Proceedings of the 5th WSEAS/IASME Int. Conf. on SYSTEMS THEORY and SCIENTIFIC COMPUTATION, Malta, September 15-17, 2005 (pp152-156)

An Accurate Fourier-Spectral Solver for Variable Coefficient Elliptic Equations

Moshe Israeli
Computer Science Department,
Technion-Israel Institute of Technology,
Technion city, Haifa 32000,
ISRAEL

Alexander Sherman
Applied Mathematics Department,
Technion-Israel Institute of Technology,
Technion city, Haifa 32000,
ISRAEL

Abstract: - We develop a solver for nonseparable, self adjoint elliptic equations with a variable coefficient. If the coefficient is the square of a harmonic function,a transformation of the dependent variable, results in a constant coefficient Poisson equation. A highly accurate, fast, Fourier-spectral algorithm can solve this equation. When the square root of the coefficient is not harmonic, we approximate it by a harmonic function. A small number of correction steps are then required to achieve high accuracy. The procedure is particularly efficient when the approximation error is small. For a given function this error becomes smaller as the size of the domain decreases. A highly parallelizable, hierarchical procedure allows a decomposition into small sub-domains. Numerical experiments illustrate the accuracy of the approach even at very coarse resolutions.

Key-Words: - Fast spectral direct solver, Poisson equation, nonseparable elliptic equations, correction steps.

1 Introduction

Variable coefficient elliptic equations are ubiquitous in many scientific and engineering applications the most important case being that of the self-adjoint operator appearing for example in diffusion processes in non uniform media. Many repeated solutions of such problems are required when solving variable coefficient or non linear time dependent problems by implicit marching methods.

Application of high-order (pseudo) spectral methods, which are based on global expansions into orthogonal polynomials (Chebyshev or Legendre polynomials), to the solution of elliptic equations, results in full (dense) matrix problems. The cost of inverting a full $N \times N$ matrix without using special properties is $O(N^3)$ operations [2]. The spectral element method allows for some sparsity. The Fourier method for the solution of the Poisson equation gives rise to diagonal matrices and has an exponential rate of convergence but looses accuracy due to the Gibbs phenomenon for non-periodic boundary conditions.

Our method to resolve the Gibbs phenomenon presents the RHS as a sum of a smooth periodic function and another function which can be integrated analytically. We were able to implement this idea to two or three dimensions.This approach is sometimes called ”subtraction”. Other approaches to the resolution of such problems appear in [4, 5, 8] and [1, 7].

The subtraction technique for the reduction of the Gibbs phenomenon in the Fourier series solution of the Poisson equation goes back to Sköllermo [9] who considered,

$$\Delta u = f \tag{1}$$

in the rectangle $[0, 1] \times [0, 1]$ with non periodic boundary conditions. We note that the subtraction algorithm in [9] was of limited applicability because of a technical difficulty which we resolved.

The subtraction technique (in the physical space) has the following advantages:

a) After subtraction, the Fast Fourier Transform can be applied to the remaining part of RHS with a high convergence rate.

b) The algorithm preserves the diagonal representation of the Laplace operator, (unlike Chebyshev and Legendre expansions), hence the invertion of the matrix is trivial.

c) The computation of the subtraction functions is even less time consuming than the application of the FFT.

In the framework of the present paper:

1. We develop first a fast direct algorithm for the solution of Eq. (2) for any function $a(x, y)$, such that $a(x, y)^{1/2}$ is harmonic. It is based on an improvment of the fast direct solver of [1] and a transformation described in [3].

2. If $a(x, y)^{1/2}$ is not harmonic, we approximate it by a harmonic function. The numerical scheme then applies the basic algorithm in a sequence of correction steps

Proceedings of the 5th WSEAS/IASME Int. Conf. on SYSTEMS THEORY and SCIENTIFIC COMPUTATION, Malta, September 15-17, 2005 (pp152-156)

(the procedure is described in Section 2 and tested in Section 3).

3. An adaptive domain decomposition approach was implemented which improves the approximation for any function $a(x, y)$ and reduces the number of corrections step required (to about three).

2 Outline of the Algorithm

We solve an elliptic equation

$$\nabla \cdot (a(x, y) \, \nabla u(x, y)) = f(x, y), \quad (x, y) \in \mathbf{D}, \qquad (2)$$

where $\mathbf{D}$ is a rectangular domain, with the Dirichlet boundary conditions

$$u(x, y) = g(x, y), \quad (x, y) \in \partial\mathbf{D}. \qquad (3)$$

We assume $a(x, y) > 0$ for any $(x, y) \in \mathbf{D}$.
Following [3] we make the change of variable

$$w(x, y) = a(x, y)^{1/2} u(x, y), \qquad (4)$$

then Eq. (2) takes the form

$$\Delta w(x, y) - p(x, y) w(x, y) = q(x, y), \qquad (5)$$

where

$$\begin{aligned} p(x, y) &= \Delta(a(x, y)^{1/2}) \cdot a(x, y)^{-1/2}, \\ q(x, y) &= f(x, y) \cdot a(x, y)^{-1/2}. \end{aligned} \qquad (6)$$

In case $a(x, y)^{1/2}$ is a harmonic function, Eq. (5) becomes the Poisson equation in w:

$$\Delta w(x, y) = q(x, y) \qquad (7)$$

This leads to the fast direct algorithm for the numerical solution of Eq. (2), where $a(x, y)^{1/2}$ is a harmonic function.

Algorithm A

1. Using the modified spectral subtractional algorithm which was described in the introduction, we solve Eq. (7) with the boundary conditions $\tilde{g}(x, y) = a(x, y)^{1/2} \cdot g(x, y)$.

2. The solution of Eq. (2) is $u(x, y) = w(x, y) \cdot a(x, y)^{-1/2}$.

Thus, this algorithm enables the solution of Eq. (2)

$$\nabla \cdot (\widetilde{a}(x, y) \, \nabla u) = f(x, y), \qquad (8)$$

where

$$\Delta(\widetilde{a}(x, y)^{1/2}) = 0$$

as a constant coefficient problem. Let us now consider the case where $a(x, y)^{1/2}$ is not exactly harmonic but can be well approximated by a harmonic function $\widetilde{a}(x, y)^{1/2}$. This means that the difference

$$\varepsilon(x, y) = a(x, y) - \widetilde{a}(x, y) \qquad (9)$$

is small. Denote by $u_0(x, y)$ the solution of (8) with boundary conditions (3) and introduce
$\widetilde{u}(x, y) = u(x, y) - u_0(x, y)$, where $u(x, y)$ is an exact solution of Eq. (2). Then (2) can be rewritten as

$$\nabla \cdot [(\widetilde{a}(x, y) + \varepsilon(x, y)) \, \nabla(u_0 + \widetilde{u})] = f(x, y). \qquad (10)$$

Taking into account (8), we obtain

$$\nabla \cdot (\widetilde{a}(x, y) \, \nabla\widetilde{u}) = -\nabla \cdot (\varepsilon(x, y) \, \nabla(u_0 + \widetilde{u})), \qquad (11)$$

where $\widetilde{u}(x, y)$ satisfies the zero boundary conditions as the difference of two functions $u(x, y)$ and $u_0(x, y)$, which both satisfy (3). Since $\widetilde{u}$ is unknown, the following correction procedure is suggested:

$$\nabla \cdot (\widetilde{a}(x, y) \, \nabla u_1) = -\nabla \cdot (\varepsilon(x, y) \, \nabla u_0) \qquad (12)$$

$$\begin{aligned} \nabla \cdot (\widetilde{a}(x, y) \, \nabla u_{n+1}) = \\ -\nabla \cdot (\varepsilon(x, y) \, \nabla(u_0 + u_n)), \quad n \geq 1. \end{aligned} \qquad (13)$$

Subtracting (11) from (13) we have

$$\begin{aligned} \nabla \cdot (\widetilde{a}(x, y) \, \nabla(u_{n+1} - \widetilde{u})) = \\ -\nabla \cdot (\varepsilon(x, y) \, \nabla(u_n - \widetilde{u})). \end{aligned} \qquad (14)$$

Suppose $\|\varepsilon\| \leq s\|a\|$ in some norm, where s is small, and denoting by u^n the corrected solution after n correction steps ($u^n = u_0 + u_n$). Since the exact solution is $u = u_0 + \widetilde{u}$, it follows that the error decreases according to:

$$\|u^{n+1} - u\| \leq s\|u^n - u\| \qquad (15)$$

Thus the algorithm for the solution of (2) will be:

1. The coefficient $a(x, y)$ in (2) is approximated by $\widetilde{a}(x, y)$ such that $\widetilde{a}(x, y)^{1/2}$ is a harmonic function in the domain $\mathbf{D}$. Equation (8) is solved using Algorithm A.

2. Several correction steps are made using (13) until the desired accuracy is attained.

To approximate $a(x,y)^{1/2}$ by a harmonic function we consider first the bilinear function

$$\widetilde{b}(x,y) = c_{11} + c_{12}x + c_{21}y + c_{22}xy, \qquad (16)$$

which takes on the corner values of $a(x,y)^{1/2}$ This approximation can be improved by matching more points on the boundary of the square $[0,1] \times [0,1]$ by the addition of functions of the type

$$\varphi(x,y) = d_k \sin(\pi kx) \sinh(\pi ky) \qquad (17)$$

which do not affect corner points. Dividing the domain to smaller squares improves the error according to the square of the size of the subdomain.

Note that the simplest approximation with bilinear functions has the following advantage: if we use domain decomposition (see Section 4), then the collection of $\widetilde{a}(x,y)$ approximated in subdomains while not smooth is a continuous function, this simplifies very much the inter domain matching process. This is true also for (17). Approximations involving interior values of $a(x,y)^{1/2}$ do not enjoy this useful property.

3 Numerical results

First let us demonstrate the rate of convergence of the improved subtraction algorithm in the case where $a(x,y)^{1/2}$ is a harmonic function.
Assume that u is the exact solution of Eq. (2) and u' is the computed solution. We will use the following measure to estimate the errors:

$$\varepsilon_{MAX} = \max |u'_i - u_i| \qquad (18)$$

Example 1. Consider the equation with $a(x,y) = (x+1)^2(y+0.5)^2$, the right hand side and the boundary conditions correspond to the exact solution $u(x,y) = (x+1)^2 + (y+0.5)^2$ in the domain $[0,1] \times [0,1]$. The results are presented in Table 1.

$N_x \times N_y$	$\varepsilon_{MAX}(4)$	$\varepsilon_{MAX}(6)$
8×8	1.03e-5	1.56e-08
16×16	8.08e-7	3.03e-10
32×32	5.67e-8	5.39e-12
64×64	3.76e-9	2.05e-13

Table 1: MAX error for for the fourth order (4) and sixth order (6) subtraction methods.

Now let us proceed to an example, where $a(x,y)^{1/2}$ is not harmonic.

Example 2. Consider the equation with $a(x,y) = (x+1+r\sin x)^2(y+0.5)^2$, the right hand side and the boundary conditions correspond to the same exact solution as in Example 2. Here we need to apply a few correction steps in order to get the desired accuracy. We used (16) for the approximation of $a(x,y)^{1/2}$ by a harmonic function. The results for $r = 0.5$ are presented in Table 2.

If we insist to get the same accuracy as in the previous example, it is necessary to apply from 4 correction steps for 8×8 points to 6 correction steps for 64×64 points. It is expected that with the growth of r more correction steps would be required.

$N_x \times N_y$	$CS = 1$	$CS = 2$	$CS = 3$
8×8	7.16e-4	1.98e-5	5.72e-7
16×16	7.26e-4	2.02e-5	5.87e-7
32×32	7.29e-4	2.03e-5	5.86e-7
64×64	7.30 -4	2.03e-5	5.88e-7

$N_x \times N_y$	$CS = 4$	$CS = 5$	$CS = 6$
8×8	2.16e-8	1.33e-08	1.33e-08
16×16	1.75e-8	3.69e-10	3.08e-10
32×32	1.74e-8	5.22e-10	1.98e-11
64×64	1.74e-8	5.35e-10	5.66e-11

Table 2: MAX error for the same exact solution and $a(x,y) = (x+1+0.5\sin x)^2(y+0.5)^2$ for various numbers of correction steps in the domain $[0,1] \times [0,1]$. CS is the number of Correction Steps.

4 Domain Decomposition

The present algorithm incorporates the following novel elements:

1. It extends our previous fast Poisson solvers [1, 7] as it provides an essentially direct solution for equations (2) where $a(x,y)^{1/2}$ is an arbitrary harmonic function, in particular, a bilinear function

$$a(x,y)^{1/2} = c_{11} + c_{12}x + c_{21}y + c_{22}xy.$$

2. In the case where $a(x,y)^{1/2}$ is not harmonic, we approximate it by $\widetilde{a}(x,y)^{1/2}$ and apply several correction steps to improve the accuracy.

However high accuracy for the solution of (2) requires an accurate approximation of $a(x,y)^{1/2}$ by a harmonic function. Such an approximation is not always easy to derive in the global domain, however it can be achieved in smaller subdomains. In this case we suggest the following Domain Decomposition algorithm.

1. The domain is decomposed into smaller rectangular subdomains. Where the boundary of the subdomains coincides with full domain boundary we take on the original boundary conditions. For other interfaces we introduce some initial boundary conditions which do not contradict the equation at the corners, where the left hand side of (2) can be computed. The function $a(x, y)$ is approximated by $\tilde{a}(x, y)^{1/2}$ in each subdomain such that $\tilde{a}(x, y)^{1/2}$ is harmonic. An auxiliary equation (8) is solved in each subdomain.

2. The collection of solutions obtained at Step 1 is continuous but doesn't have continuous derivatives at domain interfaces. To further match subdomains, a hierarchical procedure can be applied similar to the one described in [6]. For example, if we have four subdomains 1,2,3 and 4, then 1 can be matched with 2, 3 with 4, while at the final step the merged domain 1,2 is matched with 3,4.

Example 3. In Fig. 1 we illustrate the effectiveness of the domain decomposition approach by solving the one dimensional variable coefficient equation where the coefficient function is not harmonic. We solve the equation with $a(x) = (2x+3+\sin(\pi x))^2$ with exact solution $u(x) = \sin(\pi x)$. We change the number of domains (D) from 1 to 8. Obviously the correction procedure works much better then the subdomains become smaller. With 4 domains and with only 2 correction steps we reach an error of order 10^{-6}, with 8 domains we get 10^{-8}. Thus the present approach behaves essentially as direct fast method for the self adjoint PDE. The Domain Decomposition of course has the further advantage of easy parallelization on massively parallel computers.

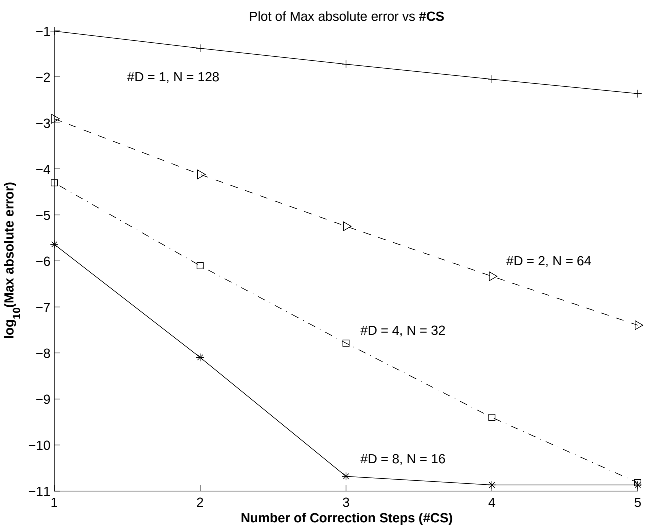

Figure 1: MAX error for the $a(x, y) = (x + 1 + 0.5\sin x)^2(y + 0.5)^2$ for various numbers of subdomains. The total number of points in the domain [0,1] remains unchanged and equals to 128. Number of subdomains used (D) is followed by the actual number of points used in each subdomain (N).

Example 4. In Table 3 and 4 we illustrate the results of applying the domain decomposition approach by solving the two dimensional variable coefficient equation where the coefficient function is not harmonic. We consider the equation with $a(x, y) = (1/2 + x^2 + y)\sin(1/3 + x + 2y) + 1$, corresponding to the exact solution $u(x, y) = (1/4 + x + y^2)\sin(1/3 + 2x + y)$. In the domain $[0, 1] \times [0, 1]$, the bilinear approximation to $a(x, y)^{1/2}$ isn't good. As a result, a large number of correction steps is needed. When we solve the same equation but in smaller subdomains, the $a(x, y)$ could be well approximated by $\tilde{a}(x, y)$ and, as a consequence, we achieve better results. A prime factor that determines a numerical approximation is the ratio $\varepsilon(x, y)/\tilde{a}(x, y)^{1/2}$. Other factor that has an influence on the numerical results is the order of the subtraction method used for solution of (1).

$Domain$	$N_x \times N_y$	$CS = 2$	$CS = 5$	$CS = 10$
$[0\ 1] \times [0\ 1]$	64×64	9.85e-2	4.58e-2	2.03e-2
$[0\ \frac{1}{2}] \times [0\ \frac{1}{2}]$	32×32	2.09e-6	1.73e-8	1.73e-8
$[0\ \frac{1}{4}] \times [0\ \frac{1}{4}]$	16×16	3.08e-8	3.00e-8	3.00e-8

Figure 3: MAX error. Fourth(4) order subtraction method was used trough calculation of the correction steps.

$Domain$	$N_x \times N_y$	$CS = 2$	$CS = 5$	$CS = 10$
$[0\ 1] \times [0\ 1]$	64×64	9.85e-2	4.58e-2	2.03e-2
$[0\ \frac{1}{2}] \times [0\ \frac{1}{2}]$	32×32	2.10e-6	1.48e-10	3.02e-12
$[0\ \frac{1}{4}] \times [0\ \frac{1}{4}]$	16×16	6.31e-9	1.31e-11	1.31e-11

Figure 4: MAX error. Sixth(6) order subtraction method was used trough calculation of the correction steps.

References

[1] E. Braverman, B. Epstein, M. Israeli and A. Averbuch, A fast spectral subtractional solver for elliptic equations, *J. Sci. Comput.* **21** (2004), No. 1, 91-128.

[2] C. Canuto, M.Y. Hussaini, A. Quarteroni, T. A. Zang, *Spectral Methods in Fluid Dynamics*, Springer-Verlag, 1989.

[3] P. Concus and G.H. Golub, Use of fast direct methods for the efficient numerical solution of nonseparable elliptic equations, *SIAM J. Numer. Anal.* **10** (1973), No. 6, 1103-1120.

[4] D. Gottlieb and C.W. Shu, On the Gibbs phenomenon and its resolution, *SIAM Review* **39** (1997), no. 4, 644-668.

[5] M.S. Min and D. Gottlieb, On the convergence of the Fourier approximation for eigenvalues and eigenfunctions of discontinuous problems, *SIAM J. Numer. Anal.* **40** (2002), No. 6, 2254-2269.

[6] M. Israeli, E. Braverman and A. Averbuch, A hierarchical domain decomposition method with low communication overhead, *Domain decomposition methods in science and engineering* (Lyon, 2000), 395-403, *Theory Eng. Appl. Comput. Methods*, Internat. Center Numer. Methods Eng. (CIMNE), Barcelona, 2002.

[7] O.F. Næss and K.S. Eckhoff, A modified Fourier-Galerkin method for the Poisson and Helmholtz equations, *J. Sci. Comput.* **17** (2002), no. 1-4, 529–539.

[8] B.D. Shizgal and J.H. Jung, Towards the resolution of the Gibbs phenomena, *J. Comput. Appl. Math.* **161** (2003), No. 1, 41-65.

[9] G. Skölermo, A Fourier method for numerical solution of Poisson's equation, *Math. Comput. (U.S.A.)* **29**, (1975) 697-711.