Fast Iterative Solution of the Time-Harmonic Elastic Wave Equation at Multiple Frequencies

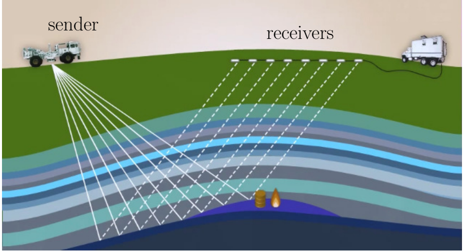
Manuel M. Baumann

Email: m.m.baumann@tudelft.nl
Delft Institute of Applied Mathematics
Delft University of Technology
Delft, The Netherlands

Royal Institute of Technology in Stockholm - February 1, 2018

Nice to be back!

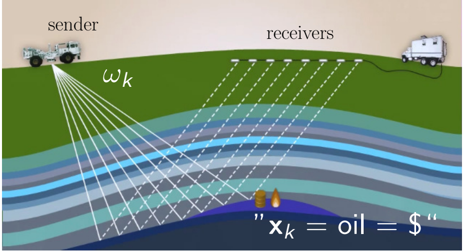
Seismic Full-Waveform Inversion



Inverse problem:

- least-squares fit of measurements and simulations,
- requires *fast* forward solves.

Seismic Full-Waveform Inversion



Inverse problem:

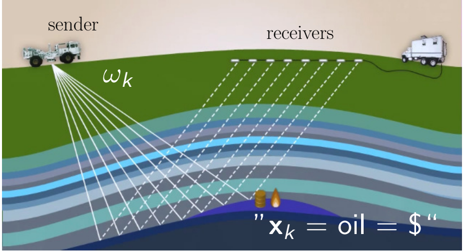
- least-squares fit of measurements and simulations,
- requires *fast* forward solves.

"Solve the linear systems of equations,

$$(K + i\omega_k C - \omega_k^2 M)\mathbf{x}_k = \mathbf{b},$$

efficiently (= fast and at low memory) for multiple frequencies."

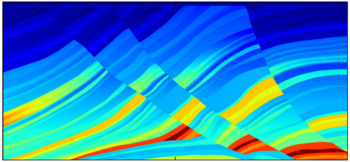
Seismic Full-Waveform Inversion



Inverse problem:

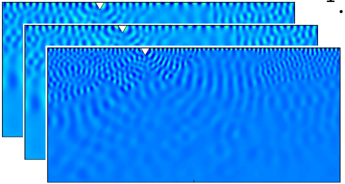
- least-squares fit of measurements and simulations,
- requires *fast* forward solves.

Density distribution



$\leftrightharpoons$

Simulations



The time-harmonic elastic wave equation

Continuous setting

Elastic wave equation

$$-\omega_k^2 \rho(\mathbf{x}) \mathbf{u}_k - \nabla \cdot \sigma(\mathbf{u}_k) = \mathbf{s}$$

with boundary conditions

$$\begin{aligned} i\omega_k \rho(\mathbf{x}) B \mathbf{u}_k + \sigma(\mathbf{u}_k)\hat{\mathbf{n}} &= \mathbf{0}, \\ \sigma(\mathbf{u}_k)\hat{\mathbf{n}} &= \mathbf{0}, \end{aligned}$$

on $\partial\Omega_a \cup \partial\Omega_r$.

Discrete setting

Solve

$$(K + i\omega_k C - \omega_k^2 M)\mathbf{u}_k = \mathbf{s}$$

with FEM matrices

$$\begin{aligned} K_{ij} &= \int_\Omega \sigma(\varphi_i) : \nabla \varphi_j \, d\Omega, \\ M_{ij} &= \int_\Omega \rho(\mathbf{x}) \varphi_i \cdot \varphi_j \, d\Omega, \\ C_{ij} &= \int_{\partial\Omega_a} \rho(\mathbf{x}) B \varphi_i \cdot \varphi_j \, d\Gamma. \end{aligned}$$

TU Delft

Shifted systems vs. matrix equation

Two main approaches for solving,

$$\boxed{(K + i\omega_k C - \omega_k^2 M)\mathbf{x}_k = \mathbf{b}, \quad k > 1.}$$

Shifted systems vs. matrix equation

Two main approaches for solving,

$$\boxed{(K + i\omega_k C - \omega_k^2 M)\mathbf{x}_k = \mathbf{b}, \quad k > 1.}$$

Shifted systems

$$\left(\begin{bmatrix} iC & K \\ I & 0 \end{bmatrix} - \omega_k \begin{bmatrix} M & 0 \\ 0 & I \end{bmatrix} \right) \begin{bmatrix} \omega_k \mathbf{x}_k \\ \mathbf{x}_k \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix}$$

- Most work for $\mathbf{x}_0$ (at $\omega = 0$)
- Requires preconditioning

Shifted systems vs. matrix equation

Two main approaches for solving,

$$\boxed{(K + i\omega_k C - \omega_k^2 M)\mathbf{x}_k = \mathbf{b}, \quad k > 1.}$$

Shifted systems

$$\left(\begin{bmatrix} iC & K \\ I & 0 \end{bmatrix} \text{-}\ \omega_k \begin{bmatrix} M & 0 \\ 0 & I \end{bmatrix}\right) \begin{bmatrix} \omega_k \mathbf{x}_k \\ \mathbf{x}_k \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix}$$

- Most work for $\mathbf{x}_0$ (at $\omega = 0$)
- Requires preconditioning

Matrix equation

$$K\mathbf{X} + iC\mathbf{X}\Omega - M\mathbf{X}\Omega^2 = \mathbf{B}$$

- Solve for $\mathbf{X} = [\mathbf{x}_1, ..., \mathbf{x}_N]$ *all-at-once*
- Requires preconditioning

TUDelft

Shifted systems vs. matrix equation

Two main approaches for solving,

$$\boxed{(K + i\omega_k C - \omega_k^2 M)\mathbf{x}_k = \mathbf{b}, \quad k > 1.}$$

Shifted systems

$$\left(\begin{bmatrix} iC & K \\ I & 0 \end{bmatrix} - \omega_k \begin{bmatrix} M & 0 \\ 0 & I \end{bmatrix} \right) \begin{bmatrix} \omega_k \mathbf{x}_k \\ \mathbf{x}_k \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix}$$

- Most work for $\mathbf{x}_0$ (at $\omega = 0$)
- Requires preconditioning

Matrix equation

$$K\mathbf{X} + iC\mathbf{X}\Omega - M\mathbf{X}\Omega^2 = \mathbf{B}$$

- Solve for $\mathbf{X} = [\mathbf{x}_1, ..., \mathbf{x}_N]$ *all-at-once*
- Requires preconditioning

Table of Content

TUDelft

What's a shifted linear system?

Definition

Shifted linear systems are of the form

$$(A - \omega I)\mathbf{x}^{(\omega)} = \mathbf{b},$$

where $\omega \in \mathbb{C}$ is the *shift*.

For the simultaneous solution, **Krylov methods** are well-suited because of the *shift-invariance* property:

$$\mathcal{K}_m(A, \mathbf{b}) \equiv \text{span}\{\mathbf{b}, A\mathbf{b}, ..., A^{m-1}\mathbf{b}\} = \mathcal{K}_m(A - \omega I, \mathbf{b}).$$

"Proof" (shift-invariance)

For $m = 2$: $\mathcal{K}_2(A, \mathbf{b}) = \text{span}\{\mathbf{b}, A\mathbf{b}\}$

$\mathcal{K}_2(A - \omega I, \mathbf{b}) = \text{span}\{\mathbf{b}, A\mathbf{b} - \omega\mathbf{b}\} = \text{span}\{\mathbf{b}, A\mathbf{b}\}$

What's a shifted linear system?

Definition

Shifted linear systems are of the form

$$(A - \omega I)\mathbf{x}^{(\omega)} = \mathbf{b},$$

where $\omega \in \mathbb{C}$ is the *shift*.

For the simultaneous solution, **Krylov methods** are well-suited because of the *shift-invariance* property:

$$\mathcal{K}_m(A, \mathbf{b}) \equiv \text{span}\{\mathbf{b}, A\mathbf{b}, ..., A^{m-1}\mathbf{b}\} = \mathcal{K}_m(A - \omega I, \mathbf{b}).$$

"Proof" (shift-invariance)

For $m = 2$: $\mathcal{K}_2(A, \mathbf{b}) = \text{span}\{\mathbf{b}, A\mathbf{b}\}$

$$\mathcal{K}_2(A - \omega I, \mathbf{b}) = \text{span}\{\mathbf{b}, A\mathbf{b} - \omega\mathbf{b}\} = \text{span}\{\mathbf{b}, A\mathbf{b}\}$$

What's a shifted linear system?

Definition

Shifted linear systems are of the form

$$(A - \omega I)\mathbf{x}^{(\omega)} = \mathbf{b},$$

where $\omega \in \mathbb{C}$ is the *shift*.

For the simultaneous solution, **Krylov methods** are well-suited because of the *shift-invariance* property:

$$\mathcal{K}_m(A, \mathbf{b}) \equiv \text{span}\{\mathbf{b}, A\mathbf{b}, ..., A^{m-1}\mathbf{b}\} = \mathcal{K}_m(A - \omega I, \mathbf{b}).$$

"Proof" (shift-invariance)

For $m = 2$: $\mathcal{K}_2(A, \mathbf{b}) = \text{span}\{\mathbf{b}, A\mathbf{b}\}$

$$\mathcal{K}_2(A - \omega I, \mathbf{b}) = \text{span}\{\mathbf{b}, A\mathbf{b} - \omega\mathbf{b}\} = \text{span}\{\mathbf{b}, A\mathbf{b}\}$$

Multi-shift GMRES

[Frommer/Glässner, 1998]

After m steps of Arnoldi, we have,

$$AV_m = V_{m+1}\underline{H}_m,$$

and the approximate solution yields:

$$\mathbf{x}_m \approx V_m\mathbf{y}_m, \quad \text{where } \mathbf{y}_m = \operatorname*{argmin}_{\mathbf{y}\in\mathbb{C}^m} \left\|\underline{H}_m\mathbf{y} - \|\mathbf{b}\|\mathbf{e}_1\right\|.$$

For shifted systems, we get

$$(A - \omega I)V_m = V_{m+1}(\underline{H}_m - \omega\underline{I}_m),$$

and, therefore,

$$\mathbf{x}_m^{(\omega)} \approx V_m\mathbf{y}_m^{(\omega)}, \quad \text{where } \mathbf{y}_m^{(\omega)} = \operatorname*{argmin}_{\mathbf{y}\in\mathbb{C}^m} \left\|\underline{H}_m^{(\omega)}\mathbf{y} - \|\mathbf{b}\|\mathbf{e}_1\right\|.$$

Multi-shift GMRES

[Frommer/Glässner, 1998]

After m steps of Arnoldi, we have,

$$AV_m = V_{m+1}\underline{\mathrm{H}}_m,$$

and the approximate solution yields:

$$\mathbf{x}_m \approx V_m \mathbf{y}_m, \quad \text{where } \mathbf{y}_m = \underset{\mathbf{y}\in\mathbb{C}^m}{\operatorname{argmin}} \left\| \underline{\mathrm{H}}_m \mathbf{y} - \|\mathbf{b}\| \mathbf{e}_1 \right\|.$$

For shifted systems, we get

$$(A - \omega I)V_m = V_{m+1}(\underline{\mathrm{H}}_m - \omega \underline{\mathrm{I}}_m),$$

and, therefore,

$$\mathbf{x}_m^{(\omega)} \approx V_m \mathbf{y}_m^{(\omega)}, \quad \text{where } \mathbf{y}_m^{(\omega)} = \underset{\mathbf{y}\in\mathbb{C}^m}{\operatorname{argmin}} \left\| \underline{\mathrm{H}}_m^{(\omega)} \mathbf{y} - \|\mathbf{b}\| \mathbf{e}_1 \right\|.$$

Multi-shift GMRES

[Frommer/Glässner, 1998]

After m steps of Arnoldi, we have,

$$AV_m = V_{m+1}\underline{\mathrm{H}}_m,$$

and the approximate solution yields:

$$\mathbf{x}_m \approx V_m \mathbf{y}_m, \quad \text{where } \mathbf{y}_m = \underset{\mathbf{y}\in\mathbb{C}^m}{\operatorname{argmin}} \left\| \underline{\mathrm{H}}_m \mathbf{y} - \|\mathbf{b}\|\mathbf{e}_1 \right\|.$$

For shifted systems, we get

$$(A - \omega I)V_m = V_{m+1}(\underline{\mathrm{H}}_m - \omega \underline{\mathrm{I}}_m),$$

and, therefore,

$$\mathbf{x}_m^{(\omega)} \approx V_m \mathbf{y}_m^{(\omega)}, \quad \text{where } \mathbf{y}_m^{(\omega)} = \underset{\mathbf{y}\in\mathbb{C}^m}{\operatorname{argmin}} \left\| \underline{\mathrm{H}}_m^{(\omega)} \mathbf{y} - \|\mathbf{b}\|\mathbf{e}_1 \right\|.$$

The shift-and-invert preconditioner

$$\mathcal{A} := \begin{bmatrix} iC & K \\ I & 0 \end{bmatrix}, \quad \mathcal{B} := \begin{bmatrix} M & 0 \\ 0 & I \end{bmatrix}$$

The multi-frequency problem reads,

$$(\mathcal{A} - \omega_k \mathcal{B})\mathbf{x}_k = \mathbf{b}, \quad k = 1, ..., N_\omega,$$

with a single preconditioner $\mathcal{P}(\tau) := (\mathcal{A} - \tau\mathcal{B})$:

$$(\mathcal{A} - \omega_k \mathcal{B})\mathcal{P}_k^{-1}\mathbf{y}_k = \mathbf{b} \quad \Leftrightarrow \quad \boxed{(\mathcal{A}(\mathcal{A} - \tau\mathcal{B})^{-1} - \eta_k I)\mathbf{y}_k = \mathbf{b}}$$

The shift-and-invert preconditioner

$$\mathcal{A} := \begin{bmatrix} iC & K \\ I & 0 \end{bmatrix}, \quad \mathcal{B} := \begin{bmatrix} M & 0 \\ 0 & I \end{bmatrix}$$

The multi-frequency problem reads,

$$(\mathcal{A} - \omega_k \mathcal{B})\mathbf{x}_k = \mathbf{b}, \quad k = 1, ..., N_\omega,$$

with a single preconditioner $\mathcal{P}(\tau) := (\mathcal{A} - \tau\mathcal{B})$:

$$(\mathcal{A} - \omega_k \mathcal{B})\mathcal{P}_k^{-1}\mathbf{y}_k = \mathbf{b} \quad \Leftrightarrow \quad \boxed{(\mathcal{A}(\mathcal{A} - \tau\mathcal{B})^{-1} - \eta_k I)\mathbf{y}_k = \mathbf{b}}$$

The shift-and-invert preconditioner

$$\mathcal{A} := \begin{bmatrix} iC & K \\ I & 0 \end{bmatrix}, \quad \mathcal{B} := \begin{bmatrix} M & 0 \\ 0 & I \end{bmatrix}$$

The multi-frequency problem reads,

$$(\mathcal{A} - \omega_k \mathcal{B})\mathbf{x}_k = \mathbf{b}, \quad k = 1, ..., N_\omega,$$

with a single preconditioner $\mathcal{P}(\tau) := (\mathcal{A} - \tau\mathcal{B})$:

$$(\mathcal{A} - \omega_k \mathcal{B})\mathcal{P}_k^{-1}\mathbf{y}_k = \mathbf{b} \quad \Leftrightarrow \quad \boxed{(\quad \mathcal{C} \quad - \eta_k I)\mathbf{y}_k = \mathbf{b}}$$

- $\mathcal{C} := \mathcal{A}(\mathcal{A} - \tau\mathcal{B})^{-1}$
- $\eta_k := \omega_k / (\omega_k - \tau)$

TUDelft

Table of Content

Optimization of seed frequency

$$\left(\mathcal{A}(\mathcal{A} - \tau \mathcal{B})^{-1} - \frac{\omega_k}{\omega_k - \tau} I \right) \mathbf{y}_k = \mathbf{b}$$

Theorem: GMRES convergence bound [Saad, Iter. Methods]

Let the eigenvalues of a matrix be enclosed by a circle with radius R and center c. Then the GMRES-residual norm after i iterations $\|\mathbf{r}^{(i)}\|$ satisfies,

$$\frac{\|\mathbf{r}^{(i)}\|}{\|\mathbf{r}^{(0)}\|} \leq c_2(X) \left(\frac{R(\tau)}{|c(\tau)|} \right)^i,$$

where X is the matrix of eigenvectors, and $c_2(X)$ its condition number in the 2-norm.

Optimization of seed frequency

$$\left(\mathcal{A}(\mathcal{A} - \tau\mathcal{B})^{-1} - \frac{\omega_k}{\omega_k - \tau} I \right) \mathbf{y}_k = \mathbf{b}$$

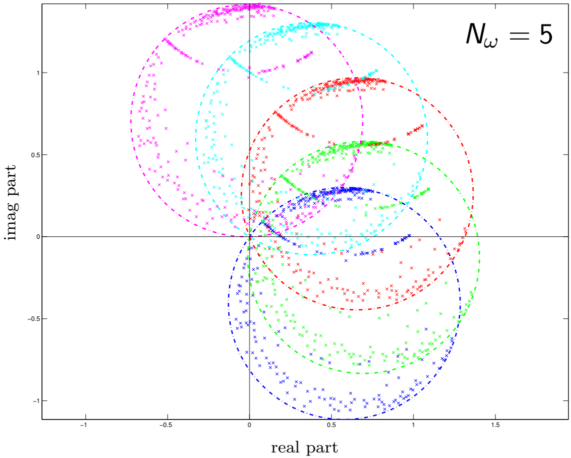
Theorem: msGMRES convergence bound [Saad, Iter. Methods]

Let the eigenvalues of a matrix be enclosed by a circle with radius R_k and center c_k. Then the GMRES-residual norm after i iterations $\|\mathbf{r}_k^{(i)}\|$ satisfies,

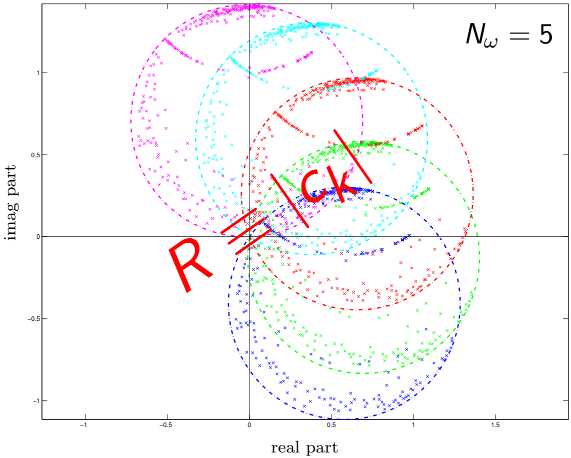
$$\frac{\|\mathbf{r}_k^{(i)}\|}{\|\mathbf{r}^{(0)}\|} \leq c_2(X) \left(\frac{R_k(\tau)}{|c_k(\tau)|} \right)^i, \quad k = 1, \ldots, N_\omega,$$

where X is the matrix of eigenvectors, and $c_2(X)$ its condition number in the 2-norm.

The preconditioned spectra – no damping

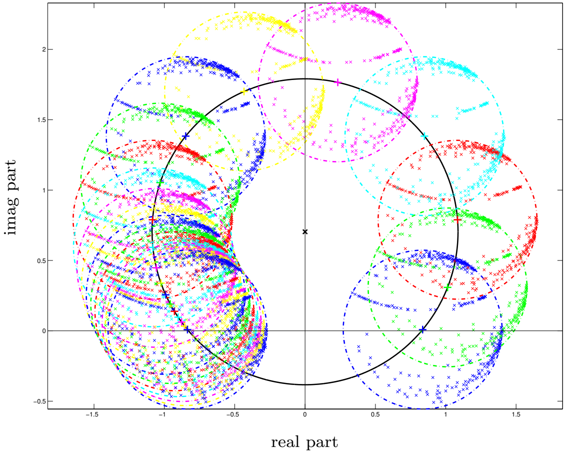


The preconditioned spectra – no damping



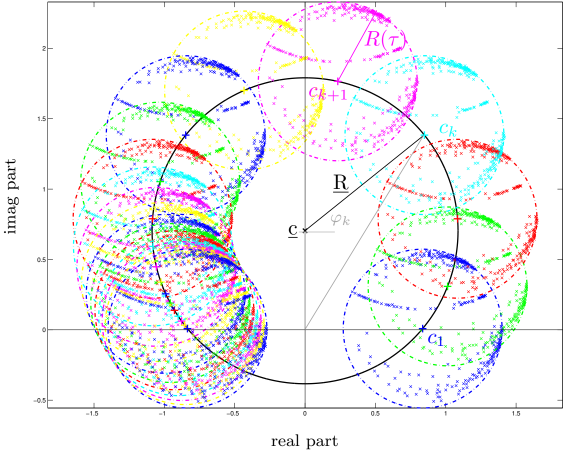
The preconditioned spectra – with damping $\epsilon > 0$

$$\hat{\omega}_k := (1 - \epsilon i)\omega_k$$



The preconditioned spectra – with damping $\epsilon > 0$

$$\hat{\omega}_k := (1 - \epsilon i)\omega_k$$



The preconditioned spectra

Lemma: Optimal seed shift for ms-GMRES [B/vG, 2016]

(i) For $\lambda \in \mathbf{\Lambda}[\mathcal{A}\mathcal{B}^{-1}]$ it holds $\Im(\lambda) \geq 0$.

(ii) The preconditioned spectra are enclosed by circles of radii R_k and center points c_k.

(iii) The points $\{c_k\}_{k=1}^{N_\omega} \subset \mathbb{C}$ described in statement (ii) lie on a circle with center $\underline{c}$ and radius $\underline{R}$.

(iv) Consider the preconditioner $\mathcal{P}(\tau^*) = \mathcal{A} - \tau^*\mathcal{B}$. An optimal seed frequency τ^* for preconditioned multi-shift GMRES is given by,

$$\tau^*(\epsilon) = \min_{\tau \in \mathbb{C}} \max_{k=1,..,N_\omega} \left(\frac{R_k(\tau)}{|c_k|} \right) = \ldots =$$

$$= \frac{2\omega_1 \omega_{N_\omega}}{\omega_1 + \omega_{N_\omega}} - i \frac{\sqrt{\left[\epsilon^2(\omega_1 + \omega_{N_\omega})^2 + (\omega_{N_\omega} - \omega_1)^2\right] \omega_1 \omega_{N_\omega}}}{\omega_1 + \omega_{N_\omega}}$$

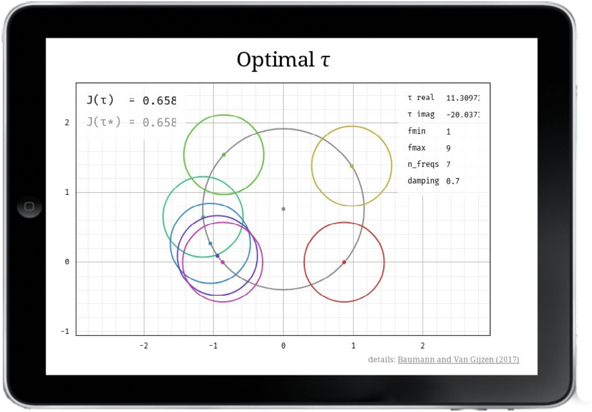
TU Delft

The preconditioned spectra

Proof: Maybe later.

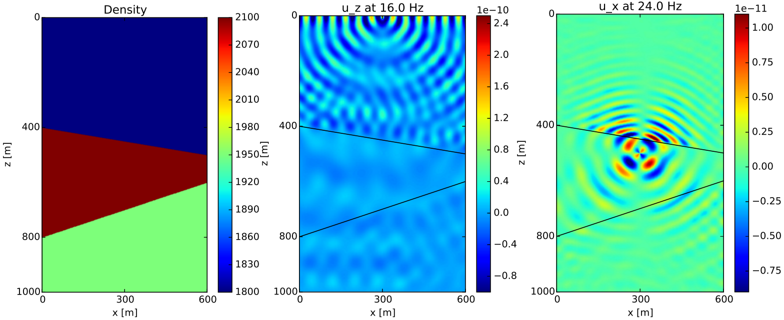
The preconditioned spectra

Proof: There is an App for that.



Numerical experiments I

Set-up: An *elastic* wedge problem.

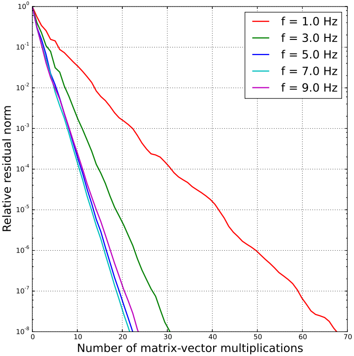


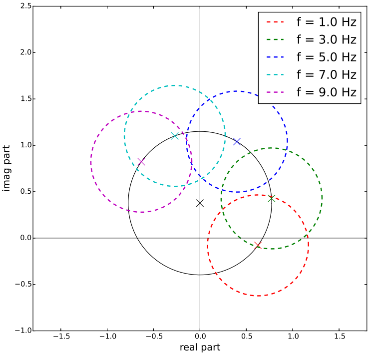
Reference

M. Baumann and M.B. van Gijzen (2017). *An Efficient Two-Level Preconditioner for Multi-Frequency Wave Propagation Problems.* DIAM Technical Report **17-03**, TU Delft.

Convergence behavior and spectral bounds

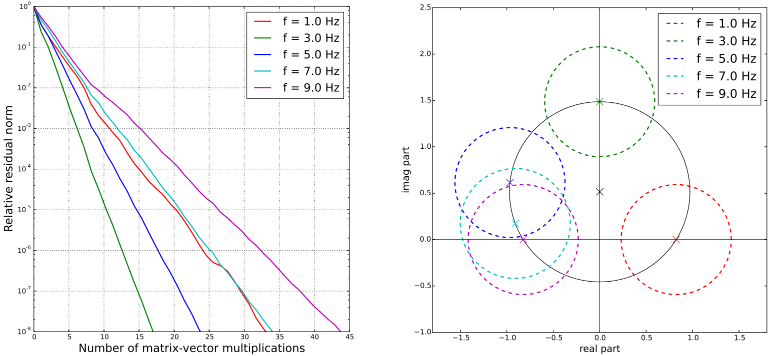
For any τ...





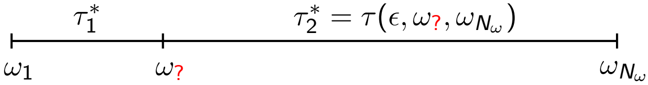
Convergence behavior and spectral bounds

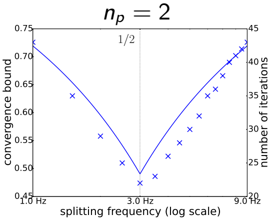
For the optimal τ^*...

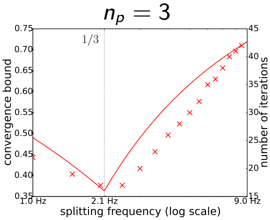


An interval splitting strategy

Suppose $n_p \geq 2$ parallel processors are available.







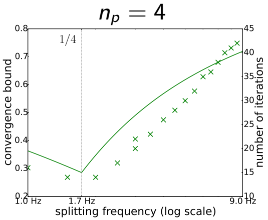


Table of Content

TUDelft

Nested (inner-outer) multi-shift Krylov methods

Solve the shifted linear systems

$$(A - \omega_k I)\, \mathbf{x}_k = \mathbf{b}, \quad k = 1, \ldots, N_\omega,$$

with a nested Krylov method, and preserve,

$$\mathcal{K}_m(A, \mathbf{r}_0) = \mathcal{K}_m(A - \omega I, \mathbf{r}_0) \quad \forall \omega.$$

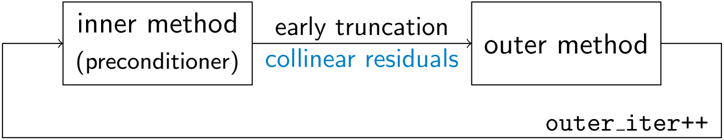
Nested (inner-outer) multi-shift Krylov methods

Solve the shifted linear systems

$$(A - \omega_k I)\,\mathbf{x}_k = \mathbf{b}, \quad k = 1, \ldots, N_\omega,$$

with a nested Krylov method, and preserve,

$$\mathcal{K}_m(A, \mathbf{r}_0) = \mathcal{K}_m(A - \omega I, \mathbf{r}_0) \quad \forall \omega.$$



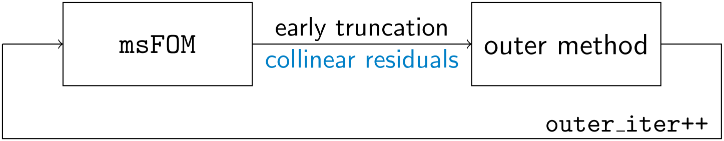
Nested (inner-outer) multi-shift Krylov methods

Solve the shifted linear systems

$$(A - \omega_k I)\,\mathbf{x}_k = \mathbf{b}, \quad k = 1, \ldots, N_\omega,$$

with a nested Krylov method, and preserve,

$$\mathcal{K}_m(A, \mathbf{r}_0) = \mathcal{K}_m(A - \omega I, \mathbf{r}_0) \quad \forall \omega.$$



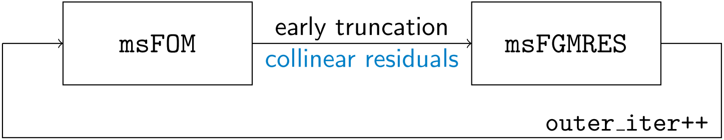
Nested (inner-outer) multi-shift Krylov methods

Solve the shifted linear systems

$$(A - \omega_k I)\,\mathbf{x}_k = \mathbf{b}, \quad k = 1, \ldots, N_\omega,$$

with a nested Krylov method, and preserve,

$$\mathcal{K}_m(A, \mathbf{r}_0) = \mathcal{K}_m(A - \omega I, \mathbf{r}_0) \quad \forall \omega.$$



Multi-shift FOM as inner method

Classical result: In FOM, the residuals are

$$\mathbf{r}_j = \mathbf{b} - A\mathbf{x}_j = \ldots = -h_{j+1,j}\mathbf{e}_j^T\mathbf{y}_j\mathbf{v}_{j+1}.$$

Thus, for the shifted residuals it holds:

$$\mathbf{r}_j^{(\omega)} = \mathbf{b} - (A - \omega I)\mathbf{x}_j^{(\omega)} = \ldots = -h_{j+1,j}^{(\omega)}\mathbf{e}_j^T\mathbf{y}_j^{(\omega)}\mathbf{v}_{j+1},$$

which gives $\gamma = y_j^{(\omega)}/y_j$.

Reference

V. Simoncini, *Restarted full orthogonalization method for shifted linear systems.* BIT Numerical Mathematics, 43 (2003).

Flexible multi-shift GMRES as outer method

Use flexible GMRES in the outer loop,

$$(A - \omega I)\widehat{V}_m = V_{m+1}\underline{\mathrm{H}}_m^{(\omega)},$$

where one column yields

$$(A - \omega I)\underbrace{\mathcal{P}(\omega)_j^{-1}\mathbf{v}_j}_{\text{inner loop}} = V_{m+1}\underline{\mathbf{h}}_j^{(\omega)}, \quad 1 \le j \le m.$$

The "inner loop" is the truncated solution of $(A - \omega I)$ with right-hand side $\mathbf{v}_j$ using msFOM.

Flexible multi-shift GMRES as outer method

Use flexible GMRES in the outer loop,

$$(A - \omega I)\widehat{V}_m = V_{m+1}\underline{\mathrm{H}}_m^{(\omega)},$$

where one column yields

$$(A - \omega I)\underbrace{\mathcal{P}(\omega)_j^{-1}\mathbf{v}_j}_{\text{inner loop}} = V_{m+1}\underline{\mathbf{h}}_j^{(\omega)}, \quad 1 \leq j \leq m.$$

The "inner loop" is the truncated solution of $(A - \omega I)$ with right-hand side $\mathbf{v}_j$ using `msFOM`.

TUDelft

Flexible multi-shift GMRES

Thus, the inner residuals are:

$$\mathbf{r}_j^{(\omega)} = \mathbf{v}_j - (A - \omega I)\mathcal{P}(\omega)_j^{-1}\mathbf{v}_j,$$
$$\mathbf{r}_j = \mathbf{v}_j - A\mathcal{P}_j^{-1}\mathbf{v}_j,$$

Imposing $\mathbf{r}_j^{(\omega)} = \gamma \mathbf{r}_j$ yields:

$$(A - \omega I)\mathcal{P}(\omega)_j^{-1}\mathbf{v}_j = \gamma A\mathcal{P}_j^{-1}\mathbf{v}_j - (\gamma - 1)\mathbf{v}_j \qquad (*)$$

Note that the right-hand side in $(*)$ is a preconditioned shifted system!

Flexible multi-shift GMRES

Thus, the inner residuals are:

$$\begin{aligned}\mathbf{r}_j^{(\omega)} &= \mathbf{v}_j - (A - \omega I)\mathcal{P}(\omega)_j^{-1}\mathbf{v}_j, \\ \mathbf{r}_j &= \mathbf{v}_j - A\mathcal{P}_j^{-1}\mathbf{v}_j,\end{aligned}$$

Imposing $\mathbf{r}_j^{(\omega)} = \gamma \mathbf{r}_j$ yields:

$$(A - \omega I)\mathcal{P}(\omega)_j^{-1}\mathbf{v}_j = \gamma A\mathcal{P}_j^{-1}\mathbf{v}_j - (\gamma - 1)\mathbf{v}_j \qquad (*)$$

Note that the right-hand side in $(*)$ is a preconditioned shifted system!

Flexible multi-shift GMRES

Altogether,

$$(A - \omega I)\mathcal{P}(\omega)_j^{-1}\mathbf{v}_j = V_{m+1}\underline{\mathbf{h}}_j^{(\omega)}$$

$$\gamma A\mathcal{P}_j^{-1}\mathbf{v}_j - (\gamma - 1)\mathbf{v}_j = V_{m+1}\underline{\mathbf{h}}_j^{(\omega)}$$

$$\gamma V_{m+1}\underline{\mathbf{h}}_j - V_{m+1}(\gamma - 1)\underline{\mathbf{e}}_j = V_{m+1}\underline{\mathbf{h}}_j^{(\omega)}$$

$$V_{m+1}\left(\gamma\underline{\mathbf{h}}_j - (\gamma - 1)\underline{\mathbf{e}}_j\right) = V_{m+1}\underline{\mathbf{h}}_j^{(\omega)}$$

which yields:

$$\underline{\mathrm{H}}_m^{(\omega)} = (\underline{\mathrm{H}}_m - \underline{\mathrm{I}}_m)\,\Gamma_m + \underline{\mathrm{I}}_m,$$

with $\Gamma_m := diag(\gamma_1, ..., \gamma_m)$.

Flexible multi-shift GMRES

Altogether,

$$
\begin{aligned}
(A - \omega I)\mathcal{P}(\omega)_j^{-1}\mathbf{v}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)} \\
\gamma A\mathcal{P}_j^{-1}\mathbf{v}_j - (\gamma - 1)\mathbf{v}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)} \\
\gamma V_{m+1}\underline{\mathbf{h}}_j - V_{m+1}(\gamma - 1)\underline{\mathbf{e}}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)} \\
V_{m+1}\left(\gamma\underline{\mathbf{h}}_j - (\gamma - 1)\underline{\mathbf{e}}_j\right) &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)}
\end{aligned}
$$

which yields:

$$\underline{\mathrm{H}}_m^{(\omega)} = (\underline{\mathrm{H}}_m - \underline{\mathrm{I}}_m)\,\Gamma_m + \underline{\mathrm{I}}_m,$$

with $\Gamma_m := diag(\gamma_1, ..., \gamma_m)$.

Flexible multi-shift GMRES

Altogether,

$$
\begin{aligned}
(A - \omega I)\mathcal{P}(\omega)_j^{-1}\mathbf{v}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)} \\
\gamma A\mathcal{P}_j^{-1}\mathbf{v}_j - (\gamma - 1)\mathbf{v}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)} \\
\gamma V_{m+1}\underline{\mathbf{h}}_j - V_{m+1}(\gamma - 1)\underline{\mathbf{e}}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)} \\
V_{m+1}\left(\gamma\underline{\mathbf{h}}_j - (\gamma - 1)\underline{\mathbf{e}}_j\right) &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)}
\end{aligned}
$$

which yields:

$$
\underline{\mathrm{H}}_m^{(\omega)} = (\underline{\mathrm{H}}_m - \underline{\mathrm{I}}_m)\,\Gamma_m + \underline{\mathrm{I}}_m,
$$

with $\Gamma_m := diag(\gamma_1, ..., \gamma_m)$.

Flexible multi-shift GMRES

Altogether,

$$
\begin{aligned}
(A-\omega I)\mathcal{P}(\omega)_j^{-1}\mathbf{v}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)}\\
\gamma A\mathcal{P}_j^{-1}\mathbf{v}_j - (\gamma-1)\mathbf{v}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)}\\
\gamma V_{m+1}\underline{\mathbf{h}}_j - V_{m+1}\,(\gamma-1)\,\underline{\mathbf{e}}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)}\\
V_{m+1}\left(\gamma\underline{\mathbf{h}}_j - (\gamma-1)\,\underline{\mathbf{e}}_j\right) &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)}
\end{aligned}
$$

which yields:

$$
\underline{\mathrm{H}}_m^{(\omega)} = (\underline{\mathrm{H}}_m - \underline{\mathrm{I}}_m)\,\Gamma_m + \underline{\mathrm{I}}_m,
$$

with $\Gamma_m := diag(\gamma_1, ..., \gamma_m)$.

Flexible multi-shift GMRES

Altogether,

$$
\begin{aligned}
(A - \omega I)\mathcal{P}(\omega)_j^{-1}\mathbf{v}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)} \\
\gamma A\mathcal{P}_j^{-1}\mathbf{v}_j - (\gamma - 1)\mathbf{v}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)} \\
\gamma V_{m+1}\underline{\mathbf{h}}_j - V_{m+1}(\gamma - 1)\underline{\mathbf{e}}_j &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)} \\
V_{m+1}\left(\gamma\underline{\mathbf{h}}_j - (\gamma - 1)\underline{\mathbf{e}}_j\right) &= V_{m+1}\underline{\mathbf{h}}_j^{(\omega)}
\end{aligned}
$$

which yields:

$$
\underline{\mathrm{H}}_m^{(\omega)} = (\underline{\mathrm{H}}_m - \underline{\mathrm{I}}_m)\,\Gamma_m + \underline{\mathrm{I}}_m,
$$

with $\Gamma_m := diag(\gamma_1, ..., \gamma_m)$.

TUDelft

Summary: Nested FOM-FGMRES

In nested FOM-FGMRES, we solve the following (small) optimization problems,

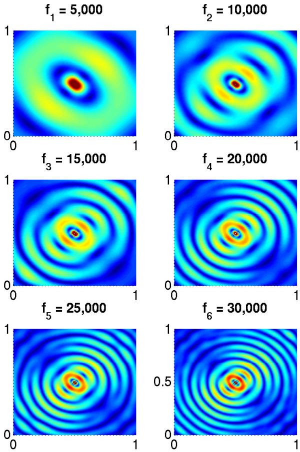
$$
\begin{aligned}
\mathbf{x}_m^{(\omega)} &= \underset{\mathbf{x}\in\widehat{\mathcal{V}}_m}{\operatorname{argmin}} \left\|\mathbf{b} - (A - \omega I)\mathbf{x}\right\| \\
&= \underset{\mathbf{y}\in\mathbb{C}^m}{\operatorname{argmin}} \left\|\mathbf{b} - (A - \omega I)\widehat{V}_m\mathbf{y}\right\| \\
&= \underset{\mathbf{y}\in\mathbb{C}^m}{\operatorname{argmin}} \left\|\mathbf{b} - V_{m+1}\underline{\mathrm{H}}_m^{(\omega)}\mathbf{y}\right\| \\
&= \underset{\mathbf{y}\in\mathbb{C}^m}{\operatorname{argmin}} \left\|\beta\mathbf{e}_1 - \left(\left(\underline{\mathrm{H}}_m - \underline{\mathrm{I}}_m\right)\Gamma_m^{(\omega)} + \underline{\mathrm{I}}_m\right)\mathbf{y}\right\|,
\end{aligned}
$$

where the entries of $\Gamma_m^{(\omega)}$ are collinearity factors of inner FOM.

Numerical experiments II

Test case from literature:

- $\Omega = [0, 1] \times [0, 1]$
- $h = 0.01$ implying
  $n = 10.201$ grid points
- system size:
  $4n = 40.804$
- $N = 6$ frequencies
- point source at center



Reference

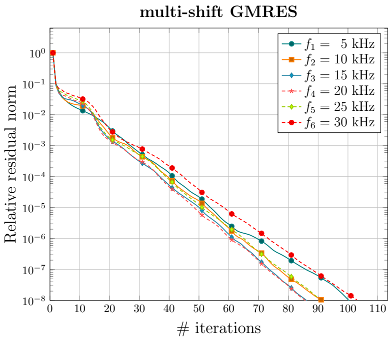
M. Baumann and M.B. van Gijzen (2015). *Nested Krylov methods for shifted linear systems.* SIAM J. Sci. Comput., **37**(5), S90-S112.

Numerical experiments II

as presented in [B./vG., 2015]

Preconditioned **multi-shift GMRES**:

- simultaneous solve
- linear convergence rates
- $\tau = \tau^* = (0.3 - 0.7i)\omega_{max}$
- CPU time: 17.71s

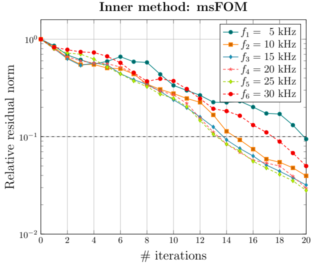


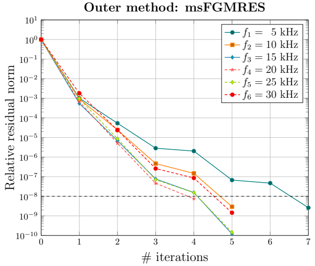
Numerical experiments II

as presented in [B./vG., 2015]

Preconditioned **nested FOM-FGMRES**:

- 20 inner iterations
- truncate when inner residual norm ~ 0.1
- very few outer iterations
- CPU time: 9.12s





Numerical experiments II

as presented in [B./vG., 2015]

Various combinations of inner/outer methods *possible*:

	multi-shift Krylov methods			
	msGMRES	msGMRESr	QMRIDR(4)	msIDR(4)
# inner iterations	-	20	-	-
# outer iterations	103	7	136	134
CPU time	17.71s	6.13s	22.35s	22.58s
	nested multi-shift Krylov methods			
	FOM-FGMRES	IDR(4)-FGMRES	FOM-FQMRIDR(4)	IDR(4)-FQMRIDR(4)
# inner iterations	20	25	30	30
# outer iterations	7	9	5	15
CPU time	9.12s	32.99s	8.14s	58.36s

Different combination (CMRH-FCMRH) is exploited in:

X.-M. Gu, T.-Z. Huang, B. Carpentieri, A. Imakura, K. Zhang, L. Du. *Variants of the CMRH method for solving multi-shifted non-Hermitian linear systems.* Technical Report, University of Groningen (2016).

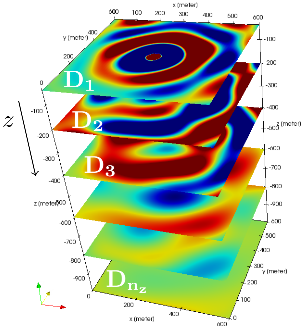
Conclusions

What I have shown today:

- ✓ Optimal τ^* for multi-shift GMRES
- ✓ Nested multi-shift FOM-FGMRES

What I have not shown today:

- ✓ MSSS preconditioning techniques
- ✓ matrix equation (global GMRES)
- ✓ large-scale 3D examples
- ✗ deflation, model-order reduction, ...



References

- M. Baumann and M.B. van Gijzen (2015). *Nested Krylov methods for shifted linear systems.* SIAM J. Sci. Comput., **37**(5), S90-S112.

- M. Baumann, R. Astudillo, Y. Qiu, E. Ang, M.B. van Gijzen, and R.-E. Plessix (2017). *An MSSS-Preconditioned Matrix Equation Approach for the Time-Harmonic Elastic Wave Equation at Multiple Frequencies.* Springer Computat. Geosci., DOI: 10.1007/s10596-017-9667-7.

- M. Baumann and M.B. van Gijzen (2017). *Efficient iterative methods for multi-frequency wave propagation problems: A comparison study.* Procedia Comput. Sci., Vol. **108**, pp. 645-654.

- M. Baumann and M.B. van Gijzen (2017). *An Efficient Two-Level Preconditioner for Multi-Frequency Wave Propagation Problems.* DIAM Technical Report **17-03**, TU Delft. [under review]

The preconditioned spectra

Lemma: Optimal seed shift for msGMRES [B/vG, 2016]

(i) For $\lambda \in \mathbf{\Lambda}[\mathcal{A}\mathcal{B}^{-1}]$ it holds $\Im(\lambda) \geq 0$.

(ii) The preconditioned spectra are enclosed by circles of radii R_k and center points c_k.

(iii) The points $\{c_k\}_{k=1}^{N_\omega} \subset \mathbb{C}$ described in statement (ii) lie on a circle with center $\underline{c}$ and radius $\underline{R}$.

(iv) Consider the preconditioner $\mathcal{P}(\tau^*) = \mathcal{A} - \tau^*\mathcal{B}$. An optimal seed frequency τ^* for preconditioned multi-shift GMRES is given by,

$$\tau^*(\epsilon) = \min_{\tau \in \mathbb{C}} \max_{k=1,..,N_\omega} \left(\frac{R_k(\tau)}{|c_k|} \right) = ... =$$

$$= \frac{2\omega_1\omega_{N_\omega}}{\omega_1 + \omega_{N_\omega}} - i\frac{\sqrt{\left[\epsilon^2(\omega_1 + \omega_{N_\omega})^2 + (\omega_{N_\omega} - \omega_1)^2\right]\omega_1\omega_{N_\omega}}}{\omega_1 + \omega_{N_\omega}}$$

TUDelft

The preconditioned spectra – Proof (1/4)

Proof. (i) We have to show $\Im(\omega) \geq 0$ for,

$$\begin{bmatrix} iC & K \\ I & 0 \end{bmatrix} x = \omega \begin{bmatrix} M & 0 \\ 0 & I \end{bmatrix} x$$

or, alternatively ($\lambda = i\omega$), consider the QEP,

$$(K + \lambda C + \lambda^2 M)v = 0.$$

§3.8		come in pairs $(\lambda, \bar{\lambda})$	λ then x is a left eigenvector of $\bar{\lambda}$
P5 §3.8	M Hermitian positive definite, C, K Hermitian positive semidefinite	$\mathrm{Re}(\lambda) \leq 0$	
P6 §3.9	M, C symmetric positive definite, K symmetric	λs are real and negative, gap between n largest and	n linearly independent eigenvectors associated with

$$\Re(\lambda) \leq 0 \;\Rightarrow\; \Im(\omega) \geq 0$$

The preconditioned spectra – Proof (1/4)

Proof. (i) We have to show $\Im(\omega) \geq 0$ for,

$$\begin{bmatrix} iC & K \\ I & 0 \end{bmatrix} x = \omega \begin{bmatrix} M & 0 \\ 0 & I \end{bmatrix} x$$

or, alternatively $(\lambda = i\omega)$, consider the QEP,

$$(K + \lambda C + \lambda^2 M)v = 0.$$

§3.8		come in pairs $(\lambda, \bar{\lambda})$	λ then x is a left eigenvector of $\bar{\lambda}$
P5 §3.8	M Hermitian positive definite, C, K Hermitian positive semidefinite	$\mathrm{Re}(\lambda) \leq 0$	
P6 §3.9	M, C symmetric positive definite, K symmetric	λs are real and negative, gap between n largest and	n linearly independent eigenvectors associated with

$$\Re(\lambda) \leq 0 \;\Rightarrow\; \Im(\omega) \geq 0$$

The preconditioned spectra – Proof (2/4)

(ii) *The preconditioned spectra are enclosed by circles.*

Factor out $\mathcal{A}\mathcal{B}^{-1}$,

$$\mathcal{C} - \eta_k I = \mathcal{A}(\mathcal{A} - \tau\mathcal{B})^{-1} - \eta_k I = \mathcal{A}\mathcal{B}^{-1}(\mathcal{A}\mathcal{B}^{-1} - \tau I)^{-1} - \eta_k I,$$

and note that

$$\mathbf{\Lambda}[\mathcal{A}\mathcal{B}^{-1}] \ni \lambda \mapsto \frac{\lambda}{\lambda - \tau} - \frac{\omega_k}{\omega_k - \tau},$$

is a Möbius transformation$^{(*)}$.

Reference

M.B. van Gijzen, Y.A. Erlangga, C. Vuik. *Spectral Analysis of the Discrete Helmholtz Operator Preconditioned with a Shifted Laplacian.* SIAM J. Sci. Comput., **29**(5), 1942–1958 (2007)

TU Delft

$(*) : \; z \mapsto \frac{az+b}{cz+d}$

The preconditioned spectra – Proof (3/4)

(iii) *Spectra are bounded by circles (c_k, R). These center point $\{c_k\}_{k=1}^{N_\omega}$ lie on a 'big circle' $(\underline{c}, \underline{R})$.*

1. Construct center:

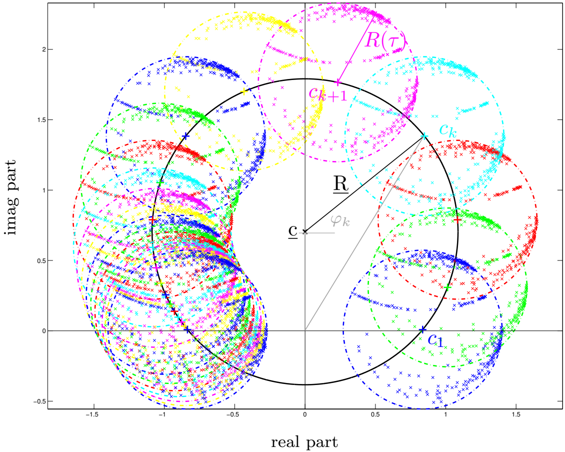
$$\underline{c} = \left(0, \frac{\epsilon|\tau|^2}{2\Im(\tau)(\Im(\tau) + \epsilon\Re(\tau))}\right) \in \mathbb{C}$$

2. A point c_k has constant distance to $\underline{c}$:

$$\underline{R}^2 = \|c_k - \underline{c}\|_2^2 = \frac{|\tau|^2(\epsilon^2 + 1)}{4(\Im(\tau) + \epsilon\Re(\tau))^2} \quad \text{(independent of } \omega_k\text{)}$$

The preconditioned spectra – with damping $\epsilon > 0$

$$\hat{\omega}_k := (1 - \epsilon i)\omega_k$$

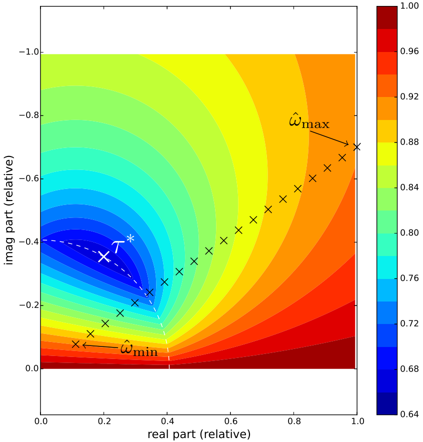


The preconditioned spectra – Proof (4/4)

(iv) *Find optimal τ^*.*

$$\tau^* = \min_{\tau \in \mathbb{C}} \max_{k=1,..,N_\omega} \left(\frac{R}{|c_k|} \right)$$

1. $|c_k| = f(\underline{c}, \underline{R}, \varphi_k)$
2. polar coordinates
3. $\frac{\partial \tau}{\partial \varphi} = 0$ (optimize along φ)

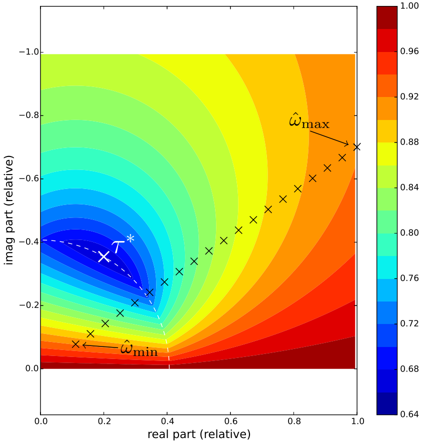


The preconditioned spectra – Proof (4/4)

(iv) *Find optimal τ^*.*

$$\tau^* = \min_{\tau\in\mathbb{C}} \max_{k\in\{1,N_\omega\}} \left(\frac{R}{|c_k|}\right)$$

1. $|c_k| = f(\underline{c}, \underline{R}, \varphi_k)$
2. polar coordinates
3. $\frac{\partial \tau}{\partial \varphi} = 0$ (optimize along φ)

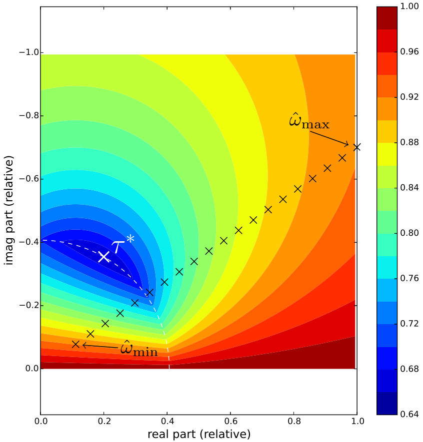


The preconditioned spectra – Proof (4/4)

(iv) *Find optimal τ^*.*

$$\tau^* = \min_{\tau\in\mathbb{C}} \max_{k\in\{1,N_\omega\}} \left(\frac{R}{|c_k|}\right)$$

1. $|c_k| = f(\underline{c}, \underline{R}, \varphi_k)$
2. polar coordinates
3. $\frac{\partial\tau}{\partial\varphi} = 0$ (optimize along φ)



$$\tau^* = \frac{2\omega_1\omega_{N_\omega}}{\omega_1+\omega_{N_\omega}} - i\,\frac{\sqrt{[\epsilon^2(\omega_1+\omega_{N_\omega})^2+(\omega_{N_\omega}-\omega_1)^2]\omega_1\omega_{N_\omega}}}{\omega_1+\omega_{N_\omega}}$$