Spectral theory and approximation of Koopman operators in chaos

Part 1: smooth functional analysis of dynamical operators

Caroline Wormell

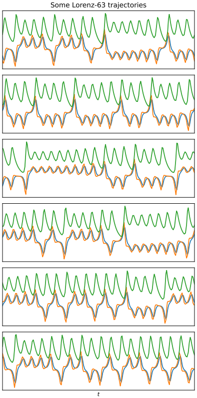
The University of Sydney

July 24, 2024

Trying to make sense of dynamical systems

The basic object of study in “dynamics” is a trajectory $\{x_t\}_{t=0,1,2,\ldots}$ in a state space D.

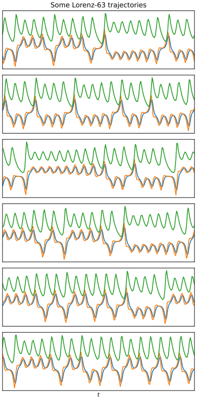
- The set of trajectories of a dynamical system can be very complicated (like people)
- Tendency to study mean behaviour (or expected behaviour with respect to some loss. . .)
- “Mean”,“expectation” implies study with respect to probability. In dynamics, this is *ergodic theory*



Smooth ergodic theory

$$\frac{1}{T}\sum_{t=0}^{T}\psi(x_t) \xrightarrow{T\to\infty} \int_D \psi(x_t)\,\mathrm{d}\mu(x) \text{ for } \mu\text{-almost all } x$$

- Birkhoff ergodic theorem (true in many systems): time averages over orbit = spatial average
  - Means almost every orbit looks like every other orbit at some point in time ("almost all orbits are dense").
- We want to compute so want, e.g. stability to error, ideally with quantitative assurances.
- This suggests studying *smooth* ergodic theory.



Operators

So, we are interested in studying

- The Koopman operator
- The transfer operator ($\approx$ Perron-Frobenius operator)

on function spaces involving differentiability.

It is mostly a separate beast to the $L^p(\mu)$ theory.

Stochastic vs deterministic

Broadly are going to consider three kinds of dynamical system (on compact manifolds):

- **Noisy dynamics** (SDEs. . .): theoretically easy, good starting point for comparison
- **Deterministic contractions**: already known, but explains some of the questions in deterministic dynamics. . .
- **Deterministic chaos**: harder and a bit obscure, main goal of this minicourse

Stochastic vs deterministic

Broadly are going to consider three kinds of dynamical system (on compact manifolds):

- **Noisy dynamics** (SDEs. . .): theoretically easy, good starting point for comparison
- **Deterministic contractions**: already known, but explains some of the questions in deterministic dynamics. . .
- **Deterministic chaos**: harder and a bit obscure, main goal of this minicourse

Objection! All real systems have a *bit* of noise in them!

Stochastic vs deterministic

Broadly are going to consider three kinds of dynamical system (on compact manifolds):

- **Noisy dynamics** (SDEs. . .): theoretically easy, good starting point for comparison
- **Deterministic contractions**: already known, but explains some of the questions in deterministic dynamics. . .
- **Deterministic chaos**: harder and a bit obscure, main goal of this minicourse

Objection! All real systems have a *bit* of noise in them!

Then you commit to resolving down to the scale of the noise!

Stochastic vs deterministic

Broadly are going to consider three kinds of dynamical system (on compact manifolds):

- **Noisy dynamics** (SDEs. . .): theoretically easy, good starting point for comparison
- **Deterministic contractions**: already known, but explains some of the questions in deterministic dynamics. . .
- **Deterministic chaos**: harder and a bit obscure, main goal of this minicourse

Objection! All real systems have a *bit* of noise in them!

Then you commit to resolving down to the scale of the noise! How can we believe lower-resolution numerics?

Dual notions

(Markov, time-autonomous) stochastic dynamics $x_t \in D$ are typically studied using two linear operators, acting on functions on the state space $\psi, \varphi : D \to \mathbb{R}$:

- The Chapman-Kolmogorov operator: predicting the expected future value of "observable" functions

$$(\mathcal{K}\psi)(x) = \mathbb{E}[\psi(x_{t+1}) | x_t = x]$$

- The Fokker-Planck operator: evolution of probabilities into future (= push-forward of measure density)

$$(\mathcal{L}\varphi)(x) = \int \varphi(y) \, \frac{\mathrm{d}\mathbb{P}}{\mathrm{d}x}[x_{t+1} = x | x_t = y] \, \mathrm{d}y$$

Dual notions

(Markov, time-autonomous) stochastic dynamics $x_t \in D$ are typically studied using two linear operators, acting on functions on the state space $\psi, \varphi : D \to \mathbb{R}$:

- The Chapman-Kolmogorov operator: predicting the expected future value of "observable" functions

$$(\mathcal{K}^{n}\psi)(x) = \mathbb{E}[\psi(x_{t+n})|x_t = x]$$

- The Fokker-Planck operator: evolution of probabilities into future (= push-forward of measure density)

$$(\mathcal{L}^{n}\varphi)(x) = \int \varphi(y)\,\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}x}[x_{t+n} = x|x_t = y]\,\mathrm{d}y$$

Dual notions

(Markov, time-autonomous) stochastic dynamics $x_t \in D$ are typically studied using two linear operators, acting on functions on the state space $\psi, \varphi : D \to \mathbb{R}$:

- The Chapman-Kolmogorov operator: predicting the expected future value of "observable" functions

$$(\mathcal{K}^{n}\psi)(x) = \mathbb{E}[\psi(x_{t+n})|x_t = x]$$

- The Fokker-Planck operator: evolution of probabilities into future (= push-forward of measure density)

$$(\mathcal{L}^{n}\varphi)(x) = \int \varphi(y)\,\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}x}[x_{t+n} = x|x_t = y]\,\mathrm{d}y$$

Dual notions

Define our transition kernel $k(x,y) = \frac{\mathrm{d}\mathbb{P}}{\mathrm{d}y}[x_{t+1} = y | x_t = x]$, then

$$(\mathcal{L}\varphi)(y) = \int \varphi(x)\, k(x,y)\, \mathrm{d}x$$

whereas

$$\begin{aligned}(\mathcal{K}\psi)(x) &= \mathbb{E}[\psi(x_{t+1}) | x_t = x] \\ &= \int \psi(y) \mathrm{d}\mathbb{P}[x_{t+1} = y | x_t = x] \\ &= \int \psi(y)\, k(x,y)\, \mathrm{d}y.\end{aligned}$$

So, these operators are dual:

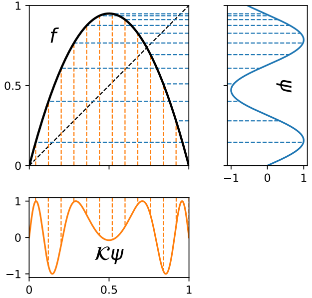
$$\int (\mathcal{L}\varphi)(y)\, \psi(y)\, \mathrm{d}y = \int \varphi(y)\, (\mathcal{K}\psi)(y)\, \mathrm{d}y.$$

Dual notions

These actually work with deterministic maps $f : D \to D$ as well (modulo "some" intricacies. . .)

- The **Koopman operator**: expected future value of functions (aka "observables")

$$(\mathcal{K}\psi)(x) = \psi(f(x))$$

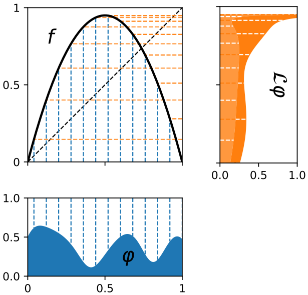


Dual notions

- The **transfer operator**, which gives you the evolution of probabilities into the future (= push-forward of measure density)

$$(\mathcal{L}\varphi)(x) = \int \text{“}\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}x}[f(y) = x]\text{”}\, \varphi(y)\,\mathrm{d}y$$

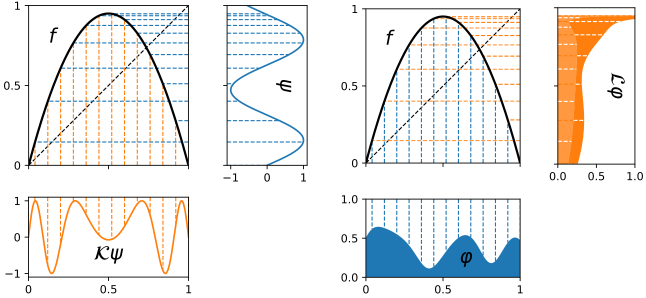
$$= \sum_{f(y)=x} \frac{\varphi(y)}{|\det Df(y)|}$$



Dual notions

These two operators are also still dual!

$$\int \varphi(x)\,(\mathcal{K}\psi)(x)\,\mathrm{d}x = \int (\mathcal{L}\varphi)(x)\,\psi(x)\,\mathrm{d}x$$



Dynamics

dynamics, n. *The study of trajectories as time goes to infinity.*

(You tell me if that's a good thing.)

Dynamics

dynamics, n. *The study of trajectories as time goes to infinity.*

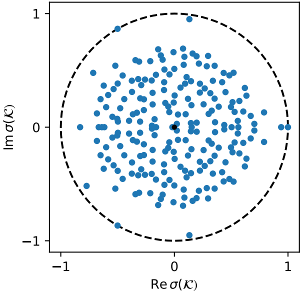
(You tell me if that's a good thing.)

- We can describe dynamics over long times by $\mathcal{K}^n$, $\mathcal{L}^n$, for n large.
- These are best described by the *spectrum* of $\mathcal{K}, \mathcal{L}$.
- Consequently, Koopman/transfer spectra can:
  - Provide reductions for the dynamics
  - Give you statistical information about the system
  - Make sense of the emergent dynamical geometry...

Koopman spectra

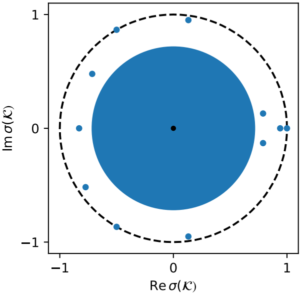
What are the spectra of these (infinite-dimensional, weirdly-posed) operators?

Generally computer approximations of the Koopman spectrum look like:



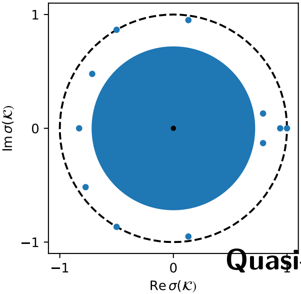
Koopman spectra

What are the spectra of these (infinite-dimensional, weirdly-posed) operators?
Generally computer approximations of the Koopman spectrum seem to be approximating



Koopman spectra

What are the spectra of these (infinite-dimensional, weirdly-posed) operators?
Generally computer approximations of the Koopman spectrum seem to be approximating

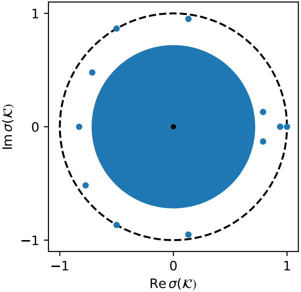


Quasi-compact operator

Structure

A lot of numerics around dynamical systems come down to studying the spectra of these operators numerically.
This lecture series will talk about

1. A mathematical framework that explains quasi-compact operators
2. How different kinds of dynamics fit into the framework
3. How this translates to computation

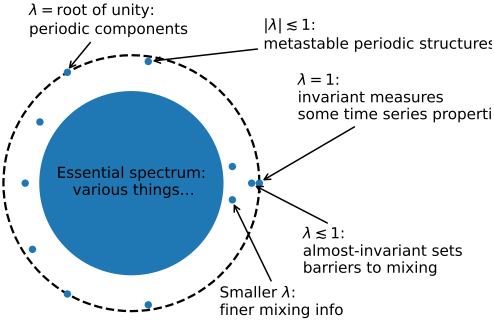


Meaning of the spectrum

Quasicompactness (and $\sigma_{\mathrm{d}}(\mathcal{K})$ with no Jordan blocks) gives:

$$\mathcal{K}^n \psi = \sum_{\substack{\lambda_k \in \sigma_{\mathrm{d}}(\mathcal{K}) \\ \mathcal{K}\psi_k = \lambda_k \psi_k}} c_k(\psi)\, \lambda_k^n\, \psi_k + \mathcal{O}(\rho_{\mathrm{ess}}(\mathcal{K})^n)$$

Different parts of the spectrum have various interpretations:



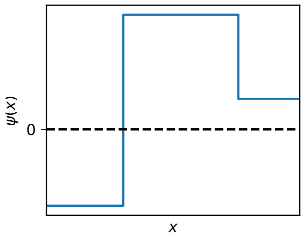
$\lambda = 1$: ergodic components

Proposition

Suppose a **function** $\psi : D \to \mathbb{R}$ *satisfies* $\psi \circ f = \psi$. *Then the level sets of* ψ *are invariant sets.*

Proof.

If $x \in \psi^{-1}(c)$, then $\psi(f(x)) = \psi(x) = c$, so $f(x) \in \psi^{-1}(c)$. □



$|\lambda| = 1$: periodicity

Proposition

Suppose a **function** $\psi : D \to \mathbb{R}$ *satisfies* $\psi \circ f = e^{i\theta\psi}$ *for some* $\theta \in [0, 2\pi]$. *Let* $E_z = \{x \in D : \psi(x) = z\}$.
Then f *maps* E_z *into* $E_{e^{i\theta}z}$.

Proof.

If $x \in E_z$, then $\psi(f(x)) = e^{i\theta}\psi(x) = e^{i\theta}z$, so $f(x) \in E_{e^{i\theta}z}$. □

$|\lambda| = 1$: periodicity

Proposition

Suppose a **function** $\psi : D \to \mathbb{R}$ *satisfies* $\psi \circ f = e^{i\theta\psi}$ *for some* $\theta \in [0, 2\pi]$. *Let* $E_z = \{x \in D : \psi(x) = z\}$.
Then f^n *maps* E_z *into* $E_{e^{in\theta}z}$.

Proof.

If $x \in E_z$, then $\psi(f(x)) = e^{i\theta}\psi(x) = e^{i\theta}z$, so $f(x) \in E_{e^{i\theta}z}$. □

$\lambda \lesssim 1$: almost invariant sets

Proposition (à la Froyland and Stancevic '10)

Suppose ψ *satisfies* $\mathcal{K}\psi = \lambda\psi$ *and* $\sup|\psi| \leq 1$ *for some* $\lambda \in (0,1)$. *Let*

$$E_+ = \{x \in D : \psi(x) > 0\}.$$

Then for some $C > 0$,

$$\int_D \mathbb{P}\left(f^t(x) \in E_+ \text{ for } t = 1, \ldots, n\right) \mathrm{d}x \geq C\lambda^{n-1}$$

That is, the Lebesgue measure of the set of points that don't leave E_+ within n steps decays as $\mathcal{O}(\lambda^n)$.
In particular $n \sim 1/(1-\lambda)$, this set is of $\mathcal{O}(1)$ Lebesgue measure.

$\lambda \lesssim 1$: almost invariant sets

Proof.

We are interested in

$$P_n = \int_D \mathbb{P}\left(x_t \in E_+ \text{ for } t = 0\dots, n-1 | x_t = x\right) \mathrm{d}x.$$

The interior probability we can rewrite as

$$\mathbb{E}\left(\prod_{t=0}^{n-1} \mathbb{1}_{E_+}(x_t) | x_0 = x\right)$$

$$\mathbb{1}_{E_+}(x) = \begin{cases} 1 & x \in E_+ \\ 0 & x \notin E_+ \end{cases}$$

- If $n = 1$, this is $\mathbb{E}[\mathbb{1}(x_0 | x_0 = x] = \mathbb{1})(x)$.
- If $n = 2$, this is $\mathbb{E}[\mathbb{1}_{E_+}(x_0)\mathbb{1}_{E_+}(x_1) | x_0 = x] = \mathbb{1}_{E_+}(x)\mathcal{K}[\mathbb{1}_{E_+}](x)$.
- If $n = 3$, this is $\mathbb{E}[\mathbb{1}_{E_+}(x_0)\mathbb{1}_{E_+}(x_1)\mathbb{1}_{E_+}(x_2) | x_0 = x] = \mathbb{1}_{E_+}(x)\mathcal{K}[\mathbb{1}_{E_+}\mathcal{K}[\mathbb{1}_{E_+}]](x)$.
- By induction, we have $(\mathbb{1}_{E_+}\mathcal{K})^{n-1}[\mathbb{1}_{E_+}](x)$.

$\lambda \lesssim 1$: almost invariant sets

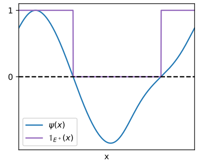
Proof (continued).

Now, $\mathcal{K}$ is a positive operator (i.e. $a \geq b \implies \mathcal{K}a \geq \mathcal{K}b$), and $\mathbb{1}_{E_+} \geq \psi$, so

$$(\mathcal{K}\mathbb{1}_{E_+})^{n-1}[\mathbb{1}_{E_+}](x) \geq (\mathcal{K}\mathbb{1}_{E_+})^{n}[\psi](x)$$

Furthermore, $\mathbb{1}_{E_+}\mathcal{K}\psi = \lambda\mathbb{1}_{E_+}\psi \geq \lambda\psi$.
So, inductively,

$$(\mathbb{1}_{E_+}\mathcal{K})^{n-1}[\psi](x) \geq \lambda^{n-1}\mathbb{1}_{E_+}\psi(x).$$



So for some $C > 0$,

$$P_n \geq \int_D \lambda^{n-1}\mathbb{1}_{E_+}(x)\psi(x)\,\mathrm{d}x = C\lambda^{n-1}.$$

□

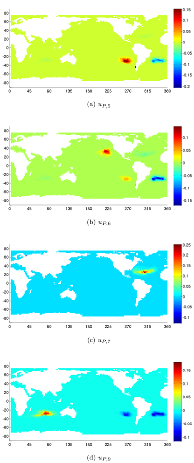
$\lambda \lesssim 1$: almost invariant sets

- Comparable results for complex λ. If

$$E_\theta = \{x \in D : \Re[e^{i\theta}\psi(x)] > 0\},$$

  you have that E_θ mostly maps to $E_{\theta+\arg\lambda}$.
- Same result for transfer operator $\mathcal{L}$ (again if bounded eigenfunctions).

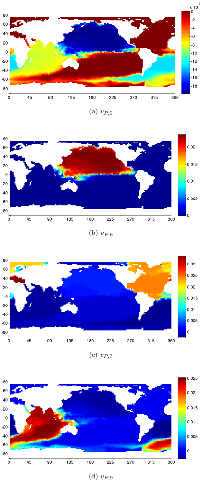
$\lambda \lesssim 1$: garbage patch example



(a) $u_{P,5}$ (b) $u_{P,6}$ (c) $u_{P,7}$ (d) $u_{P,9}$

FIG. 4: Maps of selected left eigenvectors of P showing the locations of the five great ocean garbage patches.

Transfer operator eigenfunctions



(a) $v_{P,5}$ (b) $v_{P,6}$ (c) $v_{P,7}$ (d) $v_{P,9}$

FIG. 5: Maps of right eigenvectors $\{v_{P,r}\}$ of P.

Koopman eigenfunctions

λ	P
λ_1	1.0000
λ_2	1.0000
λ_3	1.0000
λ_4	1.0000
λ_5	0.9999
λ_6	0.9999
λ_7	0.9996
λ_8	0.9991
λ_9	0.9975
λ_{10}	0.9913
λ_{11}	0.9852
λ_{12}	0.9838
λ_{13}	0.9826
λ_{14}	0.9680
λ_{15}	0.9645

Note: in a Lebesgue-orthogonal basis, $\mathcal{L}$ is the transpose of $\mathcal{K}$

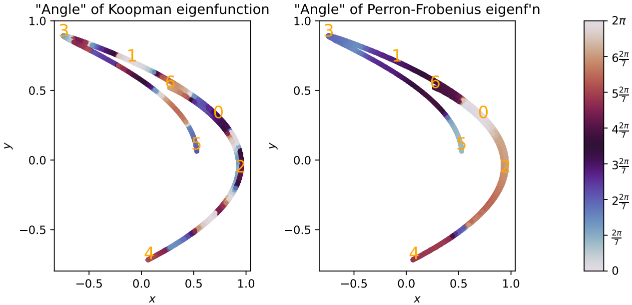
Same spectrum, different eigenfunctions.

Transfer: attractors

Koopman: basins of attraction

Koopman and transfer operator eigenfunctions

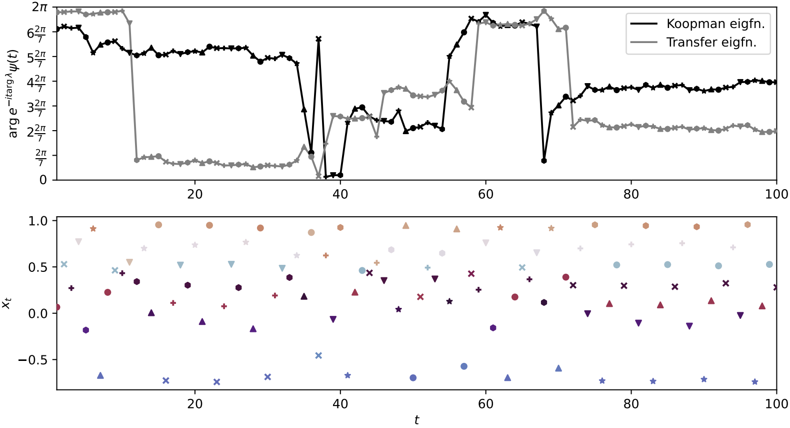
Using EDMD I just computed that a Hénon-like map has an eigenvalue $\lambda \approxeq 0.92 e^{-6i\pi/7}$.
This suggests some sort of 7-periodic behaviour persisting over timescale $\sim 1/-\log 0.92 = 13$.



Koopman and transfer operator eigenfunctions

Using EDMD I just computed that a Hénon-like map has an eigenvalue $\lambda \approxeq 0.92e^{-6i\pi/7}$.

This suggests some sort of 7-periodic behaviour persisting over timescale $\sim 1/-\log 0.92 = 13$.



A motivating conundrum

This lecture, let's try and get a theoretical grip on the spectrum of the Koopman operator for a deterministic contraction ($\kappa < 1$):

$$f(x) = \kappa x, \; x \in [-1, 1]$$

So the Koopman operator is

$$(\mathcal{K}\psi)(x) = \psi(\kappa x)$$

What does its spectrum look like?
Let's try and find some eigenfunctions.

A motivating conundrum

Let's sub in a power series $\psi(x) = \sum_{k=0}^{\infty} a_k x^k$:

$$0 = \mathcal{K}\psi - \lambda\psi$$
$$= \sum_{k=0}^{\infty} a_k \kappa^k x^k - \lambda \sum_{k=0}^{\infty} a_k x^k$$

Equating terms we get

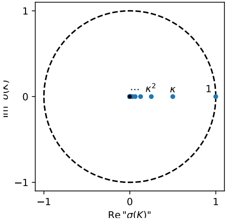
$$(\kappa^k - \lambda) a_k = 0, k \in \mathbb{N}$$

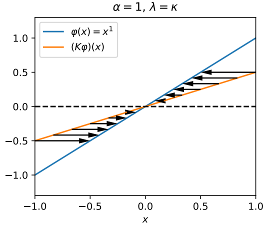
suggesting that our eigenvalues are $\{1, \kappa, \kappa^2, \kappa^3, \ldots\}$ with the respective eigenfunctions $\{1, x, x^2, x^3, \ldots\}$

A motivating conundrum

These eigenfunctions broadly give us what we expect:

- The leading eigenfunction 1 is simple (i.e. no separate basins of attraction*)
- The next eigenfunction has a root at $x = 0$, suggesting a dynamical barrier here (in fact, it is precisely the fixed point—woohoo)
- Some other spectrum accumulating at 0

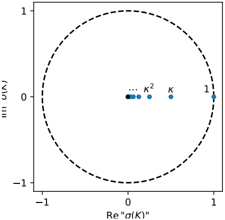


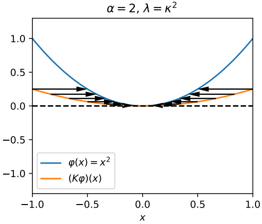


A motivating conundrum

These eigenfunctions broadly give us what we expect:

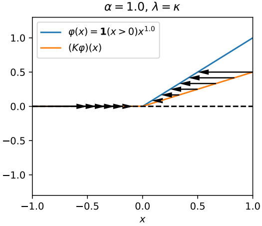
- The leading eigenfunction 1 is simple (i.e. no separate basins of attraction*)
- The next eigenfunction has a root at $x = 0$, suggesting a dynamical barrier here (in fact, it is precisely the fixed point—woohoo)
- Some other spectrum accumulating at 0





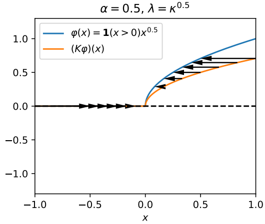
A motivating conundrum

But, by the same token, we could try $\mathcal{K}\mathbb{1}(x > 0)x = \mathbb{1}(x > 0)x$, and so on...



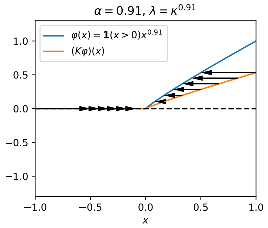
A motivating conundrum

But, by the same token, we could try
$\mathcal{K}\mathbb{1}(x > 0)x^{1/2} = \sqrt{\kappa}\mathbb{1}(x > 0)x^{1/2}$, and so on...



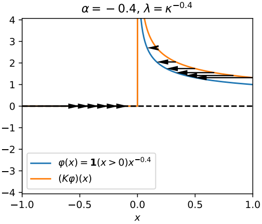
A motivating conundrum

But, by the same token, we could try
$\mathcal{K}\mathbb{1}(x > 0)x^{1/2} = \sqrt{\kappa}\mathbb{1}(x > 0)x^{1/2}$, and so on...



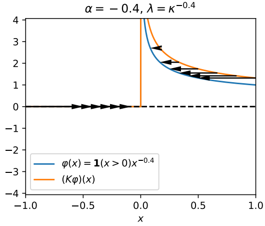
A motivating conundrum

But, by the same token, we could try
$\mathcal{K}\mathbb{1}(x>0)x^{1/2} = \sqrt{\kappa}\mathbb{1}(x>0)x^{1/2}$, and so on...



A motivating conundrum

But, by the same token, we could try
$\mathcal{K}\mathbb{1}(x>0)x^{1/2} = \sqrt{\kappa}\mathbb{1}(x>0)x^{1/2}$, and so on...



In fact, for any $\alpha \in \mathbb{C}$, we can set $\psi(x) = x^{\alpha} = e^{-\alpha \log x}$ for $x > 0$ and get

$$(\mathcal{K}\phi)(x) = e^{-\alpha(\log x - \log \kappa)} = \kappa^{-\alpha}\psi(x)$$

So, $\mathcal{K}$'s spectrum could cover the complex plane unless we are a bit careful about what functions we allow.

A motivating conundrum

In fact, this comes from a bigger fact which is that Koopman eigenfunctions/eigenvalues are multiplicative in deterministic dynamics.

In general, if

- $\mathcal{K}\psi = \lambda\psi$
- f is deterministic;
- ψ^α is well-defined;

then $\mathcal{K}[\psi^\alpha] = \lambda^\alpha\psi^\alpha$.

A motivating conundrum

How do we allow/banish functions from our linear operator $\mathcal{K}$? We **set a function space as the domain of $\mathcal{K}$.**
Crucial properties of this function space $\mathcal{B}$:

- It is a vector space.
- It has a norm $\|\cdot\|$, with respect to which it is complete (i.e. it's a *Banach space*)
- $\mathcal{K}$ maps $\mathcal{B}$ to itself.
- It doesn't have to contain *only* functions, but should contain all sufficiently nice functions (e.g. C_c^∞)

Note that to do theory it isn't very helpful to have a Hilbert space, except in some cases.

Spectrum of an infinite-dimensional operator

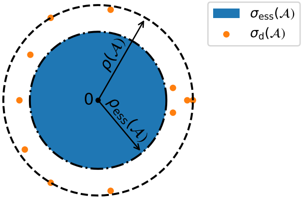
Define the resolvent of an operator $\mathcal{A} : \mathcal{B} \to \mathcal{B}$:

$$R_\lambda(\mathcal{A}) = (\mathcal{A} - \lambda I)^{-1} : \mathcal{B} \to \mathcal{B}$$

The spectrum $\sigma(\mathcal{A})$ is the set of $\lambda \in \mathbb{C}$ where $R_\lambda(\mathcal{A})$ is either not well-defined, or unbounded. It is always closed.
The spectrum includes:

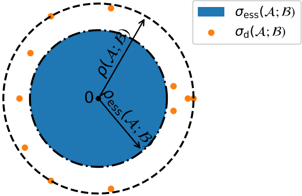
- The discrete spectrum $\sigma_{\mathrm{d}}(\mathcal{A})$, i.e. isolated eigenvalues λ of $\mathcal{A}$ with finite "algebraic multiplicity".
  The nice normal stuff we love from finite-dimensional operators.
- The rest $\sigma_{\mathrm{ess}}(\mathcal{A})$—the "essential spectrum". For Koopman operators in discrete time it is *usually* a ball around 0.

Spectral radii



- Spectral radius $\rho(\mathcal{A}) = \max|\sigma(\mathcal{A})|$
- Essential spectral radius $\rho_{\mathrm{ess}}(\mathcal{A}) = \max|\sigma_{\mathrm{ess}}(\mathcal{A})|$

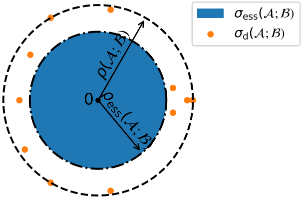
Spectral radii



- Spectral radius $\rho(\mathcal{A};\mathcal{B}) = \max|\sigma(\mathcal{A};\mathcal{B})|$
- Essential spectral radius $\rho_{\mathrm{ess}}(\mathcal{A};\mathcal{B}) = \max|\sigma_{\mathrm{ess}}(\mathcal{A};\mathcal{B})|$

Important to remember these depend on the function space $\mathcal{B}$...

Spectral radii



- Spectral radius $\rho(\mathcal{A};\mathcal{B}) = \max|\sigma(\mathcal{A};\mathcal{B})| \leq \|\mathcal{A}\|_{\mathcal{B}})$
- Essential spectral radius $\rho_{\mathrm{ess}}(\mathcal{A};\mathcal{B}) = \max|\sigma_{\mathrm{ess}}(\mathcal{A};\mathcal{B})|$

Important to remember these depend on the function space $\mathcal{B}$...

Compact operators

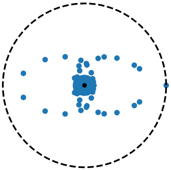
An operator $\mathcal{A} : \mathcal{B}_1 \to \mathcal{B}_2$ is *compact* if $\overline{\mathcal{A}(B_{\mathcal{B}_1}(0,1))}$ is a compact subset of $\mathcal{B}_2$.

If there exist some operators $\mathcal{A}_N : \mathcal{B}_1 \to \mathcal{B}_2$ such that

$$\|\mathcal{A}_N - \mathcal{A}\|_{\mathcal{B}_1 \to \mathcal{B}_2} \xrightarrow{N \to \infty} 0,$$

then $\mathcal{A}$ is compact. In most reasonable cases (e.g. Hilbert spaces, $\mathcal{B}_1 = \mathcal{B}_2$ has a countable Schauder basis. . .) this is iff.

- Compact operators' only essential spectrum is at $\lambda = 0$. So $\rho_{\text{ess}} = 0$.
- However, there can be countably discrete eigenvalues, which then accumulate at zero.



Stochastic systems

Remember that for most nice stochastic systems (e.g. SDE maps), the Koopman operator is a kernel operator:

$$(\mathcal{K}\psi)(y) = \mathbb{E}[\psi(x_{t+1})|x_t = y] = \int \psi(x)\, k(y,x)\, \mathrm{d}x.$$

Usually k is reasonably regular: for example,
$\int_D |\nabla_y k(y,x)|\, \mathrm{d}x \leq C$ for all x.
In this case, we find that for all $y \in D$,

$$|\nabla_y (\mathcal{K}\psi)(y)| = \int_D |\nabla_y k(y,x)||\psi(x)|\, \mathrm{d}x \leq C \sup_{x\in D} |\psi(x)|.$$

This is a nice bound, which we can translate into functional analysis as follows. . .

An estimate that is always true

Let's take $\mathcal{B}$ to be the space of bounded functions on D with the norm:

$$\|\psi\|_{\mathcal{B}} = \sup_{x \in D} |\psi(x)|.$$

Then, $\mathcal{K}$ always maps bounded functions to bounded functions by virtue of the following:

$$\|\mathcal{K}\psi\|_{\mathcal{B}} = \sup_{x \in D} |\mathbb{E}[\psi(x_{t+1})|x_t = x]| \leq \sup_{y \in D} |\psi(y)| = \|\psi\|_{\mathcal{B}}$$

And you can see it has norm (so spectral radius) bounded by 1!

Stochastic systems

In our stochastic system, let's also define a "strong" space C^1, of all the continuously differentiable functions on $[-1, 1]$, with the norm

$$\|\psi\|_{C^1} = \sup_{x \in D} |\nabla \psi(x)| + \sup_{x \in D} |\psi(x)| = \|\nabla \psi\|_{\mathcal{B}} + \|\psi\|_{\mathcal{B}}.$$

Then, our bound from before translates to saying

$$\|\mathcal{K}\psi\|_{C^1} \leq C\|\psi\|_{\mathcal{B}}.$$

So $\mathcal{K}$ makes our functions smoother!

Compact embedding

Can we use this to say anything about the compactness of $\mathcal{K}$ in $\mathcal{B}$?

Proposition

The product of a bounded operator and a compact operator (resp. approximable by finite rank) is compact (resp. approximable by finite rank).

Imagine $\mathcal{K} : \mathcal{B} \to \mathcal{B}$ as the following chain:

$$\mathcal{B} \xrightarrow{\mathcal{K}} C^1 \xrightarrow{\text{id}} \mathcal{B}$$

If we can show that $\text{id} : C^1 \to \mathcal{B}$ is compact (aka C^1 embeds compactly into $\mathcal{B}$, which we notate $C^1 \Subset \mathcal{B}$)...

then $\mathcal{K} : \mathcal{B} \to \mathcal{B}$ is compact.

Compact embedding

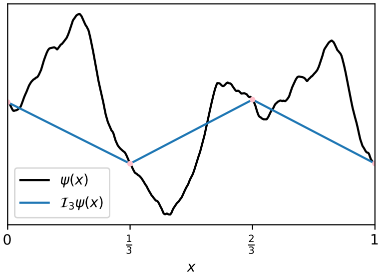
Let's try and construct some finite-dimensional operators that approach $\mathrm{id} : C^1 \to \mathcal{B}$ in norm.

- That is, let's find some finite-dimensional operators that give uniformly good approximations of differentiable functions.

For simplicity, we'll do it on the interval $[0, 1]$.

Compact embedding

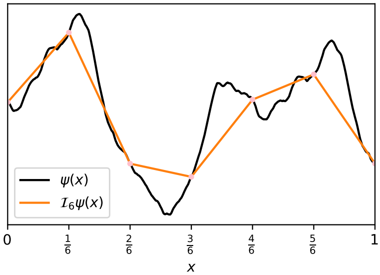
For every $\psi \in C^1$, let's define $\mathcal{P}_N \psi$ to linearly interpolate ψ at $S_N = \{0, \frac{1}{N}, \frac{2}{N}, \ldots, 1\}$:



(Exercise: show $\mathcal{P}_N$ is a linear operator.)

Compact embedding

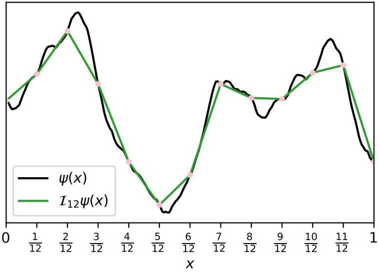
For every $\psi \in C^1$, let's define $\mathcal{P}_N\psi$ to linearly interpolate ψ at $S_N = \{0, \frac{1}{N}, \frac{2}{N}, \ldots, 1\}$:



(Exercise: show $\mathcal{P}_N$ is a linear operator.)

Compact embedding

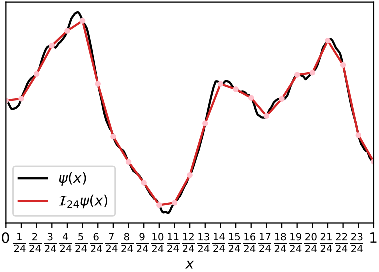
For every $\psi \in C^1$, let's define $\mathcal{P}_N \psi$ to linearly interpolate ψ at $S_N = \{0, \frac{1}{N}, \frac{2}{N}, \ldots, 1\}$:



(Exercise: show $\mathcal{P}_N$ is a linear operator.)

Compact embedding

For every $\psi \in C^1$, let's define $\mathcal{P}_N\psi$ to linearly interpolate ψ at $S_N = \{0, \frac{1}{N}, \frac{2}{N}, \ldots, 1\}$:



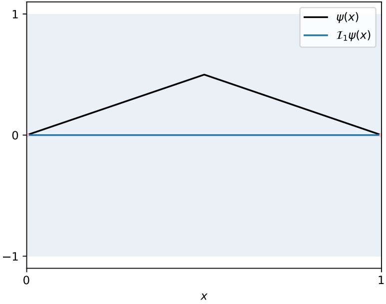
(Exercise: show $\mathcal{P}_N$ is a linear operator.)

Compact embedding

$\mathcal{P}_N$ is finite rank, and for all $\psi \in C^1$,

$$\|\mathcal{P}_N\psi - \psi\|_{\mathcal{B}} \leq \frac{1}{N}\|\psi\|_{C^1}$$

(Exercise: prove this)

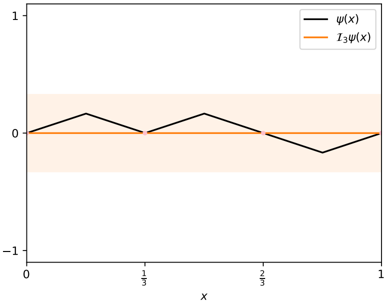


Compact embedding

$\mathcal{P}_N$ is finite rank, and for all $\psi \in C^1$,

$$\|\mathcal{P}_N \psi - \psi\|_{\mathcal{B}} \leq \frac{1}{N} \|\psi\|_{C^1}$$

(Exercise: prove this)

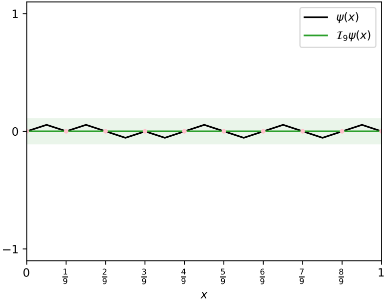


Compact embedding

$\mathcal{P}_N$ is finite rank, and for all $\psi \in C^1$,

$$\|(\mathcal{P}_N - \mathrm{id})\psi\|_{\mathcal{B}} \leq \frac{1}{N}\|\psi\|_{C^1}$$

(Exercise: prove this)



Compact embedding

So $C^1 \Subset \mathcal{B}$.

- The Koopman operator $\mathcal{K} : \mathcal{B} \to \mathcal{B}$ is compact!
- So it only has point spectrum!

Thus, we have proven that all stochastic systems on compact manifolds with differentiable kernels have compact Koopman operators!

Computing with compact operators

- In proving compactness, we came up with a nice approximation scheme (interpolation).
- We could try and approximate our Koopman operator $\mathcal{K}$ by $\mathcal{K}_N := \mathcal{P}_N \mathcal{K}$, perhaps restricting to $\operatorname{im} \mathcal{P}_N$, i.e. piecewise linear functions.
- This approximation $\mathcal{K}_N$ is $\mathcal{O}(1/nN$-close in norm to $\mathcal{K}$, so its simple eigenvalues should be $\mathcal{O}(1/N)$ error...

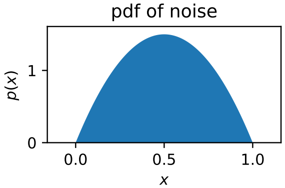
Theorem

Suppose that $\lambda \in \sigma_d(\mathcal{A}; \mathcal{B})$ with algebraic multiplicity L.
Suppose $\|\mathcal{A}_N - \mathcal{A}\|_{\mathcal{B}} \to 0$.
Then each $\mathcal{A}_N$ has L eigenvalues (counting multiplicity) $\lambda_N^1, \dots \lambda_N^L$ such that for large enough N, each

$$|\lambda_N^{(l)} - \lambda_N| \leq C \|\mathcal{A}_N - \mathcal{A}\|_{\mathcal{B}}^{1/L}.$$

Computational example

Let's set $x_{t+1} = 3.54x_t(1 - x_t) + 0.08\Xi$, where Ξ is i.i.d. noise with the following pdf $p(\xi) = \mathbb{1}_{[0,1]}(\xi)6\xi(1 - \xi)$:



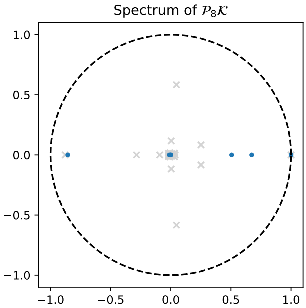
Then the kernel defining the Koopman operator is

$$k(x, y) = 0.08^{-1} p\left(\frac{y - 3.54x(1 - x)}{0.08}\right)$$

and we can try and compute $\mathcal{K}$ on $\operatorname{im} \mathcal{I}_N$.

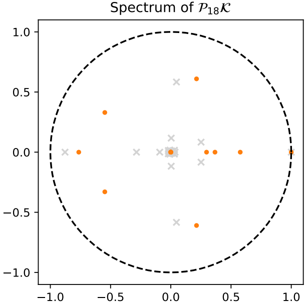
Computational example

Spectrum converges (as $\mathcal{O}(1/N)$, eventually).



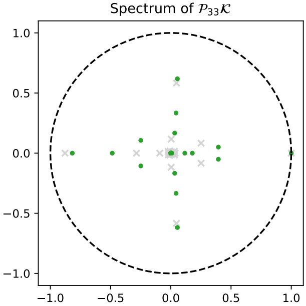
Computational example

Spectrum converges (as $\mathcal{O}(1/N)$, eventually).



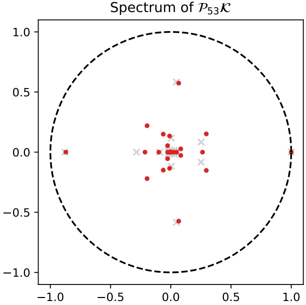
Computational example

Spectrum converges (as $\mathcal{O}(1/N)$, eventually).



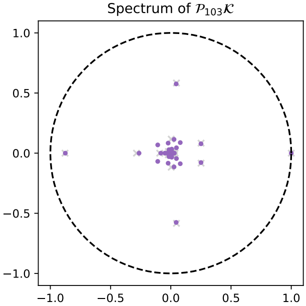
Computational example

Spectrum converges (as $\mathcal{O}(1/N)$, eventually).



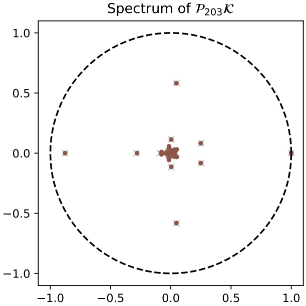
Computational example

Spectrum converges (as $\mathcal{O}(1/N)$, eventually).



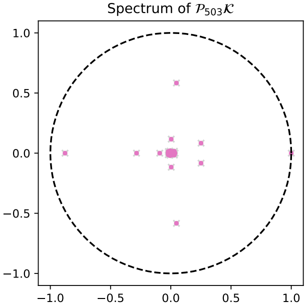
Computational example

Spectrum converges (as $\mathcal{O}(1/N)$, eventually).



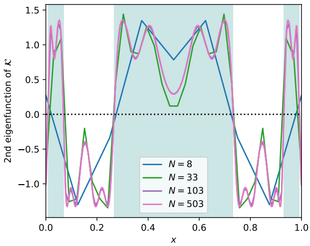
Computational example

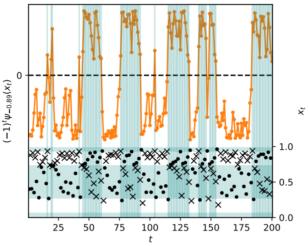
Spectrum converges (as $\mathcal{O}(1/N)$, eventually).



Computational example

Let's look at the Koopman eigenfunction for $\lambda = -0.878$ (so some set for which the period-two map is almost-invariant):





Compactness and Koopmanism

Saddish news: for most deterministic systems, the Koopman operator isn't expected to be compact on any reasonable Banach spaces.
We will see why later.

Compactness and Koopmanism

Saddish news: for most deterministic systems, the Koopman operator isn't expected to be compact on any reasonable Banach spaces.
We will see why later.

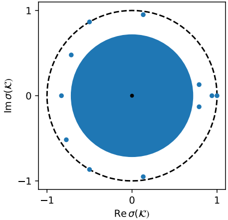
Next best option (should be possible 97%[1] of the time):

A **quasi-compact operator**

[1] 100% minus the probability of having a circle map, essentially.

Quasi-compactness

An operator is quasi-compact if it has this spectral picture:



Suppose we are only thinking about positive operators with spectral radius $= 1$ (e.g. Koopman/transfer). Then

- A quasi-compact operator has $\rho_{\text{ess}}(\mathcal{A}) < 1$.
- A quasi-compact Koopman operator has some discrete spectrum.
- Quasicompact operators are the sum of a compact operator and an operator with an iterate that is a contraction.
  - Why? $\sigma_{\text{ess}}(\mathcal{A} + \mathcal{C}) = \sigma_{\text{ess}}(\mathcal{A})$ when $\mathcal{C}$ is compact.

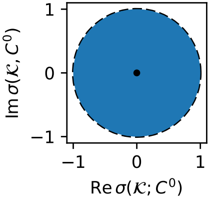
Contraction on C^0

Let's go back to $f(x) = \kappa x, x \in [-1, 1]$.
Let's take $\mathcal{B} = C^0$, the space of bounded, continuous functions on $[-1, 1]$ with the sup-norm. Then $\mathcal{K} : \mathcal{B} \to \mathcal{B}$ since f is continuous, and

$$\|\mathcal{K}\psi\| = \sup_{x\in[-1,1]} |\psi(f(x))| \leq \sup_{x\in[-1,1]} \psi(x) = \|\psi\|$$

so $\rho(\mathcal{K}; C^0) \leq 1$.

Then, eigenfunctions $\psi_\alpha(x) := \mathbb{1}(x > 0)e^{\alpha \log x}$ are in C^0 for $\Re\alpha \leq 0$.
Corresponding eigenvalues are κ^α, so $\sigma(\mathcal{A}; C^0([-1, 1]))$ fills the whole (closed) unit ball.



Just cts. spectrum!

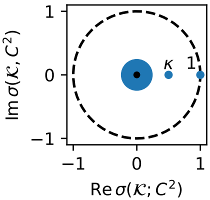
Contraction on C^r spaces

What about some spaces that remove more of the ψ_α? Let's try C^r, the space of r-times continuously differentiable functions on $[-1,1]$. The following norm on C^r works:

$$\|\psi\|_{C^r} = \|\psi^{(r)}\|_{C^0} + \psi(0) + \psi'(0) + \ldots + \psi^{(r-1)}(0).$$

We have that $\psi_\alpha^{(r)}(x) = \alpha(\alpha-1)\cdots(\alpha-r+1)\psi_{\alpha-r}(x)$, so ψ_α is in C^r if either:

- $\psi_{\alpha-r}$ is in C^0, i.e. $\Re\alpha > r$. So $B(0,\kappa^r)$ is in the spectrum.
- α is one of $0,1,2,\ldots,r-1$, i.e. $\psi_\alpha = 1, x, x^2, \ldots, x^{r-1}$. So $1, \kappa, \kappa^2, \ldots, \kappa^{r-1}$ are in the spectrum.



Contraction on C^r spaces

Is there anything else?
Well, let's try and do an eigendecomposition. Recalling that every function in C^r can be written as

$$\psi(x) = \psi(0) + \psi'(0)x + \ldots + \frac{\psi^{(r-1)}(0)}{(r-1)!}x^{r-1} + \mathcal{O}(x^r),$$

we can decompose

$$C^r = \langle 1 \rangle \oplus \langle x \rangle \oplus \cdots \oplus \langle x^{r-1} \rangle \oplus \underbrace{\left\{\psi \in C^r : \psi^{(l)}(0) = 0 \text{ for } l < r\right\}}_{=:V}.$$

All these subspaces are $\mathcal{K}$-invariant, and $\sigma(\mathcal{K})$ is the union of the spectrum of $\mathcal{K}$ restricted to these subspaces.
Only what happens on V we are uncertain of.

Contraction on C^r spaces

$$\|\psi\|_{C^r} = \|\psi^{(r)}\|_{C^0} + \psi(0) + \psi'(0) + \ldots + \psi^{(r-1)}(0).$$

$$\left\{\psi \in C^r : \psi^{(l)}(0) = 0 \text{ for } l < r\right\}$$

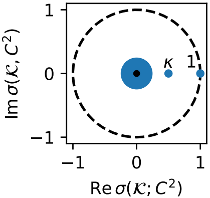
For $\psi \in V$ we have

$$\begin{aligned} \|\mathcal{K}\psi\|_{C^r} = \|(\mathcal{K}\psi)^{(r)}\|_{C^0} &= \sup_{x\in[-1,1]} \|\kappa^r \psi^{(r)}(\kappa x)\|_{C^0} \\ &= \kappa^r \sup_{y\in[-\kappa,\kappa]} |\psi^{(r)}(y)| \leq \kappa^r \|\psi\|_{C^r} \end{aligned}$$

so $\sigma(\mathcal{K}|_V)$ is a subset of $B(0,\kappa^r)$.

This means the spectrum of $\mathcal{K}$ on C^r is

$$\sigma(\mathcal{K}, C^r) = \underbrace{\overline{B(0,\kappa^r)}}_{\text{essential}} \cup \underbrace{\{\kappa^{r-1}, \kappa^{r-2}, \ldots \kappa, 1\}}_{\text{discrete}}.$$

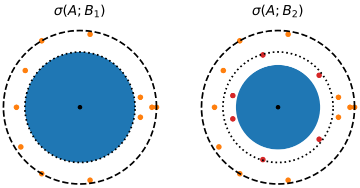


Spectrum vs function space

In general, essential spectrum will vary by function space, but the discrete eigenvalues are more canonical:

Lemma (simplified from Baladi and Tsujii, '08)

Suppose Banach space $\mathcal{B}_2$ is a dense *subset of Banach space $\mathcal{B}_1$, and $\mathcal{A}$ is bounded on both $\mathcal{B}_2$ and $\mathcal{B}_1$.*
Then, the discrete spectrum of $\mathcal{A}$ with absolute value greater than $\max\{\rho_{ess}(\mathcal{A};\mathcal{B}_1), \rho_{ess}(\mathcal{A};\mathcal{B}_2)\}$ *matches in $\mathcal{B}_1$ and $\mathcal{B}_2$ (ditto multiplicity, eigenfunctions).*



Sidenote: spaces of fractional differentiability

We can continuously interpolate between C^r spaces by looking, e.g., at Hölder continuity. The β-Hölder constant of a function is given by

$$H_\beta(\psi) = \sup_{x,y\in[-1,1]} \frac{|\psi(x) - \psi(y)|}{|x-y|^\beta}, \beta \in (-1, 1]$$

Then the $C^{r+\beta}$ norm of ψ is given by

$$\|\psi\|_{C^{r+\beta}} = \|\psi\|_{C^r} + H_\beta(\psi^{(r)}).$$

i.e. $C^{r+\beta}$ consists of functions whose the rth derivative is β-Hölder.

The essential spectral radius is $\rho_{\mathrm{ess}}(\mathcal{K}, C^{r+\beta}) = \kappa^{r+\beta}$, with discrete eigenvalues $\{1, \kappa, \ldots, \kappa^r\}$.