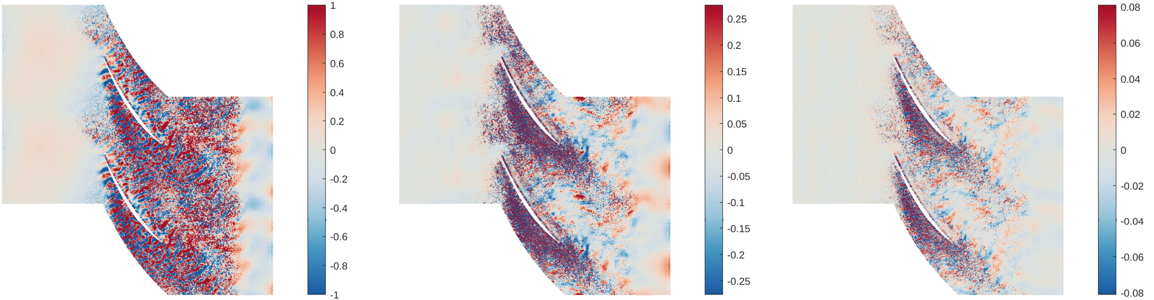
Koopman operators and the computation of spectral properties in infinite dimensions

A foundational framework for Numerical meets Data Analysis

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

Based on: Matthew Colbrook and Alex Townsend, "*Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems*" (available on arXiv)



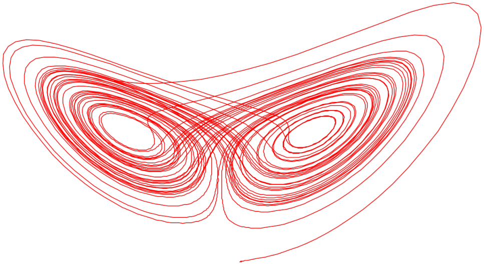
The setup: discrete-time dynamical system

Dynamical system: State $\boldsymbol{x} \in \Omega \subset \mathbb{R}^d, \quad F : \Omega \to \Omega, \quad \boldsymbol{x}_{n+1} = F(\boldsymbol{x}_n).$

Given snapshot data: $\{\boldsymbol{x}^{(m)}, \boldsymbol{y}^{(m)}\}_{m=1}^{M}$ with $\boldsymbol{y}^{(m)} = F(\boldsymbol{x}^{(m)})$.

Broad goal: Learn properties of the dynamical system.

Applications: Biochemistry, classical mechanics, climate, electronics, epidemiology, finance, fluids, molecular dynamics, neuroscience, robotics, ... (anything evolving in time).



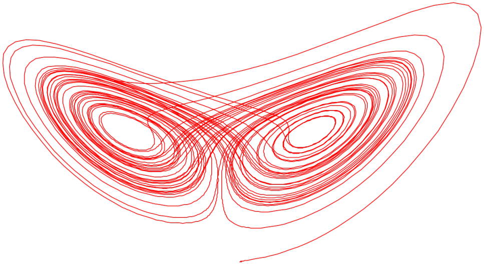
The setup: discrete-time dynamical system

Dynamical system: State $\boldsymbol{x} \in \Omega \subset \mathbb{R}^d$, $\quad F : \Omega \to \Omega, \quad \boldsymbol{x}_{n+1} = F(\boldsymbol{x}_n)$.

Given snapshot data: $\{\boldsymbol{x}^{(m)}, \boldsymbol{y}^{(m)}\}_{m=1}^{M}$ with $\boldsymbol{y}^{(m)} = F(\boldsymbol{x}^{(m)})$.

Broad goal: Learn properties of the dynamical system.

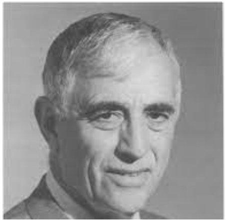
Applications: Biochemistry, classical mechanics, climate, electronics, epidemiology, finance, fluids, molecular dynamics, neuroscience, robotics, ... (anything evolving in time).

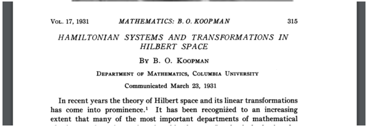


Immediate difficulties:

- F is **unknown**
- F is typically **nonlinear**
- system could be **chaotic**

Koopman operators





VOL. 17, 1931 MATHEMATICS: B. O. KOOPMAN 315

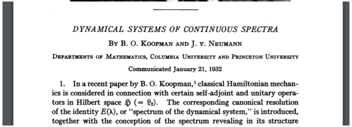
HAMILTONIAN SYSTEMS AND TRANSFORMATIONS IN HILBERT SPACE

BY B. O. KOOPMAN

DEPARTMENT OF MATHEMATICS, COLUMBIA UNIVERSITY

Communicated March 23, 1931

In recent years the theory of Hilbert space and its linear transformations has come into prominence.[1] It has been recognized to an increasing extent that many of the most important departments of mathematical



DYNAMICAL SYSTEMS OF CONTINUOUS SPECTRA

BY B. O. KOOPMAN AND J. V. NEUMANN

DEPARTMENTS OF MATHEMATICS, COLUMBIA UNIVERSITY AND PRINCETON UNIVERSITY

Communicated January 21, 1932

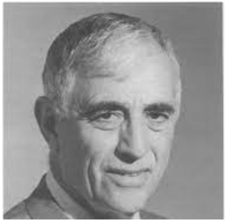
1. In a recent paper by B. O. Koopman,[1] classical Hamiltonian mechanics is considered in connection with certain self-adjoint and unitary operators in Hilbert space $\mathfrak{H}$ ($= \mathfrak{L}_2$). The corresponding canonical resolution of the identity $E(\lambda)$, or "spectrum of the dynamical system," is introduced, together with the conception of the spectrum revealing in its structure

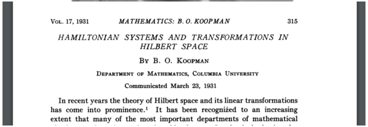
Observable $g : \Omega \to \mathbb{C}$

$$[\mathcal{K}g](\boldsymbol{x}) = g(F(\boldsymbol{x})), \qquad \boldsymbol{x} \in \Omega.$$

$\mathcal{K} : \mathcal{D}(\mathcal{K}) \subset L^2(\Omega, \omega) \to L^2(\Omega, \omega)$ is **linear**, but **infinite-dimensional**!

Koopman operators





VOL. 17, 1931 MATHEMATICS: B. O. KOOPMAN 315

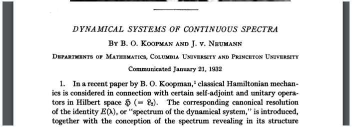
HAMILTONIAN SYSTEMS AND TRANSFORMATIONS IN HILBERT SPACE

BY B. O. KOOPMAN

DEPARTMENT OF MATHEMATICS, COLUMBIA UNIVERSITY

Communicated March 23, 1931

In recent years the theory of Hilbert space and its linear transformations has come into prominence.[1] It has been recognized to an increasing extent that many of the most important departments of mathematical



DYNAMICAL SYSTEMS OF CONTINUOUS SPECTRA

BY B. O. KOOPMAN AND J. V. NEUMANN

DEPARTMENTS OF MATHEMATICS, COLUMBIA UNIVERSITY AND PRINCETON UNIVERSITY

Communicated January 21, 1932

1. In a recent paper by B. O. Koopman,[1] classical Hamiltonian mechanics is considered in connection with certain self-adjoint and unitary operators in Hilbert space $\mathfrak{H}$ ($= \mathfrak{L}_2$). The corresponding canonical resolution of the identity $E(\lambda)$, or "spectrum of the dynamical system," is introduced, together with the conception of the spectrum revealing in its structure

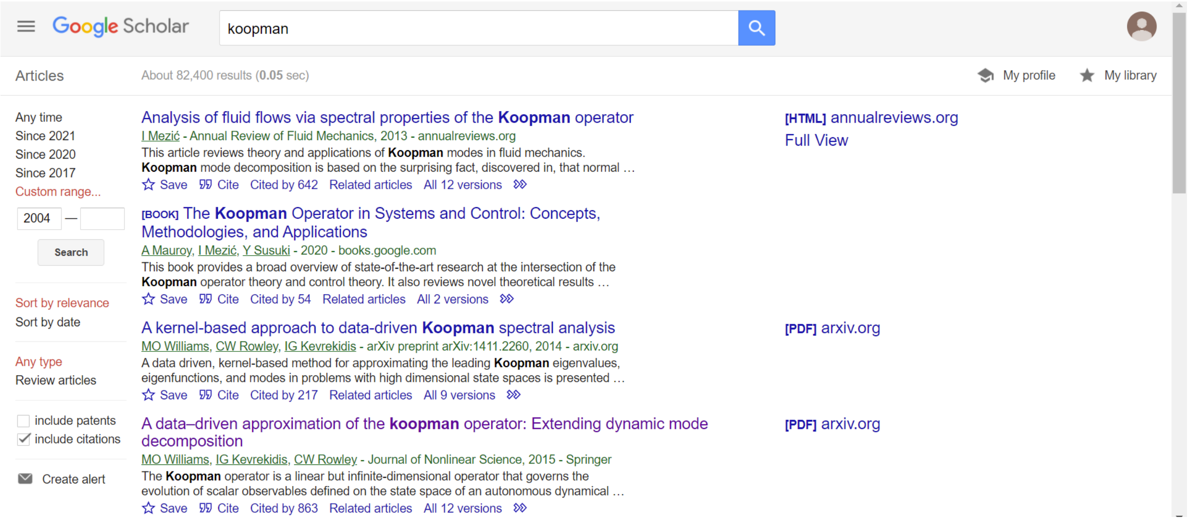
Observable $g : \Omega \to \mathbb{C}$

$$[\mathcal{K}g](\boldsymbol{x}) = g(F(\boldsymbol{x})), \qquad \boldsymbol{x} \in \Omega.$$

$\mathcal{K} : \mathcal{D}(\mathcal{K}) \subset L^2(\Omega, \omega) \to L^2(\Omega, \omega)$ is **linear**, but **infinite-dimensional**!

GOAL: Learn spectral properties of $\mathcal{K}$. Spectrum, $\sigma(\mathcal{K}) = \{z \in \mathbb{C} : \mathcal{K} - z \text{ not invertible}\}$.

Koopmania and dynamics in the big data era - a revolution parallel to deep learning



- I. Mezić, A. Banaszuk "*Comparison of systems with complex behavior*," Physica D, 2004.
- I. Mezić "*Spectral properties of dynamical systems, model reduction and decompositions*," Nonlin. Dyn., 2005.

Why spectra? (Answer: determines properties of the system)

Suppose $(\lambda, \varphi_\lambda)$ is an eigenfunction-eigenvalue pair of $\mathcal{K}$, then

$$\varphi_\lambda(\boldsymbol{x}_n) = [\mathcal{K}^n \varphi_\lambda](\boldsymbol{x}_0) = \lambda^n \varphi_\lambda(\boldsymbol{x}_0).$$

Why spectra? (Answer: determines properties of the system)

Suppose $(\lambda, \varphi_\lambda)$ is an eigenfunction-eigenvalue pair of $\mathcal{K}$, then

$$\varphi_\lambda(\boldsymbol{x}_n) = [\mathcal{K}^n \varphi_\lambda](\boldsymbol{x}_0) = \lambda^n \varphi_\lambda(\boldsymbol{x}_0).$$

Suppose system is measure-preserving (e.g., Hamiltonian, ergodic,...), $\forall g \in L^2(\Omega, \omega)$

$$g = \underbrace{\sum_{\text{e-vals } \lambda} c_\lambda \varphi_\lambda}_{\text{discrete spectral part}} + \underbrace{\int_{[-\pi,\pi]_{\mathrm{per}}} \phi_{\theta,g} \, d\theta}_{\text{continuous spectral part}}.$$

φ_λ are eigenfunctions of $\mathcal{K}$, $c_\lambda \in \mathbb{C}$, $\phi_{\theta,g}$ are "continuously parametrised" eigenfunctions.

Why spectra? (Answer: determines properties of the system)

Suppose $(\lambda, \varphi_\lambda)$ is an eigenfunction-eigenvalue pair of $\mathcal{K}$, then

$$\varphi_\lambda(\boldsymbol{x}_n) = [\mathcal{K}^n \varphi_\lambda](\boldsymbol{x}_0) = \lambda^n \varphi_\lambda(\boldsymbol{x}_0).$$

Suppose system is measure-preserving (e.g., Hamiltonian, ergodic,...), $\forall g \in L^2(\Omega, \omega)$

$$g = \underbrace{\sum_{\text{e-vals } \lambda} c_\lambda \varphi_\lambda}_{\text{discrete spectral part}} + \underbrace{\int_{[-\pi,\pi]_{\mathrm{per}}} \phi_{\theta,g} \, d\theta}_{\text{continuous spectral part}}.$$

φ_λ are eigenfunctions of $\mathcal{K}$, $c_\lambda \in \mathbb{C}$, $\phi_{\theta,g}$ are "continuously parametrised" eigenfunctions.
Koopman mode decomposition

$$g(\boldsymbol{x}_n) = [\mathcal{K}^n g](\boldsymbol{x}_0) = \sum_{\text{e-vals } \lambda} c_\lambda \lambda^n \varphi_\lambda(\boldsymbol{x}_0) + \int_{[-\pi,\pi]_{\mathrm{per}}} e^{in\theta} \phi_{\theta,g}(\boldsymbol{x}_0) \, d\theta.$$

Numerical analysis has the tools!

Global understanding of nonlinear dynamics in state-space:

"*a mathematical grand challenge of the 21st century*"

· S. Brunton, J. N. Kutz "*Data-driven Science and Engineering: Machine learning, Dynamical systems, and Control*," CUP, 2019.

Numerical analysis has the tools!

Global understanding of nonlinear dynamics in state-space:

"*a mathematical grand challenge of the 21st century*"

Challenges:

· S. Brunton, J. N. Kutz "*Data-driven Science and Engineering: Machine learning, Dynamical systems, and Control*," CUP, 2019.

Numerical analysis has the tools!

Global understanding of nonlinear dynamics in state-space:

"*a mathematical grand challenge of the 21st century*"

Challenges:

(C1) Continuous spectra.

· S. Brunton, J. N. Kutz "*Data-driven Science and Engineering: Machine learning, Dynamical systems, and Control*," CUP, 2019.

Numerical analysis has the tools!

Global understanding of nonlinear dynamics in state-space:

"*a mathematical grand challenge of the 21st century*"

Challenges:

(C1) Continuous spectra.

(C2) No finite-dimensional invariant subspaces.

· S. Brunton, J. N. Kutz "*Data-driven Science and Engineering: Machine learning, Dynamical systems, and Control*," CUP, 2019.

Numerical analysis has the tools!

Global understanding of nonlinear dynamics in state-space:

"*a mathematical grand challenge of the 21st century*"

Challenges:

(C1) Continuous spectra.

(C2) No finite-dimensional invariant subspaces.

(C3) Spectral pollution.

· S. Brunton, J. N. Kutz "*Data-driven Science and Engineering: Machine learning, Dynamical systems, and Control*," CUP, 2019.

Numerical analysis has the tools!

Global understanding of nonlinear dynamics in state-space:

"*a mathematical grand challenge of the 21st century*"

Challenges:

(C1) Continuous spectra.

(C2) No finite-dimensional invariant subspaces.

(C3) Spectral pollution.

(C4) Chaotic behaviour.

· S. Brunton, J. N. Kutz "*Data-driven Science and Engineering: Machine learning, Dynamical systems, and Control*," CUP, 2019.

Numerical analysis has the tools!

Global understanding of nonlinear dynamics in state-space:

"a mathematical grand challenge of the 21st century"

Challenges:

(C1) Continuous spectra.

(C2) No finite-dimensional invariant subspaces.

(C3) Spectral pollution.

(C4) Chaotic behaviour.

Solutions in this talk:

· S. Brunton, J. N. Kutz "*Data-driven Science and Engineering: Machine learning, Dynamical systems, and Control*," CUP, 2019.

Numerical analysis has the tools!

Global understanding of nonlinear dynamics in state-space:

"*a mathematical grand challenge of the 21st century*"

Challenges:

(C1) Continuous spectra.

(C2) No finite-dimensional invariant subspaces.

(C3) Spectral pollution.

(C4) Chaotic behaviour.

Solutions in this talk:

(S1) Compute smoothed approximations of spectral measures with explicit high-order convergence rates.

· S. Brunton, J. N. Kutz "*Data-driven Science and Engineering: Machine learning, Dynamical systems, and Control*," CUP, 2019.

Numerical analysis has the tools!

Global understanding of nonlinear dynamics in state-space:

"a mathematical grand challenge of the 21st century"

Challenges:

(C1) Continuous spectra.

(C2) No finite-dimensional invariant subspaces.

(C3) Spectral pollution.

(C4) Chaotic behaviour.

Solutions in this talk:

(S1) Compute smoothed approximations of spectral measures with explicit high-order convergence rates.

(S2) Compute spectral properties of $\mathcal{K}$ directly, as opposed to restrictions of $\mathcal{K}$ to finite-dimensional subspaces.

· S. Brunton, J. N. Kutz "*Data-driven Science and Engineering: Machine learning, Dynamical systems, and Control*," CUP, 2019.

Numerical analysis has the tools!

Global understanding of nonlinear dynamics in state-space:

"*a mathematical grand challenge of the 21st century*"

Challenges:

(C1) Continuous spectra.

(C2) No finite-dimensional invariant subspaces.

(C3) Spectral pollution.

(C4) Chaotic behaviour.

Solutions in this talk:

(S1) Compute smoothed approximations of spectral measures with explicit high-order convergence rates.

(S2) Compute spectral properties of $\mathcal{K}$ directly, as opposed to restrictions of $\mathcal{K}$ to finite-dimensional subspaces.

(S3) Compute residuals associated with the spectrum with error control, providing convergence without spectral pollution.

· S. Brunton, J. N. Kutz "*Data-driven Science and Engineering: Machine learning, Dynamical systems, and Control*," CUP, 2019.

Numerical analysis has the tools!

Global understanding of nonlinear dynamics in state-space:

"*a mathematical grand challenge of the 21st century*"

Challenges:

(C1) Continuous spectra.

(C2) No finite-dimensional invariant subspaces.

(C3) Spectral pollution.

(C4) Chaotic behaviour.

Solutions in this talk:

(S1) Compute smoothed approximations of spectral measures with explicit high-order convergence rates.

(S2) Compute spectral properties of $\mathcal{K}$ directly, as opposed to restrictions of $\mathcal{K}$ to finite-dimensional subspaces.

(S3) Compute residuals associated with the spectrum with error control, providing convergence without spectral pollution.

(S4) Handle chaotic systems using single time steps.

· S. Brunton, J. N. Kutz "*Data-driven Science and Engineering: Machine learning, Dynamical systems, and Control*," CUP, 2019.

Part 1: Computing residuals and spectra.

General Koopman operators.

Work in $L^2(\Omega, \omega)$ with inner product $\langle \cdot, \cdot \rangle$.

EDMD: a Galerkin approach

Subspace $\mathrm{span}\{\psi_j\}_{j=1}^{N_K} \subset L^2(\Omega,\omega)$, $\Psi(\boldsymbol{x}) = \begin{bmatrix}\psi_1(\boldsymbol{x}) & \cdots & \psi_{N_K}(\boldsymbol{x})\end{bmatrix} \in \mathbb{C}^{1\times N_K}$.

$$\text{For } \{\boldsymbol{x}^{(m)}, \boldsymbol{y}^{(m)} = F(\boldsymbol{x}^{(m)})\}_{m=1}^{M}, \quad \Psi_X = \begin{pmatrix}\Psi(\boldsymbol{x}^{(1)}) \\ \vdots \\ \Psi(\boldsymbol{x}^{(M)})\end{pmatrix} \in \mathbb{C}^{M\times N_K}, \quad \Psi_Y = \begin{pmatrix}\Psi(\boldsymbol{y}^{(1)}) \\ \vdots \\ \Psi(\boldsymbol{y}^{(M)})\end{pmatrix} \in \mathbb{C}^{M\times N_K}.$$

· M. Williams, I. Kevrekidis, C. Rowley "*A data–driven approximation of the Koopman operator: Extending dynamic mode decomposition*," J. Nonlin. Sci., 2015.

EDMD: a Galerkin approach

Subspace $\mathrm{span}\{\psi_j\}_{j=1}^{N_K} \subset L^2(\Omega,\omega)$, $\Psi(\boldsymbol{x}) = \begin{bmatrix}\psi_1(\boldsymbol{x}) & \cdots & \psi_{N_K}(\boldsymbol{x})\end{bmatrix} \in \mathbb{C}^{1\times N_K}$.

$$\text{For } \{\boldsymbol{x}^{(m)}, \boldsymbol{y}^{(m)} = F(\boldsymbol{x}^{(m)})\}_{m=1}^{M}, \quad \Psi_X = \begin{pmatrix}\Psi(\boldsymbol{x}^{(1)})\\ \vdots \\ \Psi(\boldsymbol{x}^{(M)})\end{pmatrix} \in \mathbb{C}^{M\times N_K}, \quad \Psi_Y = \begin{pmatrix}\Psi(\boldsymbol{y}^{(1)})\\ \vdots \\ \Psi(\boldsymbol{y}^{(M)})\end{pmatrix} \in \mathbb{C}^{M\times N_K}.$$

$$\text{Given} \quad g = \sum_{j=1}^{N_K} \psi_j \boldsymbol{g}_j, \quad \text{seek } K_{\mathrm{EDMD}} \in \mathbb{C}^{N_K\times N_K} \text{ with } \quad \mathcal{K}g \approx \sum_{j=1}^{N_K} \psi_j [K_{\mathrm{EDMD}}\boldsymbol{g}]_j.$$

· M. Williams, I. Kevrekidis, C. Rowley *"A data–driven approximation of the Koopman operator: Extending dynamic mode decomposition,"* J. Nonlin. Sci., 2015.

EDMD: a Galerkin approach

Subspace $\mathrm{span}\{\psi_j\}_{j=1}^{N_K} \subset L^2(\Omega,\omega)$, $\Psi(\boldsymbol{x}) = \begin{bmatrix}\psi_1(\boldsymbol{x}) & \cdots & \psi_{N_K}(\boldsymbol{x})\end{bmatrix} \in \mathbb{C}^{1\times N_K}$.

$$\text{For } \{\boldsymbol{x}^{(m)}, \boldsymbol{y}^{(m)} = F(\boldsymbol{x}^{(m)})\}_{m=1}^{M}, \quad \Psi_X = \begin{pmatrix}\Psi(\boldsymbol{x}^{(1)}) \\ \vdots \\ \Psi(\boldsymbol{x}^{(M)})\end{pmatrix} \in \mathbb{C}^{M\times N_K}, \quad \Psi_Y = \begin{pmatrix}\Psi(\boldsymbol{y}^{(1)}) \\ \vdots \\ \Psi(\boldsymbol{y}^{(M)})\end{pmatrix} \in \mathbb{C}^{M\times N_K}.$$

$$\text{Given} \quad g = \sum_{j=1}^{N_K} \psi_j \boldsymbol{g}_j, \quad \text{seek } K_{\mathrm{EDMD}} \in \mathbb{C}^{N_K\times N_K} \text{ with } \quad \mathcal{K}g \approx \sum_{j=1}^{N_K} \psi_j [K_{\mathrm{EDMD}}\boldsymbol{g}]_j.$$

$$\min_{B\in\mathbb{C}^{N_K\times N_K}} \int_\Omega \max_{\|\boldsymbol{g}\|_{\ell^2}=1} \left|\mathcal{K}g - \sum_{j=1}^{N_K} \psi_j [B\boldsymbol{g}]_j\right|^2 d\omega(\boldsymbol{x}) \approx \sum_{m=1}^{M} w_m \left\|\Psi(\boldsymbol{y}^{(m)}) - \Psi(\boldsymbol{x}^{(m)})B\right\|_2^2.$$

· M. Williams, I. Kevrekidis, C. Rowley *"A data–driven approximation of the Koopman operator: Extending dynamic mode decomposition,"* J. Nonlin. Sci., 2015.

EDMD: a Galerkin approach

Subspace $\mathrm{span}\{\psi_j\}_{j=1}^{N_K} \subset L^2(\Omega,\omega)$, $\Psi(\boldsymbol{x}) = \begin{bmatrix}\psi_1(\boldsymbol{x}) & \cdots & \psi_{N_K}(\boldsymbol{x})\end{bmatrix} \in \mathbb{C}^{1\times N_K}$.

$$\text{For } \{\boldsymbol{x}^{(m)}, \boldsymbol{y}^{(m)} = F(\boldsymbol{x}^{(m)})\}_{m=1}^{M}, \quad \Psi_X = \begin{pmatrix}\Psi(\boldsymbol{x}^{(1)}) \\ \vdots \\ \Psi(\boldsymbol{x}^{(M)})\end{pmatrix} \in \mathbb{C}^{M\times N_K}, \quad \Psi_Y = \begin{pmatrix}\Psi(\boldsymbol{y}^{(1)}) \\ \vdots \\ \Psi(\boldsymbol{y}^{(M)})\end{pmatrix} \in \mathbb{C}^{M\times N_K}.$$

$$\text{Given} \quad g = \sum_{j=1}^{N_K} \psi_j \boldsymbol{g}_j, \quad \text{seek } K_{\mathrm{EDMD}} \in \mathbb{C}^{N_K\times N_K} \text{ with} \quad \mathcal{K}g \approx \sum_{j=1}^{N_K} \psi_j [K_{\mathrm{EDMD}}\boldsymbol{g}]_j.$$

$$\min_{B\in\mathbb{C}^{N_K\times N_K}} \int_\Omega \max_{\|\boldsymbol{g}\|_{\ell^2}=1} \left|\mathcal{K}g - \sum_{j=1}^{N_K} \psi_j [B\boldsymbol{g}]_j\right|^2 d\omega(\boldsymbol{x}) \approx \sum_{m=1}^{M} w_m \left\|\Psi(\boldsymbol{y}^{(m)}) - \Psi(\boldsymbol{x}^{(m)})B\right\|_2^2.$$

Solution: $K_{\mathrm{EDMD}} = (\Psi_X^* W \Psi_X)^\dagger (\Psi_X^* W \Psi_Y) \quad (W = \mathrm{diag}(w_1, \ldots, w_M))$

Large data limit: $\lim_{M\to\infty} [\Psi_X^* W \Psi_X]_{jk} = \langle \psi_k, \psi_j \rangle$ and $\lim_{M\to\infty} [\Psi_X^* W \Psi_Y]_{jk} = \langle \mathcal{K}\psi_k, \psi_j \rangle$

· M. Williams, I. Kevrekidis, C. Rowley "*A data–driven approximation of the Koopman operator: Extending dynamic mode decomposition*," J. Nonlin. Sci., 2015.

Residual DMD (ResDMD): a new matrix

If $g=\sum_{j=1}^{N_K}\psi_j\boldsymbol{g}_j\in\mathrm{span}\{\psi_j\}_{j=1}^{N_K}$ and λ are a candidate eigenfunction-eigenvalue pair then

$$\begin{aligned}\|\mathcal{K}g-\lambda g\|_{L^2(\Omega,\omega)}^2&=\sum_{j,k=1}^{N_K}\boldsymbol{g}_k\overline{\boldsymbol{g}_j}\left[\langle\mathcal{K}\psi_k,\mathcal{K}\psi_j\rangle-\lambda\langle\psi_k,\mathcal{K}\psi_j\rangle-\overline{\lambda}\langle\mathcal{K}\psi_k,\psi_j\rangle+|\lambda|^2\langle\psi_k,\psi_j\rangle\right]\\&\approx\sum_{j,k=1}^{N_K}\boldsymbol{g}_k\overline{\boldsymbol{g}_j}\left[\Psi_Y^*W\Psi_Y-\lambda[\Psi_X^*W\Psi_Y]^*-\overline{\lambda}\Psi_X^*W\Psi_Y+|\lambda|^2\Psi_X^*W\Psi_X\right]_{jk}\\&=\boldsymbol{g}^*\left[\Psi_Y^*W\Psi_Y-\lambda[\Psi_X^*W\Psi_Y]^*-\overline{\lambda}\Psi_X^*W\Psi_Y+|\lambda|^2\Psi_X^*W\Psi_X\right]\boldsymbol{g}\end{aligned}$$

Residual DMD (ResDMD): a new matrix

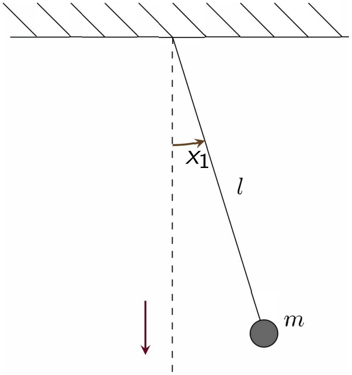
If $g=\sum_{j=1}^{N_K}\psi_j\boldsymbol{g}_j\in\mathrm{span}\{\psi_j\}_{j=1}^{N_K}$ and λ are a candidate eigenfunction-eigenvalue pair then

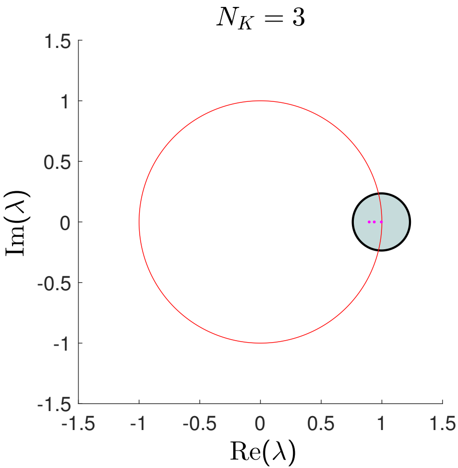
$$
\begin{aligned}
\|\mathcal{K}g-\lambda g\|^2_{L^2(\Omega,\omega)} &= \sum_{j,k=1}^{N_K}\boldsymbol{g}_k\overline{\boldsymbol{g}_j}\left[\langle\mathcal{K}\psi_k,\mathcal{K}\psi_j\rangle-\lambda\langle\psi_k,\mathcal{K}\psi_j\rangle-\overline{\lambda}\langle\mathcal{K}\psi_k,\psi_j\rangle+|\lambda|^2\langle\psi_k,\psi_j\rangle\right]\\
&\approx \sum_{j,k=1}^{N_K}\boldsymbol{g}_k\overline{\boldsymbol{g}_j}\left[\Psi_Y^*W\Psi_Y-\lambda[\Psi_X^*W\Psi_Y]^*-\overline{\lambda}\Psi_X^*W\Psi_Y+|\lambda|^2\Psi_X^*W\Psi_X\right]_{jk}\\
&= \boldsymbol{g}^*\left[\Psi_Y^*W\Psi_Y-\lambda[\Psi_X^*W\Psi_Y]^*-\overline{\lambda}\Psi_X^*W\Psi_Y+|\lambda|^2\Psi_X^*W\Psi_X\right]\boldsymbol{g}
\end{aligned}
$$

New matrix: $\Psi_Y^*W\Psi_Y$ with $\lim\limits_{M\to\infty}[\Psi_Y^*W\Psi_Y]_{jk}=\langle\mathcal{K}\psi_k,\mathcal{K}\psi_j\rangle$

Example: nonlinear pendulum

$$\dot{x_1} = x_2, \quad \dot{x_2} = -\sin(x_1), \quad \text{with} \quad \Omega = [-\pi, \pi]_{\text{per}} \times \mathbb{R}.$$





Computed pseudospectra ($\epsilon = 0.25$). Eigenvalues of K_{EDMD} shown as dots (spectral pollution).

ResDMD: avoiding spectral pollution

$$\mathrm{res}(\lambda, g)^2 = \frac{\boldsymbol{g}^* \left[\Psi_Y^* W \Psi_Y - \lambda[\Psi_X^* W \Psi_Y]^* - \overline{\lambda}\Psi_X^* W \Psi_Y + |\lambda|^2 \Psi_X^* W \Psi_X\right] \boldsymbol{g}}{\boldsymbol{g}^* \left[\Psi_X^* W \Psi_X\right] \boldsymbol{g}}.$$

Algorithm:

1. Compute K_{EDMD}, its eigenvalues and eigenvectors.
2. For each eigenpair (λ, g), compute $\mathrm{res}(\lambda, g)$.
3. Discard eigenpairs with $\mathrm{res}(\lambda, g) > \epsilon$, for accuracy tolerance $\epsilon > 0$.

Theorem (No spectral pollution, compute residuals from above.)

Let Λ_M denote the eigenvalue output of above algorithm. Then

$$\limsup_{M\to\infty} \max_{\lambda\in\Lambda_M} \|(\mathcal{K} - \lambda)^{-1}\|^{-1} \leq \epsilon.$$

BUT: typically does not capture all of spectrum!

ResDMD: computing pseudospectra and spectra

$$\sigma_\epsilon(\mathcal{K}) := \cup_{\|\mathcal{B}\|\leq\epsilon}\sigma(\mathcal{K}+\mathcal{B}), \quad \lim_{\epsilon\downarrow 0}\sigma_\epsilon(\mathcal{K}) = \sigma(\mathcal{K})$$

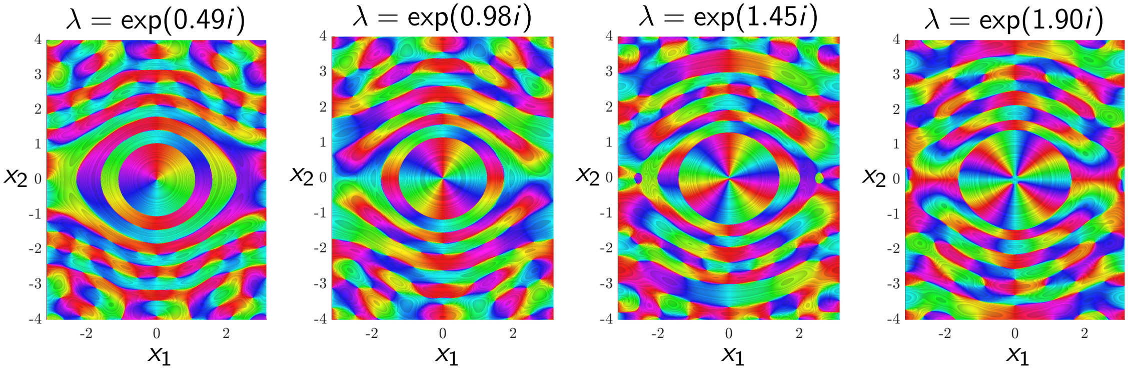
Algorithm:

1. Compute $\Psi_X^* W \Psi_X$, $\Psi_X^* W \Psi_Y$, and $\Psi_Y^* W \Psi_Y$.
2. For each z_j in a computational grid, compute $\tau_j = \min_{\boldsymbol{g}\in\mathbb{C}^{N_K}} \mathrm{res}(z_j, \sum_{k=1}^{N_K} \psi_k \boldsymbol{g}_k)$ and the corresponding singular vectors $\boldsymbol{g}_{(j)}$ (generalised SVD problem).
3. Output: $\{z_j : \tau_j < \epsilon\}$ (estimate of $\sigma_\epsilon(\mathcal{K})$) and ϵ-pseudo-eigenfunctions $\{\boldsymbol{g}_{(j)} : \tau_j < \epsilon\}$.

Theorem

No spectral pollution: $\{z_j : \tau_j < \epsilon\} \subset \sigma_\epsilon(\mathcal{K})$ *(as $M \to \infty$).*
Spectral inclusion: *Converges uniformly to $\sigma_\epsilon(\mathcal{K})$ on bounded subsets of $\mathbb{C}$ as $N_K \to \infty$.*

Example: pseudo-eigenfunctions of nonlinear pendulum



Colour represents complex argument, lines of constant modulus shown as shadowed steps.
All residuals smaller than $\epsilon = 0.05$ (can be made smaller by increasing N_K).

Part 2: Dealing with continuous spectra - computing spectral measures.

In this part, we assume that dynamics are measure-preserving.
E.g., Hamiltonian system, ergodic system, ...

This is equivalent to $\mathcal{K}$ being an isometry[a]:

$$\|\mathcal{K}g\|_{L^2(\Omega,\omega)} = \|g\|_{L^2(\Omega,\omega)}, \quad \forall g \in L^2(\Omega,\omega).$$

Spectrum lives inside the **unit disk**.

[a]For analysts: we actually consider unitary extensions of $\mathcal{K}$ with 'canonical' spectral measures.

Diagonalising infinite-dimensional operators

Finite-dimensional: $A \in \mathbb{C}^{n\times n}$ with $A^*A = AA^*$ has orthonormal basis of e-vectors $\{v_j\}_{j=1}^n$

$$v = \left(\sum_{j=1}^n v_j v_j^*\right) v, \quad v \in \mathbb{C}^n \qquad Av = \left(\sum_{j=1}^n \lambda_j v_j v_j^*\right) v, \quad v \in \mathbb{C}^n.$$

Infinite-dimensional: Operator $\mathcal{L} : \mathcal{D}(\mathcal{L}) \to \mathcal{H}$, ($\mathcal{H} =$ Hilbert space). Typically, no longer a basis of e-vectors. Spectral Theorem: Projection-valued spectral measure $\mathcal{E}$

$$g = \left(\int_{\sigma(\mathcal{L})} d\mathcal{E}(\lambda)\right) g, \quad g \in \mathcal{H} \qquad \mathcal{L}g = \left(\int_{\sigma(\mathcal{L})} \lambda\, d\mathcal{E}(\lambda)\right) g, \quad g \in \mathcal{D}(\mathcal{L}).$$

Scalar-valued spectral measures: $\nu_g(U) = \langle \underbrace{\mathcal{E}(U)}_{\text{projection}} g, g\rangle$.

Example: $\mathcal{L} = -\frac{d^2}{dx^2}$ and Fourier transform

$$\mathcal{L} = -\frac{d^2}{dx^2} \quad \xleftrightarrow{\text{spectral theorem}} \quad \text{projection-valued measure } \mathcal{E}$$

$$x \in [-\pi, \pi]_{\mathrm{per}}$$

$$\sigma(\mathcal{L}) = \{n^2 : n \in \mathbb{Z}_0\}$$

discrete spectrum

$$\hat{g}_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(x) e^{-ikx}\, dx$$

$$[\mathcal{E}([a,b])g](x) = \sum_{a \le k^2 \le b} \hat{g}_k e^{ikx}$$

$$\nu_g([a,b]) = \sum_{a \le k^2 \le b} |\hat{g}_k|^2$$

$$-\infty < x < \infty$$

$$\sigma(\mathcal{L}) = [0, +\infty)$$

continuous spectrum

$$\hat{g}(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} g(x) e^{-ikx}\, dx$$

$$[\mathcal{E}([a,b])g](x) = \int_{a \le k^2 \le b} \hat{g}(k) e^{ikx}\, dk$$

$$\nu_g([a,b]) = \int_{a \le k^2 \le b} |\hat{g}(k)|^2\, dk$$

Koopman mode decomposition

ν_g are spectral measures on $[-\pi,\pi]_{\mathrm{per}}$

Lebesgue's decomposition theorem:

$$d\nu_g(\lambda) = \underbrace{\sum_{\text{e-vals } \lambda_j} \langle \mathcal{P}_{\lambda_j} g, g\rangle \, \delta(\lambda-\lambda_j) d\lambda}_{\text{discrete part}} + \underbrace{\rho_g(\lambda)\, d\lambda + d\nu_g^{(\mathrm{sc})}(\lambda)}_{\text{continuous part}}$$

$$g = \sum_{\text{e-vals } \lambda_j} c_{\lambda_j} \underbrace{\varphi_{\lambda_j}}_{\text{e-functions}} + \underbrace{\int_{[-\pi,\pi]_{\mathrm{per}}} \phi_{\theta,g}\, d\theta}_{\text{ctsly param e-functions}}.$$

$$g(\boldsymbol{x}_n) = [\mathcal{K}^n g](\boldsymbol{x}_0) = \sum_{\text{e-vals } \lambda_j} c_{\lambda_j} \lambda_j^n \varphi_{\lambda_j}(\boldsymbol{x}_0) + \int_{[-\pi,\pi]_{\mathrm{per}}} e^{in\theta} \phi_{\theta,g}(\boldsymbol{x}_0)\, d\theta.$$

Computing ν_g provides diagonalisation of non-linear dynamical system!

Plemelj-type formula

$$\underbrace{K_\epsilon(\theta) = \frac{1}{2\pi} \cdot \frac{(1+\epsilon)^2 - 1}{1 + (1+\epsilon)^2 - 2(1+\epsilon)\cos(\theta)}}_{\text{Poisson kernel for unit disc}}, \quad \underbrace{\mathtt{C}_{\nu_g}(z) := \frac{1}{2\pi} \int_{[-\pi,\pi]_{\mathrm{per}}} \frac{e^{i\theta}\, d\nu_g(\theta)}{e^{i\theta} - z}}_{\text{generalised Cauchy transform}}$$

Plemelj-type formula

$$\underbrace{K_\epsilon(\theta) = \frac{1}{2\pi} \cdot \frac{(1+\epsilon)^2 - 1}{1 + (1+\epsilon)^2 - 2(1+\epsilon)\cos(\theta)}}_{\text{Poisson kernel for unit disc}}, \quad \underbrace{\mathtt{C}_{\nu_g}(z) := \frac{1}{2\pi}\int_{[-\pi,\pi]_{\mathrm{per}}} \frac{e^{i\theta}\, d\nu_g(\theta)}{e^{i\theta} - z}}_{\text{generalised Cauchy transform}}$$

$$\begin{aligned}
\nu_g^\epsilon(\theta_0) &= \underbrace{\int_{[-\pi,\pi]_{\mathrm{per}}} K_\epsilon(\theta_0 - \theta)\, d\nu_g(\theta)}_{\text{smoothed measure}} \\
&= \mathtt{C}_{\nu_g}\left(e^{i\theta_0}(1+\epsilon)^{-1}\right) - \mathtt{C}_{\nu_g}\left(e^{i\theta_0}(1+\epsilon)\right) \\
&= \underbrace{\frac{-1}{2\pi}\left[\langle (\mathcal{K} - e^{i\theta_0}(1+\epsilon))^{-1} g, \mathcal{K}^* g\rangle + e^{-i\theta_0}\langle g, (\mathcal{K} - e^{i\theta_0}(1+\epsilon))^{-1} g\rangle\right]}_{\text{approximate using matrices } \Psi_X^* W \Psi_X, \Psi_X^* W \Psi_Y, \Psi_Y^* W \Psi_Y}
\end{aligned}$$

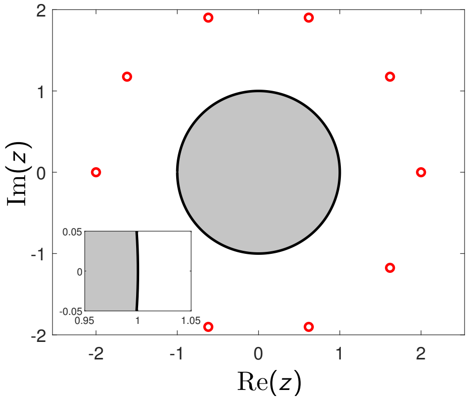
Compute smoothed approximations using ResDMD discretisations of size N_K.

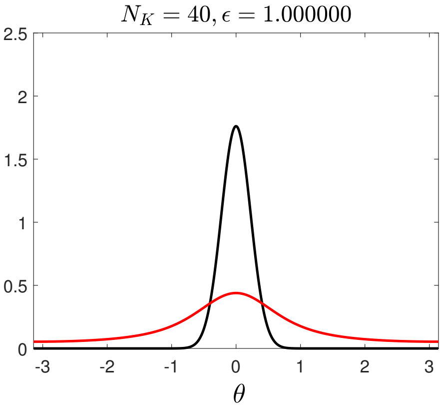
Example on $\ell^2(\mathbb{N})$ with known spectral measure

$$\mathcal{K} = \begin{bmatrix} \overline{\alpha_0} & \overline{\alpha_1}\rho_0 & \rho_1\rho_0 & & & \\ \rho_0 & -\overline{\alpha_1}\alpha_0 & -\rho_1\alpha_0 & 0 & & \\ 0 & \overline{\alpha_2}\rho_1 & -\overline{\alpha_2}\alpha_1 & \overline{\alpha_3}\rho_2 & \rho_3\rho_2 & \\ & \rho_2\rho_1 & -\rho_2\alpha_1 & -\overline{\alpha_3}\alpha_2 & -\rho_3\alpha_2 & \ddots \\ & & 0 & \overline{\alpha_4}\rho_3 & -\overline{\alpha_4}\alpha_3 & \ddots \\ & & & \ddots & \ddots & \ddots \end{bmatrix}, \alpha_j = (-1)^j 0.95^{(j+1)/2}, \rho_j = \sqrt{1-|\alpha_j|^2}.$$

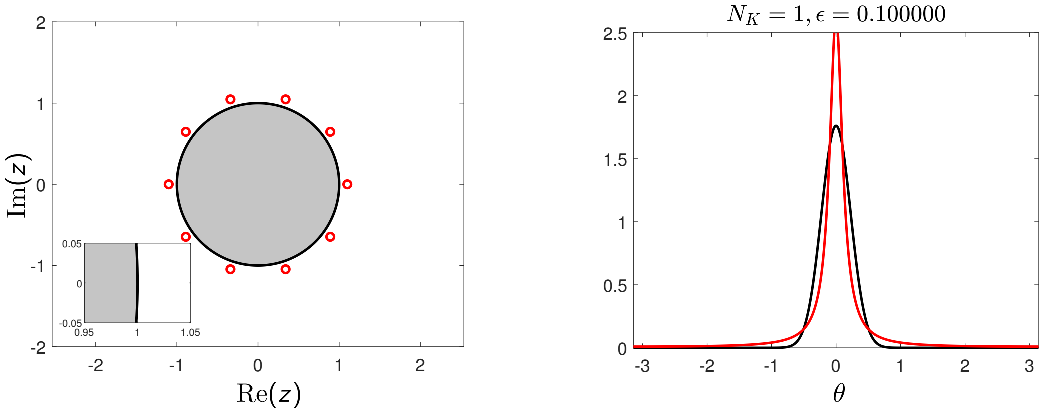
Generalised shift, typical building block of many dynamical systems (e.g., Bernoulli shifts).

Fix N_K, vary ϵ

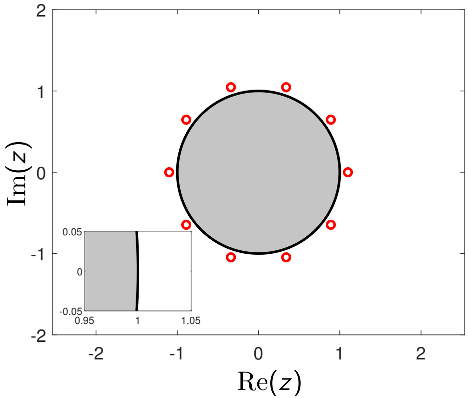


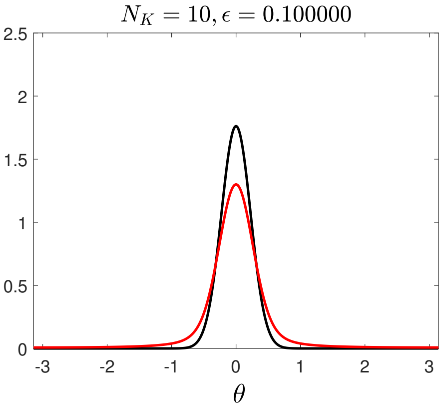


Fix ϵ, vary N_K



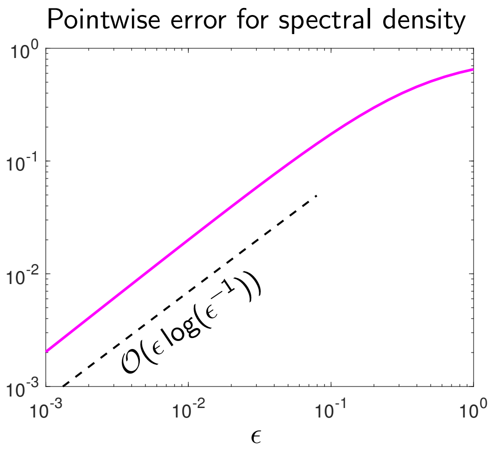
Adaptive $N_K(\epsilon)$ (or $\epsilon(N_K)$): New matrix $\Psi_Y^* W \Psi_Y$ key!

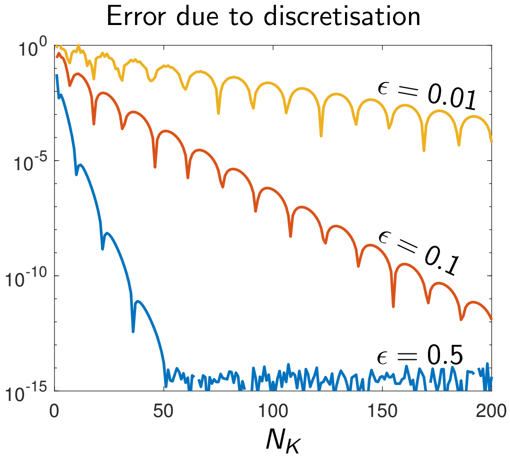




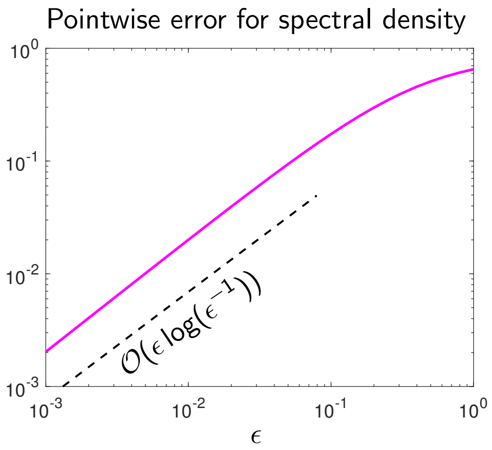
Slow convergence!

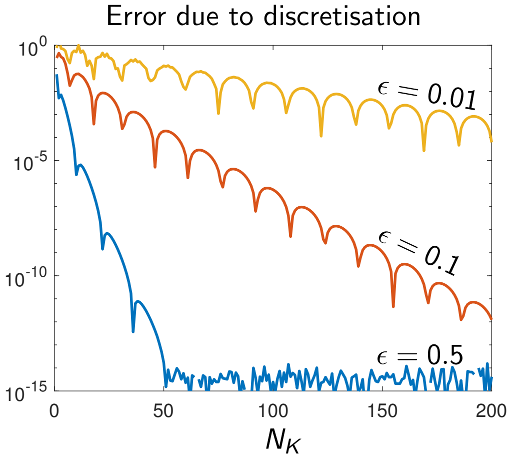
Problem: As $\epsilon \downarrow 0$, error is $\mathcal{O}(\epsilon \log(\epsilon^{-1}))$ and $N_K(\epsilon) \to \infty$.





Problem: As $\epsilon \downarrow 0$, error is $\mathcal{O}(\epsilon \log(\epsilon^{-1}))$ and $N_K(\epsilon) \to \infty$.





Critical in data-driven computations where we want N_K to be as small as possible.

Question: Can we improve the convergence rate in ϵ?

High-order kernels

Idea: Replace the Poisson kernel by

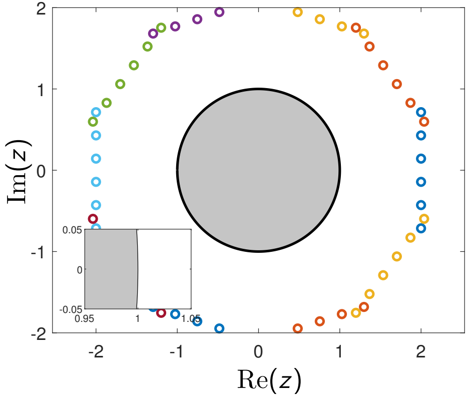
$$K_\epsilon(\theta) = \frac{e^{-i\theta}}{2\pi} \sum_{j=1}^{m} \left[\frac{c_j}{e^{-i\theta} - (1+\epsilon\overline{z_j})^{-1}} - \frac{d_j}{e^{-i\theta} - (1+\epsilon z_j)} \right]$$

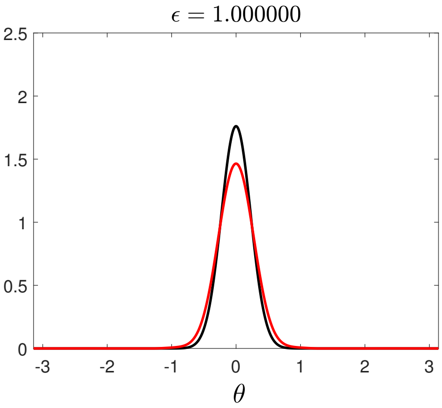
Simple way to select suitable z_j, c_j and d_j to achieve high-order kernel.

$$\nu_g^\epsilon(\theta_0) = \int_{[-\pi,\pi]_{\mathrm{per}}} K_\epsilon(\theta_0 - \theta)\, d\nu_g(\theta) = \sum_{j=1}^{m} \left[c_j \mathtt{C}_{\nu_g}\left(e^{i\theta_0}(1+\epsilon\overline{z_j})^{-1} \right) - d_j \mathtt{C}_{\nu_g}\left(e^{i\theta_0}(1+\epsilon z_j) \right) \right]$$

$\mathtt{C}_{\nu_g}(z)$ **computed using ResDMD.**

High-order kernels





Convergence

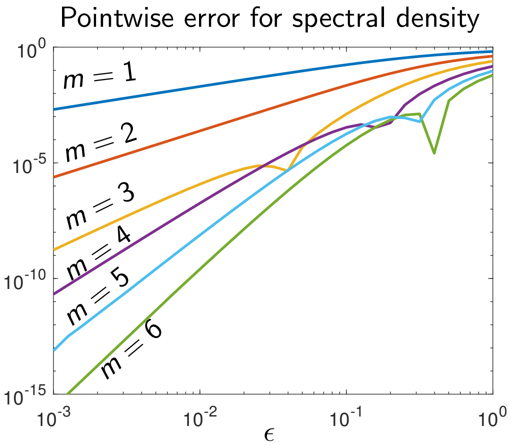
$\mathcal{O}(\epsilon^m \log(\epsilon^{-1}))$ convergence for:

- Pointwise recovery of the density ρ_g
- L^p recovery of ρ_g
- Weak convergence

$$\lim_{\epsilon \downarrow 0} \int_{[-\pi,\pi]_{\mathrm{per}}} \phi(\theta)\nu_g^{\epsilon}(\theta)\, d\theta = \int_{[-\pi,\pi]_{\mathrm{per}}} \phi(\theta)\, d\nu_g(\theta),$$

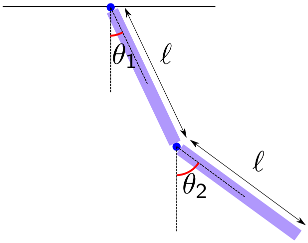
for periodic continuous ϕ.

Also recover discrete part of measure.
(i.e., eigenvalues of $\mathcal{K}$)



Evaluate at P values of θ: Parallelisable $\mathcal{O}(N_K^3 + PN_K)$ computation.

Example: double pendulum (chaotic)



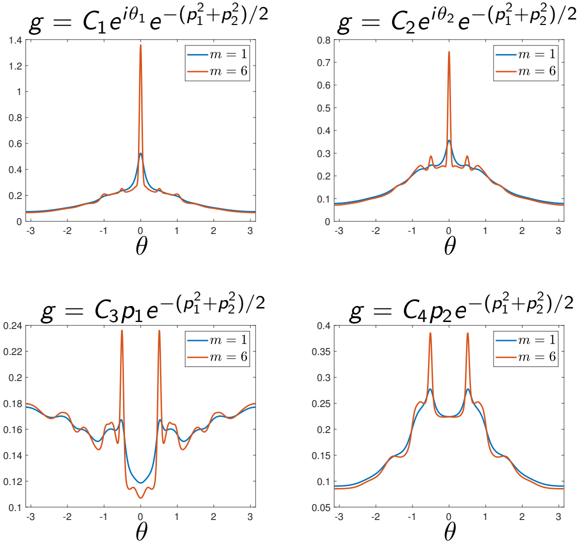
$$\dot{\theta}_1 = \frac{2p_1 - 3p_2\cos(\theta_1-\theta_2)}{16 - 9\cos^2(\theta_1-\theta_2)},$$

$$\dot{\theta}_2 = \frac{8p_2 - 3p_1\cos(\theta_1-\theta_2)}{16 - 9\cos^2(\theta_1-\theta_2)},$$

$$\dot{p}_1 = -3(\dot{\theta}_1\dot{\theta}_2\sin(\theta_1-\theta_2) + \sin(\theta_1)),$$

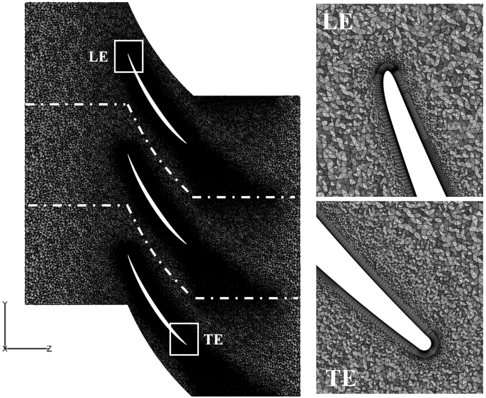
$$\dot{p}_2 = -3(-\dot{\theta}_1\dot{\theta}_2\sin(\theta_1-\theta_2) + \frac{1}{3}\sin(\theta_2)),$$

where $p_1 = 8\dot{\theta}_1 + 3\dot{\theta}_2\cos(\theta_1-\theta_2)$,
$p_2 = 2\dot{\theta}_2 + 3\dot{\theta}_1\cos(\theta_1-\theta_2)$



Part 3: High-dimensional dynamical systems and learned dictionaries.

Curse of dimensionality



Scalar field

$\Omega \subset \mathbb{R}^d$, d = number of grid/mesh points

E.g., polynomial dictionary up to tot. deg. 5.

Small grid: $d = 5 \times 5 \Rightarrow N_K \approx 50,000$.

Example later: $d \approx 300,000 \Rightarrow N_K \approx 2 \times 10^{25}$
$\gg$ **number of stars in known universe!!!!**

Conclusion: Infeasible to use hand-crafted dictionary when $d \gtrsim 25$.

Verified learned dictionaries

- Kernelized EDMD: $\mathcal{O}(d)$ cost using "kernel trick".
- Forms $\widetilde{K}_{\mathrm{EDMD}} \in \mathbb{C}^{M\times M}$ with subset of eigenvalues of $K_{\mathrm{EDMD}} \in \mathbb{C}^{N_K\times N_K}$.
- Implicitly learns dictionary: eigenfunctions of $\widetilde{K}_{\mathrm{EDMD}} \in \mathbb{C}^{M\times M}$.

· M. Williams, C. Rowley, and I. Kevrekidis "*A kernel-based method for data-driven Koopman spectral analysis*," J. Comput. Dyn., 2015.

Verified learned dictionaries

- Kernelized EDMD: $\mathcal{O}(d)$ cost using "kernel trick".
- Forms $\widetilde{K}_{\mathrm{EDMD}} \in \mathbb{C}^{M\times M}$ with subset of eigenvalues of $K_{\mathrm{EDMD}} \in \mathbb{C}^{N_K\times N_K}$.
- Implicitly learns dictionary: eigenfunctions of $\widetilde{K}_{\mathrm{EDMD}} \in \mathbb{C}^{M\times M}$.

When can you trust learning methods?

· M. Williams, C. Rowley, and I. Kevrekidis "*A kernel-based method for data-driven Koopman spectral analysis*," J. Comput. Dyn., 2015.

Verified learned dictionaries

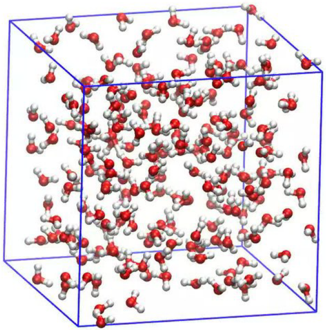
- Kernelized EDMD: $\mathcal{O}(d)$ cost using "kernel trick".
- Forms $\widetilde{K}_{\mathrm{EDMD}} \in \mathbb{C}^{M\times M}$ with subset of eigenvalues of $K_{\mathrm{EDMD}} \in \mathbb{C}^{N_K\times N_K}$.
- Implicitly learns dictionary: eigenfunctions of $\widetilde{K}_{\mathrm{EDMD}} \in \mathbb{C}^{M\times M}$.

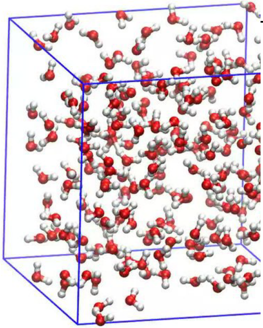
When can you trust learning methods?

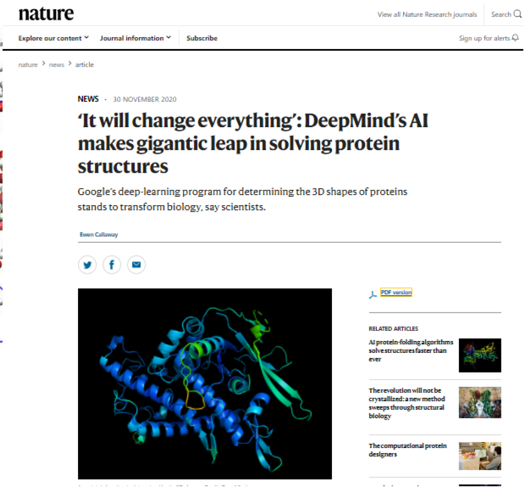
Combine with ResDMD: Convergence theory and a posterior verification of dictionary!

· M. Williams, C. Rowley, and I. Kevrekidis "*A kernel-based method for data-driven Koopman spectral analysis*," J. Comput. Dyn., 2015.

Molecular dynamics







nature

View all Nature Research journals | Search

Explore our content | Journal information | Subscribe | Sign up for alerts

nature > news > article

NEWS · 30 NOVEMBER 2020

‘It will change everything’: DeepMind’s AI makes gigantic leap in solving protein structures

Google’s deep-learning program for determining the 3D shapes of proteins stands to transform biology, say scientists.

Ewen Callaway

PDF version

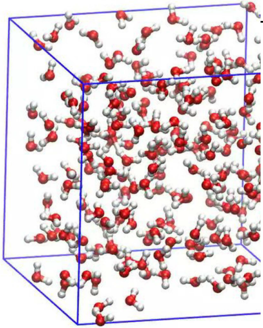
RELATED ARTICLES

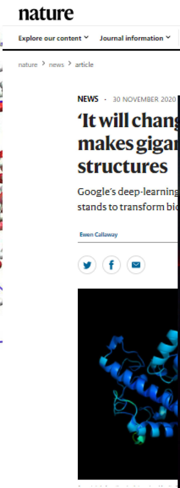
AI protein-folding algorithms solve structures faster than ever

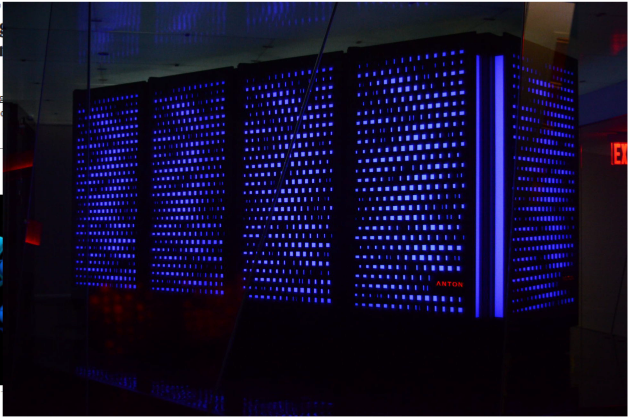
The revolution will not be crystallized: a new method sweeps through structural biology

The computational protein designers

Molecular dynamics

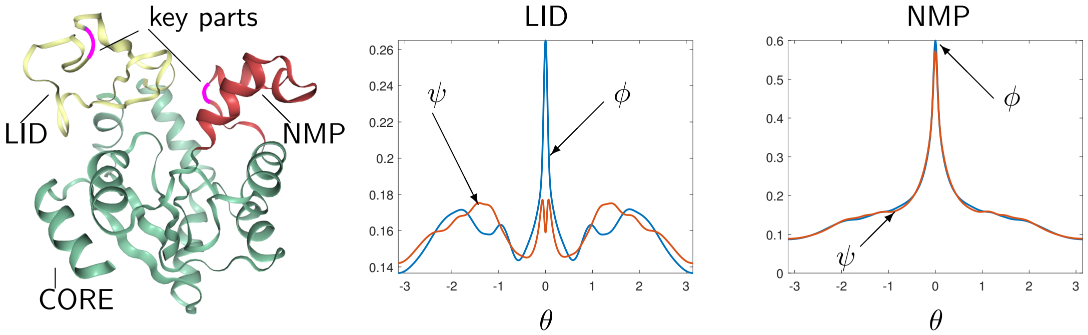






www.mdanalysis.org/MDAnalysisData/adk_equilibrium.html

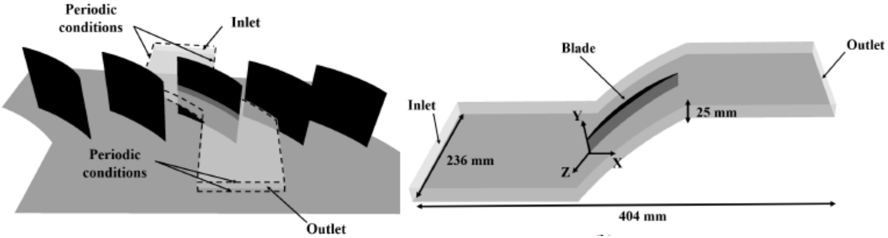
Spectral measures in molecular dynamics, $d = 20{,}046$



Left: ADK with three domains: CORE (green), LID (yellow) and NMP (red).
Middle and right: Spectral measures with respect to the dihedral angles of the selected parts.

Turbulent flow past a cascade of aerofoils, $d = 295{,}122$

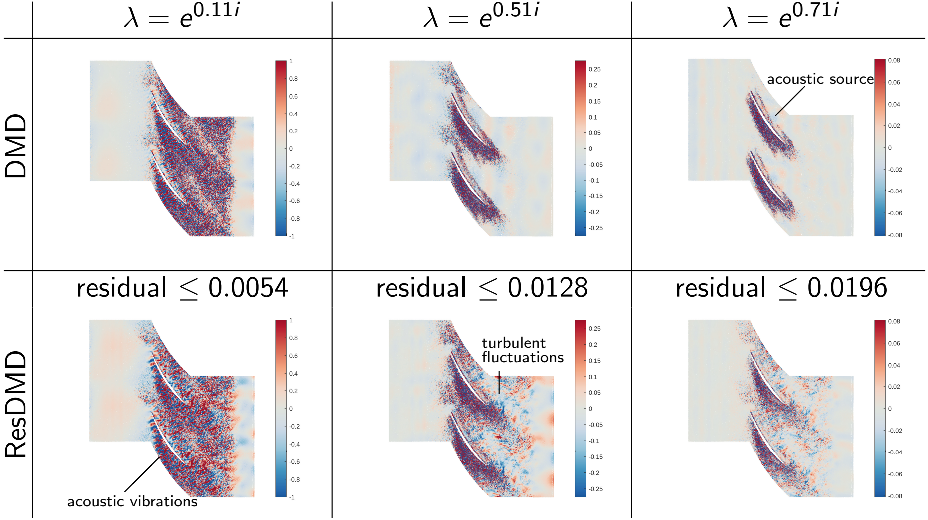
(Reynolds number 3.88×10^5.)



Motivation: Reduce noise sources (e.g., turbines, wings etc.).

· R. Koch, M. Sanjosé, and S. Moreau "*Large-Eddy Simulation of a Linear Compressor Cascade with Tip Gap: Aerodynamic and Acoustic Analysis*," AIAA Aviation, 2021.

Turbulent flow past a cascade of aerofoils, $d = 295{,}122$



Top row: Modes computed by DMD. **Bottom row:** Modes computed by ResDMD with residuals. Each column corresponds to different physical frequencies of noise pollution.

Wider programme: classification of computational problems

Example Question: What is possible in infinite-dimensional spectral computations?

How: Replace 'truncate-then-solve' with **infinite-dimensional numerical linear algebra.**

$\Rightarrow$ Compute many spectral properties for the first time.

Framework: Classify problems in **Solvability Complexity Index hierarchy**, measuring intrinsic difficulty.

$\Rightarrow$ Algorithms realise the boundaries of what computers can achieve.

Framework extends to: Barriers and foundations of AI (e.g., do there exist algorithms that train stable and accurate neural networks?), PDEs (e.g., solving the time-dependent Shrödinger equation on $L^2(\mathbb{R}^d)$), optimisation and precision analysis, computer-assisted proofs (e.g., which computations can be verified?), ...

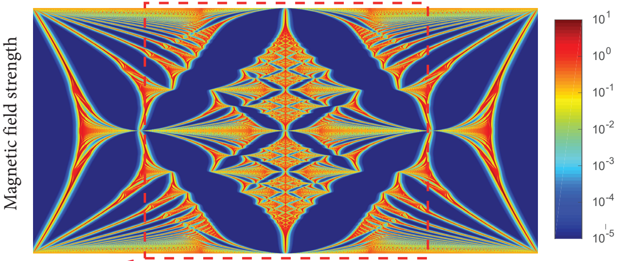
· M. Colbrook, "*The Foundations of Infinite-Dimensional Spectral Computations*," PhD diss., 2020.
· M. Colbrook, V. Antun , A. Hansen "*The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale's 18th problem*," Proc. Natl. Acad. Sci. USA, 2022.
· M. Colbrook, "*Computing semigroups with error control*," SINUM, 2022.
· M. Colbrook, A. Hansen "*The foundations of spectral computations via the solvability complexity index hierarchy*," JEMS, under revisions.
· M. Colbrook, "*On the computation of geometric features of spectra of linear operators on Hilbert spaces*," FOCM, under revisions.

Computing spectra with error control

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots \\ a_{21} & a_{22} & a_{23} & \cdots \\ a_{31} & a_{32} & a_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \qquad \left[A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \end{pmatrix} \right]_j = \sum_{k=1}^{\infty} a_{jk} x_k, \quad x \in l^2(\mathbb{N})$$

- Computes spectra with error control, i.e., output $\Gamma_n(A)$ and computed bound E_n such that
  - $\Gamma_n(A) \to \sigma(A)$ **(converges to spectrum)**
  - $\sup_{z \in \Gamma_n(A)} \mathrm{dist}(z, \sigma(A)) \leq E_n \downarrow 0.$ **(error control)**
- Avoids spectral pollution (spurious eigenvalues due to truncation)
- Rigorously compute approximate states.
- Extends to (certain) non-normal ($AA^* \neq A^*A$) operators.
- Extends to partial differential operators.

· M.Colbrook, B. Roman, A. Hansen "How to compute spectra with error control" Physical Review Letters, 2019.



Horizontal slice = spectral measure at constant magnetic field strength.

Software package:

`SpecSolve` available at https://github.com/SpecSolve
Current capabilities include: ODEs on real line & half-line, integral operators, and discrete operators.

- M. Colbrook, "*Computing spectral measures and spectral types*" Communications in Mathematical Physics, 2021.
- M. Colbrook, A. Horning, A. Townsend "*Computing spectral measures of self-adjoint operators*" SIREV, 2021.

Barriers of deep learning: stability and accuracy

PNAS

RESEARCH ARTICLE | APPLIED MATHEMATICS | FULL ACCESS

The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale's 18th problem

Matthew J. Colbrook, Vegard Antun, and Anders C. Hansen Authors Info & Affiliations

March 16, 2022 | 119 (12) e2107151119 | https://doi.org/10.1073/pnas.2107151119

Significance

Instability is the Achilles' heel of modern artificial intelligence (AI) and a paradox, with training algorithms finding unstable neural networks (NNs) despite the existence of stable ones. This foundational issue relates to Smale's 18th mathematical problem for the 21st century on the limits of AI. By expanding methodologies initiated by Gödel and Turing, we demonstrate limitations on the existence of (even randomized) algorithms for computing NNs. Despite numerous existence results of NNs with great approximation properties, only in specific cases do there also exist algorithms that can compute them. We initiate a classification theory on which NNs can be trained and introduce NNs that—under suitable conditions—are robust to perturbations and exponentially accurate in the number of hidden layers.

Mathematical paradox demonstrates the limits of AI

Humans are usually pretty good at recognising when they get things wrong, but artificial intelligence systems are not. According to a new study, AI generally suffers from inherent limitations due to a century-old mathematical paradox.

Like some people, AI systems often have a degree of confidence that far exceeds their actual abilities. And like an overconfident person, many AI systems don't know when they're making mistakes. Sometimes it's even more difficult for an AI system to realise when it's making a mistake than to produce a correct result.

Researchers from the University of Cambridge and the University of Oslo say that instability is the Achilles' heel of modern AI and that a mathematical paradox shows AI's limitations. Neural networks, the state-of-the-art tool in AI

> *There are fundamental limits inherent in mathematics and, similarly, AI algorithms can't exist for certain problems*
>
> — *Matthew Colbrook*

Concluding remarks

Summary: Rigorous and practical algorithms that overcome the challenges of (C1) Continuous spectra, (C2) Lack of finite-dimensional invariant subspaces, (C3) Spectral pollution, and (C4) Chaotic behaviour.

Part 1: Computed spectra, pseudospectra and residuals of general Koopman operators.

Idea: New matrix for residual $\Rightarrow$ ResDMD.

Part 2: Computed spectral measures of measure-preserving systems with high-order convergence. Density of continuous spectrum, discrete spectrum and weak convergence.

Idea: Convolution with rational kernels through the resolvent and ResDMD.

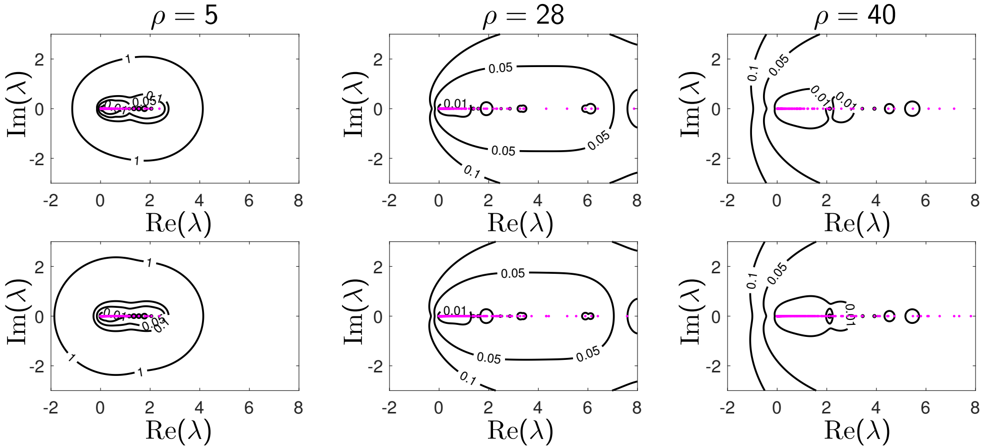
Part 3: Dealt with high-dimensional dynamical systems.

Idea: ResDMD to verify learned dictionaries.

Part of a wider programme on foundations of computation and numerical analysis.

Example: Lorenz and extended Lorenz systems

$$\dot{X} = 10\,(Y - X)\,, \quad \dot{Y} = X\,(\rho - Z) - Y, \quad \dot{Z} = XY - 8Z/3.$$



Top row: Lorenz system. **Bottom row:** Extended 11-dimensional Lorenz system.

· S. Moon et al. "*Periodicity and chaos of high-order Lorenz systems*," Inter. J. Bifur. Chaos, 2017.

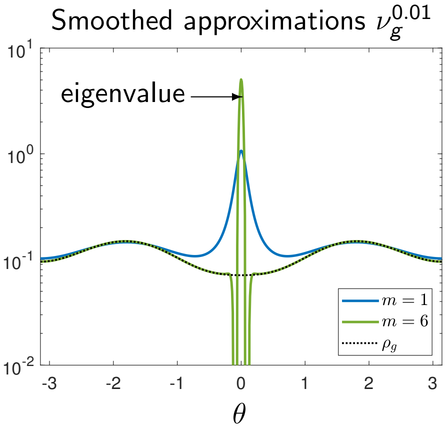
Example: Lorenz and extended Lorenz systems

$\rho = 5$				$\rho = 28$				$\rho = 40$			
$d = 3$		$d = 11$		$d = 3$		$d = 11$		$d = 3$		$d = 11$	
λ_j	r_j	λ_j	r_j	λ_j	r_j	λ_j	r_j	λ_j	r_j	λ_j	r_j
1.0108	4.9E-7	1.0108	8.6E-5	1.0423	5.1E-6	1.0346	2.6E-4	1.0689	4.6E-4	1.0046	6.2E-04
1.0217	3.8E-4	1.1550	1.1E-6	1.0712	7.9E-4	1.0423	1.9E-5	1.2214	2.9E-6	1.0868	1.1E-04
1.1550	5.1E-8	1.3339	1.0E-5	1.0862	6.3E-4	1.0472	4.8E-4	1.4191	9.9E-4	1.2214	1.3E-05
1.1675	7.6E-5	1.3380	5.2E-4	1.3839	7.5E-5	1.0594	7.7E-5	1.4823	4.9E-4	1.2419	8.3E-07
1.3340	1.3E-6	1.5410	4.0E-4	1.5810	4.4E-7	1.0598	2.0E-6	1.4916	4.8E-4	1.2452	6.7E-04
1.3385	6.9E-4			1.8065	7.4E-8	1.0685	9.8E-4	1.6216	5.2E-5	1.2526	1.2E-04
1.5410	3.1E-4			1.8829	5.8E-4	1.0707	9.4E-4	1.8527	1.7E-7	1.3498	1.7E-04
				2.8561	7.2E-5	1.0862	8.2E-4	2.1170	7.5E-8	1.3541	9.6E-04
				3.2633	2.9E-7	1.1964	2.4E-4	2.5857	3.7E-4	1.4251	1.5E-04
				5.8954	3.1E-4	1.3675	1.3E-6	3.9223	6.2E-5	1.4788	6.9E-04

Eigenvalues computed using Algorithm 1 with $\epsilon = 0.001$ along with the computed residuals r_j.

Example: tent map, $F(x) = 2\min\{x, 1-x\}$, $\Omega = [0,1]$

$$g(\theta) = C|\theta - 1/3| + C\sin(20\theta) + \begin{cases} C, & \theta > 0.78, \\ 0, & \theta \le 0.78. \end{cases}$$



Added benefit: Avoid oversmoothing, and have better localisation of singular parts.