MSC: 35JO5; 65T50
Keywords: Fourier method; Poisson equation; Helmholtz equation

1. Introduction

The problem this paper addresses is the computation of the solution of the Helmholtz equation

$$\nabla \cdot (a(z)\nabla\phi) + b(z)\phi = f(x,y,z) \tag{1}$$

in a two- or three-dimensional rectangular domain when the coefficients $a(z)$ and $b(z)$ are piecewise constant. One area where this problem arises, and that which motivates the present work, is in the modeling of layered semi-conductor devices [4,5]. This equation, with $a(z)$ being the dielectric constant and $b(z) \equiv 0$, is used to determine the electrostatic potential in the device. The solution of (1) is also used as a preconditioner for the iterative solution of the linear systems arising when inverse subspace iteration [13] is used to find the eigenvalues and eigenvectors of the one particle Schrödinger operator for the device (the latter

* Corresponding author. Tel.: +1 310 825 1298; fax: +1 310 206 2679.
E-mail addresses: anderson@math.ucla.edu (C.R. Anderson), tcecil@ices.utexas.edu (T.C. Cecil).

0021-9991/$ - see front matter
doi:10.1016/j.jcp.2004.11.028

equation being of the form (1), but with b a function of all three independent variables). A characteristic feature of the semi-conductor devices being modeled is the presence of very thin layers; layers that are one to two orders of magnitude thinner than the total device thickness. A sample device geometry is shown in Fig. 1.

Due to the nature of the coupling of the electrostatic potential and the Schrödinger operator, the solution of (1) is required only in a region about these thin layers. Thus, in order for the solution technique to be useful for modeling semi-conductors, it must be capable of creating solutions on a highly non-uniform mesh. Typical boundary conditions imposed on the solution of (1) consist of periodic conditions in the lateral (x–y) directions, Dirichlet conditions on the top, and Dirichlet, Neumann, or "infinite", boundary conditions on the bottom. By an "infinite" boundary condition, we mean the specification of a boundary condition that yields a solution in the computational domain that is the restriction of the solution in the infinite domain. In this paper, we describe an efficient, high order accurate, direct (non-iterative) method for solving (1). High order accuracy can be obtained even when the coefficients are discontinuous (with arbitrarily sized jumps) or when extreme mesh refinement is used.

One component of our procedure is the use of a discrete Fourier basis in the lateral (x–y) directions. In this regard, we are following the work of others [1,6–8,11,15] who have shown that very fast, high order accurate, methods for the solution of Poisson's equation can be accomplished using a discrete Fourier basis along with the fast Fourier transform to carry out the required change of basis. One of the principle features of recent work concerns the development of Fourier-based methods for problems with non-periodic boundary conditions. While, in this paper we only discuss the case of periodic boundary conditions in the lateral direction, the results in [1,6–8] could certainly be applied to create a method for handling more general lateral direction boundary conditions. When one uses a Fourier basis in the lateral directions, for the vertical (z-component) one is led to the problem of computing high order accurate solutions to linear two point boundary value problems with piecewise constant coefficients. The problem of creating highly accurate solutions to general linear two point boundary value problems has been well studied and methods that achieve spectral accuracy have been developed [9,12]. However, in the case of piecewise constant coefficients, one can extend an "early" patching [3] method described by Wachspress [16] to provide an efficient

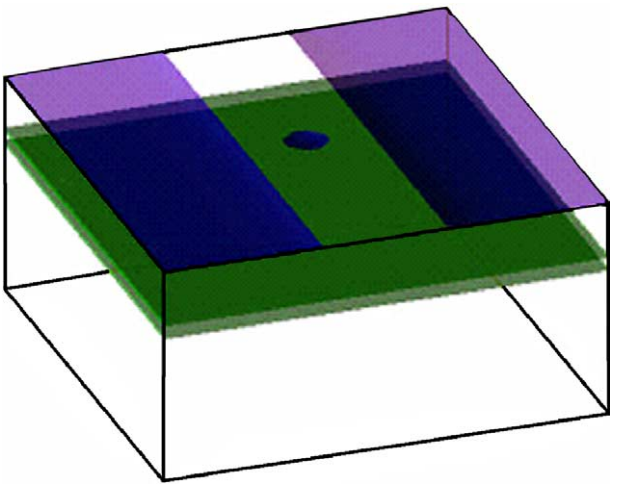

Fig. 1. Sample device geometry. Upper shaded region indicates location of applied potential (gates). Inner shaded region indicates material with non-bulk dielectric constant.

numerical procedure that can provide high accurate solutions on adaptive meshes and is thus ideally suited to the computational task.

In addition to the method's high accuracy, the method we present also has two other aspects that give it merit. First, the implementation of the method is a small modification of that which would be used to created a standard fast Helmholtz solver based upon a finite difference or finite volume discretization (e.g., use of FFT's for the lateral directions and tridiagonal solves for the vertical direction) and second, the use of a rectangular Cartesian grid allows one to easily construct computationally efficient implementations.

There exist numerous high order methods designed to work well despite the presence of discontinuous coefficients. Spectral penalty methods [10] and discontinuous Galerkin schemes [2] are designed to preserve spectral accuracy in spite of these discontinuities. These techniques are similar to ours in that they exploit mesh refinement, but they approach the boundary layer oscillations that can occur near the discontinuities by penalizing certain numerical coefficients, resulting in schemes that avoid the unwanted oscillations. Our method differs in that there is no penalty enforced to gain smoothness, but rather we solve the boundary layer problems exactly to avoid instabilities introduced by numerical approximation.

In Section 1, we outline the general numerical approach and summarize the computational steps. In Section 2, we review the Wachspress procedure and describe an extension of it that enables the procedure to be incorporated into the method we are presenting to solve (1). In Section 3, we present computational examples that demonstrate the accuracy and efficacy of our procedure.

2. Derivation of the computational method

We are concerned with the solution of

$$\nabla \cdot (a(z)\nabla\phi) + b(z)\phi = f(x,y,z) \tag{2}$$

for $(x,y,z) \in [0,L_x]\times[0,L_y]\times[0,L_z]$ with Dirichlet boundary conditions at $z = 0$,

$$\phi(x,y,0) = g(x,y), \quad (x,y) \in [0,L_x]\times[0,L_y],$$

Neumann, Dirichlet, or "infinite" boundary conditions at $z = L_z$, and $\phi(x,y,z)$ periodic in x and y for all $z \in [0,L_z]$.

The coefficients $a(z)$ and $b(z)$ are assumed to be piecewise constant and define the layered structure of the domain. Specifically, if $\{z_i\}_{i=1}^{P+1}$ is the partition of $[0,L_z]$ into the P intervals where both $a(z)$ and $b(z)$ are constant, e.g.,

$$a(z) = a_i \quad \text{for } z \in [z_i, z_{i+1}], \ \ i = 1,\ldots,P,$$
$$b(z) = b_i \quad \text{for } z \in [z_i, z_{i+1}], \ \ i = 1,\ldots,P$$

then we refer to the regions $[0,L_x]\times[0,L_y]\times[z_i,z_{i+1}]$ for $i = 1,\ldots,P$ as the layers of the domain.

Under the assumption on the coefficients and the rectangular nature of the domain, the first step in deriving the computational method is to use separation of variables to reduce the problem to that of solving of a collection of two- and one-dimensional problems. If a Fourier basis is used for the (x,y) dependence, then we seek solutions of the form

$$\phi(x,y,z) = \sum_{k_1,k_2} e^{2\pi i k_1 (x/L_x)} e^{2\pi i k_2 (y/L_y)} \gamma_{(k_1,k_2)}(z). \tag{3}$$

Formally, such a function will be a solution of (2) if the Fourier coefficient functions, $\gamma_{(k_1,k_2)}(z)$, satisfy

$$\frac{d}{dz}\left(a(z)\frac{d}{dz}\gamma_{(k_1,k_2)}(z)\right) + \left(b(z) - 4\pi^2 k_1^2\left(\frac{x}{L_x}\right)^2 - 4\pi^2 k_2^2\left(\frac{y}{L_y}\right)^2\right)\gamma_{(k_1,k_2)}(z) = \hat{f}_{(k_1,k_2)}(z), \tag{4}$$

where

$$\hat{f}_{(k_1,k_2)}(z) = \int_0^{L_x} \int_0^{L_y} f(x,y,z) e^{-2\pi i k_1 (x/L_x)} e^{-2\pi i k_2 (y/L_y)} \, dx \, dy.$$

A numerical method is obtained by using a finite number of Fourier basis functions and computing the solutions to (4) numerically at a finite number of z coordinate values $\{z_r\}_{r=1}^{N_z+1}$. The approximate solution to be computed, $\tilde{\phi}$, thus has the representation

$$\tilde{\phi}(x,y,z_r) = \sum_{k_1} \sum_{k_2} e^{2\pi i k_1 (x/L_x)} e^{2\pi i k_2 (y/L_y)} \tilde{\gamma}_{(k_1,k_2)r}, \tag{5}$$

where $\tilde{\gamma}_{(k_1,k_2)r}$ is the approximate solution $\tilde{\gamma}_{(k_1,k_2)}(z)$ evaluated at z_r and $-[N_x/2] \leqslant k_1 \leqslant [N_x/2]$ and $-[N_y/2] \leqslant k_2 \leqslant [N_y/2]$.

In order to create the requisite approximate solutions of (4), it is necessary to have a means of computing $\hat{f}_{(k_1,k_2)}(z_r)$. These function values are efficiently evaluated by using a two-dimensional fast Fourier transform (FFT). Specifically, for a given value $z_r \in \{z_j\}_{j=1}^{N_z+1}$, the function $f(x,y,z_r)$ is evaluated at the nodes of a uniform mesh in the x–y plane (e.g., a mesh with N_x panels in the x-direction and N_y panels in the y-direction). The application of the forward FFT to these values yields the values of $\hat{f}_{k_1,k_2}(z_r)$ for $-[N_x/2] \leqslant k_1 \leqslant [N_x/2]$ and $-[N_y/2] \leqslant k_2 \leqslant [N_y/2]$. Similarly, once the approximate solutions $\tilde{\gamma}_{(k_1,k_2)}(z_r)$ have been computed, the evaluation of $\tilde{\phi}$ at z_r using (5) can be accomplished by applying the inverse FFT to the values of $\tilde{\gamma}_{(k_1,k_2)}(z_r)$. The result yields the values $\tilde{\phi}(x_p,y_q,z_r)$ with $x_p = p(L_x/N_x)$ and $y_p = p(L_y/N_y)$.

For each set of values (k_1,k_2), Eq. (4) that determines the Fourier coefficients $\gamma_{(k_1,k_2)}(z)$ is a linear two-point boundary value problem with piecewise constant coefficients. As mentioned in Section 1, there are efficient high order methods for solving general linear two-point boundary value problems [9,12]. However, an extension of an "early" method due to Wachspress that takes advantage of the piecewise constant coefficient nature of the coefficients yields exact solutions in the case when $\hat{f}_{(k_1,k_2)}(z) \equiv 0$, and an accuracy when $\hat{f}_{(k_1,k_2)}(z) \neq 0$ that is only limited by the accuracy that $\hat{f}_{(k_1,k_2)}(z)$ can be approximated by polynomials over each layer of the domain. The linear system that must be solved for this solution is always tridiagonal, and hence does not suffer from the increased bandwidth problems associated with standard high order finite difference methods. Before we describe this method (in Section 3), we give a summary of the computational steps of the procedure for solving (2).

(1) Choose the points $\{z_j\}_{j=1}^{N_z+1}$ that define the nodes of the computational mesh in the z-coordinate direction. This set of points must minimally contain the points where the coefficients $a(z)$ and $b(z)$ are discontinuous. Choose the number of panels N_x and N_y that determine the computational mesh used in the x–y plane.
(2) For each $z_r \in \{z_j\}_{j=1}^{N_z+1}$ evaluate $f(x,y,z_r)$ at the nodes of the x–y plane computational mesh. Apply the forward FFT to obtain, $\hat{f}_{(k_1,k_2)}(z_r)$, the right-hand sides of the two-point boundary-value problems (4).
(3) Using the method described in Section 2, create a high order numerical approximation to $\gamma_{(k_1,k_2)}(z)$ at the points $\{z_j\}_{j=1}^{N_z+1}$.
(4) For each $z_r \in \{z_j\}_{j=1}^{N_z+1}$ apply the inverse FFT to the values $\tilde{\gamma}_{(k_1,k_2)}(z_r)$ and thus obtain a high order approximation to ϕ at the nodes of mesh that is the tensor product of the mesh defined by the points $\{z_j\}_{j=1}^{N_z+1}$ and a uniform mesh in the x–y plane.

Since there are $N_z + 1$ points in the set $\{z_j\}$, then the method requires $N_z + 1$ applications of the forward and inverse two-dimensional FFT and the solution of $N_x N_y$ tridiagonal systems of equations of size $N_z + 1$; thus the operation count for the method is formally $O(N_x N_y N_z \log(N_x N_y)) + O(N_x N_y N_z)$ operations. While a detailed mathematical analysis has not yet been performed, computational experiments indicate that the order of accuracy is dictated by the order of accuracy of the numerical procedure for solving Eq. (4) (this is

to be expected since we are using a Fourier basis in the x–y coordinate directions). As will be discussed below, the accuracy with which equations of the form (4) can be solved is limited only by the smoothness of the right-hand side in each layer; thus for problems where $f(x, y, z)$ possesses sufficiently many derivatives so that $\hat{f}_{(k_1,k_2)}(z)$ is smooth within each layer, the method can be used to compute spectrally accurate solutions.

3. Wachspress's method revisited

The one-dimensional problem that must be solved for the Fourier coefficients has the form

$$a_k \frac{\mathrm{d}^2 u}{\mathrm{d}z^2} + b_k u = f(z), \quad z \in [z_k, z_{k+1}], \quad k = 1, \ldots, P-1 \tag{6}$$

with Dirichlet, Neumann, or "infinite" boundary conditions at $z = z_1$ and $z = z_P$. We assume that $a_k > 0$ and $b_k < 0$ (other cases are handled similarly). The intervals $[z_k, z_{k+1}]$ are intervals over which the coefficients a_k and b_k are constant (see Fig. 2). These intervals need not be the largest intervals over which a_k or b_k have a particular constant value; all that is required is that the set of interval endpoints contains those points where the coefficients are discontinuous.

We first consider the case when $f(z) \equiv 0$. Wachspress's method [16] is based upon the observation that an exact solution of (6) can be created by combining locally exact solutions, $u_k(z)$, defined as solutions of (6) over each interval $[z_k, z_{k+1}]$. In order for the locally exact solutions to combine to create a global solution they must be continuous at the internal interval boundaries:

$$u_{k-1}(z_k) = u_k(z_k), \quad k = 2, \ldots, P-1 \tag{7}$$

and their derivatives must satisfy the conditions

$$a_{k-1} \left.\frac{\mathrm{d}u_{k-1}}{\mathrm{d}z}\right|_{z_k} = a_k \left.\frac{\mathrm{d}u_k}{\mathrm{d}z}\right|_{z_k}, \quad k = 2, \ldots, P-1. \tag{8}$$

If Dirichlet or Neumann boundary conditions are specified, then at the endpoints, z_1 and z_P, the values and/or the derivatives of the locally exact solutions u_1 and/or u_P are chosen to satisfy the prescribed boundary conditions. For problems on infinite or semi-infinite domains (e.g., when "infinite" boundary conditions are specified) one requires that (7) and (8) also hold when $k = 1$ and $k = P$ with $u_0(z)$ and $u_P(z)$ being bounded solutions of (6) over the semi-infinite intervals $-\infty < z \leqslant z_1$ and $z_P \leqslant z \leqslant \infty$, respectively. As discussed in [3], this has come to be known as "patching," which falls into the category of a domain decomposition

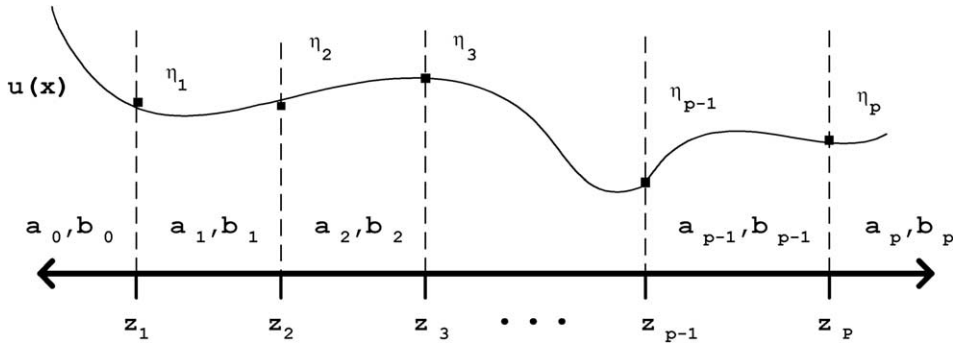

Fig. 2.

method. In [3], a thorough discussion of the general framework is given. Here, we focus on the solution method of patching together locally exact solution, with the particular solutions arising from the use of polynomial approximations to the right-hand side forcing term.

There are a variety of ways to formulate equations determining the locally exact solutions that satisfy (7) and (8). The formulation we present was selected because it results in a symmetric tridiagonal system of equations and can be easily extended to the case when $f(z) \neq 0$. We first assume Dirichlet boundary conditions and let η_k denote the function values of the global solution at the endpoints of the intervals, e.g., $\eta_k = u(z_k)$ for $k = 1, \ldots, P$. These values are well-defined since the solution is continuous at the interval boundaries. Over each interval, $[z_k, z_{k+1}]$, we construct the locally exact solutions, $u_k(z)$, in terms of these values;

$$u_k(z) = \eta_k \left[\frac{e^{\gamma_k(z-z_k)} - e^{\gamma_k h_k} e^{\gamma_k(z_{k+1}-z)}}{1 - e^{2\gamma_k h_k}} \right] + \eta_{k+1} \left[\frac{e^{\gamma_k(z_{k+1}-z)} - e^{\gamma_k h_k} e^{\gamma_k(z-z_k)}}{1 - e^{2\gamma_k h_k}} \right], \tag{9}$$

where $h_k = z_{k+1} - z_k$ and $\gamma_k = -\sqrt{|b_k/a_k|}$. With this construction, the locally exact solutions automatically satisfy the continuity condition (7). The derivative conditions (8) determine the equations that the solution values η_k must satisfy. Specifically, at interior interval boundary points (8) implies

$$\left[a_{k-1}\gamma_{k-1} \frac{2e^{\gamma_{k-1}h_{k-1}}}{(1 - e^{2\gamma_{k-1}h_{k-1}})} \right] \eta_{k-1} - \left[a_{k-1}\gamma_{k-1} \frac{(1 + e^{2\gamma_{k-1}h_{k-1}})}{(1 - e^{2\gamma_{k-1}h_{k-1}})} + a_k\gamma_k \frac{(1 + e^{2\gamma_k h_k})}{(1 - e^{2\gamma_k h_k})} \right] \eta_k + \left[a_k\gamma_k \frac{2e^{\gamma_k h_k}}{(1 - e^{2\gamma_k h_k})} \right] \eta_{k+1} = 0 \tag{10}$$

for $k = 2, \ldots, P - 1$. These equations, together with the specification of $\eta_1 = u(z_1)$ and $\eta_P = u(z_P)$ constitute a linear, symmetric, tridiagonal, positive definite system of equations for the solution values η_k. This system of equations can be solved using a standard tridiagonal solver to obtain the solution of (6) at the interval endpoints $\{z_k\}_{k=1}^{P}$. If one requires the solution at points $z \in [z_k, z_{k+1}]$ one merely evaluates $u_k(z)$ using the formula (9) and the computed values of η_k and η_{k+1}.

The formulas (9), (10) and others described below are susceptible to significant round-off errors when $\gamma_k h_k$ is small (e.g., computing $1 - e^{2\gamma_k h_k}$ directly suffers from "catastrophic cancellation"). However, this difficulty can be avoided by replacing the troublesome expressions by Taylor series expansions about $\gamma_k h_k = 0$.

In the case of Neumann boundary conditions, the values η_1 and η_P are additional unknowns, and one adjoins to the system (10) two equations that are obtained by requiring the local solution over the intervals $[z_1, x_2]$ and $[z_{P-1}, z_P]$ satisfy the Neumann boundary conditions

$$\begin{aligned} &\left[a_1\gamma_1 \frac{2e^{\gamma_1 h_1}}{(1 - e^{2\gamma_1 h_1})} \right] \eta_1 - \left[a_1\gamma_1 \frac{(1 + e^{2\gamma_1 h_1})}{(1 - e^{2\gamma_1 h_1})} \right] \eta_2 = \frac{du}{dz}(z_1), \\ &\left[a_{P-1}\gamma_{P-1} \frac{2e^{\gamma_{P-1}h_{P-1}}}{(1 - e^{2\gamma_{P-1}h_{P-1}})} \right] \eta_{P-1} - \left[a_{P-1}\gamma_{P-1} \frac{(1 + e^{2\gamma_{P-1}h_{P-1}})}{(1 - e^{2\gamma_{P-1}h_{P-1}})} \right] \eta_P = \frac{du}{dz}(z_P). \end{aligned} \tag{11}$$

For the "infinite" boundary condition case, the locally exact solutions $u_0(z)$ and $u_P(z)$ are given by

$$\begin{aligned} u_0(z) &= \left[e^{\gamma_0(z_0 - z)} \right] \eta_0, \\ u_P(z) &= \left[e^{\gamma_P(z - z_P)} \right] \eta_P \end{aligned}$$

and the derivative condition when $k = 1$ and $k = P$ gives rise to the equations

$$\begin{aligned} &- \left[a_0\gamma_0 + a_1\gamma_1 \frac{(1 + e^{2\gamma_1 h_1})}{(1 - e^{2\gamma_1 h_1})} \right] \eta_0 + \left[a_1\gamma_1 \frac{2e^{\gamma_1 h_1}}{(1 - e^{2\gamma_1 h_1})} \right] \eta_1 = 0, \\ &\left[a_{P-1}\gamma_{P-1} \frac{2e^{\gamma_{P-1}h_{P-1}}}{(1 - e^{2\gamma_{P-1}h_{P-1}})} \right] \eta_{P-1} - \left[a_{P-1}\gamma_{P-1} \frac{(1 + e^{2\gamma_{P-1}h_{P-1}})}{(1 - e^{2\gamma_{P-1}h_{P-1}})} + a_P\gamma_P \right] \eta_P = 0. \end{aligned}$$

These equations together with Eq. (10) for $k = 2, \ldots, P-1$ also comprise a linear, symmetric, tridiagonal set of equations for the solution values η_k.

The above construction requires only O(P) work and yields an exact solution. Typically one requires the solution at equispaced points inside the intervals where the constants have a particular value. In this case, one has the choice of either using the formula (9) or introducing extra intervals whose endpoints are the locations where the solution is required. Since the evaluation of exponential functions is relatively expensive, we found it computationally more efficient to introduce new intervals. The extra computational work is only algebraic in nature since only the size of the tridiagonal matrix is increased.

One can extend the above construction to the case of non-homogeneous equations. Again, assume Dirichlet boundary conditions. The first step in constructing a solution is to find a polynomial approximation $g_k(z)$ to $f(z)$ for $z \in [z_k, z_{k+1}]$. This approximating polynomial can be found in any number of ways; in our computational examples we simply use a local interpolation polynomial of modest degree. Let $\alpha_k(z)$ be an exact solution to

$$a_k \frac{\mathrm{d}^2\alpha_k}{\mathrm{d}z^2} + b_k\alpha_k = g_k(z), \quad \alpha_k(z_k) = \alpha_k(z_{k+1}) = 0 \tag{12}$$

(one method for doing this is given below). We then define $u_k(z)$ for $z \in [z_k, z_{k+1}]$ by

$$u_k(z) = \alpha_k(z) + \eta_k\left[\frac{\mathrm{e}^{\gamma_k(z-z_k)} - \mathrm{e}^{\gamma_k h_k}\mathrm{e}^{\gamma_k(z_{k+1}-z)}}{1-\mathrm{e}^{2\gamma_k h_k}}\right] + \eta_{k+1}\left[\frac{\mathrm{e}^{\gamma_k(z_{k+1}-z)} - \mathrm{e}^{\gamma_k h_k}\mathrm{e}^{\gamma_k(z-z_k)}}{1-\mathrm{e}^{2\gamma_k h_k}}\right].$$

Here, as in the case of the homogeneous equation, the values η_k are the values of the solution at z_k, the endpoints of the intervals. With this construction the locally exact solutions, $u_k(z)$, automatically satisfy the continuity condition (7). The derivative condition (8) implicitly determines equations that the solution values η_k must satisfy

$$\begin{aligned}&\left[a_{k-1}\gamma_{k-1}\frac{2\mathrm{e}^{\gamma_{k-1}h_{k-1}}}{(1-\mathrm{e}^{2\gamma_{k-1}h_{k-1}})}\right]\eta_{k-1} - \left[a_{k-1}\gamma_{k-1}\frac{(1+\mathrm{e}^{2\gamma_{k-1}h_{k-1}})}{(1-\mathrm{e}^{2\gamma_{k-1}h_{k-1}})} + a_k\gamma_k\frac{(1+\mathrm{e}^{2\gamma_k h_k})}{(1-\mathrm{e}^{2\gamma_k h_k})}\right]\eta_k + \left[a_k\gamma_k\frac{2\mathrm{e}^{\gamma_k h_k}}{(1-\mathrm{e}^{2\gamma_k h_k})}\right]\eta_{k+1}\\ &\quad = a_{k-1}\frac{\mathrm{d}\alpha_{k-1}}{\mathrm{d}z}\bigg|_{z_k} - a_k\frac{\mathrm{d}\alpha_k}{\mathrm{d}z}\bigg|_{z_k}\end{aligned} \tag{13}$$

for $k = 2, \ldots, P-1$.

The specification of the boundary conditions at z_1 and z_P determines η_1 and η_P and thus (13) constitute a linear symmetric tridiagonal system of equations for the solution values η_k. For non-Dirichlet boundary conditions, the construction of the equations is similar to that for the homogeneous equations. If Neumann or "infinite" boundary conditions are specified, then again, η_1 and η_P are additional unknowns and one adjoins equations that specify the local solution to the non-homogeneous problem over the intervals $[z_1, x_2]$ and $[z_{P-1}, z_P]$ satisfy the Neumann boundary conditions or, in the case of infinite boundary conditions, have derivatives that match with a solution over $-\infty < z \leqslant z_1$ and $z_P \leqslant z \leqslant \infty$. In the case of "infinite" boundary conditions, it is assumed that $f(z) = 0$ for $-\infty < z \leqslant z_1$ and $z_P \leqslant z \leqslant \infty$ so that $\alpha_0(z) = 0$ and $\alpha_P(z) = 0$. Additionally, it is worth noting that the tridiagonal matrix that occurs in the system of equations for the η_k's is identical to that which occurs in the homogeneous problem – only the right-hand side of the linear equations are changed.

The error in the computed solution to (6) can be bounded by the error in the local approximations to $f(z)$. Specifically, if we let $g(z)$ be the function defined by the local approximations $g_k(z)$, $g(z) = g_k(z)$ for $z \in [z_k, z_{k+1}]$, then the computed solution $\tilde{u}$ is an exact solution of

$$a_k\frac{\mathrm{d}^2\tilde{u}}{\mathrm{d}z^2} + b_k\tilde{u} = g(z), \quad z \in [z_k, z_{k+1}], \quad k = 1, \ldots, P-1.$$

Thus the error in the solution $u(z) - \tilde{u}(z)$ satisfies the differential equation

$$a_k \frac{d^2(u(z) - \tilde{u}(z))}{dz^2} + b_k(u(z) - \tilde{u}(z)) = f(z) - g(z), \quad z \in [z_k, z_{k+1}], \ k = 1, \ldots, P-1. \tag{14}$$

Assuming that $f(z) - g(z) \in L^2([z_1, z_{P+1}])$, then by combining the results of Theorems 8.3 and 8.16 of [14], one can obtain an a priori error bound

$$\sup_{z \in [z_1, z_{P+1}]} \|u(z) - \tilde{u}(z)\| \leqslant C \sup_{z \in [z_1, z_{P+1}]} \|f(z) - g(z)\|$$

associated with solutions of (14). Here C depends on the size of the interval $[z_1, z_{P+1}]$ and the values of the coefficients a_k and b_k in each layer. From this bound, one can conclude that it is the order of accuracy in the approximation of $f(z)$ that dictates the order of accuracy of the method.

One possible concern about using solutions of (12) to create local particular solutions is that the other component of the solution, the local homogeneous solution, can exhibit very steep gradients if the ratio of the coefficients a_k/b_k is small. For standard numerical methods, this can lead to inaccurate, often oscillatory, solutions unless a sufficiently fine mesh is used. However, by using an exact solution of the homogeneous equation, the gradients are completely resolved and an accurate non-oscillatory approximation is created.

To complete our discussion, we describe a procedure that can be used to construct the solutions to (12) under the assumption that $g_k(z)$ is a polynomial. Since the equation is linear and has constant coefficients over each interval, it is sufficient to find solutions to the non-homogeneous equation of the form

$$a \frac{d^2\alpha}{dz^2} + b\alpha(z) = z^n, \tag{15}$$

with $n = 0, 1, 2, \ldots$ The solution for a general polynomial right-hand side and specific boundary conditions will be a super-position of such solutions and a solution of the homogeneous equation with the required boundary conditions. For a fixed n, one can determine the solution recursively. Let

$$\alpha(z) = \frac{1}{b} z^n + \alpha_1(z)$$

then $\alpha_1(x)$ must satisfy

$$a \frac{d^2\alpha_1}{dz^2} + b\alpha_1(z) = -\frac{a}{b}(n)(n-2)z^{n-2}.$$

Let

$$\alpha_1(z) = -\frac{a}{b^2}(n)(n-2)z^{n-2} + \alpha_2(z)$$

then $\alpha_2(z)$ must satisfy

$$a \frac{d^2\alpha_2}{dz^2} + b\alpha_2(z) = \frac{a^2}{b^3}(n)(n-2)(n-3)(n-4)z^{n-4}.$$

One continues this process with each $\alpha_j(z)$ chosen to satisfy

$$a \frac{d^2\alpha_j}{dz^2} + b\alpha_j(z) = (-1)^j \frac{a^j}{b^{j+1}}(n)(n-2)(n-3)(n-4)\cdots(n-2j)z^{n-2j}.$$

Since the degree of the right-hand side is $n - 2j$, when $j > n/2$ the process will stop and a solution is obtained as the sum of the $\alpha_j(z)$'s. Using this construction, one can express the solution of (15) as

$$\alpha(z) = \frac{1}{b}\sum_{j=0}^{j\leqslant n/2}(-1)^j\left(\frac{a}{b}\right)^j\frac{n!}{(n-2j)!}z^j.$$

4. Numerical results

Since a key ingredient of the procedure described in this paper is the use of Wachspress's method to create highly accurate solutions of two-point boundary value problems, the first numerical example demonstrates its capabilities on the problem

$$\nabla\cdot(a(x)\nabla\phi) + b(x)\phi = f(x), \quad x\in[0,640],$$

where $\phi(0) = 1$, $\phi(640) = 0$, and $a(x)$ and $b(x)$ are given by

$$a(x) = \begin{cases} 0.01, & 0\leqslant x < 288,\\ 10.0, & 288\leqslant x\leqslant 352,\\ 0.01, & 352 < x\leqslant 640,\end{cases} \quad b(x) = \begin{cases} 1.0, & 0\leqslant x < 288,\\ 0.0, & 288\leqslant x\leqslant 352,\\ 1.0, & 352 < x\leqslant 640.\end{cases}$$

The source term is a Gaussian function of the form

$$f(x) = \gamma e^{(x-\bar{x})^2/\sigma_x^2},$$

with $\gamma = 1/10$, $\bar{x} = 320$ (e.g., centered in middle layer) and $\sigma_x = 5\sqrt{2}$.

Fig. 3 shows the behavior of the error versus the number of uniform discretization points for a finite volume discretization (a "standard" second-order finite difference discretization), a Wachspress discretization using a linear right-hand side approximation and a Wachspress discretization using a cubic right-hand side approximation. The errors reported are the maximal value of the difference between the computed

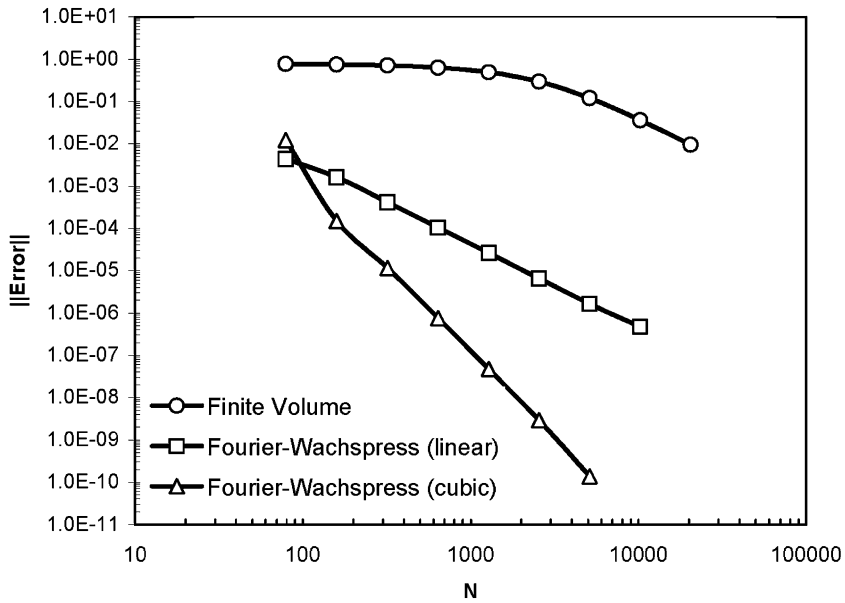

Fig. 3. Approximate solution error versus the number of uniform grid points used in the discretization.

solution, ϕ_c, and an "exact" solution, ϕ_{exact}, that was obtained by computing the solution using an exceptionally fine grid. The results in Fig. 3 clearly reveal the utility of using a Wachspress discretization when the jump in the coefficients is large. While the finite volume discretization and Wachspress (linear) are both second-order accurate discretizations, the errors associated with the finite volume discretization are from two to three orders of magnitude larger. The results also demonstrate that the higher order Wachspress method is not adversely affected by large jumps in the coefficients.

The second test problem demonstrates the capabilities of a Fourier–Wachspress method for computing solutions of Poisson's equation in a three-dimensional layered domain. Specifically, we consider the problem of computing the solution to

$$\nabla \cdot (a(z)\nabla\phi) = f(x,y,z), \quad (x,y,z) \in [-128,128] \times [-128,128] \times [0,640], \tag{16}$$

with $a(z)$ given by

$$a(z) = \begin{cases} 14.1, & 0 \leqslant z < 288, \\ 12.6, & 288 \leqslant z \leqslant 352, \\ 14.1, & 352 < z \leqslant 640. \end{cases}$$

(These coefficients are those associated with the dielectric constants of indium phosphide and gallium arsenide.) The domain was periodic in the x and y directions and Dirichlet conditions were specified at the top and bottom of the device. On the top of the domain ($z = 0$)

$$\phi(x,y,0) = \begin{cases} (x^2+y^2)^2, & x^2+y^2 \leqslant 48, \\ 0, & x^2+y^2 > 48 \end{cases}$$

and on the bottom of the domain ($z = 640$) $\phi(x,y,640) = 0$. The source term was a Gaussian function of the form

$$f(x,y,z) = \gamma \exp\left\{\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} + \frac{(z-\bar{z})^2}{\sigma_z^2}\right\} \tag{17}$$

with $\gamma = 1/20$, $\bar{z} = (0,0,320)$ (e.g., centered in middle layer) and $(\sigma_x, \sigma_y, \sigma_z) = (20\sqrt{2}, 20\sqrt{2}, 5\sqrt{2})$. (This latter function models the electron density of a single electron trapped in the middle layer.)

In Fig. 4, we give results obtained with a uniform mesh with N_z panels in the vertical direction and $N_x = N_y = N_z/10$ panels in the lateral directions. The errors reported are the maximal value of the difference between the computed solution, ϕ_c, and an "exact" solution, ϕ_{exact}, that was obtained by computing the solution using an exceptionally fine grid. As expected, the methods based upon a finite volume discretization and the Fourier–Wachspress method with linear right-hand side approximation demonstrate second-order rates of convergence. The Fourier–Wachspress method with cubic right-hand side approximation demonstrates a fourth-order rate of convergence. The Fourier–Wachspress (linear) method is slightly better than the finite volume discretization, but the difference is not as great as that exhibited by the results of the first test problem. The similarity between the results for the second-order methods can be attributed to the fact that the jumps of the coefficients in this problem are not large. For a given number of discretization points, the computational time was similar for all three methods ranging from 0.02 s for $N_z = 80$, $N_x = N_y = 8$, to 469 s when $N_z = 1280$, $N_x = N_y = 128$. (The computations were performed on a computer with a 2.8 GHz Pentium IV processor). For larger numbers of grid points, with each successive doubling of the number of points used in the vertical direction the computation time increased by a factor close to 16, indicating a total computation time of approximately $(N_x N_y N_z)^{4/3}$. This is not quite as good as formal operation count estimates would indicate, but not unexpected, since factors other than arithmetic operations (e.g., memory access issues) become significant.

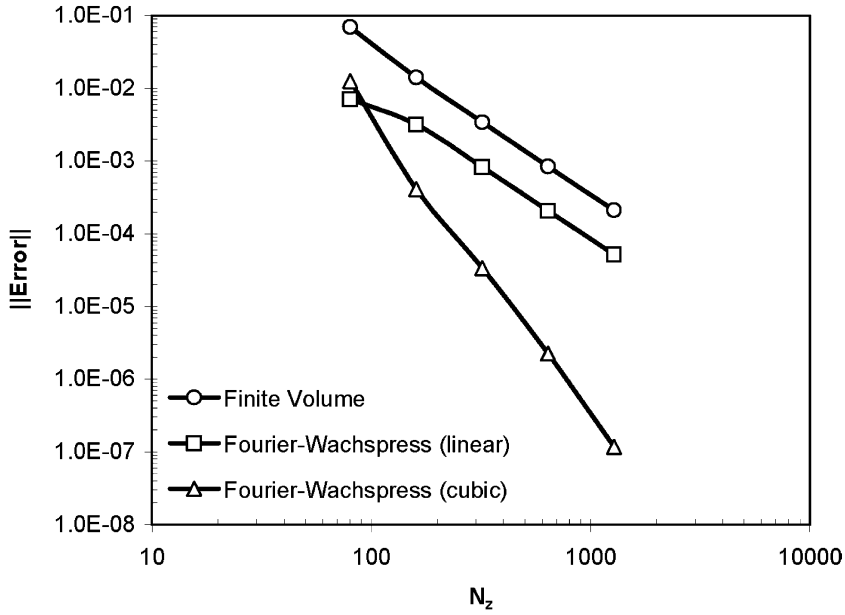

Fig. 4. Approximate solution error versus N_z, the number of panels used in the vertical direction discretization. $N_x = N_y = N_z/10$.

As a demonstration of the value of using an adaptive mesh, solutions to (16) were computed using a fixed mesh size of 16 in the upper and lower regions ($0 < z < 256$ and $384 < z < 640$) and an increasingly refined mesh was used in the middle region ($384 < z < 640$). In Fig. 5, we show the computational time required to obtain solutions with accuracies in a range from 10^{-2} to 10^{-6}. The important result revealed by this data is

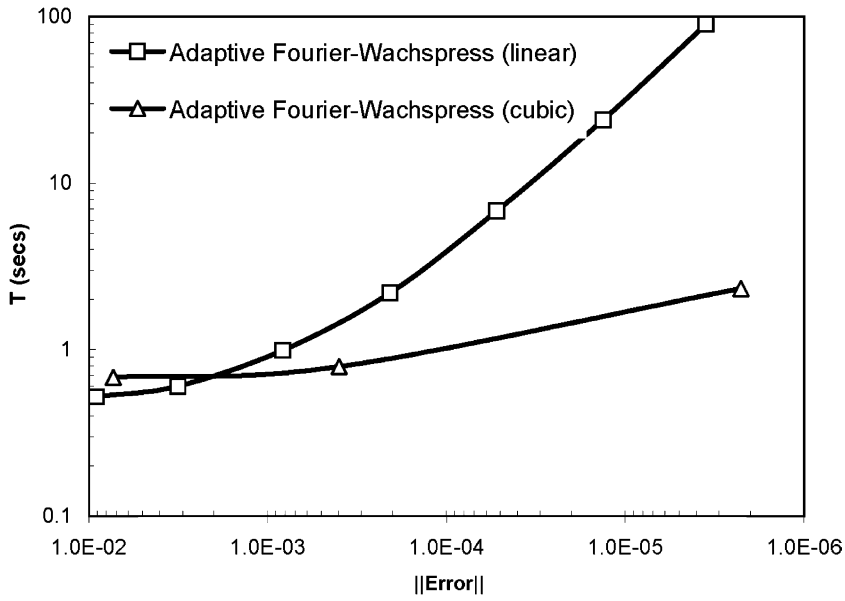

Fig. 5. Computational time versus approximate solution error when using an adaptive discretization.

the significant reduction in computational time that occurs. For example, to obtain an accuracy of approximately 3×10^{-6} requires just 2 s using the Fourier–Wachspress (cubic) method; about 1/15 the time that is required to obtain the solution using a uniform mesh. The results for the Fourier–Wachspress (linear) method are even more dramatic. To obtain a solution with an accuracy of 3×10^{-6} using a Fourier–Wachspress (linear) method requires about 90 s with an adaptive mesh, whereas a similar solution created with a uniform mesh would require 1×10^{5} s (est.). In addition to the reduction in computation time, there is also a considerable reduction in the amount of memory required to obtain the solution. In particular, when using a uniform mesh an array of size 10 MB is required to store the function values, whereas for the non-uniform mesh an array of size 0.5 MB is required.

5. Conclusions

For the Helmholtz or Poisson equation in a three-dimensional rectangular domain, Fourier-based methods provide a computationally efficient means of creating highly accurate solutions. The method presented in this paper can be viewed as a generalization of such methods to layered three-dimensional domains. Through the use of Fourier methods in two dimensions and a variant of Wachspress's method in the third dimension, the resulting scheme is capable of efficiently creating highly accurate solutions – even in the case when the coefficients defining the layered structure are discontinuous, or extreme mesh refinement is used. The method presented in this paper is focussed on layered domains whose layered structure is defined by piecewise constant coefficients, however, one can easily extend the ideas to more general coefficient variation in the vertical direction. In such cases, the use of Wachspress's method may not be optimal, and one might substitute in another high order accurate method for solving two-point boundary value problems [12] or a more general patch based method [3].

Acknowledgment

The authors thank the reviewers of this paper for their numerous constructive comments.

References

[1] A. Averbuch, M. Israeli, L. Vozovoi, A fast Poisson solver of arbitrary order accuracy in rectangular regions, SIAM J. Sci. Comput. 19 (3) (1998) 933–952.
[2] D.N. Arnold, F. Brezzi, B. Cockburn, D. Marini, Discontinuous Galerkin methods for elliptic problems, in: Discontinuous Galerkin Methods, Springer Lecture Notes Computer Science Engineering, Springer, Berlin, 2000, pp. 89–101.
[3] J.P. Boyd, Chebyshev and Fourier Spectral Methods, Dover, New York, 2001.
[4] T. Kerkhoven, A. Galick, U. Ravaloli, J. Arends, Y. Saad, Efficient numerical simulation of electron states in quantum wires, J. Appl. Phys. 68 (7) (1990) 3461–3469.
[5] I.-H. Tan, G.L. Snider, L.D. Cheng, E.L. Hu, A self-consistent solution of Schrödinger–Poisson equations using a nonuniform mesh, Appl. Phys. 68 (8) (1990) 4071–4076.
[6] E. Braverman, M. Israeli, A. Averbuch, L. Vozovoi, A fast 3D Poisson solver of arbitrary order accuracy, J. Comput. Phys 144 (1) (1998) 109–136.
[7] E. Braverman, M. Israeli, A. Averbuch, A fast spectral solver for a 3D Helmholtz equation, SIAM J. Sci. Comput. 20 (6) (1999) 2237–2260.
[8] O.F. Naess, K.S. Eckhoff, A modified Fourier–Galerkin method for the Poisson and Helmholtz equations, in: Proceedings of the Fifth International Conference on Spectral and High Order Methods (ICOSAHOM-01) (Uppsala), J. Sci. Comput. 17 (1–4) (2002) 529–539.
[9] L. Greengard, Spectral integration and two-point boundary value problems, SIAM J. Numer. Anal. 28 (4) (1991) 1071–1080.

[10] J.S. Hesthaven, Spectral penalty methods, in: Proceedings of the Fourth International Conference on Spectral and High Order Methods (ICOSAHOM 1998) (Herzliya), Appl. Numer. Math. 33 (1–4) (2000) 23–41.
[11] R.W. Hockney, A fast direct solution of Poisson's equation using Fourier analysis, J. Assoc. Comput. Mach. 12 (1965) 95–113.
[12] June-Yub Lee, Leslie Greengard, A fast adaptive numerical method for stiff two-point boundary value problems, SIAM J. Sci. Comput. 18 (2) (1997) 403–429.
[13] B.N. Parlett, The Symmetric Eigenvalue Problem, Prentice-Hall SIAM Classics in Applied Mathematics Series, Prentice-Hall Inc., 1998.
[14] D. Gilbarg, N.S.R Trudinger, Elliptic Partial Differential Equations of Second Order, Springer, Berlin, 2001 (Reprint of the 1998 edition).
[15] Gunilla Sköllermo, A Fourier method for the numerical solution of Poisson's equation, Math. Comput. 29 (1975) 697–711.
[16] E.L. Wachspress, The numerical solution of boundary value problems, in: A. Ralston, H. Wilf (Eds.), Mathematical Methods for Digital Computers, Wiley, New York, 1964.