Multigrid-based preconditioners for the heterogeneous Helmholtz equation

Yogi A. Erlangga, Cornelis Vuik, Cornelis W. Oosterlee
Department of Applied Mathematical Analysis
Delft University of Technology, The Netherlands
Email: Y.A.Erlangga@EWI.TUDelft.nl,

The research is financially supported by the Dutch Ministry of Economic Affairs under the project BTS01044
Objective of the talk:

A fast, robust iterative solver for the 2D Helmholtz equation in inhomogeneous media with high wave numbers (of typically 300–600 wave numbers in a unit square domain)

Outline:

- Problem definition: Helmholtz equation, discretization
- Krylov subspace method and multigrid
- Multigrid based preconditioner
- Numerical examples
- Conclusion
Helmholtz problem: Given $k = k(x,y)$, the wave number, in a domain $\Omega \subset \mathbb{R}^2$. Find $u \in \mathbb{C}$ such that the following PDE is satisfied:

$$ \mathcal{L}u \equiv \left( -\partial_{xx} - \partial_{yy} - (1 - i\alpha)k^2 \right) u = f \quad \text{in } \Omega $$

$$ \frac{\partial u}{\partial \nu} - ik u - i \frac{\partial^2 u}{2k \partial \tau^2} = 0 \quad \text{on } \Gamma = \partial \Omega $$

with:

- $\nu$ the outward normal direction to $\Gamma$,
- $\tau$ the tangential direction to $\Gamma$
- $k$ the wavenumber with $k = 2\pi f / c$. $f$ and $c$ are the frequency and the speed of sound, respectively.
- $0 \leq \alpha \ll 1$

Example of applications: aeroacoustics (sound produced by an engine), electromagnetics (lithophotography), geophysics (subsurface mapping).

Typical in geophysical applications:

- High wave number
- Extreme constrast of $k$
Few discretization methods:

- Finite difference
- Finite element

In our application, finite difference is used. This is common practice in geophysics in which simple domain is usually considered (e.g. rectangular).

- compact, 5 (or 9)-point finite difference stencil [Harari and Turkel, 1995]
- boundary: one-sided finite difference, $O(h)$ (or central difference, $O(h^2)$)

This leads to the linear system

$$Ax = b, \; A \in \mathbb{C}^{N \times N}, \; b, x \in \mathbb{C}^{N},$$

where $A$ is a sparse, large, highly indefinite symmetric matrix.

For high resolution solutions, very fine grid is required.
Numerical methods for solving $Ax = b$

We solve the linear system iteratively using a Krylov subspace method.

**Iterative methods**: Krylov subspace methods

**Definition**

$$\mathcal{K}^j(A, r_0) = \text{span}\{r_0, Ar_0, A^2r_0, \ldots, A^{j-1}r_0\}$$

where $\mathcal{K}^j(A, r_0)$ is the $j$-th Krylov subspace and $r_0 = b - Ax_0$ the initial residual related to the starting vector $x_0$.

- requires only matrix/vector multiplications (of order $O(N)$) (exploiting sparsity of $A$)

**Computational constraint:**

- The Krylov method is chosen such that the iteration can be performed within limited computer storage
- The Krylov method with constant amount of work per iteration is the method of choice
- With respect to the preconditioner, the method should be less stringent to the symmetry property of the preconditioning matrix

Currently we choose BiCGSTAB. [However, if $M^{-1}A$ is symmetric, COCG or SQMR may be the best choice]
Illustrative problem: The Helmholtz problem in homogeneous medium. \( \Omega = (0, 1)^2 \), \( k = 50 \), \( N = 250^2 \), with Sommerfeld’s condition at \( x = 0, 1 \) and \( y = 1 \), and Dirichlet condition at \( y = 0 \). Without residual smoother.

\[ x \text{-axis: number of iteration, } y \text{-axis: } \log_{10}(\|r\|_2/\|b\|_2) \]

For unpreconditioned case, COCG outperforms BiCGSTAB. GMRES is found to be less effective as compared to Bi-CGSTAB.
Multigrid is known as a good and efficient method for elliptic PDE. For Poisson equation, e.g., the complexity can be up to order $O(N)$.

Multigrid principles:

- Error smoothing → some iterative methods may have smoothing property of the error
- Coarse grid correction → a smooth error can be well approximated on the coarse grid

Multigrid for the Helmholtz equation: the problem is indefinite,

- standard iteration (e.g. Jacobi) does not converge. Multigrid only converges if the coarsest grid is fine enough to represent smooth frequencies → the convergence is limited by the coarsest grid. Therefore, the method is no longer $O(N)$.
- eigenvalues close to the origin may change signs → convergence degradation

**Example:** Standard multigrid, constant $k$, $N = 256^2$, 0.8-JAC, V(1,1)

<table>
<thead>
<tr>
<th>$k$</th>
<th>1</th>
<th>2.5</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iter</td>
<td>17</td>
<td>14</td>
<td>div</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.535</td>
<td>0.44</td>
<td>&gt;1.0</td>
</tr>
</tbody>
</table>
The idea: use multigrid as the preconditioner to accelerate the Krylov subspace iteration.

Meaning: solve

\[ AC^{-1} \tilde{x} = b, \quad \tilde{x} = Cx, \]

where \( C \in \mathbb{C}^{N \times N} \) is the preconditioner and \( C^{-1} \) is approximated by multigrid.

Example: \( C = A \). \( C^{-1} \) is approximated with one \( V(1,1) \) multigrid iteration

<table>
<thead>
<tr>
<th></th>
<th>k</th>
<th>1</th>
<th>2.5</th>
<th>5</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>GMRES</td>
<td>8</td>
<td>9</td>
<td>14</td>
<td>93</td>
<td></td>
</tr>
<tr>
<td>Bi-CGSTAB</td>
<td>5</td>
<td>5</td>
<td>11</td>
<td>&gt;100</td>
<td></td>
</tr>
</tbody>
</table>

Choosing \( C = A \) does not result in a good method. As the problem becomes more indefinite (more eigenvalues change signs) for larger \( k \), the method diverges.
Find $C$ such that:

- $AC^{-1}$ better conditioned than $A$
- $C^{-1}$ easy to solve using multigrid

Complex shift of the Laplace operator as the preconditioner:

$$M := -\partial_{xx} - \partial_{yy} - (\beta_1 - i \beta_2)k^2, \quad \beta_1, \beta_2 \in \mathbb{R}, \quad i = \sqrt{-1}.$$  

$C$ is built based on discretization of $M$.

For operator based preconditioner:

- the convergence of CG-like iteration can be made $h$-independent
- for $AC^{-1}$, $h$-independent convergence is obtained if the same boundary condition is used in $L$ and $M$. 
Case 1: \( \mathcal{M} \) is a complex, symmetric positive definite operator (CSPD). In this case \( \beta_1 \leq 0 \).

We find for \( AC^{-1} \) that [E, Vuik, Oosterlee, 2004]:

- \( |\lambda|_{\text{max}} \rightarrow 1 \)
- \( |\lambda|_{\text{min}} \sim \mathcal{O}(\varepsilon/k^2) \)
- the condition number \( \kappa \) is minimal if \( (\beta_1, \beta_2) = (0, \pm 1) \)

Example: Multigrid as solver for \( C_{(0,1)} \). Constant \( k, N = 256^2 \), 0.8-JAC, V(1,1)

<table>
<thead>
<tr>
<th>( k )</th>
<th>1</th>
<th>2.5</th>
<th>5</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iter</td>
<td>10</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>( \rho )</td>
<td>0.310</td>
<td>0.269</td>
<td>0.201</td>
<td>0.201</td>
</tr>
</tbody>
</table>
Case 2: \( \mathcal{M} \) is a more general operator, but is still suitable for multigrid convergence. For example: \( \beta_1 = 1 \) and \( \beta_2 \in \mathbb{R} \).

We investigate the following pairs: \((\beta_1, \beta_2) = (1, 1), (1, 0.5)\).
Problem 1: Homogeneous medium. \( \Omega = (0, 1)^2, k = 50, N = 250^2. \)

Convergence with respect to \( k \): Bi-CSGTAB, F(1,1) multigrid cycle. Second order absorbing condition.

Multigrid components:

- for \((\beta_1, \beta_2) = (0, 1)\): 0.8-JAC
- for \((\beta_1, \beta_2) = (1, 1)\): 0.7-JAC
- for \((\beta_1, \beta_2) = (1, 0.5)\): 0.5-JAC
- Coarse grid operator: Galerkin
- Prolongator: bi-linear interpolation/matrix dependent
- Restrictor: full weighting

<table>
<thead>
<tr>
<th>((\beta_1, \beta_2))</th>
<th>40</th>
<th>50</th>
<th>80</th>
<th>100</th>
<th>150</th>
<th>200</th>
<th>500</th>
<th>600</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0,1)</td>
<td>57 (0.44)</td>
<td>73 (0.92)</td>
<td>112 (4.3)</td>
<td>126 (7.7)</td>
<td>188 (28.5)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>(1,1)</td>
<td>36 (0.30)</td>
<td>39 (0.51)</td>
<td>54 (2.2)</td>
<td>74 (4.5)</td>
<td>90 (13.9)</td>
<td>114 (30.8)</td>
<td>291 (515)</td>
<td>352 (890)</td>
</tr>
<tr>
<td>(1,0.5)</td>
<td>26 (0.21)</td>
<td>31 (0.40)</td>
<td>44 (1.8)</td>
<td>52 (3.3)</td>
<td>73 (10.8)</td>
<td>92 (25.4)</td>
<td>250 (425)</td>
<td>298 (726)</td>
</tr>
</tbody>
</table>
Problem 2: Wedge problem. \( \Omega = (0, 600) \times (0, 1000) \) m\(^2\). Bi-CSSTAB. Second order absorbing condition.

Multigrid components:

- prolongator: matrix dependent
- restrictor: full weighting

<table>
<thead>
<tr>
<th>f (Hz)</th>
<th>Grid</th>
<th>((0,1))</th>
<th>((1,1))</th>
<th>((1,0.5))</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>75 \times 125</td>
<td>52 (1.2)</td>
<td>30 (0.67)</td>
<td>19 (0.42)</td>
</tr>
<tr>
<td>20</td>
<td>149 \times 249</td>
<td>91 (8.8)</td>
<td>45 (4.5)</td>
<td>27 (2.8)</td>
</tr>
<tr>
<td>40</td>
<td>301 \times 501</td>
<td>161 (66.1)</td>
<td>80 (33.5)</td>
<td>49 (20.8)</td>
</tr>
<tr>
<td>60</td>
<td>481 \times 801</td>
<td>232 (247.3)</td>
<td>118 (127.6)</td>
<td>66 (71.9)</td>
</tr>
</tbody>
</table>

In [Plessix, Mulder, 2004], Separation of Variables based preconditioner does not converge for \( f = 50 \) Hz after 2000 iterations.
Solution of wedge problem, $f = 50$ Hz.

![Velocities and solutions](image-url)
Numerical Example

Problem 3: Marmousi. $\Omega = (0,6000) \times (0,1600) \text{ m}^2$. Bi-CGSTAB and F(1,1) multigrid. Second order absorbing condition.

<table>
<thead>
<tr>
<th>f (Hz)</th>
<th>Grid</th>
<th>damping</th>
<th>(0,1)</th>
<th>(β₁, β₂)</th>
<th>(1,0.5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>751 × 201</td>
<td>0.0%</td>
<td>74 (37.5)</td>
<td>54 (27.6)</td>
<td>38 (19.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.0%</td>
<td>64 (32.5)</td>
<td>53 (27.3)</td>
<td>31 (16.5)</td>
</tr>
<tr>
<td>10</td>
<td>751 × 201</td>
<td>0.0%</td>
<td>180 (89.2)</td>
<td>84 (42.4)</td>
<td>47 (24.2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.0%</td>
<td>96 (48.3)</td>
<td>48 (24.8)</td>
<td>28 (15.0)</td>
</tr>
<tr>
<td>20</td>
<td>1501 × 401</td>
<td>0.0%</td>
<td>414 (832.3)</td>
<td>168 (308.7)</td>
<td>104 (212.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.0%</td>
<td>145 (294.7)</td>
<td>64 (133.4)</td>
<td>37 (79.3)</td>
</tr>
<tr>
<td>30</td>
<td>2001 × 534</td>
<td>0.0%</td>
<td>458 (1724.8)</td>
<td>211 (799.4)</td>
<td>136 (519.4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.0%</td>
<td>119 (455.3)</td>
<td>61 (238.4)</td>
<td>38 (151.9)</td>
</tr>
</tbody>
</table>

In [Plessix, Mulder, 2004], Separation of Variables based preconditioner does not converge for $f = 30$ Hz after 2000 iterations.
Numerical Example

Solution of Marmousi problem, $f = 30$ Hz. Figure (b): 0 % damping. (c): 5 % damping.

Workshop on the frequency domain wave equation. Delft, 24 September 2004 (slide 16)
• A new class of preconditioner for the Helmholtz equation based on Shifted-Laplace operator is proposed and analyzed.
• This class of preconditioners clusters the eigenvalues around the origin, which give some very small eigenvalues.
• Numerical tests show the effectiveness of the preconditioners.
• Extension to varying $k$: can be done easily.