Numerical methods for nonlinear eigenvalue problems

Daniel Kressner
Seminar for applied mathematics
ETH Zürich

kressner@math.ethz.ch
`http://www.math.ethz.ch/~kressner`

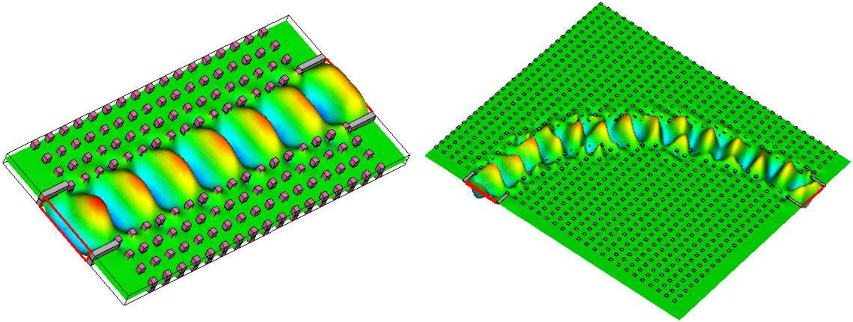
29.06.09, Karlsruhe

Outline

- Motivating application
  - Photonic band gap structure calculation
- Structured quadratic eigenvalue problems
  - Linearizations and structured linearizations
  - Structure-preserving Krylov subspace methods
- Genuinely nonlinear eigenvalue problems
  - Dealing with several eigenvalues
  - A block Newton method
  - Continuation of eigenvalues
  - Future work

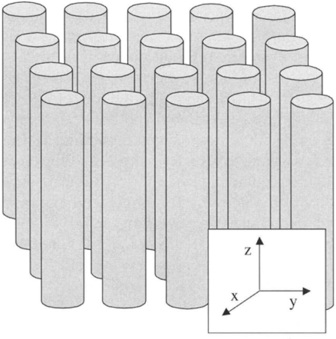
Motivating Application

Photonic crystals



- photonic crystal = lattice of mixed dielectric media (e.g., air and silicon)
- control light by designing media that prohibits propagation of electromagnetic waves in certain frequency range
- complete photonic band gap = frequency range with no propagation of electromagnetic waves of *any* polarization travelling in *any* direction.

2D periodic crystal



- material periodic along x- and y-direction; homogeneous along z-direction
- consider only electromagnetic waves with propagation in xy-plane

Mathematical model

Time-harmonic modes of electronicmagnetic wave (E, H) (E electric field, H magnetic field) decompose:

- transverse electric (TE) polarized modes $(E_x, E_y, 0, 0, 0, H_z)$
- transverse magnetic (TM) polarized modes $(0, 0, E_z, H_x, H_y, 0)$

Macroscopic Maxwell equations $\rightsquigarrow$ scalar equation for E_z of TM-mode at frequency ω:

$$-\Delta E_z = \omega^2 \varepsilon(r, \omega) E_z,$$

where $r = (x, y)$ and ε denotes relative permittivity.

Mathematical model

Time-harmonic modes of electronicmagnetic wave (E, H) (E electric field, H magnetic field) decompose:

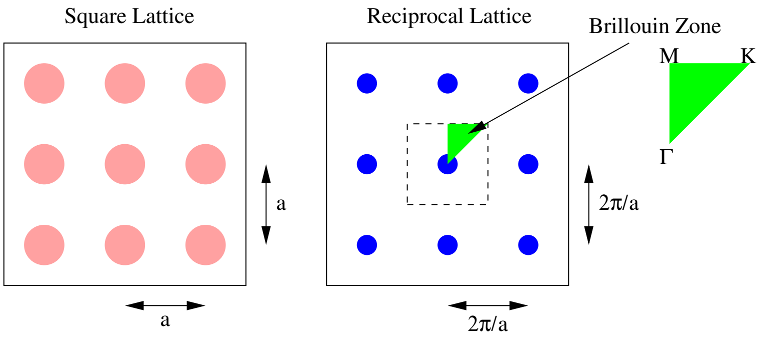
- transverse electric (TE) polarized modes $(E_x, E_y, 0, 0, 0, H_z)$
- transverse magnetic (TM) polarized modes $(0, 0, E_z, H_x, H_y, 0)$

Macroscopic Maxwell equations $\rightsquigarrow$ scalar equation for E_z of TM-mode at frequency ω:

$$-\Delta E_z = \omega^2 \varepsilon(r, \omega) E_z,$$

where $r = (x, y)$ and ε denotes relative permittivity.

Bloch solutions



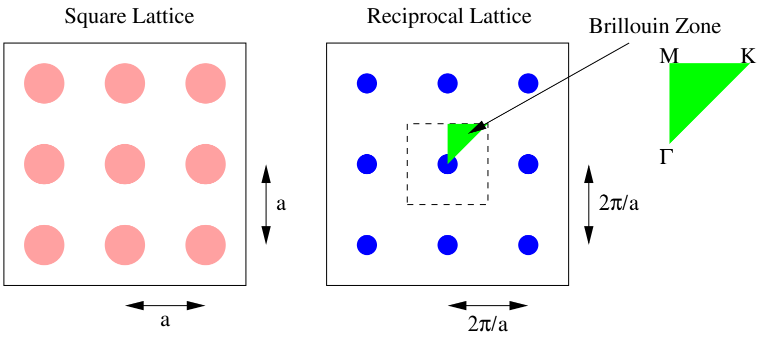
By Bloch's theorem, E_z takes the form

$$E_z(r) = e^{ik\cdot r}u(r),$$

where k is a wave vector $\in$ Brillouin zone, $u(r)$ periodic on lattice $\rightsquigarrow$

$$-(\nabla + ik)\cdot(\nabla + ik)u(r) = \omega^2\varepsilon(r,\omega)u(r)$$

Bloch solutions



By Bloch's theorem, E_z takes the form

$$E_z(r) = e^{ik \cdot r} u(r),$$

where k is a wave vector $\in$ Brillouin zone, $u(r)$ periodic on lattice $\rightsquigarrow$

$$-(\nabla + ik) \cdot (\nabla + ik) u(r) = \omega^2 \varepsilon(r, \omega) u(r)$$

Finding band gaps

The goal is to find frequency ranges $[\omega_{\text{low}}, \omega_{\text{high}}]$ for which

$$-(\nabla + \mathrm{i}k) \cdot (\nabla + \mathrm{i}k)u(r) = \omega^2 \varepsilon(r, \omega)u(r) \qquad (1)$$

has *no* solution.

Two approaches:

1. Fix frequency ω and wave vector direction $\hat{k}$ $\rightsquigarrow$ quadratic eigenvalue problem in wave vector length $\lambda = \|k\|$.
   FE discretization of (1) $\rightsquigarrow$

   $$(A_0 + \lambda A_1 + \lambda^2 A_2)x = 0, \quad x \neq 0.$$

   with A_0, A_2 real symmetric and $\sqrt{-1}A_1$ real skew-symmetric.

2. Fix wave vector k $\rightsquigarrow$ nonlinear eigenvalue problem in ω.
   FE discretization of (1) $\rightsquigarrow$

   $$\big(- \omega^2 \varepsilon_1(\omega)A_1 - \omega^2 \varepsilon_2(\omega)A_2 + A_3\big)x = 0, \quad x \neq 0.$$

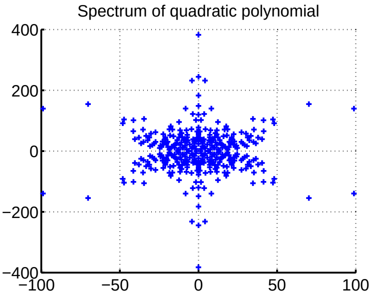
   with A_1, A_2 Hermitian pos semidef, A_3 Hermitian pos def.

Quadratic Eigenvalue Problems

Setting

$$(A_0 + \lambda A_1 + \lambda^2 A_2)x = 0, \quad x \neq 0.$$

with A_0, A_2 real symmetric and $\sqrt{-1}A_1$ real skew-symmetric.

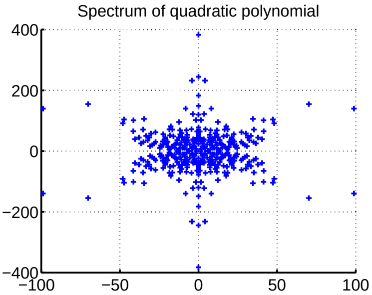


Only nonnegative real eigenvalues of interest!

Setting

$$(A_0 + \lambda A_1 + \lambda^2 A_2)x = 0, \quad x \neq 0.$$

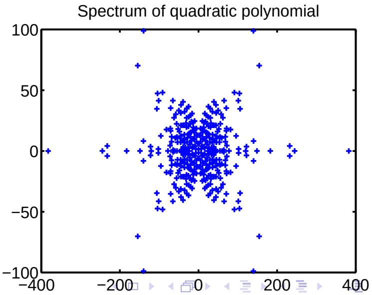
with A_0, A_2 real symmetric and $\sqrt{-1}A_1$ real skew-symmetric.



Substituting $\lambda \to \sqrt{-1}\lambda$ yields

$$(\tilde{A}_0 + \lambda \tilde{A}_1 + \lambda^2 \tilde{A}_2)x = 0, \quad x \neq 0.$$

with A_0, A_2 real symmetric and A_1 real skew-symmetric.



Standard approach to solving QEPs

1. Linearization. Introducing "velocity" $y = \lambda x$ the equation

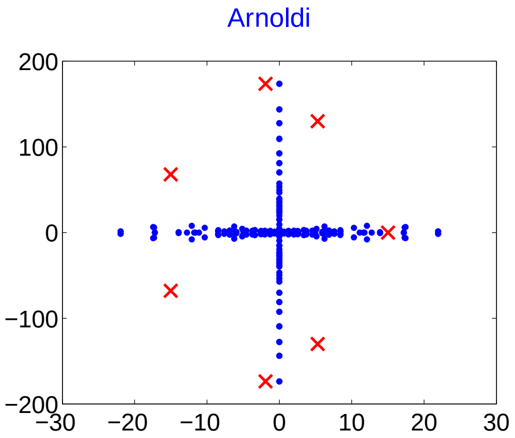
$$(A_0 + \lambda A_1 + \lambda^2 A_2)x = 0$$

   becomes a linear eigenvalue problem:

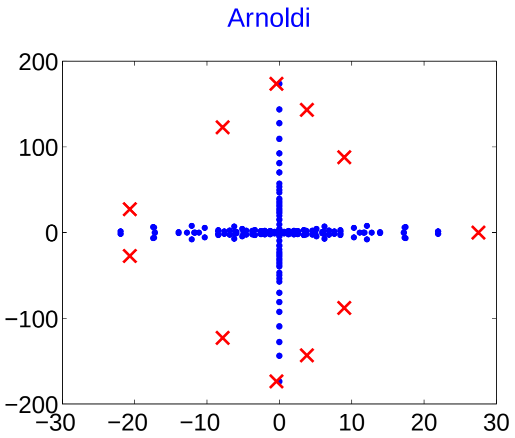
$$\left(\lambda \begin{bmatrix} 0 & A_2 \\ I & 0 \end{bmatrix} + \begin{bmatrix} A_0 & A_1 \\ -I & 0 \end{bmatrix}\right) \begin{bmatrix} x \\ \lambda x \end{bmatrix} = 0. \qquad (2)$$

2. Solution. Apply standard eigenvalue solver (e.g., a Krylov subspace method) to (2).

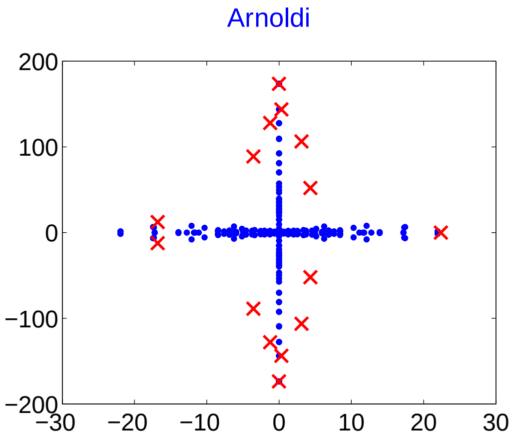
Krylov subspace method applied to linearization



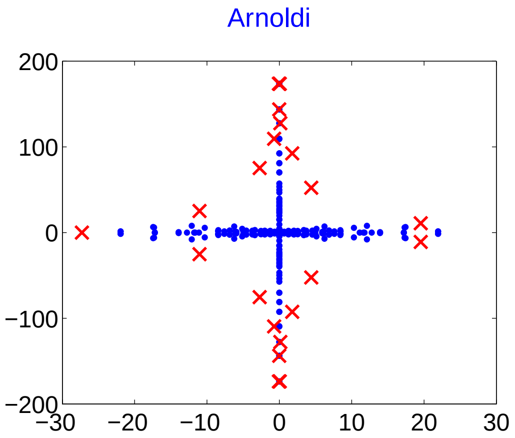
Krylov subspace method applied to linearization



Krylov subspace method applied to linearization



Krylov subspace method applied to linearization



- No preservation of spectral symmetries $\rightsquigarrow$ difficult to detect purely imaginary eigenvalues reliably.
- Fundamental problem: Linearization

$$\left(\lambda \begin{bmatrix} 0 & A_2 \\ I & 0 \end{bmatrix} + \begin{bmatrix} A_0 & A_1 \\ -I & 0 \end{bmatrix}\right) \begin{bmatrix} x \\ \lambda x \end{bmatrix} = 0.$$

  does not reflect matrix symmetries.
- Way out: Use more suitable linearization.

Rich class of linearizations described in

- D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Vector spaces of linearizations for matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 28(4):971–1004, 2006.
- D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Structured polynomial eigenvalue problems: good vibrations from good linearizations. *SIAM J. Matrix Anal. Appl.*, 28(4):1029–1051, 2006.

- No preservation of spectral symmetries $\rightsquigarrow$ difficult to detect purely imaginary eigenvalues reliably.
- Fundamental problem: Linearization

$$\left(\lambda \begin{bmatrix} 0 & A_2 \\ I & 0 \end{bmatrix} + \begin{bmatrix} A_0 & A_1 \\ -I & 0 \end{bmatrix}\right) \begin{bmatrix} x \\ \lambda x \end{bmatrix} = 0.$$

  does not reflect matrix symmetries.
- Way out: Use more suitable linearization.

Rich class of linearizations described in

- D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Vector spaces of linearizations for matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 28(4):971–1004, 2006.
- D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Structured polynomial eigenvalue problems: good vibrations from good linearizations. *SIAM J. Matrix Anal. Appl.*, 28(4):1029–1051, 2006.

Structured linearization

Use the following linearization:

$$\lambda \begin{bmatrix} A_1 & A_0 \\ -A_0 & 0 \end{bmatrix} + \begin{bmatrix} A_2 & 0 \\ 0 & A_0 \end{bmatrix} \tag{3}$$

This is a skew-symmetric/symmetric matrix pencil $\rightsquigarrow$ preserves spectral symmetries.

By congruence transformations: (4) $\Leftrightarrow$

$$\lambda \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} + \begin{bmatrix} \tilde{A}_2 & \tilde{A}_1 \\ \tilde{A}_1^T & I \end{bmatrix}, \quad \tilde{A}_2 = \tilde{A}_2^T$$

Structured linearization

Use the following linearization:

$$\lambda \begin{bmatrix} A_1 & A_0 \\ -A_0 & 0 \end{bmatrix} + \begin{bmatrix} A_2 & 0 \\ 0 & A_0 \end{bmatrix} \tag{3}$$

This is a skew-symmetric/symmetric matrix pencil $\rightsquigarrow$ preserves spectral symmetries.

By congruence transformations: (4) $\Leftrightarrow$

$$\lambda \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} + \begin{bmatrix} \tilde{A}_2 & \tilde{A}_1 \\ \tilde{A}_1^T & I \end{bmatrix}, \quad \tilde{A}_2 = \tilde{A}_2^T$$

Structured linearization

Use the following linearization:

$$\lambda \begin{bmatrix} A_1 & A_0 \\ -A_0 & 0 \end{bmatrix} + \begin{bmatrix} A_2 & 0 \\ 0 & A_0 \end{bmatrix} \tag{4}$$

This is a skew-symmetric/symmetric matrix pencil $\rightsquigarrow$ preserves spectral symmetries.

By congruence transformations: (4) $\Leftrightarrow$

$$\lambda \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} + \begin{bmatrix} -\tilde{A}_1^T & -I \\ \tilde{A}_2 & \tilde{A}_1 \end{bmatrix}, \quad \tilde{A}_2 = \tilde{A}_2^T$$

This is a Hamiltonian eigenvalue problem $\rightsquigarrow$ preserves spectral symmetries.

To exploit structure, apply Hamiltonian Arnoldi [DK'05] (Variant of SHIRA [Mehrmann/Watkins'01]).

Hamiltonian Arnoldi

Basic Ideas:

$$W = \begin{bmatrix} -\tilde{A}_1^T & -I \\ \tilde{A}_2 & \tilde{A}_1 \end{bmatrix}^2 = \begin{bmatrix} (A_1^T)^2 + A_2 & A_1 - A_1^T \\ A_1A_2 - A_2A_1^T & A_1^2 + A_2 \end{bmatrix}$$

is skew-Hamiltonian.

- Krylov subspace

$$\mathcal{K}_k(W, b) = \text{span}\{b, Wb, \ldots, W^{k-1}b\}$$

  is isotropic: $\mathcal{K}_k(W, b)^T \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \mathcal{K}_k(W, b) = 0.$

- Compression of W to an isotropic subspace ($U^T WU$ where cols of U span subspace) is again skew-Hamiltonian $\rightsquigarrow$ structure preservation.

SHIRA works explicitly with W; Hamiltonian Arnoldi works with original Hamiltonian matrix.

Hamiltonian Arnoldi

Basic Ideas:

$$W = \begin{bmatrix} -\tilde{A}_1^T & -I \\ \tilde{A}_2 & \tilde{A}_1 \end{bmatrix}^2 = \begin{bmatrix} (A_1^T)^2 + A_2 & A_1 - A_1^T \\ A_1A_2 - A_2A_1^T & A_1^2 + A_2 \end{bmatrix}$$

is skew-Hamiltonian.

- Krylov subspace

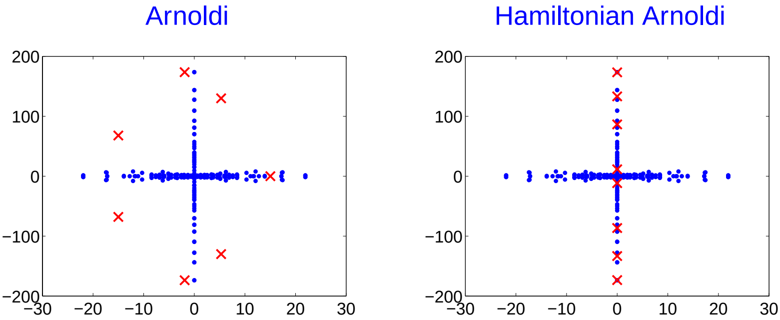
$$\mathcal{K}_k(W, b) = \text{span}\{b, Wb, \ldots, W^{k-1}b\}$$

  is isotropic: $\mathcal{K}_k(W, b)^T \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \mathcal{K}_k(W, b) = 0.$

- Compression of W to an isotropic subspace ($U^T WU$ where cols of U span subspace) is again skew-Hamiltonian $\rightsquigarrow$ structure preservation.

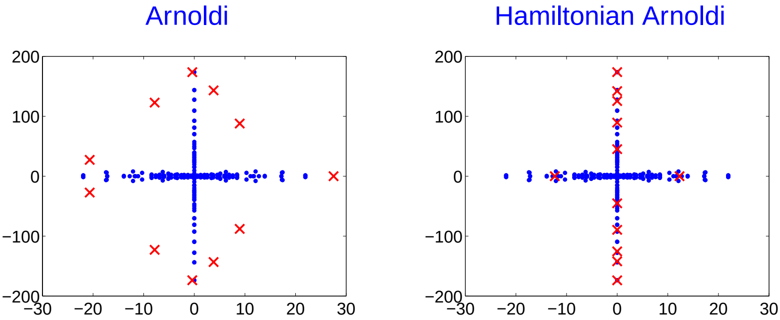
SHIRA works explicitly with W; Hamiltonian Arnoldi works with original Hamiltonian matrix.

Hamiltonian Arnoldi applied to structured linearization



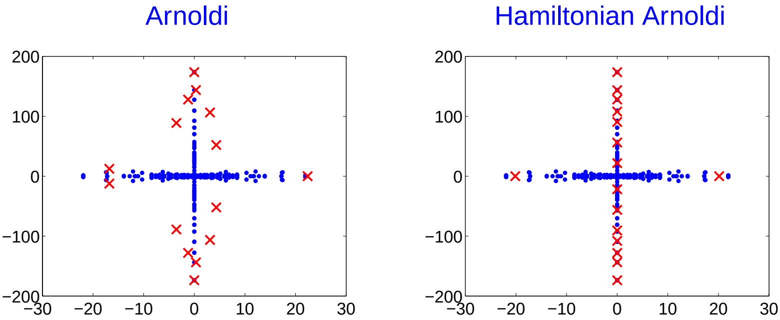
Structure-preserving Krylov subspace method preserves spectral symmetries.

Hamiltonian Arnoldi applied to structured linearization



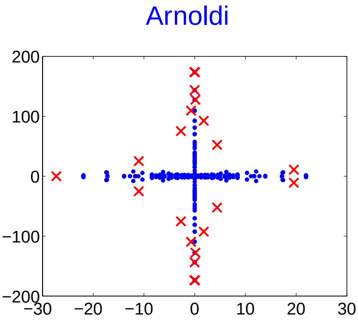
Structure-preserving Krylov subspace method preserves spectral symmetries.

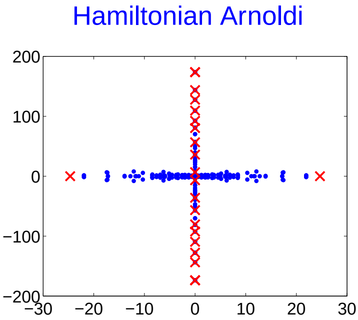
Hamiltonian Arnoldi applied to structured linearization



Structure-preserving Krylov subspace method preserves spectral symmetries.

Hamiltonian Arnoldi applied to structured linearization





Structure-preserving Krylov subspace method preserves spectral symmetries.

Summary

Use of quadratic eigenvalue problems in electronic band structure calculation fairly new and suggested in [Leminger'02; Engström/Richter'08]:

- Fix frequency ω and direction $\hat{k}$ of $k = \lambda \hat{k}$.
- For each (ω, k_{dir}), need to check $\nexists$ *purely imaginary eigenvalues* of quadratic EVP

$$(\tilde{A}_0 + \lambda \tilde{A}_1 + \lambda^2 \tilde{A}_2)x = 0, \quad x \neq 0.$$

with $\tilde{A}_0, \tilde{A}_2$ real symmetric, and $\tilde{A}_1$ real skew-symmetric.

+ always quadratic EVP independent of nature of permittivity
+ T-even polynomial, nice spectral structures
- large sample space
- finding all eigenvalues on imaginary axis is expensive
- unclear how to use continuation, eigenvalues might creep from anywhere into the imaginary axis

Nonlinear Eigenvalue Problems

Setting

- Fix wave vector k.
- For each k, need to solve nonlinear EVP

$$T(\lambda)x := \big(-\omega^2\varepsilon_1(\omega)A_1 - \omega^2\varepsilon_2(\omega)A_2 + A_3\big)x = 0, \quad x \neq 0.$$

- A_1, A_2 Hermitian pos semidef, A_3 Hermitian pos def.
- $\varepsilon_j(\omega)$ chosen by engineering intuition
- Lossless material $\rightsquigarrow \omega \in \mathbb{R}$.
- Usually only lowest frequencies of interest

Nonlinear eigenvalue problems

Consider nonlinear eigenvalue problems (NLEVPs)

$$T(\lambda)x := \big(f_1(\lambda)A_1 + f_2(\lambda)A_2 + \cdots + f_m(\lambda)A_m\big)x = 0, \quad x \neq 0$$

with $A_1, \ldots, A_m \in \mathbb{C}^{n\times n}$, analytic functions $f_1, \ldots, f_m : \Omega \to \mathbb{C}$.

Simple examples:

- Linear eigenvalue problems:

$$f_1(\lambda) = 1,\ f_2(\lambda) = -\lambda.$$

- Polynomial eigenvalue problems:

$$f_1(\lambda) = 1,\ f_2(\lambda) = \lambda,\ \ldots,\ f_m(\lambda) = \lambda^{m-1}.$$

Nonlinear eigenvalue problems

Consider nonlinear eigenvalue problems (NLEVPs)

$$T(\lambda)x := \big(f_1(\lambda)A_1 + f_2(\lambda)A_2 + \cdots + f_m(\lambda)A_m\big)x = 0, \quad x \neq 0$$

with $A_1, \ldots, A_m \in \mathbb{C}^{n\times n}$, analytic functions $f_1, \ldots, f_m : \Omega \to \mathbb{C}$.

Simple examples:

- Linear eigenvalue problems:

$$f_1(\lambda) = 1,\ f_2(\lambda) = -\lambda.$$

- Polynomial eigenvalue problems:

$$f_1(\lambda) = 1,\ f_2(\lambda) = \lambda,\ \ldots,\ f_m(\lambda) = \lambda^{m-1}.$$

Numerical methods for one eigenvalue

- Most obvious: Apply Newton to

$$T(\lambda)x = 0, \quad w^H x = 1$$

  for some normalization vector $w \in \mathbb{C}^n$. Requires solution of linear system with varying system matrix in each iteration ($T(\lambda)^{-1}T'(\lambda)v$).
- Variant: Neumaier's (1985) residual inverse iteration chooses fixed shift σ and requires solution of linear system with fixed system matrix ($T(\sigma)^{-1}T'(\lambda)v$)
- Subspace acceleration by Ruhe (1973), Hager and Wiberg (2000), Voss (2003).
- See [Mehrmann and Voss (2004), Nonlinear eigenvalue problems: A challenge for modern eigenvalue methods, GAMM Mitteilungen] for an overview of methods aimed at computing individual eigenvalues.

Dealing with several eigenvalues

For simplicity, assume $m = 2$.
Let λ_1, λ_2 be eigenvalues with eigenvectors x_1, x_2:

$$\begin{aligned}\big(f(\lambda_1)A_1 + f(\lambda_1)A_2\big)x_1 &= 0\\ \big(f(\lambda_2)A_1 + f(\lambda_2)A_2\big)x_2 &= 0\end{aligned}$$

Rearranging terms...

$$\begin{aligned}A_1x_1f(\lambda_1) + A_2x_1f(\lambda_1) &= 0\\ A_1x_2f(\lambda_2) + A_2x_2f(\lambda_2) &= 0\end{aligned}$$

...and merging both equations...

$$A_1\left[x_1, x_2\right]\begin{bmatrix} f(\lambda_1) & 0 \\ 0 & f(\lambda_2) \end{bmatrix} + A_2\left[x_1, x_2\right]\begin{bmatrix} f(\lambda_1) & 0 \\ 0 & f(\lambda_2) \end{bmatrix} = 0$$

Set $X = \left[x_1, x_2\right]$, $S = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ $\rightsquigarrow$ $A_1X\,f_1(S) + A_2X\,f_2(S) = 0$.

Dealing with several eigenvalues

For simplicity, assume $m = 2$.
Let λ_1, λ_2 be eigenvalues with eigenvectors x_1, x_2:

$$\begin{aligned}\big(f(\lambda_1)A_1 + f(\lambda_1)A_2\big)x_1 &= 0\\ \big(f(\lambda_2)A_1 + f(\lambda_2)A_2\big)x_2 &= 0\end{aligned}$$

Rearranging terms...

$$\begin{aligned}A_1x_1f(\lambda_1) + A_2x_1f(\lambda_1) &= 0\\ A_1x_2f(\lambda_2) + A_2x_2f(\lambda_2) &= 0\end{aligned}$$

...and merging both equations...

$$A_1\left[x_1,\, x_2\right]\begin{bmatrix} f(\lambda_1) & 0 \\ 0 & f(\lambda_2)\end{bmatrix} + A_2\left[x_1,\, x_2\right]\begin{bmatrix} f(\lambda_1) & 0 \\ 0 & f(\lambda_2)\end{bmatrix} = 0$$

Set $X = \left[x_1,\, x_2\right]$, $S = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ $\rightsquigarrow$ $A_1X\, f_1(S) + A_2X\, f_2(S) = 0$.

Invariant pairs

$(X, S) \in \mathbb{C}^{n\times k} \times \mathbb{C}^{k\times k}$ is called an invariant pair if

$$A_1 X f_1(S) + A_2 X f_2(S) + \cdots + A_m X f_m(S) = 0$$

Remarks:

- For linear eigenvalue problems: $A_1 X - XS = 0 \rightsquigarrow$ span(X) is invariant subspace belonging to $\Lambda(S)$.
- Theory in [Gohberg/Lancaster/Rodman'82] for $k = n$ and polynomial eigenvalue problems.
- For arbitrary k introduced and analyzed in [Beyn/Thümmler'08] for quadratic EVPs with invertible A_1.
- Extended to general polynomial EVPs in [Betcke/K.'09] and to nonlinear EVPs in [K.'09].

Avoiding degeneracies

Require extra conditions on invariant pairs to avoid degenerate situations, such as $X = 0$.

Is rank$(X) = k$ a reasonable condition?

Example [Dennis/Traub/Weber'76]:

$$\begin{bmatrix} 0 & 12 \\ -2 & 14 \end{bmatrix} + \lambda \begin{bmatrix} -1 & -6 \\ 2 & -9 \end{bmatrix} + \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

The eigenvalues 3 and 4 share the *same* eigenvector $\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$.

No!

Avoiding degeneracies

Require extra conditions on invariant pairs to avoid degenerate situations, such as $X = 0$.

Is rank$(X) = k$ a reasonable condition?

Example [Dennis/Traub/Weber'76]:

$$\begin{bmatrix} 0 & 12 \\ -2 & 14 \end{bmatrix} + \lambda \begin{bmatrix} -1 & -6 \\ 2 & -9 \end{bmatrix} + \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

The eigenvalues 3 and 4 share the *same* eigenvector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

No!

Minimal invariant pairs

Invariant pair $(X, S) \in \mathbb{C}^{n\times k} \times \mathbb{C}^{k\times k}$ is called minimal (of index ℓ) if

$$V_\ell(X,S) := \begin{bmatrix} X \\ XS \\ \vdots \\ XS^{\ell-1} \end{bmatrix}$$

has full column rank.

For previous example:

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad S = \begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix}$$

Then

$$V_1(X,S) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad V_2(X,S) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 3 & 4 \\ 3 & 4 \end{bmatrix},$$

$V_2(X,S)$ has full column rank $\rightsquigarrow$ (X,S) is minimal.

Minimal invariant pairs

Invariant pair $(X, S) \in \mathbb{C}^{n \times k} \times \mathbb{C}^{k \times k}$ is called minimal (of index ℓ) if

$$V_\ell(X, S) := \begin{bmatrix} X \\ XS \\ \vdots \\ XS^{\ell-1} \end{bmatrix}$$

has full column rank.

For previous example:

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad S = \begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix}$$

Then

$$V_1(X, S) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad V_2(X, S) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 3 & 4 \\ 3 & 4 \end{bmatrix},$$

$V_2(X, S)$ has full column rank $\rightsquigarrow$ (X, S) is minimal.

Minimal invariant pairs

Fundamental properties (polynomial: [Beyn/Thümmler'08], [Betcke/K.'09]; nonlinear [K.'09]):

- For pairwise distinct eigenvalues $\lambda_1, \dots, \lambda_k$ with eigenvectors $x_1, \dots, x_k$,

$$(X, S) = \Big(\big[x_1, \ \dots, \ x_m \big], \mathrm{diag}\big(\lambda_1, \dots, \lambda_k\big) \Big)$$

  is minimal invariant.
- (X, S) minimal invariant $\rightsquigarrow$ $(XP, P^{-1}SP)$ minimal invariant.
- (X, S) minimal invariant $\rightsquigarrow$ eigenvalues of S are eigenvalues of NLEVP.
- It is always possible to choose $\ell \leq k$.
- Nonminimal pairs (X, S) can be reduced: $\exists$ minimal invariant pair $(\widetilde{X}, \widetilde{S})$ s.t. $\mathrm{span}(\widetilde{X}) = \mathrm{span}(X)$ and $\Lambda(\widetilde{S}) = \Lambda(S)$.

Newton method for invariant pairs

To develop Newton method for computing/continuing, need operator equations of which (X, S) is a regular value.

$$\mathbb{T}(X, S) = 0 \tag{5}$$

with

$$\begin{aligned} \mathbb{T} : \mathbb{C}^{n\times k} \times \mathbb{C}_{\Omega}^{k\times k} &\to \mathbb{C}^{n\times k}, \\ (X, S) &\mapsto XA_1 f_1(S) + \cdots + XA_m f_m(S), \end{aligned}$$

(5) not sufficient to characterize (X, S).

Normalization condition: Choose $W^H = V_\ell(X, S)^\dagger \rightsquigarrow$

$$\mathbb{V}(X, S) = 0 \tag{6}$$

$$\begin{aligned} \mathbb{V} : \mathbb{C}^{n\times k} \times \mathbb{C}_{\Omega}^{k\times k} &\to \mathbb{C}^{k\times k}, \\ (X, S) &\mapsto W^H V_\ell(X, S) - I_k. \end{aligned}$$

Newton method for invariant pairs

To develop Newton method for computing/continuing, need operator equations of which (X, S) is a regular value.

$$\mathbb{T}(X, S) = 0 \tag{5}$$

with

$$\begin{array}{rcl} \mathbb{T} : \mathbb{C}^{n\times k} \times \mathbb{C}^{k\times k}_{\Omega} & \to & \mathbb{C}^{n\times k}, \\ (X, S) & \mapsto & XA_1 f_1(S) + \cdots + XA_m f_m(S), \end{array}$$

(5) not sufficient to characterize (X, S).

Normalization condition: Choose $W^H = V_\ell(X, S)^\dagger \rightsquigarrow$

$$\mathbb{V}(X, S) = 0 \tag{6}$$

$$\begin{array}{rcl} \mathbb{V} : \mathbb{C}^{n\times k} \times \mathbb{C}^{k\times k}_{\Omega} & \to & \mathbb{C}^{k\times k}, \\ (X, S) & \mapsto & W^H V_\ell(X, S) - I_k. \end{array}$$

Newton method for invariant pairs

To develop Newton method for computing/continuing, need operator equations of which (X, S) is a regular value.

$$\mathbb{T}(X, S) = 0 \tag{5}$$

with

$$\begin{array}{rcl} \mathbb{T} : \mathbb{C}^{n\times k} \times \mathbb{C}^{k\times k}_{\Omega} & \to & \mathbb{C}^{n\times k}, \\ (X, S) & \mapsto & XA_1 f_1(S) + \cdots + XA_m f_m(S), \end{array}$$

(5) not sufficient to characterize (X, S).

Normalization condition: Choose $W^H = V_\ell(X, S)^\dagger \rightsquigarrow$

$$\mathbb{V}(X, S) = 0 \tag{6}$$

$$\begin{array}{rcl} \mathbb{V} : \mathbb{C}^{n\times k} \times \mathbb{C}^{k\times k}_{\Omega} & \to & \mathbb{C}^{k\times k}, \\ (X, S) & \mapsto & W^H V_\ell(X, S) - I_k. \end{array}$$

Linearizing $\mathbb{T}$ and $\mathbb{V}$

Fréchet derivatives of $\mathbb{T}$ and $\mathbb{V}$ at (X, S):

$$\mathbb{D}\mathbb{T} : (\triangle X, \triangle S) \mapsto \mathbb{T}(\triangle X, S) + \sum_{j=1}^{m} A_j X \, [\mathbb{D}f_j(S)](\triangle S),$$

$$\mathbb{D}\mathbb{V} : (\triangle X, \triangle S) \mapsto W_0^H \triangle X + \sum_{j=1}^{\ell-1} W_j^H \big(\triangle X S^j + X \, \mathbb{D}S^j(\triangle S)\big).$$

Note that the Fréchet derivative of f_j at S can be computed using [Mathias'96, Higham'08]

$$f_j\left(\begin{bmatrix} S & \triangle S \\ 0 & S \end{bmatrix}\right) = \begin{bmatrix} f_j(S) & [\mathbb{D}f_j(S)](\triangle S) \\ 0 & f_j(S) \end{bmatrix}.$$

Is $\mathbb{L} = (\mathbb{D}\mathbb{T}, \mathbb{D}\mathbb{V})$ invertible at a minimal invariant pair (X, S)?

Linearizing $\mathbb{T}$ and $\mathbb{V}$

Fréchet derivatives of $\mathbb{T}$ and $\mathbb{V}$ at (X, S):

$$\mathbb{DT} : (\triangle X, \triangle S) \mapsto \mathbb{T}(\triangle X, S) + \sum_{j=1}^{m} A_j X \, [\mathbb{D} f_j(S)](\triangle S),$$

$$\mathbb{DV} : (\triangle X, \triangle S) \mapsto W_0^H \triangle X + \sum_{j=1}^{\ell-1} W_j^H \big(\triangle X S^j + X \, \mathbb{D} S^j(\triangle S)\big).$$

Note that the Fréchet derivative of f_j at S can be computed using [Mathias'96, Higham'08]

$$f_j\left(\begin{bmatrix} S & \triangle S \\ 0 & S \end{bmatrix}\right) = \begin{bmatrix} f_j(S) & [\mathbb{D} f_j(S)](\triangle S) \\ 0 & f_j(S) \end{bmatrix}.$$

Is $\mathbb{L} = (\mathbb{DT}, \mathbb{DV})$ invertible at a minimal invariant pair (X, S)?

Theorem (K.'09)

Let (X, S) be minimal invariant. The "Jacobian" $\mathbb{L}$ of $(\mathbb{T}, \mathbb{V})$ at (X, S) is invertible if and only if (X, S) is simple.

Remarks:

1. (X, S) is called simple if the algebraic multiplicities of S match those of the NLEVP.
2. Theorem implies local quadratic convergence of Newton iteration applied to $\big(\mathbb{T}(X, S), \mathbb{V}(X, S)\big) = (0, 0)$.

Newton method for computing invariant pairs

Input: Initial pair (X_0, S_0).
Output: Approximate solution (X_{p+1}, S_{p+1}).

1: $p \leftarrow 0$, $W \leftarrow V_l(X_0, S_0)$
2: **repeat**
3:     Res $\leftarrow \mathbb{T}(X_p, S_p)$
4:     Solve linear matrix equation $\mathbb{L}_p(\triangle X, \triangle S) = (\text{Res}, 0)$.
5:     $\widetilde{X}_{p+1} \leftarrow X_p - \triangle X, \quad \widetilde{S}_{p+1} \leftarrow S_p - \triangle S$
6:     Compute compact QR decomposition $V_l(X_p, S_p) = WR$.
7:     $X_{p+1} \leftarrow \widetilde{X}_p R^{-1}, \quad S_{p+1} \leftarrow R\widetilde{S}_{p+1}R^{-1}$
8: **until** convergence

Remarks:

- If no good initial guess available, use variant of inverse iteration to create one.
- Add simple line search to enhance global convergence properties.
- Step 4 is very expensive, $\mathcal{O}(k^3(n+k)^3)$ flops to solve linear system! After a Schur decomposition of S_p, block lower triangular structure of $\mathbb{L}_p(\triangle X, \triangle S) \rightsquigarrow \mathcal{O}(k(n+k)^3)$ flops.

Newton method for computing invariant pairs

Input: Initial pair (X_0, S_0).
Output: Approximate solution (X_{p+1}, S_{p+1}).

1: $p \leftarrow 0$, $W \leftarrow V_l(X_0, S_0)$
2: **repeat**
3: $\quad \text{Res} \leftarrow \mathbb{T}(X_p, S_p)$
4: $\quad$ Solve linear matrix equation $\mathbb{L}_p(\triangle X, \triangle S) = (\text{Res}, 0)$.
5: $\quad \widetilde{X}_{p+1} \leftarrow X_p - \triangle X, \quad \widetilde{S}_{p+1} \leftarrow S_p - \triangle S$
6: $\quad$ Compute compact QR decomposition $V_l(X_p, S_p) = WR$.
7: $\quad X_{p+1} \leftarrow \widetilde{X}_p R^{-1}, \quad S_{p+1} \leftarrow R\widetilde{S}_{p+1}R^{-1}$
8: **until** convergence

Remarks:

- If no good initial guess available, use variant of inverse iteration to create one.
- Add simple line search to enhance global convergence properties.
- Step 4 is very expensive, $\mathcal{O}(k^3(n+k)^3)$ flops to solve linear system! After a Schur decomposition of S_p, block lower triangular structure of $\mathbb{L}_p(\triangle X, \triangle S) \rightsquigarrow \mathcal{O}(k(n+k)^3)$ flops.

Electronic band structure calculcation

Example setup:

- Photonic crystals with cylindrical holes of diameter 0.6 a.
- $\varepsilon_1(\omega) = 1 + \frac{5.34}{1-\omega^2}, \varepsilon_2 \equiv 1$
- FE discretization of

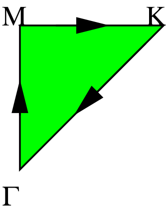
$$-(\nabla + \mathrm{i}k) \cdot (\nabla + \mathrm{i}k)u(r) = \omega^2 \varepsilon(r, \omega)u(r)$$

  on a conforming quadrilateral mesh with curvilinear quadrilateral cells. Polynomial basis functions of degree 7 using C++ library Concepts.[1]

Apply Newton method to solve resulting 980 $\times$ 980 NLEVP

$$-\omega^2 \varepsilon_1(\omega) A_1 - \omega^2 A_2(\omega) + A_3$$

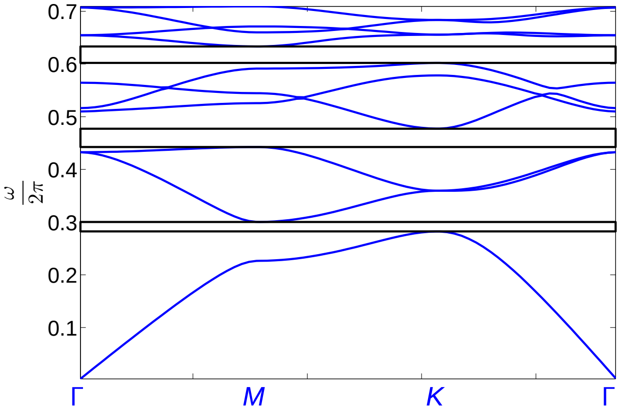
for wave vectors on boundary of Brillouin zone.

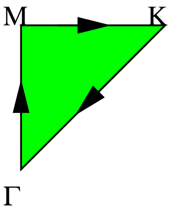


[1] Thanks to Holger Brandsmeier.

Electronic band structure calculcation

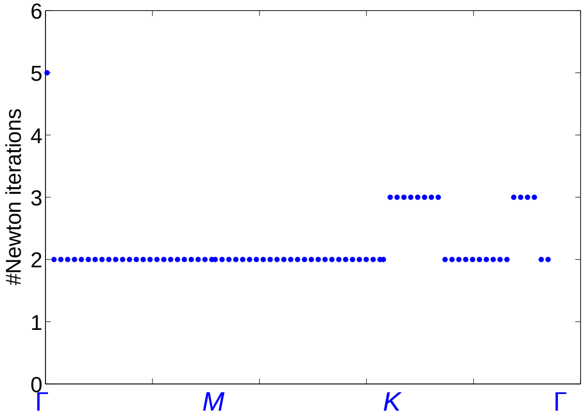
Trace 10 smallest frequencies for 75 equally distributed points on boundary of the Brillouin zone:

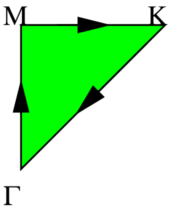




Newton iterations

The computed invariant pair for one wave vector is used to initialize the Newton method for the next wave vector.





Future work on NLEVP

Current implementation of block Newton method is

- rather expensive: requires solution of k linear systems in each iteration.
- not very robust: little control on location of converged eigenvalues (unless good initial guess available).

Aim: Development of less expensive and more methods.

- Lossless case: There usually exists a Rayleigh functional, i.e., there is a function $\rho : \mathbb{R}^n \to \mathbb{R}$

$$\frac{x^T T(\rho(x))x}{x^T x} = 0, \quad \forall x \in \mathbb{R}^n.$$

  $\rightsquigarrow$ Algorithms by Voss et al. can be applied.
  Under development: preconditioned inverse subspace iteration using a preconditioner for $T(0)$.
- Lossy case: Combination of invariant pairs with subspace expansion methods (e.g., nonlinear Jacobi-Davidson).

Selected References

Photonic crystals and QEPs:

- C. Engström and M. Richter. *On the spectrum of an operator pencil with applications to wave propagation in periodic and frequency dependent materials.* 2008, to appear (arXiv:0803.0456).

Polynomial/nonlinear EVPs:

- T. Betcke and D. Kressner. *Perturbation, computation and refinment of invariant pairs for matrix polynomials.* In preparation, 2009.
- W.-J. Beyn and V. Thümmler. *Continuation of invariant subspaces for parameterized quadratic eigenvalue problems.* Technical report, University of Bielefeld, Department of Mathematics, 2008.
- D. Kressner. *A block Newton method for nonlinear eigenvalue problems*. TR 2009-5, Seminar for applied mathematics, ETH Zurich, 2009.