## Latent Modes of Nonlinear Flows - a Koopman Theory Analysis

Ido Cohen & Guy Gilboa

DOI: 10.xxxx/xxxxxxx (do not change) First published online: MMM dd YYYY (do not change)

Ido Cohen Faculty of Electrical And Computer Engineering - Technion

Guy Gilboa Faculty of Electrical And Computer Engineering - Technion

**Abstract:** Extracting the latent underlying structures of complex nonlinear local and nonlocal flows is essential for their analysis and modeling. In this work we attempt to provide a consistent framework through Koopman theory and its related popular discrete approximation – dynamic mode decomposition (DMD). We investigate the conditions to perform appropriate linearization, dimensionality reduction and representation of flows in a highly general setting.

The essential elements of this framework are *Koopman Eigenfunction* (KEF), for which existence conditions are formulated. This is done by viewing the dynamic as a curve in state-space. These conditions lay the foundations for system reconstruction, global controllability, and observability for nonlinear dynamics.

We examine the limitations of DMD through the analysis of Koopman theory and propose a new mode decomposition technique based on the typical time profile of the dynamics. An overcomplete dictionary of decay profiles is used to sparsely approximate the flow. This analysis is also valid in the full continuous setting of Koopman theory, which is based on variational calculus. We demonstrate applications of this analysis, such as finding KEFs and their multiplicities, dynamics reconstruction and global linearization.

**Keywords:** nonlinear decomposition, dynamic mode decomposition, homogeneous operators, gradient flows, nonlinear spectral theory, Koopman eigenfunctions, Koopman mode decomposition,

© Ido Cohen, Guy Gilboa, 2021

ISBNs: xxxxxxxxx(PB) xxxxxxxx(OC) ISSNs: xxxx-xxxx (online) xxxx-xxxx (print)

# Contents

List of abbreviations		1
1	Introduction	1
2	Preliminary	3
3	Motivation for this work	9
4	Koopman Eigenfunctions and Modes	11
5	Koopman Theory for PDE	28
6	Mode Decomposition based on Time State-Space Mapping	31
7	Examples	36
8	Conclusion	40
List of Symbols		41
	References	48

## List of abbreviations

DMD	Dynamic Mode Decomposition
EDMD	Extended DMD
KDMD	kernel DMD
KEF	Koopman Eigenfunction
KEFal	Koopman Eigenfunctional
KMD	Koopman Mode Decomposition
PDE	Partial Differential Equation
ROA	Region of Attraction
S-DMD	Symmetric DMD
SVD	Singular Value Decomposition
TV	Total Variation

## 1 Introduction

Knowing the latent space of certain data allows one to represent it concisely and to differentiate between signal and clutter parts. Recovering this space in a data-driven manner is a long-standing research problem. Data resulting from dynamical systems is represented commonly as spatial structures (modes) that are attenuated or enhanced with time. A common technique in linear flows is *separation of variables*. It is assumed that a solution u(x, t) of a linear flow can be expressed as,

$$u(x,t) = X(x)T(t).$$
 (1.1)

That is, the solution is a multiplication of a function of the spatial variable x and a function of the temporal variable t. In this study we examine, from various angles, the following paradigm: a *nonlinear flow* can be well approximated (or even exactly expressed) by a linear combination of variable separated functions,

$$u(x,t) \approx \sum_{i=1}^{m} X_i(x) T_i(t).$$
 (1.2)

In this context, the spatial structures  $X_i$  are referred to as *modes* and  $T_i$  are *time-profiles*. For such an approximation, if the error is negligible and *m* is small, we obtain a significant simplification of the system. This enables better understanding and modeling, allowing accurate interpolation and prediction of the dynamics.

The theory of Koopman argues that for many nonlinear systems data measurements evolve as if the dynamical system is linear (in some infinite dimensional space). A well-known algorithm to approximate these measurements is *Dynamic Mode Decomposition* (DMD) of Schmid (2010). In this work, we formulate sufficient and necessary conditions for the existence of these measurements. These findings highlight certain flaws of DMD. Finally, we suggest a new mode decomposition to overcome some of these problems, originated in an algorithm for general spectral decomposition of Gilboa (2018).

In many dynamical processes, there are measurements of the observations that evolve linearly, or approximately so, see Otto and Rowley (2021). A theoretical justification for that can be traced back to the seminal work of Koopman (1931). These measurements are referred to as *Koopman Eigenfunctions* (KEFs). An algorithm was proposed by Mezić (2005), *Koopman Mode Decomposition* (KMD), to reconstruct the dynamics using spatial structures, termed as modes, which are the coefficients of Koopman eigenfunctions. Since KEFs evolve as if they were observations in a linear dynamical system, KMD can interpret the original dynamics as a linear one.

This decomposition might be infinite-dimensional. In Schmid (2010) DMD it was suggested to approximate KMD in a finite domain. If the KEFs measurements are linear combinations of the observations then DMD yields the Koopman mode accurately. As noted in Kutz et al. (2016a), DMD can be interpreted as an exponential data fitting algorithm. In the more general nonlinear case, DMD may not reveal well the underlying modes and the dynamics.

Recently the authors and colleagues have formalized this insight in Cohen et al. (2021a), in the context of homogeneous flows, referring to it as the DMD paradox. As the step-size approaches zero, dynamic reconstruction with DMD results in positive mean squared error, but, paradoxically, with zero DMD error. In general, this paradox exists in any dynamical system whose KEFs are not linear combinations of the observations. This phenomenon becomes extreme when the system is zero homogeneous, as shown in Cohen et al. (2021b). Such cases are common in gradient flows of one-homogeneous functionals, such as local or nonlocal TV-flows, Andreu et al. (2001), Gilboa and Osher (2009). In that case, the dynamics is only in  $C^0$  almost everywhere and exponential decay is a very crude and inaccurate approximation. For such flows, lifting the observations to a finite higher dimensional space does not solve the problem (see for example Korda and Mezić (2018); Williams et al. (2015a)).

This alleged contradiction between KMD and DMD leads us to examine the fundamentals of Koopman theory. We follow the general solution of a KEF with respect to time and analyze the mapping between the state-space and the time variable. The existence of this mapping depends on the smoothness properties of the dynamics.

As a direct result, we introduce a new method that overcomes the DMD

limitations for smoothing-type processes. These findings, with some adaptations, are valid in the full continuous settings, as discussed by Kutz et al. (2016b); Mauroy (2021).

#### **Main Contributions**

We formulate the conditions for the existence of a KEF. If it exists, there is an infinite set of KEFs. We distinguish between different types of eigenfunction groups and analyze their multiplicity. We show that certain multiplicities are crucial to obtain dynamics reconstruction, controllability, and observability (Section 4). These conclusions are extended to the full continuous setting. Conditions for the existence of *Koopman Eigenfunctionals* (KEFals) are presented (Section 5). Following these insights, we suggest an alternative algorithm for finding Koopman modes induced by fitting time profiles that best characterize the dynamics. This algorithm overcomes some inherent limitations of DMD (Section 6). We attempt to bridge between nonlinear spectral decomposition and KMD. Specifically, we show that spectral *Total Variation* (TV) of Gilboa (2014) and its generalizations yield Koopman modes. Throughout this work, we illustrate the theory with simple toy examples. Additional examples and experiments are given in Section 7. In the following section, we provide the essential definitions and notations.

### 2 Preliminary

In this section, we present some background on Koopman operators, its eigenfunctions and eigenfunctionals and the related DMD framework. We note certain properties of variational calculus which are relevant to Section 5. In addition, we outline the work of Gilboa (2018) and Katzir (2017), where nonlinear flows are decomposed through a dictionary of decay profiles. We adapt this method for the extraction of Koompan modes in Section 6.

### 2.1 Koopman theory

#### 2.1.1 Discrete spatial setting

We consider a dynamical system in a semi-discrete setting, expressed as,

$$\frac{d}{dt}\mathbf{x}(t) = P(\mathbf{x}(t)), \quad \mathbf{x}(0) = \mathbf{x}_0, \quad t \in I,$$
(2.1)

where  $\mathbf{x} \in \mathbb{R}^N$  is a state vector,  $P : \mathbb{R}^N \to \mathbb{R}^N$  is a (nonlinear) operator, and  $I = [a, b] \subseteq \mathbb{R}^+$  is a time interval. Let  $g : \mathbb{R}^N \to \mathbb{R}$  be a measurement of  $\mathbf{x}$ . The Koopman operator  $K_P^{\tau}$  is a linear operator that acts on the infinite-dimensional space of measurements  $g(\mathbf{x})$  of the state, defined by Koopman (1931); Mezić (2005),

$$K_P^{\tau}(g(\boldsymbol{x}(s))) = g(\boldsymbol{x}(s+\tau)), \quad s, s+\tau \in I,$$
(KO)

where  $\tau > 0$ . The Koopman operator is linear, namely it admits,

$$K_P^{\tau}(\alpha g(\boldsymbol{x}(s)) + \beta f(\boldsymbol{x}(s))) = \alpha K_P^{\tau}(g(\boldsymbol{x}(s))) + \beta K_P^{\tau}(f(\boldsymbol{x}(s))),$$

for all measurements g and f and for all constants  $\alpha$  and  $\beta$ . In addition, the Koopman operators  $\{K_P^{\tau}\}_{\tau \ge 0}$  admits a semigroup property, more formally,

$$K_P^{\tau_2} \circ K_P^{\tau_1} = K_P^{\tau_1 + \tau_2},$$

where  $\circ$  denotes the composition operator. An eigenfunction of the Koopman operator,  $\varphi(\mathbf{x})$ , admits,

$$K_P^{\tau}(\varphi(\boldsymbol{x}(s))) = \varphi(\boldsymbol{x}(s+\tau)) = \eta^{\tau}\varphi(\boldsymbol{x}(s)), \quad s, s+\tau \in I,$$
(2.2)

for some  $\eta \in \mathbb{C}$ . Due to the semigroup attribute of the Koopman operator, the following limit exists,

$$\lim_{\tau \to 0} \frac{K_P^{\tau}(\varphi(\mathbf{x}(s))) - \varphi(\mathbf{x}(s))}{\tau} = \lim_{\tau \to 0} \frac{\varphi(\mathbf{x}(s+\tau)) - \varphi(\mathbf{x}(s))}{\tau} = \frac{d}{dt}\varphi(\mathbf{x}(t)) \Big|_{t=s}.$$
(2.3)

This limit can be explained by the relations of the Koopman operator and Lie derivatives, see Brunton et al. (2021). It can be shown (see for instance Mauroy et al. (2020), p. 10) that a KEF admits,

$$\frac{d}{dt}\varphi(\boldsymbol{x}(t)) = \lambda \cdot \varphi(\boldsymbol{x}(t)), \quad \forall t \in I,$$
(2.4)

for some  $\lambda \in \mathbb{C}$ . The relation between  $\eta$  in Eq. (2.2) and  $\lambda$  in Eq. (2.4) is  $\eta = e^{\lambda}$ . The solution of this linear ODE is given by,

$$\varphi(\mathbf{x}(t)) = \varphi(\mathbf{x}(a))e^{\lambda t}, \quad \forall t \in I.$$
 (KEF)

#### Koopman Mode Decomposition

Koopman Mode Decomposition (KMD) is a spatiotemporal mode decomposition of dynamical systems based on KEFs. Namely, the state space x can be expressed as (Mezić (2005)),

$$\mathbf{x}(t) = \sum_{i=1}^{\infty} \mathbf{v}_i \varphi_i(t), \qquad (2.5)$$

where  $\varphi_i(t)$  is a KEF and  $v_i$  is the corresponding vector, referred to as Koopman mode. When the dynamic is nonlinear the decomposition may have infinite elements. In practice, a finite approximation method is used. The most common one is DMD, as explained in Section 2.2.

#### 2.1.2 Full continuous setting

Let  $u : L \subset \mathbb{R} \times I \subseteq \mathbb{R}^+$  be the solution of the following PDE,

$$u_t(x,t) = \mathcal{P}(u(x,t)), \quad u(x,0) = f(x).$$
 (2.6)

We assume that *u* belongs to a Hilbert space  $\mathcal{H}$  with an inner product,  $\langle v, u \rangle$ and its associated norm  $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$ ,  $\mathcal{P} : \mathcal{H} \to \mathcal{H}$  is a (nonlinear) operator. Let  $Q : \mathcal{H} \to \mathbb{R}$  be a proper, lower-semicontinuous functional. The Koopman operator,  $K_{\mathcal{P}}^t$ , in the sense of PDE, is defined by Nakao and Mezić (2020),

$$K^{\tau}_{\varphi}(Q(u(x,s))) = Q(u(x,s+\tau)), \quad s, s+\tau \in I.$$
(2.7)

An eigenfunctional,  $\phi$ , of the Koopman operator is a functional admitting the following,

$$K_{\mathcal{P}}^{\tau}(\phi(u(x,s))) = \phi(u(x,s+\tau)) = \eta^{\tau}\phi(u(x,s)), \quad s, s+\tau \in I.$$
(2.8)

By letting  $\tau \to 0$  an eigenfunctional of the Koopman operator admits the following ODE,

$$\frac{d}{dt}\phi(u(x,t)) = \lambda\phi(u(x,t)), \qquad (2.9)$$

for some  $\lambda \in \mathbb{C}$ . The relation between  $\eta$  in Eq. (2.8) and  $\lambda$  in Eq. (2.9) is  $\eta = e^{\lambda}$ . Thus, a *Koopman Eigenfunctional* (KEFal) is of the form,

$$\phi(u(x,t)) = \phi(u(x,a))e^{\lambda t}, \quad \forall t \in I.$$
 (KEFal)

*Koopman Mode Decomposition.* In the same manner as in the semi-discrete setting, we formulate the solution of the PDE, Eq. (2.6), with KEFals (Nakao and Mezić (2020)). Namely, the solution u(x, t) can be expressed as,

$$u(x,t) = \sum_{i=1}^{\infty} d_i(x)\phi_i(u),$$
(2.10)

where  $\phi_i(u)$  is a KEFal and  $d_i(x)$  is the spatial mode. One way to approximate these spatial modes, is by the method introduced by Nathan Kutz et al. (2018).

### 2.2 Dynamic Mode Decomposition (DMD)

DMD extracts the main spatial structures in the dynamics, Schmid (2010). Backed by Koopman theory, DMD is a principal method to approximate the Koopman modes. It is a data driven method, based on snapshots (mostly, uniformly in time) of the dynamics,  $x_k = x(t_k)$ . The main steps in DMD and its extensions (e.g. Exact DMD Tu et al. (2013), tlsDMD Hemati et al. (2017), fbDMD Dawson et al. (2016), S-DMD Cohen et al. (2021a), and optimized DMD Askham and Kutz (2018)) are:

- 1. Coordinates representation finding the main structures in the dynamics.
- 2. Dimensionality reduction choosing the dominant parts of the dynamics.
- 3. *Linear mapping* finding a linear mapping in the reduced dimensional space.

We describe these steps in detail in Appendix A. The result of DMD and its variants is sets of *modes*,  $\{\phi_i\}$ , *eigenvalues*  $\{\mu_i\}$ , *and coefficients*  $\{\alpha_i\}$ , where i = 1, ..., r and r is the reduced dimension. In the DMD framework, the dynamics is approximated by,

$$\tilde{\boldsymbol{x}}_k \approx \sum_{i=1}^r \alpha_i \mu_i^k \boldsymbol{\phi}_i.$$
(2.11)

## 2.3 General Spectral Decomposition

One of the main goals of signal analysis is to represent a signal sparsely, yet precisely. We focus here on approximating a solution to a PDE, (2.6), by a decomposition of the form,

$$u(x,t) \approx \sum_{i=1}^{L} h_i(x) a_i(t),$$
 (2.12)

where  $\{h_i(\cdot)\}_{i=1}^L$  are spatial functions and  $\{a_i(t)\}_{i=1}^L$  are their respective time profiles. The time profiles are typical to the operator  $\mathcal{P}$  and for homogeneous operators can be expressed analytically, see Cohen and Gilboa (2018). In the semi-discrete setting, the approximate solution of Eq. (2.1) can be expressed as,

$$\mathbf{x}(t) \approx \sum_{i=1}^{L} \mathbf{v}_i a_i(t), \qquad (2.13)$$

where  $\{\mathbf{v}_i\}_{i=1}^{L}$  are spatial structures and  $\{a_i(t)\}_{i=1}^{L}$  are the corresponding time profiles. Note that in some cases (e.g. linear diffusion or TV flow, as shown in Burger et al. (2016)) Eqs. (2.12) and (2.13) reach equality for finite or infinite *L*.

This is the basis of the general spectral decomposition suggested in the thesis of Katzir (2017) and summarized in the book of Gilboa (2018) (chapter 9). The initial condition of Eq. (2.6) is reconstructed with spatial structures that decay according to a known time profile. More formally, given the solution, u(x, t), the spatial structures are the vectors of the minimizer of the following optimization problem,

$$\min_{\mathcal{H}} \left\| \mathcal{U} - \mathcal{H} \mathcal{D} \right\|_{\mathcal{F}}^2 \tag{2.14}$$

where  $\mathcal{U}$  is a matrix of the sampled solution in time and space,  $\mathcal{H}$  is a matrix containing (in its columns) the main spatial structures, and  $\mathcal{D}$  is a dictionary of decay profiles. One can formulate these matrices as

$$\mathcal{U} = \begin{bmatrix} u(x_1, t_0) & \cdots & u(x_1, t_M) \\ \vdots & & \vdots \\ u(x_N, t_0) & \cdots & u(x_N, t_M) \end{bmatrix}, \mathcal{H} = \begin{bmatrix} h_1(x_1) & \cdots & h_r(x_1) \\ \vdots & & \vdots \\ h_1(x_N) & \cdots & h_r(x_N) \end{bmatrix},$$

$$\mathcal{D} = \begin{bmatrix} a_1(t_0) & \cdots & a_1(t_M) \\ \vdots & & \vdots \\ a_r(t_0) & \cdots & a_r(t_M) \end{bmatrix},$$
(2.15)

where  $\mathcal{U} \in \mathbb{R}^{N \times (M+1)}$ ,  $\mathcal{H} \in \mathbb{R}^{N \times r}$ , and  $\mathcal{D} \in \mathbb{R}^{r \times (M+1)}$ . The optimization problem, Eq. (2.14), fits also the form the of semi-discrete setting in the dynamics of Eq. (2.1), where it is sampled in the time axis. We can formulate the following optimization problem,

$$\|X - \mathcal{VD}\|_{\mathcal{F}}^2, \tag{2.16}$$

where the matrix X contains the samples of the dynamics

$$X = \begin{bmatrix} \boldsymbol{x}_0 & \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_M \end{bmatrix} \in \mathbb{R}^{N \times (M+1)}, \qquad (2.17)$$

the matrix  $\mathcal{V}$  contains the main spatial structure of the dynamic (Eq. (2.13))

$$\mathcal{V} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_r \end{bmatrix} \in \mathbb{R}^{N \times r}, \tag{2.18}$$

and the dictionary,  $\mathcal{D}$ , remains unchanged.

## 2.4 Variational Calculus

#### Brezis chain rule

Let Q be a functional over some Banach space and  $\partial Q$  be its variational derivative. Under the regime of the PDE, Eq. (2.6), we can formulate the time derivative of the functional, Q(u(t)), through the "chain rule of Brezis" Brezis (1973) as,

$$\frac{d}{dt}Q(u(x,t)) = \langle \partial Q(u), \frac{d}{dt}u(x,t) \rangle = \langle \partial Q(u), \mathcal{P}(u(x,t)) \rangle.$$
(2.19)

## **Fréchet Differentiability**

The operator  $\mathcal{P} : \mathcal{H} \to \mathcal{H}$  is Fréchet differentiable at *u* if there exists a bounded linear operator  $\mathcal{L}$ , such that,

$$\lim_{\|h\|\to 0} \frac{\|\mathcal{P}(u+h) - \mathcal{P}(u) - \mathcal{L}(h)\|}{\|h\|} = 0$$
(2.20)

holds from any  $h \in \mathcal{H}$ . In this case,  $\mathcal{P}(u+h)$  can be expanded in the Landau notation as

$$\mathcal{P}(u+h) = \mathcal{P}(u) + \mathcal{L}(h) + o(h), \qquad (2.21)$$

where  $\lim_{\|h\|\to 0} \|o(h)\| / \|h\| = 0.$ 

## **Proper Operator**

The operator  $\mathcal{P}(f(x))$  is proper if it gets a finite value for any  $f(x) \in \mathcal{H}$  and for any  $x \in [0, L]$ .

## Region of Attraction (ROA)

Let  $x^*$  be an equilibrium point of the dynamical system in Eq. (2.1). The region of attraction is the largest set in  $\mathbb{R}^N$  that admits the following property: if the initial condition of the dynamics is from the set, then the system converges to  $x^*$  (see e.g. Valmorbida and Anderson (2017)). More formally,

$$\mathcal{RA}(\boldsymbol{x}^*) = \{ x_{init} \in \mathbb{R}^N | \boldsymbol{x}(t=0) = x_{init}, \lim_{t \to \infty} \boldsymbol{x}(t) = \boldsymbol{x}^* \}.$$
(2.22)

## 3 Motivation for this work

This monograph follows an earlier research, attempting to directly apply Koopman operator theory for homogeneous smoothing flows. In Cohen et al. (2021a) we investigated the use of *Dynamic Mode Decomposition* (DMD) for common nonlinear flows emerging in image processing, such as TV-flow and *p*-Laplacian flows. We found out that DMD cannot be naively applied to decompose these flows and presented in detail certain flaws of this procedure.

DMD has become a very common tool in dynamical system analysis. This decomposition provokes interest in many domains of research, such as fluid dynamics, video processing, epidemiology, neuroscience, and finance, see Kutz et al. (2016a). A main advantage is its simplicity and its ability to simplify complex processes by a few modes, in many cases.

DMD invokes well-established tools of dimensionality reduction, and can often reveal the main spatial components of the dynamic. However, the algorithm entails some fundamental problems in recovering nonlinear systems. Moreover, its drawbacks are emphasized when the dynamic is stable and we use the DMD expansions such as, Azencot et al. (2019), where the inverse dynamic is taken into account. Below we show some examples where DMD is failing.

#### 3.1 DMD paradox

The DMD paradox was firstly introduced in Cohen et al. (2021a). We recap here the findings about this DMD flaw. Let  $P(\cdot)$  be a  $\gamma$ -homogeneous operator  $(\gamma \in \mathbb{R})$  over some Banach space  $\mathcal{B}$ , i.e.  $P(av) = a |a|^{\gamma-1} P(v)$  for any  $a \in \mathbb{R}$ and  $v \in \mathcal{B}$ . Let  $\phi \in \mathcal{B}$  be an eigenfunction of P, admitting  $P(\phi) = \lambda \phi$  for a real valued  $\lambda$ . Then, the solution of the PDE (Cohen and Gilboa (2018, 2020))

$$\frac{d}{dt}u = P(u), \quad u(t=0) = \phi, \tag{3.1}$$

is given by

$$u(t) = \begin{cases} \phi \left[ (1 + \lambda(1 - \gamma)t)^+ \right]^{\frac{1}{1 - \gamma}} & \gamma \neq 1\\ \phi e^{\lambda t} & \gamma = 1 \end{cases}$$
(3.2)

where  $(\cdot)^+ = \max\{\cdot, 0\}$ . Under some conditions, the eigenvalue  $\lambda$  is negative for any non-trivial eigenfunction  $\phi$ . Therefore, this solution gets the steady state in finite time when  $\gamma \in [0, 1)$ . The time  $T_{ext}$  for which the dynamic vanishes is

$$T_{ext} = \frac{1}{\lambda(\gamma - 1)}.$$
(3.3)

The decay profile is a fundamental characteristic of signal processing frameworks related to eigenfunctions of  $\gamma$ -homogeneous operators,  $\gamma \in [0, 1)$ . This decomposition generalizes the one based on gradient flows of one-homogeneous functionals, see Bungert et al. (2019a); Burger et al. (2016); Cohen and Gilboa (2020); Gilboa (2013, 2014). The decay profile depends on the homogeneity order  $\gamma$  (see Fig. 3.1). The decay varies from a truncated linear function for zero-homogeneous operators through truncated polynomial functions when  $\gamma \in [0, 1)$ , to exponential function for one-homogeneous operators and finally to hyperbolic functions when  $\gamma > 2$ .



Figure 3.1 decay profile.

The finite extinction time, inherent in flows where  $\gamma \in [0, 1)$ , reveals an unavoidable error in DMD reconstruction. Sampling the solution with fixed step size dt, we get a one dimensional data matrix. Thus, the only valid DMD is when the dimensionality reduction is maximal (r = 1). In that case, the DMD error (for details, see Appendix A, Eq. (A-5)) converges to zero as  $dt \rightarrow 0$ . However, the reconstruction error (Eq. (A-13) is bounded from below. A solution to this problem, as suggested in Cohen et al. (2021a), is to formulate a time rescaled PDE by homogeneity normalization and to apply DMD on that flow. However, this solution is not valid for zero homogeneous flows other than very simple cases, leading to another flaw in DMD.

### 3.2 Discontinuous Dynamical Systems

The analytic solution of (3.1) for  $\gamma = 0$  is known when the Banach space is  $\mathbb{R}^N$ . Applying the homogeneity normalization on zero homogeneous flow,

we find discontinuity in the dynamical modes. Thus, DMD is not valid when the modes vanishes in finite time. The time rescale of (3.1) and the relation of DMD to zero-homogeneous decomposition is detailed in Cohen et al. (2021b).

#### 3.3 Eigenvalue Multiplication

DMD is an exponential data fitting algorithm, Askham and Kutz (2018). Thus, DMD can recover precisely the dynamics only when the typical decay profile of the system is exponential. However, even for the limited case of exponential decays, DMD is not guaranteed to recover the dynamics. Let us consider a dynamic with a solution of the form

$$u(t) = v\left(e^{\lambda_1 t} + e^{\lambda_2 t}\right).$$
(3.4)

This solution cannot be reconstructed by a linear decomposition since the mode v is associated with two eigenvalues  $\lambda_1$  and  $\lambda_2$ .

The rest of this monograph attempts to propose a comprehensive solution to the aforementioned problems. We analyze the conditions for the existence of Koopman eigenfunctions and formulate the KMD modes. Since DMD is an approximation of KMD, if the KEFs do not exist the approximation with DMD is meaningless. After formulating the DMD limitations we propose an alternative mode decomposition, which coincides with KMD modes in a much broader setting.

## 4 Koopman Eigenfunctions and Modes

Koopman theory provides a linear representation to nonlinear dynamics by defining a new coordinate system. These coordinates are the measurements in the state-space termed as *Koopman eigenfunctions*. Necessary and sufficient conditions for their existence are formulated here. Since the eigenfunctions are not unique, we define the *Koopman family*, an infinite set of Koopman eigenfunctions. We also define a useful notion, referred to as the *ancestors of a Koopman family*. This allows the reconstruction of the dynamical system, under certain conditions. Moreover, it allows to considerably enlarge the *Region of Attraction* (ROA) of the system. The above conclusions are consequences of the attributes of the dynamics, *P*, and its solution,  $\mathbf{x}(t)$ , discussed and analyzed below.

### 4.1 Koopman Eigenfunctions

We first set the necessary degree of smoothness of P, required to develop the theory. This setting is highly non-restrictive and accommodates most useful linear and nonlinear dynamics, for both local and nonlocal settings. We refer to the operator P in the dynamical system (2.1).

**Assumption 4.1** (Piecewise Continuous *P*). The operator  $P : \mathbb{R}^N \to \mathbb{R}^N$  is in  $C^0$  a.e. with zero Dirac measures.

This leads to the following Lemma.

**Lemma 4.2** (Continuous solution x(t)). If the operator P in Eq. (2.1) admits Assumption 4.1 then the solution, x, is in  $C^1$  a.e.

Proof. The solution of the dynamics is

$$\mathbf{x}(t) = \mathbf{x}(a) + \int_{a}^{t} P(\mathbf{x}(\tau)) d\tau.$$
(4.1)

The solution,  $\mathbf{x}(t) \in C^1 a.e.$  since  $P \in C^0 a.e.$  and does not contain Dirac measures.

The solution,  $\mathbf{x}(t), t \in I \subset \mathbb{R}^+$ , maps from the time range *I* to  $\mathbb{R}^N$ . It can be interpreted as a parametric curve in  $\mathbb{R}^N$ , where its tangential velocity is  $P(\mathbf{x})$ . Let us denote the image of  $\mathbf{x}(t)$  as *X*. The image is the path in  $\mathbb{R}^N$  where the system passes along the interval *I*. In Fig. 4.2 an illustration of the solution of a dynamical system is shown. Using the Kinematics analogy, we can say the dynamics is a mass going from  $\mathbf{x}(a)$  to  $\mathbf{x}(b)$  with the instantaneous velocity,  $P(\mathbf{x}(t))$ , for every  $t \in I$ . We note that Lemma 4.2 holds also if Assumption 4.1 is limited to *X*.

A Koopman eigenfunction is a measurement of the solution x that admits Eq. (2.4) on the curve X. As recently was stated in Bollt (2021), a Koopman eigenfunction can be formulated as an exponential function, where its argument is the inverse mapping from X to I. The formal definition of the mapping is as follows.

**Definition 4.3** (Time state-space mapping). Let x(t) be the solution of the dynamical system (2.1) where  $t \in I$ . Let  $\xi : X \to I$  be a time state-space mapping from x to t,

$$t = \xi(\mathbf{x}). \tag{4.2}$$

<sup>&</sup>lt;sup>1</sup>The curve image is taken from https://mathinsight.org/definition/simple\_curve



Figure 4.2 The dynamic solution is represented as a curve<sup>1</sup>

This mapping is possible if the curve X is simple and open. Necessary conditions of a curve to be simple are discussed, for instance, in Chuaqui (2018) and the references therein.

**Lemma 4.4** (Differentiation of time state-space mapping). Let the conditions of Lemma 4.2 hold. If the time state-space mapping,  $t = \xi(\mathbf{x})$ , exists then it admits the following,

$$\nabla \xi(\mathbf{x})^T P(\mathbf{x}) = 1 \quad a.e. \text{ in } \mathcal{X}.$$
(4.3)

*Proof.* The mapping  $\xi(\mathbf{x})$  is in  $C^1$  a.e. in  $\mathcal{X}$  since  $\mathbf{x}(t) \in C^1 a.e.$  in *I*. The time derivative of the mapping is,

$$1 = \frac{d}{dt}t = \frac{d}{dt}\xi(\mathbf{x}) = \nabla\xi(\mathbf{x})^T \frac{d\mathbf{x}}{dt} = \nabla\xi(\mathbf{x})^T P(\mathbf{x}).$$
(4.4)

This expression is valid almost everywhere.

We now turn to discuss necessary and sufficient conditions for the existence of a nontrivial Koopman eigenfunction (that is, a nonzero function which admits Eq. (2.4) with  $\lambda \neq 0$ ).

**Proposition 4.5** (Condition for the inexistence of a Koopman eigenfunction). *If there is an equilibrium point in I then a nontrivial Koopman eigenfunction does not exist.* 

*Proof.* Let  $t_0 \in I$  be an equilibrium point and  $\varphi(\mathbf{x}(t))$  be a Koopman eigenfunction. Then,  $\mathbf{x}(t) = const$ ,  $\forall t \in [t_0, b]$ . Therefore, Eq. (2.4) does not hold for nontrivial  $\varphi$  for any  $\lambda \neq 0$ .

**Remark 4.6** (Finite support time dynamics). Let  $P(\mathbf{x})$  define a dynamic for which the solution has a finite support in time. Namely, there is an extinction time point,  $T_{ext}$ , for which  $P(\mathbf{x}(t)) = \mathbf{0}$ ,  $\forall t \ge T_{ext}$ . Then, if  $T_{ext} \in I$ , a Koopman operator  $K_P^{\tau}$  has no eigenfunctions. We observe here that the time interval I is crucial for the existence or inexistence of eigenfunctions.

From a differential geometry perspective, as noted above,  $\mathbf{x}(t)$  forms a curve where its tangential velocity is  $P(\mathbf{x})$ . The absence of an equilibrium point is equivalent to nonzero velocity. This type of parametric curves, where the velocity is always nonzero, is called *regular*. The Koopman eigenfunction does not exist for non-regular curves.

**Lemma 4.7** (Koopman Eigenfunctions induced by a time state-space mapping). Let the conditions of Lemma 4.2 hold and  $\mathbf{x}(t)$  be the solution of Eq. (2.1). If there exists a time state-space mapping,  $t = \xi(\mathbf{x})$ , then a Koopman eigenfunction exists a.e. in I.

*Proof.* The mapping,  $t = \xi(\mathbf{x})$ , is in  $C^1$  a.e. in X since  $\mathbf{x}(t)$  is in  $C^1$  a.e. in I. Given that mapping, we define the following function,

$$\varphi(\mathbf{x}) = e^{\alpha \xi(\mathbf{x}) + \beta}.$$
(4.5)

This function is in  $C^1$  a.e. in X. The time derivative of this function is,

$$\frac{d}{dt}\varphi(\mathbf{x}(t)) = \frac{d}{d\xi}e^{\alpha\cdot\xi(\mathbf{x})+\beta}\nabla\xi(\mathbf{x})^T\frac{d}{dt}\mathbf{x} = \alpha\varphi(\mathbf{x}(t))\nabla\xi(\mathbf{x})^T P(\mathbf{x}(t)).$$
(4.6)

According to Lemma 4.4,  $\nabla \xi(\mathbf{x})^T P(\mathbf{x}(t)) = 1$  a.e.. Thus, the function in Eq. (4.5) admits Eq. (2.4) for any value of  $\beta$ , where the corresponding eigenvalue is  $\lambda = \alpha$ .

**Theorem 4.8** (Sufficient condition for the existence of a Koopman eigenfunction). Let the conditions of Lemma 4.2 hold and one of the entries of the vector  $P(\mathbf{x}(t))$  is either positive or negative  $\forall t \in I$ . Then, Koopman eigenfunctions exist a.e. in the time interval I.

*Proof.* If one of the entries in P(x(t)) is either positive or negative for all  $t \in I$  then this entry is monotone and therefore injective. Then, the curve X is simple and open (see Courant and John (2012) pages 45, 177 and 207). Therefore, the time state-space mapping,  $\xi(x)$ , exists. Following Lemma 4.7, Koopman eigenfunctions can be expressed by (4.5).

The simple example below illustrates the connections between the equilibrium point, finite time dynamics and time state-space mapping.

Example 4.9 (Finite time support). Let us consider the following dynamics,

$$\frac{d}{dt}x = -2x^{\frac{1}{2}}, \quad x(0) = 1.$$
(4.7)

The solution is

$$x(t) = \begin{cases} (1-t)^2 & t \in [0,1] \\ 0 & t > 1 \end{cases}.$$
 (4.8)

For I = [0, 1], the time state-space mapping is,

$$t(x) = 1 - \sqrt{x}, \quad I = [0, 1],$$
 (4.9)

and using (4.5) with  $\alpha = 1$ ,  $\beta = 0$ , we can express a Koopman eigenfunction by,

$$\varphi(x) = e^{1 - \sqrt{x}}.\tag{4.10}$$

Now, let us repeat this example with a different time interval. Let I = [0, 1.5], containing the extinction time  $T_{ext} = 1$ . Note that, first, the time mapping, Eq. (4.9), does not hold in the entire interval, and the eigenfunction  $\varphi$  does not admit  $\frac{d}{dt}\varphi(x) = \varphi(x)$  since  $\varphi(x)$  is a nonzero constant for  $t \in [1, 1.5]$ .

### 4.1.1 Extended DMD induced from time state-space mapping

One of the methods to increase the accuracy of the classic DMD is by enriching the state-space vector with nonlinear measurements of the coordinates x, see Williams et al. (2015b). It is shown that this approach indeed improves accuracy, however - the theoretical justification is vague. In addition - the enriching method may become somewhat heuristic. We can interpret this approach as the Taylor expansion of Koopman eigenfunctions. This provides both justification and a clear method for supplying additional measurements. Let us expand the Koopman eigenfunction,  $\varphi(x) = e^{\xi(x)}$ , by a Taylor series,

$$\varphi(\mathbf{x}) = e^{\xi(\mathbf{x})} = \sum_{j=0}^{\infty} \frac{\xi(\mathbf{x})^j}{j!}.$$

We can approximate this expression by taking only finite number of elements from this sum,

$$e^{\xi(\mathbf{x})} \approx \sum_{j=0}^{M} \frac{\xi(\mathbf{x})^j}{j!}.$$

Thus, Eq. (2.4) can be approximated as

$$\frac{d}{dt} \sum_{j=0}^{M} \frac{\xi(\mathbf{x})^j}{j!} \approx \sum_{j=0}^{M} \frac{\xi(\mathbf{x})^j}{j!}.$$
(4.11)

In matrix notation, this approximation can be reformulated as,

$$\frac{d}{dt} \begin{bmatrix} 1\\ \xi(\mathbf{x})\\ \vdots\\ \frac{\xi(\mathbf{x})^M}{M!} \end{bmatrix} \approx A \begin{bmatrix} 1\\ \xi(\mathbf{x})\\ \vdots\\ \frac{\xi(\mathbf{x})^M}{M!}, \end{bmatrix}$$
(4.12)

where any matrix A with a left-eigenvector  $\begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}$  can be an optional solution to Eq. (4.12) for which Eq. (4.11) holds. In addition, taking M to infinity, A gets the form

$$[A]_{i,j} = \begin{cases} 1 & i = j+1 \\ 0 & else \end{cases}$$

where  $[A]_{i,j}$  is the *i*, *j*th entry of *A*.

## 4.2 Koopman Family

The KEF is of the form  $\varphi(t) = e^{\alpha t + \beta}$ , Eq. (4.5). This form of solution is unique following a standard existence and uniqueness theorem of ODE's. The exponential parameters,  $\alpha$  and  $\beta$ , are dictated by the eigenvalue and the initial condition. Without these restrictions, there are infinite KEFs for any dynamical system.

From a different angle, viewing the state-space x as a curve gives a compelling interpretation of the multiplicity of Koopman eigenfunctions. A curve can be reparameterized in different manners. Changing the parameters,  $\alpha$  and  $\beta$ , amounts to reparameterization by translation and scaling. This insight leads us to the following lemma, which extends the identities presented in Bollt (2021). We show that any mathematical manipulation on a KEF which maintains the form of Eq. (4.5) generates a new KEF. **Lemma 4.10** (Multiplicities of Koopman eigenfunctions). If  $\varphi_1, \varphi_2$  are Koopman eigenfunctions with the corresponding eigenvalues  $\lambda_1, \lambda_2$  then:

- 1. The function  $a \cdot \varphi_1$ ,  $a \in \mathbb{R}$   $(a \neq 0)$  is an eigenfunction with eigenvalue  $\lambda_1$ .
- 2. The function  $(\varphi_1)^{\alpha}$ ,  $\alpha \in \mathbb{C}$   $(\alpha \neq 0)$  is an eigenfunction with eigenvalue  $\alpha \lambda_1$ .
- 3. For any  $n, m \in \mathbb{R}$  the function  $(\varphi_1)^n (\varphi_2)^m$  is an eigenfunction with eigenvalue  $n\lambda_1 + m\lambda_2$ .
- 4. The function  $(\varphi_1)^{\frac{\lambda}{\lambda_1}} + (\varphi_2)^{\frac{\lambda}{\lambda_2}}$  is an eigenfunction with eigenvalue  $\lambda$ .

Proof. This can be shown by,

- 1. Using the linearity of the Koopman operator.
- 2. Writing the time derivative of  $(\varphi_1)^{\alpha}$  explicitly we get,

$$\frac{d}{dt}\left[\varphi_{1}^{\alpha}\right] = \alpha(\varphi_{1})^{\alpha-1}\lambda_{1}\varphi_{1} = \alpha\lambda_{1}\varphi_{1}^{\alpha}.$$
(4.13)

3. Similarly,

$$\frac{d}{dt} \left[ (\varphi_1)^n (\varphi_2)^m \right] = (\varphi_2)^m n (\varphi_1)^{n-1} \lambda_1 \varphi_1 + (\varphi_1)^n m (\varphi_2)^{m-1} \lambda_2 \varphi_2$$
$$= (n\lambda_1 + m\lambda_2) (\varphi_1)^n (\varphi_2)^m.$$
(4.14)

4. Finally,

$$\frac{d}{dt} \left[ (\varphi_1)^{\frac{\lambda}{\lambda_1}} + (\varphi_2)^{\frac{\lambda}{\lambda_2}} \right] = \frac{\lambda}{\lambda_1} (\varphi_1)^{\frac{\lambda}{\lambda_1} - 1} \lambda_1 \varphi_1 + \frac{\lambda}{\lambda_2} (\varphi_2)^{\frac{\lambda}{\lambda_2} - 1} \lambda_2 \varphi_2$$
$$= \lambda \left[ (\varphi_1)^{\frac{\lambda}{\lambda_1}} + (\varphi_2)^{\frac{\lambda}{\lambda_2}} \right].$$
(4.15)

#### Discussion

The multiplicities presented in Lemma 4.10 are translation and scaling of the time variable. Case 1 in this Lemma is a translation of the time axis and the rest of the cases are scaling. To distinguish between eigenfunctions which are generated from other eigenfunctions and "new" independent ones, we introduce the concepts of *Koopman family* and its *ancestors*.

**Definition 4.11** (Koopman family). Let  $\{\varphi_i\}_{i=1}^n$  be a finite set of KEFs. Let  $k_P(\{\varphi_i\}_{i=1}^n)$  be the infinity uncountable set of KEFs generated by the finite set, recursively, according to the four options stated in Lemma 4.10. Let us define  $k_P^m(\{\varphi_i\}_{i=1}^n) = k_P(k_P^{m-1}(\{\varphi_i\}_i^n))$ . We term  $\mathcal{K}_P(\{\varphi_i\}_i^n) = \bigcup_{m=1}^\infty k_P^m(\{\varphi_i\}_{i=1}^n)$  as the Koopman family of  $\{\varphi_i\}_{i=1}^n$ .

**Definition 4.12** (Ancestors of a Koopman family). Let  $\{\varphi_i^*\}_{i=1}^m$  be a finite set of KEFs. This set is an ancestor set of the Koopman family  $\mathcal{K}_P(\{\varphi_i\}_{i=1}^n)$  if the following conditions hold:

1.  $\varphi \in \mathcal{K}_P(\{\varphi_i\}_i^n) \iff \varphi \in \mathcal{K}_P(\{\varphi_i^*\}_i^m).$ 2.  $\varphi_j^* \notin \mathcal{K}_P(\{\varphi_i^*\}_{i=1, i \neq j}^m)$  for any  $j = 1, 2, \cdots, m.$ 

Note that the subscript P is for the dynamical system.

4.2.1 Koopman Eigenfunction Vector

A vector of Koopman eigenfunctions is denoted by,

$$\boldsymbol{\varphi}(\boldsymbol{x}) = \begin{bmatrix} \varphi_1(\boldsymbol{x}) & \cdots & \varphi_L(\boldsymbol{x}) \end{bmatrix}^T,$$
 (4.16)

where L can be finite or infinite. The Jacobian matrix of this vector is,

$$\frac{\partial}{\partial \boldsymbol{x}}\boldsymbol{\varphi}(\boldsymbol{x}) = \begin{bmatrix} \nabla \varphi_1(\boldsymbol{x})^T \\ \vdots \\ \nabla \varphi_L(\boldsymbol{x})^T \end{bmatrix} = \mathcal{J}(\boldsymbol{\varphi}).$$
(4.17)

**Theorem 4.13** (Linear dynamic in Koopman family). Let the conditions of Theorem 4.8 hold. The dynamical system P can be represented as a linear one with a vector of Koopman eigenfunctions, where the time derivative of this vector is,

$$\frac{d}{dt}\boldsymbol{\varphi}(\boldsymbol{x}) = \mathcal{J}(\boldsymbol{\varphi})P(\boldsymbol{x}) = \Lambda \boldsymbol{\varphi}(\boldsymbol{x}), \quad a.e.$$
(4.18)

where  $\Lambda$  is a diagonal matrix with the corresponding eigenvalues.

*Proof.* We would like to prove first the existence of a L dimensional KEF. From Theorem 4.8 there exists a KEF. From Lemma 4.10 if there exists a KEF, there are infinite set of KEFs, therefore, at least L eigenfunctions. According to the definition of the Koopman eigenfunction, the time derivative is,

$$\frac{d}{dt}\boldsymbol{\varphi}(\boldsymbol{x}) = \left[\frac{d}{dt}\varphi_1(\boldsymbol{x}), \cdots, \frac{d}{dt}\varphi_L(\boldsymbol{x})\right]^T = \left[\lambda_1\varphi_1(\boldsymbol{x}), \cdots, \lambda_L\varphi_L(\boldsymbol{x})\right]^T = \Lambda \boldsymbol{\varphi}(\boldsymbol{x}).$$
(4.19)

On the other hand, applying the chain rule we get,

$$\frac{d}{dt}\boldsymbol{\varphi}(\boldsymbol{x}) = \begin{bmatrix} \nabla \varphi_1(\boldsymbol{x})^T \\ \vdots \\ \nabla \varphi_L(\boldsymbol{x})^T \end{bmatrix} \frac{d}{dt} \boldsymbol{x}(t) = \mathcal{J}(\boldsymbol{\varphi}) P(\boldsymbol{x}).$$
(4.20)

Note that  $\varphi_i(\mathbf{x}) \in C^1$ , (*a.e.*), so the expressions above are valid only almost everywhere.

#### 4.2.2 Reconstructing the dynamics

The ability to reconstruct the dynamics is based on the relations between the vectors  $\varphi$  and x. In classical control theory, this is referred to as *observability*. Here, we suggest to examine the notion of observability by computing the rank of the Jacobian matrix  $\mathcal{J}(\varphi)$ , Eq. (4.17). The rows of this matrix are the gradients of the KEFs. The following lemma shows that the gradient of a member of a Koopman family originates with its ancestors.

**Lemma 4.14** (KEF gradients of a family). Let  $\mathcal{K}_P(\{\varphi_i^*\}_{i=1}^m)$  be the Koopman family of an ancestor set,  $\{\varphi_i^*\}_{i=1}^m$ . Let  $\varphi$  be a KEF in  $\mathcal{K}_P(\{\varphi_i^*\}_{i=1}^m)$ . Then, the gradient of  $\varphi$ ,  $\nabla \varphi$ , is a linear combination of the gradients of the ancestor set for any  $t \in I$ .

*Proof.* Let  $\mathcal{KG}$  be the linear span, defined by

$$\mathcal{KG} = span\left(\left\{\nabla\varphi_i^*\right\}_{i=1}^m\right) = \left\{\sum_{i=1}^m a_i \nabla\varphi_i^*, \ \forall a_i \in \mathbb{C}\right\}.$$
(4.21)

Let  $\varphi$  be in  $\mathcal{K}_P(\{\varphi_i^*\}_{i=1}^m)$ . According to Definition 4.11, there exist recursive steps leading from the ancestors  $\{\varphi_i^*\}_{i=1}^m$  to  $\varphi$ . Now, by induction we show that  $\nabla \varphi \in \mathcal{K}\mathcal{G}$ . Let us assume that from the ancestors to  $\varphi$  there is one step. Namely,  $\varphi$  is generated using  $\varphi_i^*, \varphi_j^*$ , according to the four cases of Lemma 4.10. The gradient of  $\varphi$  is a linear combination of the gradients of  $\varphi_i^*$  and  $\varphi_j^*$ . For cases 1, 2 and 4, the linearity is straightforward. For case 3,  $\varphi = (\varphi_i^*)^n (\varphi_j^*)^l$ , we have,

$$\nabla \varphi = \nabla (\varphi_i^*)^n \left(\varphi_j^*\right)^l = n \left(\varphi_i^*\right)^{n-1} \left(\varphi_j^*\right)^l \nabla (\varphi_i^*) + l \left(\varphi_i^*\right)^n \left(\varphi_j^*\right)^{l-1} \nabla \left(\varphi_j^*\right)$$

$$= \left[\nabla (\varphi_i^*) \quad \nabla \left(\varphi_j^*\right)\right] \begin{bmatrix} n \left(\varphi_i^*\right)^{n-1} \left(\varphi_j^*\right)^l \\ l \left(\varphi_i^*\right)^n \left(\varphi_j^*\right)^{l-1} \end{bmatrix}.$$
(4.22)

For any  $t \in I$  the vector  $\left[n\left(\varphi_{i}^{*}\right)^{n-1}\left(\varphi_{j}^{*}\right)^{l} \quad l\left(\varphi_{i}^{*}\right)^{n}\left(\varphi_{j}^{*}\right)^{l-1}\right]^{T}$  is constant. There-

fore, the gradient of  $\varphi$  is in  $\mathcal{KG}$ . Now we assume there exist *k* steps from the ancestors to  $\varphi$ . Let  $\varphi_1$  and  $\varphi_2$  be generated by k - 1 steps. The induction assumption holds, meaning, their gradients are in  $\mathcal{KG}$ . Now, there is one step from  $\varphi_1$  and  $\varphi_2$  to  $\varphi$ . As shown,  $\nabla \varphi$  is a linear combination of the gradients of its generators,  $\nabla \varphi_1$ ,  $\nabla \varphi_2$ . But these vectors belong to  $\mathcal{KG}$  by the induction assumption. Therefore,  $\nabla \varphi \in \mathcal{KG}$ .

The multiplicity of Koopman eigenfunctions results from either arithmetical manipulations (Def. 4.11) or the existence of several time state-space mappings (Def. 4.3). The main difference is the rank of the Jacobian,  $\mathcal{J}(\varphi)$ . Given a vector of KEFs, adding another Koopman eigenfunction from the Koopman family of the KEFs in the vector – does not increase the rank of the Jacobian. However, adding a Koopman eigenfunction from another time state-space mapping does. The Jacobian matrix rank is related to system controllability and observability (see for example Brunton and Kutz (2019); Evangelisti (2011)). In the following, we formulate the connections between the rank of the Jacobian matrix, the size of the ancestor set, and time state-space mappings.

**Definition 4.15** (Full observability in the context of Koopman theory). Consider the dynamical system Eq. (2.1) where  $x \in \mathbb{R}^N$ . The system is fully observable if the state-space can be revealed from the KEFs and the initial condition.

**Proposition 4.16** (Sufficient condition for full observability). Consider the dynamical system (2.1) where  $\mathbf{x} \in \mathbb{R}^N$ . Let us denote the Koopman family of all Koopman eigenfunctions of the dynamics as  $\mathcal{K}_P$ . An ancestor set of  $\mathcal{K}_P$  is denoted as  $\{\varphi_i^*\}_{i=1}^n$ . The system is fully observable if  $N \leq n$ .

*Proof.* According to Lemma 4.13, for any vector of KEFs the following equation holds,

$$\mathcal{J}(\boldsymbol{\varphi})P(\boldsymbol{x}) = \Lambda \boldsymbol{\varphi}(\boldsymbol{x}), \quad (a.e.). \tag{4.23}$$

Let us choose a vector of ancestors, i.e.

$$\boldsymbol{\varphi}^*(\boldsymbol{x}) = \begin{bmatrix} \varphi_1^*(\boldsymbol{x}) & \cdots & \varphi_n^*(\boldsymbol{x}) \end{bmatrix}^T .$$
(4.24)

According to Lemma 4.14 the rank of the Jacobian matrix is full and equal to *N*. Since the matrix  $\mathcal{J}(\boldsymbol{\varphi})^T \mathcal{J}(\boldsymbol{\varphi})$  is invertible, the dynamics, *P*, can be revealed according to the following relation,

$$P(\mathbf{x}) = \left(\mathcal{J}(\boldsymbol{\varphi})^T \mathcal{J}(\boldsymbol{\varphi})\right)^{-1} \mathcal{J}(\boldsymbol{\varphi})^T \Lambda \boldsymbol{\varphi}(\mathbf{x}), \quad (a.e.).$$
(4.25)

That is, we use the Moore-Penrose inverse. The state-space can now be calculated

as,

$$\boldsymbol{x}(t) = \boldsymbol{x}_0 + \int_a^t \left( \mathcal{J}(\boldsymbol{\varphi}(\tau))^T \mathcal{J}(\boldsymbol{\varphi}(\tau)) \right)^{-1} \mathcal{J}(\boldsymbol{\varphi}(\tau))^T \Lambda \boldsymbol{\varphi}(\tau) d\tau.$$
(4.26)

**Corollary 4.17** (Full observability for a monotone dynamics). *If each entry in P is either positive or negative for any t in I then each entry of the state-space is monotone (and injective). We can formulate N different time state-space mappings from X to I (Theorem 4.8). These mappings induce N different KEFs and according to Proposition 4.16 the system is fully observable.* 

**Remark 4.18** (Sufficient condition for dynamic reconstruction). If each of the entries of P is either positive or negative for all t in I then the dynamics can be reconstructed as

$$P(\mathbf{x}) = \mathcal{J}^{-1}(\boldsymbol{\varphi})\Lambda \boldsymbol{\varphi}(\mathbf{x}). \tag{4.27}$$

According to Corollary 4.17, if each of the entries of *P* is either positive or negative in *I* then the Jacobian matrix is  $N \times N$  and is full rank, therefore – invertible. Using Theorem 4.13 we reach Eq. (4.27).

**Remark 4.19** (Global controllability). Reconstructing the dynamical system enables us to enlarge the *Region of Attraction* (ROA), Eq. (2.22). Given the nonlinear dynamics,

$$\frac{d}{dt}\boldsymbol{x}(t) = P(\boldsymbol{x}(t)) + \boldsymbol{u}, \qquad (4.28)$$

we can cancel the nonlinearity with the ancestors of a Koopman family  $\varphi(x)$  if the dynamics is fully observable. In order to reach a stable system for any point x we define the following input u,

$$\boldsymbol{u} = \mathcal{J}^{-1}(\boldsymbol{\varphi})\Lambda\boldsymbol{\varphi}(\boldsymbol{x}) + \boldsymbol{w}, \tag{4.29}$$

where the first element cancels the nonlinearity of the system (Remark 4.18) and the second term brings the system to any desired point in  $\mathbb{R}^N$ . Note that we assume there are no singular points in *P*.

**Remark 4.20** (Reconstructing the dynamic - limitations). The recovery of the system, as described by Eq. (4.25) and (4.27), is valid for a given initial condition  $x_0$ . In order to obtain a full recovery of the system, the properties of the initial condition should be taken into account. These subject exceeds the frame of this work and requires further research.

#### 4.2.3 Reconstruction conservation laws

Dynamic reconstruction and conservation laws (such as energy, momentum etc.) are perhaps the most crucial tasks in dynamical system analysis, in general, and controlling systems in particular. Data driven algorithms to reveal the dynamic (governing laws) and the conservation laws, based on the Koopman operator theory have been studied by Rudy et al. (2017); Brunton et al. (2016); Schmidt and Lipson (2009); Kaiser et al. (2018); Langley et al. (1981). The common approach argues that the conservation laws are related to the null part of the Koopman spectrum. Namely, the Koopman eigenfunctions related to eigenvalue zero are or may be formulations of the conservation laws. In a similar manner, the dynamic can be reconstructed based on the nontrivial Koopman eigenfunctions.

We propose an alternative view. As discussed above, the relevant KEFs to system reconstruction are indeed not in the null part of the Koopman spectrum. However, from our perspective, after recovering the dynamical system (the governing laws) via KEFs the conservation laws naturally emerge from these nontrivial KEFs.

Let  $\varphi(x)$  be an eigenfunction, namely Eq. (2.4) holds for some  $\lambda$ . We consider the measurement  $\ln(\varphi(x))$ . By using Eq. (2.2), we can express the time derivative of this measurement as,

$$\frac{d}{dt}\ln\left(\varphi(\boldsymbol{x})\right) = \frac{\lambda\varphi(\boldsymbol{x})}{\varphi(\boldsymbol{x})} = \lambda.$$

On the other hand, with the chain rule we get,

$$\frac{d}{dt}\ln\left(\varphi(\boldsymbol{x})\right) = \frac{\frac{d}{dt}\varphi(\boldsymbol{x})}{\varphi(\boldsymbol{x})} = \frac{\nabla\varphi(\boldsymbol{x})^T\frac{d}{dt}\boldsymbol{x}}{\varphi(\boldsymbol{x})} = \frac{\nabla\varphi(\boldsymbol{x})^TP(\boldsymbol{x})}{\varphi(\boldsymbol{x})}$$

Then, for any Koopman eigenfunction (an intrinsic coordinate) we can formulate the following conservation law,

$$\frac{\nabla \varphi(\mathbf{x})^T P(\mathbf{x})}{\varphi(\mathbf{x})} = \lambda.$$
(4.30)

This result coincides with Eq. (4.3), by using Lemma 4.4 and formulating a KEF by a time state-space mapping, Eq. (4.5), with  $\alpha = \lambda$ . In other words, when a time state-space mapping  $\xi$  exists, an alternative formulation to the conservation law of Eq. (4.30) is,

$$1 = \frac{d}{dt}t = \frac{d}{dt}\xi(\mathbf{x}) = \nabla\xi(\mathbf{x})^T \frac{d\mathbf{x}}{dt} = \nabla\xi(\mathbf{x})^T P(\mathbf{x}).$$
(4.31)

We illustrate this with the following two examples.

**Example 4.21** (Free Fall). Let  $x_1(t)$  and  $x_2(t)$  be the height and the velocity of a mass in a free fall, respectively. The dynamical system is,

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -g \end{bmatrix},$$

where the initial condition is  $\begin{bmatrix} h & 0 \end{bmatrix}^T$ . The solution is

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} h - \frac{1}{2}gt^2 \\ -gt \end{bmatrix}.$$

The time state-space mappings are,

$$\begin{bmatrix} t_1 \\ t_2 \end{bmatrix} = \begin{bmatrix} \sqrt{\frac{2(h-x_1)}{g}} \\ -\frac{x_2}{g} \end{bmatrix}.$$

The induced conservation laws, using Eq. (4.30) are as follows. Conservation law #1

$$1 = \frac{d}{dx_1} t_1(x_1) \cdot \frac{d}{dt} x_1 = -\frac{1}{2\sqrt{\frac{2(h-x_1)}{g}}} \frac{2}{g} \cdot x_2$$

One can reformulate this to the energy conservation law,

$$gx_1 + \frac{1}{2}x_2^2 = hg.$$

#### **Conservation law #2**

$$\frac{d}{dx_2}t_2(x_2) \cdot \frac{d}{dt}x_2 = -\frac{1}{g} \cdot (-g) = 1$$

The conservation law #2 is due to the constant acceleration, g.

**Example 4.22** (Pure rolling down an incline). On an inclined plane with a slope of angle  $\alpha$ , a solid cylinder with mass *m*, radius *R*, and rotational inertia  $I_{cm}$  is released from rest. The location along the plane is denoted by  $x_1$  and its velocity by  $x_2$ . The dynamical system is,

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ \frac{g \sin \alpha}{1 + \frac{l c m}{mR^2}} \end{bmatrix}$$
(4.32)

23

with the initial condition  $\mathbf{x} = [0, 0]^T$ . The solution is,

$$x_{1}(t) = \frac{1}{2} \frac{gsin\alpha}{1 + \frac{I_{cm}}{mR^{2}}} t^{2},$$

$$x_{2}(t) = \frac{gsin\alpha}{1 + \frac{I_{cm}}{mR^{2}}} t.$$
(4.33)

The time mappings are

$$t_{1}(x_{1}) = \sqrt{2 \frac{1 + \frac{I_{cm}}{mR^{2}}}{g \sin \alpha} x_{1}},$$

$$t_{2}(x_{2}) = \frac{1 + \frac{I_{cm}}{mR^{2}}}{g \sin \alpha} x_{2}.$$
(4.34)

### **Conservation law #1**

$$1 = \frac{d}{dx_1} t_1(x_1) \frac{dx_1}{dt} = \sqrt{2 \frac{1 + \frac{I_{cm}}{mR^2}}{gsin\alpha}} \frac{1}{2\sqrt{x_1}} x_2$$
(4.35)

We can reformulate it as,

$$\underbrace{\frac{1}{2}mx_2^2}_{E_K} + \underbrace{\frac{1}{2}I_{cm}\left(\frac{x_2}{R}\right)^2}_{E_R} \underbrace{-mgx_1sin\alpha}_{E_P} = 0$$
(4.36)

getting, as expected, that the sum of the energies (Kinetic, Rotational, and Potential) is zero.

**Conservation law #2** In the same manner as in Example 4.21, conservation law #2 is a result of constant acceleration.

## 4.3 Koopman Mode Decomposition

The Koopman mode decomposition leverages this infinite family to reconstruct the observations from the Koopman eigenfunctions (Mezić (2005)). The reconstruction is a linear combination of Koopman eigenfunctions. For instance, the *i*th entry of x is assumed to be reconstructed as (Brunton et al. (2021)),

$$x_i(t) = \sum_{j=1}^{\infty} v_{i,j} \varphi_j(\boldsymbol{x}(t)), \qquad (4.37)$$

where  $v_{i,j}$  is a scalar. Then, the state-space can be written as,

$$\boldsymbol{x}(t) = \sum_{j=1}^{\infty} \boldsymbol{v}_j \varphi_j(\boldsymbol{x}(t)), \qquad (4.38)$$

where  $v_j$  is an *N* dimensional vector whose entries are the coefficients of the *j*th Koopman eigenfunction, namely  $v_j = \begin{bmatrix} v_{1,j} & \cdots & v_{N,j} \end{bmatrix}^T$ . Substituting the solution of  $\varphi(\mathbf{x})$ , Eq. (KEF), we get,

$$\mathbf{x}(t) = \sum_{j=1}^{\infty} \mathbf{v}_j \varphi_j(\mathbf{x}(a)) e^{\lambda_j t}.$$
(4.39)

The infinite triplet  $\{v_j, \varphi_j, \lambda_j\}_{j=1}^{\infty}$  is the Koopman mode decomposition, where  $\{v_j\}_{j=1}^{\infty}$  are the Koopman modes,  $\{\varphi_j\}_{j=1}^{\infty}$  are the KEFs, and  $\{\lambda_j\}_{j=1}^{\infty}$  are the Koopman eigenvalues. Note that the maximal index argument in the sum of Eq. (4.38) is not necessarily infinity. For example, it is enough to have one mode to reconstruct the linear dynamics initiated with one of its eigenvectors. In matrix notations, let *V* be a matrix whose column vectors are the corresponding Koopman modes. The state-space can be expressed as,

$$\boldsymbol{x}(t) = V\boldsymbol{\varphi}(\boldsymbol{x}(t)). \tag{4.40}$$

Thus, the dynamical system has a linear representation with the measurements  $\{\varphi_j(\mathbf{x})\}_{i=1}^{\infty}$ , Kaiser et al. (2021).

**Example 4.23** (KMD of Zero Homogeneous Dynamics). Let us consider the following dynamical system

$$\frac{d}{dt}\mathbf{x} = P(\mathbf{x}), \quad \mathbf{x}(t=0) = \mathbf{v}, \ I = [0, -1/\lambda)$$
(4.41)

where *P* is a zero homogeneous operator (admitting  $P(a \cdot \mathbf{x}) = sign(a)P(\mathbf{x}), \forall a \in \mathbb{R}$ ),  $\mathbf{v}$  and  $\lambda$  are a nonlinear eigenvector and the corresponding eigenvalue of *P*, respectively, i.e. they admit the nonlinear eigenvalue problem  $P(\mathbf{v}) = \lambda \mathbf{v}$ . We assume a stable system, where  $\lambda < 0$ . More background on such problems is presented in Gilboa (2018). Then, the solution of the ODE is,

$$\boldsymbol{x}(t) = \boldsymbol{v}\left(1 + \lambda t\right), \quad t \in I. \tag{4.42}$$

A KEF can be formulated by the time state-space mapping as,

$$\varphi(t) = e^t = e^{\frac{\langle \mathbf{x}, \mathbf{v} \rangle}{\|\mathbf{y}\|^2 - 1}}.$$
(4.43)

We would like now to express the solution (4.42) with Koopman eigenfunctions.

To express the function t we have to apply the natural logarithm, ln, on the Koopman eigenfunction. With Taylor series one can express it as,

$$t = \ln(\varphi(\mathbf{x})) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{(\varphi(\mathbf{x}) - 1)^n}{n}.$$
 (4.44)

Then, the solution of (4.42) can be written as,

$$\mathbf{x}(t) = \mathbf{v} \left( 1 + \lambda \sum_{n=1}^{\infty} (-1)^{n+1} \frac{(\varphi(\mathbf{x}) - 1)^n}{n} \right).$$
(4.45)

By expanding the terms  $(\varphi - 1)^n$  we get an infinite polynomial with respect to the KEF  $\varphi$ . KMD emerges naturally.

#### Discussion

According to this example, since there is only one mode and its decay profile is not exponential, there can be many KEFs for one Koopman mode. The multiplicity of eigenvalues for one mode is related to the limitations of DMD. Since DMD recovers only linear dynamics it cannot handle well one eigenvector with multiple eigenvalues.

We can now formulate the relation between Koopman modes and the dynamical system.

**Proposition 4.24** (The Jacobian and Koopman modes). Let  $\varphi(\mathbf{x})$  be a vector of Koopman eigenfunctions and  $\mathcal{J}(\varphi(\mathbf{x}))$  be its Jacobian matrix. In addition, let V be defined as in (4.40). Then,  $P(\mathbf{x})$  is a right eigenvector of the matrix  $V \cdot \mathcal{J}(\varphi(\mathbf{x}))$  with eigenvalue one.

*Proof.* The time derivative of Eq. (4.40) is given by,

$$\frac{d}{dt}\mathbf{x} = V \frac{d}{dt}\boldsymbol{\varphi}(\mathbf{x})$$

$$P(\mathbf{x}) = V \mathcal{J}(\boldsymbol{\varphi}(\mathbf{x}))P(\mathbf{x}).$$
(4.46)

Example 4.25 (Nonlinear system). Given the following system,

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 - x_1^2 \end{bmatrix}, \quad \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$
(4.47)

The solution is,

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} e^t \\ e^{2t} \end{bmatrix}.$$
 (4.48)

The time state-space mappings are,

$$t = \ln(x_1),$$
  

$$t = \frac{1}{2}\ln(2x_1 - x_2).$$
(4.49)

By choosing  $\alpha = 1$ ,  $\beta = 0$ , the Koopman eigenfunctions, following (4.5), are,

$$\varphi_1(\mathbf{x}) = x_1,$$
  

$$\varphi_2(\mathbf{x}) = \sqrt{2x_1 - x_2}.$$
(4.50)

The state-space,  $\begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$ , can be reconstructed by these eigenfunctions as,

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \varphi_1(\boldsymbol{x}) \\ 2\varphi_1(\boldsymbol{x}) - \varphi_2(\boldsymbol{x})^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} \varphi_1(\boldsymbol{x}) \\ \varphi_2(\boldsymbol{x})^2 \end{bmatrix} = V \boldsymbol{\varphi}(\boldsymbol{x}).$$
(4.51)

We observe there are two modes,  $[1, 2]^T$  and  $[0, -1]^T$ , which evolve linearly under the nonlinear system (4.47). In addition,  $V\mathcal{J}(\varphi(\mathbf{x})) = I_{2\times 2}$ , hence,  $P(\mathbf{x})$  is an eigenvector for any  $\mathbf{x}$ .

We have shown above the strong relation between time state-space mapping and Koopman eigenfunctions. The following proposition states a limitation between the two notions.

**Proposition 4.26** (Existence of Koopman eigenfunctions with no time state-space mapping). *The state-space mapping is not a necessary condition for the existence of Koopman eigenfunctions.* 

*Proof.* This can be shown by the following simple example. Let us consider the linear system,

$$\frac{d}{dt}\boldsymbol{x} = A\boldsymbol{x}, \quad \boldsymbol{x}(t=0) = \boldsymbol{x}_0, \tag{4.52}$$

where *A* is an  $N \times N$  matrix. For simplicity, we assume the eigenvalues,  $\{\lambda_i\}_{i=1}^N$ , are unique and the eigenvector set,  $\{\nu_i\}_{i=1}^N$ , is orthonormal. Then, the solution of this system of equations can be written as,

$$\mathbf{x}(t) = \sum_{i=1}^{N} b_i \mathbf{v}_i e^{\lambda_i t}, \qquad (4.53)$$

27

where the vector  $\boldsymbol{b} = \begin{bmatrix} b_1 & \cdots & b_N \end{bmatrix}^T$  is chosen according to the initial condition. To form the Koopman eigenfunctions and, correspondingly, the Koopman mode, one should formulate the time state-space mapping. For each eigenvector and eigenvalue of *A* there is a mapping, expressed as,

$$t_i(x) = \frac{1}{\lambda_i} \ln\left(\frac{\boldsymbol{v}_i^T \boldsymbol{x}}{b_i}\right).$$
(4.54)

Thus, the Koopman eigenfunctions are,

$$\varphi_i(\mathbf{x}) = e^{t_i(\mathbf{x})} = \left(\frac{\mathbf{v}_i^T \mathbf{x}}{b_i}\right)^{\frac{1}{\lambda_i}}.$$
(4.55)

This expression can be simplified by applying Def. 4.11, yielding the following system,

$$\frac{d}{dt} \begin{bmatrix} \mathbf{v}_1^T \mathbf{x} \\ \vdots \\ \mathbf{v}_n^T \mathbf{x} \end{bmatrix} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T \mathbf{x} \\ \vdots \\ \mathbf{v}_N^T \mathbf{x} \end{bmatrix}.$$
(4.56)

Note that if the eigenvector,  $v_i$ , is complex then the time state-space mapping, Eq. (4.54), does not exist since it is not well defined. In this case, to create a time state-space mapping, we have to choose one branch from the ln function. However, the Koopman eigenfunction, Eq. (4.55), has a unique value since the exponent cancels the ambiguity of the ln function. It shows that a Koopman eigenfunction can exist in cases where the time state-space mapping does not.

## 5 Koopman Theory for PDE

Let us generalize the results above to the continuous setting of Koopman theory, following Nakao and Mezić (2020). We consider the solution of Eq. (2.6), based on the following assumptions.

**Assumption 5.1** (Proper Operator). The operator  $\mathcal{P}(f(x))$  in Eq. (2.6) is proper.

**Lemma 5.2** (Continuous *u*). *If the operator*  $\mathcal{P}$  *in Eq.* (2.6) *admits Assumption* 5.1 *then the solution is continuous in t.* 

This is quite standard in the theory of PDEs. Basically, letting u(x, t) to be the solution of Eq. (2.6), we can write a first order Taylor expansion for the

variable t as,

$$u(x,t+dt) = u(x,t) + \mathcal{P}(u(x,t)) \cdot dt + o(dt).$$
(5.1)

Since the value of  $\mathcal{P}(u(x,t))$  is finite, we get  $|u(x,t+dt) - u(x,t)| \to 0$  as  $dt \to 0$ .

Assumption 5.3 (Fréchet Differentiability). The operator  $\mathcal{P}$  is Fréchet differentiable a.e. in  $\mathcal{H}$ .

If  $\mathcal{P}$  admits Assumption 5.3 then the solution u(x, t) is in  $C^1$  a.e. with respect to t (see e.g. Venturi and Dektor (2021)).

**Definition 5.4** (Time mapping). Let u(x, t) be the solution of the dynamical system (2.6) where  $t \in I$ . Let  $\Xi(u)$  ( $\Xi$  is capital  $\xi$ ) be a functional mapping from the solution u to t, i.e.

$$t = \Xi(u). \tag{5.2}$$

**Lemma 5.5** (Differentiation of time mapping ). *Let the Assumptions 5.1 and 5.3 hold. If the time mapping, t* =  $\Xi(u)$ *, exists then it admits the following,* 

$$\langle \partial \Xi(u(x,t)), \mathcal{P}(u(x,t)) \rangle = 1 \quad a.e. \text{ in } t \in I.$$
(5.3)

*Proof.* The mapping  $\Xi(u(x))$  is in  $C^1$  a.e. in  $t \in I$  since  $u(x,t) \in C^1$ , *a.e.* with respect to t in I. Based on the Brezis chain rule, the time derivative of the mapping is,

$$1 = \frac{d}{dt}t = \frac{d}{dt}\Xi(u) = \langle \partial \Xi(u(x,t)), \frac{d}{dt}u(x,t) \rangle = \langle \partial \Xi(u(x,t)), \mathcal{P}(u(x,t)) \rangle.$$
(5.4)

And this expression is valid almost everywhere.

**Proposition 5.6** (Condition for the inexistence of a Koopman eigenfunctional). *If there is an equilibrium point in I then a nontrivial Koopman eigenfunctional does not exist.* 

*Proof.* Let  $t_0 \in I$  be an equilibrium point and  $\phi(u(x, t))$  be a Koopman eigenfunctional. Then, u(x, t) = const,  $\forall t \in [t_0, b]$ . Therefore, Eq. (2.9) does not hold for nontrivial  $\phi$  for any  $\lambda \neq 0$ .

#### Remark on dynamics with finite time support

Remark 4.6 is valid also for dynamics of the form of Eq. (2.6). Namely, if there exits a time point,  $T_{ext} \in I$ , for which  $\mathcal{P}(u(x,t)) = 0, \forall t > T_{ext}$ , then there is no Koopman eigenfunctional for this dynamics.

**Lemma 5.7** (Koopman eigenfunctionals induced by a time state-space mapping). Let the Assumptions 5.1 and 5.3 hold and u(x, t) be the solution of Eq. (2.6). If there exists a time mapping,  $t = \Xi(u)$ , then a Koopman eigenfunctional exists.

*Proof.* Given the mapping,  $t = \Xi(u)$ , we define the following functional,

$$\phi(u) = e^{\alpha \Xi(u) + \beta}.$$
(5.5)

The time derivative of this functional is,

$$\frac{d}{dt}\phi(u(x,t)) = \frac{d}{d\Xi}e^{\alpha \cdot \Xi(u(x,t)) + \beta} \frac{d}{dt} \Xi(u(x,t)) = \alpha \phi(u(x,t)) \langle \partial \Xi(u(x,t)), \frac{d}{dt}u(x,t) \rangle$$
$$= \alpha \phi(u(x,t)) \langle \partial \Xi(u(x,t)), \mathcal{P}(u(x,t)) \rangle.$$
(5.6)

According to Lemma 5.5,  $\langle \partial \Xi(u(x,t)), \mathcal{P}(u(x,t)) \rangle = 1$  a.e.. Thus, the function in Eq. (5.5) admits Eq. (**KEFal**) for any value of  $\beta$ , where the corresponding eigenvalue is  $\lambda = \alpha$ .

**Theorem 5.8** (Sufficient condition for the existence of a Koopman eigenfunctional). Let the Assumptions 5.1 and 5.3 hold, let u(x,t) be the solution of Eq. (2.6), and let there be a real function  $f : I \to L$ , for which u(f(t), t) is monotonic with respect to t. Then, Koopman eigenfunctionals exist in the time interval I.

Proof. Let us define the following monotonic function,

$$g(t) = \int_0^L u(x,t)\delta(x - f(t))dx,$$
 (5.7)

where  $\delta$  is the Dirac delta. Then, the time mapping is

$$t = \Xi(u) = g^{-1} \left( \int_0^L u(x, t) \delta(x - f(t)) dx \right).$$
 (5.8)

According to Lemma 5.7 there exits a eigenfunctional, which can be expressed by Eq. (5.5).

## 6 Mode Decomposition based on Time State-Space Mapping

## 6.1 Bridging between nonlinear spectral decomposition and KMD

Let us recall the dynamical system and its suggested form of solution. We consider the following PDE,

$$u_t = P(u), \tag{6.1}$$

where *P* is a nonlinear operator, u(t = 0) = f. The solution of this PDE is approximated as

$$u(x,t) \approx \sum_{i=1}^{m} X_i(x) T_i(t).$$
(6.2)

We would like to mention two principal PDEs for which this approximation is precise (reaches equality). The first one is linear diffusion and the second is TV-flow (see the studies on spectral TV of Gilboa (2014), Burger et al. (2016), Bungert et al. (2019b)). In both cases, the temporal term  $T_i(t)$  are the typical decay profiles of the operator which is dictated by its homogeneity. Whereas the decay profile of linear diffusion is exponential, that of TV-flow is linear. This was generalized by Cohen and Gilboa (2018, 2020), where it is shown there is a smooth transition between exponential and linear decay for  $\gamma$ -homogeneous operators,  $\gamma \in [0, 1)$ ), see Fig. 3.1. These profiles can be calculated by analyzing an evolution initiated with a single (nonlinear) eigenfunction f, admitting  $P(f) = \lambda f$ . In this case it is simple to check that the evolution is structure preserving. That is, the spatial structure of f is maintained and only its contrast changes throughout the evolution. We thus get a separation of variables and can deduce the time profile. It was shown in Bungert and Burger (2019) that the typical decay profile is also the asymptotic behavior of the dynamic (at a time point just before extinction).

In Gilboa (2014), Burger et al. (2016) it was suggested to perform a decomposition of the signal f by identifying phase transitions of the piecewise linear dynamics of TV, or of gradient flows of one-homogeneous functionals in general. This was performed simply by taking the second time derivative of the flow, where the time-weighted expression  $\phi(x, t) = tu_{tt}(x, t)$  was referred to as a spectral component, admitting a simple reconstruction formula,  $f = \int_0^\infty \phi(t) dt$ . In Gilboa (2014) it was shown that not only the initial condition but the entire solution u(x, t) can be expressed as a weighted integration of the spectral

components,

$$u(x,t) = \int_0^\infty H(t,\tau)\phi(x,\tau)d\tau,$$

where  $H(t, \tau) = ((\tau - t)/\tau)^+$ . Comparing  $\phi(x, \tau)$  to  $X_i$  and  $H(t, \tau)$  to  $T_i(t)$  we get an expression similar to (6.2), in an integral form. In Burger et al. (2016) it was shown that for the discrete one dimensional TV-flow the number of components is finite and we can express the solution *u* by a sum of weighted spectral components. One can expand the linear decay profile to an infinite some of Koopman eigenfunctions, as done in Eq. (4.45). Hence we can observe that the nonlinear spectral components  $\phi$  are actually Koopman modes! These relations and connections are planned to be further investigated in a future work.

When the evolution is TV-flow, the set  $\{\phi\}$  is referred to as spectral TV decomposition. In Cohen and Gilboa (2020) the idea was generalized to nonlinear decompositions of  $\gamma$ -homogeneous functionals,  $\gamma \in [1, 2)$ . The typical decay profile is a truncated polynomial with fractional degree almost for every value of  $\gamma$ . Thus, the decomposition was based on fractional calculus, which made this process less accessible numerically.

To bypass the use of fractional calculus it was suggested to apply DMD on the gradient descent of the respective homogeneous functional. As discussed earlier, it was shown that recovering the dynamic with DMD yields an inherent error, Cohen et al. (2021a). A time rescaling method was proposed to improve the DMD decomposition. It was shown theoretically that an evolution of a single eigenfunction is constructed accurately and for general signals improvement in the decomposition was achieved. However, a major problem of phase changes in the flow, due to extinction of modes, was not addressed. This is most inherent in flows based on zero-homogeneous operators, common in signal and image processing. Alternative recent methods were suggested to improve DMD, however none of them tackles well phase transitions in the flow. These methods use machine learning principles in the design of advanced DMD algorithms, such as Extended DMD (EDMD) Williams et al. (2016, 2015a,b) and kernel DMD (KDMD) Kawahara (2016). Several learning-based approaches suggested to build a data-driven dictionary to reconstruct the dynamics sparsely Bollt (2021); Li et al. (2017); Pan et al. (2021); Rudy et al. (2017). These works focus on learning the spatial structures that approximate Koopman modes. In other words, these algorithms aim at finding measurements that evolve linearly under the dynamical system.

Since DMD is primarily investigated in the context of fluid dynamics, oscillatory flows are more common, and less attention was directed to smoothing or decaying flows, which are most common in image and signal processing. We thus aim at extending the Koopman tools to this type of processes. System reconstruction based on finding spatial structures has some limitations, most notably for processes with finitely decaying modes, since the reconstruction of KEFs may be infinite-dimensional. The reconstruction of a KEF as a polynomial of the observation, as in Example 4.9, contains an infinite vector of measurements, which is highly intractable numerically.

Our approach is based on the assumption that the observed dynamic has a typical monotone decay profile within a given time interval. Thus, instead of focusing on measurements that decay exponentially, the focus of our algorithm is on finding spatial structures that decay according to a predefined family of profiles. Let us recall the generalized spectra which was introduced by Katzir (2017) and Gilboa (2018). This work focused on a decomposition induced by the typical decay profile of the respective operator. The spatial structures are deduced from a dictionary containing an overcomplete set of decay profiles. More formally, given a nonlinear dynamic,

$$\frac{d\boldsymbol{x}}{dt} = P(\boldsymbol{x}),$$

with a typical decay profile, a(t), we extract the spatial structure from the solution, x(t), with the following optimization problem,

$$\min \arg_{\mathcal{V}} \{ \| X - \mathcal{VD} \|_{\mathcal{F}} \}, \quad s.t. \min \| \mathcal{V} \|_{0}$$

...

where X,  $\mathcal{V}$  and  $\mathcal{D}$  are defined in Eqs. (2.17), (2.18), and (2.15), respectively. In the rest of this section we show that if the decay profile is monotone then the spatial structures resulting from the general spectral decomposition are the Koopman modes of KMD.

## 6.2 Generalized dynamic mode decomposition

#### Spatiotemporal mode decomposition based on a monotone decay profile

Let us assume the dynamics induces a known typical monotone profile for different spatial structures in the data. The profile, denoted as  $a_{\lambda_i}(t)$ , varies according to the spatial structure,  $v_i$ , and depends on a parameter  $\lambda_i$ . In addition, we assume that the solution can be approximate as,

$$\mathbf{x}(t) = \sum_{i=1}^{N} \mathbf{v}_i \cdot a_{\lambda_i}(t) + e \tag{6.3}$$

where e is a small error term.

Given the time sampling point set  $\{t_i\}_0^M$  (not to be confused with time state-space mapping), we define the overcomplete dictionary,

$$\mathcal{D} = \begin{bmatrix} a_{\lambda_0} (t_0) & \cdots & a_{\lambda_0} (t_M) \\ \vdots & \vdots \\ a_{\lambda_L} (t_0) & \cdots & a_{\lambda_L} (t_M) \end{bmatrix},$$
(6.4)

where *L* is large enough. An atom of this dictionary is a row. Since the time profile  $a_{\lambda_i}(t)$  is monotone there exists an inverse function for each atom, denoted as,

$$t = \xi(a_{\lambda_i}(t)). \tag{6.5}$$

In matrix formulation, for a discrete time setting, this can be written as,

$$\boldsymbol{t} = \boldsymbol{\xi}(\mathcal{D}), \tag{6.6}$$

where  $t \in \mathbb{R}^{(L+1) \times M}$ . It is assumed that there exists a (sparse) mode matrix *V* which can approximate the samples of the system *X* using the dictionary by,

$$X = V\mathcal{D} + e, \tag{6.7}$$

where  $X = \begin{bmatrix} x_0 & \cdots & x_M \end{bmatrix}$  and *e* is a small error term.

#### **Dimensionality Reduction**

Following the assumption of DMD, we would like to obtain a sparse representation of modes. This problem has been thoroughly investigated and can be formulated as Mairal et al. (2014b),

$$\min_{V} \|X - V\mathcal{D}\|_{\mathcal{F}}^{2}, \quad s.t. \|V\|_{0} \le r,$$

$$(6.8)$$

where  $||V||_0 < r$  indicates the requirement that only up to *r* columns in *V* are not zero. This problem is NP-hard and the sparsity constraint is relaxed to solving the following minimization problem,

$$\min_{V} \|X - V\mathcal{D}\|_{\mathcal{F}}^2 + \lambda \|V\|_1.$$
(6.9)

The solution of (6.9) is the minimizer of the left term when the nonzero entries in each mode are at least  $\lambda$  (see algorithm 6 p. 153 in Mairal et al. (2014b)).

In general, there are several well known algorithms to recover the modes when the dictionary is known (see Elad (2010)). We note that our problem is somewhat more difficult than the common signal processing case since the atoms in the dictionary are highly coherent (strongly correlated). Here, we apply the implementation from Mairal et al. (2014a) for the Lasso algorithm (Eq. (6.9)) with a fine-tuning post-processing stage (B) to solve this problem. The output of this algorithm is  $\hat{V}$  and  $\hat{D}$ , where each column in the matrix  $\hat{V}$  contains a mode and  $\hat{D}$  has the corresponding atoms, taken from the dictionary  $\mathcal{D}$ . The entire dynamics can be approximated as,

$$X \approx \hat{V}\hat{\mathcal{D}},\tag{6.10}$$

where  $\approx$  denotes equality in the sense of Eq. (6.9).

### Approximation of Koopman eigenfunctions

Given the modes  $\hat{V}$  and the data matrix X and assuming  $\hat{V}^T \hat{V}$  is invertible, one can express the dictionary as

$$\hat{\mathcal{D}} \approx (\hat{V}^T \hat{V})^{-1} \hat{V}^T X.$$
(6.11)

This reconstruction of the dictionary is necessary to be in the argument of the time state-space mapping, Eqs. (6.5) and (6.6), as follows,

$$\boldsymbol{t} = \boldsymbol{\xi}(\hat{\mathcal{D}}) = \boldsymbol{\xi}((\hat{V}^T \hat{V})^{-1} \hat{V}^T X).$$
(6.12)

Thus, we can express with the dynamic measurements an exponential function. According to Eq. (4.5), the KEFs are given by,

$$\varphi(X) = e^{t(X)} = e^{\xi((\hat{V}^T \hat{V})^{-1} \hat{V}^T X)}.$$
(6.13)

We summarize this algorithm in Algo. 1.

## Algorithm 1 Koopman Mode Approximation

#### 1: Inputs:

Data sequence  $\{x_k\}_0^N$  and typical profile  $a_\lambda(t)$ 

- 2: Find modes  $\hat{V}$  and dictionary  $\hat{D}$  (for example invoke Algo 2).
- 3: Formulate the decay profiles with the modes  $\hat{V}$  and the data *X*, Eq. (6.11).
- 4: Formulate the time state-space mapping, Eq. (6.12).
- 5: Outputs:

Extract KEFs from the observations by Eq. (6.13).

#### Relation between spatiotemporal mode decomposition and KMD

The definition of KMD is to express the state-space vector as spatiotemporal mode decomposition where the temporal terms are exponential functions (KEFs).

This can be done easily by extracting the time variable *t* from Eq. (6.13) and plugging it in Eq. (6.3). Then, the typical decay profile  $a_{\lambda_i}(t)$  can be expressed using a Taylor series (under sufficient smoothness conditions). By variation of parameter, the KMD is obtained (see Example 4.23).

Note that the above presentation is only intended to show a possible algorithmic path that is implied by our analysis. We limit the scope of our discussion here and leave for future work important issues, such as spectrum and system reconstruction accuracy, dimensionality reduction, robustness to noise, and prediction capacity, for more details on these concepts see Gavish and Donoho (2014); Lu and Tartakovsky (2020).

### 7 Examples

In this section, we apply the theory to a few examples. We examine the following: system reconstruction; global controllability; mode decomposition based on a dictionary of monotone profiles; and finding eigenfunctionals in partial differential equations.

**Example 7.1** (System Reconstruction and Global Controllability). This example is based on Mauroy et al. (2020) (p. 10). Given the system,

$$\frac{d}{dt}x(t) = P(x) + u = x - x^3 + u,$$
(7.1)

we would like to obtain global controllability via a Koopman eigenfunction according to Remark 4.19. Note that there are three equilibrium points -1, 0 and 1 with ROAs:  $\mathcal{RA}(-1) = (-\infty, 0)$ ,  $\mathcal{RA}(0) = \{0\}$ , and  $\mathcal{RA}(1) = (0, \infty)$ , respectively. The solution of this equation is,

$$t(x) = \ln\left(\frac{x}{\sqrt{1-x^2}}\right) + C.$$
 (7.2)

According to Theorem 4.8 one of the Koopman eigenfunctions is,

$$\varphi(x) = e^{t(x)} = \frac{x}{\sqrt{1 - x^2}}.$$
 (7.3)

We set the input *u* to,

$$u = -\mathcal{J}(\varphi)^{-1}\varphi + w, \tag{7.4}$$

where w is the input after feedback linearization. The Jacobian matrix is simply

the derivative of  $\varphi$  with respect to x,

$$\mathcal{J}(\varphi) = \left(1 - x^2\right)^{-\frac{3}{2}},\tag{7.5}$$

yielding,

$$u = -\mathcal{J}(\varphi)^{-1}\varphi + w = -\left(1 - x^2\right)^{\frac{3}{2}} \frac{x}{\sqrt{1 - x^2}} + w = -x(1 - x)^2 + w.$$
(7.6)

Substituting this input in the dynamical system, Eq. (7.1), we get the following,

$$\frac{d}{dt}x(t) = P(x) + u = x - x^3 + u = x - x^3 - x(1 - x)^2 + w = w.$$
(7.7)

This system is linear and controllable.

**Example 7.2** (*Total Variation* eigenfunctional). A very common PDE in image processing is the gradient descent flow with respect to the total-variation (TV) functional Bellettini et al. (2002), which for smooth functions u can be expressed as,

$$J_{TV}(u(x)) = \langle |\nabla u(x)|, 1 \rangle.$$
(7.8)

The gradient descent flow for this non-smooth convex functional is defined by,

$$u_t = \mathcal{P} \in -\partial J_{TV}(u), \qquad u(t=0) = u_0, \tag{7.9}$$

where  $\partial J_{TV}(u)$  denotes the subdifferential of TV at u. The flow is known also as the 1-Laplacian flow. When  $x \in \mathbb{R}$  the solution is piece-wise linear, at any time interval  $I_i$  the solution admits, Cohen et al. (2021b),

$$u(x,t) = h_{1,i}(x) + h_{2,i}(x)\lambda_i t.$$
(7.10)

In addition, it was shown by Burger et al. (2016); Cohen et al. (2021b) that the two modes are orthogonal,  $h_{1,j} \perp h_{2,j}$ . Thus, at each interval there are two eigenfunctionals, the trivial one and the second one, corresponding to the linearly evolving mode,

$$\phi(u) = e^{\frac{(h_{2,j},u)}{\|h_{2,j}\|^2 \lambda_j}}.$$
(7.11)

**Example 7.3** (Nonlinear PDE # 2). Let the solution of Eq. (2.6) be,

$$u(x,t) = v_1(x) \cdot a_1(t) + v_2(x) \cdot a_2(t).$$
(7.12)

The solution u(x, t) and the spatial structures  $v_i(x)$ ,  $i \in \{1, 2\}$ , are depicted in Fig. 7.3. The decay profile is of the form of  $a_i(t) = (1 + \lambda_i t)^+$ , where  $\lambda_1 = 1/10$  and  $\lambda_2 = 1/30$ . DMD yields the decomposition depicted in Fig. 7.4. The

37



**Figure 7.3** (a) The solution of Eq. (7.12). On the right, the spatial structures (modes),  $v_1$  (b) and  $v_2$  (c). They evolve with linear decay at a rate of  $\lambda_1 = 1/10$  and  $\lambda_2 = 1/30$ , respectively.

modes are complex and each of them is depicted in two graphs, the real and the imaginary parts (Fig. 7.4a). It demonstrates the limitations of DMD in systems with typical dynamics which are not exponential.



(a) *Dynamic Mode Decomposition*. Two plots on the left: real and imaginary values of the first DMD mode, compared to  $v_1$  (dashed). Two plots on the right: real and imaginary values of the second DMD mode, compared to  $v_2$  (dashed).



**Figure 7.4 Dynamic Mode Decomposition and Reconstruction.** Top row: First two DMD modes, compared to  $v_i$ . Bottom row: Reconstruction through Eq. (2.11) (left) and the corresponding error (right). We can observe the dynamics is not reconstructed well and the error is significant.

The decomposition resulted from Algo. 2 is depicted in Fig. 7.5. The modes are shown in Fig. 7.5a and recover the modes accurately. The entire dynamics reconstruction is given in Fig 7.5b with the corresponding error in Fig. 7.5c. Having the modes, we can find the eigenfunctionals,



(a) Sparse Mode Decomposition - The modes resulting from Algo. 2 (blue) and the actual modes of the dynamics (dashed red).



Figure 7.5 Sparse Mode Decomposition and Reconstruction (Algo. 2) - (a) Sparse mode decomposition compared to the modes,  $v_1$  and  $v_2$ . (b) Dynamic reconstruction with Algo. 2 (Eq. (6.10)) (c) The corresponding error. Correct modes are obtained yielding close to perfect reconstruction of the dynamics.

$$\boldsymbol{\phi}(t) = \begin{bmatrix} \langle v_1, v_1 \rangle & \langle v_1, v_2 \rangle \\ \langle v_2, v_1 \rangle & \langle v_2, v_2 \rangle \end{bmatrix}^{-1} \begin{bmatrix} \langle v_1(x), u(x, t) \rangle \\ \langle v_2(x), u(x, t) \rangle \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$
(7.13)

They are depicted in Fig. 7.6. One can see that the eigenfunctionals are valid until the vanishing points. The first mode vanishes at t = 10 and the second at t = 30.



**Figure 7.6 Eigenfuntionals** - based on the monotone decay profile dictionary. These are the eigenfunctionals formulated in Eq. (7.13).

## 8 Conclusion

This work investigates ways to broaden the use of tools from Koopman theory for the analysis of local and nonlocal PDE's emerging in image and signal processing. We focus on evolution of smoothing processes with possible phase transitions in the dynamics, inherent in zero-homogeneous operators. We discuss necessary and sufficient conditions for the existence of Koopman eigenfunctions. We examine KMD, system reconstruction, global linearity, controllability, and observability through Koopman theory. These insights highlight some limitations of DMD. With the technique of time state-space mapping, we show how conservation laws emerge naturally from any KEF. In addition, we justify the approximation of EDMD based on this mapping.

The classical DMD accurately evaluates KMD as long as KEFs are linear combinations of the observations and KMD is finite-dimensional. However, DMD has clear limitations in four different settings: 1) The typical decay profile of the system is not exponential; 2) One Koopman mode is associated with multiple eigenvalues; 3) There is an equilibrium point in the time interval I; 4) Koopman modes do not exist for all t in I. Another limitation emerges when the dynamic P is in  $C^0$  almost everywhere. In this case, some of the modes might vanish at different times, as we see in the total-variation flow.

We suggest a new type of decomposition to overcome these fundamental problems. It is based on inverse time state-space mapping of injective curves. We implement this method using overcomplete dictionaries of monotone profiles, typical to the dynamics. This decomposition coincides with a basic assumption of DMD – a flow can be sparsely represented by a few dominant modes. We show our decomposition yields Koopman modes. This work can lead to many interesting connections between decomposition, signal representation, nonlinear PDE's and their relation to Koompan theory.

## List of Symbols

$\boldsymbol{x}_i$	The <i>i</i> the sample of the state vector belongs to $\mathbb{R}^N$
Χ	Contains the samples of the dynamics $X =$
	$[\mathbf{x}_0 \cdots \mathbf{x}_M]$ belongs to $\mathbb{R}^{N \times (M+1)}$
U	A matrix where $\mathcal{U}_{i,i} = u(x_i, t_i)$
Н	A auto-correlation matrix of the set $\{h_i\}_{i=1}^M$
Hu	A vector where $Hu_i = \langle h_i(x), u(x, t) \rangle$
h(x)	A vector $\boldsymbol{h}_i(x) = h_i(x)$
V	Contains the main spatial structures $\{v_i\}$
${\mathcal D}$	A dictionary of a family of a decay profile
Р	A (nonlinear) function $P : \mathbb{R}^N \to \mathbb{R}^N$ in $C^1$
	a.e.
t	Time index where $t \in \mathbb{R}^+$
<i>g</i>	This is an observation function of the state vector
	$\boldsymbol{x}, g: \mathbb{R}^N \to \mathbb{R}$
$K_P^{\tau}$	The Koopman operator. The superscript denotes
	the time parameter and the subscript denotes the
	dynamical system
Ι	An interval $[a, b]$ in the time axis
$\varphi(\boldsymbol{x}(t))$	A Koopman eigenfunction
λ	A Koopman eigenvalue
$\nabla$	The gradient of a function
Т	denotes the transform
${\cal H}$	A Hilbert space
$\mathcal P$	An (nonlinear) operator $\mathcal{P}: \mathcal{H} \to \mathcal{H}$
Q	A (nonlinear) proper, lower-semicontinuous
	functional $Q: \mathcal{H} \to \mathbb{R}$
$\phi(\cdot)$	A Koopman eigenfunctional
$\boldsymbol{v}_i$	A preserved spatial shape under the dynamics <i>P</i>
$h_i(x)$	A preserved spatial shape under the dynamics
	$\mathcal{P}$
$a_i(t)$	The time profile corresponding to the <i>i</i> th pre-
	served spatial shape
$\gamma, \gamma - 1$	Denote the homogeneity degrees of a functional
	and its variational derivative, respectively.
$X_0^{M-1}, X_1^M$	Data matrices $[\mathbf{x}_0, \cdots, \mathbf{x}_{M-1}], [\mathbf{x}_1, \cdots, \mathbf{x}_M]$
$U, \Sigma, V$	Singular Value Decomposition (SVD) of $\mathbf{x}_0^{N-1}$

## Ido Cohen & Guy Gilboa

Sub-matrices of $U, V$ containing the first $r$
columns
Sub-matrix of $\boldsymbol{\Sigma}$ containing the most significant
r eigenvalues of the SVD which are the diagonal
of $\Sigma$
The curve in $\mathbb{R}^N$ representing the solution $x$
A mapping from the curve $\mathbf{x}(t)$ to the time
variable <i>t</i>
A Koopman mode
The Jacobian of Koopman mode $\varphi(x)$
A functional mapping from $u(x, t)$ to $t$
An eigenfunctional

## **Appendix A**

## The Dynamic Mode Decomposition steps

#### Coordinate representation

Given N observations of the dynamical system, Eq. (2.1), we form the data matrices as

$$X_0^{M-1} = [x_0, x_1, \cdots, x_{M-1}], \quad X_1^M = [x_1, x_2, \cdots, x_M] \in \mathbb{R}^{N \times M}$$
(A-1)

where  $\mathbf{x}_k = \mathbf{x}(t_k)$ . To find the spatial structures the SVD is applied on the data matrix,

$$X_0^{M-1} = U\Sigma V^*. \tag{A-2}$$

where  $V^*$  is the conjugate transpose of *V*. The columns of *U* span the column space of  $X_0^{M-1}$ . Thus, the spatial structures are represented by its coordinates

$$\boldsymbol{c}_k = U^* \boldsymbol{x}_k. \tag{A-3}$$

#### Dimensionality reduction

Assuming the data is embedded in subspace spanned by the first r columns of U. Then, the coordinates related to that subspace is

$$\boldsymbol{c}_{r,k} = \boldsymbol{U}_r^* \cdot \boldsymbol{x}_k. \tag{A-4}$$

#### Linear mapping

Following the second assumption of the DMD, there is a linear mapping, F, from  $c_{r,k}$  to  $c_{r,k+1}$ . The linear mapping, F, minimizes the DMD error, given by

$$F = \arg\min_{F} \|F \cdot C_{r,0}^{M-1} - C_{r,1}^{M}\|_{\mathcal{F}}^{2},$$
(A-5)

where  $\|\cdot\|_{\mathcal{F}}$  denotes the Frobenius norm and

$$C_{r,0}^{M-1} = U_r^* X_0^{M-1}, \quad C_{r,1}^M = U_r^* X_1^M.$$
 (A-6)

The linear mapping, F, is the optimal linear mapping in the sense of the DMD error, Eq. (A-4), and we write the coordinate dynamic as

$$\boldsymbol{c}_{r,k+1} \approx F \cdot \boldsymbol{c}_{r,k}. \tag{A-7}$$

Then, we can write the dynamic for all k as

$$\boldsymbol{c}_{r,k} \approx F^{\kappa} \cdot \boldsymbol{c}_{r,0}. \tag{A-8}$$

#### Modes, eigenvalues, and coefficients

Now, we would like to summarize the discussion above and to depict the dynamics as a linear system. In general, we can reconstruct a sample at step k from its coordinates as

$$\tilde{\boldsymbol{x}}_k = U_r \boldsymbol{c}_{r,k}.\tag{A-9}$$

In addition, if F is diagonalizable it can be formulated as

$$F = WDW^{-1}, \tag{A-10}$$

where W contains the right eigenvectors of F, and D is a diagonal matrix whose entries are the eigenvalues of F.

Then, the dynamic can be simplified as

$$\tilde{\boldsymbol{x}}_k \approx U_r \cdot F^k \cdot U_r^* \boldsymbol{x}_0 = U_r \cdot W D^k W^{-1} \cdot U_r^* \boldsymbol{x}_0 \tag{A-11}$$

Now, let us define the modes,  $\{\phi_i\}_{i=1}^r$ , eigenvalues,  $\{\mu_i\}_{i=1}^r$ , and coefficients,  $\{\alpha_i\}_{i=1}^r$ .

*Modes* are defined as  $\Phi = \begin{bmatrix} \phi_1 & \cdots & \phi_r \end{bmatrix} = U_r W$ .

*Eigenvalues* are the diagonal entries of the matrix D,  $\{\mu_i\}_{i=1}^r$ .

*Coefficients* are defined by  $\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 & \cdots & \alpha_r \end{bmatrix}^T = W^{-1}U_r^*\boldsymbol{x}_0.$ 

We can now reconstruct the approximate dynamics as,

$$\tilde{\boldsymbol{x}}_k \approx \Phi D^k \boldsymbol{\alpha} = \sum_{i=1}^r \alpha_i \mu_i^k \boldsymbol{\phi}_i.$$
 (A-12)

#### Reconstruction error

Many applications are satisfied with the above step for recovering the spatial structures in the dynamics. However, for recovering the dynamic with DMD another measurement must be considered. To assess the accuracy, not only the "moving" from one sample to the next one should be taken under considerations but also the dynamic in general. Namely, the criterion should be the summation over the distance between  $x_k$  and  $\hat{x}_k$ . For example, the summation over squared

Euclidean distances is resulted in

$$E_{rec} = \sum_{k=0}^{M} \|x_k - \hat{x}_k\|^2 = \|X - \hat{X}\|_{\mathcal{F}}^2$$
(A-13)

which is Frobenius norm of the error.

## **Appendix B**

## Sparse Representation

The main focus should be put on the time profile of the dynamic since the Koopman theory is based on that. In addition, we assume the dynamics induces a family of monotonic time profiles,  $\mathcal{D}$ , where they differ by their parameters. For example, in linear systems, these functions are exponential, in zero-homogeneous dynamical systems the functions are linear with different slops.

We assume the typical decay profile is known and we find the nonzero mode for example with the Lasso algorithm Mairal et al. (2014b). Then, we remove the not relevant modes and the corresponding atoms in the dictionary. We elaborate the algorithm in Algo. 2

## Algorithm 2 Sparse Representation

## 1: Inputs:

Data sequence  $\{x_k\}_0^M$  and decay dictionary  $\mathcal{D}$ 

## 2: Initialize:

 $S\mathcal{R} = \emptyset$ 

- 3: Find the sparse representation V according to Mairal et al. (2014a).
- 4: Let I be the set of indices of the atoms in  $\mathcal{D}$  sorted (from low to high) according to the norm of the modes (column vectors of V).
- 5: Remove from  $\mathcal{I}$  the indices for which the modes are zeros.
- 6: while True do
- 7: Define  $\hat{D}$  as a new dictionary containing the atoms with indices I.

8: Compute 
$$\hat{V} = X\hat{D}^T (\hat{D}\hat{D}^T)^{-1}$$

- 9: Compute the error  $||X \hat{V}\hat{\mathcal{D}}||_{\mathcal{F}}^2$
- 10: Add the set I and its corresponding error to SR
- 11: Remove the first index in I.
- 12: **if** I is empty **then**
- 13: Break
- 14: **end if**
- 15: end while
- 16: Find in SR the set of indices I that yields the minimum error
- 17: Define  $\hat{D}$  as a new dictionary containing the atoms with indices I.
- 18: Compute  $\hat{V} = X\hat{D}^T (\hat{D}\hat{D}^T)^{-1}$
- 19: Outputs:

## References

- Andreu, Fuensanta, Ballester, Coloma, Caselles, Vicent, and Mazón, José M. 2001. Minimizing total variation flow. *Differential and integral equations*, 14(3), 321–360.
- Askham, Travis, and Kutz, J Nathan. 2018. Variable projection methods for an optimized dynamic mode decomposition. *SIAM Journal on Applied Dynamical Systems*, **17**(1), 380–416.
- Azencot, Omri, Yin, Wotao, and Bertozzi, Andrea. 2019. Consistent dynamic mode decomposition. SIAM Journal on Applied Dynamical Systems, 18(3), 1565–1585.
- Bellettini, Giovanni, Caselles, Vicent, and Novaga, Matteo. 2002. The total variation flow in RN. *Journal of Differential Equations*, **184**(2), 475–525.
- Bollt, Erik M. 2021. Geometric Considerations of a Good Dictionary for Koopman Analysis of Dynamical Systems: Cardinality, "Primary Eigenfunction," and Efficient Representation. *Communications in Nonlinear Science and Numerical Simulation*, 105833.
- Brezis, Haim. 1973. Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert. Elsevier.
- Brunton, Steven L, and Kutz, J Nathan. 2019. *Data-driven science and engineering: Machine learning, dynamical systems, and control.* Cambridge University Press. Pages 276–320.
- Brunton, Steven L, Proctor, Joshua L, and Kutz, J Nathan. 2016. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the national academy of sciences*, **113**(15), 3932–3937.
- Brunton, Steven L, Budišić, Marko, Kaiser, Eurika, and Kutz, J Nathan. 2021. Modern Koopman theory for dynamical systems. *arXiv preprint arXiv:2102.12086*.
- Bungert, Leon, and Burger, Martin. 2019. Asymptotic profiles of nonlinear homogeneous evolution equations of gradient flow type. *Journal of Evolution Equations*, 1–32.

- Bungert, Leon, Burger, Martin, and Tenbrinck, Daniel. 2019a. Computing nonlinear eigenfunctions via gradient flow extinction. Pages 291–302 of: *International Conference on Scale Space and Variational Methods in Computer Vision*. Springer.
- Bungert, Leon, Burger, Martin, Chambolle, Antonin, and Novaga, Matteo. 2019b. Nonlinear spectral decompositions by gradient flows of one-homogeneous functionals. *arXiv preprint arXiv:1901.06979*.
- Burger, Martin, Gilboa, Guy, Moeller, Michael, Eckardt, Lina, and Cremers, Daniel. 2016. Spectral decompositions using one-homogeneous functionals. *SIAM Journal on Imaging Sciences*, **9**(3), 1374–1408.
- Chuaqui, Martin. 2018. General criteria for curves to be simple. *Journal of Mathematical Analysis and Applications*, **464**(1), 955–963.
- Cohen, Ido, and Gilboa, Guy. 2018 (Oct.). *Shape Preserving Flows and the p*–*Laplacian Spectra*. working paper or preprint.
- Cohen, Ido, and Gilboa, Guy. 2020. Introducing the p-Laplacian spectra. *Signal Processing*, **167**, 107281.
- Cohen, Ido, Azencot, Omri, Lifshits, Pavel, and Gilboa, Guy. 2021a. Modes of homogeneous gradient flows. *SIAM Journal on Imaging Sciences*, **14**(3), 913–945.
- Cohen, Ido, Berkov, Tom, and Gilboa, Guy. 2021b. Total-Variation Mode Decomposition. Pages 52–64 of: Elmoataz, Abderrahim, Fadili, Jalal, Quéau, Yvain, Rabin, Julien, and Simon, Loïc (eds), *Scale Space and Variational Methods in Computer Vision*. Cham: Springer International Publishing.
- Courant, Richard, and John, Fritz. 2012. *Introduction to calculus and analysis I*. Springer Science & Business Media.
- Dawson, Scott TM, Hemati, Maziar S, Williams, Matthew O, and Rowley, Clarence W. 2016. Characterizing and correcting for the effect of sensor noise in the dynamic mode decomposition. *Experiments in Fluids*, 57(3), 42.
- Elad, Michael. 2010. Sparse and redundant representations: from theory to applications in signal and image processing. Springer Science & Business Media.

- Evangelisti, E. 2011. Controllability and Observability: Lectures given at a Summer School of the Centro Internazionale Matematico Estivo (CIME) held in Pontecchio (Bologna), Italy, July 1-9, 1968. Vol. 46. Springer Science & Business Media.
- Gavish, Matan, and Donoho, David L. 2014. The optimal hard threshold for singular values is  $4/\sqrt{3}$ . *IEEE Transactions on Information Theory*, **60**(8), 5040–5053.
- Gilboa, Guy. 2013. A spectral approach to total variation. Pages 36–47 of: *International Conference on Scale Space and Variational Methods in Computer Vision*. Springer.
- Gilboa, Guy. 2014. A total variation spectral framework for scale and texture analysis. *SIAM journal on Imaging Sciences*, **7**(4), 1937–1961.
- Gilboa, Guy. 2018. Nonlinear Eigenproblems in Image Processing and Computer Vision. Springer.
- Gilboa, Guy, and Osher, Stanley. 2009. Nonlocal operators with applications to image processing. *Multiscale Modeling & Simulation*, **7**(3), 1005–1028.
- Hemati, Maziar S, Rowley, Clarence W, Deem, Eric A, and Cattafesta, Louis N. 2017. De-biasing the dynamic mode decomposition for applied Koopman spectral analysis of noisy datasets. *Theoretical and Computational Fluid Dynamics*, **31**(4), 349–368.
- Kaiser, Eurika, Kutz, J Nathan, and Brunton, Steven L. 2018. Discovering conservation laws from data for control. Pages 6415–6421 of: 2018 IEEE Conference on Decision and Control (CDC). IEEE.
- Kaiser, Eurika, Kutz, J Nathan, and Brunton, Steven. 2021. Data-driven discovery of Koopman eigenfunctions for control. *Machine Learning: Science and Technology*.
- Katzir, Oren. 2017 (March). On the scale-space of filters and their applications.M.Phil. thesis, Technion Israel Institute of Technology, Haifa 3200003.
- Kawahara, Yoshinobu. 2016. Dynamic mode decomposition with reproducing kernels for Koopman spectral analysis. *Advances in neural information processing systems*, **29**, 911–919.

- Koopman, Bernard O. 1931. Hamiltonian systems and transformation in Hilbert space. *Proceedings of the national academy of sciences of the united states of america*, **17**(5), 315.
- Korda, Milan, and Mezić, Igor. 2018. Linear predictors for nonlinear dynamical systems: Koopman operator meets model predictive control. *Automatica*, 93, 149–160.
- Kutz, J Nathan, Brunton, Steven L, Brunton, Bingni W, and Proctor, Joshua L. 2016a. *Dynamic mode decomposition: data-driven modeling of complex systems*. SIAM.
- Kutz, J Nathan, Proctor, Joshua L, and Brunton, Steven L. 2016b. Koopman theory for partial differential equations. *arXiv preprint arXiv:1607.07076*.
- Langley, Pat, Bradshaw, Gary L, and Simon, Herbert A. 1981. BACON. 5: The discovery of conservation laws. Pages 121–126 of: *IJCAI*, vol. 81. Citeseer.
- Li, Qianxiao, Dietrich, Felix, Bollt, Erik M, and Kevrekidis, Ioannis G. 2017. Extended dynamic mode decomposition with dictionary learning: A datadriven adaptive spectral decomposition of the Koopman operator. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, **27**(10), 103111.
- Lu, Hannah, and Tartakovsky, Daniel M. 2020. Prediction accuracy of dynamic mode decomposition. SIAM Journal on Scientific Computing, 42(3), A1639– A1662.
- Mairal, Julien, Bach, F, Ponce, J, Sapiro, G, Jenatton, R, and Obozinski, G. 2014a. Spams: A sparse modeling software, v2. 6. URL http://spams-devel. gforge. inria. fr/downloads. html.
- Mairal, Julien, Bach, Francis, Ponce, Jean, et al. 2014b. Sparse Modeling for Image and Vision Processing. *Foundations and Trends® in Computer Graphics and Vision*, **8**(2-3), 85–283.
- Mauroy, Alexandre. 2021. Koopman Operator Theory for Infinite-Dimensional Systems: Extended Dynamic Mode Decomposition and Identification of Nonlinear PDEs. *arXiv preprint arXiv:2103.12458*.
- Mauroy, Alexandre, Susuki, Y, and Mezić, I. 2020. *The Koopman Operator in Systems and Control*. Springer.
- Mezić, Igor. 2005. Spectral properties of dynamical systems, model reduction and decompositions. *Nonlinear Dynamics*, **41**(1-3), 309–325.

- Nakao, Hiroya, and Mezić, Igor. 2020. Spectral analysis of the Koopman operator for partial differential equations. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, **30**(11), 113131.
- Nathan Kutz, J, Proctor, Joshua L, and Brunton, Steven L. 2018. Applied Koopman theory for partial differential equations and data-driven modeling of spatio-temporal systems. *Complexity*, **2018**.
- Otto, Samuel E, and Rowley, Clarence W. 2021. Koopman operators for estimation and control of dynamical systems. *Annual Review of Control, Robotics, and Autonomous Systems*, **4**.
- Pan, Shaowu, Arnold-Medabalimi, Nicholas, and Duraisamy, Karthik. 2021. Sparsity-promoting algorithms for the discovery of informative Koopmaninvariant subspaces. *Journal of Fluid Mechanics*, 917.
- Rudy, Samuel H, Brunton, Steven L, Proctor, Joshua L, and Kutz, J Nathan. 2017.Data-driven discovery of partial differential equations. *Science Advances*, 3(4), e1602614.
- Schmid, Peter J. 2010. Dynamic mode decomposition of numerical and experimental data. *Journal of fluid mechanics*, **656**, 5–28.
- Schmidt, Michael, and Lipson, Hod. 2009. Distilling free-form natural laws from experimental data. *science*, **324**(5923), 81–85.
- Tu, Jonathan H, Rowley, Clarence W, Luchtenburg, Dirk M, Brunton, Steven L, and Kutz, J Nathan. 2013. On dynamic mode decomposition: Theory and applications. *arXiv preprint arXiv:1312.0041*.
- Valmorbida, Giorgio, and Anderson, James. 2017. Region of attraction estimation using invariant sets and rational Lyapunov functions. *Automatica*, 75, 37–45.
- Venturi, Daniele, and Dektor, Alec. 2021. Spectral methods for nonlinear functionals and functional differential equations. *Research in the Mathematical Sciences*, 8(2), 1–39.
- Williams, Matthew O, Kevrekidis, Ioannis G, and Rowley, Clarence W. 2015a. A data–driven approximation of the koopman operator: Extending dynamic mode decomposition. *Journal of Nonlinear Science*, 25(6), 1307–1346.

- Williams, Matthew O, Rowley, Clarence W, Mezić, Igor, and Kevrekidis, Ioannis G. 2015b. Data fusion via intrinsic dynamic variables: An application of data-driven Koopman spectral analysis. *EPL (Europhysics Letters)*, **109**(4), 40007.
- Williams, Matthew O, Hemati, Maziar S, Dawson, Scott TM, Kevrekidis, Ioannis G, and Rowley, Clarence W. 2016. Extending data-driven Koopman analysis to actuated systems. *IFAC-PapersOnLine*, **49**(18), 704–709.

# Acknowledgements

We would like to thank Prof. Gershon Wolansky from Department of Mathematics, Technion and Dr. Eli Appelboim from Electrical and Computer Engineering Department, and Dan Glaubach for stimulating discussions.