Topological techniques for inference on periodic functions with phase variation

Wojciech Reise

Dec 6, 2023, Orsay

TopAI chair
ANR-19-CHIA-0001

Data with phase variation

Phase variation

Data with phase variation is of the form

$$S_n = f(\gamma_n) + W_n, \tag{1}$$

where $f : [0,1] \to \mathbb{X}$ and $\gamma_1, \ldots, \gamma_N : [0,1] \to [0,1]$ are increasing homeomorphisms.

Problems

1. Computing a representative of f (Su et al. 2014)

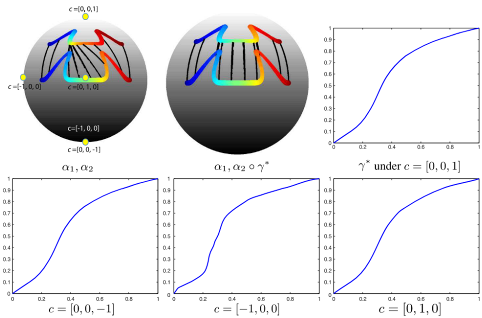


FIG. 2. *Registration of trajectories on* $\mathbb{S}^2$.

When the data is periodic, what can we do?

1. Instantaneous phase estimation
2. Zero-crossings

Data with phase variation

Phase variation

Data with phase variation is of the form

$$S_n = f(\gamma_n) + W_n, \tag{1}$$

where $f : [0,1] \to \mathbb{X}$ and $\gamma_1, \ldots, \gamma_N : [0,1] \to [0,1]$ are increasing homeomorphisms.

Problems

1. Computing a representative of f (Su et al. 2014)
2. Clustering of $S_1, \ldots, S_n$ (Srivastava et al. 2011)

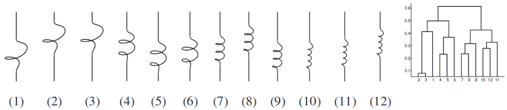


Fig. 4. A set of helices with different numbers and placements of spirals and their clustering using the elastic distance function.

When the data is periodic, what can we do?

1. Instantaneous phase estimation
2. Zero-crossings

Data with phase variation

Phase variation

Data with phase variation is of the form

$$S_n = f(\gamma_n) + W_n, \tag{1}$$

where $f : [0,1] \to \mathbb{X}$ and $\gamma_1, \dots, \gamma_N : [0,1] \to [0,1]$ are increasing homeomorphisms.

Problems

1. Computing a representative of f (Su et al. 2014)
2. Clustering of $S_1, \dots, S_n$ (Srivastava et al. 2011)
3. Synchronising curves: estimating $\gamma_n \circ \gamma_{n'}^{-1}$ (Zhao and Itti 2018)

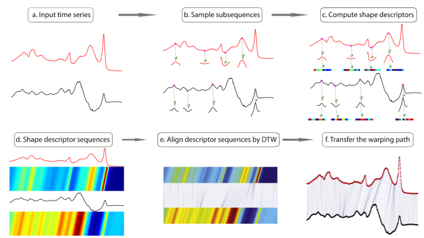


Fig. 2. Pipeline of shapeDTW. shapeDTW consists of two major steps: encode local structures by shape descriptors and align descriptor sequences by DTW. Concretely, we sample a subsequence from each temporal point, and further encode it by some shape descriptor. As a result, the original time series is converted into a descriptor sequence of the same length. Then we align two descriptor sequences by DTW and transfer the found warping path to the original time series.

When the data is periodic, what can we do?

1. Instantaneous phase estimation
2. Zero-crossings

Data with phase variation

Phase variation

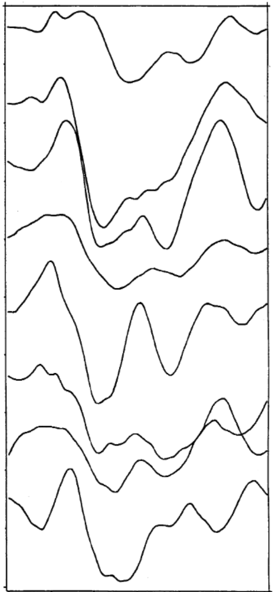
Data with phase variation is of the form

$$S_n = f(\gamma_n) + W_n, \tag{1}$$

where $f : [0,1] \to \mathbb{X}$ and $\gamma_1, \ldots, \gamma_N : [0,1] \to [0,1]$ are increasing homeomorphisms.

Problems

1. Computing a representative of f (Su et al. 2014)
2. Clustering of $S_1, \ldots, S_n$ (Srivastava et al. 2011)
3. Synchronising curves: estimating $\gamma_n \circ \gamma_{n'}^{-1}$ (Zhao and Itti 2018)
4. Instantaneous phase estimation (Boashash 1992)



Topological data analysis for time series

Outline

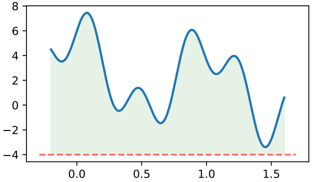
Additivity of persistence diagrams of periodic functions

Segmentation of periodic signals

Signatures of periodic signals with phase variation

Intuition

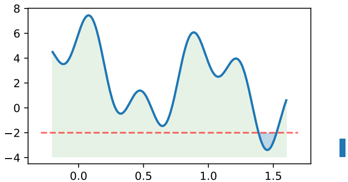
The persistence diagram $D(f)$ of a continuous function $f : [0, T] \to \mathbb{R}$ is a multi-set of points in $\mathbb{R}^2$.



[1] Frédéric Chazal et al. (2016). *The Structure and Stability of Persistence Modules*. SpringerBriefs in Mathematics 2191-8198. Springer, Cham. ISBN: 978-3-319-42543-6.

Intuition

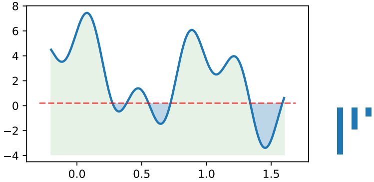
The persistence diagram $D(f)$ of a continuous function $f : [0, T] \to \mathbb{R}$ is a multi-set of points in $\mathbb{R}^2$.



[1] Frédéric Chazal et al. (2016). *The Structure and Stability of Persistence Modules*. SpringerBriefs in Mathematics 2191-8198. Springer, Cham. ISBN: 978-3-319-42543-6.

Intuition

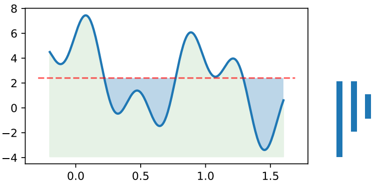
The persistence diagram $D(f)$ of a continuous function $f : [0, T] \to \mathbb{R}$ is a multi-set of points in $\mathbb{R}^2$.



[1] Frédéric Chazal et al. (2016). *The Structure and Stability of Persistence Modules*. SpringerBriefs in Mathematics 2191-8198. Springer, Cham. ISBN: 978-3-319-42543-6.

Intuition

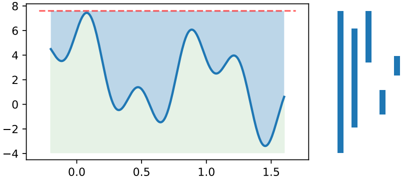
The persistence diagram $D(f)$ of a continuous function $f : [0, T] \to \mathbb{R}$ is a multi-set of points in $\mathbb{R}^2$.



[1] Frédéric Chazal et al. (2016). *The Structure and Stability of Persistence Modules*. SpringerBriefs in Mathematics 2191-8198. Springer, Cham. ISBN: 978-3-319-42543-6.

Intuition

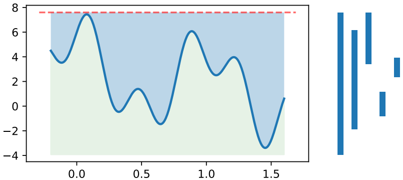
The persistence diagram $D(f)$ of a continuous function $f : [0, T] \to \mathbb{R}$ is a multi-set of points in $\mathbb{R}^2$.

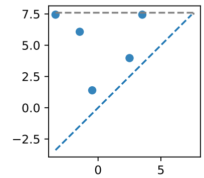


[1] Frédéric Chazal et al. (2016). *The Structure and Stability of Persistence Modules*. SpringerBriefs in Mathematics 2191-8198. Springer, Cham. ISBN: 978-3-319-42543-6.

Intuition

The persistence diagram $D(f)$ of a continuous function $f : [0, T] \to \mathbb{R}$ is a multi-set of points in $\mathbb{R}^2$.





[1] Frédéric Chazal et al. (2016). *The Structure and Stability of Persistence Modules*. SpringerBriefs in Mathematics 2191-8198. Springer, Cham. ISBN: 978-3-319-42543-6.

Intuition

The persistence diagram $D(f)$ of a continuous function $f : [0, T] \to \mathbb{R}$ is a multi-set of points in $\mathbb{R}^2$.

Theorem (Chazal et al. 2016[1])

The persistence diagram of sub level sets of a continuous function $f : \mathbb{X} \to \mathbb{R}$ on a compact topological space $\mathbb{X}$ is well-defined.

[1] Frédéric Chazal et al. (2016). *The Structure and Stability of Persistence Modules.* SpringerBriefs in Mathematics 2191-8198. Springer, Cham. ISBN: 978-3-319-42543-6.

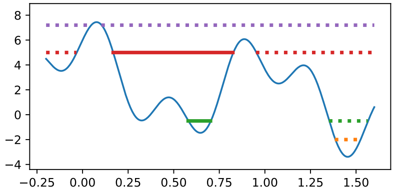
Definition

1. Persistence module

For each $t \in \mathbb{R}$, we calculate $H_0(X_t)$, where $X_t := f^{-1}(]-\infty, t])$. For any $s \leq t$, the inclusion $X_s \to X_t$ gives a map $\iota_s^t : H_0(X_s) \to H_0(X_t)$.

2. Rectangle measure

3. Persistence diagram



Definition

1. Persistence module

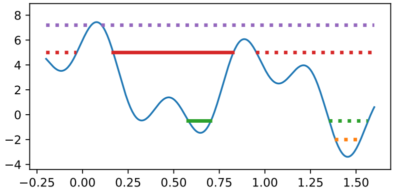
For each $t \in \mathbb{R}$, we calculate $H_0(X_t)$, where $X_t := f^{-1}(]-\infty, t])$. For any $s \leq t$, the inclusion $X_s \to X_t$ gives a map $\iota_s^t : H_0(X_s) \to H_0(X_t)$.

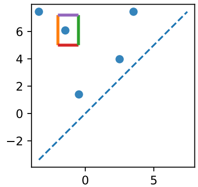
2. Rectangle measure

A measure m on rectangles of $\mathbb{R}^2$.

$$m(]a, b] \times [c, d[) = \dim \left(\frac{im(\iota_b^c) \cap \ker(\iota_c^d)}{im(\iota_a^c) \cap \ker(\iota_c^d)} \right),$$

3. Persistence diagram





Definition

1. Persistence module

For each $t \in \mathbb{R}$, we calculate $H_0(X_t)$, where $X_t := f^{-1}(]-\infty, t])$. For any $s \leq t$, the inclusion $X_s \to X_t$ gives a map $\iota_s^t : H_0(X_s) \to H_0(X_t)$.

2. Rectangle measure

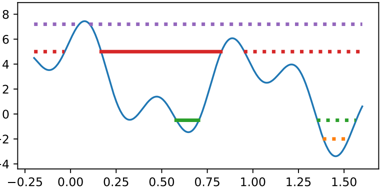
A measure m on rectangles of $\mathbb{R}^2$.

$$m(]a, b] \times [c, d[) = \dim \left(\frac{im(\iota_b^c) \cap \ker(\iota_c^d)}{im(\iota_a^c) \cap \ker(\iota_c^d)} \right),$$

3. Persistence diagram

The persistence diagram $D(f)$ is a multi-set in $\mathbb{R}^2$, where $(s, t) \in \mathbb{R}^2$ has multiplicity

$$m(s, t) = \lim_{\delta \to 0^+} m(]s - \delta, s + \delta] \times [t - \delta, t + \delta[).$$



Stability: bottleneck distance

Definition (Herbert Edelsbrunner and John Harer 2010, p. VIII.2)

We call a ϵ-matching between two persistence diagrams D and D' a bijection $\Gamma : A \to A'$ between some subsets of $A \subset D$ and $A' \subset D'$, considered with multiplicity, if

$$\begin{aligned} d_\infty(a, \Gamma(a)) &\leq \epsilon, \qquad \text{for any } a \in A, \\ d_\infty(a, \Delta) &\leq \epsilon, \qquad \text{for any } a \in (D \setminus A) \cup (D' \setminus A'). \end{aligned}$$

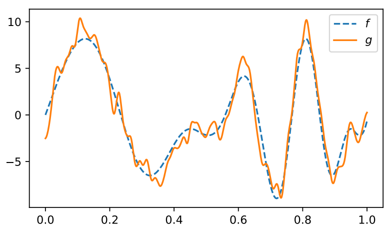
where $\Delta = \{(x, x) \in \mathbb{R}^2\}$ denotes the diagonal.

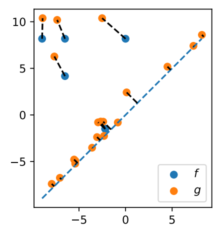
$$d_B(D, D') \coloneqq \inf\{\epsilon > 0 \mid \Gamma \text{ is an } \epsilon\text{-matching between } D' \text{ and } D'\}.$$

Theorem (Bottleneck stability of diagrams)

Let $f, g : \mathbb{X} \to \mathbb{R}$ be two continuous functions on a compact space $\mathbb{X}$. Then,

$$d_B(D(f), D(g)) \leq \|f - g\|_\infty.$$





Total p-persistence

Definition

The total p-persistence of a diagram D is

$$\mathrm{pers}_p(D) := \left(\sum_{(b,d)\in D} (d-b)^p \right).$$

Proposition (Plonka and Zheng 2016, Perez 2022)

For $p = 1$, $\mathrm{pers}_1(D(f)) + \mathrm{pers}_1(D(-f)) = TV(f)$.
If f is α-Hölder for $p > 1 + 1/\alpha$, then, $\mathrm{pers}_p(D(f)) < \infty$.

[2] Olympio Hacquard et al. (2021). "Topologically Penalized Regression on Manifolds". In: *arXiv:2110.13749 [cs, math, stat]*. arXiv: 2110.13749 [cs, math, stat]. (Visited on 01/04/2022), Section 3.2.

Total p-persistence

Definition

The total p-persistence of a diagram D is

$$\mathrm{pers}_p(D) := \left(\sum_{(b,d)\in D} (d-b)^p \right).$$

Proposition (Plonka and Zheng 2016, Perez 2022)

For $p = 1$, $\mathrm{pers}_1(D(f)) + \mathrm{pers}_1(D(-f)) = TV(f)$. (Only dim $= 1$!)[2]
If f is α-Hölder for $p > 1 + 1/\alpha$, then, $\mathrm{pers}_p(D(f)) < \infty$.

[2] Olympio Hacquard et al. (2021). "Topologically Penalized Regression on Manifolds". In: *arXiv:2110.13749 [cs, math, stat]*. arXiv: 2110.13749 [cs, math, stat]. (Visited on 01/04/2022), Section 3.2.

Persistence diagram of sub level sets of periodic functions

Let $\phi : \mathbb{R} \to \mathbb{R}$ be a 1-periodic function and denote by $\phi|_{[a,b]}$ the restriction of ϕ to an interval $[a, b]$.

Theorem (Additivity of persistence diagrams for periodic functions)

For $R > 1$, there exists $c \in [0, 1]$ such that

$$D(\phi|_{[0,R]}) = \lfloor R - 1 \rfloor D(\phi|_{[c,c+1]}) + D', \quad \textit{with} \ \ \mathrm{pers}_p(D') \leq 2\mathrm{pers}_p(D(\phi|_{[c,c+1]})). \tag{2}$$

Persistence diagram of sub level sets of periodic functions

Let $\phi : \mathbb{R} \to \mathbb{R}$ be a 1-periodic function and denote by $\phi|_{[a,b]}$ the restriction of ϕ to an interval $[a, b]$.

Theorem (Additivity of persistence diagrams for periodic functions)

For $R > 1$, there exists $c \in [0, 1]$ such that

$$D(\phi|_{[0,R]}) = \lfloor R - 1 \rfloor D(\phi|_{[c,c+1]}) + D', \quad \textit{with} \ \ \mathrm{pers}_p(D') \leq 2\mathrm{pers}_p(D(\phi|_{[c,c+1]})). \tag{2}$$

If $R \in \mathbb{N}^*$, then

$$D(\phi|_{[0,R]}) = RD(\phi|_{[c,c+1]}). \tag{3}$$

Persistence diagram of sub level sets of periodic functions

Let $\phi : \mathbb{R} \to \mathbb{R}$ be a 1-periodic function and denote by $\phi|_{[a,b]}$ the restriction of ϕ to an interval $[a, b]$.

Theorem (Additivity of persistence diagrams for periodic functions)

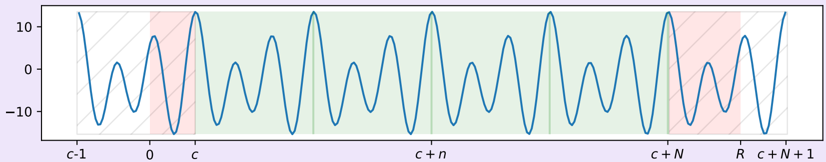
For $R > 1$, there exists $c \in [0, 1]$ such that

$$D(\phi|_{[0,R]}) = \lfloor R - 1 \rfloor D(\phi|_{[c,c+1]}) + D', \quad \textit{with} \ \ \mathrm{pers}_p(D') \leq 2\mathrm{pers}_p(D(\phi|_{[c,c+1]})). \tag{2}$$

If $R \in \mathbb{N}^*$, then

$$D(\phi|_{[0,R]}) = RD(\phi|_{[c,c+1]}). \tag{3}$$

Idea of the proof of (2)



1. For X, Y disjoint topological spaces, $X \cap Y = \emptyset$, we have $H_0(X \cup Y) \simeq H_0(X) \cup H_0(Y)$
2. If $f : X \to Y$ is a homeomorphism, $H_0(Y) \simeq H_0(X)$.

Persistence diagram of sub level sets of periodic functions

Let $\phi : \mathbb{R} \to \mathbb{R}$ be a 1-periodic function and denote by $\phi|_{[a,b]}$ the restriction of ϕ to an interval $[a, b]$.

Theorem (Additivity of persistence diagrams for periodic functions)

For $R > 1$, there exists $c \in [0, 1]$ such that

$$D(\phi|_{[0,R]}) = \lfloor R - 1 \rfloor D(\phi|_{[c,c+1]}) + D', \quad \text{with} \ \ \mathrm{pers}_p(D') \leq 2\mathrm{pers}_p(D(\phi|_{[c,c+1]})). \tag{2}$$

If $R \in \mathbb{N}^*$, then

$$D(\phi|_{[0,R]}) = RD(\phi|_{[c,c+1]}). \tag{3}$$

Corollary

Let $\gamma : [0, T] \to [0, R]$ be bijective and increasing. Let $W : [0, T] \to \mathbb{R}$ be a continuous function. Then,

$$d_B(D(\phi \circ \gamma + W), \lfloor R - 1 \rfloor D(\phi|_{[c,c+1]}) + D') \leq \|W\|_\infty.$$

Conclusion and perspective

Finish; Or even better, move to the end? Establishing better properties would allow us to improve the bounds on both fronts

Conclusion

1. $D(\phi)$ captures information about the value and height of local extrema of ϕ
2. $D(\phi)$ reflects the number of periods of ϕ.

Perspectives

1. Wasserstein stability?
2. Extension to dim ≥ 2?

Plan

Segmenting a periodic curve with phase variation

Based on work with T.Bonis, F.Chazal, B.Michel

Problem

Given $S = (\phi \circ \gamma) + W$, estimate γ.
Setting:

1. ϕ is unknown.
2. The number of periods N is an integer: $\gamma : [0, T] \rightarrow [0, N]$ with $N \in \mathbb{N}$.

Segmenting a periodic curve with phase variation

Based on work with T.Bonis, F.Chazal, B.Michel

Problem

Given $S = (\phi \circ \gamma) + W$, estimate γ.
Setting:

1. ϕ is unknown.
2. The number of periods N is an integer: $\gamma : [0, T] \to [0, N]$ with $N \in \mathbb{N}$.

Solution: odometry

1. Estimate N.
2. Find $t_1 < \ldots < t_N$ such that $\gamma(t_n) - \gamma(t_{n-1}) = 1$ for all $n = 2, \ldots, N$.

Let $\hat{\gamma} : [0, T] \to \mathbb{R}^*$ be such that $\hat{\gamma}(t_n) = n$ and interpolate linearly.

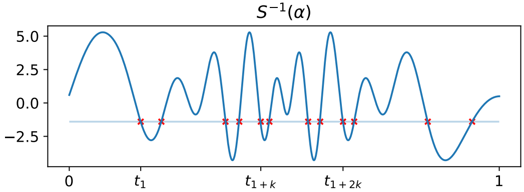
Figure with example (ZC or homology?)

Zero-crossings

Let $K \coloneqq |\phi^{-1}(\alpha) \cap [0,1[|$ and assume that $0 < K < \infty$, for some $\alpha \in \mathbb{R}$.

Estimation of N

If K is known, $N_\alpha(S) \coloneqq \frac{|S^{-1}(\alpha)|}{K}$ is an estimator of N.



Segmentation of the signal

If $S^{-1}(\alpha) = \{t_1, \ldots, t_{NK}\}$, then $\gamma(t_{n+k}) - \gamma(t_n) = 1$ for $1 \leq k \leq K$ and $n \leq NK - k$.

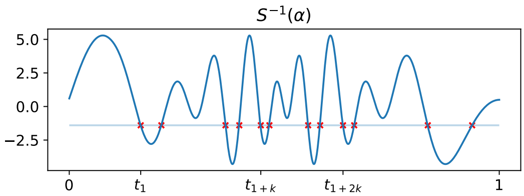
[3] Sunia Tanweer, Firas A. Khasawneh, and Elizabeth Munch (2023). *Robust Zero-crossings Detection in Noisy Signals Using Topological Signal Processing*. arXiv: 2301.07703 [cs, eess]. (Visited on 10/17/2023).

Zero-crossings

Let $K \coloneqq |\phi^{-1}(\alpha) \cap [0,1[|$ and assume that $0 < K < \infty$, for some $\alpha \in \mathbb{R}$.

Estimation of N

If K is known, $N_\alpha(S) \coloneqq \frac{|S^{-1}(\alpha)|}{K}$ is an estimator of N.



Segmentation of the signal

If $S^{-1}(\alpha) = \{t_1, \dots, t_{NK}\}$, then $\gamma(t_{n+k}) - \gamma(t_n) = 1$ for $1 \le k \le K$ and $n \le NK - k$.

Issues

- K is not known (and not necessarily finite)
- N_α is not stable[3]

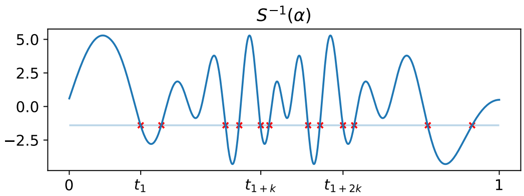
[3] Sunia Tanweer, Firas A. Khasawneh, and Elizabeth Munch (2023). *Robust Zero-crossings Detection in Noisy Signals Using Topological Signal Processing*. arXiv: 2301.07703 [cs, eess]. (Visited on 10/17/2023).

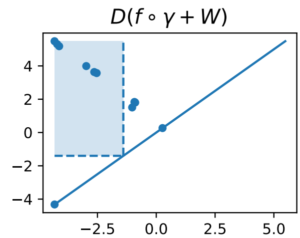
Zero-crossings

Let $K := |\phi^{-1}(\alpha) \cap [0,1[|$ and assume that $0 < K < \infty$, for some $\alpha \in \mathbb{R}$.

Estimation of N

If K is known, $N_\alpha(S) := \frac{|S^{-1}(\alpha)|}{K}$ is an estimator of N.





Segmentation of the signal

If $S^{-1}(\alpha) = \{t_1, \ldots, t_{NK}\}$, then $\gamma(t_{n+k}) - \gamma(t_n) = 1$ for $1 \leq k \leq K$ and $n \leq NK - k$.

Issues

- K is not known (and not necessarily finite)
- N_α is not stable[3]

[3] Sunia Tanweer, Firas A. Khasawneh, and Elizabeth Munch (2023). *Robust Zero-crossings Detection in Noisy Signals Using Topological Signal Processing*. arXiv: 2301.07703 [cs, eess]. (Visited on 10/17/2023).

Estimation of N: stability

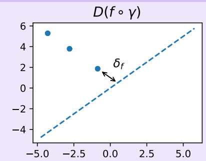
Estimator

For $\tau > 0$, we define

$$\hat{N}_\tau(S) := \gcd\{|D(S) \cap B(x,\tau)| \mid x \in D(S), \mathrm{pers}(x) > \tau\}. \tag{4}$$

Separation constant

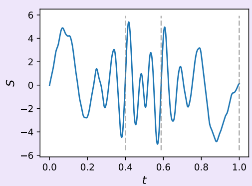
$$\delta_\phi := \min(d(x_1,x_2), d(x_1,\Delta) \mid x_1, x_2 \in D(\phi))$$

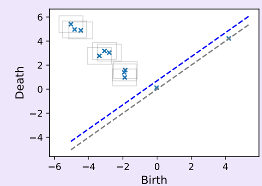


Proposition (Stability)

For $\tau > 0$ satisfying $2\|W\|_\infty < \tau < \delta/3$, we have that

$$\hat{N}_\tau(S) = \hat{N}_\tau(\phi \circ \gamma).$$





Correctness

Non-degeneracy

We say that $\phi|_{[0,1]}$ is **non-degenerate** if $\hat{N}(\phi|_{[0,1]}) = 1$,

$$\hat{N}(\phi|_{[0,1]}) := \gcd \left\{ \lim_{\tau \to 0^+} |D(\phi|_{[0,1]}) \cap B(x, \tau)| \mid x \in D(\phi) \right\}. \tag{5}$$

Sufficient condition

If ϕ has at least one unique critical value, it is non-degenerate.

Corollary (Stability)

If ϕ is non-degenerate, then for any $\tau > 0$ such that $2\|W\|_\infty < \tau < \delta/3$, we have

$$\hat{N}_\tau(S) = N.$$

Correctness

Non-degeneracy

We say that $\phi|_{[0,1]}$ is **non-degenerate** if $\hat{N}(\phi|_{[0,1]}) = 1$,

$$\hat{N}(\phi|_{[0,1]}) := \gcd\left\{ \lim_{\tau \to 0^+} |D(\phi|_{[0,1]}) \cap B(x,\tau)| \mid x \in D(\phi) \right\}. \tag{5}$$

Sufficient condition

If ϕ has at least one unique critical value, it is non-degenerate.

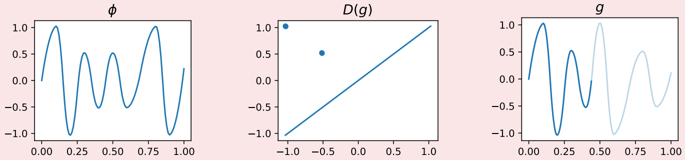
Corollary (Stability)

If ϕ is non-degenerate, then for any $\tau > 0$ such that $2\|W\|_\infty < \tau < \delta/3$, we have

$$\hat{N}_\tau(S) = N.$$

Identifiability with the diagram

There exists a 1-periodic function g such that $D(g|_{[0,1]}) = D(\phi|_{[0,1]})/\hat{N}(\phi|_{[0,1]})$!

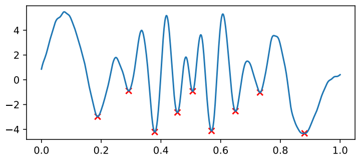


Odometric sequence

Proposition

Let $\tau > 0$ and $\hat{\mathcal{C}}_\tau$ be the set of local minima of S, corresponding to points in the diagram with persistence more than τ. If $\tau \in]2\epsilon, \delta/3[$, then

$$|\hat{\mathcal{C}}_\tau| = NK.$$

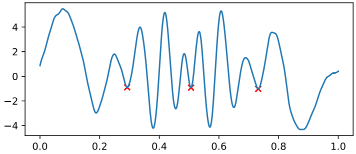


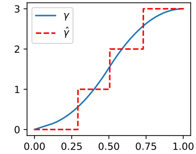
Odometric sequence

Proposition

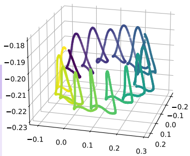
Let $\tau > 0$ and $\hat{\mathcal{C}}_\tau$ be the set of local minima of S, corresponding to points in the diagram with persistence more than τ. If $\tau \in]2\epsilon, \delta/3[$, then

$$|\hat{\mathcal{C}}_\tau| = NK.$$





Application: magnetic odometry



Problem

Using the magnetic signal recorded in a moving car, **B**, estimate the cars' trajectory.

Proposed solution

1. $S := \langle \mathbf{S}, v \rangle$, project **S** along a suitable direction $v \in \S^2$
2. $\hat{N}_\tau(S)$, for an appropriate scale τ,
3. Derive an odometric sequence $t_1, \ldots, t_{\hat{N}_\tau(S)}$ from $\mathcal{C}_\tau$.
4. Construct $\hat{\gamma} : [0, T] \to \mathbb{R}$.

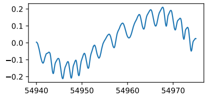
Results

	E_O		E_I	
Method	$\mathbf{S}_{v_1}$	$(\nabla \mathbf{S})_{v_1}$	$\mathbf{S}_{v_1}$	$(\nabla \mathbf{S})_{v_1}$
Hom	15.75	16.66	3.02	3.01
Hom_0	15.75	16.66	3.02	3.01
ZC	9.91	5.62	6.35	16.51

Application: magnetic odometry

Problem

Using the magnetic signal recorded in a moving car, **B**, estimate the cars' trajectory. The angular position $t \mapsto \gamma(t)$ of a wheel in time is visible through $\mathbf{S}(t) = \mathbf{B}(\gamma(t), \gamma_h(t))$



Proposed solution

1. $S := \langle \mathbf{S}, v \rangle$, project **S** along a suitable direction $v \in \S^2$
2. $\hat{N}_\tau(S)$, for an appropriate scale τ,
3. Derive an odometric sequence $t_1, \dots, t_{\hat{N}_\tau(S)}$ from $\mathcal{C}_\tau$.
4. Construct $\hat{\gamma} : [0, T] \to \mathbb{R}$.

Results

	E_O		E_I	
Method	$\mathbf{S}_{v_1}$	$(\nabla \mathbf{S})_{v_1}$	$\mathbf{S}_{v_1}$	$(\nabla \mathbf{S})_{v_1}$
Hom	15.75	16.66	3.02	3.01
Hom_0	15.75	16.66	3.02	3.01
ZC	9.91	5.62	6.35	16.51

Limitations and perspectives

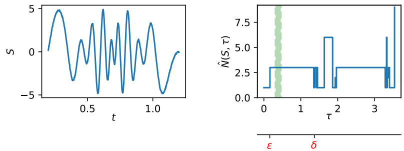
1. Identifiability

2. More robust estimators

3. The method is applicable only to $N \in \mathbb{N}$: boundary effects when $\gamma(1) - \gamma(0) \notin \mathbb{N}$.

Limitations and perspectives

1. Identifiability
   - Use merge trees to verify that the segmentation is correct
2. More robust estimators
   - Extend the guarantees to $\hat{N}_c$ and $\hat{N}_T$
   - Choose the sets A differently



3. The method is applicable only to $N \in \mathbb{N}$: boundary effects when $\gamma(1) - \gamma(0) \notin \mathbb{N}$.

Outline

Problem statement

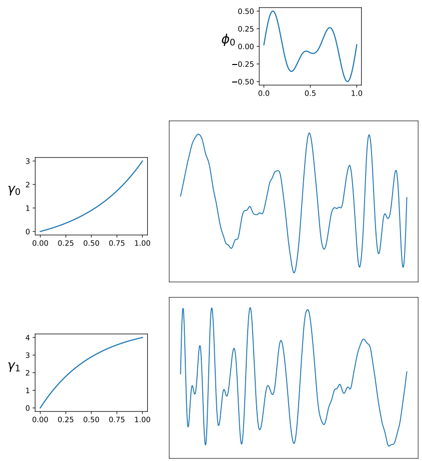
Data

Consider $S = \phi \circ \gamma + W$, where

- $\phi : \mathbb{R} \to \mathbb{R}$ is 1–periodic,
- $\gamma : [0, T] \to \mathbb{R}$ an increasing bijection,
- $W : [0, T] \to \mathbb{R}$ is a cont. stoch. proc.

Aim

Given S, construct a signature of ϕ.



Problem statement

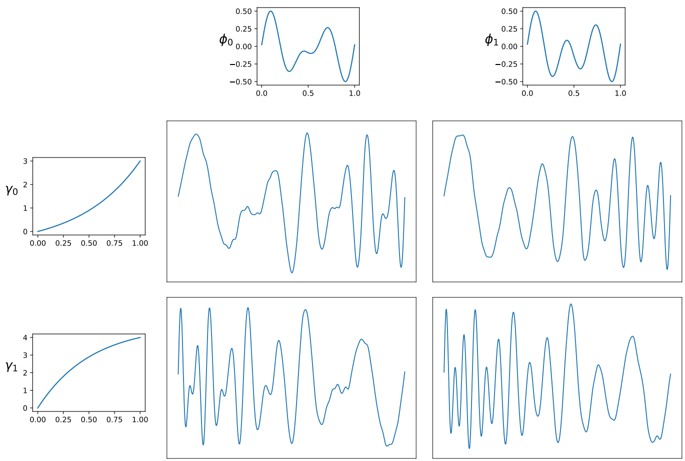
Data

Consider $S = \phi \circ \gamma + W$, where

- $\phi : \mathbb{R} \to \mathbb{R}$ is 1–periodic,
- $\gamma : [0, T] \to \mathbb{R}$ an increasing bijection,
- $W : [0, T] \to \mathbb{R}$ is a cont. stoch. proc.

Aim

Given S, construct a signature of ϕ.



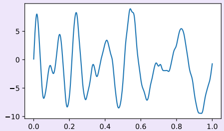
Topological descriptors

Review?

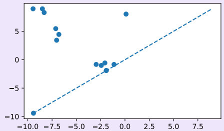
Proposed approach: normalized functionals of persistence

Pipeline

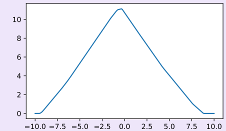
Specify that it is a classic procedure



$\xrightarrow{D}$



$\xrightarrow{\bar{\rho}}$



$$F(S) := \mathbb{E}[\bar{\rho}(D(S))]. \tag{6}$$

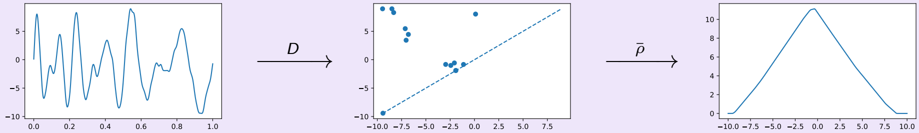
Properties

1. Consistency
2. Stability
3. Estimation

Proposed approach: normalized functionals of persistence

Pipeline

Specify that it is a classic procedure



$$F(S) := \mathbb{E}[\bar\rho(D(S))]. \tag{6}$$

Properties

1. Consistency

$$\bar\rho(D(\phi|_{[0,R]})) \xrightarrow{\|\cdot\|_{\mathcal{H}}} \bar\rho(D(\phi|_{[c,c+1]})), \qquad \text{as } r \to \infty. \tag{7}$$

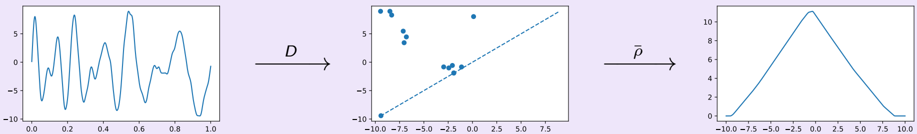
2. Stability

3. Estimation

Proposed approach: normalized functionals of persistence

Pipeline

Specify that it is a classic procedure



$$F(S) := \mathbb{E}[\bar{\rho}(D(S))]. \tag{6}$$

Properties

1. Consistency

$$\bar{\rho}(D(\phi|_{[0,R]})) \xrightarrow{\|\cdot\|_{\mathcal{H}}} \bar{\rho}(D(\phi|_{[c,c+1]})), \qquad \text{as } r \to \infty. \tag{7}$$

2. Stability

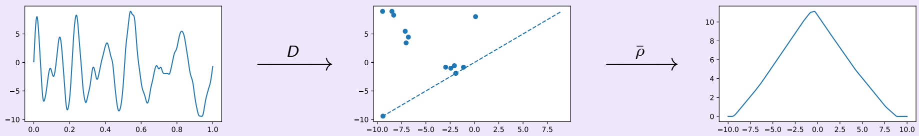
$$\|\mathbb{E}_{\gamma\sim\mu_1,W}[\bar{\rho}(\phi\circ\gamma + W)] - \mathbb{E}_{\gamma\sim\mu_2,W}[\bar{\rho}(\phi\circ\gamma + W)]\|_{\mathcal{H}} \leq CW_1(\mu_1,\mu_2)^{\alpha}. \tag{8}$$

3. Estimation

Proposed approach: normalized functionals of persistence

Pipeline

Specify that it is a classic procedure



$$F(S) := \mathbb{E}[\bar{\rho}(D(S))]. \tag{6}$$

Properties

1. Consistency
$$\bar{\rho}(D(\phi|_{[0,R]})) \xrightarrow{\|\cdot\|_{\mathcal{H}}} \bar{\rho}(D(\phi|_{[c,c+1]})), \qquad \text{as } r \to \infty. \tag{7}$$

2. Stability
$$\|\mathbb{E}_{\gamma\sim\mu_1,W}[\bar{\rho}(\phi\circ\gamma + W)] - \mathbb{E}_{\gamma\sim\mu_2,W}[\bar{\rho}(\phi\circ\gamma + W)]\|_{\mathcal{H}} \leq CW_1(\mu_1,\mu_2)^{\alpha}. \tag{8}$$

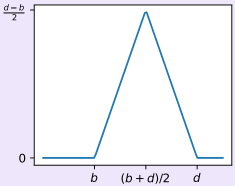
3. Estimation
When γ is generated from a Markov Chain, with $T, R \to \infty$, the empirical mean is a consistent estimator of $\mathbb{E}[\bar{\rho}(\phi\circ\gamma + W)]$.

Normalized functionals of persistence

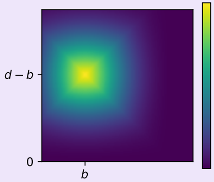
Functional representation

Let $\mathcal{H}$ be a functional Banach space

$$\begin{array}{cccccccc} \kappa: & \mathbb{R}^2 & \to & \mathcal{H} & & & \\ & (b,d) & \mapsto & \kappa_{(b,d)}: & \mathbb{T} & \to & \mathbb{R} \\ & & & & x & \mapsto & \kappa_{(b,d)}(x). \end{array}$$



Persistence silhouette[4]



Persistence image[5]

Normalized functionals of persistence diagrams

For some $p \geq 1$ and $\epsilon > 0$,

$$\bar{\rho}(D) := \frac{\sum_{(b,d)\in D} w(d-b)\kappa_{(b,d)}}{\sum_{(b,d)\in D} w(d-b)}, \qquad \text{where } w(d-b) = \max(d-b-\epsilon, 0)^p. \tag{9}$$

[4] Peter Bubenik (2015). "Statistical Topological Data Analysis Using Persistence Landscapes". In: *Journal of Machine Learning Research* 16.1, pp. 77–102

[5] Henry Adams et al. (2017). "Persistence Images: A Stable Vector Representation of Persistent Homology". In: *The Journal of Machine Learning Research* 18.1, pp. 218–252

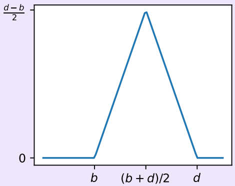
Normalized functionals of persistence

Functional representation

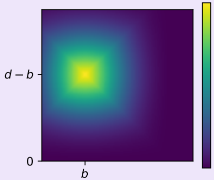
Let $\mathcal{H}$ be a functional Banach space

$$\begin{array}{rccccc} \kappa: & \mathbb{R}^2 & \to & \mathcal{H} & & \\ & (b,d) & \mapsto & \kappa_{(b,d)}: & \mathbb{T} \to & \mathbb{R} \\ & & & & x \mapsto & \kappa_{(b,d)}(x). \end{array}$$

1. $\operatorname{supp}(\kappa_{(b,d)}) \subset K$, K bounded,
2. $x \mapsto \kappa_{(b,d)}(x)$ (uniformly) Lipschitz,
3. $\|\kappa_{(b,d)} - \kappa_{(b',d')}\|_{\mathcal{H}} \leq L_\kappa \|(b,d) - (b',d')\|$,
4. $\|\kappa_{(b,b)}\|_{\mathcal{H}} \leq C$.



Persistence silhouette[4]



Persistence image[5]

Normalized functionals of persistence diagrams

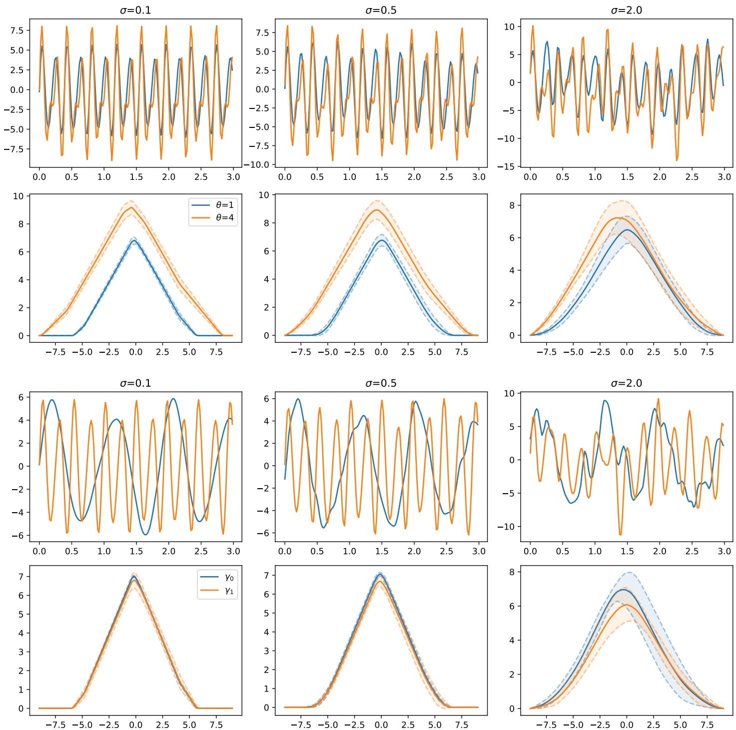
For some $p \geq 1$ and $\epsilon > 0$,

$$\bar{\rho}(D) := \frac{\sum_{(b,d)\in D} w(d-b)\kappa_{(b,d)}}{\sum_{(b,d)\in D} w(d-b)}, \qquad \text{where } w(d-b) = \max(d-b-\epsilon, 0)^p. \tag{9}$$

[4] Peter Bubenik (2015). "Statistical Topological Data Analysis Using Persistence Landscapes". In: *Journal of Machine Learning Research* 16.1, pp. 77–102

[5] Henry Adams et al. (2017). "Persistence Images: A Stable Vector Representation of Persistent Homology". In: *The Journal of Machine Learning Research* 18.1, pp. 218–252

Numerical example



Estimation of signatures: introduction

Can we estimate the signature in practice, where we observe a single time series $(S_n)_{n=1}^N \subset \mathbb{R}$,

$$S_n = \phi(\gamma(t_n)) + W(t_n)?$$

Proposition (Chazal et al. 2014[6], Berry et al. 2018[7])

Let be $D_1, \ldots, D_N$ i.i.d.persistence diagrams. Under assumptions on $(\bar{\rho}_x)_{x\in\mathbb{T}}$,

$$\sqrt{N}\left(\frac{1}{N}\sum_{n=1}^{N}\bar{\rho}(D_n) - \bar{\rho}^*\right) \xrightarrow{d} \mathbb{G},$$

for a zero-mean stochastic process $\mathbb{G}$.

Challenges

- $\bar{\rho}$ is calculated on a window
- $S_1, \ldots, S_N$ are not independent

[6] Frédéric Chazal et al. (2014). "Stochastic Convergence of Persistence Landscapes and Silhouettes". In: *Annual Symposium on Computational Geometry - SOCG'14*. Kyoto, Japan: ACM Press, pp. 474–483. ISBN: 978-1-4503-2594-3. DOI: 10.1145/2582112.2582128. (Visited on 03/05/2021)

[7] Eric Berry et al. (2018). *Functional Summaries of Persistence Diagrams*. arXiv: 1804.01618. (Visited on 03/21/2019)

Estimation of signatures: introduction

Can we estimate the signature in practice, where we observe a single time series $(S_n)_{n=1}^{N} \subset \mathbb{R}$,

$$S_n = \phi(\gamma(t_n)) + W(t_n)?$$

Proposition (Chazal et al. 2014[6], Berry et al. 2018[7])

Let be $D_1, \dots, D_N$ i.i.d.persistence diagrams. Under assumptions on $(\bar{\rho}_x)_{x \in \mathbb{T}}$,

$$\sqrt{N}\left(\frac{1}{N}\sum_{n=1}^{N} \bar{\rho}(D_n) - \bar{\rho}^*\right) \xrightarrow{d} \mathbb{G},$$

for a zero-mean stochastic process $\mathbb{G}$.

Challenges

- $\bar{\rho}$ is calculated on a window
  - $\bar{\rho}(\mathbf{S})$, where $\mathbf{S} := (S_1, \dots, S_M)$ for some $M \in \mathbb{N}$
- $S_1, \dots, S_N$ are not independent
  - Dependence in the window $\mathbf{S}$.
  - Dependence in γ_n and W_n.

[6] Frédéric Chazal et al. (2014). "Stochastic Convergence of Persistence Landscapes and Silhouettes". In: *Annual Symposium on Computational Geometry - SOCG'14*. Kyoto, Japan: ACM Press, pp. 474–483. ISBN: 978-1-4503-2594-3. DOI: 10.1145/2582112.2582128. (Visited on 03/05/2021)

[7] Eric Berry et al. (2018). *Functional Summaries of Persistence Diagrams*. arXiv: 1804.01618. (Visited on 03/21/2019)

Estimation of signatures: procedure

Model

- $\gamma_{n+1} = \gamma_n + hV_n$, where $(V_n)_{n\in\mathbb{N}}$ is a stationary Markov chain supported on $[v_{\min}, v_{\max}] \subset \mathbb{R}^*$,
- $(W_n)_{n\in\mathbb{N}}$ a stationary, real-valued noise process.

Procedure

1. Fix $M \in \mathbb{N}$,
2. Generate a sample $(\mathbf{S}_n)_{n=1}^{N-M-1}$, where $\mathbf{S}_n = (S_n, \ldots, S_{n+M-1})$,
3. Calculate $\hat{F} := \frac{1}{N-M-1}\sum_{n=1}^{N-M-1} \bar{\rho}(\mathbf{S}_n)$.

Theorem

Assume that W is exponentially β-mixing. Then,

$$\sqrt{N-M+1}(\hat{F} - \mathbb{E}[\bar{\rho}(\mathbf{S})]) \to G \tag{10}$$

where G is a zero–mean Gaussian process with covariance

$$(s,t) \mapsto \lim_{k\to\infty} \sum_{n=1}^{\infty} \mathrm{cov}(\bar{\rho}(\mathbf{S}_k)(s), \bar{\rho}(\mathbf{S}_n)(t)).$$

- Adams, Henry et al. (2017). "Persistence Images: A Stable Vector Representation of Persistent Homology". In: *The Journal of Machine Learning Research* 18.1, pp. 218–252.
- Berry, Eric et al. (2018). *Functional Summaries of Persistence Diagrams*. arXiv: 1804.01618. (Visited on 03/21/2019).
- Boashash, B. (1992). "Estimating and Interpreting the Instantaneous Frequency of a Signal. I. Fundamentals". In: *Proceedings of the IEEE* 80.4, pp. 520–538. ISSN: 1558-2256. DOI: 10.1109/5.135376.
- Bubenik, Peter (2015). "Statistical Topological Data Analysis Using Persistence Landscapes". In: *Journal of Machine Learning Research* 16.1, pp. 77–102.
- Chazal, Frédéric et al. (2014). "Stochastic Convergence of Persistence Landscapes and Silhouettes". In: *Annual Symposium on Computational Geometry - SOCG'14*. Kyoto, Japan: ACM Press, pp. 474–483. ISBN: 978-1-4503-2594-3. DOI: 10.1145/2582112.2582128. (Visited on 03/05/2021).
- Chazal, Frédéric et al. (2016). *The Structure and Stability of Persistence Modules*. SpringerBriefs in Mathematics 2191-8198. Springer, Cham. ISBN: 978-3-319-42543-6.
- Hacquard, Olympio et al. (2021). "Topologically Penalized Regression on Manifolds". In: *arXiv:2110.13749 [cs, math, stat]*. arXiv: 2110.13749 [cs, math, stat]. (Visited on 01/04/2022).
- Herbert Edelsbrunner and John Harer (2010). *Computational Topology: An Introduction*. American Mathematical Society. ISBN: 978-0-8218-4925-5.
- Perez, Daniel (2022). *On C0-persistent Homology and Trees*. DOI: 10.48550/arXiv.2012.02634. arXiv: 2012.02634v3.
- Plonka, Gerlind and Yi Zheng (2016). "Relation between Total Variation and Persistence Distance and Its Application in Signal Processing". In: *Advances in Computational Mathematics* 42.3, pp. 651–674. ISSN: 1572-9044. DOI: 10.1007/s10444-015-9438-8. (Visited on 05/05/2021).

Srivastava, A et al. (2011). "Shape Analysis of Elastic Curves in Euclidean Spaces". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33.7, pp. 1415–1428. ISSN: 0162-8828. DOI: 10.1109/TPAMI.2010.184. (Visited on 05/05/2021).

Su, Jingyong et al. (2014). "Statistical Analysis of Trajectories on Riemannian Manifolds: Bird Migration, Hurricane Tracking and Video Surveillance". In: *The Annals of Applied Statistics* 8.1, pp. 530–552. ISSN: 1932-6157. DOI: 10.1214/13-AOAS701. (Visited on 06/13/2020).

Tanweer, Sunia, Firas A. Khasawneh, and Elizabeth Munch (2023). *Robust Zero-crossings Detection in Noisy Signals Using Topological Signal Processing*. arXiv: 2301.07703 [cs, eess]. (Visited on 10/17/2023).

Zhao, Jiaping and Laurent Itti (2018). "shapeDTW: Shape Dynamic Time Warping". In: *Pattern Recognition* 74, pp. 171–184. ISSN: 00313203. DOI: 10.1016/j.patcog.2017.09.020. (Visited on 11/19/2020).

Thank you!

Proof of additivity of sub level sets: details

Proof.

Let $c := \inf\{x \in [0,1[\mid \phi(x) = \max \phi\}$, $N = \max\{n \in \mathbb{N} \mid c + n \leq R\}$ and denote by $\mathbb{X}_t := \phi^{-1}([-\infty, t[)$.
Step 1: For any $t < M$, $\mathbb{X}_t \cap [0, c] \cap [c, c+1] = \emptyset$, so

$$H_0(\mathbb{X}_t \cap [0, R]) \simeq H_0(\mathbb{X}_t \cap [0, c]) \oplus H_0(\mathbb{X}_t \cap [c, c+N]) \oplus H_0(\mathbb{X}_t \cap [c+N, R])), \tag{11}$$

Step 2: similarly,

$$H_0(\mathbb{X}_t \cap [c, c+N]) \simeq \bigsqcup_{n=1}^{N} H_0(\mathbb{X}_t \cap [c+(n-1), c+n]) \tag{12}$$

$$(x \mapsto x+n) \quad \simeq \bigsqcup_{n=1}^{N} H_0(\mathbb{X}_t \cap [c, c+1]) \tag{13}$$

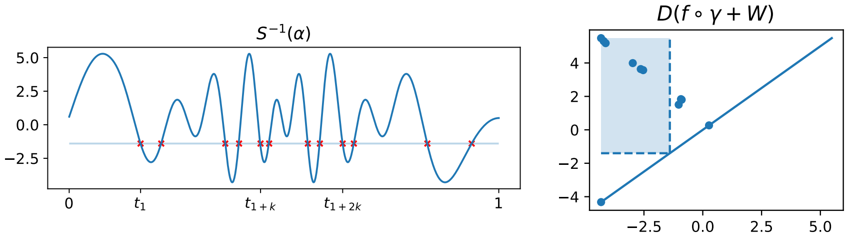
Step 3: The inclusion $[0, c] \subset [c-1, c]$ induces an injective morphism

$$H_0(\mathbb{X}_t \cap [0, c]) \hookrightarrow H_0(\mathbb{X}_t \cap [c-1, c]).$$

□

Zero-crossings from the persistence diagram

$$|S^{-1}(\alpha)| = 2 \lim_{\delta \to 0^+} |D(S) \cap (]-\infty, \alpha - \delta] \times [\alpha + \delta, \infty[)|.$$



Counting measure

The persistence diagram D is also a counting measure on rectangles $A \subset \Delta_+ = \{(b, d) \in \mathbb{R}^2 \mid x < y\}$. By (3),

$$|D(\phi \circ \gamma) \cap A| = N \, |D(\phi|_{[0,1]}) \cap A|$$