Topological Data Analysis

2022–2023

Lecture 11

Stability for Functions

15 December 2022

Let $f:[a,b]\to\mathbb{R}$ be a continuous function. For every $t\in\mathbb{R}$, let

$$L_t(f)=\{x\in[a,b]\mid f(x)\le t\}$$

$$V_t(f)=H_0(L_t(f)).$$

Then $V(f)$ is a persistence module with $\pi_{s,t}:V_s(f)\to V_t(f)$ induced by the inclusion $L_s(f)\subseteq L_t(f)$ if $s\le t$.

If f has only finitely many critical points, then $V(f)$ is of finite type.

Our goal is to prove the inequality

$$\boxed{d_{int}(V(f),V(g))\le\|f-g\|_\infty}$$

where $\|f-g\|_\infty=\sup\{|f(x)-g(x)|:a\le x\le b\}$.

Proof: Given $f,g:[a,b]\to\mathbb{R}$, pick $\delta=\|f-g\|_\infty$. We need to prove that $V(f)$ and $V(g)$ are δ-interleaved. Here $\dim V(f)_\infty=1=\dim V(g)_\infty$

Note that $V(f)[\delta]=V(f-\delta)$ and $V(g)[\delta]=V(g-\delta)$.

$$V(f-\delta)_t = H_0(L_t(f-\delta)) = H_0(L_{t+\delta}(f)) = V_{t+\delta}(f) = V(f)[\delta]_t$$

$$L_t(f-\delta) = \{x \in [a,b] \mid f(x)-\delta \leq t\} =$$

$$= \{x \in [a,b] \mid f(x) \leq t+\delta\} = L_{t+\delta}(f) \checkmark$$

Since $\delta = \|f-g\|_\infty$, we have that $|f(x)-g(x)| \leq \delta$ for all $a \leq x \leq b$.

Hence
$$g(x)-\delta \leq f(x) \leq g(x)+\delta$$
$$f(x)-\delta \leq g(x) \leq f(x)+\delta \quad \text{for all } x.$$

Therefore
$$f(x)-2\delta \leq g(x)-\delta \leq f(x)$$
$$g(x)-2\delta \leq f(x)-\delta \leq g(x) \quad \text{for all } x.$$

This tells us that
$$L_t(f) \subseteq L_{t+\delta}(g) \qquad f(x) \leq t \Rightarrow g(x) \leq t+\delta$$
$$L_t(g) \subseteq L_{t+\delta}(f) \quad \text{for all } t.$$

These inclusions yield morphisms of persistence modules

$$V(f) \xrightarrow{F} V(g)[\delta] \quad \text{and} \quad V(g) \xrightarrow{G} V(f)[\delta]$$

by passing to homology.

Moreover $G[\delta] \circ F$ is induced by the inclusion $L_t(f) \subseteq L_{t+2\delta}(f)$, which is precisely $\sigma_{2\delta}(f)$.

$(\sigma_{2\delta})_t = \pi_{t,t+\delta}$

By symmetry, we also have $F[\delta] \circ G = \sigma_{2\delta}(g)$.

This proves that $V(f)$ and $V(g)$ are δ-interleaved. ✓

Morse functions

Let M be a closed smooth manifold.

For a smooth function $f: M \to \mathbb{R}$, a critical point of f is a point $p \in M$ such that $\frac{\partial f}{\partial x_i}(p) = 0$ for all i, where $(x_1, \dots, x_n)$ are local coordinates in a chart around p.

If f has finitely many critical points, then f yields a persistence module of finite type, namely

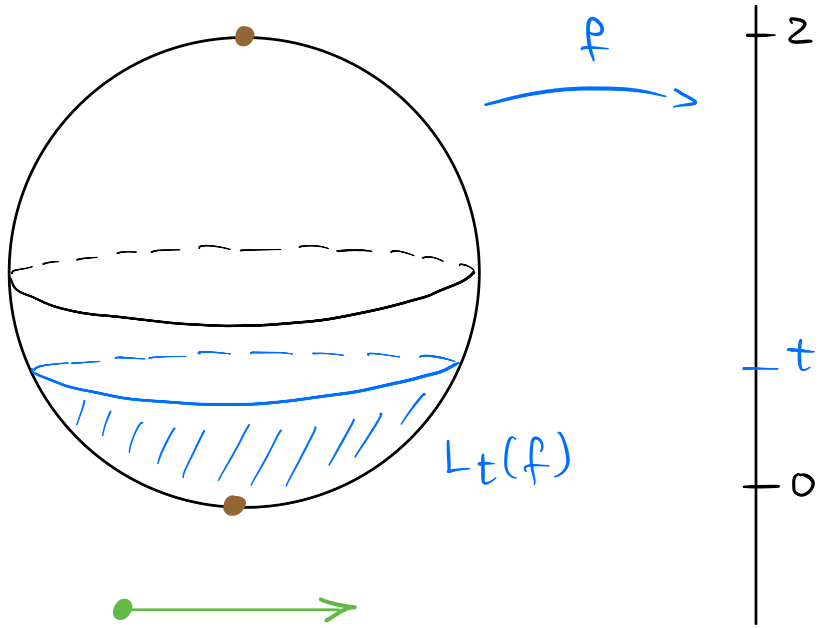
$$V_t(f) = H_*(L_t(f))$$

Now H_* denotes singular homology

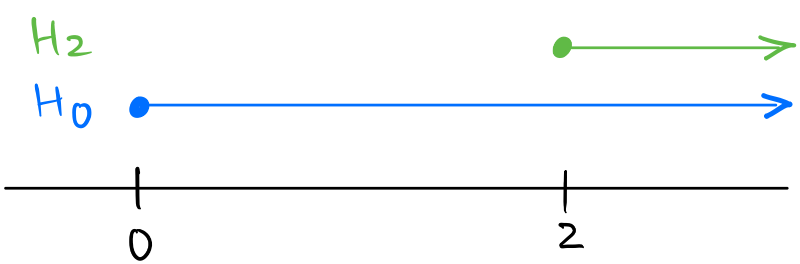
where $L_t(f) = \{x \in M \mid f(x) \leq t\}$.

Examples

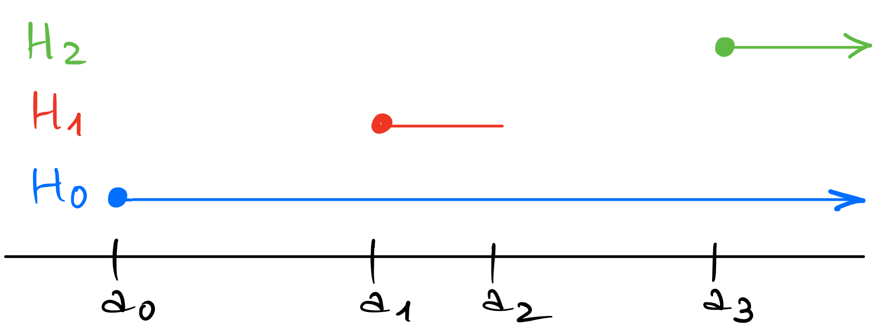
① $M = S^2$:

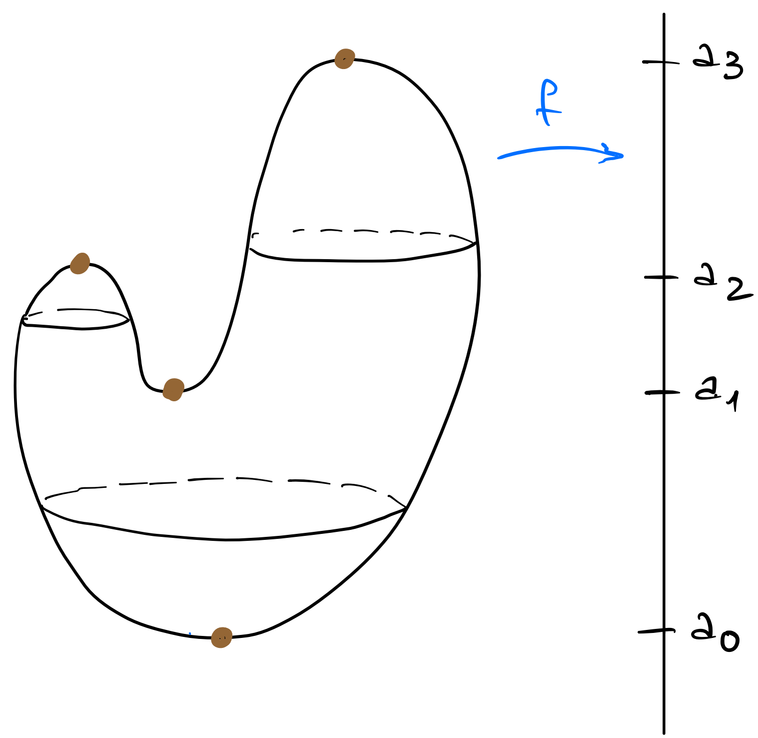


The spectrum of $V(f)$ is contained in the set of critical values of the function f.

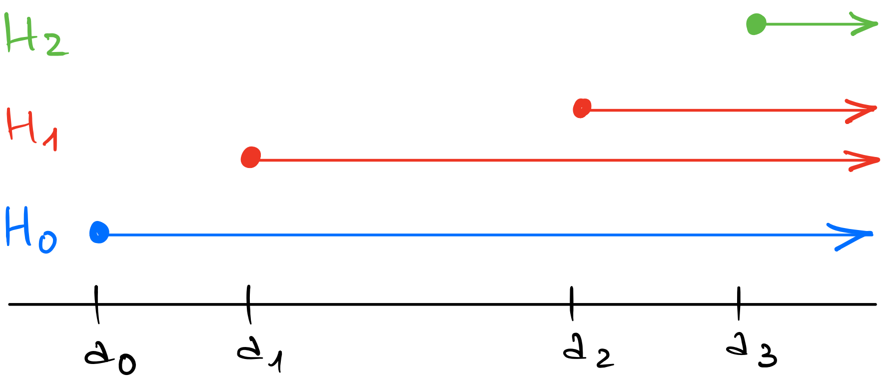


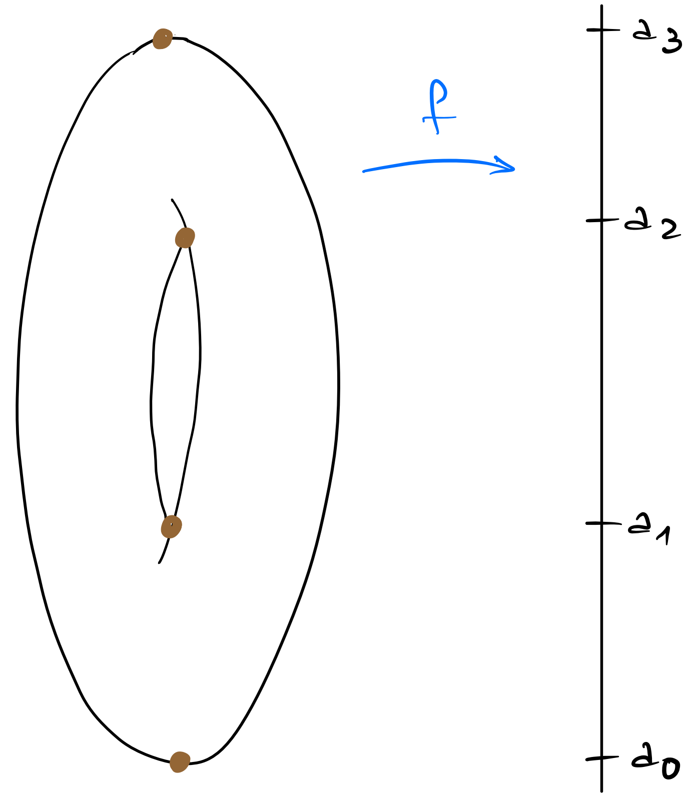
② M another sphere





3 $\quad M = S^1 \times S^1$





The Stability Theorem holds with the same proof.

A smooth function $f : M \to \mathbb{R}$ is called a <u>Morse</u> function if the Hessian $\left(\dfrac{\partial^2 f}{\partial x_i \partial x_j}\right)$ is nondegenerate at critical points.

Then critical points are isolated and therefore there is only a finite number of critical points, since M is compact.

Čech complexes

Let X and Y be point clouds in $\mathbb{R}^N$. Define $f, g : \mathbb{R}^N \longrightarrow \mathbb{R}$ as

$$f(p) = d(p, X), \qquad g(p) = d(p, Y).$$

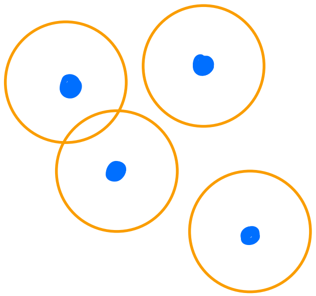
Then f and g yield persistence modules $V(f)$ and $V(g)$ of finite type in the same way as before:

$$L_t(f) = \{ p \in \mathbb{R}^n \mid f(p) \leq t \}$$

$$V_t(f) = H_*(L_t(f)).$$

The Stability Theorem holds with the same proof.

Now let $\tilde{V}_t(X) = H_*(C_t(X))$, $\tilde{V}_t(Y) = H_*(C_t(Y))$, where C_t are Čech complexes.



Note that, if $X = \{x_i\}_{i \in I}$, then

$$L_t(f) = \{ p \in \mathbb{R}^n \mid d(p, X) \leq t \} = \bigcup_{i \in I} \overline{B}_t(x_i).$$

Hence $L_t(f) \simeq |C_{2t}(X)|$ by the Nerve Theorem.

geometric realization

Consequently, $V_t(f) = H_*(L_t(f)) = H_*(C_{2t}(X)) = \tilde{V}_{2t}(X)$

or equivalently $\tilde{V}_t(X) = V_{t/2}(f) = V_t(2f)$.

Similarly, $\tilde{V}_t(Y) = V_t(2g)$.

On the other hand,

$$\|f-g\|_\infty = \sup\{|d(p,X)-d(p,Y)| : p\in\mathbb{R}^N\} =$$
$$= d_H(X,Y).$$

Hausdorff distance $(*)$

This tells us that

$$d_{int}(\tilde{V}(X),\tilde{V}(Y)) = d_{int}(V(2f),V(2g)) \leq$$
$$\leq \|2f-2g\|_\infty = 2\,d_H(X,Y).$$

This is a form of stability for Čech complexes!

(*) Proof:

$$\sup\{d(p,Y) : p\in X\} = \sup\{d(p,Y) - d(p,X) : p\in X\} \leq$$
$$\leq \sup\{d(p,Y) - d(p,X) : p\in \mathbb{R}^N\}.$$

Since X is finite, for each $p\in\mathbb{R}^N$ there is a point $x_p\in X$ such that $d(p,X) = d(p,x_p)$. Then, for every $p\in\mathbb{R}^N$,

$$d(p,Y) - d(p,X) = d(p,Y) - d(p,x_p) \leq d(x_p,Y) \leq$$
$$\leq \sup\{d(q,Y) \mid q\in X\}.$$

Therefore the above inequality is an equality.

Consequently,

$$\sup\{|d(p,X) - d(p,Y)| : p\in\mathbb{R}^N\} =$$
$$= \max\{\sup\{d(p,X) - d(p,Y) : p\in\mathbb{R}^N\}, \sup\{d(p,Y) - d(p,X) : p\in\mathbb{R}^N\}\} =$$
$$= \max\{\sup\{d(p,X) : p\in Y\}, \sup\{d(p,Y) : p\in X\}\} = d_H(X,Y). \checkmark$$