Graduate Courses in Num. An., Louvain, April 1998

Jacobi-Davidson Methods for computing a few selected eigenpairs

Gerard Sleijpen

Henk van der Vorst

Diederik Fokkema

Universiteit Utrecht

Department of Mathematics

http://www.math.ruu.nl/people/sleijpen

Program

I. Standard Eigenvalue problems

- Overview of classical methods
- Jacobi-Davidson
- Deflation
- Using a preconditioner

II. Generalized Eigenvalue problems

- Jacobi-Davidson QZ
- Harmonic Ritz values
- Preconditioning

Galerkin approach

- Generate a subspace with basis $\mathbf{v}_1, \ldots, \mathbf{v}_k$.
- With $\mathbf{V}_k := [\mathbf{v}_1 | \ldots | \mathbf{v}_k]$ find y and θ such that

$$\mathbf{r} := \mathbf{A}\mathbf{V}_k\, y - \theta \mathbf{V}_k\, y \perp \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$$

$(\theta, \mathbf{V}_k\, y)$ **Ritz pair** w.r.t. **search space** span$(\mathbf{V}_k)$.

$(\theta, \mathbf{V}_k\, y)$ with **residual** $\mathbf{r}$ approximates $(\lambda, \mathbf{x}_\lambda)$.

In practice. $\quad k \ll n$

∘ Generate an orthonormal system $\mathbf{v}_1, \ldots, \mathbf{v}_k$:

$\mathbf{t}$ expansion vector $\mathbf{V}_k \;\Rightarrow\; \mathbf{v}_{k+1} \equiv \text{modGS}_{\mathbf{V}_k}(\mathbf{t})$.

∘ Solve the small problem

$$\mathbf{V}_k^* \mathbf{A}\mathbf{V}_k\, y = \theta\, y$$

Subspace iteration

1. **Start:**

   Choose an initial subspace $\mathbf{V}$.

2. **Repeat:**

   a. Compute $\mathbf{W} \leftarrow \mathbf{A}\mathbf{V}$, $H \leftarrow \mathbf{V}^*\mathbf{W}$.

   b. Compute a desired eigenpair (θ, y) of H (with $\| y \| = 1$).

   c. Compute $\mathbf{u} \leftarrow \mathbf{V} y$,

   $$\mathbf{r} \leftarrow \mathbf{A}\,\mathbf{u} - \theta\,\mathbf{u}.$$

   d. **Stop if** satisfied.

   e. Compute an expansion vector $\mathbf{t}$

   f. $\mathbf{V} \leftarrow \text{ModGS}([\,\mathbf{V}\,|\,\mathbf{t}\,])$

Some practical notes

f. $\mathbf{v} \leftarrow \text{ModGS}_{\mathbf{V}}(\mathbf{t})$, $\mathbf{V} \leftarrow [\,\mathbf{V}\,|\,\mathbf{v}\,]$

Alternative: repeated classical Gram-Schmidt

a. Only last column of $\mathbf{W}$ and last row & column of H have to be updated:

$$H \leftarrow \begin{bmatrix} H \\ \mathbf{v}^*\mathbf{W} \end{bmatrix},\; H \leftarrow [H\,|\,\mathbf{V}^*\mathbf{w}],\; \mathbf{W} \leftarrow [\,\mathbf{W}\,|\,\mathbf{w}\,]$$

b. Use QR (LAPACK) to decompose H:

$H = USU^*$ with

$U^*U = I$ & S upper triangular

Order S such that S_{11} is the wanted θ.

Then $y = U(:,1)$

c. $\mathbf{r} \leftarrow \mathbf{W}y - \theta\mathbf{u}$

Modified Gram–Schmidt

$$
\boxed{
\begin{array}{l}
\text{for} \quad i = 1, \ldots, k \quad \text{do} \\
\qquad \beta_i = \mathbf{v}_i^* \mathbf{t} \\
\qquad \mathbf{t} \leftarrow \mathbf{t} - \beta_i \, \mathbf{v}_i \\
\text{end for} \\
\mathbf{v} = \mathbf{t} / \|\mathbf{t}\|
\end{array}
}
$$

Project and subtract the components one by one: compute one component and subtract that component, . . .

Any error is projected and components are subtracted also.

Modified Gram-Schmidt is a less unstable than classical Gram-Schmidt. Its stability is satisfactory in practice for linear solvers.
As expensive as classical Gram-Schmidt (on non-parallel machines)

[Daniel-Gragg-Kaufman-Stewart (DGKS)]

Repeated Gram-Schmidt

$\rho = \mathbf{t}^*\mathbf{t}, \quad b = \mathbf{V}_k^*\mathbf{t}$

$\mathbf{t} \leftarrow \mathbf{t} - \mathbf{V}_k\, b$

$\sigma = \mathbf{t}^*\mathbf{t},$

while $\sigma < \kappa\,\rho$

$\quad c = \mathbf{V}_k^*\mathbf{t}$

$\quad \mathbf{t} \leftarrow \mathbf{t} - \mathbf{V}_k\, c$

$\quad b \leftarrow b + c$

$\quad \rho \leftarrow \sigma, \quad \sigma \leftarrow \mathbf{t}^*\mathbf{t}$

end while

$\mathbf{v} = \mathbf{t}/\sqrt{\sigma}$

e.g., $\kappa = 0.5$

In practice:

- Appears to be stable
- At most 1 refinement step needed (Only a little more expensive than Gram-Schmidt)

Theory:

Proof for "once is enough" only if $k = 1$.

Choices for the expansion vector

with initial search subspace span($\mathbf{v}_1$).

Arnoldi's method ['51]

$$\mathbf{t} = \mathbf{A}\mathbf{v}_k\,.$$

Then $H := \mathbf{V}^*\mathbf{A}\mathbf{V}$ is upper Hessenberg.

Lanczos' method if $\mathbf{A}^* = \mathbf{A}$ ['50]

$$\mathbf{t} = \mathbf{A}\mathbf{v}_k\,.$$

Then $H := \mathbf{V}^*\mathbf{A}\mathbf{V}$ is tridiagonal.

Search space span($\mathbf{V}_k$) $=$ Krylov subspace

$$\text{span}\Big(\mathbf{v}_1\,, \mathbf{A}\mathbf{v}_1\,, \ldots, \mathbf{A}^{k-1}\,\mathbf{v}_1\Big).$$

Arnoldi (Lanczos) favors extremal eigenvalues

Shift-and-Invert-Arnoldi

Select target $\tau \in \mathbb{C}$.

$$\mathbf{t} = (\mathbf{A} - \tau\,\mathbf{I})^{-1}\mathbf{v}_k\,.$$

Then $H = \mathbf{V}^*\mathbf{A}\mathbf{V}$ is upper Hessenberg.

Approach favors eigenvalues close to τ

(amplifies $\mathbf{x}_\lambda$-comp. with $\frac{1}{\lambda-\tau}$)

Rational Krylov Subspace it.

[Ruhe '84, '93]

$$\mathbf{t} = (\mathbf{A} - \tau_k\,\mathbf{I})^{-1}(\mathbf{A} - \sigma_k\mathbf{I})\mathbf{v}_k\,.$$

(amplifies $\mathbf{x}_\lambda$-comp. with $\frac{\lambda-\sigma_k}{\lambda-\tau_k}$)

However, for large problems,

"solve $(\mathbf{A} - \tau\,\mathbf{I})\mathbf{t} = \mathbf{v}_k$"

is expensive and has to be done accurately.

Alternatives

Davidson's method ['75]

$$\mathbf{t} = (\mathrm{diag}(\mathbf{A}) - \theta\,\mathbf{I})^{-1}\mathbf{r}$$

with $\quad \mathbf{r} := \mathbf{A}\mathbf{u} - \theta\,\mathbf{u}\ \ (\mathbf{u} := \mathbf{V}y,\ \mathbf{V}^*\mathbf{A}\mathbf{V}y = \theta y).$

Generalized Davidson [Morgan '90]

$$\mathbf{K} \approx \mathbf{A} - \theta\,\mathbf{I} \quad \text{and} \quad \mathbf{t} = \mathbf{K}^{-1}\mathbf{r}$$

If $\quad \mathbf{t} = (\mathbf{A} - \theta\,\mathbf{I})^{-1}\mathbf{r} \quad$ then **stagnation**.

If $\quad \mathbf{t} = \mathbf{r} = (\mathbf{A} - \theta\,\mathbf{I})\mathbf{u} \quad$ then **Arnoldi**:

$$\mathrm{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{A}\mathbf{v}_k) = \mathrm{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k, (\mathbf{A} - \theta\,\mathbf{I})\mathbf{u}).$$

Expansion with Ritz vectors

$$\mathbf{K} \approx \mathbf{A} - \theta\,\mathbf{I}$$

Inexact Shift-and-Invert,

$$\mathbf{t} = \mathbf{K}^{-1}\mathbf{v}_k\,,$$

aims for eigensolution $\quad \mathbf{K}$.

Generalized Davidson,

$$\mathbf{t} = \mathbf{K}^{-1}\mathbf{r} = \mathbf{K}^{-1}(\mathbf{A} - \theta\,\mathbf{I})\mathbf{u},$$

aims for eigensolution $\quad \mathbf{K}^{-1}(\mathbf{A} - \lambda\,\mathbf{I})\mathbf{x}_\lambda = 0$.

A Gen. Davidson $\approx$ Cayley Transform Arnoldi:

$$\mathbf{t} = (\mathbf{A} - \tau\,\mathbf{I})^{-1}(\mathbf{A} - \sigma\,\mathbf{I})\mathbf{u}.$$

Towards optimal expansion

Consider the Ritz approximation $(\theta, \mathbf{u})$

$$(\mathbf{u} := \mathbf{V}y,\ \|\mathbf{u}\| = 1)$$

$$\theta := \mathbf{u}^*\mathbf{A}\mathbf{u}, \qquad \mathbf{r} := \mathbf{A}\,\mathbf{u} - \theta\,\mathbf{u} \perp \mathbf{V}$$

Find $\mathbf{t} \perp \mathbf{u}$ such that $\mathbf{x} = \mathbf{u} + \mathbf{t}$.

$$(\mathbf{A} - \lambda\,\mathbf{I})\,\mathbf{t} = -\mathbf{r} + (\lambda - \theta)\,\mathbf{u}$$

or, equivalently,

$$\begin{cases} \mathbf{u}^*\mathbf{A}\mathbf{x} = \lambda\,\mathbf{u}^*\mathbf{x} \\ (\mathbf{I} - \mathbf{u}\,\mathbf{u}^*)(\mathbf{A} - \lambda\,\mathbf{I})(\mathbf{I} - \mathbf{u}\,\mathbf{u}^*)\,\mathbf{t} = -\mathbf{r} \end{cases}$$

[Sleijpen & van der Vorst '96]

Jacobi-Davidson

Since λ is unknown:

solve $\quad \mathbf{t} \perp \mathbf{u} \quad$ (approximately) from

the correction equation

$$\boxed{(\mathbf{I} - \mathbf{u}\,\mathbf{u}^*)\,(\mathbf{A} - \theta\,\mathbf{I})\,(\mathbf{I} - \mathbf{u}\,\mathbf{u}^*)\,\mathbf{t} = -\mathbf{r}}$$

Jacobi-Davidson

1. **Start:**

   Choose initial orth. matrix $\mathbf{V}$.

2. **Repeat:**

   a. Compute $\mathbf{W} \leftarrow \mathbf{A}\mathbf{V}$, $H \leftarrow \mathbf{V}^*\mathbf{W}$.

   b. Compute a desired eigenpair (θ, y) of H (with $\| y \| = 1$).

   c. Compute $\mathbf{u} \leftarrow \mathbf{V} y$,
   $\mathbf{r} \leftarrow \mathbf{A}\,\mathbf{u} - \theta\,\mathbf{u}$.

   d. **Stop if** satisfied.

   e. Solve (approximately)

   $$(\mathbf{I} - \mathbf{u}\,\mathbf{u}^*)\,(\mathbf{A} - \theta\,\mathbf{I})\,(\mathbf{I} - \mathbf{u}\,\mathbf{u}^*)\,\mathbf{t} = -\mathbf{r}.$$

   f. $\mathbf{V} \leftarrow \mathrm{ModGS}([\,\mathbf{V}\,|\,\mathbf{t}\,])$

$$(\mathbf{I}-\mathbf{u}\,\mathbf{u}^*)\,(\mathbf{A}-\theta\,\mathbf{I})\,(\mathbf{I}-\mathbf{u}\,\mathbf{u}^*)\,\mathbf{t}=-\mathbf{r}.$$

Any preconditioned it. method can be used.

Preconditioning: $\quad \mathbf{K}\approx\mathbf{A}-\theta\,\mathbf{I}$

Preconditioning equation:

$$(\mathbf{I}-\mathbf{u}\,\mathbf{u}^*)\,\mathbf{K}\,(\mathbf{I}-\mathbf{u}\,\mathbf{u}^*)\,\mathbf{t}=-\mathbf{r}$$

Solution:

$$\mathbf{K}\,\mathbf{t}=-\mathbf{r}+\mathbf{u}\,\alpha \quad \text{such that} \quad \mathbf{t}\perp\mathbf{u},$$

or

$$\mathbf{t}=-\mathbf{K}^{-1}\mathbf{r}+\mathbf{K}^{-1}\mathbf{u}\,\alpha \qquad \&$$

$$\alpha=(\mathbf{u}^*\mathbf{K}^{-1}\mathbf{u})^{-1}(\mathbf{u}^*\mathbf{K}^{-1}\mathbf{r}).$$

Solve approximately

$$(\mathbf{I}-\mathbf{u}\,\mathbf{u}^*)\,(\mathbf{A}-\theta\,\mathbf{I})\,(\mathbf{I}-\mathbf{u}\,\mathbf{u}^*)\,\mathbf{t}=-\mathbf{r}$$

$$\neq$$

Solve approximately

$$\left.\begin{array}{l}(\mathbf{A}-\theta\,\mathbf{I})\,\mathbf{t}_r=\mathbf{r}\\(\mathbf{A}-\theta\,\mathbf{I})\,\mathbf{t}_u=\mathbf{u}\end{array}\right\}\quad\&\quad\begin{array}{l}\mathbf{t}=-\mathbf{t}_r+\alpha\,\mathbf{t}_u\\\text{s.t.}\quad\mathbf{t}\perp\mathbf{u}\end{array}$$

except if

"solve appr." $=$ 1 step with fixed precond. $\mathbf{K}$.

Example. With static polynomial method p:

$$p\,((\mathbf{I}-\mathbf{u}\,\mathbf{u}^*)\,(\mathbf{A}-\theta\,\mathbf{I})\,(\mathbf{I}-\mathbf{u}\,\mathbf{u}^*))\neq$$

$$(\mathbf{I}-\mathbf{u}\,\mathbf{u}^*)\,p\,(\mathbf{A}-\theta\,\mathbf{I})\,(\mathbf{I}-\mathbf{u}\,\mathbf{u}^*)$$

Related methods

with $\mathbf{K}$ a fixed preconditioner

Expand the search space by the sol. $\mathbf{t}$ of

- $\mathbf{K}\,\mathbf{t} = -\mathbf{r} + \alpha\,\mathbf{u} \qquad \& \qquad \alpha = \mathbf{u}^*\mathbf{K}^{-1}\mathbf{t}/\mathbf{u}^*\mathbf{K}^{-1}\mathbf{u}$

  Olsen's method ['90 for symm. $\mathbf{A}$]

- $\mathbf{K}\,\mathbf{t} = -\mathbf{r}$

  Generalized Davidson [Morgan '90]

- $\mathbf{K}\,\mathbf{t} = \mathbf{u}$

  Inexact Shift-and-Invert.

$$\mathbf{t} = -\overbrace{\mathbf{K}^{-1}\mathbf{r}}^{\text{generalized Davidson}} + \alpha\,\overbrace{\mathbf{K}^{-1}\mathbf{u}}^{\text{inexact S\&I}}$$

Extreme choices

- $\mathbf{K} = \mathbf{I}$:

  $$\mathbf{t} = -\mathbf{r} + \alpha\,\mathbf{u} = -\mathbf{A}\,\mathbf{u} + (\theta + \alpha)\mathbf{u}.$$

  $\Rightarrow$ Effective expansion vector: $\mathbf{A}\,\mathbf{u}$

  J-D $\equiv$ Gen. Dav. $\equiv$ Arnoldi.

  Inexact Shift-and-Invert stagnates.

- $\mathbf{K} = \mathbf{A} - \theta\,\mathbf{I}$:

  $$\mathbf{t} = -\mathbf{u} + \alpha\,(\mathbf{A} - \theta\,\mathbf{I})^{-1}\mathbf{u}.$$

  $\Rightarrow$ Effective expansion v.: $(\mathbf{A} - \theta\,\mathbf{I})^{-1}\mathbf{u}$

  J-D $\equiv$ Shift-and-Invert with optimal shift.

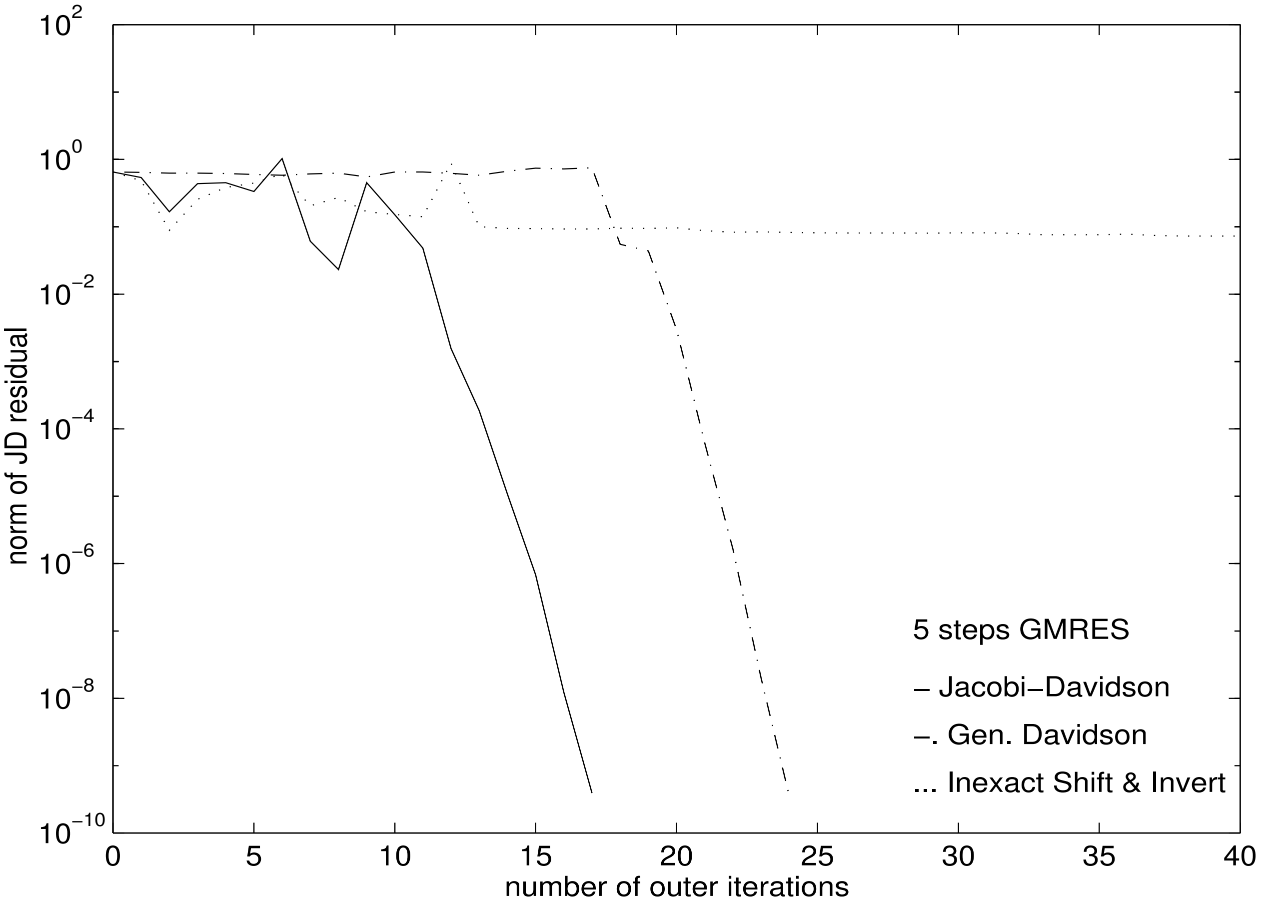
  Generalized Davidson stagnates.

$\lambda \approx 0, \quad \mathbf{A} = \text{SHERMAN4} - .5\,\mathbf{I}, \quad n = 1104$

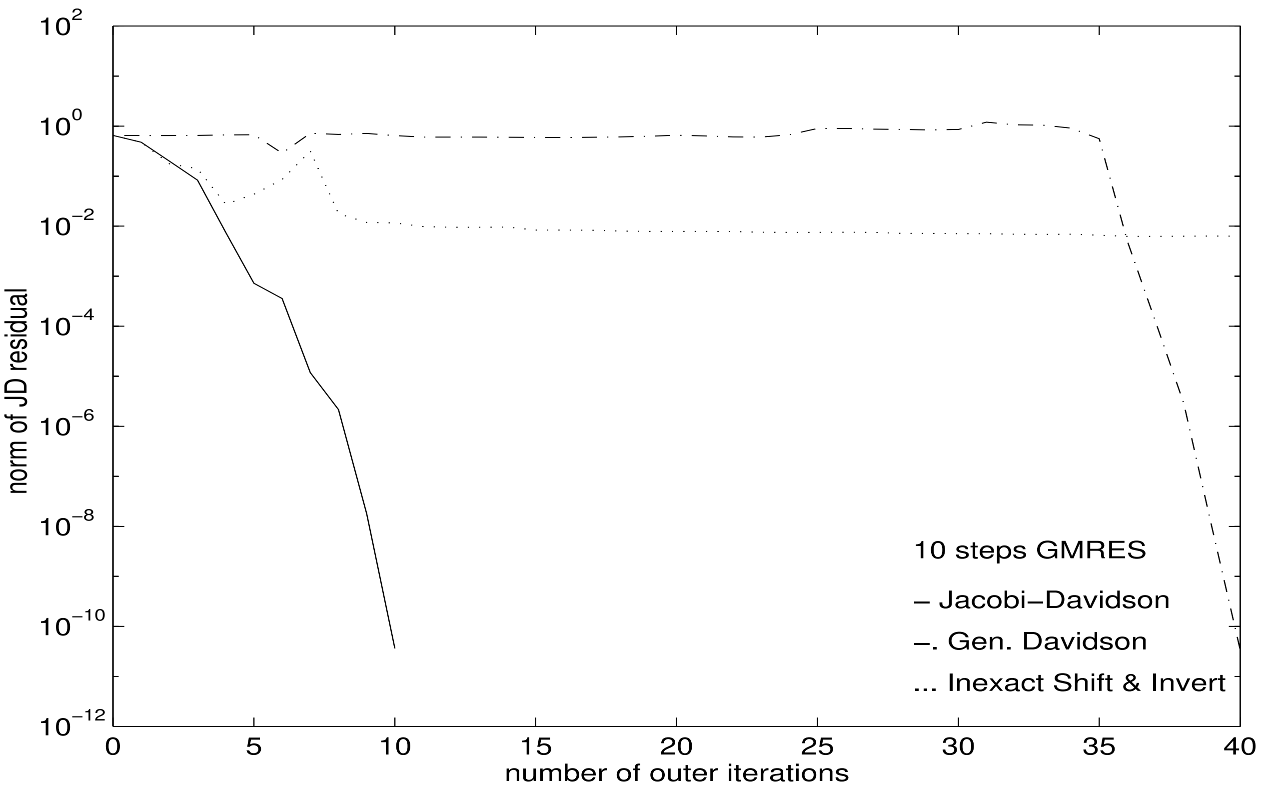
CorEq:

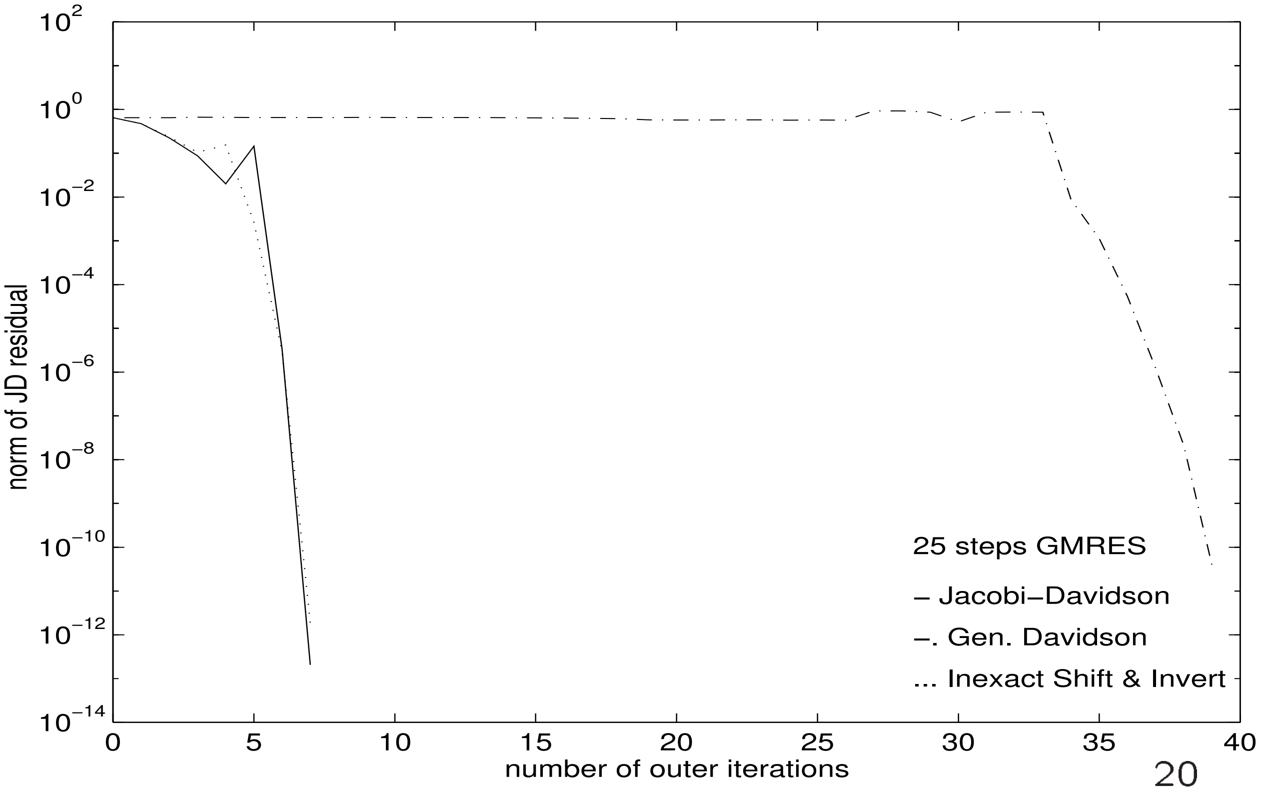
- $(\mathbf{I} - \mathbf{u}\,\mathbf{u}^*)(\mathbf{A} - \theta\,\mathbf{I})(\mathbf{I} - \mathbf{u}\,\mathbf{u}^*)\,\mathbf{t} = -\mathbf{r}$ [J-D]
- $(\mathbf{A} - \theta\,\mathbf{I})\,\mathbf{t} = -\mathbf{r}$ ["Davidson"]
- $(\mathbf{A} - \theta\,\mathbf{I})\,\mathbf{t} = \mathbf{u}$ ["Shift-and-Invert"]

m steps precond. GMRES with $\mathbf{K} = \text{ILU}(2)$.



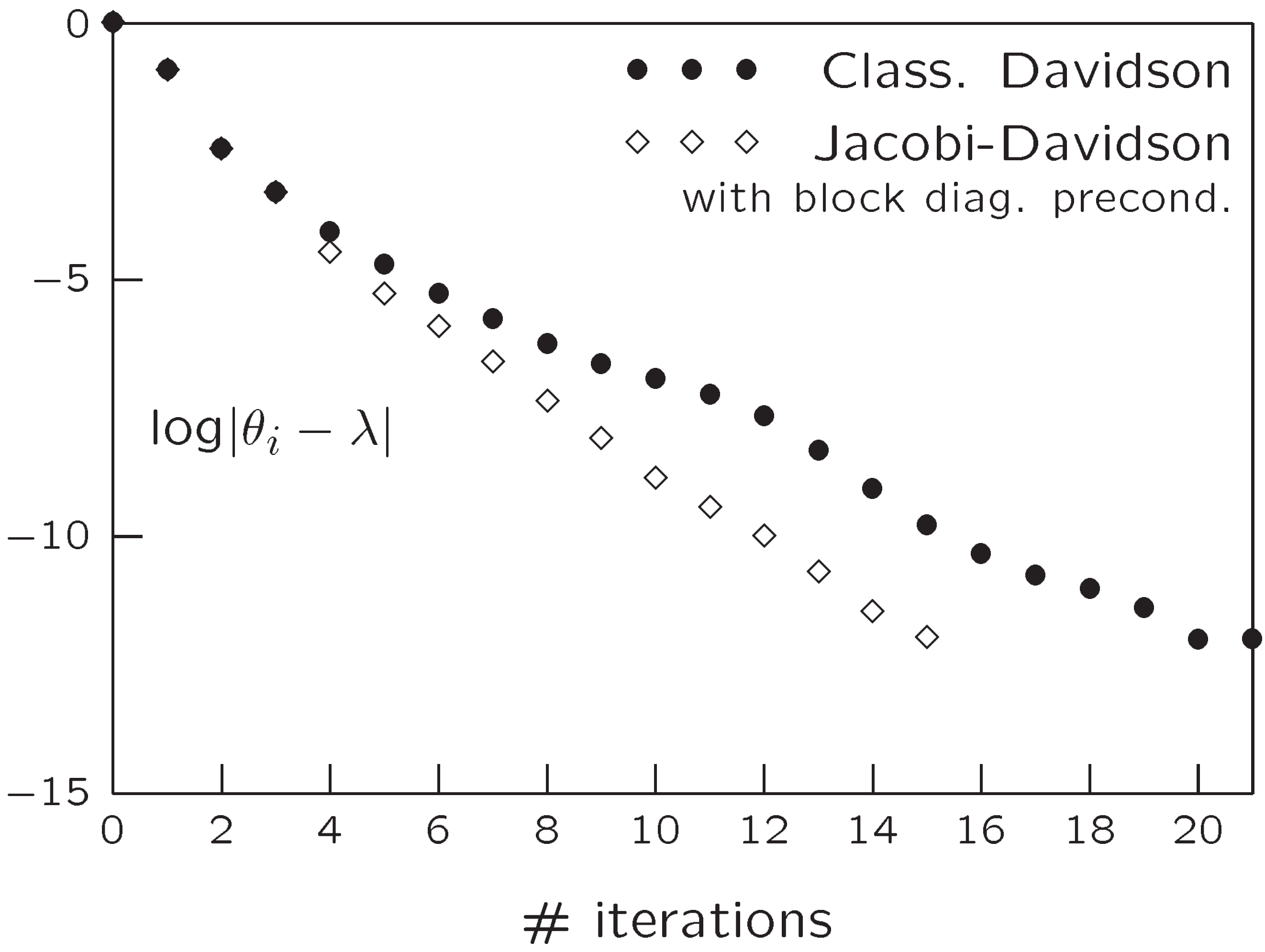
$m = 5$


norm of JD residual
number of outer iterations
10 steps GMRES
– Jacobi–Davidson
–. Gen. Davidson
... Inexact Shift & Invert


norm of JD residual
number of outer iterations
25 steps GMRES
– Jacobi–Davidson
–. Gen. Davidson
... Inexact Shift & Invert

[van Dam, van Lenthe '94]

Example [Quantum Chemistry]



$\mathbf{A}$ symmetric, $n \approx 50,000$

block diag. precond. with max. block size 66

J-D as a Newton scheme

$$\mathbf{F}(\mathbf{u},\theta) \equiv \mathbf{A}\,\mathbf{u} - \theta\,\mathbf{u}, \qquad \|\,\mathbf{u}\,\| = 1,\ \theta \in \mathbb{C}$$

$$\mathbf{F} : \{\mathbf{u}, \|\,\mathbf{u}\,\| = 1\} \times \mathbb{C} \to \mathbb{C}^n$$

$$\boxed{\mathbf{A}\,\mathbf{x} = \lambda\,\mathbf{x} \quad \& \quad \|\,\mathbf{x}\,\| = 1} \quad \Leftrightarrow \quad \boxed{\mathbf{F}(\mathbf{x},\lambda) = 0}$$

Newton:

- solve $\mathbf{t} \perp \mathbf{u}$

$$\frac{\partial\,\mathbf{F}}{\partial\,(\mathbf{u},\theta)}(\mathbf{t},\epsilon) = -\mathbf{r} \equiv -\mathbf{F}(\mathbf{u},\theta)$$

- $\mathbf{u} \leftarrow \mathbf{u} + \mathbf{t}, \qquad \theta \leftarrow \theta + \epsilon$

The sol. $\mathbf{t}$ of the Newton correction satisfies

$$\mathbf{t} \perp \mathbf{u}, \qquad (\mathbf{I} - \mathbf{u}\,\mathbf{u}^*)\,(\mathbf{A} - \theta\,\mathbf{I})\,\mathbf{t} = -\mathbf{r}$$

J-D $=$ accelerated (inexact) Newton

$\Rightarrow$ *($\approx$) quadratic convergence*

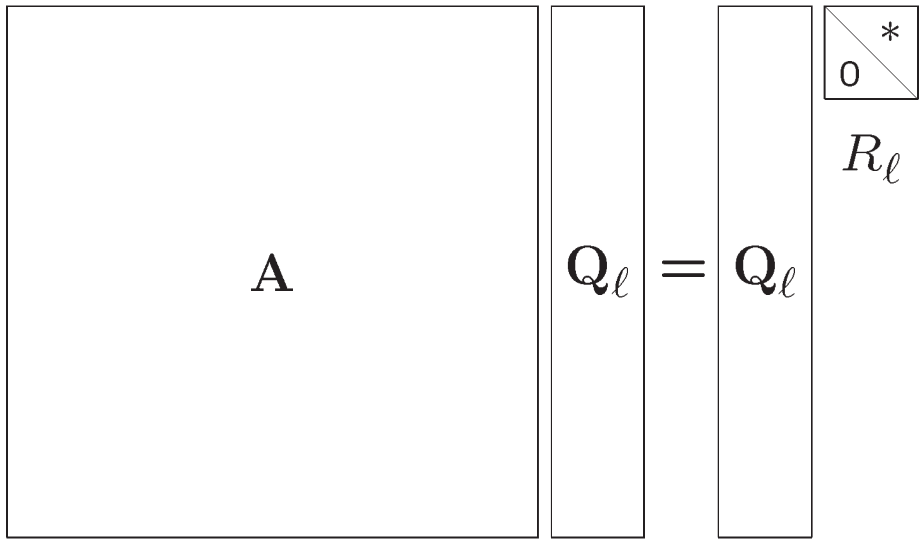
[Fokkema, S, van der Vorst '96]

JDQR

Compute ℓ eigenvalues close to some $\tau \in \mathbb{C}$.

- Deflation
- Schur vectors
- Restart

Partial Schur form



$\mathbf{A}\,\mathbf{Q}_\ell = \mathbf{Q}_\ell\,R_\ell$ with $\mathbf{Q}_\ell^*\mathbf{Q}_\ell = I$

R upper Δ

$\mathrm{diag}(R_\ell) = (\lambda_1, \ldots, \lambda_\ell)$

$|\lambda_i - \tau|$ increases

If $R_\ell\,y_i = \lambda_i\,y_i$ then $\mathbf{A}\,\mathbf{Q}_\ell\,y_i = \lambda_i\,\mathbf{Q}_\ell\,y_i$.

Finding other eigenpairs

Expansion of the partial Schur form:

$$\mathbf{A}\,[\,\mathbf{Q}_\ell\,|\,\mathbf{q}_{\ell+1}\,] = [\,\mathbf{Q}_\ell\,|\,\mathbf{q}_{\ell+1}\,]\begin{bmatrix} R_\ell & s \\ & \lambda_{\ell+1} \end{bmatrix}$$

with $\quad \mathbf{q}_{\ell+1}^*\mathbf{Q}_\ell = 0.$

Hence

$$(\mathbf{I}-\mathbf{Q}_\ell\,\mathbf{Q}_\ell^*)(\mathbf{A}-\lambda_{\ell+1}\mathbf{I})(\mathbf{I}-\mathbf{Q}_\ell\,\mathbf{Q}_\ell^*)\mathbf{q}_{\ell+1} = \mathbf{0}$$

which means
that $(\lambda_{\ell+1}, \mathbf{q}_{\ell+1})$ is an eigenpair of

$$\widetilde{\mathbf{A}} := (\mathbf{I}-\mathbf{Q}_\ell\,\mathbf{Q}_\ell^*)\mathbf{A}(\mathbf{I}-\mathbf{Q}_\ell\,\mathbf{Q}_\ell^*)$$

JDQR: apply JD for $\widetilde{\mathbf{A}}$.

Projected matrices

The projected matrices in JDQR satisfy:

$$\mathbf{V}_k^*(\mathbf{I}-\mathbf{Q}_\ell\mathbf{Q}_\ell^*)\mathbf{A}(\mathbf{I}-\mathbf{Q}_\ell\mathbf{Q}_\ell^*)\mathbf{V}_k = \mathbf{V}_k^*\mathbf{A}\mathbf{V}_k$$

Correction Equation

$(\theta,\mathbf{u})$ Ritz pair $\widetilde{\mathbf{A}}$ $\quad\Rightarrow\quad$ $\mathbf{u}\perp\mathbf{Q}_\ell$

Solve $\quad \mathbf{t}\perp\widetilde{\mathbf{Q}}\equiv[\,\mathbf{Q}_\ell\,|\,\mathbf{u}\,]$ $\quad$ (appr) from

$$(\mathbf{I}-\widetilde{\mathbf{Q}}\,\widetilde{\mathbf{Q}}^*)(\mathbf{A}-\theta\mathbf{I})(\mathbf{I}-\widetilde{\mathbf{Q}}\,\widetilde{\mathbf{Q}}^*)\mathbf{t}=-\mathbf{r},$$

with

$$\mathbf{r}=(\mathbf{I}-\mathbf{Q}_\ell\,\mathbf{Q}_\ell^*)(\mathbf{A}\mathbf{u}-\theta\mathbf{u}).$$

Implicit restart

For λ close to a given **target** $\tau \in \mathbb{C}$

- For $\mathbf{V}$, compute $H := \mathbf{V}^*\mathbf{A}\mathbf{V}$

$$\left.\begin{array}{c} H = USU^* \ \text{ with } \ U^*U = I, \\ S \text{ upper triangular} \end{array}\right\} \text{QR}$$

- Order S such that $\quad |S_{jj} - \tau| \quad$ increases.

First (few) columns of reordered $\mathbf{V}$
represent best eigenvector approx.
$\Rightarrow$ can be used for selection and restart

- Restart with $\mathbf{V} \leftarrow \mathbf{V}[u_1|\ldots|u_{j_{\text{min}}}]$ if needed.

$$\boxed{\begin{array}{l} \textbf{If } \dim(\mathbf{V}) = j_{\text{max}} \\ \bullet\ \mathbf{V} \leftarrow \mathbf{V}U(:, 1:j_{\text{min}}) \\ \bullet\ H \leftarrow S(1:j_{\text{min}}, 1:j_{\text{min}}) \end{array}}$$

JDQR

1. **Start:** Choose an initial space $\mathbf{V}$.
   - Set $\Lambda = [\,]$, $\mathbf{Q} = [\,]$
2. **Repeat:**

   a. Compute $\mathbf{W} \leftarrow \mathbf{A}\mathbf{V}$, $\quad H \leftarrow \mathbf{V}^*\mathbf{W}$

   b. Compute $\quad H = USU^*$

   with $\quad U^*U = I$, $\quad S$ upper Δ,

   $|S_{11} - \tau| \leq |S_{22} - \tau| \leq \ldots$.

   c. $\theta \leftarrow S_{11}$, $y \leftarrow Ue_1$, $\mathbf{u} \leftarrow \mathbf{V}y$,

   $\mathbf{r} \leftarrow \text{ModGS}_{\mathbf{Q}}(\mathbf{A}\,\mathbf{u} - \theta\,\mathbf{u})$.

   d. **If** $\|\,\mathbf{r}\,\| \leq$ *tol*
   - $\Lambda \leftarrow [\,\Lambda\,|\,\theta\,]$, $\qquad \mathbf{Q} \leftarrow [\,\mathbf{Q}\,|\,\mathbf{u}\,]$.
   - $\mathbf{V} \leftarrow \mathbf{V}U(:, 2:\dim(\mathbf{V}))$, $U \leftarrow I$,

     $S \leftarrow S(2:\dim(\mathbf{V}), 2:\dim(\mathbf{V}))$.
   - **Goto** 2.c.

   e. Solve (app.) $\mathbf{t} \perp \widetilde{\mathbf{Q}} \equiv [\,\mathbf{Q}\,|\,\mathbf{u}\,]$

   $$(\mathbf{I} - \widetilde{\mathbf{Q}}\,\widetilde{\mathbf{Q}}^*)\,(\mathbf{A} - \theta\,\mathbf{I})\,(\mathbf{I} - \widetilde{\mathbf{Q}}\,\widetilde{\mathbf{Q}}^*)\,\mathbf{t} = -\mathbf{r}.$$

   f. $\mathbf{V} \leftarrow \text{ModGS}([\,\mathbf{V}\,|\,\mathbf{t}\,])$

Properties JDQR

- Leads to partial Schur form.
- Employs orthonormal sets of vectors.
- Finds also multiple eigenvalues.
- Easy selection of wanted eigenvalues.
- No matrix inversion involved.
- Preconditioner can be included.
- Ingredients:

  vector updates, Dot-products, MVs,
  preconditioners
  QR for small problems.

Some practical notes

- After convergence of some eigenvectors, solution of correction equation easier.

- Number of pojections can be limited.

- Correction equation can be preconditioned.

  Preconditioner can be used more often: pays to make better preconditioning.

- For generalized eigenproblems:

  JD + QZ-reduction

Some practical notes

- ◦ Maximum dimension $\mathbf{V}$ is 15, after implicit restart $\dim(\mathbf{V}) = 10$.

- Solve cor. eq. with GMRES (or BiCGstab(2)) with residual accuracy

  $$\|\tilde{\mathbf{r}}_{\text{GMRES}}\| \leq 2^{-j}\|\mathbf{r}\|$$

  but with atmost p GMRES steps ($p = 1, 5, 10$).

  Here,

  $\mathbf{r}$ = JD residual = initial GMRES residual,

  j is # JD steps in search for λ,

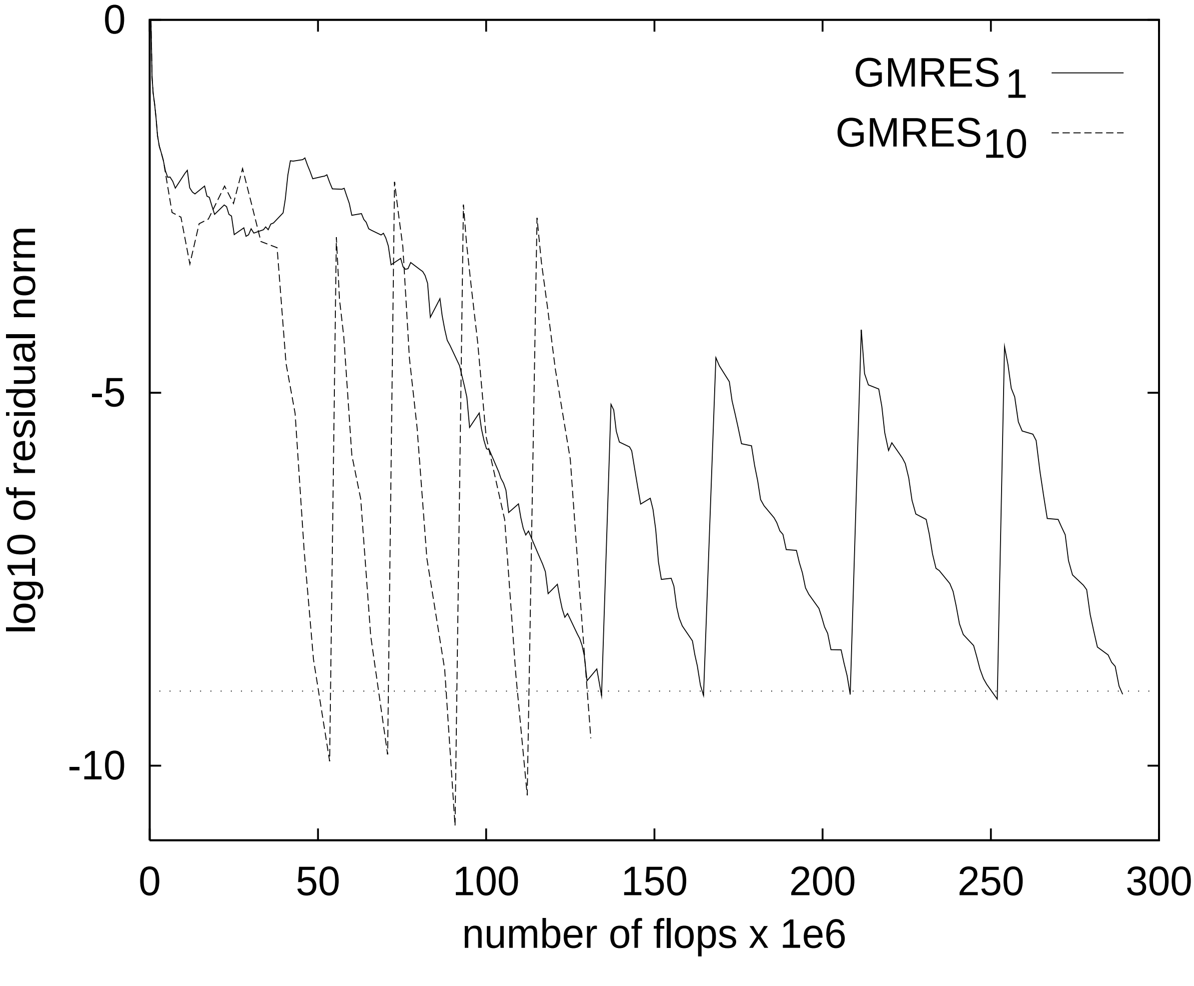
  j is reset to 0 after detection of λ.

- Start: built space $\mathbf{V}$ of dim j_{min} as quick as pos.: in first j_{min} steps (only)
  - τ instead of θ
  - GMRES_1 instead of GMRES_p

DW4096: Square Dielectric Waveguide
standard eigenproblem

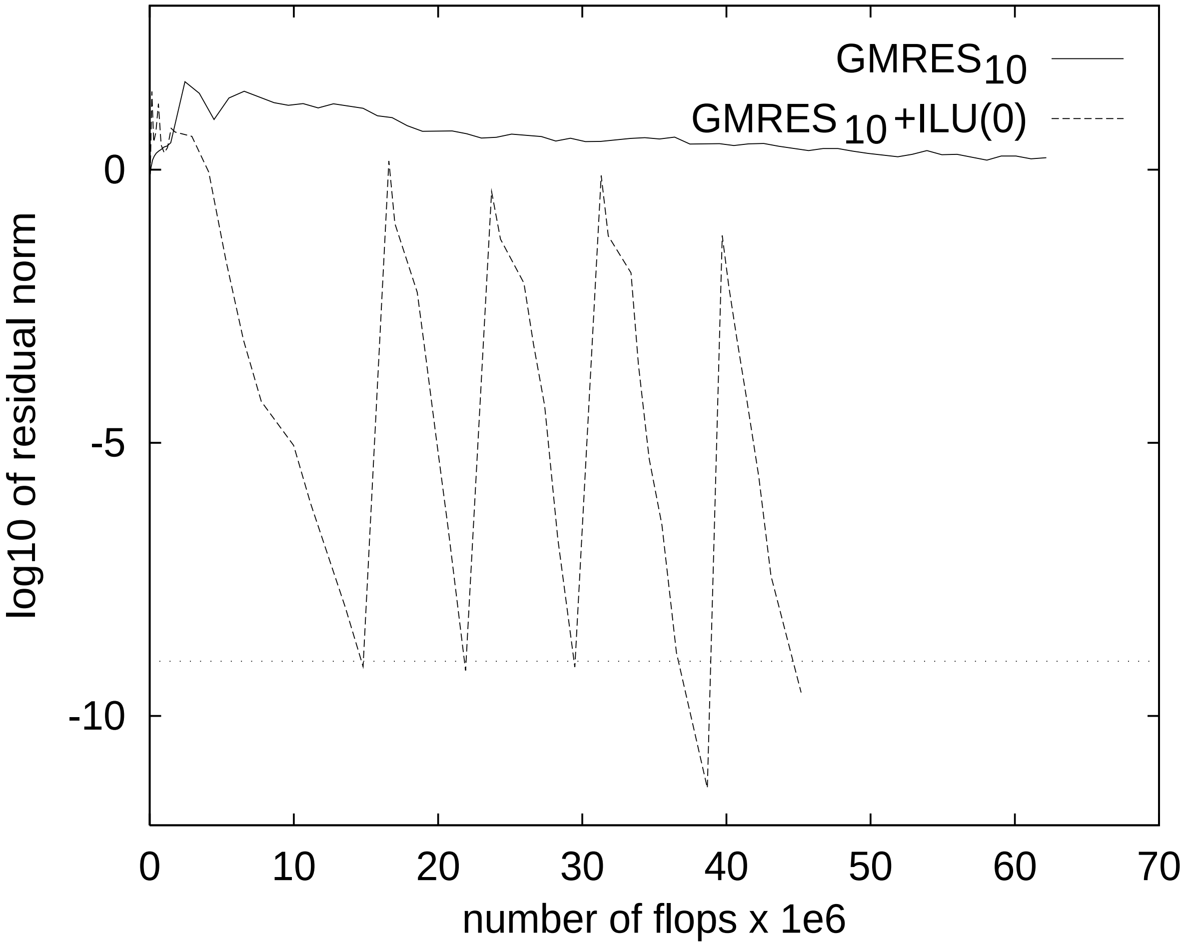
Searching for the five rightmost eigenvalues:

$$8.6479e-01, \quad 8.6528e-01, \quad 8.6564e-01,$$
$$8.6594e-01, \quad 8.6657e-01$$



BWM2000: Brusselator wave model
standard eigenproblem

Searching for the five rightmost eigenvalues.



Solve approximately

$$(\mathbf{I}-\mathbf{u}\,\mathbf{u}^*)\,(\mathbf{A}-\theta\,\mathbf{I})\,(\mathbf{I}-\mathbf{u}\,\mathbf{u}^*)\,\mathbf{t}=-\mathbf{r}.$$

Any preconditioned it. method can be used.

Preconditioning: $\mathbf{K}\approx\mathbf{A}-\theta\,\mathbf{I}$

Preconditioning equation:

$$(\mathbf{I}-\mathbf{u}\,\mathbf{u}^*)\,\mathbf{K}\,(\mathbf{I}-\mathbf{u}\,\mathbf{u}^*)\,\mathbf{c}=\mathbf{s}$$

Solution:

$\mathbf{K}\,\mathbf{c}=\mathbf{s}-\mathbf{u}\,\alpha$ such that $\mathbf{c}\perp\mathbf{u}$,

or

$\mathbf{c}=\mathbf{K}^{-1}\mathbf{s}-\mathbf{K}^{-1}\mathbf{u}\,\alpha$ where, with

$\mathbf{y}\equiv\mathbf{K}^{-1}\mathbf{u}$ & $\mu\equiv\mathbf{u}^*\mathbf{y}$, $\alpha=\mu^{-1}(\mathbf{u}^*\mathbf{K}^{-1}\mathbf{s})$.

or

$$\mathbf{c}=\mathbf{K}^{\dagger}\mathbf{s}\equiv(\mathbf{I}-\mathbf{y}\mu^{-1}\mathbf{u}^*)\mathbf{K}^{-1}\mathbf{s}$$

Solve approximately

$$\left(\mathbf{I}-\widetilde{\mathbf{Q}}\,\widetilde{\mathbf{Q}}^*\right)(\mathbf{A}-\theta\,\mathbf{I})\left(\mathbf{I}-\widetilde{\mathbf{Q}}\,\widetilde{\mathbf{Q}}^*\right)\mathbf{t}=-\mathbf{r}.$$

Deflated preconditioning

$$(\mathbf{I}-\widetilde{\mathbf{Q}}\,\widetilde{\mathbf{Q}}^*)\,\mathbf{K}\,(\mathbf{I}-\widetilde{\mathbf{Q}}\,\widetilde{\mathbf{Q}}^*)\,\mathbf{c}=\mathbf{s}.$$

Solution:

$$\mathbf{K}\,\mathbf{c}=\mathbf{s}-\widetilde{\mathbf{Q}}\,y \qquad \text{such that} \qquad \mathbf{c}\perp\mathbf{Q}$$

or

$$\mathbf{c}=\mathbf{K}^{-1}\mathbf{s}-\mathbf{K}^{-1}\widetilde{\mathbf{Q}}\,a \qquad \&$$

$$a=(\underbrace{\widetilde{\mathbf{Q}}^*\,\mathbf{K}^{-1}\,\widetilde{\mathbf{Q}}}_{\text{small matrix}})^{-1}\,\widetilde{\mathbf{Q}}^*\,\mathbf{K}^{-1}\mathbf{s}.$$

replaces μ

$\widetilde{\mathbf{Q}}$ replaces $\mathbf{u}$

$$(\mathbf{I}-\mathbf{u}\,\mathbf{u}^*)\,(\mathbf{A}-\theta\,\mathbf{I})\,(\mathbf{I}-\mathbf{u}\,\mathbf{u}^*)\,\mathbf{t}=-\mathbf{r}$$

$$\mathbf{K}\approx\mathbf{A}-\theta\,\mathbf{I},\qquad(\mathbf{I}-\mathbf{u}\,\mathbf{u}^*)\,\mathbf{K}\,(\mathbf{I}-\mathbf{u}\,\mathbf{u}^*)\,\mathbf{c}=\mathbf{s}.$$

$\mathbf{K}^\dagger\equiv(\mathbf{I}-\mathbf{y}\mu^{-1}\mathbf{u}^*)\mathbf{K}^{-1}$ with $\mathbf{y}\equiv\mathbf{K}^{-1}\mathbf{u}$, $\mu\equiv\mathbf{u}^*\mathbf{y}$

Explicit left preconditioning

$$\mathbf{K}^\dagger\,(\mathbf{I}-\mathbf{u}\,\mathbf{u}^*)\,(\mathbf{A}-\theta\,\mathbf{I})\,(\mathbf{I}-\mathbf{u}\,\mathbf{u}^*)\,\mathbf{t}=-\mathbf{K}^\dagger\mathbf{r}.$$

Note that $\mathbf{K}^\dagger\mathbf{s}\perp\mathbf{u}$. In part., $\mathbf{K}^\dagger\mathbf{r}\perp\mathbf{u}$.

The action of a precond. operator on an $\mathbf{s}$

If $\mathbf{s}\perp\mathbf{u}$

- $\mathbf{s}\leftarrow(\mathbf{A}-\theta\,\mathbf{I})\mathbf{s}$
- Solve $\mathbf{K}\mathbf{s}'=\mathbf{s}$
- $\mathbf{s}\leftarrow\mathbf{s}'-\mathbf{y}(\mu^{-1}(\mathbf{u}^*\mathbf{s}'))$

then $\mathbf{s}\perp\mathbf{u}$

Only one (skew) projection!

J-D

- is an accelerated inexact Newton scheme
  — quadratic convergence —
- relies on orthogonal projections
- allows any precond. it. method
  to solve the linear (Jacobian) system
- robust
- no matrix inversion, no exact solves
- also for interior eigenvalues

J-D is <u>not</u>

- a Krylov subspace method,
- Davidson with Jacobi-preconditioning,
- a method that produces eigenvectors for
  a preconditioned eigenproblem

What to use?

- For small problems: QR-algorithm

- If matrices can be inverted
  (systems can be solved exactly):

  Shift-&-Invert Arnoldi with implicit restart

  [Sorensen '92]

  Rational Krylov Subspace with implicit restart

  [Ruhe '84, '93]

  If $\mathbf{A}$ special structure $(\mathbf{A}^* = \mathbf{A}, \ldots)$ ...

- If matrix inversion is not feasible: JD

Generalized eigenproblems

$$\mathbf{Ax} = \lambda \mathbf{Bx}$$

Example: MHD

- Interaction plasma and magnetic field
- Fluid eqns + Maxwell
- **Applications**:
  stability tokamak plasmas, coronal loops, solar flares

Linear MHD

$$
\begin{aligned}
\frac{\partial \rho}{\partial t} &= -\nabla \cdot (\rho \mathbf{V}) \\
\rho \frac{\partial \mathbf{V}}{\partial t} &= -\rho(\mathbf{V} \cdot \nabla)\mathbf{V} - \nabla p \\
&\qquad\qquad +(\nabla \times \mathbf{B}) \times \mathbf{B} \\
\frac{\partial p}{\partial t} &= -(\mathbf{V} \cdot \nabla)p - \gamma p \nabla \cdot \mathbf{V} \\
&\qquad\qquad +(\gamma - 1)\eta(\nabla \times \mathbf{B})^2 \\
\frac{\partial \mathbf{B}}{\partial t} &= \nabla \times (\mathbf{V} \times \mathbf{B}) - \nabla \times (\eta \nabla \times \mathbf{B}) \\
\text{with} &\quad \nabla \cdot \mathbf{B} = 0
\end{aligned}
$$

- Linearise around an equilibrium

- Assume solutions of the form $e^{\lambda t}$: generalized eigenproblem

Linear MHD (2)

- Block tridiagonal matrices:

  100 – 2 000 blocks of size 80 – 320

  $\Rightarrow$ matrices of order 8 000 – 640 000

For smallest systems:

- Shift-and-Invert IRA

(Sorensen)

- Shift-and-Invert Lanczos

(Cullum et al)

- Jacobi-Davidson, no inversion

(Booten, Fokkema, Sleijpen, van der Vorst)

Problem and Approach

Problem: find one or more eigenvalues of $\mathbf{A}\mathbf{x} = \lambda\,\mathbf{B}\mathbf{x}$

$\mathbf{A}$ and $\mathbf{B}$ unsymmetric, large

CONVENTIONAL APPROACH:
Shift-and-Invert

$$(\mathbf{A} - \tau\mathbf{B})^{-1}\mathbf{B}\mathbf{x} = \tilde{\lambda}\,\mathbf{x}$$

$$\tilde{\lambda} = \frac{1}{\lambda - \tau}, \qquad \lambda = \tau + \frac{1}{\tilde{\lambda}}$$

Apply Arnoldi, convergence to eigenvalues closest to τ

Problem: $(\mathbf{A} - \tau\mathbf{B})^{-1}$

QZ Method

Compute orthogonal $\mathbf{Q}$ and $\mathbf{Z}$:

$$\mathbf{AQ} = \mathbf{ZS}, \qquad \mathbf{BQ} = \mathbf{ZT}$$

such that $\mathbf{S}$ and $\mathbf{T}$ are upper triangular.

Columns $\mathbf{q}_j$:
Generalized Schur vectors

Eigenvalue: $\mathbf{S}_{jj}/\mathbf{T}_{jj}$

Iterative approach based upon
Jacobi-Davidson method

$$(\mathbf{A} - \lambda \mathbf{B})\mathbf{x} = \mathbf{0}$$

Petrov-Galerkin approach

Find $\mathbf{u} \in \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ such that

$$\mathbf{r} = (\mathbf{A} - \theta \mathbf{B})\mathbf{u} \perp \{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$$

With $\mathbf{V}_k \equiv [\, \mathbf{v}_1 \,|\, \mathbf{v}_2 \,|\, \ldots \,|\, \mathbf{v}_k \,]$,
$\mathbf{W}_k \equiv [\, \mathbf{w}_1 \,|\, \mathbf{w}_2 \,|\, \ldots \,|\, \mathbf{w}_k \,]$,
and $\mathbf{u} = \mathbf{V}_k\, y$:

$$(\mathbf{W}_k^* \mathbf{A} \mathbf{V}_k - \theta \mathbf{W}_k^* \mathbf{B} \mathbf{V}_k)\, y = 0$$

Apply standard QZ (LAPACK) for projected matrices

$\mathbf{W}_k^* \mathbf{A} \mathbf{V}_k$ and $\mathbf{W}_k^* \mathbf{B} \mathbf{V}_k$

$\mathbf{AQ} = \mathbf{ZS} \quad \& \quad \mathbf{BQ} = \mathbf{ZT}$

Find orthogonal U_L and U_R:

$$U_L^*(\mathbf{W}_k^* \mathbf{A} \mathbf{V}_k)U_R = S_A$$

$$U_L^*(\mathbf{W}_k^* \mathbf{B} \mathbf{V}_k)U_R = S_B$$

S_A and S_B are upper-triangular
ordered such that leading elements
define eigenvalue closest to τ

Petrov pair $(\theta, \mathbf{u})$ approx. eigenpair $(\lambda, \mathbf{x})$;

$$\theta = S_B(1,1)/S_A(1,1), \qquad \mathbf{u} = \mathbf{V}_k\, U_R(:,1)$$

At convergence:

- $\mathbf{x}$ is the first column of $\mathbf{Q}$!!
- is $\mathbf{W}_k\, U_L(:,1)$ the first column $\mathbf{z}_1$ of $\mathbf{Z}$??

$\mathbf{z}_1 \in \text{span}(\mathbf{A}\mathbf{q}_1, \mathbf{B}\mathbf{q}_1)$
Therefore, $\mathbf{W}_k$ taken as $\beta\, \mathbf{A}\mathbf{V}_k + \alpha\, \mathbf{B}\mathbf{V}_k$
(we favor: $\beta = 1$, $\alpha = -\tau$)

JACOBI-DAVIDSON

How to expand $\mathbf{V}_k$ (and $\mathbf{W}_k$)?

Compute *residual vector*

$$\mathbf{r} \equiv (\mathbf{A} - \theta\mathbf{B})\mathbf{u}$$

Also compute auxiliary vector:

$$\gamma\,\mathbf{z} \equiv (\beta\mathbf{A} + \alpha\mathbf{B})\mathbf{u} \qquad (\mathbf{z} = \mathbf{W}_k U_L(:,1))$$

Find correction $\mathbf{t}$ for $\mathbf{u}$ with

$$\mathbf{t}^*\mathbf{u} = 0$$

and

$$(\mathbf{I} - \mathbf{z}\,\mathbf{z}^*)(\mathbf{A} - \theta\mathbf{B})(\mathbf{I} - \mathbf{u}\,\mathbf{u}^*)\mathbf{t} = -\mathbf{r}$$

Correction equation approximately solved
with few steps of GMRES

$\mathbf{t}$ expansion vector $\mathbf{V}_k$: $\quad \mathbf{v}_{k+1} \equiv \mathrm{modGS}_{\mathbf{V}_k}(\mathbf{t})$
$\mathbf{s} \equiv (\beta\mathbf{A}+\alpha\mathbf{B})\mathbf{t}$ exp. v. $\mathbf{W}_k$: $\mathbf{w}_{k+1} \equiv \mathrm{modGS}_{\mathbf{V}_k}(\mathbf{s})$

How to continu?

Expansion of Schur Form:

$$\mathbf{A}\left[\mathbf{Q}_\ell, \mathbf{q}_{\ell+1}\right] = \left[\mathbf{Z}_\ell, \mathbf{z}_{\ell+1}\right]\begin{bmatrix} S_\ell & s \\ & \alpha \end{bmatrix}$$

and

$$\mathbf{B}\left[\mathbf{Q}_\ell, \mathbf{q}_{\ell+1}\right] = \left[\mathbf{Z}_\ell, \mathbf{z}_{\ell+1}\right]\begin{bmatrix} T_\ell & t \\ & \beta \end{bmatrix}$$

with $\quad \mathbf{q}_{\ell+1}^*\mathbf{Q}_\ell = 0, \quad \mathbf{z}_{\ell+1}^*\mathbf{Z}_\ell = 0.$

This leads to

$$(\mathbf{I} - \mathbf{Z}_\ell\mathbf{Z}_\ell^*)(\beta\mathbf{A} - \alpha\mathbf{B})(\mathbf{I} - \mathbf{Q}_\ell\mathbf{Q}_\ell^*)\mathbf{q}_{\ell+1} = \mathbf{0}$$

$(\lambda_{\ell+1} \equiv \alpha/\beta, \mathbf{q}_{\ell+1})$ eigenpair of:

$$(\mathbf{I} - \mathbf{Z}_\ell\mathbf{Z}_\ell^*)(\mathbf{A} - \lambda_{\ell+1}\mathbf{B})(\mathbf{I} - \mathbf{Q}_\ell\mathbf{Q}_\ell^*)$$

Apply "JD - algorithm with Schur form" for this matrix pair $\Rightarrow \mathbf{q}_{\ell+1}$ and $\mathbf{z}_{\ell+1}$.

Projected matrices

The projected matrices in JD satisfy:

$$\mathbf{W}_k^*(\mathbf{I}-\mathbf{Z}_\ell\mathbf{Z}_\ell^*)\mathbf{A}(\mathbf{I}-\mathbf{Q}_\ell\mathbf{Q}_\ell^*)\mathbf{V}_k = \mathbf{W}_k^*\mathbf{A}\mathbf{V}_k$$

and

$$\mathbf{W}_k^*(\mathbf{I}-\mathbf{Z}_\ell\mathbf{Z}_\ell^*)\mathbf{B}(\mathbf{I}-\mathbf{Q}_\ell\mathbf{Q}_\ell^*)\mathbf{V}_k = \mathbf{W}_k^*\mathbf{B}\mathbf{V}_k$$

Correction Equation

With $\widetilde{\mathbf{Q}} = [\,\mathbf{Q}_\ell\,|\,\mathbf{u}\,]$ & $\widetilde{\mathbf{Z}} = [\,\mathbf{Z}_\ell\,|\,\mathbf{z}\,]$

solve $\mathbf{t} \perp \widetilde{\mathbf{Q}}$ (appr) from

$$(\mathbf{I}-\widetilde{\mathbf{Z}}\,\widetilde{\mathbf{Z}}^*)(\mathbf{A}-\theta\mathbf{B})(\mathbf{I}-\widetilde{\mathbf{Q}}\,\widetilde{\mathbf{Q}}^*)\mathbf{t} = -\mathbf{r}$$

with

$$\mathbf{r} \equiv (\mathbf{I}-\mathbf{Z}_\ell\,\mathbf{Z}_\ell^*)(\mathbf{A}\mathbf{u}-\theta\mathbf{B}\mathbf{u})$$

Implicit Restart

For projected matrices we compute the generalized Schur form

$$\mathbf{W}_k^* \mathbf{A} \mathbf{V}_k \; U_R = U_L \, S_A$$

and

$$\mathbf{W}_k^* \mathbf{B} \mathbf{V}_k \; U_R = U_L \, S_B$$

S_A and S_B are ordered so that

$$|S_B(j,j)/S_A(j,j) - \tau|$$

form a nondecreasing row

The first few columns of reordered $\mathbf{V}_k$ represent best eigenvector approx's

$\Rightarrow$ can be used for selection & restart

$$(\mathbf{I}-\mathbf{z}\,\mathbf{z}^*)(\mathbf{A}-\theta\mathbf{B})(\mathbf{I}-\mathbf{u}\,\mathbf{u}^*)\mathbf{t}=-\mathbf{r}$$

$$\mathbf{K}\approx\mathbf{A}-\theta\,\mathbf{B},\qquad(\mathbf{I}-\mathbf{z}\,\mathbf{z}^*)\mathbf{K}(\mathbf{I}-\mathbf{u}\,\mathbf{u}^*)\mathbf{c}=\mathbf{s}.$$

$\mathbf{K}^\dagger\equiv(\mathbf{I}-\mathbf{y}\mu^{-1}\mathbf{u}^*)\mathbf{K}^{-1}$ with $\mathbf{y}\equiv\mathbf{K}^{-1}\mathbf{z}$, $\mu\equiv\mathbf{u}^*\mathbf{y}$

Explicit left preconditioning

$$\mathbf{K}^\dagger(\mathbf{I}-\mathbf{z}\,\mathbf{z}^*)(\mathbf{A}-\theta\mathbf{B})(\mathbf{I}-\mathbf{u}\,\mathbf{u}^*)\mathbf{t}=-\mathbf{K}^\dagger\mathbf{r}.$$

Note that $\mathbf{K}^\dagger\mathbf{s}\perp\mathbf{z}$. In part., $\mathbf{K}^\dagger\mathbf{r}\perp\mathbf{z}$.

The action of a precond. operator on an $\mathbf{s}$

If $\mathbf{s}\perp\mathbf{z}$

- $\mathbf{s}\leftarrow(\mathbf{A}-\theta\,\mathbf{B})\mathbf{s}$
- Solve $\mathbf{K}\mathbf{s}'=\mathbf{s}$
- $\mathbf{s}\leftarrow\mathbf{s}'-\mathbf{y}(\mu^{-1}(\mathbf{u}^*\mathbf{s}'))$

then $\mathbf{s}\perp\mathbf{z}$

Only one (skew) projection!

$$\mathbf{W}_k = \alpha\,\mathbf{A}\,\mathbf{V}_k + \beta\,\mathbf{B}\,\mathbf{V}_k$$

Choosing the mix of $\mathbf{A}$ and $\mathbf{B}$, I

First approach:
aim for stable approximate "left Schur vector"

Inspiration:

$\mathbf{B} = \mathbf{I}$: choose $\quad \mathbf{w} = \mathbf{u} = 0\,\mathbf{A}\,\mathbf{u} + 1\,\mathbf{B}\,\mathbf{u}.$

$\mathbf{A} = \mathbf{I}$: choose $\quad \mathbf{w} = \mathbf{u} = 1\,\mathbf{A}\,\mathbf{u} + 0\,\mathbf{B}\,\mathbf{u}.$

$\mathbf{w} = 1\,\mathbf{A}\,\mathbf{u} + 0\,\mathbf{B}\,\mathbf{u}$ is "instable" choice

if $\mathbf{B} = \mathbf{I}$ and $\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \approx \mathbf{0}$

Select α, β such that
— $|\alpha|^2 + \beta^2 = 1$, $\beta \geq 0$, and
— $\|\alpha\,\mathbf{A}\,\mathbf{u} + \beta\,\mathbf{B}\,\mathbf{u}\|$ approximately maximal:

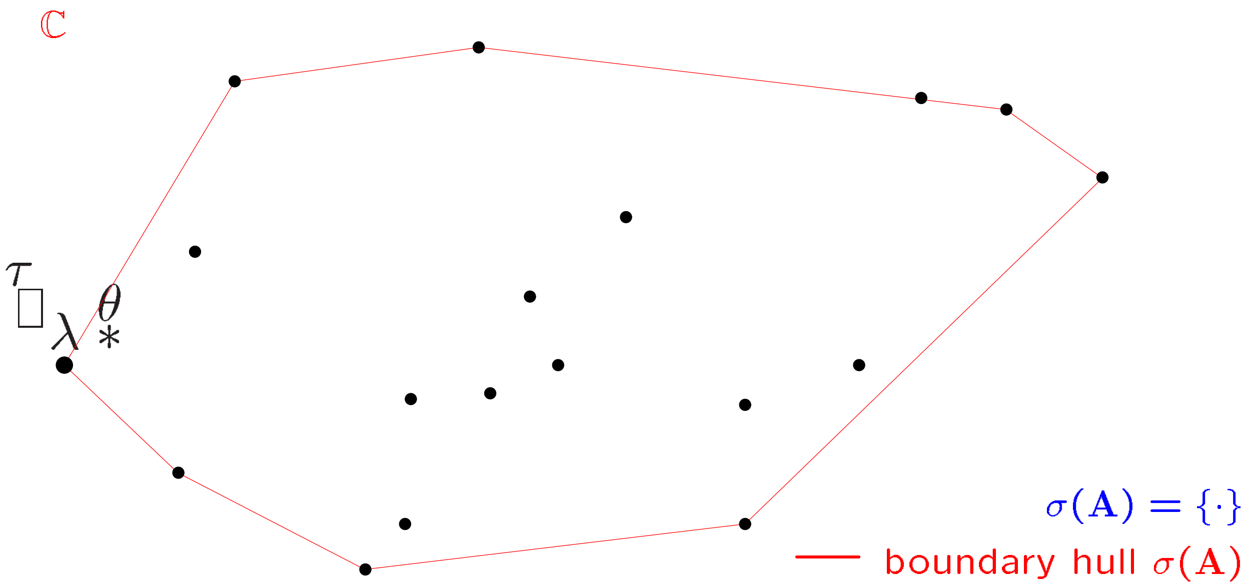
$$\mathbf{C} = \frac{\bar{\tau}}{\sqrt{1+|\tau|^2}}\mathbf{A} + \frac{1}{\sqrt{1+|\tau|^2}}\mathbf{B}$$

Selecting Ritz pairs

Suppose $\mathbf{A}$ normal & $\mathbf{B} = \mathbf{I}$.

$$\theta = \sum_i |\gamma_i|^2 \lambda_i \quad \text{with} \quad \mathbf{u} = \sum_i \gamma_i \mathbf{x}_i$$

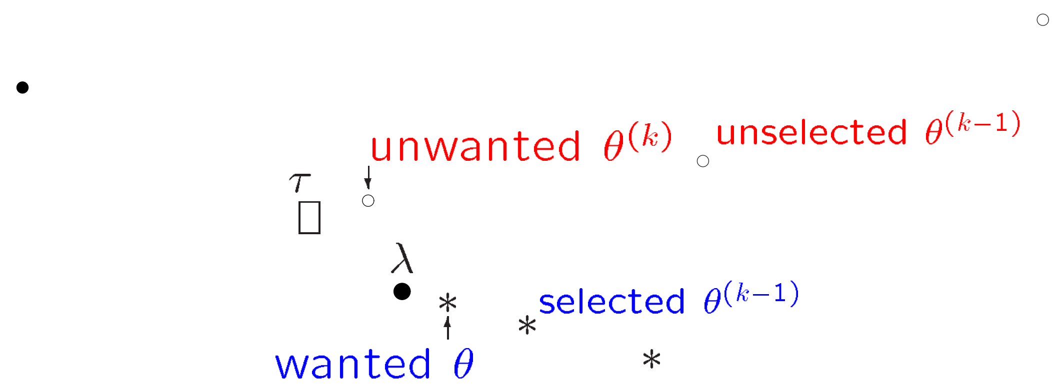
Suppose τ outside hull spectrum:



If $\theta \approx \tau$ then wanted component large & selecting extremal Ritz value is safe.

What if τ is in the interior of the hull of the spectrum?

Misselection if $\theta \approx \lambda$



Avoiding misselection

Computing all residual norms is expensive

- ◦ Shift target temporarily to
  old selected Ritz value $\theta^{(k-1)}$

- Use harmonic Ritz values

Choosing the mix of $\mathbf{A}$ and $\mathbf{B}$, II

Second approach: avoid missection

$$(\mathbf{A} - \lambda\,\mathbf{I})\mathbf{x} = \mathbf{0} \qquad \Leftrightarrow \qquad (\mathbf{A} - \tau\,\mathbf{I})^{-1}\mathbf{x} = \frac{1}{\lambda - \tau}\mathbf{x} :$$

if $\lambda \approx \tau$ then $1/(\lambda - \tau)$ extremal for $(\mathbf{A} - \tau\,\mathbf{I})^{-1}$.

Avoid inversion: use Harmonic Ritz values

Galerkin for $(\mathbf{A} - \tau\,\mathbf{I})^{-1}$, with

search space $=$ test space $=$ span$((\mathbf{A} - \tau\,\mathbf{I})\mathbf{V})$,

is equivalently to Petrov for $\mathbf{A}$, with

search space $\mathbf{V}$ & test space $\mathbf{W} \equiv (\mathbf{A} - \tau\,\mathbf{I})\mathbf{V}$.

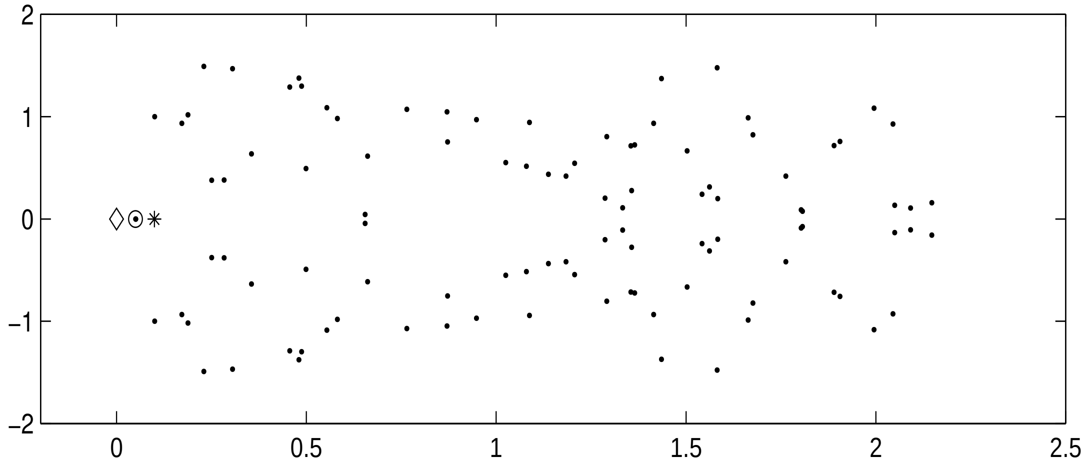
Harmonic Ritz value $\equiv$ this Petrov value

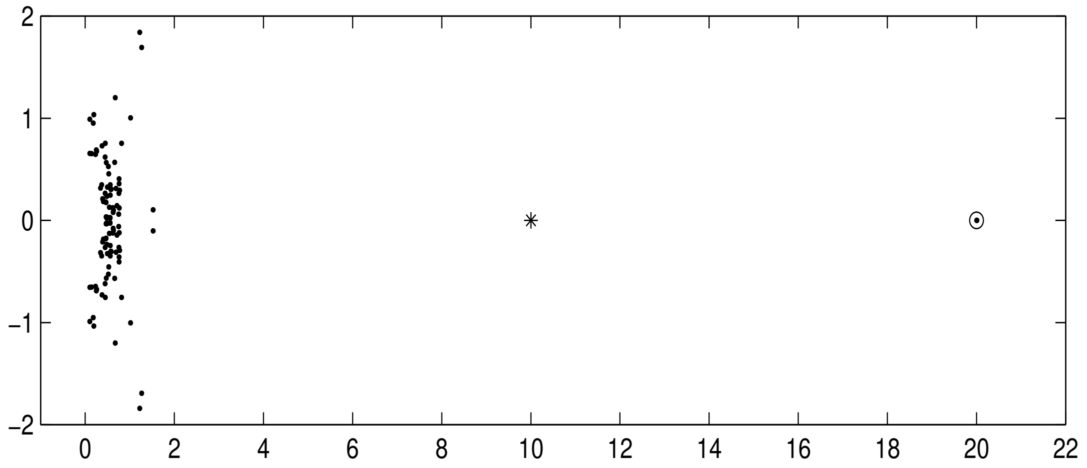
For generalized eigenproblems, take

$$\text{span}(\mathbf{W}) = \text{span}((\mathbf{A} - \tau\,\mathbf{B})\mathbf{V}),$$

$$\mathbf{C} = \frac{1}{\sqrt{1 + |\tau|^2}}\mathbf{A} - \frac{\tau}{\sqrt{1 + |\tau|^2}}\mathbf{B}.$$

Harm. R. also for extremal eigenv.





A normal

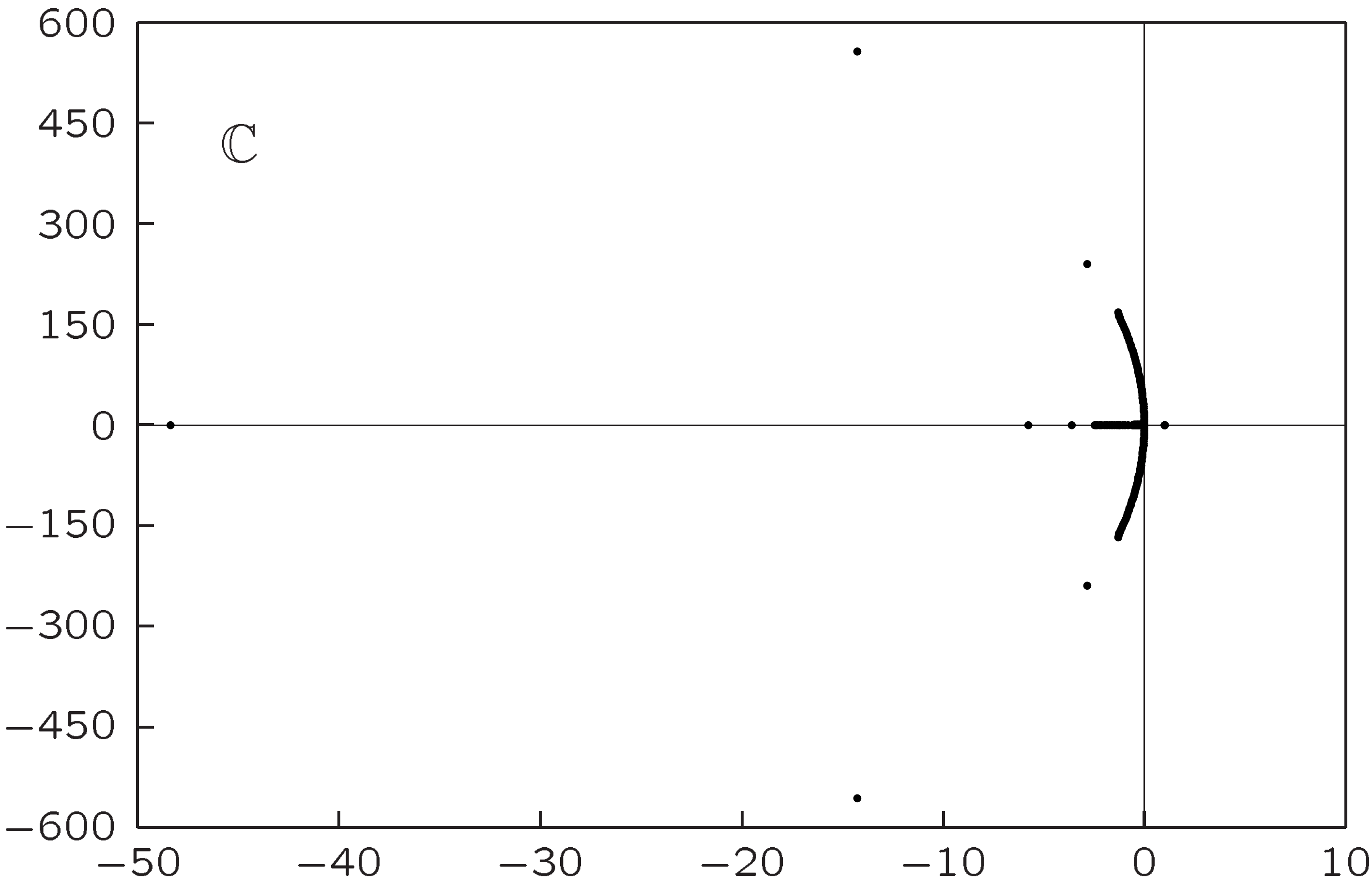
Top figure θ could be combination $2\epsilon \pm i$

$$\sigma(\mathbf{A}) = \{\cdot\} \qquad \tau = \diamond \qquad \lambda = \odot \qquad \theta = *$$

Bottom figure

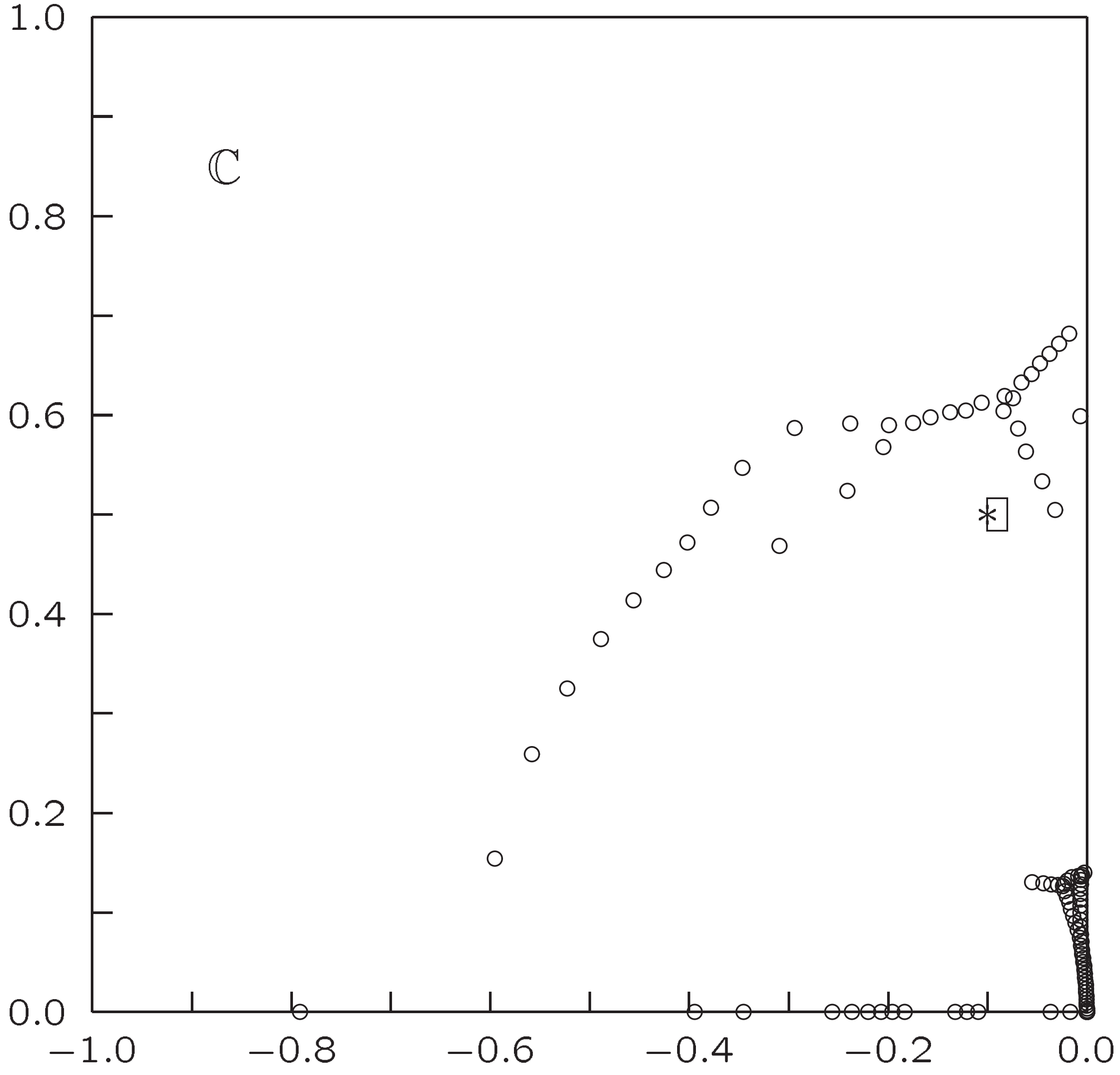
$$\sigma((\mathbf{A}-\tau\mathbf{I})^{-1}) = \{\cdot\} \quad \frac{1}{\lambda-\tau} = \odot \quad \frac{1}{\theta-\tau_{55}} = *$$

Spectrum MHD test problem



Entire spectrum MHD test problem ($n = 416$).
Correct eigenvalues (.) are shown.

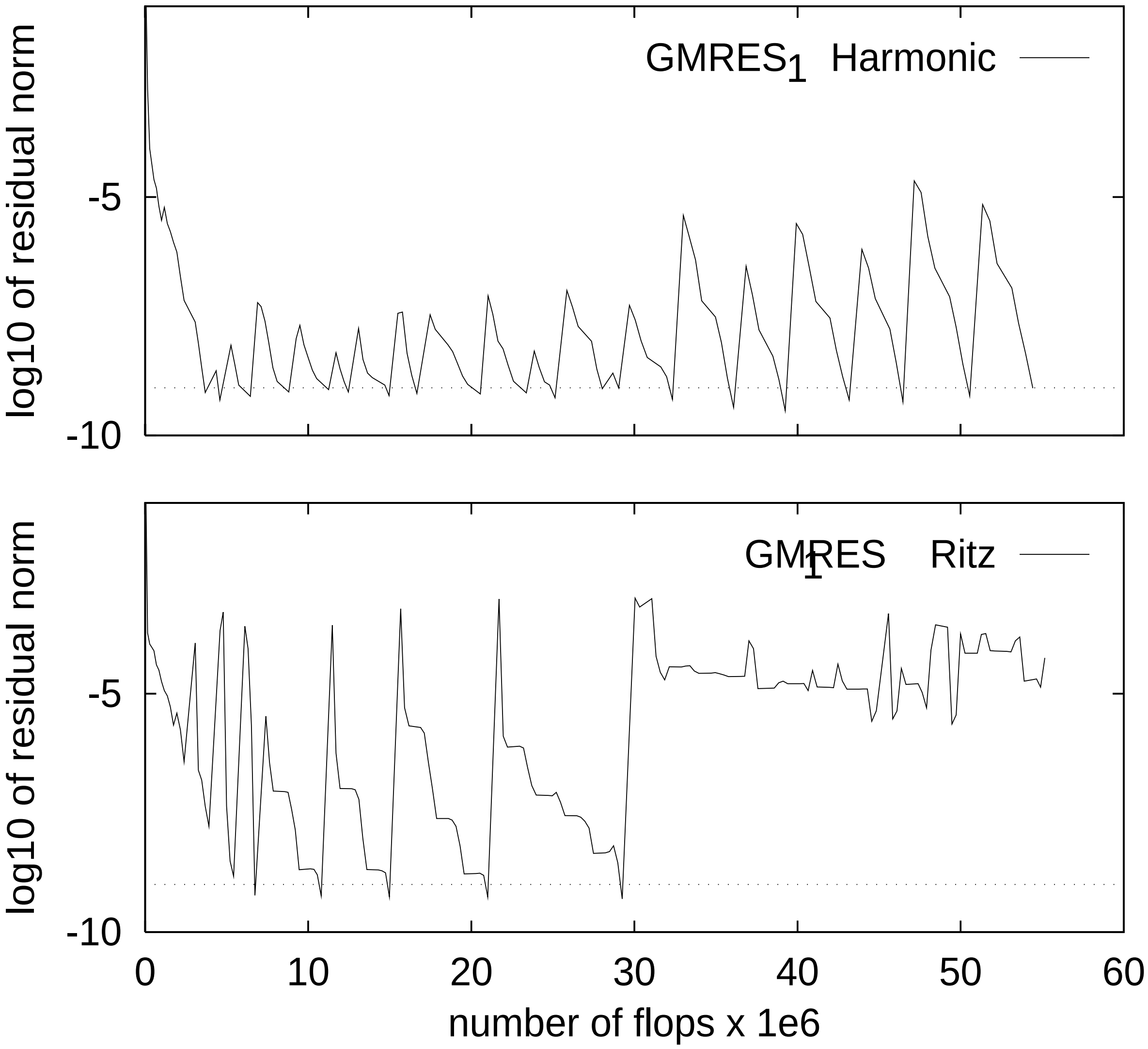
Alfvén spectrum



MHD test problem ($n = 416$).
Correct eigenvalues ($\circ$), target ($*$).

MHD416

Searching for the Alfvén branch



ILU(2) for $\mathbf{A} - \tau\mathbf{B}$ is used as preconditioner.

max dim$(\mathbf{V})$=15
min dim$(\mathbf{V})$=10

Misselection if $\theta \not\approx \lambda$

θ may be $\frac{1}{2}(\lambda_2 + \lambda_3)$

λ_1 λ_2 $*$ λ_3 $*$ $\cdots$ λ_n

θ may be
$(1 - \epsilon)\lambda_1 + \epsilon\lambda_n$

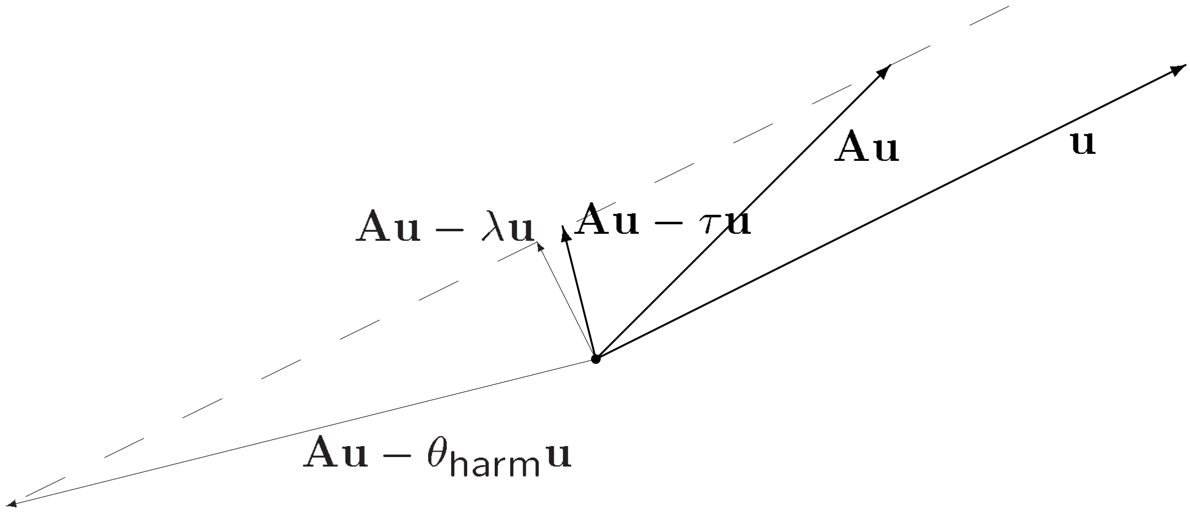
Avoiding misselection

- Scan "all" angles between new Ritz vectors and the selected old Ritz vector:
  $$(\mathbf{V}_k\, y_k)^*(\mathbf{V}_{k-1}\, y_{k-1}) = y_k(1:k-1)^* y_{k-1}$$

no large vector operations

Using accurate initial vectors

The effect of testing with harmonic Ritz values when $(\tau, \mathbf{u})$ approximates an eigenpair $(\lambda, \mathbf{x})$ moderately well.



The residual $\mathbf{r} = \mathbf{Au} - \theta_{\text{harm}}\mathbf{u}$ is orthogonal with respect to the test subspace spanned by $\mathbf{Au} - \tau\mathbf{u}$.

Note that $|\lambda - \tau| \ll |\lambda - \theta|$.

Take • $\mathbf{W} = (\mathbf{A} - \tau\mathbf{I})\mathbf{V}$ to test,

• $\mathbf{u}$ as appr. eigenvector,

• $\theta = \text{minarg}_\eta \|\mathbf{Au} - \eta\mathbf{u}\|$ not $\theta = \theta_{\text{harm}}$,

• $\mathbf{z} \in \text{span}(\mathbf{Au}, \mathbf{u})$ s.t. $\mathbf{z} \perp \mathbf{r} = \mathbf{Au} - \theta\mathbf{u}$

$$(\mathbf{I}-\mathbf{u}_k\,\mathbf{u}_k^*)\,(\mathbf{A}-\theta_k\,\mathbf{I})\,(\mathbf{I}-\mathbf{u}_k\,\mathbf{u}_k^*)\,\mathbf{t}_{k+1}=-\mathbf{r}_k.$$

Stagnation in exact JD

Stagnation if $\mathbf{t}_{k+1}\in \text{span}(\mathbf{V}_k\,)$

- $\mathbf{t}_{k+1}=\mathbf{0} \quad \Rightarrow \quad \mathbf{r}_k=\mathbf{0}$ & $(\theta,\mathbf{u}_k)=(\lambda,\mathbf{x})$
(Lucky stagnation)

- $\mathbf{t}_{k+1}\neq\mathbf{0} \quad \Rightarrow \quad \theta^{(k)}$ is an defective eigenvalue of $\mathbf{V}_k^*\mathbf{A}\mathbf{V}_k$ (non-trivial Jordan box)

Example. $\begin{bmatrix} 0 & 1 & \alpha \\ 0 & 0 & \beta \\ \epsilon & -\epsilon & \gamma \end{bmatrix}$

Initial vector $\mathbf{v}_1=e_1$. Then $\mathbf{V}_2=[e_1|e_2]=\mathbf{V}_3$

$$(\mathbf{I} - \mathbf{u}_k \, \mathbf{u}_k^*) \, \mathbf{K} \, (\mathbf{I} - \mathbf{u}_k \, \mathbf{u}_k^*) \, \mathbf{t}_{k+1} = -\mathbf{r}_k$$

Near stagnation in inexact JD

$\mathbf{K}$ favors its own extremal eigenvectors

$\mathbf{u}_k$ may have only a small components in these directions $\Rightarrow$ $\mathbf{u}_{k+1} = \mathbf{u}_k$ except for these components & $\mathbf{t}_{k+1} \approx \mathbf{t}_{k+2}$

$$(\mathbf{I} - \mathbf{u}_k\,\mathbf{u}_k^*)\,\mathbf{K}\,(\mathbf{I} - \mathbf{u}_k\,\mathbf{u}_k^*)\,\mathbf{t}_{k+1} = -\mathbf{r}_k$$

Curing near stagnation

Practical observation. Stagnation if

$$\tan\angle(\mathbf{t}_{k+1}, \mathbf{V}_k) = \frac{|\zeta|}{\|b\|} < 0.06$$

where

$$\mathbf{t}_{k+1} = \mathbf{V}_k\,b + \mathbf{v}_{k+1}\,\zeta$$

(no additional large vector operations)

Cure.
Solve (appr) $\quad \mathbf{t} = \mathbf{t}_{k+2} \perp \mathbf{V} = \mathbf{V}_{k+1} \quad$ from

$$(\mathbf{I} - \mathbf{V}\,\mathbf{V}^*)(\mathbf{A} - \theta\mathbf{I})(\mathbf{I} - \mathbf{V}\,\mathbf{V}^*)\mathbf{t} = -\mathbf{r}_{k+1}$$

[Genseberger, S.'98]

$$(\mathbf{I} - \mathbf{V}\,\mathbf{V}^*)(\mathbf{A} - \theta\mathbf{I})(\mathbf{I} - \mathbf{V}\,\mathbf{V}^*)\mathbf{t} = -\mathbf{r}$$

Advantages

- May help to overcome stagnation!! yes
- ∘ Reduces number of GMRES steps?? no
- ∘ Improves stability $(\mathbf{t} \perp \mathbf{V})$?? no

Disadvantages

- May spoil fast convergence!! yes

  $\dim(\mathbf{V})=1$ (re)start & exact $\Rightarrow$ JD = JDV

  $\dim(\mathbf{V})>1$ (re)start or inexact $\Rightarrow$ JD $\neq$ JDV

  (JDV cor. eq. is no Newton correction)!

- More costly per step. yes

Conclusion.

JDV only to be used to overcome stagnation

With $\quad \widetilde{\mathbf{Q}} = [\mathbf{Q}_\ell|\mathbf{u}] \quad$ and $\quad \tilde{\mathbf{Z}} = [\mathbf{Z}_\ell|\mathbf{z}]$

$$(\mathbf{I} - \tilde{\mathbf{Z}}\,\tilde{\mathbf{Z}}^*)(\mathbf{A} - \theta\mathbf{B})(\mathbf{I} - \widetilde{\mathbf{Q}}\,\widetilde{\mathbf{Q}}^*)\mathbf{t} = -\mathbf{r}$$

where

$$\mathbf{r} = (\mathbf{I} - \mathbf{Z}_\ell\,\mathbf{Z}_\ell^*)(\mathbf{A}\mathbf{u} - \theta\mathbf{B}\mathbf{u})$$

Near stagnation in JDQZ

Stagnation if

$$\tan\angle(\mathbf{t}_{k+1}, \mathbf{V}_k) = \frac{|\zeta|}{\|b\|} < 0.06$$

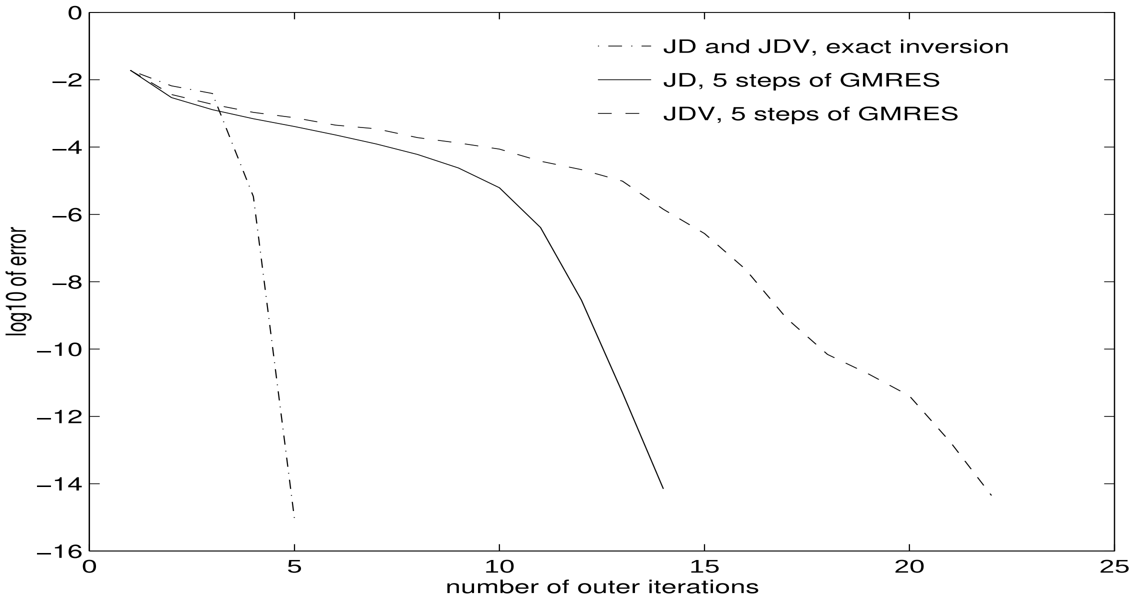
where $\quad \mathbf{t}_{k+1} = \mathbf{V}_k\,b + \mathbf{v}_{k+1}\,\zeta$

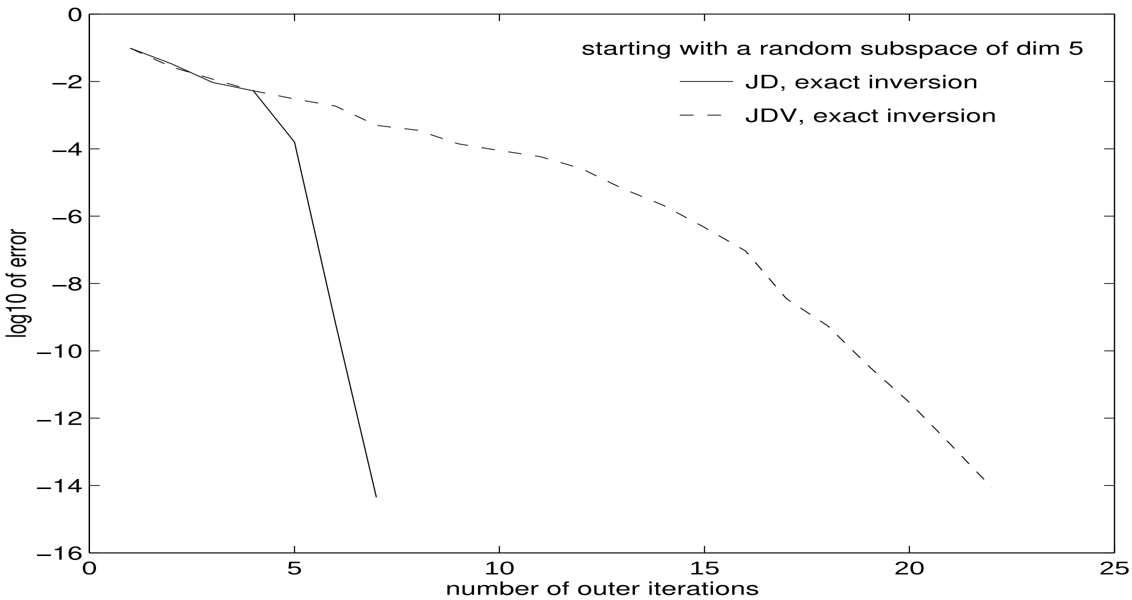
<u>**Cure.**</u> Use, for (only) the next step

$$\widetilde{\mathbf{Q}} = [\mathbf{Q}_\ell|\mathbf{V}] \quad \text{and} \quad \tilde{\mathbf{Z}} = [\mathbf{Z}_\ell|\mathbf{W}]$$

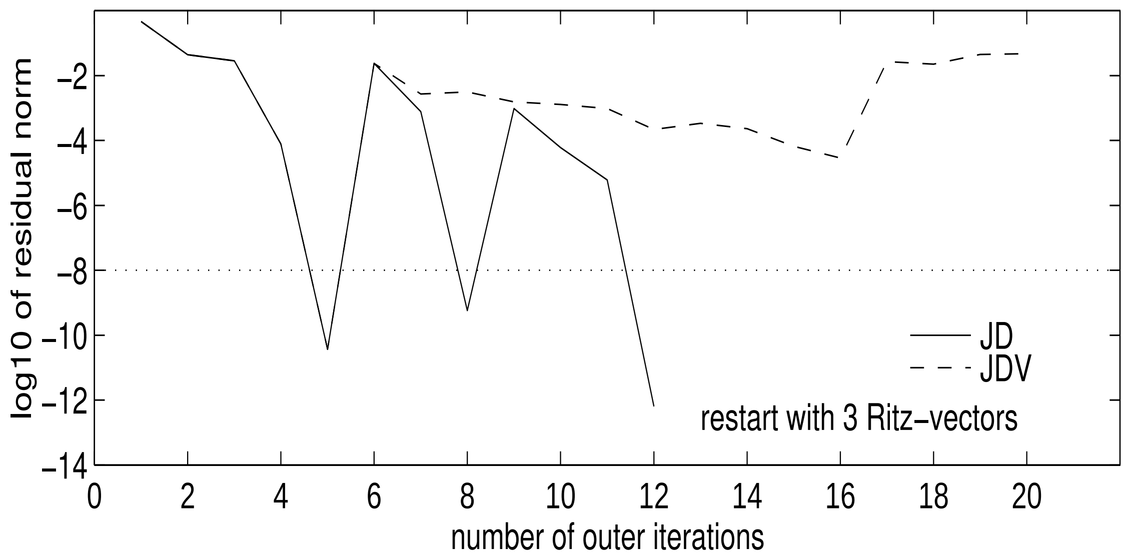
$$(\mathbf{I} - \tilde{\mathbf{Z}}\,\tilde{\mathbf{Z}}^*)(\mathbf{A} - \theta\mathbf{B})(\mathbf{I} - \widetilde{\mathbf{Q}}\,\widetilde{\mathbf{Q}}^*)\mathbf{t} = -\mathbf{r}$$

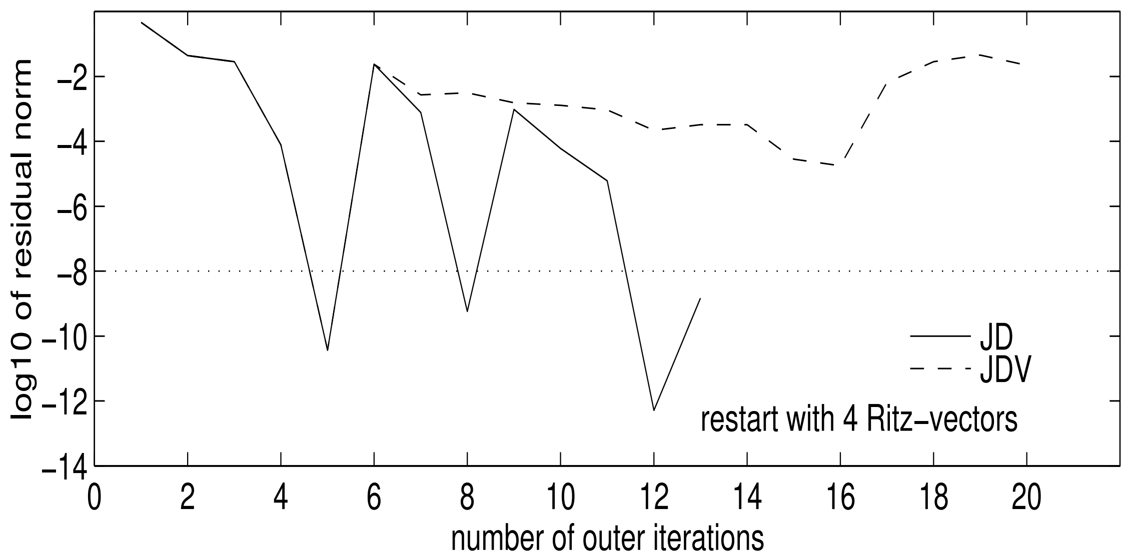
JD versus JDV





A 100-dim tridiagonal





SHERMAN1 shifted by -2.5

GMRES_{200}

Generalized eigenproblems

$$\mathbf{A}\,\mathbf{x} = \lambda\,\mathbf{B}\,\mathbf{x}$$

[Booten, Fokkema **MHD**]

$$\mathbf{P}(\lambda)\mathbf{x} := \mathbf{A}_0\,\mathbf{x} + \lambda\,\mathbf{A}_1\,\mathbf{x} + \ldots + \lambda^{\ell}\,\mathbf{A}_{\ell}\,\mathbf{x} = 0$$

[van Gijzen, S. vdV **acoustics, con. mech.]**

- $$\mathbf{V}^*\mathbf{P}(\theta)\mathbf{V}y = 0$$

$$\mathbf{u} = \mathbf{V}y, \quad \mathbf{w} = \mathbf{P}'(\theta)\mathbf{u}, \quad \mathbf{r} = \mathbf{P}(\theta)\mathbf{u}$$

- The correction equation:

$$\left(\mathbf{I} - \frac{\mathbf{w}\,\mathbf{u}^*}{\mathbf{u}^*\mathbf{w}}\right)\mathbf{P}(\theta)\left(\mathbf{I} - \frac{\mathbf{u}\,\mathbf{u}^*}{\mathbf{u}^*\mathbf{u}}\right)\mathbf{t} = -\mathbf{r}$$

- The space: $\mathbf{V} = \mathrm{modGS}\,([\,\mathbf{V}\,|\,\mathbf{t}\,])$

What if $\mathbf{K}\tilde{\mathbf{x}} = \mathbf{0}$ & $\tilde{\mathbf{x}} \approx \mathbf{x}$?

[joint work with F. Wubs, Groningen U]

Suppose

$\mathbf{K} : \tilde{\mathbf{x}}^\perp \to \tilde{\mathbf{y}}^\perp$ well-cond., Cond. nmb $\mathcal{C}$

with $\tilde{\mathbf{y}}$ such that $\mathbf{y}^*\mathbf{K} = \mathbf{0}$.

Use equivalent aumented system:

$$\begin{bmatrix} \mathbf{K} & \mathbf{z} \\ \mathbf{u}^* & 0 \end{bmatrix} \begin{bmatrix} \mathbf{t} \\ \alpha \end{bmatrix} = \begin{bmatrix} \mathbf{s} \\ 0 \end{bmatrix}$$

Condition augmented system:

$$\sim \quad \mathcal{C} \cdot \mathrm{Cond}_K(0) \cdot 1/\cos(\mathbf{u}, \tilde{\mathbf{x}}) \cdot \underbrace{1/\cos(\mathbf{z}, \tilde{\mathbf{y}})}_{\approx\ \mathrm{Cond}_{\mathbf{B}^{-1}\mathbf{A}}(\lambda)}$$

MRILU

Reorder & Partition

$$\mathbf{A} - \tau\mathbf{B} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{E}_u \\ \mathbf{E}_\ell & \mathbf{A}_2 \end{bmatrix}$$

with • $\mathbf{A}_1$ well conditioned

• $\mathbf{A}_2$ of low dimension
(with ill. cond. Schur complement)

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_1 & \mathbf{E}_u \\ \mathbf{E}_\ell & \mathbf{A}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{E}_\ell\mathbf{K}_1^{-1} & \widetilde{\mathbf{A}}_2 \end{bmatrix} \begin{bmatrix} \mathbf{K}_1 & \mathbf{E}_u \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$$

with

$\mathbf{K}_1$ block ILU $\mathbf{A}_1$

$\widetilde{\mathbf{A}}_2$ Schur complement wrt $\mathbf{K}_1$:

$$\widetilde{\mathbf{A}}_2 = \mathbf{A}_2 - \mathbf{E}_\ell\mathbf{K}_1^{-1}\mathbf{E}_u$$

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_1 & \mathbf{E}_u \\ \mathbf{E}_\ell & \mathbf{A}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{E}_\ell \mathbf{K}_1^{-1} & \widetilde{\mathbf{A}}_2 \end{bmatrix} \begin{bmatrix} \mathbf{K}_1 & \mathbf{E}_u \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{K} & \mathbf{z} \\ \mathbf{u}^* & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{K}_1 & \mathbf{E}_u & \mathbf{z}_1 \\ \mathbf{E}_\ell & \mathbf{A}_2 & \mathbf{z}_2 \\ \mathbf{u}_1^* & \mathbf{u}_2* & 0 \end{bmatrix} =$$

$$\begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{E}_\ell \mathbf{K}_1^{-1} & \widetilde{\mathbf{A}}_2 & \tilde{\mathbf{z}}_2 \\ \tilde{\mathbf{u}}_1^* \mathbf{K}_1^{-1} & \tilde{\mathbf{u}}_2^* & \beta \end{bmatrix} \begin{bmatrix} \mathbf{K}_1 & \mathbf{E}_u & \mathbf{z}_1 \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0}^* & \mathbf{0}^* & 1 \end{bmatrix}$$

$$\begin{bmatrix} \widetilde{\mathbf{A}}_2 & \tilde{\mathbf{z}}_2 \\ \tilde{\mathbf{u}}_2^* & \beta \end{bmatrix}$$ well conditioned & low dimensional

allows "exact" direct methods

[Saad 94, Botta & Wubs 97]

MRILU

Reorder & partition

$$\mathbf{A}^{(0)} = \mathbf{A} - \tau\mathbf{B} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{E} \\ \mathbf{F} & \mathbf{A}_2 \end{bmatrix}$$

such that $\mathbf{A}_1 \approx \mathbf{D} := \mathrm{diag}(\mathbf{A}_1)$

Take $\mathrm{M}^{(0)} = \mathrm{M} = \begin{bmatrix} \mathbf{D} & \mathbf{E} \\ \mathbf{F} & \mathbf{A}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{F}\mathbf{D}^{-1} & \widetilde{\mathbf{A}}_2 \end{bmatrix} \begin{bmatrix} \mathbf{D} & \mathbf{E} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$

Use Schur complement $\widetilde{\mathbf{A}}_2 = \mathbf{A}_2 - \mathbf{F}\mathbf{D}^{-1}\mathbf{E}$

for solving $\mathrm{M} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} b \\ c \end{bmatrix}$:

- solve (approximately) $\widetilde{\mathbf{A}}_2\, y = c - \mathbf{F}\mathbf{D}^{-1} b$
- solve exactly $\mathbf{D}\, x = b - \mathbf{E}\, y$

Can be repeated with
$\mathbf{A}$ replaced by $\mathbf{A}^{(1)} = \widetilde{\mathbf{A}}_2$, etc.

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_1 & \mathbf{E}_u \\ \mathbf{E}_\ell & \mathbf{A}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{E}_\ell \mathbf{K}_1^{-1} & \widetilde{\mathbf{A}}_2 \end{bmatrix} \begin{bmatrix} \mathbf{K}_1 & \mathbf{E}_u \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$$

Improved preconditioner

$$\mathbf{A} - \theta \mathbf{B} = (\mathbf{A} - \tau \mathbf{B}) - (\theta - \tau)\mathbf{B}$$

$$\approx \mathbf{K} - \delta \mathbf{B} = \begin{bmatrix} \mathbf{K}_1 - \delta \mathbf{B}_1 & \mathbf{E}_u - \delta \mathbf{F}_u \\ \mathbf{E}_\ell - \delta \mathbf{F}_\ell & \mathbf{A}_2 - \delta \mathbf{B}_2 \end{bmatrix}$$

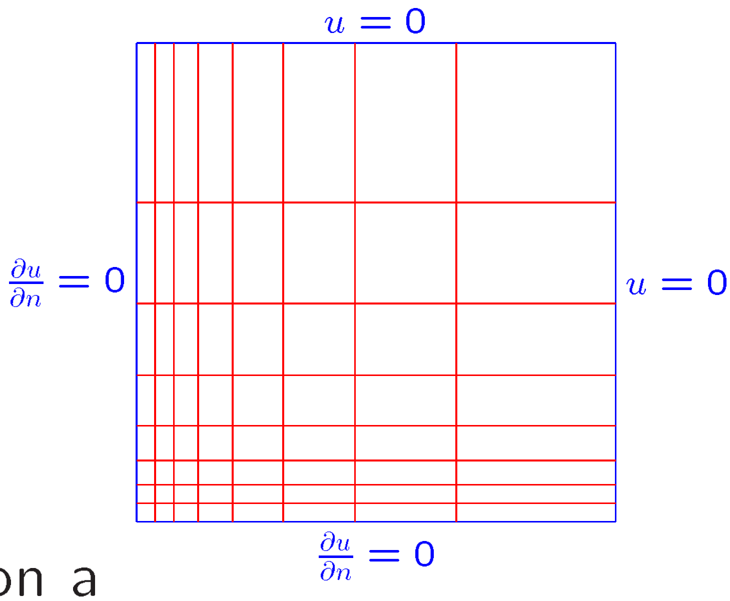
Neumann series:

$$(\mathbf{K}_1 - \delta \mathbf{B}_1)^{-1} \approx \mathbf{C}_1 \equiv \mathbf{K}_1^{-1} + \delta\, \mathbf{K}_1^{-1} \mathbf{B}_1 \mathbf{K}_1^{-1}$$

$$\mathbf{M}_\theta = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ (\mathbf{E}_\ell - \delta \mathbf{F}_\ell)\mathbf{C}_1 & \widetilde{\mathbf{A}}_2 - \delta \tilde{\mathbf{B}}_2 \end{bmatrix} \begin{bmatrix} \mathbf{C}_1^{-1} & \mathbf{E}_u - \delta \mathbf{F}_u \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$$

Note that $\quad \mathbf{M}_\theta \tilde{\mathbf{x}} = \mathbf{0} \quad \Rightarrow \quad \theta \approx \lambda$

Numerical Example



2d Laplace operator on a streched $N \times N$ regular grid.

Dirichlet on east & north boundary
Neumann on west and south boundary

Preconditioned for 3 smallest eigenvalues
$\epsilon = 10^{-8}$
(3 eigenvalues detected at machine precision)

N	Cond($\mathbf{A}$)	fill in	levels	# JD steps
7	$2.4\,10^5$	5.6	3	7
15	$7.4\,10^6$	7.3	8	10
31	$7.3\,10^7$	7.9	12	12
63	$6.1\,10^8$	8.4	17	13

Numerical Example

Linearized incompressible Navier Stokes (Re=100).
Discretized on 17×17 stag. grid, upwind

Generalized eigenproblem of dim. $3 \times 15 \times 15$

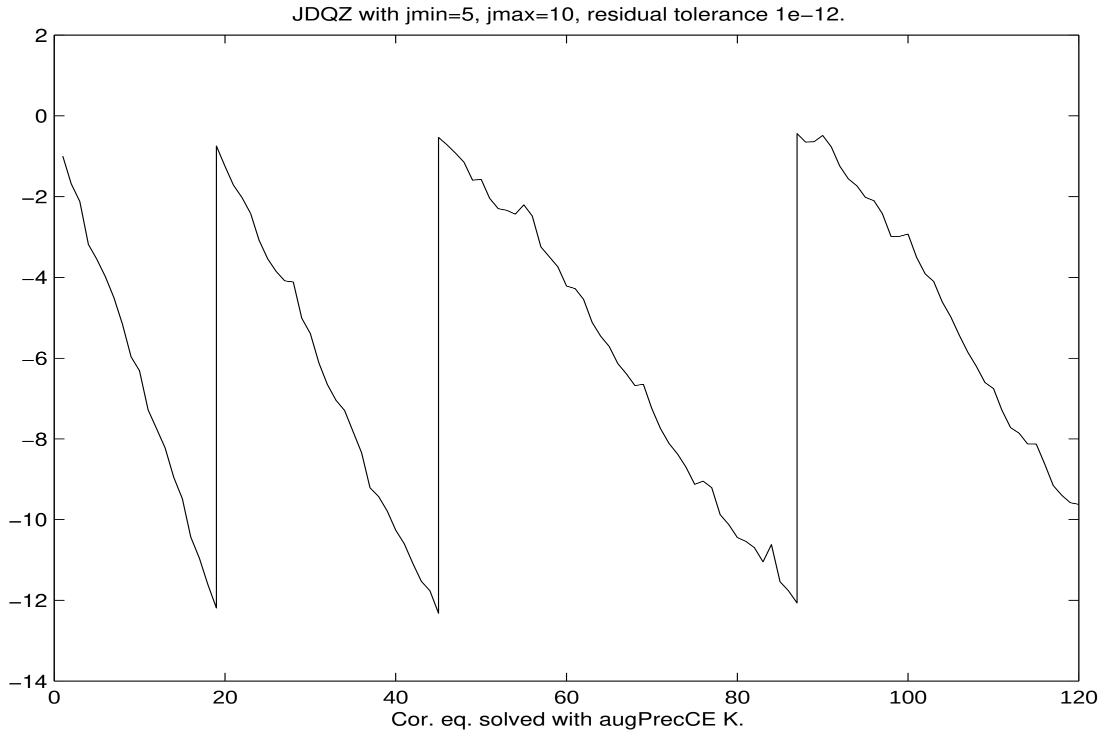
$$\mathbf{A}\mathbf{x} = \lambda\,\mathbf{B}$$

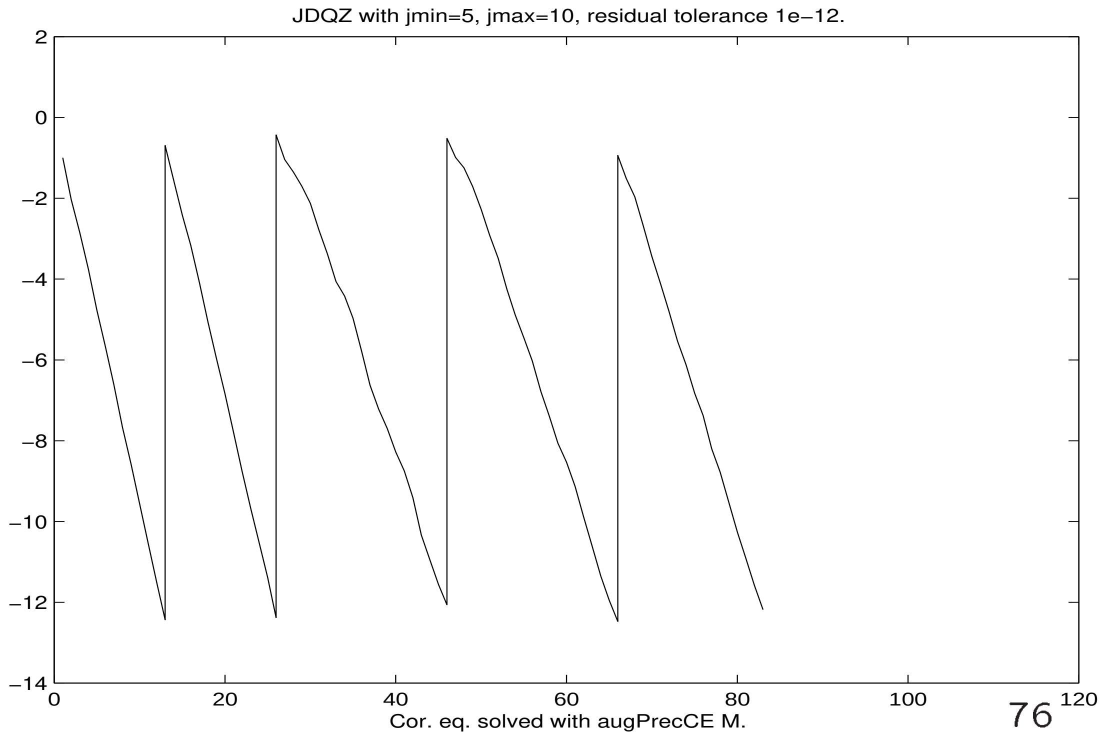
Find 5 eigenvalues with smallest real part

Corr. eq. solved with 1 step precond. GMRES

`[Olsen's method]`

Maximum dimension $\mathbf{V}_m$: 10
after restart: $m = 5$.


JDQZ with jmin=5, jmax=10, residual tolerance 1e−12.
2
0
−2
−4
−6
−8
−10
−12
−14
0
20
40
60
80
100
120
Cor. eq. solved with augPrecCE K.


JDQZ with jmin=5, jmax=10, residual tolerance 1e−12.
2
0
−2
−4
−6
−8
−10
−12
−14
0
20
40
60
80
100
120
Cor. eq. solved with augPrecCE M.

More Details

JD-algorithm:

Sleijpen and van der Vorst

SIMAX, 17 (1996), pp. 401-425

JD: general problems

Sleijpen et al

BIT, 36:3 (1996), pp. 595–633

JDQR and JDQZ

Fokkema et al,

Preprint 941, SIAM J. Sc. Comput., fall 1998

Parallel performance

Sleijpen et al, SIAM News, Aug. 96

these papers and others:

`http://www.math.ruu.nl/people/sleijpen`
`http://www.math.ruu.nl/people/vorst`