Some advances in G-invariant persistent topology and homology

Patrizio Frosini[1,2]

[1]Department of Mathematics, University of Bologna, Italy
[2]ARCES - Vision Mathematics Group, University of Bologna, Italy
patrizio.frosini@unibo.it

ACA 2013 - Special Session on Computer algebra in algebraic topology and its applications

Outline

1. **The limitations of classical Persistent Homology**
2. **G-invariant persistent homology via quotient spaces**
3. **G-invariant persistent homology via G-functionals**

1. **The limitations of classical Persistent Homology**
2. **G-invariant persistent homology via quotient spaces**
3. **G-invariant persistent homology via G-functionals**

The point of this talk

- It is well known that classical persistent homology is invariant under the action of the group $\mathrm{Homeo}(X)$ of all self-homeomorphisms of a topological space X. As a consequence, this theory is not able to distinguish two filtering functions $\varphi, \psi : X \to \mathbb{R}$ if a homeomorphism $h : X \to X$ exists, such that $\psi = \varphi \circ h$.
- However, in several applications the existence of a homeomorphism $h : X \to X$ such that $\psi = \varphi \circ h$ is not sufficient to consider φ and ψ equivalent to each other.
- How can we adapt the concept of persistence in order to get invariance just under the action of a proper subgroup of $\mathrm{Homeo}(X)$ rather than under the action of the whole group $\mathrm{Homeo}(X)$?

Example

These data are equivalent for classical Persistent Homology

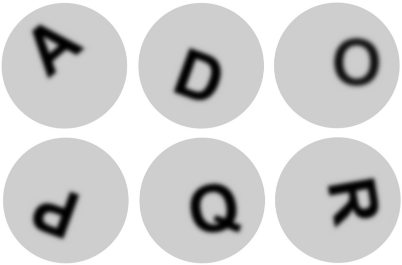


Figure: Examples of letters A, D, O, P, Q, R represented by functions $\varphi_A, \varphi_D, \varphi_O, \varphi_P, \varphi_Q, \varphi_R$ from the unit disk $D \subset \mathbb{R}^2$ to the real numbers. Each function $\varphi_Y : D \to \mathbb{R}$ describes the grey level at each point of the topological space D, with reference to the considered instance of the letter Y. Black and white correspond to the values 0 and 1, respectively (so that light grey corresponds to a value close to 1).

Example (continuation)

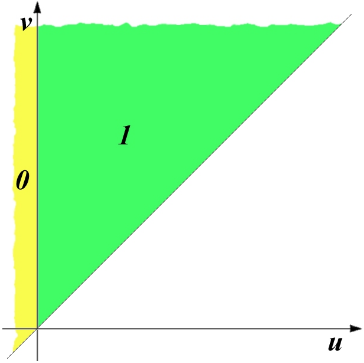


Figure: The persistent Betti number function in degree 0 for all images in the previous figure ("letters *A*, *D*, *O*, *P*, *Q*, *R*").

Example (continuation)

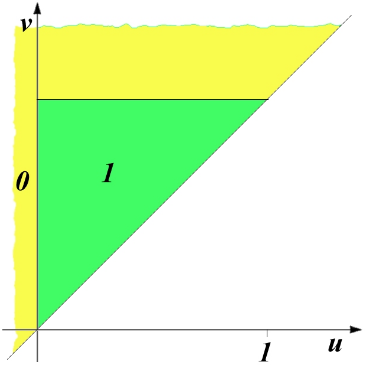
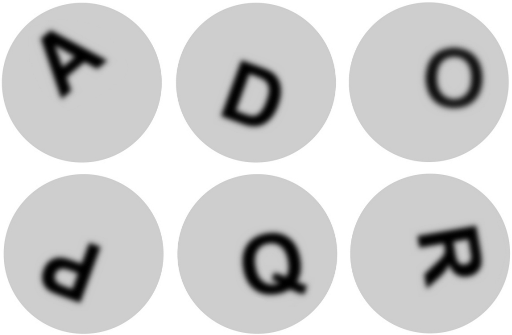


Figure: The persistent Betti number function in degree 1 for all images in the previous figure ("letters *A*, *D*, *O*, *P*, *Q*, *R*").

Example (continuation)

In our example classical persistent homology fails because it is invariant under the action of homeomorphisms, and our six images are equivalent up to homeomorphisms.



The main point

It follows that

Classical persistent homology is not tailored to study invariance with respect to a group G different from the group of all self-homeomorphisms of a topological space.

In this talk we will show two ways to adapt classical persistent homology to the group G.

Observation

One could think of solving the problem described in the previous example by using other filtering functions, possibly defined on different topological spaces. For example, we could extract the boundaries of our letters and consider the distance from the center of mass of each boundary as a new filtering function. This approach presents some problems:

1. It usually requires an extra computational cost (e.g., to extract the boundaries of the letters in our previous example).
2. It can produce a different topological space for each new filtering functions (e.g., the letters of the alphabet can have non-homeomorphic boundaries). Working with several topological spaces instead of just one can be a disadvantage.
3. It is not clear how we can translate the invariance we need into the choice of new filtering functions defined on new topological spaces.

Before proceeding we need a *"ground truth"*.

In this talk, our ground truth will be the *natural pseudo-distance*.

Definition (Natural pseudo-distance)

Let X be a topological space. Let G be a subgroup of the group $\mathrm{Homeo}(X)$ of all self-homeomorphisms $f : X \to X$. Let S be a subset of the set $C^0(X, \mathbb{R})$ of all continuous functions from X to $\mathbb{R}$. The pseudo-distance $d_G : S \times S \to \mathbb{R}$ defined by setting

$$d_G(\varphi_1, \varphi_2) = \inf_{g \in G} \|\varphi_1 - \varphi_2 \circ g\|_\infty$$

is called the natural pseudo-distance associated with the group G.

We point out that d_G is G-invariant: $d_G(\varphi_1, \varphi_2 \circ g) = d_G(\varphi_1, \varphi_2)$ for any $g \in G$, for every $\varphi_1, \varphi_2 \in C^0(X, \mathbb{R})$.

The rationale of using the natural pseudo-distance d_G

The rationale of using the natural pseudo-distance d_G consists in considering two shapes σ_1 and σ_2 equivalent to each other if a transformation exists in the group G, taking the measurements on σ_1 to the measurements on σ_2.

Example: Two gray-level pictures can be considered equivalent if a gray-level-preserving rigid motion exists, taking one picture into the other.

Remark

We need to apply persistent homology in a way that is invariant under the action of a given subgroup G of the group $\mathrm{Homeo}(X)$.

We could think of using the well known concept of Equivariant Homology. In other words, in the case that G acts freely on X, one could think of considering the topological quotient space X/G, endowed with the filtering functions $\hat{\varphi}, \hat{\psi}$ that take each orbit ω of the group G to the maximum of φ and ψ on ω, respectively.

We observe that this approach would not be of help in the case that the action of the group G is transitive (such as in the "letters example"), since the quotient of X/G is just a singleton. As a consequence, if we considered two filtering functions $\varphi, \psi : X \to \mathbb{R}$ with $\max \varphi = \max \psi$, the persistent homology of the induced functions $\hat{\varphi}, \hat{\psi} : X/G \to \mathbb{R}$ would be the same.

Therefore, we need to use a different procedure.

Our approach

Our approach is based on the choice of a suitable group H associated with the group G.

- We choose a subgroup H of Homeo(X) such that
  1. H is finite (i.e. $H = \{h_1, \dots, h_r\}$);
  2. $g \circ h \circ g^{-1} \in H$ for every $g \in G$ and every $h \in H$. (Due to the finiteness of H, property 2 implies that the restriction to H of the conjugacy action of each $g \in G$ is a permutation of H.)
- We compute the persistent homology group of the quotient space $\frac{X}{H}$, with respect to the filtering function $\hat{\varphi}$ that takes each orbit ω of the group H to the maximum of φ on ω.

We shall use the symbol $PH_n^{\hat{\varphi}}(u, v)$ to denote the persistent homology group in degree n with respect to the filtering function $\hat{\varphi} : \frac{X}{H} \to \mathbb{R}$, computed at the point (u, v).

Remark

If G is Abelian, a simple way of getting a subgroup H of Homeo(X) verifying properties 1 and 2 consists in setting H equal to a finite subgroup of G. However, we have to observe that in most of the applications, the group G is not Abelian.

If G is finite, a trivial way of getting a subgroup H of Homeo(X) verifying properties 1 and 2 consists in setting $H = G$. This choice leads to consider the quotient space X/G. However, we have to observe that in most of the applications, the group G is not finite.

The trivial choice $H = \{Id\}$ can be always made, but it leads to compute the classical persistent homology of the topological space X.

Two key properties of $PH_n^{\hat{\varphi}}$ are expressed by the following results:

Theorem (Invariance with respect to the group G)

If $g \in G$ and $u, v \in \mathbb{R}$ with $u < v$, the groups $PH_n^{\hat{\varphi} \circ g}(u, v)$ and $PH_n^{\hat{\varphi}}(u, v)$ are isomorphic.

Under suitable assumptions about the topological space X the next statement holds:

Theorem (Stability)

For every $n \in \mathbb{Z}$, let us set $\rho_n^{\hat{\varphi}}(u, v) := \mathrm{rank}\left(PH_n^{\hat{\varphi}}(u, v)\right)$ and $\rho_n^{\hat{\psi}}(u, v) := \mathrm{rank}\left(PH_n^{\hat{\psi}}(u, v)\right)$. Then

$$D_{match}(\rho_n^{\hat{\varphi}}, \rho_n^{\hat{\psi}}) \leq d_G(\varphi, \psi) \leq d_{Id}(\varphi, \psi) = \|\varphi - \psi\|_\infty.$$

D_{match} is the classical matching distance between the persistent diagrams associated with $PH_n^{\hat{\varphi}}$ and $PH_n^{\hat{\psi}}$.

Some problems with our method:

The previous method presents some drawbacks:

- It requires to find a suitable group H, associated with G. This group could be difficult to find or not exist at all.
- The computation of the persistent homology group in degree n with respect to the filtering function $\hat{\varphi} : \frac{X}{H} \to \mathbb{R}$ requires a triangulation of X that is invariant under the action of H. This triangulation could be difficult to find or not exist at all.

An alternative approach based on G-functionals

Fortunately, an alternative approach is available.

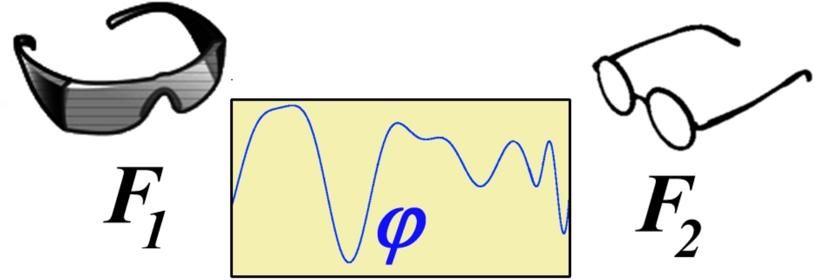
We will describe it in the second part of this talk.

This part is based on an ongoing joint research project with Grzegorz Jabłoński (Jagiellonian University).

An alternative approach based on *G*-functionals

Informal description of our idea

Instead of changing the topological space X, we can get invariance with respect to the group G by changing the “glasses” that we use to observe the filtering functions. In our approach, these “glasses” are G-functionals F_i, which act on the filtering functions.



Let us consider the following objects:

- A compact and locally connected topological space X with nontrivial homology in degree k.
- The vector space $C^0(X, \mathbb{R})$ of all continuous maps from X to $\mathbb{R}$, endowed with the max-norm $\|\cdot\|_\infty$.
- A subset S of $C^0(X, \mathbb{R})$.
- The set $\mathcal{F}_G$ of all non-expansive G-functionals from S to $C^0(X, \mathbb{R})$.

In plain words, $F \in \mathcal{F}_G$ means that

1. $\|F(\varphi_1) - F(\varphi_2)\|_\infty \leq \|\varphi_1 - \varphi_2\|_\infty$. *(Non-expansivity)*
2. $F(\varphi \circ g) = F(\varphi) \circ g$ *(i.e., if $S = C^0(X, \mathbb{R})$ then F is a G-functional, if we look at $C^0(X, \mathbb{R})$ as a G-set where the action of G on $C^0(X, \mathbb{R})$ is given by composition on the right.)*

Let us consider the following metric D_G:

$$D_G(\varphi_1, \varphi_2) := \sup_{F \in \mathcal{F}_G} d_{match}(r^k(F(\varphi_1)), r^k(F(\varphi_2)))$$

for every $\varphi_1, \varphi_2 \in S$, where $r^k(\psi)$ denotes the persistent Betti number function of ψ in degree k.

We have recently proven this statement:

Theorem

The distance D_G coincides with the natural pseudo-distance d_G on S.

Our idea

The previous theorem suggests the following approach.

Let us choose a finite subset $\mathcal{F}^*$ of $\mathcal{F}_G$, and set

$$D_G^*(\varphi_1, \varphi_2) := \max_{F \in \mathcal{F}^*} d_{match}(r^k(F(\varphi_1)), r^k(F(\varphi_2)))$$

for every $\varphi_1, \varphi_2 \in C^0(X, \mathbb{R})$.

Obviously, $D_G^* \leq D_G$. Furthermore, we can hope to have $\mathcal{F}^*$ so dense in $\mathcal{F}_G$ that the new pseudo-distance D_G^* is close to D_G.

We also underline that D_G^* is G-invariant: $D_G^*(\varphi_1, \varphi_2 \circ g) = D_G^*(\varphi_1, \varphi_2)$ for any $g \in G$, for every $\varphi_1, \varphi_2 \in C^0(X, \mathbb{R})$.

Let us check what happens in practice

We have considered

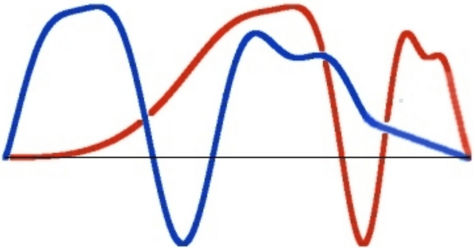
1. a dataset of 10000 functions from S^1 to $\mathbb{R}$, depending on five random parameters;
2. these three groups:
   - the group $\mathrm{Homeo}(S^1)$ of all self-homeomorphisms of S^1,
   - the group $R(S^1)$ of all rotations of S^1,
   - the trivial group Id, containing just the identity of S^1.

Obviously,

$$\mathrm{Homeo}(S^1) \supset R \supset I.$$

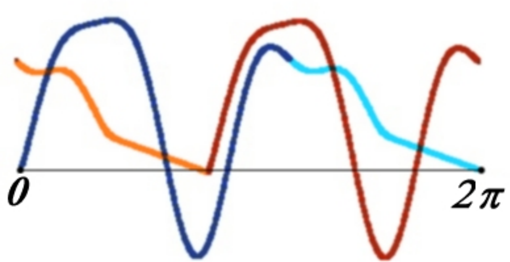
Let us check what happens in practice

The choice of Homeo(S^1) as an invariance group implies that the following two functions are considered equivalent. Their graphs are obtained from each other by applying a horizontal stretching.



Let us check what happens in practice

The choice of $R(S^1)$ as an invariance group implies that the following two functions are considered equivalent. Their graphs are obtained from each other by applying a rotation of S^1.



The choice of *Id* as an invariance group means that two functions are considered equivalent if and only if they coincide everywhere.

The results of our experiments: the group $\mathrm{Homeo}(S^1)$

If we choose $G = \mathrm{Homeo}(S^1)$, in order to proceed we need to choose a finite set of non-expansive $\mathrm{Homeo}(S^1)$-functionals.
In our experiment we have considered these three non-expansive $\mathrm{Homeo}(S^1)$-functionals:

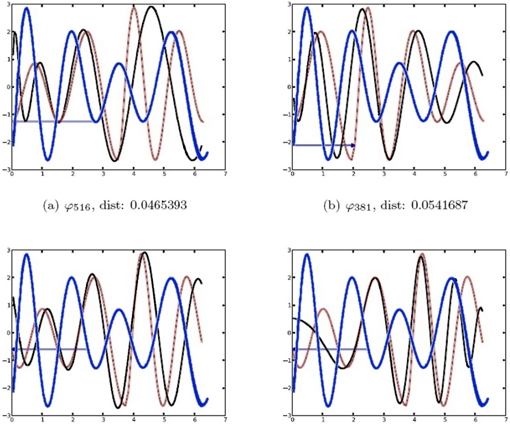
- $F_0 = id$ (i.e., $F_0(\varphi) = \varphi$);
- $F_1 = -id$ (i.e., $F_0(\varphi) = -\varphi$);
- $F_2 = \frac{1}{5} \cdot \sup\{-\varphi(x_1) + \varphi(x_2) - \frac{1}{2}\varphi(x_3) + \frac{1}{2}\varphi(x_4) - \varphi(x_5) + \varphi(x_6)\}$, $(x_1, \ldots, x_6)$ varying in the set of all 6-tuples in S^1 where the points are counterclockwise listed.

This choice produces the $\mathrm{Homeo}(S^1)$-invariant pseudo-distance

$$D^1_{\mathrm{Homeo}(S^1)}(\varphi_1, \varphi_2) := \max_{0 \le i \le 2} d_{match}(r^k(F_i(\varphi_1)), r^k(F_i(\varphi_2)).)$$

The results of our experiments: the group $\mathrm{Homeo}(S^1)$

Here is a query (in blue), and the first four retrieved functions (in black):



(a) φ_{516}, dist: 0.0465393

(b) φ_{381}, dist: 0.0541687

(c) φ_{7776}, dist: 0.0984192

(d) φ_{6214}, dist: 0.10376

dist:

Mean	Max	Standard deviation
1.648	4.831	0.934

The results of our experiments: the group $R(S^1)$

If we choose $G = R(S^1)$, in order to proceed we need to choose a finite set of non-expansive $R(S^1)$-functionals.
Obviously, since F_0, F_1 and F_2 are Homeo(S^1)-invariant, they are also $R(S^1)$-invariant. In our experiment we have added these five non-expansive $R(S^1)$-functionals (which are not Homeo(S^1)-invariant) to F_0, F_1 and F_2:

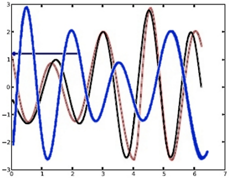
- $F_3(\varphi) := \max\{\varphi(x), \varphi(x+\pi)\}$
- $F_4(\varphi) := \frac{1}{2} \cdot (\varphi(x) + \varphi(x + \frac{\pi}{4}))$
- $F_5(\varphi) := \max\{\varphi(x), \varphi(x + \pi/10), \varphi(x + \frac{2\pi}{10}), \varphi(x + \frac{3\pi}{10})\}$
- $F_6(\varphi) := \frac{1}{3} \cdot (\varphi(x) + \varphi(x + \frac{\pi}{3}) + \varphi(x + \frac{\pi}{4}))$
- $F_7(\varphi) := \frac{1}{3} \cdot (\varphi(x) + \varphi(x + \frac{\pi}{3}) + \varphi(x + \frac{2\pi}{3}))$

This choice produces the $R(S^1)$-invariant pseudo-distance

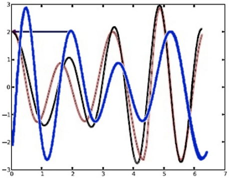
$$D^2_{R(S^1)}(\varphi_1, \varphi_2) := \max_{0 \le i \le 7} d_{match}(r^k(F_i(\varphi_1)), r^k(F_i(\varphi_2)).)$$

The results of our experiments: the group $R(S^1)$

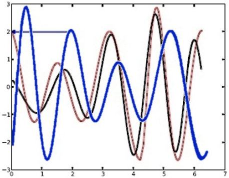
Here is a query (in blue), and the first four retrieved functions (in black):



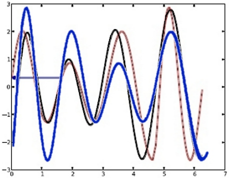
(a) φ_{5566}, dist: 0.333405



(b) φ_{8454}, dist: 0.422668



(c) φ_{8909}, dist: 0.453949



(d) φ_{4426}, dist: 0.46463

dist:

Mean	Max	Standard deviation
1.938	4.831	0.874

The results of our experiments: the group Id

If we choose $G = Id$, in order to proceed we need to choose a finite set of non-expansive functionals (obviously, every functional is a Id-functional).
In our experiment we have considered these three non-expansive functionals (which are not $R(S^1)$-functional):

- $F_8(\varphi) := \sin(x)\varphi(x)$
- $F_9(\varphi) := \frac{\sqrt{2}}{2}\sin(x)\varphi(x) + \frac{\sqrt{2}}{2}\cos(x)\varphi(x + \frac{\pi}{2})$
- $F_{10}(\varphi) := \sin(2x)\varphi(x)$

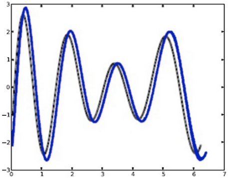
We have added F_8, F_9, F_{10} to $F_1, \ldots, F_7$.

This choice produces the pseudo-distance

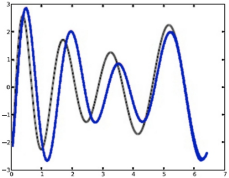
$$D_G^3(\varphi_1, \varphi_2) := \max_{0 \le i \le 10} d_{match}(r^k(F_i(\varphi_1)), r^k(F_i(\varphi_2)).)$$

The results of our experiments: the group *Id*

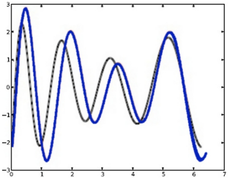
Here is a query (in blue), and the first four retrieved functions (in black):



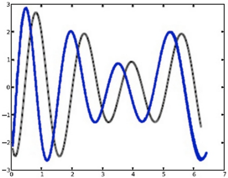
(a) φ_{7133}, dist: 0.415802



(b) φ_{7001}, dist: 0.598145



(c) φ_{389}, dist: 0.617218



(d) φ_{5723}, dist: 0.617981

	Mean	Max	Standard deviation
dist:	2.022	4.831	0.828

Conclusions

In this talk we have shown that

- Persistent homology can be adapted to proper subgroups of the group of all self-homeomorphisms of a triangulable space, in two different ways. Both of these methods are stable with respect to noise.
- In particular, the approach based on non-expansive G-functionals can be used for any subgroup G of Homeo(S^1). An experiment concerning this method has been illustrated, showing the possible use of this approach for data retrieval.

THANKS FOR YOUR ATTENTION