Technische Universität München
TUM School of Computation, Information and Technology

Topological Signatures for Analyzing Recurrence in Sampled Dynamics

David Thomas Hien

Vollständiger Abdruck der von der TUM School of Computation, Information and Technology der Technischen Universität München zur Erlangung eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz: Prof. Dr. Eric Sonnendrücker

Prüfende der Dissertation:

1. Prof. Dr. Oliver Junge
2. Prof. Dr. Konstantin Mischaikow
3. Prof. Dr. Thomas Wanner (schriftliche Beurteilung)
4. Prof. Dr. Ulrich Bauer (mündliche Prüfung)

Die Dissertation wurde am 19.02.2025 bei der Technischen Universität München eingereicht und durch die TUM School of Computation, Information and Technology am 03.06.2025 angenommen.

Acknowledgements

First of all, I want to thank Oliver Junge for his encouragement, advice, and support throughout my time at TUM. I am grateful to Ulrich Bauer for introducing me to topological data analysis and for letting me be part of his computational topology group. Special thanks go to Konstantin Mischaikow, whose passion for dynamics is truly inspiring, and for hosting my three-month stay at Rutgers.

I also want to thank my M3 colleagues Alvaro de Diego and April Herwig, as well as all members of the computational topology group, for providing a stimulating work environment. Thanks to Jannik Steinmeier for sharing the final stretch of the writing process and figuring out how to actually submit a thesis at TUM. Special thanks go to Fabian Lenzen and Fabian Roll for being great colleagues and even greater friends.

Finally, I cannot thank my grandparents, parents, and my brother enough for their continued support.

Abstract

Recurrence is a fundamental characteristic of dynamical systems with complex behavior, yet understanding its structure is challenging. We develop an algebraic topological signature to identify and classify elementary recurrent motions. These signatures are computable from sampled trajectories and their statistics provide coarse global information about the structure of recurrent behavior. We demonstrate this through several examples.

Contents

1 Introduction

Dynamical systems theory is the mathematical study of time-dependent systems. Such systems appear throughout nature, including neural activity, weather, molecular dynamics, and many other complex phenomena. At an abstract mathematical level, a dynamical system describes the evolution of *states* over *time*. A state can be simple, such as the position and velocity of a pendulum suspended from a string, or complex, like the weather on Earth at a given moment. Classically, one often assumes that the state evolution is described by a model, typically in terms of ordinary or partial differential equations.

Nowadays, data is ubiquitous in the study of dynamical systems [KS04, BPK16, Sch10, BMM12]. Furthermore, when precise models are unknown or they are not amenable to direct computation, data is all that is available. A key problem is therefore the extraction of meaningful dynamical behavior directly from data.

A fundamental property of complex dynamical systems is recurrence; typical trajectories repeatedly return close to previous states. Understanding this recurrence is difficult, especially if the system has many degrees of freedom and visualization is limited.

Recurrent systems are often studied using statistical methods, which provide insight into a system's 'typical' behavior through concepts like invariant measures and almost-invariant sets [DJ99]. Even for dynamics with highly complex trajectory structure, these methods enable coarse-grained, macroscopic descriptions of dynamics. However, powerful as they are, these statistical descriptions offer little insight into how trajectories evolve or how they are organized in phase space. This naturally leads to the question of coarse descriptors of recurrent sets which are more revealing about the dynamics.

A natural approach to characterizing recurrence is to identify trajectories which return to their initial state, i.e., periodic orbits. In fact, simple chaotic dynamics are often described in terms of a scaffolding based on periodic orbits [GH13]. The advantage of periodic orbits is that, at least abstractly, they are simple objects – closed trajectories that have the topology of a circle. However, in the context of data driven dynamics they are difficult to identify, especially if the system is high-dimensional and the data has noise [AAS22].

The starting point of our approach is a suitable weakening of the notion of periodicity using tools from topological data analysis. Topological data analysis (TDA) is a fairly recent and active research area focused on applying topological concepts to data. Persistent homology [EH08, Car09, Ghr07], its most prominent tool, is often applied by assigning a filtered geometric complex to point cloud data. This allows one to obtain a stable description of large-scale topological features in the underlying data. By now, persistent homology has successfully been applied to time-dependent data in a variety of contexts, including bifurcation analysis [GMK22], the analysis of flow field patterns [KLT+16], biological aggregation models [TZH15], and others [PH14, CS24].

Contribution In this thesis, we use methods from topological data analysis to identify global structures, which we refer to as cycling motions, in time series sampled from dynamics. More precsiely, we define the cycling signature, a topological descriptor for segments in sampled dynamics. Statistics of these signatures can be used to obtain qualitative and quantitative description of cycling motions in the data. The resulting analysis is global in nature; clusters of cycling motions are detectable and transitions between cycling motions can be analyzed. Using this method, we analyze three example datasets, one of which is a hyperchaotic attractor in a 4d system of ordinary differential equations. We see this as a promising approach to revealing coarse-grained dynamical information in high-dimensional systems.

Summary

Consider a time-continuous dynamical system given by a smooth flow $\Phi\colon \mathbb{R} \times X \to X$ on a state space $X \subset \mathbb{R}^d$, meaning that

(i) $\Phi(0, p) = p$ for all $p \in M$, and

(ii) $\Phi(t + s, p) = \Phi(t, \Phi(s, p))$ for all $t, s \in \mathbb{R}$ and $p \in M$.

Assume that we do not know Φ but only have access to a (possibly perturbed) sample from this system. This means that we are given a time series $\Gamma = (x_i)_{i=0,\ldots,N}$ together with sampling times $T = (t_i)_{i=1,\ldots,N}$ such that

$$x_i \approx \Phi(t_i - t_{i-1}, x_{i-1}), \qquad i = 1, \ldots, N,$$

where $x_0 \in X$ is an initial point, the sampling times are assumed to satisfy $t_0 < t_1 < \cdots < t_N$, and x_i are measurements from the underlying flow that may be perturbed by, e.g., measurement errors or numerical integration.

If Γ is sampled from a recurrent set with complicated dynamics, then, given a sufficiently long time span and sampling density of Γ, the time series will closely return to certain states, i.e., there are tuples (i, j) such that $x_i \approx x_j$, but also points x_ℓ with $i < \ell < j$ far away from both x_i and x_j. We will later define certain segments containing such a 'recurrence point' x_i as *cycling* (Definition 3.1.8).

A time series may have an abundance of cycling segments. Often these cycling segments can be structured into a small number of different qualitative classes or *cycling* *motions*. Our goal is to identify these cycling motions and understand their global structure.

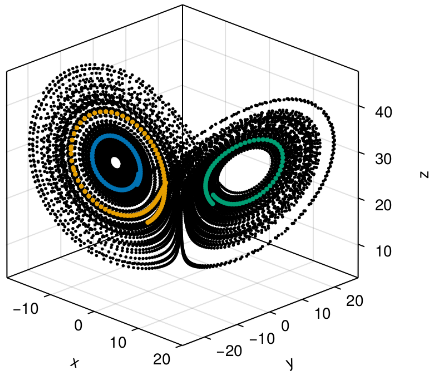


Figure 1.1: Cycling segments in the classical Lorenz attractor. The black points come from numerical integration of the system. Three segments are highlighted. The orange and blue segments belong to the same cycling motion which should be different from the cycling motion of the green segment.

Motivating Example Figure 1.1 shows a time series Γ and some segments which are colored orange, blue and green. The time series Γ is sampled from the Lorenz system with classical parameters [Lor63]. Three cycling segments are highlighted (heuristically, they are seen to be cycling since they trace an almost-closed path). The orange and blue segments wrap around the left 'hole' of the attractor, while the green segment wraps around the right 'hole'. This suggests that the orange and blue segments belong to the same cycling motion, distinct from the cycling motion of the green segment. More generally, we want to detect that, at a coarse level, there are two macroscopic cycling motions, one corresponding to segments wrapping around the left hole, one corresponding to segments wrapping around the right hole.

Cycling Signatures To capture these cycling motions, we define the cycling signature (Definition 3.3.1), a topological descriptor of a time series segment. A time series segment is called cycling if it is 'non-trivially almost-closed' in a topological sense. Its cycling signature captures the types of 'cycling' dynamics that the segment contains. More precisely, it assigns to a segment γ of Γ a filtered vector space which captures the cycling motions γ can contain. Since all these vector spaces are subspaces of the vector space assigned to the entire time series, we can compare these spaces for different segments of Γ. Of particular interest are segments that produce a cycling signature of dimension one (for suitable filtration parameters). Such segments capture a unique feature which is identified by its cycling signature. Cycling motions can then be identified through cycling signatures that often appear in the time series: Referring back to Fig. 1.1, all segments wrapping around the left hole are identified as belonging in the same cycling motion which differs from the cycling motion of all segments wrapping around the right hole. If a segment attains a higher-dimensional cycling signature, the signature does not immediately reveal which one-dimensional signatures give rise to it since vector spaces do not canonically decompose into subspaces. However, such segments can still be used to relate the aforementioned elementary cycling motions via subspace inclusion.

On a technical level, the cycling signature is the image persistence module of a map in homology whose domain is the offset filtration of the segment of interest and whose codomain is the constant filtration of a comparison space. A comparison space (see Definition 3.2.1) is a space that corresponds to the entire time series Γ. In our examples, we construct the comparison space by first lifting the time series data to a suitable space and then covering the resulting points with coarse cubes.

Time Series Analysis using Cycling Signatures We provide a computational pipeline [Hie24] to analyze cycling motions in (multivariate) time series data. The pipeline is based on a subsampling approach; for a given time series, we extract segments of varying time spans and compute their cycling signatures. We introduce summary statistics to analyze the resulting collection of signatures. We demonstrate this approach through three time series which are generated by numerical integration of three differential equations. In particular, we identify and analyze cycling motions in a four-dimensional system with a hyperchaotic attractor.

Persistent Homology of Distance Matrices of Time Series We observe a relationship between the persistent homology of a time series and the persistent homology of its distance matrix which provides an alternative method for computing cycling signatures (see Section 5.3) and also offers information which extends the cycling signature analysis (see Section 6.4).

Thesis Structure

Chapter 2 contains some background on topological data analysis. In Chapter 3, we define the cycling signature. In Chapter 4, we provide a discretization of the cycling signature and present an algorithm to compute it. In Chapter 5, we explore the connection between persistent homology of a time series and the persistent homology of its distance matrices. In Chapter 6, we introduce summary statistics and apply the cycling signature to three example systems. Finally, in Chapter 7, we offer some concluding thoughts.

Collaboration

Chapters 3, 4 and 6 of this thesis are part of [BHJM23] and are joint work with Oliver Junge, Ulrich Bauer, and Konstantin Mischaikow.

AI Disclosure Statement During the preparation of this work, ChatGPT was used in order to improve writing style, check grammar and spelling and aid in creating figures using TikZ.

2 Preliminaries on Topological Data Analysis

We introduce some notions on topological data analysis and refer to [Oud15] for a textbook reference.

2.1 Filtered Topological Spaces and Simplicial Complexes

In the following, we regard any poset P as a category with an object for each element in P and a morphism $p \to q$ whenever $p, q \in P$ and $p \leq q$. We denote by **Top** the category whose objects are topological spaces and whose morphisms are continuous maps. To study topological properties accross different scales, the notion of a topological space is augmented by one or more parameters.

Definition 2.1.1. A *P-filtered (topological) space* is a functor $X\colon P \to \mathbf{Top}$ such that if $p \leq q$, then $X_{p,q}\colon X_p \to X_q$ is an inclusion.

The collection of all P-filtered topological spaces form a category $\mathbf{Top}^P$ whose objects are functors $P \to \mathbf{Top}$ and whose morphisms are natural transformations of such functors. Explicitly, we have the following.

Definition 2.1.2. A *morphism $\phi\colon X \to Y$ of P-filtered spaces X and Y* is a collection of continuous maps $\phi_p\colon X_p \to Y_p$ for each $p \in P$ such that the following diagram commutes

$$\begin{tikzcd} X_p \arrow[r, hook, "X_{p,q}"] \arrow[d, "\phi_p"] & X_q \arrow[d, "\phi_q"] \\ Y_p \arrow[r, hook, "Y_{p,q}"] & Y_q \end{tikzcd}$$

for all $p, q \in P$ with $p \leq q$.

Example 2.1.3. Let X be a topological space. The *constant P-filtered space* ΔX has $(\Delta X)_p = X$ for all p.

Example 2.1.4 (Sublevel filtration)**.** The *sublevel filtration* of a (not necessarily continuous) function $f\colon X \to \mathbb{R}$ is $F\colon \mathbb{R} \to \mathbf{Top}$, $F_r = f^{-1}((-\infty, r])$.

Example 2.1.5 (Offset Filtration)**.** The *offset filtration* of a subset $X \subset M$ of a metric space (M, d) is the $\mathbb{R}$-filtered space $O(X)$ where

$$O_r(X) = \{m \in M \mid d(m, X) \leq r\}.$$

We sometimes write $O_r(X, d)$ to emphasize the metric. Note that the offset filtration is a sublevel filtration with respect to $f(x) = d(x, X)$.

In practice, one often consideres filtered spaces arising from discrete objects like simplicial complexes. In the following, we recall some basic notions and refer the reader to [Mun84] or [Hat02] for details.

Definition 2.1.6. An *(abstract) simplicial complex* K is a collection of finite sets such that if $\sigma \in K$ and $\tau \subset \sigma$ is nonempty, we have $\tau \in K$. An element $\sigma \in K$ is called a *simplex*. Its *dimension* is one less than its cardinality, i.e. $\dim \sigma = \operatorname{card} \sigma - 1$. The *$k$-skeleton* of K is the collection $K^{(k)}$ of all simplices σ in K with dimension $\dim \sigma \leq k$. A simplex $\tau \subset \sigma$ is a *face* of σ, we then call σ a *coface* of τ. A map $f \colon K \to L$ between simplicial complexes K and L is a *simplicial map* if for any $\sigma = \{v_0, v_1, \dots, v_k\} \in K$ the simplex $\tau = \{f(v_0), f(v_1), \dots, f(v_k)\}$ is in L and $f(\sigma) = \tau$. We denote by **Simp** the category of simplicial complexes and simplicial maps.

Every simplicial complex K has a geometric realization $|K|$, that is, there is a geometric simplicial complex with K as vertex scheme (see [Mun84] for details). Note that 'geometric realization' is functorial, i.e. it also applies to simplicial maps. We can thus use geometric realization to regard **Simp** as a subcategory of **Top**.

A flag in a simplicial complex is a chain $\sigma_0 \subsetneq \sigma_1 \subsetneq \cdots \subsetneq \sigma_k$. The collection of all flags of a simplicial complex K is again a simplicial complex. It is called the (first) barycentric subdivision of K and denoted $\operatorname{sd} K$.

Definition 2.1.7. A *P-filtered simplicial complex* is a functor $K \colon P \to$ **Simp** such that if $p \leq q$, then $X_p \subset X_q$ and $K_{p,q} \colon K_p \to K_q$ is the inclusion map.

We will sometimes slightly abuse notation and denote by K the filtered complex as well as $\bigcup_r K_r$.

Nerves and Čech complexes

A common way to associate a simplicial complex to a family of sets $\mathcal{U} = (U_i)_{i \in I}$ is the Nerve complex $\mathcal{N}(\mathcal{U})$ whose simplices are the finite collections $J \subset I$ of elements of $\mathcal{U}$ with nonempty intersection $\bigcap_{i \in J} U_i$. Typically, the collection $\mathcal{U}$ covers a topological space X of interest. Nerve theorems then yield some form of topological equivalence between X and $\mathcal{N}(\mathcal{U})$, provided the cover is 'good' enough. A recent overview of the literature is given in [BKRR23], which also includes the Nerve theorem for compact convex sets that we describe next.

The nerve of an offset filtration is equivalently expressed as follows.

Definition 2.1.8. The *Čech complex* $\check{C}_r(X)$ at radius r of a finite set of points $X \subset M$ in a metric space (M, d) is the simplicial complex

$$\check{C}_r(X) = \left\{ \emptyset \neq Q \subset X \mid Q \text{ finite and } \bigcap_{q \in Q} \overline{B(q,r)} \neq \emptyset \right\},$$

where $B(q, r)$ is the ball with center q and radius r in M. We sometimes write $\check{C}_r(X, M, d)$ to highlight the metric space. The *Čech filtration* of X is the $\mathbb{R}$-filtered simplicial complex with $\check{C}(X) \colon \mathbb{R} \to \mathbf{Simp}$, $r \mapsto \check{C}_r(X)$.

Given a finite set of points $X \subset \mathbb{R}^m$, there are now two filtrations associated to it: the filtration of metric thickenings $O(X)$ and the Čech filtration $\check{C}(X)$. In the following we state a nerve theorem from [BKRR23] that guarantees topological equivalence of these two filtrations.

Description of the homotopy equivalence. Let $K_r = \operatorname{sd} \check{C}_r(X)$ be the barycentric subdivision of the Čech complex of X and K the corresponding filtered simplicial complex. We now define a map $f \colon |K| \to O(X)$ from this filtration to the offset filtration $O(X)$. We start by defining f on the 0-skeleton of X, i.e., the filtration of 0-skeletons. For this, fix any vertex v in K. Since v is in the barycentric subdivision of $\check{C}_r(X)$, it corresponds to a flag of length one, i.e. a simplex $\sigma = \{x_0, \dots, x_n\}$ in $\check{C}_r(X)$ spanned by vertices $x_0, \dots, x_n$ in X. In particular, the birth radius of v in K is equal to the birth radius of σ in $\check{C}(X)$ which, by definition of the Čech complex, is the smallest radius $r_b > 0$ such that

$$\bigcap_{i=0}^{n} \overline{B(x_i, r_b)} \neq \emptyset.$$

Fix any point y in this intersection. We set $f_r(v) = y$ for all $r \geq r_b$, see Fig. 2.1. After repeating this for the entire 0-skeleton, we extend f_r via linear interpolation to all of $|K_r|$, i.e., for any simplex $\sigma = \{x_0, \dots, x_n\}$ in K_r, we set, in barycentric coordinates,

$$f_r(\lambda_0 x_0 + \dots + \lambda_n x_n) = \lambda_0 f(x_0) + \dots + \lambda_n f(x_n). \tag{2.1}$$

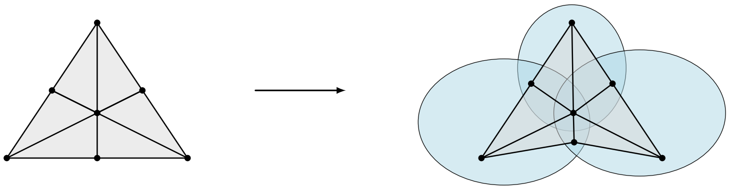


Figure 2.1: Illustration of the map f_r (cf. [BKRR23]).

Theorem 2.1.9. *For any point cloud $X \subset \mathbb{R}^m$, the map $f\colon |K| \to O(X)$ constructed above is a morphism of filtered topological spaces such that each f_r is a homotopy equivalence.*

Proof. For any fixed r, the map f_r is precisely the map Γ from [BKRR23] Theorem 3.1, therefore it is a homotopy equivalence. Furthermore, by Theorem 3.11 in the same paper, this homotopy equivalence is natural with respect to morphisms of pointed covered spaces. By Example 1.6 therein, these include the setting of barycentric subdivisions of Čech complexes. □

Vietoris–Rips Complexes

Definition 2.1.10. The Vietoris–Rips complex at radius r of a metric space (X, d) is the simplicial complex

$$\mathrm{VR}_r(X) = \{\emptyset \neq \sigma \subset X \mid \sigma \text{ finite and } \operatorname{diam} \sigma \leq r\}.$$

where $\operatorname{diam} \sigma$ denotes the diameter of σ with respect to the metric d. The Vietoris–Rips filtration is the $\mathbb{R}$-filtered simplicial complex with $\check{\mathrm{C}}(X)\colon \mathbb{R} \to \mathbf{Simp}$, $r \mapsto \mathrm{VR}_r(X)$.

The Vietoris–Rips complex [Vie27] is a popular tool in topological data analysis and commonly used to compute persistent homology of point sets [Bau21]. The Vietoris–Rips complex is a clique complex which means that it is completely determined by its 1-skeleton. To compute it, it is enough to check all pairwise intersections of balls centered at X which is is preferrable to checking all possible interesections which is required for computing the Čech complex. From a persistent point of view, for a subset $X \subset \mathbb{R}^m$, there is a chain of inclusions $\mathrm{VR}_{r'}(X) \subset \check{\mathrm{C}}_r(X) \subset \mathrm{VR}_r(X)$ for $r'/r \geq \sqrt{2m/(m+1)}$ (see [dSG07, Theorem 2.5]) which justifies its use in applications.

2.2 Persistent Homology

Persistent homology is one of the major tools of topological data analysis. In the following, we denote by **Vect** the category of vector spaces.

2.2.1 Persistence Modules

Definition 2.2.1. A *persistence module* indexed by a poset P is a functor $M: P \to \mathbf{Vect}$. The maps $M_{p,q}$ for $p \leq q$ are called structure maps. A peristence module is called *pointwise finite-dimensional* if M_p is finite-dimensional for all $p \in P$.

In case all structure maps of a persistence are inclusions, we obtain a filtered vector space.

Example 2.2.2. A *P-filtered vector space* is a functor $V: P \to \mathbf{Vect}$ such that if $p \leq q$, then $V_p \subset V_q$ and $V_{p,q}: V_p \to V_q$ is the inclusion map.

A morphism $f: M \to N$ of persistence modules M and N is a natural transformation of the functors M and N. Since $\mathbf{Vect}$ is an abelian category, so is the category $\mathbf{Vect}^P$ of P-indexed persistence modules and the required kernels, cokernels and images are constructed pointwise (see [BS14]).

Definition 2.2.3. Let $f: M \to N$ be a morphism of persistence modules. Its *kernel, image and cokernel* are the persistence modules $\ker f$, $\operatorname{im} f$ and $\operatorname{coker} f$, where $\ker_p f = \ker f_p$, $\operatorname{im}_p f = \operatorname{im} f_p$ and $\operatorname{coker}_p f = \operatorname{coker} f_p$ for all $p \in P$.

Filtered vector spaces have a canonical notion of filtered basis.

Definition 2.2.4. Let V be an I-filtered vector space. A *(filtration compatible) basis* B of V is a collection of subsets $B_r \subset V_r$, $r \in I$, such that

- $B_r \subset B_s$ for $r \leq s$, and
- B_r is a basis for V_r for all $r \in I$.

2.2.2 Structure

From now on, we restrict to the setting of one-parameter persistent homology, meaning that $P \subset \mathbb{R}$.

Definition 2.2.5. Let be $I \subset P$ be an interval. The *interval module* of I is the persistence module

$$(\mathbb{F}^I)_x = \begin{cases} \mathbb{F} & \text{if } x \in I, \\ 0 & \text{else,} \end{cases} \qquad (\mathbb{F}^I)_{x,y} = \begin{cases} \mathrm{id}_k & \text{if } x \leq y \in I, \\ 0 & \text{else.} \end{cases}$$

In the special case that $I = [a, b)$, we write $I(a, b) = \mathbb{F}^I$.

Persistence modules often decompose into interval modules.

Theorem 2.2.6 (Structure of Persistence Modules [BCB20, CB15])**.** *If M is p.f.d. P-indexed persistence module where P is totally ordered, then there exists a unique multiset $\mathcal{B}(M)$, called the* barcode of M*, of intervals in P, such that*

$$M \cong \bigoplus_{I \in \mathcal{B}(M)} \mathbb{F}^I.$$

Some persistence modules arising naturally in practice are not p.f.d. but satisfy the weaker notion of finiteness called $\mathfrak{q}$-tameness. A persistence module $M\colon P \to \mathbf{Vect}$ is called $\mathfrak{q}$-tame, if $\operatorname{im} M_{p,q}$ is finite dimensional for all $p, q \in P$. While these persistence modules do not have a barcode in the sense of the preceeding theorem, one can pass to the observable category to get a slightly weaker notion of a barcode. A persistence module is called *ephemeral* if all its structure maps are zero. Given a $\mathfrak{q}$-tame persistence module M, one can find a p.f.d. persistence module N together with a map $M \mapsto N$ which has ephemeral kernel and cokernel (i.e. is a weak isomorphism in the *observable category* [CCBdS16]). One can then define the observable barcode for M as the barcode for N where the information about the endpoints of the intervals is ignored.

2.2.3 Stability

A foundational result in the theory of persistent homology is the stability of barcodes [CSEH06]. We now recall some relevant notions of distances and stability. The Hausdorff-distance of a metric space (X, d) is a metric for the compact subsets of a metric space. For subsets $A, B \subset X$, it is given by

$$d_H(A, B) = \max\{\sup_{x \in A} d(x, B), \sup_{x \in B} d(x, A)\}.$$

Algebraic objects that arise in the study of persistent homology can be related using interleavings. To define these, we denote the ϵ-shift of a functor $F\colon \mathbb{R} \to \mathbf{C}$ by $F_{\bullet+\epsilon}\colon \mathbb{R} \to \mathbf{C}$, $(F_{\bullet+\epsilon})_r = F_{r+\epsilon}$. Let $G\colon P \to \mathbf{C}$ be another functor. Furthermore, fix $\epsilon > 0$ and define $f\colon F_\bullet \to F_{\bullet+2\epsilon}$, $f_r = F_{r,r+2\epsilon}$ and $g\colon G_\bullet \to G_{\bullet+2\epsilon}$, $g_r = G_{r,r+2\epsilon}$. An ϵ-*interleaving* of F and G consists of natural transformations $\alpha\colon F_\bullet \to G_{\bullet+\epsilon}$ and $\beta\colon G_\bullet \to F_{\bullet+\epsilon}$ such that $\alpha \circ \beta = f$ and $\beta \circ \alpha = g$. The interleaving distance of two functor is then given by

$$d_I(F, G) = \inf_\epsilon\{F \text{ and } G \text{ are } \epsilon\text{-interleaved}\}.$$

The Hausdorff-distance of two metric spaces can equivalently be expressed using the interleaving distance of their offset filtrations.

Lemma 2.2.7. *Let (X, d) be a metric space, $A, B \subset X$ closed and $f(x) = d(x, A)$, $g(x) = d(x, B)$. Then the following are equivalent:*

i) $d_H(A, B) \leq \epsilon$,

ii) $\|f-g\|_\infty \leq \epsilon$,
iii) $d_I(O(A), O(B)) \leq \epsilon$.

Now, given two subsets A, B of a metric space (X, d), we can apply the homology functor to their offset filtrations. Clearly, the composition of a functor with an ϵ-interleaving gives another ϵ-interleaving, we thus obtain $d_I(H(O(A)), O(B)) \leq d_H(A, B)$. The last step in showing stability of the persistent homology pipeline is the algebraic stability theorem which was introduced in [CdSGO16]. It states that passing through the structure theorem is a 1-Lipschitz map with respect to the interleaving distance for persistence modules and the bottleneck distance on barcodes. We omit a statement here, since this will not be used in the remainder of this thesis.

Rather, we state the following stability result for kernels, images and cokernels which generalizes a stability result from [CSEHM09].

Theorem 2.2.8 (Stability theorem for kernels, images and cokernels [BS14])**.** *Let $h\colon X \to Y$ be a continuous map of topological spaces. Let $f, f'\colon Y \to \mathbb{R}$ and $g, g'\colon X \to \mathbb{R}$ be (not necessarily continuous) maps, such that*

$$g(h(y)) \leq f(h(y)) \text{ and } g'(h(y)) \leq f'(h(y)), \text{ for all } y \in Y. \tag{2.2}$$

Let $F\colon \mathbb{R} \to \mathbf{Top}$ be the filtered space given by $F_r = f^{-1}(-\infty, a]$, and define F', G, G' similarly. By 2.2, the map h induces maps $\alpha\colon F \to G$ and $\beta\colon F' \to G'$ of filtered spaces. Let $\mathbf{A}$ be an abelian category and $H\colon \mathbf{Top} \to \mathbf{A}$ be some functor. Finally, let $\epsilon = \max\{\|f-f'\|_\infty, \|g-g'\|_\infty\}$. Then,

$$d_I(\ker H\alpha, \ker H\beta), d_I(\operatorname{im} H\alpha, \operatorname{im} H\beta), d_I(\operatorname{coker} H\alpha, \operatorname{coker} H\beta) \leq \epsilon.$$

2.3 Cubical Complexes

In the following, we introduce cubical (grid) complexes, a type of complex that is well-suited for computations [KMM04].

An elementary interval is an interval of the form $I = [k, k]$ or $I = [k, k+1]$ where $k \in \mathbb{Z}$. It is called non-degenerate if it has length one and degenerate otherwise. We sometimes write $[k]$ instead of $[k, k]$ for degenerate intervals. An elementary cube in $\mathbb{R}^m$ is a set of the form

$$Q = I_1 \times I_2 \times \cdots \times I_d$$

where the I_k are elementary intervals. The dimension of Q is the number of non-degenerate intervals I_k. The set of all k-dimensional unit cubes in $\mathbb{R}^m$ is denoted $\mathrm{Cub}_k(\mathbb{R}^m)$ and $\mathrm{Cub}(\mathbb{R}^m) = \bigcup_{k=0}^{d} \mathrm{Cub}_k(\mathbb{R}^m)$ is the set of all elementary cubes in $\mathbb{R}^m$.

Given $P, Q \in \mathrm{Cub}(\mathbb{R}^m)$, we call P a face of Q if $P \subset Q$ and write $\downarrow Q$ for the set of all faces of Q. For $K \subset \mathrm{Cub}(\mathbb{R}^m)$ we set $\downarrow K = \bigcup_{Q \in K} \downarrow Q$. A set $K \subset \mathrm{Cub}(\mathbb{R}^m)$ is called a *cubical (grid) complex* if $K = \downarrow K$. Its geometric realization is $|K| = \bigcup_{Q \in K} Q$.

Fix a coefficient ring R and let K be a cubical complex. We denote by $C(K)$ the free R-module generated by the k-dimensional elementary cubes in K. To define the boundary operator, fix a k-dimensional elementary cube $Q = I_1 \times \cdots \times I_d$ with $I_k = [a_k, b_k]$. Let $i_1 < i_2 < \cdots < i_k$ be the indices such that I_{i_j} is non-degenerate and let

$$Q_j^- = \prod_{i<i_j} I_i \times [a_{i_j}] \times \prod_{i_j<i} I_i \quad \text{and} \quad Q_j^+ = \prod_{i<i_j} I_i \times [b_{i_j}] \times \prod_{i_j<i} I_i.$$

be the cubes where the j-th nondegenerate elementary interval of Q is replaced by its lower/upper boundary. The boundary operator $\partial\colon C_k(K) \to C_{k-1}(K)$ is the linear extension of

$$\partial Q = \sum_{j=1}^{k} (-1)^j \left(Q_j^- - Q_j^+\right).$$

Together with the $C_k(K)$, this defines the cubical chain complex of K. We note that the homology of a cubical complex is canonically isomorphic to cellular homology of $|K|$.

3 A Topological Signature for Sampled Dynamics

The cycling signature is a topological descriptor for a time series segment. In Section 3.1 we define the notion of cycling which is a suitable generalization of periodicity. In Section 3.2, we explain how cycling segments can be classified using homology. This leads to the notion of the cycling signature which we define in Section 3.3.

This chapter is joint work with Ulrich Bauer, Oliver Junge and Konstantin Mischaikow and is part of [BHJM23].

3.1 Cycling Segments

The definition of cycling is motivated by the observation periodic orbits can be identified using homology.

Proposition 3.1.1. *Let $\phi\colon \mathbb{R} \times X \to X$ be a dynamical system in a metric space (X, d) and $\zeta = \phi([a, b], x)$ an orbit segment. Then the following are equivalent:*

(i) ζ contains a full periodic orbit,

(ii) ζ is homeomorphic to a circle,

(iii) $H_1(\zeta) \neq 0$ for homology in an arbitrary coefficient ring.

Proof. *(i) ⇒ (ii):* Let $T > 0$ be the (minimal) period of the periodic orbit. Then $h\colon \mathbb{R}/\mathbb{Z} \to X$, $h(s) = \phi(x, a + sT)$ is well-defined, bijective and continuous. Let $I \subset \mathbb{R}$ be open and note that $K = h(S^1 \setminus I)$ is compact as an image of a compact set. As a metric space, X is Hausdorff. Thus, for any $x \in h(I)$ we can find disjoint neighborhoods U_x of x and V_x of K. Since $U = \bigcup\{U_x \mid x \in h(I)\}$ is a neighborhood of $h(I)$ that is disjoint from K, $h(I)$ is relatively open. Since I was arbitrary, h is a homeomorphism.

(ii) ⇒ (iii): This is a basic fact from algebraic topology.

(iii) ⇒ (i): If ζ does not contain a full periodic orbit, it deformation retracts onto x_0 via $h_s(t) = \phi(x, a + st(b - a))$. This implies that ζ is homotopy equivalent to a point and therefore $H_1(\zeta) = 0$ in any coefficient ring. □

We want to generalize this criterion to capture how far a segment is from being periodic. Furthermore, we want a notion that is applicable to sampled trajectories, i.e. multivariate time series.

Definition 3.1.2. By a *time series* Γ in a metric space (X, d), we mean a collection of points $x_0, \ldots, x_N \in X$ with sampled times $T = (t_i)_{i=1,\ldots,N}$ such that $t_0 < t_1 < \cdots < t_N$. We often identify Γ with the point cloud $\{x_0, \ldots, x_N\}$. A *segment* γ of Γ is a consecutive subsequence x_k, $x_{k+1}, \ldots,$ x_ℓ for some $k < \ell$. The *time span* of a segment is $\tau(\gamma) = t_\ell - t_k$.

For time series segments $\gamma = (x_k, x_{k+1}, \ldots, x_\ell)$, the criterion in Proposition 3.1.1 cannot be applied directly, even if $x_\ell = x_k$, since a finite sample carries the discrete topology.

Using ideas from topological data analysis, we now derive a meaningful generalization of the criterion $H_1(\zeta) \neq 0$. For this, we consider *thickenings* of γ. By the *r-thickening* of a subset G of a metric space (M, d) we mean

$$O_r(G) = \{x \in M \mid d(x, G) < r\}, \qquad 0 \leq r. \tag{3.1}$$

The thickened set $O_r(\gamma)$ in X can be interpreted as the set that contains all segments $\tilde{\gamma}$ that are 'r-perturbations' of γ with respect to the metric d.

a **b**

Figure 3.1: a. A periodic orbit is a topological circle. **b.** Almost periodic time series segment (black) with thickening (blue). Only the thickening has nontrivial H_1.

Example 3.1.3. Let γ be the time series segment shown as black dots in Fig. 3.1b. This segment represents what we want to detect as cycling. Clearly, $H_1(\gamma) = 0$. However, thickening γ with respect to the Euclidean metric (indicated by the blue balls in Fig. 3.1b), we detect nontrivial homology $H_1(O_r(\gamma)) \neq 0$ for r in some interval (r_0, r_1). □

The example demonstrates how thickening almost-periodic segments generalizes the criterion $H_1(\gamma) \neq 0$. However, in general, a metric which incorporates only spatial information may not be able to distinguish relevant homological features of a time series from features that arise due to sampling. This is illustrated in the following example.

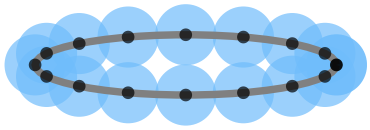


Figure 3.2: Periodic segment (black) with thickening (blue) sampled from a squished periodic orbit (gray).

Example 3.1.4 (Squished periodic orbit)**.** Suppose we have an ellipse-shaped periodic orbit $E = \{x \in \mathbb{R}^2 \mid (x_1/a)^2 + (x_2/b)^2 = 1\}$ with $b < a$, as the grey curve in Fig. 3.2.

Using the metric space $\mathbb{R}^2$ with the metric induced by the Euclidean norm, we have

$$H_1(O_r(E)) \cong I(0, b).$$

Even though we consider a periodic orbit, for b much smaller than a only correspondingly small thickening radii r lead to nontrivial H_1 since the center of the ellipse quickly closes up with increasing r. This poses a problem when passing to a time series Γ sampled from the periodic orbit (black points in Fig. 3.2), as it may not be possible to distinguish a relevant homological feature from a feature induced by sampling. Many thickened segments γ of Γ have holes and thus fulfill the criterion $H_1(O_r(\gamma)) \neq 0$. For example, we would erroneously detect the segment consisting of the seven leftmost points in the sampling as 'cycling'. This should be regarded as sampling artifact, since, in this time series, only segments that wrap around most of the ellipse should be detected as 'cycling'.

In order to be able to detect the relevant homological features, we lift the sampling data to the unit tangent bundle.

The *tangent bundle* $\mathrm{T}(X)$ of $X \subset \mathbb{R}^m$ is the set of all tuples $(x, v) \in X \times \mathbb{R}^m$, where x is a point in X and v is the velocity of some curve through x. Any point $x \in X$ that is not an equilibrium point of Φ has an associated nonzero tangent vector $v(x) = \frac{d}{dt}\Phi(t, x)|_{t=0}$. For our purpose, it is more appropriate to consider the direction of motion at each non-equilibrium point (and ignore the speed), thus we normalize the second component and pass to the *unit tangent bundle*

$$\mathrm{UT}(X) = \{(x, v) \in \mathrm{T}(X) \mid \|v\|_2 = 1\}.$$

Let X_0 be the set of equilibria of Φ and consider the section

$$\rho \colon X \setminus X_0 \to \mathrm{UT}(X \setminus X_0), \quad x \mapsto (x, v(x)/\|v(x)\|_2) \tag{3.2}$$

of the unit tangent bundle induced by the flow. For time series data, we recover v through finite differences (which imposes a condition on the sampling density). We use the (approximated) section ρ in order to lift the time series data $x_0, \dots, x_N$ from X to $\mathrm{UT}(X \setminus X_0) \subset \mathrm{UT}(X)$, i.e., we consider the lifted time series $\rho(\Gamma) := \{\rho(x_0), \dots, \rho(x_N)\}$

in $\mathrm{UT}(X)$. Then, instead of considering distances between points $x, y \in X$ directly, we measure the distance between $\rho(x)$ and $\rho(y)$ in $\mathrm{UT}(X)$ which incorporates spatial as well as directional information.

Definition 3.1.5. The *dynamic UT-metric* d_C^{dyn} with parameter $C > 0$ is given by

$$d_C^{dyn}((x,v),(y,w)) = \max\left\{\|x-y\|_2, C\|v-w\|_2\right\}, \quad \text{where } (x,v),(y,w) \in \mathrm{UT}(X). \tag{3.3}$$

It is imporant to note that C needs to be chosen with respect to the scales present in the system. If C is too small with respect to relevant thickening radii r, the lift to the unit tangent bundle does not affect the distance between points that move into different directions. Indeed, if $r \geq 2C$, the tangent component does not affect the thickening with respect to d_C^{dyn} since the inequality $\|w - v\|_2 \leq r$ is fulfilled for any $v, w \in S^d$. Conversely, if C is too large, the spatial information may be lost altogether.

The following example demonstrates that by this lifting to the unit tangent bundle, we recover natural dynamical features of segments.

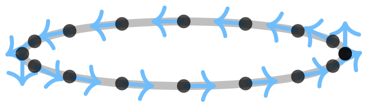


Figure 3.3: Periodic time series segment (black) with unit tangent vectors (blue) sampled from a squished periodic orbit.

Example 3.1.6 (Squished periodic orbit revisited)**.** The situation in Example 3.1.4 changes if we lift to the unit tangent bundle (see Fig. 3.3). Again, we first consider the periodic orbit E. Consider the lifts (x, v) and (y, w) of two points that are opposite on the ellipse. Their distance d_C^{dyn} is the maximum of $\|x - y\|_2$ and $C\|v - w\|_2 = 2C$. While the first term may be small if $x \approx (0, \pm a)$, the second term is independent of a, therefore any thickening radius $r < 2C$ leads to nontrivial H_1.

This change has important consequences for the sampled segment. Now, the distance between two consecutive points is small compared to the distance between points opposite on the ellipse. A computation shows that suitable C yields

$$H_1(O_r(\gamma)) = I(\tfrac{s}{2}, 2C)$$

where s is the smallest distance such that two consecutive balls intersect. In particular, we obtain a robust criterion for detecting 'cycling' segments. Furthermore, no segments that wrap around less than half the ellipse are detected. □

The importance of the lift to the unit tangent bundle is further illustrated in the Lorenz dataset.

Example 3.1.7 (Cycling segments in the Lorenz system)**.** Consider segments in the time series in Fig. 1.1 that wrap around the left wing and can be thickened, using the Euclidean distance, to have nontrivial degree 1 homology. Segments that come close to the equilibrium at the center of the left wing have nontrivial H_1 only for a small range of thickening parameters. Moreover, one can generate cycling segments close to the equilibrium with an arbitrarily small range of parameters where H_1 is nontrivial.

In contrast, thickenings in the unit tangent bundle take into account the different tangent directions close to the equilibrium. This allows robust detection of segments that cycle around the two equilibria at the center of the wings of the Lorenz attractor.

The lift to $\mathrm{UT}(X)$ and the metric d_C^{dyn} on $\mathrm{UT}(x)$ for some suitable $C > 0$ allow us to make the following definition.

Definition 3.1.8. A segment γ is *r-cycling* if $H_1(O_r(\rho(\gamma))) \neq 0$, where $\rho(\gamma)$ is the r-thickening of $\rho(\gamma)$ in $\mathrm{UT}(X)$ with respect to the metric d_C^{dyn}.

3.2 Classification of Cycling Segments

While the highlighted segments in Fig. 1.1 are detected to be r-cycling for suitable r, we have not yet identified the qualitative similarity and dissimilarity of these segments. In the following, we explain how a comparison space can be used to classify cycling segments.

Definition 3.2.1. A *comparison space* for a time series $\Gamma \subset X$ is a neighborhood $Y \subset \mathrm{UT}(X)$ of $\rho(\Gamma)$ where ρ is the section of $\mathrm{UT}(X)$ defined in (3.2).

Note that there exists a possibly infinite $r_0 = r_0(Y) > 0$ such that for any $r < r_0$ and for any segment γ of Γ, we have an inclusion map $i_{\gamma,r} \colon O_r(\rho(\gamma)) \to Y$. By functoriality of homology, the map $i_{\gamma,r}$ induces a linear map

$$H_1(i_{\gamma,r}) \colon H_1(O_r(\rho(\gamma))) \to H_1(Y)$$

in homology. The image of this map is a subspace $\operatorname{im} H_1(i_{\gamma,r})$ of $H_1(Y)$. We compare different cycling segments by their induced subspaces of the vector space $H_1(Y)$.

Definition 3.2.2. The *r-cycling space of a segment* γ with respect to the comparison space Y is the subspace

$$\mathrm{Cyc}_r(\gamma, Y) = \operatorname{im} H_1(i_{\gamma,r})$$

of $H_1(Y)$. We write $\mathrm{Cyc}_r(\gamma)$ for $\mathrm{Cyc}_r(\gamma, Y)$ in case the comparison space is clear from the context. The *r-cycling rank* is the dimension of $\mathrm{Cyc}_r(\gamma)$. We call $H_1(Y)$ the *homological comparison space.*

In the following, we illustrate the concept of comparison spaces and the cycling space using a time series from a stochastically perturbed double well system.

Example 3.2.3 (Double well system without unit tangent bundle)**.** To illustrate this idea, consider the time series in Fig. 3.4. For clarity, we initially disregard the unit tangent bundle (i.e., $C = 0$ in the metric (3.3)) and let the union of the boxes, denoted by Z, take the role of the comparison space. Since Z has two holes, we have $\dim H_1(Z) = 2$.

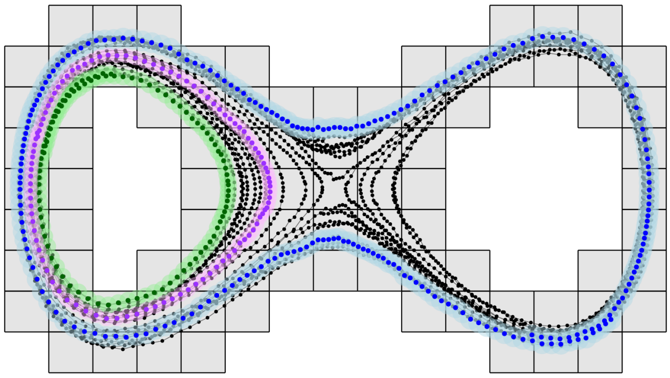


Figure 3.4: Time series with (simplified) comparison space. The time series (black dots) is covered with boxes, the union of the boxes is the comparison space Y. Three cycling segments are highlighted (dark blue, dark green dark purple) together with a thickening (light blue, light green, light purple). The comparison space classifies these into two different classes, one containing the blue, the other containing the green and purple segments.

Three segments of the times series (black) are highlighted: a green segment γ_g and a purple one γ_p which only loop around the left hole, as well as a blue one γ_b which orbits the right and the left hole.

For a suitable range of r, the thickenings of these segments have the topology of a circle and therefore one-dimensional H_1, where the homology is generated by a curve that loops exactly once around the circle. By inclusion, these curves also give rise to homology classes in Z and, with slight abuse of notation, we have

$$\mathrm{Cyc}_r(\gamma_g, Z) = \mathrm{Cyc}_r(\gamma_p, Z) \neq \mathrm{Cyc}_r(\gamma_b, Z),$$

Therefore, we detect that the green and purple segments are qualitatively similar while the blue segment is qualitatively different. □

Example 3.2.4 (Double well system with unit tangent bundle)**.** As in the previous example, we consider the time series in Fig. 3.4 though this time, we lift the time series

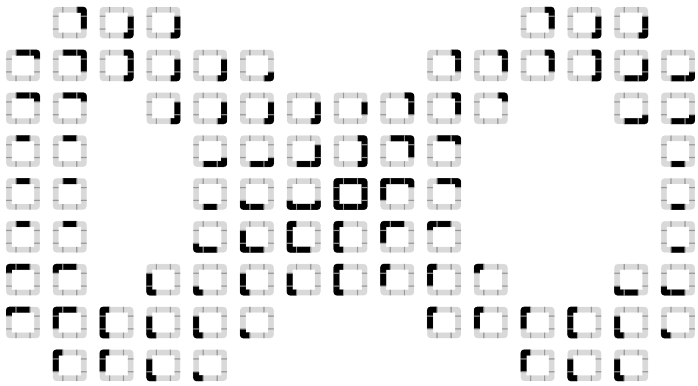


Figure 3.5: Comparison space in the unit tangent bundle. We use the S^1_∞ unit tangent bundle instead of S^1_2 in this figure (and for computations), since S^1_∞ is easier to cover with boxes than S^1_2. Each cluster of boxes corresponds to the cover of $Q \times S^1_\infty$ where Q is a box in Fig. 3.4. The comparison space is the union all the black boxes.

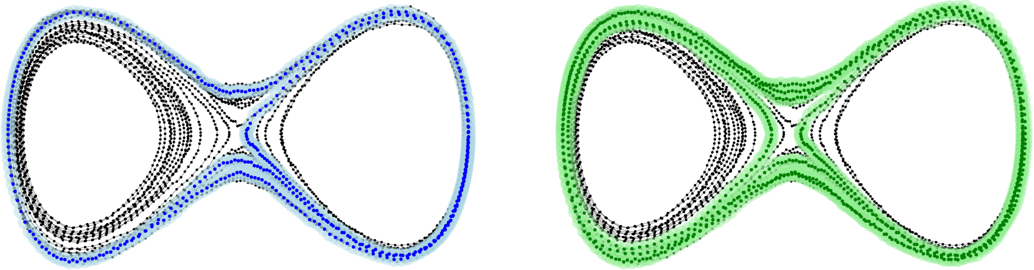
to the unit tangent bundle before constructing a cubical cover. In the tangent space, we then consider boxes centered on a grid and take these to cover the lifted data. In our example, we consider boxes of the form $\overline{B_\infty(p, r/2)} \times \overline{B_\infty(q, 1/2)}$ where $p \in r\mathbb{Z}^d$ and $q \in \mathbb{Z}^d$.

The resulting comparison space Y in the tangent bundle is visualized in Fig. 3.5. Every black box in this figure is to be understood as the product of a box from Fig. 3.4 with a box in the tangent component. As we can see, over all boxes but one of the boxes from Fig. 3.4, a contractible segment of the sphere is covered. Only in the center of the picture, the entire sphere is covered with boxes, since the center box in Fig. 3.4 contains points moving in all directions. This additional “hole” leads to $\dim H_1(Y) = 3$. The double well system (without perturbation) has a hyperbolic equilibrium at the origin, which is precisely where this extra ‘hole’ in the comparison spaces Y is located. The appearance of this extra ‘hole’ is therefore expected, since close to a hyperbolic equilibrium, all tangent directions occur. Having explained the difference between Y and Z, we now use Y to compare the three highlighted segments. Writing γ_g, γ_p and γ_b as in the previous example, we have three 1-dimensional cycling spaces with

$$\mathrm{Cyc}_r(\gamma_g, Y) = \mathrm{Cyc}_r(\gamma_p, Y) \neq \mathrm{Cyc}_r(\gamma_b, Y).$$

As expected, the segments are classified in accordance with our intuition. □

In the previous example, we only considered one-dimensional cycling spaces. In general, the dimension of a cycling space may be as large as the dimension of the homological comparison space.



a **b**

Figure 3.6: Segments with higher-dimensional cycling spaces.

Example 3.2.5 (Higher-dimensional cycling space)**.** We consider the segments shown in blue and green in Fig. 3.6, which we denote by γ_{b2} and γ_{g2}, respectively.

For exposition, we again start in the simplified setting where the union of boxes Z in Fig. 3.4 takes the role of the comparison space. In this case, the cycling spaces of both segments in Fig. 3.6 are two-dimensional and we have

$$\mathrm{Cyc}_r(\gamma_{b2}, Z) = \mathrm{Cyc}_r(\gamma_{g2}, Z) = H_1(Z)$$

Note that we do not detect the qualitative difference of the segments: γ_{b2} does not contain a subsegment which wraps only around the left hole in the dataset, while γ_{g2} does. This is an instance of the general fact that a vector spaces does not have a canonical decomposition into lower-dimensional subspaces.

However, with the comparison space Y in the unit tangent bundle, the cycling space $\mathrm{Cyc}_r(\gamma_{b2}, Y)$ has dimension 2 while $\dim \mathrm{Cyc}_r(\gamma_{g2}, Y) = 3$. Indeed, the cycling space of γ_{b2} does not contain the cycling space of segment γ_p from Fig. 3.4 as a subspace since no curve in the blue set wraps around the left hole but not the center hole.

While the cycling spaces alone do not allow us to conclude that γ_{g2} contains a subsegment with cycling space $\mathrm{Cyc}_r(\gamma_p, Y)$, we can infer from $\mathrm{Cyc}_r(\gamma_p, Y) \not\subset \mathrm{Cyc}_r(\gamma_{2b}, Y)$ that γ_{b2} does not contain a subsegment with that cycling space. □

Remark 3.2.6 (Comparison Space for the Lorenz time series)*.* We return to the Lorenz time series from Fig. 1.1. A comparison space to detect and distinguish the cycling segments on the wings should have a hole at the center of each wing. Covering the time series directly, i.e., without lifting to the unit tangent bundle, would require impractically small boxes that result in a box cover with a very large amount of boxes. Such a cover would furthermore form a hole on the outside of the right wing as an artifact of the time series not reaching this part of the attractor.

In contrast, the lifted dataset can be covered with rather coarse boxes without losing the holes at the centers of the wings. This is because a closed curve can have arbitrarily small length in space, while its length in the unit tangent bundle is bounded from below by $2\pi C$ (as a consequence of Fenchel's theorem [Fen29], which states that the total curvature of any closed curve is at least 2π). The unit tangent bundle is therefore crucial in order to construct a comparison space that allows us to distinguish the different classes of cycling segments.

3.3 The Cycling Signature

A priori, it is not clear how to appropriately choose the thickening radius r in the construction of the cycling space. Clearly, the cycling spaces of the segments under consideration will depend on r, and different choices of r may be relevant depending on the scales in the system. We adopt the viewpoint of persistent homology and argue that we should be interested in the entire range of possible thickening radii.

From a persistent homology perspective, the cycling signature can be expressed more conceptually as follows. By collecting the maps $i_{\gamma,r}$ for $r \in I = [0, r_0(Y))$, we obtain a morphism

$$i_\gamma \colon O(\rho(\gamma)) \to \Delta Y$$

of I-filtered topological spaces, where $\Delta Y = (\Delta Y_r)_{r \in I}, \Delta Y_r = Y$ for all $r \in I$, is the constant filtration of Y. Passing to homology induces a morphism

$$H_1(i_\gamma) \colon H_1(O(\rho(\gamma))) \to H_1(\Delta Y) \tag{3.4}$$

of persistence modules. The cycling signature is then the image of this morphism. Since the structure maps of $H_1(\Delta Y)$ are monomorphisms, so are the structure maps of $H_1(i_\gamma)$. Since furthermore $\operatorname{im} H_1(i_{\gamma,r}) \subset H_1(i_\gamma)$ for all r, $\operatorname{im} H_1(i_\gamma)$ is a filtered vector space.

Definition 3.3.1. The *cycling signature* of a segment γ of Γ in Y is the I-filtered vector space

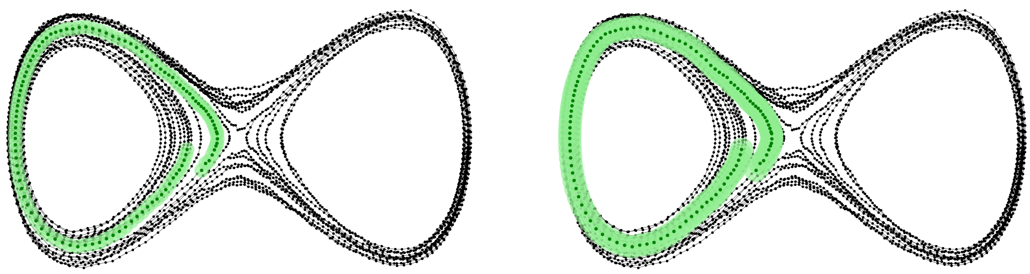
$$\operatorname{Cyc}(\gamma, Y) = \operatorname{im} H_1(i_\gamma)$$

where $I = [0, r_0(Y))$ is the *interval of admissible radii.* We write $\operatorname{Cyc}(\gamma)$ for $\operatorname{Cyc}(\gamma, Y)$ in case the comparison space is clear from the context. The *cycling rank function* of γ is the function

$$r \mapsto \dim \operatorname{Cyc}_r(\gamma, Y).$$

We note that this strategy of using image persistence to match features has been employed successfully in several other contexts [CSEHM09, SPS+23, RB23].

Example 3.3.2 (Rank and thickening radius)**.** Consider the segment in Fig. 3.7 and the comparison space Y from Example 3.2.5. From the figure, it is clear that small thickenings lead to rank 0, while larger thickenings lead to rank 1. This is captured in



a **b**

Figure 3.7: Segment with different thickenings. The light green set in **a** has rank 0, and the one in **b** has rank 1.

the cycling rank: there is a threshold $\bar{r}$ such that

$$\dim \mathrm{Cyc}_r(\gamma, Y) = \begin{cases} 0, & \text{if } r < \bar{r}, \\ 1, & \text{if } \bar{r} \leq r < r_0(Y). \end{cases}$$

$\square$

Remark 3.3.3. As of now, the cycling signature is defined as degree one homology of a map of filtered topological spaces. However, it can be extended to include higher homological degrees. Invariant tori in Hamiltonian systems provide one example where trajectories are organized in an object which has nontrivial homology in degree greater one.

3.4 Stability

We show a basic stability result for cycling signatures. Let $\mathrm{TS}(X)$ denote the set of all time series (see 3.1.2) in the metric space (X, d) and let $\mathrm{Int}(\mathbb{R})$ denote the poset of all intervals in $\mathbb{R}$, ordered by inclusion. Given a time series $\Gamma = (x_i)_{i=1}^N$ with sampling times $T = (t_i)_{i=1}^N$ and an interval $J \in \mathrm{Int}(\mathbb{R})$, we write Γ_J for the segment $(x_i)_{i=k}^{\ell}$ where $k = \min\{i \mid t_i \in J\}$ is the first, and $\ell = \max\{i \mid t_i \in J\}$ is the last index corresponding to times in J.

Definition 3.4.1. For $\Gamma, \tilde{\Gamma} \in \mathrm{TS}(X)$ define

$$d_{\mathrm{ts}}(\Gamma, \tilde{\Gamma}) = \max_{J \in \mathrm{Int}(\mathbb{R})} d_H(\Gamma_J, \tilde{\Gamma}_J)$$

where d_H is the Hausdorff distance.

Lemma 3.4.2. $(\mathrm{TS}(X), d_{\mathrm{ts}})$ *is a metric space.*

Proof. Let Γ and $\tilde{\Gamma}$ be time series indexed by T and $\tilde{T}$, respectively. For one point intervals $\bar{t} = \{t\}$, $\Gamma_{\bar{t}}$ is the empty set if $t \notin T$, or a single point otherwise. Therefore, $d_{ts}(\Gamma, \tilde{\Gamma}) = 0$ if and only if $T = \tilde{T}$ and $\Gamma_t = \tilde{\Gamma}_t$ for all $t \in T$. Symmetry and the triangle inequality follow readily from the fact that d_H is a metric. □

The metric d_{ts} captures pointwise perturbations of a given time series. Let $\Gamma = \{x_0, \dots, x_n\}$ be a T-indexed time series. Suppose $\tilde{\Gamma} = \{y_0, \dots, y_N\}$ is a pointwise perturbation, by which we mean $\max_i d(x_i, y_i) \leq \epsilon$ for an $\epsilon > 0$. Then $d_{ts}(\Gamma, \tilde{\Gamma}) \leq \epsilon$.

In the following, we investigate stability of the cycling signature as a map from $\mathrm{TS}(X)$ to filtered vector spaces. To compare cycling signatures of two time series Γ and $\tilde{\Gamma}$, we want some more flexibility with comparison spaces.

Definition 3.4.3. A *generalized comparison space* for a time series $\Gamma \subset X$ is a map $f\colon \Gamma \to Y$ where (X', d') is some metric space and $Y \subset X'$ is a neighborhood of $f(\Gamma)$. The *interval of admissible radii* for J is $[0, r_0(f))$, where $r_0(f)$ is the supremum over radii r such that $O_r s(f(\Gamma)) \subset Y$. In case $Y \subset X$ and f is the inclusion map, we call Y a *simplified comparison space.* Let γ be a segment of Γ. Its *cycling signature of with respect to a generalized comparison space* $f\colon \Gamma \to Y$ is defined as

$$\mathrm{Cyc}(\gamma, f) = \mathrm{im}\, H_1(f).$$

Note that a simplified comparison space is basically a neighborhood. We say that two time series have a joint (generalized) comparison space, if it is a (generalized) comparison space for both time series. Its interval of admissible radii is the intersection of the intervals of the individual time series.

Theorem 3.4.4. *Let* $\Gamma, \tilde{\Gamma}$ *be time series in a metric space* (X, d) *with joint simplified comparison space* Y*. Let* I *be its interval of admissible radii. Fix* $J \in \mathrm{Int}(\mathbb{R})$ *and set* $\gamma = \Gamma_J$ *and* $\tilde{\gamma} = \tilde{\Gamma}_J$*. Then*

$$d_I(\mathrm{Cyc}(\gamma, \mathrm{id}_Y), \mathrm{Cyc}(\tilde{\gamma}, \mathrm{id}_Y)) \leq d_H(\gamma, \tilde{\gamma}) \leq d_{ts}(\Gamma, \tilde{\Gamma}) \tag{3.5}$$

where we regard the cycling signatures as I*-indexed filtered vector spaces.*

Proof. The left inequality in (3.5) follows from Theorem 2.2.8 as follows. Define $f, g\colon Y \to \mathbb{R}$ via $f(x) = d(x, \gamma)$ and $g(x) = d(x, \tilde{\gamma})$. We have the sublevel filtrations $F_r = f^{-1}(-\infty, r]$ and $G_r = g^{-1}(-\infty, r]$. The identity map $i\colon Y \to Y$ induces maps $\alpha\colon F \to \Delta Y$ and $\beta\colon G \to \Delta Y$ of filtered spaces. This is precisely the setting of Theorem 2.2.8 which implies

$$d_I(\mathrm{im}\, H_1(\alpha), \mathrm{im}\, H_1(\beta)) \leq \|f - g\|_\infty . \tag{3.6}$$

Now, we have $O_r(\gamma) = F_r$ and $O_r(\tilde{\gamma}) = G(r)$ and thus $\mathrm{Cyc}(\gamma, \mathrm{id}_Y) = \mathrm{im}\, H_1(\alpha)$ and $\mathrm{Cyc}(\gamma, \mathrm{id}_Y) = \mathrm{im}\, H_1(\beta)$. Thus we have that the lefthand side of (3.6) is equal to

$d_I(\mathrm{Cyc}(\gamma, \mathrm{id}_Y), \mathrm{Cyc}(\tilde{\gamma}, \mathrm{id}_Y))$. Furthermore, by Lemma 2.2.7, we have $\|f-g\|_\infty = d_H(\gamma, \tilde{\gamma})$. This proves the left inequality in (3.5). The right inequality in (3.5) directly follows from the definition of d_{ts}. ☐

Stability for (ordinary) comparison spaces require us to take into account the lift into the unit tangent bundle. We will need the following lemma.

Lemma 3.4.5. *Let A and B be finite sets, $f_1, f_2 \colon A \to \mathbb{R}$ and $g_1, g_2 \colon B \to \mathbb{R}$. Suppose for every $a \in A$, there is a $b \in B$ such that $|f_i(a) - g_i(b)| \le \epsilon$ for $i = 1, 2$. Symmetrically, suppose for every $b \in B$, there is an $a \in A$ such that $|f_i(a) - g_i(b)| \le \epsilon$ for $i = 1, 2$. Then*

$$T \coloneqq \min_{a \in A} \max\{f_1(a), f_2(a)\} - \min_{b \in B} \max\{g_1(b), g_2(b)\}$$

satisfies $|T| \le \epsilon$.

Proof. By symmetry, we can assume $T \ge 0$. Suppose $b = \mathrm{argmin}_{b \in B} \max\{g_1(b), g_2(b)\}$. By symmetry, we can furthermore assume $g_1(b) \ge g_2(b)$. By assumption, we have $a_b \in A$ such that $|f_i(a_b) - g_i(b)| \le \epsilon$ for $i = 1, 2$. Now

$$0 \le T \le \max\{f_1(a_b), f_2(a_b)\} - g_1(b)$$

The choice of a_b implies, $f_1(a_b) \le g_1(b) + \epsilon$ and $f_2(a_b) \le g_2(b) + \epsilon \le g_1(b) + \epsilon$ which finishes the proof. ☐

We can now present a stability result for cycing signatures

Theorem 3.4.6. *Let Γ and $\tilde{\Gamma}$ be time series in $\mathbb{R}^m$. Let Y be a joint comparison space for $\Gamma, \tilde{\Gamma}$ and I its interval of admissible radii. Let ρ and $\tilde{\rho}$ be the lifts to the unit tangent bundle $\mathrm{UT}(\mathbb{R}^m)$ and L, $\tilde{L}$ their Lipschitz constants. Furthermore, let C be the constant in the dynamic metric. Let $J \in \mathrm{Int}(\mathbb{R})$ be an interval and define $\gamma = \Gamma_J$ and $\tilde{\gamma} = \tilde{\Gamma}_J$. Then*

$$d_I(\mathrm{Cyc}(\gamma, Y), \mathrm{Cyc}(\tilde{\gamma}, Y)) \le C \min\{L, \tilde{L}\}\, d_H(\gamma, \tilde{\gamma}) + C\|\rho - \tilde{\rho}\|_\infty \tag{3.7}$$

and thus also

$$d_I(\mathrm{Cyc}(\gamma, Y), \mathrm{Cyc}(\tilde{\gamma}, Y)) \le C \min\{L, \tilde{L}\}\, d_{\mathrm{ts}}(\Gamma, \tilde{\Gamma}) + C\|\rho - \tilde{\rho}\|_\infty \tag{3.8}$$

where we regard the cycling signatures as I-indexed filtered vector spaces.

Proof. Without loss of generality, can assume $L \le \tilde{L}$. Set $f(y) = d_C^{dyn}(y, \rho(\gamma))$ and

$g(y) = d_C^{dyn}(y, \tilde{\rho}(\tilde{\gamma}))$. Since γ and $\tilde{\gamma}$ are finite, we have

$$\|f - g\|_\infty = \sup_{y \in Y} \left| d_C^{dyn}(y, \rho(\gamma)) - d_C^{dyn}(x, \tilde{\rho}(\tilde{\gamma})) \right| = \sup_{y \in Y} \left| \min_{x \in \rho(\gamma)} d_C^{dyn}(y, x) - \min_{\tilde{x} \in \tilde{\rho}(\tilde{\gamma})} d_C^{dyn}(y, \tilde{x}) \right|$$

$$= \sup_{(q,w) \in Y} \left| \min_{(p,v) \in \rho(\gamma)} \max \left\{ \|p - q\|_2, C\|v - w\|_2 \right\} - \min_{(\tilde{p},\tilde{v}) \in \tilde{\rho}(\tilde{\gamma})} \max \left\{ \|\tilde{p} - q\|_2, C\|\tilde{v} - w\|_2 \right\} \right|.$$

Now fix $(q, w) \in Y$ and note that Lemma 3.4.5 applies to the term inside the maximum in the last line. Fix $x = (p, v) \in \rho(\gamma)$. Choose $\tilde{x} = (\tilde{p}, \tilde{w}) \in \tilde{\rho}(\tilde{\gamma})$ such that $d_C^{dyn}(x, \tilde{x}) \leq d_H(\rho(\gamma), \tilde{\rho}(\tilde{\gamma}))$. We have

$$\left| \|p - q\|_2 - \|\tilde{p} - q\|_2 \right| \leq \|p - \tilde{p}\|_2 \leq d_C^{dyn}(x, \tilde{x}).$$

Note that $v = \rho(p)$, $\tilde{v} = \tilde{\rho}(\tilde{p})$. Thus

$$\begin{aligned}
&\left| C\|v - w\|_2 - C\|\tilde{v} - w\|_2 \right| \\
= &\left| C\left\|\rho(p) - w\right\|_2 - C\left\|\tilde{\rho}(\tilde{p}) - w\right\|_2 \right| \\
\leq &\, C\left\|\rho(p) - \tilde{\rho}(\tilde{p})\right\|_2 \leq C\left\|\rho(p) - \rho(\tilde{p})\right\|_2 + C\left\|\rho(\tilde{p}) - \tilde{\rho}(\tilde{p})\right\|_2 \\
\leq &\, CL\|p - \tilde{p}\|_2 + C\|\rho + \tilde{\rho}\|_\infty \leq CL\, d_C^{dyn}(x, \tilde{x}) + C\|\rho + \tilde{\rho}\|_\infty =: \bar{\epsilon}.
\end{aligned}$$

We claim that we can now apply Lemma 3.4.5. For this, set $A = \gamma$, $B = \tilde{\gamma}$. As in the lemma, define $f_1(x) = \|p - q\|$, $f_2(x) = C\|v - w\|$, $g_1(\tilde{x}) = \|\tilde{p} - q\|$, $f_2(\tilde{x}) = C\|\tilde{v} - w\|$. We have shown that for every $a \in A$, there is $b \in B$ such that $|f_i(a) - g_i(b)| \leq \bar{\epsilon}$ for $i = 1, 2$. The analogous statement for $b \in B$ follows by exchanging γ and $\tilde{\gamma}$ in the previous argument.

In summary, we now have

$$d_H(\rho(\gamma), \tilde{\rho}(\tilde{\gamma})) = \|f - g\|_\infty \leq CL\, d_C^{dyn}(x, \tilde{x}) + C\|\rho + \tilde{\rho}\|_\infty \tag{3.9}$$

We can now regard $\rho(\gamma)$ and $\tilde{\rho}(\tilde{\gamma})$ as time series in the unit tangent bundle with simplified comparison space Y. It now follows from Theorem 3.4.4, that

$$d_I(\mathrm{Cyc}(\rho(\gamma), \mathrm{id}_Y), \mathrm{Cyc}(\tilde{\rho}(\tilde{\gamma}), \mathrm{id}_Y)) \leq d_H(\rho(\gamma), \tilde{\rho}(\tilde{\gamma})). \tag{3.10}$$

Combining (3.9) and (3.10) with our assumption $L \leq \tilde{L}$ finishes the proof. □

4 Computation of Cycling Signatures

In this chapter, we describe an algorithm to compute the cycling signature of a segment γ of a time series Γ. The cycling signature is a filtered vector space over a finite field $\mathbb{F}$ that we fix throughout this section. Given a segment γ, we compute a basis of $\mathrm{Cyc}_r(\gamma)$ in the sense of Definition 2.2.4.

Our algorithm for computing cycling signatures requires *(i)* the construction of a comparison space Y, *(ii)* a discretization of the map $H_1(i_\gamma)$ that appears in the definition of the cycling signature (see Eq. (3.4)), and, *(iii)* an algorithm to compute the image of this map. Before going into details, we first give a brief overview.

Cubical Comparison Spaces A cubical comparison space (see Definition 4.1.3) for a time series $\Gamma \subset \mathbb{R}^d$ is constructed by first lifting all points of Γ to the unit tangent bundle $\mathrm{UT}_\infty(\mathbb{R}^d)$ with respect to the l_∞ norm and then covering the resulting point cloud with boxes on a specified grid. Strictly speaking, a cubical comparison space Y_∞ is not quite a comparison space in the sense of Definition 3.2.1 since it is a neighborhood of a lift in $\mathrm{UT}_\infty(\mathbb{R}^d)$ and not in $\mathrm{UT}_2(\mathbb{R}^d)$. However, we show that Y_∞ is homotopy equivalent to a comparison space Y through radial projection. The reasons for working in the modified setting is that cubical grid complexes in UT_∞ are convenient to work with and cubical covers in $\mathrm{UT}_\infty(\mathbb{R}^d)$ lead to cubical complexes with fewer cubes than cubical covers in $\mathrm{UT}_2(\mathbb{R}^d)$.

Discretization of the Cycling Signature The cycling signature is the image of the map in the top row of the following diagram, the bottom row contains the corresponding discretizations which we introduce in the following.

$$\begin{tikzcd} H_1(O(\rho(\gamma))) \arrow[r, "H_1(i_\gamma)"] & H_1(\Delta Y) \\ H_1(\check{\mathrm{C}}(\rho(\gamma))) \arrow[r, "H_1(\phi)"] \arrow[u, "\simeq"] & H_1(\Delta Y_\infty) \arrow[u, "\simeq"] \end{tikzcd} \tag{4.1}$$

In the previous paragraph, we outlined the construction of the comparison space Y using a cubical comparison space Y_∞. Since Y_∞ is a cubical complex, we use it directly as the discretization of Y.

As often in topological data analysis, the offset filtration is discretized using the Čech filtration. We show equivalence of these two filtrations using a functorial Nerve theorem 2.1.9 for closed convex subsets of Euclidean space. The reason for using this theorem is that it yields an explicit and simple description of the homotopy equivalence $|\check{\mathrm{C}}(\rho(\gamma))| \to O(\rho(\gamma))$. To apply this theorem, we have to ensure convexity of the cover elements in the offset filtration. We do this by passing from internal thickenings (where the thickened subsets are contained in the unit tangent bundle) to external thickenings where the thickened subsets are subsets of the tangent bundle.

Using the discretizations, we now have a sequence of maps

$$|\check{\mathrm{C}}(\rho(\gamma))| \longrightarrow O(\rho(\gamma)) \longrightarrow \Delta Y \longrightarrow \Delta|Y_\infty|$$

where the first and last maps are the homotopy equivalences from the discretization and the middle map comes from the definition of the cycling signature. To obtain the map on the bottom of Eq. (4.1), we apply an acyclic carrier theorem to this sequence of maps. We explicitly construct an acyclic carrier from the Čech filtration to the cubical comparison space and use this to compute a chain map $\phi\colon C_1(\check{\mathrm{C}}(\rho(\gamma))) \to C_1(\Delta Y_\infty)$ which induces the map in homology on the bottom of Eq. (4.1).

Images of Maps of Filtered Vector Spaces This can be done using a simple variation of the standard algorithm from persistent homology.

This chapter is structured as follows. In Section 4.1, we introduce the bottom row of diagram 4.1. In Section 4.2, we explain how the image of a map of filtered vector spaces can be computed. In Section 4.3, we describe our algorithm for computing cycling signatures and discuss our implementation.

4.1 Discretization of the Cycling Signature

In this section, we introduce the bottom row of diagram 4.1 and the respective homotopy equivalences.

4.1.1 Internal and External Thickenings

We now treat the left vertical arrow of the diagram (4.1). In the definition of the cycling signature, the thickenings of segments γ were taken internally in the unit tangent bundle, i.e. we considered

$$O_r^{I,C}(\rho(\gamma)) := O_r(\rho(\gamma)) = \{p \in \mathrm{UT}(X) \mid d_C(p, \rho(\gamma)) \leq r\},$$

where I indicates that the thickening is internal and C is the metric parameter. To apply Theorem 2.1.9, we want to work with external thickenings, where the balls are taken in $\mathrm{T}(\mathbb{R}^d) = \mathbb{R}^{2d}$, i.e.

$$O_r^{E,D}(\rho(\gamma)) = \{p \in \mathrm{T}(X) \mid d_D(p, \rho(\gamma)) \leq r\}.$$

As it turns out, the two filtrations are equivalent up to a rescaling.

Proposition 4.1.1. *Let $\gamma \subset \mathbb{R}^d$ be finite. With $C \geq 0$, $r \in [0, C)$ and $D_r = r/(C\theta(r/C))$, the map*

$$\alpha_r \colon O_r^{E,D_r}(\rho(\gamma)) \to O_r^{I,C}(\rho(\gamma)), \quad (p, v) \mapsto (p, v/\|v\|_2)$$

is a homotopy equivalence. The collection of maps α is furthermore a map of filtered topological spaces.

Proof. See Proposition 4.4.6. □

Corollary 4.1.2. *Let $\gamma \subset \mathbb{R}^d$ finite, $C \geq 0$ and $D_r = r/(C\theta(r/C))$. There is a commutative diagram of $[0, C)$-indexed filtered topological spaces*

$$\begin{tikzcd}
O^{I,C}(\rho(\gamma)) & O^{E,D_\bullet}(\rho(\gamma)) \arrow[l, "\alpha"'] \\
|\,\check{\mathrm{C}}(\gamma, \mathrm{UT}, d_C)| \arrow[u, "g"] \arrow[r, "h"] & |\,\check{\mathrm{C}}(\gamma, \mathrm{T}, d_{D_\bullet})| \arrow[u, "f"']
\end{tikzcd}$$

where for fixed filtration value all maps are homotopy equivalences. We write $D_\bullet$ to emphasize that the parameter depends on r.

Proof. Note that $\check{\mathrm{C}}_r(\gamma, \mathrm{UT}, d_C) = \check{\mathrm{C}}_r(\gamma, \mathrm{T}, d_{D_r})$ for $r \in [0, C)$, since intersections of internal balls with radius r and metric d_C are nonempty if and only if external balls with radius r and metric $d_{D(r)}$ are nonempty (see Corollary 4.4.7). Thus, we can choose the same geometric realization for both of them.

Since $O^{E,D_\bullet}(\rho(\gamma))$ is a family of subsets of $\mathbb{R}^{2d}$, by Theorem 2.1.9, the right vertical map can be chosen to be a map $f = (f_r)_r$ of filtered topological spaces where each component is a homotopy equivalence. For the map α, this follows from Proposition 4.1.1.

We can thus construct the left vertical map as $g_r = \alpha_r \circ f_r \circ \mathrm{id}_r$. □

Applying homology to the map g from the corollary, we get an isomorpism for the left vertical arrow in diagram (4.1).

4.1.2 Cubical Comparison Spaces

We now treat the right vertical arrow in diagram (4.1) and explain how we construct the comparison space Y. For this, we introduce the unit tangent bundle with respect to the l_∞-norm

$$\mathrm{UT}_\infty(\mathbb{R}^d) = \{(p, v) \in \mathrm{T}(\mathbb{R}^d) : \|v\|_\infty = 1\}$$

and note that there is a homeomorphism $\eta\colon \mathrm{UT}_2(\mathbb{R}^d) \to \mathrm{UT}_\infty(\mathbb{R}^d), (p, v) \mapsto (p, v/\|v\|_\infty)$. We can thus obtain a comparison space Y by finding a neighborhood U of $\eta(\rho(\Gamma))$ and then taking $Y = \eta^{-1}(U)$. Computationally, this is convenient, since spheres S^d_∞ are much easier to cover with boxes than spheres S^d_2.

Given a point $p \in \mathbb{R}^d$ and a radius $r \in \mathbb{R}_{>0}$, we write $Q_r(p) = \prod_{i=1}^d \left[p_i - \frac{r}{2}, p_i + \frac{r}{2}\right]$ for the box with side lengths r centered at p. To construct the comparison space, we use boxes of the form $Q_{r,k}(p, q) = Q_r(p) \times Q_{1/k}(q) \subset T_\infty(\mathbb{R}^d)$ where $r, k \in \mathbb{R}_{>0}$, $p \in r\mathbb{Z}^d$ and $q \in 1/k\mathbb{Z}^d$.

Note that the collection of these boxes covers $\mathbb{R}^{2d}$. To cover the unit tangent bundle we fix two parameters: the space box size $r \in \mathbb{R}_{>0}$ and the sphere box size $1/k$, where $k \in \mathbb{N}_{>0}$, and consider the following cover of $\mathrm{UT}_\infty(\mathbb{R}^d)$

$$\mathcal{Q}^d_{r,k} = \{Q_{r,k}(p, q) \mid p \in r\mathbb{Z}^d,\ q \in 1/k\mathbb{Z}^d,\ \|q\|_\infty = 1\}. \tag{4.2}$$

We can obtain comparison spaces by covering a neighborhood of $\eta(\rho(\Gamma))$, e.g.

$$Y_\infty = \{Q \in \mathcal{Q}^d_{r,k} \mid Q \cap \eta(O_\epsilon(\rho(\Gamma))) \neq \emptyset\}, \quad \text{for some } \epsilon \geq 0,$$

and then taking $\eta^{-1}(p_\infty(|Y_\infty|))$, where $p_\infty\colon \mathrm{T}(\mathbb{R}^d) \setminus (\mathbb{R}^d \times \{0\}) \to \mathrm{UT}_\infty(\mathbb{R}^d), (p, v) \mapsto (p, v/\|v\|_\infty)$ and we write $|Y_\infty| = \bigcup Y_\infty$.

Definition 4.1.3. A *cubical comparison space* is a set $Y_\infty \subset \mathcal{Q}^d_{r,k}$ for some $r > 0$ and $k \in \mathbb{N}$ such that $Y = \eta^{-1}(p_\infty(|Y_\infty|))$ is a comparison space for $\Gamma \subset \mathbb{R}^d$.

Proposition 4.1.4. *For any $K \subset \mathcal{Q}^d_{r,k}$, where $r > 0$ and $k \in \mathbb{N}_{>0}$ we have a homotopy equivalence $p_\infty(|K|) \simeq |K|$.*

Proof. See Proposition 4.4.13 □

Any $Q \in \mathcal{Q}^d_{r,k}$ is a polytope and thus has faces of all dimensions $\leq d$, we denote the collection of all faces by $\downarrow Q$. Given $K \subset \mathcal{Q}^d_{r,k}$, we write $\downarrow K = \bigcup_{Q \in K} \downarrow Q$.

4.1.3 Discretization of the Map

In this section, we introduce the map on the bottom row of the diagram (4.1).

Throughout this section, let $\Gamma \subset \mathbb{R}^d$ be a time series, γ a segment of Γ and Y_∞ a cubical comparison space for Γ with respect to d_C and admissible radii $I \subset [0, C)$. We furthermore denote by $R \in I$ the largest critical point of $(\check{\mathrm{C}}_r(\rho(\gamma)))_{r \in I}$ in I.

Proposition 4.1.5. *Consider the map*

$$\Theta\colon |\operatorname{sd}\check{\mathrm{C}}(\rho(\gamma))| \to \Delta Y_\infty, \quad \Theta = \eta \circ \alpha \circ f,$$

where $\alpha \circ f$ is as in Corollary 4.1.2 and η as in Definition 4.1.3. There is a subdivision K_R of $\operatorname{sd}\check{\mathrm{C}}_R(\rho(\gamma))$ such that

$$F(\sigma) = {\downarrow}\{Q \in Y_\infty \mid Q \cap \Theta_R(\sigma) \neq \emptyset\}$$

is an acyclic carrier (see [Mun84], Theorem 13.3) from K_R to Y_∞.

Proof. Step 1: We claim that there is a subdivision K_R of $\operatorname{sd}\check{\mathrm{C}}_R(\rho(\gamma))$ with the property that for any face H of the l_∞ unit ball $B_\infty(0,1) \subset \mathbb{R}^d$, the set $\Theta_R^{-1}(\mathbb{R}^d \times H)$ is a subcomplex of K_R.

The set $\gamma_E^R := O_R^{E,D_R}(\rho(\gamma))$ is compact and does not intersect the zero section $\mathbb{R}^d \times \{0\} \subset \mathrm{T}(X)$ since, with respect to d_C, the distance between $\mathbb{R}^d \times \{0\}$ and $\mathrm{UT}(\mathbb{R}^d)$ is exactly C while $R < C$ by assumption. Thus there are constants r_1, r_2, r_3 such that

$$\gamma_E^R \subset \overline{B_\infty(0,r_1)} \times (\overline{B_\infty(0,r_2)} \setminus B_\infty(0,r_3)) =: A \subset \mathrm{T}(X).$$

We claim that A is the geometric realization of a simplicial complex L. In fact, the set A can be decomposed into sets $A_H = A \cap (\mathbb{R}^d \times \{rH \mid r > 0\})$ where H is a face of $B_\infty(0,1) \subset \mathbb{R}^d$. Each of the A_H is a convex polyhedron since, pictorially, it is the product of the cube $B_\infty(0,r_1)$ with the pyramid $\{rH \mid r > 0\}$ that has the corner at the origin chopped off. Furthermore, the intersection of two such sets is either empty, or a face of either of the polyhedra. Therefore, A has a natural cell structure that can be subdivided into a simplicial complex L such that $A = |L|$, see e.g. Prop. 2.9 in [RS72].

By (2.1), $f_R\colon |\operatorname{sd}\check{\mathrm{C}}_R(\rho(\gamma))| \to \gamma_E^R \subset |L|$ is a linear map on each simplex. Thus, there are simplicial subdivisions K_R of $\operatorname{sd}\check{\mathrm{C}}_R(\rho(\gamma))$ and L' of L such that f_R is a simplicial map with respect to these simplicial complexes (Lemma 2.13 in [RS72]). We turn K_R into a filtered simplicial complex in the obvious way, i.e. by $K_r = \{\sigma \in K_R \mid |\sigma| \subset |\check{\mathrm{C}}_r(\rho(\gamma))|\}$.

Step 2: We need to show that $F(\sigma)$ is nonempty and acyclic, and that $F(\hat{\sigma}) \subset F(\sigma)$ if $\hat{\sigma}$ is a face of σ. Since the latter property is clear by definition, we now focus on the former.

Let $\sigma \in K_R$ and set $\tau = \Theta_R(\sigma)$. By above subdivision, we have $\tau \subset \mathbb{R}^d \times H$ for a facet H of $B_\infty(0,1)$. Therefore, restricted to σ, the map Θ_R is given by the composition of the linear map f_R with $\eta \circ \alpha_R$, which in the tangent component is the radial projection onto H. Both of these maps preserve compactness and convexity, thus τ is compact and convex. It thus follows from Corollary 4.4.16 that $F(\sigma)$ is acyclic. □

Corollary 4.1.6. *There is a commutative diagram*

$$\begin{tikzcd} H_1(|\operatorname{sd}\check{\mathrm{C}}(\rho(\gamma))|) \arrow[r, "H_1(\Theta)"] & H_1(\Delta|Y_\infty|) \arrow[d, "H_1(\psi_2)"'] \\ H_1(K) \arrow[u, "H_1(\psi_1)"] \arrow[r, "H_1(\nu)"] & H_1(\Delta Y_\infty) \end{tikzcd}$$

where K is a subdivision of $\operatorname{sd}\check{\mathrm{C}}(\rho(\gamma))$ and ψ_1 and ψ_2 are chain maps that induce the isomorphism between singular and cellular homology.

Proof. Let K be the subdivision and F be the acyclic carrier from Proposition 4.1.5. We get an algebraic acyclic carrier by setting $G(\sigma) = C(F(\sigma))$ to be the cellular chain complex of $F(\sigma)$. By the acyclic carrier theorem, there is a chain map $\phi_R \colon C(K_R) \to C(Y_\infty)$ carried by G. We turn this into a filtered map by the appropriate restrictions.

Let $C(\Theta_R)$ denote the chain map induced by Θ_R. We now claim that the chain map $\psi_2 \circ C(\Theta_R) \circ \psi_1$ is also carried by G. To see this, recall that the isomorphism between singular and cellular homology has the property that for any subcomplex L_R of K_R, and any subcomplex Z of Y_∞ we have $\psi_1(L_R) \subset \psi_1(|L_R|)$, $\psi_2(|Z|) \subset \psi_2(Z)$. By construction, we have $\Theta_R(|\sigma|) \subset |F(\sigma)|$ for any simplex $\sigma \in K_R$ and thus

$$\psi_2 \circ C(\Theta_R) \circ \psi_1(C(\sigma)) \subset \psi_2 \circ C(\Theta_R)(C(|\sigma|)) \subset \psi_2(\Theta_R(|\sigma|)) \subset \psi_2(|F(\sigma)|) \subset C(F(\sigma)).$$

Thus, $\psi_2 \circ C(\Theta_R) \circ \psi_1$ is chain homotopic to ϕ_R. We can turn $\psi_2 \circ C(\Theta_R) \circ \psi_1$ into a filtered chain map using the appropriate restrictions to obtain the top map of the diagram. Furthermore, we note that the chain homotopy between $\psi_2 \circ C(\Theta_R) \circ \psi_1$ and ϕ_R respects the acyclic carrier, which implies that appropriate restrictions of those give a chain homotopy between $\psi_2 \circ C(\Theta_r) \circ \psi_1$ and ϕ_r for every r. Therefore, the above diagram is a commutative diagram of persistence modules. $\square$

Corollary 4.1.7. *Let $Y = \eta^{-1}(p_\infty(|Y_\infty|))$. There is a commutative diagram*

$$\begin{tikzcd} H_1(O(\rho(\gamma))) \arrow[r, "H_1(i_\gamma)"] & H_1(\Delta Y) \arrow[d, "\simeq"'] \\ H_1(\check{\mathrm{C}}(\rho(\gamma))) \arrow[u, "\simeq"] \arrow[r, "H_1(\phi)"] & H_1(\Delta Y_\infty) \end{tikzcd}$$

Proof. Each space in this diagram is isomorphic to the corresponding space in the preceding corollary. For the top left and top right spaces, this follows from the Nerve theorem 2.1.9 and Proposition 4.1.4. For the bottom row, this follows from topological invariance of barycentric subdivision. □

4.2 Maps of Filtered Vector Spaces

We now explain how to compute a basis for the image of a map from a persistence module into a filtered vector space. Throughout this section, all persistence modules are pointwise finite dimensional (i.e. all V_i are finite dimensional) and indexed by $T = \{0, 1, \ldots, q\}$.

Definition 4.2.1 (c.f. [BS23])**.** Let V be a persistence module. Its *immortal part* is the persistence module $V^\infty = (V_i^\infty)_{i \in T}$ where

$$V_i^\infty = \{v \in V_i \mid v = 0 \text{ or } V_{i,q}(v) \neq 0\},$$

its *mortal part* is the persistence module $V^\dagger = (V_i^\dagger)_{i \in T}$ where

$$V_i^\dagger = \{v \in V_i \mid V_{i,q}(v) = 0\},$$

and the structure maps are the appropriate restrictions.

A filtered vector space is equal to its immortal part and has trivial mortal part. More generally, the immortal part of any persistence module V is isomorphic to a filtered vector space.

The cycling signature is expressed as the image of a morphism form a persistence module into a filtered vector space. In the following, we show that, computationally, it is enough to consider the case of a map between filtered vector spaces.

Proposition 4.2.2. *Let $\phi \colon V \to W$ be a morphism of persistence modules. If W is a filtered vector space, then*

1. $\operatorname{im} \phi$ is a filtered vector space, and
2. $\operatorname{im} \phi = \operatorname{im} \phi|_{V^\infty}$.

Proof. For the first claim, we need to show $\operatorname{im} \phi_i \subset \operatorname{im} \phi_j$ for $i \leq j$. Note that post-composition with an inclusion $W_{i,j}$ does not change the image, thus $\operatorname{im} \phi_i = \operatorname{im} W_{i,j} \circ \phi_i$. The claim now follows from

$$\operatorname{im} \phi_i = \operatorname{im} W_{i,j} \circ \phi_i = \operatorname{im} \phi_j \circ V_{i,j} \subset \operatorname{im} \phi_j.$$

For the second claim, note that a persistence module is a direct sum of its mortal and immortal part. It is therefore enough to show that $\operatorname{im} \phi_{|V^\dagger} = 0$. For this, fix $i \in T$ and note that for an arbitrary $v \in V_i^\dagger$

$$\phi_i(v) = W_{i,q} \circ \phi_i(v) = \phi_q \circ V_{i,q}(v) = 0$$

where we again use that $W_{i,q}$ is an inclusion map. □

As a consequence of the above proposition, the image of a morphism $\phi\colon V \to W$ of a persistence module into a filtered vector space can be computed by considering the restriction $\phi|_{V^\infty}$ which is a map of filtered vector spaces. In the setting of vector spaces, the image of a linear map can be computed using matrix reduction as described in the following. Recall that $\mathbb{F}$ is some fixed field.

Definition 4.2.3. The pivot $\mathrm{pivot}(v) \in \mathbb{N}$ of a vector $v \in \mathbb{F}^m$ is the largest $i \in 1 : m$ such that $v_i \neq 0$ or zero if no such i exists. A matrix is called *reduced* if all non-zero pivots of its columns are distinct.

Any matrix $M \in \mathbb{F}^{m\times n}$ can be reduced, i.e. there is an invertible and upper triangular matrix $U \in \mathrm{Gl}_n(\mathbb{F})$ such that $R = MU \in \mathbb{F}^{m\times n}$ is reduced (see e.g. [EH08]) Suppose we have a decomposition $R = MU$. Since R is reduced, a basis for $\operatorname{im} M$ is given by $\{R_i \mid \mathrm{pivot}\, R_i \neq 0\}$. A different basis can be obtained from the columns of M.

Lemma 4.2.4. *Suppose $R = MU$ with $M \in \mathbb{F}^{m\times n}$, $R \in \mathbb{F}^{m\times n}$ reduced and $U \in \mathrm{Gl}_n(\mathbb{F})$ upper triangular. Then a basis of $\operatorname{im} M$ is given by $\{M_i \mid \mathrm{pivot}(R_i) \neq 0\}$.*

Proof. Let $R_{1:i}$ and $M_{1:i}$ be the matrices consisting of the first i columns of R and M, respectively. Since U is invertible and upper triangular, $\operatorname{im} R_{1:i} = \operatorname{im} M_{1:i}$ for all i. In particular, a column R_{i+1} is linearly independent from $\operatorname{im} R_{1:i}$ if and only if a column M_{i+1} is linearly independent from $\operatorname{im} M_{1:i}$. The claim now follows, since $\{R_i \mid \mathrm{pivot}(R_i) \neq 0\}$ is a basis for $\operatorname{im} M$. □

Given a linear map $\phi\colon V \to W$ of (unfiltered) vector spaces, a basis $v_1, \dots, v_n$ of V and a basis $\alpha_1, \dots, \alpha_m$ of the dual space W^* of W, a basis of $\operatorname{im} \phi$ can be computed by reducing the matrix $M = [\alpha_i(\phi(v_j))]_{i,j}$. If $R = MU$ as above, a basis for $\operatorname{im} \phi$ is given by

$$\{\phi(v_j) \mid \mathrm{pivot}(R_j) \neq 0\},$$

since the map $w \mapsto (\alpha_1(w), \dots, \alpha_m(w))$ induces an isomorphism $W \cong \mathbb{F}^m$.

This algorithm readily generalizes to the setting of a map $\phi : V \to W$ of filtered vector spaces by reducing the matrix of $\phi_q\colon V_q \to W_q$, i.e. at the final index. The only added requirement is that the basis elements need to be ordered by birth.

Proposition 4.2.5. *Let $\phi\colon V \to W$ be a map of filtered vector spaces. Let furthermore B^V and B^{W^*} be bases for V and the dual space W^* of W, respectively, and suppose*

1. $B^V_q = \{v_1, \dots, v_n\}$, with $\mathrm{birth}(v_i) \leq \mathrm{birth}(v_j)$ for $i \leq j$,

2. $B_q^{W^*} = \{\alpha_1, \ldots, \alpha_m\}$, with $\mathrm{birth}(\alpha_i) \leq \mathrm{birth}(\alpha_j)$ for $i \leq j$.

Let $M = [\alpha_i(\phi_q(v_j))]_{i,j}$ and suppose $R = MU$ with R reduced and U invertible and upper triangular. Then, a basis B for $\operatorname{im}\phi$ *is given by*

$$B_i = \{\phi_i(v_j) \mid j \leq i \textit{ and } R_j \neq 0\}.$$

Proof. Clearly, $B_i \subset B_j$ for $i \leq j$. Now fix $i \in T$ and let $R_{1:i}$ and $M_{1:i}$ be the matrices consisting of the first i columns of R and M, respectively. Since the map $w \mapsto (\alpha_1(w), \ldots, \alpha_m(w))$ induces an isomorphism $W \cong \mathbb{F}^m$, it is enough to show that $\tilde{B}_i = \{M_j \mid j \leq i \text{ and } R_j \neq 0\}$ is a basis of $\operatorname{im} M_{1:i}$. This follows from Lemma 4.2.4 since we have $R_{1:i} = M_{1:i}U_{1:i,1:i}$. □

Remark 4.2.6. In the context of persistent homology, the use of the pairing between homology and cohomology is sometimes referred to as annotation [BCC+12, DFW14].

We can now use this result in order to compute the cycling signature in the following way: Let Γ be a time series, Y a comparison space for Γ and γ a segment of Γ. Consider the Čech filtration $\check{\mathrm{C}}(\gamma)$. Since γ is a finite point cloud, there is a finite number of critical points $0 = r_0 \leq r_1 \leq r_2 \leq \ldots$ where the filtration changes. Let q be the largest integer with $r_q < r_0(Y)$ and define a filtered vector space via

$$V_k = \left(H_1(\check{\mathrm{C}}(\gamma))^\infty\right)_{r_k},$$

i.e. by evaluating the filtered vector space $H_1(\check{\mathrm{C}}(\gamma))^\infty$ at the critical point r_k. We then get a map $\phi\colon V \to \Delta H_1(Y)$ of filtered vector spaces indexed over $T = \{0, 1, \ldots, q\}$ to which we can apply Proposition 4.2.5 to get a basis $\tilde{B}$ of $\operatorname{im}\phi = \mathrm{Cyc}(\gamma, Y)$. We then set $B_r = \tilde{B}_i$, where i is the largest integer such that $r_i \leq r$. This yields a basis B of $\mathrm{Cyc}_r(\gamma, Y)$.

4.3 Algorithm and Implementation

To compute a cycling signature $\mathrm{Cyc}(\gamma, Y_\infty)$ using Proposition 4.2.5, we need to compute and reduce the matrix $M_{i,j} = [\alpha_i(\phi(v_j))]_{i,j}$ where

- the α_i are a basis of $H^1(Y_\infty)$ for some cubical comparison space Y_∞ of Γ, and
- the v_j are a basis for $H_1(\check{\mathrm{C}}(\rho(\gamma)))$.

In the following, we explain how we these steps are implemented in [Hie24]. The computation furthermore requires a choice of metric d_C and knowledge of the maximal filtration radius r_0 which determines the interval of admissible radii $I = [0, r_0)$. In Section 4.3.5 we provide heuristics for choosing these two values.

4.3.1 Computation of Cubical Comparison Spaces and their Cohomology

In practice, we construct the cubical comparison space Y_∞ by fixing parameters $r, k \in \mathbb{N}$ and collecting all boxes in $\mathcal{Q}_{r,k}$ that contain a point from Γ. While this does not guarantee that the map Θ_r from Proposition 4.1.5, and thus also the chain map ϕ from Corollary 4.1.7, is well-defined for $r > 0$, we have implemented computational heuristics that enable us to construct the necessary map in homology (see Section 4.3.6). We then compute a basis $\alpha_1, \dots, \alpha_m$ of $H^1(Y_\infty)$ using a version of the coreduction algorithm [MB08]. Since the computation of the basis only depends on Γ, but not on any specific segment, we can do this once and reuse it for all segments in Γ.

4.3.2 Computing the Persistent Homology of Segments

Instead of the Čech filtration, it is common practice in computational topology, to compute using the Vietoris-Rips filtration

$$\mathrm{VR}_r(P) = \{\sigma \subset P \mid \operatorname{diam} \sigma \leq r\}.$$

We adopt this heuristic in the following way. First, a persistence code (in our case Ripserer [Ču20], a Julia implementation of Ripser [Bau21]) is used to compute a barcode $(b_i, d_i)_i$ and generators $G = (g_i)_i$ for $H_1(\mathrm{VR}(\rho(\gamma)))^\infty$. Later, in Section 5.3, we present an alternative to obtain such a generating set. Second, we make use of the fact that the 1-skeleta of Čech and Vietoris-Rips complexes are identical. Therefore, any chain in G is also a chain in $C_1(\check{\mathrm{C}}(\rho(\gamma)))$. We can thus regard the set G as an approximate generating set for $H_1(\check{\mathrm{C}}(\rho(\gamma)))$. We note that this is a heuristic, since in general the map $C_1(\mathrm{VR}(\rho(\gamma))) \to C_1(\check{\mathrm{C}}(\rho(\gamma))$ does not induce a well-defined map in homology.

We then let G_∞ to be all generators in G that correspond to bars $(b_i, d_i]$ that contain the maximal filtration radius r_0, i.e. with $b_i < r_0 < d_i$, to obtain an 'approximate basis´ for the immortal part of $H_1(\check{\mathrm{C}}(\rho(\gamma))$.

4.3.3 Computation of the Map in Homology

We now explain how we compute $\phi(c)$ for some chain $c \in C_1(\check{\mathrm{C}}(\rho(\gamma))$ where ϕ is the map from Corollary 4.1.7. By definition, $c = \sum_k \lambda_k e_k$, $\lambda_k \in \mathbb{F}$, for some edges e_k in $\check{\mathrm{C}}(\rho(\gamma))$. By linearity, it is thus sufficient to compute the images $\phi(e_k)$ for all k.

So, fix some edge $e \in \check{\mathrm{C}}(\rho(\gamma))$. Formally, the map ϕ is the composition of a subdivision and the map induced by the acyclic carrier from Proposition 4.1.5. Computationally, we subdivide e such that two consecutive vertices of the subdivision lie in adjacent boxes and each segment lies on the same face of $\partial B_\infty(0,1)$ in the unit tangent component. We note that this takes care of the subdivision that is necessary because of the nerve theorem 2.1.9. We then construct a 1-chain as follows: each vertex in the subdivision is mapped to a distinguished point of a box in Y_∞ that contains it. Each edge $e = [v, w]$ in the subdivision is then mapped to a chain that is contained in the union of the boxes and has $\phi(w) - \phi(v)$ as its boundary (this is how a chain map is induced by an acyclic carrier, see [Mun84]). The sum of all these chains is then the image of the edge in $C_1(Y_\infty)$.

4.3.4 Computation of the Image Barcode

Both, the previously computed basis $\alpha_1, \ldots, \alpha_m$ of $H^1(Y_\infty)$ and the chains $\phi(c_1), \ldots, \phi(c_n)$, where $G_\infty = \{c_1, \ldots, c_n\}$, are represented as vectors in $\mathbb{F}^k$ where k is the number of edges in Y_∞. The entries of $M_{i,j} = [\alpha_i(\phi(v_j))]_{i,j}$ are then computed simply by taking the inner product of the previously computed vectors corresponding to α_i and $\phi(v_j)$. In our examples, the matrix dimensions of M are less than 5, thus we reduce M using a basic implementation of the standard reduction algorithm and then read off a basis of the cycling signature as explained in Proposition 4.2.5 and the paragraph following it.

4.3.5 Heuristics for the metric d_C and the maximal filtration radius r_0

To compute a cycling signature, we need to specify values for the constant C of d_C and the maximal filtration radius $r_0(Y)$. Given a cubical comparison space $Y_\infty \subset \mathcal{Q}_{r,k}$, we suggest $C = rk$ and $r_0 = r$ as default values. We now give a heuristic explanation for these choices. Our choices are motivated by Proposition 4.4.17 which states that a cubical complex consisting of radius R cubes can be thickened by $r < R$ without changing its homotopy type.

Suppose Y_∞ is constructed as above, i.e. consists of all boxes in $\mathcal{Q}^d_{r,k}$ that intersect with $\rho(\Gamma)$. Similar to Proposition 4.4.17, we can consider

$$Y_\infty^{s,C} = \{O_s(Q, d_C) \mid Q \in Y_\infty\}$$

where all the individual cubes are thickened with respect to d_C. Recall, that the cubes in $\mathcal{Q}^d_{r,k}$ are of the form $Q_r(p) \times Q_{1/k}(q)$. Thus, to have a thickening that preserves the nerve, the first component should be thickened less than r, the second component should be thickened by less than $1/k$. For the two components of the metric d_C from Eq. (3.3), this yields the equations $s < r$ and $Cs < 1/k$. Thus, with $C = rk$, we have that any thickening with $s < r$ preserves the nerve of $Y_\infty^{s,C}$. Furthermore, any other thickening that has this guarantee is contained in $|Y_\infty^{s,C}|$, in this sense our choices are therefore optimal.

4.3.6 Heuristics for the Map in Homology

As described in Section 4.1.3, the construction of the chain selector does not take into account the thickened boxes but only works with the original box cover. It thus may happen in the computation of the map in homology that a segment e of an edge in

$\check{\mathrm{C}}(\rho(\gamma))$ is contained in a box that is not contained in Y_∞. By our choice of parameters, this segment is contained in the union of the thickened boxes. As a heuristic, we then use the collection of these boxes in the chain map construction.

4.4 Auxiliary Results

4.4.1 Covers of Convex Compact Sets

The outer cubical cover of a set $X \subset \mathbb{R}^d$ is the cubical set

$$\mathcal{K}(X) = \downarrow \left\{ Q \in \mathrm{Cub}_d(\mathbb{R}^d) \mid \sigma \cap X \neq \emptyset \right\}.$$

We write $\mathcal{K}_p(X)$ for the set of p-dimensional cubes in $\mathcal{K}(X)$.

Proposition 4.4.1. *The outer cubical cover $\mathcal{K}(\sigma)$ of a convex compact set σ is acyclic.*

The proof of Proposition 4.4.1 will proceed by induction on the number of maximal cubes in $K(\sigma)$. In the induction step, a layer of cubes will be cut off via a Mayer-Vietoris argument.

We need several lemmas in preparation for proving the proposition.

Definition 4.4.2. A convex, compact set $\sigma \subset \mathbb{R}^d$ is in *general position* with respect to $\mathrm{Cub}_k(\mathbb{R}^d)$ if $\sigma \cap \mathrm{Int}\, Q \neq \emptyset$ for all $Q \in K_d(\sigma)$.

As an example, $\sigma = \{0\} \subset \mathbb{R}$ is not in general position, since it does not intersect the interior of the interval $[0,1] \in \mathcal{K}(\sigma)$. As it turns out, any compact convex set can be enlarged slightly to another compact convex set that is in general position and has the same outer cover. We prove this in the next two lemmas.

Lemma 4.4.3. *Let $\sigma \subset \mathbb{R}^d$ be convex and compact. Then there is $\bar{\epsilon} > 0$ such that for all $\epsilon \in [0, \bar{\epsilon})$,*

$$\mathcal{K}(\sigma) = \mathcal{K}(\sigma^\epsilon),$$

Proof. Let $L = \downarrow(\mathrm{Cub}(\mathbb{R}^d) \setminus \mathcal{K}(\sigma))$ be the cubical complex of all boxes that do not intersect σ. Any box $Q \in L$ has positive distance from σ, i.e., it is $d_2(Q, \sigma) > 0$. Furthermore, given any $C > 0$, there are only finitely many boxes with $d(Q, \sigma) < C$. In particular, there is a minimal distance from σ to L, i.e. the minimum

$$\bar{\epsilon} = \min\{\mathrm{dist}(\sigma, Q) \mid Q \in L\}$$

exists. Let $\epsilon \in [0, \bar{\epsilon})$. By the triangle inequality, any box in L has positive distance to σ^ϵ, thus $\mathcal{K}(\sigma) = \mathcal{K}(\sigma^\epsilon)$. $\square$

Lemma 4.4.4. *Let $\sigma \subset \mathbb{R}^d$ be convex and compact. Then there is an $\epsilon > 0$ such that*

1. $\mathcal{K}(\sigma) = \mathcal{K}(\sigma^\epsilon)$, and
2. σ^ϵ is in general position.

Proof. The previous lemma implies the existence of a $\bar{\epsilon}$ such that that the first condition is satisfied for all $\epsilon \in [0, \bar{\epsilon})$. We now fix any such ϵ.

For the second condition, let $Q \in \mathcal{K}(\sigma)$ be arbitrary. We pick any point $x \in \sigma \cap Q \neq \emptyset$ and note that $\emptyset \neq B(x, \epsilon) \cap \operatorname{Int} Q \subset \sigma^\epsilon \cap \operatorname{Int} Q$. Thus, σ^ϵ is in general position. □

Proof of Proposition 4.4.1. By Lemma 4.4.4, we can assume that σ is in general position.

We prove the claim by induction on the number of top-dimensional cubes in $\mathcal{K}(\sigma)$. The case $n = 1$ is simple, since a single cube is contractible. Now suppose $\mathcal{K}(\sigma)$ contains $n + 1$ top-dimensional cubes. Let d denote the dimension of the top-dimensional cubes. Since $n > 1$, there must be a $k \in \mathbb{N}$ such that projection on the k-th coordinate satisfies $p_k(K(\sigma)) = [m_1, m_2]$ with $m_2 - m_1 \geq 2$. Now let $l = m_2 - 1$.

We will now 'cut' $\mathcal{K}(\sigma)$ at the hyperplane $H = p_k^{-1}(\{l\})$. Setting $A = \downarrow\{Q \in \mathcal{K}_d(\sigma) \mid p_k(Q) \subset (-\infty, l]\}$ and $B = \downarrow\{Q \in \mathcal{K}_d(\sigma) \mid p_k(Q) \subset [l, \infty)\}$, we have the short exact sequence of chain complexes

$$0 \to C(A \cap B) \to C(A) \oplus C(B) \to C(A \cup B) \to 0$$

where the last term is equal to $C(\mathcal{K}(\sigma))$. Its long exact sequence is the Mayer-Vietoris sequence in homology. It thus suffices to show that the left and middle term in this sequence are acyclic.

We start by showing that A and B are acyclic. In order to apply the induction hypothesis, we argue that A and B are outer covers of compact convex sets. By general position, for any $Q \in K_d(\sigma)$, we can fix a point $x_Q \in \sigma \cap \operatorname{Int} Q$. Define

$$\delta_A = \frac{1}{2} \max\{l - p_k(x_Q) \mid Q \in A\}, \text{ and } \delta_B = \frac{1}{2} \min\{p_k(x_Q) - l \mid Q \in B\}.$$

Since the x_Q are contained in the interior of maximal cubes, we have $p_k(x_Q) \notin \mathbb{Z}$ and thus $\delta_A, \delta_B > 0$. We set $\sigma_A = \sigma \cap p_k^{-1}((-\infty, l - \delta_A])$ and $\sigma_B = \sigma \cap p_k^{-1}([l + \delta_B, \infty)$ and claim $A = \mathcal{K}(\sigma_A)$ and $B = \mathcal{K}(\sigma_B)$.

To show $A \subset \mathcal{K}(\sigma_A)$, fix a d-dimensional cube $Q \in A$. We have previously fixed a point $x_Q \in \sigma \cap \operatorname{Int} Q$. An elementary estimate shows $p_k(x_Q) < l - \delta_A$, thus we have $x_Q \in \sigma_A$. Since Q is the unique cube with x_Q in its interior, we must have $Q \in \mathcal{K}(\sigma_A)$. Conversely, if $Q \in \mathcal{K}(\sigma_A)$, we have $Q \in \mathcal{K}(\sigma)$, by the definition of outer cubical covers. Since $\sigma_A \subset p_k^{-1}((-\infty, l - \delta_A])$, we must have $Q \subset p_k^{-1}((-\infty, l])$ and thus $Q \in A$. This shows that also $A \supset \mathcal{K}(\sigma_A)$. Similar arguments work to show $B = \mathcal{K}(\sigma_B)$.

We now have that A and B are outer covers of compact convex sets. As we show next, both sets have less than $n+1$ top-dimensional cubes. To see this, note that any d-dimensional cube $Q \in K(\sigma)$ is either contained in A or in B and that, by the definition of l, both sets contain at least one d-dimensional cube. By induction, $C(A)$ and $C(B)$ are therefore acyclic.

It is left to show that $A \cap B$ is acyclic. For this, note that $\hat{\sigma} = \sigma \cap H$ is compact and convex. Furthermore, $\hat{\sigma} \subset A \cap B$. Now let

$$\pi_k \colon \mathbb{R}^d \to \mathbb{R}^{d-1}, \quad (x_1, x_2, \ldots, x_d) \mapsto (x_1, \ldots, \hat{x}_k, \ldots, x_d)$$

be the projection omitting the k-th component. Then $\pi(\hat{\sigma})$ is compact, convex, and covered by

$$C = \{\pi_k(Q) \mid Q \in A \cap B\}.$$

We claim $C = \mathcal{K}(\pi(\hat{\sigma}))$. This follows from convexity of σ as follows: Fix $Q \in \mathcal{K}(C)$. To it corresponds a unique $d-1$-dimensional elementary cube P in $\mathbb{R}^d$ that is in $A \cap B$. In particular there are d-dimensional elementary cubes $Q_A \in A$ and $Q_B \in B$ that have P as their face. This means that there are points $x_A \in Q_A \cap \sigma$ and $x_B \in Q_B \cap \sigma$. By convexity of σ, the line segment between them lies in σ and intersects P at some point x. This point satisfies $\pi_k(x) \in Q$. Thus the claim follows.

Furthermore, since any $d-1$ cube of $K(C)$ is adjacent to a d-dimensional cube in each of A and B, we have that $2\,\#\,\mathcal{K}_{d-1}(C) \leq n$ where $\#\,\mathcal{K}_{d-1}(C)$ denotes the cardinality of $\mathcal{K}_{d-1}(C)$. Thus, by induction, C is acyclic, and since C and $A \cap B$ are homeomorphic, so is $A \cap B$. With the Mayer-Vietoris argument from above, this finishes the proof. $\square$

4.4.2 Internal and External Covers of Spheres

Let $X \subset S_2^{d-1}$ be a point cloud on the ℓ^2-unit sphere in $\mathbb{R}^d$. Any point $p \in S_2^{d-1}$ has an external (closed) ball

$$B(p, r) = \{x \in \mathbb{R}^d \mid \|x - p\|_2 \leq r\}$$

and an internal (closed) ball

$$B^S(p, r) = \{x \in S_2^{d-1} \mid \|x - p\|_2 \leq r\}.$$

The goal of this section is to show that the external and internal thickenings of X are equivalent, or more precisely, that we have

$$\bigcup_{p \in X} B^S(p, r) \simeq \bigcup_{p \in X} B(p, \theta(r)), \text{ where } \theta(r) = \sqrt{1 - \left(1 - \frac{r^2}{2}\right)^2}. \tag{4.3}$$

Fig. 4.1a shows that the rescaling is indeed necessary.

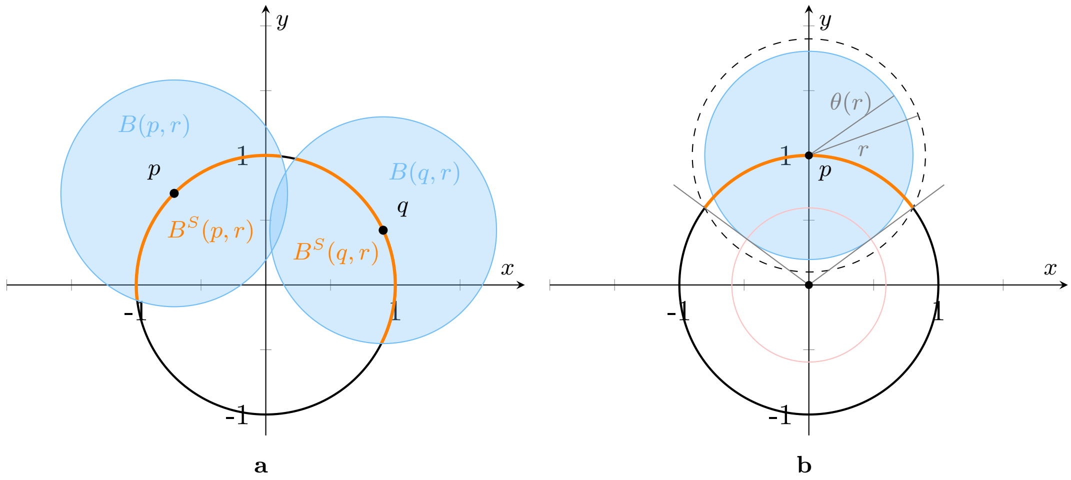


Figure 4.1: a) Internal and external balls at p and q with radius r. Since the external balls intersect while the internal balls do not, the two sets are not homotopy equivalent. b) Illustration for π and ψ. The map π projects $B(p, \theta(r))$, indicated in blue, radially onto S^1. The map ψ projects projects $B^S(p, r)$, indicated in orange, radially onto the pink sphere with radius $\frac{2-r^2}{2}$. The resulting arc is a deformation retract of $B(p, \theta(r))$ via a straight-line homotopy.

Proposition 4.4.5. *Let* $0 \leq r < 1$. *The maps*

$$\pi\colon \bigcup_{p \in X} B(p, \theta(r)) \to \bigcup_{p \in X} B^S(p, r), \quad x \mapsto \frac{x}{\|x\|_2} \tag{4.4}$$

and

$$\psi\colon \bigcup_{p \in X} B^S(p, r) \to \bigcup_{p \in X} B(p, \theta(r)), \quad y \mapsto \frac{2 - r^2}{2} y. \tag{4.5}$$

are homotopy equivalences and homotopy inverses to each other.

Proof. We first consider the setting of a single ball, i.e., $X = \{p\}$ with $p \in S_2^{d-1}$.

Claim: π is well-defined: Fix $x \in B(p, \theta(r))$. Let $\lambda = \langle x, p\rangle / \|x\|_2^2$ and note that $\langle x, p - \lambda x\rangle = 0$. This implies

$$\begin{aligned}
\|x - p\|_2^2 &= \left\|(1-\lambda)x + \lambda x - p\right\|_2^2 \\
&= \left\|(1-\lambda)x\right\|_2^2 - 2\langle (1-\lambda)x, \lambda x - p\rangle + \|\lambda x - p\|_2^2 \\
&= \left\|(1-\lambda)x\right\|_2^2 + \|\lambda x - p\|_2^2 \geq \|\lambda x - p\|_2^2 .
\end{aligned}$$

Using $\|p\|_2 = 1$ and again $\langle x, p - \lambda x\rangle = 0$, we furthermore have

$$\|\lambda x - p\|_2^2 = \|\lambda x\|_2^2 - 2\langle \lambda x, p\rangle + \|p\|_2^2 = \|\lambda x\|_2^2 - 2\langle \lambda x, p - \lambda x + \lambda x\rangle + 1 = 1 - \|\lambda x\|_2^2 .$$

Combining the last two calculations, we get

$$\|\lambda x\|_2^2 = 1 - \|\lambda x - p\|_2^2 \geq 1 - \|x - p\|_2^2 \geq 1 - \theta(r)^2 = \left(1 - \frac{r}{2}\right)^2 .$$

We estimate the distance of $\pi(x)$ to p using the preceding inequality

$$\begin{aligned} \left\|\pi(x) - p\right\|_2^2 = \left\|\pi(x)\right\|_2^2 - 2\langle \pi(x), p\rangle + \|p\|_2^2 &= 1 - 2\frac{\langle x, p\rangle}{\|x\|_2} + 1 \\ &= 2 - \lambda\|x\|_2 \leq 2 - 2\left(1 - \frac{r^2}{2}\right) = r^2 \end{aligned}$$

which implies $\pi(x) \in B^S(p, r)$.

Claim: ψ is well-defined. Fix $y \in B^S(p, r)$. Note that

$$\|y - p\|_2^2 = \|y\|_2^2 - 2\langle y, p\rangle + \|p\|_2^2 = 2 - 2\langle y, p\rangle.$$

We calculate

$$\begin{aligned} \left\|\psi(y) - p\right\|_2^2 &= \left\|\frac{2 - r^2}{2} y - p\right\|_2^2 = \left(\frac{(2 - r^2)}{2}\right)^2 \|y\|_2^2 - (2 - r^2)\langle y, p\rangle + \langle p, p\rangle \\ &= \frac{(2 - r^2)^2}{4} - (2 - r^2)\langle y, p\rangle + 1 = \frac{(2 - r^2)^2}{4} - (2 - r^2)\frac{(2 - \|y - p\|_2^2)}{2} + 1 \\ &\leq \frac{(2 - r)^2}{4} - (2 - r^2)\frac{2 - r^2}{2} + 1 = 1 + \frac{(2 - r^2)^2}{4} - \frac{(2 - r^2)^2}{2} = 1 - \frac{(2 - r^2)^2}{4} \\ &= \theta(r)^2. \end{aligned}$$

Claim: $\pi \circ \psi$ is the identity. It is clear by definition that $\pi \circ \psi = \mathrm{id}$.

Claim: $\psi \circ \pi \cong \mathrm{id}$. The straight-line homotopy $H\colon B(p, \theta(r)) \times [0, 1] \to B(p, \theta(r))$

$$H(x, t) = (1 - t)x + t\psi \circ \pi(x) = (1 - t)x + t\frac{2 - r^2}{2\|x\|_2} x$$

is well-defined since Euclidean balls are convex.

The general case: All constructions so far extend to the setting of multiple balls.

□

4.4.3 Equivalence of Internal and External Thickenings

To construct the cycling signature, we consider metric thickenings of segments γ in the unit tangent bundle $\mathrm{UT}(\mathbb{R}^d)$. These thickenings are with respect to a metric

$$d_C((p,v),(q,w)) = \max\{\|p-q\|_2\,, C\|v-w\|_2\}.$$

With an internal cover, we mean the sets $O_r(\gamma)^r$. This is comprised of *internal balls* of the form $\overline{B^I_C(x,r)} = \{y \in \mathrm{UT}(\mathbb{R}^d) \mid d_C(y,x) \le r\}$ and has the form

$$O_r(\gamma) = \bigcup_{x\in\gamma} \overline{B^I_C(x,r)}.$$

The external cover consists of balls $\overline{B^E_D(x,r)} = \{y \in \mathrm{T}(\mathbb{R}^d) \mid d_D(y,x) \le r\}$ and has the form

$$\bigcup_{x\in\gamma} \overline{B^E_D(x,r)}.$$

The goal of this section is to show that internal and external covers are equivalent. This will use the results from the previous appendix. Recall the definition of θ in Eq. (4.3).

Proposition 4.4.6. *Let $C > 0$, $r \in [0, C)$ and $D = r/(C\theta(r/C))$. Then for $\gamma \in \mathrm{UT}(\mathbb{R}^d)$ finite, the map*

$$\alpha\colon \bigcup_{x\in\gamma} \overline{B^E_D(x,r)} \to \bigcup_{x\in\gamma} \overline{B^I_C(x,r)}, \quad (p,v) \mapsto (p, v/\|v\|_2)$$

is a homotopy equivalence.

Proof. Recall the maps π and ψ from Eqs. (4.4) and (4.5). Using these, we have $\alpha((p,v)) = (p,\pi(v))$. Define furthermore

$$\beta\colon \bigcup_{x\in\gamma} \overline{B^E_C(x,r)} \to \bigcup_{x\in\gamma} \overline{B^I_D(x,r)}, \quad (p,v) \mapsto (p,\psi(v)).$$

We now claim that α and β are well-defined and homotopy inverse to each other. By definition of the d_C, we have

$$\overline{B^E_D((p,v),r)} = \overline{B(p,r)} \times \overline{B(v,r/D)} = \overline{B(p,r)} \times \overline{B(v,\theta(r/C))}$$

and similarly

$$\overline{B^I_C((p,v),r)} = \overline{B(p,r)} \times \overline{B^S(v,r/C)}.$$

Thus, by Proposition 4.4.5, both maps are well-defined. We furthermore have $\beta \circ \alpha(p,v) = (p, \pi\circ\psi(v)) = (p,v)$ and $\psi\circ\pi \cong \mathrm{id}$ where a homotopy is given by the straight line homotopy in the second component, i.e.

$$H((p,v)),t) = (p,(1-t)v + t\psi \circ \pi(v)).$$

$\square$

Corollary 4.4.7. *Let $C > 0$, $r \in [0, C)$ and $D = r/(C\theta(r/C))$. Then for $\gamma \in \mathrm{UT}(\mathbb{R}^d)$ finite,*

$$\bigcap_{x\in\gamma} \overline{B^E_D(x,r)} \neq \emptyset \iff \bigcap_{x\in\gamma} \overline{B^I_C(x,r)} \neq \emptyset$$

Proof. Fix $x \in \gamma$. It follows from the proposition and its proof that

$$y \in \overline{B^E_D(x,r)} \implies \alpha(y) \in \overline{B^I_C(x,r)}, \quad \text{and} \quad y \in \overline{B^I_C(x,r)} \implies \beta(y) \in \overline{B^E_D(x,r)}.$$

The equivalence now follows by mapping the intersection points using α and β. $\square$

4.4.4 Cubical Comparison Space Induces Comparison Space

Throughout this section, fix $d, k \in \mathbb{N}_{>0}$ and let $\pi_\infty \colon \mathbb{R}^d \setminus \{0\} \to S^1_\infty$, $x \mapsto x/\|x\|_\infty$. The goal of this appendix is to prove Proposition 4.1.4. The proof will make use of the following cover of the relevant cubes. For any cube $Q = Q_{1/k}(p)$, where $p \in 1/k\mathbb{Z}^d$ with $\|k\|_\infty = 1$, and any $i = 1, \dots, d$, we define

$$\mathrm{Dec}_i(Q) = \begin{cases} \{\{x \in Q \mid |x_i| \geq |x_j| \text{ for } j \neq i\}\} & \text{if } |p_i| = 1, \\ \emptyset & \text{else.} \end{cases}$$

and $\mathrm{Dec}(Q) = \bigcup_{i=1}^d \mathrm{Dec}_i(Q)$. For an example, see Fig. 4.2 where $\mathrm{Dec}(Q)$ is visualized for all cubes with radius $1/2$ centered on $1/2\mathbb{Z}^d$.

Lemma 4.4.8. *For any cube $Q = Q_{1/k}(p) \subset \mathbb{R}^d$, where $p \in 1/k\mathbb{Z}^d$ and $\|p\|_\infty = 1$, we have $Q \subset |\mathrm{Dec}(Q)|$.*

Proof. Note that any $x \in Q$ can be written as $x = p + v$ with $\|v\|_\infty = \frac{1}{2k}$. Thus, if $|p_i| = 1$, it is $|x_i| = |p_i + v_i| \geq 1 - \frac{1}{2k}$. Otherwise, $|p_i| \neq 1$ and we have $|x_i| = |p_i + v_i| \leq (k-1)/k + \frac{1}{2k} = 1 - \frac{1}{2k}$.

Thus, for any $x \in Q$, there is an i such that $|x_i| \geq |x_j|$ for $j \neq i$ and $|p_i| = 1$. For this index, we have $x \in |\mathrm{Dec}_i(Q)|$. $\square$

Definition 4.4.9. Let $H = \{x \in \mathbb{R}^d \mid x_i = a\}$ with $a \neq 0$ be a hyperplane. The *radial projection (from the origin)* to H is the map

$$p \colon \mathbb{R}^d \setminus \{x \in \mathbb{R}^d \mid x_i = 0\} \to H, \quad x \mapsto a\frac{x}{x_i}.$$

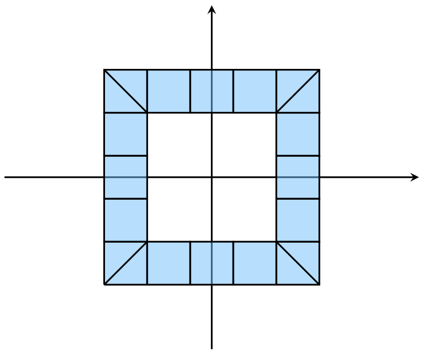


Figure 4.2: Illustration of Dec for a cubical cover of S^1_∞.

Lemma 4.4.10. *Let $H = \{x \in \mathbb{R}^d \mid x_i = a\}$ with $a \neq 0$ be a hyperplane and p the radial projection from the origin onto H. Let K be a convex set with $K \cap H = \emptyset$. Then $p(K)$ is convex.*

Proof. Fix $x, y \in \mathbb{R}^d$ and note that for $\lambda \in [0, 1]$

$$p(\lambda x + (1-\lambda)y) = \frac{\lambda x_i}{\lambda x_i + (1-\lambda)y_i} \frac{ax}{x_i} + \frac{(1-\lambda)y_i}{\lambda x_i + (1-\lambda)y_i} \frac{ay}{y_i} =: a(\lambda)p(x) + b(\lambda)p(y),$$

where we define $a(\lambda)$ and $b(\lambda)$. Note that a and b are well-defined, since either $x_i, y_i > 0$ or $x_i, y_i < 0$. Furthermore, $a(\lambda) + b(\lambda) = 1$ and $a(0) = b(1) = 0$, $a(1) = b(0) = 1$. The image of $\lambda \mapsto p(\lambda x + (1-\lambda)y)$ is thus the straight line between $p(x)$ and $p(y)$. From this observation, we see that $p(K)$ is convex. □

Lemma 4.4.11. *Let $Q = Q_{1/k}(p) \subset \mathbb{R}^d$ be a cube where $p \in 1/k\mathbb{Z}^d$ with $\|p\|_\infty = 1$. For any $i = 1, \ldots, d$, and any $R \in \mathrm{Dec}_i(Q)$, it is*

1. R_i is compact and convex,
2. $\pi_\infty(R) \subset \{x \in \mathbb{R}^d \mid x_i = p_i/|p_i|\}$, and
3. $\pi_\infty(R)$ is compact convex.

Proof. (1): R_i is compact since it is the intersection of a compact cube with a (closed) halfspace. Convexity follows from the observation that r_i is the intersection of the convex sets Q and $\{x \mid |x_i| \geq |x_j| \text{ for } j \neq i\}$.

(2): Let $x = \pi(y) \in \pi_\infty(R_i)$. We need to show that $x_i = p_i/|p_i|$. We have

$$x = \pi_\infty(y) = \frac{y}{\|y\|_\infty} = \frac{y}{|y_i|}.$$

Note that $y = p + v$ for a $v \in [-1/(2k), 1/(2k)]^d$. Furthermore, since $R \in \mathrm{Dec}_i(Q)$, we have $|p_i| = 1$. Therefore, $p_i + v_i$ and p_i have the same sign and thus $x_i = (p_i + v_i)/|p_i + v_i| = p_i/|p_i|$.

(3): By (2), the restriction of p_∞ to R is a radial projection onto a hyperplane. Thus, convexity is preserved and R_i is compact as the image of a compact set. □

Lemma 4.4.12. *Fix* $P \subset \{p \in 1/k\mathbb{Z}^d \mid \|p\|_\infty = 1\}$. *Then*

1. $\bigcap_{p\in P} Q_{1/k}(p) \neq \emptyset$ if and only if $\|p - q\|_\infty \leq \frac{1}{k}$ for all $p, q \in P$,
2. If $\bigcap_{p\in P} Q_{1/k}(p) \neq \emptyset$, then $\bigcap_{A\in\mathcal{A}} A \neq \emptyset$ for $\mathcal{A} = \bigcup_{p\in P} \mathrm{Dec}(Q_{1/k}(p))$,
3. For any $\mathcal{A} \subset \bigcup_{p\in P} \mathrm{Dec}(Q_{1/k}(p))$ it is

$$\bigcap_{A\in\mathcal{A}} A = \emptyset \implies \bigcap_{A\in\mathcal{A}} p_\infty(A) = \emptyset.$$

Proof. (1): This is a simple calculation.

(2): Set $a_i = \min\{p_i \mid p \in P\}$ and $b_i = \max\{p_i \mid p \in P\}$. By (1), either $a_i = b_i$ or $a_i = b_i - 1/k$. Define

$$x_i = \begin{cases} \frac{a_i+b_i}{2} & \text{if } a_i \neq b_i \text{ and } |b_i| \neq 1, \\ b_i - \frac{\mathrm{sign}(b_i)}{2k} & \text{else.} \end{cases}$$

We claim $x \in \bigcap_{A\in\mathcal{A}} A$. Fix $A \in \mathcal{A}$. Then $A \in \mathrm{Dec}(Q_{1/k}(p))$ for some $p \in P$. For all i, it is either $p_i = a_i$ or $p_i = b_i$. Using this, it is easy to check that $|x_i - p_i| \leq 1/(2k)$ and therefore $x \in Q_{1/k}(p)$. Furthermore, if $|p_i| = 1$, it is $|x_i| = 1 - \frac{1}{2k}$. If otherwise $|p_i| < 1$, it is $|x_i| \leq 1 - \frac{1}{2k}$. From this, we have that $|p_i| = 1$ implies that $|x_i| \geq |x_j|$ for $i \neq j$ and thus, $x \in A$ for any $A \in \mathrm{Dec}(Q_{1/k}(p))$.

(3): Suppose $\bigcap_{A\in\mathcal{A}} A = \emptyset$. By (1) and (2), there are $A, B \in \bigcap_{A\in\mathcal{A}} A$ with $A \in \mathrm{Dec}(Q_{1/k}(p))$ and $B \in \mathrm{Dec}(Q_{1/k}(q))$ such that $p, q \in P$ and $\|p - q\|_\infty > 1/k$.

Fix i such that $|p_i - q_i| \geq 2/k$. We will show

$$\max\{a_i \mid a \in \pi_\infty(A)\} < \min\{b_i \mid b \in \pi_\infty(B)\} \tag{4.6}$$

which implies that $\pi_\infty(A)$ and $\pi_\infty(B)$ can be separated by a hyperplane and therefore $\pi_\infty(A) \cap \pi_\infty(B) = \emptyset$.

For $x \in Q_{1/k}(p)$ and $y = \pi_\infty(x) = x/\|x\|_\infty$, elementary estimates yield

$$y_i \in \begin{cases} \left[\frac{2kp_i-1}{2k+1}, \frac{2kp_i+1}{2k-1}\right] & \text{if } p_i \geq 1/k, \\ \left[\frac{-1}{2k-1}, \frac{1}{2k+1}\right] & \text{if } p_i = 0, \\ \left[\frac{2kp_i+1}{2k-1}, \frac{2kp_i-1}{2k+1}\right] & \text{if } p_i \leq -1/k. \end{cases}$$

By relabelling p and q if necessary, we can assume $p_i + 2/k \leq q_i$. After substituting the relevant interval bounds into Eq. (4.6), we check that the inequality is satisfied for all relevant values of p_i and q_i: See here if $p_i, q_i \geq 1$ (or, by symmetry, if $p_i, q_i \leq -1$), here if $p_i = 0$ (or, by symmetry, $q_i = 0$). In case $p_i < 0$ and $0 < q_i$, the inequality can easily seen to be true by comparing signs. □

We are finally able to prove the relevant proposition.

Proposition 4.4.13. *For any $K \subset \mathcal{Q}^d_{r,k}$, where $r > 0$ and $k \in \mathbb{N}_{>0}$ we have a homotopy equivalence $p_\infty(|K|) \simeq |K|$.*

Proof. Let $S = \{(p, k) \mid Q_r(p) \times Q_{1/k}(q) \in K\}$. We let $\mathcal{C} = \{Q_r(p) \times A \mid (p, q) \in S,\ A \in \mathrm{Dec}(Q_{1/k}(q))\}$. By Lemma 4.4.8, each cube in K is covered by its corresponding elements in $\mathcal{C}$. Thus, the union of these cubes $|K|$ is covered by $\mathcal{C}$. This implies that $\mathcal{D} = \{\pi_\infty(X) \mid X \in \mathcal{C}\}$ is a cover of $\pi_\infty(|K|)$. By Lemma 4.4.11, the covers $\mathcal{C}$ and $\mathcal{D}$ consist of compact convex sets.

We now claim that these two covers have the same Nerve. For this, we need to show for any $X_i, \ldots, X_n \in \mathcal{C}$

$$\bigcap_i X_i \neq \emptyset \iff \bigcap_i p_\infty(X_i) \neq \emptyset$$

The implication '⇒' is obvious. The implication '⇐' follows from (3) in Lemma 4.4.12 together with the observation that the product of two cubes intersects if and only if the component cubes intersect.

We want to apply the Nerve theorem from [BKRR23], Theorem 3.11 to above covers. For this, we turn the covers into pointed covers by choosing any point $x_A \in A$ for $A \in \mathcal{C}$ and the point $p_\infty(x_A)$ for $p_\infty(A)$. We then have that p_∞ induces a morphism of pointed covers, each of which satisfies the assumptions of the Nerve theorem. This yields that p_∞ is a homotopy equivalence between the covered spaces. □

4.4.5 Covering Compact Convex Sets in the Unit Tangent Bundle

We define outer cubical covers with respect to grids that are slightly more general than $\mathbb{Z}^d$. For $A \in \mathrm{GL}_d(\mathbb{R})$, we set

$$\mathcal{Q}_A = \{AQ \mid Q \in \mathrm{Cub}(\mathbb{R}^d)\} \tag{4.7}$$

where $AQ = \{Av \mid v \in Q\}$ and note that all notions from Section 2.3 carry over to this generalization.

Definition 4.4.14. The *outer cubical cover of* $X \subset \mathbb{R}^d$ *with respect to* $\mathcal{Q}_A(\mathbb{R}^d)$ is the cubical complex

$$\mathcal{K}^A(X) = \downarrow \left\{ Q \in \mathcal{Q}_A(\mathbb{R}^d) \mid Q \cap X \neq \emptyset \right\}.$$

Proposition 4.4.15. *Let* $A \in \mathrm{GL}(d)$ *and suppose* σ *is convex and compact. Then* $K_A(\sigma)$ *is acyclic.*

Proof. Linear automorphisms preserve compactness, convexity and homology. By Proposition 4.4.1, the outer cubical cover $K(A^{-1}\sigma)$ is acyclic. The claim now follows from $|K_A(\sigma)| = A|K(A^{-1}\sigma)|$. □

The following is the corollary that is used in the construction of the acyclic carrier in Proposition 4.1.5. Recall the notation $\mathcal{Q}^d_{r,k}$ from Section 4.1.2.

Corollary 4.4.16. *Let* $\sigma \subset \mathrm{UT}_\infty(\mathbb{R}^d)$ *be compact and convex. For any* $r \in \mathbb{R}, k \in \mathbb{Z}$*, the set*

$$L = \downarrow\{Q \in \mathcal{Q}^d_{r,k} \mid Q \cap \sigma \neq \emptyset\}$$

is acyclic.

Proof. Let $v \in \mathbb{R}^{2d}$ be the vector with first d entries r followed by d entries $\frac{1}{k}$. Let furthermore $A \in \mathbb{R}^{2d \times 2d}$ be the diagonal matrix with v on the diagonal. Define $\tau = \{w - v/2 \mid w \in \sigma\}$. By Proposition 4.4.15, the outer cubical cover $K_A(\tau)$ is acyclic.

We now claim that $|L|$ and $|K_A(\tau)|$ differ only by a translation by $v/2$. For this, fix $Q \in K_A(\tau)$ and set $\tilde{Q} = \{w + v/2 \mid w \in Q\}$. Clearly, $\tilde{Q} \cap \sigma \neq \emptyset$. We thus only need to show that $\tilde{Q}$ is contained in $\mathcal{Q}^d_{r,k}$. It is easy to see that $\tilde{Q} = Q_r(p) \times Q_{1/k}(q)$ for some $p \in r\mathbb{Z}^d$ and $q \in 1/k\mathbb{Z}^d$, thus it is only left to show that $\|q\|_\infty = 1$. To see this, it is enough to show that whenever $\|q\|_\infty \neq 1$, we have $Q_{1/k}(q) \cap S^1_\infty = \emptyset$. Pictorially, the box $Q_{1/k}(q)$ is a box of side length $1/k$ centered at $q \in 1/k\mathbb{Z}^d$, thus we have $q - 1/(2k) \leq \|x\|_\infty \leq q + 1/(2k)$ for all $x \in Q_{1/k}(q)$. In particular, $\|x\|_\infty \neq 1$ if $q \in 1/k\mathbb{Z}^d$ with $\|q\|_\infty \neq 1$.

Since translation is a homeomorphism, it now follows that $|L|$ is acyclic. □

4.4.6 Thickenings of Cubical Complexes

Recall the notation $\mathcal{Q}_Q$ from Eq. (4.7).

Proposition 4.4.17. *Let* A *be a positive multiple of the identity matrix, i.e.* $A = rI_d$ *for some* $r > 0$*, and* $K \subset \mathcal{Q}^d_A(\mathbb{R}^d)$*. Set*

$$K^s = \{O_s(Q) \mid Q \in K\}$$

where $0 < s < r/2$ and the thickening is with respect to the Euclidean metric. Then $\bigcup_{Q \in K^s} Q \simeq |K|$.

Proof. Note that thickenings of two non-adjacent boxes in $\mathcal{Q}^d_A(\mathbb{R}^d)$ only intersect if $s \geq r/2$. Thus, for $s < r/2$, we have that $Q_0, \ldots, Q_k \in K$ have nonempty intersection if and only if $O_s(Q_0), \ldots, O_s(Q_k) \in K^s$ have nonempty intersection. Therefore, the nerves of the covers K^s of $\bigcup_{Q \in K^s} Q$ and K of $|K|$ are isomorphic. By the nerve theorem for closed convex sets ([BKRR23]), the spaces are therefore homotopy equivalent. $\square$

5 Time Series, Distance Matrices and Persistent Homology

An entry $D_{i,j} = d(x_i, x_j)$ in the distance matrix of a point cloud $X = \{x_0, \dots, x_n\}$ naturally corresponds to an edge of length $D_{i,j}$ between x_i and x_j in the Vietoris–Rips or Čech complex of X. In an ordered point cloud, like a time series, this edge can be completed to a cycle by following the order of the point cloud, i.e. if $i \leq j$, we have a cycle $[x_i, x_{i+1}] + [x_{i+1}, x_{i+2}] + \cdots + [x_{j-1}, x_j] + [x_j, x_i]$ (see Section 5.1 for a detailed example). Now suppose $\mathrm{DC}(X)$ is a filtered complex whose vertices are precisely pairs (i, j) with birth $D_{i,j}$. We can then assign to any such vertex, and thus to any 0-cycle of $\mathrm{DC}(X)$, a 1-cycle of the corresponding ordered point cloud.

Carrying out this construction yields a chain map $\psi\colon C(\mathrm{DC}(X)) \to C(\mathrm{VR}(X))$ of degree 1 with the property that the induced map $H_0(\psi)$ is an epimorphism. In the context of cycling signatures, the map ψ is useful for two reasons: Not only does it provides an alternative algorithm to compute cycling signatures, it also offers information complementary to the cycling signature. Constructions similar to ψ also work with the Čech complex as codomain and for the offset filtration of smooth curves, though in the latter case we only show that the cokernel of the resulting map is ephemeral.

The chapter is structured as follows: In Section 5.1, we present a detailed example to illustrate the connection between distance matrix and Vietoris–Rips complex. In Section 5.2 we define the complex $\mathrm{DC}(X)$ and the aformentioned chain map ψ, as well as a variant which has the Čech complex as its codomain, and study their properties. A result of interest is that $H_0(\psi)$ is an epimorphism, the consequences of which are explored in Section 5.3. In Section 5.4, we present analogous constructions for the setting of continuous curves and use an interleaving argument to show that the induced map of persistene modules has ephemeral cokernel.

5.1 Motivating Example

We start with an example that illustrates the relationship between persistent homology of a curve and sublevelset persistence of its distance matrix. Let $\Gamma = \{x_1, \dots, x_{41}\}$ be the time series in Fig. 5.1a; it traces a curve that comes close to itself after some time. Its distance matrix M is visualized as a heatmap in Fig. 5.1b.

The degree 1 persistent homology of Γ consists of one feature that captures the 'hole' in the right part of the time series. More precisely, it is $H_1(\mathrm{VR}(\Gamma)) \cong I(b, c)$ where b

and c are as indicated, a corresponding generator is drawn in green.

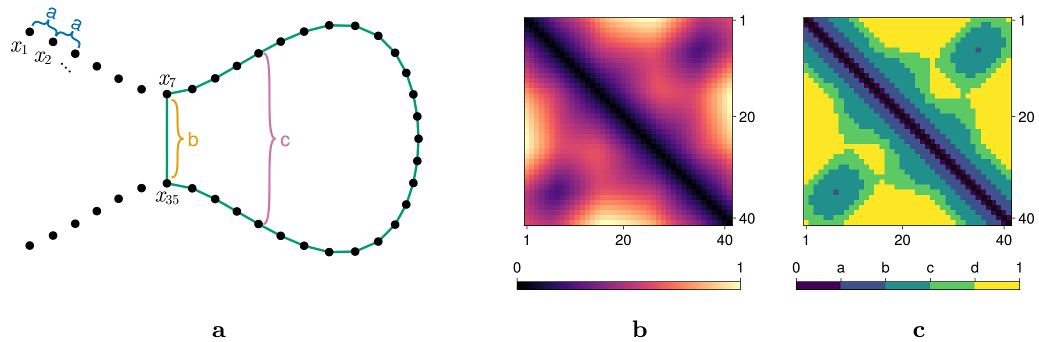


Figure 5.1: Time series with distance matrix. **a.** Time series. The distance between consecutive points is a. **b.** Heat map of the (normalized) distance matrix of the time series. The points in the distance matrix are ordered as in the time series. **c.** Some level sets of the heatmap in **b**. The colormap is as follows: The 'zeroth' level set is colored black and corresponds to distance 0. Distances in $(0, a]$ are purple, distances in $(a, b]$ are dark blue, etc.

We now turn to the distance matrix M of Γ. For this, let $G = (G_r)_r$ denote the sublevelset filtration of the heatmap in Fig. 5.1b. We make this precise later, for now, we let G_r be the union of the 'pixels' in Fig. 5.1b with value less or equal r. Its connected components (i.e. homology in degree zero) evolve roughly as follows (see also Fig. 5.1c):

1. For $0 \leq r < b$, G_r has one connected component around the diagonal.
2. For $b \leq r < d$, where $d > c$ is as indicated in Fig. 5.1c, connected components emerge at $(7, 35)$ and $(35, 7)$.
3. For $d \leq r$, all connected components have merged to a single component.

We relate this to $H_1(\mathrm{VR}(\Gamma))$: The first point is due to the fact that Γ is sampled from a curve, i.e. a discrete point set where consecutive points are close. The second point is where things get interesting: In the Vietoris–Rips filtration, $r = b$ is the birth index for the non-trivial 1-cycle (green in Fig. 5.1a). At the same time, the edge of length b corresponds precisely to the entries $(7, 35)$ and $(35, 7)$ where the new connected components in G appear. Furthermore, the location of this new connected component *identifies* the non-trivial 1-cycle in the Vietoris–Rips complex: Provided the curve is sampled densely enough, we can start with the edge $[x_{35}, x_7]$ and then follow the time series via $[x_7, x_8]$, $[x_8, x_9]$, $\ldots$, $[x_{34}, x_{35}]$ to x_{35}. Applying the same procedure to neighboring entries (e.g. $(7, 34)$), or generally, entries in the same connected component yields 1-cycles that are easily seen to be homologous to the previously constructed one. The same procedure also works for pixels in the diagonal component, this will produce cycles that are boudaries. Increasing the filtration parameter further, we see that the cycle specified by $(7, 35)$ dies at filtration value c in the Vietoris–Rips filtration, while its connected component in the distance matrix filtration merges with the diagonal at filtration value d.

In summary, we observed that pixels define 1-cycles with the property such that pixels of the distance matrix that are in the same connected component give cycles in the same homology class. However, cycles may in $\mathrm{VR}(\Gamma)$ can die before their corresponding connected components in the distance matrix.

In the remainder of this chapter, we formalize versions of this connection in the form of a degree 1 chain map and study its properties.

5.2 The Distance Complex of a Time Series

In this section we describe the first instance of the connection between persistent homology of a time series Γ and persistent homology of its distance matrix. We assign a filtration to a distance matrix using a construction that, in the context of image analysis using persistent homology is sometimes called the $\mathcal{V}$-construction [BGH+22, SPS+23, RWS11]. Throughout this section, we fix a metric space (X, d).

Definition 5.2.1. The *distance complex* $\mathrm{DC}(\Gamma)$ of a time series $\Gamma = \{x_0, \ldots, x_n\} \subset X$ is the filtered cubical complex where $\mathrm{DC}_r(\Gamma)$ is the full subcomplex of $\mathrm{Cub}(\mathbb{R}^2)$ spanned by the vertices $V_r = \{(i, j) \mid 0 \leq i \leq j \leq n \text{ with } d(x_i, x_j) \leq r\}$.

To relate the distance complex and the Vietoris–Rips complex of a time series, we need the following notion that captures how 'curve-like' a time series is.

Definition 5.2.2. A time series $\Gamma = \{x_i\}_{i=0}^n$ is *R-tight* if $R = \max_i d(x_i, x_{i-1})$.

Suppose $\Gamma = \{x_0, x_1, \ldots, x_n\}$ is an R-tight time series. A *curve chain* of Γ is a chain

$$c_\Gamma(i, j) = \sum_{k=i+1}^{j} [x_{k-1}, x_k].$$

By R-tightness, we have $c_\Gamma(i, j) \in C(\mathrm{VR}_r(\Gamma))$ for all $r \geq R$. Note that $\partial c_\Gamma(i, j) = [x_j] - [x_i]$.

In the following, for any $r \geq R$, we define a chain map of degree 1

$$\psi_r \colon C(\mathrm{DC}_r(\Gamma)) \to C(\mathrm{VR}_r(\Gamma)). \tag{5.1}$$

For any vertex $v = (i, j)$ in $\mathrm{DC}_r(\Gamma)$, we set

$$\psi_r(v) = c_\Gamma(i, j) - [x_i, x_j], \tag{5.2}$$

for any edge e in $\mathrm{DC}_r(\Gamma)$, we set

$$\psi_r(e) = \begin{cases} [x_i, x_j, x_{i+1}], & \text{if } e = [i, i+1] \times [j, j], \\ [x_i, x_j, x_{j+1}], & \text{if } e = [i, i] \times [j, j+1], \end{cases} \tag{5.3}$$

for any square $q = [i, i+1] \times [j, j+1] \in \mathrm{DC}_r(\Gamma)$, we set

$$\psi(q) = [x_i, x_{i+1}, x_j, x_{j+1}], \tag{5.4}$$

In these formulas, we make the implicit assumption that any degenerate simplex is identified with 0. We then extent ψ_r via linearity.

Lemma 5.2.3. *ψ is a well-defined degree 1 morphism of graded $[R, \infty)$-filtered modules.*

Proof. It follows straight from the definitions that ψ respects the grading. To show compatibility with the filtrations, we first show that ψ_r is well-defined for fixed $r \geq R$, i.e. that we have $\psi_r(C_k(\mathrm{DC}_r(\Gamma))) \subset C_{k+1}(\mathrm{VR}_r(\Gamma))$ for all $k \geq 0$. By linearity, it is enough to show $\psi_r(c) \in C(\mathrm{VR}_r(\Gamma))$ for all elementary chains c in $C(\mathrm{DC}_r(\Gamma))$.

Let $v = (i, j)$ be a vertex in $\mathrm{DC}_r(\Gamma)$. This implies $d(x_i, x_j) \leq r$ and thus $[x_i, x_j] \in \mathrm{VR}_r(\Gamma)$. As remarked earlier, by R-tightness, $c_\Gamma(j, i)$ is an elment of $C(\mathrm{VR}_r(\Gamma))$, thus $\psi_r(v) = c_\Gamma(i, j) - [x_i, x_j]$ is a sum of elements in $C(\mathrm{VR}_r(\Gamma))$.

By 5.3 and 5.4, if τ is an edge or square of $\mathrm{DC}_r(\Gamma)$, then $\psi_r(\tau)$ is a chain supported on a single simplex σ. The edges of σ are between points (x_i, x_j) such that (i, j) is a vertex of $\mathrm{DC}_r(\Gamma)$. Thus, their lengths are bounded by r and all edges of σ are in $\mathrm{VR}(\Gamma)$. Since Vietoris–Rips complexes are clique complexes, this implies $\sigma \in \mathrm{VR}_r(\Gamma)$ and thus $\psi_r(\tau) \in C(\mathrm{VR}_r(\Gamma))$.

Lastly, to show that $(\psi_r)_{r \geq R}$ is a map of filtered modules, we need to show that for $r < s$, the restriction of ψ_s to $C(\mathrm{DC}_r(\Gamma))$ is ψ_r. This is immediate from the definition, since the right hand sides of Eqs. (5.2) to (5.4) are independent of r. ☐

As an immediate consequence, we can omit the subscript of ψ_r as long as we tacitly assume $r \geq R$.

Lemma 5.2.4. *We have $\psi\partial = \partial\psi$ and ψ is a chain map.*

Proof. Any vertex $v = (i, j)$ in $\mathrm{DC}_r(\Gamma)$ trivially has $\psi(\partial(v)) = 0$. We calculate

$$\partial(\psi(v)) = \partial(c_\Gamma(i, j)) - \partial([x_i, x_j]) = ([x_j] - [x_i]) - ([x_j] - [x_i]) = 0,$$

thus, by linearity, $\psi(\partial c) = \partial\psi(c)$ for all 0-chains.

For 1-chains, we start with an edge $e = [i, i+1] \times [j, j]$ in $\mathrm{DG}_r(\Gamma)$. With $v = (i, j)$ and $w = (i+1, j)$, we get

$$\psi(\partial e) = \psi(w) - \psi(v) = c_\Gamma(i+1, j) - [x_{i+1}, x_j] - c_\Gamma(i, j) + [x_i, x_j]$$

Similarly, if $e = [i,i] \times [j, j+1]$ we get, with $u = (i, j+1)$,

$$\psi(\partial e) = \psi(u) - \psi(v) = c_\Gamma(i, j+1) - [x_i, x_{j+1}] - c_\Gamma(i,j) + [x_i, x_j].$$

Simplifying the resulting terms yields

$$\psi(\partial e) = \begin{cases} [x_j, x_{i+1}] - [x_i, x_{i+1}] + [x_i, x_j], & \text{if } e = [i, i+1] \times [j, j], \\ [x_j, x_{j+1}] - [x_i, x_{j+1}] + [x_i, x_j], & \text{if } e = [i, i] \times [j, j+1]. \end{cases}$$

Above expression is seen to be equal to

$$\partial(\psi(e)) = \begin{cases} \partial[x_i, x_j, x_{i+1}], & \text{if } e = [i, i+1] \times [j, j], \\ \partial[x_i, x_j, x_{j+1}], & \text{if } e = [i, i] \times [j, j+1], \end{cases}$$

which, by linearity, shows $\psi(\partial c) = \partial\psi(c)$ for all 1-chains.

For squares $q = [i, i+1] \times [j, j+1] \in \mathrm{DC}_r(\Gamma)$, we calculate

$$\begin{aligned} \psi(\partial q) &= \psi([i+1] \times [j, j+1] - [i] \times [j, j+1] - [i, i+1] \times [j+1] + [i, i+1] \times [j]) \\ &= [x_{i+1}, x_j, x_{j+1}] - [x_i, x_j, x_{j+1}] - [x_i, x_{j+1}, x_{i+1}] + [x_i, x_j, x_{i+1}] \\ &= [x_{i+1}, x_j, x_{j+1}] - [x_i, x_j, x_{j+1}] + [x_i, x_{i+1}, x_{j+1}] - [x_i, x_{i+1}, x_j] \end{aligned}$$

which is equal to

$$\begin{aligned} \partial(\psi(q)) &= \partial[x_i, x_{i+1}, x_j, x_{j+1}] \\ &= [x_{i+1}, x_j, x_{j+1}] - [x_i, x_j, x_{j+1}] + [x_i, x_{i+1}, x_{j+1}] - [x_i, x_{i+1}, x_j]. \end{aligned}$$

Since all other chain groups of $\mathrm{DC}(\Gamma)$ are trivial, ψ commutes with ∂. This, together with Lemma 5.2.3, implies that ψ is a chain map. □

We now relate the connected components of $\mathrm{DC}(\Gamma)$ with $H_1(\mathrm{VR}(\Gamma))$.

Theorem 5.2.5. *Let $\Gamma = \{x_i\}_{i=0}^n$ be an R-tight time series in a metric space (X, d) and $I = [R, \infty)$. The map ψ in Eq.* (5.1) *is a degree 1 morphism of I-filtered chain complexes. It induces an epimorphism $H_0(\psi)\colon H_0(\mathrm{DC}(\Gamma)) \to H_1(\mathrm{VR}(\Gamma))$ of I-indexed persistence modules.*

Proof. By Lemma 5.2.4, ψ is a degree 1 chain map of filtered simplicial complexes. To show surjectivity, let

$$c = \sum_{i=1}^m \alpha_i [x_{k_i}, x_{l_i}] \in C_1(\mathrm{VR}_r(\Gamma))$$

be a cycle. By changing the signs of the α_i, we can assume $k_i < l_i$ for all i. Set $v_i = (k_i, l_i)$ and $a = \sum_{i=1}^m \alpha_i v_i$. Since $[x_{k_i}, x_{l_i}] \in \mathrm{VR}_r(\Gamma)$, we have $d(x_{k_i}, x_{l_i}) \leq r$ and therefore $v_i \in \mathrm{DC}_r(\Gamma)$ which implies $a \in C(\mathrm{DC}_r(\Gamma))$. We calculate

$$\psi(a) = \sum_{i=1}^{m} \alpha_i \psi(v_i) = \sum_{i=1}^{m} \alpha_i (c_\Gamma(k_i, l_i) - [x_{k_i}, x_{l_i}]) = \sum_{i=1}^{m} \alpha_i c_\Gamma(k_i, l_i) - c = b - c.$$

where we set $b = \sum_{i=1}^{m} \alpha_i c_\Gamma(k_i, l_i)$. Note that $\partial b = \partial c - \partial\psi(a) = \partial c - \psi(\partial(a)) = 0$. Thus b is a cycle that is supported on the line graph with nodes $\{0, \ldots, n\}$. Since the line graph has no nontrivial 1-cycles, we must have $b = 0$. Thus $\psi(a) = c$ which shows that $H_0(\psi)$ is surjective. $\square$

Clearly, $H_0(\psi)$ is not injective. Indeed, the essential homology class corresponding to the diagonal of the distance matrix is always in the kernel of $H_0(\psi)$. Furthermore, even the map $\tilde{H}_0(\psi)$ in relative homology is not necessarily injective. This can be observed in the introductory example where $\tilde{H}_0(\psi)$ is a map from $\tilde{H}_0(\mathrm{DC}(\Gamma)) = \mathrm{k}_{[b,d)}$ to $\tilde{H}_1(\mathrm{VR}(\Gamma)) = \mathrm{k}_{[b,c)}$. Since $c < d$, this map fails to be injective for filtration values in $[c, d)$.

For $k \geq 2$, the maps $H_k(\psi)$ are trivial because the homology of the distance complex is trivial in dimension two and greater. Thus, besides $k = 0$, the only case of interest is $k = 1$. We now present two examples. The first demonstrate that $H_1(\psi)$ can be nontrivial, the second that, contrary to $H_0(\psi)$, it is not always surjective.

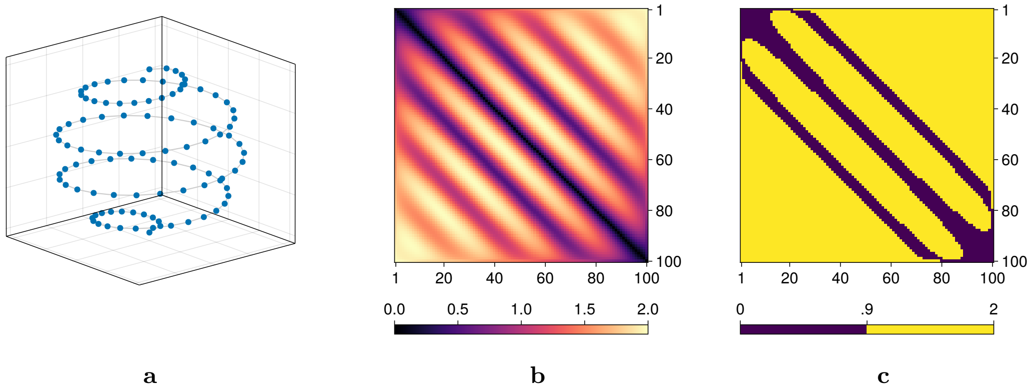


Figure 5.2: Spiral time series with distance matrix. **a.** Time series sampled from a curve that spirals upwards on the surface of the unit 2-sphere. **b.** Distance matrix as heatmap. Note the ridges parallel to the diagonal. **c.** Sublevel set of **b** with non-trivial H_1.

Example 5.2.6. Consider the time series γ in Fig. 5.2a. It lies on the surface of a unit 2-sphere and spirals from south pole to north pole. A persistence computation shows

$H_2(\mathrm{VR}(\gamma)) \cong I(0.899, 1.650)$[1].

The distance matrix of γ is visualized as a heatmap in Fig. 5.2b; the levelset in Fig. 5.2c reveals that the distance complex has nontrivial homology in degree 1. A computation shows that

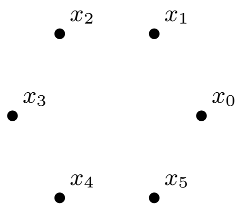
$$H_1(\mathrm{DC}(\gamma)) \cong I(0.899, 2.00) \oplus I(1.35, 2.00) \oplus I(1.67, 2.00) \oplus I(1.94, 2.00).$$

We furthermore compute a cohomology generator α corresponding to the bar in $H_2(\mathrm{VR}(\gamma))$ and homology generators $v_1, \ldots, v_4$ corresponding to the bars in $H_1(\mathrm{DC}(\gamma))$. We compute the time series to be R-tight with $R = 0.254$, thus ψ_r is defined for $r \geq R$. A computation (using coefficients in $\mathbb{F}_{43}$) yields

$$\alpha(\psi(v_1)) = -1, \quad \alpha(\psi(v_2)) = -1, \quad \alpha(\psi(v_3)) = 2, \quad \alpha(\psi(v_4)) = 0.$$

If $H_1(\psi)$ were trivial, all Kronecker pairings above would be zero. Therefore $H_1(\psi)$ is nontrivial.

Example 5.2.7. Let $\gamma = \{x_k = (\cos(2\pi k/6), \sin(2\pi k/6)) \mid k = 0, \ldots, 5\}$ be the vertices of a regular hexagon. The distance matrix of γ is



$$D = \begin{pmatrix} 0 & 1 & \sqrt{3} & 2 & \sqrt{3} & 1 \\ 1 & 0 & 1 & \sqrt{3} & 2 & \sqrt{3} \\ \sqrt{3} & 1 & 0 & 1 & \sqrt{3} & 2 \\ 2 & \sqrt{3} & 1 & 0 & 1 & \sqrt{3} \\ \sqrt{3} & 2 & \sqrt{3} & 1 & 0 & 1 \\ 1 & \sqrt{3} & 2 & \sqrt{3} & 1 & 0 \end{pmatrix}$$

It is not hard to see that

$$H_0(\mathrm{DC}(\gamma)) \cong I(0, \infty) \oplus I(0, 1)^{\oplus 5} \oplus I(1, 2) \quad \text{and} \quad H_1(\mathrm{DC}(\gamma)) \cong 0.$$

The map ψ_r is defined for $r \geq \sqrt{3}$. Furthermore, $H_1(\psi)$ is trivial since its domain is $H_1(\mathrm{DC}(\gamma)) = 0$. Since $\mathrm{VR}_r(\gamma) \cong S^2$ for $r \in [\sqrt{3}, 2)$ by [AA17], Theorem 7.6, the map $H_1(\psi)$ fails to be surjective. Furthermore, this is another example where the kernel of $H_0(\psi)$ contains more than the essential class of $H_0(\mathrm{DC}(\gamma))$.

The construction of ψ can be modified to have the Čech filtration as its codomain.

Definition 5.2.8. The *Čech distance complex* $\check{\mathrm{C}}\mathrm{DC}(\Gamma)$ of a time series Γ is the cubical complex

$$\check{\mathrm{C}}\mathrm{DC}(\Gamma) = \{[i, k] \times [j, l] \in \mathrm{DC}(\Gamma) \mid \{x_i, x_j, x_k, x_l\} \in \check{\mathrm{C}}(\Gamma)\}.$$

[1]In this example, all computed values are rounded to 3 decimals

We have the following version of Theorem 5.2.5.

Theorem 5.2.9. *Let* $\Gamma = \{x_i\}_{i=0}^n$ *be an* R*-tight time series and* $I = [R, \infty)$*. The restriction* $\tilde{\psi} = \psi_{|\check{\mathrm{C}}\mathrm{DC}(\Gamma)}$ *of the map from Eq.* (5.1) *is a degree 1 morphism* $C(\check{\mathrm{C}}\mathrm{DC}(\Gamma)) \to C(\check{\mathrm{C}}(\Gamma))$*. It induces an epimorphism* $H_0(\psi)\colon H_0(\check{\mathrm{C}}\mathrm{DC}(\Gamma)) \to H_1(\check{\mathrm{C}}(\Gamma))$ *of* I*-indexed persistence modules.*

Proof. It follows straight from the definition that $\operatorname{im} \psi_{|\check{\mathrm{C}}\mathrm{DC}(\Gamma)} \subset \check{\mathrm{C}}(\Gamma)$, thus $\tilde{\phi}$ is well-defined. Since $\check{\mathrm{C}}\mathrm{DC}^{(0)}(\Gamma) = \mathrm{DC}^{(0)}(\Gamma)$ as well as $\check{\mathrm{C}}^{(1)}(\Gamma) = \mathrm{VR}^{(1)}(\Gamma)$, the construction of preimages in the proof of Theorem 5.2.5 carries over to this setting. □

5.3 Application to Cycling Signature Computations

The relationship between degree zero homology of a distance matrix and degree one homology of a time series can be used to compute (most of) the cycling signature of a time series segment.

Distance Matrix Algorithm Recall that the cycling signature of a segment γ is computed as the image of the map $H_1(\phi)\colon H_1(\check{\mathrm{C}}(\gamma)) \to H_1(Y)$, which arises from discretizing the map in the definition of the cycling signature (see Definition 3.3.1 and the last paragraph of Section 4.2). Given an R-tight segment γ, the map $\tilde{\psi}$ from Theorem 5.2.9, which relates the Čech distance complex with the Čech complex of the curve, is an epimorphism of $[R, \infty)$-indexed persistence modules. For these filtration values, computing $\operatorname{im} H_1(\phi)$ is therfore equivalent to computing the image of

$$H_0(\check{\mathrm{C}}\mathrm{DC}(\gamma)) \xrightarrow{\tilde{\psi}} H_1(\check{\mathrm{C}}(\gamma)) \xrightarrow{\phi} H_1(Y).$$

While the information on $\mathrm{Cyc}_r(\gamma)$ for $r \in [0, R)$ is lost with this approach, this loss is minor if R is small enough. Additionally, the cycles in the image of $\tilde{\psi}$ contain additional information on the dynamics (see Section 6.4).

Implementation Using the notation from Section 4.3, the distance matrix algorithm amounts to applying the standard reduction algorithm to the matrix $\tilde{M}$ with entries $\tilde{M}_{i,j} = [\alpha_i(\phi(\tilde{\psi}(w_j)))]$ where the w_i are a basis for the immortal part of $H_0(\check{\mathrm{C}}\mathrm{DC}(\gamma))$. In our implementation, we employ the heuristic of using Vietoris–Rips filtrations, meaning that we compute a filtration compatible generating set of $H_1(\mathrm{VR}(\Gamma))^\infty$ using ψ and then proceed as described in Section 4.3.2.

Runtime Considerations From a runtime perspective, the distance matrix algorithm has the advantage that no reduction of a degree 1 coboundary matrix is necessary. Given a time series with n points, this reduces the worst-case complexity of the persistence algorithm, which is cubic in n [Mor], to the worst-case complexity of finding connected components in a graph with $n(n+1)/2$ nodes, which is quadratic in n. The direct method has the advantage that $\phi\colon H_1(\check{C}(\gamma)) \to H_1(Y)$ is evaluted less often, since a basis for $H_1(\check{C}(\gamma))$ contains at most as many elements as a generating set.

Using our implementation, we observe that the distance matrix algorithm is faster than the direct method for lower dimensional systems and shorter segments. For higher-dimensional spaces, the situation is less clear and the direct method sometimes outperforms the distance matrix algorithm even though, especially for longer segments, the computation of $H_0(\mathrm{DC}(\Gamma))$ is significantly faster than the computation of $H_1(\mathrm{VR}(\Gamma))$. Some of the overhead can potentially be mitigated since the current implementation computes $\phi(\psi(w))$ for all w in a basis of $H_0(\mathrm{DC}(\Gamma))^\infty$. An improvement would be to stop the computation, once the computed image spans the homological comparison space, since the remaining elements do not influence the cycling signature.

5.4 Distance Landscape of a Curve

In this section, we consider (smooth) curves $\gamma\colon [0,1] \to \mathbb{R}^d$ in Euclidean space. We will slightly abuse notation and write γ for $\operatorname{im}\gamma$. As before, we relate the topology of γ to the topology of its ‘distance matrix’. The latter is made precise in the following definition.

Definition 5.4.1. The *distance profile* of a curve $\gamma\colon [0,1] \to \mathbb{R}^d$ is the function

$$f\colon \mathcal{T} \to \mathbb{R}, \quad (s,t) \mapsto \frac{1}{2}\big\|\gamma(s) - \gamma(t)\big\|_2$$

where $\mathcal{T} = \{(s,t) \in [0,1]^2 \mid s \leq t\}$. Its subslevel sets define a filtered topological space $\mathrm{DL}_r(\gamma) = f^{-1}((-\infty, r])$ which we call the *distance landscape* of f.

The chain map relating the distance landscape with the offset filtration can be expressed conveniently using cubical singular homology. We recall some basic definitions and refer to [Mas80] for a comprehensive treatment.

Let I denote the closed interval $[0,1]$. A singular n-cube in a topological space X is a continuous map $\sigma\colon I^n \to X$. It is degenerate if it does not depend on one of its coordinates, i.e. for all $x_1, \dots, x_n$, the map $t \mapsto \sigma(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n)$ is constant. Otherwise, it is nondegenerate. The n-th singular cubical chain group of $C_n^{Cub}(X)$ is the quotient of the free abelian group generated by all singular n-cubes in X by the subgroup generated by all degenerate singular n-cubes in X. We note that this normalization by degenerate cubes is necessary, since otherwise even a single point would have nontrivial homology in all degrees.

To define the boundary operator, for any singular k-cube $\sigma\colon I^k \to X$, we define its front i-face $A_i\sigma\colon I^{n-1} \to X$ and its back i-face $B_i\sigma\colon I^{n-1} \to X$ via

$$A_i\sigma(x_1, \dots, x_{n-1}) = \sigma(x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_{n-1}), \text{ and}$$
$$B_i\sigma(x_1, \dots, x_{n-1}) = \sigma(x_1, \dots, x_{i-1}, 1, x_{i+1}, \dots, x_{n-1})$$

The boundary operator $\partial_n\colon C_n^{Cub}(X) \to C_{n-1}^{Cub}(X)$ is induced by the linear extension of

$$\partial\sigma = \sum_{i=1}^{n} (-1)^i (A_i\sigma - B_i\sigma).$$

We define a chain map $\theta\colon C_k^{Cub}(\mathrm{DL}(\gamma)) \to C_{k+1}^{Cub}(O(\Gamma))$. Let $\sigma\colon [0,1]^k \to \mathrm{DL}(\gamma)$ be a singular k-cube in $C_k(\mathrm{DL}(\gamma))$. Its birth is at filtration value $R = \max_{x\in[0,1]^k} f \circ \sigma$. Writing $\sigma(x) = (\sigma_1(x), \sigma_2(x))$, we define singular $k+1$-cubes

- $\rho_\sigma\colon [0,1]^k \times [0,1] \to O_r(\Gamma)$, $(x,\lambda) \mapsto \gamma((1-\lambda)\sigma_1(x) + \lambda\sigma_2(x))$, and
- $\eta_\sigma\colon [0,1]^k \times [0,1] \to O_r(\Gamma)$, $(x,\lambda) \mapsto (1-\lambda)\gamma(\sigma_1(x)) + \lambda\gamma(\sigma_2(x))$.

We now set $\theta_r(\sigma) = \rho_\sigma - \eta_\sigma$ for all $r \geq R$ and extend the map by linearity. Note that ρ_σ is a boundary for $\dim \sigma > 0$.

Proposition 5.4.2. *The map θ defined above is a chain map of filtered chain complexes.*

Proof. We first ensure that θ is filtration compatible. Let σ be a singular n-cube in $C(\mathrm{DL}_r(\Gamma))$ and note that this implies

$$r \geq \max_{x\in[0,1]^n} f(\sigma(x)) = \max_{x\in[0,1]^n} \frac{1}{2} \left\| \gamma(\sigma_1(x)) - \gamma(\sigma_2(x)) \right\|.$$

Therefore the distance between any pair $\gamma(\sigma_1(x))$ and $\gamma(\sigma_2(x))$ is at most $2r$. Because of this, the line segment connecting these points is contained in the r-thickening of Γ and we have $\eta_\sigma \in C(O_r(\Gamma))$. Since, trivially, $\rho_\sigma \in C(O_r(\Gamma))$, we get $\theta(\sigma) \in C(O_r(\Gamma))$.

We now show that θ commutes with the boundary operator. Again, fix a singular n-cube $\sigma \in C(X)$. For $\theta\partial$, we calculate

$$\theta(\partial\sigma) = \theta\left(\sum_{k=1}^{n} (-1)^k (A_k\sigma - B_k\sigma)\right) = \sum_{k=1}^{n} (-1)^k \left(\rho_{A_k\sigma} - \eta_{A_k\sigma} - \rho_{B_k\sigma} + \eta_{B_k\sigma}\right).$$

The term $\partial\theta(\sigma) = \partial(\rho_\sigma - \eta_\sigma)$ is slightly more complicated. We have

$$\partial\rho_\sigma - \partial\eta_\sigma = \sum_{k=1}^{n+1} (-1)^k \left(A_k\rho_\sigma - B_k\rho_\sigma\right) - \sum_{k=1}^{n+1} (-1)^k \left(A_k\eta_\sigma - B_k\eta_\sigma\right).$$

In both sums, the last term simplifies to $\gamma(\sigma_1(x)) - \gamma(\sigma_2(x))$ while for $1 \leq k \leq n$ we have $A_k\eta_\sigma - B_k\eta_\sigma = \eta_{A_k\sigma} - \eta_{B_k\sigma}$ and $A_k\rho_\sigma - B_k\rho_\sigma = \rho_{A_k\sigma} - \rho_{B_k\sigma}$. Collecting all terms, we therefore get

$$\partial\theta_\sigma = \partial\rho_\sigma - \partial(\eta_\sigma) = \sum_{k=1}^{n}(-1)^k \left(\rho_{A_k\sigma} - \rho_{B_k\sigma} - \eta_{A_k\sigma} + \eta_{B_k\sigma}\right).$$

This completes the proof that $\partial\theta = \theta\partial$. □

We now relate the distance landscape of a curve to the distance complex of a sample. A *sample* of γ is a time series $\Gamma = \gamma(\bar{t}) = \{\gamma(t_i) \mid i = 0, \dots, N\}$ with sampling times $\bar{t}$ where $0 \leq t_0 \leq t_1 \cdots \leq t_N$. Its *mesh size* is $\Delta\bar{t} := \max_i |t_i - t_{i-1}|$.

We construct a geometric realization of $\check{\mathrm{C}}\mathrm{DC}(\Gamma)$ that will be useful in the following. For this, let $p_{\bar{t}} \colon \mathbb{R}^2 \to \mathbb{R}^2$ denote the piecewise linear homeomorphism with $p_{\bar{t}}(i,j) = (t_i, t_j)$ for all $0 \leq i, j \leq N$. The Čech distance landscape $\check{\mathrm{C}}\mathrm{DL}(\Gamma)$ is the union

$$\check{\mathrm{C}}\mathrm{DL}_r(\Gamma) = \bigcup \left\{ p_{\bar{t}}(Q) \mid Q \in \check{\mathrm{C}}\mathrm{DC}_r(\Gamma) \right\}.$$

Lemma 5.4.3. $\check{\mathrm{C}}\mathrm{DL}_r(\Gamma)$ *is homeomorphic to* $|\check{\mathrm{C}}\mathrm{DC}_r(\Gamma)|$.

Proof. A homeomorphism is given by the appropriate restriction of $p_{\bar{t}}$. □

Lemma 5.4.4. *Let* $\gamma \colon [0,1] \to \mathbb{R}^d$ *be a curve. Then*

$$d_I\left(\check{\mathrm{C}}\mathrm{DL}(\gamma(\bar{t})), \mathrm{DL}(\gamma)\right) \to 0, \text{ as } \Delta\bar{t} \to 0.$$

Proof. Fix $\epsilon > 0$. By uniform continuity of the distance profile, there is $\delta > 0$ such that $|f(x) - f(y)| < 2\epsilon/3$ for all $x, y \in \mathcal{T}$ with $\|x - y\|_2 \leq \sqrt{2}\delta$. Assume $\Delta\bar{t} \leq \delta$ and set $\Gamma = \gamma(\bar{t})$.

Claim: $\check{\mathrm{C}}\mathrm{DL}_r(\Gamma) \subset \mathrm{DL}_{r+\epsilon}(\gamma)$. For any $x \in \check{\mathrm{C}}\mathrm{DL}_r(\Gamma)$, there is a cube $Q \in \mathrm{DC}_r(\Gamma)$ with $x \in p_{\bar{t}}(Q)$. Let v be any vertex of that cube. Since $\Delta\bar{t} \leq \delta$, we have $|x_1 - v_1| \leq \delta$, $|x_2 - v_2| \leq \delta$ and therefore $\|x - v\| \leq \sqrt{2}\delta$. Now

$$f(x) \leq |f(x) - f(v) + f(v)| \leq |f(x) - f(v)| + |f(v)| \leq 2\epsilon/3 + r \leq r + \epsilon$$

where the last step uses uniform continuity and that $v \in p_{\bar{t}}(Q) \in \check{\mathrm{C}}\mathrm{DC}_r(\Gamma)$. This shows $x \in \mathrm{DL}_{r+\epsilon}(\gamma)$.

Claim: $\mathrm{DL}_r(\gamma) \subset \check{\mathrm{C}}\mathrm{DL}_{r+\epsilon}(\Gamma)$. For any $x \in \mathrm{DL}_r(\gamma)$, there are i, j such that $t_i \leq x_1 \leq t_{i+1}$ and $t_j \leq x_2 \leq t_{j+1}$. Now, $Q = [t_i, t_{i+1}] \times [t_j, t_{j+1}]$ is an elementary cube (not necessarily of dimension 2), such that any vertex v of Q satisfies $\|x - v\| \leq \sqrt{2}\delta$. By uniform continuity, we therefore have

$$f(v) \leq |f(v) - f(x) + f(x)| < |f(v) - f(y)| + |f(y)| \leq 2\epsilon/3 + r \leq r + \epsilon. \tag{5.5}$$

Let $p_1 = \gamma(t_i)$, $p_2 = \gamma(t_{i+1})$, $p_3 = \gamma(t_j)$, and $p_4 = \gamma(t_{j+1})$. To show $p_{\bar{t}}(Q) \in \mathrm{DC}_{t+\epsilon}(\Gamma)$, we need to show $\bigcap_{k=1}^4 \overline{B(p_k, r+\epsilon)} \neq \emptyset$. Let $z = (p_1 + p_2 + p_3 + p_4)/4$. We have

$$\begin{aligned}
\|z - p_1\| = \|z - \gamma(t_i)\| &\leq \frac{1}{4}\sum_{k=1}^{4} \|p_k - \gamma(t_i)\| \\
&= \frac{1}{4}\left(\|\gamma(t_{i+1}) - \gamma(t_i)\| + \|\gamma(t_j) - \gamma(t_i)\| + \|\gamma(t_{j+i}) - \gamma(t_i)\|\right) \\
&= \frac{1}{2}\left(f(t_{i+1}, t_i) + f(t_j, t_i) + f(t_{j+1}, t_i)\right) \leq \frac{1}{2}\left(\frac{2\epsilon}{3} + (r + \frac{2\epsilon}{3}) + (r + \frac{2\epsilon}{3})\right) = r + \epsilon.
\end{aligned}$$

The last step uses uniform continuity of f for the first term and Eq. (5.5) for the other two. Similar arguments show $\|z - p_k\| \leq r + \epsilon$ for $k = 2, 3, 4$. We thus have $p_{\bar{t}}(Q) \in \check{\mathrm{D}}\mathrm{C}_{r+\epsilon}(\Gamma)$ and thus $x \in Q \in \check{\mathrm{D}}\mathrm{L}_{r+\epsilon}(\Gamma)$. $\square$

Lemma 5.4.5. *Let $\gamma\colon [0,1] \to \mathbb{R}^d$ be a curve. For any $\epsilon > 0$ and $R > 0$, there is a $\delta > 0$, such that for all samplings Γ of γ with sampling times $\bar{t}$ satisfying $\Delta\bar{t} \leq \delta$, there is a commutative diagram of $[R, \infty)$-indexed persistence modules*

$$\begin{tikzcd}
H_0(\check{\mathrm{D}}\mathrm{C}_\bullet(\Gamma)) \arrow[r, "\psi"] \arrow[d, "\alpha"] & H_1(\check{\mathrm{C}}_\bullet(\Gamma)) \arrow[d, "\beta"] \\
H_0(\mathrm{DL}_{\bullet+\epsilon}(\gamma)) \arrow[r, "\rho"] & H_1(O_{\bullet+\epsilon}(\gamma))
\end{tikzcd}$$

where α and β are ϵ-interleaving morphisms.

Proof. We construct a diagram

$$\begin{tikzcd}
C(\check{\mathrm{D}}\mathrm{C}_\bullet(\Gamma)) \arrow[r, "\psi"] \arrow[d, "q"] \arrow[ddr, "\alpha"] & C(\check{\mathrm{C}}_\bullet(\Gamma)) \arrow[r, "s"] \arrow[ddr, "\beta"] & C(O_\bullet(\Gamma)) \arrow[d, "i"] \\
C(|\check{\mathrm{D}}\mathrm{C}_\bullet(\Gamma)|) \arrow[d, "p_{\bar{t}}"] & & C(O_{\bullet+\epsilon}(\gamma)) \arrow[d, "f"] \\
C^{Cub}(\check{\mathrm{D}}\mathrm{L}_\bullet(\Gamma)) \arrow[r, "a"] & C^{Cub}(\mathrm{DL}_{\bullet+\epsilon}(\gamma)) \arrow[r, "\theta"] & C^{Cub}(O_{\bullet+\epsilon}(\gamma))
\end{tikzcd}$$

and show that it commutes for 0- and 1-chains in $\check{\mathrm{D}}\mathrm{C}_\bullet(\Gamma)$.

By Lemma 5.4.4, there is $\delta_1 > 0$ such that all samplings Γ of γ with sampling times $\bar{t}$ such that $\Delta\bar{t} \leq \delta_1$, satisfy $d_I\left(\check{\mathrm{C}}\mathrm{DL}(\gamma(\bar{t})), \mathrm{DL}(\gamma)\right) < \epsilon$. By compactness of γ, there is a $\delta_2 > 0$ such that $\bar{t} \leq \delta_2$ implies $d_I(O(\Gamma), O(\gamma)) < \epsilon$. By uniform continuity, there is δ_3 such that $\Delta\bar{t} \leq \delta_3$ implies R'-tightness for $R' \leq R$. Fix furthermore $\delta_4 > 0$ such that $\left\|\gamma'\right\|_\infty \delta_4 < \epsilon$ and $\left\|\gamma''\right\|_\infty \delta_4^2 < \epsilon$.

Set $\delta = \min \delta_i$ and fix a sample Γ of γ with sampling times $\bar{t}$ satisfying $\Delta\bar{t} < \delta$. By the choice of δ, we can find ϵ-interleaving morphisms a and b as in the diagram above.

We describe the remaining maps:

- q maps a k-dimensional elementary cube Q to the singular cube $I^k \to Q$ where the i-th copy of I is mapped linearly and increasing onto the i-th non-degenerate interval $[a, b]$. In the following, we will slightly abuse notation and identify an elementary cube Q with the singular cube $q(Q)$.
- $p_{\bar{t}}$ is given by the homeomorphism in Lemma 5.4.3.
- s is given by a composition $C(\check{\mathrm{C}}_\bullet(\Gamma)) \to C(|\check{\mathrm{C}}_\bullet(\Gamma)|) \to C(O_\bullet(\Gamma))$. The first map is the equivalence of simplicial and singular homology (see [Mun84, Theorem 34.3]). Since this isomorphism passes through ordered simplicial homology, its description is somewhat convoluted: Let $\sigma = [X_0, \ldots, X_n]$ be a simplex in $\check{\mathrm{C}}(\Gamma)$ such that $X_k = \{x_{i_k}\}$ and $i_0 \leq i_1 \leq \cdots \leq i_n$. Then σ is mapped to the singular simplex $(\lambda_0, \ldots, \lambda_n) \mapsto \sum_i \lambda_i |X_i|$. The first chain map is then the linear extension of this map. Note that it is important that the vertices of σ are ordered, we choose the total order on Γ induced by the time series. The second map is induced by the homotopy equivalence Theorem 2.1.9 and maps a singular simplex $(\lambda_0, \ldots, \lambda_n) \mapsto \sum_k \lambda_i |X_k|$ to the singular simplex $(\lambda_0, \ldots, \lambda_n) \mapsto \sum_k \lambda_k x_{i_k}$, where, again the order is important. In the following, we abuse notation and identify the 0-simplex $\{x_k\}$ with x_k as well as the simplicial simplex $\sigma = [x_0, x_1, \ldots, x_k]$ with the singular chain $s(\sigma)$. This allows us express s succinctly as $s([x_{i_0}, \ldots, x_{i_k}]) = [x_{i_0}, \ldots, x_{i_k}]$ if $i_1 \leq \cdots \leq i_k$.
- i is the inclusion map.
- f is a chain equivalence (see [EM53, Section 8]) between simplicial singular and cubical singular homology where a singular k-simplex is mapped to

$$f\sigma(\mu_1, \ldots, \mu_k) = \sigma(\lambda_0, \ldots, \lambda_k)$$

where

$$\lambda_0 = 1 - \mu_1, \quad \lambda_i = \mu_1 \ldots \mu_i(1 - \mu_{i+1}) \text{ for } 1 \leq i < k, \quad \lambda_k = \mu_1 \ldots \mu_k.$$

We define α and β via composition as indicated in the diagram. Since these are compositions of an ϵ-interleaving morphisms with chain equivalences, the maps α and β are again ϵ-interleaving morphisms (the inverse interleaving morphisms can be constructed by the corresponding composition of inverses).

We now construct a chain homotopy $H\colon C(\check{\mathrm{C}}\mathrm{DC}_\bullet(\Gamma)) \to C^{Cub}(O_{\bullet+\epsilon}(\gamma))$ in degrees 0 and 1 between $\theta \circ \alpha$ and $\beta \circ \psi$. Note that this is a map of degree 2. For a vertex $[i] \times [j] \in \check{\mathrm{C}}\mathrm{DC}_r(\Gamma)$, we define the singular 2-cube

$$\tau_{[i]\times[j]}(\mu_1, \mu_2) = \begin{cases} (1-\mu_1)\rho_{[i]\times[j-1]}(\mu_2) + \mu_1\rho_{[i]\times[j]}(\mu_2) & \text{if } i \neq j \\ 0 & \text{else.} \end{cases}$$

We show $\tau_v \in C^{Cub}(O_{r+\epsilon}(\gamma))$ for all vertices in $\check{\mathrm{C}}\mathrm{DC}_r^{(1)}(\Gamma)$. It is enough to consider vertices $[i] \times [j]$ with $i \neq j$. We have

$$\begin{aligned}
&\max_{\mu\in[0,1]^2} \left\| \tau_{[i]\times[j]}(\mu_1, \mu_2) - \rho_{[i]\times[j]}(\mu_2) \right\|_\infty \\
&= \max_{\mu\in[0,1]^2} \left\| (1-\mu_1)\rho_{[i]\times[j-1]}(\mu_2) + \mu_1\rho_{[i]\times[j]}(\mu_2) - \rho_{[i]\times[j-1,j]}(\mu_2) \right\|_\infty \\
&= \max_{\mu\in[0,1]^2} \mu_1 \left\| \rho_{[i]\times[j]}(\mu_2) - \rho_{[i]\times[j-1]}(\mu_2) \right\|_\infty \\
&= \max_{\mu\in[0,1]} \left\| \gamma((1-\mu)x_i + \mu x_j) - \gamma((1-\mu)x_i + \mu x_{j+1}) \right\|_\infty \\
&\leq (t_{j+1} - t_j)\left\|\gamma'\right\|_\infty \leq \delta\left\|\gamma'\right\|_\infty \leq \epsilon,
\end{aligned}$$

where the inequality at the start of the last line is obtained by applying the intermediate value theorem to each component of γ and observing that the arguments are at most $(t_{j+1} - t_j)$ apart. Since $\rho_{[i]\times[j]}$ is supported in $O_r(\gamma)$, this shows that τ_v is supported in $O_{r+\epsilon}(\gamma)$.

We set $H([i] \times [j]) = \sum_{k=i+1}^{j} \tau_{[i]\times[k]}$ for all vertices and extend the map by linearity. To show $\theta \circ \alpha - \beta \circ \psi = \partial H + H\partial$, we fix a vertex $v = [i] \times [j]$ and calculate

$$\theta(\alpha(v)) - \beta(\psi(v)) = \rho_v - \eta_v - c_\Gamma(i,j) + [x_i, x_j] = \rho_v - \sum_{k=i}^{j-1} [x_k, x_{k+1}],$$

where we use $[x_i, x_j] = \eta_v$. To calculate $\partial\tau_{[i]\times[k]}$, we note

$$\begin{aligned}
A_1\tau_{[i]\times[k]}(\mu) &= \rho_{[i]\times[k]}(\mu), & B_1\tau_{[i]\times[k]}(\mu) &= \rho_{[i]\times[k+1]}(\mu), \\
A_2\tau_{[i]\times[k]}(\mu) &= (1-\mu)x_i - \mu x_i = x_i, & B_2\tau_{[i]\times[k]}(\mu) &= (1-\mu)x_{k-1} - \mu x_k = [x_{k-1}, x_k].
\end{aligned}$$

Since A_2 is degenerate, this yields $\partial\tau_{[k]\times[k+1]} = \rho_{[i]\times[k+1]}(\mu) - \rho_{[i]\times[k]}(\mu) - [x_k, x_{k+1}]$. Using this, we obtain

$$\partial Hv = \partial\tau_{[i]\times[j]} = \sum_{k=i+1}^{j} \partial\tau_{[i]\times[k]} = \rho_v - \sum_{k=i}^{j-1} [x_k, x_{k+1}]$$

which implies the desired equality for all vertices and by linearity for all 0-chains.

For an edge $e \in \check{\mathrm{C}}\mathrm{DC}_r^{(1)}(\Gamma)$ we set

$$\tau_e(\mu_1, \mu_2, \mu_3) = \begin{cases} (1-\mu_3)\eta_e(\mu_1,\mu_2) + \mu_3\left((1-\mu_2)x_i + \mu_2((1-\mu_1)x_j + \mu_1 x_{j+1})\right) & \text{if } e = [i] \times [j, j+1], \\ (1-\mu_3)\eta_e(\mu_1,\mu_2) + \mu_3\left((1-\mu_2)((1-\mu_1)x_i + \mu_1 x_{i+1}) + \mu_2 x_j\right) & \text{if } e = [i, i+1] \times [j]. \end{cases}$$

We verify $\tau_e \in C^{Cub}(O_{r+\epsilon}(\gamma))$ for all edges e in $\check{\mathrm{C}}\mathrm{DC}_r(\Gamma)$. For an edge $e = [i] \times [j, j+1]$, we have

$$\begin{aligned}
&\|\tau_e - \eta_e\|_\infty = \max_{\mu \in [0,1]^3} \mu_3 \left\|\eta_e(\mu_1,\mu_2) - (1-\mu_2)x_i + \mu_2((1-\mu_1)x_j + \mu_1 x_{j+1})\right\|_\infty \\
&= \max_{\mu\in[0,1]^2} \left\|(1-\mu_2)\gamma(t_i) + \mu_2\gamma((1-\mu_1)t_j + \mu_1 t_{j+1}) - (1-\mu_2)x_i + \mu_2((1-\mu_1)x_j + \mu_1 x_{j+1})\right\|_\infty \\
&= \max_{\mu\in[0,1]^2} \left\|\mu_2\gamma((1-\mu_1)t_j + \mu_1 t_{j+1}) - \mu_2((1-\mu_1)\gamma(t_j) + \mu_1\gamma(t_{j+1}))\right\|_\infty \\
&= \max_{\mu\in[0,1]} \left\|\gamma((1-\mu)t_j + \mu t_{j+1}) - ((1-\mu)\gamma(t_j) + \mu\gamma(t_{j+1}))\right\|_\infty \\
&\leq \frac{(t_{j+1} - t_j)^2}{2}\|\gamma''\|_\infty \leq \frac{\delta^2}{2}\|\gamma''\|_\infty \leq \epsilon,
\end{aligned}$$

where the first inequality in the last line follows from the approximation error for polynomial interpolation (see e.g. [SB02, Theorem 2.1.4.1]). A similar calulation works for edges $e = [i, i+1] \times [j]$. Since η_e is supported on $O_r(\gamma)$, the support of τ_e must be contained in $O_{r+\epsilon}(\gamma)$.

Since ρ_e is supported on γ and $H_1^{Cub}(\gamma) = 0$, there is a chain c_e such that $\partial c_e = \rho_e$. We now set $He = c_e + \tau_e - d_e$ for all edges in $\check{\mathrm{C}}\mathrm{DC}(\Gamma)$, where d_e will be specified later, and extend the map by linearity. To verify $\theta \circ \alpha - \beta \circ \psi = \partial H + H\partial$, we fix an edge $e = [i] \times [j, j+1]$ and note that

$$\begin{aligned}
\tau_e(\mu_1,\mu_2,\mu_3) &= (1-\mu_3)\eta_e(\mu_1,\mu_2) + \mu_3\left((1-\mu_2)x_i + \mu_2((1-\mu_1)x_j + \mu_1 x_{j+1})\right) \\
&= (1-\mu_3)\left((1-\mu_2)\gamma(t_i) + \mu_2\gamma((1-\mu_1)t_j + \mu_1 t_{j+1})\right) \\
&\quad + \mu_3\left((1-\mu_2)x_i + \mu_2((1-\mu_1)x_j + \mu_1 x_{j+1})\right)
\end{aligned}$$

Therefore,

$$\begin{aligned}
A_1\tau_e(\mu_2,\mu_3) &= (1-\mu_3)\big((1-\mu_2)\gamma(t_i)+\mu_2\gamma(t_j)\big)+\mu_3\big((1-\mu_2)\gamma(t_i)+\mu_2\gamma(t_j)\big)\\
&= \big((1-\mu_2)\gamma(t_i)+\mu_2\gamma(t_j)\big)\\
B_1\tau_e(\mu_1,\mu_3) &= (1-\mu_3)\big((1-\mu_2)\gamma(t_i)+\mu_2\gamma(t_{j+1})\big)+\mu_3\big((1-\mu_2)\gamma(t_i)+\mu_2\gamma(t_{j+1})\big)\\
&= \big((1-\mu_2)\gamma(t_i)+\mu_2\gamma(t_{j+1})\big)\\
A_2\tau_e(\mu_1,\mu_3) &= (1-\mu_3)\gamma(t_i)+\mu_3\gamma(t_i)=\gamma(t_i)\\
B_2\tau_e(\mu_1,\mu_3) &= (1-\mu_3)\gamma((1-\mu_1)t_j+t_{j+1})+\mu_3\big((1-\mu_1)\gamma(t_j)+\mu_1\gamma(t_{j+1})\big)=\tau_{[j]\times[j+1]}(\mu_1,\mu_3)\\
A_3\tau_e(\mu_1,\mu_2) &= \eta_e(\mu_1,\mu_2)\\
B_3\tau_e(\mu_1,\mu_2) &= (1-\mu_2)\gamma(t_i)+\mu_2((1-\mu_1)\gamma(t_j)+\mu_1\gamma(t_{j+1})).
\end{aligned}$$

Note that the first three singular cubes are degenerate. We furthermore have $H\partial e = H([i]\times[j]) - H([i]\times[j+1]) = -\tau_{[j]\times[j+1]}$, thus

$$\partial He + H\partial e = (\rho_e - \tau_{[j]\times[j+1]} - \eta_e + B_3\tau_e - \partial d_e) + \tau_{[j]\times[j+1]} = \rho_e - \eta_e + B_3\tau_e - \partial d_e.$$

Starting with the other side of the equation, we have

$$\theta\circ\alpha(e) - \beta\circ\psi(e) = \rho_e - \eta_e - \beta([x_i,x_j,x_{j+1}]),$$

where $\beta([x_i,x_j,x_{j+1}])(\mu_1,\mu_2) = (1-\mu_1)x_i + \mu_1(1-\mu_2)x_j + \mu_1\mu_2 x_{j+1}$. Since

$$\begin{aligned}
A_1B_3\tau_e(\mu) &= (1-\mu)x_i+\mu x_j = A_2q(\mu),\\
B_1B_3\tau_e(\mu) &= (1-\mu)x_i+\mu x_{j+1} = B_2q(\mu),\\
A_2B_3\tau_e(\mu) &= x_i = A_1q(\mu),\\
B_2B_3\tau_e(\mu) &= (1-\mu)x_j+\mu x_{j+1} = B_1q(\mu).
\end{aligned}$$

We have $\partial(B_3\tau_e+\beta([x_i,x_j,x_{j+1}]))=0$. Since both simplices are supported in the acyclic set $|[x_i,x_j,x_{j+1}]|$, we can find a chain d_e such that $\partial d_e = B_3\tau_e+\beta([x_i,x_j,x_{j+1}])$. With this choice, we have $\theta\circ\alpha(e)-\beta\circ\psi(e)=\partial He+H\partial e$. Similar calculations work for edges of the form $[i,i+1]\times[j]$, we thus have constructed the desired chain homotopy. □

Theorem 5.4.6. *The cokernel of $H_0(\theta)\colon H_0(\mathrm{DL}(\gamma))\to H_1(O(\gamma))$ is ephemeral.*

Proof. To show that $\mathbb{V} = \operatorname{coker} H_0(\theta)$ is ephemeral, we need to show that all structure maps are zero. We show $\mathbb{V}_{r,r+2\epsilon}=0$ for $r>0$, $\epsilon>0$ and treat the case $r=0$ later. By

Lemma 5.4.5 we can find a sample Γ of γ such that there is a commutative diagram

$$\begin{tikzcd}
H_0(\check{\mathrm{C}}\mathrm{DC}_{r+\epsilon}(\Gamma)) \arrow[r, two heads, "\phi"] \arrow[d, "\alpha"] & H_1(\check{\mathrm{C}}_{r+\epsilon}(\Gamma)) \arrow[d, "\beta"] & H_1(O_r(\gamma)) \arrow[l, "\beta'"'] \arrow[dl, "H_1(O_{r,r+2\epsilon}(\gamma))"] \\
H_0(\mathrm{DL}_{r+2\epsilon}(\gamma)) \arrow[r, "\psi"] & H_1(O_{r+2\epsilon}(\gamma)). &
\end{tikzcd}$$

Now fix $h \in H_1(O_r(\gamma))$. Since ψ is surjective, there is $\tilde{h} \in H_0(\check{\mathrm{C}}\mathrm{DC}_{r+\epsilon}(\Gamma))$ such that $\psi_{r+\epsilon}(\tilde{h}) = \beta'_r(h)$. Thus, with $k = \alpha_{r+\epsilon}(\tilde{h})$, we have $H_1(O_{r,r+2\epsilon}(\gamma))(h) = \theta_{r+2\epsilon}(k)$. Therefore, $H_1(O_{r,r+2\epsilon}(\gamma))(h)$ is an element in $\operatorname{im}\theta_{r+2\epsilon}$ and therefore trivial in $\mathbb{V}$. This shows that all structure maps $\mathbb{V}_{r,r+\epsilon} = 0$ for $r > 0$. In case $r = 0$, we choose $s \in (0, \epsilon)$ and observe $\mathbb{V}_{0,\epsilon} = \mathbb{V}_{s,\epsilon} \circ \mathbb{V}_{0,s} = 0$ since $\mathbb{V}_{s,\epsilon} = 0$ by the previous argument. $\square$

6 Time Series Analysis using Sampled Cycling Signatures

The notions on cycling introduced in Chapter 3 apply to a single segment of a given time series. In order to obtain an overview on the cycling properties of different segments from a (long) time series, we carry out a sampling approach.

This chapter is structured as follows. In Section 6.1, we introduce the sampling setup and three example times series. In Section 6.2, we define summary statistics to analyze the resulting collection of sampled segments. This allows us to give a heuristic definition of cycling motions as 'prevalent' one-dimensional cycling signatures. In Section 6.3, we experimentally explore stability of this analysis. The analysis of cycling signatures is augmented in Section 6.4 using the observations in Chapter 5. In particular, we analyze the relationship between different cycling spaces and transitions. Finally, in Section 6.5 we explore the relationship between cycling signatures and unstable periodic orbits in a chaotic attractor.

Sections 6.1 to 6.3 are joint work with Ulrich Bauer, Oliver Junge and Konstantin Mischaikow and part of [BHJM23].

6.1 Example Systems: Lorenz, double well, Dadras

We generate example datasets (see Fig. 6.1) by numerically integrating three systems of differential equations. We then choose a (finite) set of time spans $\mathcal{T} \subset [0, \infty)$, a sample size $N \in \mathbb{N}$ and randomly sample N segments of time span τ for each $\tau \in \mathcal{T}$. We then compute the cycling signature for these $N \cdot |\mathcal{T}|$ segments as described in Chapter 4. All cycling computations are carried out using coefficients in $\mathbb{F}_{43}$.

6.1.1 Lorenz System

The Lorenz system [Lor63] is the system of differential equations

$$\begin{aligned}
\dot{x} &= \sigma(y - x), \\
\dot{y} &= x(\rho - z) - y, \\
\dot{z} &= xy - \beta z.
\end{aligned}$$

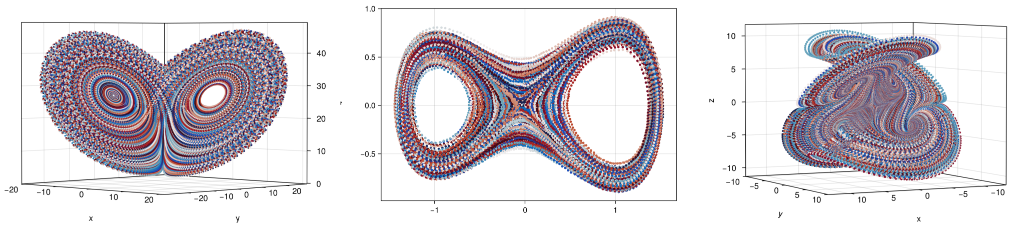


Figure 6.1: Time series for the Lorenz system (left), a stochastically perturbed Hamiltonian system with a double well potential (center) and the Dadras system projected to the first three coordinates (right). The color of the x_i changes gradually with time t_i to visualize time evolution.

We consider the system at the classical values $\sigma = 10$, $\rho = 28$ and $\beta = \frac{8}{3}$ where it is known to exhibit a chaotic attractor that resembles the shape of a butterfly.

Time Series Generation We integrate the system numerically, from time 0 to 10000, using the Tsit5 scheme [Tsi11] with initial condition $(0, 10, 0)$ and step size 0.1. For each time series point, we compute the lift to the unit tangent bundle using the explicit formula for the vector field. We sample additional points to ensure that the dynamic UT-metric d_8 of two consecutive points is less than 1. The resulting dataset Γ_L set is shown on the left of Fig. 6.1.

Signature Computations We construct a cubical comparison space Y_L for Γ_L using boxes with box size $(r, k) = (8, 1)$. It has $\dim H_1(Y_L) = 2$. We sample $N = 1000$ segments for each time span $\tau \in \mathcal{T} = \{10, 20, \ldots, 1000\}$ and for each segment γ, we compute the cycling signature $\mathrm{Cyc}(\gamma, Y_L)$ using the metric d_C with $C = 8$ and maximal filtration value 8.

6.1.2 Double Well System

Consider the stochastic differential equation

$$dx = f(x)dt + \sigma dW, \tag{6.1}$$

$x = (q, p) \in \mathbb{R}^2$, with W denoting the Wiener process. For the vector field $f : \mathbb{R}^2 \to \mathbb{R}^2$, we choose an asymmetric double well Hamiltonian system with two homoclinic orbits, given by

$$f(x) = \begin{pmatrix} H_p(x) \\ -H_q(x) \end{pmatrix} - ah(H(x)) \begin{pmatrix} H_q(x) \\ H_p(x) \end{pmatrix}$$

where $h(z) = (z^3 - z)/2$ and

$$H(q, p) = p^2/2 + q^4/8 - q^2/2 - q^3/15 - q/10.$$

Time Series Generation We generate two time series by integrating the system numerically using the SRIW1 scheme [Rö10] with initial value $x = (1, 0.75)$, step size 0.01 and noise level $\sigma = 0.015$. One, with $a = 0.02$, is used throughout the remainder of this chapter (see the middle figure in Fig. 6.1). The other, with $a = 0$, is used in Chapter 3. In both cases, we lift the data to the unit tangent bundle using forward finite differences, i.e. we approximate the unit tangent vector at $x_i \in \Gamma$ by normalizing $x_{i+1} - x_i$.

Signature Computations We construct a cubical comparison space Y_{dw} for Γ_{dw} using boxes with box size $(0.2, 1)$. It has $\dim H_1(Y_{dw}) = 3$. We sample $N = 1000$ segments for each time span $\tau \in \mathcal{T} = \{10, 20, \ldots, 1000\}$ and for each segment γ, we compute the cycling signature $\mathrm{Cyc}(\gamma, Y_{dw})$ using the metric d_C with $C = 0.2$ and maximal filtration value 0.2.

6.1.3 Dadras System

In [DMQW12], a 4d system of differential equations with a 4 wing (hyper)-chaotic attractor is introduced. The equations are

$$\begin{aligned}
\dot{x} &= ax - yz + w, \\
\dot{y} &= xz - by, \\
\dot{z} &= xy - cz + xw \\
\dot{w} &= -y,
\end{aligned}$$

where we set $(a, b, c) = (8, 40, 14.9)$ as suggested in the original article. The system posesses a unique equilibrium at the origin, the eigenvalues of the Jacobian at the origin are 0, a, $-b$, $-c$.

Time Series Generation We integrate the system numerically, from time 0 to 100000 using the Tsit5 scheme [Tsi11] with initial condition $(10, 1, 10, 1)$ and step size 0.01. In the resulting time series, the wings are quite large compared to the distance of their centers to the origin. To ensure that the cubical comparison space does not have too many boxes, we shrink the wings by applying the nonlinear rescaling $x \mapsto x/\sqrt{\|x\|_2}$ to all time series points. As before, we compute the lift to the unit tangent bundle by explicitly using the vector field (and the rescaling). We then sample additional points such that the Euclidean distance of two consecutive points lifted to the unit tangent bundle is less than 0.8. The resulting dataset Γ_{Da} is shown on the right of Fig. 6.1.

Signature Computations We construct a cubical comparison space Y_{da} for Γ_{da} using boxes with box size $(4, 1/3)$. It has $\dim H_1(Y_{dw}) = 5$. We sample $N = 1000$ segments for each time span $\tau \in \mathcal{T} = \{10, 20, \ldots, 1000\}$ and for each segment γ, we compute the cycling signature $\mathrm{Cyc}(\gamma, Y_{da})$ using the metric d_C with $C = 12$ and maximal filtration value 4.

6.2 Summary Statistics for Cycling Signatures

Throughout this section, let Γ be a time series with sampling times T. Furthermore, let $\mathcal{T} \subset \mathbb{R}$ denote a (finite) set of time spans and $S = (S_\tau)_{\tau\in\mathcal{T}}$ be a collection of segments with time span τ sampled from Γ.

In the following we introduce summary statistics for cycling segments and illustrate them using the previously introduced example systems.

6.2.1 Cycling Rank Distribution

A simple way to analyze sampled segments is to focus on their cycling ranks.

Definition 6.2.1. We define the *(empirical) rank distribution* of $S = \{S_\tau\}_{\tau\in\mathcal{T}}$ to be

$$\mathrm{rkDist}_k \colon \mathcal{T} \times \mathbb{R} \to \mathbb{N}_0, \mathrm{rkDist}_k(\tau, r) = \#\{\gamma \in S_\tau \mid \dim \mathrm{Cyc}_r(\gamma) = 1\}.$$

The empirical cycling rank distributions for the example datasets are shown in Fig. 6.2. Rank 0 is detected if either the thickening radius r the segment time span τ is small. For small thickening radii, this is expected since a small thickening of a (discrete) point cloud has no degree 1 homology. Regarding the segment time span, we actually gain insight into the systems: the time span at which the number of rank 0 segments begins to decrease suggests the minimal time span of a cycling segment (Lorenz ~60, double well ~80, Dadras ~40). Furthermore, in all three systems, for large enough thickening radius and segment time span, all segments have positive cycling rank, i.e. any sufficiently long segment in the time series is cycling. This means that in all systems cycling is a typical phenomenon.

6.2.2 Cycling Space Distribution

Cycling spaces offer a significantly more fine-grained view than the cycling rank.

Definition 6.2.2. Let $V \subset H_1(Y)$ be a subspace of the homological comparison space. The *(empirical) cycling space distribution* of V with respect to $S = \{S_\tau\}_{\tau\in\mathcal{T}}$ is

$$\mathrm{VDist}_V \colon \mathcal{T} \times \mathbb{R} \to \mathbb{N}_0, \quad \mathrm{VDist}_V(\tau, r) = \#\{\gamma \in S_\tau \mid \mathrm{Cyc}_r(\gamma) = V\}. \tag{6.2}$$

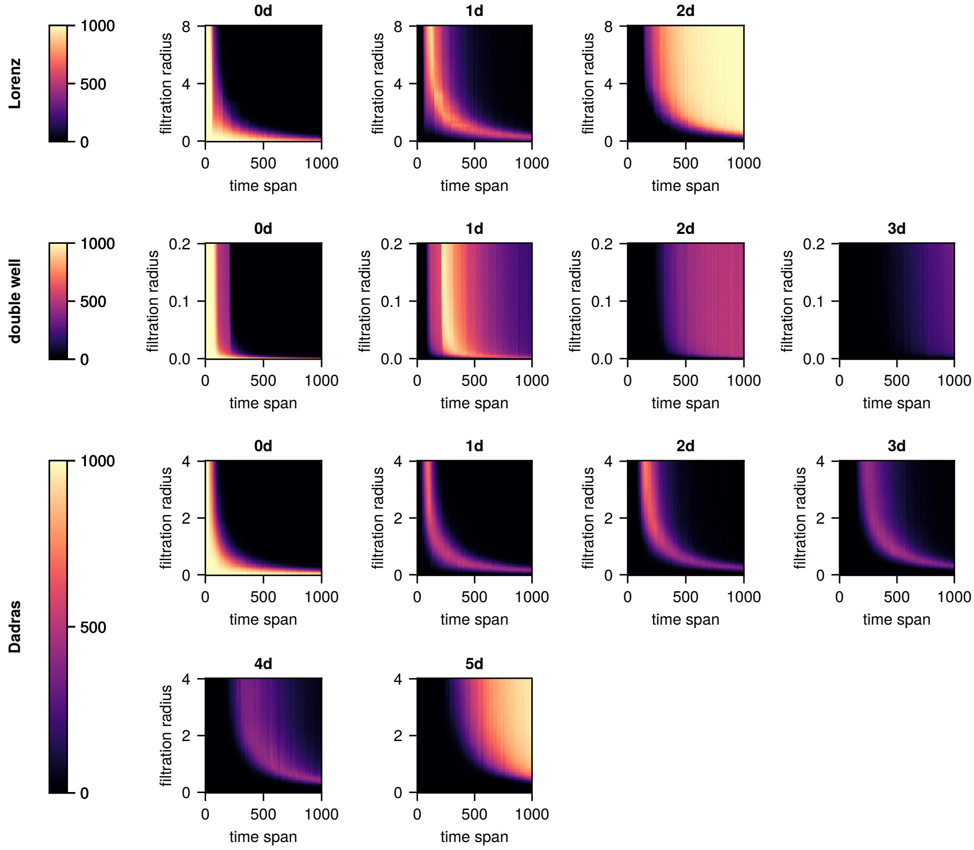


Figure 6.2: Empirical cycling rank distributions for the Lorenz, double well and Dadras datasets. For each system, the rank distributions for all ranks between zero and the dimension of the homological comparison space are shown.

Often, it is easier to analyze the *(empirical) radius distribution* of V with respect to S, i.e.

$$\mathrm{rVDist}_V\colon \mathbb{R} \to \mathbb{N}_0, \quad \mathrm{rVDist}_V(r) = \sum_{\tau\in\mathcal{T}} \mathrm{sigDist}_V(\tau, r)$$

which counts the total number of segment that exhibit V at a given radius, and the *(empirical) time span distribution* of V with respect to S

$$\mathrm{tVDist}_V\colon \mathcal{T} \to \mathbb{N}_0, \quad \mathrm{tVDist}_V(\tau) = \#\{\gamma \in S_\tau \mid \exists r : \mathrm{Cyc}_r(\gamma) = V\},$$

that, for each time span, counts the total number of segments that exhibit V. We furthermore define the *total persistence* of V with respect to S to be $\mathrm{tpV(V)} = \int_0^\infty \mathrm{rVDist}_V(r)\, dr$.

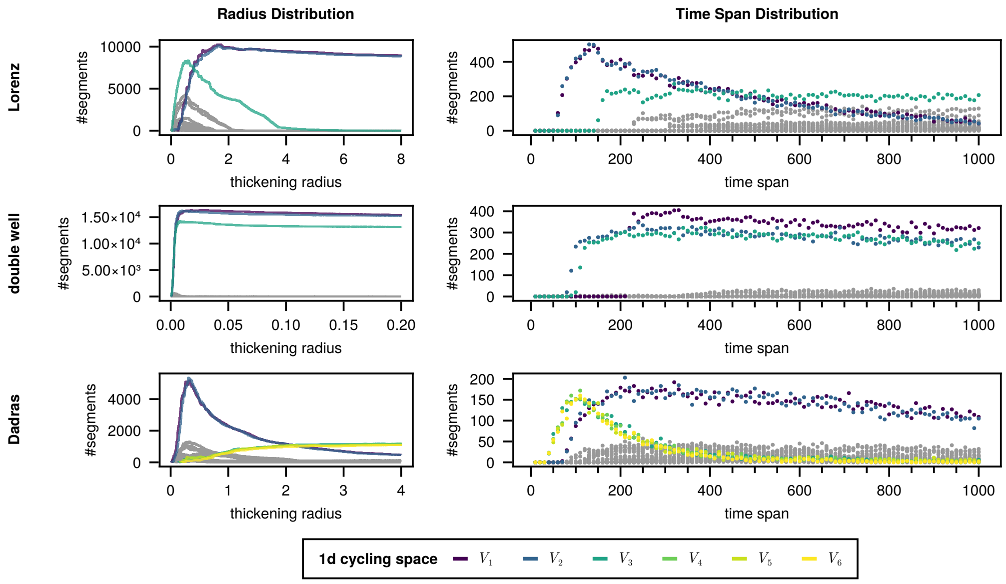


Figure 6.3: Radius distribution (left column) and time span distribution (right column) for 1d cycling spaces in the example systems. We highlight the 3, 3 and 6 cycling spaces in Lorenz, double well and Dadras, respectively, that have the largest total persistence. In Dadras, we only show the 40 most persistent of the 1395 cycling spaces that appear in the sample. Note that the time span distribution is less smooth than the radius distribution.

Remark 6.2.3. In general, $\mathrm{tVDist}_V(\tau)$ is not equal to $\int_0^\infty \mathrm{VDist}_V(\tau, r)\, dr$ since the former counts the number of segments that express a given cycling space, while the latter can be interpreted as a weighted count where the weight is given by the measure of the interval where the cycling space is expressed. It follows that generally also tpV(V) is not equal to $\sum_{\tau \in S} \mathrm{tVDist}_V(\tau)$.

1d Cycling Spaces

One-dimensional cycling spaces capture a single type of cycling, which is uniquely identified by its cycling space. In total, we detect 37 1d cycling spaces in the Lorenz dataset, 8 in the double well dataset and 1395 in the Dadras dataset. Note that this observation does not conflict with the fact that the comparison spaces for the Lorenz and the Dadras system have dimensions 2 and 5, respectively, as the rank 1 cycling spaces may be linearly dependent.

Radius distribution of cycling spaces In the left column of Fig. 6.3, the radius distribution of cycling spaces is visualized. This gives immediate information about the cycling spaces that can be expected for given thickening parameters. In Lorenz, we see that for $r \gtrapprox 4$ only V_1 and V_2 appear. For $2 \lessapprox r \lessapprox 4$, we additionally detect V_3. The other cycling spaces are detected only for even smaller r. In double well, all frequent subspaces appear for $r \approx 0.02$ and appear with similar frequency for all greater r. In the Dadras time series, for large thickening radii, there appear to be six cycling spaces that are detected frequently, though we detect two qualitatively different distributions: While V_3 to V_6 are approximately monotonically increasing, all other cycling spaces peak for some radius and then fall off quickly. This is particularly noticeable for V_1 and V_2 and the sharp peark suggests that many of these segments have cycling rank greater than one for larger thickening radius. In all systems, this suggests that are there are a small number of 'macroscopic' cycling motions despite the larger number of 1d cycling spaces that are detected.

Time span distribution of cycling spaces The time span distribution of cycling spaces (see right column of Fig. 6.3) is particular interesting for small time spans since it provides insight into the minimal time span of a cycling space to be detected. Both, the Lorenz and the double well system possess two cycling spaces with small τ_{min} compared to a third cycling space, whereas in Dadras, four spaces V_3, V_4, V_5, V_6 have small τ_{min} compared to V_1 and V_2, see Table 6.1.

System	$\tau_{min}(V_1)$	$\tau_{min}(V_2)$	$\tau_{min}(V_3)$	$\tau_{min}(V_4)$	$\tau_{min}(V_5)$	$\tau_{min}(V_6)$
Lorenz	50	60	150	(230)	(230)	(300)
Double Well	210	70	90	(130)	(30)	(110)
Dadras	70	70	40	40	40	40

Table 6.1: Values of τ_{min} for cycling spaces in the example systems.

In the Lorenz system this is expected, since the third type of cycling, the longer one, is the "sum" of the two shorter ones on the individual wings and traces out both wings as shown in Fig. 6.4.

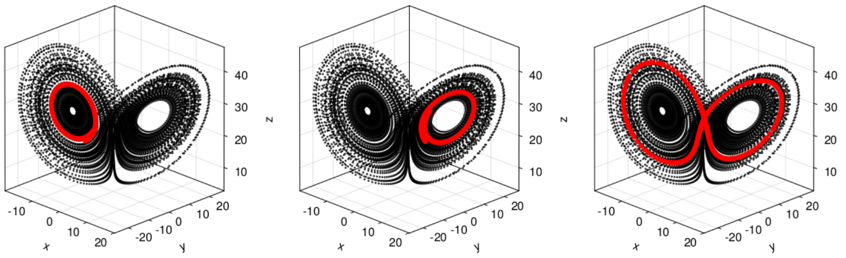


Figure 6.4: Rank 1 segments in the Lorenz system. The full time series Γ is shown in black. The red points in each plot highlight a segment with cycling space V_1, V_2 and V_3, respectively.

Especially in Lorenz and Dadras, new cycling spaces appear for larger time span. Referring back to the rank distribution, segments with longer time span typically have cycling rank 1 only for small thickening radii. This suggests additional structure in the data that is smaller in scale than the ‘macroscopic’ cycling spaces we highlighted so far. In Section 6.5, we link these cycling spaces to unstable periodic orbits.

Cycling Motions

Using the summary statistics for one-dimensional cycling spaces, we now give a heuristic definition of cycling motions.

Definition 6.2.4. A *cycling motion* in a time series is a ‘prevalent’ one-dimensional cycling space.

Higher-dimensional Cycling Spaces

More complicated cycling behavior is captured by segments with cycling rank greater than one. As mentioned in Example 3.2.5, the cycling space itself does not reveal which one-dimensional cycling spaces give rise to it, since vector spaces do not canonically decompose into one-dimensional subspaces. However, cycling spaces with dimension two or greater can still be used to relate lower dimensional cycling spaces with each other via subspace inclusion. For example, prominent higher-dimensional cycling spaces can reveal clusters of 1d cycling spaces where the time series stays for extended periods.

A phenomenon of interest are transitions between one-dimensional cycling spaces, i.e. segments that consecutively visit different one-dimensional cycling spaces. Any segment with a transition has a signature with dimension two or greater that contains the corresponding one-dimensional cycling spaces as subspaces. Conversely, there cannot be transitions between two specific 1d cycling spaces if there is no segment with a higher-dimensional cycling space that contains these two 1d spaces. These observations allow

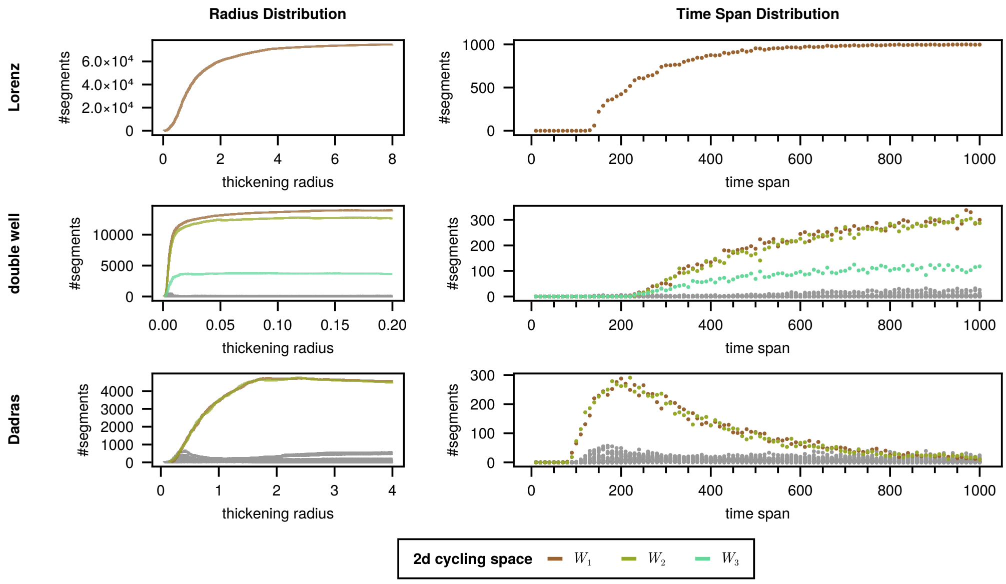


Figure 6.5: Radius distribution of 2d cycling spaces in the example systems.

us to make preliminary statements about transitions between certain cycling spaces. We present a more thorogh analysis of transitions in Section 6.4.

2d Cycling Spaces In the Lorenz system, the homological comparison space is of dimension two, therefore all rank 2 segments have the same cycling space, and all three of the frequent one-dimensional cycling spaces are subspaces of this space.

In the other two systems, rank 2 segments provide insight into the transitions between cycling spaces: We cannot rule out any transition between the three prominent cycling spaces in the double well system. To see this, note that the statistics of the 2d cycling spaces, depicted in Fig. 6.5, reveal three prominent 2d cycling spaces. It can be seen in Fig. 6.6 that any pair of 1d cycling spaces is contained in a 2d cycling space (for example V_1 and V_2 are contained in W_3).

For transitions that occur we obtain quantitative statements: As can be seen from Fig. 6.5, the spaces W_1 and W_2 appear much more frequently than W_3. The inclusion graph in Fig. 6.6 (center) reveals that W_1 and W_2 contain V_3, the space with large τ_{min}, and one of the others V_1 and V_2, whereas W_3 contains both of the spaces with small τ_{min}. An interpretation of this is that transitions between V_1 and V_2 are less likely than transitions involving V_3.

In the Dadras system, we detect two prominent 2d cycling spaces which organize the six one-dimensional cycling spaces into two clusters. These are defined by W_1 and W_2, and consist of $V_1, V_4, V_5 \subset W_1$ and $V_2, V_3, V_6 \subset W_2$, i.e., each contains two subspaces with smaller, and one with larger τ_{min}. Note that W_1 and W_2 are disjoint, therefore direct transitions between cycling spaces in the same cluster are frequent while direct transitions between cycling spaces from different clusters are rare.

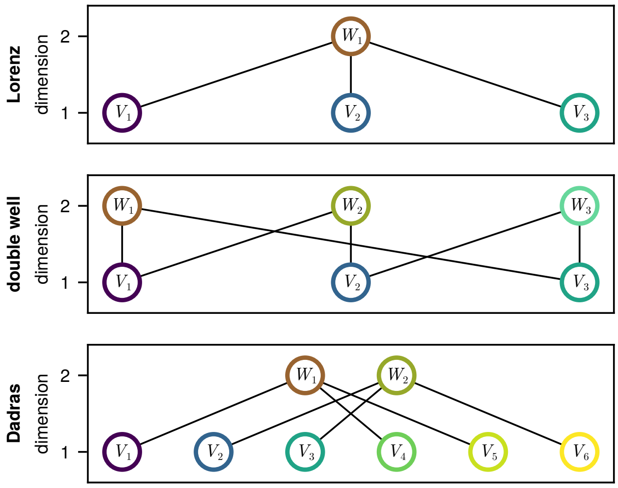
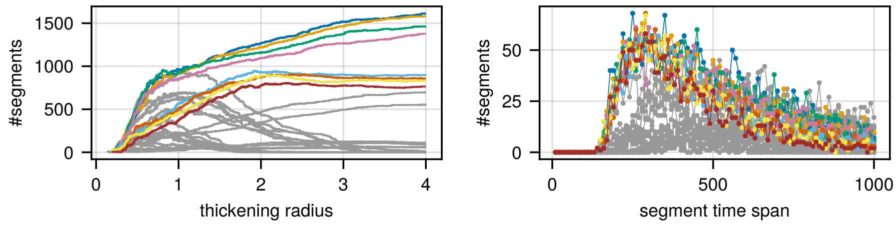
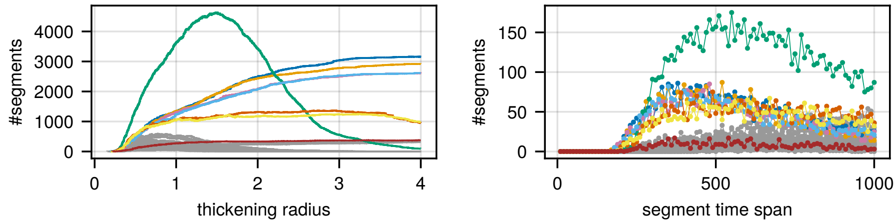


Figure 6.6: Hasse diagram of the inclusion order of cycling spaces. In each diagram, the nodes in the bottom and top rows correspond to frequent rank 1 and 2 cycling spaces, respectively. Node colors correspond to the ones in Fig. 6.3 for 1d spaces and Fig. 6.5 for 2d spaces. An edge indicates that the rank 1 cycling space is a subspace of the rank 2 cycling space.

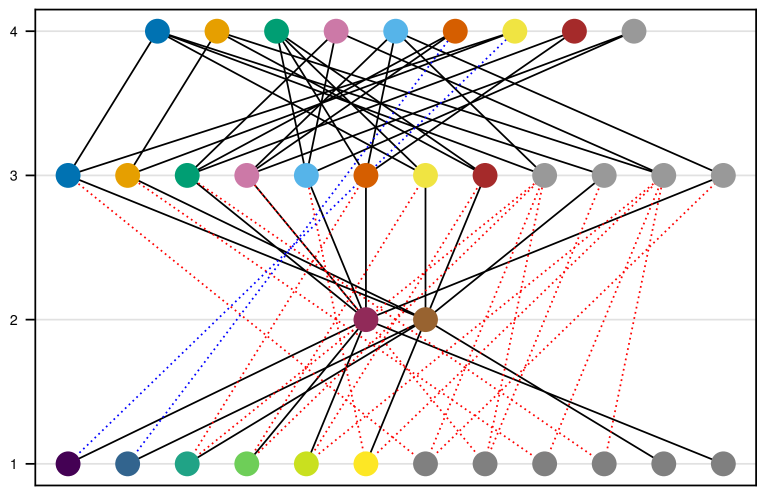
3d/4d/etc. Cycling Spaces Since our comparison spaces for the Lorenz and double-well time series are two-dimensional and three-dimensional, respectively, we detect no cycling space with dimension three or greater in the Lorenz dataset and only one in the double well dataset. In contrast, in the Dadras dataset, a large number of such cycling spaces are detected. The radius and length distributions of the 3d and 4d cycling spaces are depicted in Fig. 6.7a and Fig. 6.7b, respectively. These figures reveal that, compared to the 1d and 2d cycling spaces, there is no separation of frequent and infrequent cycling spaces. Some structure becomes apparent when looking at the Hasse diagram of the subspace relation in Fig. 6.7c: Most 3d cycling spaces are a sum of W_1 or W_2 and some 1d cycling space. This can be interpreted as an indicator for chaos in the underlying system since, after spending time in a 2d cycling space, it seems unpredictable which 1d cycling space is visited afterwards. Lastly, it is interesting to note that in the distribution of 4d cycling spaces, a green curve stands out in radius and length distribution. By Fig. 6.7c, this curve corresponds to the cycling space which is the direct sum of W_1 and W_2.



a 3d cycling spaces.



b 4d cycling spaces.



c Hasse diagram for 1- to 4-dimensional cycling spaces with subspace relation.

Figure 6.7: Radius distribution, time span distribution and subspace inclusion for 3d and 4d cycling spaces in the Dadras time series. **a**, **b** The 30 cycling spaces of dimension 3 and 4, respectively, with the largest total persistence. In both figures, each cycling space has the same color in radius and time span distribution and the most persistent 8 have a color different from gray. In the time span distributions, points corresponding to the same cycling space are connected for better visibility. **c** Hasse diagram of the subspace relation. The colors of the vertices are consistent with the colors in Figs. 6.6, 6.7a and 6.7b. Edges between non-consecutive dimensions are dotted and colored for better visibility only and have no other meaning.

Comparing Lorenz, double well and Dadras

Visually, there are similarities between the time series of the Lorenz and the double well system (see Fig. 6.1); they both have two centers ('holes') around which trajectories cycle. The dimensions of the comparison spaces differ (2 for Lorenz, 3 for double well), but in and of itself this does not necessarily imply significant differences in the dynamics. However, we can quantify differences using the statistics of cycling signatures.

In both systems, three 1d cycling spaces stand out (see Fig. 6.3). Furthermore, the time span distribution reveals that, in both systems, one cycling space has τ_{min} approximately twice as large as the other two. In Lorenz, the two shorter 1d cycling spaces have almost identical radius and time span distribution; the longer one is different in both distributions. Its radius distribution increases for smaller radii than the other two but also falls off much more quickly which suggests, that at the coarsest level the Lorenz system only has two cycling motions. Furthermore, its time span distribution is fairly constant once it is detected while the other two are monotonically falling.

The double well system also has two prominent 1d cycling spaces with similar τ_{min}, though the slight shift in the time span distribution suggests that V_3 is slightly longer than V_2. Furthermore, the radius distribution shows that V_3 is detected less often than V_2. This shows that the shorter cycling motions are less symmetric than the ones in the Lorenz time series. The structure of the third cycling motion is different in the systems. In the Lorenz system, this third space is observed infrequently, whereas in the double well system, it dominates in frequency once it appears at a time scale of $\tau_{min} \approx 210$.

Cycling behavior in the Dadras time series is very different from the other examples since we find six 1d cycling spaces while the other two systems only have three.

Cycling Signatures in the Dadras system

The Dadras system is a 4d system of ordinary differential equations, which appears to exhibit a (hyper)chaotic attractor [DMQW12]. As the system is four-dimensional, visualization is limited to projections that necessarily omit some spatial information. Cycling signatures allow us to analyze the dynamics in the Dadras system without visualizations.

As stated above, Fig. 6.3 indicate six 1d cycling spaces. Our analysis of rank two segments (Fig. 6.6) reveals that these 1d cycling spaces are organized in two Lorenz-like clusters, each of which comprises two spaces with smaller τ_{min} than the third which they span.

While the 1d and 2d cycling spaces indicate a degree of order in the system, the 3d and 4d cycling spaces show quite the opposite. Indeed, the large number of these cycling spaces indicates that transitions into and out of the 2d 'clusters' are unpredictable through cycling signatures. This is an indicator for chaos or randomness in the system.

Together, the V_i, $i = 1, \ldots, 6$, span a four-dimensional subspace of the homological comparison space. Similar to the double well system, there is a hyperbolic equilibrium at the origin that generates a nontrivial homology class in the homological comparison space. This generator, together with the four 1d cycling spaces V_1, V_2, V_3, V_4, generates the 5 dimensional homological comparison space.

6.3 Stability

The computation of cycling signatures depends on the choice of comparison space and unit tangent bundle metric. From a practical point of view it is important that the computation of cycling signatures is stable with respect to these choices. Lacking a thorough theoretical understanding of the stability of cycling signatures, we experimentally investigate the effect of small parameter changes on the presented results.

To experimentally investigate stability of our results, we repeat the experiment with different parameters. More precisely, we choose $(r, k, C) = (8, 2, 16)$ as reference parameter values and vary the parameters r, k, C as follows:

- $r = 7$, 8, 9 with $k = 2$ and $C = 16$,
- $k = 1$, 2, 3 with $r = 8$ and $C = 16$,
- $C = 16$, 17, 18 with $r = 8$ and $k = 2$.

For each parameter combination, we sample 1000 segments for segment time spans $\{10, 20, \ldots, 500\}$. Thus, changes in the cycling signatures arise not only from changes in parameters but also due to sampling. To simplify visualization, we evaluate all cycling signatures at radius 2 which should adequately capture the three most prominent cycling motions. As we can see in Fig. 6.8 and Fig. 6.9, changes in the statistics of rank distribution and are small in the tested range of parameters.

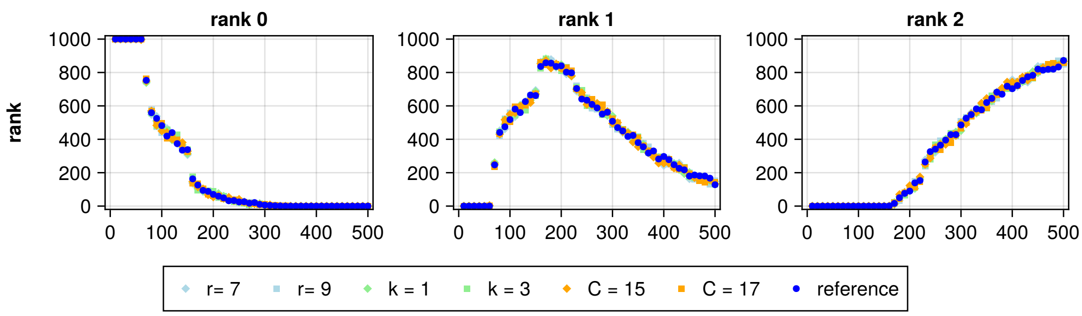


Figure 6.8: Rank distribution versus segment time span for a variety of parameter choices. All cycling signatures are evaluated at $r = 4$.

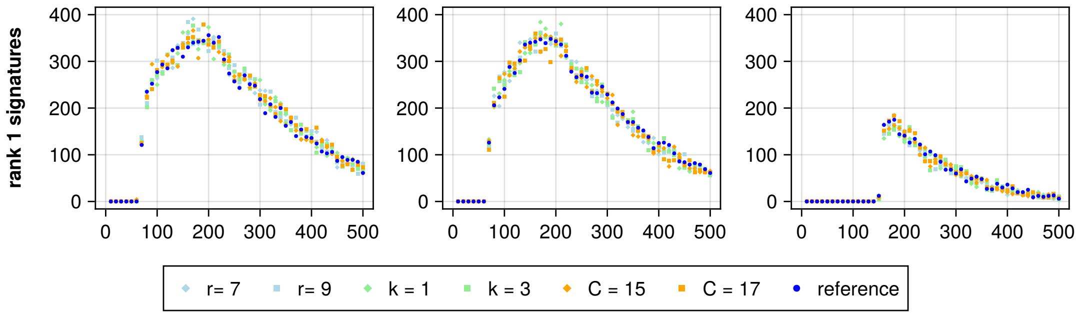


Figure 6.9: Distribution of 1d cycling spaces versus segment time span for a variety of parameter choices. All cycling signatures are evaluated at $r = 4$.

6.4 Curve Cycles

Higher-dimensional cycling spaces allow us to relate lower-dimensional cycling spaces using subspace inclusion. However, this is not sufficient to determine for an individual segment the elementary oscillations that it is composed of and the transitions that happen in the segment. In the following, we refine the analysis of cycling signatures using our observations on persistent homology and distance matrices from Chapter 5.

Let Γ be a time series with sampling times $\bar{t} = (t_k)_{k=1}^{N}$ and let ψ be the chain map defined in Eq. (5.1).

Definition 6.4.1. A chain $\psi(v)$ where $v = (i, j)$ is a vertex of $\mathrm{DC}(\gamma)$ is called a *curve cycle* for γ. We call the interval $[t_i, t_j]$ its *time domain* and the homology class $[c]$ a *curve homology class* for γ.

As illustrated in Fig. 5.1a, curve cycles have the property that they follow the time series for all but a single edge. If this 'extra' edge is short, then, intuitively, the curve cycle represents an almost-periodic segment of the time series.

Given a tight enough time series, by Theorem 5.2.5, any homology class in the cycling signature is a linear combination of curve homology classes. The collection of curve cycles therefore contains at least as much information as the cycling signature. In the following, we use additional information provided by the curve cycles carry to further study the relationship of different cycling motions.

6.4.1 Curve Cycles and Topological Sums

As a first application of curve cycles, we revisit the observation that the longer cycling motion in Lorenz is the 'sum' of the shorter ones. In Fig. 6.4 this seems like a plausible claim, though so far we were able to deduce that the cycling space V_3 is contained in

V_1+V_2, which is a vacuous statement since the dimension of the homological comparison space is two.

Using curve cycles, we strengthen this to a statement on the level of homology. More precisely, Table 6.2 contains computational results confirming the following: For each $i = 1, 2, 3$ there is a nontrivial homology class $v_i \in V_i$ such that: most segments γ with $\mathrm{Cyc}_{r_0}(\gamma) = V_i$ have a curve cycle $c \in v_i$ with birth b_c such that

$$B_r = \begin{cases} \{[c]\} & \text{if } b_c \leq r \leq r_0, \\ \emptyset & \text{else,} \end{cases}$$

is a filtration compatible basis for $\mathrm{Cyc}(\gamma)$. Furthermore, we have $v_3 = v_1 + v_2$.

Note that $r_0 = r_0(Y_L)$ is the maximal radius for which the cycling signature is defined meaning that we only consider segments that have at most cycling rank 1.

V_i	v_i	**#segments with** v_i	**#segments with** V_i	**percentage**
V_1	$\begin{bmatrix} 0 \\ 1 \end{bmatrix} \mod 43$	19353	19353	100%
V_2	$\begin{bmatrix} -1 \\ 0 \end{bmatrix} \mod 43$	19127	19167	99.8%
V_3	$\begin{bmatrix} -1 \\ 1 \end{bmatrix} \mod 43$	12722	17620	72.2%

Table 6.2: Output of computations to experimentally verify that most occurences of V_i are generated by a curve class v_i and that $v_1 + v_2 = v_3$.

6.4.2 Transition Analysis

Intuitively, a transition from a cycling space into another cycling space in a time series should mean that a segment that generates the first signature is 'followed' by a segment that generates the second signature. We now define such a notion of transition between cycling spaces using curve cycles.

We use the following order relation for intervals.

Definition 6.4.2. Let I and J be intervals. If for all $s \in I$ there is a $t \in J$ such that $s \leq t$, then I *bounds* J *above*. If for all $t \in J$ there is a $s \in I$ such that $s \leq t$, then J *bounds* I *below*.

Suppose γ is a segment of Γ with curve cycles c and c' that have time domains I and I', respectively.

Definition 6.4.3. The segment γ has a *transition* from c to c', if I bounds I' from below and I' bounds I from above.

We now use transitions of curve cycles to define transitions of 1d cycling spaces.

Definition 6.4.4. A segment γ has a transition from cycling space V_1 to V_2 if it has a transition of curve cycles c_1 and c_2 such that $[c_1]$ generates V_1 and $[c_2]$ generates V_2.

A segment γ that attains a cycling rank greater or equal two can have any number of transitions between 1d cycling spaces. A naive (and computationally expensive) way to compute all transitions of γ would be to compute all curve cycles of a segment and then, for each pair, check if it is a transition between 1d cycling spaces. In our examples, we used the following heuristic approach:

1. Compute a filtration-minimal vertex for each connected component in $\mathrm{DC}_{r_0}(\gamma)$.
2. Compute the 1d cycling space for each such vertex.
3. For each pair of cycling spaces, check if there is a transition using the previously computed curve cycles.

While this may not find all possible transitions, it proved to be sufficient for our examples.

Transitions in the Double Well System

We strengthen the previous claim on transitions in the double well system. For each of the cycling spaces W_1, W_2 and W_3, we want to confirm that they 'typically' contain a transition between the respective 1d cycling spaces from Fig. 6.6.

For each of the cycling spaces W_i, $i = 1, 2, 3$, we therefore compute all segments γ with $\mathrm{Cyc}_{r_0}(\gamma) = W_i$ and check if there are transitions between the relevant 1d cycling spaces using the simplified procedure described above.

As can be seen in Fig. 6.10, almost all segments with $\mathrm{Cyc}_{r_0}(\gamma) = W_i$ for $i = 1, 2, 3$ contain a transition of the respective frequent 1d subspaces. Furthermore, transitions in both directions occur with similar frequency, e.g. for W_1, we have transitions from V_1 to V_3 and vice versa. Interestingly, with increasing segment length we see that both transitions start to occur for W_1 and W_2, whereas for W_3 this is not the case.

Transitions in the Lü system

Transitions between cycling spaces provide an additional way to qualitatively distinguish time series. As another example, we consider the following system of differential equations

$$\begin{aligned}
\dot{p} &= \frac{cp}{3} + \frac{1}{3M}((c+a)(q^2 - p^2) + 2pq(a - z)) - \frac{a}{3}(p - q) + \frac{qz}{3} \\
\dot{q} &= \frac{1}{3}((c - a)q - (a + z)p) + \frac{1}{3M}((a - z)(p^2 - q^2) + 2pq(a + c)) \\
\dot{z} &= \frac{3}{2}p^2 q - \frac{1}{2}q^3 - bz
\end{aligned}$$

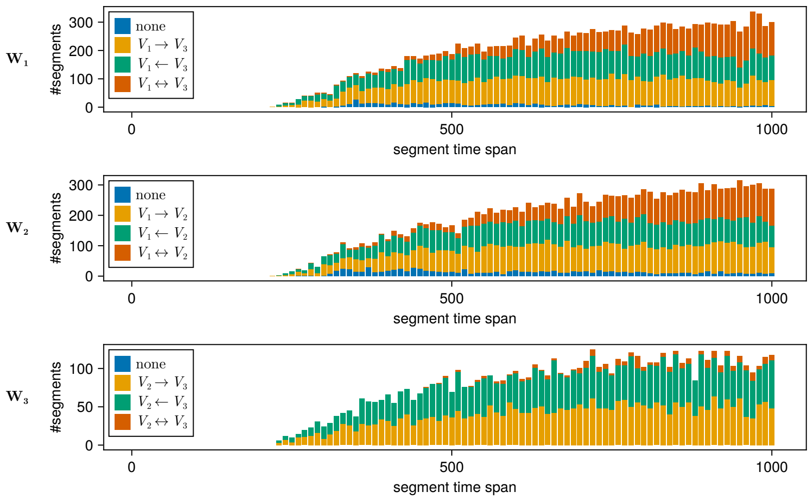


Figure 6.10: Transitions in the double well dataset.

where $M = \sqrt{p^2 + q^2}$ and the parameters are $(a, b, c) = (36, 3, 13)$. This system, which we call the Lü3 system, was derived in [ABP08] by replacing the twofold symmetry of the two wing Lü system [LC02] with a threefold symmetry, as was previously done for the Lorenz system [MS93].

We obtain a time series by numerically integrating this system with initial condition $(1, 0, 4.5)$ and step size 0.01. As before, we generate cycling signatures by sampling 1000 segments for time spans $\{10, 20, \ldots, 500\}$ using a cubical comparison space $Y_{Lu3} \subset \mathcal{Q}_{5,1}$ with $\dim H^1(Y_{Lu3}) = 3$.

An analysis of cycling signatures reveals that, similar to the double well dataset, we detect three frequent 1d cycling spaces and three frequent 2d cycling spaces such that each pair of frequent 1d cycling spaces generates one of the 2d cycling spaces (see left side of Fig. 6.11).

As for the double well system, for each of the cycling spaces W_i, $i = 1, 2, 3$, we compute all segments γ with $\mathrm{Cyc}_{r_0}(\gamma) = W_i$ and check if there are transitions between the relevant 1d cycling spaces using the procedure described above.

The transition analysis reveals that, similar to the double well system, most of these segments contain a transition between the respective 1d cycling spaces. However, we also see that transitions tend to happen in one direction only, namely $V_1 \to V_2$, $V_2 \to V_3$ and $V_3 \to V_1$. This is a qualitative difference from the double well system where transitions in both directions can be found.

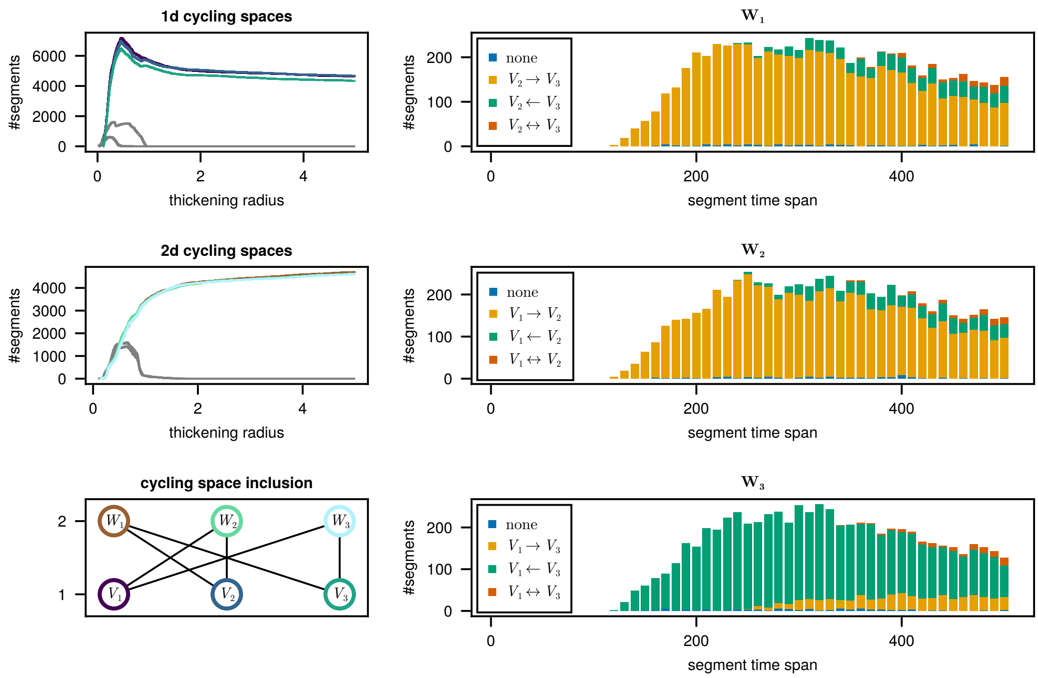


Figure 6.11: Cycling and transitions in the Lü3 system.

6.5 Cycling Signatures and Unstable Periodic Orbits

Unstable periodic orbits (UPOs) play a central role in the study of chaotic attractors. Especially in the physics literature, the set of unstable periodic orbits in a chaotic attractor is often referred to as the skeleton of the attractor. The evolution of a typical trajectory within a chaotic attractor is described as following along unstable periodic orbits and occasionally hopping between them [ACE+87, Cvi91, BM02, ZS06].

Now, suppose γ is a sufficiently densely sampled time series segment that follows a given UPO. Then γ and the UPO are close in interleaving distance, and we expect their cycling signatures to be similar. In particular, the homology class induced by the periodic orbit should generate the cycling space $\mathrm{Cyc}_r(\gamma)$ for a range of small $r > 0$. In a long time series, we expect the UPO to be visited repeatedly. Therefore, we expect these periodic orbits to be revealed through 1d cycling spaces that appear frequently for small thickening radii.

In Fig. 6.12 we plot a segment of minimal time span for the 16 1d cycling spaces with the largest total persistence in the Lorenz dataset. For each segment, we furthermore compute a curve cycle and express its homology basis in terms of the curve cycles corresponding to V_1 and V_2.

It is well known that in the Lorenz system, a Poincaré section can be used to associate a word in the free group with two generators to any periodic orbit. A basic description of this symbolic dynamic is that looping around the left wing corresponds to the symbol A, looping around the right wing corresponds to the symbol B. Every periodic orbit can then be described by a finite sequence of the A and B.

We visually compare Fig. 6.12 with the rigorously computed short periodic orbits from [GT08]. Every segment in Fig. 6.12, other than the ones corresponding to $[0, 1]$ and $[1, 0]$, appears to match a short periodic orbit in the Lorenz system. However, not all UPOs are captured. This can be explained as follows. Passing from symbolic dynamics to homology amounts to passing from the free group with two generators to the free abelian group with two generators. As a result, the order of the symbols is lost. In [GT08], distinct UPOs for the words $AAABB$ and $AABAB$ are approximated. Both map to the vector $[3, 2]$ in the comparison space. Thus, our approach cannot distinguish them.

One could attempt to remedy this using a more thorough transition analysis along the lines of Section 6.4, though we leave this for future work.

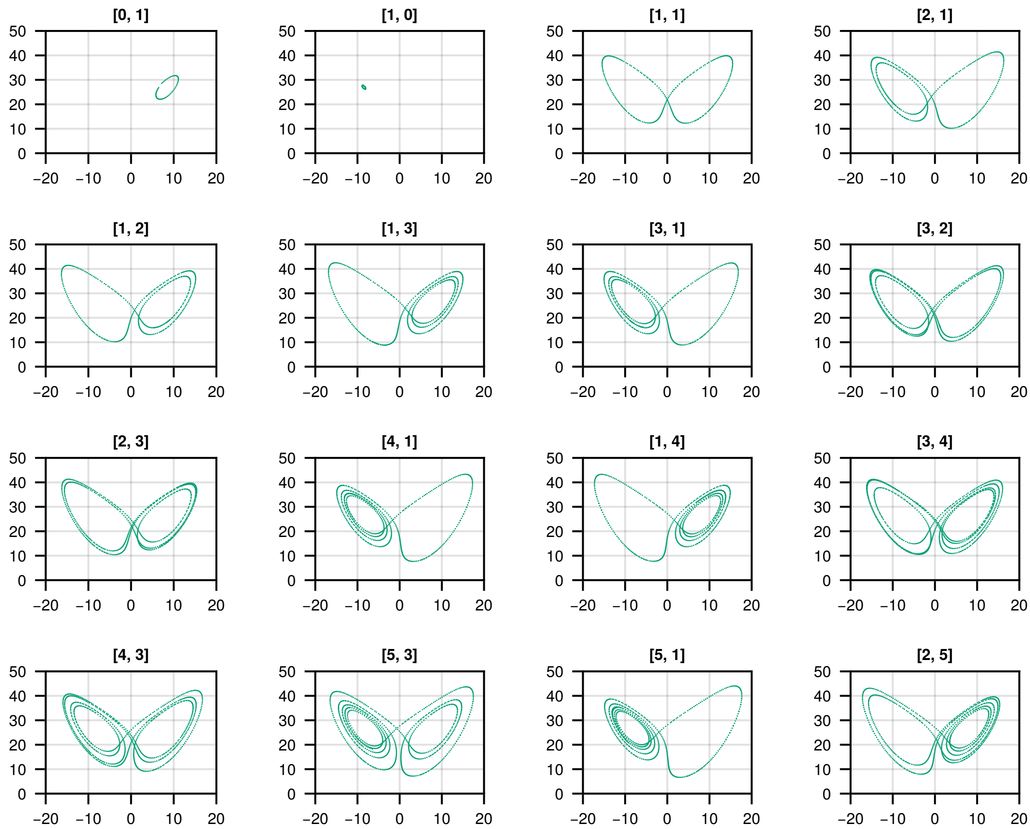


Figure 6.12: Segments for the 16 most persistent 1d cycling spaces. The title of each plot is the homology class of a curve cycle in the comparison space, expressed in the basis indicated by the first two plots.

7 Conclusion

We presented an approach to identify coarse clusters of cycling motions and transitions between these coarse clusters in sampled dynamics. This approach is based on the cycling signature, a topological descriptor for time series segments.

A key ingredient in computing cycling signatures is the construction of the comparison space Y. Here, we constructed Y from an extrinsic cubical covering of a thickening of the given time series. The advantage of this is that typically the number of cubes is much smaller than the number of points in the given time series so that the computation of $H_1(Y)$ becomes feasible. Furthermore, this construction is applicable to general time series data and provides a certain amount of flexibility through the box size parameters. Of course, other constructions are possible and some applications might need comparison spaces which are adapted to the system and data of interest.

In this work, we presented a preliminary stability analysis of the cycling signatures approach. While the cycling signature of a segment is stable with respect to small perturbations of the segment, a comprehensive stability analysis in the context of sampled dynamics is still an open problem.

The cycling signature can be viewed as a three parameter persistence module, where two parameters specify start and end of a segment in a given time series, and the third parameter is for the thickening. Such a three-parameter filtration was already considered in the work of [KM20] in the context of dynamic metric spaces. A high-level difference to our work is that their focus is on multi-agent systems and their structure at a single time point, while we are interested in interactions of a single 'agent' at disjoint time periods. We think that the three-parameter persistence view on the cycling signature offers several paths for interesting future research. For one, it could provide a suitable setting in which to investigate stability of our statistical invariants with respect to small perturbations of the data. Furthermore, preliminary computations using the two-parameter persistence software [BLL23] for the time parameters with fixed radius suggest that finer invariants can be computed using this approach.

Bibliography

[AA17] Michał Adamaszek and Henry Adams. The vietoris–rips complexes of a circle. *Pacific Journal of Mathematics*, 290(1):1–40, July 2017.

[AAS22] Sajjad Azimi, Omid Ashtari, and Tobias M. Schneider. Constructing periodic orbits of high-dimensional chaotic systems by an adjoint-based variational method. *Physical Review E*, 105(1):014217, January 2022.

[ABP08] S Anastassiou, T Bountis, and Y G Petalas. On the topology of the lü attractor and related systems. *Journal of Physics A: Mathematical and Theoretical*, 41(48):485101, October 2008.

[ACE+87] Ditza Auerbach, Predrag Cvitanović, Jean-Pierre Eckmann, Gemunu Gunaratne, and Itamar Procaccia. Exploring chaotic motion through periodic orbits. *Physical Review Letters*, 58(23):2387–2389, June 1987.

[Bau21] Ulrich Bauer. Ripser: efficient computation of vietoris–rips persistence barcodes. *Journal of Applied and Computational Topology*, 5(3):391–423, June 2021.

[BCB20] Magnus Botnan and William Crawley-Boevey. Decomposition of persistence modules. *Proceedings of the American Mathematical Society*, 148(11):4581–4596, August 2020.

[BCC+12] Oleksiy Busaryev, Sergio Cabello, Chao Chen, Tamal K Dey, and Yusu Wang. Annotating simplices with a homology basis and its applications. In *Scandinavian workshop on algorithm theory*, pages 189–200. Springer, 2012.

[BGH+22] Bea Bleile, Adélie Garin, Teresa Heiss, Kelly Maggs, and Vanessa Robins. *The Persistent Homology of Dual Digital Image Constructions*, pages 1–26. Springer International Publishing, 2022.

[BHJM23] Ulrich Bauer, David Hien, Oliver Junge, and Konstantin Mischaikow. Cycling signatures: Identifying cycling motions in time series using algebraic topology, 2023.

[BKRR23] Ulrich Bauer, Michael Kerber, Fabian Roll, and Alexander Rolle. A unified view on the functorial nerve theorem and its variations. *Expositiones Mathematicae*, 41(4):125503, 2023.

[BLL23] Ulrich Bauer, Fabian Lenzen, and Michael Lesnick. Efficient two-parameter persistence computation via cohomology. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023.

[BM02] Elizabeth Bradley and Ricardo Mantilla. Recurrence plots and unstable periodic orbits. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 12(3):596–600, September 2002.

[BMM12] Marko Budišić, Ryan Mohr, and Igor Mezić. Applied koopmanism. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 22(4), December 2012.

[BPK16] Steven L. Brunton, Joshua L. Proctor, and J. Nathan Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15):3932–3937, March 2016.

[BS14] Peter Bubenik and Jonathan A. Scott. Categorification of persistent homology. *Discrete & Computational Geometry*, 51(3):600–627, January 2014.

[BS23] Ulrich Bauer and Maximilian Schmahl. Lifespan functors and natural dualities in persistent homology. *Homology, Homotopy and Applications*, 25(2):297–327, 2023.

[Car09] Gunnar Carlsson. Topology and data. *Bull. Amer. Math. Soc. (N.S.)*, 46(2):255–308, 2009.

[CB15] William Crawley-Boevey. Decomposition of pointwise finite-dimensional persistence modules. *Journal of Algebra and Its Applications*, 14(05):1550066, March 2015.

[CCBdS16] Frédéric Chazal, William Crawley-Boevey, and Vin de Silva. The observable structure of persistence modules. *Homology, Homotopy and Applications*, 18(2):247–265, 2016.

[CdSGO16] Frédéric Chazal, Vin de Silva, Marc Glisse, and Steve Oudot. *The Structure and Stability of Persistence Modules.* Springer International Publishing, 2016.

[CS24] Sunil Chaudhari and Sanjay Kumar Singh. *A Review on Topological Data Analysis in Time Series*, pages 495–503. Springer Nature Singapore, 2024.

[CSEH06] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. *Discrete & Computational Geometry*, 37(1):103–120, December 2006.

[CSEHM09] David Cohen-Steiner, Herbert Edelsbrunner, John Harer, and Dmitriy Morozov. Persistent homology for kernels, images, and cokernels. In *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*. Society for Industrial and Applied Mathematics, January 2009.

[Cvi91] Predrag Cvitanović. Periodic orbits as the skeleton of classical and quantum chaos. *Physica D: Nonlinear Phenomena*, 51(1–3):138–151, August 1991.

[DFW14] Tamal K. Dey, Fengtao Fan, and Yusu Wang. Computing topological persistence for simplicial maps. In *Proceedings of the thirtieth annual symposium on Computational geometry*, SOCG'14, pages 345–354. ACM, June 2014.

[DJ99] Michael Dellnitz and Oliver Junge. On the approximation of complicated dynamical behavior. *SIAM Journal on Numerical Analysis*, 36(2):491–515, January 1999.

[DMQW12] Sara Dadras, Hamid Reza Momeni, Guoyuan Qi, and Zhong-lin Wang. Four-wing hyperchaotic attractor generated from a new 4d system with one equilibrium and its fractional-order form. *Nonlinear Dynamics*, 67(2):1161–1173, 2012.

[dSG07] Vin de Silva and Robert Ghrist. Coverage in sensor networks via persistent homology. *Algebraic & Geometric Topology*, 7(1):339–358, April 2007.

[EH08] Herbert Edelsbrunner and John Harer. Persistent homology—a survey. In *Surveys on discrete and computational geometry*, volume 453 of *Contemp. Math.*, pages 257–282. Amer. Math. Soc., Providence, RI, 2008.

[EM53] Samuel Eilenberg and Saunders MacLane. Acyclic models. *American Journal of Mathematics*, 75(1):189, January 1953.

[Fen29] Werner Fenchel. über Krümmung und Windung geschlossener Raumkurven. *Math. Ann.*, 101(1):238–252, 1929.

[GH13] John Guckenheimer and Philip Holmes. *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, volume 42. Springer Science & Business Media, 2013.

[Ghr07] Robert Ghrist. Barcodes: The persistent topology of data. *Bulletin of the American Mathematical Society*, 45:61–75, 2007.

[GMK22] İsmail Güzel, Elizabeth Munch, and Firas A. Khasawneh. Detecting bifurcations in dynamical systems with crocker plots. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 32(9), September 2022.

[GT08] Zbigniew Galias and Warwick Tucker. Short periodic orbits for the lorenz system. In *2008 International Conference on Signals and Electronic Systems*, pages 285–288. IEEE, 2008.

[Hat02] Allen Hatcher. *Algebraic topology*. Cambridge University Press, Cambridge, 18. printing edition, 2002.

[Hie24] David Hien. CyclingSignatures.jl. `https://github.com/davidhien/CyclingSignatures.jl`, 2024.

[KLT+16] Miroslav Kramár, Rachel Levanger, Jeffrey Tithof, Balachandra Suri, Mu Xu, Mark Paul, Michael F. Schatz, and Konstantin Mischaikow. Analysis of kolmogorov flow and rayleigh–bénard convection using persistent homology. *Physica D: Nonlinear Phenomena*, 334:82–98, November 2016.

[KM20] Woojin Kim and Facundo Mémoli. Spatiotemporal persistent homology for dynamic metric spaces. *Discrete & Computational Geometry*, 66(3):831–875, February 2020.

[KMM04] T. Kaczynski, K. Mischaikow, and M. Mrozek. *Computational Homology*. Applied Mathematical Sciences. Springer New York, 2004.

[KS04] Holger Kantz and Thomas Schreiber. *Nonlinear time series analysis*, volume 7. Cambridge University Press, 2004.

[LC02] Jinhu Lü and Guanrong Chen. A new chaotic attractor coined. *International Journal of Bifurcation and Chaos*, 12(03):659–661, March 2002.

[Lor63] Edward N. Lorenz. Deterministic nonperiodic flow. *J. Atmos. Sci.*, 20(2):130–141, 1963.

[Mas80] William S. Massey. *Singular Homology Theory*. Springer New York, 1980.

[MB08] Marian Mrozek and Bogdan Batko. Coreduction homology algorithm. *Discrete & Computational Geometry*, 41(1):96–118, April 2008.

[Mor] Dmitriy Morozov. Persistence algorithm takes cubic time in the worst case. BioGeometry News, Department of Computer Science, Duke University, Durham, NC, 2005.

[MS93] Rick Miranda and Emily Stone. The proto-lorenz system. *Physics Letters A*, 178(1–2):105–113, July 1993.

[Mun84] James R. Munkres. *Elements of Algebraic Topology*. Addison Wesley Publishing Company, 1984.

[Oud15] Steve Y. Oudot. *Persistence theory: from quiver representations to data analysis*. Applied mathematics. American Mathematical Society, Providence, Rhode Island, 2015. Angabe zur ungezählten Reihe auf dem Umschlag.

[PH14] Jose A. Perea and John Harer. Sliding windows and persistence: An application of topological methods to signal analysis. *Foundations of Computational Mathematics*, 15(3):799–838, May 2014.

[RB23] Yohai Reani and Omer Bobrowski. Cycle registration in persistent homology with applications in topological bootstrap. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(5):5579–5593, May 2023.

[RS72] Colin P. Rourke and Brian J. Sanderson. *Introduction to Piecewise-Linear Topology*. Springer Berlin Heidelberg, 1972.

[RWS11] V Robins, P J Wood, and A P Sheppard. Theory and algorithms for constructing discrete morse complexes from grayscale digital images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(8):1646–1658, August 2011.

[Rö10] Andreas Rößler. Runge–Kutta methods for the strong approximation of solutions of stochastic differential equations. *SIAM Journal on Numerical Analysis*, 48(3):922–952, January 2010.

[SB02] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis*. Springer New York, 2002.

[Sch10] Peter J. Schmid. Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, 656:5–28, July 2010.

[SPS+23] Nico Stucki, Johannes C. Paetzold, Suprosanna Shit, Bjoern Menze, and Ulrich Bauer. Topologically faithful image segmentation via induced matching of persistence barcodes. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett, editors, *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 32698–32727. PMLR, 23–29 Jul 2023.

[Tsi11] Ch. Tsitouras. Runge–Kutta pairs of order 5(4) satisfying only the first column simplifying assumption. *Computers & Mathematics with Applications*, 62(2):770–775, 2011.

[TZH15] Chad M. Topaz, Lori Ziegelmeier, and Tom Halverson. Topological data analysis of biological aggregation models. *PLOS ONE*, 10(5):e0126383, May 2015.

[Vie27] L. Vietoris. Über den höheren zusammenhang kompakter räume und eine klasse von zusammenhangstreuen abbildungen. *Mathematische Annalen*, 97(1):454–472, December 1927.

[ZS06] J. Zhang and M. Small. Complex network from pseudoperiodic time series: Topology versus dynamics. *Physical Review Letters*, 96(23):238701, June 2006.

[Ču20] Matija Čufar. Ripserer.jl: flexible and efficient persistent homology computation in julia. *Journal of Open Source Software*, 5(54):2614, October 2020.