Iterative Helmholtz Solvers

Scalability and Accuracy of Helmholtz solvers

Delft University of Technology

Kees Vuik, Vandana Dwarka

July 2nd, 2019

Outline

Aim and Impact

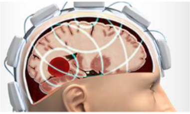
- Joint-work with PhD candidtate Vandana Dwarka
- Contribute to broad research on Helmholtz solvers
- Understand inscalability (convergence) and accuracy (pollution)
- Improve convergence properties
- Alert solver developers to keep accuracy of solution in mind

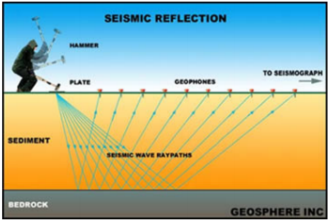
Introduction - The Helmholtz Equation

- Inhomogeneous Helmholtz equation + BC's

$$(-\nabla^2 - k^2)\, u(\mathbf{x}) = f(\mathbf{x}), \mathbf{x} \in \Omega \subseteq \mathbb{R}^n$$

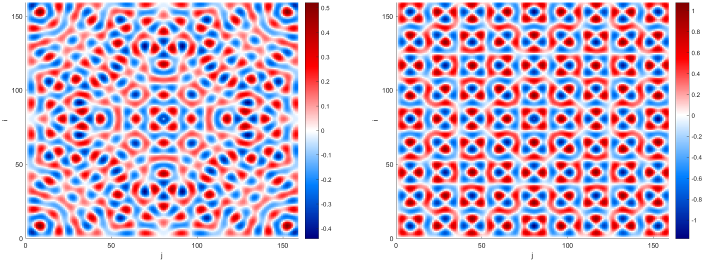
- k is the wave number: $k = \frac{2\pi}{\lambda}$
- Practical applications in seismic and medical imaging





Introduction - Challenges

- Negative & positive eigenvalues ⇒ limits Krylov based solvers
- Fast near-origin moving eigenvalues ⇒ slows convergence
  - CSLP: A with a complex shift
  - Deflation + CSLP
  - Despite improvements problem remains
- Numerical solution is often inaccurate!



- Problems exacerbate in 2D & 3D

Introduction - Numerical Model

- Analytical 1D model problem

$$-\frac{d^2u}{dx^2} - k^2\,u = \delta(x - \frac{1}{2}),$$
$$u(0) = 0, u(1) = 0,$$
$$x \in \Omega = [0,1] \subseteq \mathbb{R},$$

- Discretization using second-order FD with at least 10 gpw
- We obtain a linear system $A\hat{u} = f$

$$A = \frac{1}{h^2}\text{tridiag}[-1 \;\; 2-(kh)^2 \;\; -1],$$

- A is real, symmetric, normal, indefinite and sparse
- Using Radiation BC's A additionally becomes non-Hermitian

Introduction - Spectral Properties

- Near-null eigenvalues near intersection with origin
- The index where this happens for the analytical case is

$$\lambda_{j_{\min}} = 0 \Rightarrow j^2\pi^2 \approx k^2\pi^2,$$
$$\Rightarrow j_{\min} = \lfloor \frac{k}{\pi} \rfloor \text{ or } \lceil \frac{k}{\pi} \rceil.$$

- The index where this happens for the numerical case is

$$\hat{\lambda}_{\hat{j}_{\min}} = 0 \Rightarrow \frac{1}{h^2}\left[2 - 2\cos(j\pi h)\right] \approx k^2$$
$$\Rightarrow \hat{j}_{\min} = \text{round}\left[\frac{\arccos 1 - \kappa^2}{\pi h}\right]$$

- Our research:
  - Near-null eigenvalues ⇒ accuracy issues ⇒ Pollution
  - Near-null eigenvalues ⇒ convergence issues ⇒ Deflation

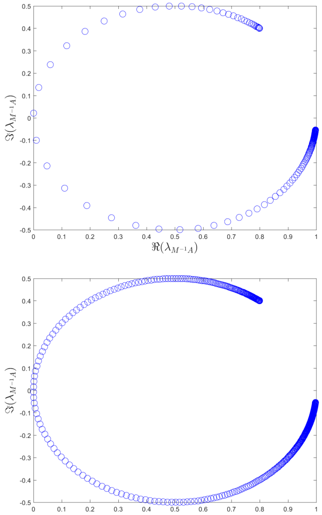
Preconditioning - CSLP

- Preconditioning to speed up convergence of Krylov subspace methods
- Solve $M^{-1}Au = M^{-1}f$, M is CSLP-preconditioner.

$$M = A - (\beta_1 + \beta_2 i)k^2 I,$$
$$(\beta_1, \beta_2) \in [0, 1]$$

- Increasing $k \Rightarrow$ inscalable CSLP-solver
- Project unwanted eigenvalues onto zero = Deflation

Figure: $\sigma(M^{-1}A)$ for $k = 50$ (top) and $k = 150$ bottom.



Preconditioning - Deflation

- Projection principle: solve $PAu = Pf$

$$\tilde{P} = AQ \text{ where } Q = ZE^{-1}Z^T \text{ and } E = Z^TAZ,$$
$$P = I - \tilde{P},\ Z \in \mathbb{R}^{m \times n},\ m < n$$

- Columns of Z span deflation subspace
- Ideally Z contains eigenvectors
- In practice approximations

DEF - I

- Main focus on DEF-preconditioner (Sheikh, A., 2014)

$$P = I - AQ \text{ where } Q = Z{A_{2h}}^{-1}Z^T \text{ and } A_{2h} = Z^TAZ$$

- Inter-grid vectors from multi-grid as deflation vectors
- Inter-grid approximation based on linear interpolation
- Use DEF + CSLP combined $\Rightarrow$ spectral improvement

$$P^TM^{-1}Au = P^TM^{-1}f$$

- Monitor eigenvalues using rigorous Fourier analysis
- Effect aggravates in higher-dimensions
- Near-null eigenvalues unless $\#gpw$ increases along

DEF - II

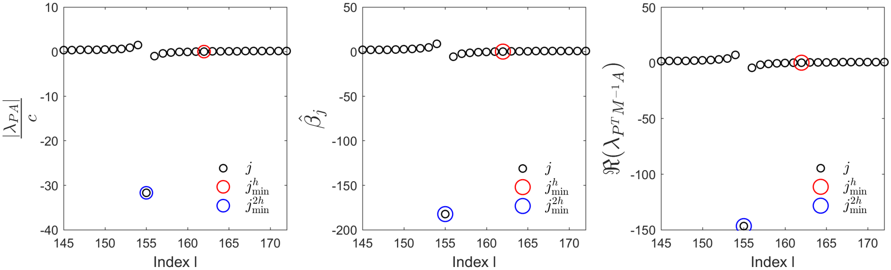
- Near-null eigenvalues arise at projection level
- Block-diagonalize $P \Rightarrow$ eigenvalues

$$\begin{aligned}
\lambda^l(P) &= s^l + c^l, \\
s^l &= \left(1 - \frac{\lambda^l(A)\cos(l\pi\frac{h}{2})^4}{\lambda^l(A_{2h})}\right) = \frac{\lambda^{n+1-l}(A)\sin(l\pi\frac{h}{2})^4}{\lambda^l(A_{2h})}, \\
c^l &= \left(1 - \frac{\lambda^{n+1-l}(A)\sin(l\pi\frac{h}{2})^4}{\lambda^l(A_{2h})}\right) = \frac{\lambda^l(A)\cos(l\pi\frac{h}{2})^4}{\lambda^l(A_{2h})}, \\
\lambda^l(PA) &= \lambda^l(A)s^l + \lambda^{n+1-l}(A)c^l, \\
l &= 1, 2, \ldots, \frac{n}{2}.
\end{aligned}$$

DEF - II

- Investigate near-null eigenvalue of <u>all</u> operators involved

Figure: $\lambda_j(PA), \beta^j, \lambda_j(P^T M^{-1} A)$ for $k = 500$



- Eigenvalues of PA and $P^T M^{-1} A$ behave like $\hat{\beta} = \frac{\lambda^l(A)}{\lambda^l(A_{2h})}$
- If near-kernel of A and A_{2h} misaligned $\Rightarrow$ near-null eigenvalues reappear!
- Equivalent to $j^h_{\min} \neq j^{2h}_{\min}$

DEF - III

- Recall: deflation space spanned by linear approximation basis vectors
- Transfer coarse-fine grid ⇒ interpolation error ⇒ near-kernel A_{2h} shifts
- Measure effect by projection error E

$$E(kh) = \|(I - P)\phi_{j_{\min},h}\|^2,$$

$$P = Z(Z^TZ)^{-1}Z^T$$

Figure: Restricted & interpolated eigenvectors (left $kh = 0.625$, right $k^3h^2 = 0.625$

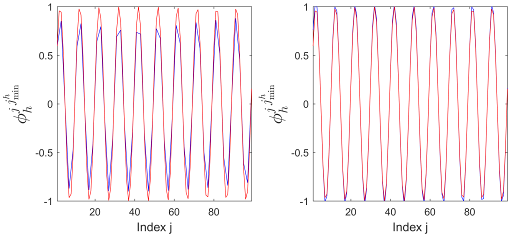


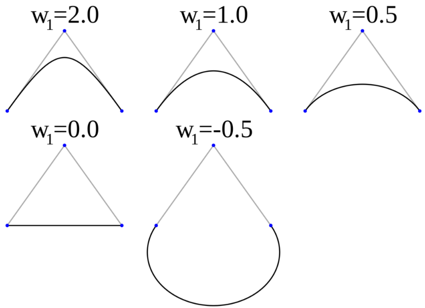
Table: Projection error DEF-scheme

k	$E(0.625)$	$E(0.3125)$
10^2	0.8818	0.1006
10^3	9.2941	1.0062
10^4	92.5772	10.0113
10^5	926.135	100.1382
10^6	9261.7129	1001.3818

Our Approach - Introduction

- Higher-order deflation vectors
- Rational quadratic Bezier curve $\Rightarrow$ one control-point
- Weight-parameter w to adjust control-point

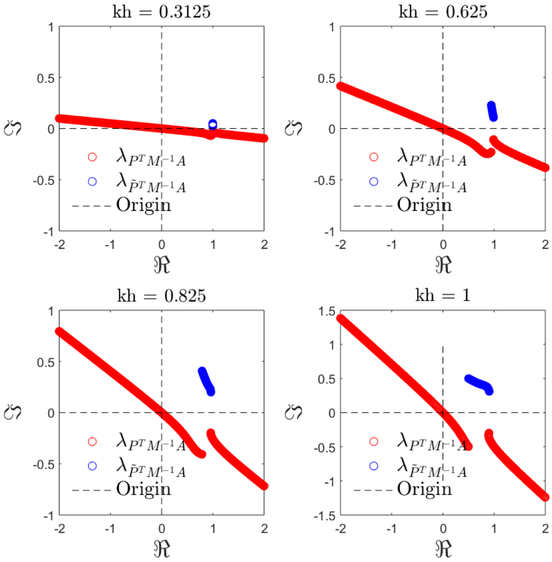
Figure: Effect of changing weight



- w determined such that projection error minimized
- Rigorous Fourier analysis confirms favourable spectrum

Our Approach - Spectral Analysis (1D)

Figure: Spectrum of old (red) and new (blue) method for $k = 10^6$



Numerical Experiments - 1D

Table: GMRES-iterations with tol $= 10^{-7}$ using the new scheme $APD(w)$ and CSLP(1,0.5).

k	APD(0.1250)	APD(0.0575)	APD(0.01875)	APD(0.00125)
	$kh = 1$	$kh = 0.825$	$kh = 0.625$	$kh = 0.3125$
10^2	6	5	4	3
10^3	6	5	4	3
10^4	6	5	4	3
10^5	6	5	4	3
10^6	6	5	4	3

- DEF + CSLP takes 367 its. and 16.1104 sec. for $k = 10^4$
- We obtain wave number independence for all kh
- Weight-parameter w less important as kh decreases

Projection Error - 1D

Table: Projection error $E(kh)$ for $APD(w)$ and CSLP(1,0.5)

k	APD(0.1250)	APD(0.0575)	APD(0.01875)	APD(0.00125)
	$kh = 1$	$kh = 0.825$	$kh = 0.625$	$kh = 0.3125$
10^2	0.0127	0.0075	0.0031	0.0006
10^3	0.0233	0.0095	0.0036	0.0007
10^4	0.0246	0.0095	0.0038	0.0007
10^5	0.0246	0.0095	0.0038	0.0007
10^6	0.0246	0.0095	0.0038	0.0007

- Weight-parameter w chosen to minimize projection error
- In all cases projection error *strictly* < 1

Numerical Experiments - 2D

Table: GMRES-iterations with tol $= 10^{-7}$ using the new scheme $APD(w)$ and CSLP(1,0.5). AD contains no CSLP.

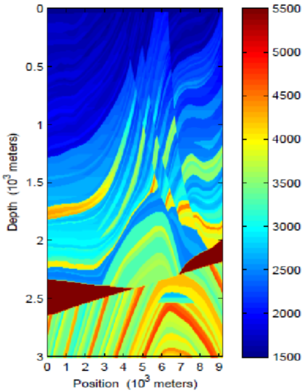
k	APD(0.1250)	APD(0.0575)	AD(0)
	$kh = 0.625$	$kh = 0.3125$	$kh = 0.3125$
100	4	4	3
250	5	4	4
500	5	5	5
750	7	5	5
1000	8	8	7

- DEF + CSLP takes 471 iterations and 1195.9730 sec. for $k = 250$
- Close to wave number independence
- Weight-parameter w and CSLP less important as kh decreases

Marmousi - 2D

Table: Solve time (s) and GMRES-iterations for 2D Marmousi

	DEF-TL	APD-TL	DEF-TL	APD-TL
	10 gpw			
f	Solve time (s)		Iterations	
1	1.72	4.08	3	4
10	7.20	3.94	16	6
20	77.34	19.85	31	6
40	1175.99	111.78	77	6
	20 gpw			
1	9.56	3.83	3	5
10	19.64	15.45	7	5
20	155.70	122.85	10	5
40	1500.09	1201.45	15	5



Numerical Experiments - 3D

Table: GMRES-iterations with tol $= 10^{-7}$ using the new scheme $APD(w)$ and CSLP(1,0.5). $kh = 0.625$.

k	APD(0.125)	APD(0)
	Iterations	Iterations
10	4	4
25	4	5
50	4	5
75	4	5

- DEF + CSLP takes 66 iterations for $k = 40$
- Two-level method memory constrained $\Rightarrow$ multilevel

Conclusion

- Deflation projects unwanted eigenvalues to zero
- Large $k \Rightarrow$ near-null eigenvalues reappear
- Near-kernel alignment of A and A_{2h}
- Misalignment affects both accuracy and convergence
- New deflation scheme: higher-order approximation
- Obtain faster convergence

References - I

- Upcoming articles
  http://ta.twi.tudelft.nl/nw/users/vuik/pub_it_helmholtz.html
- Further reading

C. Vuik et. al.
Modern Solvers for Helmholtz Problems.
Birkhuser, 2017.

F. Ihlenburg, I. Babuska.
Dispersion analysis and error estimation of Galerkin finite element methods for the Helmholtz equation.
International Journal for Numerical Methods in Engineering, 38(22):3745-3774, 1995.

I. Babuska, S. Sauter.
Is the Pollution Effect of the FEM Avoidable for the Helmholtz Equation Considering High Wave Numbers?
SIAM Journal on Numerical Analysis, 34(6):23922423, 1997.

References - II

- Further reading

A. Sheikh
Development of the Helmholtz Solver based on a Shifted Laplace Preconditioner and a Multigrid Deflation technique.
PhD Thesis, Delft University of Technology, 2014

L. Garcia Ramos, R. Nabben
On the Spectrum of Deflated Matrices with Applications to the Deflated Shifted Laplace Preconditioner for the Helmholtz Equation
SIAM Journal on Matrix Analysis and Applications, 39(1):262286, 2018

O. Ernst, M. Gander
Why it is difficult to Solve Helmholtz Problems with Classical Iterative methods
Numerical Analysis of Multiscale Problems. Lecture Notes in Computational Science and Engineering, 83(1):325-363, 2012