INF 556: Topological Data Analysis **Fall 2018**

Lecture 3: Introduction to homology

Lecturer: Steve Oudot *T.A.: Théo Lacombe*

Disclaimer:
Some typo and errors may remain. Please mention them at `theo.lacombe@polytechnique.edu`. Use these notes with caution, especially during the exam (we decline all responsibility linked with the use of these notes during the exam session).

Reminder: These notes are a concise summary of the lectures. They do not intend in any case to substitute to your personal notes and are just an additional support in order to clarify or insist on some points.

Some references: As a complement for these lecture notes, you can check for two books:

- *Element of Algebraic Topology*, by J.Munkres (1984), especially chapter 1.
- *Introduction to Computational Topology*, by H.Edelsbrunner and J.Harer (you can find an extract relative to Homology theory on the website for this course).

Keywords: Homotopy, Homology group, Simplicial complexes.

3.1 Context - Topological spaces

General idea: Classify (regroup) items which are *equivalent* up to an isomorphism (the class of admissible transformations depends on the topic). Basically, you consider that X should be consider the same as Y if you can transform, in some sense, X into Y.

Example: classification of surfaces.

Definition 1. *h is an* homeomorphism *if:*

- *h is $\mathcal{C}^0$*
- *h is bijective*
- *h^{-1} is also bijective.*

Two items X and Y are said to be homeomorph if there is an homeomorphism h such as $h(X) = Y$.

Remark: One can easily check that *being homeomorph* defines an *equivalence relation*, that is:

- X is always homeomorph to X itself.
- if X is homeomorph to Y, then Y is homeomorph to X
- if X is homeomorph to Y, and Y to Z, then X is homeomorph to Z.

Then we have the following theorem (this is out of the scope of this course and is just a cultural example. See on-line for illustrations and details.):

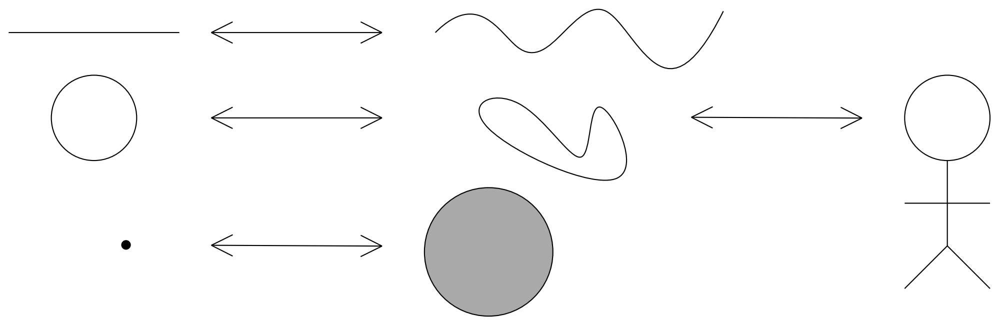

Figure 3.1: Some example of homotopic items. Left-right-arrows mean "are homotopic". Grey means "filled".

Theorem 1 (Classification of surfaces up to homeomorphism - admitted, out of the scope of this course)**.** *Every connected, compact, boundary-free surface is homeomorph to one of:*

- *the sphere* $\mathcal{S}^2$
- *A connected sum of torus*
- *A connected sum of projective plan.*

This theorem states that up to homeomorphism, there exists three different types of connected compact surfaces. That's a standard way to summarize and classify you data. We can look for other way to classify our data:

Definition 2. $f, g : X \to Y$ *are homotopic if there exist an application* $\varphi : [0,1] \times X \to Y$ *such as* φ *is continuous and* $\varphi(0,\cdot) = f, \varphi(1,\cdot) = g$.

Definition 3. X *and* Y *are said to have the same type of homotopy (or shortly are homotopic) if there exist* f, g *such has* $f \circ g = id_Y, g \circ f = id_X$.

The idea is that you can continuously transform f into g by looking at $t \mapsto \varphi(t,\cdot)$.

Proposition 1. X *and* Y *are homeomorph* $\Rightarrow$ X *and* Y *are homotopic.*

Proof. Exercise. □

3.2 Homology

3.2.1 Intuition

Let X be a topological space.

Definition 4. *A* path *is a continuous map* $[0,1] \to X$

Definition 5. $x \sim y$ *in* X *if and only if it exists a path* $\gamma : [0,1] \to X$ *such that* $\gamma(0) = x, \gamma(1) = y$.

Proposition 2. *This is an equivalence relation over* X.

Proof. Exercise. □

Definition 6 (Path connected-components)**.** *The path connected components of* X *is the space defined as* $X/\sim$*, that is the equivalent classes of the relation* $\sim$.

Alternative formulation: $x, y \in X$ are in the same connected component if it exist a path γ such that $\gamma(\underbrace{\partial[0,1]}_{\{0,1\}}) = \{x, y\}$.

Definition 7 (In higher dimension:). *Similarly, a* loop *is a map $\gamma : \mathcal{S}^1 \to X$ that is continuous ($\mathcal{S}^1$ designs the sphere in $\mathbb{R}^2$, that is the unit circle).*

Two loops γ, γ' are equivalent if there exists a surface Σ and a map $\nu : \Sigma \to X$ such that $\nu(\partial\Sigma) = \gamma(\mathcal{S}^1) \cup \gamma'(\mathcal{S}^1)$

More generally:

- *a r-*cycle *is a map $\gamma : \Sigma \to X$ where $\dim \Sigma = r$ and $\partial\Sigma = \emptyset$.*
- *$\gamma \sim \gamma'$ if exists $\nu : \Omega \to X$ with $\dim \Omega = r + 1$ such that $\nu(\partial\Omega) = \gamma(\Sigma) \cup \gamma'(\Sigma')$.*

See fig 3.2 for a visual representation of it.

Since such theoretical notions are not handy (and also because in practice, Data Analysis often start with a point cloud), we are going to introduce a discrete formulation of these notions.

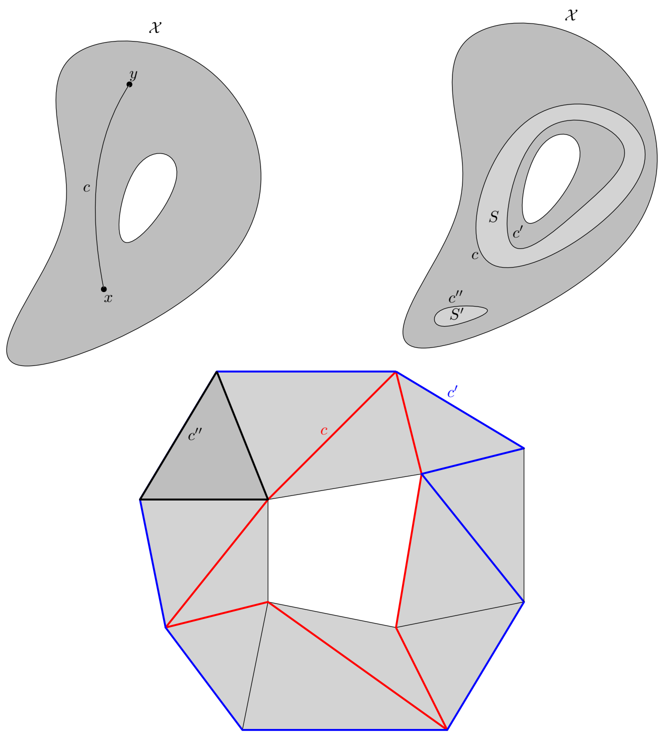

Figure 3.2: *(left)* Here, points x and y are equivalent in $\mathcal{X}$ in sense that $\{x\} \cup \{y\}$ is the boundary of a (continuous) path c (included in $\mathcal{X}$). *(right)* Here, path c and c' are equivalent because $c \cup c'$ is the boundary of a continuous surface S included in $\mathcal{X}$. Besides, c'' is *trivial* because $c \cup \emptyset$ is the boundary of a continuous surface in $\mathcal{X}$. *(bottom)* Similar notions on a simplicial complex.

3.3 Simplicial complexes

The general idea is that simplicial complexes extend the standard notion of graph (vertices and edges, that is items with dimension zero and one) by adding some higher dimensional components to it.

Definition 8 (Combinatorial definition of simplicial complexes). *Let V be some finite set (the vertices). A* simplicial complex *on V is a set $K \subset \mathcal{P}(V)$ such that: $\forall \tau \subset \sigma \subset V, \sigma \in K \Rightarrow \tau \in K$. Such a τ is called a* face *of σ.*

With a geometrical approach (but keep in mind that the basic definition is a purely combinatorial one), a simplex is a generalization of the notion of triangle (dim $p = 2$) or tetrahedron (dim $p = 3$) to arbitrary dimensions. A p-simplex is a p-dimensional polytope defined by the convex hull of $p+1$ affinely independent points in $\mathbb{R}^p$. A *k-face* of a p-simplex ($k \leqslant p$) is the convex hull of any subset with cardinality k of the $p+1$ points defining the simplex.

A simplicial complex K is a set of simplices verifying some structural properties:

- For any simplex $\sigma \in K$ and any face σ' of σ, we have $\sigma' \in K$.
- The intersection of two simplices $\sigma_1, \sigma_2 \in K$ is either $\emptyset$ or a face of both σ_1 and σ_2.

Definition 9 (Triangulability). *A topological space X is* triangulable *if there is a simplicial complex K and an homeomorphism $h : X \to K$.*

3.4 Simplicial homology

3.4.1 Chains space

Let X be a (finite) simplicial complex. Let $k \in \mathbb{N}$ (represents the dimension). We are interested in linear combinations over a given field $\mathbb{K}$ (called *field of coefficients*) between all k-simplices in X (denoted by X_k).

Definition 10 (k-chains). *The space of k-chains of a simplicial complex X over a field $\mathbb{K}$ is defined by:*

$$C_k(X, \mathbb{K}) := \left\{ \sum_{i=0}^{|X_k|} \alpha_i \sigma_i : \alpha_i \in \mathbb{K}, \sigma_i \in X_k \right\}$$

$C_k(X, \mathbb{K})$ can easily be endowed with a $\mathbb{K}$-vector space structure, by defining, for $c = \sum_i \alpha_i \sigma_i, c' = \sum_i \beta_i \sigma_i$ and $\lambda \in \mathbb{K}$,

$$\lambda c + c' := \sum_{i=1}^{|X_k|} \underbrace{(\lambda \alpha_i + \beta_i)}_{\in \mathbb{K}} \sigma_i$$

Remark: Since everything is finite here, $C_k(X, \mathbb{K})$ is a finite dimensional vector space with dimension given by the number of different k-simplices in X, that is $|X_k|$. Therefore, $C_k(X, \mathbb{K})$ is isomorphic (as a vector space) to $\mathbb{K}^{|X_k|}$.

About the field: For some theoretical reasons which are out of the scope of this course, we will work with finite fields which are of the form $\mathbb{Z}/p\mathbb{Z}$ for some prime number p. In practice, we will generally take $p = 2$, which makes computation easier since in this field, $+1 = -1$ (i.e. there is no orientation, etc.).

3.4.2 Boundary operator

Remark: During the lecture, this subsection was first presented with only non-oriented complexes. For sake of concision, we directly provide the oriented formulation (to get the non-oriented one, just consider $\mathbb{K} = \mathbb{Z}/2\mathbb{Z}$ and use the fact that $(-1)^i = +1$ to simplify things).

As said previously, we basically interested in *finding cycle that are not boundaries*. More precisely, we introduce the following operator:

Definition 11. *Given a k-simplex $[v_0 \dots v_k] \in X_k$, we denote by $[v_0 \dots, \widehat{v_j}, \dots v_k] \in X_{k-1}$ the $(k-1)$-simplex (which is a face of $[v_0 \dots v_k]$) where the vertex v_j got removed.*

The boundary operator over k-dimensional chains ∂_k for $k \geqslant 1$ is defined as:

$$\begin{aligned}
\partial_k : C_k(X, \mathbb{K}) &\to C_{k-1}(X, \mathbb{K}) \\
\sigma = [v_0 \dots v_k] &\mapsto \sum_{j=0}^{k} (-1)^j \underbrace{[v_0 \dots \widehat{v_j}, \dots, v_k]}_{\in X_{k-1}} \\
(\lambda\sigma + \sigma') &\mapsto \lambda\partial_k\sigma + \partial_k\sigma' && \textit{(linear extension)}
\end{aligned}$$

By convention, we set $\partial_0 = 0$.

Proposition 3. $\forall k > 0, \partial_{k-1} \circ \partial_k = 0$

Interpretation: "The boundary of a boundary is null".

Proof. Exercise. □

3.4.3 Homology group

We have the following chain of application:

$$C_k(X, \mathbb{K}) \xrightarrow{\partial_k} C_{k-1}(X, \mathbb{K}) \xrightarrow{\partial_{k-1}} C_{k-2}(X, \mathbb{K}) \xrightarrow{\partial_{k-2}} \dots$$

We thus define:

- *k-cycles* : $Z_k := \ker(\partial_k) = \{\sigma \in X_k | \partial_k\sigma = 0\}$
- *k-boundaries* : $B_k := \mathrm{Im}(\partial_{k+1}) = \{\partial_{k+1}\sigma : \sigma \in X_{k+1}\}$

With Prop 3, we know that $\mathrm{Im}(\partial_{k+1}) \subset \ker(\partial_k)$, which basically states that boundaries of $(k+1)$-simplices are k-cycles.

We can finally define the homology group:

Definition 12. *The k-th homology group of X for the field $\mathbb{K}$ is the following quotient:*

$$H_k(X, \mathbb{K}) := \frac{\ker(\partial_k)}{\mathrm{Im}(\partial_{k+1})} \tag{3.1}$$

Interpretation: It represent the "space of cycles modulo the boundaries".

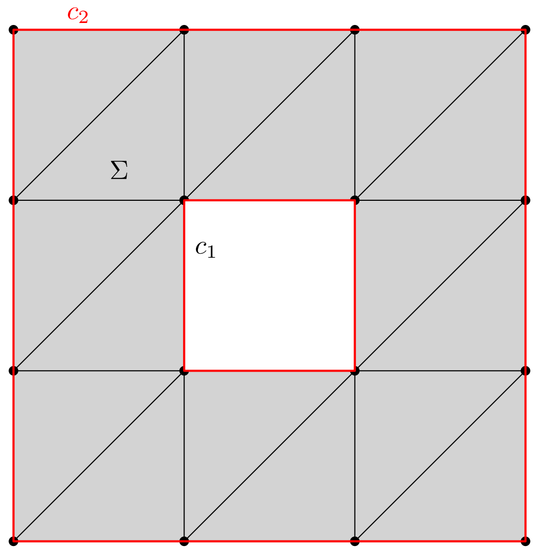

Figure 3.3: The two red cycles are equivalent in H_1 because there is a 2-dimensional area (the grey one, denoted as Σ) such that $\partial\Sigma = c_1 - c_2$. Therefore, when taking the quotient Z_1/B_2, which states that $\partial\Sigma = 0$, you basically observe that homologically, $c_1 = c_2$, thus are in particular linearly dependent.