

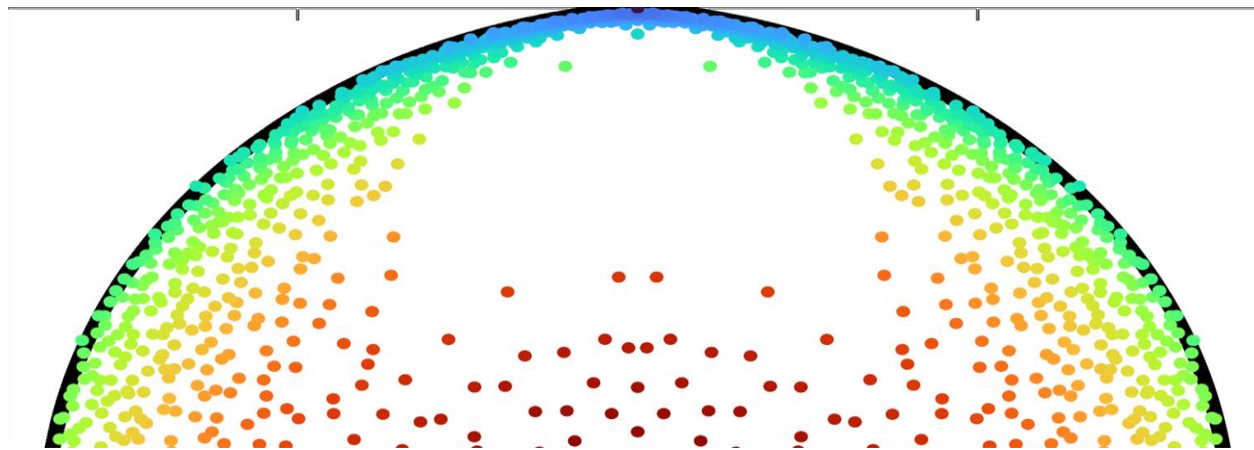
Residual Dynamic Mode Decomposition

Robust and verified data-driven Koopmanism
for nonlinear dynamical systems

Matthew Colbrook (m.colbrook@damtp.cam.ac.uk)

University of Cambridge

Work with **Lorna Ayton** (Cambridge), **Máté Szőke** (Virginia Tech) and **Alex Townsend** (Cornell)



C., Townsend, “Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems,” preprint.

C., Ayton, Szőke, “Residual dynamic mode decomposition: robust and verified Koopmanism,” **Journal of Fluid Mechanics**, 2023.

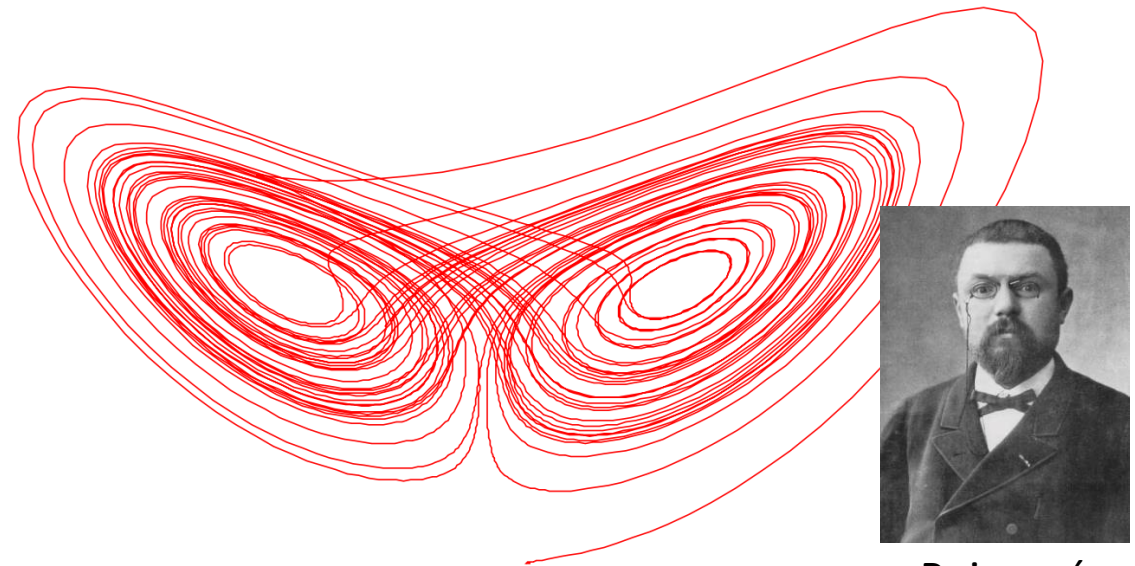
C., “The mpEDMD algorithm for data-driven computations of measure-preserving dynamical systems,” **SINUM**, to appear.

Data-driven dynamical systems

- State $x \in \Omega \subseteq \mathbb{R}^d$, **unknown** function $F: \Omega \rightarrow \Omega$ governs dynamics

$$x_{n+1} = F(x_n)$$

- **Goal:** Learn about system from data $\{x^{(m)}, y^{(m)} = F(x^{(m)})\}_{m=1}^M$
 - **Data:** experimental measurements or numerical simulations
 - E.g., **used for** forecasting, control, design, understanding
- **Applications:** chemistry, climatology, electronics, epidemiology, finance, fluids, molecular dynamics, neuroscience, plasmas, robotics, video processing, etc.



Poincaré

Operator viewpoint

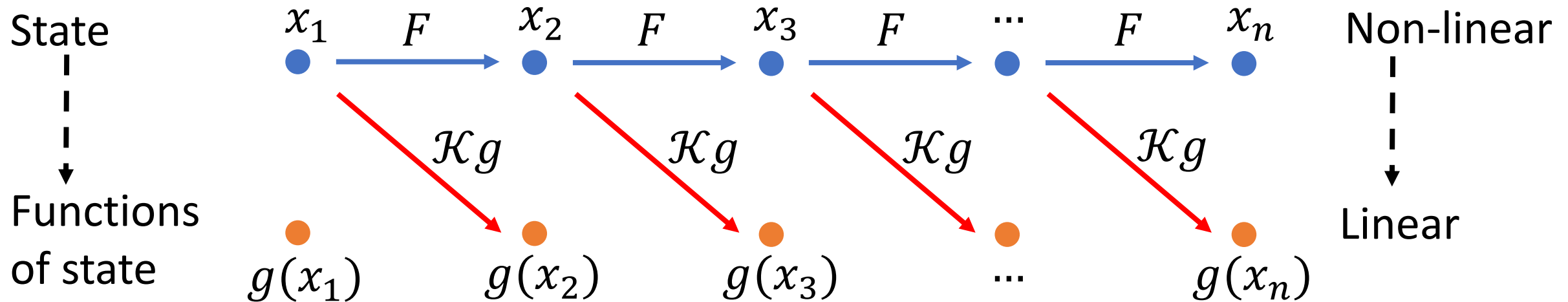
Koopman

von Neumann



- **Koopman operator** \mathcal{K} acts on functions $g: \Omega \rightarrow \mathbb{C}$

$$[\mathcal{K}g](x_n) = g(F(x_n)) = g(x_{n+1})$$
- \mathcal{K} is **linear** but acts on an **infinite-dimensional** space.



- Work in $L^2(\Omega, \omega)$ for positive measure ω , with inner product $\langle \cdot, \cdot \rangle$.

• Koopman, "Hamiltonian systems and transformation in Hilbert space," *Proc. Natl. Acad. Sci. USA*, 1931.

• Koopman, v. Neumann, "Dynamical systems of continuous spectra," *Proc. Natl. Acad. Sci. USA*, 1932.

Why is linear (much) easier?

$$x_{n+1} = F(x_n)$$

- Suppose $F(x) = Ax, A \in \mathbb{R}^{d \times d}, A = V\Lambda V^{-1}$.
- Set $\xi = V^{-1}x$,

$$\xi_n = V^{-1}x_n = V^{-1}A^n x_0 = \Lambda^n V^{-1}x_0 = \Lambda^n \xi_0$$

- Let $w^T A = \lambda w$, set $\varphi(x) = w^T x$,

$$[\mathcal{K}\varphi](x) = w^T Ax = \lambda \varphi(x)$$

Long-time dynamics
become trivial!



Eigenfunction

Much more general (**non-linear** and even **chaotic** F).

Koopman mode decomposition

generalised
eigenfunction of \mathcal{K}

eigenfunction of \mathcal{K}

$$g(x) = \sum_{\text{eigs } \lambda_j} c_{\lambda_j} \varphi_{\lambda_j}(x) + \int_{[-\pi, \pi]_{\text{per}}} \phi_{\theta, g}(x) \, d\theta$$

$$g(x_n) = [\mathcal{K}^n g](x_0) = \sum_{\text{eigs } \lambda_j} c_{\lambda_j} \lambda_j^n \varphi_{\lambda_j}(x_0) + \int_{[-\pi, \pi]_{\text{per}}} e^{in\theta} \phi_{\theta, g}(x_0) \, d\theta$$

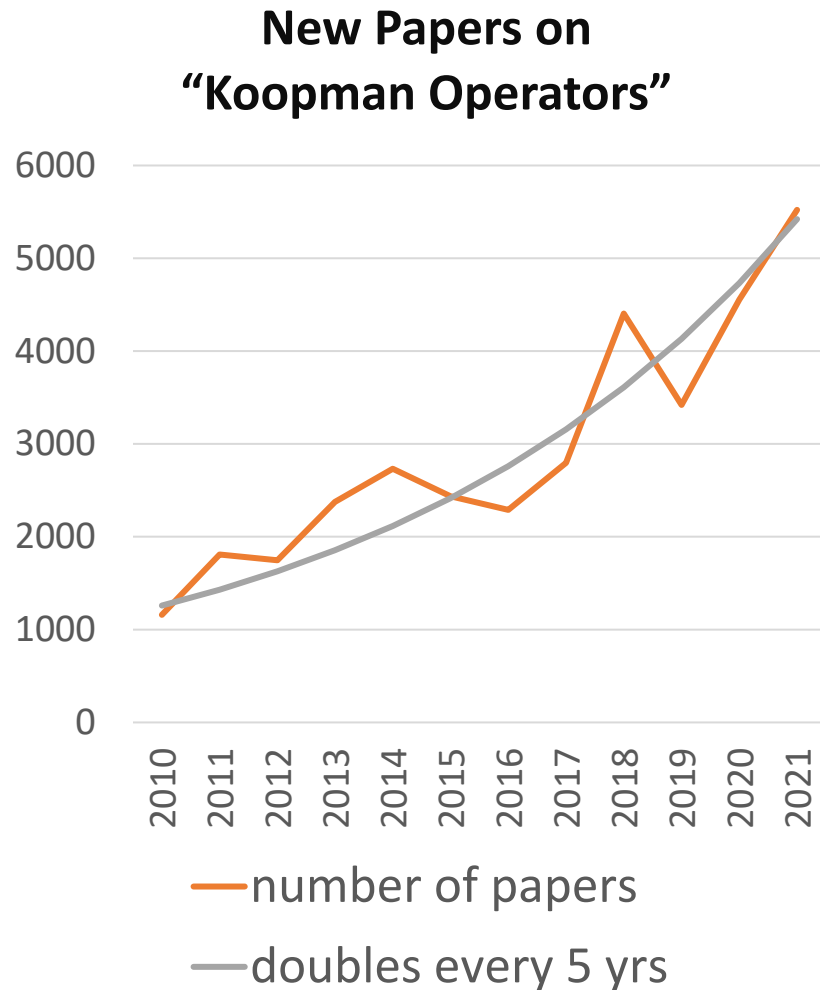
$$[\mathcal{K}g](x) = g(F(x))$$

Encodes: geometric features, invariant measures, transient behaviour, long-time behaviour, coherent structures, quasiperiodicity, etc.

GOAL: Data-driven approximation of \mathcal{K} and its spectral properties.

- Mezić, “Spectral properties of dynamical systems, model reduction and decompositions,” **Nonlinear Dynam.**, 2005.

Koopmania*: A revolution in the big data era?



≈35,000 papers over last decade!

BUT: Very little on verified methods!

Computing spectra in infinite dimensions is notoriously hard!

*Wikipedia: "its wild surge in popularity is sometimes jokingly called 'Koopmania'"

Challenges of computing

$$\text{Spec}(\mathcal{K}) = \{\lambda \in \mathbb{C}: \mathcal{K} - \lambda I \text{ is not invertible}\}$$

Truncate: $\mathcal{K} \longrightarrow \mathbb{K} \in \mathbb{C}^{N_K \times N_K}$

- 1) **“Too much”:** Approximate spurious modes $\lambda \notin \text{Spec}(\mathcal{K})$
- 2) **“Too little”:** Miss parts of $\text{Spec}(\mathcal{K})$
- 3) **Continuous spectra.**

Verification: Is it right?

Build the matrix: Dynamic Mode Decomposition (DMD)

Given dictionary $\{\psi_1, \dots, \psi_{N_K}\}$ of functions $\psi_j: \Omega \rightarrow \mathbb{C}$,

$$\{x^{(m)}, y^{(m)} = F(x^{(m)})\}_{m=1}^M$$

$$\langle \psi_k, \psi_j \rangle \approx \sum_{m=1}^M w_m \overline{\psi_j(x^{(m)})} \psi_k(x^{(m)}) = \left[\underbrace{\begin{pmatrix} \psi_1(x^{(1)}) & \dots & \psi_{N_K}(x^{(1)}) \\ \vdots & \ddots & \vdots \\ \psi_1(x^{(M)}) & \dots & \psi_{N_K}(x^{(M)}) \end{pmatrix}}_{\Psi_X} \right]^* \underbrace{\begin{pmatrix} w_1 & & \\ & \ddots & \\ & & w_M \end{pmatrix}}_W \underbrace{\begin{pmatrix} \psi_1(x^{(1)}) & \dots & \psi_{N_K}(x^{(1)}) \\ \vdots & \ddots & \vdots \\ \psi_1(x^{(M)}) & \dots & \psi_{N_K}(x^{(M)}) \end{pmatrix}}_{\Psi_X} \right]_{jk}$$

$$\langle \mathcal{K}\psi_k, \psi_j \rangle \approx \sum_{m=1}^M w_m \overline{\psi_j(x^{(m)})} \underbrace{\psi_k(y^{(m)})}_{[\mathcal{K}\psi_k](x^{(m)})} = \left[\underbrace{\begin{pmatrix} \psi_1(x^{(1)}) & \dots & \psi_{N_K}(x^{(1)}) \\ \vdots & \ddots & \vdots \\ \psi_1(x^{(M)}) & \dots & \psi_{N_K}(x^{(M)}) \end{pmatrix}}_{\Psi_X} \right]^* \underbrace{\begin{pmatrix} w_1 & & \\ & \ddots & \\ & & w_M \end{pmatrix}}_W \underbrace{\begin{pmatrix} \psi_1(y^{(1)}) & \dots & \psi_{N_K}(y^{(1)}) \\ \vdots & \ddots & \vdots \\ \psi_1(y^{(M)}) & \dots & \psi_{N_K}(y^{(M)}) \end{pmatrix}}_{\Psi_Y} \right]_{jk}$$

$$\mathcal{K} \longrightarrow \mathbb{K} = (\Psi_X^* W \Psi_X)^{-1} \Psi_X^* W \Psi_Y \in \mathbb{C}^{N_K \times N_K}$$

Recall open problems: too much, too little, continuous spectra, verification

- Schmid, "Dynamic mode decomposition of numerical and experimental data," **J. Fluid Mech.**, 2010.
- Rowley, Mezić, Bagheri, Schlatter, Henningson, "Spectral analysis of nonlinear flows," **J. Fluid Mech.**, 2009.
- Kutz, Brunton, Brunton, Proctor, "Dynamic mode decomposition: data-driven modeling of complex systems," **SIAM**, 2016.
- Williams, Kevrekidis, Rowley "A data-driven approximation of the Koopman operator: Extending dynamic mode decomposition," **J. Nonlinear Sci.**, 2015.

Residual DMD (ResDMD): Approx. \mathcal{K} and $\mathcal{K}^*\mathcal{K}$

$$\langle \psi_k, \psi_j \rangle \approx \sum_{m=1}^M w_m \overline{\psi_j(x^{(m)})} \psi_k(x^{(m)}) = \left[\underbrace{\Psi_X^* W \Psi_X}_G \right]_{jk}$$

$$\langle \mathcal{K}\psi_k, \psi_j \rangle \approx \sum_{m=1}^M w_m \overline{\psi_j(x^{(m)})} \underbrace{\psi_k(y^{(m)})}_{[\mathcal{K}\psi_k](x^{(m)})} = \left[\underbrace{\Psi_X^* W \Psi_Y}_{K_1} \right]_{jk}$$

$$\langle \mathcal{K}\psi_k, \mathcal{K}\psi_j \rangle \approx \sum_{m=1}^M w_m \overline{\psi_j(y^{(m)})} \psi_k(y^{(m)}) = \left[\underbrace{\Psi_Y^* W \Psi_Y}_{K_2} \right]_{jk}$$

Residuals: $g = \sum_{j=1}^{N_K} \mathbf{g}_j \psi_j, \quad \|\mathcal{K}g - \lambda g\|^2 \approx \mathbf{g}^* [K_2 - \lambda K_1^* - \bar{\lambda} K_1 + |\lambda|^2 G] \mathbf{g}$

-
- C., Townsend, “Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems,” preprint.
 - C., Ayton, Szőke, “Residual Dynamic Mode Decomposition,” **J. Fluid Mech.**, 2023.
 - Code: <https://github.com/MColbrook/Residual-Dynamic-Mode-Decomposition>

ResDMD: avoiding “too much”

$$\text{res}(\lambda, \mathbf{g})^2 = \frac{\mathbf{g}^* [K_2 - \lambda K_1^* - \bar{\lambda} K_1 + |\lambda|^2 G] \mathbf{g}}{\mathbf{g}^* G \mathbf{g}}$$

eigenvectors

eigenvalues

Algorithm 1:

1. Compute $G, K_1, K_2 \in \mathbb{C}^{N_K \times N_K}$ and eigendecomposition $K_1 V = G V \Lambda$.
2. For each eigenpair (λ, \mathbf{v}) , compute $\text{res}(\lambda, \mathbf{v})$.
3. **Output:** subset of e-vectors $V_{(\varepsilon)}$ & e-vals $\Lambda_{(\varepsilon)}$ with $\text{res}(\lambda, \mathbf{v}) \leq \varepsilon$ ($\varepsilon = \text{input tol}$).

Theorem (no spectral pollution): Suppose quad. rule converges. Then

$$\limsup_{M \rightarrow \infty} \max_{\lambda \in \Lambda^{(\varepsilon)}} \|(\mathcal{K} - \lambda)^{-1}\|^{-1} \leq \varepsilon$$

ResDMD: avoiding “too much”

$$\text{res}(\lambda, \mathbf{g})^2 = \frac{\mathbf{g}^* [K_2 - \lambda K_1^* - \bar{\lambda} K_1 + |\lambda|^2 G] \mathbf{g}}{\mathbf{g}^* G \mathbf{g}}$$

eigenvectors

eigenvalues

Algorithm 1:

1. Compute $G, K_1, K_2 \in \mathbb{C}^{N_K \times N_K}$ and eigendecomposition $K_1 V = G V \Lambda$.
2. For each eigenpair (λ, \mathbf{v}) , compute $\text{res}(\lambda, \mathbf{v})$.
3. **Output:** subset of e-vectors $V_{(\varepsilon)}$ & e-vals $\Lambda_{(\varepsilon)}$ with $\text{res}(\lambda, \mathbf{v}) \leq \varepsilon$ ($\varepsilon = \text{input tol}$).

Theorem (no spectral pollution): Suppose quad. rule converges. Then

$$\limsup_{M \rightarrow \infty} \max_{\lambda \in \Lambda^{(\varepsilon)}} \|(\mathcal{K} - \lambda)^{-1}\|^{-1} \leq \varepsilon$$

BUT: Typically, does not capture all of spectrum! (“too little”)

ResDMD: avoiding “too little”

$$\text{Spec}_\varepsilon(\mathcal{K}) = \bigcup_{\|\mathcal{B}\| \leq \varepsilon} \text{Spec}(\mathcal{K} + \mathcal{B}), \quad \lim_{\varepsilon \downarrow 0} \text{Spec}_\varepsilon(\mathcal{K}) = \text{Spec}(\mathcal{K})$$

Algorithm 2:

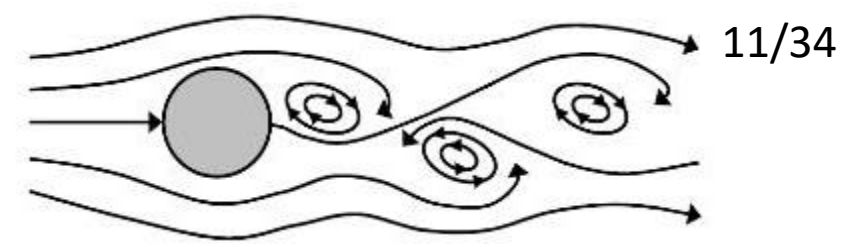
First convergent method for general \mathcal{K}

1. Compute $G, K_1, K_2 \in \mathbb{C}^{N_K \times N_K}$.
2. For z_k in comp. grid, compute $\tau_k = \min_{g = \sum_{j=1}^{N_K} \mathbf{g}_j \psi_j} \text{res}(z_k, g)$, corresponding g_k (gen. SVD).
3. **Output:** $\{z_k: \tau_k < \varepsilon\}$ (approx. of $\text{Spec}_\varepsilon(\mathcal{K})$), $\{g_k: \tau_k < \varepsilon\}$ (ε -pseudo-eigenfunctions).

Theorem (full convergence): Suppose the quadrature rule converges.

- **Error control:** $\{z_k: \tau_k < \varepsilon\} \subseteq \text{Spec}_\varepsilon(\mathcal{K})$ (as $M \rightarrow \infty$)
- **Convergence:** Converges locally uniformly to $\text{Spec}_\varepsilon(\mathcal{K})$ (as $N_K \rightarrow \infty$)

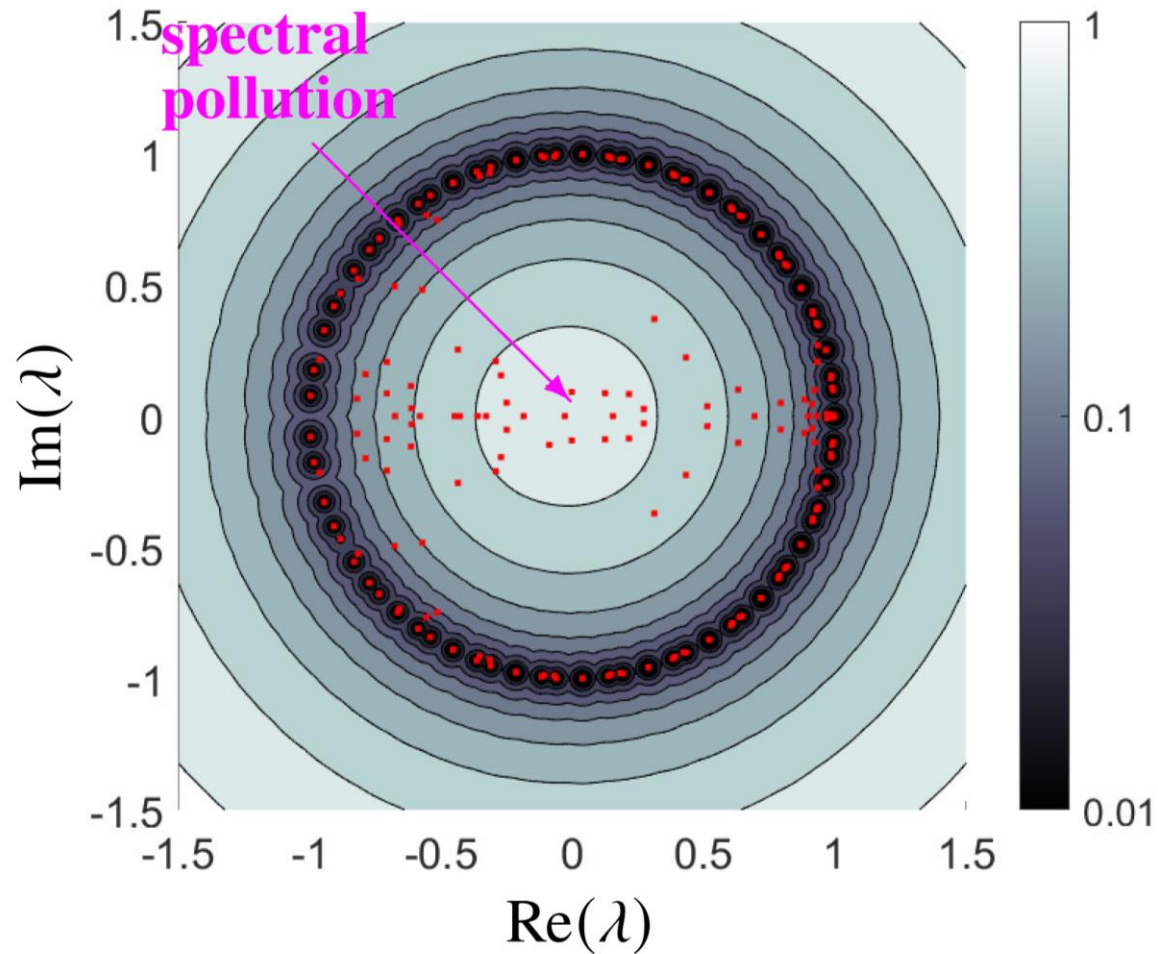
Example: Flow past a cylinder wake



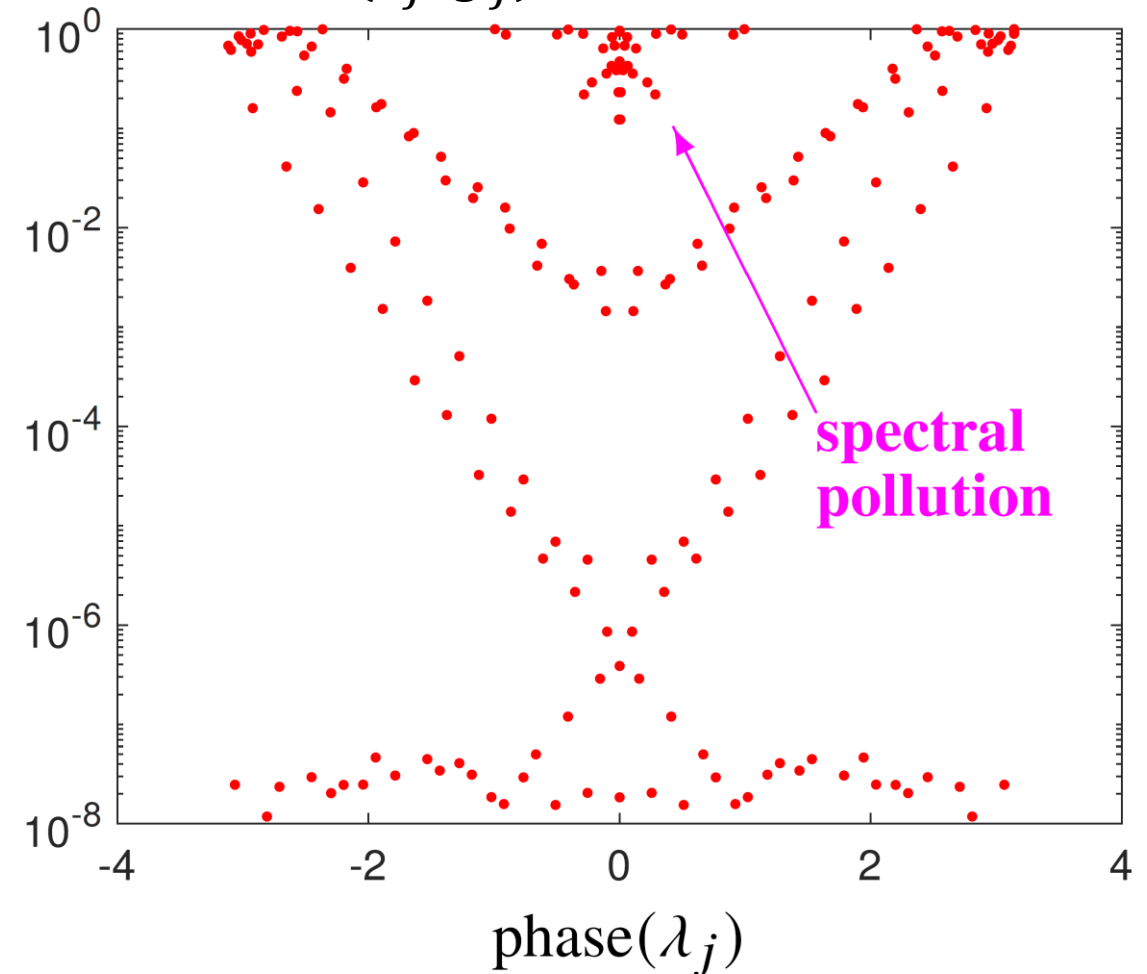
11/34

$Re = 100$, Dimension (d) = 80,000 (vel. at grid points)

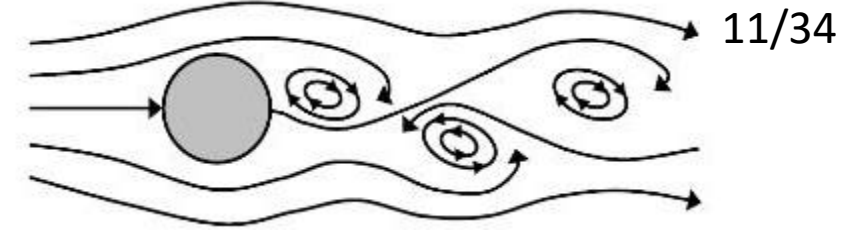
Pseudospectra, linear dictionary



$\text{res}(\lambda_j, g_j)$, linear dictionary



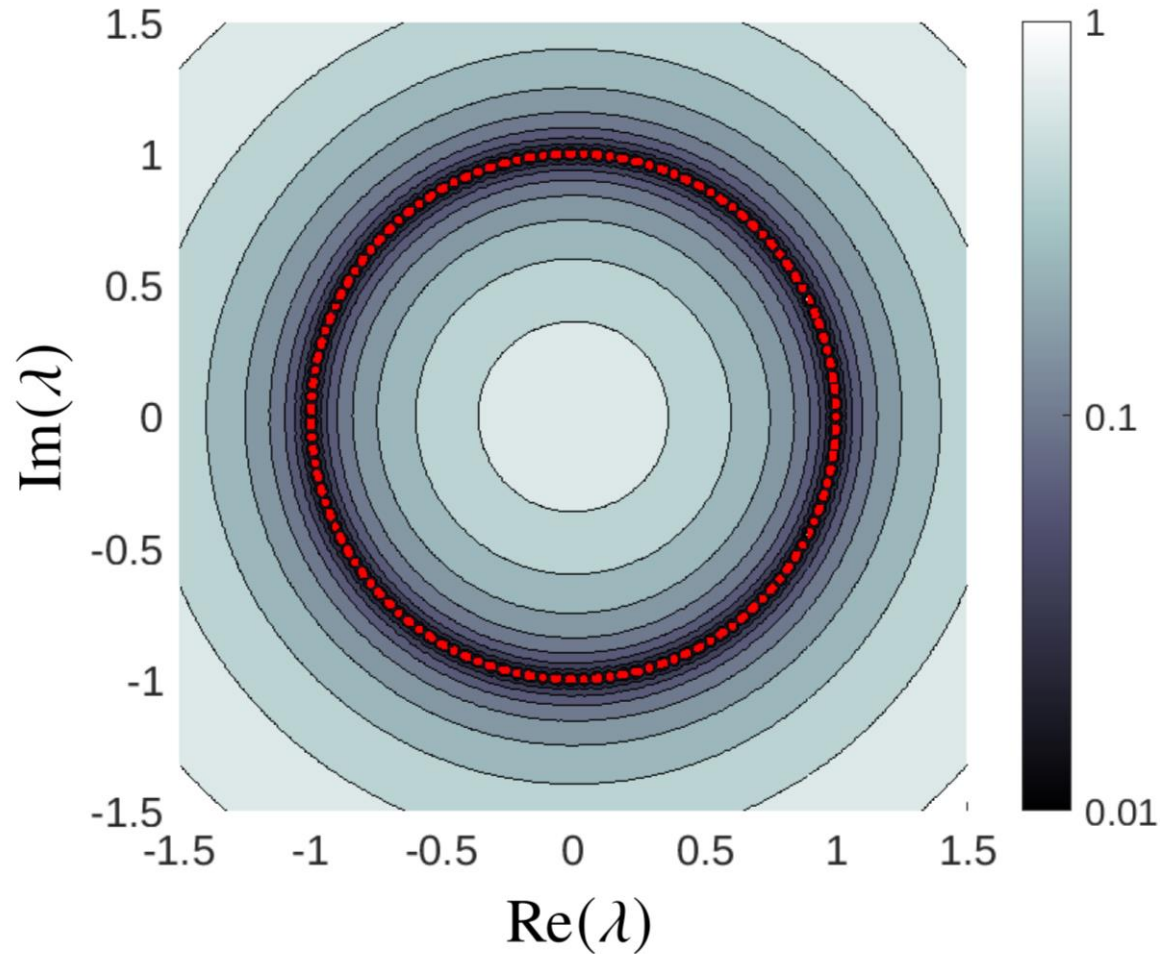
Example: Flow past a cylinder wake



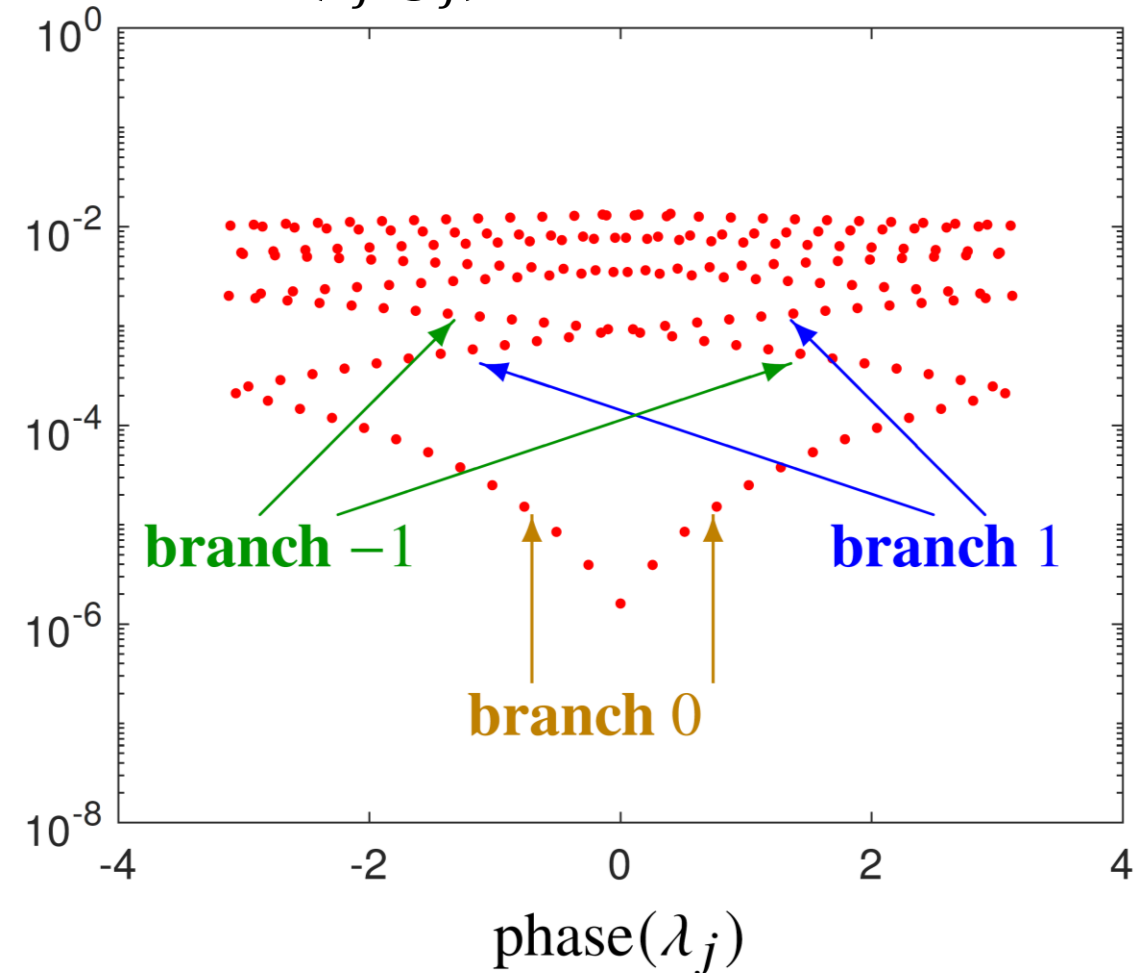
11/34

$Re = 100$, Dimension (d) = 80,000 (vel. at grid points)

Pseudospectra, nonlinear dictionary

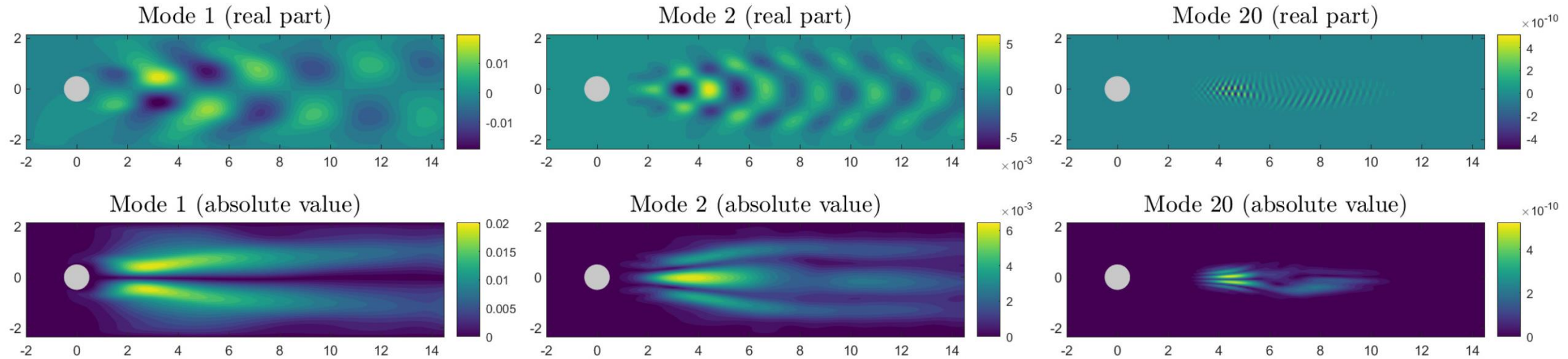


$res(\lambda_j, g_j)$, nonlinear dictionary

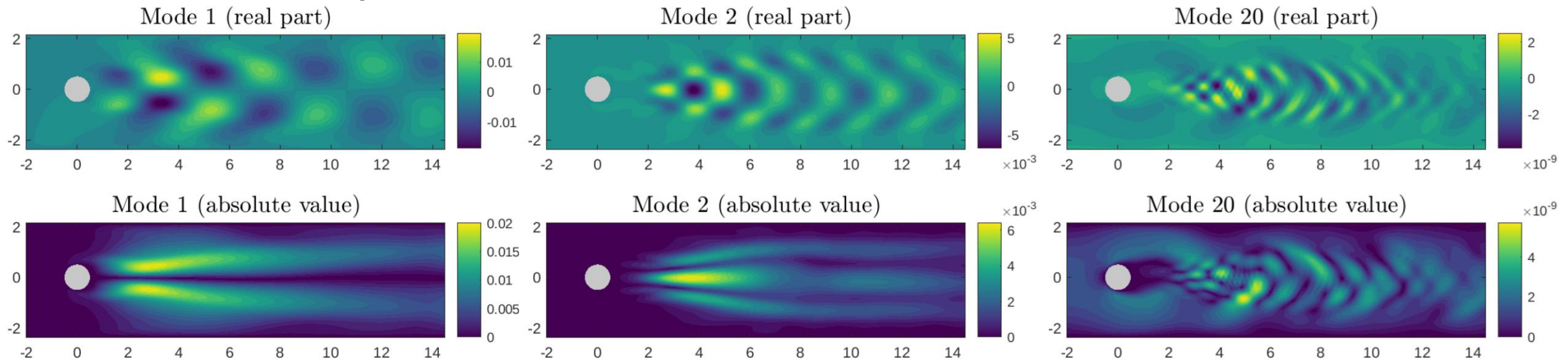


Koopman Modes

Linear dictionary



Nonlinear dictionary



Quadrature with trajectory data

$$\text{E.g., } \langle \mathcal{K}\psi_k, \psi_j \rangle = \lim_{M \rightarrow \infty} \sum_{m=1}^M w_m \overline{\psi_j(x^{(m)})} \underbrace{\psi_k(y^{(m)})}_{[\mathcal{K}\psi_k](x^{(m)})}$$

Three examples:

- **High-order quadrature:** $\{x^{(m)}, w_m\}_{m=1}^M$ M -point quadrature rule.
 Rapid convergence. Requires free choice of $\{x^{(m)}\}_{m=1}^M$ and small d .
- **Random sampling:** $\{x^{(m)}\}_{m=1}^M$ selected at random.
 Large d . Slow Monte Carlo $O(M^{-1/2})$ rate of convergence. ← Most common
- **Ergodic sampling:** $x^{(m+1)} = F(x^{(m)})$.
 Single trajectory, large d . Requires ergodicity, convergence can be slow. ↘

The Challenges

- ~~1) “Too much”: Approximate spurious modes $\lambda \notin \text{Spec}(\mathcal{K})$~~ ✓
- ~~2) “Too little”: Miss parts of $\text{Spec}(\mathcal{K})$~~ ✓
- 3) Continuous spectra.

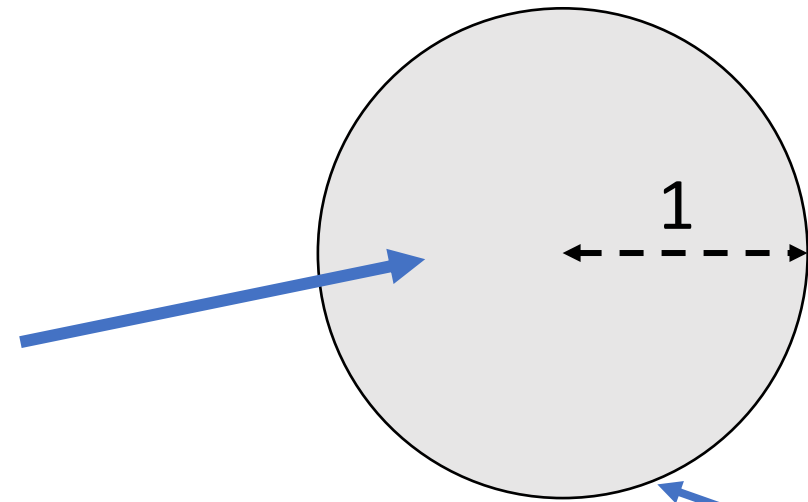
Verification: Is it right?

Setup for continuous spectra

Suppose system is measure preserving (e.g., Hamiltonian, ergodic, post-transient etc.)

$$\Leftrightarrow \mathcal{K}^* \mathcal{K} = I \text{ (isometry)}$$

$$\Rightarrow \text{Spec}(\mathcal{K}) \subseteq \{z: |z| \leq 1\}$$



(NB: we consider unitary extensions via Wold decomposition.)

spectral
measure
supp. on
boundary

Spectral decomposition of operators

$A \in \mathbb{C}^{n \times n}$ normal \Rightarrow O.N. basis of eigenvectors v_1, \dots, v_n :

$$v = \left(\sum_{k=1}^n \underset{\substack{\uparrow \\ \text{Projector onto Span}(v_k)}}{v_k v_k^*} \right) v, \quad Av = \left(\sum_{k=1}^n \underset{\substack{\uparrow \\ \text{eigenvalues}}}{\lambda_k} v_k v_k^* \right) v, \quad v \in \mathbb{C}^n$$

Spectral decomposition of operators

$A \in \mathbb{C}^{n \times n}$ normal \Rightarrow O.N. basis of eigenvectors v_1, \dots, v_n :

$$v = \left(\sum_{k=1}^n \underset{\substack{\uparrow \\ \text{Projector onto Span}(v_k)}}{v_k v_k^*} \right) v, \quad Av = \left(\sum_{k=1}^n \underset{\substack{\uparrow \\ \text{eigenvalues}}}{\lambda_k} v_k v_k^* \right) v, \quad v \in \mathbb{C}^n$$

Energy of “v” in each eigenvector: $\mu_v(\lambda_j) = \langle v_j v_j^* v, v \rangle = |v_j^* v|^2$

This is called the spectral measure with respect to a vector v .

Spectral decomposition of operators

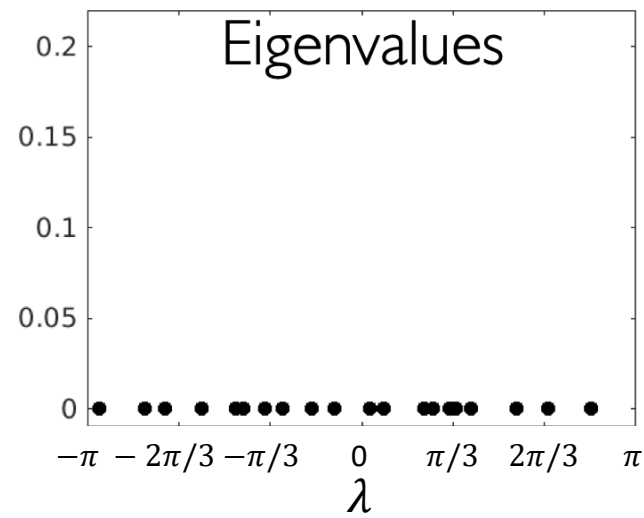
$A \in \mathbb{C}^{n \times n}$ normal \Rightarrow O.N. basis of eigenvectors v_1, \dots, v_n :

$$v = \left(\sum_{k=1}^n v_k v_k^* \right) v, \quad Av = \left(\sum_{k=1}^n \lambda_k v_k v_k^* \right) v, \quad v \in \mathbb{C}^n$$

↑
Projector onto $\text{Span}(v_k)$
↑
eigenvalues

Energy of “v” in each eigenvector: $\mu_v(\lambda_j) = \langle v_j v_j^* v, v \rangle = |v_j^* v|^2$

This is called the spectral measure with respect to a vector v .



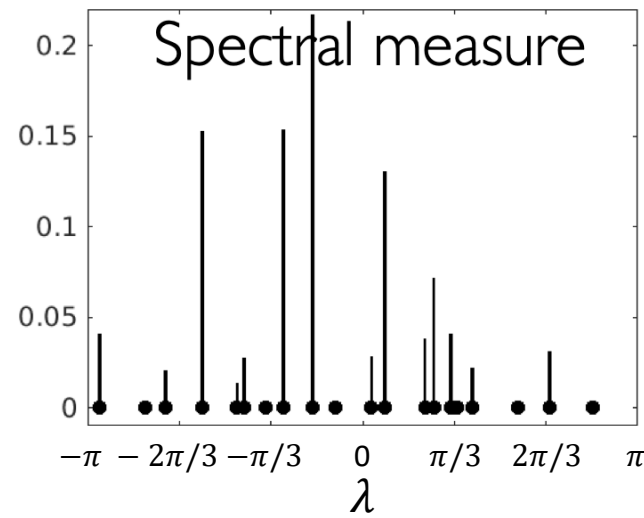
Spectral decomposition of operators

$A \in \mathbb{C}^{n \times n}$ normal \Rightarrow O.N. basis of eigenvectors v_1, \dots, v_n :

$$v = \left(\sum_{k=1}^n \underset{\substack{\uparrow \\ \text{Projector onto Span}(v_k)}}{v_k v_k^*} \right) v, \quad Av = \left(\sum_{k=1}^n \underset{\substack{\uparrow \\ \text{eigenvalues}}}{\lambda_k} v_k v_k^* \right) v, \quad v \in \mathbb{C}^n$$

Energy of “v” in each eigenvector: $\mu_v(\lambda_j) = \langle v_j v_j^* v, v \rangle = |v_j^* v|^2$

This is called the spectral measure with respect to a vector v .



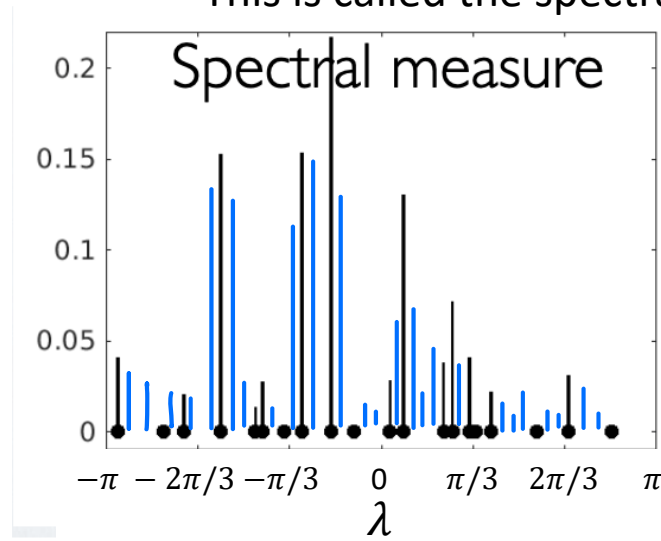
Spectral decomposition of operators

$A \in \mathbb{C}^{n \times n}$ normal \Rightarrow O.N. basis of eigenvectors v_1, \dots, v_n :

$$v = \left(\sum_{k=1}^n \underset{\substack{\uparrow \\ \text{Projector onto Span}(v_k)}}{v_k v_k^*} \right) v, \quad Av = \left(\sum_{k=1}^n \underset{\substack{\uparrow \\ \text{eigenvalues}}}{\lambda_k} v_k v_k^* \right) v, \quad v \in \mathbb{C}^n$$

Energy of “v” in each eigenvector: $\mu_v(\lambda_j) = \langle v_j v_j^* v, v \rangle = |v_j^* v|^2$

This is called the spectral measure with respect to a vector v .



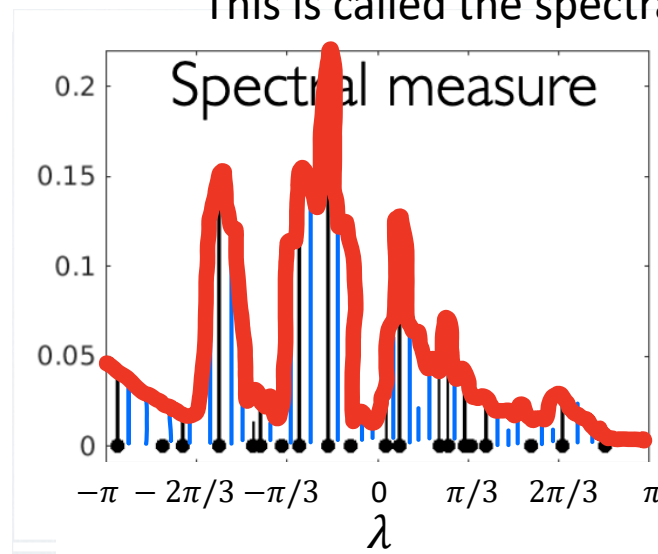
Spectral decomposition of operators

$A \in \mathbb{C}^{n \times n}$ normal \Rightarrow O.N. basis of eigenvectors v_1, \dots, v_n :

$$v = \left(\sum_{k=1}^n \underset{\substack{\uparrow \\ \text{Projector onto Span}(v_k)}}{v_k v_k^*} \right) v, \quad Av = \left(\sum_{k=1}^n \underset{\substack{\uparrow \\ \text{eigenvalues}}}{\lambda_k v_k v_k^*} \right) v, \quad v \in \mathbb{C}^n$$

Energy of “v” in each eigenvector: $\mu_v(\lambda_j) = \langle v_j v_j^* v, v \rangle = |v_j^* v|^2$

This is called the spectral measure with respect to a vector v .



Spectral decomposition of operators

$A \in \mathbb{C}^{n \times n}$ normal \Rightarrow O.N. basis of eigenvectors v_1, \dots, v_n :

$$v = \left(\sum_{k=1}^n \underset{\substack{\uparrow \\ \text{Projector onto Span}(v_k)}}}{v_k v_k^*} \right) v, \quad Av = \left(\sum_{k=1}^n \underset{\substack{\uparrow \\ \text{eigenvalues}}}{\lambda_k} v_k v_k^* \right) v, \quad v \in \mathbb{C}^n$$

Energy of “v” in each eigenvector: $\mu_v(\lambda_j) = \langle v_j v_j^* v, v \rangle = |v_j^* v|^2$

This is called the spectral measure with respect to a vector v .

\mathcal{K} is unitary \Rightarrow projection-valued measure ξ

$$g = \left(\int_{\mathbb{T}} d\xi(y) \right) g, \quad \mathcal{K}g = \left(\int_{\mathbb{T}} y d\xi(y) \right) g$$

Spectral measure $\nu_g(B) = \langle \xi(B)g, g \rangle$

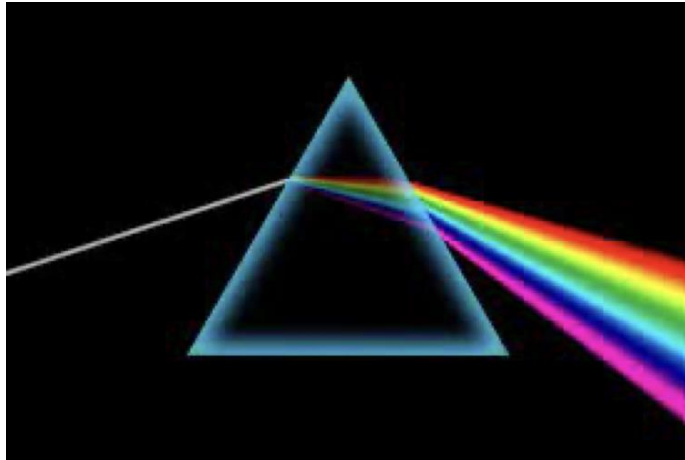
Spectral decomposition of operators

$A \in \mathbb{C}^{n \times n}$ normal \Rightarrow

\Rightarrow

O

White light contains a continuous spectra



$v,$

A

$\text{an}(v_k)$

eigenvector:

This is ca

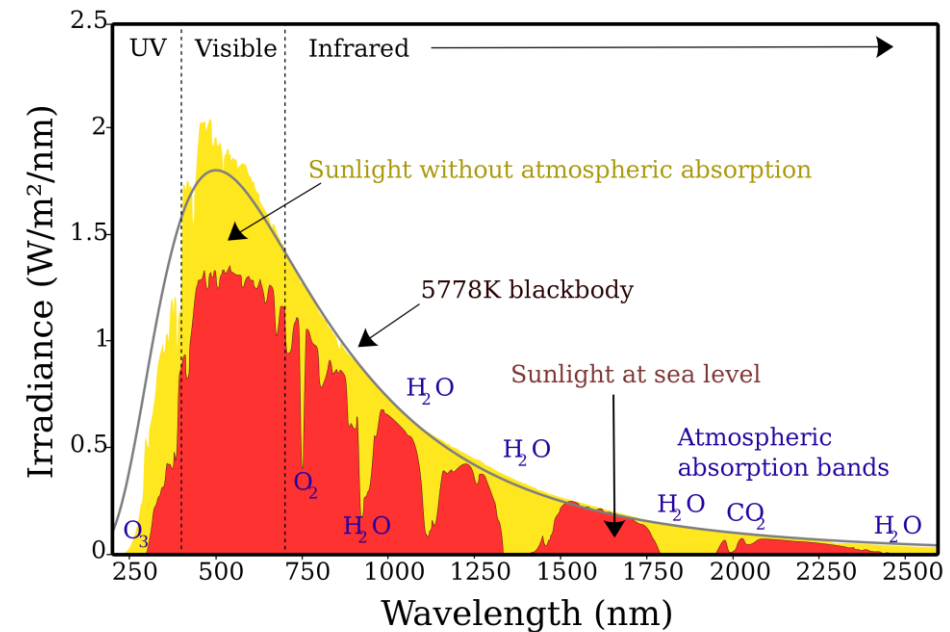
\Rightarrow

$$g = \left(\int_{\mathbb{T}} d\xi(y) \right) g,$$

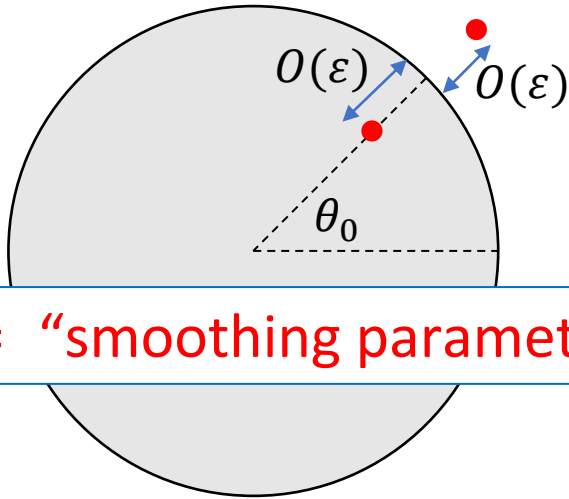
Spectral measure $\nu_g(B) = \langle \xi(B)g, g \rangle$

Often interesting to look at
the intensity of each wavelength

Spectrum of Solar Radiation (Earth)



Evaluating spectral measure



$\varepsilon =$ “smoothing parameter”

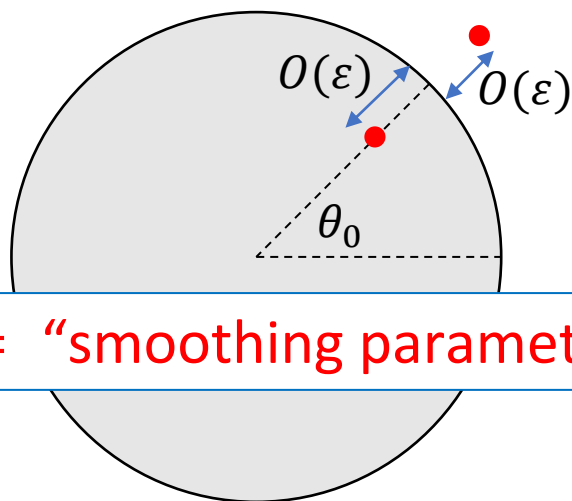
$$[P_\varepsilon * \nu_g](\theta_0) = \int_{[-\pi, \pi]_{\text{per}}} P_\varepsilon(\theta_0 - \theta) d\nu_g(\theta)$$

Smoothing convolution

Poisson kernel for
unit disk

$$P_\varepsilon(\theta_0) = \frac{1}{2\pi} \frac{(1 + \varepsilon)^2 - 1}{1 + (1 + \varepsilon)^2 - 2(1 + \varepsilon)\cos(\theta_0)}$$

Evaluating spectral measur



$\varepsilon = \text{"smoothing parameter"}$

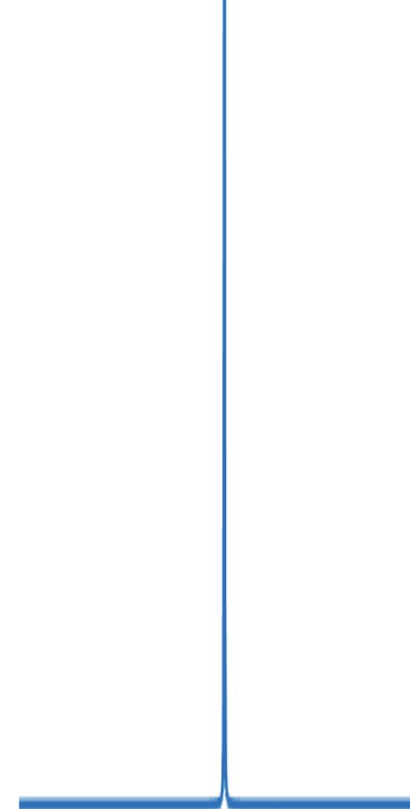
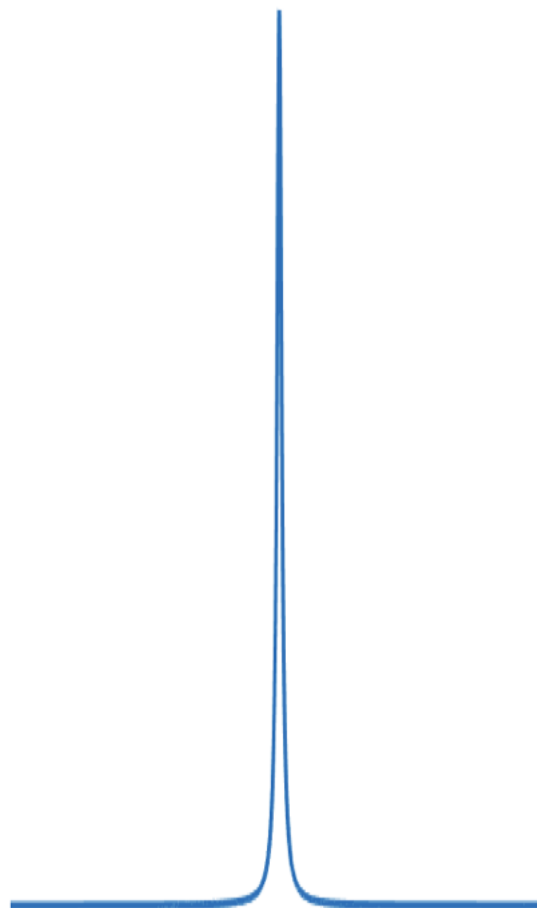
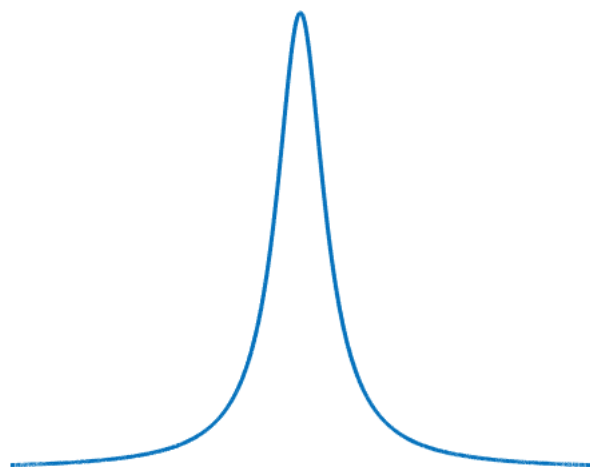
Poisso
unit di

$$[P_\varepsilon * \nu_g](\theta_0) = \int_{\mathbb{T}} P_\varepsilon(\theta_0 - \theta) d\nu_g(\theta)$$

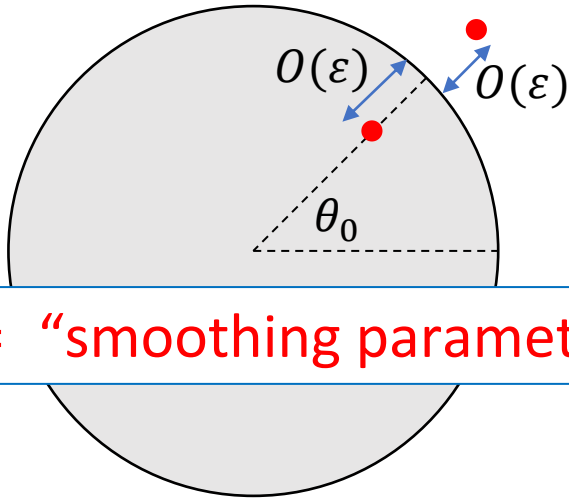
Smoothing co

$$\frac{1}{1 + \varepsilon}$$

$\overline{0}$



Evaluating spectral measure



$\varepsilon =$ “smoothing parameter”

Smoothing convolution

$$[P_\varepsilon * \nu_g](\theta_0) = \int_{[-\pi, \pi]_{\text{per}}} P_\varepsilon(\theta_0 - \theta) d\nu_g(\theta)$$

Poisson kernel for
unit disk

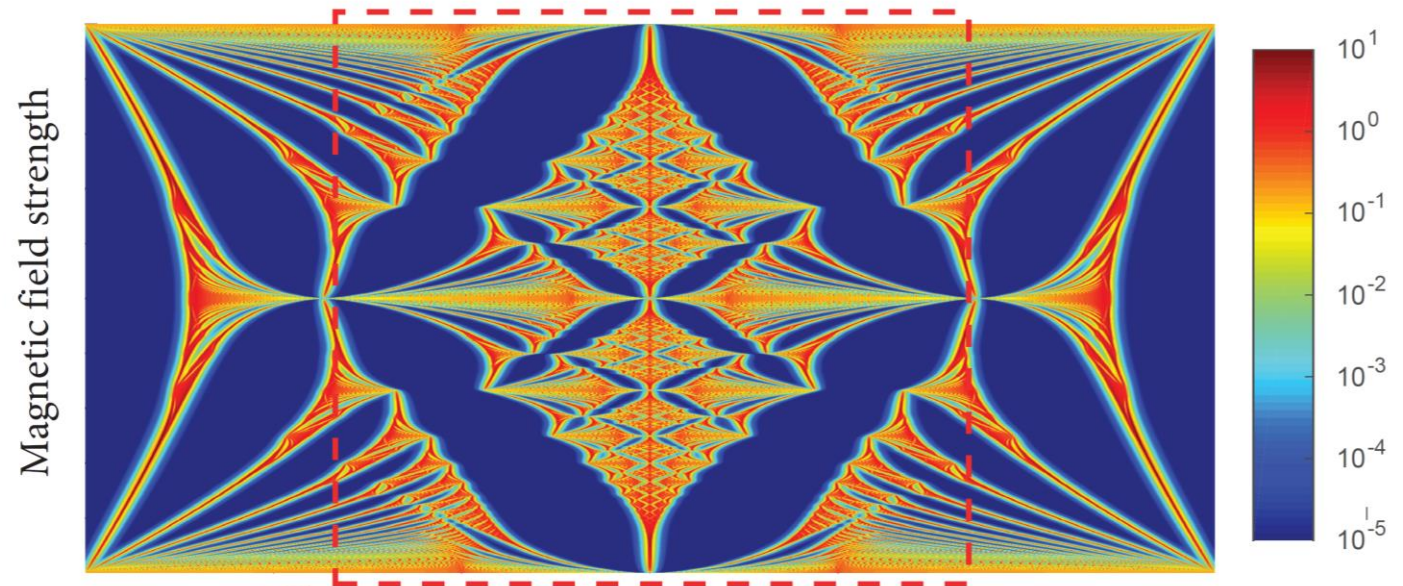
$$P_\varepsilon(\theta_0) = \frac{1}{2\pi} \frac{(1 + \varepsilon)^2 - 1}{1 + (1 + \varepsilon)^2 - 2(1 + \varepsilon)\cos(\theta_0)}$$

$$[P_\varepsilon * \nu_g](\theta_0) = \mathcal{C}_g(e^{i\theta_0}(1 + \varepsilon)^{-1}) - \mathcal{C}_g(e^{i\theta_0}(1 + \varepsilon))$$

$$\mathcal{C}_g(z) = \int_{[-\pi, \pi]_{\text{per}}} \frac{e^{i\theta} d\nu_g(\theta)}{e^{i\theta} - z} = \begin{cases} \langle (\mathcal{K} - zI)^{-1}g, \mathcal{K}^*g \rangle, & \text{if } |z| > 1 \\ -z^{-1} \langle g, (\mathcal{K} - \bar{z}^{-1}I)^{-1}g \rangle, & \text{if } 0 < |z| < 1 \end{cases}$$

ResDMD computes
with error control

Spectral measures of self-adjoint operators



Horizontal slice = spectral measure at constant magnetic field strength.

Software package

SpecSolve available at <https://github.com/SpecSolve>
 Capabilities: ODEs, PDEs, integral operators, discrete operators.

Example

$$\mathcal{K} = \begin{pmatrix} \overline{\alpha_0} & \overline{\alpha_1}\rho_0 & \rho_0\rho_1 & & & \\ \rho_0 & -\overline{\alpha_1}\alpha_0 & -\alpha_0\rho_1 & & & \\ & \overline{\alpha_2}\rho_1 & -\overline{\alpha_2}\alpha_1 & \overline{\alpha_3}\rho_2 & \rho_3\rho_2 & \\ & \rho_2\rho_1 & -\alpha_1\rho_2 & -\overline{\alpha_3}\alpha_2 & -\rho_3\alpha_2 & \ddots \\ & & & \overline{\alpha_4}\rho_3 & -\overline{\alpha_4}\alpha_3 & \ddots \\ & & & \ddots & \ddots & \ddots \end{pmatrix}$$

$$\alpha_j = (-1)^j 0.95^{(j+1)/2}, \quad \rho_j = \sqrt{1 - |\alpha_j|^2}$$

Generalised shift, typical building block of many dynamical systems.

Fix N_K , vary ε : unstable!

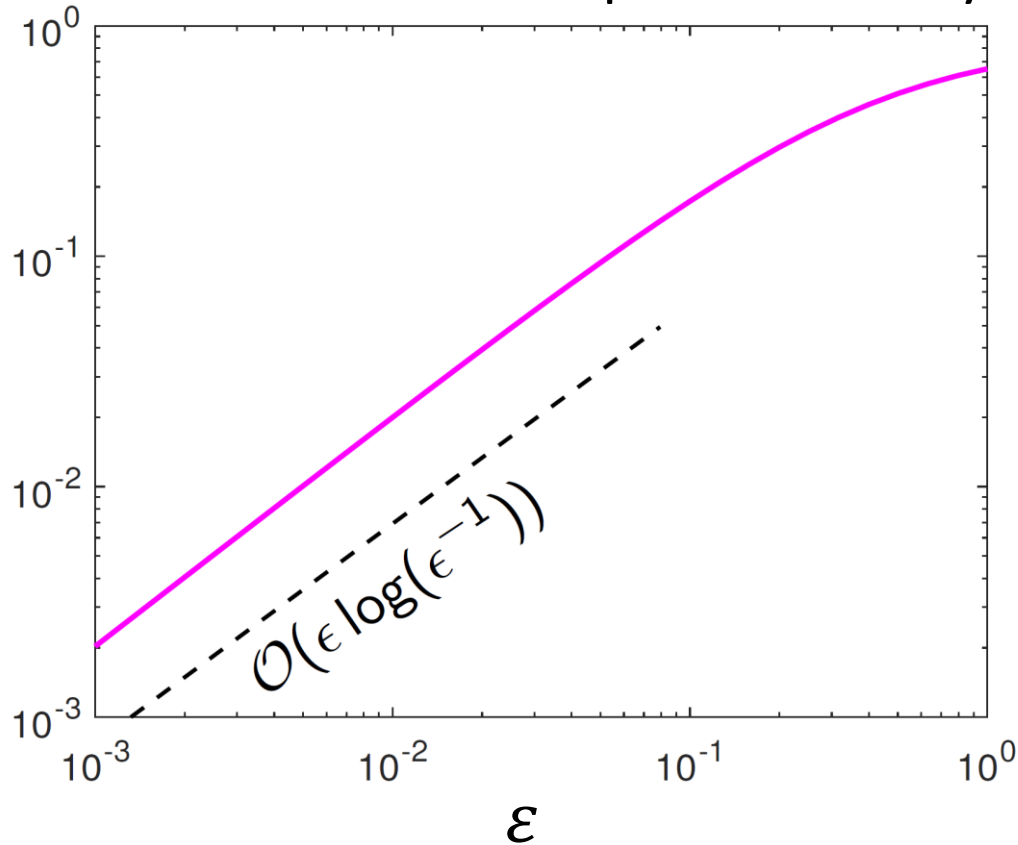
Fix ε , vary N_K : too smooth!

Adaptive: new matrix to compute residuals crucial

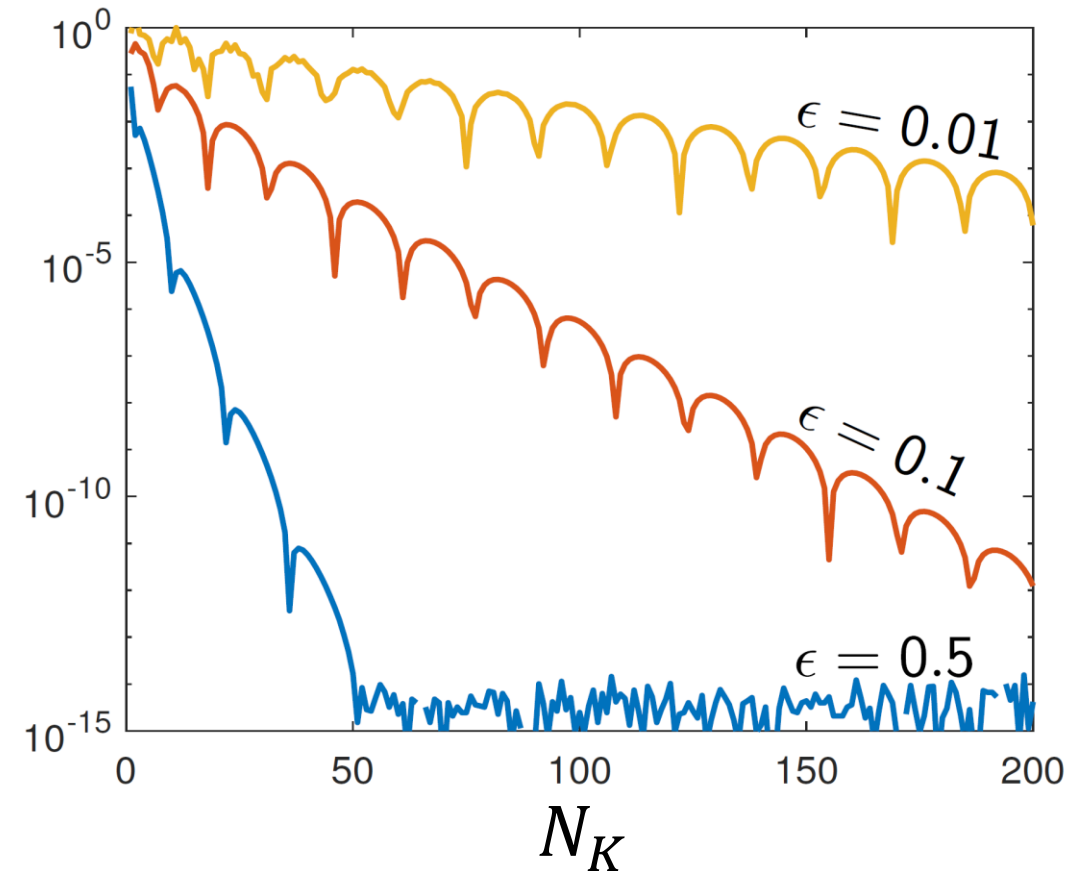
But ... slow convergence

Problem: As $\varepsilon \downarrow 0$, error is $O(\varepsilon \log(1/\varepsilon))$ and $N_K(\varepsilon) \rightarrow \infty$.

Pointwise error for spectral density



Error due to discretization

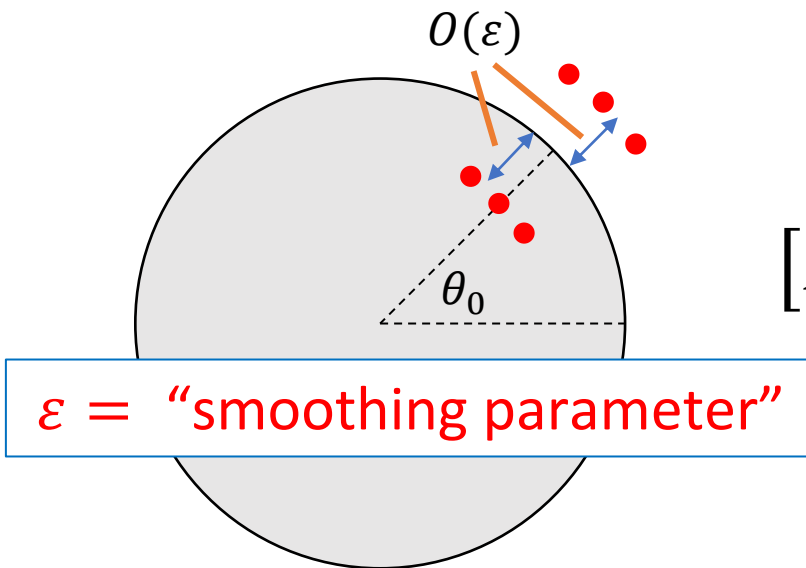


Small N_K critical in data-driven computations. Can we improve convergence rate?

High-order rational kernels

m th order rational kernels:

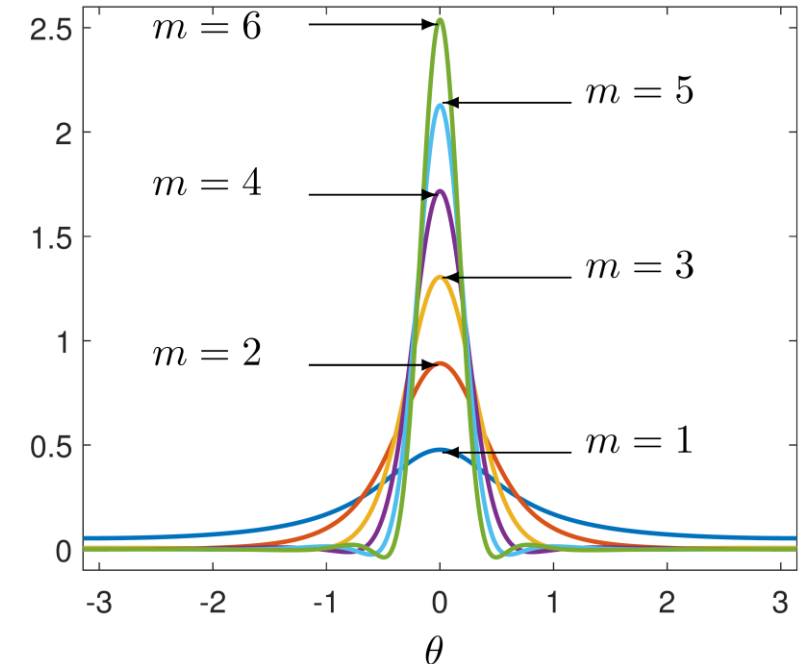
$$K_\varepsilon(\theta) = \frac{e^{-i\theta}}{2\pi} \sum_{j=1}^m \left[\frac{c_j}{e^{-i\theta} - (1 + \varepsilon \bar{z}_j)^{-1}} - \frac{d_j}{e^{-i\theta} - (1 + \varepsilon z_j)} \right]$$



ResDMD computes
with error control

$$[K_\varepsilon * v_g](\theta_0) = \sum_{j=1}^m \left[c_j \mathcal{C}_g(e^{i\theta_0}(1 + \varepsilon \bar{z}_j)^{-1}) - d_j \mathcal{C}_g(e^{i\theta_0}(1 + \varepsilon z_j)) \right]$$

Kernels



Smaller N_K (larger ε)

Convergence

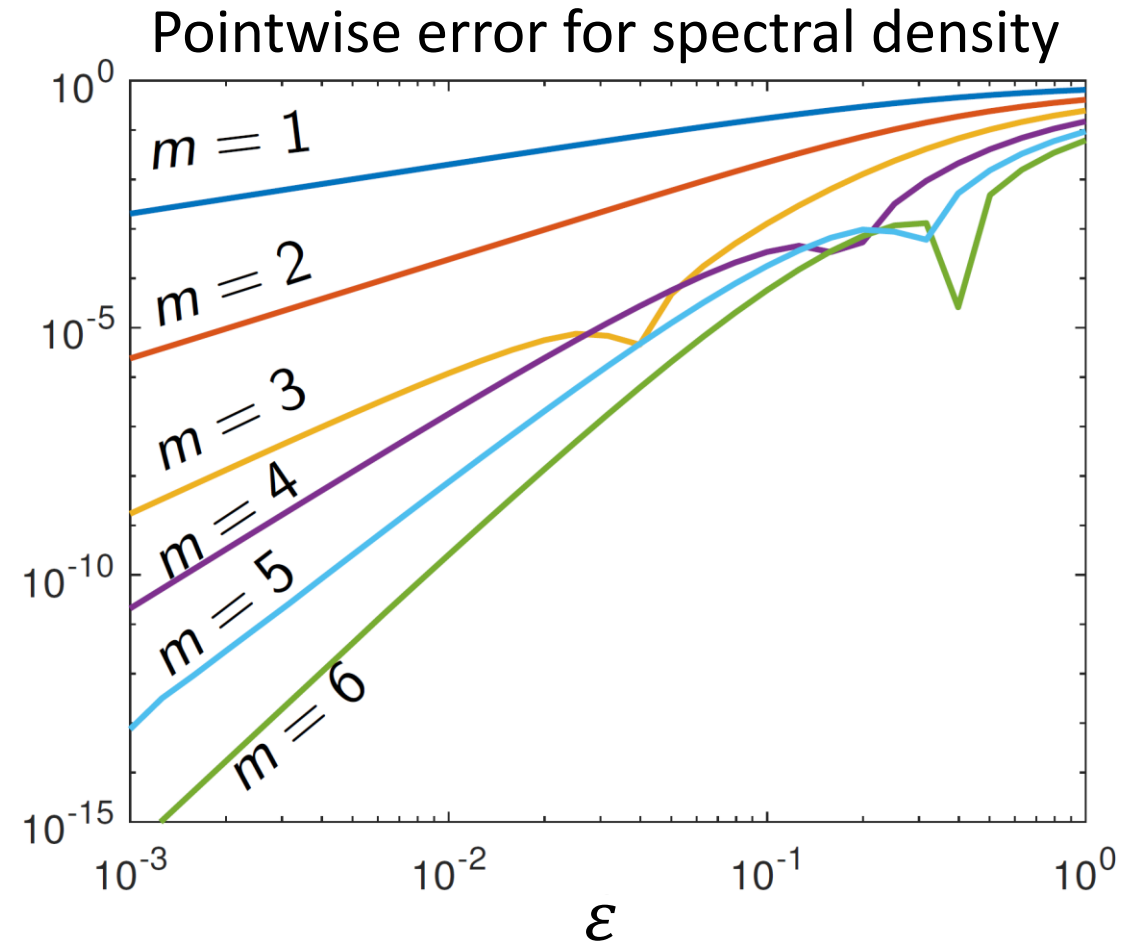
Theorem: Automatic selection of $N_K(\varepsilon)$ with $O(\varepsilon^m \log(1/\varepsilon))$ convergence:

- Density of continuous spectrum ρ_g .
(pointwise and L^p)
- Integration against test functions.
(weak convergence)

$$\int_{[-\pi, \pi]_{\text{per}}} h(\theta) [K_\varepsilon * \nu_g](\theta) \, d\theta$$

$$= \int_{[-\pi, \pi]_{\text{per}}} h(\theta) \, d\nu_g(\theta) + O(\varepsilon^m \log(1/\varepsilon))$$

Also recover discrete spectrum.



The Challenges

~~1) “Too much”: Approximate spurious modes $\lambda \notin \text{Spec}(\mathcal{K})$~~



~~2) “Too little”: Miss parts of $\text{Spec}(\mathcal{K})$~~

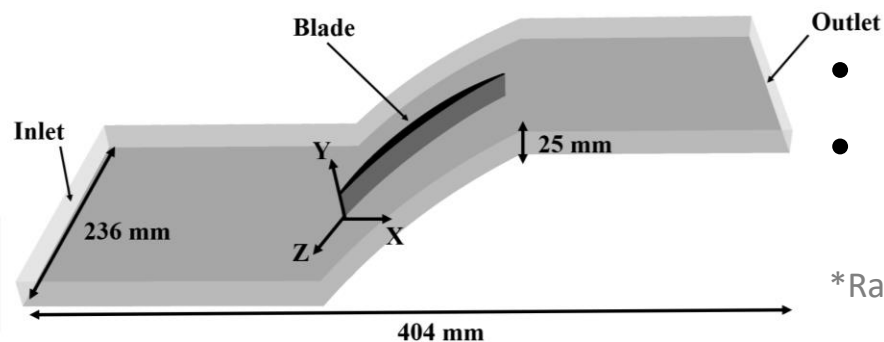
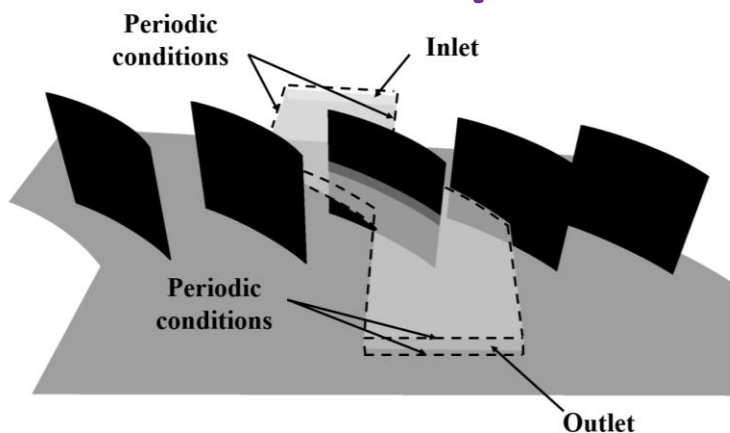


~~3) Continuous spectra.~~



Verification: Is it right?

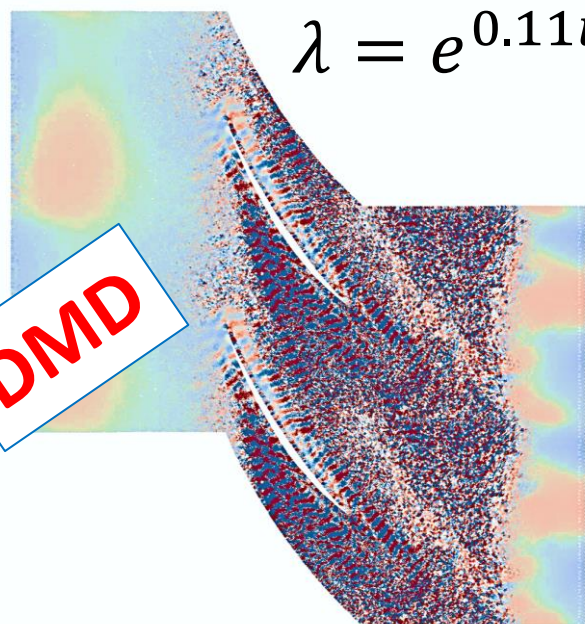
Example: Trustworthy computation for large d



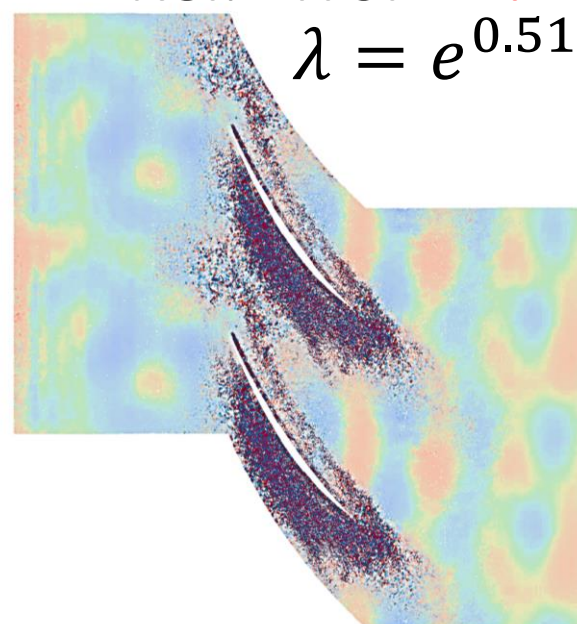
- Reynolds number $\approx 3.9 \times 10^5$
- Ambient dimension (d) $\approx 300,000$ (number of measurement points)

*Raw measurements provided by Stephane Moreau (Sherbrooke)

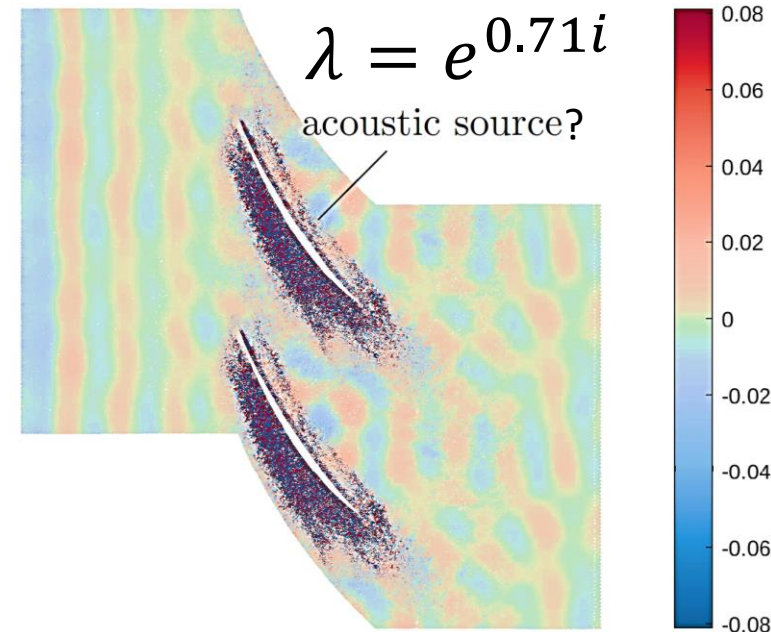
Rel. Error = ?
 $\lambda = e^{0.11i}$



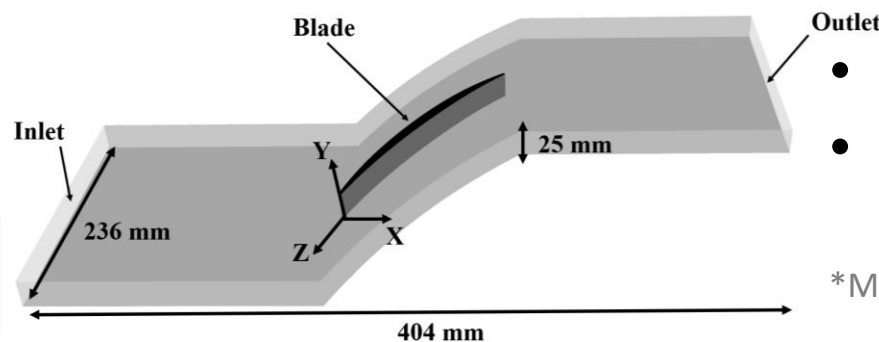
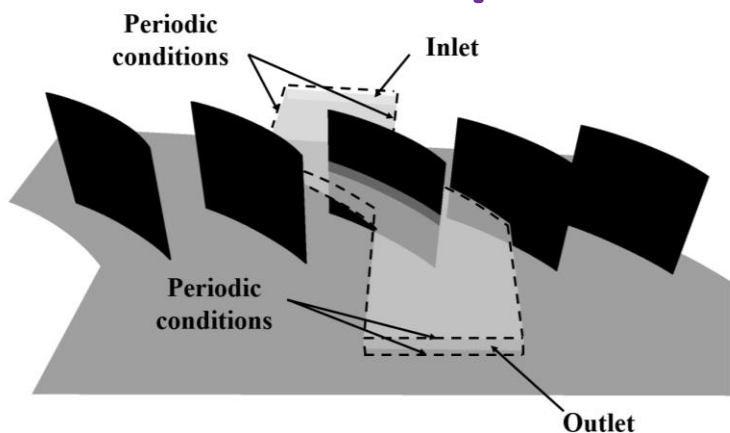
Rel. Error = ?
 $\lambda = e^{0.51i}$



Rel. Error = ?
 $\lambda = e^{0.71i}$



Example: Trustworthy computation for large d



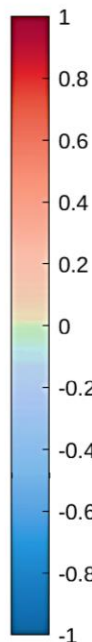
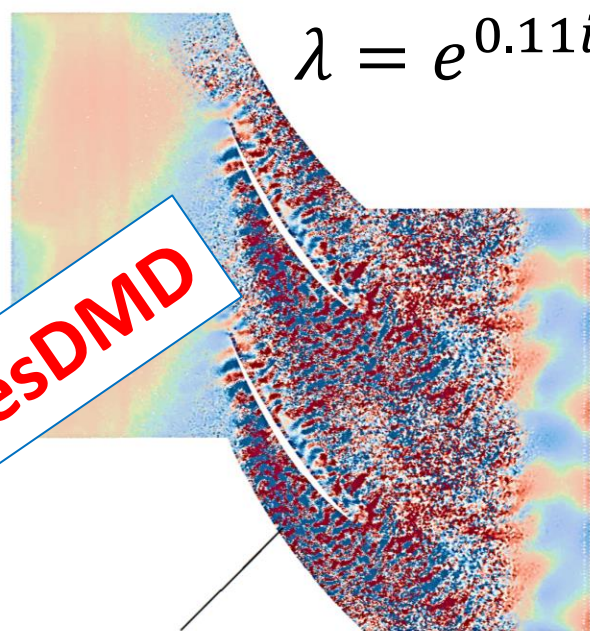
- Reynolds number $\approx 3.9 \times 10^5$
- Ambient dimension (d) $\approx 300,000$ (number of measurement points)

*Measurements provided by Stephane Moreau (Sherbrooke)

Rel. Error ≤ 0.0054

$$\lambda = e^{0.11i}$$

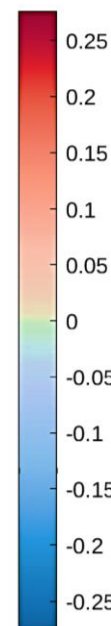
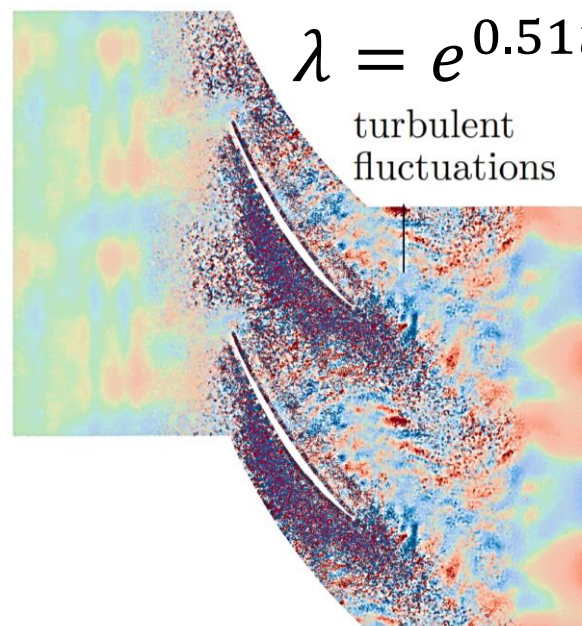
ResDMD



Rel. Error ≤ 0.0128

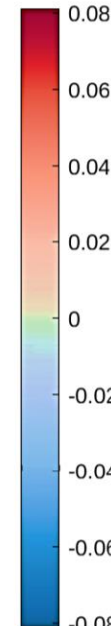
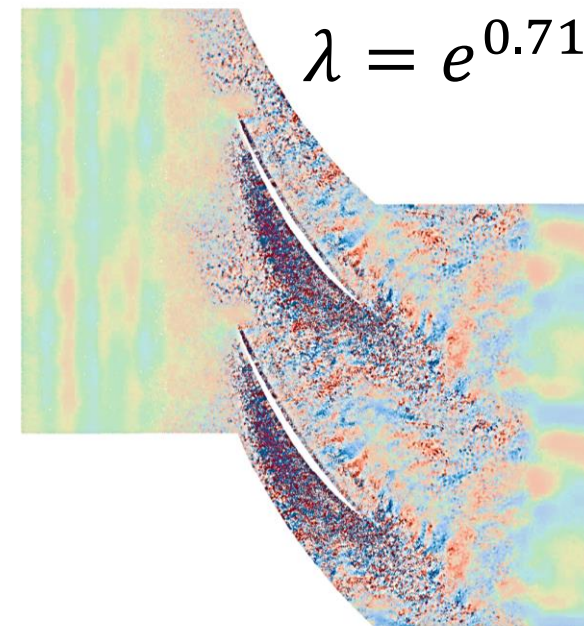
$$\lambda = e^{0.51i}$$

turbulent
fluctuations



Rel. Error ≤ 0.0196

$$\lambda = e^{0.71i}$$



acoustic vibrations

- C., Townsend, "Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems," preprint.

Large d ($\Omega \subseteq \mathbb{R}^d$): robust and scalable

Popular to learn dictionary $\{\psi_1, \dots, \psi_{N_K}\}$

E.g., DMD with truncated SVD (linear dictionary, most popular),
kernel methods (this talk), neural networks, etc.

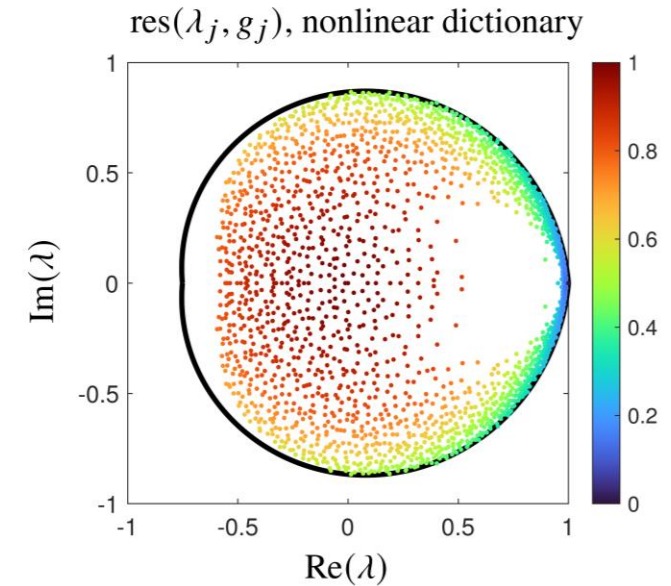
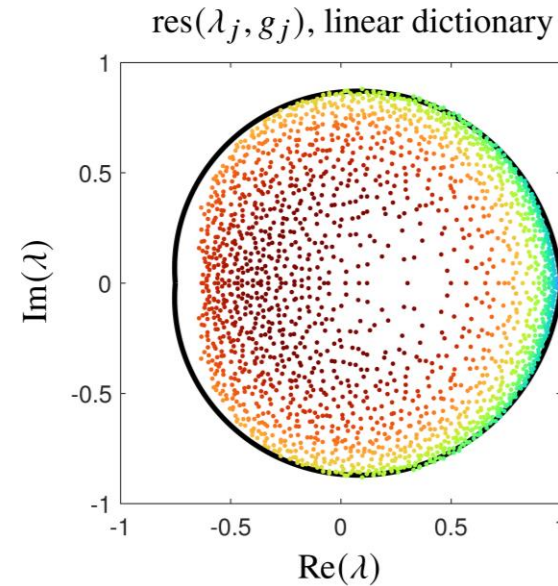
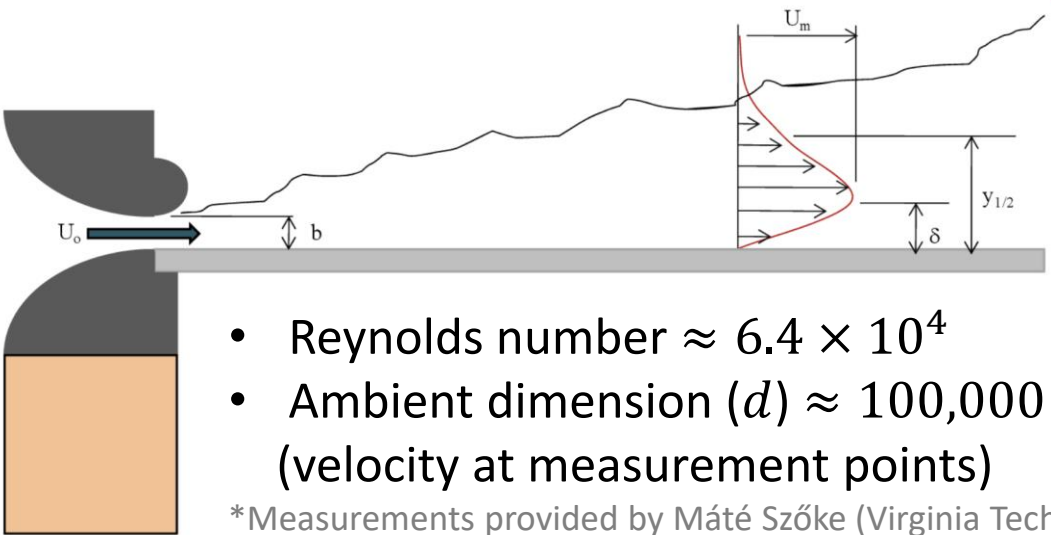
Q: Is discretisation $\text{span}\{\psi_1, \dots, \psi_{N_K}\}$ large/rich enough?

Above algorithms:

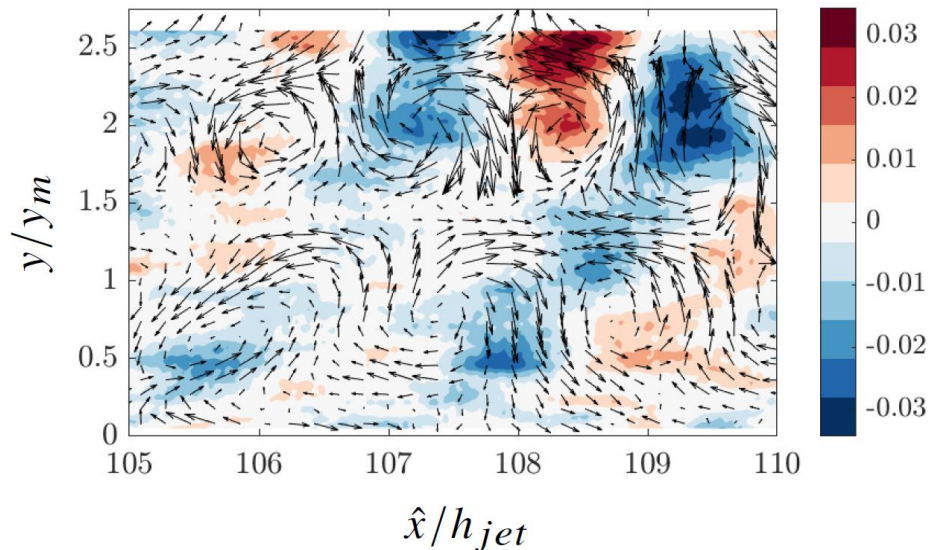
- Pseudospectra: $\{z_k : \tau_k < \varepsilon\} \subseteq \text{Spec}_\varepsilon(\mathcal{K})$ **error control**
- Spectral measures: $\mathcal{C}_g(z)$ and smoothed measures **adaptive check**

\Rightarrow Rigorously **verify** learnt dictionary $\{\psi_1, \dots, \psi_{N_K}\}$

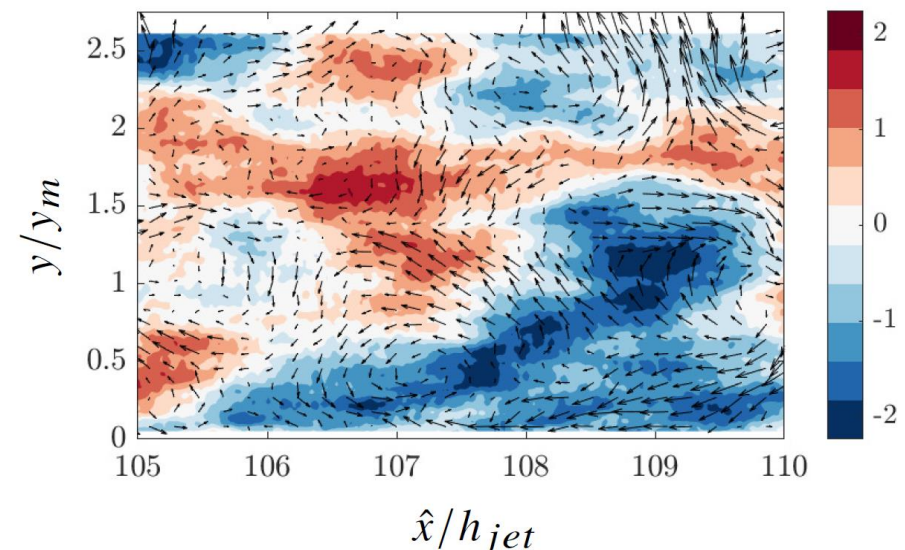
Example: Verify the dictionary



$$\lambda = 0.9439 + 0.2458i, \text{ error} \leq 0.0765$$

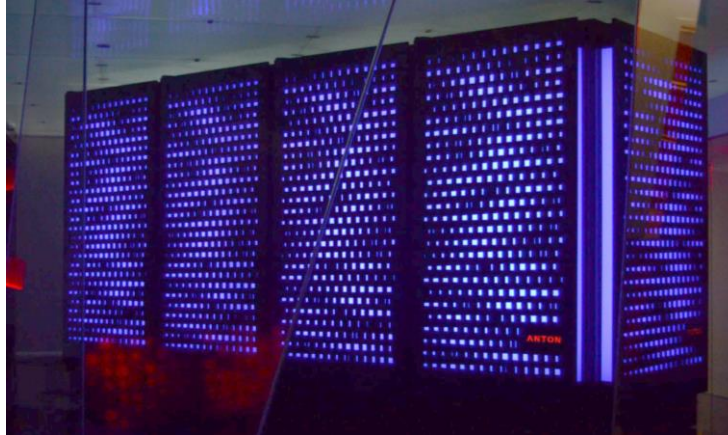
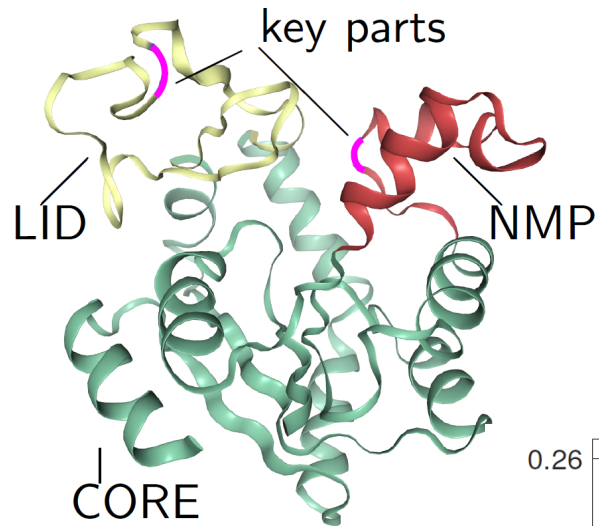


$$\lambda = 0.8948 + 0.1065i, \text{ error} \leq 0.1105$$



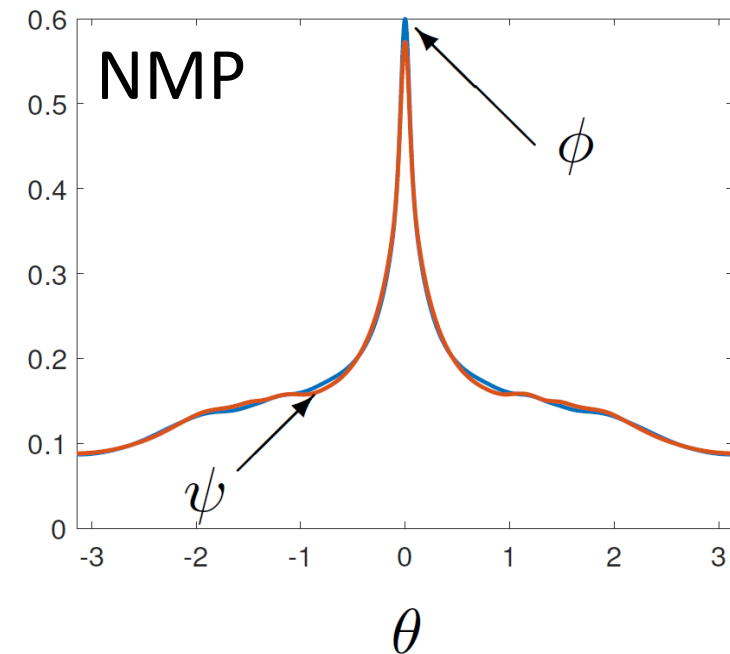
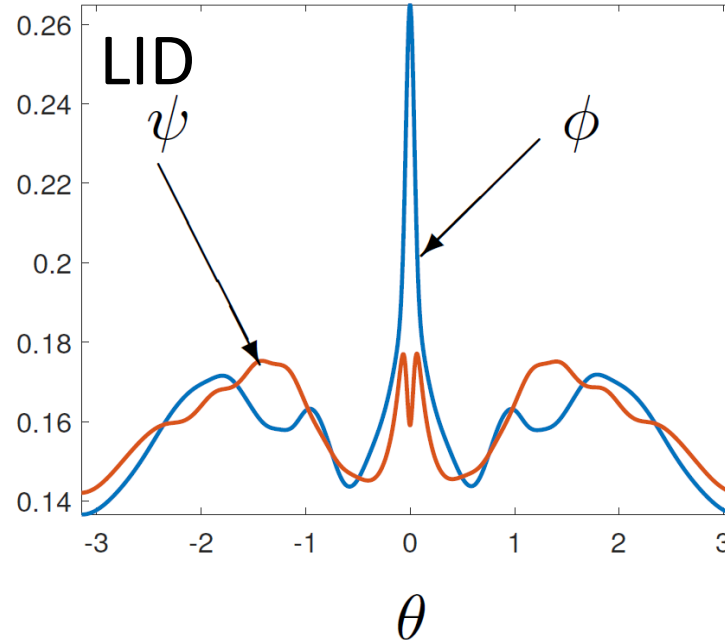
Example: Spectral measures in large d

Adenylate Kinase

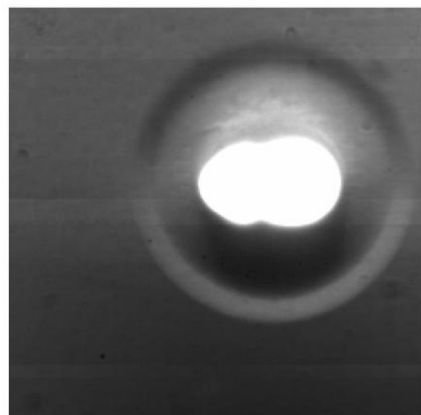
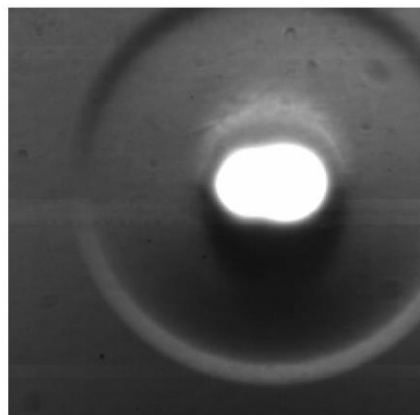
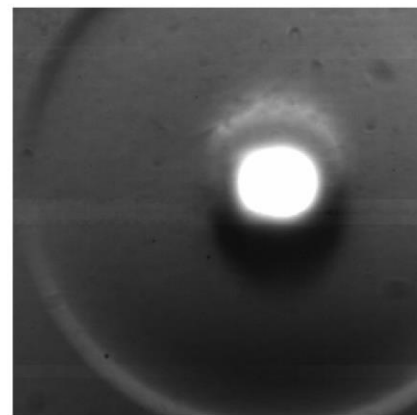
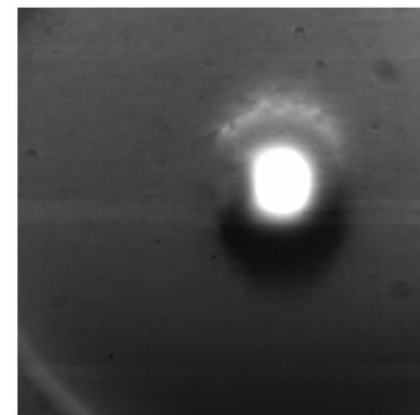
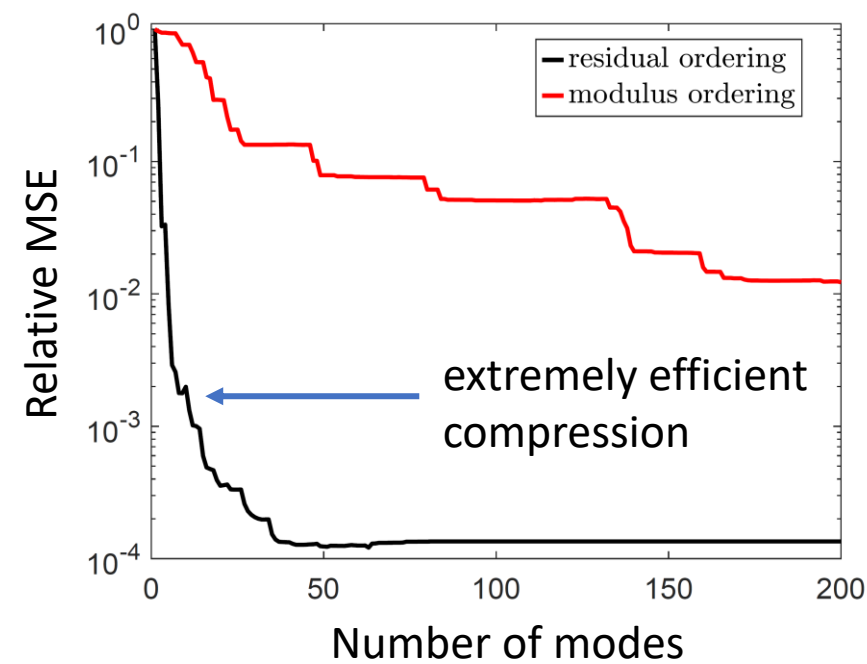
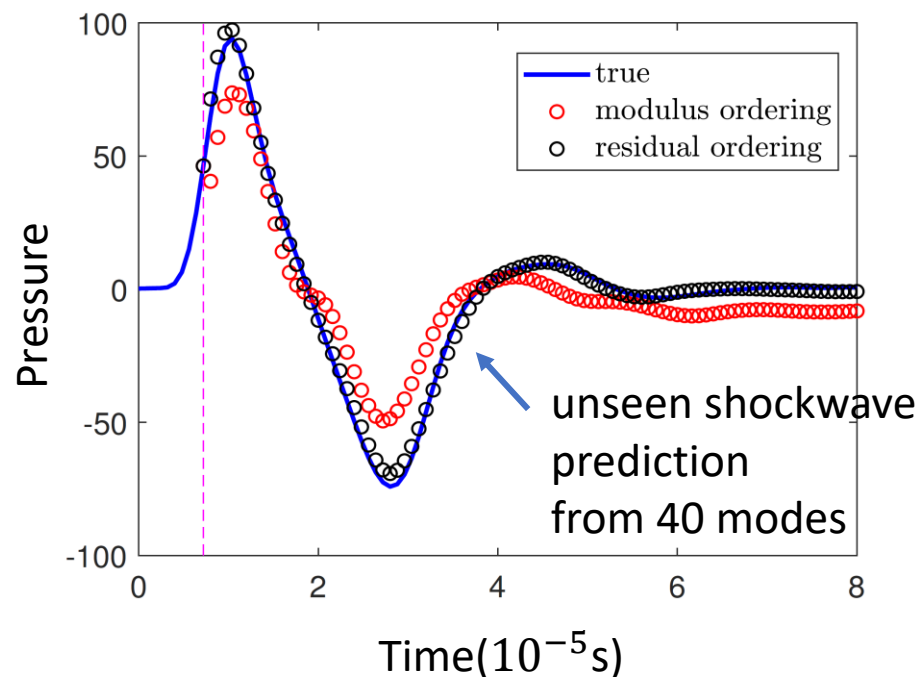


- Ambient dimension (d) $\approx 20,000$ (positions and momenta of atoms)
- 6th order kernel (spec res 10^{-6})

*Dataset: www.mdanalysis.org/MDAnalysisData/adk_equilibrium.html



Example: Trustworthy Koopman mode decomposition

a) $t = 5 \mu\text{s}$ b) $t = 10 \mu\text{s}$ c) $t = 15 \mu\text{s}$ d) $t = 20 \mu\text{s}$ 

- C., Ayton, Szőke, "Residual Dynamic Mode Decomposition," *J. Fluid Mech.*, 2023.

Wider programme

- Inf.-dim. computational analysis \Rightarrow **Compute spectral properties rigorously.**
- Continuous linear algebra \Rightarrow **Avoid the woes of discretization**
- Solvability Complexity Index hierarchy \Rightarrow **Classify diff. of comp. problems, prove algs are optimal.**
- **Extends to:** Foundations of AI, optimization, computer-assisted proofs, and PDE learning.

-
- C., “On the computation of geometric features of spectra of linear operators on Hilbert spaces,” **Found. Comput. Math.**, 2023.
 - C., Horning, Townsend “Computing spectral measures of self-adjoint operators,” **SIAM Rev.**, 2021.
 - C., Hansen, “The foundations of spectral computations via the solvability complexity index hierarchy,” **J. Eur. Math. Soc.**, 2022.
 - C., Antun, Hansen, “The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale’s 18th problem,” **Proc. Natl. Acad. Sci. USA**, 2022.
 - C., “Computing spectral measures and spectral types,” **Comm. Math. Phys.**, 2021.
 - C., Roman, Hansen, “How to compute spectra with error control,” **Phys. Rev. Lett.**, 2019.
 - C., “Computing semigroups with error control,” **SIAM J. Numer. Anal.**, 2022.
 - Boullé, Townsend, “Learning elliptic partial differential equations with randomized linear algebra”, **Found. Comput. Math.**, 2022.
 - Boullé, Kim, Shi, Townsend, “Learning Green’s functions associated with parabolic partial differential equations”, **JMLR**, to appear.
 - Gilles, Townsend, “Continuous analogues of Krylov methods for differential operators,” **SIAM J. Numer. Anal.**, 2019.
 - Horning, Townsend, “FEAST for Differential Eigenvalue Problems,” **SIAM J. Numer. Anal.**, 2020.
 - Ben-Artzi, C., Hansen, Nevanlinna, Seidel, “On the solvability complexity index hierarchy and towers of algorithms,” arXiv, 2020.
 - Smale, “The fundamental theorem of algebra and complexity theory,” **Bull. Amer. Math. Soc.**, 1981.
 - McMullen, “Families of rational maps and iterative root-finding algorithms,” **Ann. of Math.**, 1987.

Rigorous data-driven Koopmanism!

“Too much” or “Too little”

Idea: New matrix for residual

⇒ **ResDMD** for computing spectra.

Continuous spectra

Idea: Smoothing via resolvent and **ResDMD**.

Is it right?

ResDMD verifies computations.

E.g., learned dictionaries.

Rigorous data-driven Koopmanism!

“Too much” or “Too little”

Idea: New matrix for residual

⇒ **ResDMD** for computing spectra.

Continuous spectra

Idea: Smoothing via resolvent and **ResDMD**.

Is it right?

ResDMD verifies computations.

E.g., learned dictionaries.

Optimization and Learning with Zeroth-order Stochastic Oracles

By Stefan M. Wild

Mathematical optimization is a foundational technology for machine learning and the solution of design, decision, and control problems. In most optimization applications, the principal assumption is the availability of at least the

sequence is that material properties are only available via *in situ* and *in operando* characterization. In the context of optimization, this scenario is called a “zeroth-order oracle”—our knowledge about a particular system or property is data driven and limited by the black-box nature of measurement procurement. An additional challenge is

An optimization solver specifies a particular composition of solvents and bases, an operating temperature, and reaction times; this combination is then run through a continuous flow reactor. The material that exits the reactor is then automatically characterized

through an inline nuclear magnetic resonance detector that illuminates properties of the synthesized materials. These stochastic, zeroth-order oracle outputs return to the solver in a closed-loop setting that

See *Optimization* on page 3

Read more about these breakthroughs in SIAM News!

search for novel materials for energy storage. In order to create viable new materials, we must move beyond pure theory and account for the actual processes that occur during materials synthesis. A necessary

ers—such as the laboratory environment in Figure 1—doing so is impossible. Figure 1 displays an instantiation of a data-driven optimization setting in a chemical lab at Argonne National Laboratory

Figure 1. A continuous flow reactor at Argonne National Laboratory's Materials Engineering Research Facility uses the optimization solver PyMDO to perform autonomous discovery (2).

Nonprofit Org.
 U.S. Postage
 Permit No. 360
 Baltimore, MD

Resilient Data-driven Dynamical Systems with Koopman: An Infinite-dimensional Numerical Analysis Perspective

By Steven L. Brunton
 and Matthew J. Colbrook

Dynamical systems, which describe the evolution of systems in time, are ubiquitous in modern science and engineering. They find use in a wide variety of applications, from mechanics and circuits to climatology, neuroscience, and epidemiology. Consider a discrete-time dynamical system with state x in a state space $\Omega \subset \mathbb{R}^d$ that is governed by an unknown and typically nonlinear function $F: \Omega \rightarrow \Omega$:

$$x_{n+1} = F(x_n), \quad n \geq 0. \quad (1)$$

The classical, geometric way to analyze such systems—which dates back to the seminal work of Henri Poincaré—is based

on the local analysis of fixed points, periodic orbits, stable or unstable manifolds, and so forth. Although Poincaré's framework has revolutionized our understanding of dynamical systems, this approach has at least two challenges in many modern applications: (i) Obtaining a global understanding of the nonlinear dynamics and (ii) handling systems that are either too complex to analyze or offer incomplete information about the evolution (i.e., unknown, high-dimensional, and highly nonlinear F).

Koopman operator theory, which originated with Bernard Koopman and John von Neumann [6, 7], provides a powerful alternative to the classical geometric way of dynamical systems because it addresses nonlinearity: the fundamental issue that underlies the aforementioned challenges.

We lift the nonlinear system (1) into an infinite-dimensional space of observable functions $g: \Omega \rightarrow \mathbb{C}$ via a Koopman operator K :

$$K(g(x_n)) = g(x_{n+1}).$$

The evolution dynamics thus become linear, allowing us to utilize generic solution techniques that are based on spectral decompositions. In recent decades, Koopman operators have captivated researchers because of emerging data-driven and numerical implementations that coincide with the rise of machine learning and high-performance computing [2].

One major goal of modern Koopman operator theory is to find a coordinate transformation with which a linear system may approximate even strongly nonlinear dynamics; this coordinate system relates to the spectrum of the Koopman operator. In 2005, Igor Mezic introduced the Koopman mode decomposition [8], which provided a theoretical basis for connecting the dynamic mode decomposition (DMD) with the Koopman operator [9, 10]. DMD quickly became the workhorse algorithm for computational approximations of the Koopman operator due to its simple and highly extensible formulation in terms of linear algebra, and the fact that it applies equally well to data-driven modeling when no governing equations are available. However, researchers soon realized that simply building linear models in terms of the primitive measured variables cannot sufficiently capture nonlinear dynamics beyond periodic and quasi-periodic phenomena. A major breakthrough occurred with the introduction of extended DMD (EDMD), which generalizes DMD to a broader class of basis functions in which to expand eigenfunctions of the Koopman operator [11].

See *Dynamical Systems* on page 4

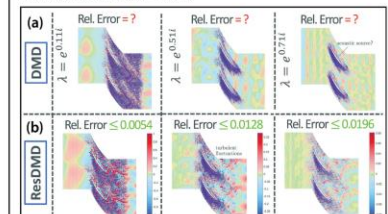


Figure 1. Koopman modes of a turbulent flow (Reynolds number 3.0×10^3) past a cascade of airfoils that are computed from trajectory data (a) or 300,000 Koopman modes are projections of the physical field onto eigenfunctions of K ; they provide the collective motion of the fluid that occurs at the same spatial frequency, growth, or decay rate according to an approximate eigenvalue λ . **1a**, Koopman modes that were computed via existing state-of-the-art techniques. Note the lack of error bounds. **1b**, Koopman modes that were computed using residual dynamic mode decomposition (ResDMD). The physical picture in 1b is different from 1a, but we know that it is correct because of the guaranteed relative error bounds (green text). This outcome illustrates the importance of verification. Figure courtesy of Matthew Colbrook.

Code: <https://github.com/MColbrook/Residual-Dynamic-Mode-Decomposition>

Rigorous data-driven Koopmanism!

“Too much” or “Too little”

Idea: New matrix for residual

⇒ **ResDMD** for computing spectra.

Continuous spectra

Idea: Smoothing via resolvent and **ResDMD**.

Is it right?

ResDMD verifies computations.

E.g., learned dictionaries.



Read more about these breakthroughs in SIAM News!



Resilient Data-driven Dynamical Systems with Koopman: An Infinite-dimensional Numerical Analysis Perspective

By Steven L. Brunton and Matthew J. Colbrook

Dynamical systems, which describe the evolution of systems in time, are ubiquitous in modern science and engineering. They find use in a wide variety of applications, from mechanics and circuits to climatology, neuroscience, and epidemiology. Consider a discrete-time dynamical system with state x in a state space $\Omega \subset \mathbb{R}^d$ that is governed by an unknown and typically nonlinear function $F: \Omega \rightarrow \Omega$:

$$x_{n+1} = F(x_n), \quad n \geq 0. \quad (1)$$

The classical, geometric way to analyze such systems—which dates back to the seminal work of Henri Poincaré—is based

on the local analysis of fixed points, periodic orbits, stable or unstable manifolds, and so forth. Although Poincaré's framework has revolutionized our understanding of dynamical systems, this approach has at least two challenges in many modern applications: (i) Obtaining a global understanding of the nonlinear dynamics and (ii) handling systems that are either too complex to analyze or offer incomplete information about the evolution (i.e., unknown, high-dimensional, and highly nonlinear F).

Koopman operator theory, which originated with Bernard Koopman and John von Neumann [6, 7], provides a powerful alternative to the classical geometric view of dynamical systems because it addresses nonlinearity: the fundamental issue that underlies the aforementioned challenges.

We lift the nonlinear system (1) into an infinite-dimensional space of observable functions $g: \Omega \rightarrow \mathbb{C}$ via a Koopman operator K :

$$K(g(x)) = g(F(x)).$$

The evolution dynamics thus become linear, allowing us to utilize generic solution techniques that are based on spectral decompositions. In recent decades, Koopman operators have captivated researchers because of emerging data-driven and numerical implementations that coincide with the rise of machine learning and high-performance computing [2].

One major goal of modern Koopman operator theory is to find a coordinate transformation with which a linear system may approximate even strongly nonlinear dynamics; this coordinate system relates to the spectrum of the Koopman operator. In 2005, Igor Mezic introduced the Koopman mode decomposition [8], which provided a theoretical basis for connecting the dynamic mode decomposition (DMD) with the Koopman operator [9, 10]. DMD quickly became the workhorse algorithm for computational approximations of the Koopman operator due to its simple and highly extensible formulation in terms of linear algebra, and the fact that it applies equally well to data-driven modeling when no governing equations are available. However, researchers soon realized that simply building linear models in terms of the primitive measured variables cannot sufficiently capture nonlinear dynamics beyond periodic and quasi-periodic phenomena. A major breakthrough occurred with the introduction of extended DMD (EDMD), which generalizes DMD to a broader class of basis functions in which to expand eigenfunctions of the Koopman operator [11].

See Dynamical Systems on page 4

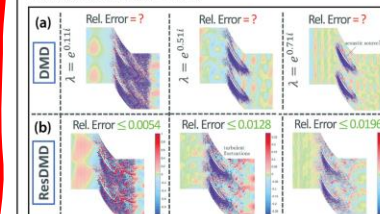
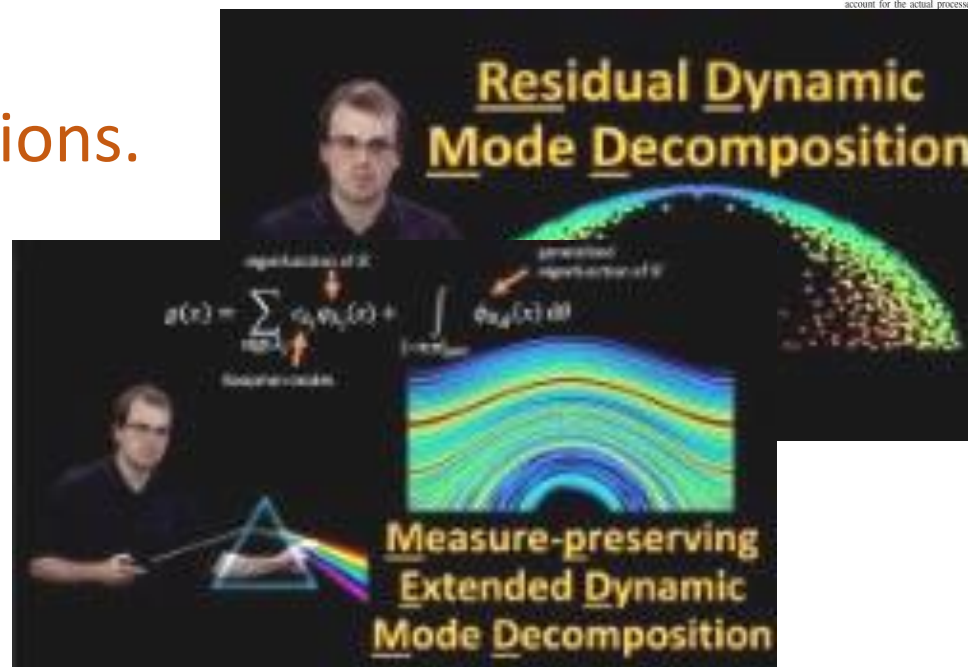


Figure 1. Koopman modes of a turbulent flow (Reynolds number 3.0×10^3) past a cascade of airfoils that are computed from trajectory data (4 s, 300,000). Koopman modes are projections of the physical field onto eigenfunctions of K ; they provide the collective motion of the fluid that occurs at the same spatial frequency, growth, or decay rate according to an approximate eigenvalue λ . **1a.** Koopman modes that were computed via existing state-of-the-art techniques. Note the lack of error bounds. **1b.** Koopman modes that were computed using residual dynamic mode decomposition (ResDMD). The physical picture in 1b is different from 1a, but we know that it is correct because of the guaranteed relative error bounds (green text). This outcome illustrates the importance of verification. Figure courtesy of Matthew Colbrook.



Short video summaries available on YouTube:

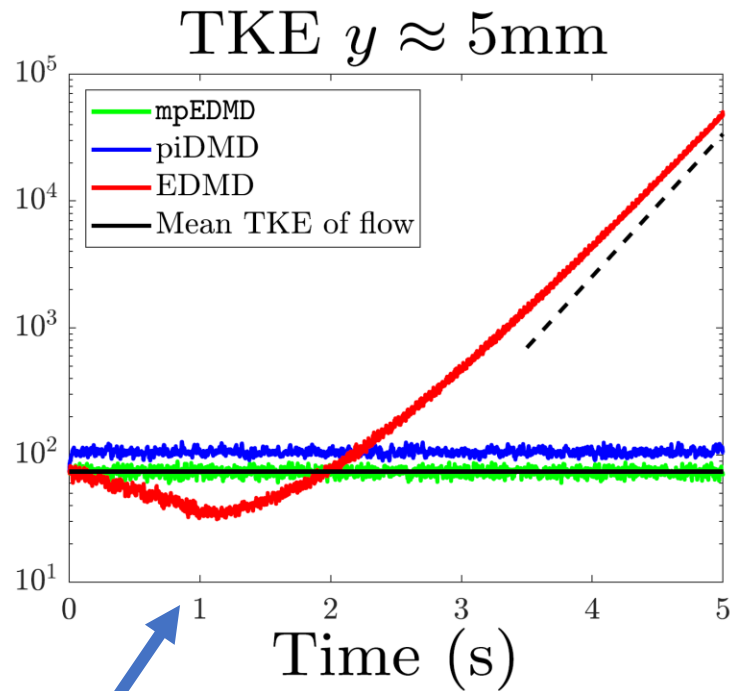
(Thank you Steve Brunton for letting me use your channel!)

Code: <https://github.com/MColbrook/Residual-Dynamic-Mode-Decomposition>

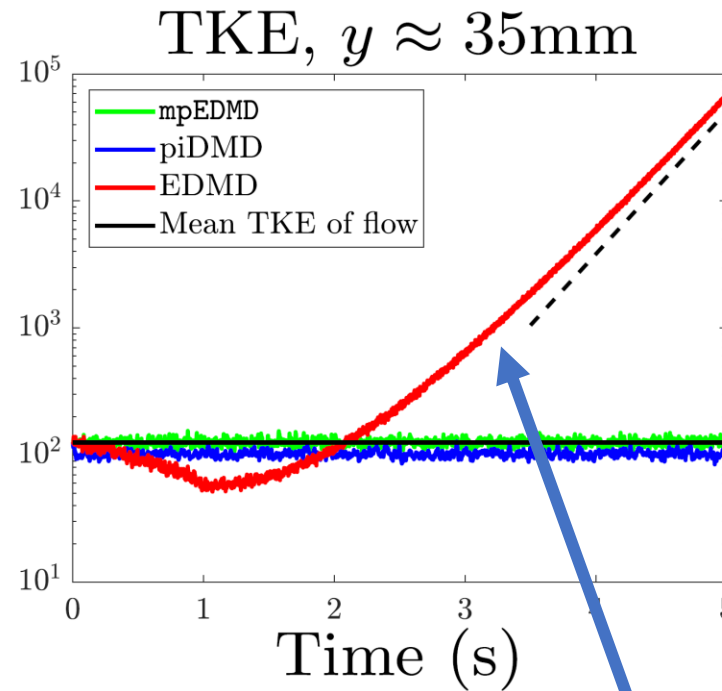
Additional slides...

measure-preserving EDMD...

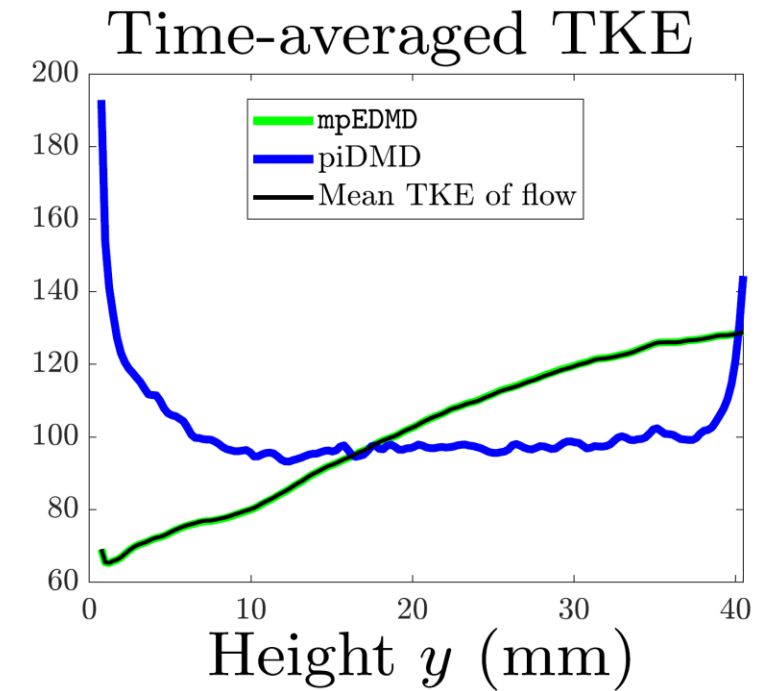
- Polar decomposition of \mathcal{K} . Easy to combine with any DMD-type method!
- Converges for spectral measures, spectra, Koopman mode decomposition.
- Measure-preserving discretization for arbitrary measure-preserving systems.




Snapshots collected over 1s



EDMD unstable!



Solvability Complexity Index Hierarchy

Class $\Omega \ni A$, want to compute $\Xi: \Omega \rightarrow (\mathcal{M}, d)$  metric space

- Δ_0 : Problems solved in finite time (v. rare for cts problems).

- Δ_1 : Problems solved in “one limit” with full error control:

$$d(\Gamma_n(A), \Xi(A)) \leq 2^{-n}$$

- Δ_2 : Problems solved in “one limit”:

$$\lim_{n \rightarrow \infty} \Gamma_n(A) = \Xi(A)$$

- Δ_3 : Problems solved in “two successive limits”:

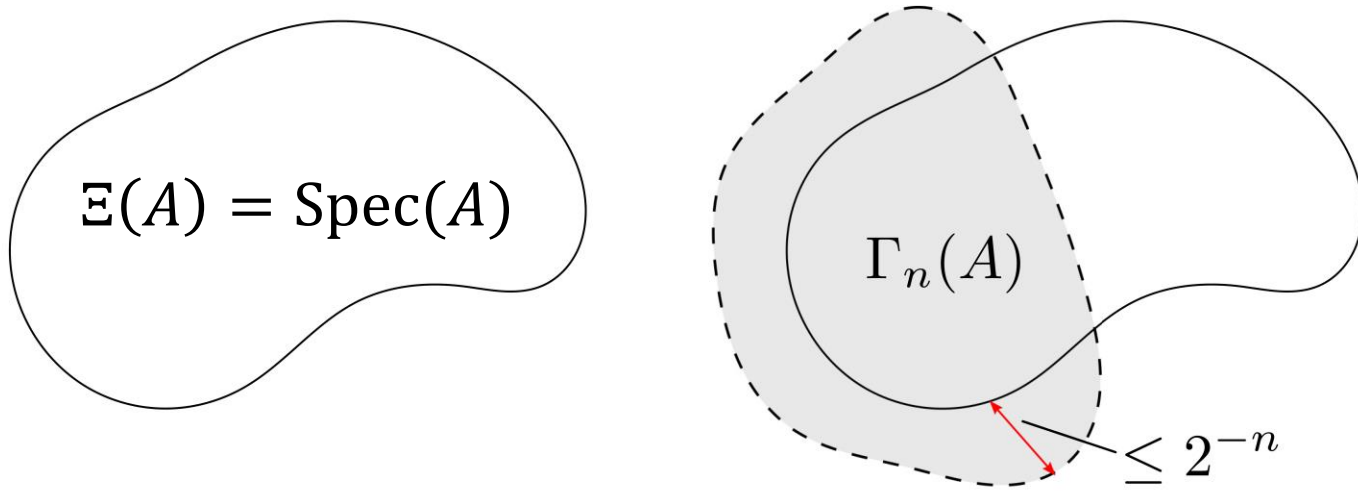
$$\lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \Gamma_{n,m}(A) = \Xi(A)$$

⋮

-
- Ben-Artzi, C., Hansen, Nevanlinna, Seidel, “*On the solvability complexity index hierarchy and towers of algorithms*,” preprint.
 - Hansen, “*On the solvability complexity index, the n -pseudospectrum and approximations of spectra of operators*,” **J. Amer. Math. Soc.**, 2011.
 - McMullen, “*Families of rational maps and iterative root-finding algorithms*,” **Ann. of Math.**, 1987.
 - Doyle, McMullen, “*Solving the quintic by iteration*,” **Acta Math.**, 1989.
 - Smale, “*The fundamental theorem of algebra and complexity theory*,” **Bull. Amer. Math. Soc.**, 1981.

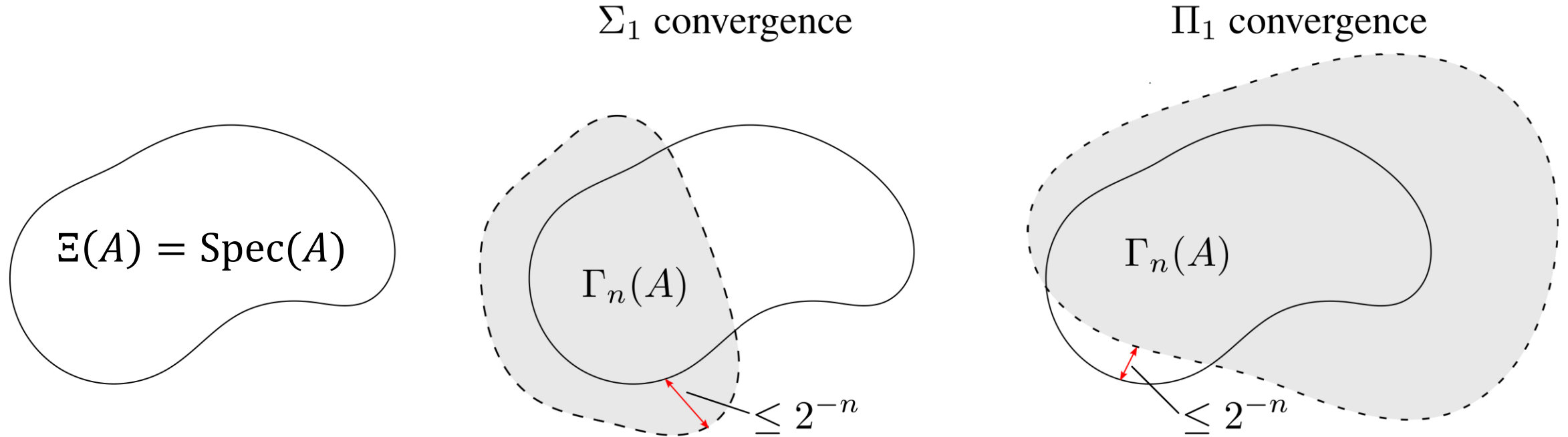
Error control for spectral problems

Σ_1 convergence



- $\Sigma_1: \exists \text{ alg. } \{\Gamma_n\} \text{ s.t. } \lim_{n \rightarrow \infty} \Gamma_n(A) = \Xi(A), \max_{z \in \Gamma_n(A)} \text{dist}(z, \Xi(A)) \leq 2^{-n}$

Error control for spectral problems



- $\Sigma_1: \exists \text{ alg. } \{\Gamma_n\} \text{ s.t. } \lim_{n \rightarrow \infty} \Gamma_n(A) = \Xi(A), \max_{z \in \Gamma_n(A)} \text{dist}(z, \Xi(A)) \leq 2^{-n}$
- $\Pi_1: \exists \text{ alg. } \{\Gamma_n\} \text{ s.t. } \lim_{n \rightarrow \infty} \Gamma_n(A) = \Xi(A), \max_{z \in \Xi(A)} \text{dist}(z, \Gamma_n(A)) \leq 2^{-n}$

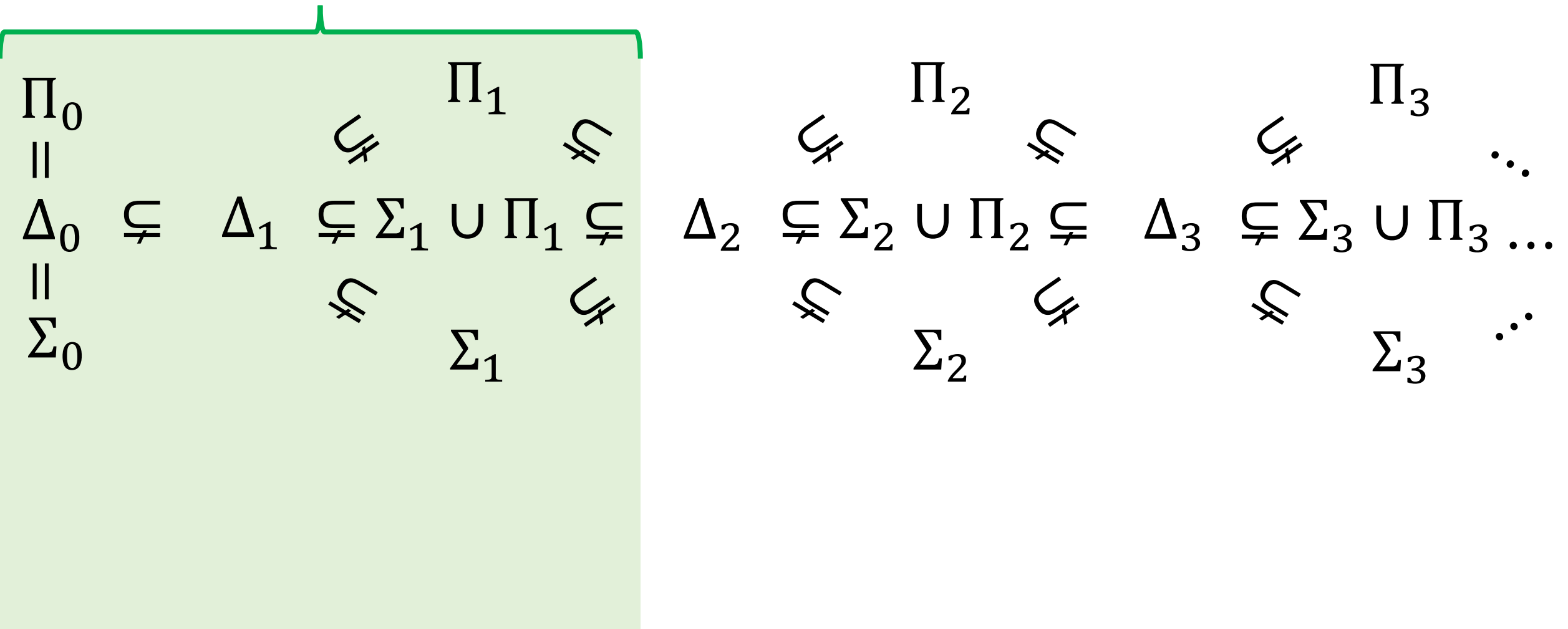
Such problems can be used in a proof!

Small sample of classification theorems

Increasing difficulty



Error control

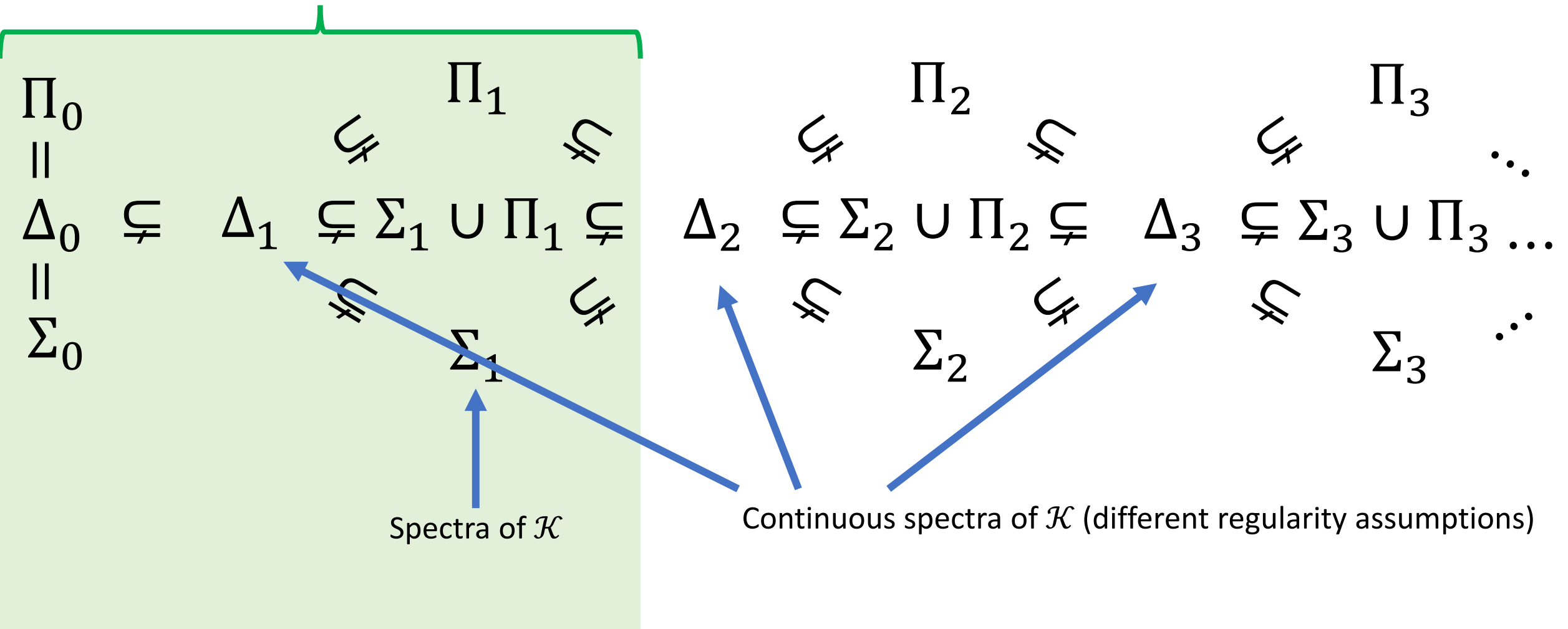


Small sample of classification theorems

Increasing difficulty



Error control



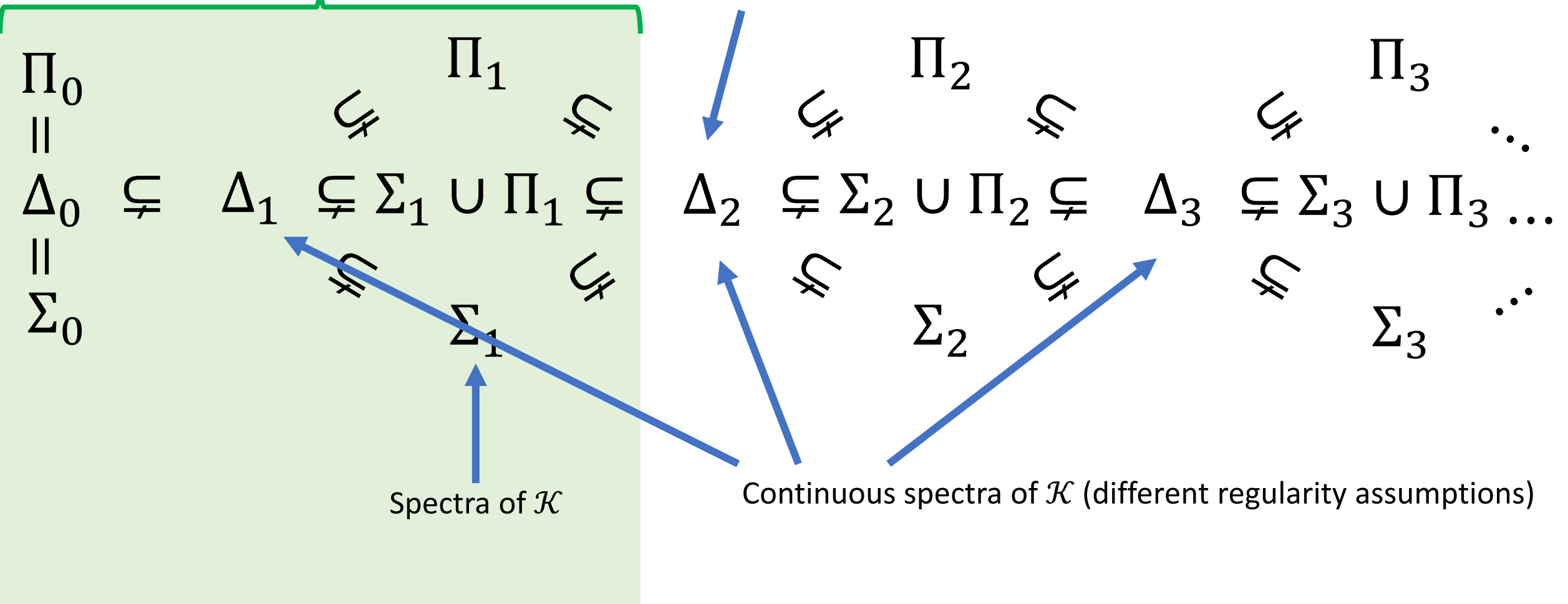
Small sample of classification theorems

Increasing difficulty



Error control

Spectra of compact operators

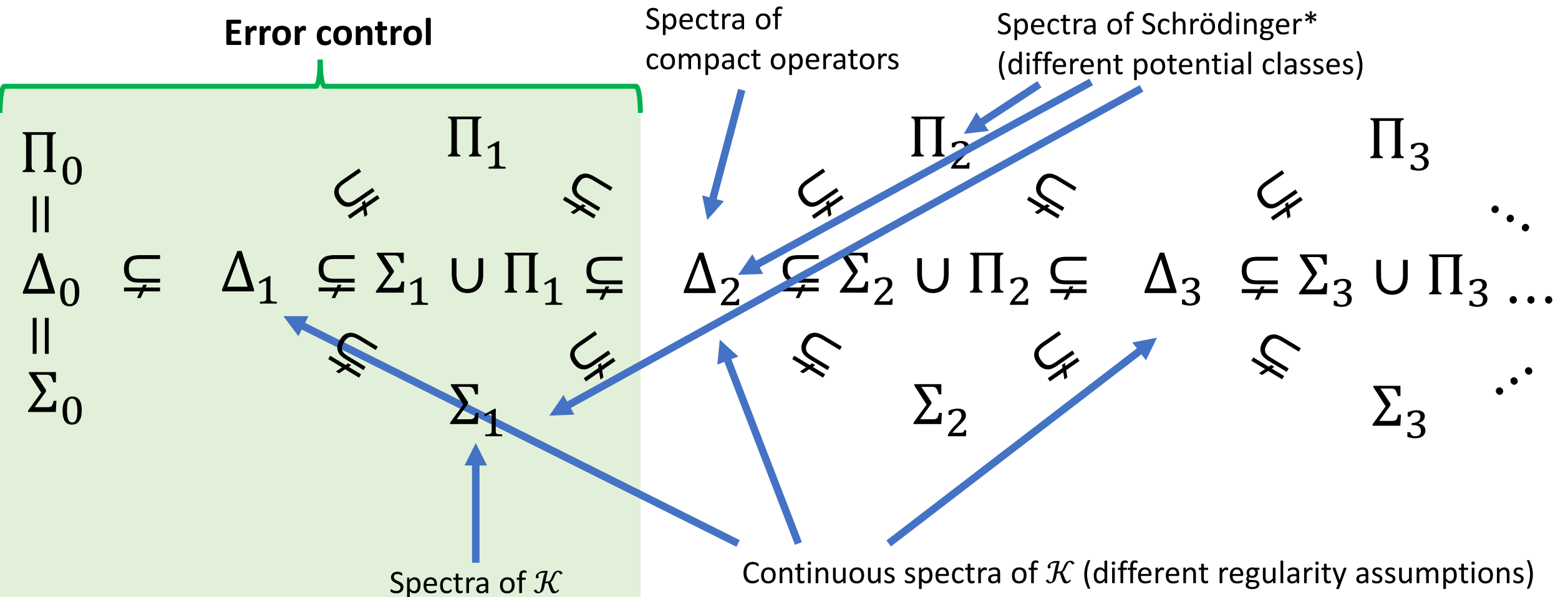


Small sample of classification theorems

Increasing difficulty



Error control



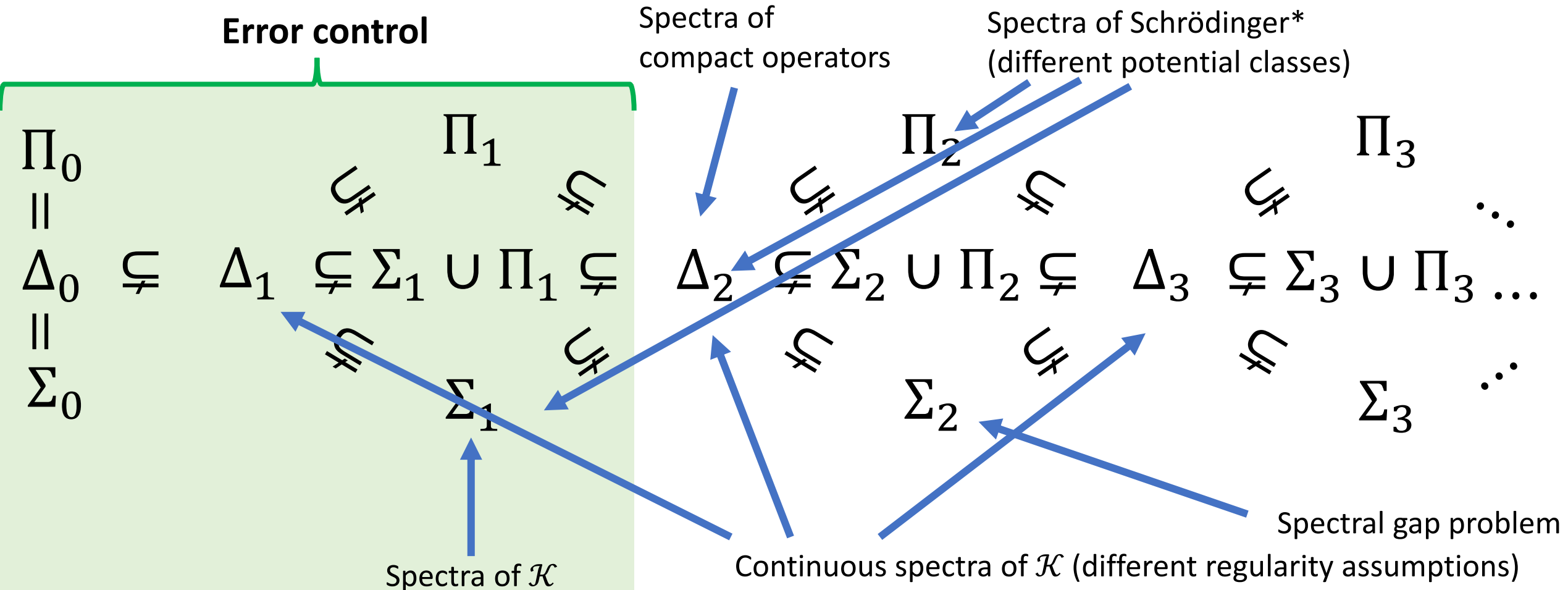
*Open problem of Schwinger: "The special canonical group," "Unitary operator bases," PNAS, 1960.

Small sample of classification theorems

Increasing difficulty



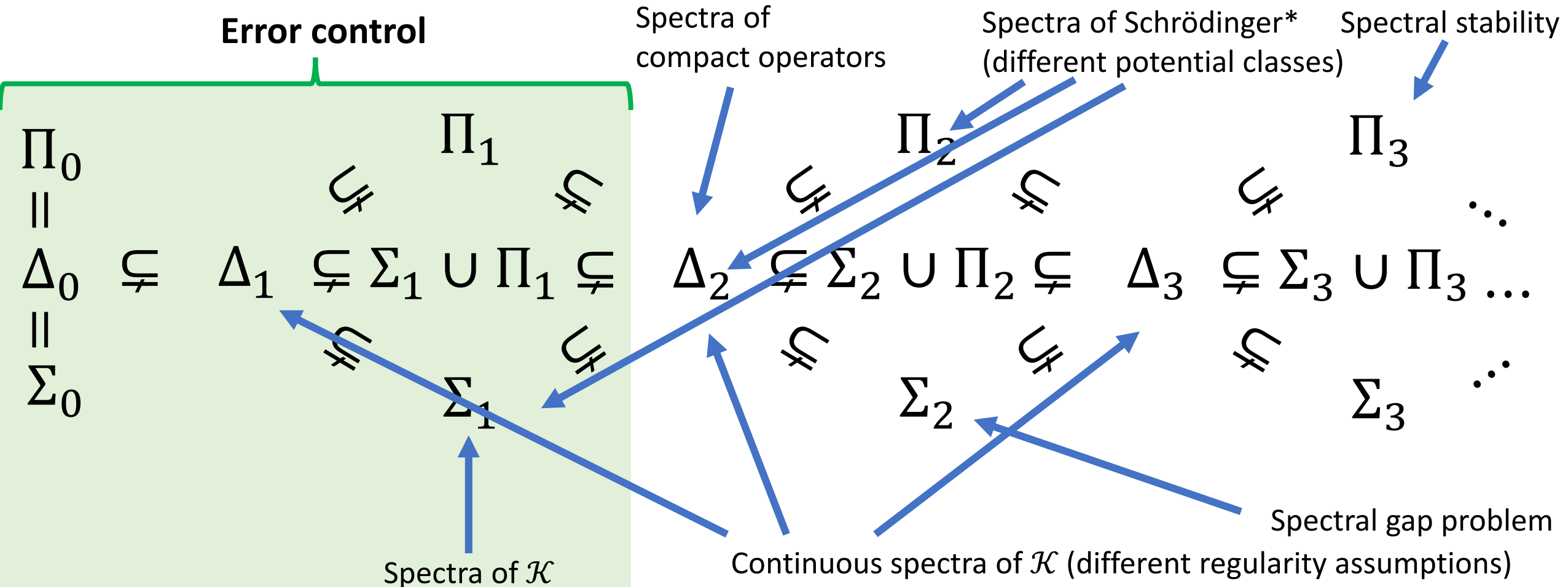
Error control



*Open problem of Schwinger: "The special canonical group," "Unitary operator bases," PNAS, 1960.

Small sample of classification theorems

Increasing difficulty



*Open problem of Schwinger: "The special canonical group," "Unitary operator bases," PNAS, 1960.