Outline

1. Problem Formulation
2. Iterative Methods
   - Iterative Solvers
   - Preconditioning Techniques
   - Multilevel Krylov Multigrid Method
3. Fourier Analysis
   - Theory
   - Analysis of the Preconditioning
   - Multigrid Analysis
   - Multigrid Convergence
4. Numerical Experiments
5. Future Work
Helmholtz Problem

Helmholtz equation

\[-\Delta u(x) - k^2 u(x) = f(x) \text{ in } \Omega \in \mathbb{R}^3\]

Boundary condition

- Dirichlet / Neumann / Sommerfeld

Discretization

- finite difference method / finite element method

Linear system

- sparse
- symmetric but non-Hermitian
Problem Formulation

Thesis Work

Objective

spectral properties $\implies$ convergence behaviour

Task

1. Preconditioning techniques
   - shifted Laplacian preconditioner $M$
   - deflation operator $P$ and $Q$

2. Iterative solver
   - multigrid method for $M^{-1}$
   - Krylov subspace method for $Ax = b / AM^{-1}x = b / AM^{-1}Qx = b$

3. Fourier analysis
   - spectrum distribution
   - convergence factor

4. Numerical solution
Objective

spectral properties $\rightarrow$ convergence behaviour

Task

1. Preconditioning techniques
   - shifted Laplacian preconditioner $M$
   - deflation operator $P$ and $Q$

2. Iterative solver
   - multigrid method for $M^{-1}$
   - Krylov subspace method for $Ax = b / AM^{-1}x = b / AM^{-1}Qx = b$

3. Fourier analysis
   - spectrum distribution
   - convergence factor

4. Numerical solution
Model Problem

1D dimensionless Helmholtz problem with homogeneous Dirichlet boundary condition

\[
\begin{cases}
-\Delta u(x) - k^2 u(x) = f(x) \quad \text{for } x \in (0, 1), \\
u(0) = u(1) = 0.
\end{cases}
\]

The resulting linear system

\[
Ax = b \quad \text{where } A = \frac{1}{h^2} \begin{bmatrix}
2 & -1 & & \\
-1 & 2 & -1 & \\
& \ddots & \ddots & \ddots \\
& -1 & 2 & -1
\end{bmatrix} - k^2 I
\]

Wave resolution

\[
gw \cdot h = \frac{2\pi}{k}
\]
Model Problem

Eigenvalue

$$\lambda_l = \frac{4}{h^2} \sin^2 \left( \frac{l\pi h}{2} \right) - k^2 \quad \text{for } l = 1, 2, \cdots, n$$

Difficulty in solving Helmholtz problem
Multigrid Method

The solver for inverting the shifted Laplacian preconditioner $M$

- The coarsening strategy is done by doubling the mesh size, i.e. $\Omega_h \rightarrow \Omega_{2h}$.
- The smoother is $\omega$-Jacobi iteration operator.
  - $\omega$ is chosen as the optimal one $\omega_{\text{opt}}$.
- The intergrid transfer
  - restriction by full weighting operator
  - prolongation by linear interpolation operator

The failure of MG in solving $Ax = b$

- The coarse grid cannot cope with high wavenumber problem.
- The $\omega$-Jacobi iteration does not converge.
Multigrid Method

The solver for inverting the shifted Laplacian preconditioner $M$

- The coarsening strategy is done by doubling the mesh size, i.e. $\Omega_h \rightarrow \Omega_{2h}$.
- The smoother is $\omega$-Jacobi iteration operator.
  - $\omega$ is chosen as the optimal one $\omega_{opt}$.
- The intergrid transfer
  - restriction by full weighting operator
  - prolongation by linear interpolation operator

The failure of MG in solving $Ax = b$

- The coarse grid cannot cope with high wavenumber problem.
- The $\omega$-Jacobi iteration does not converge.
Krylov Subspace Methods

The solver for solving the linear system $Ax = b$

- **GMRES**, used in the thesis work
- CG
- BiCGStab
- GCR, IDR(s),...
Approximated Inversion

Given the iteration operator $G$, there is the approximated inversion

$$A^{-1} = (I - G)A^{-1}.$$  

For the stationary iteration, there is

$$A^{-1}_m = (I - G^m)A^{-1}.$$  

For the multigrid iteration, there is

$$A^{-1}_{MG} = (I - T_1^m)A^{-1}.$$
Shifted Laplacian Preconditioner

\[ M := -\Delta_h - (\beta_1 + i\beta_2)k^2I \]

- preconditioned system
  \[ \hat{A} := AM^{-1} = M^{-1}A \quad \text{and} \quad \sigma(AM^{-1}) = \sigma(M^{-1}A) \]
- preservation of symmetry
  \[ (AM^{-1})^T = AM^{-1} \]
- circular spectrum distribution
  \[ \left(\lambda_r - \frac{1}{2}\right)^2 + \left(\lambda_i - \frac{\beta_1 - 1}{2\beta_2}\right)^2 = \frac{\beta_2^2 + (1 - \beta_1)^2}{(2\beta_2)^2} \]
Shifted Laplacian Preconditioner

\[ M := -\Delta_h - (\beta_1 + i\beta_2)k^2I \]

- preconditioned system

\[ \hat{A} := AM^{-1} = M^{-1}A \quad \text{and} \quad \sigma(AM^{-1}) = \sigma(M^{-1}A) \]

- preservation of symmetry

\[ (AM^{-1})^T = AM^{-1} \]

- circular spectrum distribution

\[ (\lambda_r - \frac{1}{2})^2 + (\lambda_i - \frac{\beta_1 - 1}{2\beta_2})^2 = \frac{\beta_2^2 + (1 - \beta_1)^2}{(2\beta_2)^2} \]
\[ \beta_1 = 1 \implies \text{the most compact distribution} \]

The spectrum distributions of the preconditioned matrix \( AM^{-1} \) with respect to several typical shifts when \( k = 100 \)
Deflation Operator

For an invertible $\hat{A}$, take any $n \times r$ full rank matrices $Y$ and $Z$

\[
\begin{align*}
\text{left} & \quad P := I - \hat{A} Z \hat{E}^{-1} Y^T + \lambda_d Z \hat{E}^{-1} Y^T, \\
\text{right} & \quad Q := I - Z \hat{E}^{-1} Y^T \hat{A} + \lambda_d Z \hat{E}^{-1} Y^T,
\end{align*}
\]

where $\hat{E} = Y^T \hat{A} Z$.

The spectrum distributions of the deflated matrix $\hat{A}Q$ towards $\lambda_d = 0.2$

where $k = 100$, shift = 1 − $\nu1$, $AZ = Z \Lambda_r$ and $Y^T \hat{A} = \Lambda_r Y^T$
Deflation Operator

\[ \sigma(\hat{A}) = \{\lambda_1, \cdots, \lambda_n\} \text{ with } |\lambda_1| \leq \cdots \leq |\lambda_n| \]

- Projector in case of \( \lambda_d = 0 \)
  \[
P_D \cdot P_D = P_D \quad \text{and} \quad Q_D \cdot Q_D = Q_D.
  \]

- Preservation of symmetry in case of \( \lambda_d = 0 \),

- Spectrum distribution
  \[
  \sigma(P\hat{A}) = \sigma(\hat{AQ}) = \{\lambda_d, \cdots, \lambda_d, \mu_{r+1}, \cdots, \mu_n\}.
  \]

- Condition number
  \[
  \kappa(P\hat{A}) = \frac{|\mu_n|}{\min\{|\lambda_d|, |\mu_{r+1}|\}} \quad \text{in case of } \lambda_d \neq 0,
  \]
Deflation Operator

Inaccuracy in $\hat{E}^{-1}$

Assume $AZ = Z\Lambda_r$ and $Y^T\hat{A} = \Lambda_r Y^T$, then $\hat{E} = Y^T\hat{A}Z = \Lambda_r$.

$$\hat{E}^{-1} = \text{diag}\left(\frac{1 - \epsilon_1}{\lambda_1}, \ldots, \frac{1 - \epsilon_r}{\lambda_r}\right)$$
Deflation Operator

Inaccuracy in $\hat{E}^{-1}$

Assume $AZ = Z \Lambda_r$ and $Y^T \hat{A} = \Lambda_r Y^T$, then $\hat{E} = Y^T \hat{A} Z = \Lambda_r$.

$$\hat{E}^{-1} = \text{diag}\left( \frac{1 - \epsilon_1}{\lambda_1}, \ldots, \frac{1 - \epsilon_r}{\lambda_r} \right)$$

$$\sigma(P\hat{A}) = \{(1 - \epsilon_1)\lambda_d + \lambda_1\epsilon_1, \ldots, (1 - \epsilon_r)\lambda_d + \lambda_r\epsilon_r, \lambda_{r+1}, \ldots, \lambda_n\}$$
Deflation Operator

Inaccuracy in $\hat{E}^{-1}$

Assume $AZ = Z \Lambda_r$ and $Y^T \hat{A} = \Lambda_r Y^T$, then $\hat{E} = Y^T \hat{A}Z = \Lambda_r$.

\[
\hat{E}^{-1} = diag\left(\frac{1 - \epsilon_1}{\lambda_1}, \ldots, \frac{1 - \epsilon_r}{\lambda_r}\right)
\]

\[
\Rightarrow \sigma(P\hat{A}) = \{(1 - \epsilon_1)\lambda_d + \lambda_1\epsilon_1, \ldots, (1 - \epsilon_r)\lambda_d + \lambda_r\epsilon_r, \lambda_{r+1}, \ldots, \lambda_n\}
\]

\[
\begin{cases}
\lambda_d = 0 & \Rightarrow \sigma(P\hat{A}) = \{\lambda_1\epsilon_1, \ldots, \lambda_r\epsilon_r, \lambda_{r+1}, \ldots, \lambda_n\}, \\
\lambda_d = \lambda_n & \Rightarrow \sigma(P\hat{A}) \approx \{\lambda_n, \ldots, \lambda_n, \lambda_{r+1}, \ldots, \lambda_n\}.
\end{cases}
\]
Multilevel Krylov Multigrid Method

A recursive Krylov solution of $\hat{E}^{-1}$

1. Use the approximation

$$M^{-1} \approx Z(Y^T M Z)^{-1} Y^T.$$  

2. Take the replacement

$$\hat{E} := Y^T \hat{A} Z = Y^T A M^{-1} Z \approx Y^T A Z (Y^T M Z)^{-1} Y^T Z$$

$$\hat{E}^{-1} \approx \left( A(2) M^{-1} B(2) \right)^{-1}$$

3. Solve $A^{-1}_{(2)}$ in the same way as $A^{-1}$
The illustration of multilevel Krylov multigrid method in a five-level grid
Principles of Fourier Analysis

Find out a subspace $E = \text{span}\{\phi_1, \cdots, \phi_m\}$ such that

$$KE \subset E \implies K \Phi = \Phi \tilde{K}.$$  

For any $v = \Phi c \in E$, there is

$$Kv = K \Phi c = \Phi \tilde{K} c \quad \text{where} \quad \tilde{K} \text{ amplifies} \ c.$$  

Assume $E$ is the union of several disjoint subspaces. Then, there is a diagonal block matrix

$$K \triangleq [\tilde{K}^l] \quad \text{with} \ l \text{ as the block index.}$$
Principles of Fourier Analysis

Find out a subspace $E = \text{span}\{\phi_1, \cdots, \phi_m\}$ such that 

$$KE \subset E \implies K\Phi = \Phi \tilde{K}.$$ 

For any $v = \Phi c \in E$, there is 

$$K v = K \Phi c = \Phi \tilde{K} c \quad \text{where } \tilde{K} \text{ amplifies } c.$$ 

Assume $E$ is the union of several disjoint subspaces. Then, there is a diagonal block matrix 

$$K : \overset{\Delta}{=} [\tilde{K}^l] \quad \text{with } l \text{ as the block index.}$$
Fourier Analysis for Multigrid Analysis

In a two-level grid, the invariance subspace is given by

\[ E^l_{h} := \text{span}\{\phi^l_{h}, \phi^{n-l}_{h}\} \text{ in } \Omega_{h} \quad \implies \quad E^l_{2h} := \text{span}\{\phi^l_{2h}\} \text{ in } \Omega_{2h}. \]

\[
\begin{align*}
A_1, M_1, S & : E^l_{h} \to E^l_{h} \\
A_2, M_2 & : E^l_{2h} \to E^l_{2h} \\
R^2_1 & : E^l_{h} \to E^l_{2h} \\
P^1_2 & : E^l_{2h} \to E^l_{h}
\end{align*}
\]

\[ \implies \quad T^2_1 : E^l_{h} \to E^l_{h} \]

In a multilevel grid, there is

\[ \tilde{T}^m_k = \tilde{S}^{\nu_2}_k (I - \tilde{P}^k_{k+1} (I - \tilde{T}^m_{k+1}) \tilde{M}^{-1}_{k+1} \tilde{R}^{k+1}_{k} \tilde{M}_k) \tilde{S}^{\nu_1}_k \quad \text{with } \tilde{T}^m_m = 0, \]
Fourier Analysis for Multigrid Analysis

In a two-level grid, the invariance subspace is given by

$$E_h^l := \text{span}\{\phi^l_h, \phi^{n-l}_h\} \text{ in } \Omega_h$$ \implies $$E_{2h}^l := \text{span}\{\phi_{2h}^l\} \text{ in } \Omega_{2h}.$$  

\[
\begin{align*}
A_1, M_1, S &: E_h^l \to E_h^l \\
A_2, M_2 &: E_{2h}^l \to E_{2h}^l \\
R_1^2 &: E_h^l \to E_{2h}^l \\
P_2^1 &: E_{2h}^l \to E_h^l
\end{align*}
\] \implies $$T_1^2 : E_h^l \to E_h^l$$

In a multilevel grid, there is

$$\tilde{T}_k^m = \tilde{S}_k^{\nu_2} \left( I - \tilde{P}_{k+1}^k (I - \tilde{T}_{k+1}^m) \tilde{M}_{k+1}^{-1} \tilde{R}_{k+1}^{k+1} \tilde{M}_k \right) \tilde{S}_k^{\nu_1} \text{ with } \tilde{T}_m^m = 0,$$
Application to Preconditioning

Shifted Laplacian preconditioner

\[ \tilde{A} = \tilde{A}\tilde{M}^{-1} \]

Deflation operator

\[ \tilde{Q} = I - \tilde{P}_2^1 \tilde{E}_2 \tilde{R}_2^1 (\lambda_n I - \tilde{A}) \quad \text{with} \quad \tilde{E}_2 = \tilde{R}_2^2 \tilde{A}\tilde{P}_2^1 \]

Advantage of Fourier analysis

- computational time
- memory requirement
- accuracy
Application to Preconditioning

Shifted Laplacian preconditioner

\[ \tilde{A} = \tilde{A} \tilde{M}^{-1} \]

Deflation operator

\[ \tilde{Q} = I - \tilde{P}_2 \tilde{E}_2 \tilde{R}_2 (\lambda_n I - \tilde{A}) \quad \text{with} \quad \tilde{E}_2 = \tilde{R}_1 \tilde{A} \tilde{P}_2 \]

Advantage of Fourier analysis

- computational time
- memory requirement
- accuracy
Basic Preconditioning Effect

- The spectrum of $AM^{-1}$ is restricted to a circular distribution.

- The spectrum of $AM^{-1}Q$ is clustered around $(1, 0)$.
Choice of Shift $\beta_1 + \iota \beta_2$

1. $\beta_2$ determines the length of arc on which the eigenvalues of $AM^{-1}$ are located.

2. $\beta_2$ has the indirect influence on the tightness of spectrum distribution of $AM^{-1}Q$. 
Influence of Wave resolution $gw$

1. High resolution exerts little negative influence on the spectrum distribution of $AM^{-1}$.

2. High resolution results in a more favourable spectrum distribution of $AM^{-1}Q$.
Approximated Shifted Laplacian Preconditioning

\[ A(M^{-1} = A(I - T_1^m)M^{-1} \]

- The multigrid introduces disturbance to the preconditioning effect.

- The disturbance can be easily corrected by several iterations at a cheap cost.
Approximated Deflation Preconditioning

\( AM^{-1}Q \) where the construction of \( Q \) is based on the \( IM^{-1} \)

- The preconditioning \( AM^{-1}Q \) is much more sensitive to the accuracy in the approximation of \( M^{-1} \).
Multigrid Convergence Factor

- independence of $k$
- independence of the sign of $\beta_2$

$k = 30$ and $gw = 30$

$k = 30$ and $gw = 120$

High resolution is favourable for the convergence.
Optimal Shift for the Preconditioner

- A small shift is favourable for the Krylov convergence of $AM^{-1}$.
- A large shift is favourable for the multigrid convergence of $M^{-1}$.

To find out

$$(\beta_1 + i\beta_2)_{\text{opt}} := \arg\min\{|\beta_1 + i\beta_2| : \max_{1 \leq l \leq n-1} G(l, \beta_1, \beta_2) \leq c < 1\}.$$ 

<table>
<thead>
<tr>
<th>gw = 10</th>
<th>gw = 30</th>
<th>gw = 60</th>
<th>gw = 120</th>
<th>gw = 240</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m = 2$</td>
<td>0.1096</td>
<td>0.0126</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$m = 3$</td>
<td>0.3228</td>
<td>0.0616</td>
<td>0.0150</td>
<td>0</td>
</tr>
<tr>
<td>$m = 4$</td>
<td>0.3931</td>
<td>0.2002</td>
<td>0.0632</td>
<td>0.0155</td>
</tr>
<tr>
<td>$m = 5$</td>
<td>0.3931</td>
<td>0.2886</td>
<td>0.2012</td>
<td>0.0636</td>
</tr>
</tbody>
</table>

The optimal $\beta_2$ in the shift $1 + i\beta_2$ for $\rho(T_1^m) \leq c = 0.9$
Numerical Experiments

Basic Convergence Behaviour

Overview of the convergence behaviour by different preconditioning

- verify
  - the different preconditioning effect
  - the advantage of $AM^{-1}Q$ over $AM^{-1}$
  - the influence of wavenumber and wave resolution
Influence of Orthogonalization

- Householder reflection outperforms modified Gram-Schmidt in convergence behaviour.
- GMRES using modified Gram-Schmidt fails to converge in the very small system.

very small system

normal size system
Influence of Approximated Preconditioning

- The inaccuracy in $M^{-1}$ has little influence on the convergence behaviour

<table>
<thead>
<tr>
<th>$gw$</th>
<th>$k = 10$</th>
<th>$k = 50$</th>
<th>$k = 100$</th>
<th>$k = 200$</th>
<th>$k = 300$</th>
<th>$k = 400$</th>
<th>$k = 500$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>11/11</td>
<td>36/36</td>
<td>60/60</td>
<td>108/105</td>
<td>153/149</td>
<td>193/188</td>
<td>265/258</td>
</tr>
<tr>
<td>30</td>
<td>12/12</td>
<td>36/36</td>
<td>60/58</td>
<td>114/108</td>
<td>161/152</td>
<td>209/196</td>
<td>255/240</td>
</tr>
<tr>
<td>60</td>
<td>12/12</td>
<td>36/36</td>
<td>63/62</td>
<td>113/111</td>
<td>161/158</td>
<td>207/204</td>
<td>255/250</td>
</tr>
</tbody>
</table>

Number of iterations with respect to different degrees of approximations i.e. $AM^{-1} / AIM^{-1}$

- The inaccuracy in $Q$ slows down the convergence.

<table>
<thead>
<tr>
<th>$gw$</th>
<th>$k = 10$</th>
<th>$k = 50$</th>
<th>$k = 100$</th>
<th>$k = 200$</th>
<th>$k = 300$</th>
<th>$k = 400$</th>
<th>$k = 500$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>6/9/9</td>
<td>11/18/18</td>
<td>14/26/27</td>
<td>21/43/44</td>
<td>28/59/61</td>
<td>33/71/70</td>
<td>39/98/111</td>
</tr>
</tbody>
</table>

Number of iterations with respect to different degrees of approximations i.e. $AM^{-1}Q / AM^{-1}Q / AIM^{-1}Q$
Internal Iteration in MKMG

\((\Diamond, \#_2, \cdots, \#_{m-1}, \circ)\)

- It is worth doing more iterations on the higher levels.
- The convergence behaviour will be slowed by more iterations on the lower levels.

Number of iterations with respect to different MKMG setup in a six-level grid
Suggestion on Future Work

- higher dimensional problems
- local Fourier analysis
- different Krylov solvers
Thank you for watching!