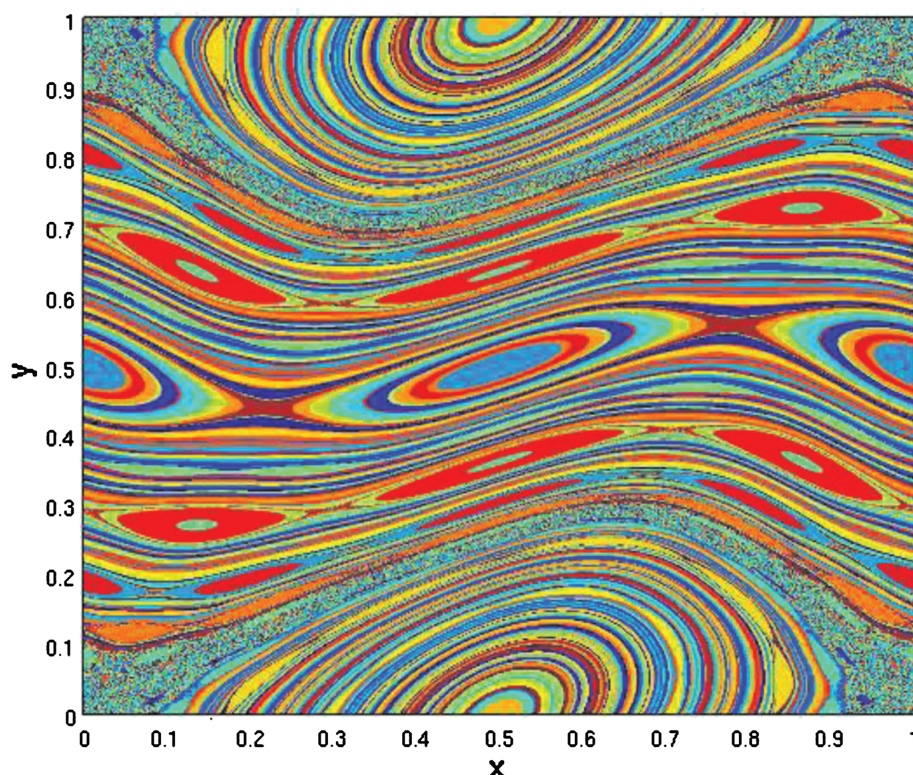# Koopman Operator, Geometry, and Learning of Dynamical Systems

Igor Mezić

Introduction

The end of the 20th century and the beginning of the 21st century has seen a revolutionary increase in the availability of data. Indeed, we are in the middle of the *sensing* revolution, where sensing is used in the broadest meaning of data acquisition. Most of this data goes unprocessed, unanalyzed, and, consequently, unused. This causes missed opportunities, in domains of vast societal importance—

Igor Mezić is a professor of mechanical engineering and of mathematics at the University of California, Santa Barbara. His email address is mezic@ucsb.edu.

health, commerce, technology, network security, just to mention some.

A variety of mathematical methods have emerged out of that need. Perhaps the most popular, the methodology of Deep Neural Networks has as the underlying learning elements "neuron functions" that are modeled after biological neurons. The algorithms based on deep learning have achieved substantial success in image recognition, speech recognition, and natural language processing, deploying the "supervised" machine learning philosophy. Convolutional neural networks provided a superstructure to the deep neural network architecture that resembles the organization of the animal visual cortex. This led to an enormous success in image recognition, and even in realistic image generation, via Generative Adversarial Networks (GANs). All of these examples are essentially static pattern

recognition or generation tasks. Deep learning methodologies are less successful in dynamically changing contexts, present, for example, in autonomous driving. This is because the learning architectures are not adapted to physical properties of the time variable. In contrast, the symmetry associated with translation in time naturally occurs in the Koopman operator framework and imbues it with the fundamental group structure. It is interesting that, in contrast to biological modules responsible for vision, even the basic issue of finding specific brain structures that are responsible for perception of time, and thus understanding of dynamics, is still being investigated.

Koopman operator theory has recently emerged as one of the main candidates for machine learning of dynamical processes. In this paper, we briefly describe its history, emerging from efforts to extend the methodology used in quantum mechanics, and describe the current focus, setting it within the new concept of dynamic process *representation*, and connecting along the way to the geometric dynamical systems theory methods that enable data-driven discovery of essential elements of the theory, for example stable and unstable manifolds. What emerges is a powerful framework for unsupervised learning from small amounts of data, enabling self-supervised learning that is much more in line with the theory of human learning than the machine learning methods of the second wave.

History

Driven by the success of the operator-based framework in quantum theory, Bernard Koopman proposed in his 1931 paper [Koo31] to treat classical mechanics in a similar way, using the spectral properties of the composition operator associated with dynamical system evolution. The work, restricted to Hamiltonian dynamical systems, did not attract much attention originally, as evidenced by the fact that between 1931 and 1990, the Koopman paper [Koo31] was cited 100 times, according to Google Scholar. This can be attributed largely to the major success of the geometric picture of dynamical systems theory in its state-space realization advocated by Poincaré. In fact, with Lorenz's discovery of a strange attractor in 1963, the dynamical systems community turned to studying dissipative systems and much progress has been made since. Within the current research in dynamical systems, some of the crucial roadblocks are associated with high-dimensionality of the problems and necessity of understanding behavior globally (away from the attractors) in the state space. However, the weaknesses of the geometric approach are related exactly to its locality—as it often relies on perturbative expansions around a known geometrical object—and low-dimensionality, as it is hard to make progress in higher dimensional systems using geometry tools.

Out of today's 1200+ citations of Koopman's original work, [Koo31], about 80% come from the last 20 years. It was only in the 1990s and 2000s that potential for wider applications of the Koopman operator-theoretic approach has been realized [Mez94, Mez05, RMB+09]. In the past decade the trend of applications of this approach has continued. This is partially due to the fact that strong connections have been made between the spectral properties of the Koopman operator for dissipative systems and the geometry of the state space. In fact, the hallmark of the work on the operator-theoretic approach in the last two decades is the linkage between geometrical properties of dynamical systems—whose study has been advocated and strongly developed by Poincaré and followers—with the geometrical properties of the level sets of Koopman eigenfunctions [Mez94, MM12, MMM13]. The operator-theoretic approach has been shown capable of detecting objects of key importance in geometric study, such as invariant sets, but doing so globally, as opposed to locally as in the geometric approach. It also provides an opportunity for study of high-dimensional evolution equations in terms of dynamical systems concepts [Mez05, RMB+09] via a spectral decomposition, and links with associated numerical methods for such evolution equations [Sch10, RMB+09].

Even the early work in [Mez94] and its continuation in [MB04, Mez05, RMB+09] already led to the realization that spectral properties, and thus geometrical properties, can be *learned* from data, thus initiating a strong connection that is forming today between machine learning and dynamical systems communities [LDBK17, YKH19, LKB18, TKY17]. The key notion driving these developments is that of representation of a—possibly nonlinear—dynamical system as a linear operator on a typically infinite-dimensional space of functions. This then leads to search for linear, finite-dimensional invariant subspaces. In this paper I formalize the concept of *dynamical system representation* enabling the study of finite-dimensional linear and nonlinear representations, learning, and the geometry of state-space partitions.

Dynamical System Representations

State space vs. observables space. It is customary, since Poincaré, to start the discussion of mathematics of dynamical systems with the notion of the state space, which already includes a *numerical representation* of' the state of a system. However, to set the operator-theoretic approach properly, it is useful to start with just the primitive notion of a set M of (nonnumerically described) states of a given system. Elements $\mathbf{m} \in M$ are abstract to start with, and the dynamics is given by a rule that assigns $\mathbf{m}^t \in M$ to $\mathbf{m}$ for any element t of the time set $\mathcal{T}$. The time set can be $\mathbb{R}$, $\mathbb{Z}$, but more complicated cases such as $\mathcal{T} = \mathbb{Z}^2$ can

be considered as well.[1] As we are interested in framing the process of learning and modeling dynamics from data in the Koopman (composition) operator framework, we begin by describing the basic notions of representation of dynamics using functions.

Discrete dynamical systems. The set $\mathcal{O}$ of all complex functions $f : M \to \mathbb{C}$ is called the space of observables. It is a linear vector space over the field of complex numbers. A discrete deterministic dynamical system T on M is a map $M \to M$, and the time set is $\mathcal{T} = \mathbb{Z}$. For $\mathbf{m} \in M$ the iteration of the map $\mathbf{m}$ is defined by $\mathbf{m}' = T\mathbf{m}$. Any such map defines an operator $U : \mathcal{O} \to \mathcal{O}$ by

$$Uf(\mathbf{m}) = f \circ T(\mathbf{m}) = f(T\mathbf{m}). \tag{1}$$

The operator U is linear, as composition distributes over addition:

$$\begin{aligned} U(c_1 f_1 + c_2 f_2)(\mathbf{m}) &= (c_1 f_1 + c_2 f_2)(T\mathbf{m}) \\ &= c_1 f_1(T\mathbf{m}) + c_2 f_2(T\mathbf{m}) \\ &= c_1 U f_1(\mathbf{m}) + c_2 U f_2(\mathbf{m}). \end{aligned} \tag{2}$$

Finite-dimensional representations. Ultimately, data is about numbers. We can understand a lot about the map T on M by collecting data on observables. To formalize this, we need the notion of *representation*.

Definition 1. A finite-dimensional representation $(\mathbf{f}, \mathbf{F})$ of T in $\mathcal{O}$ is a set of functions $\mathbf{f} = (f_1, ..., f_n)$ and a mapping $\mathbf{F}$ such that

$$U\mathbf{f}(\mathbf{m}) = \mathbf{f}(T\mathbf{m}) = \mathbf{F}(\mathbf{f}(\mathbf{m})), \tag{3}$$

where $\mathbf{F} : \mathbb{C}^n \to \mathbb{C}^n$ and n is the dimension of the representation. If $\mathbf{f} : M \to \mathbb{R}^n$ is a real set of functions, then the representation is real.

The image of M in $\mathbb{C}^n$ under $\mathbf{f}$—the space $\mathbf{f}(M)$—is called the *state space*.[2] The simplest examples of state spaces are Euclidean spaces of n-tuples of real numbers $\mathbb{R}^n$. Consider $\mathbf{f} = (f_1, ..., f_n)$ such that $f_j(\mathbf{m}) = m_j, \mathbf{m} = (m_1, ..., m_n) \in \mathbb{R}^n$. Any mapping $\mathbf{m}' = T\mathbf{m}$ on $\mathbb{R}^n$ has a real representation $\mathbf{f}' = \mathbf{F}(\mathbf{f})$, where $\mathbf{F}_j(\mathbf{f}) = f_j(T\mathbf{m})$. The representation (3) is called *linear* provided $\mathbf{F} : \mathbb{C}^n \to \mathbb{C}^n$ is a linear mapping. Finite-dimensional representations are key to learning dynamical systems from data.

Example 2. Let $M = \mathbb{T}^2$, the two-dimensional torus, and let $T : \mathbb{T}^2 \to \mathbb{T}^2$ be the mapping that translates points on the torus by angle ω_1 in the direction of rotation around the symmetry axis, and ω_2 in the direction of the cross-sectional circle.[3] Consider the representation

$$\begin{aligned} \mathbf{f} &= (z_1, z_2)^T, \\ \mathbf{F}(z_1, z_2) &= (e^{i\omega_1} z_1, e^{i\omega_2} z_2), \end{aligned} \tag{4}$$

where $z_1 = e^{i\theta_1}$ and $z_2 = e^{i\theta_2}$, θ_1 being the angle along the rotational symmetry and θ_2 the angle along the cross-sectional circle. We have

$$\mathbf{f}' = A\mathbf{f}, \tag{5}$$

where A is a diagonal matrix with $(e^{i\omega_1}, e^{i\omega_2})$ as diagonal elements. Thus, $(\mathbf{f}, \mathbf{F})$ is a complex, linear representation of T. Note that $\theta_1' = \theta_1 + \omega_1, \theta_2' = \theta_1 + \omega_1$ is **not** a linear representation.

Mathematically, one of the key questions in this context is whether a finite-dimensional representation exists. Namely, a set of functions $\mathbf{f} : M \to \mathbb{C}$ does not necessarily satisfy $\mathbf{f} \circ T = \mathbf{F}(\mathbf{f})$ for any $\mathbf{F}$. If one considers trajectories $\{\mathbf{m}_j, j \in \mathbb{Z}\}$ of T, then it is easy to see that there isn't necessarily an $\mathbf{F} : \mathbb{C}^{\{J, -\infty\}} \to \mathbb{C}, \{J, -\infty\} = \{J, J-1, ..., J-N, ...\}$, such that

$$\mathbf{f}(\mathbf{m}_{J+1}) = \mathbf{f} \circ T(\mathbf{m}_J) = \mathbf{F}(\mathbf{f}(\{\mathbf{m}_j, j \le J\})), \tag{6}$$

i.e., the next value of $\mathbf{f}$ cannot always be obtained uniquely even if we know the whole history of the evolution of $\mathbf{f}$ on the trajectory. The representation relationship $\mathbf{f} \circ T = \mathbf{F}(\mathbf{f})$ requires that the next value of $\mathbf{f}$ is uniquely determined by the current value. This is the Markov property. If a representation does not satisfy the Markov property, but its dynamics depends only on a finite number of previous trajectory points, i.e.,

$$\begin{aligned} \mathbf{f}(\mathbf{m}_{J+1}) &= \mathbf{f} \circ T(\mathbf{m}_J) \\ &= \mathbf{F}(\mathbf{f}(\{\mathbf{m}_j, J - N \le j \le J\})) \\ &= \mathbf{F}(\mathbf{f}(\mathbf{m}_J), \mathbf{f}(\mathbf{m}_{J-1}), ..., \mathbf{f}(\mathbf{m}_{J-N})) \\ &= \mathbf{F}(\mathbf{f}(T^N \mathbf{m}_{J-N}), \mathbf{f}(T^{N-1} \mathbf{m}_{J-N}), \\ &\quad ..., \mathbf{f}(\mathbf{m}_{J-N})), \end{aligned} \tag{7}$$

then the so-called time-delay embedding can be used to make it Markovian: Let

$$\tilde{\mathbf{f}}(\mathbf{m}) = (\mathbf{f}, \mathbf{f} \circ T, ..., \mathbf{f} \circ T^N)(\mathbf{m}) = (\mathbf{f}, \mathbf{f}_1, ..., \mathbf{f}_N)(\mathbf{m}). \tag{8}$$

Then, setting $\mathbf{m} = \mathbf{m}_{J-N}$

$$\begin{aligned} \tilde{\mathbf{f}}' &= (\mathbf{f} \circ T(\mathbf{m}), \mathbf{f} \circ T^2(\mathbf{m}), ..., \mathbf{f} \circ T^{N+1}(\mathbf{m})) \\ &= (\mathbf{f}_1(\mathbf{m}), \mathbf{f}_2(\mathbf{m}), ..., \mathbf{f}_N(\mathbf{m}), \mathbf{f} \circ T^{N+1}(\mathbf{m})) \\ &= (\mathbf{f}_1(\mathbf{m}), \mathbf{f}_2(\mathbf{m}), ..., \mathbf{f}_N(\mathbf{m}), \mathbf{f}(\mathbf{m}_{N+1})) \\ &= (\mathbf{f}_1(\mathbf{m}), \mathbf{f}_2(\mathbf{m}), ..., \mathbf{f}_N(\mathbf{m}), \\ &\quad \mathbf{F}(\mathbf{f}(T^N \mathbf{m}), \mathbf{f}(T^{N-1} \mathbf{m}), ..., \mathbf{f}(\mathbf{m}))) \\ &= \tilde{\mathbf{F}}(\tilde{\mathbf{f}}). \end{aligned} \tag{9}$$

---

[1] *For example, we could be given two transformations T, S on M and the dynamics starting from an initial point $\mathbf{m}$ can be given by sequences of transformations $S^{n_2} T^{n_1}$, where $(n_1, n_2) \in \mathbb{Z}^2$. This case, that is of interest in control theory, will not be expanded on further here.*

[2] *We might be used to thinking about M itself as the "state space." But the original notion in Poincaré's work refers to $\mathbf{f}(M)$, where $\mathbf{f}$ is comprised of observables that are positions and momenta of a mechanical system. Perhaps M can be called the "abstract state space."*

[3] *I intentionally do not start with the notation $\theta_1' = \theta_1 + \omega_1, \theta_2' = \theta_1 + \omega_1$ to emphasize that we can start with abstract points on the torus and physically describe the transformation on it.*

Physically, there is the additional problem of whether experimental observations can provide all the information necessary to describe the finite-dimensional representation. Historically, the analysis of the problem of finite representations has often been considered in the context of the Mori-Zwanzig theory that has close links to Koopman operator theory. We discuss this connection later in the paper.

Representations and conjugacies. There are representations that are capable of separating points on M. We call these *faithful*.

Definition 3. A representation $(\mathbf{f}, \mathbf{F})$ is called faithful provided $\mathbf{f} : M \to \mathbf{f}(M)$ is injective:

    (i) $\mathbf{f}(\mathbf{m}) = \mathbf{f}(\mathbf{n}) \Rightarrow \mathbf{m} = \mathbf{n}$, or equivalently

    (ii) $\mathbf{m} \neq \mathbf{n} \Rightarrow \mathbf{f}(\mathbf{m}) \neq \mathbf{f}(\mathbf{n}) \,\forall \mathbf{m}, \mathbf{n} \in M$.

In terms of representations, the Takens embedding theorem shows that a faithful representation of dynamics on an m-dimensional Riemannian manifold can be obtained by using a sufficiently large set of time-delayed observables for generic pairs of smooth real functions and dynamical systems (f, T).

Theorem 4 (Takens). *Let M be a compact Riemannian manifold of dimension m, $T : M \to M$ a $C^r, r \geq 1$, diffeomorphism, and $f : M \to \mathbb{R}$ a C^r function. For generic (f, T) the map $\tilde{\mathbf{f}} : M \to \mathbb{R}^{2m+1}$ given componentwise by*

$$\tilde{\mathbf{f}}(x) = (f, f \circ T, f \circ T^2, ..., f \circ T^{2m})$$

is an embedding and $\tilde{\mathbf{f}}(M)$ is a compact submanifold of $\mathbb{R}^{2m+1}$. Thus, for generic (f, T), $(\tilde{\mathbf{f}}, \mathbf{F})$ is a faithful real representation of T.

Time-delayed observables have been used in approximations of Koopman operators since [MB04].

A representation might provide redundant information: for example, it might contain two functions f_j and f_k such that $f_j = F(f_k)$ for some function F. If it does not, it is called efficient.

Definition 5. An n-dimensional faithful representation is called efficient provided there are no $H : \mathbb{C}^{n-1} \to \mathbb{C}$ and $j \in \{1, ..., n\}$ such that

$$f_j = H(f_1, ..., f_{j-1}, f_{j+1}, ..., f_n). \tag{10}$$

An n-dimensional efficient faithful representation is called minimal provided any other m-dimensional efficient faithful representation satisfies $m \geq n$.

Example 6. In Example 2 $(\mathbf{f}, \mathbf{F})$ is a minimal, faithful, efficient representation of T.

It is clear that all minimal efficient faithful representations have the same dimension. Thus, the dimension of the system can be defined as the dimension of its minimal efficient representation.[4] Additionally, different faithful representations of the underlying mapping T play nicely with each other as they are related by a conjugacy.

Proposition 7. *Let $(\mathbf{f}, \mathbf{F}), (\mathbf{g}, \mathbf{G})$ be two different faithful n-dimensional representations. Then there is a bijection $\mathbf{h} : \mathbf{g}(M) \to \mathbf{f}(M)$ such that*

$$\mathbf{f} = \mathbf{h}(\mathbf{g}), \tag{11}$$

and $\mathbf{h}$ is a conjugacy of representations, i.e.,

$$\mathbf{h}(\mathbf{G}(\mathbf{g})) = \mathbf{F}(\mathbf{h}(\mathbf{g})). \tag{12}$$

Proof. Since $\mathbf{f}$ and $\mathbf{g}$ are faithful, $\mathbf{f}(\mathbf{m})$ and $\mathbf{g}(\mathbf{m})$ are unique for every $\mathbf{m} \in M$, and thus for any $\mathbf{g} \in \mathbf{g}(M)$ there is a unique $\mathbf{f} \in \mathbf{f}(M)$. The resulting mapping $\mathbf{h} : \mathbf{g}(M) \to \mathbf{f}(M)$ is a bijection. Further, we know

$$\mathbf{h}(\mathbf{g} \circ T) = \mathbf{h}(\mathbf{G}(\mathbf{g})). \tag{13}$$

Now,

$$\begin{aligned} \mathbf{h}(\mathbf{g} \circ T) &= \mathbf{f} \circ T \\ &= \mathbf{F}(\mathbf{f}) \\ &= \mathbf{F}(\mathbf{h}(\mathbf{g})), \end{aligned} \tag{14}$$

and thus

$$\mathbf{h}(\mathbf{G}(\mathbf{g})) = \mathbf{F}(\mathbf{h}(\mathbf{g})). \tag{15}$$

                                                           $\square$

The concept of conjugacy has classically been used in dynamical systems for local linearization theorems. Since the Koopman operator description is global, extensions of the local theory are needed (see [Mez20] and the preceding Lan and Mezić (2013) reference therein).

Faithful representations are capable of describing all of the dynamics of T. However, that dynamics is often high dimensional and has components that are irrelevant for understanding the problem at hand. In this context, the notion of the *reduced representation* is useful.

Definition 8. A representation $(\mathbf{f}, \mathbf{F})$ is called reduced provided it is not faithful.

Note that, by the definition of the concept of representation, even for reduced representations we have

$$\mathbf{f}(\mathbf{m}) = \mathbf{f}(\mathbf{n}) \Rightarrow \mathbf{f}(T\mathbf{m}) = \mathbf{f}(T\mathbf{n}) \,\forall \mathbf{m}, \mathbf{n} \in M,$$

since, if $\mathbf{f}(\mathbf{m}) = \mathbf{f}(\mathbf{n})$, then

$$\mathbf{f}(T\mathbf{m}) = \mathbf{F}(\mathbf{f}(\mathbf{m})) = \mathbf{F}(\mathbf{f}(\mathbf{n})) = \mathbf{f}(T\mathbf{n}). \tag{16}$$

We note here the hierarchy of the introduced forms of representations: a faithful representation might not be efficient. However, an efficient representation is "minimally" faithful. A reduced representation is not faithful.

---

[4]*The underlying space M can have a fractal dimension—e.g., in the case of the Lorenz attractor—but the representation is integer dimensional.*

The concept of reduced representations is exemplified in the notion of factors in ergodic theory, for which we need to equip M with a measure μ.

Definition 9. Let $T : M \to M$ be a measure-preserving dynamical system with respect to a measure μ on M. Then a map $S : N \to N$ is a factor of T provided it preserves the measure $\nu = \mu \circ h^{-1}$ on N, where $h : M \to N$ is a measurable mapping such that $h(T(\mathbf{m})) = Sh(\mathbf{m})$.

Let $(\mathbf{f}, \mathbf{F})$ be a reduced representation of $T : M \to M$, where the components of $\mathbf{f}$ are measurable functions on M. Then we have the following.

Proposition 10. *The dynamical system* $\mathbf{f}' = \mathbf{F}(\mathbf{f})$ *on* $\mathbf{f}(M)$ *equipped with the measure* ν *defined by* $\nu(A) = \mu(\mathbf{f}^{-1}(A))$ *is a factor of* T.

Proof. We have
$$\mathbf{f}(T\mathbf{m}) = \mathbf{F}(\mathbf{f}(\mathbf{m})). \tag{17}$$
Since $\mathbf{f}$ is measurable and $\nu(A) = \mu(\mathbf{f}^{-1}(A))$, $\mathbf{F}$ is a factor of T. □

In the context of factors, $\mathbf{f}$ is required to be measurable, in contrast with the notion of semiconjugacy in topological dynamics, where the representation is required to be continuous.

Proposition 11. *Let* $\mathbf{f} : M \to \mathbf{f}(M)$ *be a continuous (proper) surjection, i.e., there are two points in M that map to a single point in $\mathbf{f}(M)$, and let $(\mathbf{f}, \mathbf{F})$ be a (nonfaithful) representation. Then $\mathbf{F}$ is semiconjugate to T.*

Proof. We again have
$$\mathbf{f}(T\mathbf{m}) = \mathbf{F}(\mathbf{f}(\mathbf{m})), \tag{18}$$
and thus $\mathbf{F}$ and T are semiconjugate. □

Both of these concepts—factors and semiconjugacies—are key in model reduction of dynamical systems [Mez05, Mez20]. In the larger context of machine learning, factors and semiconjugacies—that, as shown below, can be realized using eigenfunctions—play the role of autoencoders, helping reduce the dimension and reduce "noise" in the dynamic dataset. We discuss continuous time evolutions next.

Representations of continuous time evolution. In the case of continuous time $t \in \mathbb{R}$, the evolution on M consists of a group of transformations T^t, satisfying
$$T^{t+s}\mathbf{m} = T^t T^s \mathbf{m} \, \forall t, s \in \mathbb{R}. \tag{19}$$
Any such evolution group defines an operator group $U^t : \mathcal{O} \to \mathcal{O}$ by
$$U^t f(\mathbf{m}) = f(T^t \mathbf{m}), f \in O. \tag{20}$$
A representation $(\mathbf{f}, \mathbf{F}^t)$ of T^t then consists of a set of real or complex functions $\mathbf{f}$ and a group of transformations $\mathbf{F}^t$ that satisfy
$$\mathbf{f}(T^t\mathbf{m}) = \mathbf{F}^t(\mathbf{f}(\mathbf{m})). \tag{21}$$

For fixed τ, U^τ is a linear composition operator associated with T^τ.

Definition 12. A representation $(\mathbf{f}, \mathbf{F}^t)$ is called ordinary differential if it is finite and
$$\begin{aligned} \mathbf{v}(\mathbf{m}) &= \frac{d[U^t\mathbf{f}(\mathbf{m})]}{dt} = \lim_{t \to 0} \frac{U^t\mathbf{f}(\mathbf{m}) - \mathbf{f}(\mathbf{m})}{t} \\ &= \lim_{t \to 0} \frac{\mathbf{F}^t(\mathbf{f}(\mathbf{m})) - \mathbf{f}(\mathbf{m})}{t} \end{aligned} \tag{22}$$
exists. In this case, the evolution is represented by a finite set of ordinary differential equations
$$\dot{\mathbf{f}} = \mathbf{v}(\mathbf{f}). \tag{23}$$

Example 13. Consider the set of all the states $\mathbf{m}$ of a mass-spring system, and the real representation $\mathbf{f} = (x(\mathbf{m}), p(\mathbf{m}))$, where x is a numerical function that represents the deviation of the mass position from the unstretched length of the spring and p is the linear momentum, $p = \mathsf{m}v$, where m is the mass parameter, assumed constant, and v is an observable representing change of x with time t. Derivatives with respect to time are labeled by $\dot{()}$. Then
$$\begin{aligned} \dot{x} &= p/\mathsf{m}, \\ \dot{p} &= -kx \end{aligned} \tag{24}$$
is a two-dimensional, faithful, efficient, ordinary differential representation. Setting $\omega^2 = k/m, z = x + ip$, we have a one-dimensional, faithful, efficient, minimal representation
$$\dot{z} = -i\omega z. \tag{25}$$
On the other hand, using the energy observable $E = p^2/2m + kx^2/2$, we obtain a one-dimensional, real reduced representation
$$\dot{E} = 0. \tag{26}$$

As the next example shows, simple representations can exist even for strange M.

Example 14 (Lorenz representation). Let M be the Lorenz attractor, which is a subset of $\mathbb{R}^3$. Let $f(\mathbf{m}) = (x(\mathbf{m}), y(\mathbf{m}), z(\mathbf{m}))$, where M is viewed as embedded in $\mathbb{R}^3$ and $(x(\mathbf{m}), y(\mathbf{m}), z(\mathbf{m}))$, $\mathbf{m} \in M$, are projections of $\mathbf{m}$ on the x, y, z axes. Then
$$\begin{aligned} \dot{x} &= \sigma(y - x), \\ \dot{y} &= x(\rho - z) - y, \\ \dot{z} &= xy - \beta z \end{aligned} \tag{27}$$
is a 3-dimensional efficient ordinary differential equation representation. Note that the underlying set M is fractal, and yet the dynamics on it possesses a differential representation. It is of interest to note that the ordinary differential equations (27) are valid off the set M when it is viewed as embedded in $\mathbb{R}^3$, but from the current point of view, the representation itself is only valid when restricted to M.

As far as the full dynamic process the Lorenz representation is supposed to represent, it is reduced and is in fact inexact: the dynamics it models is that of a Boussinesq approximation of thermal convection dynamics, reduced by truncating Fourier series expansion of the solution.

Remark 15. In the case where the representation of T is not finite, and thus involves a field of observables, e.g., $\mathbf{f} \in L^2(P)$ for some continuous space P (an example is $P = \mathbb{R}$), we speak of a field representation. The scalar vorticity field

$$\omega(\mathbf{x}), \mathbf{x} = (x, y) \in A \subset \mathbb{R}^2, \qquad (28)$$

of a two-dimensional, incompressible, inviscid fluid satisfies the equation

$$\dot{\omega} = \mathcal{N}\omega, \qquad (29)$$

where $\mathcal{N}$ is a nonlinear operator.

Representations for control systems. The relationship between Koopman operator theory and control theory has been explored intensely over the last decade [MMS20]. Control systems in discrete time are defined on the product space $M \times N$,

$$\begin{aligned} \mathbf{m}' &= T(\mathbf{m}, \mathbf{n}), \\ \mathbf{n}' &= W(\mathbf{n}), \end{aligned} \qquad (30)$$

where $\mathbf{m} \in M, \mathbf{n} \in N$. The system (30) is a skew-product system. The N system is the control system. Physically, the assumption is that the N system is separated from the M system, and it possesses its own internal dynamics described by W. A simple additive, linear structure for the representation is obtained if we can find $\mathbf{f} : M \to \mathbb{C}$ and $\mathbf{u} : N \to \mathbb{C}$ such that

$$\mathbf{f}' = A\mathbf{f} + B\mathbf{u}. \qquad (31)$$

Note that $\mathbf{u}$ is physically the input to the system. Feedback relationship of the type

$$\mathbf{u} = \mathbf{H}(\mathbf{f}) \qquad (32)$$

couples the two systems.

Representations and data. In general, it is not easy to find faithful representations, and their existence has to be validated experimentally. Namely, for a finite set of functions to comprise a representation, equation (16) needs to hold for every $\mathbf{m}, \mathbf{n} \in M$. It is also clear that such validation is only possible for a finite set of points $\mathbf{m}, \mathbf{n} \in M$, and thus there is always uncertainty present. It is easier to show that a representation is *not* faithful. For example, consider an object moving along a straight line, observed at time 0. It is impossible to predict where it will be at some time τ later if we do not know its momentum p. Thus, the representation that includes only the observable x (we could call it Galileo's) that measures the position along the straight line is not faithful. In contrast, the representation (x, p) that contains both the position and momentum observables is faithful (as long as the object does not interact

with any others and thus proceeds moving with constant momentum). This is Newton's representation, and it leads to ordinary differential equations $\dot{x} = p/m, \dot{p} = 0$.

Another aspect of representation that is important is its accuracy. Namely, a set of functions $\mathbf{f} : M \to \mathbb{R}^m$ might be such that $|\mathbf{f}(T\mathbf{m}) - \mathbf{F}(\mathbf{m})| \leq \epsilon$ for some ϵ that is small with respect to the average value of $\mathbf{f}$. This was the case, for example, with Newton's representation of the motion under the law of gravity, where the orbit of planet Mercury, the closest to the sun and thus experiencing the largest force of all planets, has a small but measurable deviation from the motion predicted by the inverse-square law. This was rectified by Einstein's geodesic representation. But there are measurable deviations of motion of galaxies from Einstein's representation, too. Like faithfulness, accuracy of representation is also measurable only up to experimental uncertainty.

Representations and geometry of state-space partitions. The key connection between the notion of representations and geometrical objects associated with a dynamical system—for example invariant sets—is that of a partition. The collection of disjoint sets $A_\alpha, \alpha \in \mathbb{C}$, forms a partition ζ of M if and only if

$$\bigcup_{\alpha \in \mathbb{C}} A_\alpha = M. \qquad (33)$$

The partition ζ_f induced by an observable f is defined by

$$\zeta_f = \{B_\alpha^f \subseteq M | B_\alpha^f = \{\mathbf{m} \in M | f(\mathbf{m}) = \alpha, \alpha \in \mathbb{C}\}\}, \qquad (34)$$

i.e., the sets B_α^f are level sets of f on M, indexed over values α in the image of f. The *product* $\vee$ of two partitions is defined by

$$\zeta \vee \xi = \{A \cap B | A \in \zeta, B \in \xi\}. \qquad (35)$$

The finest partition π is the partition into individual elements of the set M.

Proposition 16. *A representation* $(\mathbf{f}, \mathbf{F})$*, where* $\mathbf{f} = (f_1, ..., f_n)$*, is faithful if and only if the partition*

$$\zeta_{\mathbf{f}} = \bigvee_{j=1,...,n} \zeta_{f_j} \qquad (36)$$

is the finest partition π.

Proof. Assume $(\mathbf{f}, \mathbf{F})$ is faithful and $\pi \neq \bigvee_{j=1,...,n} \zeta_{f_j}$. Then there are two points in M that have the same values of $\mathbf{f}$ associated with them and we get a contradiction. Conversely, assume that $\pi = \bigvee_{j=1,...,n} \zeta_{f_j}$. Then, the partition is faithful as every point of M gets assigned a unique set of values of observables $f_1, ..., f_n$. □

Eigenfunctions and linear representations. An eigenfunction ϕ of the composition operator associated with T and T^t, respectively, satisfies

$$U\phi = \lambda\phi, \qquad (37)$$
$$U^t\phi = e^{\lambda t}\phi, \qquad (38)$$

where λ is the associated eigenvalue.[5] In discrete time, let $\mathbf{F}(\phi) = \lambda\phi$. Thus, $(\phi, \mathbf{F})$ is a (possibly reduced) representation of T.

Remark 17. Note that, at this stage, we do not put a constraint on ϕ being in a specific functional space. This is intentional, as the eigenfunction definition does not need the definition of the space the function lives in. However, when we describe spectral expansions below, the precise definitions of such functional spaces are needed and will be given.

More generally, let $(\mathbf{f}, \mathbf{F})$ be an m-dimensional representation of T such that $\mathbf{F}(\mathbf{f}) = A\mathbf{f}$, where A is an $m \times m$ matrix. We denote such a linear representation by $(\mathbf{f}, A)$. Since we have

$$U\mathbf{f}(\mathbf{m}) = \mathbf{f} \circ T(\mathbf{m}) = A\mathbf{f}(\mathbf{m}), \qquad (39)$$

we call A an eigenmatrix of U. A similar definition holds for the continuous-time case, where we require

$$U^t\mathbf{f}(\mathbf{m}) = \mathbf{f} \circ T^t(\mathbf{m}) = e^{At}\mathbf{f}(\mathbf{m}) \qquad (40)$$

for A to be an eigenmatrix of U^t. In the differentiable case, we get

$$\frac{d[U^t\mathbf{f}(\mathbf{m})]}{dt}\Big|_{t=0} = A\mathbf{f}(\mathbf{m}). \qquad (41)$$

Assume that A has distinct eigenvalues. Let $\langle \cdot, \cdot \rangle$ denote the standard complex inner product on $\mathbb{C}^m$, defined by

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_i \mathbf{v}_i \cdot \mathbf{w}_i^c.$$

Then we get the following.

Proposition 18. *The eigenvalues of A, $\lambda_1, ..., \lambda_m$, are eigenvalues of U, and the associated eigenfunctions of U are given by*

$$\phi_j = \langle \mathbf{f}, \mathbf{w}_j \rangle, \qquad (42)$$

where w_j is the eigenvector of A^T.

Proof. In discrete time, we have

$$
\begin{aligned}
U\phi_j(\mathbf{m}) &= \phi_j(T(\mathbf{m})) = \langle \mathbf{f}(T(\mathbf{m})), \mathbf{w}_j \rangle \\
&= \langle A\mathbf{f}(\mathbf{m}), \mathbf{w}_j \rangle = \langle \mathbf{f}(\mathbf{m}), A^T\mathbf{w}_j \rangle \\
&= \langle \mathbf{f}(\mathbf{m}), \lambda_j^c \mathbf{w}_j \rangle = \lambda_j \langle \mathbf{f}(\mathbf{m}), \mathbf{w}_j \rangle \\
&= \lambda_j \phi_j(\mathbf{m}).
\end{aligned}
$$

The proof for continuous time proceeds in a similar way. □

Example 19. The system $\dot{x} = f(x), x \in \mathbb{R}^1, f \in C^1$, always has a smooth eigenfunction ϕ satisfying $\dot\phi = \lambda\phi$ provided $\phi = \exp(\lambda \int_{x_0}^x f^{-1}(x)dx) \in C^1$ for some $\lambda < \infty$. The representation $(\phi, \mathbf{G}^t)$ is linear, where $\mathbf{G}^t(\phi) = e^{\lambda t}\phi$ and $\dot\phi = \lambda\phi$.

Perhaps the most important and simplest representation, if it exists, is the eigenfunction representation, given in discrete time by

$$\mathbf{f}(T\mathbf{m}) = \Lambda(\mathbf{f}(\mathbf{m})), \qquad (43)$$

where Λ is a diagonal $m \times m$ matrix and the diagonal elements $\Lambda_{jj} = \lambda_j, j = 1, ..., m$, are eigenvalues of U, satisfying

$$Uf_j = \lambda_j f_j, \qquad (44)$$

where f_j is the eigenfunction of U associated with the eigenvalue λ_j.

Let A be an efficient faithful representation. Let P be a diagonalizing matrix such that $P^{-1}AP = \Lambda$, where Λ is a diagonal matrix, and let $\phi = (\phi_1, ..., \phi_m)$. Then

$$\phi' = P^{-1}\mathbf{f}' = P^{-1}A\mathbf{f} = P^{-1}AP\phi = \Lambda\phi, \qquad (45)$$

and (ϕ, Λ) is a linear representation conjugate to $(\mathbf{f}, A)$. This is of interest because it leads to the following corollary.

Corollary 20. *For any linear diagonalizable finite-dimensional representation $(\mathbf{f}, \mathbf{F})$ (respectively, $(\mathbf{f}, \mathbf{F}^t)$) of T (respectively, T^t), $\mathbf{f}$ is in the span of n eigenfunctions $\phi = (\phi_1, \phi_2, ..., \phi_n)$ of U (respectively, U^t), where n is the dimension of the representation.*

Proposition 21. *Consider a linear, diagonalizable, finite-dimensional representation $(\mathbf{f}, \mathbf{F})$ (respectively, $(\mathbf{f}, \mathbf{F}^t)$) of T (respectively, T^t), and another finite-dimensional representation $(\mathbf{g}, \mathbf{G})$ (respectively, $(\mathbf{g}, \mathbf{G}^t)$) of T (respectively, T^t). Let $\mathbf{h}$ be a homeomorphism between them, $\mathbf{f} = \mathbf{h}(\mathbf{g})$. Then $\mathbf{g} = \mathbf{h}^{-1}(P\phi)$ for an invertible matrix P.*

Proof. We have that $\mathbf{f}$ is in the span of ϕ, namely

$$\mathbf{f} = P\phi, \qquad (46)$$

and P is invertible. Since $\mathbf{g} = \mathbf{h}^{-1}\mathbf{f} = \mathbf{h}^{-1}(P\phi)$ we are done. □

Algebra of eigenfunctions. Eigenfunctions of U form an algebra: Let ϕ, ψ be eigenfunctions of U with the associated eigenvalues λ, β. Then $\phi\psi$ is also an eigenfunction, with eigenvalue $\lambda\beta$:

$$U\phi\psi = (\phi\psi) \circ T = \phi \circ T\psi \circ T = \lambda\beta\phi\psi. \qquad (47)$$

Thus, any efficient representation using n eigenfunctions, (ϕ, Λ), leads to many nonefficient representations that can be obtained by adding products of eigenfunctions into the set of representation functions.

Eigenfunctions, Geometry, and Stability

The above discussion leads to the conclusion that eigenfunctions of the Koopman family of operators U^t are useful in the context of representations—not only is the representation consisting of eigenfunctions linear, it is also fully decoupled, as each eigenfunction ϕ satisfies

$$\dot{\phi} = \lambda\phi, \qquad (48)$$

where λ is the associated eigenvalue. In the discrete-time case, eigenfunctions of U similarly satisfy

$$\phi' = \phi \circ T = \lambda\phi. \qquad (49)$$

The more general notion is that of a function ϕ^k, such that for discrete maps $(U - \lambda I)^k \phi^k = 0$, and for continuous time $(U^t - e^{\lambda t}I)^k \phi^k = 0$. Such a function is called a generalized eigenfunction [Mez20]. Clearly, eigenfunctions satisfy such equations for $k = 1$. Using generalized eigenfunctions, for a linear representation $(\mathbf{f}, A)$ we get a linear system

$$\dot{\phi} = J\phi, \qquad (50)$$

where J is the complex Jordan canonical form matrix and $\phi = (\phi_1^1, ..., \phi_1^{k_1}, ..., \phi_l^1, ..., \phi_l^{k_l})$, l being the number of distinct eigenvalues and k_j the algebraic multiplicity of eigenvalue λ_j.

Level sets of eigenfunctions and invariant partitions. Level sets of eigenfunctions provide insight into the geometry of the state space of a dynamical system. Consider an eigenfunction ϕ for $T : M \to M$ at eigenvalue 1. It satisfies, in discrete and continuous time, respectively,

$$\begin{aligned} \phi \circ T(\mathbf{m}) &= \phi(\mathbf{m}), \\ \dot{\phi}(\mathbf{m}) &= 0. \end{aligned} \qquad (51)$$

Therefore, ϕ is invariant under the dynamics of T, and its level sets, defined by $\phi = c$ for some constant $c \in \mathbb{C}$, are *invariant sets*. Thus, learning of linear representations from data enables learning of invariant sets of the underlying system. The partition ζ_ϕ into the level sets of ϕ is an example of a *fixed partition*, since for any set $D \in \zeta_\phi$, in discrete time $TD = D$. The finest such partition is the ergodic partition that has interesting additional metric properties [Mez94]. The mapping $\phi : M \to \mathbb{C}$ defines a *fixed factor* of T, whose domain is $\phi(M)$ and the dynamics is trivial, given in discrete time by $\phi' = \phi$.

Example 22. Consider the set of states (positions and angular velocities) $\mathbf{m} \in M = S^1 \times \mathbb{R}$ of a pendulum of mass m. Let $\theta \in [0, 2\pi)$ and define representation functions $\theta(\mathbf{m}), \omega(\mathbf{m}) = \dot{\theta}(\mathbf{m}) \in \mathbb{R}$. We have the ordinary differential representation

$$\begin{aligned} \dot{\theta} &= \omega, \\ \dot{\omega} &= -\frac{g}{l}\sin\theta, \end{aligned} \qquad (52)$$

with g the acceleration of gravity and l the length of the pendulum. Let

$$H = \omega^2/2 - \frac{g}{l}\cos(\theta), \qquad (53)$$

and $\dot{H} = 0$. Thus, the Hamiltonian H is an eigenfunction of the Koopman operator associated with pendulum dynamics. Its level sets are invariant. The level sets of the Hamiltonian for 1 degree of freedom systems form the ergodic partition, but this is not the case for N degrees of freedom Hamiltonian systems, since, e.g., tori with irrational rotation dynamics can have half the dimension of the state space.

The eigenfunction ϕ_ω of T corresponding to an eigenvalue $\lambda = e^{i\omega} \neq 1$ on the unit circle yields level sets that form an *invariant partition*. Namely if $D \in \zeta_{\phi_\omega}$, then $TD = E \in \zeta_{\phi_\omega}$. If $n\omega = 2\pi m$, where $m, n \in \mathbb{Z}$, then $T^n D = D$ for every $D \in \zeta_{\phi_\omega}$. In that case ζ_{ϕ_ω} is a *periodic partition*. The same analysis holds for continuous-time systems in the case when the eigenvalue is $\lambda = i\omega$, on the imaginary axis. For limit cycling systems, with limit cycling frequency ω, there exists an eigenfunction ϕ_ω, the level sets of which satisfy

$$\dot{\phi}_\omega = i\omega\phi_\omega. \qquad (54)$$

Such level sets are called isochrons. The notion of generalized isochrons in dynamical systems with a toroidal attractor with diophantine irrational rotation dynamics stems from further examination of partitions induced by imaginary axis eigenvalues [MM12].

More generally, consider an eigenfunction of U that satisfies

$$\phi' = \lambda\phi \qquad (55)$$

with $|\lambda| < 1$, or an eigenfunction of U^t that satisfies

$$\dot{\phi} = \lambda\phi \qquad (56)$$

for $\lambda \in \mathbb{C}^-$, the left half-plane (excluding the imaginary axis). Then necessarily, $\phi(t) \to 0$ as $t \to \infty$. The level sets of ϕ again form a *partition* of M that is invariant. Namely the collection of disjoint sets $A_\alpha^\phi, \alpha \in \mathbb{C}$, such that the sets $A_\alpha^\phi = \{\mathbf{m} \in M | \phi(\mathbf{m}) = \alpha, \alpha \in \mathbb{C}\}$ form a partition of M, such that

$$\begin{aligned} \bigcup_{\alpha \in \mathbb{C}} A_\alpha^\phi &= M, \\ T^t A_\alpha^\phi &= A_{\alpha e^{\lambda t}}^\phi. \end{aligned} \qquad (57)$$

The second property indicates the sets A_α^ϕ are elements of an invariant partition under T^t. Level sets of Koopman eigenfunctions always provide us with invariant partitions of the state space.

Proposition 23. *Let ϕ be a Koopman eigenfunction of a continuous-time system $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$ on $M \subset \mathbb{R}^n$ with the flow*

$T^t : M \to M$, or of a map $T : M \to M$. Then the level sets A_α^ϕ of ϕ,

$$A_\alpha^\phi = \{\mathbf{x} \in M | \phi(\mathbf{x}) = \alpha\}, \qquad (58)$$

where $\alpha \in \mathbb{C}$, are elements of the invariant partition $\zeta_\phi = \{A_\alpha^\phi | \phi(\mathbf{x} \in A_\alpha^\phi) = \alpha\}$.

Proof. We will do the proof for the continuous-time case. By definition, ϕ satisfies

$$\dot{\phi} = \lambda\phi, \qquad (59)$$

where λ is the eigenvalue associated with ϕ. Therefore

$$\phi(T^t\mathbf{x}) = e^{\lambda t}\phi(\mathbf{x}), \qquad (60)$$

and letting $\phi(\mathbf{x}) = \alpha$, we get $\phi(T^t\mathbf{x}) = e^{\lambda t}\alpha$. Since $e^{\lambda t}\alpha$ does not depend on $\mathbf{x}$, all the points in the set A_α^ϕ get mapped into $A_{e^{\lambda t}\alpha}^\phi$ by T^t, $t \in \mathbb{R}$, and thus ζ_ϕ is an invariant partition. The proof in the discrete-time case is similar. $\qquad\square$

Invariant partitions can be recurrent and nonrecurrent.

Definition 24. A recurrent invariant partition is an invariant partition such that for any set D in it there is no τ such that $d(T^tD, D) \geq \epsilon > 0$ for some ϵ and all $t > \tau > 0$. Here $d(E, F)$ is the Hausdorff distance of sets E and F. An invariant partition that is not recurrent is called nonrecurrent.

In other words, given an $\epsilon > 0$, for any set A in a recurrent partition and for any τ, there is a time $t > \tau$ such that $d(T^tA, A) < \epsilon$. Fixed and periodic partitions are clearly recurrent. So are partitions associated with an eigenvalue on the unit circle (discrete time) or imaginary axis (continuous time) where $\omega \neq 2\pi m/n$ for any $m, n \in \mathbb{Z}$.

Example 25. Consider the system

$$\begin{aligned} \dot{r} &= r(1 - r^2), \\ \dot{\theta} &= \omega, \end{aligned} \qquad (61)$$

where $(r, \theta) \in [0, \infty) \times S^1$. The level sets of the coordinate function $r : [0, \infty) \times S^1 \to \mathbb{R}$ comprise an invariant nonrecurrent partition for the system. However, r is not an eigenfunction of the system. If we map every level set of r into a single point we obtain the quotient space $Q = [0, \infty)$. However the dynamics "induced" on it by the mapping $r : M \to R$ from the state space to r is nonlinear. Using the eigenfunction of the system given by $\phi = (r^2 - 1)/r^2$ corresponding to the eigenvalue -2 (which is also the Floquet stability exponent for the limit cycle) we obtain linear dynamics $\dot{\phi} = -2\phi$ on Q. The invariant partition of $\mathbb{R}^2$ into level sets of ϕ is nonrecurrent. The invariant partition into level sets of $\phi_\omega = e^{i2\pi\theta}$ corresponding to eigenvalue $\lambda = i\omega$ is recurrent (periodic).

The numerical study of invariant sets and invariant partitions using Koopman operator theory was initiated in [Mez94] for measure-preserving systems, where joint level

sets of time averages of continuous functions—that are elements of the eigenspace of U at 1—were used to compute invariant partitions. Figure 1 (from [LM15]) shows numerical approximation of the ergodic partition of the so-called standard map on the 2-torus using such time averages (these are a part of the Generalized Laplace Analysis, the larger computational framework in Koopman operator theory; see below for more details). We consider the standard map in the form:

$$\begin{aligned} x' &= x + y + \epsilon\sin(2\pi x) && [\text{mod } 1], \\ y' &= y + \epsilon\sin(2\pi x) && [\text{mod } 1], \end{aligned} \qquad (62)$$

where $(x, y) \in [0, 1] \times [0, 1] \equiv [0, 1]^2$. It is an area-preserving (symplectic) map which exhibits a variety of invariant sets, both regular (composed of periodic or quasi-periodic orbits) and chaotic. These sets evolve in size and structure as the parameter ϵ is varied.
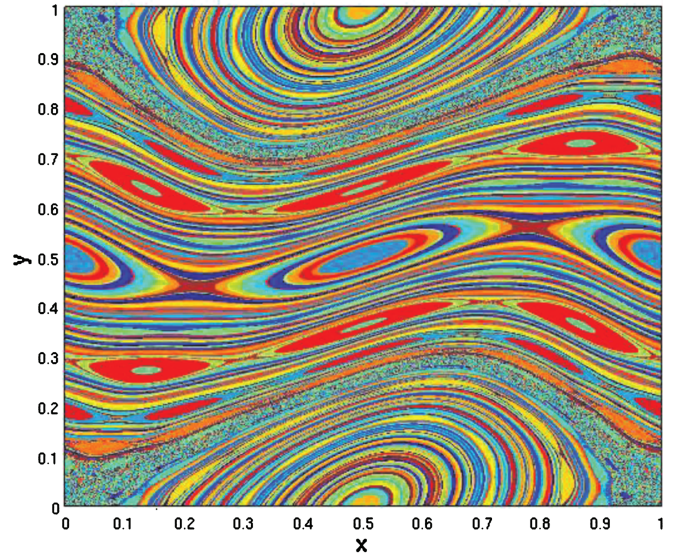

Figure 1. Two-function approximation of the ergodic partition of the standard map (62).

The above example indicates that joint level sets of several (or, in the ergodic partition case, countably infinite) eigenfunctions also provide invariant partitions. This concept can be used to compute stable, unstable, and center manifolds [Mez20].

Proposition 26. *Let $T : M \to M$ have a fixed point at $\mathbf{m}$. Let $\lambda_1, ..., \lambda_u$ be positive real part eigenvalues, let $\lambda_{u+1}, ..., \lambda_{u+c}$ be 0 real part eigenvalues, and let $\lambda_{u+c+1}, ..., \lambda_s$ be negative real part eigenvalues of a linear faithful efficient representation $(\mathbf{x}, A)$ with $\mathbf{x}(\mathbf{m}) = 0$. Let*

$$\phi_1, ..., \phi_{u+c+s} \qquad (63)$$

be the (generalized) eigenfunctions of the associated Koopman operator. Then the joint level set of (generalized) eigenfunctions

$$L_s = \{\mathbf{x} \in \mathbb{R}^n | \phi_1(\mathbf{x}) = 0, ..., \phi_{u+c}(\mathbf{x}) = 0\} \qquad (64)$$

is the stable subspace E^s,

$$L_c = \{\mathbf{x} \in \mathbb{R}^n | \phi_1(\mathbf{x}) = 0, ..., \phi_u(\mathbf{x}) = 0,$$
$$..., \quad \phi_{u+c+1}(\mathbf{x}) = 0,$$
$$..., \quad \phi_{u+c+s}(\mathbf{x}) = 0\} \qquad (65)$$

is the center subspace E^c, and

$$L_u = \{\mathbf{x} \in \mathbb{R}^n | \phi_{u+c+1}(\mathbf{x}) = 0, ..., \phi_{u+c+s}(\mathbf{x}) = 0\} \qquad (66)$$

is the unstable subspace E^u.

In turn, $\mathbf{x}^{-1}(L_s), \mathbf{x}^{-1}(L_c)$, and $\mathbf{x}^{-1}(L_u)$ are the stable subset, the center subset, and the unstable subset of $\mathbf{m} = \mathbf{x}^{-1}(0)$, the fixed point of T.

Proof. The proof follows directly from Proposition 3.1 in [Mez20]. □

In the case when the fixed point is stable, the *inertial manifolds* can also be computed as joint zero level sets of a subset of Koopman operator eigenfunctions. Namely, if one can order eigenvalues by the magnitude of their real parts, $Re(\lambda_1) < Re(\lambda_2) < \cdots < Re(\lambda_n)$, then the *m*-dimensional inertial manifold is defined as the joint zero level set of $\phi_j, j \geq m$.

Eigenfunctions and stability. Provided we find eigenfunctions that compose a faithful representation, we can use them to examine stability properties of dynamical systems.

Theorem 27. *Let $(\phi = (\phi_1, ..., \phi_n), \Lambda)$, where Λ is a diagonal matrix, be a faithful representation of T^t such that $\lambda_j \in \mathbb{C}^-, j = 1, ..., n$, and let $0 \in \phi(M)$. Then the point $\mathbf{m}$ that satisfies $\phi(\mathbf{m}) = 0$ is a globally stable fixed point of T^t.*

Proof. Clearly $\lim_{t \to \infty} \phi(\mathbf{m}_0) = 0$ for any $\mathbf{m}_0 \in M$.[6] But $\phi(M)$ contains 0, and the representation is faithful. Thus assuming $\lim_{t \to \infty} T^t \mathbf{m}_0 \neq \mathbf{m}$ leads to a contradiction. □

An analogous statement holds for discrete time T, and for more general attractors, with possibly a smaller number of "stable" eigenfunctions involved—in the case of a limit cycling system in two dimensions, the limit cycle attractor is the zero level set of a single globally supported eigenfunction associated with an eigenvalue with a negative real part, and existence of such an eigenfunction serves to determine the global stability. The condition of faithfulness can be checked near the fixed point or the limit cycle in the case of ordinary differential representations, and more general attractors can be considered [MMS20].

---

[6]*Interestingly, this follows directly from the eigenfunction equation $\phi' = \lambda\phi$ and does not require continuity of ϕ.*

Nonlinear Representations

We have discussed linear representations that are based on finding eigenfunctions of the Koopman operator, and lead to linear dynamics (reduction of the full Koopman operator) on a subspace of observables. Finite nonlinear representations also lead to a reduction since the space of observables they operate on is the space of all observables that are constant on joint level sets of $\mathbf{f}$—as in that case knowing $\mathbf{f}$ leads to knowing $\mathbf{F}(\mathbf{f})$. The following simple lemma holds.

Lemma 28. *The space $\mathcal{O}_\mathbf{f}$ of observables that are constant on joint level sets of $\mathbf{f}$ is a linear subspace of $\mathcal{O}$.*

Definition 29. *A subspace of $\mathcal{O}$ is said to be generated by a finite set of functions $\mathbf{f}$ if it is the subspace $\mathcal{O}_\mathbf{f}$ containing all observables $g(\mathbf{f}) \in \mathcal{O}$.*

Corollary 30. *A finite-dimensional invariant subspace of U spanned by observables in a linear representation of dimension n is a span of generalized eigenfunctions. Let $(\mathbf{f}, \mathbf{F})$ be a finite-dimensional, nonlinear representation. Then the subspace $\mathcal{O}_\mathbf{f}$ generated by $\mathbf{f}$ is an invariant subspace.*

Thus, the search for finite-dimensional linear representations can be reduced to search for spans of generalized eigenfunctions. The search for nonlinear representations is the search for invariant subspaces generated by finite sets of observables. It becomes clear that the eigenvalue-eigenfunction problem for the Koopman operator of finding ϕ and λ such that

$$U\phi = \phi \circ T = \lambda\phi \qquad (67)$$

is just a particular example of finding solutions of the *representation eigenproblem* for a finite set of functions $\mathbf{f}$ and a map $\mathbf{F}$ that satisfy

$$\boxed{U\mathbf{f} = \mathbf{f} \circ T = \mathbf{F}(\mathbf{f})} \qquad (68)$$

in the particular case when $\mathbf{F}(\mathbf{f}) = A\mathbf{f}$, where A is an $n \times n$ matrix; A is called an eigenmatrix. In the same vein, we could call $\mathbf{F}$ an *eigenmap*.

Nonlinear representations can be reduced to linear representations provided a conjugacy exists.

Proposition 31. *Assume $(\mathbf{f}, \mathbf{F})$ is an n-dimensional representation of T and $\mathbf{h}$ is a conjugacy of $(\mathbf{f}, \mathbf{F})$ to a linear representation $(\mathbf{g}, A)$. Then $\mathbf{f}$ is in a subspace of $\mathcal{O}$ generated by the set of (generalized) eigenfunctions $\phi = (\phi_1, ..., \phi_n)$.*

Proof. Since A is linear, there are eigenfunctions $\phi = C\mathbf{g}$, where C is an $n \times n$ invertible matrix such that

$$\phi' = \phi \circ T = J\phi, \qquad (69)$$

where J is the Jordan normal form matrix for A. Since $\mathbf{h}$ is a conjugacy, $\mathbf{h}(\mathbf{F}(\mathbf{f} \circ T)) = A\mathbf{g}(\mathbf{h}(\mathbf{f} \circ T))$ and $\mathbf{f} = \mathbf{h}^{-1}\mathbf{g} = \mathbf{h}^{-1}C^{-1}\phi$, proving that $\mathbf{f}$ are in the invariant subspace generated by ϕ. □

The above result has a profound consequence on the issue of which systems can be rendered linear, that is in turn related to spectral properties of U. We discuss this next.

The Spectral Triple

So far, we have discussed the eigenvalues and eigenfunctions of the Koopman operator, and their connection to linear representations. Let U act on some Banach space of observables. Then, the eigenvalues are part of the spectrum $\sigma(U)$, the complement of the residual set $\rho(U)$ defined as

$$\rho(U) = \{\lambda \in \mathbb{C} | (U - \lambda I)^{-1} \text{exists}\}. \tag{70}$$

The operator $(U - \lambda I)^{-1}$ is called the resolvent operator. The residual set, and thus the spectrum, are dependent on the functional space on which U operates [Mez20]. The operator U is called scalar provided

$$U = \int_{\sigma(U)} \beta dE, \tag{71}$$

where E is a family of spectral projections forming resolution of the identity, and spectral provided

$$U = S + N, \tag{72}$$

where S is scalar and N is quasi-nilpotent. Examples of functional spaces in which Koopman operators are scalar and spectral are given in [Mez20]. The notion of Koopman modes (originally called shape modes when defined in [Mez05] for the measure-preserving, on-attractor part of the dynamics) is given by the following definition.

Definition 32. Let $\mathbf{f} \in \mathcal{O}^n$ be a vector of observables. For a scalar Koopman operator U, the Koopman mode of $\mathbf{f}$ associated with an eigenvalue λ is given by

$$\mathbf{s}_\lambda = \mathbf{f}_\lambda ./ \phi, \tag{73}$$

where ./ is componentwise division, ϕ is the unit norm eigenfunction associated with λ, and

$$\mathbf{f}_\lambda = \mathbf{f} - \int_{\sigma(U)\backslash\{\lambda\}} \beta dE(\mathbf{f}). \tag{74}$$

We assume that the dynamical system T has a Milnor attractor $\mathcal{A}$ such that for every continuous function g, for almost every $\mathbf{m} : M \to M$ with respect to an a priori measure ν on M (without loss of generality as we can replace M with the basin of attraction of $\mathcal{A}$) the limit

$$g^*(\mathbf{m}) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} U^i g(\mathbf{m}) \tag{75}$$

exists. This is the case, e.g., for smooth systems on subsets of $\mathbb{R}^n$ with a Sinai-Bowen-Ruelle (or physical) measure μ, where ν is the Lebesgue measure. Let $\mathbf{g}(\mathbf{z}, \mathbf{m})$ be a field of observables indexed by the field variable $\mathbf{z}$ (e.g., $\mathbf{g}(\mathbf{z}, \mathbf{m})$ could be temperature at spatial position $\mathbf{z}$ when the system

is at $\mathbf{m} \in M$). The *spectral expansion* of the action of U on $\mathbf{g}$ in $\mathcal{O} = L^2(\mu)$ is given by [Mez05]

$$
\begin{aligned}
U\mathbf{g}(\mathbf{z}, \mathbf{m}) =\ & \mathbf{g}^*(\mathbf{z}, \mathbf{m}) \\
& + \sum_{j=1}^{k} \exp(i\omega_j)\phi_j(\mathbf{m})\mathbf{s}_j(\mathbf{z}) \\
& + \int_0^1 \exp(i2\pi\alpha)dE(\alpha)(\mathbf{g}(\mathbf{z}, \mathbf{m})),
\end{aligned}
$$

where $\mathbf{g}^*(\mathbf{z}, \mathbf{m})$ is the time average (75), $\lambda_j = e^{i\omega_j}$ is an eigenvalue with the associated eigenfunction ϕ_j, k can be ∞, and $\mathbf{s}_j(\mathbf{z})$ is the jth Koopman mode, i.e., the projection $\int_M \mathbf{g}\phi_j^c d\mu$ of $\mathbf{g}$ on the eigenspace of the eigenfunction ϕ_j. The triple $(\lambda_j, \phi_j, \mathbf{s}_j)$ is called the spectral triple. From the previous discussion, any finite set of ϕ_j's provides for a (diagonal!) linear representation of T.

The term

$$\int_0^1 \exp(i2\pi\alpha)dE(\alpha)(\mathbf{g}(\mathbf{z}, \mathbf{m})) \in \mathcal{O}_{cont} \subset \mathcal{O} \tag{76}$$

is the projection of $\mathbf{g}$ on the continuous part of the spectrum that is orthogonal to

$$\mathcal{O}_{eig} = \text{cl}(\text{span}\{\phi_j, j = 1, ..., \infty\}), \tag{77}$$

where cl stands for closure. Any finite-dimensional representation of T in $\mathcal{O}_{cont}$ must be nonlinear.

Corollary 33 (To Proposition 31). *A finite-dimensional representation $(\mathbf{f}, \mathbf{F})$ is not smoothly conjugate to a linear representation provided $\mathbf{f} \not\perp \mathcal{O}_{cont}$.*

Proof. Assume the conjugacy to a linear representation exists. Then U has point spectrum [Mez20] and therefore $\mathbf{f} \perp \mathcal{O}_{cont}$. $\square$

Dynamical systems can consist of linearizable and nonlinearizable parts, as the following example shows.

Example 34. Consider the following coupling of the limit cycling system in Example 25 and the Lorenz system in Example 14:

$$
\begin{aligned}
\dot{r} &= (1 + f(x, y, z))r(1 - r^2), \\
\dot{\theta} &= \omega, \\
\dot{x} &= \sigma(y - x), \\
\dot{y} &= x(\rho - z) - y, \\
\dot{z} &= xy - \beta z,
\end{aligned}
\tag{78}
\tag{79}
$$

where f is some positive bounded real function $0 < f < c < \infty$. The representation of this system is part linear and part nonlinear, where the complex function $\psi = e^{i\theta}$ constitutes the linear part

$$\dot{\psi} = i\omega\psi. \tag{80}$$

Since the Lorenz system is mixing, it has no point spectrum, and thus is not linearizable. This is therefore an example in which the system is not linearizable as a whole, but there exists a semiconjugacy where a part of it is linearizable.

Similar results are available for spectral expansions of a large class of systems with Milnor attractors; see [Mez20].

Types of spectra. There are two elements that determine the spectrum of a given dynamical system: the function space and the type of the attractor determined by its dynamics [Mez20]. Interestingly, these are commingled: a linear dynamical system in a complex plane can have a fixed point, where on a subset of linear observables the spectrum is discrete, but in $L^2(\nu)$, where ν is Lebesgue, will have a very large spectrum, for example filling the entire unit disk of the complex plane. The "on-attractor" space can always be chosen to be $L^2(\mu)$. But the transient dynamics requires spaces *adapted* to the dynamics, as described in [Mez20]. In such spaces, the on-attractor spectrum and the off-attractor spectrum combine by multiplication to provide the full spectrum of the Koopman operator: For a scalar Koopman operator of a dynamical system with a Milnor attractor with μ being a Borel measure, define the tensor product space

$$\mathcal{H} = \mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_{\mathcal{B}}, \tag{81}$$

where $\mathcal{H}_{\mathcal{A}} = L^2(\mu)$. Further, define $\mathcal{H}_{\mathcal{B}} = \tilde{\mathcal{H}}_{\mathcal{B}} \cup \mathbf{1}$, where $\mathbf{1}$ is the constant unit function on $\mathcal{D}$ and $\tilde{\mathcal{H}}_{\mathcal{B}} \subset C(\mathcal{D})$ is a Hilbert space of functions $f : \mathcal{D} \to \mathbb{C}$ that vanish on the attractor $\mathcal{A}$. Clearly, $U^t = U^t|_{\mathcal{H}_{\mathcal{A}}} \otimes U^t|_{\mathcal{H}_{\mathcal{B}}}$ on $\mathcal{H}$. Define $P(a, b) = a \cdot b, \ a, b \in \mathbb{C}$, to be the scalar product of a and b, and

$$P(A, B) = \bigcup_{a \in A, b \in B} P(a, b), \quad A, B \subset \mathbb{C}. \tag{82}$$

We have the following [Mez20].

Theorem 35. *Consider the composition operator $U^t : \mathcal{H} \to \mathcal{H}$, and let $\sigma(U^t|_{\mathcal{H}_{\mathcal{A}}}), \sigma(\mathcal{K}|_{\mathcal{H}_{\mathcal{B}}})$ be the spectra of its restrictions to $\mathcal{H}_{\mathcal{A}}$ and $\mathcal{H}_{\mathcal{B}}$ with the associated projection-valued spectral measures $P_\omega, \ \omega \in S^1$, and $P_z, \ z \in \mathbb{C}$. Then $\sigma(U^t) = \mathrm{cl}(P(\sigma(U^t|_{\mathcal{H}_{\mathcal{A}}}), \sigma(U^t|_{\mathcal{H}_{\mathcal{B}}})))$ and*

$$U^t = \int_{\mathbb{C}} \int_{\mathbb{R}} e^{zt} e^{i2\pi\omega t} dP_\omega dP_z. \tag{83}$$

Example 36. For a continuous-time, globally stable limit cycle system in $\mathbb{R}^{n+1}$ with limit cycle frequency ω, the on-attractor spectrum is $in\omega, n \in \mathbb{Z}$. The off-attractor spectrum in Modulated Fock Space [Mez20] consists of positive integer combinations $\lambda_{\mathbf{m},k} = \mathbf{m} \cdot \beta, \mathbf{m} \in \mathbb{N}^n$, of Floquet stability exponents $\beta = (\beta_1, ..., \beta_n)$. Thus, the full spectrum in $\mathcal{H}$ is given by $in\omega + \mathbf{m} \cdot \beta, n \in \mathbb{Z}, \mathbf{m} \in \mathbb{N}^n$.

Consider the three-dimensional, limit cycling system

$$\dot{x} = y, \tag{84}$$
$$\dot{y} = x - x^3 - cy, \tag{85}$$
$$\dot{\theta} = \omega. \tag{86}$$

The two fixed points of equations (84)–(85) are $y = 0, x = \pm 1$. The linearization matrix at those is

$$A = \begin{bmatrix} 0 & 1 \\ -2 & -c \end{bmatrix}, \tag{87}$$

and thus the eigenvalues are determined by

$$(-\lambda)(-c - \lambda) + 2 = \lambda^2 + c\lambda + 2 = 0, \tag{88}$$

leading to

$$\lambda_{1,2} = \frac{-c \pm \sqrt{c^2 - 8}}{2}. \tag{89}$$

For $c = \sqrt{7}, \omega = 1$, the eigenvalues read $\lambda_{3,4} = -1.3228756 \pm .5i$. Setting $\omega = 1$, the other two principal eigenvalues are $\pm i$. In Figure 2 (from [Mez20]) we show a subset of the eigenvalues of the Koopman operator on $L^2(S^1) \times \mathcal{A}$, where $\mathcal{A}$ is the space of analytic functions on the plane, in the basin of attraction of either of the limit cycles (since they are symmetric) of (84)–(86).

Eigenvalues of the Koopman operator, for $\{0, 1, .., 4\} \times \{0, 1, .., 4\}$ lattice of integers.

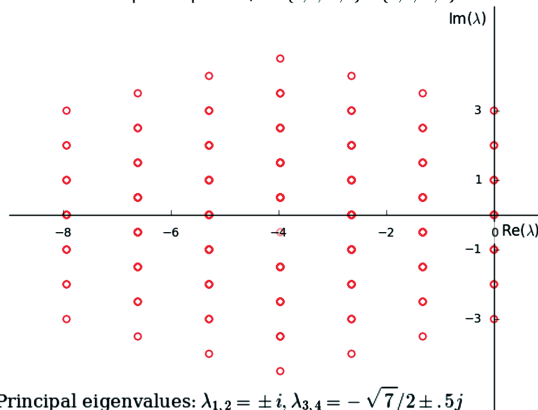Principal eigenvalues: $\lambda_{1,2} = \pm i, \lambda_{3,4} = -\sqrt{7}/2 \pm .5j$

Figure 2. Eigenvalues of the Koopman operator, for the system (84)–(86) and the lattice of integers $\{0, 1, ..., 4\} \times \{0, 1, ..., 4\}$.

While the computations of point spectrum were already available in [MB04] using GLA, and [RMB+09] using DMD, computations of continuous spectrum are more recent [KPM20, Gia21]. They have been used to identify coherent pseudoeigenfunctions in the (mixing!) Lorenz system. The contour plot of such eigenfunction is shown in Figure 3 (from [KPM20]). Such observables have much longer prediction horizons than a typical observable on the Lorenz system.
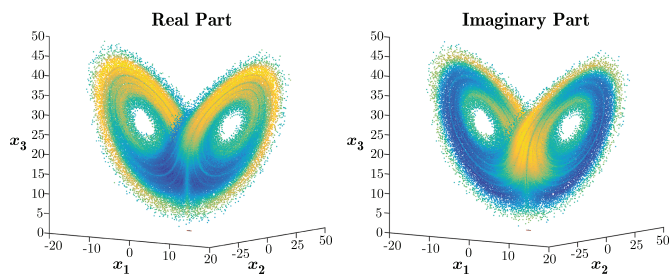
Figure 3. Lorenz system: Approximation of the spectral projection $P_{[a,b]}f$ with $f(x) = x_3$ and $[a,b] = [0.24, 0.28]$. The contours of the real and imaginary parts of the resulting pseudoeigenfunction are indicated by color, showing substantial regularity.

Learning Dynamical Systems from Data

Historically, the dynamical equations of motion, such as Newton's, Einstein's, and Schrödinger's, were obtained using depth of human intuition guided by small amounts of, or no, data (Einstein's case). Classical automatized approaches to learning dynamical systems from data arose in control theory. The goal was to connect system inputs $\mathbf{u}$ to system outputs $\mathbf{y}$ via analysis of a structured model connecting these. The most commonly used structure of the model is linear:

$$\begin{aligned} \dot{\mathbf{x}} &= A\mathbf{x} + B\mathbf{u}, \\ \mathbf{y} &= C\mathbf{x}, \end{aligned} \tag{90}$$

where $\mathbf{x}$ is the state of the system and $\mathbf{y}$ is a linear vector of observables.

In *static* machine learning problems there are also "inputs" and "outputs," in the simplest case $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^m$, although both input and output spaces can be more complicated, say manifolds. The key objective is to connect inputs and outputs by a map $f : \mathbb{R}^n \to \mathbb{R}^m$ learned from a measured subset of input-output pairs

$$(\mathbf{x}_j, \mathbf{y}_j), j = 1, ..., N.$$

Let $Im(f)$ be the image of the map f, and $Dom(f)$ its domain. Provided $Im(f) \subset Dom(f)$, f could be considered a dynamical system, since $f^k(\mathbf{x})$ is well defined for any $\mathbf{x} \in Dom(f)$. In this case the data pairs $(\mathbf{x}_j, \mathbf{y}_j)$ can be obtained as successive points along the trajectory of f:

$$\mathbf{y}_j = f(\mathbf{x}_j) = f^j(\mathbf{x}). \tag{91}$$

The learning problem in both the cases of static and dynamical systems f is the same: given the data pairs, approximate f for any input point. The dynamics does provide an advantage though, as data can be sampled along a trajectory advancing in time. Assume a discrete-time dynamical system T has an n-dimensional linear representation $(\mathbf{f}, A)$, such that

$$\mathbf{f}' = A\mathbf{f}. \tag{92}$$

If we take samples of $\mathbf{f}$ along a trajectory $\mathbf{f}(\mathbf{m}), \mathbf{f}(T\mathbf{m}), ..., \mathbf{f}(T^m\mathbf{m})$, obtaining a sequence of *snapshots*

$$\mathbf{f}_1 = \mathbf{f}(\mathbf{m}), \mathbf{f}(T\mathbf{m}) = \mathbf{f}_2, ..., \mathbf{f}(T^m\mathbf{m}) = \mathbf{f}_m, \tag{93}$$

we have

$$\mathbf{f}(T^k) = A\mathbf{f}(T^{k-1}), k = 1, ..., m. \tag{94}$$

We can form data matrices

$$X = \begin{bmatrix} f_1(\mathbf{m}) & f_1(T\mathbf{m}) & \cdots & f_1(T^{m-1}\mathbf{m}) \\ f_2(\mathbf{m}) & f_2(T\mathbf{m}) & \cdots & f_2(T^{m-1}\mathbf{m}) \\ \vdots & & & \\ f_n(\mathbf{m}) & f_n(T\mathbf{m}) & \cdots & f_n(T^{m-1}\mathbf{m}) \end{bmatrix} \tag{95}$$

and

$$X' = \begin{bmatrix} f_1(T\mathbf{m}) & f_1(T^2\mathbf{m}) & \cdots & f_1(T^m\mathbf{m}) \\ f_2(T\mathbf{m}) & f_2(T^2\mathbf{m}) & \cdots & f_2(T^m\mathbf{m}) \\ \vdots & & & \\ f_n(T\mathbf{m}) & f_n(T^2\mathbf{m}) & \cdots & f_n(T^m\mathbf{m}) \end{bmatrix}. \tag{96}$$

Note that each row j of data matrices X, X' is an evaluation of the function f_j on the trajectory of T starting at M. Setting

$$X' = XC, \tag{97}$$

we see that C is the companion matrix

$$\tilde{U} = C = \begin{pmatrix} 0 & 0 & ... & 0 & c_1 \\ 1 & 0 & ... & 0 & c_2 \\ 0 & 1 & ... & 0 & c_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & ... & 1 & c_m \end{pmatrix}, \tag{98}$$

where $\mathbf{c} = (c_1, ..., c_m) = \mathbf{c}_m$. The solution of this equation, provided $m = n$ and X is nonsingular, is

$$C = X^{-1}X'. \tag{99}$$

The matrix C would then be thought of as an approximation to the Koopman operator acting on the space $\mathcal{O}_m = \mathbb{R}^m$ of functions on the set of points $Tr = (\mathbf{m}, T\mathbf{m}, ..., T^m\mathbf{m})$. There are several caveats here. Typically we do not know in advance that a linear representation exists. Thus, we need to specify the dimension n and choose $f_1, ..., f_n$. In the case when $\mathbf{f}$ does not span an n-dimensional invariant subspace of U, we can as well identify C from (99), but we can have

$$f_j(T^m) \neq \sum_{k=1}^{m-1} c_k f_j(T^k\mathbf{m}). \tag{100}$$

In other words, it might be that the mth element of the Krylov sequence $f_j(\mathbf{m}), f_j(T\mathbf{m}), ...$ does not belong to the same subspace as the first $m - 1$ elements. In addition, all the functions $f_k, k = 1, ..., n$, would have the same relationship between the last element of their own Krylov sequence and the prior elements. If the trajectory Tr is periodic with period m, then $\mathbf{c}_m = (1, 0, ..., 0)$ provides an exact reduction of the Koopman operator to $\mathcal{O}_m$. In principle, for such trajectories the number of *snapshots* $\mathbf{f}_1, ..., \mathbf{f}_m$

can be smaller than the number of functions in each snapshot, and

$$C = X^+ X, \qquad (101)$$

where

$$X^+ = (X^\dagger X)^{-1} X^\dagger \qquad (102)$$

is the Moore-Penrose pseudoinverse of X, and $X^\dagger$ is the hermitian transpose (we allow for complex observables) of X. This was the reason behind the initial success utilizing the methodology of Dynamic Mode Decomposition [Sch10] to complex fluid flows in [RMB$^+$09], as the Koopman modes uncovered there were the consequence of the quasi-periodic nature of a portion of the underlying attractor [Mez05].

A more structured approach to finding linear representations is that of finite-section methods, the first version of which was Extended Dynamic Mode Decomposition (EDMD) [WKR15]. Consider the Koopman operator acting on an observable space $\mathcal{O}$ of functions on the state space M, equipped with the complex inner product $\langle \cdot, \cdot \rangle$,[7] and let $\mathbf{f} = \{f_j\}$, $j \in \mathbb{N}$, be an orthonormal basis on $\mathcal{O}$, such that, for any function $f \in \mathcal{O}$ we have

$$f = \sum_{j \in \mathbb{N}} c_j f_j. \qquad (103)$$

Let

$$u_{kj} = \langle U f_j, f_k \rangle. \qquad (104)$$

Then,

$$(Uf)_k = \langle Uf, f_k \rangle = \sum_{j \in \mathbb{N}} c_j \langle U f_j, f_k \rangle = \sum_{j \in \mathbb{N}} u_{kj} c_j. \qquad (105)$$

The basis functions do not necessarily need to be orthogonal. Consider the action of U on an individual, basis function f_j:

$$U f_j = \sum_{k \in \mathbb{N}} u_{kj} f_k, \qquad (106)$$

where u_{kj} are now just coefficients of $U f_j$ in the basis. We obtain

$$Uf = \sum_{j \in \mathbb{N}} c_j U f_j = \sum_{j \in \mathbb{N}} c_j \sum_{k \in \mathbb{N}} u_{kj} f_k = \sum_{k \in \mathbb{N}} \left(\sum_{j \in \mathbb{N}} u_{kj} c_j \right) f_k, \qquad (107)$$

and we again have

$$(Uf)_k = \sum_{j \in \mathbb{N}} u_{kj} c_j. \qquad (108)$$

As in the previous section, associated with any linear subspace $\mathcal{G}$ of $\mathcal{O}$, there is a projection onto it, denoted $P = P^2$, that we can think of as projection "along" the space $(I - P)\mathcal{O}$, since, for any $f \in \mathcal{O}$ we have

$$P(I - P)f = (P - P^2)f = 0, \qquad (109)$$

---

[7]*Note that we are using the complex inner product linear in the first argument here. The physics literature typically employs the so-called Dirac notation, where the inner product is linear in its second argument.*

and thus any element of $(I - P)\mathcal{O}$ has projection 0. We denote by $\tilde{U}$ the infinite-dimensional matrix with elements u_{kj}, $k, j \in \mathbb{N}$. Thus, the finite-dimensional section of the matrix

$$\tilde{U}_n = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & & u_{2n} \\ \vdots & & \ddots & \vdots \\ u_{n1} & u_{n2} & \cdots & u_{nn} \end{bmatrix} \qquad (110)$$

is the so-called compression of $\tilde{U}$ that satisfies

$$\tilde{U}_n = P_n \tilde{U} P_n, \qquad (111)$$

where P_n is the projection "along" $(I - P_n)\mathcal{O}$ to the span of the first n basis functions, span$(f_1, ..., f_n)$. To apply the finite-section methodology of approximation of the Koopman operator, we need to estimate coefficients u_{kj} from data.

In the case of a nonorthonormal basis, denote by $\hat{f}_k$ the dual basis vectors, such that

$$\langle f_j, \hat{f}_k \rangle = \delta_{jk}, \qquad (112)$$

where $\delta_{jj} = 1$ for any j, and $\delta_{jk} = 0$ if $j \neq k$.

For the infinite-dimensional Koopman matrix coefficients we get

$$u_{kj} = \langle U f_j, \hat{f}_k \rangle. \qquad (113)$$

Let's consider the *finite* set of independent functions $\tilde{\mathbf{f}} = \{f_1, ..., f_N\}$ and the associated dual set $\{\hat{g}_1, ..., \hat{g}_N\}$ in the span $\tilde{\mathcal{O}}$ of $\tilde{\mathbf{f}}$, that satisfy

$$\langle f_j, \hat{g}_k \rangle = \delta_{jk}. \qquad (114)$$

Under the ergodicity condition, in the case $\mathcal{O} = L^2(\mu)$ it is possible to obtain $\tilde{u}_{kj}$ from data.

Theorem 37. *Let $\{f_1, ..., f_N\}$ be a set of functions in $L^2(M, \mu)$, and let T be ergodic on M with respect to an invariant measure μ. Let $\mathbf{x}_l, l \in \mathbb{N}$, be a trajectory on M. Then, for almost any $\mathbf{x}_1 \in M$*

$$\tilde{u}_{kj} = \lim_{m \to \infty} \frac{1}{m} \sum_{l=1}^{m} f_j \circ T(\mathbf{x}_l) \hat{g}_k^c(\mathbf{x}_l)$$

$$= \lim_{m \to \infty} \frac{1}{m} \sum_{l=1}^{m} f_j(\mathbf{x}_{l+1}) \hat{g}_k^c(\mathbf{x}_l), \qquad (115)$$

where, for any finite m, $\hat{g}_k^c(\mathbf{x}_l), l = 1, ..., m$, are obtained as rows of the matrix $(F^\dagger F)^{-1} F^\dagger$, where

$$F = [f_1(\mathbf{X}) \, f_2(\mathbf{X}) \cdots f_N(\mathbf{X})], \qquad (116)$$

$F^\dagger = (F^c)^T$ *is the conjugate (Hermitian) transpose of F, and $f_j(\mathbf{X})$ is the column vector $(f_j(\mathbf{x}_1) \cdots f_j(\mathbf{x}_m))^T$. The limit in* (115) *exists due to Birkhoff's ergodic theorem.*

The above result is convenient as a single trajectory of an ergodic system can be used to estimate the inner product. But the formulation is restricted to measure-preserving systems. Alternatively, the above formula is valid in any case

where the function space is a Hilbert space, and inner product can be defined as a weighted sum over sample points $\mathbf{x}_l$,

$$\tilde{u}_{kj} = \lim_{m \to \infty} \frac{1}{m} \sum_{l=1}^{m} w(\mathbf{x}_l) f_j(\mathbf{x}_{l+1}) \hat{g}_k^c(\mathbf{x}_l). \tag{117}$$

Proposition 38. *Let (M, μ) be a measure space and $T : M \to M$. Let $\mathbf{f} = (f_{j_1}, ..., f_{j_n})$ be a subset of a basis in a Hilbert space of observables $\mathcal{O} = L^2(\mu)$. Let $u_{j_k l} = 0$ whenever $k \in 1, ..., n, l \notin \{j_1, ..., j_n\}$. Then T admits a finite-dimensional, linear representation $(\mathbf{f}, \tilde{U}_{\mathbf{f}})$, where $\tilde{U}_{\mathbf{f}}$ is the matrix which is restriction of $\tilde{U}$ to $\mathbf{f}$.*

Proof. The condition "$u_{j_k l} = 0$ whenever $k \in 1, ..., n, l \notin \{j_1, ..., j_n\}$" assures that

$$\mathbf{f} \circ T = A\mathbf{f}, \tag{118}$$

where $A = \tilde{U}_{\mathbf{f}}$. Namely, the time evolution of functions in $\mathbf{f}$ projected on any subspace that does not contain any of the functions in $\mathbf{f}$ is 0. □

It is interesting that the finite section method can be used to find nonlinear solutions $(\mathbf{f}, \mathbf{F})$ of the equation (68) and reveal nonlinear representations.

Proposition 39. *Let (M, μ) be a measure space and $T : M \to M$. Let $\mathbf{f} = (f_{j_1}, ..., f_{j_n})$ be a subset of a basis in a Hilbert space of observables $\mathcal{O} = L^2(\mu)$. Let $u_{j_k l} = 0$ whenever $k \in 1, ..., n$, $f_l \neq F(\mathbf{f})$ for some function F. Further assume that there are $l_k \in \mathbb{N}, k = 1, ..., K$, such that $f_{l_k} = F_k(\mathbf{f})$, where F_k is nonlinear. Then T admits a finite-dimensional, nonlinear representation $(\mathbf{f}, \mathbf{F})$, where $\mathbf{F}$ is given by*

$$\mathbf{F}(\mathbf{f}) = A\tilde{\mathbf{f}}, \tag{119}$$

where $\mathbf{f} \subset \tilde{\mathbf{f}}$ and A is an $n \times (n + K)$ matrix.

Proof. The condition "$u_{j_k l} = 0$ whenever $k \in 1, ..., n$, $f_l \neq F(\mathbf{f})$" assures that the time evolution of $\mathbf{f}$ under T is a (nonlinear) function F of $\mathbf{f}$ only, and $\mathbf{F}(\mathbf{f}) = A\tilde{\mathbf{f}}$, because $f_{j_k} \circ T = \sum_{m \in \{j_1, ..., j_n, l_1, ..., l_K\}} u_{j_k m} f_m$. □

The key to the DMD-type approximations to the Koopman operator are the predetermined basis functions. In the case of EDMD [WKR15], these are selected a priori, and in the case of Hankel-DMD [AM17] they are generated using an initial choice of observables supplemented by time-delayed observables generated by dynamics. While EDMD generally (exceptions are kernel-based methods) suffers from the curse of dimensionality, Hankel-DMD does not, as in any dimension the generated functions fill up an invariant subspace of the Koopman operator. However, there is no guarantee that there is a (linear or nonlinear) finite representation amongst the observables in either case. The framework of Generalized Laplace Analysis [MM14] solves that problem by computing the spectrum and then computing the eigenfunctions by weighted time averages.

As we have seen, eigenfunctions provide us with linear representations. The deep neural network formalism has been used to compute linear representations, where both the observables and the eigenmatrix A are learned [LDBK17, YKH19, LKB18, TKY17].

The neural network formulation for the solution of the (nonlinear) representation eigenproblem (68) is

$$\min_{(\theta, \psi)} \sum_{j=1}^{m} \|\mathbf{n}_{\mathcal{O}}(\mathbf{m}_{j+1}, \theta) - \mathbf{n}_{\mathbf{F}}(\mathbf{n}_{\mathcal{O}}(\mathbf{m}_j, \theta), \psi)\|, \tag{120}$$

where $\mathbf{n}_{\mathcal{O}}(\mathbf{m}_{j+1}, \theta)$ is the neural network representing the observables, with parameters θ, and $\mathbf{n}_{\mathbf{F}}(\mathbf{n}, \psi)$ is the neural network representing the eigenmap, with parameters ψ. The dimension of the vector $\mathbf{n}$ (and thus the dimension of $\mathbf{F}$) is a hyperparameter.

Extensions

"Static" composition operator. Instead of $T : M \to M$ whose iterations produce dynamical trajectories, we now consider a "static" map T between different spaces $T : M \to N$, where no notion of iteration of the map is available. The set $\mathcal{O}_M$ of all complex functions $f : M \to \mathbb{C}$ is the space of observables on M, while the set $\mathcal{O}_N$ of all complex functions $g : N \to \mathbb{C}$ is the space of observables on N. Both are linear vector spaces over the field of complex numbers. A finite-dimensional representation $(\mathbf{f}, \mathbf{g}, \mathbf{F})$ of T is a double set of functions $\mathbf{f} = (f_1, ..., f_n), \mathbf{g} = (g_1, ..., g_m)$ and a mapping $\mathbf{F}$ such that

$$\mathbf{g}(T\mathbf{m}) = \mathbf{F}(\mathbf{f}(\mathbf{m})), \tag{121}$$

where $\mathbf{F} : \mathbb{C}^n \to \mathbb{C}^m$ and (n, m) is the dimension of the representation. We again have the notion of a *faithful* representation.

Definition 40. *A representation $(\mathbf{f}, \mathbf{g}, \mathbf{F})$ of $T : M \to N$ is called faithful provided $\mathbf{f} : M \to \mathbf{f}(M), \mathbf{g} : N \to \mathbf{g}(N)$ are injective.*

The previous notion of representation is recovered when $M = N$ and $\mathbf{f} = \mathbf{g}$.

Any map $T : M \to N$ defines the *pullback* composition operator $U : \mathcal{O}_N \to \mathcal{O}_M$ by

$$Ug(\mathbf{m}) = g(T\mathbf{m}). \tag{122}$$

This is the linear composition operator associated with $T : M \to N$. The image $\mathcal{I}(U)$ of U is the set of functions $f : \mathcal{O}_M \to \mathbb{C}$ that are constant on level sets $L_{\mathbf{n}}$ of T:

$$L_{\mathbf{n}} = \{\mathbf{m} \in M | \mathbf{m} \in T^{-1}(\mathbf{n})\}. \tag{123}$$

The space $\mathcal{I}(U)$ is a linear subspace of $\mathcal{O}_M$.

Provided $\mathcal{O}_M$ is a separable Hilbert space and $\mathcal{I}(U)$ is closed, we can also define the pushforward operator $P : \mathcal{O}_M \to \mathcal{O}_N$ by

$$P(f) = Id(\Pi f) \circ T, \tag{124}$$

where $\Pi : \mathcal{O}_M \to \mathcal{J}(U)$ is the orthogonal projection operator onto $\mathcal{J}(U)$, and $Id : \mathcal{J}(U) \to \mathcal{O}(N)$ is the operator that identifies functions in $\mathcal{J}(U)$ with those in $\mathcal{O}(N)$. The singular value decomposition is valid for bounded operators between separable Hilbert spaces and thus we have the following characterization of $\mathcal{J}(U)$.

Proposition 41. *Let $\mathcal{O}_N, \mathcal{O}_M$ be separable Hilbert spaces, and let $\mathcal{J}(U)$ be closed. Then the space $\mathcal{J}(U)$ is orthogonal to the subspace at the singular value 0 of P. In addition, U is the pseudoinverse of P.*

Proof. The kernel of P, consisting of functions orthogonal to $\mathcal{J}(U)$, is the subspace of $\mathcal{O}_M$ corresponding to the singular value 0 of P. We also have

$$UPf = \Pi f, \tag{125}$$

proving that U is the pseudoinverse of P. $\qquad\square$

With a little bit of topological and measure-theoretic infrastructure, we can characterize the projection operator Π.

Theorem 42. *Let (M, N) be two Radon spaces—separable metric spaces on which every probability measure is a Radon measure. Assume that M is endowed with a Borel measure μ and T is a measurable map. Let N be endowed with the measure $\nu(A) = \mu(T^{-1}(A))$. Then*

$$\Pi(f) = \mathbb{E}(f|\mathcal{B}_{T^{-1}}), \tag{126}$$

where $\mathbb{E}(f|\mathcal{B}_{T^{-1}})$, the conditional expectation of f with respect to the sigma algebra induced by T, is the orthogonal projection of f on $\mathcal{J}(U)$.

Proof. First observe that $\Pi(f)$ as defined in (126) is in $\mathcal{J}(U)$ and defines a projection, since applying conditional expectation twice yields the same result as applying it once. We need to prove that $f - \mathbb{E}(f|\mathcal{B}_{T^{-1}})$ is orthogonal to $\mathcal{J}(U)$, i.e.,

$$\int_M h^c(f - \mathbb{E}(f|\mathcal{B}_{T^{-1}}))d\mu = 0, \tag{127}$$

where $h \in \mathcal{J}(U)$. Any function in $\mathcal{J}(U)$ is constant on level sets of T, and by the disintegration of measure theorem

$$\int_M h^c(f - \mathbb{E}(f|\mathcal{B}_{T^{-1}}))d\mu$$
$$= \int_M \int_{T^{-1}(\mathbf{n})} h^c(f - \mathbb{E}(f|\mathcal{B}_{T^{-1}}))d\mu_{\mathbf{n}}d\nu(\mathbf{n})$$
$$= \int_M h^c \left(\int_{T^{-1}(\mathbf{n})} (f - \mathbb{E}(f|\mathcal{B}_{T^{-1}}))d\mu_{\mathbf{n}} \right) d\nu(\mathbf{n})$$
$$= 0, \tag{128}$$

since

$$\int_{T^{-1}(\mathbf{n})} f d\mu_{\mathbf{n}} = \mathbb{E}(f|\mathcal{B}_{T^{-1}}). \tag{129}$$

$\qquad\square$

The above two results constitute the beginning of the analysis of nonlinear maps between different spaces using an operator-theoretic approach that is similar to the one deployed in dynamical systems using Koopman operators. Now we have the following proposition.

Proposition 43. *The set of functions $(\mathbf{f}, \mathbf{g}, \mathbf{F}) : M \times N \times \mathbb{C}^k \to \mathbb{C}^k \times \mathbb{C}^l \times \mathbb{C}^l$ is a finite-dimensional representation of T iff*

$$U\mathbf{g}(\mathbf{m}) = \mathbf{F}(\mathbf{f}(\mathbf{m})). \tag{130}$$

The representation is linear provided $\mathbf{F}$ is a finite-dimensional $k \times l$ matrix:

$$U\mathbf{g}(\mathbf{m}) = A\mathbf{f}(\mathbf{m}). \tag{131}$$

To get an approximation to a finite-dimensional linear representation, we may select bases $(f_1, ..., f_m, ...)$ on $\mathcal{O}_M$ and $(g_1, ..., g_n, ...)$ on $\mathcal{O}_N$, and construct the representation of P. We assume we have access to N realizations of data pairs corresponding to $(\mathbf{m}_j, \mathbf{n}_j = T(\mathbf{m}_j)), j = 1, ..., N$. The data points are $\mathbf{f}_j = (f_1(\mathbf{m}_j), ..., f_m(\mathbf{m}_j))^T$ and $\mathbf{g}_j = (g_1(\mathbf{n}_j), ..., g_n(\mathbf{n}_j))^T$. We form matrices

$$X = [\mathbf{f}(\mathbf{m}_1) \cdots \mathbf{f}(\mathbf{m}_N)],$$
$$Y = [\mathbf{g}(\mathbf{n}_1) \cdots \mathbf{g}(\mathbf{n}_N)]. \tag{132}$$

The solution to

$$\min_A ||Y - AX|| \tag{133}$$

is

$$A = YX^+, \tag{134}$$

where X^+ is the pseudoinverse of X.

Relationship to Mori-Zwanzig formalism. The Mori-Zwanzig formalism describes evolution of a subset of observables $\mathbf{f} \in \mathcal{O}$, where $\mathcal{O}$ is a Hilbert space, using the Koopman operator and orthogonal projection P on the subspace spanned by $\mathbf{f} = \{f_1, ..., f_n\}$. Using P and $Q = I - P$ we get

$$U^t\mathbf{f} = (P + Q)U^t\mathbf{f} = \tilde{U}_n^t\mathbf{f} + QU^t\mathbf{f}, \tag{135}$$

where $\tilde{U}_n^t$ is the finite section matrix (110), and $Q_n U^t\mathbf{f}$ is the projection of the evolution of $\mathbf{f}$ on the space orthogonal to the span of $\mathbf{f}$ in $\mathcal{O}$. It is immediately clear that, provided $(\mathbf{f}, \tilde{U}_n^t)$ is a linear representation, $QU^t\mathbf{f}$ is zero, and thus we get

$$U^t\mathbf{f} = \tilde{U}_n^t\mathbf{f}. \tag{136}$$

Assuming U_n^t is diagonalizable, we get

$$U^t\mathbf{f} = \Phi D^t \Phi^{-1}\mathbf{f}, \tag{137}$$

where D^t is a diagonal matrix containing $\lambda_j^t, j = 1, ..., n$, on the diagonal, where λ_j is an eigenvalue of $\tilde{U}_n^t$.

In discrete time, the evolution reads

$$U^2 = (P + Q)U(PU + QU)$$
$$= (PU)^2 + PUQU + QUPU + (QU)^2$$

and, by induction,

$$
\begin{aligned}
U^n &= (PU)^n \\
&\quad + \sum_{\substack{k_1,\dots,k_n=0 \\ \sum_j k_j = n}}^{n-1} (PU^{k_1}QU^{k_2}\cdots PU^{k_1}QU^{k_2}) \\
&\quad + (QU)^n \\
&= \tilde{U}_n \\
&\quad + \sum_{\substack{k_1,\dots,k_n=0 \\ \sum_j k_j = n}}^{n-1} [(PU)^{k_1}(QU)^{k_2}\cdots(PU)^{k_{n-1}}(QU)^{k_n}] \\
&\quad + (QU)^n .
\end{aligned}
$$

The second term is often interpreted as the "memory term" but in fact it contains a total of n applications of U just like the first and the last term. Rather, it describes the part of the evolution that depends on the evolution in both the span of $\mathbf{f}$ and its orthogonal complement. The following result clarifies the point, and identifies the situation in which the evolution in projected variables is Markovian.

Proposition 44. *If the evolution in the orthogonal complement of $\mathbf{f}$ is dependent on $\mathbf{f}$ only, but is not 0, i.e., $QU\mathbf{f} \neq 0$, then T admits a nonlinear representation $(\mathbf{f}, \mathbf{F})$.*

Proof. Since $QU\mathbf{f} = \mathbf{G}(\mathbf{f})$ for some $\mathbf{G} : \mathbb{C}^n \to \mathbb{C}^n$ and $PU\mathbf{f} = \tilde{U}_n\mathbf{f}$, then

$$
U\mathbf{f} = PU\mathbf{f} + QU\mathbf{f} = \tilde{U}_n\mathbf{f} + \mathbf{G}(\mathbf{f}) = \mathbf{F}(\mathbf{f}), \qquad (138)
$$

where $\mathbf{F} : \mathbb{C}^n \to \mathbb{C}^n$. □

The following example of the result in the above proposition also indicates the perils of modeling the $(QU)^n$ term in (138) as noise as it is commonly done in Mori-Zwanzig literature.

Example 45. Consider the irrational circle rotation $T : S^1 \to S^1$ defined by $z' = e^{i\omega}z$, where $z = e^{i\theta}$ and $\omega/2\pi$ is irrational. This is an ergodic system on S^1. We denote the complex inner product with respect to Haar measure μ on S^1 by $\langle \cdot, \cdot \rangle$. Consider an analytic $L^2(\mu)$ observable $f : S^1 \to \mathbb{C}$ that separates points on S^1, namely $f(z_1) = f(z_2) \Rightarrow z_1 = z_2$. The Taylor expansion of f is given by

$$
f = \sum_{n=0}^{\infty} c_n z^n, \qquad (139)
$$

where $c_n \in \mathbb{C}$. We have

$$
Uf = f' = \sum_{n=0}^{\infty} c_n e^{in\omega} z^n. \qquad (140)
$$

Denoting the complex conjugate by $(\cdot)^c$, we define the

complex scalar λ by

$$
\lambda = \langle f', f \rangle = \sum_{n=0}^{\infty} c_n c_n^c e^{in\omega} \langle z^n, (z^n)^c \rangle = \sum_{n=0}^{\infty} |c_n|^2 e^{in\omega}. \qquad (141)
$$

We have

$$
f' = \lambda f + QUf. \qquad (142)
$$

Note that

$$
QUf = \sum_{n=0}^{\infty} c_n(1-\lambda)e^{in\omega}z^n = (1-\lambda)Uf. \qquad (143)
$$

Since f separates and T is a bijection, $f(z_1) = f(z_2) \Rightarrow z_1 = z_2 \Rightarrow Uf(z_1) = Uf(z_2)$, and thus $QUf = (1-\lambda)Uf = G(f)$, where $G : f(S^1) \to \mathbb{C}$. Thus,

$$
f' = \lambda f + Gf \qquad (144)
$$

is Markovian, i.e., contains no memory terms, and is "closed," i.e., the term QUf should not be modeled as noise. The result can hold even if f does not separate points on S^1. Namely, it is easy to see that for any n, $f = z^n$ leads to $c_n = 1, c_{j\neq n} = 0, \lambda = e^{in\omega}, G(f) = 0$ and the evolution reads

$$
f' = e^{in\omega}f, \qquad (145)
$$

reflecting the fact that f is an eigenfunction of U with eigenvalue $e^{in\omega}$.

The main result (144) in the above example is also true for dynamical systems in Hilbert functional spaces in which the Koopman operator has point spectrum [Mez20], and there is a set of separating *principal eigenfunctions* [MM14] of U or U^t. Provided the set of observables $\mathbf{f}$ separates, we can use the algebraic property of eigenfunctions (47); then the evolution of $\mathbf{f}$ under U reads

$$
\mathbf{f}' = A\mathbf{f} + \mathbf{G}(\mathbf{f}), \qquad (146)
$$

and there are no noise and memory terms.

From the foregoing analysis it becomes clear that the noise and memory terms in the Mori-Zwanzig framework arise due to (1) the fact that the chosen representation $\mathbf{f}$ is not faithful and (2) the spectrum of the Koopman operator associated with a dynamical system has a continuous part.

Conclusions and Futures

We provided a framework for learning of dynamical systems rooted in the concept of representations and Koopman operators. The interplay between the two leads to the full description of systems that can be represented linearly in a finite dimension, based on the properties of Koopman operator spectrum. This is achieved using eigenfunctions, in the case when the operator spectrum is discrete. A reduced set of eigenfunctions can be viewed as an autoencoder. As shown here, while systems with continuous spectrum do not admit finite-dimensional linear representations, they might be represented in the finite-dimensional

nonlinear setting, within an infinite-dimensional invariant subspace of the operator. Thus, even nonlinear representations can be learned using the Koopman operator framework, and functional approximations using neural networks might be well suited for this task, encoded in the *representation eigenproblem*.

The essential difference in the type of learning happening in our brains and the type of supervised learning dominant in machine learning is the notion of time. Time is also at the core of understanding causal relationships. Namely, without time only correlation between observables is possible. The Koopman operator theory remedies this by explicitly taking time into account and providing this dimension of learning with the explicit mathematical structure. Moreover, based on techniques such as Generalized Laplace Analysis, that naturally yield themselves to adjustments using streaming data, unsupervised learning, leading to generative models, is achievable, where future data is adapted organically into the learned structure. The approach thus provides a suitable setting for unsupervised learning of dynamical systems.

References

[AM17] Hassan Arbabi and Igor Mezić, *Ergodic theory, dynamic mode decomposition, and computation of spectral properties of the Koopman operator*, SIAM J. Appl. Dyn. Syst. **16** (2017), no. 4, 2096–2126, DOI 10.1137/17M1125236. MR3720364

[Gia21] Dimitrios Giannakis, *Delay-coordinate maps, coherence, and approximate spectra of evolution operators*, Res. Math. Sci. **8** (2021), no. 1, Paper No. 8, 33, DOI 10.1007/s40687-020-00239-y. MR4208217

[Koo31] Bernard O. Koopman, *Hamiltonian systems and transformation in Hilbert space*, Proc. Natl. Acad. Sci. USA **17** (1931), no. 5, 315.

[KPM20] Milan Korda, Mihai Putinar, and Igor Mezić, *Data-driven spectral analysis of the Koopman operator*, Appl. Comput. Harmon. Anal. **48** (2020), no. 2, 599–629, DOI 10.1016/j.acha.2018.08.002. MR4047538

[LDBK17] Qianxiao Li, Felix Dietrich, Erik M. Bollt, and Ioannis G. Kevrekidis, *Extended dynamic mode decomposition with dictionary learning: a data-driven adaptive spectral decomposition of the Koopman operator*, Chaos **27** (2017), no. 10, 103111, 10, DOI 10.1063/1.4993854. MR3709448

[LKB18] Bethany Lusch, J. Nathan Kutz, and Steven L. Brunton, *Deep learning for universal linear embeddings of nonlinear dynamics*, Nature Communications **9** (2018), no. 1, 1–10.

[LM15] Zoran Levnajić and Igor Mezić, *Ergodic theory and visualization. II. Fourier mesochronic plots visualize (quasi)periodic sets*, Chaos **25** (2015), no. 5, 053105, 12, DOI 10.1063/1.4919767. MR3389869

[MB04] Igor Mezić and Andrzej Banaszuk, *Comparison of systems with complex behavior*, Phys. D **197** (2004), no. 1-2, 101–133, DOI 10.1016/j.physd.2004.06.015. MR2092096

[Mez05] Igor Mezić, *Spectral properties of dynamical systems, model reduction and decompositions*, Nonlinear Dynam. **41** (2005), no. 1-3, 309–325, DOI 10.1007/s11071-005-2824-x. MR2157184

[Mez20] Igor Mezić, *Spectrum of the Koopman operator, spectral expansions in functional spaces, and state-space geometry*, J. Nonlinear Sci. **30** (2020), no. 5, 2091–2145, DOI 10.1007/s00332-019-09598-5. MR4163461

[Mez94] Igor Mezić, *On the geometrical and statistical properties of dynamical systems: Theory and applications*, ProQuest LLC, Ann Arbor, MI, 1994. Thesis (Ph.D.)–California Institute of Technology. MR2691109

[MM12] Alexandre Mauroy and Igor Mezić, *On the use of Fourier averages to compute the global isochrons of (quasi)periodic dynamics*, Chaos **22** (2012), no. 3, 033112, 9, DOI 10.1063/1.4736859. MR3388606

[MM14] Ryan Mohr and Igor Mezić, *Construction of eigenfunctions for scalar-type operators via laplace averages with connections to the Koopman operator*, arXiv:1403.6559 (2014).

[MMM13] Alexandre Mauroy, Igor Mezić, and Jeff Moehlis, *Isostables, isochrons, and Koopman spectrum for the action-angle representation of stable fixed point dynamics*, Phys. D **261** (2013), 19–30, DOI 10.1016/j.physd.2013.06.004. MR3144005

[MMS20] Alexandre Mauroy, Igor Mezić, and Yoshihiko Susuki, *Koopman operator in systems and control*, Springer, 2020.

[RMB+09] Clarence W. Rowley, Igor Mezić, Shervin Bagheri, Philipp Schlatter, and Dan S. Henningson, *Spectral analysis of nonlinear flows*, J. Fluid Mech. **641** (2009), 115–127, DOI 10.1017/S0022112009992059. MR2577895

[Sch10] Peter J. Schmid, *Dynamic mode decomposition of numerical and experimental data*, J. Fluid Mech. **656** (2010), 5–28, DOI 10.1017/S0022112010001217. MR2669948

[TKY17] Naoya Takeishi, Yoshinobu Kawahara, and Takehisa Yairi, *Learning Koopman invariant subspaces for dynamic mode decomposition*, Advances in Neural Information Processing Systems, 2017, pp. 1130–1140.

[WKR15] Matthew O. Williams, Ioannis G. Kevrekidis, and Clarence W. Rowley, *A data-driven approximation of the Koopman operator: extending dynamic mode decomposition*, J. Nonlinear Sci. **25** (2015), no. 6, 1307–1346, DOI 10.1007/s00332-015-9258-5. MR3415049

[YKH19] Enoch Yeung, Soumya Kundu, and Nathan Hodas, *Learning deep neural network representations for Koopman operators of nonlinear dynamical systems*, 2019 American Control Conference (ACC), 2019, pp. 4832–4839.

Igor Mezić

Credits

Opening image and Figure 1 are reproduced from Zoran Levnajić and Igor Mezić, *Ergodic theory and visualization. II. Fourier mesochronic plots visualize (quasi)periodic sets*, Chaos 25 (2015), no. 5, with permission of AIP Publishing.

Figure 2 is courtesy of Igor Mezić. Originally appeared in [Mez20].

Figure 3 is reprinted from Milan Korda, Mihai Putinar, and Igor Mezić, *Data-driven spectral analysis of the Koopman operator*, Appl. Comput. Harmon. Anal. 48 (2020), no. 2, 599–629, with permission from Elsevier.

Photo of Igor Mezić is courtesy of Matt Perko, UC Santa Barbara.