A new approach to biparameter persistence based on varying the associated group-equivariant non-expansive operator

Patrizio Frosini

Department of Computer Science, University of Pisa
patrizio.frosini@unipi.it

Topology of Arrangements with an Eye to Applications
September 1-5, 2025, Pisa

Outline

Persistence diagrams

Some basics on the theory of GENEOs

Modifying GENEOs to derive new distances in TDA

Persistence diagrams

Some basics on the theory of GENEOs

Modifying GENEOs to derive new distances in TDA

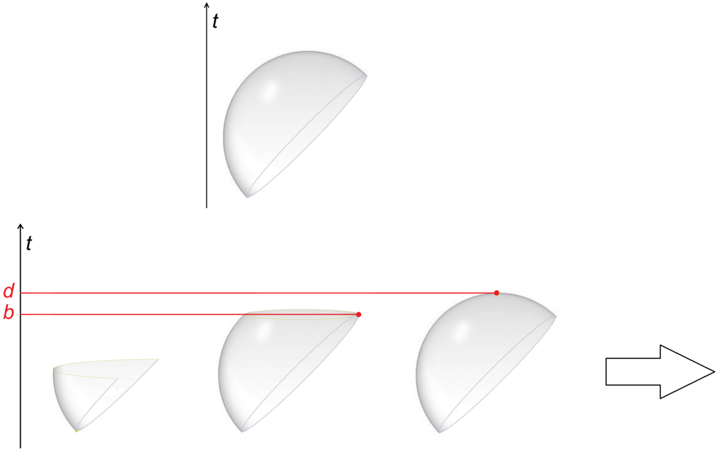
Let us recall what a persistence diagram is

Let $f : X \to \mathbb{R}$ be a continuous function. Let us consider the sublevel sets $X_t := \{x \in X : f(x) \leq t\}$ for t varying in $\mathbb{R}$, where the parameter t is seen as the time.

Informally speaking, the persistence diagram $\mathrm{Dgm}_k(f)$ of f in degree k is the multiset of pairs (b_i, d_i), where b_i and d_i denote the birth and death times of the i-th homological feature (or “hole”) of dimension k in the filtration (X_t) as t increases.

- The points of the persistence diagram are endowed with multiplicity;
- Each point of the diagonal $u = v$ is assumed to be a point of the persistence diagram, endowed with infinite multiplicity;
- We can replace the points on the diagonal with just one point Δ, counted with infinite multiplicity.

An example in degree 1

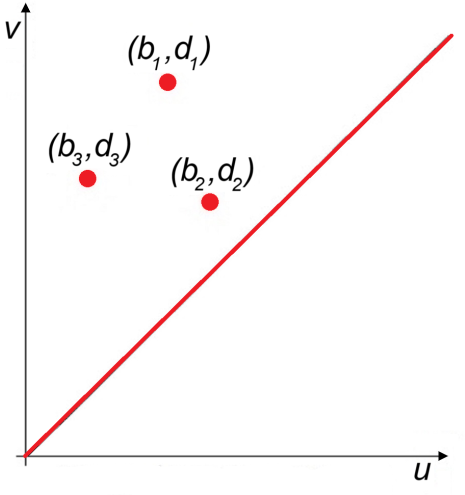


homological event
in degree 1(birth)

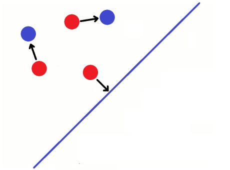
homological event
in degree 1(death)

(b,d)

An example of persistence diagram



The bottleneck distance



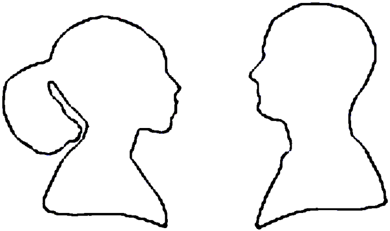
$$\boxed{d_B(D_1, D_2)}$$

One way to transform D_1 into D_2

Persistence diagrams can be compared by means of the bottleneck distance d_B. The bottleneck distance $d_B(D_1, D_2)$ between two persistence diagrams D_1, D_2 is the minimum cost of matchings between the two diagrams. The cost of a matching is the maximum displacement of the points of D_1 to the points of D_2 defined by that matching, where each displacement is given by the sup norm distance in $\mathbb{R}^2$ between the points and their images. Moving a point onto the diagonal means "deleting" it.

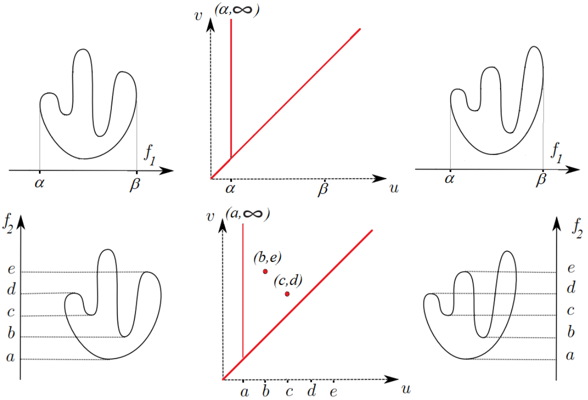
What can we do for vector-valued functions?

In topological data analysis we often have to consider vector-valued functions instead of scalar functions. As an example, we could have to compare two closed curves in $\mathbb{R}^2$, i.e., two functions $f_1, f_2 : S^1 \to \mathbb{R}^2$. How can we extend the concept of persistence diagram to this setting?



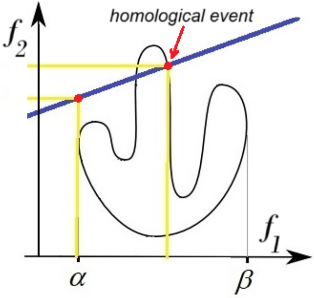
The foliation method

A not very useful idea is to study the two components f_1 and f_2 separately. In this case, the two curves below cannot be distinguished.



The foliation method

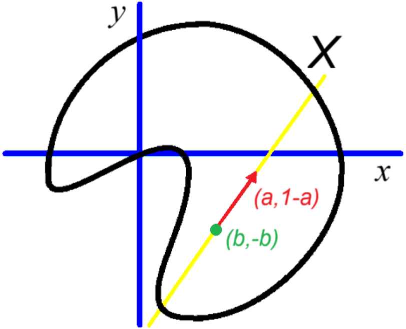
A better idea is to study each filtration associated with a line of positive slope.



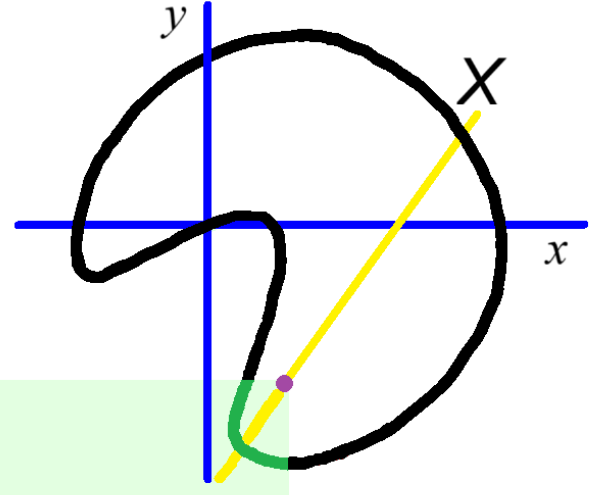
The chosen line allows to distinguish between these curves.

Bifiltrations and positive slope lines

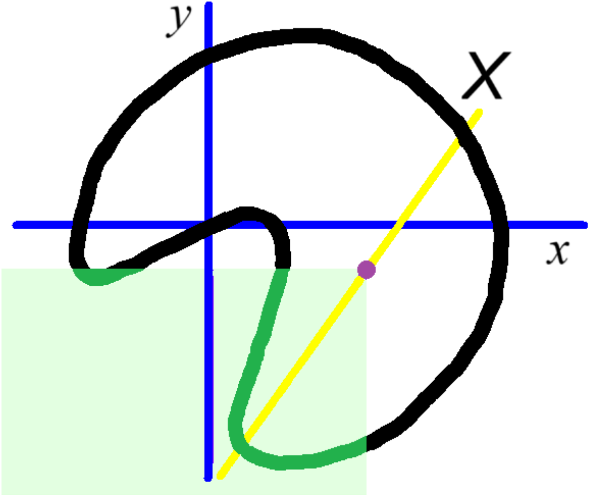
If we have a bifiltration given by a function $\boldsymbol{f} = (f_1, f_2) : X \to \mathbb{R}^2$, we can consider a unit vector (w.r.t. $\|\cdot\|_1$) $w = (a, 1-a)$ with a positive slope, and a point $P = (b, -b)$. Every choice of P and w defines a filtration $\{X_t\}$ of X, where X_t is the set of points of X whose image by $\boldsymbol{f}$ is both under and on the left of the point $P + tw$. As a consequence, each choice of P and w defines a persistence diagram.



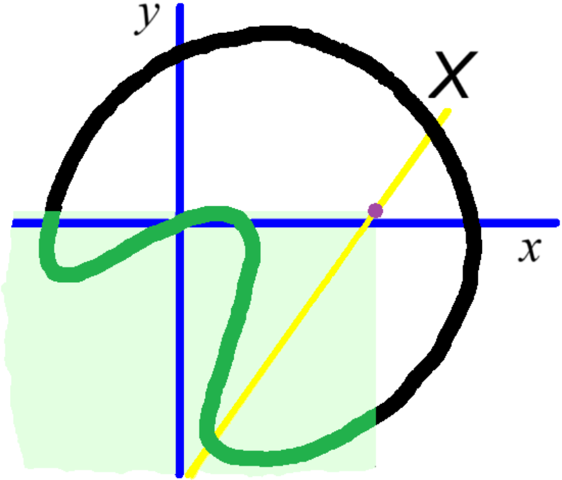
Bifiltrations and positive slope lines



Bifiltrations and positive slope lines



Bifiltrations and positive slope lines



The normalized function $f^*_{(a,b)}$

If we set $(x,y) = P + tw = \big(at + b, (1-a)t - b\big)$ and define the function $f_{(a,b)}(p) := \max\left\{\frac{f_1(p)-b}{a}, \frac{f_2(p)+b}{1-a}\right\}$, then we can write $X_{(x,y)} = \{p \in X : f_1(p) \leq x, f_2(p) \leq y\}$ as the set $X_t = \{p \in X : f_{(a,b)}(p) \leq t\}$.

As a consequence, the bifiltration $\{X_{(x,y)}\}$ of X leads us to consider the persistence diagram $\mathrm{Dgm}_k(f_{(a,b)})$ of the function $f_{(a,b)}$.

In order to get a stability theorem we have to normalize $f_{(a,b)}$ by setting

$$f^*_{(a,b)}(p) := \min\{a, 1-a\} \cdot f_{(a,b)}(p).$$

The persistence diagram $\mathrm{Dgm}_k(f^*_{(a,b)})$ can be obtained by multiplying the persistence diagram $\mathrm{Dgm}_k(f_{(a,b)})$ by the value $\min\{a, 1-a\}$.

Stability of $\mathscr{D}_{\mathrm{match}}$

We can define a 2D matching distance $\mathscr{D}_{\mathrm{match}}(\boldsymbol{f},\boldsymbol{f}')$ by setting $\mathscr{D}_{\mathrm{match}}(\boldsymbol{f},\boldsymbol{f}') := \sup_{(a,b)\in]0,1[\times\mathbb{R}} d_{\mathrm{B}}\left(\mathrm{Dgm}_k(f^*_{(a,b)}),\mathrm{Dgm}_k(f'^*_{(a,b)})\right)$.

Theorem (Stability Theorem for the matching distance)

$\mathscr{D}_{\mathrm{match}}(\boldsymbol{f},\boldsymbol{f}') \leq \|\boldsymbol{f}-\boldsymbol{f}'\|_\infty$.

The distance $\mathscr{D}_{\mathrm{match}}$ has been introduced in the paper *[S. Biasotti, A. Cerri, P. Frosini, D. Giorgi, C. Landi, Multidimensional size functions for shape comparison, Journal of Mathematical Imaging and Vision, vol. 32 (2008), n. 2, 161-179.]*

Some key remarks

- We have seen that biparameter persistence can be compared by a method that is based on these operators $F_{a,b}$:

$$F_{a,b}(f_1, f_2) := \mathrm{Dgm}_k \left(\min\{a, 1-a\} \max\left\{ \frac{f_1(p) - b}{a}, \frac{f_2(p) + b}{1-a} \right\} \right).$$

- The operators $F_{a,b}$ are invariant under reparametrization: if $h : X \to X$ is a homeomorphism, then

$$F_{a,b}(f_1, f_2) = F_{a,b}(f_1 \circ h, f_2 \circ h).$$

- The operators $F_{a,b}$ are non-expansive:

$$d_{\mathrm{B}} \left(F_{a,b}(f_1, f_2), F_{a,b}({f'}_1, {f'}_2) \right) \leq \|(f_1, f_2) - ({f'}_1, {f'}_2)\|_\infty.$$

Operators of this kind fall under the definition of group equivariant non-expansive operators (GENEOs).

Persistence diagrams

Some basics on the theory of GENEOs

Modifying GENEOs to derive new distances in TDA

Let's start by defining perception spaces

We recall that a pseudo-metric is just a metric d without the property $d(x_1, x_2) = 0 \implies x_1 = x_2$.

Definition

Let us consider

1. A nonempty set Φ endowed with a pseudo-metric D_Φ.
2. Let us denote by the symbol $*$ the left action of a group $(G, \circ)$ on Φ, and endow G with the pseudo-metric D_G defined by setting $D_G(g_1, g_2) := \sup_{f \in \Phi} D_\Phi(g_1 * f, g_2 * f)$ for any $g_1, g_2 \in G$. We will also assume that the action of the group G on the metric space (Φ, D_Φ) is isometric, i.e., for every $f_1, f_2 \in \Phi$ and every $g \in G$, $D_\Phi(g * f_1, g * f_2) = D_\Phi(f_1, f_2)$.

We say that (Φ, G) is a perception space.

Perception spaces

The set Φ represents the data we may get from our measuring tools (functions, graphs, cloud of points,...). The group G represents the possible invariances of data the observer may be interested in.
For example, Φ can be a set of grey-level images represented as functions from $\mathbb{R}^2$ to $[0,1]$, while G can be the group of isometries of the real plane.
Another simple example can be given by the set of electrocardiograms represented as functions of the time variable, while G can be the group of time translations.
In any case, the following statement holds.

Proposition

$(G, \circ)$ *is a topological group and the action of* G *on* Φ *is continuous.*

GEOs and GENEOs

Definition

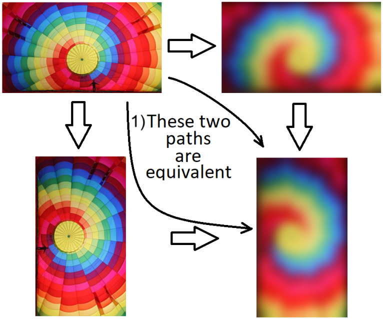
- Let (Φ, G), (Ψ, K) be two perception spaces. If a map $F : \Phi \rightarrow \Psi$ and a group homomorphism $T : G \rightarrow K$ are given, such that $F(g * f) = T(g) * F(f)$ for every $f \in \Phi$, $g \in G$, we say that (F, T) is an (extended) *group equivariant operator* (GEO).
- If (F, T) is non-expansive (i.e. $D_{\Psi}(F(f_1), F(f_2)) \leq D_{\Phi}(f_1, f_2)$ for every $f_1, f_2 \in \Phi$, and $D_K(T(g_1), T(g_2)) \leq D_G(g_1, g_2)$ for every $g_1, g_2 \in G$), we say that (F, T) is an (extended) *group equivariant non-expansive operator* (GENEO).

See *[Jacopo Joy Colombini, Filippo Bonchi, Francesco Giannini, Fosca Giannotti, Roberto Pellungrini and Patrizio Frosini, Mathematical Foundation of Interpretable Equivariant Surrogate Models, World Conference on Explainable Artificial Intelligence (XAI-2025), 09-11 July, 2025 - Istanbul, Turkey.]*

An example of GENEO

When we blur an image by applying a **convolution** with a rotationally symmetric kernel whose mass is less than 1 in L^1, we apply a GENEO:

- $T : \mathrm{Isom}(\mathbb{R}^2) \to \mathrm{Isom}(\mathbb{R}^2)$ is the identity homomorphism;
- $(F, T) : \left(C_c\left(\mathbb{R}^2, [0,1]^3\right), \mathrm{Isom}(\mathbb{R}^2)\right) \to \left(C_c\left(\mathbb{R}^2, [0,1]^3\right), \mathrm{Isom}(\mathbb{R}^2)\right)$ is a GENEO.



1)These two paths are equivalent

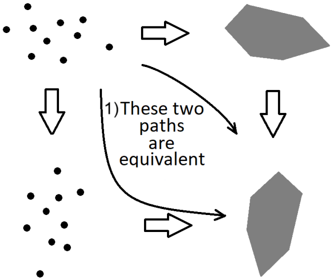
Here, the max-norm distance between functions is used.

2) Blurring does not increase the distance between the images.

Another example of GENEO

When we compute the **convex hull** of a cloud of points, we apply a GENEO:

- $\mathscr{F}$ = the collection of all finite nonempty subsets of the real plane;
- $\mathscr{C}$ = the collection of all convex nonempty subsets of the real plane;
- $T : \mathrm{Isom}(\mathbb{R}^2) \to \mathrm{Isom}(\mathbb{R}^2)$ is the identity homomorphism;
- $(F, T) : (\mathscr{F}, \mathrm{Isom}(\mathbb{R}^2)) \to (\mathscr{C}, \mathrm{Isom}(\mathbb{R}^2))$ is a GENEO.



1)These two paths are equivalent

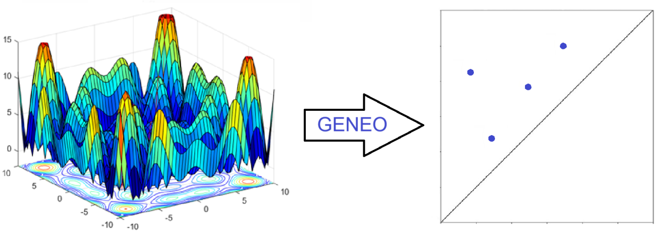
Here, the Hausdorff distance between compact sets is used.

2) The operation of taking the convex hull does not increase the Hausdorff distance between sets.

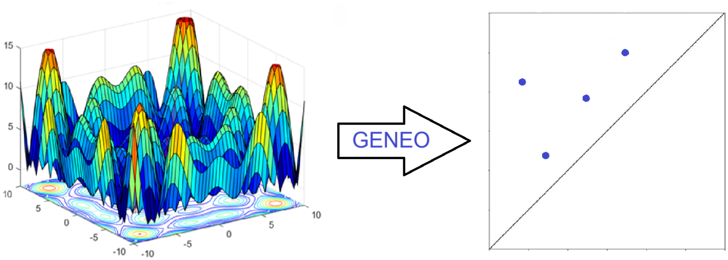
Another example of GENEO

The operator taking each filtering function $f : X \to \mathbb{R}$ to its persistence diagram is another example of GENEO:

- DGM is the metric space of all persistence diagrams of real-valued continuous functions defined on a topological space X;
- $\{\mathrm{id}\}$ is the trivial group acting on DGM, consisting only of the identity map;
- T is the trivial homomorphism from $\mathrm{Homeo}(X)$ to $\{\mathrm{id}\}$;
- $(F, T) : (C_c(X, \mathbb{R}), \mathrm{Homeo}(X)) \to (DGM, \{\mathrm{id}\})$ is a GENEO.



Another example of GENEO



- Equivariance of (F, T) = invariance of persistence diagrams under reparameterization of the domain.
- Nonexpansiveness of (F, T) = stability of persistence diagrams.

GENEOs and Machine Learning

If interested, you can find more details about the theory of GENEOs in these papers:

- M. G. Bergomi, P. Frosini, D. Giorgi, N. Quercioli, *Towards a topological-geometrical theory of group equivariant non-expansive operators for data analysis and machine learning*, **Nature Machine Intelligence**, vol. 1(9) (2019), 423–433.
- G. Bocchi, P. Frosini, M. Ferri, *A novel approach to graph distinction through GENEOs and permutants*, **Scientific Reports**, 15 (2025), 6259.

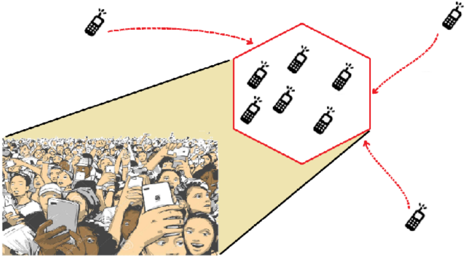
GENEOs and Machine Learning

- A. Micheletti, *A new paradigm for artificial intelligence based on group equivariant non-expansive operators*, **European Mathematical Society Magazine**, 128 (2023), 4-12.
- G. Bocchi, P. Frosini, A. Micheletti, A. Pedretti, G. Palermo, D. Gadioli, C. Gratteri, F. Lunghini, A. D. Biswas, P. F.W. Stouten, A. R. Beccari, A. Fava, C. Talarico, *GENEOnet: A breakthrough in protein binding pocket detection using group equivariant non-expansive operators*, **Scientific Reports**, (2025) (to appear).

Research projects (I)

CNIT / WiLab - Huawei Joint Innovation Center (JIC)

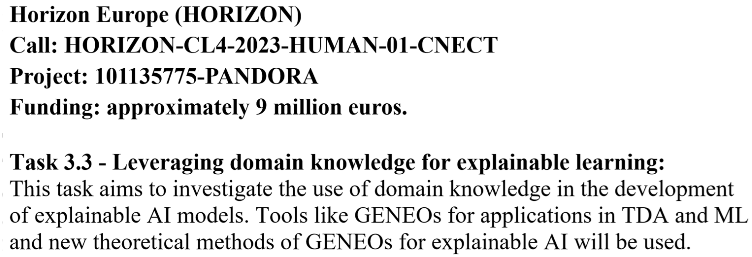
Project on GENEOs for 6G



WILAB HUAWEI

Research projects (II)

PANDORA



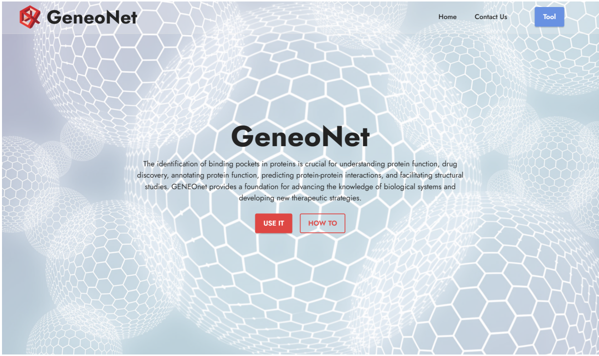
Horizon Europe (HORIZON)
Call: HORIZON-CL4-2023-HUMAN-01-CNECT
Project: 101135775-PANDORA
Funding: approximately 9 million euros.

Task 3.3 - Leveraging domain knowledge for explainable learning: This task aims to investigate the use of domain knowledge in the development of explainable AI models. Tools like GENEOs for applications in TDA and ML and new theoretical methods of GENEOs for explainable AI will be used.

The project has received funding form the
European Union's Horizon Europe
Framework Programme (Horizon) under
grant agreement No 101135775

https://pandora-heu.eu/consortium/

Research projects (III)



GeneoNet

The identification of binding pockets in proteins is crucial for understanding protein function, drug discovery, annotating protein function, predicting protein-protein interactions, and facilitating structural studies. GENEOnet provides a foundation for advancing the knowledge of biological systems and developing new therapeutic strategies.

The GeneoNet webservice represents the outcome of our partnership with Italian Pharmaceutical Company Dompé Farmaceutici S.p.A.: `https://geneonet.exscalate.eu/`

TAKE-AWAY MESSAGE (I)

- GENEOs are useful tools for TDA and Machine Learning;
- The computation of persistence diagrams is a special case of a GENEO;
- The classical matching distance is based on the GENEOs $F_{a,b}$.

Persistence diagrams

Some basics on the theory of GENEOs

Modifying GENEOs to derive new distances in TDA

Contributors to this research

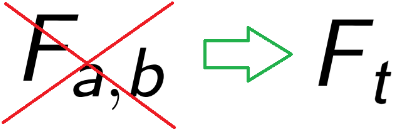
Joint work with:

- Ulderico Fugacci (IMATI-CNR, Genoa, Italy)
- Eloy Mósig García (University of Pisa, Italy)
- Nicola Quercioli (University of Bologna and WiLab-CNIT, Italy)
- Sara Scaramuccia (University of Rome Tor Vergata, Italy)
- Francesca Tombari (University of Oxford, UK)

The GENEOs $F_{a,b}$ have some drawbacks

- They depend on two parameters, which makes the computation non-trivial.
- The map taking (f_1, f_2) to $\min\{a, 1-a\} \cdot \max\left\{\frac{f_1(p)-b}{a}, \frac{f_2(p)+b}{1-a}\right\}$ is not differentiable.

Why not replace them with different GENEOs that depend on a single parameter and are more regular?



Let's consider a new GENEO

For each $t \in [0,1]$, let us consider the operator

$$F_t(f_1, f_2) := \mathrm{Dgm}_k\Big((1-t)f_1 + tf_2\Big).$$

We observe that

$$(F_t, T) : \Big(C(X, \mathbb{R}^2), \mathrm{Homeo}(X)\Big) \to (\mathrm{DGM}, \{\mathrm{id}\})$$

is a GENEO for any $t \in [0,1]$, provided that

- DGM is the metric space of all persistence diagrams of real-valued continuous functions defined on the compact topological space X;
- $\{\mathrm{id}\}$ is the trivial group acting on DGM, consisting only of the identity map;
- $T : \mathrm{Homeo}(M) \to \{\mathrm{id}\}$ is the trivial homomorphism.

Remark

This approach can be easily extended to functions taking values in $\mathbb{R}^n$.

Let's make a comparison

Let us compare $F_{a,b}$ and F_t:

$$\boxed{F_{a,b}(f_1, f_2) := \mathrm{Dgm}_k\left(\min\{a, 1-a\}\max\left\{\tfrac{f_1-b}{a}, \tfrac{f_2+b}{1-a}\right\}\right)}$$

$$\boxed{F_t(f_1, f_2) := \mathrm{Dgm}_k\Big((1-t)f_1 + tf_2\Big)}$$

Remark: Infinitely many other examples of parametric families of GENEOs can be given. For example, the following:

$$\hat{F}_t(f_1, f_2) := \mathrm{Dgm}_k\Big((1-t)\max(f_1, f_2) + t\min(f_1, f_2)\Big) \text{for } t \in [0, 1];$$

$$\tilde{F}_t(f_1, f_2) := \mathrm{Dgm}_k\left(\left(\frac{1}{2}|f_1|^t + \frac{1}{2}|f_2|^t\right)^{\frac{1}{t}}\right) \text{ for } t \geq 1.$$

Why prefer one GENEO over another?

Naturally, the choice to replace $F_{a,b}$ with F_t has both advantages and disadvantages:

- F_t can be computed more easily than $F_{a,b}$;
- F_t is based on a smooth function, whereas $F_{a,b}$ is not;

however:

- $F_{a,b}(f_1, f_2)$ contains more information than $F_t(f_1, f_2)$ about the pair of functions (f_1, f_2).

Consequently, the choice between $F_{a,b}$ and F_t (or other GENEOs) is subjective and depends on the application at hand. In this talk, we aim only to illustrate some of the possibilities offered by choosing F_t.

The simplicial matching distance

Just as the operator $F_{a,b}$ allows the introduction of the classical matching distance, the operator F_t likewise allows the introduction of a new pseudo-metric.

Assume that X is a compact and locally path-connected metric space. If $\boldsymbol{f} = (f_1, f_2) : X \to \mathbb{R}^2$, let us set $\boldsymbol{f}^t := (1-t)f_1 + tf_2$ for any $t \in [0,1]$. If $\boldsymbol{f}, \boldsymbol{f}' : X \to \mathbb{R}^2$ are two continuous functions, let us define

$$\boxed{D_s(\boldsymbol{f}, \boldsymbol{f}') := \max_{t\in[0,1]} d_{\mathrm{B}}\left(\mathrm{Dgm}_k(\boldsymbol{f}^t), \mathrm{Dgm}_k(\boldsymbol{f}'^t)\right).}$$

The distance D_s is called the *simplicial matching distance*.

Proposition

D_s is a stable pseudo-metric (i.e., $D_s(\boldsymbol{f}, \boldsymbol{f}') \leq \|\boldsymbol{f} - \boldsymbol{f}'\|_\infty$).

A related concept

The *persistent homology transform* (PHT) is a topological transform which takes as input a subset of a Euclidean space, and to each unit vector assigns the persistence module of the height function over that subset with respect to that direction. A distance between two subsets is defined by integrating over the sphere the distance between the respective extended persistence modules.

- K. Turner, S. Mukherjee, and D. M. Boyer, *Persistent homology transform for modeling shapes and surfaces*, Inf. Inference 3 (2014), no. 4, 310–344.
- K. Turner, V. Robins, and J. Morgan, *The extended persistent homology transform of manifolds with boundary*, J Appl. and Comput. Topology 8, 2111–2154 (2024).

Differences between these two approaches for $n = 2$

PHT

1. PHT acts on a subset X of the real plane;
2. Knowing the coordinate functions from X to $\mathbb{R}^2$ is equivalent to knowing X;
3. PHT considers unit vectors with respect to the Euclidean metric.

GENEOs

1. The GENEO F_t acts on a vector-valued function $f = (f_1, f_2) : X \to \mathbb{R}^2$;
2. Knowing the input f of the GENEO F_t does not allow reconstruction of X;
3. The GENEO F_t is based on convex combinations, i.e., unit vectors with respect to the L^1-norm, with non-negative components.

The last one is a key point: it implies non-expansiveness and allows us to use the structure illustrated in the next slides: the **Pareto grid**.

Some technical assumptions

To define the Pareto grid, we need some technical assumptions.

First, we assume that X is a closed smooth manifold M of dimension $r \geq 2$. Then, we assume that $\boldsymbol{f} = (f_1, f_2)$ is a smooth map from M to the real plane $\mathbb{R}^2$. We choose a Riemannian metric on M so that we can define gradients for f_1 and f_2.

The Jacobi set $\mathbb{J}(\boldsymbol{f})$ is the set of all points $p \in M$ at which the gradients of f_1 and f_2 are linearly dependent.

If $p \in \mathbb{J}(\boldsymbol{f})$ and $\nabla f_1(p) \cdot \nabla f_2(p) \leq 0$, we say that the point p is a critical Pareto point for $\boldsymbol{f}$. The set of all critical Pareto points of $\boldsymbol{f}$ is denoted by $\mathbb{J}_P(\boldsymbol{f})$.

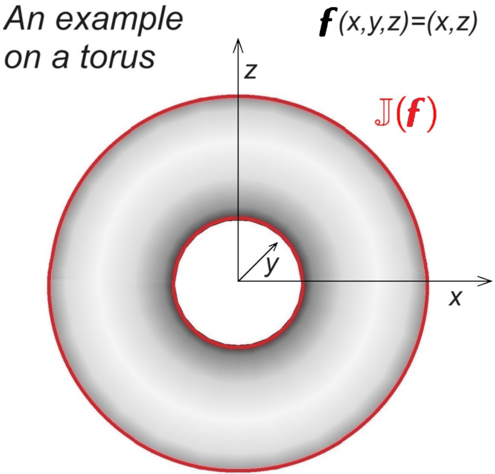
Some technical assumptions

If we assume that $\boldsymbol{f} : M \to \mathbb{R}^2$ is regular enough in a suitable sense (here we skip the technical details), then the Jacobi set is a smooth 1-submanifold of M, consisting of finitely many components, each one diffeomorphic to a circle.

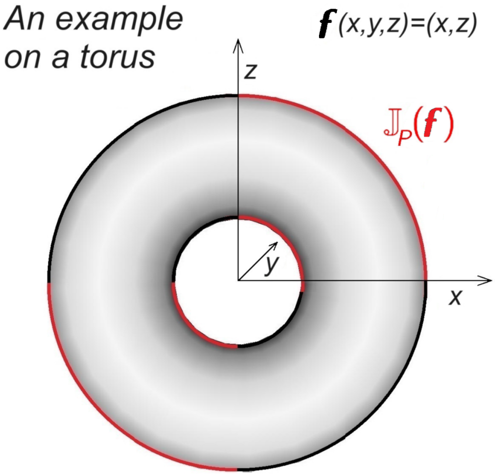
Furthermore, the set of critical Pareto points at which the gradients of f_1 and f_2 are not orthogonal to the Jacobi set is made of a finite family $\{\alpha^i\}$ of arcs. Along these arcs, one of f_1 and f_2 is strictly increasing and the other is strictly decreasing. Each arc α^i can meet critical points for f_1, f_2 only at its endpoints.

For more details: *[Y.H. Wan, Morse theory for two functions, Topology 14 (1975), no. 3, 217-228.]*

The Jacobi set



The set of critical Pareto points



The Pareto grid $\Gamma(\boldsymbol{f})$

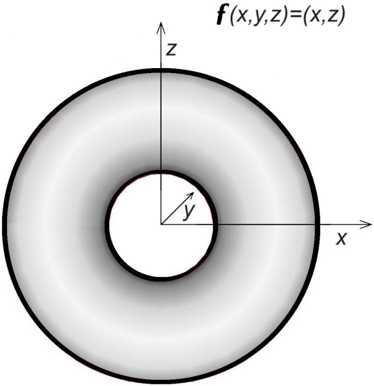
Our goal is to establish a formal connection between the position of the points in $\mathrm{Dgm}_k(\boldsymbol{f}^t)$ and the orthogonal intersections of all lines with direction $(1-t, t)$ with a particular subset of the plane $\mathbb{R}^2$, referred to as the Pareto grid of $\boldsymbol{f}$.

Definition

The Pareto grid $\Gamma(\boldsymbol{f})$ of $\boldsymbol{f}$ is the image under $\boldsymbol{f}$ of the set $\mathbb{J}_P(\boldsymbol{f})$ of all critical Pareto points.

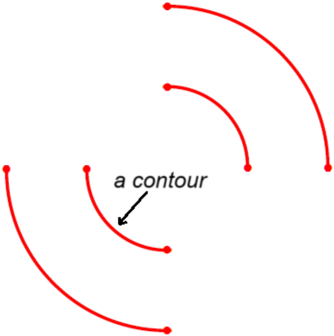
The concept of a Pareto grid is closely related to the concept of the extended Pareto grid introduced in *[A. Cerri, M. Ethier, P. Frosini, On the geometrical properties of the coherent matching distance in 2D persistent homology, Journal of Applied and Computational Topology, vol. 3 (2019), n. 4, 381–422.]*

The Pareto grid: An example



The torus endowed with the filtering function $\boldsymbol{f}(p) := (x(p), z(p))$.

The Pareto grid: An example



The Pareto grid for the torus endowed with the filtering function $\boldsymbol{f}(p) := (x(p), z(p))$. The closures of the images of the previously cited arcs α^i will be called contours of $\boldsymbol{f}$.

Contour-arcs

We can endow the points of $\Gamma(\boldsymbol{f})$ with a suitable concept of multiplicity.

Let $\mathscr{D}(\boldsymbol{f})$ be the set of double points in $\Gamma(\boldsymbol{f})$. Each connected component of $\Gamma(\boldsymbol{f}) \setminus \mathscr{D}(\boldsymbol{f})$ is called a contour-arc of $\boldsymbol{f}$.

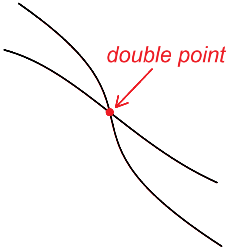


Figure: Two contours intersect and form four contour arcs.

The Position Theorem

Theorem (Position Theorem per $\boldsymbol{f}^t = (1-t)f_1 + tf_2$)

Let w be a finite coordinate of a point in $\mathrm{Dgm}_k(\boldsymbol{f}^t)$. There exist a contour-arc $\alpha : [0,1] \to \mathbb{R}^2$ and a $\bar{\tau} \in [0,1]$ such that

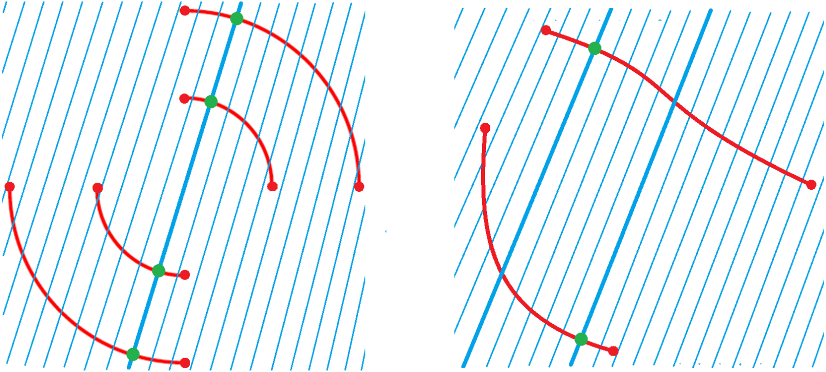
1. $\frac{d\alpha}{d\tau}(\bar{\tau}) \cdot (1-t,t) = 0$
2. $\alpha(\bar{\tau}) \cdot (1-t,t) = w$.

At the extrema of $[0,1]$ we consider the left and right derivative. The symbol $\cdot$ denotes the inner product.

The Position Theorem allows us to compute the possible values for the coordinates of the points in $\mathrm{Dgm}_k(\boldsymbol{f}^t)$.

The Position Theorem

We conjecture that, for any fixed $t \in [0,1]$, the set of points at which lines of direction $(1-t,t)$ intersect the Pareto grid **orthogonally** is finite, provided that the function $\boldsymbol{f}$ is generic.



Special values of t

Definition

Let $\mathrm{CA}(\boldsymbol{f},\boldsymbol{f}')$ be the set of all contour-arcs of $\boldsymbol{f}$ and $\boldsymbol{f}'$. The *special set* of $(\boldsymbol{f},\boldsymbol{f}')$, denoted by $\mathrm{Sp}(\boldsymbol{f},\boldsymbol{f}')$, is the union of $\{0,1\}$ with the set of all $t \in]0,1[$ for which at least one of the following properties holds:

1. There exist four lines (possibly not distinct) r_1, r_2, r_3, r_4 with direction $(1-t,t)$ and four contour-arcs (possibly not distinct) $\alpha^1, \alpha^2, \alpha^3, \alpha^4 \in \mathrm{CA}(\boldsymbol{f},\boldsymbol{f}')$, such that r_i intersects orthogonally α^i at a point P_i for each $i \in \{1,2,3,4\}$, $(P_1,P_2) \neq (P_3,P_4)$, and the two values $|(P_1-P_2)\cdot(1-t,t)|$, $|(P_3-P_4)\cdot(1-t,t)|$ are either equal or one is half the other.

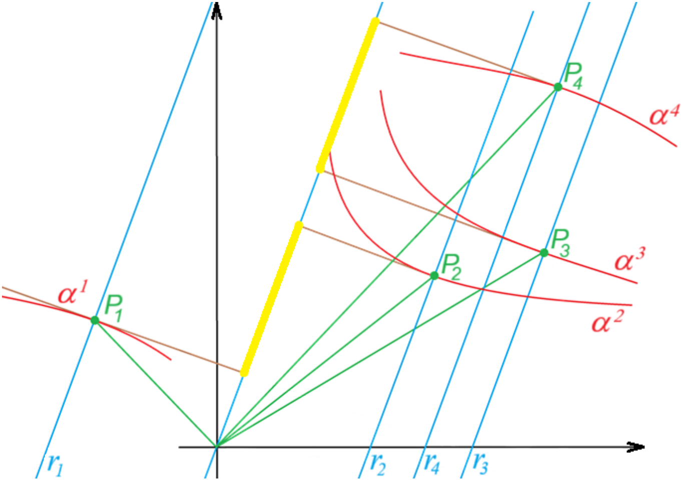
Special values of t

Definition (continued)

2. There exist two lines r_1, r_2 with direction $(1-t,t)$ and two contour-arcs $\alpha^1, \alpha^2 \in \mathrm{CA}(\boldsymbol{f}, \boldsymbol{f}')$, such that r_i intersects orthogonally α^i at the point P_i for $i \in \{1,2\}$, $P_1 \neq P_2$, and— denoting respectively by ρ^i and (x^i, y^i) the radius and the center of the osculating circle of α^i at P_i— we have either $\rho^1 = \rho^2$, or $\rho^1 = \infty$, or $\rho^2 = \infty$, or $t = \frac{(\rho^1-\rho^2)-(y^1-y^2)+(x^1-x^2)}{2(\rho^1-\rho^2)}$.

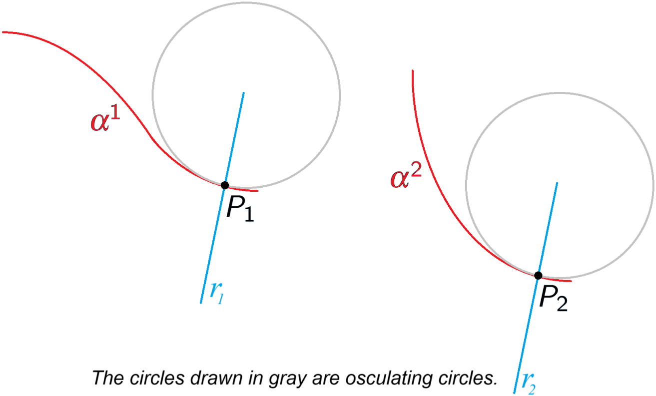
Special values of t

Example of case 1): The two yellow segments have the same length.



Special values of t

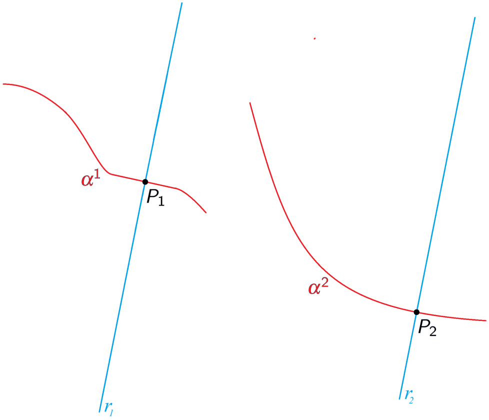
Example of case 2): The curvature of α^1 at P_1 equals the curvature of α^2 at P_2.



The circles drawn in gray are osculating circles.

Special values of t

Another example of case 2): The curvature of α^1 at P_1 vanishes (i.e., $\rho^1 = \infty$).



Our main result

The use of the Pareto grid allows us to prove the following result:

Theorem

If the simplicial matching distance $D_s(\boldsymbol{f}, \boldsymbol{f}')$ is realized at $\bar{t}$, i.e.,

$$D_s(\boldsymbol{f}, \boldsymbol{f}') = d_{\mathrm{B}}\left(\mathrm{Dgm}_k(\boldsymbol{f}^{\bar{t}}), \mathrm{Dgm}_k(\boldsymbol{f}'^{\bar{t}})\right),$$

then $\bar{t}$ belongs to the special set of $(\boldsymbol{f}, \boldsymbol{f}')$.

It is expected that the special set is finite; consequently, the previous result should significantly reduce the complexity of computing the simplicial matching distance.

Future research

Several lines of research concerning the matching distance have corresponding counterparts for the simplicial matching distance

$$D_s\left(f, f'\right) := \max_{t\in[0,1]} d_{\mathrm{B}}\left(\mathrm{Dgm}_k(\boldsymbol{f}^t), \mathrm{Dgm}_k(\boldsymbol{f}'^t)\right).$$

In particular, it would be interesting to develop algorithms to efficiently compute the pseudo-metric D_s and to assess its practical usefulness in applications.

TAKE-AWAY MESSAGE (II)

In this talk, we introduced a new distance for biparameter persistence, called the simplicial matching distance, since it is based on the simplex generated by the components of the vector-valued filtering function.

We also presented:

- The concept of the Pareto grid;
- The Position Theorem for the simplicial matching distance;
- A theorem stating that the simplicial matching distance is always realized at special values of the variable t.

Thanks for your attention