Hamiltonian and Symplectic Lanczos Processes

David S. Watkins

Department of Mathematics
Washington State University
Pullman, WA 99164-3113
U.S.A.

Collaborators: V. Mehrmann, T. Apel,
P. Benner, H. Faßbender, . . .

Problem: Linear Elasticity

- $(\lambda^2 M + \lambda G + K)v = 0$

  $M^T = M > 0 \quad G^T = -G \quad K^T = K \leq 0$

- quadratic eigenvalue problem

- large, sparse (finite elements)

- Find few eigenvalues nearest imaginary axis (and corresponding eigenvectors).

Problem: Optimal Control

- $\begin{bmatrix} A & BB^T \\ C^TC & -A^T \end{bmatrix} - \lambda \begin{bmatrix} E & 0 \\ 0 & E^T \end{bmatrix}$

  (large and sparse)

- Hamiltonian/skew-Hamiltonian

- multiply by $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$

- $\begin{bmatrix} C^TC & -A^T \\ -A & -BB^T \end{bmatrix} - \lambda \begin{bmatrix} 0 & E^T \\ -E & 0 \end{bmatrix}$

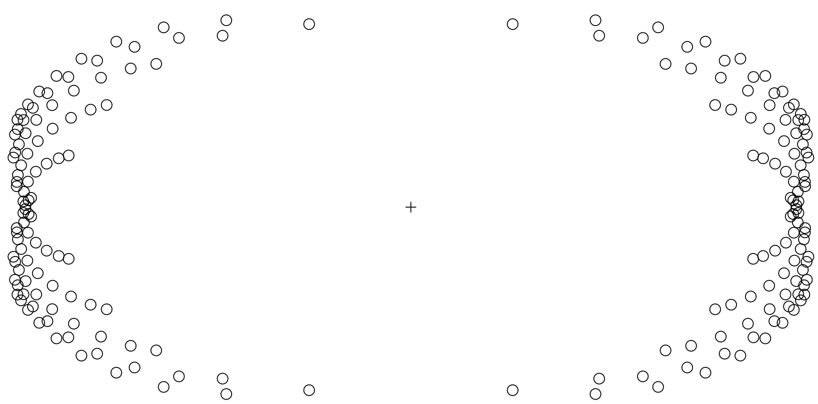
- symmetric/skew-symmetric

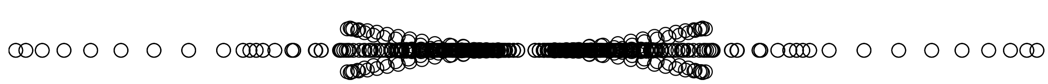
Hamiltonian Structure

- Our matrices are real.

- λ, $\overline{\lambda}$, $-\overline{\lambda}$, $-\lambda$ occur together.

- seen also in Hamiltonian matrices





Hamiltonian Matrices

- $H \in \mathbb{R}^{2n \times 2n}$

- $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \in \mathbb{R}^{2n \times 2n}$

- H is *Hamiltonian* iff JH is symmetric.

- $H = \begin{bmatrix} A & K \\ N & -A^T \end{bmatrix}$,

  where $K = K^T$ and $N = N^T$

Linearization

- $\lambda^2 Mv + \lambda Gv + Kv = 0$

- $w = \lambda v, \qquad Mw = \lambda Mv$

- $\begin{bmatrix} -K & 0 \\ 0 & -M \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} - \lambda \begin{bmatrix} G & M \\ -M & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = 0$

- $Ax - \lambda Nx = 0$

- symmetric/skew-symmetric

Reduction to Hamiltonian Matrix

- $A - \lambda N$ (symmetric/skew-symmetric)

- $N = R^T J R$ $\left(J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \right)$

  sometimes easy, always possible

- Transform:

$$A - \lambda R^T J R$$

$$R^{-T} A R^{-1} - \lambda J$$

$$J^T R^{-T} A R^{-1} - \lambda I$$

- $H = J^T R^{-T} A R^{-1}$ is Hamiltonian.

Example

- $N = \begin{bmatrix} G & M \\ -M & 0 \end{bmatrix}$

- $N = R^T J R = \begin{bmatrix} I & -\frac{1}{2}G \\ 0 & M \end{bmatrix} \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ \frac{1}{2}G & M \end{bmatrix}$

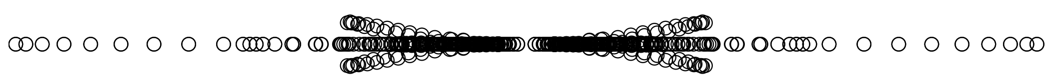
- $$\begin{aligned} H &= J^T R^{-T} A R^{-1} \\ &= \begin{bmatrix} I & 0 \\ -\frac{1}{2}G & I \end{bmatrix} \begin{bmatrix} 0 & M^{-1} \\ -K & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ -\frac{1}{2}G & I \end{bmatrix} \end{aligned}$$

Sparse Representation of H

- Krylov subspace methods

- We just need to apply the operator. $(M = LL^T)$

$$H = \begin{bmatrix} I & 0 \\ -\frac{1}{2}G & I \end{bmatrix} \begin{bmatrix} 0 & M^{-1} \\ -K & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ -\frac{1}{2}G & I \end{bmatrix}$$



$$H^{-1} = \begin{bmatrix} I & 0 \\ \frac{1}{2}G & I \end{bmatrix} \begin{bmatrix} 0 & (-K)^{-1} \\ M & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ \frac{1}{2}G & I \end{bmatrix}$$

Exploitable Structures

- Hamiltonian

$$H^{-1}$$

$$H^{-1}(H - \tau I)^{-1}(H + \tau I)^{-1}$$

- skew-Hamiltonian

$$H^{-2}$$

$$(H - \tau I)^{-1}(H + \tau I)^{-1}$$

- symplectic

$$(H - \tau I)^{-1}(H + \tau I)$$

$\tau =$ target shift

Note: $(H - \tau I)^{-1}$ has none of these structures.

Unsymmetric Lanczos Process

- Standard unsymmetric Lanczos effects a (partial) similarity transformation

$$A \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \begin{bmatrix} \ddots & \ddots & \\ \ddots & \ddots & \ddots \\ & \ddots & \ddots \end{bmatrix}$$

$$U^{-1}AU = \begin{bmatrix} \ddots & \ddots & \\ \ddots & \ddots & \ddots \\ & \ddots & \ddots \end{bmatrix}$$

- partial similarity transformation:

$$A \begin{bmatrix} u_1 & \cdots & u_k \end{bmatrix} = \begin{bmatrix} u_1 & \cdots & u_k \end{bmatrix} \begin{bmatrix} \ddots & \ddots & \\ \ddots & \ddots & \ddots \\ & \ddots & \ddots \end{bmatrix}$$

$$+u_{k+1}\beta_k e_k^T$$

- short Lanczos runs
  (breakdowns!!, no look-ahead)

$$A \begin{bmatrix} u_1 & \cdots & u_k \end{bmatrix} = \begin{bmatrix} u_1 & \cdots & u_k \end{bmatrix} \begin{bmatrix} \diagdown \end{bmatrix} + u_{k+1}\beta_k e_k^T$$

- Get eigenvalues of $\begin{bmatrix} \diagdown \end{bmatrix}$

- Restart (implicitly)

  IRA (Sorensen 1991), ARPACK

  Restart Lanczos with HR
  (Grimme/Sorensen/Van Dooren 1996)

Structured Lanczos Methods

- Similarity transformation: $S^{-1}AS = \hat{A}$

- S symplectic $\Rightarrow$ structure preserved

  - symplectic (Lie group)
  - Hamiltonian (Lie algebra)
  - skew-Hamiltonian (Jordan algebra)

- Conclusion: A "Lanczos" process that builds a symplectic similarity transformation will preserve structure.

  Vectors produced should be columns of a symplectic matrix.

Symplectic Matrices

- $S \in \mathbb{R}^{2n \times 2n}$

- $S^T J S = J \qquad \left(J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \right)$

- $S = \begin{bmatrix} U & V \end{bmatrix}$

- $\begin{bmatrix} U^T \\ V^T \end{bmatrix} J \begin{bmatrix} U & V \end{bmatrix} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$

- $U^T J U = 0$, $V^T J V = 0$, $U^T J V = I$

- Subspaces are isotropic.

Isotropic Subspaces

- $y^T Jx = 0$ for all x, $y \in \mathcal{U}$
- $U = \left[\begin{array}{ccc} u_1 & \cdots & u_k \end{array} \right]$
- $U^T JU = 0$
- Structured methods build isotropic subspaces.

Skew-Hamiltonian Case

Theorem: B skew Hamiltonian, $x \neq 0 \Rightarrow$

$$\text{span}\{x, Bx, \ldots, B^{j-1}x\} \text{ is isotropic.}$$

- Conclusion: Krylov subspace methods preserve skew-Hamiltonian structure automatically.
- Examples: Arnoldi, unsymmetric Lanczos
- exact vs. floating-point arithmetic

Skew-Hamiltonian Arnoldi Process

- Isotropic Arnoldi process

$$\tilde{q}_{j+1} = Bq_j - \sum_{i=1}^{j} q_i h_{ij} - \sum_{i=1}^{j} Jq_i t_{ij}$$

- produces *isotropic* subspaces:
  $Jq_1, \ldots, Jq_k$ are orthogonal to $q_1, \ldots, q_k$.

- Theory $t_{ij} = 0$

- Practice $t_{ij} = \epsilon$ (roundoff)

- Enforcement of isotropy is crucial.

- Consequence: get each eigenvalue only once.

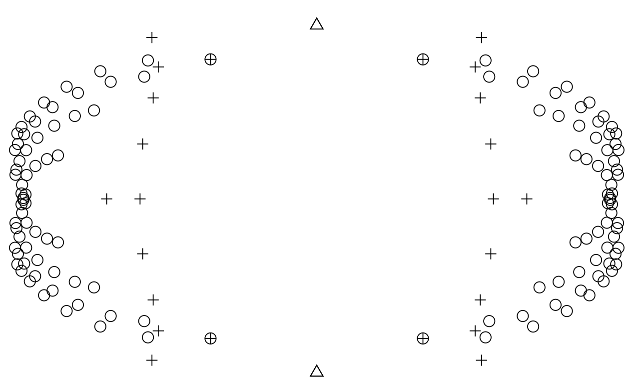
Example

- Method: Implicitly Restarted Arnoldi (effective combination of Arnoldi and subspace iteration)

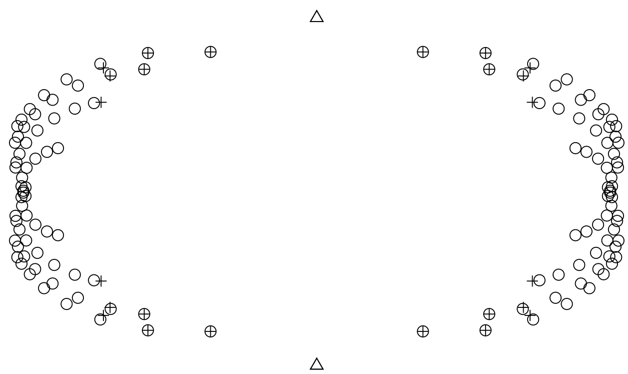
- Toy problem ($n = 64$); asking for 8 eigenvalues (right half-plane).

- Target $\tau = i$ (not particularly good)

- After 12 Arnoldi steps (no restart) . . .



- After one restart (12 more Arnoldi steps)



- Errors: 10^{-14}, 10^{-7}, 10^{-6}, 10^{-2}

- After 7 iterations (restarts) algorithm stops with 8 eigenvalues correct to ten decimal places.

- Residuals: $\|(\lambda^2 M + \lambda G + K)v\| \leq 10^{-12}$ $(\|v\| = 1)$

Further Experience

- Fortran/C code

- $n \approx 2 \times 10^5$

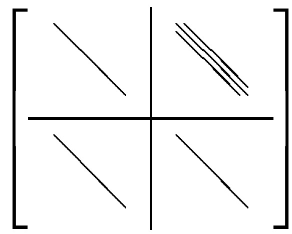
- Disadvantage: Eigenvectors cost extra.
  (eigenvectors of H^2 vs. H)

- We haven't done
  skew-Hamiltonian Lanczos.

Hamiltonian Case

- Bunse-Gerstner/Mehrmann 1986:

$$S^{-1}HS = \begin{bmatrix} E & T \\ D & -E \end{bmatrix} =$$

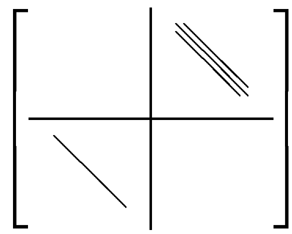


- Further condensation: $E = 0$, $D = \text{diag}\{\pm 1 \cdots \pm 1\}$.

- $S = \begin{bmatrix} U & V \end{bmatrix}$

- $H \begin{bmatrix} U & V \end{bmatrix} = \begin{bmatrix} U & V \end{bmatrix} \begin{bmatrix} 0 & T \\ D & 0 \end{bmatrix}$

Condensed Hamiltonian Lanczos Process

- $H \begin{bmatrix} U & V \end{bmatrix} = \begin{bmatrix} U & V \end{bmatrix}$ 

  $U = \begin{bmatrix} u_1 & u_2 & \cdots \end{bmatrix} \qquad V = \begin{bmatrix} v_1 & v_2 & \cdots \end{bmatrix}$

- $Hu_k = v_k d_k$

  $Hv_k = u_{k-1}b_{k-1} + u_k a_k + u_{k+1}b_k$

- $u_{k+1}b_k \quad = Hv_k - u_k a_k - u_{k-1}b_{k-1}$

  $v_{k+1}d_{k+1} = Hu_{k+1}$

- Coefficients are chosen so that $S = \begin{bmatrix} U & V \end{bmatrix}$ is symplectic.

- Collect coefficients.

Equivalence

- H^2 is skew-Hamiltonian.

- Condensed Hamiltonian Lanczos applied to H is theoretically equivalent to ordinary Lanczos applied to H^2.

- Hamiltonian algorithm costs half as many matrix-vector multiplies.

Isotropy

- $S^T J S = J, \qquad J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$

- $\begin{bmatrix} U^T \\ V^T \end{bmatrix} J \begin{bmatrix} U & V \end{bmatrix} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$

- $U^T J U = 0, \qquad$ (isotropic subspaces)
  $V^T J V = 0,$

- In (floating-point) practice, isotropy must be enforced by J-reorthogonalization.

- All vectors must be retained.

- short Lanczos runs, restarts

Implicitly Restarted Hamiltonian Lanczos Process

- use SR, not QR
  (Benner/Fassbender 1997)

- In condensed case, $SR = HR$
  (Benner/Fassbender/W 1998)

- Use of HR yields significant simplification.

Symplectic Case

Structure

- Eigenvalues of S appear in quartets
  μ, μ^{-1}, $\overline{\mu}$, $\overline{\mu}^{-1}$.

- Symplectic Lanczos process must extract these simultaneously.

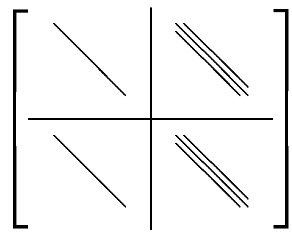
- This is accomplished
  by using both S and S^{-1}.

- $S^{-1} = -JS^TJ$

Symplectic Similarity

- symplectic butterfly form:
  (Banse/Bunse-Gerstner 1994)

$$W^{-1}SW = \begin{bmatrix} D_1 & T_1 \\ D_2 & T_2 \end{bmatrix} =$$

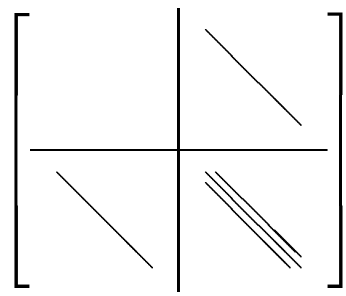


- Further condensation: $D_1 = 0$,
  $D_2 = \text{diag}\{\pm 1 \cdots \pm 1\}$, $T_1 = -D_2$, ...

- $W = \begin{bmatrix} U & V \end{bmatrix}$

- $S \begin{bmatrix} U & V \end{bmatrix} = \begin{bmatrix} U & V \end{bmatrix} \begin{bmatrix} 0 & -D \\ D & DT \end{bmatrix}$

Condensed Symplectic Lanczos Process

- $S \begin{bmatrix} U & V \end{bmatrix} = \begin{bmatrix} U & V \end{bmatrix}$ 

- $Su_k = v_k d_k$

  $Sv_k = -u_k d_k + v_{k-1}\tilde{b}_{k-1} + v_k \tilde{a}_k + v_{k+1}\tilde{b}_k$

- $v_{k+1}\tilde{b}_k \quad = Sv_k - v_k \tilde{a}_k - v_{k-1}\tilde{b}_{k-1} + u_k d_k$

  $u_{k+1} d_{k+1} = S^{-1} v_{k+1}$

- Coefficients are chosen so that $\begin{bmatrix} U & V \end{bmatrix}$ is symplectic.

- Collect coefficients.

Equivalence

- $S + S^{-1}$ is skew-Hamiltonian.

- Condensed symplectic Lanczos applied to S is theoretically equivalent to ordinary Lanczos applied to $S + S^{-1}$.

- Symplectic algorithm costs half as many matrix-vector multiplies.

Isotropy (rerun)

- $\begin{bmatrix} U^T \\ V^T \end{bmatrix} J \begin{bmatrix} U & V \end{bmatrix} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$

- $U^T J U = 0$, (isotropic subspaces)
  $V^T J V = 0$,

- In (floating-point) practice, isotropy must be enforced by J-reorthogonalization.

- All vectors must be retained.

- short Lanczos runs, restarts

Implicitly Restarted Symplectic Lanczos Process

- use symplectic SR, not QR

- In condensed case, $SR = HR$
  (Benner/Fassbender/W 1998)

- Use of HR yields **significant** simplification.

Remarks on Stability

- Both Hamiltonian and symplectic Lanczos processes are potentially unstable.

- Breakdowns can occur.

- Are the answers worth anything?

- right and left eigenvectors

- residuals

- condition numbers for eigenvalues

- Don't skip these tests.

Example

- $\lambda^2 Mv + \lambda Gv + Kv = 0$

- $n = 3423$

$$H = \begin{bmatrix} I & 0 \\ -\frac{1}{2}G & I \end{bmatrix} \begin{bmatrix} 0 & M^{-1} \\ -K & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ -\frac{1}{2}G & I \end{bmatrix}$$

$$H^{-1} = \begin{bmatrix} I & 0 \\ \frac{1}{2}G & I \end{bmatrix} \begin{bmatrix} 0 & (-K)^{-1} \\ M & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ \frac{1}{2}G & I \end{bmatrix}$$

Compare various approaches:

- Hamiltonian(1) H^{-1}
- Hamiltonian(3) $H^{-1}(H-\tau I)^{-1}(H+\tau I)^{-1}$
- symplectic $(H-\tau I)^{-1}(H+\tau I)$
- unstructured $(H-\tau I)^{-1}$
  $+$ ordinary Lanczos with implicit restarts

Get 6 smallest eigenvalues in right half-plane.

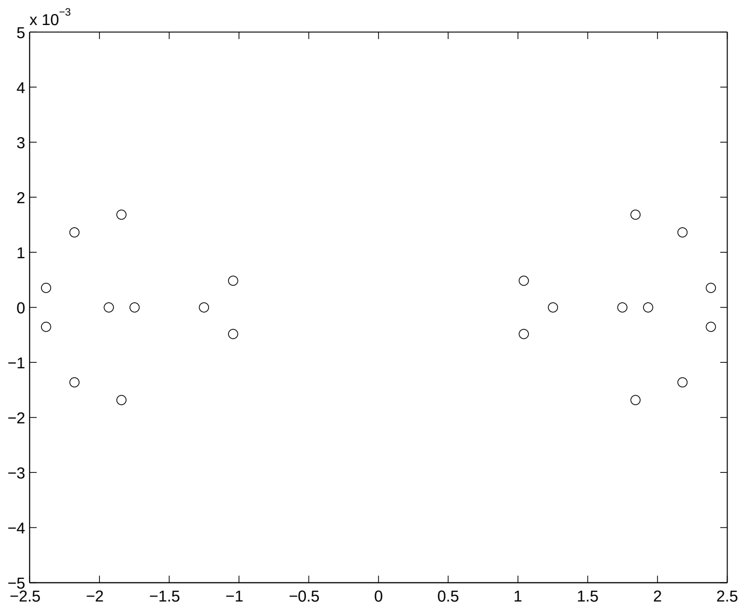
Tolerance $= 10^{-8}$

Take 20 steps and restart with 10.

No-Clue Case ($\tau = 0$)

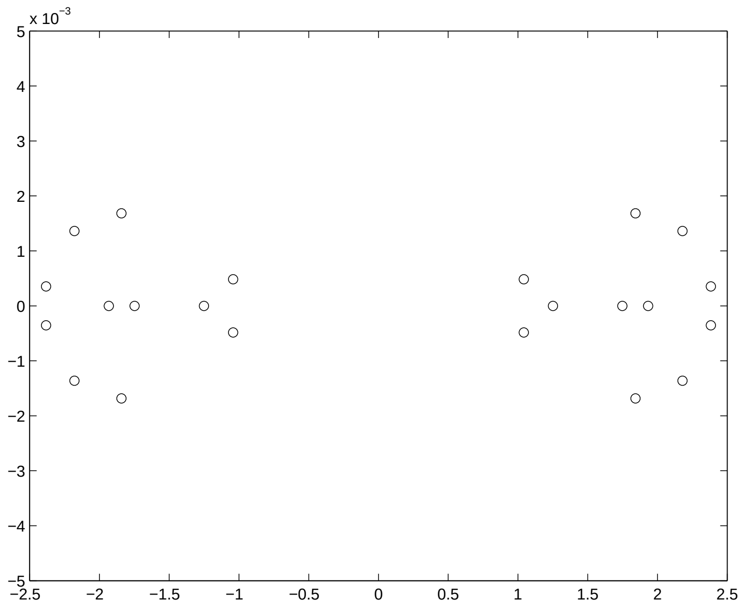
Method	Solves	Eigvals Found	Max. Resid.
Hamiltonian(1)	78	9	2×10^{-10}
Unstructured	158	$7 + 7$	5×10^{-7}

Unstructured code must find everything twice.



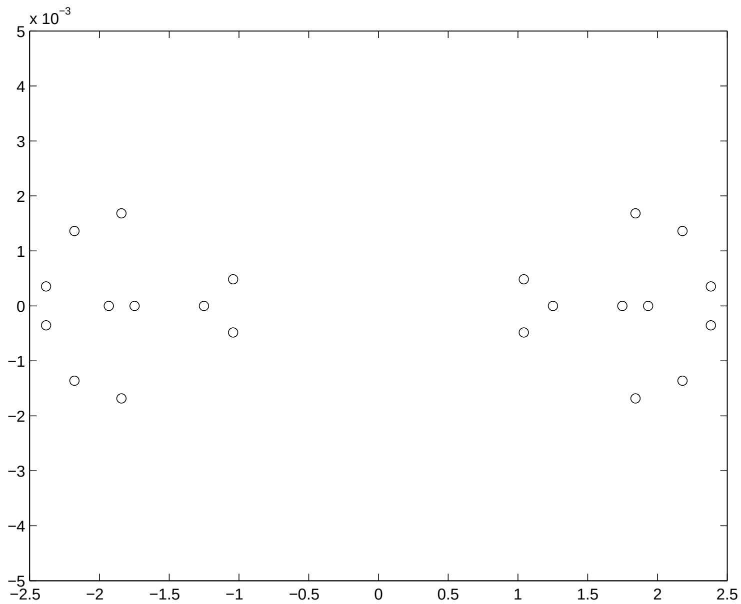
Conservative Shift ($\tau = 0.5$)

Method	Solves	Eigvals Found	Max. Resid.
Hamiltonian(1)	78	9	2×10^{-10}
Unstructured	138	7 + 2	3×10^{-5}
Hamiltonian(3)	174	11	3×10^{-13}
Symplectic	156	11	2×10^{-8}



Aggressive Shift ($\tau = 1.5$)

Method	Solves	Eigvals Found	Max. Resid.
Hamiltonian(1)	78	9	2×10^{-10}
Unstructured	96	9	1×10^{-7}
Hamiltonian(3)	120	9	2×10^{-12}
Symplectic	156	11	2×10^{-11}



The Last Slide

- We have developed structure-preserving implicitly-restarted Lanczos methods for Hamiltonian and symplectic eigenvalue problems.

- The structure-preserving methods are more accurate than a comparable non-structured method.

- By exploiting structure we can solve our problems more economically.