1978

High Order Discretizations of the Helmholtz Problem Which Admit Interactive Solution Techniques

Ronald F. Boisvert

Report Number:

78-296

HIGH ORDER DISCRETIZATIONS OF THE HELMHOLTZ PROBLEM WHICH ADMIT ITERATIVE SOLUTION TECHNIQUES

Ronald F. Boisvert

Department of Computer Sciences
Purdue University

Abstract

Two one-parameter families of fourth order finite difference discretizations of the Dirichlet problem for the Helmholtz equation $\partial^2 u/\partial x^2 + \partial^2 u/\partial y^2 + Fu = G$ on a rectangle are presented. The general suitability of several standard iterative techniques for the solution to the resulting systems of linear equations are discussed. Some numerical results are presented which indicate that such high order schemes may be preferable to the usual low order methods for use with iterative solution techniques for such problems.

1. Introduction

We consider the numerical solution to the the Dirichlet problem on a rectangle for the Helmholtz equation

$$(1) \qquad \nabla^2 u + Fu = G, \qquad \nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$$

where $F \leq 0$ is a constant, by the use of high order difference methods of the HODIE type introduced by Lynch and Rice [8].

A nine-point HODIE discretization of the operator $Lu = G$ at the point (x_0, y_0) takes the form

$$L_h U = (1/h^2) \sum \alpha_i U_i = \sum \beta_j G_j = I_h G$$

where U_i is an estimate of $u(x_i, y_i)$, $G_j = G(x_j, y_j)$ and h is the grid spacing. The points (x_i, y_i) are the grid points adjacent to (x_0, y_0) labelled as

```
6----2----5
|    |    |
3----0----1
|    |    |
7----4----8
```
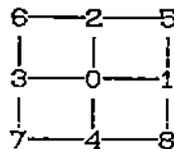
and the points (x_j, y_j) are additional points, called auxiliary points, which are located near (x_0, y_0). The coefficients α_i and β_j are chosen to make the discretization exact for the space $\mathbb{P}_n$ of polynomials of degree at most n, that is, $L_h s - I_h L s = 0$ for all $s \in \mathbb{P}_n$.

A key idea of the HODIE method is that high order is attained by the use of auxiliary points and not by the use of grid points more than one grid line away from the central point. The resulting linear system for the U_i has the same block tridiagonal form as the usual nine-point discretization of the Laplacian; in fact, the coefficients α_i are $O(h)$ perturbations of these.

For an elliptic operator with variable coefficients, the difference equation coefficients α_i and β_j are determined by solving a small linear system for each grid point. The size of this system is fixed independant of h and thus the time to solve the system of difference equations for U dominates for h small.

2. Families of Fourth Order Methods for the Helmholtz Problem

In the case of constant coefficient operators, the coefficients α_i and β_j are constants independant of (x_0, y_0). In addition, for the Helmholtz problem, symmetry can be used to reduce the number of auxiliary points required to get fourth order accuracy from the usual 13 to 5. We define two sets of auxiliary points,

$$P(h, \xi) = \{(0,0), (\xi, 0), (0, \xi), (-\xi, 0), (0, -\xi)\},$$
$$Q(h, \xi) = \{(0,0), (\xi, \xi), (-\xi, \xi), (-\xi, -\xi), (\xi, -\xi)\},$$

where $0<\xi\leq1$ is a parameter, and label them as

$P(h,\xi)$: $Q(h,\xi)$:
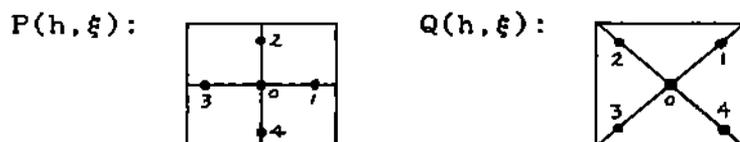
Two one-parameter families of fourth order HODIE discretizations of the form $L_hU = I_hG$, where

$$L_hU = (1/(6h^2)) \ (\alpha_0U_0 + \alpha_1(U_1+U_2+U_3+U_4) + \alpha_5(U_5+U_6+U_7+U_8))$$

$$I_hG = \beta_0G_0 + \beta_1(G_1+G_2+G_3+G_4)$$

can be constructed using the point sets $P(h,\xi)$ and $Q(h,\xi)$. The coefficients of these schemes are:

For the set $Q(h,\xi)$

$$\alpha_0 = -20 + (8-\xi^2)Fh^2 - (1-\xi^2)F^2h^4/2$$
$$\alpha_1=\alpha_2=\alpha_3=\alpha_4 = 4$$
$$\alpha_5=\alpha_6=\alpha_7=\alpha_8 = 1 + \xi^2Fh^2/4$$
$$\beta_0 = (6\xi^2-1)/(6\xi^2) - (1-\xi^2)Fh^2/12$$
$$\beta_1=\beta_2=\beta_3=\beta_4 = 1/(24\xi^2)$$

For the set $P(h,\xi)$

$$\alpha_0 = -20 + 2(3-\xi^2)Fh^2 - (1-\xi^2)F^2h^4/2$$
$$\alpha_1=\alpha_2=\alpha_3=\alpha_4 = 4 + \xi^2Fh^2/2$$
$$\alpha_5=\alpha_6=\alpha_7=\alpha_8 = 1$$
$$\beta_0 = (3\xi^2-1)/(3\xi^2) - (1-\xi^2)Fh^2/12$$
$$\beta_1=\beta_2=\beta_3=\beta_4 = 1/(12\xi^2).$$

The fourth order convergence of these schemes is proven in [1].

A number of specific discretizations in these families have been considered previously for the case of the Poisson equation (F=0). The method with auxiliary point set $P(h,1)$ appears in a number of sources, see [3] for example. Similar families for the Poisson equation were derived by Esch [4] using Fourier series methods.

In [1], the discretization error of members of these families are considered in more detail. In particular, it is shown that the method based on $Q(h,\sqrt{0.3})$ is optimal with respect to a certain bound on the discretization error. The method $Q(h,1/2)$ is found to be a nearly optimal method, and the method $Q(h,1)$ is seen to give the worst overall performance in both theory and practice. All these methods were found to compare favorably with the collocation finite element method using bi-cubic Hermite elements.

3

3. The Use of Iterative Methods

Once a problem is discretized using one of the methods of the previous section, we are left with a linear system of equations to solve for the values of the solution at the grid points. Traditionally, such linear systems have been very large (from $10^3 \times 10^3$ to $10^6 \times 10^6$, for example) and hence iterative solution techniques have been preferred, since storage requirements are typically much smaller for such methods. Although high order methods have the advantage that for smooth problems a relatively coarse mesh can deliver the required accuracy (e. g. for h sufficiently small, if h is halved, the error in a second order method goes down by a factor of 1/4, whereas the error in a fourth order method goes down by a factor of 1/16), it is still concievable that large matrix problems might need to be solved. Hence we consider the applicability of iterative methods.

We start with some standard definitions.

Definitions

Let A be a real matrix of order N>1. Then,
a) A is <u>irreducible</u> if and only if there does not exist a permutation matrix P such that

$$P^{-1}AP = \begin{bmatrix} B & 0 \\ C & D \end{bmatrix}$$

where B and D are square matrices.

b) A is <u>weakly diagonally dominant</u> if $|a_{ii}| \geq \sum |a_{ij}|$ for each i, where the sum is taken over all $j \neq i$ and strict inequality holds for at least one i.

c) A is an <u>L-matrix</u> if $a_{ii} > 0$ for $i=1,2,\ldots,N$ and $a_{ij} \leq 0$ for $i,j=1,2,\ldots,N$ with $i \neq j$.

d) A is a <u>Stieltjes matrix</u> if it is a positive definite L-matrix.

We now summarize some of the basic properties of the matrices generated by the HODIE discretizations of the previous section.

Theorem 3.1

Let $-A$ be the NxN matrix generated by any of the HODIE discretizations of (1) given above. We assume that the natural ordering of equations and unknowns is used. Then, for h sufficiently small, A is real and
a) symmetric,
b) irreducible,
c) weakly diagonally dominant,
d) a block tridiagonal matrix, each block being itself tridiagonal and strictly diagonally dominant,
e) an L-matrix, and
f) positive definite.

Note that an immediate consequence of e) and f) is that A is a Stieltjes matrix.

For a matrix with these properties, much of the existing theory on iterative method applies, which we summarize here. We know, for example, that any of the point Jacobi, Gauss–Seidel and successive overrelaxation (SOR, for relaxation parameter $0<\omega<2$) converge when applied to the matrix A of Theorem 3.1 (this follows from irreducibility and weak diagonal dominance, see [10], pp. 107–108). In addition, since A is both an L-matrix and a Stieltjes matrix, we know that the asymptotic rate of convergence of the Gauss–Seidel method is greater than that of the Jacobi method.

Since the discretizations given here are based on nine-point formulas, the resulting matrix does not have "Property A" (see [10], pp. 41–42),and hence the SOR theory which yields an optimum relaxation factor does not apply. However, since A is Stieltjes, the SOR theory holds approximately, that is, the relaxation parameter called for by the theory for matrices with Property A is nearly optimal, and the SOR method might still be effective. Also, when considered as a block tridiagonal matrix, A does have "Property A(π)" (see [10], p. 445) and thus the SOR theory does apply to relaxation by lines. Such block iterative techniques have been found to be quite effective in both theory and practice.

A number of techniques for accelerating the convergence of the basic iterative methods have been proposed that apply in the case we consider here. For example, since A is symmetric positive definite, the Chebyshev semi–iterative methods apply (see [10], pp. 344–385). When used in conjunction with the Jacobi method, optimum values of the acceleration parameters are known, yielding the so–called Jacobi–SI method, which can result in an order–of–magnitude improvement in convergence rate when compared to the point Jacobi method. Other acceleration methods are possible, notably the conjugate gradient acceleration developed in [5].

The use of semi–iterative methods has not been found to yield substantial gains when applied to the SOR method for general problems. However, a variant of SOR, the symmetric SOR (SSOR) method, is particularly suited to this technique. At each step of an SSOR iteration, two sweeps are made over the equations, one forward and one backward. This method converges whenever A is symmetric and positive definite, as is the case here, and the SOR relaxation parameter is chosen $0<\omega<2$ (see [10], pp. 461–465). In addition, a particularly good choice of ω is known for matrices that also satisfy the SSOR condition: $\rho(LU)\leq 1/4$, where L and U are the strictly lower and upper triangular parts of the matrix $\text{diag}(A)^{-1}\cdot A$ respectively an $\rho(M)$ denotes the spectral radius of M.

The matrices we consider here do, in fact, satisfy the SSOR condition. Consider the $Q(h,\xi)$ case (the $P(h,\xi)$ case is similar). Since A is symmetric, $L=U^T$ and hence it is sufficient to show that $\rho(L)\leq 1/2$. Also, $\rho(L)\leq\|L\|_\infty$, and so evidently, $\rho(L)\leq |\alpha_0|(\ |\alpha_1| + |\alpha_2| + |\alpha_5| + |\alpha_6|\ ) = (10 + \xi^2 Fh^2)/(20 - (6-\xi^2)Fh^2 + (1-\xi^2)F^2h^4)/2)$, so long as h is chosen so small that $\xi^2|F|h^2\leq 4$. Thus $\rho(L)\leq 1/2$ for h sufficiently small.

Although SSOR requires twice the work per iteration of SOR, implementations using semi–iteration have proven highly successful.

4. Numerical Tests

A number of test problems have been discretized with both the second order five-point star and the fourth order HODIE method based on auxiliary point set Q(h,1/2). In each case, five iterative algorithms in the ITPACK package of programs [7] were applied: SOR, Jacobi-SI, Jacobi-CG, SSOR-SI and SSOR-CG (SI = Chebyshev semi-iterative acceleration, CG = conjugate gradient acceleration). These programs are based on the adaptive methods of Hageman and Young [5] in which relaxation and acceleration parameters are chosen automatically. Starting vectors were obtained by linearly interpolating the boundary conditions.

The tests were performed on the CDC 6500 computer system at Purdue University with the aid of the system for performance evaluation of software for partial differential equations described in [2]. All modules tested are included in ELLPACK [9].

Six problems from the PDE population study [6] were chosen which exhibit a variety of behaviors. They are (problem numbers are from [6]):

● Problem 3
    Operator : Poisson
    Solution : $3\exp(x+y)xy(1-x)(1-y)$
    Features : Entire solution

● Problem 4
    Operator : Poisson
    Solution : A function of $x^{3/4}$ and $y^{3/4}$
    Features : Solution has singular first derivative

● Problem 8
    Operator : Poisson, G is a delta-type function
    Solution : $\phi(x)\phi(y)$, where $\phi(x)=1$ for $x \leq 0.35$, $\phi(x)=p(x)$ for
                $0.35 \leq x \leq 0.65$, and $\phi(x)=0$ for $0.65 \leq x$; and $p(x)$ is a
                quintic polynomial with continuous second derivatives.
    Features : Solution has a wavefront at a right angle through the
                center of the domain connecting two flat areas.
                Solution has a discontinuous third derivative.

● Problem 41
    Operator : Helmholtz, $\nabla^2 u - 10u = G$
    Solution : $\cos(10y) + \sin(10(x-y))$
    Features : Oscillatory solution

● Problem 7
    Operator : Helmholtz, $\nabla^2 u - 100u = g$
    Solution : $[\cosh(10x)/\cosh(10) + \cosh(20y)/\cosh(20)]/2$
    Features : Nearly singular solution with boundary layer

● Problem 6
    Operator : Helmholtz, $\nabla^2 u - (100+\cos(3\pi x)+\sin(2\pi y))u = G$
    Solution : $0.31(5.4-\cos(4\pi x))\sin(\pi x)y(y-1)(5.4-\cos(4\pi y))(1/(1+$

$\phi^4)-(1/2))$, where $\phi = 4(x-(1/2))^2+4(y-(1/2)^2)$

Features : Non-constant coefficient operator, oscillatory solution .

Note that problem 6 has a variable coefficient F and hence the analysis given above does not completely apply. However, as shown in [1], the HODIE discretization using auxiliary point set P(h,1) yields a fourth order method in this case. The resulting matrix is no longer symmetric, although for sufficiently small h it remains irreducible, weakly diagonally dominant and an L-matrix. We shall see that the observed behavior of the iterative methods remains the same as in the other cases for the problem considered here.

Each problem was solved on the region $0 \leqq x \leqq 1$, $0 \leqq y \leqq 1$ with square grids giving h = 1/4, 1/8, 1/12, and 1/16. The results are summarized in Table 1 and Figure 1.

For each problem tested, the high order discretization yielded the most efficient method, that is, for a given accuracy, the smallest execution time for all methods tested occurred with the high order discretization. The superiority of the HODIE methods extends to the case where u has nearly singular behavior or even discontinuous derivatives. In this case the rate of convergence is reduced, perhaps until it is the same as for a lower order method (as in Problem 4), but we have found that the absolute error given by the high order discretization is almost always smaller.

The adaptive iterative methods tested here appear to be as effective in solving the nine-point fourth order difference equations as in solving the second order five-point star difference equations. We observe that the number of iterations required to solve a linear system of a given order for a specific problem was roughly the same regardless of whether the system was generated by the second or fourth order discretization. Since the matrices generated by HODIE have twice as many off-diagonal elements it is not surprising that the execution times were greater in this case. The increase in work was never more than 50 percent, however, and the greater accuracy of the fourth order discretization was always enough to offset this.

Of the adaptive iterative codes tested, the SSOR-CG method seemed to show the best overall performance in solving the five-point star equations, its execution time being smallest, or nearly so, in each case. For solving the nine-point HODIE equations, both the Jacobi-CG and the SSOR-CG codes appeared to yield the best performance. The Jacobi-SI code showed the longest running time in each case.

References

[1] Boisvert, R. F., The effect of accuracy of the placement of auxiliary points in the HODIE method for the Helmholtz problem, Purdue Univ. C. S. Dept. CSD-TR 266, June 1978.

[2] Boisvert, R. F., Houstis, E. N., and Rice, J. R., A system for performance evaluation of partial differential equations

software, Purdue Univ. C. S. Dept. CSD-TR 278, July 28, 1978.

[3] Collatz, L., *The Numerical Treatment of Differential Equations*, 3rd edition, Springer-Verlag, New York, 1966.

[4] Esch, R., High-order difference approximations to Poisson's equation, in *Proceedings of a Harvard Symposium on Digital Computers and their Applications*, Annals Comp. Lab. Harvard U., 31 (1961), pp. 81-102.

[5] Hageman, L. A., and Young, D. ,M., *On the Effective Use of Iterative Methods for Solving Large Linear Systems*, a monograph, in preperation.

[6] Houstis, E. N., and Rice, J. R., A population of partial differential equations for evaluating methods, Purdue Univ. C. S. Dept. CSD-TR 263, May 15, 1978.

[7] Kincaid, D. R., and Grimes. R. G., Numerical studies of several adaptive iterative algorithms, CNA 126, Center for Numerical Analysis, Univ. of Texas at Austin, Aug. 1977.

[8] Lynch, R. E., and Rice, J. R., High accuracy finite difference approximations to solutions of elliptic partial differential equations, Proc. Nat. Acad. Sci., 75 (1978), pp. 2541-2544.

[9] Rice, J. R., ELLPACK: A research tool for elliptic partial differential equations software, in *Mathematical Software III* (J. Rice, ed.), Academic Press, (1977), pp. 65-118.

[10] Young, D. M., *Iterative Solution of Large Linear Systems*, Academic Press, New York, 1971.

Problem 3   max solution = 0.57

five-point star discretization

h	error	disc	SOR	J-SI	J-CG	SSOR-SI	SSOR-CG
1/4	1.0-2	0.02	0.06 13	0.10 21	0.06 5	0.03 7	0.08 5
1/8	4.8-3	0.07	0.39 24	0.84 43	0.53 18	0.44 11	0.47 9
1/12	1.9-3	0.17	1.58 47	2.19 67	1.45 26	1.44 15	1.39 12
1/16	1.1-3	0.31	3.15 50	6.59 99	2.92 33	2.33 17	2.33 14

high order discretization

h	error	disc	SOR	J-SI	J-CG	SSOR-SI	SSOR-CG
1/4	2.1-4	0.03	0.07 13	0.11 20	0.08 6	0.10 8	0.11 6
1/8	1.3-0	0.12	0.01 30	0.66 44	0.63 17	0.70 12	0.65 9
1/12	2.3-6	0.27	2.07 44	2.71 60	1.91 26	2.25 16	2.04 13
1/16	5.4-7	0.47	5.05 68	8.41 76	3.92 34	4.31 17	3.97 13

Problem 41   max solution = 1.98

five-point star discretization

h	error	disc	SOR	J-SI	J-CG	SSOR-SI	SSOR-CG
1/4	1.6-0	0.03	0.06 11	0.09 15	0.06 5	0.02 6	0.09 5
1/8	2.8-1	0.10	0.44 29	0.64 34	0.61 17	0.48 10	0.49 9
1/12	1.2-1	0.22	1.52 46	1.77 62	1.49 27	1.32 13	1.52 12
1/16	6.6-2	0.39	3.09 49	4.20 63	3.30 38	3.08 16	3.38 15

high order discretization

h	error	disc	SOR	J-SI	J-CG	SSOR-SI	SSOR-CG
1/4	6.2-2	0.05	0.06 11	0.10 16	0.07 5	0.10 7	0.10 6
1/8	5.4-3	0.21	0.66 24	0.71 34	0.61 16	0.66 11	0.71 10
1/12	1.6-3	0.47	1.71 53	2.26 47	1.87 25	2.26 16	2.20 13
1/16	5.9-4	0.81	4.42 48	5.41 62	4.17 34	4.46 16	4.86 15

Problem 4   max solution = 0.026

five-point star discretization

h	error	disc	SOR	J-SI	J-CG	SSOR-SI	SSOR-CG
1/4	1.1-3	0.03	0.06 12	0.10 21	0.06 5	0.08 8	0.12 65
1/8	4.9-4	0.11	0.39 25	0.65 42	0.53 16	0.45 11	0.43 8
1/12	2.9-4	0.27	1.53 48	2.24 69	1.44 28	1.38 16	1.31 11
1/16	2.0-4	0.40	4.05 69	5.05 86	2.97 38	2.99 17	2.73 13

high order discretization

h	error	disc	SOR	J-SI	J-CG	SSOR-SI	SSOR-CG
1/4	1.6-3	0.07	0.07 14	0.11 20	0.09 8	0.11 8	0.11 6
1/8	3.5-4	0.30	0.69 20	0.81 41	0.63 17	0.65 12	0.69 10
1/12	1.6-4	0.72	2.30 60	2.75 61	1.83 25	1.86 13	1.82 11
1/16	7.0-5	1.29	5.30 62	6.43 78	3.92 34	4.33 17	3.86 13

Problem 7   max solution = 1.00

five-point star discretization

h	error	disc	SOR	J-SI	J-CG	SSOR-SI	SSOR-CG
1/4	6.4-2	0.03	0.04 8	0.08 9	0.06 5	0.08 8	0.06 3
1/8	3.4-1	0.12	0.27 14	0.36 13	0.39 13	0.33 5	0.36 6
1/12	3.3-2	0.??	0.88 21	1.17 29	1.19 21	1.15 10	1.03 8
1/16	7.0-2	0.43	2.10 29	2.74 39	2.62 28	2.37 11	2.26 9

high order discretization

h	error	disc	SOR	J-SI	J-CG	SSOR-SI	SSOR-CG
1/4	9.6-3	0.07	0.06 8	0.07 9	0.07 5	0.10 7	0.08 4
1/8	3.7-3	0.26	0.40 14	0.48 18	0.51 13	0.49 7	0.45 6
1/12	1.1-3	0.56	1.30 21	1.64 29	1.62 19	1.76 11	1.46 8
1/16	3.8-4	0.96	3.16 28	3.87 37	3.42 25	3.69 12	3.44 10

Problem 5   max solution = 1.00

five-point star discretization

h	error	disc	SOR	J-SI	J-CG	SSOR-SI	SSOR-CG
1/4	5.0-1	0.07	0.05 13	0.09 18	0.06 5	0.08 7	0.09 6
1/8	1.2-1	0.11	0.44 30	0.65 44	0.59 20	0.51 12	0.55 10
1/12	5.0-2	0.22	1.37 40	2.23 70	1.65 32	1.67 16	1.51 13
1/16	7.2-2	0.41	3.55 59	5.34 93	3.41 43	3.23 18	3.38 16

high order discretization

h	error	disc	SOR	J-SI	J-CG	SSOR-SI	SSOR-CG
1/4	5.0-1	0.07	0.07 15	0.12 20	0.06 8	0.11 9	0.13 7
1/8	1.1-2	0.30	0.72 37	0.95 44	0.68 18	0.61 14	0.67 10
1/12	2.2-2	0.55	1.97 48	2.67 65	2.06 28	2.30 17	2.11 13
1/16	3.2-2	0.87	5.07 65	6.81 96	4.32 36	4.65 19	4.62 16

Problem 6   max solution = 3.30

five-point star discretization

h	error	disc	SOR	J-SI	J-CG	SSOR-SI	SSOR-CG
1/4	1.1-0	0.06	0.04 8	0.06 9	0.04 3	0.07 6	0.06 3
1/8	3.7-2	0.32	0.26 13	0.34 18	0.28 9	0.35 7	0.36 6
1/12	9.1-2	0.60	0.85 19	1.15 29	0.92 15	1.01 9	0.88 7
1/16	5.1-2	1.46	2.04 26	2.66 38	2.01 20	2.09 10	1.97 8

high order discretization

h	error	disc	SOR	J-SI	J-CG	SSOR-SI	SSOR-CG
1/4	7.1-1	0.26	0.05 8	0.07 7	0.07 5	0.10 7	0.07 3
1/8	351-2	1.28	0.36 13	0.46 17	0.43 10	0.62 9	0.46 6
1/12	7.6-3	3.42	1.27 19	1.53 26	1.30 15	1.42 8	1.46 6
1/16	2.3-3	6.23	3.22 22	3.84 37	2.98 20	3.69 12	2.97 6

Table 1 : Performance data for six test problems
Legend: max solution = max absolute value of solution at nodes; error = max absolute value of error at nodes; disc = time (sec.) to perform discretization; other columns give time (sec.) and number of iterations for iterative methods. An entry of the form n.m-k is to be interpreted as $n.m \times 10^{-k}$.

<u>Figure 1</u> : Plots of performance data

The following plots show log(max error at nodes) versus log(time in sec.) for each test problem whose data appears in Table 1. Time includes both discretization and solution time. For each method the best least squares line is fit through the data. The symbols used for each method are

second order discretization:

Y   SOR
Z   Jacobi-SI
X   Jacobi-CG
   SSOR-SI
◇   SSOR-CG

fourth order discretization:

X   SOR
+   Jacobi-SI
△   Jacobi-CG
⊙   SSOR-SI
⊡   SSOR-CG

PROBLEM 3

second order methods

fourth order methods

PROBLEM 4

second order methods

fourth order methods