Nonlinear Eigenvalue Problems: Theory and Applications

David Bindel[1]

Department of Computer Science
Cornell University

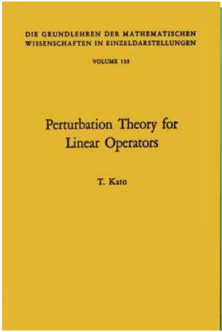
29 January 2016

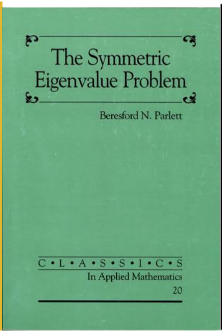
[1] Joint work with Amanda Hood

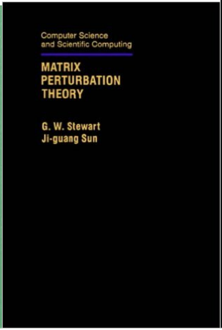
Why eigenvalues?

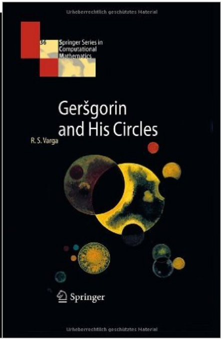
Why eigenvalues?

- The polynomial connection
- The optimization connection
- The approximation connection
- The Fourier/quadrature/special function connection
- The **dynamics** connection









Vibrations are everywhere, and so too are the eigenvalues associated with them. As mathematical models invade more and more disciplines, we can anticipate a demand for eigenvalue calculations in an ever richer variety of contexts.

– *Beresford Parlett,* The Symmetric Eigenvalue Problem

Why eigenvalues?

As a student in a first ODE course:

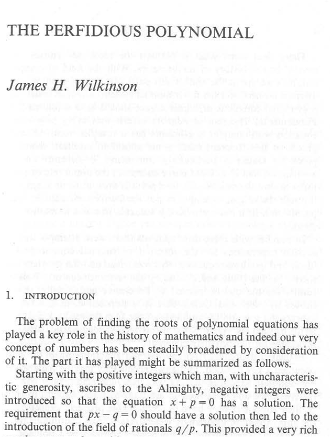
$$y' - Ay = 0 \xrightarrow{y=e^{\lambda t}v} (\lambda I - A)v = 0.$$

Me: “How do I compute this?”

$$p(\lambda) = \det(\lambda I - A) = 0.$$

And then I learned better...

Eigenvalues or polynomial roots?



THE PERFIDIOUS POLYNOMIAL

James H. Wilkinson

1. INTRODUCTION

The problem of finding the roots of polynomial equations has played a key role in the history of mathematics and indeed our very concept of numbers has been steadily broadened by consideration of it. The part it has played might be summarized as follows.

Starting with the positive integers which man, with uncharacteristic generosity, ascribes to the Almighty, negative integers were introduced so that the equation $x + p = 0$ has a solution. The requirement that $px - q = 0$ should have a solution then led to the introduction of the field of rationals q/p. This provided a very rich

The cosy relationship that mathematicians enjoyed with polynomials suffered a severe setback in the early fifties when electronic computers came into general use. Speaking for myself I regard it as the most traumatic experience in my career as a numerical analyst.

— J. H. Wilkinson

Representation matters.

Why nonlinear eigenvalues?

$$y' - Ay = 0 \xrightarrow{y=e^{\lambda t}v} (\lambda I - A)v = 0$$

$$y'' + By' + Ky = 0 \xrightarrow{y=e^{\lambda t}v} (\lambda^2 I + \lambda B + K)v = 0$$

$$y' - Ay - By(t-1) = 0 \xrightarrow{y=e^{\lambda t}v} (\lambda I - A - Be^{-\lambda})v = 0$$

$$T(d/dt)y = 0 \xrightarrow{y=e^{\lambda t}v} T(\lambda)v = 0$$

- Higher-order ODEs
- Dynamic element formulations
- Delay differential equations
- Boundary integral equation eigenproblems
- Radiation boundary conditions

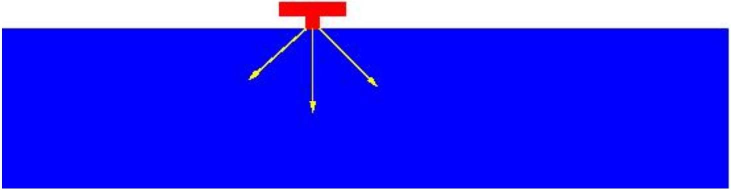
Big and little

$$y'' + By' + Ky = 0 \xrightarrow{v=y'} \begin{bmatrix} v \\ y \end{bmatrix}' + \begin{bmatrix} B & K \\ -I & 0 \end{bmatrix} \begin{bmatrix} v \\ y \end{bmatrix} = 0$$

$$(\lambda^2 I + \lambda B + K)u = 0 \xrightarrow{v=\lambda u} \left(\lambda I + \begin{bmatrix} B & K \\ -I & 0 \end{bmatrix} \right) \begin{bmatrix} v \\ u \end{bmatrix} = 0$$

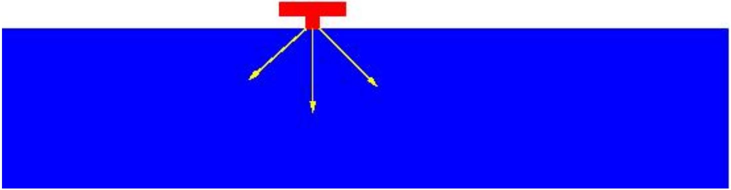
Tradeoff: more variables = more linear.

My motivation



$$T(\omega)v \equiv \left(K - \omega^2 M + G(\omega)\right) v = 0$$

My motivation



$$T(\omega)v \equiv \left(K - \omega^2 M + G(\omega)\right) v = 0$$

Wanted: Perturbation theory justifying a terrible estimate of $G(\omega)$

The big picture

$$T(\lambda)v = 0, \quad v \neq 0.$$

where

- $T : \Omega \to \mathbb{C}^{n \times n}$ analytic, $\Omega \subset \mathbb{C}$ simply connected
- Regularity: $\det(T) \not\equiv 0$

Nonlinear spectrum: $\Lambda(T) = \{z \in \Omega : T(z) \text{ singular}\}$.

What do we want?

- Qualitative information (e.g. no eigenvalues in right half plane)
- Error bounds on computed/estimated eigenvalues
- Assurances that we know all the eigenvalues in some region

Solver strategies

$$T(\lambda)v = 0, \quad v \neq 0.$$

A common approach:

1. **Approximate** T locally (linear, rational, etc)
2. Solve **approximate** problem.
3. Repeat as needed.

How should we choose solver parameters? What about global behavior? What can we trust? What might we miss?

Analyticity to the rescue

$$T(\lambda)v = 0, \quad v \neq 0.$$

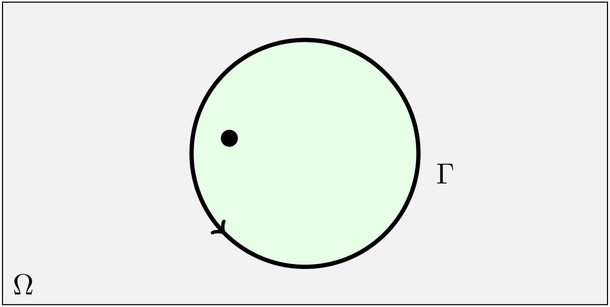
where

- $T : \Omega \to \mathbb{C}^{n \times n}$ **analytic**, $\Omega \subset \mathbb{C}$ simply connected
- Regularity: $\det(T) \not\equiv 0$

Nonlinear spectrum: $\Lambda(T) = \{z \in \Omega : T(z)$ singular$\}$.

Goal: Use analyticity to *compare* and to *count*

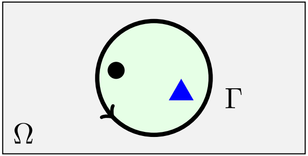
Winding and Cauchy's argument principle

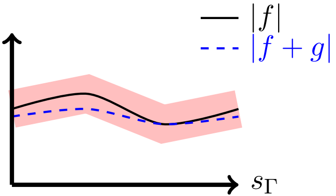


$$W_\Gamma(f) = \frac{1}{2\pi i}\int_\Gamma \frac{f'(z)}{f(z)}\,dz$$

$= \#$ zeros $- \#$ poles

Winding, Rouché, and Gohberg-Sigal





Analytic	$f, g : \Omega \to \mathbb{C}$	$T, E : \Omega \to \mathbb{C}^{n\times n}$
Winding #	$\frac{1}{2\pi i} \int_\Gamma \frac{f'(z)}{f(z)}\, dz$	$\operatorname{tr}\left(\frac{1}{2\pi i} \int_\Gamma T(z)^{-1} T'(z)\, dz\right)$
Theorem	Rouché (1862): $\lvert g\rvert < \lvert f\rvert$ on $\Gamma \implies$ same # zeros of $f, f+g$	Gohberg-Sigal (1971): $\lVert T^{-1}E\rVert < 1$ on $\Gamma \implies$ same # eigs of $T, T+E$

Comparing NEPs

Suppose

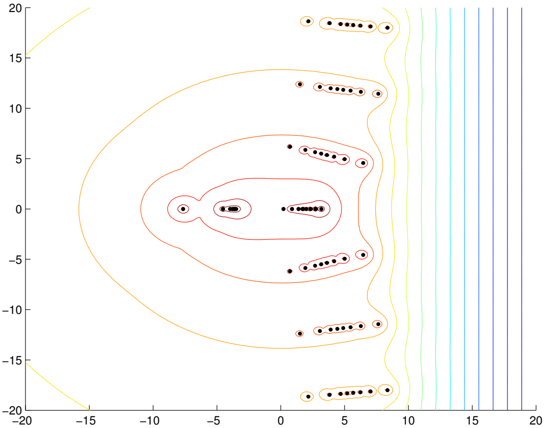
$$T, E : \Omega \to \mathbb{C}^{n\times n} \text{ analytic}$$
$$\Gamma \subset \Omega \text{ a simple closed contour}$$
$$T(z) + sE(z) \text{ nonsingular } \forall s \in [0,1], z \in \Gamma$$

Then T and $T + E$ have the same number of eigenvalues inside Γ.

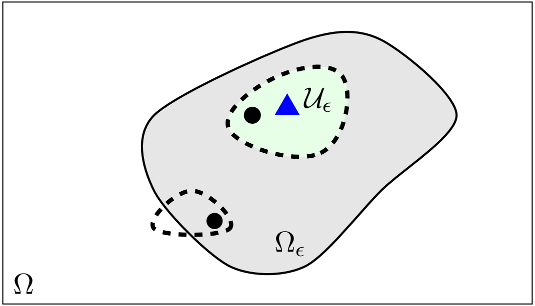
Pf: Constant winding number around Γ.

Nonlinear pseudospectra

$$\Lambda_\epsilon(T) \equiv \{z \in \Omega : \|T(z)^{-1}\| > \epsilon^{-1}\}$$



Pseudospectral comparison



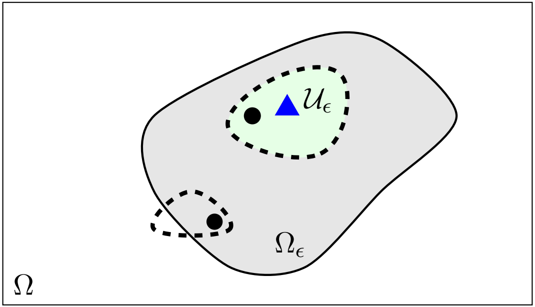
E analytic, $\|E(z)\| < \epsilon$ on Ω_ϵ. Then

$$\Lambda(T+E) \cap \Omega_\epsilon \quad \subset \quad \Lambda_\epsilon(T) \cap \Omega_\epsilon$$

Also, if $\mathcal{U}_\epsilon$ a component of Λ_ϵ and $\bar{\mathcal{U}}_\epsilon \subset \Omega_\epsilon$, then

$$|\Lambda(T+E) \cap \mathcal{U}_\epsilon| \quad = \quad |\Lambda(T) \cap \mathcal{U}_\epsilon|$$

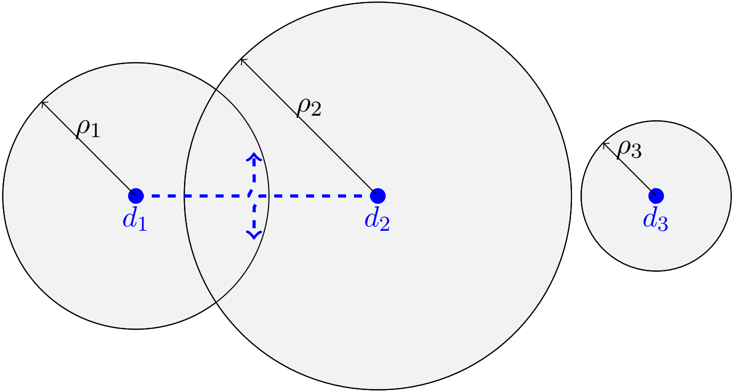
Pseudospectral comparison



- Most useful when T is linear
- Even then, can be expensive to compute!
- What about related tools?

The Gershgorin picture (linear case)

$$A = D + F, \qquad D = \operatorname{diag}(d_i), \qquad \rho_i = \sum_j |f_{ij}|$$



Gershgorin $(+\epsilon)$

Write $A = D + F$, $D = \mathrm{diag}(d_1, \ldots, d_n)$. Gershgorin disks are:

$$G_i = \left\{ z \in \mathbb{C} : |z - d_i| \leq \sum_j |f_{ij}| \right\}.$$

Useful facts:

- Spectrum of A lies in $\bigcup_{i=1}^m G_i$
- $\bigcup_{i \in \mathcal{I}} G_i$ disjoint from other disks $\implies$ contains $|\mathcal{I}|$ eigenvalues.

Pf:

$A - zI$ strictly diagonally dominant outside $\bigcup_{i=1}^m G_i$.
Eigenvalues of $D - sF$, $0 \leq s \leq 1$, are continuous.

Nonlinear Gershgorin

Write $T(z) = D(z) + F(z)$. Gershgorin *regions* are

$$G_i = \left\{ z \in \mathbb{C} : |d_i(z)| \leq \sum_j |f_{ij}(z)| \right\}.$$

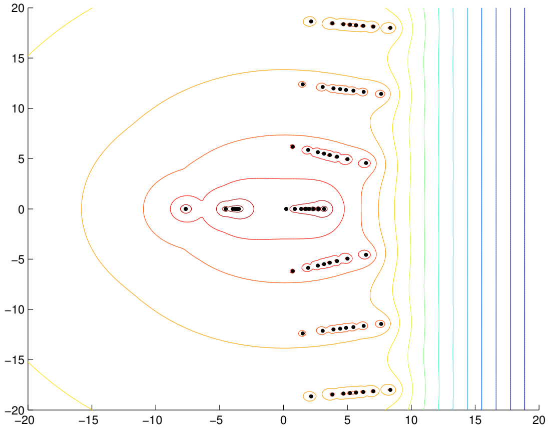
Useful facts:

- Spectrum of T lies in $\bigcup_{i=1}^m G_i$
- Bdd connected component of $\bigcup_{i=1}^m G_i$ strictly in Ω
  $\implies$ same number of eigs of D and T in component
  $\implies$ at least one eig per component of G_i involved

Pf: Strict diag dominance test + continuity of eigs

Example I: Hadeler

$$T(z) = (e^z - 1)B + z^2 A - \alpha I, \quad A, B \in \mathbb{R}^{8\times 8}$$



Comparison to simplified problem

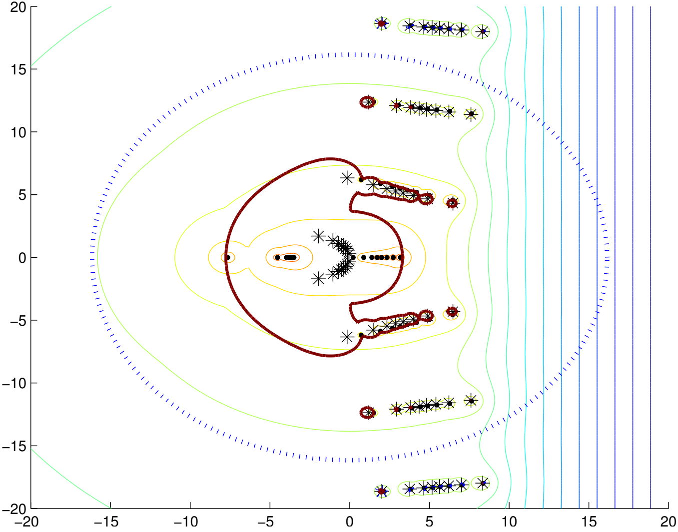
Bauer-Fike idea: apply a similarity!

$$T(z) = (e^z - 1)B + z^2 A - \alpha I$$

$$\begin{aligned}\tilde{T}(z) &= U^T T(z) U \\ &= (e^z - 1)D_B + z^2 I - \alpha E \\ &= D(z) - \alpha E\end{aligned}$$

$$G_i = \{z : |\beta_i(e^z - 1) + z^2| < \rho_i\}.$$

Gershgorin regions



A different comparison

Approximate $e^z - 1$ by a Chebyshev interpolant:

$$
\begin{aligned}
T(z) &= (e^z - 1)B + z^2 A - \alpha I \\
\tilde{T}(z) &= q(z)B + z^2 A - \alpha
\end{aligned}
$$

$$T(z) = \tilde{T}(z) + r(z)B$$

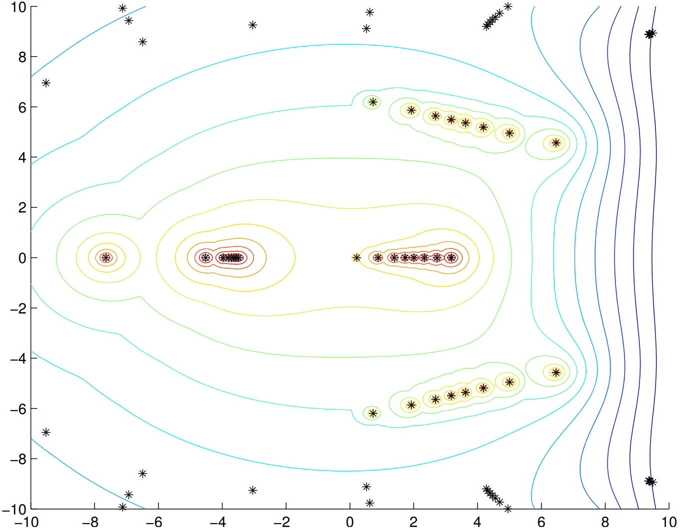
Linearize $\tilde{T}$ and transform both:

$$
\begin{aligned}
\tilde{T}(z) &\mapsto D_C - zI \\
T(z) &\mapsto D_C - zI + r(z)E
\end{aligned}
$$

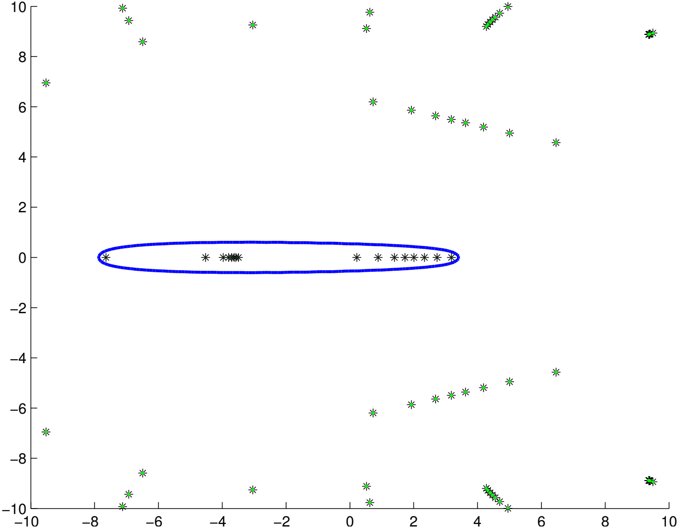
Restrict to $\Omega_\epsilon = \{z : |r(z)| < \epsilon\}$:

$$G_i \subset \hat{G}_i = \{z : |z - \mu_i| < \rho_i \epsilon\}, \qquad \rho_i = \sum_j |e_{ij}|$$

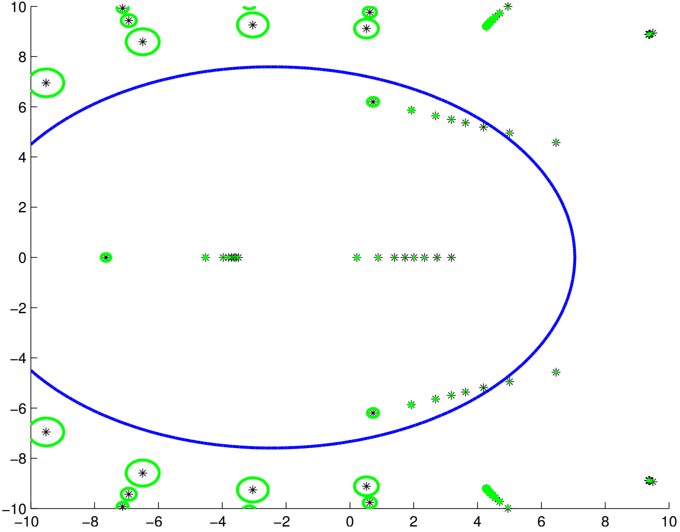
Spectrum of $\tilde{T}$



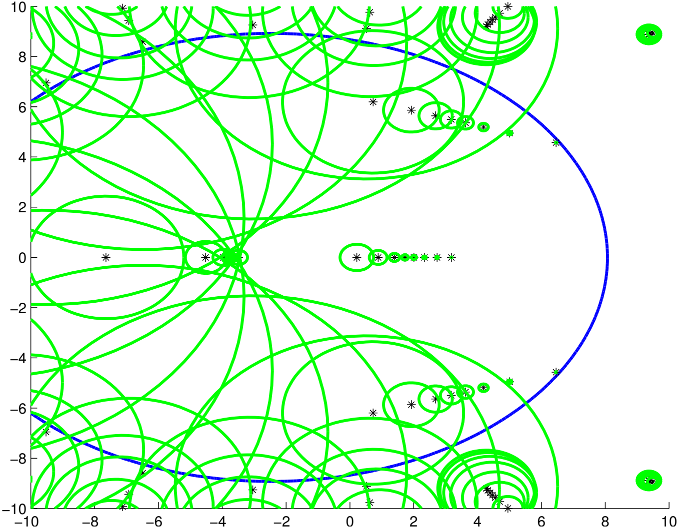
$\hat{G}_i$ for $\epsilon < 10^{-10}$



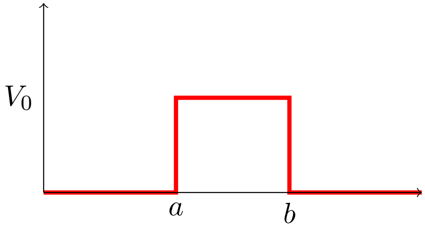
$\hat{G}_i$ for $\epsilon = 0.1$



$\hat{G}_i$ for $\epsilon = 1.6$

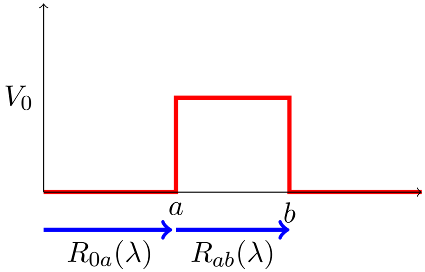


Example II: Resonance problem



$$\psi(0) = 0$$
$$\left(-\frac{d^2}{dx^2} + V - \lambda\right)\psi = 0 \qquad \text{on } (0, b),$$
$$\psi'(b) = i\sqrt{\lambda}\psi(b),$$

Reduction via shooting



$$\psi(0) = 0,$$

$$R_{0a}(\lambda)\begin{bmatrix}\psi(0)\\ \psi'(0)\end{bmatrix} = \begin{bmatrix}\psi(a)\\ \psi'(a)\end{bmatrix},$$

$$R_{ab}(\lambda)\begin{bmatrix}\psi(b)\\ \psi'(b)\end{bmatrix} = \begin{bmatrix}\psi(b)\\ \psi'(b)\end{bmatrix},$$

$$\psi'(b) = i\sqrt{\lambda}\psi(b)$$

Reduction via shooting

First-order form:

$$\frac{du}{dx} = \begin{bmatrix} 0 & 1 \\ V - \lambda & 0 \end{bmatrix} u, \text{ where } u(x) \equiv \begin{bmatrix} \psi(x) \\ \psi'(x) \end{bmatrix}.$$

On region (c, d) where V is constant:

$$u(d) = R_{cd}(\lambda)u(c), \quad R_{cd}(\lambda) = \exp\left((d - c) \begin{bmatrix} 0 & 1 \\ V - \lambda & 0 \end{bmatrix} \right)$$

Reduce resonance problem to 6D NEP:

$$T(\lambda)u_{\text{all}} \equiv \begin{bmatrix} R_{0a}(\lambda) & -I & 0 \\ 0 & R_{ab}(\lambda) & -I \\ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} & 0 & \begin{bmatrix} 0 & 0 \\ -i\sqrt{\lambda} & 1 \end{bmatrix} \end{bmatrix} \begin{bmatrix} u(0) \\ u(a) \\ u(b) \end{bmatrix} = 0.$$

Expansion via rational approximation

Consider the equation

$$\left(\begin{bmatrix} A & B \\ C & D \end{bmatrix} - \lambda I\right) \begin{bmatrix} u \\ v \end{bmatrix} = 0$$

Partial Gaussian elimination gives the **spectral Schur complement**

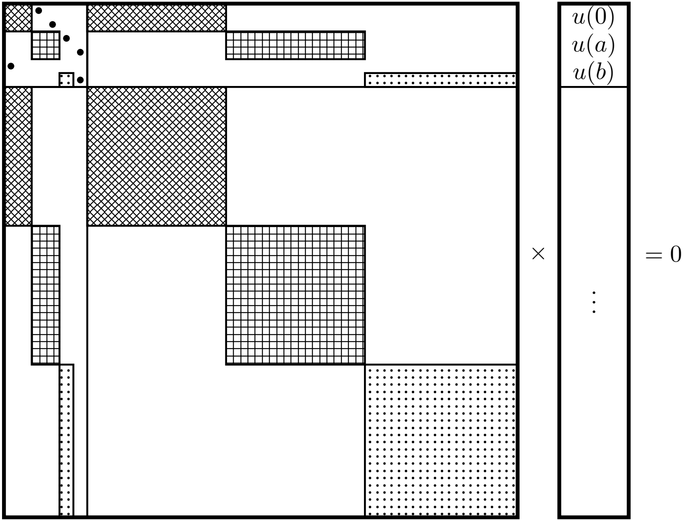
$$\left(A - \lambda I - B(D - \lambda I)^{-1} C\right) u = 0$$

Idea: Given

$$T(\lambda) = A - \lambda I + F(\lambda),$$

find a rational approximation $F(\lambda) \approx B(D - \lambda I)^{-1} C$.

Expansion via rational approximation



Analyzing the expanded system

- $\hat{T}(z)$ is a Schur complement in $K - zM$
  - So $\Lambda(\hat{T})$ is easy to compute.
- Or: think $T(z)$ is a Schur complement in $K - zM + E(z)$
- Compare $\hat{T}(z)$ to $T(z)$ or compare $K - zM + E(z)$ to $K - zM$

Analyzing the expanded system

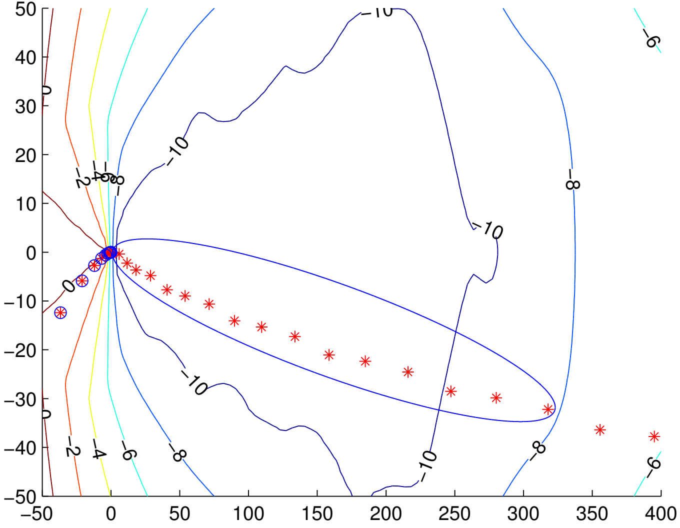
Q: Can we find all eigs in a region *not missing anything*?

Concrete plan ($\epsilon = 10^{-8}$)

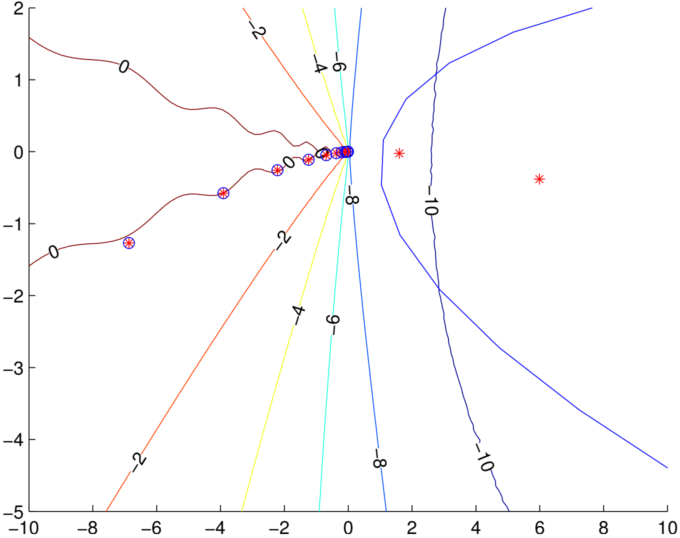
- T = shooting system
- $\hat{T}$ = rational approximation
- Find region D with boundary Γ s.t.
  - $D \subset \Omega_\epsilon$ (i.e. $\|T - \hat{T}\| < \epsilon$ on D)
  - Γ does not intersect $\Lambda_\epsilon(T)$
- $\implies$ Same eigenvalue counts for T, $\hat{T}$
- $\implies$ Eigs of $\hat{T}$ in components of $\Lambda_\epsilon(T)$
  - Converse holds if $D \subset \Omega_{\epsilon/2}$

Can refine eigs of $\hat{T}$ in D via Newton.

Resonance approximation



Resonance approximation



For more

Localization theorems for nonlinear eigenvalues.
David Bindel and Amanda Hood, SIAM Review 57(4), Dec 2015