Designing low rank methods for matrices with displacement structure

Heather Wilber
Oden Institute
University of Texas at Austin

October 23, 2021

Bernhard Beckermann

Daniel Kressner

Daniel Rubin

Alex Townsend

Designing low rank methods for matrices with displacement structure

A matrix $X \in \mathbb{C}^{m\times n}$ is said to have (A, B) displacement structure if

$$AX - XB = F,$$

where $A \in \mathbb{C}^{m\times m}$, $B \in \mathbb{C}^{n\times n}$, and $F \in \mathbb{C}^{m\times n}$.

Sylvester matrix equations appear in:

stability analysis for dynamical systems • discretizations of PDEs • signal processing and time series analysis • eigenvalue assignment problems • iterative solvers for continuous algebraic Ricattti matrix equation • analyses/computations involving special structured matrices (e.g., Toeplitz, Cauchy, Vandermonde)

In practical settings:

- A and B are sparse, banded, or structured, so that fast shifted inverts/matrix-vector products are available.
- F is often a low rank matrix (rank 1 or 2).
- X is dense.

[(Simoncini, 2016)]

Designing low rank methods for matrices with displacement structure

$$AX - XB = F$$

When the spectra of A and B are well-separated and F is low rank, X is well-approximated by low rank matrices.

1. Why is this true?
2. When is this true?
   - Only in the above circumstances or in greater generality?
   - Can we be precise about how the low rank properties of X depend on A, B, and F?
3. How can we take advantage of it?

The ADI method and Zolotarev rational functions

$$AX - XB = F \qquad A(ZDY^*) - (ZDY^*)B = USV^* \qquad (\text{S of size } \rho \times \rho)$$

(factored) ADI: A recipe for low rank approximations

$$Z^{(k)} = \left[\ \hat{Z}^{(1)}\ \middle|\ \hat{Z}^{(2)}\ \middle|\ \cdots\ \middle|\ \hat{Z}^{(k)}\ \right], \qquad \begin{cases} \hat{Z}^{(1)} = (A - \beta_1 I)^{-1} US, \\ \hat{Z}^{(i+1)} = (A - \alpha_i I)(A - \beta_{i+1} I)^{-1} Z^{(i)} \end{cases}$$

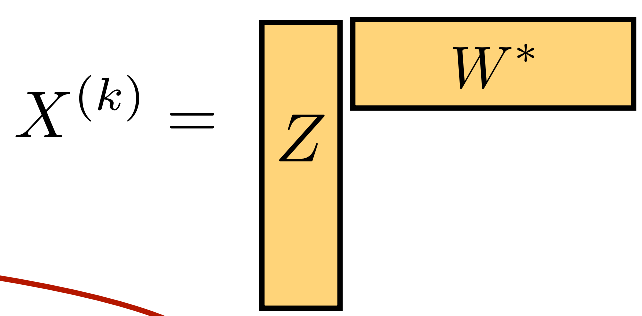
$$Y^{(k)} = \left[\ \hat{Y}^{(1)}\ \middle|\ \hat{Y}^{(2)}\ \middle|\ \cdots\ \middle|\ \hat{Y}^{(k)}\ \right], \qquad \begin{cases} \hat{Y}^{(1)} = (B^* - \alpha_1 I)^{-1} V, \\ \hat{Y}^{(i+1)} = (B^* - \beta_i I)(B^* - \alpha_{i+1} I)^{-1} Y^{(i)} \end{cases}$$

$$D^{(k)} = \mathrm{diag}\left((\beta_1 - \alpha_1) I_\rho, \cdots, (\beta_k - \alpha_k) I_\rho\right)$$

$$X^{(k)} = Z^{(k)} D^{(k)} Y^{(k)^*}$$

After k iterations:

- $X^{(k)} = ZW^*, \quad \mathrm{rank}(X^{(k)}) \leq k\rho, \quad \rho = \mathrm{rank}(F)$
- $X - X^{(k)} = r_k(A) X r_k(B)^{-1}, \quad r(z) = \prod_{j=1}^{k} \frac{z - \alpha_j}{z - \beta_j}$



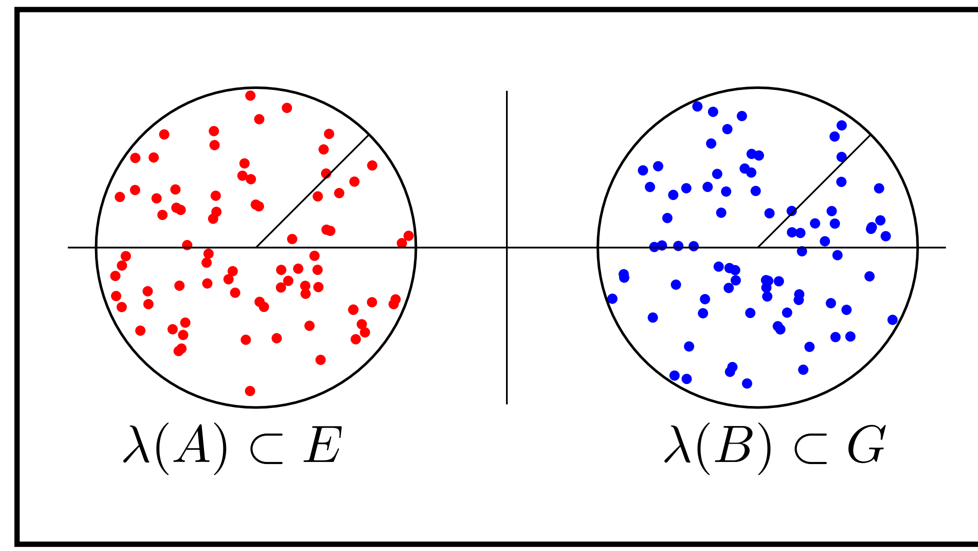
[(Beckermann & Townsend, 2017), (Sabino, 2008), (Penzl, 1999), (Benner, Truhar & Li, 2009), (Li & White, 2002), (Druskin, Knizhnerman & Simoncini, 2011), (Peaceman & Rachford, 1955), (Lu & Wachspress, 1991), (Townsend & W., 2018)]

The ADI method and Zolotarev rational functions

$$X - X^{(k)} = r_k(A) X r_k(B)^{-1}, \quad r(z) = \prod_{j=1}^{k} \frac{z - \alpha_j}{z - \beta_j}$$

$$\|X - X^{(k)}\|_2 \leq \|r_k(A) r_k(B)^{-1}\|_2 \|X\|_2 \leq \|r_k(\lambda(A))\|_2 \|r_k(\lambda(B))^{-1}\|_2 \|X\|_2$$

$$\|X - X^{(k)}\|_2 \leq \frac{\sup_{z \in E} |r_k(z)|}{\inf_{z \in G} |r_k(z)|} \|X\|_2$$



Zolotarev's third problem:

$$Z_k(E, G) := \inf_{r \in \mathcal{R}^k} \frac{\sup_{z \in E} |r(z)|}{\inf_{z \in G} |r(z)|}$$

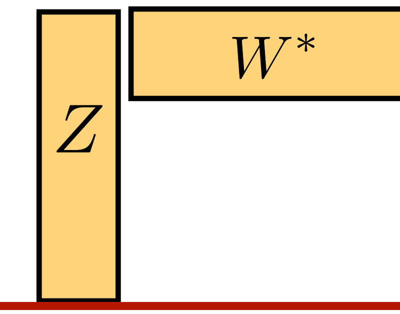
(Y. I. Zolotarev)

[(Beckermann & Townsend, 2017), (Sabino, 2008), (Penzl, 1999), (Benner, Truhar & Li, 2009), (Li & White, 2002), (Druskin, Knizhnerman & Simoninci, 2011), (Peaceman & Rachford, 1955), (Lu & Wachspress, 1991), (Townsend & W., 2018)]

The ADI method and Zolotarev rational functions

<u>After k iterations:</u>

- $X^{(k)} = ZW^*, \quad \text{rank}(X^{(k)}) \le k\rho, \quad \rho = \text{rank}(F)$
- $X - X^{(k)} = r_k(A) X r_k(B)^{-1}, \qquad r(z) = \prod_{j=1}^{k} \frac{z - \alpha_j}{z - \beta_j}$

$X^{(k)} =$ 

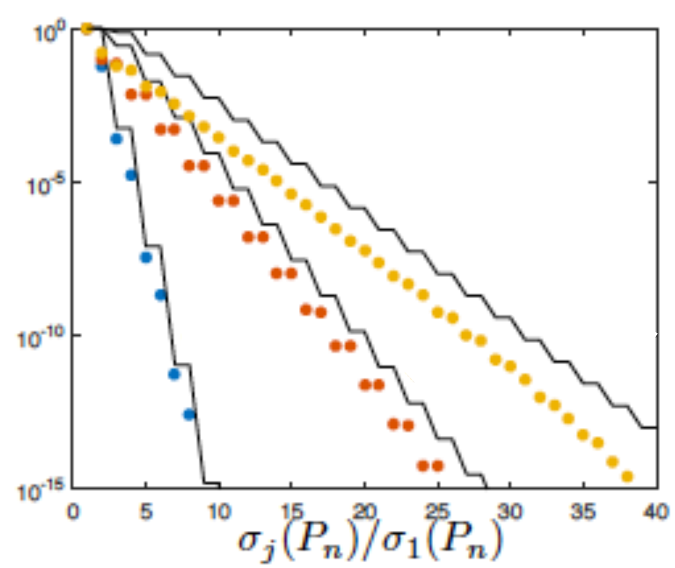
$$\sigma_{k\rho+1}(X) \le \|X - X^{(k)}\|_2 \le \|r_k(A) r_k(B)^{-1}\|_2 \|X\|_2 \le Z_k(E,G)\|X\|_2$$

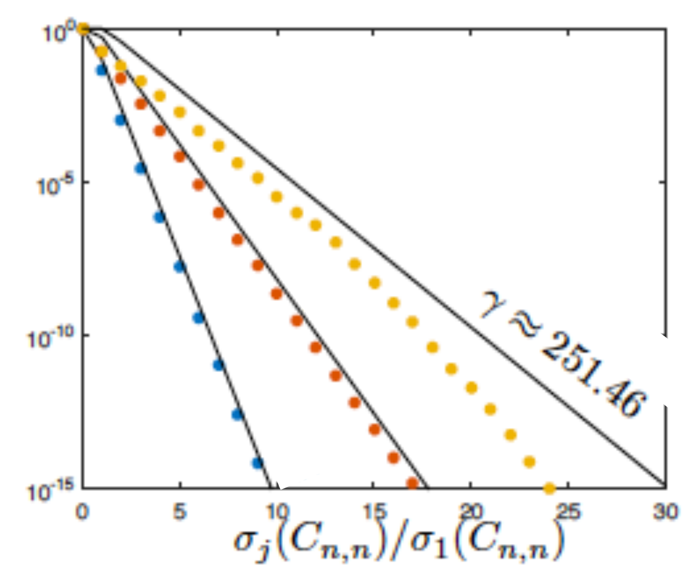
- Explicit bounds on the singular values of X
- A cheap method for constructing low rank approximations $X^{(k)} = ZW^* \approx X$

Explains low rank properties in real-valued Vandermonde, Pick, Cauchy, Loewner matrices and more...


(A. Townsend)


(B. Beckermann)





[(Beckermann & Townsend, 2017)]

The ADI method and Zolotarev rational functions

Connections to many other problems:

- Error bounds for rational Krylov methods, Cauchy skeletonization.

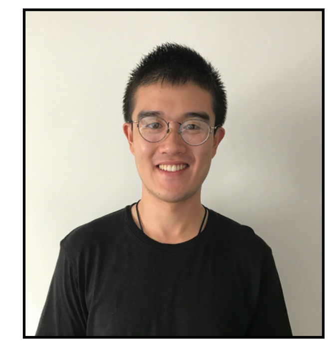
[Druskin, Knizhnerman and Simoninci (2011), Beckermann (2011)]

- Optimal complexity solvers for some elliptic PDEs.

[Olver and Townsend (2013) , Fortunato and Townsend (2018), Townsend, W., Wright (2016,2017), Boulle and Townsend (2019)]

- Compression properties in tensors/tensor train compression.

[Townsend and Shi, 2021]



(T. Shi)

- Fast solvers for certain linear systems $Xy = b$.

[Martinsson, Rokhlin, and Tygert (2005), Chandrasekaran, Gu, Xia, and Zhu (2007), Xia, Xi, and Gu (2012).]

- Efficient solvers for Ricatti (CARE) equation (rADI, qADI).

[Benner, Bujanović, Kürshcher, and Saak (2018), Wong and Balakrishnan (2005).]

- Best $\mathcal{R}^k$ approximation to the sign function (and others).

[Istace and Thiran (1995), Gawlik and Nakatsukasa (2019), (Nakatsukasa and Freund (2016)]

When can ADI-based arguments be used?

To bound singular values of X via fADI, we need...

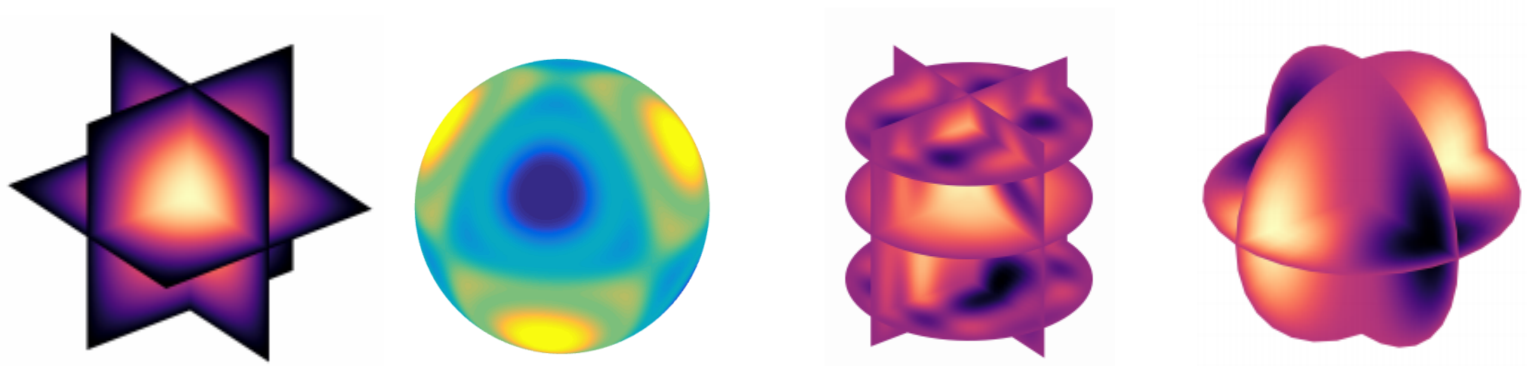
1. $\text{rank}(F)$ is small.
2. The spectra of A and B are well-separated.
3. A solution to Zolotarev's problem is known for sets E, G, where $\lambda(A) \subset E$ and $\lambda(B) \subset G$.

Problem: Many settings where these constraints are not satisfied!

Strategy 1: Make more problems ADI-friendly

ADI-friendly spectral discretizations for optimal complexity Poisson solvers

(D. Fortunato)



[Boulle and Townsend (2019), Fortunato and Townsend (2018), Olver and Townsend (2013), Townsend, W., and Wright (2016,2017)]

When can ADI-based arguments be used?

To bound singular values of X via fADI, we need...

1. $\text{rank}(F)$ is small.

2. The spectra of A and B are well-separated.

3. A solution to Zolotarev's problem is known for sets E, G, where $\lambda(A) \subset E$ and $\lambda(B) \subset G$.

Problem: **Many practical applications do not satisfy these constraints!**

Strategy 2: Make ADI friendlier for more problems!

Expanding ADI-based methods

1. ~~rank(F) is small.~~

F has decaying singular values.

Townsend, W., (2018):

ADI with high-rank right-hand sides.

- low rank solver for $AX - XB = F$, F is full rank.
- bounds on numerical ranks of matrices, e.g., multidimensional Vandermonde

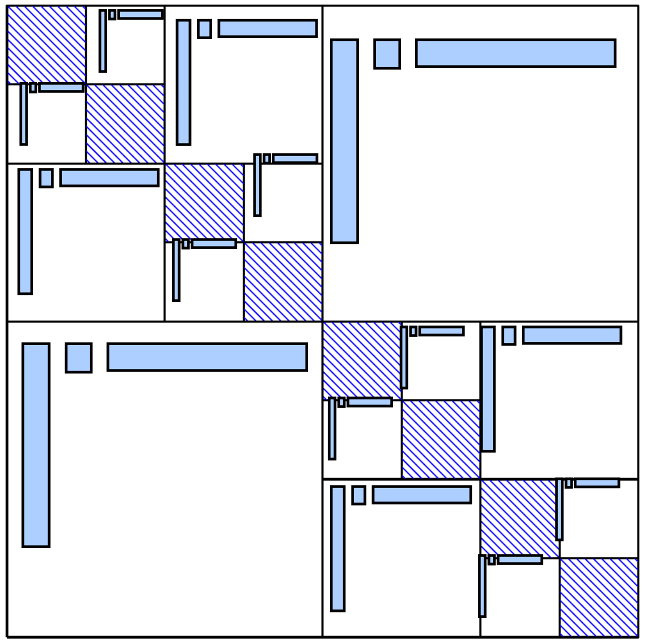
2. ~~The spectra of A and B are well separated~~.

Subsets of the spectra of A and B are well-separated.

Beckermann, Kressner, W., (2021)

superfast solvers for Toeplitz system $Tx = b$.

ADI-based hierarchical compression.



- extends to other related linear systems (e.g., NUDFT, Toeplitz+Hankel)
- explicit approx. error bounds + competitive with state-of-the-art.

[Kressner, Massei and Robol (2019), Martinsson, Rokhlin, and Tygert (2005), Chandrasekaran, Gu, Xia, and Zhu (2007), Xia, Xi, and Gu (2012)]

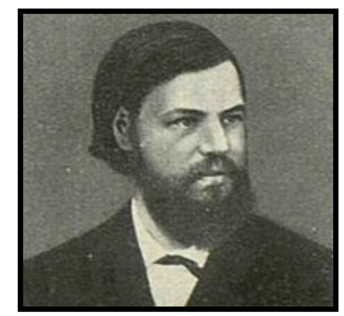
Zolotarev's problem in the complex plane

3. ~~A solution~~ to Zolotarev's problem is known for sets E, G, where $\lambda(A) \subset E$ and $\lambda(B) \subset G$.

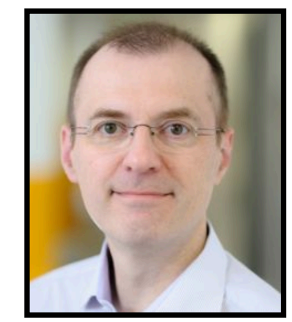
An approximate solution

Solution is known for:

- Intervals of the real line
- Disks in the complex plane



(Y. I. Zolotarev)

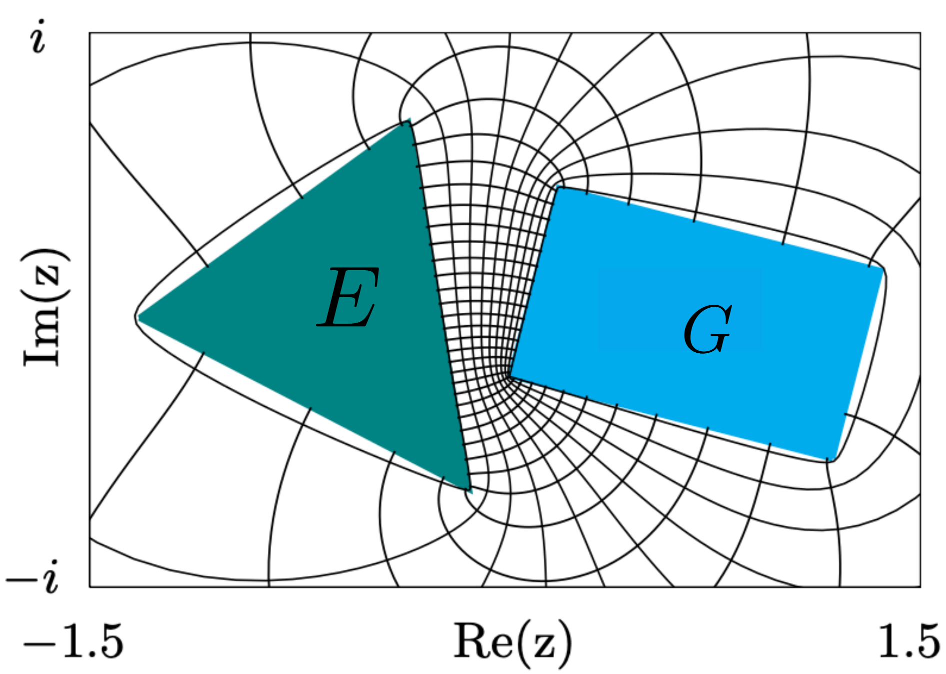


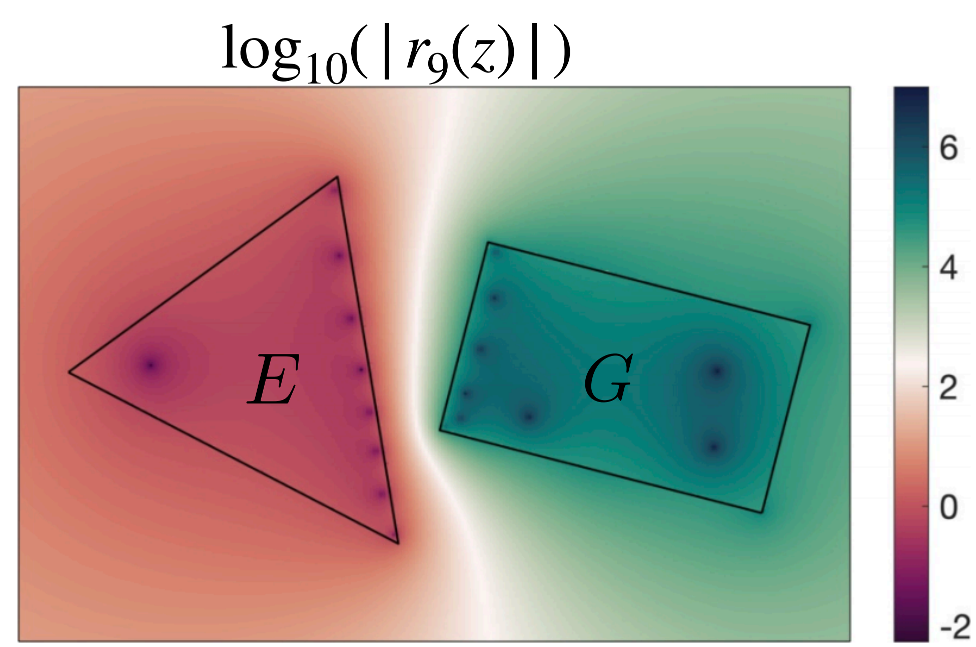
(G. Starke)

Zolotarev's third problem:

$$Z_k(E,G) := \inf_{r \in \mathcal{R}^k} \frac{\sup_{z\in E} |r(z)|}{\inf_{z \in G} |r(z)|}$$

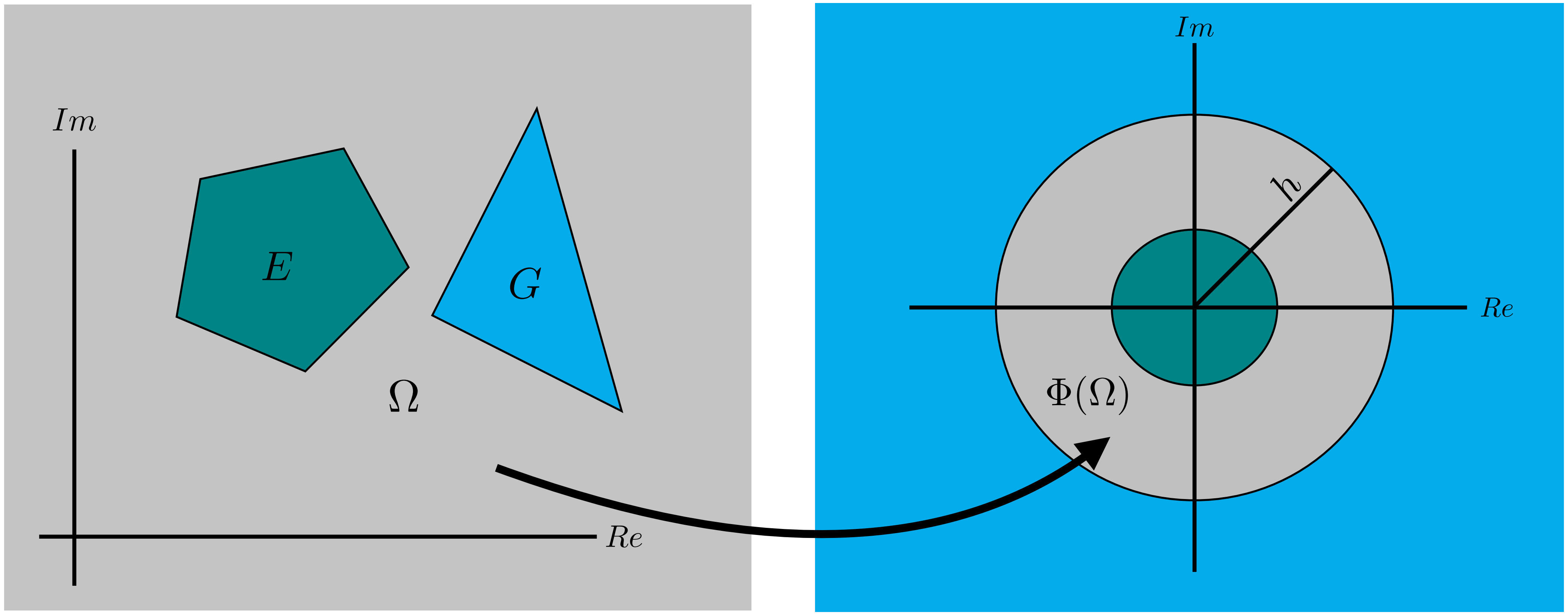
For more general sets in $\mathbb{C}$...





[(Zolotarev, 1877), (Sabino, 2008), (Beckermann & Townsend, 2017, 2019), (Starke, 1992) (Ganelius, 1976, 1979)]

Zolotarev's problem in the complex plane



$h = \exp(1/\mathrm{cap}(E,G))$

$$\Phi : \Omega \to \mathcal{A} = \{z \in \mathbb{C}, 1 \leq |z| \leq h\}$$

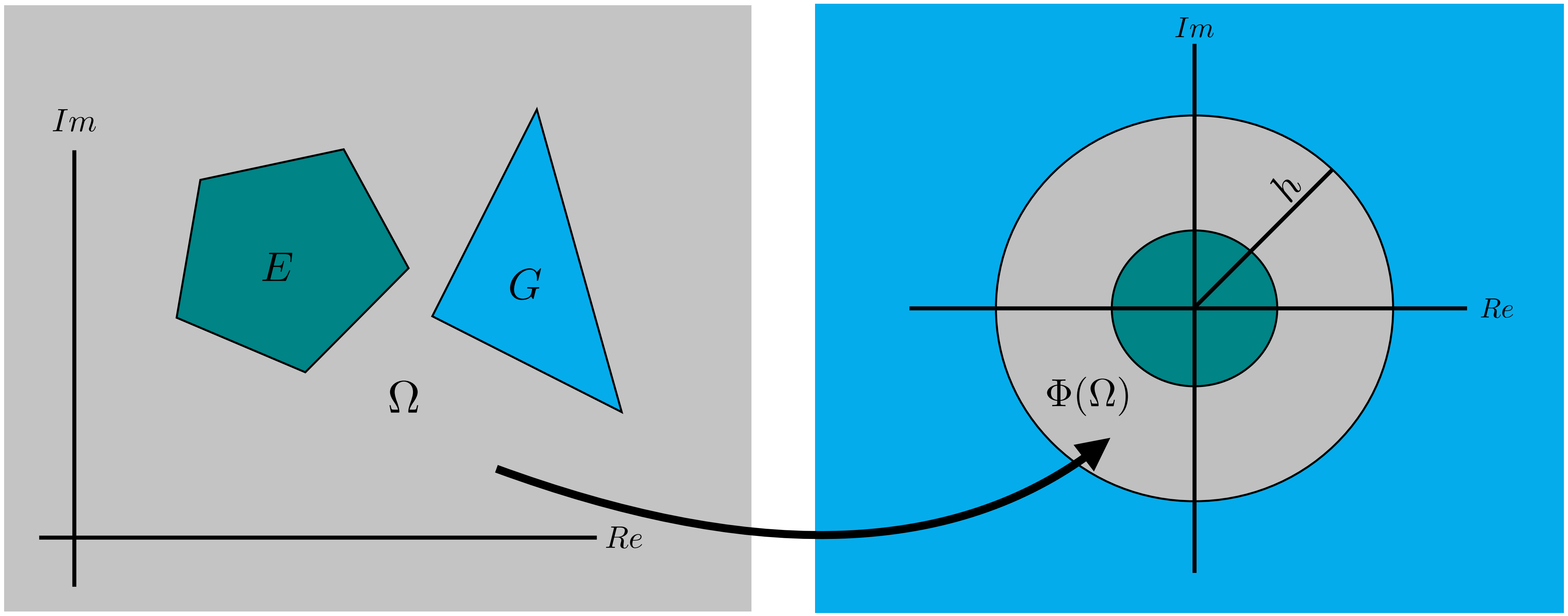
Suppose that Φ is a type $(1,1)$ rational function...

$$h^{-k} \leq Z_k(E,G) \leq \frac{\sup_{z\in E} \Phi^k(z)}{\inf_{z\in G} \Phi^k(z)} \leq \frac{1}{h^k} = h^{-k}.$$

$$\implies Z_k(E,G) = h^{-k}$$

[(Townsend, Rubin, W., 2021) (Starke, 1992) (Ganelius, 1976, 1979) (Gončar, 1969)]

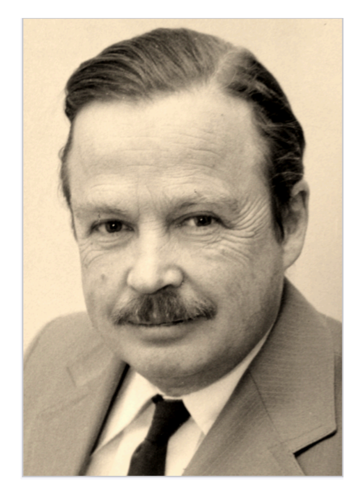
Zolotarev's problem in the complex plane



$$\Phi : \Omega \to \mathcal{A} = \{z \in \mathbb{C}, 1 \leq |z| \leq h\}$$

When Φ isn't a rational function, the story gets more complicated...

- Apply a special "filtering" process to $\Phi^k(z)$,
- Results in a type (k, k) rational $\tilde{r}(z)$ (Faber rational),
- Bound $\dfrac{\sup_{z\in E} |\tilde{r}_k(z)|}{\inf_{z\in G} |\tilde{r}_k(z)|}$ from above.


(T. Ganelius)


(G. Faber)

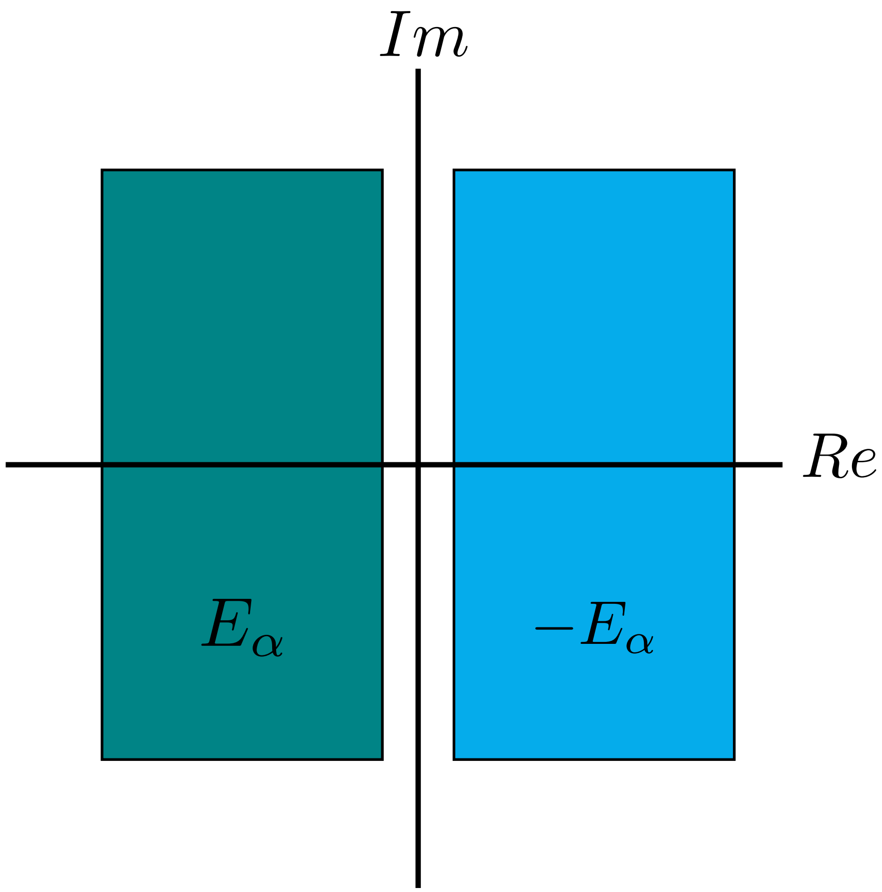
[(Townsend, Rubin, W., 2021) (Ganelius, 1976, 1979)]

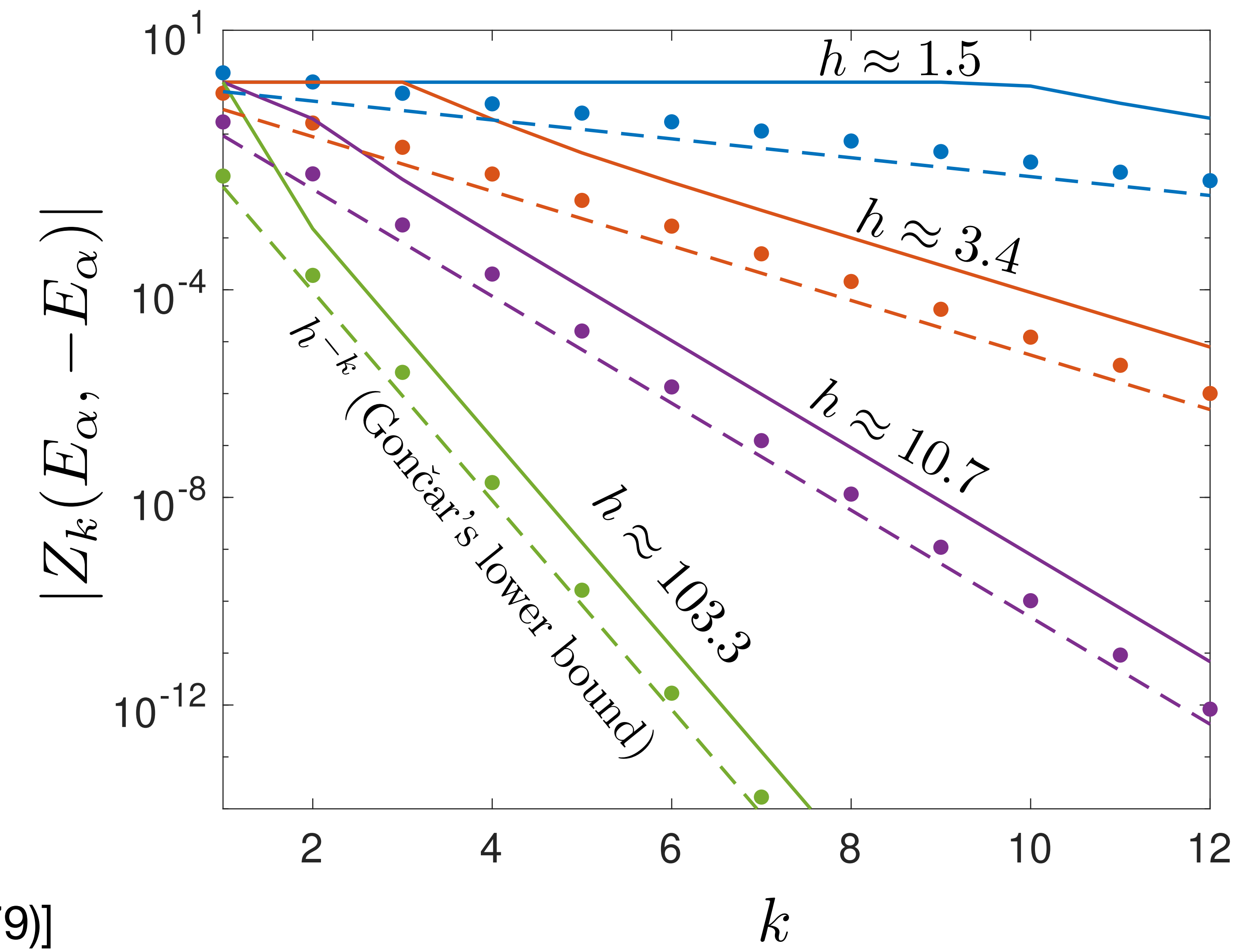
Zolotarev's problem in the complex plane

<u>Theorem</u> (Rubin, Townsend, W., 2021) If E, G are disjoint, bounded open convex sets in $\mathbb{C}$, then there is k_0 where for $k > k_0$,

$$Z_k(E,G) \leq 16h^{-k} + \mathcal{O}(h^{-2k}).$$

*We have an inelegant explicit upper bound and expression for k_0.





[(Townsend, Rubin, W., 2021) (Ganelius, 1976, 1979)]

Zolotarev's problem in the complex plane

Disjoint sets E and G	Bound	Reference
finite intervals of $\mathbb{R}$	$Z_k(E,G) \le 4h^{-k}$	Beckermann, Townsend (2017)
disks in $\mathbb{C}$	$Z_k(E,G) \le h^{-k}$	Starke (1992)
arcs on a circle $\mathbb{C}$	$Z_k(E,G) \le 4h^{-k}$	Beckermann, Kressner, W. (2021)
more general sets in $\mathbb{C}$	$Z_k(E,G) \le 16h^{-k} + \mathcal{O}(h^{-2k})$	

- Bounds on singular values for families of matrices.
- Bounds for rational approximation to $\text{sign}(z)$ on E, G.

ADI shift parameters?

[(Townsend, Rubin, W., 2021) (Ganelius, 1976, 1979)]

Revisiting old ideas (with new tools)

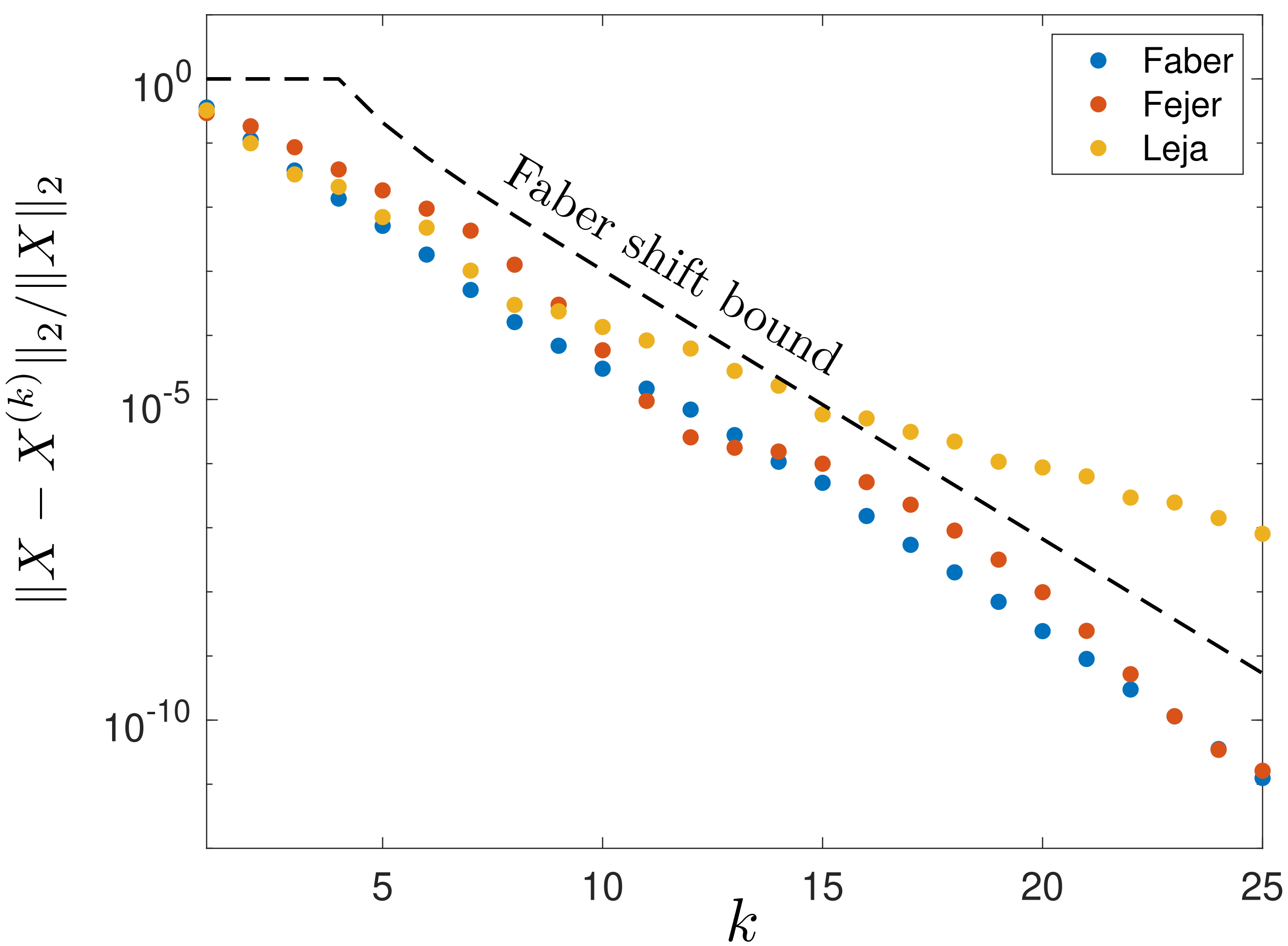
Asymptotically optimal points: $\lim_{k\to\infty}\left(\frac{\sup_{z\in E}|r_k(z)|}{\inf_{z\in G}|r_k(z)|}\right)^{1/k} = h^{-1}$

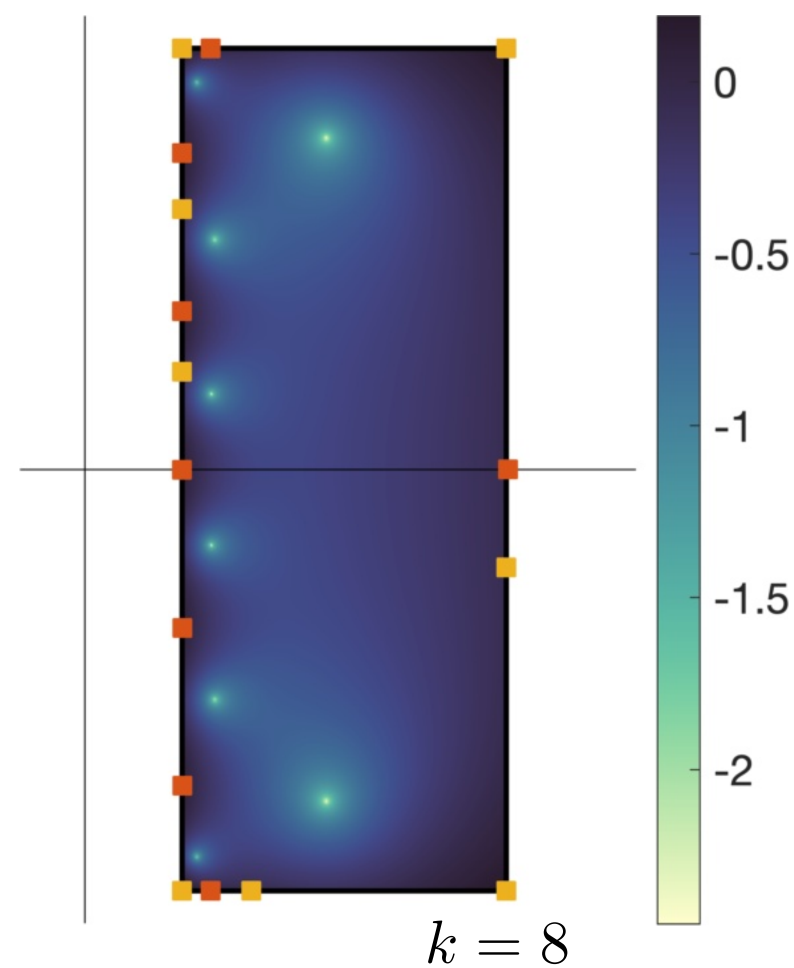
Fejer Points:

(1) Pick equally spaced points $\mathcal{P}$ on inner and outer boundaries of $\mathcal{A}$,

(2) Use $\Phi^{-1}(\mathcal{P})$ as ADI shift parameters.

Leja Points: Greedy selection process from discretization of the boundaries of E and G.





[(Starke, 1992)]

Revisiting old ideas (with new tools)

Many modern tools available to compute Φ (and h)

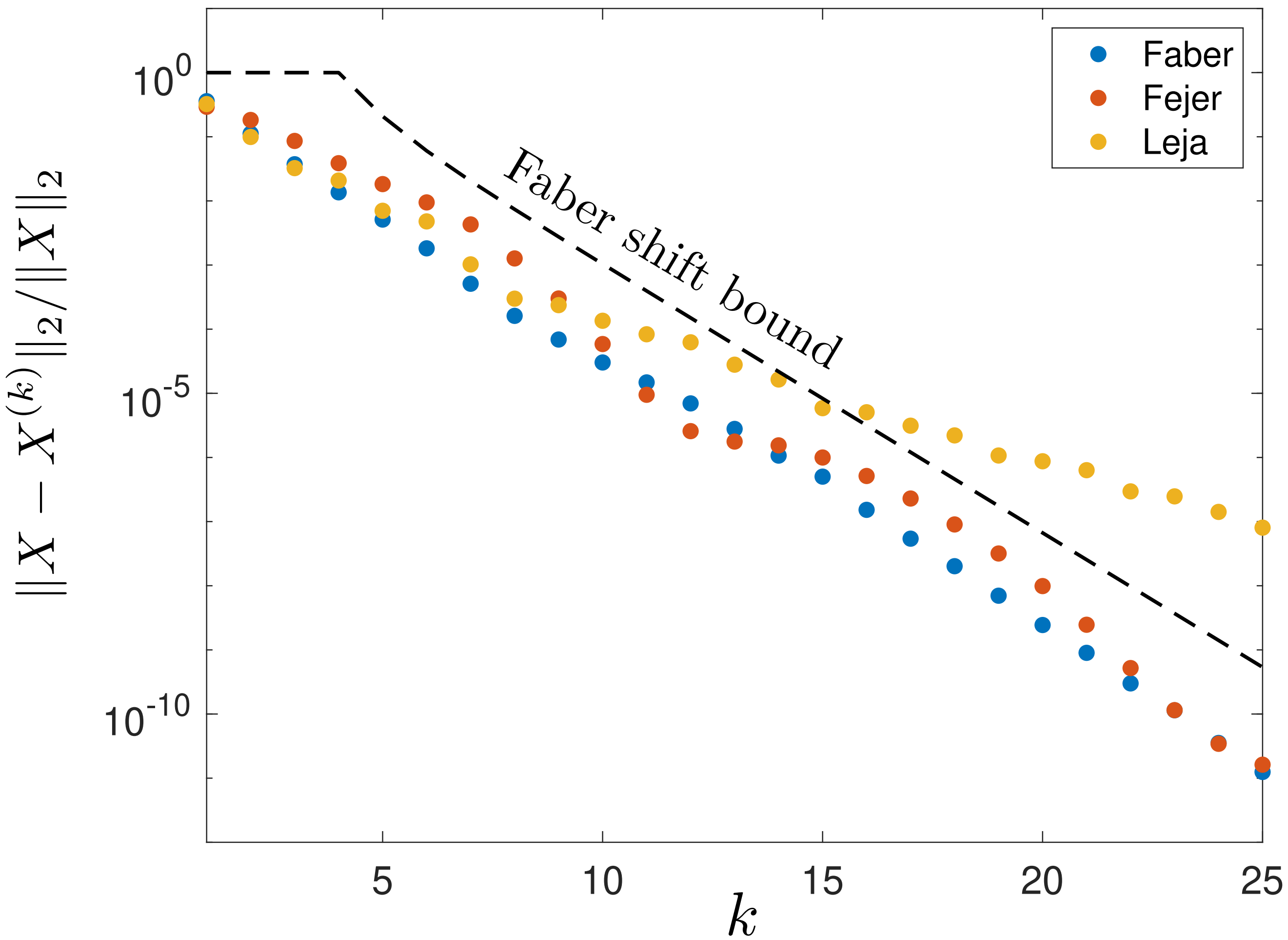
Lightning Laplace solver (Trefethen, Gopal, Baddoo)

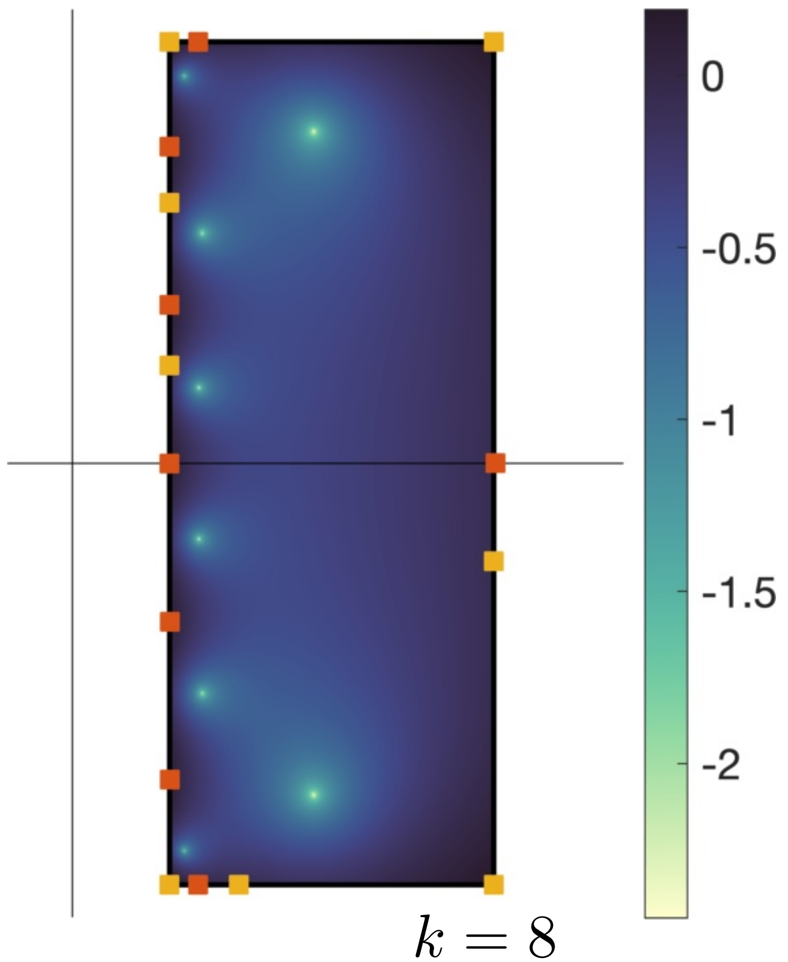
Integral formulations (Gaier, Schiffer, Nasser)

Schwarz-Christoffel methods (Delillo, Elcrat, Driscoll, Crowdy, many more…)

To compute/evaluate Φ^{-1}

Construct a complex-valued barycentric rational interpolant to samples $(\Phi(z), z)$.





[(Townsend, Rubin, & W., 2021) (Trefethen, 2020) (Nakatsukasa, Trefethen, & Sète, 2018)]

Thank you!