

# GDT PERSISTANCE – EXPOSÉ 3

## VERS DES APPLICATIONS DE L’HOMOLOGIE PERSISTANTE

JOHAN LERAY

### INTRODUCTION

Le but de cet exposé est d’exposer certains résultats d’homologie persistante "plus proche des applications". Cet exposé s’appuie notamment sur les chapitres 4 et 7 de [Oud15].

### TABLE DES MATIÈRES

Introduction	1
Notations	1
1. Complexes de Cech et Vietoris Rips	1
1.1. Weak feature size	1
1.2. Distances	3
1.3. Complexes de Cech et de Vietoris-Rips	3
1.4. Une application	5
2. Inclure de la géométrie	6
2.1. Théorèmes de stabilité	6
2.2. Espaces métriques munis de fonctions	7
2.3. "Changer d’espace"	9
Références	12

### NOTATIONS

On note  $H_*(-)$ , l’homologie à coefficient dans un corps.

#### 1. COMPLEXES DE CECH ET VIETORIS RIPS

Cette première section est basée sur [CO08] et [Oud15, Chap. 4].

**1.1. Weak feature size.** On considère  $K$  un compact de  $(\mathbb{R}^d, \|\cdot\|)$  et on note  $\delta_K$  la distance

$$\begin{aligned} \delta_K : \mathbb{R}^d &\longrightarrow \mathbb{R}_+ \\ x &\longrightarrow \min_{y \in K} \|x - y\| \end{aligned}$$

À  $x$  dans  $\mathbb{R}^d$ , on associe  $\Pi_K(x)$ , l’ensemble des  $y$  dans  $K$  qui réalisent le minimum de  $\delta_K$ . Si  $|\Pi_K(x)| = 1$ , on note  $\pi_K(x)$  le projeté de  $x$  sur  $K$ .

**Définition 1** (Axe médian). On appelle *axe médian* de  $K$ , l’ensemble

$$\mathfrak{M}(K) := \{x \in \mathbb{R}^d \mid \#\Pi_K(x) \geq 2\}$$

---

*Date:* 11 décembre 2018.

La distance  $\delta_K$  est différentiable sur  $\mathbb{R}^d - (K \cup \overline{\mathfrak{M}(K)})$ , de gradient

$$\tilde{\nabla}_K(x) = \frac{x - \pi_K(x)}{\delta_K(x)}.$$

On généralise ce gradient à  $\mathbb{R}^d - K$  par

$$\tilde{\nabla}_K(x) = \frac{x - c(\Pi_K(x))}{\delta_K(x)},$$

où  $c(\Pi_K(x))$  désigne le centre de la plus petite boule contenant l'ensemble  $\Pi_K(x)$ .

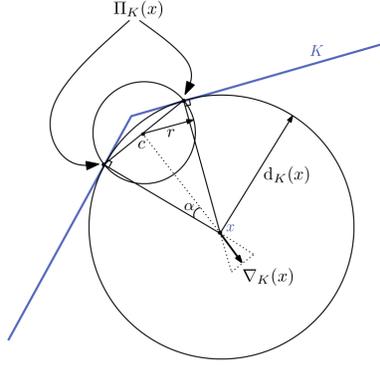


FIGURE 1. Exemple – cf. [CL05]

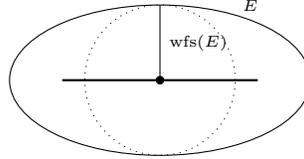
**Définition 2.** Un point  $x$  de  $\mathbb{R}^d$  est dit *critique* si  $\nabla_K(x) = 0$ . Une valeur  $v$  de  $\mathbb{R}$  est une *valeur critique* de  $\delta_K$  s'il existe un point critique  $X$  tel que  $v = \delta_K(x)$ .

**Remarque 3.** Tout point de  $K$  est critique, en particulier 0 est valeur critique.

**Définition 4** (Weak Feature Size).

$$\text{wfs}(K) := \inf \{ \alpha > 0 \mid \alpha \text{ valeur critique de } \delta_K \}$$

**Exemple 5.** On considère  $E$ , l'ellipse suivante.



Alors  $\text{wfs}(E)$  est la longueur de son petit rayon.

**Notation 6.** Pour  $\alpha > 0$ , et  $X$ , un sous-ensemble de  $\mathbb{R}^d$ , on note

$$X^\alpha := \{ x \in \mathbb{R}^D \mid \delta_K(x) \leq \alpha \}.$$

**Lemme 7** (d'isotopie). Soient  $0 < \alpha < \beta$  tel qu'il n'y a pas de valeur critique de  $\delta_K$  dans l'intervalle  $[\alpha, \beta]$ , alors  $K^\alpha$  est un retract par déformation de  $K^\beta$ , donc  $K^\alpha \hookrightarrow K^\beta$  est une équivalence d'homotopie. En particulier, pour tout  $0 < \alpha < \beta < \text{wfs}(K)$ ,  $K^\alpha$  et  $K^\beta$  ont le même type d'homotopie.

**1.2. Distances.** On définit deux distances. La première est une distance entre espaces plongés dans  $\mathbb{R}^d$ .

**Définition 8** (Distance de Hausdorff). Soient  $K$  et  $P$ , deux sous-ensembles de  $\mathbb{R}^d$ . On définit la *distance de Hausdorff* entre  $K$  et  $P$  par

$$d_{\mathcal{H}}(K, P) := \max \left\{ \sup_{x \in K} \inf_{p \in P} \|x - p\|, \sup_{p \in P} \inf_{x \in K} \|x - p\| \right\}.$$

**Remarque 9.** On a une autre formulation de la distance de Hausdorff :

$$d_H(K, P) = \min \{ \alpha \in \mathbb{R}_+ \mid K \subset P^\alpha \text{ et } P \subset K^\alpha \}.$$

On définit également la *distance de Gromov-Hausdorff* entre deux espaces métriques, qui ne dépend pas d'un plongement.

**Définition 10** (Distance de Gromov-Hausdorff). Soient  $X$  et  $Y$ , deux espaces métriques. On définit la *distance de Gromov-Hausdorff* entre  $X$  et  $Y$ , par

$$d_{\mathcal{GH}}(X, Y) = \inf_{Z, \gamma_X, \gamma_Y} d_{\mathcal{H}}^Z(\gamma_X(X), \gamma_Y(Y)).$$

**1.3. Complexes de Cech et de Vietoris-Rips.** On définit deux complexes suivants.

**Définition 11** (Complexe de Vietoris-Rips). Soit  $(X, d_X)$  un espace métrique. Pour  $\alpha$  dans  $\mathbb{R}_+$ , on définit le complexe simplicial  $R^\alpha(X)$  où  $R_0^\alpha(X) = X$  et  $\sigma = [x_0, \dots, x_n]$  est dans  $R_n^\alpha(X)$  si,

$$(1) \quad \forall i \neq j \in \llbracket 0, n \rrbracket, d_X(x_i, x_j) \leq \alpha.$$

**Définition 12** (Complexe de Cech). Soit  $(X, d_X)$  un espace métrique. Pour  $\alpha$  dans  $\mathbb{R}_+$ , on définit le complexe simplicial  $C^\alpha(X)$  où  $C_0^\alpha(X) = X$  et  $\sigma = [x_0, \dots, x_n]$  est dans  $C_n^\alpha(X)$  si

$$(2) \quad \bigcap_{i \in \llbracket 0, n \rrbracket} B(x_i, \alpha) \neq \emptyset,$$

où  $B(x_i, \alpha)$  désigne la boule fermée de centre  $x_i$  et de rayon  $\alpha$

**Remarque 13.** Le complexe de Cech est à un interet théorique alors que le complexe de Rips a un interet pratique. En effet, la condition (1) est dur à tester numériquement, tandis que la condition (2) est facile à vérifier.

Dans la suite, on démontrera les résultats théoriques sur le complexe de Cech, qui induiront des résultats similaires sur le complexe de Rips, grâce au lemme suivant.

**Lemme 14.** *Pour tout  $\alpha > 0$ , on a les inclusions*

$$C^{\alpha/2}(X) \subseteq R^\alpha(X) \subseteq C^\alpha(X).$$

**Lemme 15.** *Soient  $X$  un compact de  $\mathbb{R}^d$  et  $L$  un ensemble fini dans  $\mathbb{R}^d$  tel que  $d_{\mathcal{H}}(X, L) < \varepsilon < \frac{1}{4} \text{wfs}(X)$ . Alors, pour tout  $\alpha$  et  $\alpha'$  dans  $[\varepsilon, \text{wfs}(X) - \varepsilon]$  tels que  $\alpha' - \alpha > 2\varepsilon$  et pour tout  $\lambda$  dans  $]0, \text{wfs}(X)[$ , on a, pour tout entier  $k$ , l'isomorphisme suivant :*

$$H_k(X^\lambda) \cong \text{Im} \left( H_k(L^\alpha) \xrightarrow{i_*} H_k(L^{\alpha'}) \right).$$

*Démonstration.* On peut supposer que  $\varepsilon < \alpha < \alpha' - 2\varepsilon < \text{wfs}(X) - 3\varepsilon$ . Par hypothèse, on a les inclusions

$$(3) \quad X^{\alpha-\varepsilon} \hookrightarrow L^\alpha \hookrightarrow X^{\alpha+\varepsilon} \hookrightarrow L^{\alpha'} \hookrightarrow X^{\alpha'+\varepsilon}.$$

Par le Lemme 7, on a, pour tout  $0 < \beta < \beta' < \text{wfs}(X)$ ,  $X^\beta \xrightarrow{\sim} X^{\beta'}$ . Donc la suite d'inclusions (3) induit une suite de morphismes

$$\begin{array}{ccccccc} & & \cong & & \cong & & \\ & \curvearrowright & & \curvearrowright & & & \\ \mathbf{H}_k(X^{\alpha-\varepsilon}) & \longrightarrow & \mathbf{H}_k(L^\alpha) & \longrightarrow & \mathbf{H}_k(X^{\alpha+\varepsilon}) & \longrightarrow & \mathbf{H}_k(L^{\alpha'}) & \longrightarrow & \mathbf{H}_k(X^{\alpha'+\varepsilon}), \end{array}$$

où les isomorphismes découlent du lemme 7. Ainsi, on a

$$\text{rk}(\mathbf{H}_k(X^{\alpha-\varepsilon}) \rightarrow \mathbf{H}_k(X^{\alpha+\varepsilon})) = \dim_{\mathbb{R}} \mathbf{H}_k(X^{\alpha+\varepsilon}),$$

donc

$$\text{rk}(\mathbf{H}_k(L^\alpha) \rightarrow \mathbf{H}_k(L^{\alpha'})) = \dim_{\mathbb{R}} \mathbf{H}_k(X^{\alpha+\varepsilon}) = \dim_{\mathbb{R}} \mathbf{H}_k(X^\lambda).$$

□

**Lemme 16** (Nerf persistant). *Soient  $X \subset X'$ , deux espaces paracompacts, munis de  $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$  et  $\mathcal{U}' = \{U'_\alpha\}_{\alpha \in A}$ , deux bons recouvrements respectivement de  $X$  et de  $X'$ , tels que*

$$\forall \alpha \in A, U_\alpha \subseteq U'_\alpha.$$

*Alors, il existe des équivalences d'homotopies  $\mathcal{N}\mathcal{U} \rightarrow X$  et  $\mathcal{N}\mathcal{U}' \rightarrow X'$  tel que le diagramme suivant*

$$\begin{array}{ccc} \mathcal{N}\mathcal{U} & \hookrightarrow & \mathcal{N}\mathcal{U}' \\ \downarrow & & \downarrow \\ X & \hookrightarrow & X' \end{array}$$

*commute au niveau homologique et homotopique.*

*Idée de preuve.* On note  $N = |A| - 1$  et  $\Delta^N$  le  $N$ -simplexe standard. À tout sous-ensemble non vide  $S$  de  $A$ , on associe  $[S]$ , la face de  $\Delta^N$  engendré par les éléments de  $S$ , et

$$U_S = \bigcap_{s \in S} U_s \subset X.$$

On pose  $\Delta X \subset X \times \Delta^N$ , défini par

$$\Delta X = \bigcup_{\emptyset \neq S \subseteq A} U_S \times [S].$$

On a l'équivalence d'homotopie  $p : \Delta X \xrightarrow{\sim} X$  (voir [Hat02]). De plus, on a le diagramme commutatif suivant :

$$\begin{array}{ccccc} X & \xleftarrow{\sim} & \Delta X & \xrightarrow{\sim} & \mathcal{N}\mathcal{U} \\ \downarrow & & \downarrow & & \downarrow \\ X' & \xleftarrow{\sim} & \Delta X' & \xrightarrow{\sim} & \mathcal{N}\mathcal{U}' \end{array} .$$

(voir [Hat02, Prop, 4G.3] pour le carré de droite.)

□

On applique le Lemme 16 à  $X = L^\alpha$ ,  $X' = L^{\alpha'}$ , avec les bons recouvrements  $\mathcal{U} = \{B(l, \alpha)\}_{l \in L}$  et  $\mathcal{U}' = \{B(l, \alpha')\}_{l \in L}$  : il existe des équivalences d'homotopies  $C^\alpha(L) \rightarrow L^\alpha$  et  $C^{\alpha'}(L) \rightarrow L^{\alpha'}$  telles que le diagramme

$$\begin{array}{ccc} C^\alpha(L) & \hookrightarrow & C^{\alpha'}(L) \\ \downarrow & & \downarrow \\ L^\alpha & \hookrightarrow & L^{\alpha'} \end{array}$$

commute après passage aux groupes homologie/homotopie. Par le Lemme 15, on obtiens le théorème suivant.

**Théorème 17.** *Soient  $X$  un compact de  $\mathbb{R}^d$  et  $L$  un ensemble fini dans  $\mathbb{R}^d$  tel que  $d_{\mathcal{H}}(X, L) < \varepsilon < \frac{1}{4}\text{wfs}(X)$ . Alors, pour tout  $\alpha$  et  $\alpha'$  dans  $[\varepsilon, \text{wfs}(X) - \varepsilon]$  tels que  $\alpha' - \alpha > 2\varepsilon$  et pour tout  $\lambda$  dans  $]0, \text{wfs}(X)[$ , on a, pour tout entier  $k$ , l'isomorphisme suivant :*

$$\mathrm{H}_k(X^\lambda) \cong \mathrm{Im} \left( \mathrm{H}_k(C^\alpha(L)) \xrightarrow{i_*} \mathrm{H}_k(C^{\alpha'}(L)) \right).$$

**Théorème 18.** *Soient  $X$  un compact de  $\mathbb{R}^d$  et  $L$  un ensemble fini dans  $\mathbb{R}^d$  tel que  $d_{\mathcal{H}}(X, L) < \varepsilon < \frac{1}{9}\text{wfs}(X)$ . Alors, pour tout  $\alpha$  dans  $[2\varepsilon, \frac{1}{4}(\text{wfs}(X) - \varepsilon)]$  et pour tout  $\lambda$  dans  $]0, \text{wfs}(X)[$ , on a, pour tout entier  $k$ , l'isomorphisme suivant :*

$$\mathrm{H}_k(X^\lambda) \cong \mathrm{Im} \left( \mathrm{H}_k(R^\alpha(L)) \xrightarrow{i_*} \mathrm{H}_k(R^{4\alpha}(L)) \right).$$

*Démonstration.* On a les inclusions suivantes

$$C^{\alpha/2}(L) \hookrightarrow R^\alpha(L) \hookrightarrow C^\alpha(L) \hookrightarrow C^{2\alpha}(L) \hookrightarrow R^\alpha(L) \hookrightarrow C^{4\alpha}(L).$$

On applique le foncteur  $\mathrm{H}_k$  à ces inclusions :

$$\begin{array}{ccccccc} & & & \varphi_2 & & & \\ & & & \curvearrowright & & & \\ \mathrm{H}_k(C^{\alpha/2}(L)) & \rightarrow & \mathrm{H}_k(R^\alpha(L)) & \rightarrow & \mathrm{H}_k(C^\alpha(L)) & \xrightarrow{\varphi_1} & \mathrm{H}_k(C^{2\alpha}(L)) \rightarrow \mathrm{H}_k(R^\alpha(L)) \rightarrow \mathrm{H}_k(C^{4\alpha}(L)) . \\ & & & \curvearrowleft & & & \\ & & & \psi & & & \end{array}$$

Comme  $\alpha \geq 2\varepsilon$ , les morphismes  $\varphi_1$  et  $\varphi_2$  sont des isomorphismes, donc on a

$$\mathrm{rk}(\varphi_1) = \dim_{\mathbb{R}} \mathrm{H}_k(X^\lambda) = \mathrm{rk}(\varphi_2),$$

ce qui implique que  $\mathrm{rk}(\psi) = \dim_{\mathbb{R}} \mathrm{H}_k(X^\lambda)$ .  $\square$

**Remarque 19.** On a le même type de résultat pour les groupes d'homotopies.

**1.4. Une application.** Considérons un compact  $X$  de  $\mathbb{R}^d$  et  $W$  un ensemble fini (potentiellement avec un cardinal très grand) tel que  $d_{\mathcal{H}}(X, W) < \delta$ . Chazal et al. ont développés un algorithme qui construit point par point, un sous ensemble  $L$  de  $W$  tel que "l'homologie de  $R^{(-)}(L)$  est proche de l'homologie de  $X$ ". Plus précisément, l'algorithme prend en entrée l'ensemble  $W$ , et comprend une boucle sur  $W$  tel qu'à chaque itération, on a l'ensemble  $L(i)$  et la valeur

$$\varepsilon(i) := \max_{w \in W} \min_{v \in L} \|v - w\|$$

décroissante en  $i$ , tels que  $d_{\mathcal{H}}(X, L(i)) < \delta + \varepsilon(i)$ .

**Théorème 20.** *Si  $\delta < \frac{1}{18} \text{wfs}(X)$ . Alors, à chaque itération  $i$  telle que  $\delta < \varepsilon(i) < \frac{1}{18} \text{wfs}(X)$  et pour tout  $\lambda$  dans  $]0, \text{wfs}(X)[$ , on a, pour tout entier  $k$ , l'isomorphisme suivant :*

$$\mathbf{H}_k(X^\lambda) \cong \text{Im} \left( \mathbf{H}_k \left( R^{4\varepsilon(i)}(L(i)) \right) \xrightarrow{i_*} \mathbf{H}_k \left( R^{16\varepsilon(i)}(L(i)) \right) \right).$$

Ainsi, cet algorithme permet d'avoir une approximation discrète  $L$  du compact  $X$  avec un nombre de points raisonnable, qui retient l'information topologique de  $X$ .

## 2. INCLURE DE LA GÉOMÉTRIE

Pour cette section, on considère  $(X, d_X)$  et  $(Y, d_Y)$ , deux espaces métriques. Le but de cette section est de présenter deux méthodes (cf. sections 2.2 et 2.3) permettant de prendre en compte la géométrie de l'objet en utilisant l'homologie persistante (qui ne donne que des renseignements de nature topologique).

### 2.1. Théorèmes de stabilité.

**Définition 21** (Correspondance). Une *correspondance*  $C$  entre deux ensembles  $U$  et  $V$ , noté  $C : U \rightrightarrows V$ , est un sous-ensemble de  $U \times V$  tel que les projections  $\pi_U|_C : C \rightarrow U$  et  $\pi_V|_C : C \rightarrow V$  sont surjectives.

À une correspondance  $C : X \rightrightarrows Y$ , on associe la quantité appelée *distance de distorsion*, donnée par

$$\text{dist}_m(C) := \sup_{(x,y),(x',y') \in C} |d_X(x, x') - d_Y(y, y')|.$$

On a une reformulation de la distance de Gromov-Hausdorff en terme de correspondances.

**Proposition 22.** *On a l'égalité suivante :*

$$d_{\mathcal{GH}}(X, Y) = \frac{1}{2} \inf_{C: X \rightrightarrows Y} \text{dist}_m(C).$$

**Proposition 23.** *On a l'inégalité suivante :*

$$d_i \left( \mathbf{H}_*(R^{(-)}(X)), \mathbf{H}_*(R^{(-)}(Y)) \right) \leq 2d_{\mathcal{GH}}(X, Y),$$

où  $d_i$  désigne la distance d'entrelacement.

**Théorème 24.** *Pour tout espace métrique précompact  $X$  et  $Y$ , les filtrations  $R(X)$ ,  $R(Y)$ ,  $C(X)$  et  $C(Y)$  sont  $\mathfrak{q}$ -tames et on a les inégalités*

$$d_b(\text{dgm}(R(X)), \text{dgm}(R(Y))) \leq 2d_{\mathcal{GH}}(X, Y)$$

et

$$d_b(\text{dgm}(C(X)), \text{dgm}(C(Y))) \leq 2d_{\mathcal{GH}}(X, Y),$$

où  $\text{dgm}(\mathcal{F}(X))$  désigne le code-barre associé à la filtration  $\mathcal{F}(X)$  et  $d_b$  est la distance bottleneck sur les code-barres.

*Démonstration.* Découle de la Proposition 23 et du théorème d'isométrie (cf. exposé de Denis)  $\square$

**2.2. Espaces métriques munis de fonctions.** Dans cette section, on munit chacun des espaces  $X$  et  $Y$  d'une fonction  $f_X : X \rightarrow \mathbb{R}$  et  $f_Y : Y \rightarrow \mathbb{R}$ . On modifie la distance de Gromov-Hausdorff et le complexe de Rips en prenant en compte la donnée de cette fonction.

§ *Stabilité.* À une correspondance  $C : X \rightrightarrows Y$ , on associe la quantité appelée *distorsion fonctionnelle*, donnée par

$$\text{dist}_f(C) := \sup_{(x,y) \in C} |f_X(x) - f_Y(y)|.$$

On raffine la distance de Gromov-Hausdorff par

$$d_{\mathcal{GH}}((X, f_X), (Y, f_Y)) = \frac{1}{2} \inf_{C: X \rightrightarrows Y} \max\{\text{dist}_m(C), \text{dist}_f(C)\}.$$

Au triplet  $(X, d_X, f_X)$ , on associe le complexe de Rips :

$$R^\alpha(X, d_X, f_X) := R^\alpha(f_X^{-1}([\!-\infty, \alpha]), d_X).$$

On a des résultats similaires à la proposition 23 et au théorème 24.

**Théorème 25.** *Soient  $(X, d_X, f_X : X \rightarrow \mathbb{R})$  et  $(Y, d_Y, f_Y : Y \rightarrow \mathbb{R})$  deux espaces précompacts. Alors on a*

$$d_i(\mathbb{H}_*(R^{(-)}(X, f_X)), \mathbb{H}_*(R^{(-)}(Y, f_Y))) \leq 2d_{\mathcal{GH}}((X, f_X), (Y, f_Y)),$$

et si  $f_X$  et  $f_Y$  sont lipschitziennes, alors

$$d_b(\text{dgm}(R^{(-)}(X, f_X)), \text{dgm}(R^{(-)}(Y, f_Y))) \leq 2d_{\mathcal{GH}}((X, f_X), (Y, f_Y)).$$

§ *Exemples de fonctions et extension.* On donne quelques exemples de fonctions :

- la fonction *diamètre*, donnée comme la fonction constante égale à

$$\sup_{x, x' \in X} d_X(x, x');$$

- la fonction *excentricité*, définie par

$$e_\infty^X : X \longrightarrow \mathbb{R} \\ x \longmapsto \sup_{x' \in X} d_X(x, x')$$

Le problème est que la classe de fonctions utilisables est trop petite pour des applications. L'idée de Chazal et al. est de se placer sur des espaces métriques mesurés  $(X, d_X, \mu_X)$  avec  $\mu_X$  une mesure de probabilité borélienne telle que  $\text{supp} \mu_X = X$ . Ils munissent l'ensemble des espaces métriques mesurés munis de fonctions continues à valeur dans  $\mathbb{R}$ , de la distance dite de *Gromov-Wassertein*, notée  $d_{\mathcal{GW}}$ , qui vérifie, pour tout quadruplet  $(X, d_X, \mu_X, f_X : X \rightarrow \mathbb{R})$  et  $(Y, d_Y, \mu_Y, f_Y : Y \rightarrow \mathbb{R})$

$$d_{\mathcal{GH}}((X, d_X, f_X), (Y, d_Y, f_Y)) \leq d_{\mathcal{GW}}((X, d_X, \mu_X, f_X), (Y, d_Y, \mu_Y, f_Y)).$$

Se placer dans ce cadre des espaces métriques mesurés permet de munir nos espaces d'une classe plus large de fonction, comme par exemple, les  $p$ -excentricités : pour  $p \geq 1$ , on pose

$$e_p^X x \longmapsto \left( \int_X d_X(x, x')^p \mu_X(dx') \right)^{\frac{1}{p}}.$$

On a notamment le lemme suivant.

**Lemme 26.** *Pour tout  $p \geq 1$ ,*

$$d_b \left( \text{dgm}(R^{(-)}(X, e_p^X)), \text{dgm}(R^{(-)}(Y, e_p^Y)) \right) \leq 2d_{\mathcal{GH}} \left( (X, e_p^X), (Y, e_p^Y) \right).$$

§ *Aspects algorithmiques.* Calculer  $\text{dgm}(R(-)(X, d_X, f_X))$  est trop couteux en général. On peut simplifier le calcul de deux manière différentes.

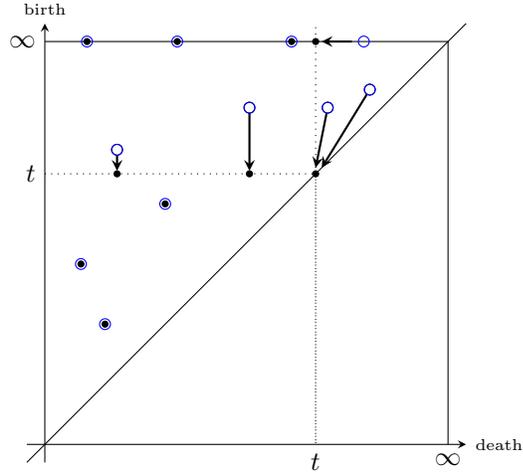
(1) On peut "couper" notre filtration

$$\forall \alpha < \beta, f_X([\!-\infty, \alpha]) \subseteq f_X([\!-\infty, \beta]).$$

À un réel  $t$  appelé *seuil*, on lui associe la filtration tronquée

$$\forall \alpha < \beta \leq t, f_X([\!-\infty, \alpha]) \subseteq f_X([\!-\infty, \beta]) \subseteq f_X^{-1}(\mathbb{R}).$$

On note  $\text{dgm}(\bar{R}^t(X, d_X, f_X))$ , le code barre associé à cette filtration.



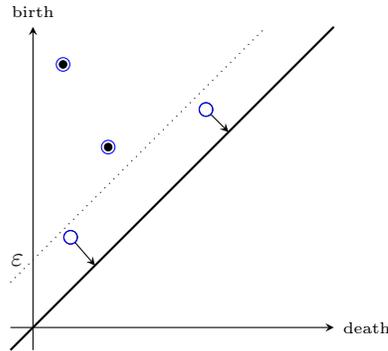
Les points  $\circ$  représentent le code-barre  $\text{dgm}(R(-)(X, d_X, f_X))$  alors que les  $\bullet$  forment le code-barre associé à la filtration tronquée.

Pour deux fonctions continues  $f_X, g_X : X \rightarrow \mathbb{R}$ , on a l'inégalité

$$d_b \left( \text{dgm}(\bar{R}^t(X, f_X)), \text{dgm}(\bar{R}^t(X, g_X)) \right) \leq d_b \left( \text{dgm}(R(X, f_X)), \text{dgm}(R(X, g_X)) \right),$$

donc la troncature de code barre ne fait pas grandir la distance bar-code.

(2) On peut également estimer que les classes du code-barre qui sont proches de la diagonale proviennent de bruit dans nos données. Soit  $\varepsilon > 0$  : on lui associe le code-barre  $\text{dgm}^\varepsilon(R(X))$  donné par le code barre  $\text{dgm}(R(X))$  où l'on projète les classes  $\varepsilon$ -proche de la diagonale sur celle-ci :



Les points  $\circ$  représentent le code-barre  $\text{dgm}(R(-))(X, d_X, f_X)$  alors que les  $\bullet$  forment le code-barre associé à  $\text{dgm}^\varepsilon(R(X))$ . Cette méthode permet notamment de ne garder que les  $N$  classes les plus représentatives.

§ *Application.* On considère  $\mathcal{X}$ , l'ensemble des 60 images 3D (voir Figure 2), et on veut reconnaître les 6 différentes classes de formes. À  $X$  et  $Y$  dans  $\mathcal{X}$ , on associe

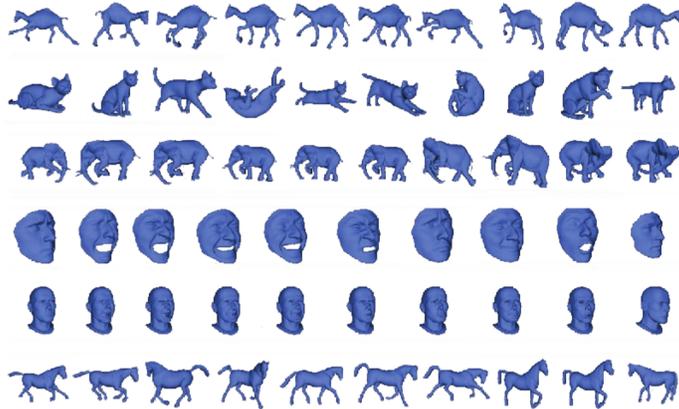


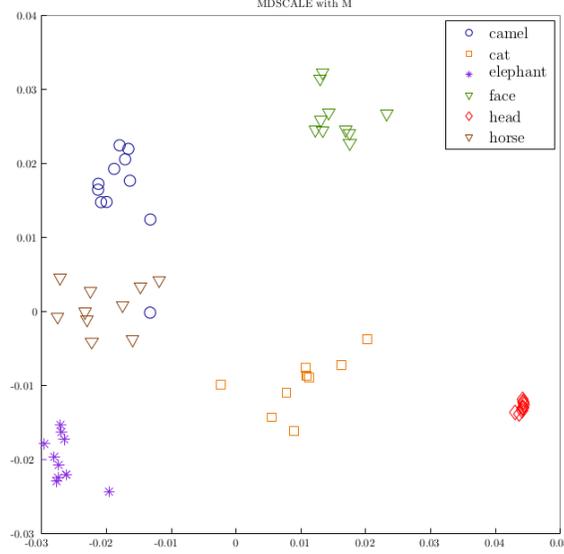
FIGURE 2. Images 3D – cf. [CCSG<sup>+</sup>09]

la quantité  $c(X, Y)$

$$\frac{1}{\text{cst}} \max_{k \in [0, 2], f \in \mathcal{F}} \left\{ d_b(\text{dgm}^{\varepsilon_k(X)}(\bar{R}^t(X, f)), \text{dgm}^{\varepsilon_k(Y)}(\bar{R}^t(Y, f))) - \max\{\varepsilon_k(X), \varepsilon_k(Y)\} \right\}$$

avec  $\mathcal{F} := \cup_{\lambda \in \Lambda} \cup_{p \in \{1, 2, 3, \infty\}} \{\lambda \cdot e_p, \lambda \cdot (\sup(e_p) - e_p)\}$ ,  $\Lambda$  un ensemble de valeur numérique comprises entre 0 et 90 et  $t$  "bien choisi" (cf. [CCSG<sup>+</sup>09, Sect. 7]). À la matrice ayant les  $c(X, Y)$  pour coefficient, on peut lui associer la Figure 3, par une procédure de *metric multidimensional scaling*.

2.3. "**Changer d'espace**". Une autre méthode afin d'inclure de l'information géométrique de  $(X, d_X)$  n'est plus de prendre  $X$  comme espace filtré mais de lui associer un autre espace filtré plus grand. De nouveau, on considère  $X$  comme un

FIGURE 3. Graphique – cf. [CCSG<sup>+</sup>09]

sous-ensemble de  $(\mathbb{R}^n, \|\cdot\|)$  et rappelons que l'on note  $\delta_X$ , la distance

$$\begin{aligned} \delta_X : \mathbb{R}^n &\longrightarrow \mathbb{R}_+ \\ x &\longrightarrow \inf_{y \in X} \|x - y\| \end{aligned}$$

§ *Complexe tangent*. Cette section est basée sur [CZCG05]. On définit une généralisation de l'espace tangent unitaire.

**Définition 27** (Complexe tangent). On pose

$$T^0(X) := \left\{ (x, \xi) \in X \times S^{n-1} \mid \lim_{t \rightarrow 0} \frac{\delta_X(x + t\xi)}{t} = 0 \right\}$$

et on appelle *complexe tangent* à  $X$ , la fermeture dans  $\mathbb{R}^n$  de  $T^0(X)$ , noté  $T(X)$ . Le complexe tangent de  $X$  est muni d'une projection

$$\pi : T(X) \longrightarrow X.$$

**Lemme 28.** Soit  $x$ , un point lisse de  $X$ , i.e. il existe un ouvert  $U$  de  $\mathbb{R}^n$  contenant  $x$  et  $f : U \rightarrow \mathbb{R}^m$  lisse tels que  $U \cap X = f^{-1}(x)$  et, pour tout  $\xi$  dans  $U$ ,  $\text{rk} Df(\xi) = m$ . Alors

$$T(X)_x := \pi^{-1}(x) \cong S^{n-m-1}.$$

**Exemple 29.** On considère  $L = \mathbb{R} \subset \mathbb{R}^2$ . On a  $T(X)_x = \{-1, 1\}$ . Alors  $T(L) \cong L \times S^0$  (voir Figure 4).

**Exemple 30.** On considère  $X = (\mathbb{R}_+ \times \{0\}) \cup (\{0\} \times \mathbb{R}_+)$ . Pour tout point  $x$  de  $X - \{(0, 0)\}$ , on a  $T(X)_x = \{-1, 1\}$ , et  $T(X)_{(0,0)} = \{(\pm 1, 0), (0, \pm 1)\}$ . Finalement, on a

$$T(X) = (\mathbb{R}_+ \times 0) \times \{(\pm 1, 0)\} \cup (0 \times \mathbb{R}_+) \times \{(0, \pm 1)\}$$

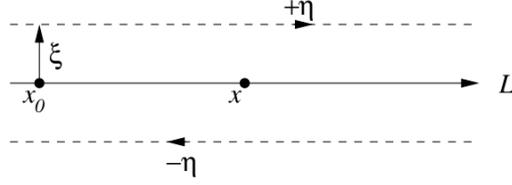


FIGURE 4. Complexe tangent d'une droite dans le plan – cf. [CZCG05]

(voir Figure 5). Il comprend 4 composantes contractiles, contrairement à celui de la droite, qui n'en comprend que deux. La topologie du complexe tangent retient donc la singularité de l'ensemble.

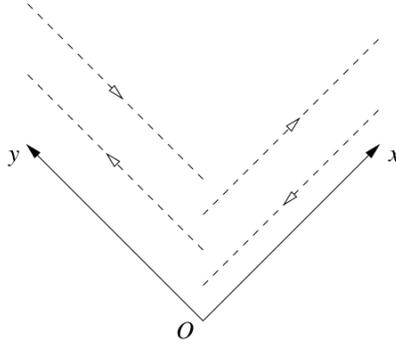


FIGURE 5. Complexe tangent de demi droite – cf. [CZCG05]

§ *Complexe tangent filtré.*

**Définition 31** (Cercle osculateur généralisé). Un point  $(x, \xi)$  de  $T_0(X)$  possède un cercle osculateur s'il existe  $x_0$  dans  $\mathbb{R}^n$  tel que, pour tout  $v$  tel que  $(x, v)$  soit un point de  $T_0(X)$   $(x - x_0) \cdot v = 0$ , et

$$\lim_{\theta \rightarrow 0} \frac{\delta_X(x + \theta + \cos \theta \cdot (x - x_0) + \sin \theta \cdot |x - x_0| \cdot \xi)}{\theta^2} = 0.$$

On note  $x_0$  est le centre du cercle et  $\rho(x, \xi) := |x - x_0|$  est le rayon du cercle

**Définition 32** (Complexe tangent filtré). On définit le complexe filtré  $T_\bullet(X)$ , défini, pour tout réel  $\delta$ , comme la fermeture  $\overline{T_\delta^0(X)} \subseteq \mathbb{R}^n \times S^{n-1}$  avec

$$T_\delta^0(X) = \left\{ (x, \xi) \in T^0(X) \mid \frac{1}{\rho(x, \xi)} \leq \delta \right\}.$$

**Exemple 33** (Cercle et ellipse). On considère  $X$ , un cercle de rayon  $R$ . On a alors

$$T_\delta(X) = \begin{cases} 0 & \text{si } \delta \leq \frac{1}{R}, \\ T(X) & \text{si } \delta \geq \frac{1}{R}. \end{cases}$$

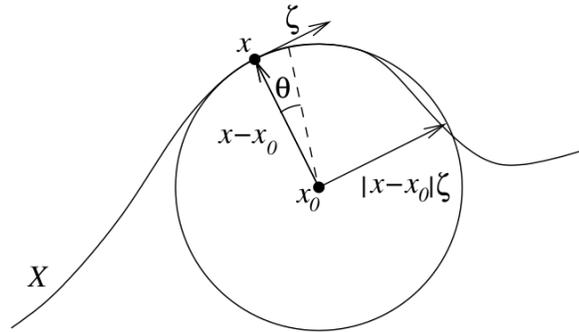


FIGURE 6. Cercle osculateur – cf. [CZCG05]

(voir Figure 6). On considère  $Y$ , l'ellipse de paramétrisation

$$(a \cos(t), b \sin(t)),$$

avec  $b > a > 0$ . Le complexe filtré  $T_{\bullet}(Y)$  est différent de celui du cercle (cf. Figure 8). Les lignes bleues représentent les différentes classes du  $H_0$  et les lignes rouges représentent les différentes classes du  $H_1$ .

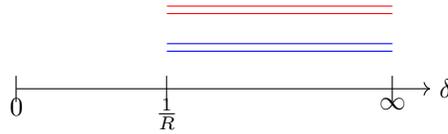


FIGURE 7. Homologie du complexe filtré du cercle

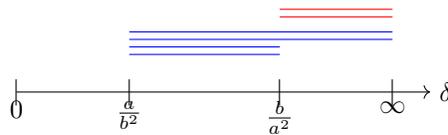


FIGURE 8. Homologie du complexe filtré de l'ellipse

#### RÉFÉRENCES

- [CCSG<sup>+</sup>09] Frédéric Chazal, David Cohen-Steiner, Leonidas J Guibas, Facundo Mémoli, and Steve Y Oudot. Gromov-hausdorff stable signatures for shapes using persistence. In *Computer Graphics Forum*, volume 28, pages 1393–1403. Wiley Online Library, 2009. 9, 10
- [CL05] Frédéric Chazal and André Lieutier. The “ $\lambda$ -medial axis”. *Graphical Models*, 67(4) :304–331, 2005. 2
- [CO08] Frédéric Chazal and Steve Yann Oudot. Towards persistence-based reconstruction in euclidean spaces. In *Proceedings of the twenty-fourth annual symposium on Computational geometry*, pages 232–241. ACM, 2008. 1

- [CZCG05] GUNNAR CARLSSON, AFRA ZOMORODIAN, ANNE COLLINS, and LEONIDAS J. GUIBAS. Persistence barcodes for shapes. *International Journal of Shape Modeling*, 11(02) :149–187, 2005. 10, 11, 12
- [Hat02] Allen Hatcher. Algebraic topology. 2002. *Cambridge UP, Cambridge*, 606(9), 2002. 4
- [Oud15] Steve Y Oudot. *Persistence theory : from quiver representations to data analysis*, volume 209. American Mathematical Society Providence, RI, 2015. 1

LAGA, UNIVERSITÉ PARIS 13, 99 AVENUE JEAN BAPTISTE CLÉMENT 93430, VILLETANEUSE,  
FRANCE

*Email address:* `leray@math.univ-paris13.fr`