A Jacobi-Davidson Method for Solving Complex Symmetric Eigenvalue Problems

Peter Arbenz

ETH Zurich

Institute of Scientific Computing

http://www.inf.ethz.ch/~arbenz

Michiel Hochstenbach

University of Utrecht

Mathematics Department

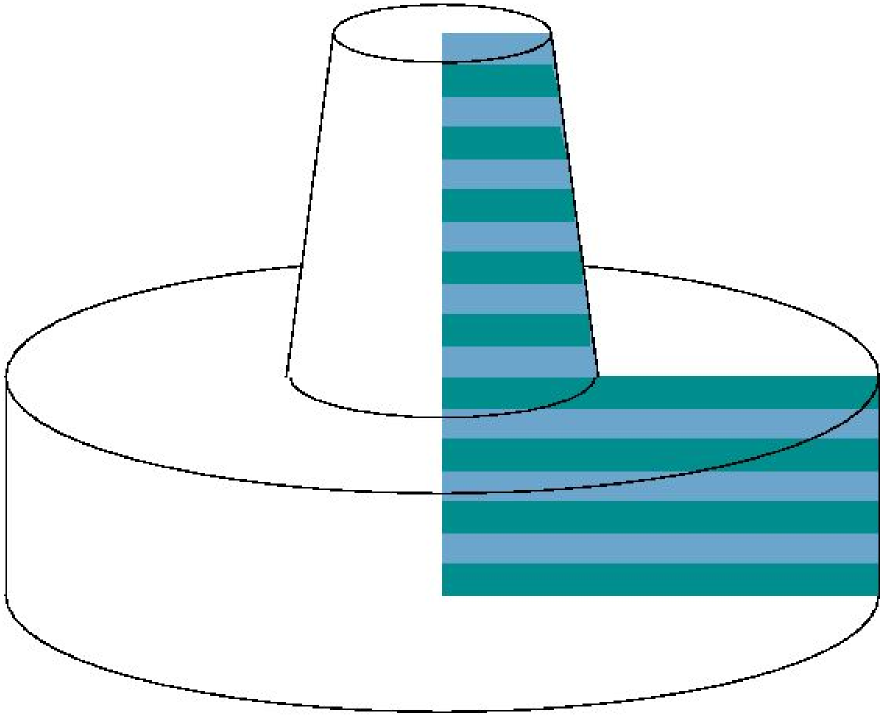
http://www.math.uu.nl/people/hochsten

SCOPES Kick-off meeting, Sofia, February 27, 2002.

The problem

Maxwell equations

$$\mathbf{curl}\,(\mu^{-1}\mathbf{curl}\,\mathbf{e}(\mathbf{x})) - k_0^2\,\varepsilon\,\mathbf{e}(\mathbf{x}) = \mathbf{0}, \qquad \operatorname{div}\mathbf{e}(\mathbf{x}) = 0, \qquad \mathbf{x}\in\Omega \qquad (1)$$



Permittivity ε complex and discontinuous.

Finite element discretization (Nédélec elements)

The problem [cnt'd]

We want to compute a few interior eigenpairs of

$$A\mathbf{x} = \lambda \mathbf{x}, \tag{2}$$

or

$$A\mathbf{x} = \lambda B\mathbf{x}, \qquad \exists\, B^{-1}, \tag{3}$$

where both A and B are large, sparse and complex-symmetric.

Eq. (3) can be transformed into (2) if a symmetric factorization $B = CC^T$ exists. (Analogous to the real-symmetric case).

What's special about the complex-symmetric evp?

Every matrix is similar to a complex-symmetric matrix as($n = 4$)

$$\begin{bmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \lambda & 1 \\ & & & \lambda \end{bmatrix} \sim \begin{bmatrix} \lambda & 1 & & \\ 1 & \lambda & 1 & \\ & 1 & \lambda & 1 \\ & & 1 & \lambda \end{bmatrix} + i \begin{bmatrix} & & -1 & \\ & -1 & & 1 \\ -1 & & 1 & \\ & 1 & & \end{bmatrix} \tag{4}$$

see Gantmacher, vol. 2 (1959) or Horn-Johnson (1985). Thus, it may be arbitrarily difficult to solve (2) or (3), respectively.

Nevertheless, there are some properties among the eigenvectors.

- $A\mathbf{x} = \lambda\mathbf{x} \quad \Longrightarrow \quad \mathbf{x}^T A = \lambda\mathbf{x}^T.$
- $A\mathbf{x} = \lambda\mathbf{x}, \quad A\mathbf{y} = \mu\mathbf{y}, \quad \lambda \neq \mu \quad \Longrightarrow \quad (\mathbf{x}, \mathbf{y})_T := \mathbf{x}^T\mathbf{y} = 0.$
- If A is diagonalizable then the diagonalization can be realized by a complex-orthogonal matrix Q, $Q^T Q = I$.
- Takagi's factorization (SVD): $A = U\Sigma U^T$

Iterative solvers for $A\mathrm{x} = \mathrm{b}$ exploiting the complex symmetric structure

- Complex orthogonal cg method, COCG (van der Vorst & Melissen, 1990)

  Construction of a basis for the m-th Krylov space $\{\mathbf{v}, A\mathbf{v}, \ldots, A^{m-1}\mathbf{v}\}$ by enforcing orthogonality w.r.t. the pseudo-inner product $(\mathbf{x}, \mathbf{y})_T := \mathbf{x}^T\mathbf{y}$. This yields a three-term recurrence among the basis (Lanczos) vectors,

$$AV_m = V_m T_m + \mathbf{v}_m \mathbf{e}_m^T, \qquad T_m \text{ tridiagonal.} \tag{5}$$

  The approximate solution after m step of the procedure is $V_m\mathbf{y}_m$ where

$$\mathbf{b} - AV_m\mathbf{y}_m \perp \mathcal{R}(V_m) \qquad \Longrightarrow \qquad V_m^T V_m T_m \mathbf{y}_m = V_m^T \mathbf{b}. \tag{6}$$

  The Lanczos procedure may break down, as $(\mathbf{x}, \mathbf{x})_T = 0$ for $\mathbf{x} \neq \mathbf{0}$ is possible.

  We actually set $\mathbf{v}_0 = \mathbf{b} = \mathbf{b} - A\mathbf{x}_0$ for $\mathbf{x}_0 = \mathbf{0}$.

  Note: This procedure can be interpreted as BiCG with initial 'shadow vector' $\mathbf{w}_0 = \bar{\mathbf{v}}_0$.

- Complex-symmetric QMR, CSYMQMR (Freund, 1992)

  Same Krylov space, but the approximate solution $V_m\mathbf{y}_m$ at the m-th iteration step is determined in a QMR fashion,

  $$\|\tilde{T}_m\mathbf{y}_m - V_{m+1}^T\mathbf{b}\| = \text{ minimal.} \tag{7}$$

  Smoothed behavior of the residual norm.

- CSYM (Bunse-Gerstner & Stöver, 1999)

  Computation of a factorization of the form

  $$QAQ^T = \text{ complex symmetric tridiagonal} \tag{8}$$

  where Q is unitary. This is done column by column.

  Note: Procedure can be considered a first step towards the Takaki factorization of A.

Want to exploit these solvers in the context of Jacobi-Davidson.

Jacobi-Davidson (JD), cook book review

(Sleijpen & van der Vorst, 1995)

Let $\mathcal{V}_m = \mathcal{R}(V_m) \equiv \mathcal{R}[\mathbf{v}_1, \ldots, \mathbf{v}_m])$. We want to improve $\mathcal{V}_m$ as our trial space for solving $A\mathbf{x} = \lambda B\mathbf{x}$.

We proceed in two steps:

1. Extraction of a suitable vector from $\mathcal{V}_m$.

   Compute a Ritzpair $(\tilde{\lambda}, \tilde{\mathbf{q}})$, $\tilde{\mathbf{q}} \in \mathcal{V}_m$.

   As $m \ll n$ this is a very small complex symmetric subproblem.

2. Expansion of V_m by a suitable vector.

   Solve the so-called correction equation

$$(I - B\tilde{\mathbf{q}}\tilde{\mathbf{q}}^T)(A - \eta B)\mathbf{t} = -\tilde{\mathbf{r}}, \qquad (I - \tilde{\mathbf{q}}\tilde{\mathbf{q}}^T B)\mathbf{t} = \mathbf{t} \qquad (9)$$

   for $\mathbf{t}$. Here, $\tilde{\mathbf{r}} = (A - \tilde{\lambda}B)\tilde{\mathbf{q}}$ and η is some shift.

   Finally, $\mathbf{t}$ is B-pseudo-orthogonalized against $\mathbf{v}_1, \ldots, \mathbf{v}_m$ to yield $\mathbf{v}_{m+1}$ and $V_{m+1} = \mathcal{R}([\mathbf{v}_1, \ldots, \mathbf{v}_{m+1}])$.

Remarks

- If $\hat{\mathbf{t}}$ is the exact solution of (9) then

$$\tilde{\mathbf{q}} + \hat{\mathbf{t}} = \gamma(A - \eta B)^{-1} B\tilde{\mathbf{q}}, \qquad \hat{\mathbf{t}}^T B\tilde{\mathbf{q}} = 0.$$

  Thus, solving the correction equation can be considered as executing one step of inverse vector iteration.

- However, the shift η is fixed only in the beginning of the iteration. Close to convergence $\eta = \tilde{\lambda}$ is set which amounts to Rayleigh quotient iteration. With this shift, we have cubic convergence if we solve (9) exactly.

- However, (9) is solved iteratively (and approximatively). (Neither A nor B must be factored.) Preconditioning possible (→ Davidson). Loss of cubic convergence rate. But quick solution of 'approximate' correction.

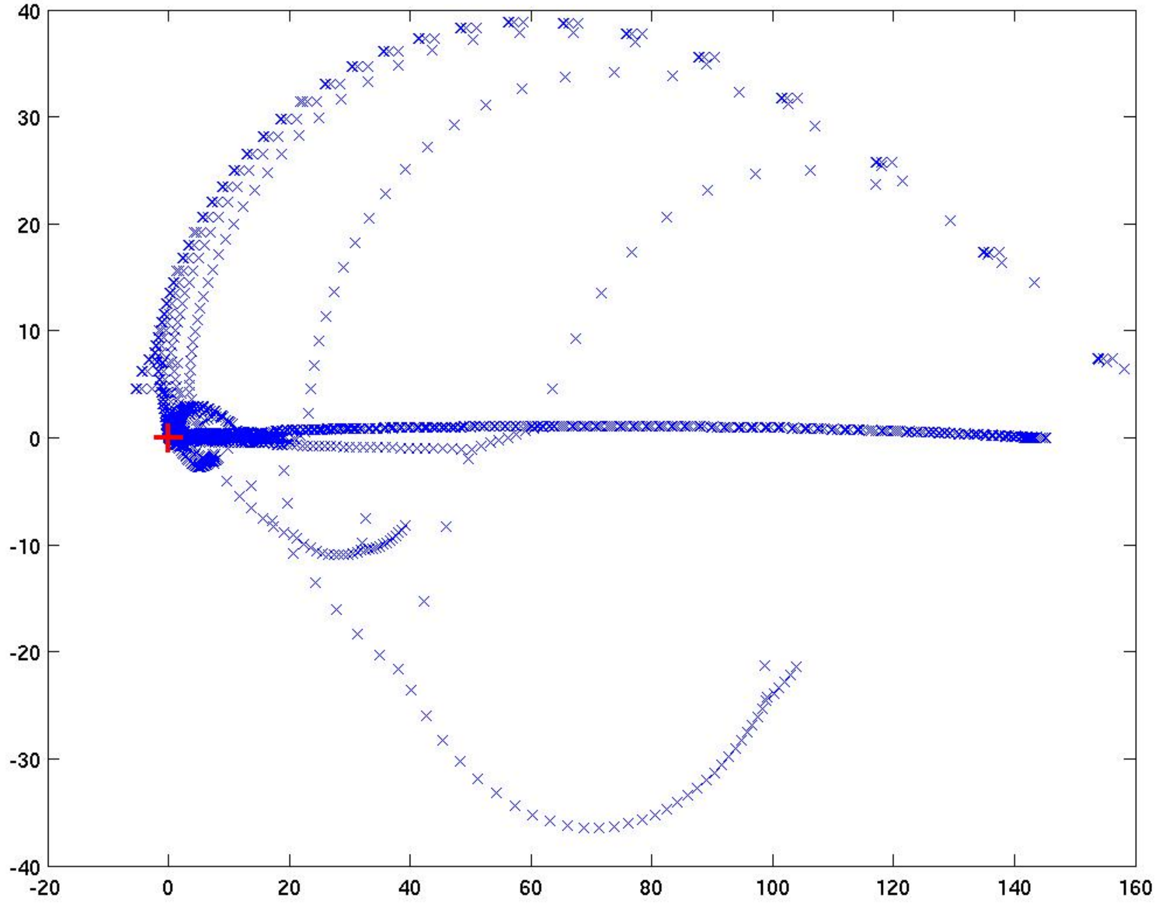
- If several eigenpairs are desired then compute one at the time.

- Restart if $m = j_{\text{max}}$: extract j_{min} best Ritz vectors from search space.

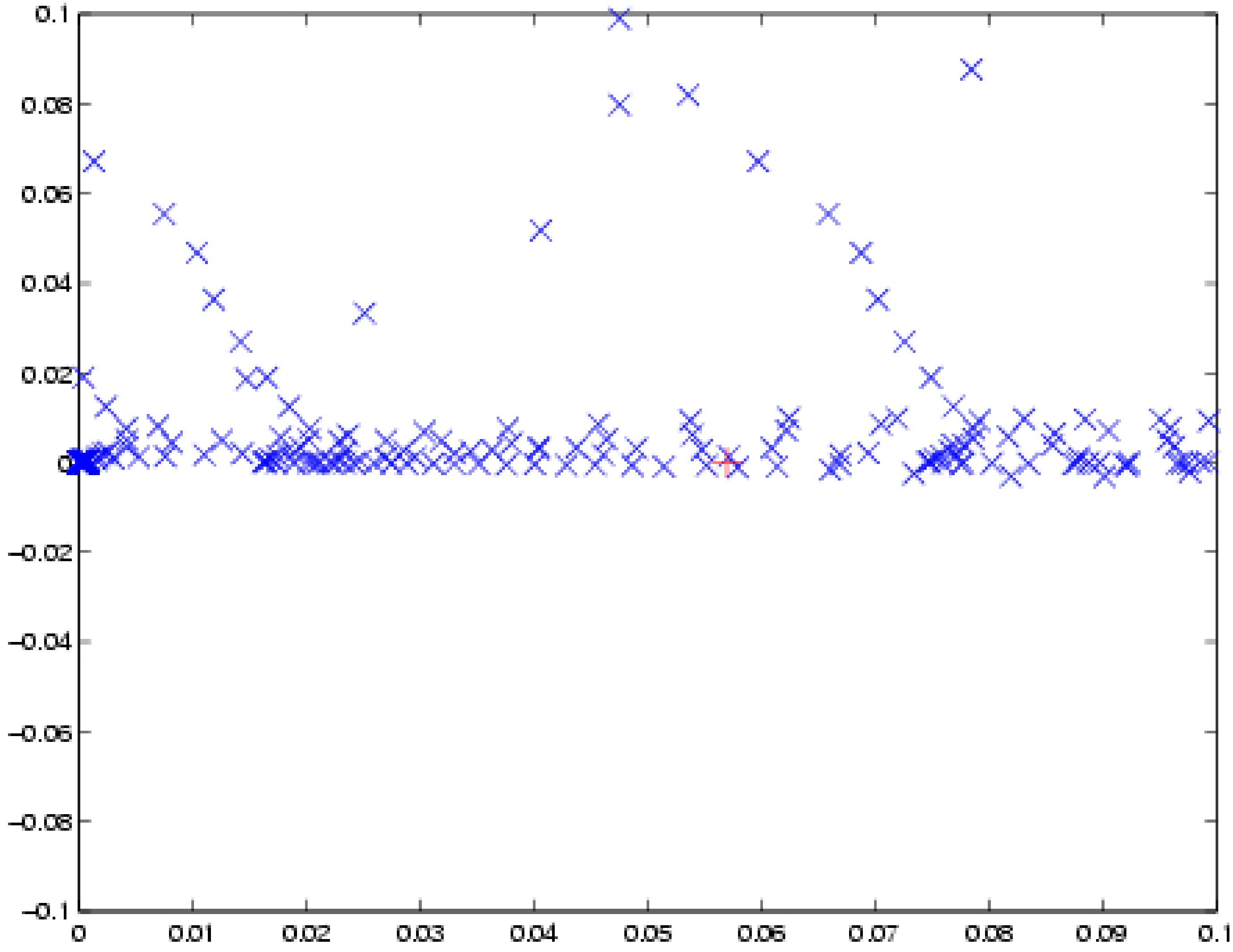
- The $\mathcal{V}_m$ aren't Krylov subspaces.

Computing interior eigenvalues

Want to compute a few eigenvalues close to some target τ.



Closer look at the spectrum close to the target.



Harmonic Ritz approach

Difficult to extract interior eigenvalues with the straightforward Ritz-Galerkin approach.

We use the harmonic Ritz approach (Bai et al. 2000):

$$(A-\tau B)^{-1}B\tilde{\mathbf{u}} - (\tilde{\theta}-\tau)^{-1}\tilde{\mathbf{u}} \perp_B \widetilde{\mathcal{U}}, \quad \widetilde{u} \in \widetilde{\mathcal{U}}. \tag{10}$$

With $\widetilde{\mathcal{U}} := B^{-1}(A-\tau B)\mathcal{V}_m$ and $\widetilde{u} = V_m\mathbf{c}$ this condition becomes

$$V_m^T(A-\tau B)V_m\mathbf{c} = (\widetilde{\theta}-\tau)^{-1}V_m^T(A-\tau B)B^{-1}(A-\tau B)V_m\mathbf{c}. \tag{11}$$

Complex-symmetric, but involves B^{-1}.

(One reason for selection JD as an eigensolver was that it does not need matrix factorizations.)

Make a Petrov-Galerkin approach. Replace

$$(A - \tau B)^{-1} B\tilde{\mathbf{u}} - (\tilde{\theta} - \tau)^{-1}\tilde{\mathbf{u}} \perp_B \widetilde{\mathcal{U}}, \quad \widetilde{u} \in \widetilde{\mathcal{U}}. \tag{10}$$

by

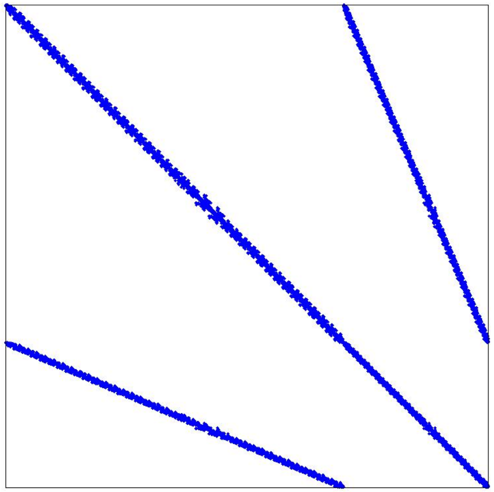
$$(A - \tau B)^{-1} B\tilde{\mathbf{u}} - (\tilde{\theta} - \tau)^{-1}\tilde{\mathbf{u}} \perp_T \widetilde{\mathcal{V}}, \quad \widetilde{u} \in \widetilde{\mathcal{U}}. \tag{12}$$

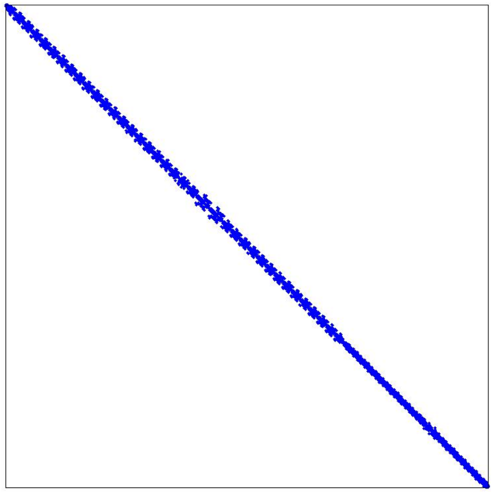
where $\widetilde{\mathcal{U}} = \mathcal{V}_m$ and $\widetilde{\mathcal{V}} = (A - \tau B)^2 \mathcal{V}_m$. This becomes

$$V_m^T (A - \tau B) B V_m \mathbf{c} = (\widetilde{\theta} - \tau)^{-1} V_m^T (A - \tau B)^2 V_m \mathbf{c}. \tag{13}$$

This eigenvalue problem is not complex-symmetric!

A numerical experiment from VCSEL design



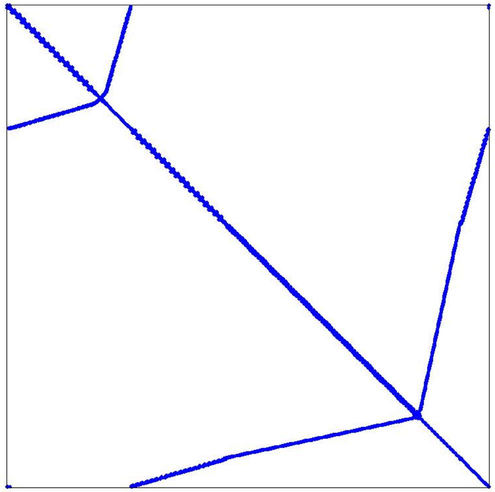


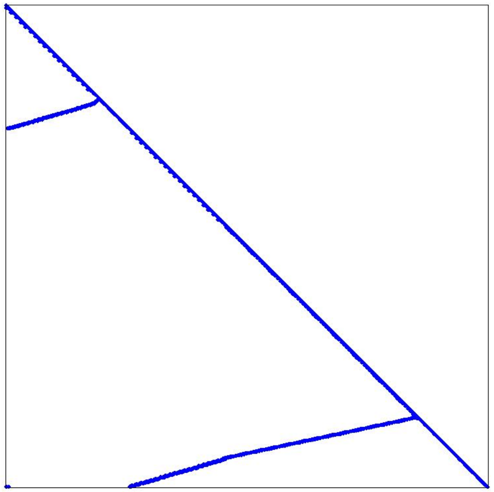
$n = 6243, nnz(A) = 79833, nnz(B) = 24945.$

Preconditioners

Choose complex symmetric preconditioners of the form LDL^T.

- diagonal
- symmetric Gauss-Seidel
- incomplete complex-symmetric 'Cholesky'
- LDL^T factorization of $A - \tau B$ (after symmetric minimum-degree reordering) (MATLAB: `[L,U]=lu(A,0)`)

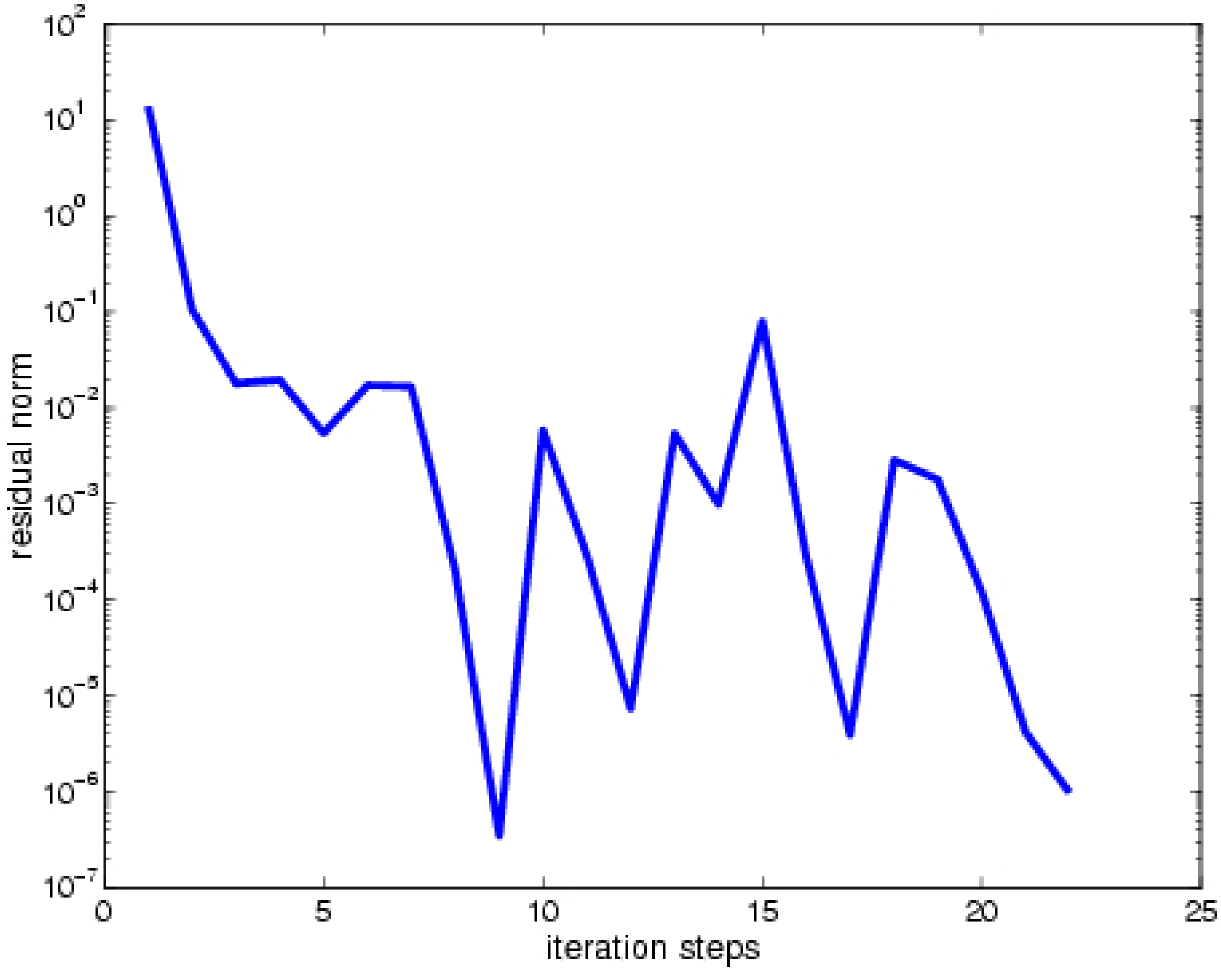




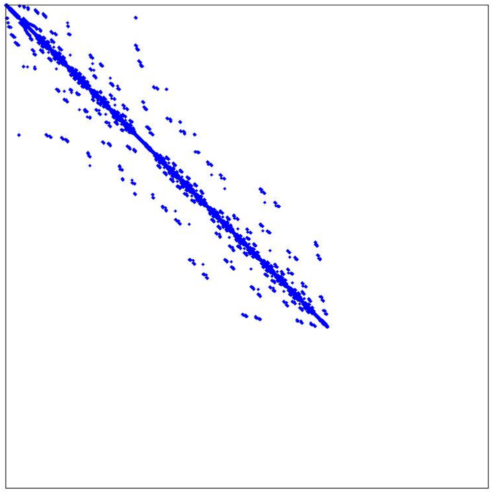
Reordered matrix and triangular factor L

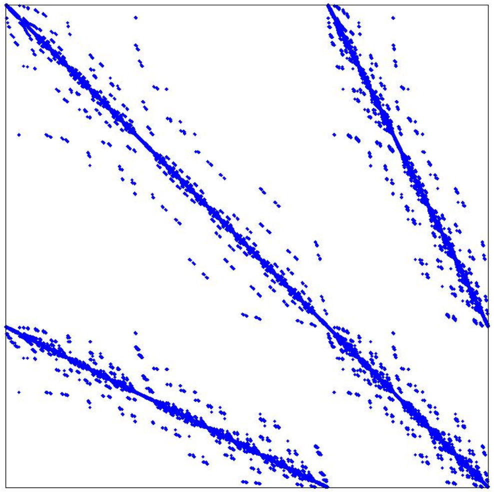
$$nnz(A - \tau B) = 79833, nnz(L) = 94399.$$

Convergence history with Jacobi-Davidson/csym QMR

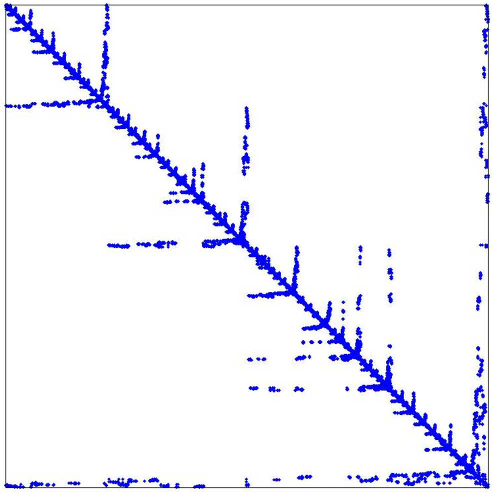


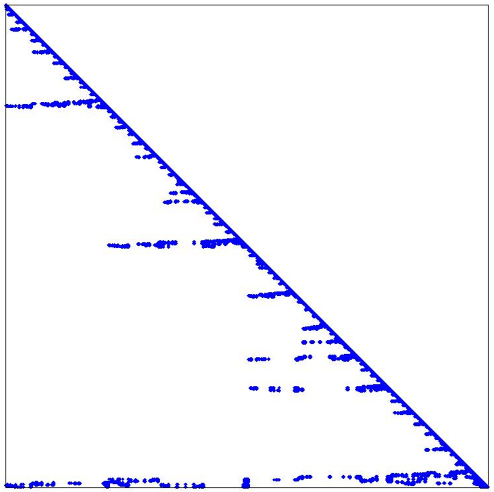
A waveguide problem





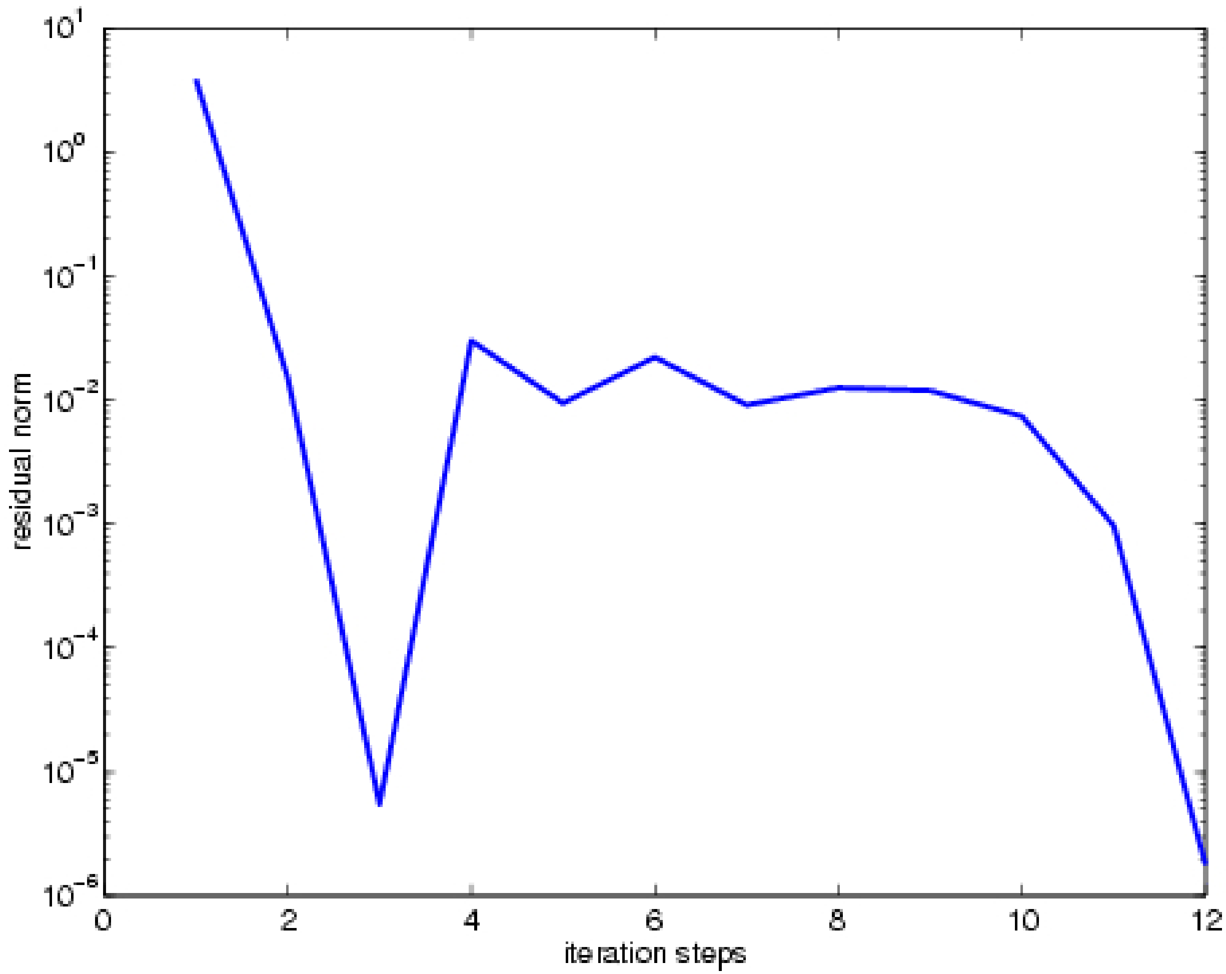
$n = 32098, nnz(A) = 148436, nnz(B) = 411622.$





$n = 32098, nnz(A - \tau B) = 411622, nnz(L) = 2595054.$

Convergence history with Jacobi-Davidson/csym QMR



Conclusions (open problems)

- We can solve our test problems (the shown and larger ones)
- But we use plain LU factorization as preconditioner. (Shift-and-invert Lanczos/Arnoldi may be a better choice than JD.)
- Other preconditioners do not work.

  Why?

  - Inner iteration does not converge (in the permitted number of steps).
  - Extraction of the harmonic Ritz pair? See (Sleijpen, van den Eshof, 2001).