Nonlinear Eigenvalue Problems: An Introduction

Cedric Effenberger

Seminar for Applied Mathematics
ETH Zurich

Pro*Doc Workshop
Disentis, August 18–21, 2010

Overview

Linear Eigenvalue Problems

Standard eigenvalue problem

Given $A \in \mathbb{C}^{n \times n}$, we seek

- eigenvalues $\lambda \in \mathbb{C}$, and
- eigenvectors $x \in \mathbb{C}^n \setminus \{0\}$

such that $Ax = \lambda x$.

Generalized eigenvalue problem

Given $A, B \in \mathbb{C}^{n \times n}$, we seek

- eigenvalues $\lambda \in \mathbb{C}$, and
- eigenvectors $x \in \mathbb{C}^n \setminus \{0\}$

such that $Ax = \lambda Bx$.

Generalization to nonlinear EVPs

Both problems can be reformulated as

$$T(\lambda)x = 0,$$

where

$$T(\lambda) = \begin{cases} A - \lambda I, & \text{for the standard EVP,} \\ A - \lambda B, & \text{for the generalized EVP.} \end{cases}$$

- $T(\lambda) = \lambda^d A_d + \cdots + \lambda A_1 + A_0 \quad \rightsquigarrow \quad$ polynomial EVP
- $T(\lambda) = f_1(\lambda)A_1 + \cdots + f_r(\lambda)A_r \quad \rightsquigarrow \quad$ nonlinear EVP

$A_j \in \mathbb{C}^{n \times n}$ constant matrices, $f_j : D \to \mathbb{C}$ analytic,
$D \subset \mathbb{C}$ open, connected

Free vibrations of mechanical systems with damping

Equations of motion

$$M\ddot{u} + C\dot{u} + Ku = 0$$

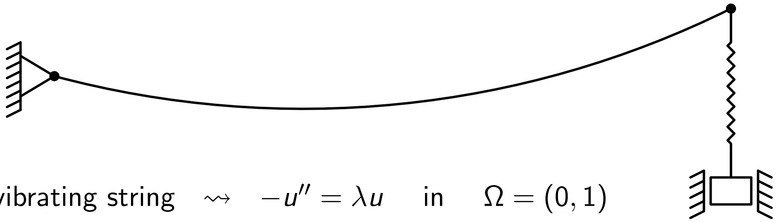
- M: mass matrix
- C: damping matrix
- K: stiffness matrix

- System of second-order, linear, homogeneous ODEs
- Eigenfrequencies ω of the mechanical system are eigenvalues of

$$(\omega^2 M + \omega C + K)u = 0.$$

$\rightsquigarrow$ Quadratic eigenvalue problem

Free vibrations of string with elastically attached mass



- vibrating string $\rightsquigarrow$ $-u'' = \lambda u$ in $\Omega = (0,1)$
- left end clamped $\rightsquigarrow$ $u(0) = 0$
- mass elastically attached to right end $\rightsquigarrow$ $u'(1) + \dfrac{\alpha\lambda}{\lambda - \alpha} u(1) = 0$

λ-dependent boundary condition!

Discretization with n piecewise linear FE leads to

$$\left[K + \frac{\alpha\lambda}{\lambda - \alpha} e_n e_n^T - \lambda M\right] x = 0.$$

Delay eigenvalue problems

Linear delay differential equation with r discrete delays

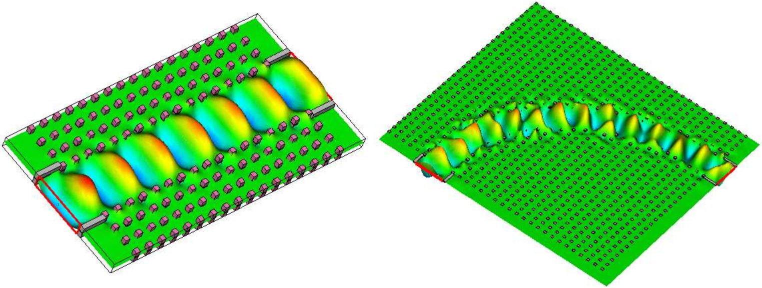
$$\dot{u}(t) = A_0 u(t) + A_1 u(t - \tau_1) + \cdots + A_r u(t - \tau_r)$$

- models influences which take effect only after some time
- famous example: Hot shower problem
- stability analysis involves nonlinear EVPs

Delay eigenvalue problem

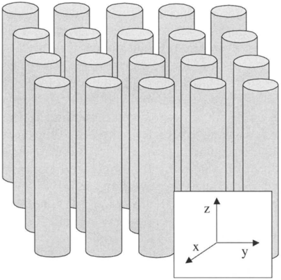
$$(-\lambda I + A_0 + \mathbf{e}^{-\tau_1 \lambda} A_1 + \cdots + \mathbf{e}^{-\tau_r \lambda} A_r) x = 0$$

Electronic bandstructure computations



- photonic crystal: lattice of mixed dielectric media
- control electromagnetic waves by designing the crystal such that it inhibits their propagation
- complete photonic band gap: frequency range with no propagation of electromagnetic waves of *any* polarization travelling in *any* direction

Electronic bandstructure computations II



- 2D crystal: periodic in x- and y-direction; homogeneous in z-direction
- consider only electromagnetic waves propagating in the xy-plane

Electronic bandstructure computations III

Time-harmonic modes of electromagnetic wave (E, H) can be decomposed into

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

For TM polarized modes, the macroscopic Maxwell Equations reduce to a scalar equation for E_z,

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

- r: spatial variable
- ω: frequency
- ε: relative permittivity

ω-dependent material parameter!

Electronic bandstructure computations IV

- Bloch ansatz: $E_z = \mathbf{e}^{\mathbf{i}k\cdot r} u(r)$

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

  - k: wave vector
  - u: periodic function on lattice

- assumption: lattice consists of 2 materials, one of which is air

$$\Omega = \Omega_1 \cup \Omega_2, \qquad \varepsilon(r,\omega) = \begin{cases} \varepsilon_1 = 1, & r \in \Omega_1 \\ \varepsilon_2(\omega), & r \in \Omega_2 \end{cases}$$

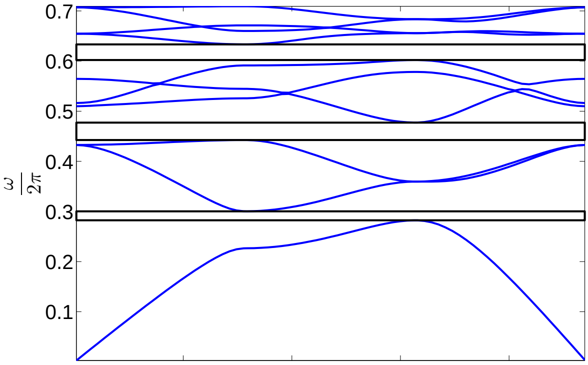
- discretization using discontinuous Galerkin with p-enhancement (Engström & Wang, 2010)

$$\left[G - \omega^2 M_1 - \omega^2 \varepsilon_2(\omega) M_2\right] u = 0$$

Electronic bandstructure computations V

- nonlinearity caused by ω-dependency of ε_2
- popular model for permittivity: Lorentz model

$$\varepsilon_2(\omega) = \alpha + \sum_{k=1}^{K} \frac{\xi_k}{\eta_k - \omega^2 - \mathbf{i}\gamma_k \omega} \quad \rightsquigarrow \quad \text{rational EVP}$$



Spectrum of nonlinear EVPs

Definition

Let $T : D \to \mathbb{C}^{n \times n}$ be holomorphic. We call

- $\rho(T) := \{\lambda \in D : T(\lambda) \text{ is invertible}\}$ the resolvent set of T,
- $\sigma(T) := D \setminus \rho(T)$ the spectrum of T.

Theorem

Either the resolvent set of T is empty or the spectrum of T consists of isolated eigenvalues.

The number of eigenvalues in $\sigma(T)$ may be infinite!

Newton-based methods

Two possibilities to apply Newton's method:

1. $\det T(\lambda) = 0$ (scalar equation in $\mathbb{C}$)
2. $T(\lambda)x = 0$ (vector equation in $\mathbb{C}^n$)

We focus on the second case.

- $x, \lambda \rightsquigarrow n + 1$ unknowns
- $T(\lambda)x = 0 \rightsquigarrow n$ constraints

We have to add one additional constraint.

$$v^H x = 1$$

Nonlinear inverse iteration

$$F(x, \lambda) := \begin{bmatrix} T(\lambda)x \\ v^H x - 1 \end{bmatrix}$$

Application of Newton's method

$$0 \stackrel{!}{=} F(x, \lambda) + \mathrm{D}F(x, \lambda) \begin{bmatrix} \hat{x} - x \\ \hat{\lambda} - \lambda \end{bmatrix} = \begin{bmatrix} T(\lambda)\hat{x} + (\hat{\lambda} - \lambda)\, T'(\lambda)x \\ v^H \hat{x} - 1 \end{bmatrix}$$

yields

$$\hat{x} = -(\hat{\lambda} - \lambda)\, T(\lambda)^{-1} T'(\lambda)x$$
$$\hat{\lambda} = \lambda - \frac{1}{v^H T(\lambda)^{-1} T'(\lambda)x}.$$

Nonlinear inverse iteration II

Algorithm

Input: approximate eigenpair (x_0, λ_0) with $v^H x_0 = 1$
for $j = 0, 1, \dots$ until convergence **do**
solve $T(\lambda_j)\tilde{x}_{j+1} = T'(\lambda_j)x_j$ for $\tilde{x}_{j+1}$
update $\lambda_{j+1} := \lambda_j - \frac{v^H x_j}{v^H \tilde{x}_{j+1}}$
normalize $x_{j+1} := \frac{1}{v^H u_{j+1}} \tilde{x}_{j+1}$
end for

- local quadratic convergence to simple eigenvalues
- main computational work lies in the solution of the linear system

Reducing the computational work

Replacing $T(\lambda_j)^{-1}$ by $T(\sigma)^{-1}$ for a fixed σ leads to misconvergence!

But ...

$$\begin{aligned}
\hat{x} &= -(\hat{\lambda} - \lambda) T(\lambda)^{-1} T'(\lambda) x \\
&= x - T(\lambda)^{-1} T(\lambda) x - (\hat{\lambda} - \lambda) T(\lambda)^{-1} T'(\lambda) x \\
&= x - T(\lambda)^{-1} \big[T(\lambda) + (\hat{\lambda} - \lambda) T'(\lambda) \big] x \\
&= x - T(\lambda)^{-1} T(\hat{\lambda}) x + \mathbf{O}\Big(|\hat{\lambda} - \lambda|^2 \Big)
\end{aligned}$$

- Now $T(\sigma)^{-1}$ can be used in place of $T(\lambda)^{-1}$ safely.
- Even inexact solution of the linear system is possible.

Residual inverse iteration

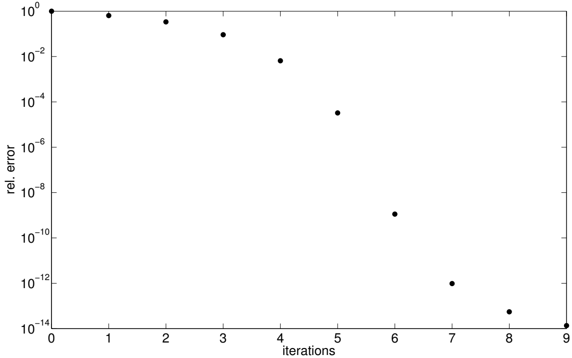
Algorithm

Input: approximate eigenpair (x_0, σ)
for $j = 0, 1, \ldots$ until convergence **do**
solve $v^H T(\sigma)^{-1} T(\lambda_{j+1}) x_j = 0$ for λ_{j+1}
solve $T(\sigma) \triangle x = -T(\lambda_{j+1}) x_j$ for $\triangle x$
update $\tilde{x}_{j+1} := x_j + \triangle x$
normalize $x_{j+1} := \frac{1}{v^H \tilde{x}_{j+1}} \tilde{x}_{j+1}$
end for

- local linear convergence to simple eigenvalues
- convergence speed dependent on distance from σ to eigenvalue

Convergence

- string with elastically attached mass
- smallest magnitude eigenvalue using residual inverse iteration
- initial guess: $\sigma = 0$, x discrete version of function $u(z) = z$



it.	λ
1	0.750000000000
2	0.611928913313 7
3	0.499216086848 5
4	0.460296260722 1
5	0.457333371895 5
6	0.457318489468 5
7	0.457318488954 6
8	0.457318488954 2
9	0.457318488954 2

Difficulties

- no global convergence
- can only handle one eigenvalue at a time
- cannot handle multiple eigenvalues
- consecutive runs may converge to the same eigenvalue

Linear eigensolvers (ARPACK, Jacobi-Davidson) exploit linear independence of eigenvectors to prevent reconvergence.

Example (loss of linear independence for NLEVPs)

The eigenvalues 1 and 2 of

$$\left(\begin{bmatrix}1 & 1\\ 2 & 0\end{bmatrix} + \lambda\begin{bmatrix}-1 & -2\\ 0 & -3\end{bmatrix} + \lambda^2\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\right)x = 0$$

share the same eigenvector $\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$.

Several eigenvalues in the linear case

$$Ax = \lambda x$$

Let (x_1, λ_1), (x_2, λ_2) be two eigenpairs:

$$\begin{aligned} Ax_1 &= \lambda_1 x_1, \\ Ax_2 &= \lambda_2 x_2. \end{aligned}$$

The above equations can be merged:

$$A \begin{bmatrix} x_1 & x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} x_1 & x_2 \end{bmatrix}}_{=:X} \underbrace{\begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix}}_{=:\Lambda}$$

Hence, we have

$$AX = X\Lambda.$$

$\rightsquigarrow$ X spans an invariant subspace of A.

Generalization to the nonlinear case

$$f_1(\lambda)A_1x + \cdots + f_r(\lambda)A_rx = 0$$

Assume $r = 2$ and let (x_1, λ_1), (x_2, λ_2) be two eigenpairs.

$$\begin{aligned} A_1x_1f_1(\lambda_1) + A_2x_1f_2(\lambda_1) &= 0, \\ A_1x_2f_1(\lambda_2) + A_2x_2f_2(\lambda_2) &= 0. \end{aligned}$$

These equations can again be merged

$$A_1 \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} f_1(\lambda_1) & \\ & f_1(\lambda_2) \end{bmatrix} + A_2 \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} f_2(\lambda_1) & \\ & f_2(\lambda_2) \end{bmatrix} = 0$$

With $X = \begin{bmatrix} x_1 & x_2 \end{bmatrix}$ and $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2)$ as before

$$A_1Xf_1(\Lambda) + A_2Xf_2(\Lambda) = 0.$$

Invariant pairs

Definition

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

$$A_1 X f_1(\Lambda) + \cdots + A_r X f_r(\Lambda) = 0.$$

- need to exclude degenerate situations, such as $X = 0$
- linear case: require X to have full column rank
- not suitable for the nonlinear case:

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \Lambda = \begin{bmatrix} 1 & \\ & 2 \end{bmatrix}$$

is a perfectly reasonable invariant pair of

$$\left(\begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} + \lambda \begin{bmatrix} -1 & -2 \\ 0 & -3 \end{bmatrix} + \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) x = 0.$$

Minimality

Definition

- An invariant pair (X, Λ) is called minimal if

$$\begin{bmatrix} X \\ X\Lambda \\ \vdots \\ X\Lambda^{\ell-1} \end{bmatrix}$$

  has full column rank for some integer ℓ.
- The smallest such ℓ is called the minimality index of (X, Λ).

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \Lambda = \begin{bmatrix} 1 & \\ & 2 \end{bmatrix} \quad \rightsquigarrow \quad \begin{bmatrix} X \\ X\Lambda \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 2 \\ 1 & 2 \end{bmatrix}$$

Minimal invariant pairs

$\rightsquigarrow$ (X, Λ) is a minimal invariant pair with minimality index 2.

Theorem

Let (X, Λ) be a minimal invariant pair of a nonlinear EVP

$$T(\lambda)x = 0.$$

Then every eigenvalue of Λ is an eigenvalue of T.

Computation of invariant pairs

Apply Newton's method to

$$\mathbf{T}(X,\Lambda) := A_1 X f_1(\Lambda) + \cdots + A_r X f_r(\Lambda) = 0.$$

- $X, \Lambda \quad \rightsquigarrow \quad n \cdot m + m^2$ unknowns
- $\mathbf{T}(X,\Lambda) = 0 \quad \rightsquigarrow \quad n \cdot m$ constraints

We have to impose m^2 additional constraints.

$$\mathbf{N}(X,\Lambda) := V^H \begin{bmatrix} X \\ X\Lambda \\ \vdots \\ X\Lambda^{\ell-1} \end{bmatrix} - I = 0, \qquad V \text{ suitably chosen}$$

Theorem

Let (X,Λ) be a minimal invariant pair of T. The Fréchet derivative $\mathrm{D}\mathbf{F}$ of $\mathbf{F} := \left[\begin{smallmatrix} \mathbf{T} \\ \mathbf{N} \end{smallmatrix}\right]$ at (X,Λ) is invertible iff the multiplicities of Λ's eigenvalues match those of T.

Ongoing work: Continuation of invariant pairs

Joint project with Daniel Kressner (SAM, ETHZ)
and Wolf-Jürgen Beyn (University of Bielefeld)

Parameter-dependent nonlinear EVP

$$T(\lambda, s)x = 0$$

- Goal: track several eigenvalues as the parameter s varies
- Idea:
  - use invariant pair $\big(X(s), \Lambda(s)\big)$ at parameter value s as an initial guess for invariant pair at $s + \triangle s$
  - apply Newton to obtain invariant pair $\big(X(s + \triangle s), \Lambda(s + \triangle s)\big)$ at parameter value $s + \triangle s$
  - repeat

Part II:
Solution of polynomial / rational EVPs

Linearization

Polynomial EVPs can be solved by transforming them into linear ones.

- Example: $(\lambda^2 M + \lambda C + K)x = 0$
- introduce auxiliary variable $y = \lambda x$
- rewrite as $\lambda M y + C y + K x = 0$

This can be written as a generalized linear EVP with the same eigenvalues:

$$\begin{bmatrix} -C & -K \\ I & 0 \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} = \lambda \begin{bmatrix} M & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix}.$$

- linearizations not unique
- could also rewrite as $\lambda M y + \lambda C x + K x = 0$

$$\rightsquigarrow \quad \begin{bmatrix} K & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} -C & -M \\ I & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Linearization II

There are entire vector spaces of linearizations (Mackey et al., 2006)!

- choice of linearization depends on underlying application
- structure preservation of interest

Example

The eigenvalues of an even polynomial eigenvalue problem

$$(\lambda^2 M + \lambda C + K)x = 0, \quad M = M^T, C = -C^T, K = K^T$$

occur in pairs $(\lambda, -\lambda)$.

- If M is invertible, the symmetric / skew-symmetric linearization

$$\begin{bmatrix} K & 0 \\ 0 & M \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} -C & -M \\ M & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

preserves this property.

Linearization III

- linearization of polynomial EVPs of degree $d > 2$ analogously by introducing the auxiliary vector

$$\begin{bmatrix} x \\ \lambda x \\ \vdots \\ \lambda^{d-1} x \end{bmatrix}$$

Conclusion 1:

- polynomial EVP of size n $\rightsquigarrow$ linearized EVP of size dn
- polynomial EVP has exactly dn eigenvalues (counting multiplicities)

Conclusion 2:

- The extended vectors above will always be linearly independent, even though the eigenvectors x themselves may not.

Solution of rational EVPs

- most straightforward way: multiplication by the common denominator of all rational terms $\rightsquigarrow$ polynomial EVP
- degree of resulting polynomial EVP may be very large
- sometimes smarter ways to solve rational EVPs:

Ongoing joint project with Daniel Kressner and Christian Engström (both SAM, ETHZ)

$$\left[G - \omega^2 M_1 - \omega^2 \left(\alpha + \frac{\xi}{\eta - \omega^2 - \mathbf{i}\gamma\omega}\right) M_2\right] u = 0$$

- $M_1 + M_2$ is a splitting of the total mass matrix of the problem.

Rational EVPs in electronic bandstructure computations

- by suitable ordering of nodes:

$$M_2 = \begin{bmatrix} 0 & \\ & \check{M}_2 \end{bmatrix} \in \mathbb{R}^{n \times n}$$

with a positive definite $\check{M}_2 \in \mathbb{R}^{m \times m}$, $m \ll n$

- by Cholesky decomposition:

$$M_2 = F^T F, \quad F \in \mathbb{R}^{m \times n}$$

Idea (Bai & Su, 2008):

- write rational term(s) in transfer function form

$$\frac{\omega^2 \xi}{\eta - \omega^2 - \mathbf{i}\gamma\omega} = -\xi + b^T (A - \omega E)^{-1} b$$

Conversion to transfer function form

- by partial fraction expansion

$$
\begin{aligned}
\frac{p(\omega)}{q(\omega)} &= \zeta + \frac{\sigma_1}{\omega - \rho_1} + \cdots + \frac{\sigma_k}{\omega - \rho_k} \\
&= \zeta + \left(\frac{\omega}{\sigma_1} - \frac{\rho_1}{\sigma_1}\right)^{-1} + \cdots + \left(\frac{\omega}{\sigma_k} - \frac{\rho_k}{\sigma_k}\right)^{-1} \\
&= \zeta + b^T (A - \omega E)^{-1} b,
\end{aligned}
$$

where

$$
b = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, \quad A = \begin{bmatrix} -\frac{\rho_1}{\sigma_1} & & \\ & \ddots & \\ & & -\frac{\rho_k}{\sigma_k} \end{bmatrix}, \quad E = \begin{bmatrix} -\frac{1}{\sigma_1} & & \\ & \ddots & \\ & & -\frac{1}{\sigma_k} \end{bmatrix}
$$

Linearization of rational EVPs

- EVP with rational term in transfer function form:

$$\hat{G}u - b^T(A-\omega E)^{-1}bF^TFu = \omega^2\hat{M}u,$$

where $\quad \hat{G} = G + \xi M_2, \quad \hat{M} = M_1 + \alpha M_2$

- rearrange:

$$\begin{aligned} b^T(A-\omega E)bF^TF &= b^T(A-\omega E)^{-1}b \otimes F^TF \\ &= \mathcal{B}^T(\mathcal{A}-\omega\mathcal{E})^{-1}\mathcal{B} \end{aligned}$$

where $\quad \mathcal{B} = b \otimes F, \quad \mathcal{A} = A \otimes I, \quad \mathcal{E} = E \otimes I$

Linearization of rational EVPs II

$$\hat{G}u - \mathcal{B}^T(\mathcal{A} - \omega\mathcal{E})^{-1}\mathcal{B}u = \omega^2\hat{M}u$$

Introducing the auxiliary variables $v := -(\mathcal{A} - \omega\mathcal{E})^{-1}\mathcal{B}u, \quad w := \omega u,$ we obtain the linearized EVP

$$\begin{bmatrix} \hat{G} & \mathcal{B}^T & \\ \mathcal{B} & \mathcal{A} & \\ & & I \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \omega \begin{bmatrix} & & \hat{M} \\ & \mathcal{E} & \\ I & & \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix}$$

problem size		standard lin.		rational lin.	
p	#DoF	size	comp. time	size	comp. time
2	288	1728	262 s	720	78 s
4	720	4320	819 s	1800	321 s
6	1344	8064	1547 s	3360	727 s
8	2160	12960	–	5400	1688 s

Summary

- introduced nonlinear EVPs
- showed several sample applications
- discussed Newton-based technique for determining one or several eigenpairs
- demonstrated linearization techniques for polynomial and rational EVPs
- outlined two ongoing projects