Geometry of the ergodic quotient reveals coherent structures in flows

Marko Budišić *, Igor Mezić

Department of Mechanical Engineering and Center for Control, Dynamical Systems and Computation, University of California, Santa Barbara, USA

ABSTRACT

Dynamical systems that exhibit diverse behaviors can rarely be completely understood using a single approach. However, by identifying coherent structures in their state spaces, i.e., regions of uniform and simpler behavior, we could hope to study each of the structures separately and then form the understanding of the system as a whole. The method we present in this paper uses trajectory averages of scalar functions on the state space to: (a) identify invariant sets in the state space, and (b) to form coherent structures by aggregating invariant sets that are similar across multiple spatial scales. First, we construct the ergodic quotient, the object obtained by mapping trajectories to the space of the trajectory averages of a function basis on the state space. Second, we endow the ergodic quotient with a metric structure that successfully captures how similar the invariant sets are in the state space. Finally, we parametrize the ergodic quotient using intrinsic diffusion modes on it. By segmenting the ergodic quotient based on the diffusion modes, we extract coherent features in the state space of the dynamical system. The algorithm is validated by analyzing the Arnold–Beltrami–Childress flow, which was the test-bed for alternative approaches: the Ulam's approximation of the transfer operator and the computation of Lagrangian Coherent Structures. Furthermore, we explain how the method extends the Poincaré map analysis for periodic flows. As a demonstration, we apply the method to a periodically-driven three-dimensional Hill's vortex flow, discovering unknown coherent structures in its state space. Finally, we discuss differences between the ergodic quotient and alternatives, propose a generalization to analysis of (quasi-)periodic structures, and lay out future research directions.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Historically, a lot of work in dynamical systems focused on understanding a single type of behavior in a model: elliptical zones, hyperbolic trajectories, chaotic attractors, mixing regions, etc. As computers become more powerful, and experimental methods, such as Particle Image/Tracking Velocimetry, more precise, we are able to study large, complicated systems, whose state spaces comprise many different coexisting coherent structures: regions of uniform and simpler behavior, which might be well understood individually. To apply our understanding to the entire system, however, we first need to identify where coherent structures lie in the state space. This task can be made difficult by the quantity of coherent structures, fractal boundaries between them, or the high dimension of the state space.

Over the past decade, several different interpretations of coherent structures have appeared, each serving as a foundation for algorithms which extract them from the flow or from trajectory data. Coherent structures have been identified as: (i) distinguished

trajectories that organize behavior of the flow nearby, (ii) regions of state space between barriers of dynamical transport, (iii) sets of initial conditions that are not dispersed by the flow, and (iv) sets of trajectories on which all flow-invariant functions take constant values. All the mentioned definitions identify sets that are dynamically invariant, in some sense, however, the choice of atomic objects, i.e., barriers of transport vs. initial conditions vs. trajectories, influences the type of approximation we obtain when using algorithms finite in time, space, and precision.

Identification of coherent structures based on barriers to transport includes both classical geometric studies of invariant manifolds and newer approaches that attempt to generalize invariant manifolds to aperiodic, transient flows. The study of Lagrangian Coherent Structures (LCS) [1,2] focuses on structures that take the roles of hyperbolic invariant manifolds in time-varying flows. As a consequence, they are dominantly studied through finite-time Lyapunov exponent fields whose local extrema are tracked to obtain proxies to LCS [3–5].

While LCS-based methods identify distinguished hyperbolic trajectories as organizing structures, a recent paper [6], coauthored by the second author, proposed to identify both hyperbolic and elliptic behavior in a finite-time, transient context. The approach, termed mesochronic velocity analysis, studies the Jacobian of the flow averaged over short segments of trajectories.

* Corresponding author. Tel.: +1 805 893 5095.
E-mail addresses: mbudisic@engr.ucsb.edu (M. Budišić), mezic@engr.ucsb.edu (I. Mezić).

Both LCS and mesochronic velocity fields identify distinguished trajectories that organize surrounding behavior. "Surrounding" is usually interpreted in terms of regions between neighboring distinguished trajectories. Such an interpretation is difficult to generalize to state spaces of higher dimension, where trajectories are not of codimension-one and cannot be used to segment the state space, without additional assumptions, e.g., symmetries of solutions.

Approaches that aggregate smaller flow features into larger coherent structures are the alternative to analyzing barriers that separate trajectories. A prominent aggregation method is the computation of invariant sets through eigenfunctions of the trans-fer (Perron–Frobenius) operator, a linear operator that captures evolutions of densities of points carried by dynamics [7–10]. Nu-merically, the approach most commonly proceeds by the Ulam's method: approximation of the system by a stochastic Markov chain, acting on a discretized state space. The eigenvectors of the Markov chain transition matrix approximate the eigenfunctions of the transfer operator. The Ulam's method requires a compu-tation of short trajectories started from a large number of initial conditions within each discretization cell. Despite its prominence, Ulam's method can be difficult to apply to, e.g., high-dimensional systems, and some experimental setups. A large number of state space dimensions may render computations infeasible, since the number of elements in Ulam's matrix grows exponentially with the dimension of the state space. On the other hand, if resetting and precise preparation of initial conditions is not possible, e.g., in certain fluids experiments, it can be more feasible to run fewer ex-periments for longer periods of time.

The method that does use long trajectories as a direct input has been pursued in [11,12]. It groups trajectories based on values of invariant functions along them. This approach interprets invariant functions as eigenfunctions of the Koopman operator, which captures how values of functions evolve along trajectories of the system. Koopman and the aforementioned Perron–Frobenius operator form a dual pair in the appropriate function spaces. Knowing eigenfunctions of the Koopman operator allows us to detect invariant sets in the state space, as level sets of Koopman eigenfunctions divide the state space into invariant partitions. It is possible to compute them without approximating the operator itself: starting from any observable, i.e., a scalar function on the state space, and averaging its values along trajectories projects the observable to the invariant eigenspace of the Koopman operator. Choosing observables from a function basis, averaging them, and forming joint level sets of averaged observables results in finer and finer stationary partitions of the state space. Ultimately, the process converges to the ergodic partition of the space: the unique, finest partition into invariant stationary sets. It was noted in [13] that even a small number of observables can be used to approximate the ergodic partition well. It remained unclear, however, how to select such a small set of observables that captures a lot of detail except by trial and error.

The algorithm proposed in this paper follows the analysis of dynamics using averaged observables. It does not aim to "guess" an observable whose time average reveals the most information about the coherent structure; rather, it *constructs* a suitable set of invariant functions from averages of a function basis. To explain the method, we expand on the representation of dynamics using the ergodic quotient [13]: each trajectory is represented by a sequence of time-averaged observables taken from a Fourier function basis. Collection of all such sequences, the ergodic quotient, is equivalent to the ergodic partition, but more practical to work with, as analytic tools on sequence spaces are very well studied. To analyze the ergodic quotient, we endow it with a distance function, adapted from [14], that identifies two trajectories as coherent if they spend similar fractions of their evolutions within any particular set in the state space. As a result, the problem of dynamical analysis of flows is converted into a geometric analysis of the ergodic quotient, with a metric structure given by the empirical distance. Our approach is similar in spirit to analysis of integrable systems through geometry of their integrals of motion [15].

To analyze geometry of the ergodic set we used the Diffusion Maps algorithm [16], which retrieves dominant modes of diffusion on objects with metric structure, e.g., branched manifolds. We introduced the use of Diffusion Maps as a tool in [17]; here we explain its results in terms of the ergodic quotient. To validate the algorithm, we apply it to the Arnold–Beltrami–Childress (*ABC*) flow, which was previously used as a test-bed for algorithms based on the Ulam's method [8] and on LCS [18]. The second example, the periodically-forced Hill's vortex flow, demonstrates the application of the algorithm to time-dependent flows.

When applied to analysis of time-periodic dynamical systems, our method extends the familiar analysis of the Poincaré map, including the averaged information about trajectories between two piercings of the Poincaré surface. Building on the application to time-dependent systems, we explain how the algorithm could be easily extended to the analysis of invariant features of periodic and quasi-periodic sets in the state space [12]. In this case, instead of ergodic averages along trajectories, we use harmonic averages along trajectories, which can also be interpreted through eigenspaces of the Koopman operator. While the ergodic averaging projected functions onto invariant eigenspace of the Koopman operator, the harmonic averaging projects functions to eigenspaces associated to other eigenvalues.

There are several benefits to the analysis based on averaging observables along trajectories. Since we use the entire trajectory evolution, we do not need to access the entire state space directly, e.g., we can place initial conditions on a lower-dimensional sur-face and still obtain the full information about coherent struc-tures intersected by that surface. Furthermore, compared to Ulam's method, we require fewer initial conditions for analysis. These fea-tures makes the method applicable to experimental setups where resetting and initialization of experiments are restricted or costly. Moreover, when there are low-dimensional surfaces of interest, such restricted initialization of trajectories makes it possible to analyze even high-dimensional systems without the exponential growth in number of initial conditions. Finally, approximation of eigenfunctions of the Koopman operator does not require passing through a stochastic approximation of the deterministic dynam-ical system; instead projections of observables to eigenfunctions are evaluated directly by averaging.

Although both of our examples in this paper originated as models of fluid flows, the ergodic quotient analysis can be applied to a broad range of dynamical systems on compact state spaces. The incompressible fluid flows and hamiltonian systems, however, are two classes of systems that commonly have many invariant sets in their state spaces, arranged in intricate structure, therefore, those classes provide the most natural targets for our analysis. Nevertheless, dissipative systems can also be studied using the ergodic quotient, in which case the method identifies how basins of attraction are arranged in state spaces.

The paper is structured as follows: the Section 2 presents the theory of the ergodic quotient, the empirical distance, and the diffusion coordinates, followed by the application to the autonomous *ABC* flow. Section 3 demonstrates how application to non-autonomous, periodically forced flows improves on the classical analysis of the Poincarè map and supports the case by a numerical analysis of the periodically-forced 3D Hill's vortex flow. In Section 4 we discuss how our method compares with alternatives, as well as present our suggestions for application and future research. The Appendices analyze numerical properties of the averaging algorithm and the empirical distance, and present details about the Diffusion Maps that are necessary for an implementation as a computer code.
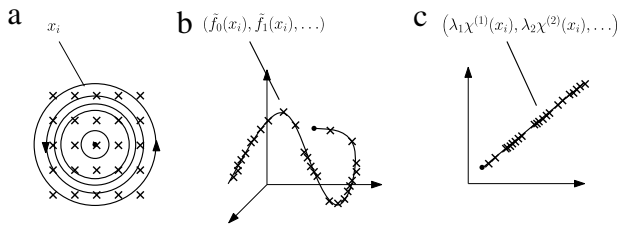
Fig. 1. A sketch of (a) the state space portrait, (b) the ergodic quotient in the space of averaged observables, and (c) the ergodic quotient in the space of diffusion coordinates. Crosses depict initial conditions used for numerical computation. Note that crosses on the ergodic quotient represent entire trajectories started from associated initial conditions on the state space.

2. The ergodic quotient

Geometry of the ergodic quotient enables comparison of trajectories of dynamical systems across multiple spatial scales. To be able to study it, we move through three different spaces: the state space, the space of averaged observables, where the ergodic quotient is constructed, and the space of diffusion coordinates, where the ergodic quotient is described by independent, intrinsic coordinates. Before delving deeper, we give an overview of the process, with italicized terms defined later in the text.

To measure how similar two trajectories are in the state space (Fig. 1.a), we use the *empirical distance*, which compares trajectories based on their residence times in spherical sets in the state space. On the state space, however, the empirical distance is not practically computable, so we map trajectories to a space where it is. By selecting observables f_k from a Fourier function basis and averaging them along trajectories, each trajectory gets mapped to a point on the *ergodic quotient*, which comprises sequences of *trajectory averages*[1] $\tilde{f}_k$ for every trajectory in the state space (Fig. 1.b). We endow the sequence space with a metric induced by a H^{-s} Sobolev space norm. The H^{-s} metric is topologically equivalent to the empirical distance, yet it has a benefit that it is easily computable as a weighted euclidean metric. The ergodic quotient is typically low-dimensional, therefore its representation in a (infinitely-dimensional) sequence space is not economical. To address this issue, we use a nonlinear change of coordinates to represent the quotient in *diffusion coordinates*, preserving the intrinsic geometry of the quotient. Diffusion coordinates are an orthogonal, efficient, scale-ordered, low-dimensional representation of the ergodic quotient that has an effect of straightening the ergodic quotient (Fig. 1.c). The euclidean distance in diffusion coordinates represents the *diffusion distance* along the ergodic quotient. Diffusion distance is an intrinsic distance, locally consistent with empirical distance; for the sake of introduction, it can be thought of as a robust version of the geodesic distance along the quotient.

Computationally, the construction starts with a finite set of trajectories x_i on the state space. They are mapped to the ergodic quotient by averaging a finite subset of the Fourier basis along them, therefore, the points on the numerical ergodic quotient are vectors of trajectory averages. Based on trajectory averages, we compute the matrix of pairwise H^{-s} distances between trajectories, which forms the input data to the Diffusion Maps algorithm. It computes the diffusion coordinates of each trajectory by solving a matrix eigenproblem. Starting trajectories are then grouped into coherent structures, using a *k*-means algorithm applied to the diffusion coordinates of trajectories. To visualize

the coherent structures in the state portrait, we either color trajectories based on their *k*-cluster membership, or we color them based on a single diffusion coordinate.

2.1. Construction: from state space portrait to the ergodic quotient

Take a compact euclidean manifold $\mathcal{M} \subset \mathbb{R}^D$ as the state space of dynamics. The flow map is a composition (semi)-group of continuous functions $\varphi_t(x) : \mathcal{M} \to \mathcal{M}$, with time $t \in \mathcal{T}$. For ODE-generated flows, the time-like parameter is a real number $\mathcal{T} = \mathbb{R}_0^+$, and the semigroup is assumed to be continuous in t. In map-generated flows, time progression over $\mathcal{T} = \mathbb{Z}_0^+$ is interpreted as the iterated application of the map $\varphi(x)$, i.e., $\varphi_t(x) = \varphi[\varphi_{t-1}(x)]$, with $\varphi_0(x) = x$. The space of observables $\mathcal{F}$ is a set of Borel-measurable functions $f : \mathcal{M} \to \mathbb{C}$; in this paper we will take $\mathcal{F}$ as the set of continuous functions on $\mathcal{M}$.

The geometric analysis of dynamics studies the state space portrait of the system, i.e., the collection of all trajectories in the state space, $\left\{ (\varphi_t(x))_{t \in \mathbb{R}} : \forall x \in \mathcal{M} \right\}$. Chaotic behavior, however, implies that individual trajectories are non-robust to errors in initial conditions x. Despite the lack of robustness of trajectories, the averages of observables along them are robust, even in chaotic regimes [19,20]. As our approach is based on averaging, we expect the computations to remain robust even in chaotic regimes, which is the main motivator for this approach. Presented introductory topics follow the exposition given in [21, Section 4].

A finite-time average of an observable f is given by

$$\frac{1}{T} \int_0^T [f \circ \varphi_t] (x) dt.$$

Let $\Sigma \subset \mathcal{M}$ be the set of points on which the limit as $T \to \infty$ exists. On Σ, we can define the infinite time average

$$\tilde{f}(x) \triangleq \lim_{T \to \infty} \frac{1}{T} \int_0^T [f \circ \varphi_t] (x) dt.$$

If the flow map preserves a measure λ, the set $\mathcal{M} \setminus \Sigma$ will be negligible in measure, $\lambda(\mathcal{M} \setminus \Sigma) = 0$, by Birkhoff's Ergodic Theorem. As hamiltonian systems and divergence-free flows preserve the Lebesgue measure, it is unlikely that an initial condition for which time averages do not converge will be selected when analyzing such systems. The same argument can be made for systems that preserve a Lebesgue-continuous measure, which contains an even larger class of systems. Consequently, we will equate Σ and $\mathcal{M}$ for purposes of this paper.

The infinite-time average $f \mapsto \tilde{f}(x)$ is a positive, linear, continuous functional, and the Riesz representation theorem asserts that the averaging functional can be represented by a Borel probability measure μ_x, the *empirical measure*, attached to the initial condition x:

$$\int_{\mathcal{M}} f d\mu_x = \tilde{f}(x). \tag{1}$$

The empirical measures of sets $\mu_x(E)$ can be interpreted as the fraction of evolution that the trajectory evolving from x spends inside the set E (residence or mean-sojourn time). For regular orbits, empirical measures are masses supported on fixed points or spread out along periodic orbits, while on chaotic sets they may have a positive-area support. It is worth noting that empirical measures μ_x are *instantaneous* invariants of the system, despite our use of an asymptotic process in their construction. They play a similar role in the measure-theoretic description of dynamics as the trajectories do in the geometric description.

Instead of working with potentially singular empirical measures directly, we will work with their weak representatives. Fix a countable basis of continuous functions f_k on $\mathcal{M}$, indexed

---

[1] Terms *time average* and *trajectory average* are used interchangeably. In literature, *ergodic average* and *Cesáro average* are also used.

by a multi-index k, e.g., Fourier harmonics with D-dimensional wavevectors $k \in \mathbb{Z}^D$. Let the *ergodic quotient map* be $\pi : \mathcal{M} \to l^\infty(\mathbb{Z}^D)$, given by

$$\pi(x) := (\ldots, \tilde{f}_k(x), \ldots)_k, \quad k \in \mathbb{Z}^D.$$

Action of π at x can be interpreted as forming a weak representation of the empirical measure μ_x. The set of all such sequences, generated by averaging from each initial condition, is the *ergodic quotient* $\xi := \pi(\mathcal{M})$. It will play the same role that the state-space portrait plays in geometric analysis, with the added benefit of robustness in all dynamical regimes.

Strictly, every point on the ergodic quotient should be thought of as the set of trajectory averages of *all* continuous observables: by selecting a basis we merely pick a representation of the ergodic quotient in the particular sequence space $l^\infty(\mathbb{Z}^D)$. It was shown in [12] that all such representations are equivalent; therefore, we will speak of any of them as *the* ergodic quotient when there is no ambiguity.

Conceptually, we have moved from points in the state space x to associated averages of observables in two steps:

$$x \xrightarrow{\text{I}} \mu_x \xrightarrow{\text{II}} (\ldots, \tilde{f}_k(x), \ldots)_{k \in \mathbb{Z}^D}. \tag{2}$$

The representation steps (I) and (II) in (2) are not bijective in general.

Step (I), at the very least, associates all points on the same trajectory with the same empirical measure; however, the pre-images of empirical measures might contain more than a single trajectory. In [17] we discussed the approximation of the ergodic partition, which is the (unique) partition of the state space into ergodic sets. An invariant subset of the state space $S \subset \mathcal{M}$ is called ergodic when the flow φ_t restricted to it, $\varphi_t : S \mapsto S$, is an ergodic system. Since trajectory averages of ergodic systems do not depend on initial conditions, it follows that all points $x \in S$ have the same empirical measure. In this sense, the pre-image of the ergodic quotient is the ergodic partition of the state space. Such an outcome is desirable, as ergodic sets are minimal robust atomic objects that contain (non-robust) trajectories.

Step (II) is bijective only if the entire function basis can be used to construct the ergodic quotient. On continuous state spaces which have infinite bases, we will need to truncate the function set to implement the algorithm on a computer: the magnitude of error introduced by basis truncation will depend on the distance structure on the ergodic quotient, discussed in the Section 2.2.

Remark. Unless we are looking for the ergodic quotient of the entire system, we can place initial conditions only in a region of space that is of a particular interest. Moreover, only one initial condition from an ergodic set is required to represent the empirical measure for the entire ergodic set, as any trajectory in the ergodic set will correctly sample the associated empirical measure. Therefore, it is sufficient to sample the initial conditions from the surface that intersects coherent features we wish to explore. After selecting the initial conditions and the set of observables to average, we compute trajectories starting with each initial condition until time averages along them converge (see Appendix A), thus obtaining the approximate mapping of the state space into the ergodic quotient.

2.2. Metric structure: empirical distance

Trajectories are aggregated into coherent structures by a criterion of similarity embodied in a distance function between trajectories. The choice of the distance function determines the type of coherent structures we obtain. In this work, we chose to work with the empirical distance, which compares two trajectories based on residence times [14]. Instead of working with the

empirical distance directly on the state space, we use a distance on the ergodic quotient that is equivalent to it, thus converting the analysis of dynamics into analysis of geometry of the ergodic quotient.

The empirical distance was originally used to quantify how closely a trajectory samples an a priori distribution on the state space, i.e., to compare an empirical measure with a fixed prior measure on the state space. The same distance can be used to compare two empirical measures and estimate how similar the processes that generate measures are. Define the empirical distance $\mathcal{D}$ between measures μ_x, μ_y by

$$\mathcal{D}(\mu_x, \mu_y)^2 = \int_0^1 \int_{\mathcal{M}} \left\{ \mu_x[B(p, r)] - \mu_y[B(p, r)] \right\}^2 dp dr,$$

where $B(p, r) \subset \mathcal{M}$ are spherical sets of center p and radius r normalized such that $B(p, 1) \supset \mathcal{M}$ and Vol $\mathcal{M} = 1$. Note that integration is over all spherical sets, without normalization by the sets' volumes, resulting in higher sensitivity to differences across larger spatial scales.

Integration over uncountably many spherical sets is practically infeasible, however, the metric induced by the norm on a negative-order Sobolev space of measures can be used instead. The distance

$$\mathcal{D}_{-s}(\mu_x, \mu_y) \triangleq \left\| \tilde{F}(x) - \tilde{F}(y) \right\|_{2, -s}$$

$$\left\| \tilde{F}(x) - \tilde{F}(y) \right\|_{2, -s}^2 = \sum_{k \in \mathbb{Z}^D} \frac{\left| \tilde{f}_k(x) - \tilde{f}_k(y) \right|^2}{\left[1 + (2\pi \|k\|_2)^2 \right]^s}, \tag{3}$$

is easily computed as a weighted sum of Fourier coefficients, $\tilde{F}(x) \triangleq \left(\ldots, \tilde{f}_k(x), \ldots \right)_{k \in \mathbb{Z}^D}$ of measures involved. The order s of the Sobolev space is selected based on the dimension D of the domain of measures $\mathcal{M}$ as $s = (D + 1)/2$. Using $\mathcal{D}_{-s}$ instead of $\mathcal{D}$ does not change the resulting topology, as the empirical metric and the H^{-s} with the order $s = (D + 1)/2$ are equivalent, i.e., there exist positive constants c and C such that for all μ_x and μ_y it holds that

$$c\mathcal{D}(\mu_x, \mu_y) \leq \mathcal{D}_{-s}(\mu_x, \mu_y) \leq C\mathcal{D}(\mu_x, \mu_y).$$

When μ_x and μ_y are empirical measures, the Fourier coefficients can be computed as trajectory averages of observables, due to (1), which is the core observation on which we base our method. Therefore, we choose observables f_k as Fourier basis functions,

$$f_k(x) = (2\pi)^{-D/2} \exp \left[i2\pi \sum_{d=1}^{D} k_d x_d \right], \tag{4}$$

with k_d, and x_d being components of the wave- and state-vectors, respectively, assuming $\mathcal{M} = \mathbb{T}^D$ for simplicity.[2] Interpreted in the ergodic quotient framework, sequences $\tilde{f}_k(x)$ and $\tilde{f}_k(y)$ are points on the ergodic quotient corresponding to empirical measures μ_x and μ_y.

For purposes of our method, it is important to capture the topology and the geometry of the ergodic quotient, not the values of involved quantities. Therefore, we will use the empirical distance $\mathcal{D}$ and the H^{-s} Sobolev space distance $\mathcal{D}_{-s}$ interchangeably: the former in explaining conceptual implications of the metric structure, the latter in computational considerations.

In practice, we cannot compute averages of the entire function basis. The direct effect of the basis truncation is difficult to

---

[2] For arbitrary domains, observables are solutions of the Helmholtz problem on the domain. Mathew and Mezić will address the general formula for the H^{-s} distance in their upcoming work.
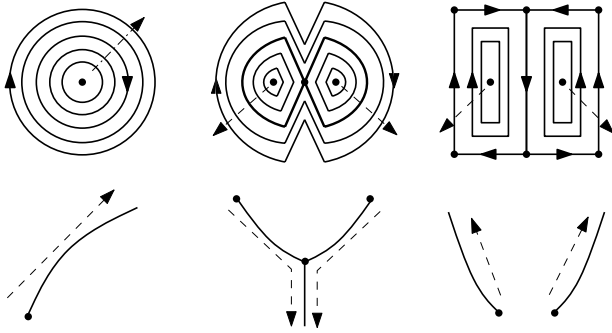
Fig. 2. Geometry of the ergodic quotient under empirical distance. Sketches of matching features in state space portraits (top) and ergodic quotients (bottom) are given from left to right for a harmonic oscillator, a Duffing-type oscillator and a double gyre system. Dots indicate equilibria and dashed arrows indicate how moving an initial condition in the state space moves its image on the ergodic quotient.

quantify; intuitively, the higher the cutoff K is set, restricting wavevectors to $k \in [-K, K]^D$, the finer are the state space features we are able to resolve through empirical distance. In Appendix B we show that truncation introduces at most $\mathcal{O}(1/\sqrt{K})$ error in the distance for any state space dimension D.

Remark. From a signal processing perspective, the H^{-s} distances correspond to a low-pass filtering of empirical distributions prior to their comparison. Since numerical errors are more likely to manifest in averaging of higher-wavenumber modes, we expect that the use of such distances will improve the numerical stability of the analysis. For a brief analysis of these effects, see Appendix A.

2.3. Geometry of the ergodic quotient

Equipping the ergodic quotient with a distance function provides a setting for studying the geometry of the space of invariant functions. The study of dynamics through ergodic quotient shows similarities with the study of dynamics of integrable systems through geometry of integrals of motion. The main difference is that the ergodic quotient is a constructive technique: we do not need to know the expressions for integrals of motion to construct the ergodic quotient or approximate it numerically. Moreover, construction of the ergodic quotient is possible even when the system is not integrable.

In Fig. 2 we demonstrate what the ergodic quotient might look like locally for several simple state-space portraits. In elliptic regions, invariant functions are a one-parameter family, mapping to a filament in the ergodic quotient. In Duffing-type dynamics, one-parameter invariants exist along libration trajectories in elliptic zones and along the revolution trajectories outside of the separatrix. Moving initial conditions close to separatrix from either side results in trajectories slowing down more and more as they pass the saddle, and consequently time averages of an observable will converge to the value that the observable attains at the saddle. As a result, filaments for librations and revolution trajectories meet in the ergodic quotient. In double-gyre flows, the filaments do not meet, since two adjacent elliptic cells share only two out of four saddles; as trajectories slow down next to all four, the time averages will converge to a weighted average of the function values at the saddles, which will be different for every adjoining elliptic cell.

Remark. Fomenko in [15] used a similar construction, the Reeb graphs, to analyze integrable systems. Reeb graphs are graphical representations of connected components in level sets of Morse functions, in particular, the integrals of motion. For an integrable

system, the joint level sets of integrals of motion are ergodic sets. Fomenko did not extend his work to systems for which we do not have expressions for integrals of motion. The ergodic quotient, therefore, can be interpreted as a setting for generalization of Fomenko's approach to non-integrable systems.

2.4. Diffusion modes as global coordinates

Although the ergodic quotient is a subset of an infinite-dimensional sequence space, its main features could be approximated well in a lower number of dimensions, as illustrated in Fig. 2. To obtain a low-dimensional representation, we will convert from the ambient coordinates, where axes are given by averaged observables, to intrinsic coordinates, where axes are independent parameters that parametrize the ergodic quotient. As intrinsic coordinates, we chose the modes of diffusion, or heat spread, along the ergodic quotient, which provide an efficient, orthogonal low-dimensional representation. Numerically, we pass from ambient coordinates to diffusion modes using the Diffusion Maps algorithm [16].

Problems of approximating high-dimensional data by low-dimensional parametrizations are common in data mining and machine learning communities, where a low-dimensional process, e.g., an object moving in the field of vision, is measured by high-dimensional observation, e.g., image consisting of pixels. The usefulness of diffusion eigenfunctions, as a prominent example of Laplacian-based methods, is well documented in such settings, e.g., [22].

On a differentiable manifold, the diffusion operator is given by $A = e^{-t\Delta}$, where Δ is the Laplace–Beltrami operator and t the time over which diffusion evolves. Spectrum of the diffusion operator is contained within the unit interval, with first eigenvalue always $\lambda_0 = 1$, and the rest of the eigenvalues ordered in decreasing order. Eigenfunctions of the diffusion operator $\chi^{(k)}$ are orthogonal and normalized to $\|\chi^{(k)}\|_\infty$ amplitude with the trivial eigenfunction $\chi^{(0)} \equiv 1$. The numerical algorithm, The Diffusion Maps, retrieves the values of eigenfunctions sampled at finite number of points on the underlying manifold.

When we apply Diffusion Maps to sequences of time averages, the ergodic quotient plays the role of the differentiable manifold. The interpretation of numerical results as discretizations of the Laplace–Beltrami eigenfunctions is strictly valid only in the case when ergodic quotient is a differentiable manifold, as it seems to be for some simple systems, but not necessarily in the general case. Despite this fact, this paper demonstrates that Diffusion Maps are practically useful in analysis of dynamical systems and we expect that a consistent interpretation will be found even for sets without a differential structure.

The properties of diffusion eigenfunctions are demonstrated on simple examples in papers [23,24], and detailed analysis of the mathematical properties can be found in [25–27]. In short, we can expect the lowest-index eigenfunctions to be combinations of indicators on disconnected components of the ergodic quotient. The eigenfunctions of higher index behave similarly to Fourier modes on each of the components: first varying over coarser features and, as the order is further increased, varying over finer features.

In the spirit of representation sequence (2), we will add embedding into the eigenfunctions of the diffusion operator as the final step:

$$x \overset{\text{I}}{\to} \mu_x \overset{\text{II}}{\to} (\ldots, \tilde{f}_k(x), \ldots)_{k \in \mathbb{Z}^D}$$
$$\overset{\text{III}}{\to} (\lambda_1 \chi^{(1)}(x), \lambda_2 \chi^{(2)}(x), \ldots),$$

and refer to embedding $x \mapsto (\lambda_1 \chi^{(1)}(x), \lambda_2 \chi^{(2)}(x), \ldots)$ as the *diffusion coordinates* of the dynamics. The sketch of the embeddings was given at the beginning as Fig. 1.
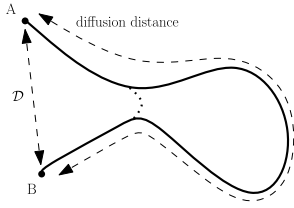
Fig. 3. A sketch of distances in the space of trajectory averages. The full line represents the ergodic quotient, the chordal distance is the empirical metric, and the dotted line indicates the point where a small amount of noise can change the geodesic distance discontinuously, but not the diffusion distance.

The euclidean distance in the diffusion coordinate representation carries a geometric meaning, which is different from, but related to the meaning of the distance in the ambient space of the ergodic quotient, spanned by averaged observables. The empirical distance $\mathcal{D}$ is the chordal distance in the ambient space, i.e., the length of a shortest path between points. Diffusion distance, like geodesic distance, takes into account only paths that lie within the ergodic quotient. However, while the geodesic distance is the length of the shortest path connecting two points, diffusion distance is, in a sense, the average length of all paths that connect the two points. This average is computed as the difference in heat flow reaching two points when heat is injected at any other point on the quotient, averaged over all possible positions of the heat source. A more detailed description can be found in [16].

For our purposes it is sufficient to state that averaging over all paths provides robustness to noise. If the noise introduces a short-cut between two parts of the manifold, as depicted in Fig. 3, the geodesic distance changes discontinuously, however, the diffusion distance does not, as averaging over all paths ensures that a small amount of noise does not change the geometry significantly [24, Section 4.2].

The Diffusion Maps algorithm approximates the operator $e^{-t\Delta}$ by a $N \times N$ matrix, where N is the number of points sampled from the ergodic quotient: in our case equal to the number of initial conditions. Discrete diffusion modes are obtained as eigenvectors of that matrix. The Appendix C provides more details about the implementation: the input to Diffusion Maps is a matrix of pairwise $\mathcal{D}_{-s}$ distances between computed trajectories, while the output is the set of diffusion coordinate vectors, assigning diffusion coordinates to each trajectory.

2.5. Example: steady ABC flow

We will demonstrate the use of the ergodic quotient on the Arnold–Beltrami–Childress (*ABC*) flow [28]. This flow was previously used to demonstrate identification of Lagrangian Coherent Structures in [18] and almost-invariant sets in [8]. It is a steady, volume-preserving flow, evolving on a 3-torus. Despite its very simple system of ODEs, the state space of the *ABC* flow contains both regular vortices and chaotic zones.

Trajectories of the *ABC* flow obey the following equations on $(x, y, z) \in \mathbb{T}^3$

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = A(x, y, z) \triangleq \frac{1}{2\pi} \begin{bmatrix} A \sin 2\pi z + C \cos 2\pi y \\ B \sin 2\pi x + A \cos 2\pi z \\ C \sin 2\pi y + B \cos 2\pi x \end{bmatrix}. \tag{5}$$

As in [18,8], we used $A = \sqrt{3}$, $B = \sqrt{2}$, $C = 1$, but the domain was rescaled to Vol $\mathbb{T}^3 = 1$. Observables are the complex Fourier harmonics (4) with wavevectors $k \in [-10, 10]^3 \cap \mathbb{Z}^3$. The trajectories were integrated using an adaptive solver (Radau) for at least $T = 500$ per initial condition, but possibly longer to achieve

tolerance ATOL $= 2 \times 10^{-4}$ in convergence of averages for each observable and trajectory (see Appendix A).[3]

In short, the steps used for identification of coherent structures are:

1. Select initial conditions x_i,
2. Compute trajectories and time averages until convergence,
3. Compute H^{-s} distance matrix

   $$(d_{ij}) = \mathcal{D}_{-s}(\mu_{x_i}, \mu_{x_j}),$$

4. Feed (d_{ij}) into Diffusion Maps to compute diffusion coordinates $(\lambda_1 \chi_1(x_i), \lambda_2 \chi_2(x_i), \ldots)$,
5. Compute k-means clusters (optional, for visualization).

Connected components in the ergodic quotient correspond to sharply distinct coherent features. Information about all distinct components will rarely be visible from a single, or a few, diffusion coordinates. To quickly recover the dominant features in the flow, we can apply a clustering algorithm. Clustering based on the euclidean distance, e.g., the k-means algorithm, is justified since the euclidean distance approximates the diffusion distance on the ergodic quotient.

The Fig. 4 demonstrates the application of k-means clustering, with $k = 7$, to the first ten diffusion coordinates on the ergodic quotient of the *ABC* flow. The six clusters correspond to six primary vortices that were identified in [28, Fig. 4]. Furthermore, these are the same structures visible in LCS analysis of the flow [18, Fig. 1] and the eigenvectors of the Ulam discretization of Perron–Frobenius operator [8, Fig. 6]. This corroborates our assertion that the components of the ergodic quotient correspond to features typically identified as coherent structures in fluid flows.

In contrast to LCS and Ulam's methods, the initial conditions do not have to be placed in the entire state space, as mentioned in Section 2.1: it is sufficient to place them on a set that intersects interesting coherent structures that we want to analyze or visualize. In Fig. 5 we show a visualization based on $N = 500$ initial conditions sourced from a rectangle $(y, z) \in [0.35, 0.8] \times [0.6, 0.9]$, on the $x = 0$ face of the unit cell. This region intersects one of the six primary vortices shown in Fig. 4a. Since we focused the initial conditions in a smaller region of state space, the same number of diffusion coordinates discerns smaller features in the state space, in this case the secondary 1:2 vortices that wrap around the central vortex.

The Fig. 5c shows that some diffusion coordinates are local-ordering functions on the elliptic zones, acting as computationally constructed local integrals of motion. The 10th ergodic coordinate acts as an energy-like integral of motion, locally on the inner primary vortex. The blue filament in upper-left corner of Fig. 5b indicates that the positive values of the 12th coordinate vary approximately continuously over the secondary vortex lobes, colored in blue in Fig. 5a. This is an indication that a plot similar to Fig. 5c could be made to demonstrate local-ordering on a part of the secondary vortex, using the 12th coordinate.

3. Extensions including time-varying features

3.1. Periodically-driven flows

Although we have presented the construction of the ergodic quotient for the autonomous flows, in practice, nothing prevents

---

[3] The precise value of the tolerance chosen did not make a big difference in data obtained by the method for the *ABC* flow. Reducing tolerance further extends the computation on a relatively small number of trajectories (see Fig. A.9) while most of the trajectories require very similar simulation times even if the tolerance is reduced by an order of magnitude.
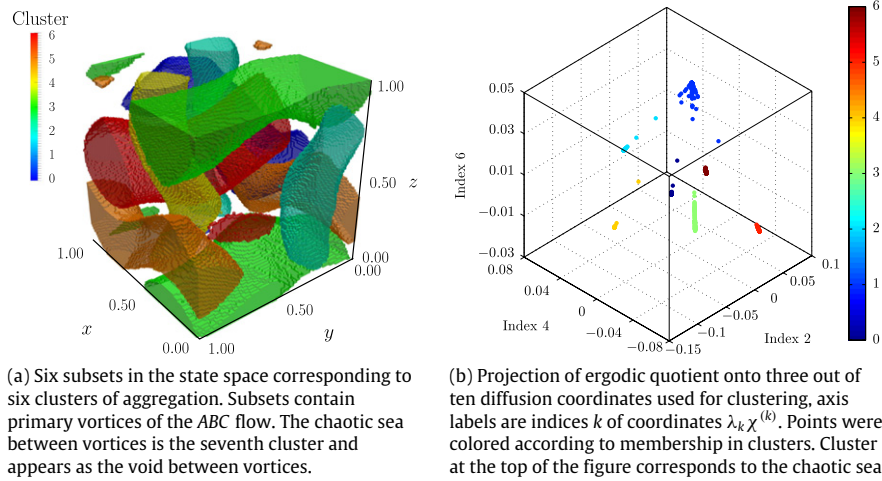
(a) Six subsets in the state space corresponding to six clusters of aggregation. Subsets contain primary vortices of the *ABC* flow. The chaotic sea between vortices is the seventh cluster and appears as the void between vortices.

(b) Projection of ergodic quotient onto three out of ten diffusion coordinates used for clustering, axis labels are indices k of coordinates $\lambda_k \chi^{(k)}$. Points were colored according to membership in clusters. Cluster at the top of the figure corresponds to the chaotic sea.

Fig. 4. Six primary vortices extracted by k-means clustering ($k = 7$) of projection of ergodic quotient onto first 10 diffusion coordinate. *ABC* flow (5) was simulated with $A = \sqrt{3}$, $B = \sqrt{2}$, $C = 1$, from $N = 1000$ initial conditions uniformly distributed in basic periodicity cell $[0, 1)^3$, with observables cut off at wavenumber $K = 10$, and convergence tolerance ATOL = 2×10^{-4}. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(a) Five subsets identified by the k-means algorithm with $k = 5$ acting on first ten diffusion coordinates.

(b) Projection of the ergodic quotient onto three diffusion coordinates, axis labels are indices k of coordinates $\lambda_k \chi^{(k)}$. Points were colored according to membership in clusters.

(c) Level sets of the 10th diffusion eigenvector. The chaotic sea has been cropped for clarity of display by thresholding the first diffusion eigenvector.

Fig. 5. A primary vortex aligned with the x-axis of the *ABC* flow ($A = \sqrt{3}$, $B = \sqrt{2}$, $C = 1$). Lobe-shaped structures that wrap around the central vortex correspond to secondary 1:2 vortices. Results are based on $N = 500$ trajectories initialized uniformly in rectangle $x = 0$, $(y, z) \in [0.35, 0.8] \times [0.6, 0.9]$, with observables cut off at wavenumber $K = 10$, and convergence tolerance ATOL = 2×10^{-4}. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
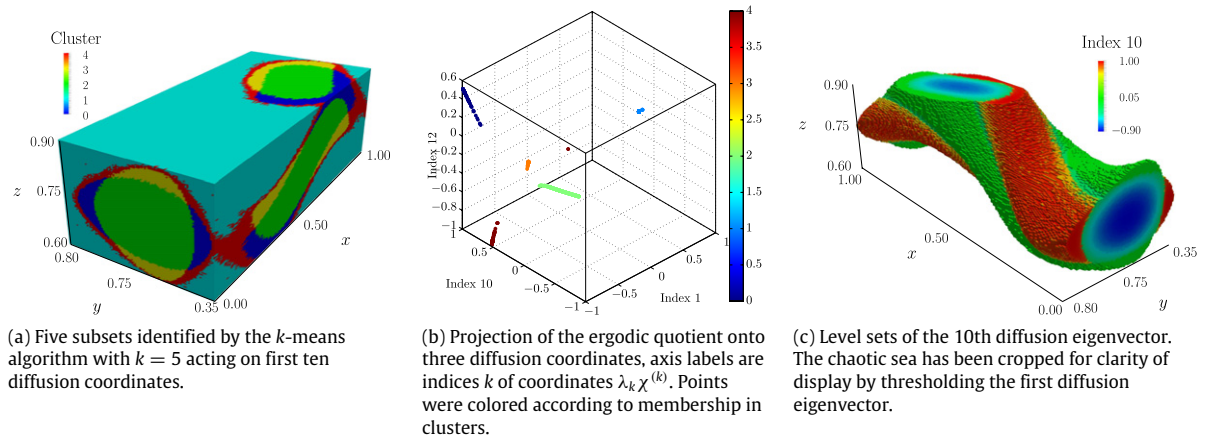
one from applying the exact same computational procedure to a non-autonomous flow. In this Section, we explain that when system is non-autonomous, yet periodically-driven, one still obtains useful and non-trivial information about the structure of invariant sets.

The non-autonomous systems can be embedded in an autonomous system by extending the state space. Let $\mathcal{M}_e = \mathcal{M} \times \mathcal{T}$ be an extension of the state space $\mathcal{M}$. For periodically driven flows, $\mathcal{T}$ can be taken as $\mathbb{T}$, while for aperiodically driven flows and finite-time flows $\mathcal{T}$ can be taken as $\mathbb{R}$ and $[t_i, t_f] \subset \mathbb{R}$, respectively.

For a non-autonomous dynamical system defined by an ODE $\dot{x} = A(x, t)$, where $x \in \mathcal{M}$, we can define the extended, autonomous system through

$$\dot{x} = A(x, \tau),$$
$$\dot{\tau} = c,$$

where $(x, \tau) \in \mathcal{M}_e$ and c is a time-rescaling constant. For periodically-driven systems on compact $\mathcal{M}$, i.e., for $\mathcal{T} = \mathbb{T}$, the extended state space $\mathcal{M}_e$ is also compact. If, additionally, the flow map $\varphi_t : \mathcal{M}_e \to \mathcal{M}_e$ of the system is continuous, the system $(\mathcal{M}_e, \varphi_t)$ fulfills all the conditions that we required for the construction of the ergodic quotient map $\pi_e : \mathcal{M}_e \to l^\infty$, given in Section 2.1.

The difference between an autonomous system on $\mathcal{M}$ and a periodically-driven system cast into an autonomous system on $\mathcal{M}_e$ is in the eye of observer. Typically, we will treat only the first factor of $\mathcal{M}_e$ as "physical" states, always keeping in mind that $\mathcal{T}$ was a formal construct accounting for the "physical" time. This is reflected in observables usually still available as functions $f : \mathcal{M} \to \mathbb{C}$.

If the observables on $\mathcal{M}$ were chosen as Fourier functions $f_k : \mathcal{M} \to \mathbb{C}$, with $k \in \mathbb{Z}^D$, we can interpret them as a subset of Fourier observables $g_w : \mathcal{M}_e \to \mathbb{C}$, $w \in \mathbb{Z}^{d+1}$,

$$g_w(x, \tau) = C f_k(x) e^{i 2\pi w_{d+1} \tau},$$

with the constant C normalizing the observable to $\|g_w\|_\infty = 1$. When wave-vector w is constrained to $w = (k, 0)$, the identity $g_w(x, \tau) \equiv f_k(x)$ holds for any τ. Consequently, the trajectory averages $\tilde{g}_w(x, \tau) = \tilde{f}_k(x)$ are then insensitive to the initial "physical" time τ.

If we proceed with such a partial set of observables g_w, where $w = (k, 0)$, $k \in \mathbb{Z}^D$, we cannot in general hope to obtain the full ergodic quotient by mapping $x \to \tilde{g}_w(x, \tau)$. Denote the partial quotient map by $\pi|_{\mathcal{M}} : \mathcal{M}_e \to l^\infty(\mathbb{Z}^D)$, defined through

$$\pi|_{\mathcal{M}}(x, \tau) := (\ldots, \tilde{g}_w(x, \tau), \ldots)_w,$$
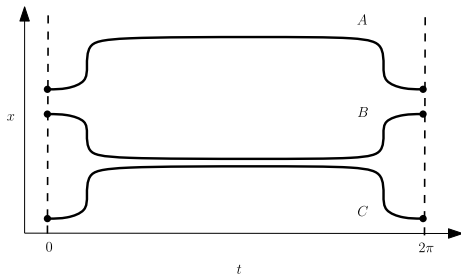
Fig. 6. Three trajectories illustrating the distinction between the direct study of a Poincaré map and coherent features identified by partial ergodic quotient analysis.

for $w = (k, 0)$ and $k \in \mathbb{Z}^D$, with the associated partial quotient $\xi|_{\mathcal{M}} \triangleq \pi|_{\mathcal{M}}(\mathcal{M})$. Looking at pre-images of any trajectory-averaged observable, it follows that $\tilde{g}_w^{-1}(c) = \tilde{f}_k^{-1}(c) \times \mathbb{T}$. Therefore, for any point $p \in \mathbb{Z}^D$, with pre-image $\pi^{-1}(p)$ formed using observables f_k, the pre-image under partial quotient map is $\pi|_{\mathcal{M}}^{-1}(p) = \pi^{-1}(p) \times \mathbb{T}$. Such pre-images are still invariant sets for the dynamics, as they are formed from joint level sets of averaged observables. The difference between the invariant partition corresponding to the ergodic quotient of the extended system ξ_e and the invariant partition corresponding to the partial quotient $\xi|_{\mathcal{M}}$ is that only ξ_e yields the ergodic partition, in general.

In plainer terms: any analysis based on the partial quotient $\xi|_{\mathcal{M}}$ will identify features distinct only in physical state-space $\mathcal{M}$, but not distinctions in physical time. This is sometimes desirable; as a trivial example consider a simple oscillator, $\dot{\theta} = 1$ for $\theta \in \mathcal{M} = \mathbb{T}$, and, even though it is already autonomous, consider its formal extension into $\mathcal{M}_e = \mathbb{T}^2$. Analysis of the partial quotient $\xi|_{\mathcal{M}}$ would identify a single invariant set $\mathcal{M}$, while the analysis of ξ_e would identify a continuum of ergodic sets, each corresponding to a trajectory with a different initial phase. Depending on the application, either of the two analysis can contain valuable information about the analyzed system.

A common technique used in analysis of periodically-driven systems is the Poincaré map $\Phi_{t_0}^n : \mathcal{M} \to \mathcal{M}$ associated to the system:

$$\Phi_{t_0}^n(x) := \varphi_{t_0 + 2\pi n}(x).$$

It is a simple result to show that for continuous flows φ_t, the choice of the initial time (or phase) t_0 results in topologically-conjugate maps, implying that each there exists a homotopy such that each invariant set of Φ_{t_1} has a homotopy-related invariant counterpart in state space of Φ_{t_2}. However, the analysis presented in this paper goes beyond topology, as we study geometry of the features in the state space.

Consider three trajectories sketched in the Fig. 6. If a Poincaré map was taken at $t \equiv 0 (\mathrm{mod}\ 2\pi)$ all of them would be represented as fixed points in state space of the map. Looking at just the Poincaré map, we would be much more likely to identify trajectories A and B as similar, since they pierce the Poincaré surface at nearby locations. However, considering the entire evolution, trajectories B and C deviate significantly only over relatively short time interval. Since time averages capture mean behavior, H^{-s} distance on $\pi|_{\mathcal{M}}$ would (correctly) identify B and C as similar, regardless of the time instance at which Poincaré section used for visualization is taken.

In summary, constructing the partial quotient, i.e., embedding the state space using trajectory averages of observables on the "physical" states $\mathcal{M}$, identifies ergodic partition of the Poincaré map. Additionally, the metric structure on the partial quotient reveals which invariant sets are similar based on the *entire trajectory evolution*, not just on the points at which trajectories pierce the Poincaré surface. Therefore, this method extends the Poincaré map analysis by providing additional information which is potentially useful in identifying coherent structures. To demonstrate, we proceed with a numerical example.

3.2. Example: periodically-driven 3D Hill's vortex flow

As an example of a periodically-driven flow, consider the following ODE, evolving states $x = (R, z, \theta) \subset \mathbb{R}^+ \times \mathbb{R} \times \mathbb{T}$:

$$\begin{bmatrix} \dot{R} \\ \dot{z} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} 2Rz \\ 1 - 4R - z^2 \\ c/2R \end{bmatrix} + \varepsilon \underbrace{\begin{bmatrix} \sqrt{2R} \sin \theta \\ z(\sqrt{2R})^{-1} \sin \theta \\ 2 \cos \theta \end{bmatrix} \sin 2\pi t}_{A(x,t)},$$

where parameters c and ε are swirl and perturbation strengths, respectively. For all values of c and ε the system preserves the state space volume, i.e., $\nabla \cdot A \equiv 0$, where $\nabla = (\partial_R, \partial_z, R^{-1}\partial_\theta)$. This system was studied analytically in [29,30] as an example of a non-hamiltonian system showing Kolmogorov–Arnold–Moser-type structures at low values of $\varepsilon = c$. In this paper, we will show the structure of invariant sets for both low- and high-valued perturbations for the purposes of illustrating the method of analysis; therefore, we will not explore here the bifurcation mechanism behind the structures in the state space. A preliminary numerical study was presented in [31].

The results of the first simulation (Fig. 7) demonstrate the low-perturbation structure of the state space, at $c = \varepsilon = 0.01$. Uniformly distributed initial conditions ($N = 1000$) were seeded on the $\theta = 0$ plane in the rectangle $(R, z) \in [0.01, 0.30] \times [-0.5, 0.5]$, which is within a bounded invariant set. The observables were chosen as Fourier basis functions on $[0.0, 0.5] \times [-1.0, 1.0] \times [0, 2\pi] \subset \mathcal{M}$, with wavevectors in the box $|k_R| \leq 8$, $|k_z| \leq 8$, $|k_\theta| \leq 4$. Simulations were run for no less than $T_{\min} = 600$, and until tolerance ATOL $= 10^{-4}$ was achieved (see Appendix A for definition), with most simulations terminating before 2000 time units expired.

At $c = \varepsilon = 0$, any fixed-θ plane is invariant and on it the system conserves the Hamiltonian $H(R, z) = Rz^2 - R + 2R^2$. The Fig. 7 demonstrates that at low perturbation values, the structure of the cross-section of invariant sets is very similar to the level sets of the Hamiltonian, despite the flow having a full three-dimensional character. At the plotted scale, χ_1 appears to be a monotonic function of the Hamiltonian; however, the non-hamiltonian KAM theory predicts that invariant tori do not form a continuum, but are interspersed with thin chaotic zones. Observables of low wavenumbers effectively coarse-grain dynamics; since the chaotic zones are so thin at low perturbation values, they are not identified as significantly different invariant sets. Plotting higher-order diffusion eigenfunctions against the χ_1 reveals the similarity of diffusion eigenfunctions to the Legendre polynomial system, which is likely due to connection of Legendre polynomials to spherical harmonics and the spherical shape of the Hamiltonian function. In this case it is particularly obvious how monotonic parametrization of the ergodic quotient can serve as a proxy for conserved quantities, possibly leading to Fomenko-type analysis.

When the perturbation is increased, the dissolution of invariant tori continues, but an additional invariant set becomes visible in the state space. For the next simulation, we have chosen $c = \varepsilon = 0.3495$. There is nothing special about this particular parameter value: all perturbations $c = \varepsilon$ between roughly 0.2 and 0.4 exhibit the same topological structure, visible on similar scales. As the region where trajectories stay bounded shrinks with the growth of the perturbation, we have seeded the initial conditions uniformly ($N = 1000$) in $(R, z) \in [0.2, 0.3] \times [-0.1, 0.1]$ in $\theta = 0$ plane, with initial time $t = 0$. The observables were Fourier functions on $[0.0, 0.5] \times [-0.5, 0.5] \times [0, 2\pi)$, with the same wave-vector bounds as before. The simulation times and tolerances also remained as in the low-perturbation case.

The Fig. 8a and b show a secondary coherent set that becomes visible within the onion layers at higher perturbation values. The primary tubular invariant set (in red, Cluster 3) remains at the
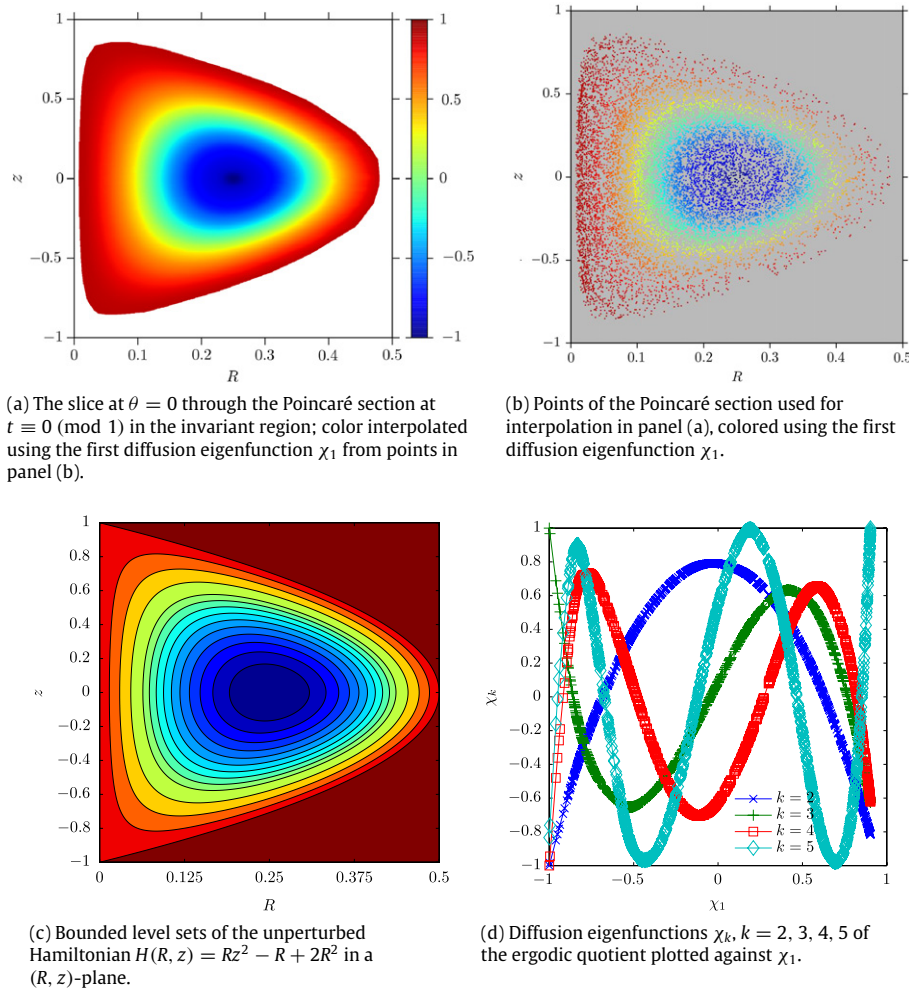
(a) The slice at $\theta = 0$ through the Poincaré section at $t \equiv 0 \pmod 1$ in the invariant region; color interpolated using the first diffusion eigenfunction χ_1 from points in panel (b).

(b) Points of the Poincaré section used for interpolation in panel (a), colored using the first diffusion eigenfunction χ_1.

(c) Bounded level sets of the unperturbed Hamiltonian $H(R, z) = Rz^2 - R + 2R^2$ in a (R, z)-plane.

(d) Diffusion eigenfunctions χ_k, $k = 2, 3, 4, 5$ of the ergodic quotient plotted against χ_1.

Fig. 7. Structure of the invariant sets in the state space of a periodically-forced 3D Hill's vortex at low perturbation $c = \varepsilon = 0.01$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

center, with the secondary invariant set (in blue, Cluster 0), sickle shaped in its cross section, wrapping around the tubular core in 1:2 resonance. Both primary and secondary sets are enclosed by the outer tubular shells (in yellow, Cluster 2). Here, we stress again, that even though we are showing sets in a three-dimensional Poincarè section, they are truly time invariant, despite the time-dependent forcing. More precisely, to each of the invariant sets $C \subset \mathcal{M}$ identified by clustering, there exists a direct extension in the extended state space $C \times \mathbb{T} \subset \mathcal{M}_e$.

The topology of the embedded quotient reflects the change in the invariant set structure, compared to the low perturbation case. While embedding in χ_1 vs. χ_2 in Fig. 7d showed a topological line interval, the high perturbation case in Fig. 8c shows that a branch splits off the interval which was once formed by segments corresponding to clusters that compose the inner and outer tubular regions. The Fig. 8d takes a closer look at diffusion eigenfunctions within only one of the identified clusters. Comparing the shapes of coordinate functions with the low perturbation case in Fig. 7d in interval $\chi_1 \in [0.4, 1]$, we can again notice the similarity to Legendre polynomials, indicating that the flow appears to locally conserve a continuous motion integral, even if that is not the case globally.

3.3. Harmonic quotients

While this paper deals primarily with time-invariant structures, the Section 3.1 describes how even time-varying flows can be

analyzed using trajectory averages, through embedding of the system into an extended state space, expanded by time variable. Fourier observables g on the extended state space $\mathcal{M}_e = \mathcal{M} \times \mathcal{T}$ were formed from Fourier observables $f_k : \mathcal{M} \to \mathbb{C}$ as $g_w(x, t) = f_k(x)e^{i2\pi\omega t}$, where $w = (k, \omega)$ was the extended wave-vector, with ω the wavenumber along time axis. Time averages of observables were given by

$$\tilde{g}_w(x, t) = \lim_{T \to \infty} \frac{1}{T} \int_0^T g[\varphi_\tau(x), \tau] d\tau$$
$$= \lim_{T \to \infty} \frac{1}{T} \int_0^T [f \circ \varphi_\tau](x) e^{i2\pi\omega\tau} d\tau.$$

When $\mathcal{T} = \mathbb{T}$, it was sufficient to consider only $\omega \in \mathbb{Z}$ to obtain a full function basis on $\mathcal{M}_e$. However, when the time is the full real line, as in the aperiodic case, we require a real-valued wavenumber $\omega \in \mathbb{R}$. The complete functional basis on $\mathcal{M}_e$ is then uncountable, and the ergodic quotient of the extended system, evolving in $\mathcal{M}_e$, would not be a subset of a sequence space anymore.

Nevertheless, even in the case of flows φ_t generated by autonomous systems, it is of interest to fix ω and to consider functions $f^{(\omega)}$ given by averages

$$\tilde{f}^{(\omega)}(x) \triangleq \lim_{T \to \infty} \frac{1}{T} \int_0^T [f \circ \varphi_\tau](x) e^{i2\pi\omega\tau} d\tau,$$

termed *harmonic averages*. The harmonic averages were considered by Wiener and Wintner under the name *Fourier averages* [32],
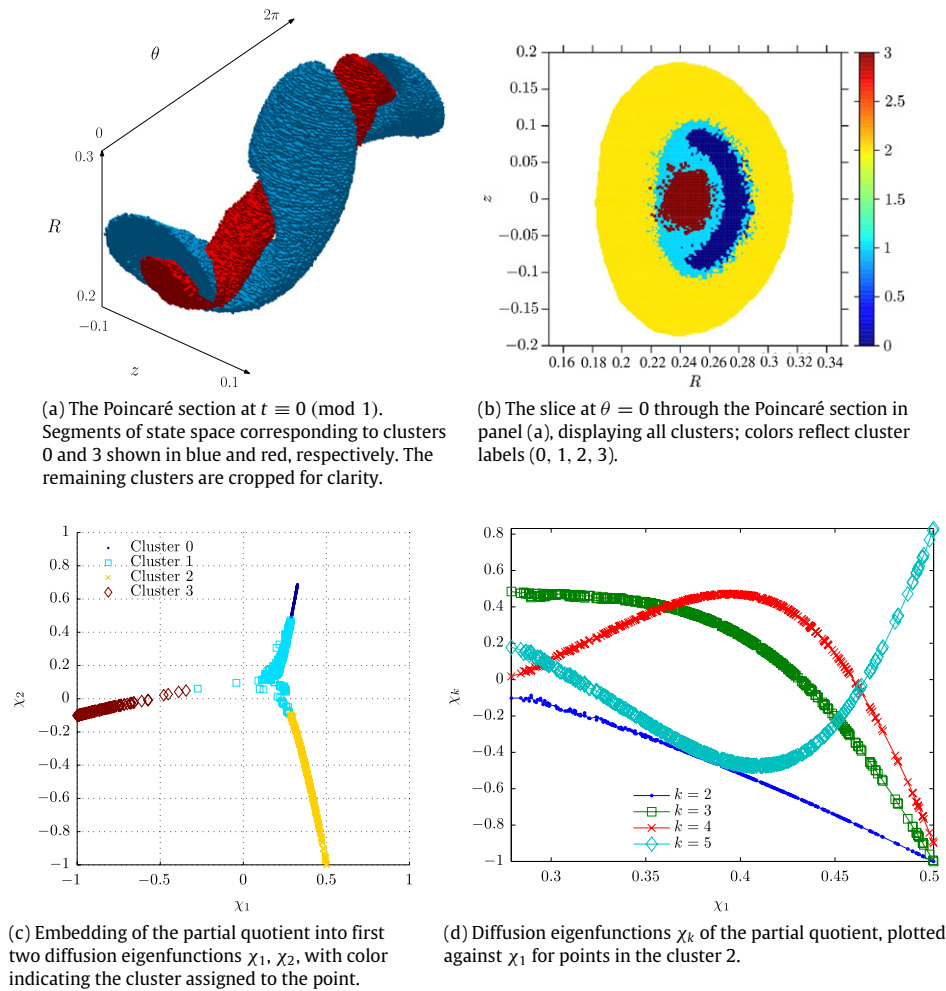
(a) The Poincaré section at $t \equiv 0$ (mod 1). Segments of state space corresponding to clusters 0 and 3 shown in blue and red, respectively. The remaining clusters are cropped for clarity.

(b) The slice at $\theta = 0$ through the Poincaré section in panel (a), displaying all clusters; colors reflect cluster labels (0, 1, 2, 3).

(c) Embedding of the partial quotient into first two diffusion eigenfunctions χ_1, χ_2, with color indicating the cluster assigned to the point.

(d) Diffusion eigenfunctions χ_k of the partial quotient, plotted against χ_1 for points in the cluster 2.

Fig. 8. Structure of the invariant sets of a periodically-forced 3D Hill's vortex at high perturbation $c = \varepsilon = 0.3495$. Diffusion coordinates were clustered using k-means algorithm into 4 clusters to extract the main features. The Poincaré section used for visualization is at $t \equiv 0$ (mod 1). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

but more recently, they have been analyzed in [12]. We use them here to provide a generalization of the ergodic quotients which can be used to analyze periodic and quasi-periodic transport.

For any fixed $\omega \in \mathbb{R}$, and Fourier observables f_k, the embedding $\pi_\omega(x) = (\ldots, \tilde{f}_k^{(\omega)}(x), \ldots)_{k \in \mathbb{Z}^D}$ generates an analog of the ergodic quotient, the *harmonic quotient at frequency ω*. The ergodic quotient is then a special case of the harmonic quotient, computed for $\omega = 0$. In understanding the ergodic quotient, the level sets of averaged observables $\tilde{f}$ played a crucial role as invariant sets. Similarly, level sets of $\tilde{f}^{(\omega)}$ provide insight into periodic sets and periodic transport when $\omega \in \mathbb{Q}$, and wandering sets and quasi-periodic transport when $\omega \notin \mathbb{Q}$ [12]. More concretely, if $\omega = 1/2$, $\tilde{f}^{(\omega)}$ will vanish over all points x that are not φ_t-periodic with period 2. Consequently, the joint level sets of $\tilde{f}_k^{(\omega)}$ away from $(0, 0, \ldots)$ identify sets in state space between which there exists dynamical transport of period 2.

The functional mapping $f \mapsto \tilde{f}^{(\omega)}$ has an interpretation connected to the Koopman operator. For systems satisfying conditions in this paper, the Koopman group consists of linear, bounded composition operators $U_t : \mathcal{F} \to \mathcal{F}$, defined as

$$U_t f \triangleq f \circ \varphi_t,$$

for time $t \in \mathcal{T}$. The space $\mathcal{F}$ is a space of observables, in literature typically chosen as $L^p(\mathcal{M}, \mu)$, for $p = 1, 2$ or ∞, where μ is a φ_t-invariant measure. The Koopman operator is the generator of the

Lie group; when time is discrete, i.e., $\mathcal{T} \subset \mathbb{Z}$, the generator is often denoted simply by U.

The harmonic averages $\tilde{f}^{(\omega)}$ are eigenfunctions of the Koopman operator, since it is simple to show that

$$U_t \tilde{f}^{(\omega)} = e^{i2\pi\omega t} \tilde{f}^{(\omega)}.$$

It follows that, for any observable f, the harmonic average $\tilde{f}^{(\omega)}$ is an eigenfunction of the Koopman group generator at eigenvalue $\lambda_\omega = e^{i2\pi\omega}$ for discrete time, and at eigenvalue $\lambda_\omega = i2\pi\omega$ for continuous time. The operator $P_\omega : \mathcal{F} \to \mathcal{F}$,

$$P_\omega f = \tilde{f}^{(\omega)}$$

is a projection operator, projecting the space of observables onto the eigenspace of the Koopman operator at λ_ω. In this context, the harmonic quotient map π_ω, applied to the Fourier basis, evaluates the functions spanning the eigenspace at λ_ω. Therefore, by endowing the harmonic quotient with a particular geometry, we are able to computationally analyze the eigenspaces of the Koopman operator, without ever computing, or approximating, the Koopman operator itself.

4. Discussion

It is difficult to give a level-field comparison of different methods for identification of "coherent structures" in state spaces

due to the lack of a method-independent, universal definition of what a coherent structure might be. While each of the methods, namely, LCS, Ulam's approximation, and ergodic quotient, has an operational understanding of a coherent structure, ultimately the experience and application is the only judge of how well a method captures what is identified as "coherent" in practical settings. For this reason, we decided to apply the analysis to the *ABC* flow; this was not because we viewed the *ABC* flow as a particularly challenging problem, but rather because it was well analyzed both theoretically and computationally, resulting in well-established knowledge about what coherent structures one should obtain. In this sense, we found that the analysis of ergodic quotient identified the same structures as the theoretical analysis and the alternative computational methods.

When it comes to the amount of work, computational or experimental, needed to compute an approximation to invariant sets, one has to asses the setting in which either of the methods is applied. An advantage of the ergodic quotient method is that it requires a single initial condition to be placed in any ergodic set sought to be identified. In practical terms, this means that simulations or experiments need to be initialized only on a small subset of the entire state space: this was clearly exemplified in Section 3.2 where simulations were initialized on the $\theta = 0$ planes and the method successfully identified coherent structures in the full (R, z, θ) state space. By contrast, methods based on Ulam's approximation require seeding the initial conditions in the entire invariant domain. While this might not be a significant distinction when working with ODEs, in experimental settings, e.g., when analyzing fluid flows using Particle Tracking Velocimetry, it might be possible to initialize the trajectories only in a specific, small region, and count on the long experimental runs to collect data about the rest of the state space. Such data is a natural input to ergodic quotient methods, whereas Ulam-type computations would require additional processing, e.g., flow interpolation, that might introduce further errors.

The number of initial conditions required to approximate the ergodic quotient is fairly low: only $N = 1000$ trajectories were used to retrieve a very detailed image ($S = 200^3$ spatial cells) of the flow structures in Section 2.5. This is a consequence of the recognition that the entire trajectory can be described to a high precision by a relatively small set of values, i.e., averages of a function basis. While the entire trajectory evolution is used for *visualization* of the structures, the *identification* of structures, where the algorithmic effort is really spent, is performed only with N data-points (samples of the ergodic quotient) as inputs.

In terms of numerical linear algebra, we do acknowledge that although the Ulam's matrix is large, $S \times S$, where S is the total number of spatial cells, it is often sparse. In contrast, the eigenproblem at the core of the Diffusion Maps is potentially dense, although on a much smaller $N \times N$ matrix, where N is the number of trajectories.[4] Nevertheless, we see the distinction as conceptual, rather than computational, signifying that the complexity of the state space is often captured by a far smaller number of parameters than needed to describe transport between pairs of cells in state space which is especially true for (near-) integrable, high-dimensional systems. This paper focuses on visualization as the first step in validation of the ergodic quotient approach; however, we expect that the ergodic quotient could be analyzed directly, for example, to detect topological changes such as those between embedding χ_1 vs. χ_2 in Figs. 7d and 8c. These changes indicate a bifurcation of the flow, without ever visualizing the state space. If

our expectations are confirmed, the high-dimensional state spaces, which cannot be visualized and requiring a large number of cells to discretize the state space, could potentially be analyzed through the study of a small number of trajectories by noting topological changes in the ergodic quotient.

Both the Ulam's approximation and the ergodic quotient construction feature exponential growth of problem size with the increase of the state space dimension D. Number of state space discretization cells, for a fixed axial resolution R, grows as R^D. Similarly, to approximate the ergodic quotient, if we truncate the function basis at a uniform cutoff wavenumber K, the number of required Fourier observables is K^D. For certain high dimensional systems, it is not always necessary to analyze the state space directly: rather a particular, low-dimensional output space might suffice. For example, in networks of power generators the variable of interest in detecting instabilities might be the mean frequency across all generators (see [33]). If instead of the state space we study observables only on the output space, the number of required observables can be significantly lower. It is therefore possible that it is sufficient to initialize the trajectories on a low-dimensional set in the state space and analyze it in a low-dimensional output space, despite the dynamics being high-dimensional. Both order reductions are a natural fit with the presented technique in analysis of trajectory averages presented here.

The application of the ergodic quotient techniques to periodically-forced systems has been presented in Section 3. The extension of the method to the case of transient systems, where long-time averages lose the information about the initial evolution, is the focus of our future studies. The case of transient systems is related to the study of finite-time systems, i.e., those defined only on a fixed time interval, which includes all data sets where experiment time cannot be extended into the steady regime, e.g., field samples of geostrophic flows. We do not intend delve deep into a discussion about how much we can predict about the future of an aperiodic flow based only on a finite time interval. Instead, we point out that the analysis presented in this paper does have a very clear interpretation even on a fixed finite-time interval: the finite-time quotient would reflect sets of initial conditions that remain together under advection by the flow map over the interval considered. In that case, the H^{-s} metric and diffusion maps result in a multi-scale aggregation of finite-time trajectories, identifying regions of space where material is transported in a homogeneous fashion. The resulting "coherent structures" present a template of the transport of material over the analyzed time interval. It is left for future research to study the relationship between transport templates corresponding to overlapping or adjoining time intervals and their invariance and predictive power.

Recently, the analysis of mesohyperbolicity and mesoellipticity has yielded success in the study of finite-time flows. Both the ergodic quotient and mesohyperbolicity depend on averages of observables along trajectory lines. The first distinction between these methods is in their purpose: mesohyperbolicity and mesoellipticity are proposed generalizations of concepts of hyperbolicity and ellipticity from autonomous, steady-state flows to unsteady, finite-time flows. In contrast, ergodic quotient is directly related to the ergodic partition, which studies invariant sets. The two concepts are potentially related, as is the case in classical theory of autonomous systems, where the relationship between analysis of invariant manifolds and analysis of Lyapunov coefficients is standard.

The second, more technical, distinction is in the objects averaged along trajectories. In construction of the ergodic quotient, the ultimate goal is to include all observables in the averaging process, in an attempt to represent the underlying empirical measures as well as it is possible. The observables chosen depend only on the topology of the state space, not on the actual flow map of the dynamical system. In studies of mesohyperbolicity/ellipticity, a very

---

[4] Due to the rapidly-decaying exponential diffusion kernel, however, the diffusion eigenproblem might also be sparsified by truncating the kernel. We have not explored the consequences of sparsification, though.

particular set of observables is chosen: they are entries in the Jacobian matrix of the flow. Therefore, while the methods use the same technique of computing averages of observables, the theory behind the averages serves different purposes, and, ultimately, demonstrates the versatility of information that can be extracted from trajectory averages of observables.

From the standpoint of the direct application of methods presented in this paper, there are still points where one could improve the presented procedure, although most of such efforts fall outside the scope of dynamical systems, e.g., into fields of machine learning or computational geometry. For example, setting the bandwidth of the exponential kernel in the Diffusion Maps algorithm (see Appendix C), although relatively simple to perform manually, is still not fully automatic, which might slow down the analysis process when dealing with a particularly complex ergodic quotient.

Furthermore, we were guided mostly by simplicity and clarity in our choice of the *k*-means algorithm for post-processing of diffusion coordinate, in an effort not to derail the main point of this paper. The Fig. 8c reflected the imperfections in the choice of *k*-means clustering as the final step in post-processing. One could expect that the four clusters identified would be the three "branches" and the point at which they merge. While the numerical results come close to such classification of points, there is room for improvement. It is possible that instead of identifying clusters, identification of connected components or one-dimensional sets, based on a proximity graph in a diffusion coordinate space, would yield even better results than those presented here.

5. Conclusion

Our results demonstrate that, by analyzing the ergodic quotient, we can infer how trajectories organize into coherent structures in the state space. To construct an approximation to the ergodic quotient, we average a subset of a function basis on the state space, and endow it with the empirical distance. In this setting, the geometry of the ergodic quotient reveals dynamically coherent regions in the state space. The geometry was studied using diffusion coordinates which rectify the ergodic quotient, enabling clear visualization and extraction of clusters corresponding to coherent structures.

To evaluate the method, we identified coherent structures for the *ABC* flow, an area-preserving, three-dimensional, steady flow. Our results correspond to those coherent structures identified by Lagrangian Coherent Structures [18] and Ulam's method [8]. Unlike those approaches, however, our method does not require placing the initial conditions in the entire state space, but rather depends on long-time integration to explore the state space. For the same reason, it requires a relatively small number of initial conditions. This makes it suitable for analyzing experimentally harvested trajectories in experiments where initial conditions can be placed in a small region of the state space, but the dynamics of the system evolves on a larger domain.

For autonomous flows, points on the ergodic quotient represent ergodic invariant sets in the state space. When the flow is periodically-dependent on time, the construction we present identifies invariant sets of the Poincaré map. In contrast to analysis of Poincaré map, however, coarse graining of invariant sets takes into the account the entire evolution of the trajectory, not only their Poincaré section, providing a more accurate method of identifying trajectories which have similar behavior. On an example of a periodically-forced 3D Hill's vortex flow we show that the method is able to identify previously undescribed coherent features in state spaces. Finally, we close the discussion of time-dependent features by generalizing the ergodic quotient to

harmonic quotients, which can provide information about periodic and quasi-periodic transport in the state space.

Although this paper discusses applications of the ergodic quotient to analysis, we believe that the proposed framework could, in future, be useful in design. Diffusion coordinates provide an orthogonal system of invariant functions for the system. They are, therefore, particularly suitable as a setting for design of invariants of motion, e.g., through optimization of linear combinations of diffusion coordinates based on a design criterion.

Appendix A. Convergence of averages

Our implementation of the averaging process computes averages of observables along with the numerical evolution of the trajectory, i.e., "on-line". A heuristic criterion uses the on-line finite averages to terminate the trajectory integration once a desired tolerance on convergence is achieved. As a result, different initial conditions will result in different lengths of trajectories, depending on the regularity of trajectory that evolves from an initial condition.

The ODE model of a system is numerically integrated producing state vectors $\varphi_{t_n}(x)$ at adaptively-determined time instances t_n. Summing a zeroth-order interpolation of evolution of f_k approximates the continuous-time average along the trajectory:

$$\frac{1}{t_N - t_0} \int_{t_0}^{t_N} [f \circ \varphi_t](x) dt$$

$$\approx \frac{1}{t_N - t_0} \sum_{n=1}^{N} (t_n - t_{n-1})[f \circ \varphi_{t_{n-1}}](x).$$

To determine the stopping time of the integration, we pause the integrator in intervals of fixed length T_e and at every time instance $t_0 + jT_e$ compare the vector of time averages with the vector obtained at the previous pause $t_0 + (j - 1)T_e$:

$$\text{ADIFF}_j(x) = \left\| \tilde{F}_j(x) - \tilde{F}_{j-1}(x) \right\|_\infty$$

where $\tilde{F}_j(x)$ indicates the vector containing all the chosen observables averaged along trajectory starting at x, computed in the $[t_0, t_0 + jT_e]$ interval. When $\text{ADIFF}_j(x) < \text{ATOL}$ for some pre-specified small tolerance ATOL, the integration is stopped, and $F_j(x)$ taken to be the value of time-averaged observables at point x. The underlying assumption is that the finite-time averages will show little variation as the infinite limit is approached. As an alternative to $\|\cdot\|_\infty$, the H^{-s} distance (3) could have been used in the stopping criterion, however, we have not explored that option here.

Theory gives some guidance on times required for convergence of finite-time averages: over regular trajectories, averages converge with T^{-1}, while over mixing trajectories, the slope is gentler, $T^{-1/2}$. A classical result asserts that for other types of trajectories the slope $T^{-\alpha}$ can be arbitrarily small [37, Section 3.2B]. A computational study of convergence with various slopes is given in [13]. From a practical standpoint, the outlook is not so bleak: slopes of small α are to be typically expected to occur along trajectories that

pass close to homoclinic and heteroclinic orbits, and get entrained around periodic orbits that exist in their vicinity (intermittency, or stickiness) [38]. Regions of intermittency are typically small, e.g., for a pendulum perturbed by a small periodic forcing of magnitude ε, the area of intermittent regions scales as $\varepsilon \log(1/\varepsilon)$ [39].

In Fig. A.9 we present several metrics which illustrate how well the time averages converge at the time the simulation is stopped by the ADIFF(x) < ATOL criterion. Time average vectors $\hat{F}_\infty(x)$ and $\hat{F}_{\text{ATOL}}(x)$ both represent finite-time averages, computed for trajectories originating at initial conditions x. The difference between them is in the moment of termination of the averaging process. The vector $\hat{F}_{\text{ATOL}}(x)$ was obtained by stopping the averaging at time $T(x)$, dependent on the initial condition, when the described stopping criterion reached ATOL = 10^{-5}. The vector $\hat{F}_\infty(x)$ mimics the "true" infinity, where averaging was terminated at the initial-condition-independent time $T_\infty = 2.5 \times 10^4$. We compared these finite-time vectors in the L^∞ norm $\|\cdot\|_\infty$ (Fig. A.9b) and in the H^{-s} norm $\|\cdot\|_{2,-s}$ (3) (Fig. A.9c), as the latter is the relevant error for the remainder of the computation.

Generated images demonstrate that the described termination criterion results in stopping times that vary by an order of magnitude (Fig. A.9a). As a result, the errors in convergence are consistently low, for regular, chaotic and intermittent trajectories. If we compare the magnitude of errors in H^{-s}- and L^∞-norms, it is clear that the use of H^{-s} norm, which discounts higher wavenumbers, has a desired side-effect of increasing the robustness of the averaging to finite stopping times.

Even though the proposed criterion is only one of many criteria which could be used for stopping, we believe that, based on the results presented, it provides a good heuristic that consistently produced low errors in time averages, while keeping the simulation times reasonable.

Appendix B. A finite approximation to empirical metric

A finite approximation to the norm on the Sobolev space H^{-s} (3) is computed by truncating the infinite sum over wavevectors $k \in \mathbb{Z}^D$ to a box $k \in [-K, K]^D$. In this appendix we show that such approximation converges with rate $\sqrt{K}$ with the rate constant depending on the dimension of the system.

Proposition 1 (*Error in Finite-Wavenumber Approximation to H^{-s} Norm*)**.** *Let the error in approximation ε_K be given by a weighted sum of coefficients, $c_k = \left| \tilde{f}_k(x) - \tilde{f}_k(y) \right|$, summed outside the $[-K, K]^D$ box:*

$$\varepsilon_K^2 = \sum_{k \notin [-K, K]^D} \frac{c_k^2}{\left[1 + (2\pi \|k\|_2)^2 \right]^s},$$

where $s = (D + 1)/2$, and $0 \le c_k \le \mathcal{C}(D)$ are differences in averages of Fourier observables scaled to unit L_2 norm. Then

$$\varepsilon_K < \frac{\mathcal{E}(D)}{\sqrt{K}}, \quad \text{with } \mathcal{E}(D) < \left(\frac{2}{\pi} \right)^{3/4} \frac{D^{1/4}}{\pi^D},$$

where $\mathcal{E}(D)$ is the error decay rate constant.

Proof. Norms on $\mathbb{Z}^D$ are ordered in scale, $\|k\|_\infty < \|k\|_2$ on $\mathbb{Z}^D$, enabling us to bound ε_K by summing over ∞-spheres in $\mathbb{Z}^D$:

$$\varepsilon_K^2 = \sum_{\kappa > K} \sum_{\|k\|_\infty = \kappa} \frac{c_k^2}{\left[1 + (2\pi \|k\|_2)^2 \right]^s}$$

$$\le \mathcal{C}(D)^2 \sum_{\kappa > K}^\infty \frac{\sigma_\kappa - \sigma_{\kappa-1}}{\left[1 + (2\pi \kappa) \right]^s},$$

where $\sigma_\kappa \triangleq \#\{k \in \mathbb{Z}^D : \|k\|_\infty \le \kappa\}$ count the number of wavevectors in each ∞-ball. To bound σ_κ, we expand the polynomials in κ:

$$\sigma_\kappa - \sigma_{\kappa-1} = (2\kappa + 1)^D - (2\kappa - 1)^D$$

$$= \sum_{d=0}^{D-1} \binom{D}{d} (2\kappa)^d [1 - (-1)^{D-d}]$$

$$\le D \binom{D}{\lfloor D/2 \rfloor} (2\kappa)^{D-1},$$

as there is less than $D/2$ nonzero coefficients, all bounded by the largest coefficient $2 \binom{D}{\lfloor D/2 \rfloor}$.

Choosing order $s = (D + 1)/2$, the error is bounded above by tails of sums of inverse squares

$$\varepsilon_K^2 \le \mathcal{C}(D)^2 D \binom{D}{\lfloor D/2 \rfloor} \sum_{\kappa > K}^\infty \frac{(2\kappa)^{D-1}}{[2\pi\kappa]^{D+1}}$$

$$\le D \binom{D}{\lfloor D/2 \rfloor} \frac{\mathcal{C}(D)^2}{4\pi^{D+1}} \sum_{\kappa > K}^\infty \kappa^{-2}.$$

As $\sum_{\kappa > K}^\infty \kappa^{-2} < \int_K^\infty u^{-2} = K^{-1}$, it follows that $\varepsilon_K \in \mathcal{O}(K^{-1/2})$.

To bound the decay constant $\mathcal{E}(\mathcal{D})$, we use an upper bound for the binomial coefficient, computed in [40]:

$$\binom{mn}{pn} < \frac{1}{\sqrt{2\pi n}} \frac{m^{mn+1/2}}{(m - p)^{(m-p)n+1/2} p^{pn+1/2}},$$

for $m, n, p \in \mathbb{N}$ and $m > p$. Setting $n = \lfloor D/2 \rfloor + 1$, $m = 2$, $p = 1$, provides a conservative estimate:

$$\binom{D}{\lfloor D/2 \rfloor} < \sqrt{\frac{8}{\pi}} \frac{2^D}{\sqrt{D + 1}}.$$

As observables are L_2-normalized Fourier functions (4), $\mathcal{C}(D) = (2\pi)^{-D/2}$, allowing us to compute a bound on $\mathcal{E}^2(D)$:

$$\mathcal{E}^2(D) = D \binom{D}{\lfloor D/2 \rfloor} \frac{\mathcal{C}(D)^2}{4\pi^{D+1}} < \sqrt{\frac{8}{\pi^3}} \frac{\sqrt{D}}{\pi^{2D}},$$

resulting in

$$\mathcal{E}(D) < \left(\frac{2}{\pi} \right)^{3/4} \frac{D^{1/4}}{\pi^D}. \quad \square$$

Appendix C. Diffusion maps

The Diffusion Maps is an algorithm that retrieves an orthogonal coordinate system, the diffusion coordinates, for a set of points in a metric space. If the set of points is taken from a differentiable submanifold in the original space, the retrieved coordinates are discretized eigenfunctions of a Laplace–Beltrami operator on the submanifold. Diffusion coordinates yield a parsimonious description of the data set, i.e., even a truncated diffusion coordinate set provides a good low-dimensional model of the original data set.

The Diffusion Maps is a part of a broader class of Laplacian manifold-learning algorithms, which rely on spectral computations to obtain information about geometry of manifolds. The algorithm was first described by Lafon and Coifman in [23,16] with more detailed discussions on numerical implementation given in [41,24]. To convey the intuition about the algorithm, we provide a short account of the general algorithm and implementation details we found relevant. At the end, we discuss how we applied it to the ergodic quotient.
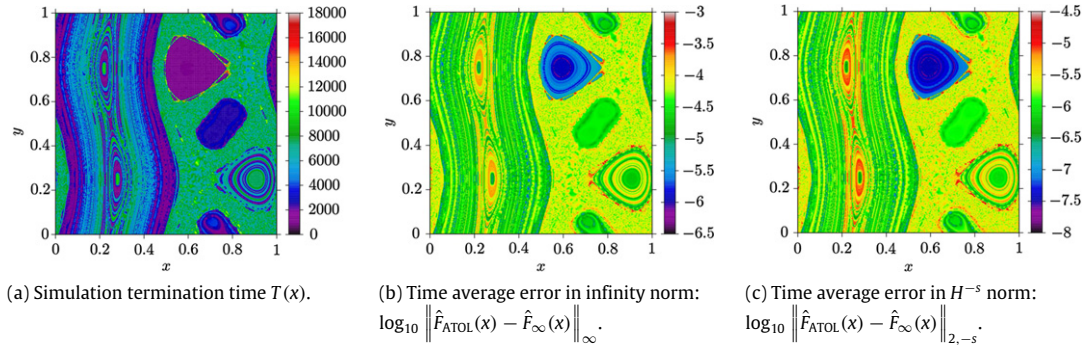
(a) Simulation termination time $T(x)$.

(b) Time average error in infinity norm:
$\log_{10}\left\|\hat{F}_{ATOL}(x) - \hat{F}_\infty(x)\right\|_\infty$.

(c) Time average error in H^{-s} norm:
$\log_{10}\left\|\hat{F}_{ATOL}(x) - \hat{F}_\infty(x)\right\|_{2,-s}$.

Fig. A.9. Convergence of averages for $N = 500$ trajectories of *ABC* flow starting at plane $z = 0$. Simulations were terminated at tolerance ATOL $= 1 \times 10^{-5}$. Figures show: A.9a the termination time $T(x)$, A.9b and c the difference between time averages at time $T(x)$ and at "true" infinity $T_\infty = 2.5 \times 10^4$, in infinity and H^{-s} norms, respectively. Observables were Fourier harmonics (4) with wavevectors $k \in [-3, 3]^3$, and simulation extension time was $T_e = 500$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The ambient space is the space $\mathbb{C}^N$ with a distance function $\mathcal{D}$. We aim to obtain a representation of a subset $\mathcal{Y} \subset \mathbb{C}^N$, with a distribution $dP(y) = p(y)dy$ supported on $\mathcal{Y}$. Averaging operators A_h are defined through the kernel k_h:

$$A_h\varphi(y) = \frac{\int_{\mathbb{C}^N} k_h(y, z)\varphi(z)dP(z)}{\int_{\mathbb{C}^N} k_h(y, z)dP(z)} \tag{C.1}$$

$$k_h(y, z) = (4\pi)^{-d/2} \exp\frac{-\mathcal{D}^2(y, z)}{4h} \tag{C.2}$$

and they represent diffusion of points on $\mathcal{Y}$, while the diffusion time-scale h specifies the width of the Gaussian kernel.[5] The $h \to 0$ limit of the flow $A_h^{t/h}$ has a simple structure when $\mathcal{Y}$ has a well-defined differential structure: it is generated by a Fokker–Planck-type operator G:

$$\lim_{h \to 0} A_h^{t/h} = e^{-tG}$$

$$G = \Delta - \frac{\nabla p}{p},$$

where diffusion evolves over time period t. The Diffusion Maps algorithm computes values of eigenfunctions $\chi^{(k)}(y)$ of e^{-tG} at a finite number of points y_i that are sampled from $\mathcal{Y}$ according to the (sampling) density $p(y)$.

Eigenfunctions of the Laplace–Beltrami operator convey the geometry of the set $\mathcal{Y}$. If understanding geometry of $\mathcal{Y}$ is the goal, as in the case of ergodic quotient, it is necessary to remove the bias in G resulting from non-uniformity in density p. The bias can be removed by replacing the heat kernel k_h by its rescaled version:

$$\hat{k}_h = \frac{k_h(y, z)}{\hat{p}(y)\hat{p}(z)}$$

$$\hat{p}(y) = \int k_h(y, z)dz,$$

where $\hat{p}$ acts as an estimate for the sampling density. In this case, the flow in the $h \to 0$ limit represents the pure Laplace–Beltrami diffusion on $\mathcal{Y}$, i.e., $G = \Delta$.

To compute the diffusion coordinates, defined as the embedding of points y_i into eigenvalue-scaled eigenfunctions of the diffusion operator $y \to (\lambda_1\chi^{(1)}(y), \lambda_2\chi^{(2)}(y), \ldots)$, Diffusion Maps treats the finite set of sampling points $y_i \in \mathbb{C}^N$ as vertices of a fully-connected graph in $\mathcal{Y}$, while weights of edges are given by the ambient distance $\mathcal{D}$ on $\mathbb{C}^D$, between pairs of points y_i. The diffusion process on the graph can be represented by a Markov chain, whose eigenvectors $\chi^{(k)}$ are discretizations of eigenfunctions of the diffusion on $\mathcal{Y}$. In other words, we construct a discrete embedding $y_i \to (\lambda_1\chi_i^{(1)}, \lambda_2\chi_i^{(2)}, \ldots, \lambda_N\chi_i^{(N)})$, assuming $\left\|\chi^{(k)}\right\|_\infty = 1$.

From the standpoint of the implementation, Diffusion Maps takes a matrix of pairwise distances between points y_i as the input, and computes diffusion coordinates $\lambda\chi$ following these steps:

$$\mathcal{D} \text{ matrix} \quad (d_{ij}^2) = \mathcal{D}^2(y_i, y_j), \tag{C.3a}$$

$$\text{Heat kernel} \quad A_{ij} = \exp\left[-\frac{d_{ij}^2}{4h}\right], \tag{C.3b}$$

$$\text{Est. sampling density} \quad \hat{p}_i = \sum_{j=1}^N A_{ij}, \tag{C.3c}$$

$$\text{Unbiased heat kernel} \quad \hat{A}_{ij} = \frac{A_{ij}}{p_i p_j}, \tag{C.3d}$$

$$\text{State transition} \quad S_{ij} = \frac{\hat{A}_{ij}}{\sum\limits_{j=1}^N \hat{A}_{ij}}, \tag{C.3e}$$

$$\text{Solve the eigenproblem} \quad S\chi = \lambda\chi, \tag{C.3f}$$

with $i, j \in 1, 2, \ldots, N$ indexing elements of matrices and vectors, and $\mathcal{D}_{-s,K}$ representing summation (3) truncated to $k \in [-K, K]^D$.

The time-scale of the heat kernel h in (C.3b) is a parameter that determines the strength of diffusion along graph that samples $\mathcal{Y}$. It influences the information about topology of $\mathcal{Y}$ that we infer from discrete data: too small of a choice of h will introduce artificial discontinuities, while too large of a choice will artificially connect disconnected components of $\mathcal{Y}$. Paper [24, Section 5.3] describes several heuristics for determining h from the data set; we chose the *Neighborhood Size Stability* approach. It computes the minimal time-scale h for which every graph vertex has N_{min} neighbors within the characteristic diffusion distance $\sqrt{2h}$, as measured by the H^{-s} distance. Number N_{min} is chosen by the user, depending on number of discrete samples analyzed.

Since S is not symmetric, complex values appear as partial results in numerical solutions of its eigenproblem, which is a source of further numerical errors. To avoid such issues, it is common to first symmetrize the matrix, using a spectrum-preserving symmetrization:

$$\hat{S}_{ij} = \frac{\hat{A}_{ij}}{\sqrt{\sum\limits_{k=1}^N \hat{A}_{ik}}\sqrt{\sum\limits_{k=1}^N \hat{A}_{kj}}}.$$

---

[5] Sometimes $1/h$ is referred to as the *bandwidth*.

Solving the eigenproblem for $\hat{S}$, we obtain eigenvectors $\hat{\chi}^{(k)}$ from which we recover the eigenvectors that sample the diffusion eigenfunctions (C.3f) by rescaling $\hat{\chi}^{(k)}$ by the zeroth eigenvector $\chi_i^{(k)} = \hat{\chi}_i^{(k)}/\hat{\chi}_i^{(0)}$. Vectors $\chi^{(k)}$ are point-wise evaluations of diffusion eigenfunctions $\chi^{(k)}(y)$ at y_i, i.e., $\chi_i^{(k)} = \chi^{(k)}(y_i)$.

In application of Diffusion Maps to the ergodic quotient, initial conditions $x_i \in \mathcal{M}$ and associated trajectories get mapped to the points y_i in the space $\mathbb{C}^N$ of averaged observables, with dimension N given by the number of chosen observables. The ambient distance $\mathcal{D}$ is the H^{-s} distance, while the set $\mathcal{Y}$ is the ergodic quotient ξ. The sampling density $p(y)$ is determined both by the distribution of initial conditions x_i of trajectories on the state space, which can be controlled, and the distribution of the ergodic sets in the state space, which is a priori unknown and cannot be controlled. As a consequence, $p(y)$ is rarely uniform and the rescaling of the diffusion kernel is a necessary step in the algorithm to obtain the intrinsic coordinate set with the correct interpretation of the euclidean distance.

References

[1] G. Haller, G. Yuan, Lagrangian coherent structures and mixing in two-dimensional turbulence, Physica D 147 (2000) 352–370.

[2] G. Haller, A variational theory of hyperbolic Lagrangian coherent structures, Physica D 240 (2011) 574–598.

[3] G. Haller, Lagrangian coherent structures from approximate velocity data, Phys. Fluids 14 (2002) 1851–1861.

[4] S.C. Shadden, F. Lekien, J.E. Marsden, Definition and properties of Lagrangian coherent structures from finite-time Lyapunov exponents in two-dimensional aperiodic flows, Physica D 212 (2005) 271–304.

[5] B. Joseph, B. Legras, Relation between kinematic boundaries, stirring, and barriers for the Antarctic polar vortex, J. Atmos. Sci. 59 (2002) 1198–1212.

[6] I. Mezić, S. Loire, V.A. Fonoberov, P.J. Hogan, A new mixing diagnostic and Gulf oil spill movement, Sci. Mag. 330 (2010) science.1194607v2.

[7] G. Froyland, M. Dellnitz, Detecting and locating near-optimal almost-invariant sets and cycles, SIAM J. Sci. Comput. 24 (2003) 1839–1863. (electronic).

[8] G. Froyland, K. Padberg, Almost-invariant sets and invariant manifolds-connecting probabilistic and geometric descriptions of coherent structures in flows, Physica D 238 (2009) 1507–1523.

[9] G. Froyland, N. Santitissadeekorn, A. Monahan, Transport in time-dependent dynamical systems: finite-time coherent sets, Chaos 20 (2010) 043116.

[10] G. Froyland, O. Junge, P. Koltai, Estimating long term behavior of flows without trajectory integration: the infinitesimal generator approach, 2011. arXiv:1101.4166v1.

[11] I. Mezić, S. Wiggins, A method for visualization of invariant sets of dynamical systems based on the ergodic partition, Chaos 9 (1999) 213–218.

[12] I. Mezić, A. Banaszuk, Comparison of systems with complex behavior, Physica D 197 (2004) 101–133.

[13] Z. Levnajić, I. Mezić, Ergodic theory and visualization, I, mesochronic plots for visualization of ergodic partition and invariant sets, Chaos 20 (2010) 19.

[14] G. Mathew, I. Mezić, Metrics for ergodicity and design of ergodic dynamics for multi-agent systems, Physica D 240 (2011) 432–442.

[15] A.V. Bolsinov, A.T. Fomenko, Integrable Hamiltonian Systems: Geometry, Topology, Classification, CRC Press, 2004.

[16] R. Coifman, S. Lafon, Diffusion maps, Appl. Comput. Harmon. Anal. 21 (2006) 5–30.

[17] M. Budišić, I. Mezić, An approximate parametrization of the ergodic partition using time averaged observables, in: Proceedings of the 48th IEEE Conference on Decision and Control, Held Jointly With the 2009 28th Chinese Control Conference, CDC/CCC 2009, 2009, pp. 3162–3168.

[18] G. Haller, Distinguished material surfaces and coherent structures in three-dimensional fluid flows, Physica D 149 (2001) 248–277.

[19] H. Sigurgeirsson, A.M. Stuart, Statistics from computations, in: Foundations of Computational Mathematics, in: LMS Lecture Note Series, vol. 284, Cambridge University Press, 2001, pp. 323–344.

[20] P. Cvitanović, R. Artuso, R. Mainieri, G. Tanner, G. Vattay, Chaos: classical and quantum, Niels Bohr Institute, Copenhagen, stable thirteenth ed., 2010.

[21] A. Katok, B. Hasselblatt, Introduction to the Modern Theory of Dynamical Systems, Cambridge University Press, New York, NY, USA, 1995.

[22] U. Von Luxburg, A tutorial on spectral clustering, Stat. Comput. 17 (2007) 395–416.

[23] R. Coifman, S. Lafon, A.B. Lee, M. Maggioni, B. Nadler, F. Warner, S.W. Zucker, Geometric diffusions as a tool for harmonic analysis and structure definition of data: diffusion maps, Proc. Nat. Acad. Sci. 102 (2005) 7426–7431.

[24] A.B. Lee, L. Wasserman, Spectral connectivity analysis, J. Amer. Stat. Assoc. 105 (2010) 1241–1255.

[25] P.W. Jones, M. Maggioni, R. Schul, Manifold parametrizations by eigenfunctions of the Laplacian and heat kernels, Proc. Nat. Acad. Sci. 105 (2008) 1803–1808.

[26] D. Jakobson, N. Nadirashvili, J. Toth, Geometric properties of eigenfunctions, Russian Math. Surveys 56 (2001) 1085.

[27] K. Uhlenbeck, Generic properties of eigenfunctions, Amer. J. Math. 98 (1976) 1059–1078.

[28] T. Dombre, U. Frisch, J.M. Greene, M. Hénon, A. Mehr, A.M. Soward, Chaotic streamlines in the *ABC* flows, J. Fluid Mech. 167 (1986) 353–391.

[29] I. Mezić, S. Wiggins, On the integrability and perturbation of 3-dimensional fluid-flows with symmetry, J. Nonlinear Sci. 4 (1994) 157–194.

[30] U. Vaidya, I. Mezić, Existence of invariant tori in action-angle–angle maps with degeneracy, 2011.

[31] M. Budišić, I. Mezić, A bifurcation in an unsteady 3d fluid flow found using the ergodic quotient, in: IUTAM Symposium on 50 Years of Chaos: Applied and Theoretical (poster presentation), Kyoto.

[32] N. Wiener, A. Wintner, Harmonic analysis and ergodic theory, Amer. J. Math. 63 (1941) 415–426.

[33] Y. Susuki, I. Mezić, T. Hikihara, Coherent swing instability of power grids, J. Nonlinear Sci. (2011).

[34] E. Jones, T. Oliphant, P. Peterson, others, SciPy: open source scientific tools for Python, Technical Report, 2001.

[35] R.H. Clewley, W.E. Sherwood, M.D. LaMar, J.G. Guckenheimer, PyDSTool dynamical systems software, Technical Report, 2004.

[36] F. Alted, I. Vilata, PyTables: hierarchical datasets in Python, Technical Report, 2002.

[37] K. Petersen, Ergodic Theory, Cambridge University Press, Cambridge, UK, 1989.

[38] A.D. Perry, S. Wiggins, KAM tori are very sticky: rigorous lower bounds on the time to move away from an invariant Lagrangian torus with linear flow, Physica D 71 (1994) 102–121.

[39] D. Treschev, Width of stochastic layers in near-integrable two-dimensional symplectic maps, Physica D 116 (1998) 21–43.

[40] P. Stanica, Good lower and upper bounds on binomial coefficients, J. Inequal. Pure Appl. Math. 2 (2001) 5. Article 30 (electronic).

[41] S. Lafon, A.B. Lee, Diffusion maps and coarse-graining: a unified framework for dimensionality reduction, graph partitioning, and data set parameterization, Pattern Anal. Mach. Intell., IEEE Trans. 28 (2006) 1393–1403.