

**Applications de l'homologie persistante pour la reconnaissance
des formes**

par

HAMDI CHAIMA

Mémoire présenté au Département de mathématiques en vue de l'obtention
du grade de maître ès sciences (M.Sc.)

FACULTÉ DES SCIENCES
UNIVERSITÉ DE SHERBROOKE

Sherbrooke, Québec, Canada, juillet 2017

Le 19 juillet 2017

*le jury a accepté le mémoire de Madame Chaima Hamdi
dans sa version finale.*

Membres du jury

Professeur Tomasz Kaczynski
Directeur de recherche
Département de mathématiques

Professeur Madjid Allili
Membre interne
Département d'informatique

Professeure Vasilisa Shramchenko
Présidente-rapporteuse
Département de mathématiques

SOMMAIRE

L'homologie persistante est un outil fondamental dans la topologie computationnelle. Cette méthode est utilisée pour reconnaître et comparer les formes. Dans ce travail nous étudions d'abord l'homologie persistante dans le cas unidimensionnel d'ordre 0 qu'on appelle aussi fonction de taille. Nous présentons une démonstration du fait que toute fonction de taille peut être représentée comme un ensemble de points et de lignes dans le plan réel, avec des multiplicités. Cela permet une approche algébrique aux fonctions de taille et la construction de nouvelles pseudo distances entre les fonctions de taille pour comparer les formes. Nous calculons ensuite l'homologie persistante unidimensionnelle d'ordre n avec différentes méthodes de filtration de l'espace correspondant à l'histoire d'un complexe croissant. Nous classons un changement topologique qui se produit pendant la croissance soit comme une caractéristique ou un bruit, en fonction de sa durée de vie ou de sa persistance dans la filtration. Une présentation avec des codes barres affiche alors la persistance de ces invariants.

L'homologie persistante multidimensionnelle nous permet de soutirer plus d'informations sur les formes en utilisant la fonction de filtration avec des valeurs dans \mathbb{R}^k . Pour fournir un descripteur de forme concis et complet dans le cas multidimensionnel nous réduisons le calcul de l'homologie persistante multidimensionnelle au calcul de l'homologie persistante ordinaire pour une famille paramétrée de fonctions à valeur dans \mathbb{R} .

REMERCIEMENTS

Je tiens d'abord à remercier mon directeur de maîtrise, le professeur Tomasz Kaczynski, pour m'avoir guidée, encouragée et conseillée tout au long de ma formation riche en rebondissements et en rencontres mûrissantes tant sur le plan personnel que sur le plan scientifique. J'ai été extrêmement sensible à ses qualités humaines de gentillesse, d'écoute et de compréhension.

Je remercie les membres de mon jury pour avoir accepté d'évaluer ce travail. Je remercie également tous mes professeurs et mes collègues des départements de Mathématiques et d'Informatique.

Mes remerciements vont aussi à mes amis et toutes les personnes qui m'ont accompagnée et soutenue tout au long de ces années d'études. Je pense tout particulièrement à mes chers parents, mes frères et soeurs.

TABLE DES MATIÈRES

SOMMAIRE	iii
REMERCIEMENTS	iv
TABLE DES MATIÈRES	v
LISTE DES FIGURES	vii
INTRODUCTION	1
CHAPITRE 1 — Fonction de taille 1-dimensionnelle et distance d'appariement	4
1.1 Fonction de taille unidimensionnelle et série formelle	4
1.2 Distance d'appariement d'une fonction de taille unidimensionnelle	9
CHAPITRE 2 — Homologie persistante	13
2.1 Homologie simpliciale	14
2.1.1 Les complexes simpliciaux	14

2.1.2	Groupes d'homologie	16
2.2	La persistance	20
2.2.1	Filtration	20
2.2.2	Méthode de filtration	21
2.2.3	Formulation algébrique d'homologie persistante	25
2.2.4	Module persistant	26
2.3	Visualisation de l'homologie persistante	37
 CHAPITRE 3 — L'homologie persistante multidimensionnelle et la stabilité		
	lité	40
3.1	Réduction du cas multidimensionnel au cas unidimensionnel	42
3.2	Distance d'appariement des invariants de rang multidimensionnels	44
3.3	Stabilité des nombres de Betti persistants par rapport aux perturbations du domaines	50
 CHAPITRE 4 — Applications		55
 CONCLUSION		69
 BIBLIOGRAPHIE		71

LISTE DES FIGURES

1.1	(a) L'objet à étudier est la courbe X représentée par une ligne solide, et ses caractéristiques sont étudiées par la fonction de mesure $\varphi(P) = d(P, B)$. (b) La fonction de taille $\ell_{\langle X, \varphi \rangle}$ [1].	6
1.2	Fonction de taille représentant des points et lignes d'angles [1].	8
1.3	(a) Fonction de taille correspondant à la série formelle $r + a + b + c$. (b) Fonction de taille correspondant à la série formelle $r' + a' + c$. (c) La distance d'appariement entre les deux fonctions de taille [22].	11
1.4	(a) Fonction de taille correspondant à la série formelle $r + a + b$. (b) Fonction de taille correspondant à la série formelle $r' + a'$. (c) La distance d'appariement entre les deux fonctions de taille [21].	11
2.1	n -simplexes, pour chaque $0 \leq n \leq 3$ [6].	14
2.2	La surface d'un tétraèdre est une triangulation d'une sphère.	15
2.3	n -simplexes, $0 \leq n \leq 3$. L'orientation sur le tétraèdre est représentée sur ses faces [6].	16
2.4	n -simplexes, $0 \leq n \leq 3$. Un complexe de chaîne pour un complexe de dimension trois.	18

2.5	Triangulation d'un disque.	19
2.6	Un tore, avec ses deux cycles d'ordre 1 indiqués en ligne verticale et horizontale.	20
2.7	Filtration d'un complexe simplicial [6].	21
2.8	Exemple d'un nuage de points (à gauche) et d'un choix correspondant de boules ($\epsilon/2$). Pour obtenir le complexe de Čech, nous ajoutons un k -simplexe chaque fois que nous voyons un sous-ensemble de k points avec intersection commune. [9].	22
2.9	L'union des disques est décomposée en régions convexes par les cellules de Voronoï. Le complexe alpha correspondant est superposé. [7].	23
2.10	Les inégalités $p \geq 0$, $l \geq i$ et $l + p < j$ définissent une région triangulaire dans le plan d'index et de persistance. Cette région définit quand le cycle est un élément de base pour l'espace vectoriel d'homologie [6].	33
2.11	La forme d'une matrice échelonnée selon les colonnes. Une * indique une valeur non nulle et les pivots sont encadrés.	35
2.12	Visualisation par code barres et triangles du complexe simplicial de la Figure 2.7. Les triangles de s et de stw sont non bornés et non dessinés. Les triangles clairs et sombres représentent respectivement les 0-cycles et 1-cycles.	38
2.13	Visualisation par digramme de persistance qui présente les points de pair (naissance, mort).	39
2.14	Visualisation par diagramme de Pranav. La région ombrée transforme en panneau à une région en forme de V déterminée par $(v + v, 0)$. Les bras du V ont une pente -1 et 1 respectivement.	39

3.1	Un espace topologique étudié par une fonction de mesure unidimensionnelle et bidimensionnelle. L'espace topologique filtré par fonction de mesure bidimensionnelle montre plus de détails dans son diagramme de persistance.	41
3.2	Réduction des nombres de Betti persistants bidimensionnels pour une fonction $\vec{\phi}$ au cas unidimensionnel [22].	43
3.3	L'espace topologique X et Y et $\beta_{(X, F_{(\vec{l}, \vec{b})}^{\vec{\phi}})}, \beta_{(Y, F_{(\vec{l}, \vec{b})}^{\vec{\phi}})}$ associé au demi-plan $\pi_{(\vec{l}, \vec{b})}$ [19].	49
3.4	Deux images binaires d'une pieuvre. L'image à droite est une version bruitée de celle de gauche [21].	52
3.5	À gauche : le diagramme de persistance correspondant à l'image originale de la pieuvre. À droite : le diagramme de persistance correspondant à l'image de la pieuvre bruitée [21].	53
3.6	(a) Les nombres de Betti persistants $\beta_{\vec{\phi}_1}$ restreints au demi plan $\pi_{(\vec{l}, \vec{b})}$. (b) Les nombres de Betti persistants $\beta_{\vec{\phi}_2}$ restreints au même demi plan. (c)-(d) Le même (a) et (b) respectivement multiplier par $\mu = \min \{l_1, l_2\}$ [21].	54
4.1	Octaèdre régulier et son diagramme de persistance en dimensions 0,1 et 2.	55
4.2	(a) Nervures opposées. (b) Diagramme de persistance. (c) Visualisation du diagramme de persistance.	57
4.3	(a) Nervures alternées. (b) Diagramme de persistance. (c) Visualisation du diagramme de persistance.	57
4.4	(a) Nervures palmées. (b) Diagramme de persistance. (c) Visualisation du diagramme de persistance.	58

4.5	(a) Nervures opposées moins une nervure. (b) Diagramme de persistance. (c) Visualisation du diagramme de persistance.	58
4.6	(a) Nervures opposées moins une nervure différente. (b) Diagramme de persistance. (c) Visualisation de diagramme de persistance.	58
4.7	(a) Nervures opposées moins deux nervures. (b) Diagramme de persistance. (c) Visualisation du diagramme de persistance.	59
4.8	(a) Nervures opposées moins trois nervures. (b) Diagramme de persistance. (c) Visualisation du diagramme de persistance.	59
4.9	(a) Plante avec deux feuilles. (b) Diagramme de persistance. (c) Visualisation de diagramme de persistance en dimension 0. (d) Visualisation du diagramme de persistance en dimension 1.	61
4.10	(a) Plante avec quatre feuilles. (b) Diagramme de persistance. (c) Visualisation du diagramme de persistance en dimension 0. (d) Visualisation du diagramme de persistance en dimension 1.	61
4.11	(a) Plante avec cinq feuilles. (b) Diagramme de persistance. (c) Visualisation du diagramme de persistance en dimension 0. (d) Visualisation du diagramme de persistance en dimension 1.	62
4.12	(a) Image d'os. (b) Diagramme de persistance. (c) Visualisation du diagramme de persistance.	64
4.13	(a) Image d'os différente. (b) Diagramme de persistance. (c) Visualisation du diagramme de persistance.	64
4.14	(a) Image de papillon. (b) Diagramme de persistance. (c) Visualisation du diagramme de persistance en dimension 0. (d) Visualisation du diagramme de persistance en dimension 1.	65

4.15	(a) Image de papillon différente. (b) Diagramme de persistance. (c) Visualisation du diagramme de persistance en dimension 0. (d) Visualisation du diagramme de persistance en dimension 1.	65
4.16	Tore et son diagramme de persistance.	66
4.17	(a) Visualisation du diagramme de persistance en dimension 0. (b) Visualisation du diagramme de persistance en dimension 1.	66
4.18	Double tore et son diagramme de persistance.	67
4.19	(a) Visualisation du diagramme de persistance en dimension 0. (b) Visualisation du diagramme de persistance en dimension 1.	67

INTRODUCTION

Les méthodes topologiques sont devenues très utiles pour résoudre des problèmes d'imagerie numérique, comme la reconnaissance des formes, l'analyse d'images et de nuages de points de tailles et de dimensions quelconques. La topologie algébrique applique des outils de l'algèbre à l'étude des espaces topologiques. Plus précisément, elle associe des objets algébriques à des structures topologiques, donnant ainsi des informations sur les invariants topologiques de ces espaces.

L'homologie persistante [7] nous permet d'extraire des informations sur la forme d'un objet en étudiant sa topologie à l'aide d'une fonction continue, dite fonction de mesure ou de filtration. En d'autres termes, on traverse un objet (espace topologique) avec une fonction continue en considérant, pour chaque valeur u , le sous-espace où la fonction prend des valeurs inférieures à u , et en étudiant combien de temps les invariants topologiques de ce sous-espace persistent.

Dans les applications, nous travaillons sur différents objets. Le calcul de l'homologie persistante avec la propriété de stabilité nous permet de détecter les différences entre les objets. En effet, un simple changement dans la fonction d'entrée produit un changement dans le diagramme de persistance. Des invariants plus persistants sont détectés sur une large gamme et sont considérés comme plus susceptibles de représenter les caractéristiques réelles de l'espace sous-jacent, plutôt que du bruit.

La comparaison des formes est reliée directement avec les méthodes de signatures. En

effet, en fonction du choix de construction de ces signatures (fonction de mesure, d'ordre d'homologie choisi,...ect) nous obtenons des résultats très différents.

Dans le chapitre 1, nous présentons une approche possible – la fonction de taille – pour résoudre le problème de description des formes [1]. Nous prouvons que toute fonction de taille peut être représentée comme un ensemble de points et de lignes dans le plan réel, avec des multiplicités. En changeant les fonctions de mesure, les fonctions de taille correspondantes fournissent différentes descriptions de la forme donnée. Enfin, nous montrons la correspondance bijective entre les fonctions de taille et une classe de séries formelles. Ces résultats nous permettent de décrire la définition de distance d'appariement [3] à la fin de ce chapitre.

Dans le chapitre 2, nous commençons par un rappel sur les complexes simpliciaux et l'homologie simpliciale, dont nous avons besoin pour définir l'homologie persistante. Nous présentons ensuite plusieurs façons pour construire un complexe simplicial (filtration) : la méthode de Rips [9], la méthode de Alpha shape [11], etc. Une formulation algébrique d'homologie persistante sera décrite dans ce chapitre afin de calculer l'homologie persistante. Une vision différente de l'homologie persistante dans le but de comprendre sa structure sera présentée dans la section module persistant [12]. Enfin, nous réalisons une visualisation de l'homologie persistante avec des codes barres, chaque intervalle représentant la durée de vie des attributs topologiques. Nous concluons ce chapitre par une visualisation des attributs dans un plan avec paire (naissance et mort), très utile dans le cas d'un nombre élevé de données.

Le chapitre 3 présente le cas de fonction de taille multidimensionnelle [19]. En effet, la fonction de filtration ou de mesure dans ce cas prend ses valeurs dans \mathbb{R}^k . Le fait que la fonction de mesure est multidimensionnelle, nous permet de tirer beaucoup d'informations sur la forme. Une conséquence très importante de cette filtration multidimensionnelle est la stabilité. Nous présentons plusieurs théorèmes pour se servir de cette stabilité. Nous finissons ce chapitre par l'étude de la stabilité par rapport aux ensembles bruités [23].

Dans le chapitre 4, nous présentons deux expériences. Dans la première nous avons appliqué la méthode de filtration de Rips à plusieurs et différentes formes présentées comme des nuages de points. Nous calculons avec le programme Javaplex [15] l'homologie persistante sur nos formes afin d'établir les codes barres. Nous visualisons chaque code barre sur un plan avec paire (naissance, mort). Enfin, nous calculons la distance entre ces formes et nous examinons la stabilité. Nous testons si un petit changement dans les formes cause des changements dans les codes barres. Dans la deuxième expérience, nous appliquons l'homologie persistante sur des images réelles. Cette fois-ci, nous choisissons la filtration avec la méthode de Lazy Witness [10] en raison de la taille importante des données. Nous concluons ce chapitre par le calcul de la stabilité entre le tore et le double tore.

CHAPITRE 1

Fonction de taille 1-dimensionnelle et distance d'appariement

Dans ce chapitre, nous considérons un outil mathématique qui nous permet de décrire et de comparer les formes des espaces topologiques.

Nous commençons par la définition de la fonction de taille et nous représentons cette dernière par un ensemble de points et de lignes dans le plan réel, avec la multiplicité [1]. Cela permet une approche algébrique des fonctions de taille et la construction d'une pseudo-distance [2] entre les fonctions de taille pour comparer les formes.

1.1 Fonction de taille unidimensionnelle et série formelle

Soit X un espace topologique, nous appellerons *fonction de mesure* la fonction continue $\varphi : X \rightarrow \mathbb{R}$.

Définition 1.1. Étant donné $y \in \mathbb{R}$ et une fonction de mesure φ , deux points P et Q dans X sont dits $\langle \varphi \leq y \rangle$ -**homotopes** si et seulement si soit $P = Q$, soit le chemin $\gamma : [0, 1] \rightarrow X$ relie P et Q dans X avec $\varphi(\gamma(t)) \leq y$ pour $t \in [0, 1]$.

Si P et Q sont $\langle \varphi \leq y \rangle$ -homotopes, nous écrivons $P \cong_{\varphi \leq y} Q$ et appelons $\gamma : \langle \varphi \leq y \rangle$ -homotopie de P à Q .

Pour chaque nombre réel x , nous définirons $X\langle \varphi \leq x \rangle := \{P \in X : \varphi(P) \leq x\}$. nous voyons que, pour $x, y \in \mathbb{R}$ la relation de $\langle \varphi \leq y \rangle$ -homotopie est une relation d'équivalence dans $X\langle \varphi \leq x \rangle$.

Définition 1.2. Soit $x \in \mathbb{R}$ et $X\langle \varphi \leq x \rangle$ l'ensemble $\{P \in X : \varphi(P) \leq x\}$. Nous considérons la fonction $\ell_{\langle X, \varphi \rangle} : \mathbb{R}^2 \rightarrow \mathbb{N} \cup \{\infty\}$ définie en mettant $\ell_{\langle X, \varphi \rangle}(x, y)$ égale au nombre de classes d'équivalences dans lesquelles $X\langle \varphi \leq x \rangle$ est divisé par la relation d'équivalence de $\langle \varphi \leq y \rangle$ -homotopie. Nous appellerons $\ell_{\langle X, \varphi \rangle}$ **fonction de taille** associée à la paire de taille (X, φ) .

Formellement : pour $x \leq y$, et $P, Q \in X\langle \varphi \leq x \rangle$, $p \sim Q \Leftrightarrow P, Q$ appartient à la composante connexe de $X\langle \varphi \leq y \rangle$

$$\ell_{\langle X, \varphi \rangle}(x, y) = \# [X\langle \varphi \leq x \rangle / \sim^y].$$

Afin de rendre cette définition claire, nous présentons l'exemple suivant :

Exemple 1.3. Nous renvoyons le lecteur à la Figure 1.1. L'objet à étudier est la courbe X représentée par une ligne continue, et la fonction de mesure φ est la distance euclidienne de chaque point par rapport au barycentre $B : \varphi(P) = d(P, B)$ pour chaque P dans X . La fonction de taille $\ell_{\langle X, \varphi \rangle}$ est illustrée à droite. Comme on peut le voir, le domaine $\{(x, y) \in \mathbb{R}^2 : x < y\}$ est divisé en régions (qui peuvent être bornées ou non bornées), dans lesquels la valeur de la fonction de taille est constante. Les nombres affichés correspondent aux valeurs prises par la fonction de taille de chaque région. Par exemple, pour $a \leq x <$

$y < b$, l'ensemble $X\langle\varphi \leq x\rangle$ est constitué de deux composantes connexes qui ne peuvent pas être jointes sous y . Donc $\ell_{\langle X, \varphi \rangle}(x, y) = 2$ pour $a \leq x < y < b$. Dans le domaine $a \leq x < b$ et $y \geq b$, les deux composantes connexes de $X\langle\varphi \leq x\rangle$ peuvent être jointes sous niveau de y , donc $\ell_{\langle X, \varphi \rangle}(x, y) = 1$.

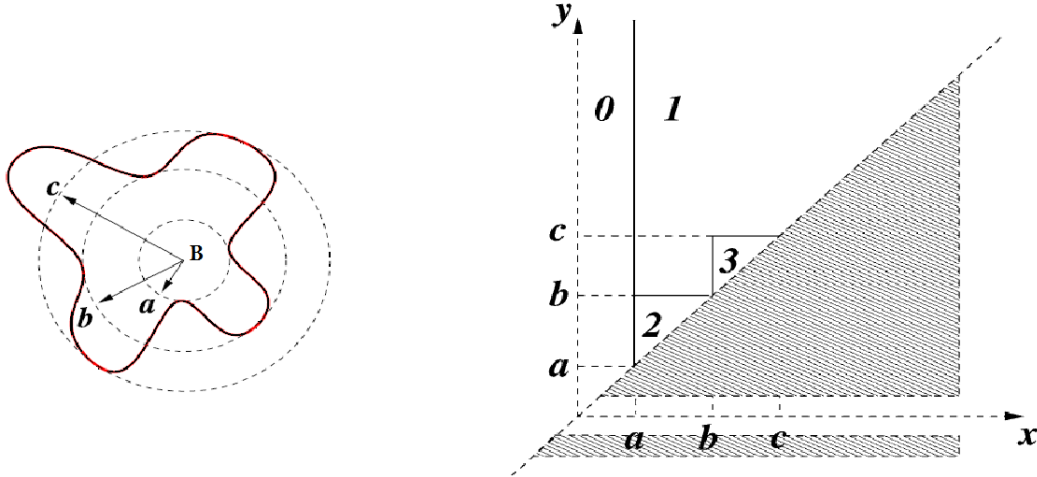


Figure 1.1 – (a) L’objet à étudier est la courbe X représentée par une ligne solide, et ses caractéristiques sont étudiées par la fonction de mesure $\varphi(P) = d(P, B)$. (b) La fonction de taille $\ell_{\langle X, \varphi \rangle}$ [1].

Nous allons maintenant fournir quelques propositions indiquant certaines propriétés simples des fonctions de taille.

Proposition 1.4. [1] $\ell_{\langle X, \varphi \rangle}(x, y)$ est non décroissante dans la variable x et non croissante dans la variable y .

Proposition 1.5. [1] $\ell_{\langle X, \varphi \rangle}(x, y) < +\infty$ pour $x < y$.

Proposition 1.6. [1] La fonction de taille ne transmet pas d’informations sur la forme en dehors de la région $\{(x, y) \in \mathbb{R}^2 : x < y, x \geq \min_{P \in X} \varphi(P)\}$.

Proposition 1.7. [1] Pour chaque $y \geq \max_{P \in X} \varphi(P)$, $\ell_{\langle X, \varphi \rangle}(x, y)$ est égale au nombre de composantes connexes par arc N de X tel que $x \geq \min_{P \in N} \varphi(P)$.

Série formelle

La fonction de taille contient des informations sur la forme étudiée. D'un point de vue algébrique, il s'agit de points et de lignes de discontinuités. L'objet formel que nous obtiendrons permettra des calculs plus facile et le stockage de l'information réduite. Pour ce faire, nous allons représenter une fonction de taille en donnant la liste des régions triangulaires en lesquelles son domaine est divisé par des discontinuités. En outre, nous allons attribuer une multiplicité à chaque triangle de sorte que la somme des multiplicités des triangles contenant un point donné donne la valeur de la fonction de taille à ce point.

On note \mathcal{S}_0 l'ensemble $\{(x, y) \in \mathbb{R}^2 : x < y\}$ et \mathcal{R} l'ensemble des lignes verticales de \mathbb{R}^2 . De plus, nous appellerons \mathcal{X} l'ensemble $\mathcal{S}_0 \cup \mathcal{R}$.

Définition 1.8. *Toute fonction $m : \mathcal{X} \rightarrow \mathbb{Z}$ est considérée comme une série formelle dans \mathcal{X} . Nous appelons l'ensemble $\text{supp}(m) = \{K \in \mathcal{X} : m(K) \neq 0\}$ support de m .*

L'ensemble des séries formelles en \mathcal{X} est un groupe commutatif par rapport à la somme habituelle de fonctions.

On dénote alors chaque série formelle m par le symbole $\sum_{K \in \mathcal{X}} m(K)K$.

Définition 1.9. *Pour tout $p = (x, y) \in \mathbb{R}^2$ et α, β nombres réels positifs avec $(x + \alpha, y - \beta) \in \mathcal{S}_0$:*

$$\mu_{\alpha, \beta}(p) = \ell_{\langle X, \varphi \rangle}(x + \alpha, y - \beta) - \ell_{\langle X, \varphi \rangle}(x + \alpha, y + \beta) - \ell_{\langle X, \varphi \rangle}(x - \alpha, y - \beta) + \ell_{\langle X, \varphi \rangle}(x - \alpha, y + \beta).$$

*Le nombre $\mu(p) := \min \{\mu_{\alpha, \beta}(p) : \alpha, \beta > 0, x + \alpha < y - \beta\}$ sera appelé **multiplicité** de p pour $\ell_{\langle X, \varphi \rangle}$. De plus, nous appellerons **point d'angle** de $\ell_{\langle X, \varphi \rangle}$ chaque point $p \in \mathcal{S}_0$ tel que le nombre $\mu(p)$ est strictement positif.*

Définition 1.10. *Pour chaque ligne verticale $[r : x = k] \in \mathcal{R}$ nous appellerons **multiplicité** de r pour $\ell_{\langle X, \varphi \rangle}$ le nombre*

$$\mu(r) := \min_{\alpha > 0, k + \alpha < y} \ell_{\langle X, \varphi \rangle}(k + \alpha, y) - \ell_{\langle X, \varphi \rangle}(k - \alpha, y).$$

Lorsque $\mu(r)$ est strictement positif, la ligne r sera dite être une ligne d'angle pour la fonction de taille.

Exemple 1.11. Dans la fonction de taille représentée sur la Figure 1.2 les seuls points d'angles (*cornerpoints*) sont p_1, p_2, p_4 et p_5 avec multiplicité $\mu(p_1) = \mu(p_4) = \mu(p_5) = 1$ et $\mu(p_2) = 3$. Le point p_3 n'est pas un point d'angle puisque $\mu(p_3) = 0$.

Les lignes d'angles (*cornerlines*) dans la fonction de taille représentée dans Figure 1.2 sont données par r_1 et r_2 avec multiplicité $\mu(r_1) = 1$ and $\mu(r_2) = 2$.

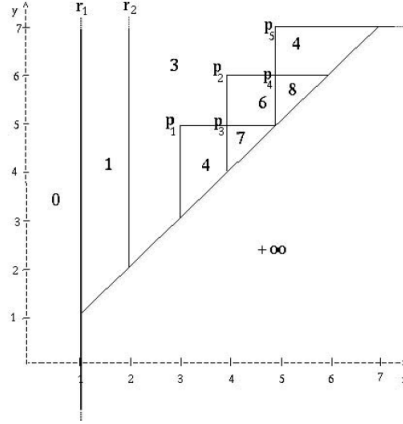


Figure 1.2 – Fonction de taille représentant des points et lignes d'angles [1].

On conclut cette section par un théorème.

Tout d'abord, pour chaque $\rho \geq 0$ nous définissons la relation d'équivalence \cong_ρ sur l'ensemble de toutes les fonctions de taille, nous mettons $\ell_{\langle X_1, \varphi_1 \rangle} \cong_\rho \ell_{\langle X_2, \varphi_2 \rangle}$ si et seulement si $\ell_{\langle X_1, \varphi_1 \rangle}$ et $\ell_{\langle X_2, \varphi_2 \rangle}$ coïncident presque partout dans \mathcal{S}_ρ (en d'autres termes, le sous-ensemble de \mathcal{S}_ρ dans lequel les deux fonctions diffèrent a une mesure de Lebesgue nulle), où

$$\mathcal{S}_\rho = \{(x, y) \in \mathbb{R}^2 : x < y - \rho\}, \forall \rho > 0.$$

Ci-après, pour chaque nombre réel $\rho \geq 0$, nous dénoterons \mathcal{L}_ρ le quotient de l'ensemble de toutes les fonctions de taille par la relation d'équivalence \cong_ρ .

Soit Ω_ρ , avec $\rho > 0$, l'ensemble de toutes les valeurs naturelles de la série formelle σ dans $\mathcal{S}_\rho \cup \mathcal{R}$, ayant un support fini et la vérification de la propriété suivante : il y a une ligne $[r : x = k]$ en $\text{supp}(\sigma)$ de telle sorte que le demi-plan $\{(x, y) \in \mathbb{R}^2 : x \geq k\}$ contient tous les autres points et lignes dans $\text{supp}(\sigma)$.

On dénotera Ω_0 comme étant l'ensemble de toutes les valeurs naturelles de la série formelle σ en $\mathcal{S}_0 \cup \mathcal{R}$ (éventuellement avec $\text{card}(\text{supp}(\sigma)) = \infty$), ayant les propriétés suivantes :

- 1) $\forall \rho > 0$ la restriction de $\sigma \in \Omega_0$ à $\mathcal{S}_\rho \cup \mathcal{R}$ est une série appartenant à Ω_ρ ;
- 2) $\text{sup} \{y : (x, y) \in \mathcal{S}_0 \cap \text{supp}(\sigma)\} < \infty$ (i.e. le suprémum y_{sup} des ordonnées des points en σ est fini).

Il est évident maintenant de définir une application $\tilde{\alpha}_\rho : \mathcal{L}_\rho \rightarrow \Omega_\rho$ pour chaque nombre réel $\rho \geq 0$.

Théorème 1.12. *Quel que soit le nombre réel $\rho \geq 0$ l'application $\tilde{\alpha}_\rho : \mathcal{L}_\rho \rightarrow \Omega_\rho$ est une bijection.*

Les descriptions de forme fournies par les fonctions de taille peuvent être comparées à l'aide d'une distance appropriée, appelée distance d'appariement. De cette façon, l'évaluation de la similitude entre des objets est obtenue en calculant la distance d'appariement entre les fonctions de taille correspondantes, tel que présenté dans [3].

1.2 Distance d'appariement d'une fonction de taille unidimensionnelle

Comme nous avons vu dans la section précédente, nous pouvons réduire le problème de comparaisons des fonctions de taille à la comparaison des séries formelles.

Soit $\rho > 0$. Considérons deux fonctions de taille ℓ_1 et ℓ_2 . L'application $\tilde{\alpha}_\rho$ du théorème 1.12 prend ces deux fonctions de taille en deux séries formelles σ_1 et σ_2 avec support fini dans $\mathcal{S}_\rho \cup \mathcal{R}$. Il en résulte que toute distance d'appariement entre σ_1 et σ_2 induit une distance d'appariement entre ℓ_1 et ℓ_2 .

$$d(\ell_1, \ell_2) = \min_\gamma \max_{p \in \sigma_1} \delta(p, \gamma(p))$$

où γ est pris parmi toutes les bijections entre les points de σ_1 et σ_2 et où

$$\begin{aligned} \delta((x, y), (x', y')) &= \min \left\{ \max \{ |x - x'|, |y - y'| \}, \max \left\{ \frac{y-x}{2}, \frac{y'-x'}{2} \right\} \right\}, \\ \delta((x, y), \Delta) &= \frac{y-x}{2}, \delta(\Delta, \Delta) = 0. \end{aligned}$$

Dans cet énoncé, les lignes avec l'équation $x = k$ sont identifiées aux points « à l'infini » avec les coordonnées (x, ∞) .

Dans le cas de $\max \left\{ \frac{y-x}{2}, \frac{y'-x'}{2} \right\}$ où $x, x' = \infty$ et $y, y' \neq \infty$ on utilise la convention notationnelle suivante : $y - \infty = \infty$ et $y' - \infty = \infty$, si $x, y = \infty$ on utilise les conventions notationnelles suivantes : $\infty - \infty = 0$ et $\infty/2 = \infty$. Dans le cas de $\max \{ |x - x'|, |y - y'| \}$ où $y' \neq \infty$ on utilise la convention notationnelle suivante : $|\infty - y'| = |\infty|$ et $|\infty| = \infty$. Pour les cas $\min \{c, \infty\}$, $\max \{c, \infty\}$ on utilise les conventions notationnelles suivantes : $\min \{c, \infty\} = c$ et $\max \{c, \infty\} = \infty$ (voir [21]).

La distance d'appariement représente le coût minimal pour déplacer tous les points de σ_1 sur ceux de σ_2 . Étant donné que les deux fonctions de taille peuvent avoir un nombre différent de points d'angles, ceux-ci peuvent également être adaptés à des points de la diagonale, comme illustré sur la Figure 1.3 (c). Notez que la définition de δ implique que la distance d'appariement de deux points situés sur la diagonale est égale à zéro ($\delta(\Delta, \Delta) = 0$).

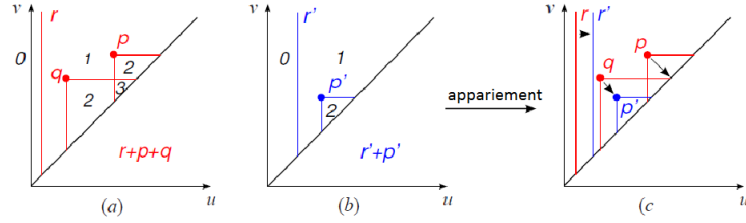


Figure 1.3 – (a) Fonction de taille correspondant à la série formelle $r + a + b + c$. (b) Fonction de taille correspondant à la série formelle $r' + a' + c$. (c) La distance d'appariement entre les deux fonctions de taille [22].

Exemple 1.13. (Pour plus d'exemples voir [3]) Dans cet exemple, nous calculons la distance d'appariement entre les deux fonctions de taille illustrées dans la Figure 1.4. La distance d'appariement entre $\ell_1 = r + a + b$ et $\ell_2 = r' + a'$ est donnée par

$$d(\ell_1, \ell_2) = \max \{ \delta(r, r'), \delta(b, a'), \delta(a, \Delta) \} = 4.$$

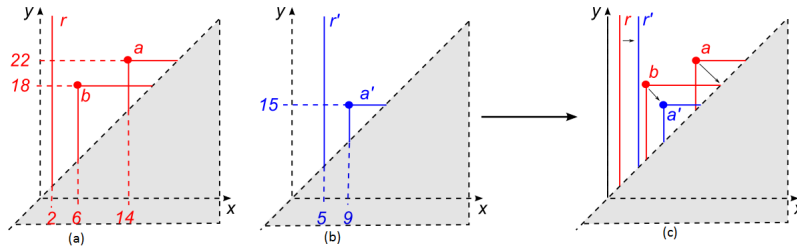


Figure 1.4 – (a) Fonction de taille correspondant à la série formelle $r + a + b$. (b) Fonction de taille correspondant à la série formelle $r' + a'$. (c) La distance d'appariement entre les deux fonctions de taille [21].

Finalement, la distance d'appariement entre les fonctions de taille définies pour des filtrations de sous-niveau, présente des propriétés de stabilité intéressantes.

Théorème 1.14. (Stabilité) [19] Soient $\varphi, \phi : X \rightarrow \mathbb{R}$ deux fonctions de mesure continues.

La propriété suivante est vraie :

$$d(\ell_{\langle X, \varphi \rangle}, \ell_{\langle X, \phi \rangle}) \leq \|\varphi(P) - \phi(P)\|_{\infty}.$$

Théorème 1.15. [19] *Si (X, ϕ) et (X, φ) sont deux fonctions de taille et $\max_{P \in X} |\phi(P) - \varphi(P)| \leq \eta$, Alors $d_{\text{match}}(\ell_{\langle X, \varphi \rangle}, \ell_{\langle X, \phi \rangle}) \leq \eta$.*

CHAPITRE 2

Homologie persistante

Dans ce chapitre, nous nous concentrons sur la définition et le calcul de l'homologie persistante [6]. Pour trouver l'homologie persistante d'un espace, l'espace doit d'abord être représenté comme un complexe simplicial. Nous commençons alors notre étude par un rappel sur les complexes simpliciaux [4, 5]. Nous passons ensuite au calcul de l'homologie sur des chaînes pour présenter la formulation algébrique de l'homologie persistante et pour ensuite la calculer. Enfin, nous montrons que l'homologie persistante d'un complexe filtré est une homologie standard d'un module gradué particulièrement sur un anneau polynomial.

On finit ce chapitre par une visualisation de l'homologie persistante avec des codes barres, et ensuite sur le plan avec la paire (naissance, mort).

2.1 Homologie simpliciale

2.1.1 Les complexes simpliciaux

Définition 2.1. (Simplexe) Soit v_0, \dots, v_n un ensemble géométriquement indépendant dans \mathbb{R}^d . On définit σ , le n -simplexe engendré par v_0, \dots, v_n , comme étant l'ensemble de tout les points x de \mathbb{R}^d tels que

$$x = \sum_{i=0}^n t_i v_i, \text{ avec } \sum_{i=0}^n t_i = 1$$

où les t_i sont les coordonnées barycentriques positives ou nulles de x relatives à v_0, \dots, v_n . Les points v_0, \dots, v_n sont appelés les sommets de σ .

La dimension de σ est n . Nous montrons des simplexes de faible dimension avec leurs noms dans la Figure 2.1.

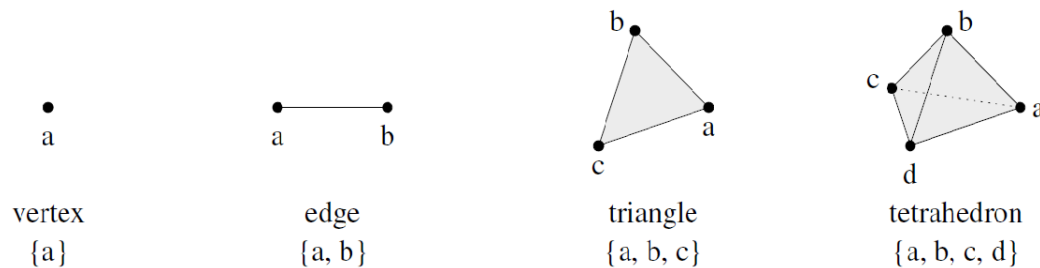


Figure 2.1 – n -simplexes, pour chaque $0 \leq n \leq 3$ [6].

Définition 2.2. (Face) Nous définissons une k -face de σ comme un k -simplexe ($k \leq n$) τ engendré par $k + 1$ points de v_0, \dots, v_n . L'ensemble vide est le simplexe engendré par un ensemble de sommets vide et est une face de tout simplexe.

Si $k < n$, nous appelons la face une face propre de σ .

Définition 2.3. (Complexe simplicial) Un complexe simplicial K dans \mathbb{R}^d est une collection de simplexes dans \mathbb{R}^d telle que :

1. toute face d'un simplexe de K est dans K ;
2. l'intersection de deux simplexes non disjoints doit exactement être une de leurs faces communes.

Nous noterons par \mathcal{K} le support de K , c'est-à-dire $\mathcal{K} = |K| = \cup K$ (le sous-ensemble de \mathbb{R}^d obtenu par union de tous les simplexes de K).

Définition 2.4. Si L est une sous-collection d'un complexe simplicial K qui contient toutes les faces de ses éléments, alors L est un complexe simplicial. On l'appelle un sous-complexe de K . Un sous-complexe de K est la collection de tous les simplexes de K de dimension d'au plus p . Il est appelé p -squelette de K et est noté K^p . Les points de la collection K^0 sont appelés les sommets de K .

Définition 2.5. (triangulation) Une triangulation d'un espace topologique X est un couple (K, f) où K est un complexe simplicial sur $|K|$ homéomorphe à X , et un homéomorphisme $f : |K| \rightarrow X$.

Exemple 2.6. Le bord d'un 3-simplexe (tétraèdre) est homéomorphe à une sphère et est une triangulation de la sphère, comme le montre la Figure 2.2.

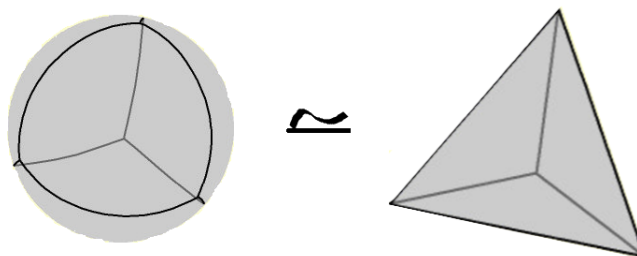


Figure 2.2 – La surface d'un tétraèdre est une triangulation d'une sphère.

Définition 2.7. Un simplexe orienté est un simplexe σ muni d'une orientation de σ . Définissons que deux ordres dans ses ensembles de sommets sont équivalents si ils diffèrent

l'un de l'autre par une permutation paire. Si la dimension de σ est strictement supérieure à 0, les ordres des sommets de σ ont deux classes d'équivalence. Chacune d'elles est appelée une orientation de σ . Si σ est un 0-simplexe, alors toutes les orientations sont équivalentes et on a donc une seule orientation de σ .

Exemple 2.8. *Considérons le complexe simplicial*

$$\Delta[a, b, c] = \{[a], [b], [c], [a, b], [a, c], [b, c], [a, b, c]\}.$$

Les simplexes de dimension 0 : $[a], [b], [c]$ n'ont qu'une seule orientation puisqu'ils n'ont qu'un seul ordre. Chaque simplexe de dimension 1 a deux orientations qui correspondent aux deux ordres : par exemple (a, b) et (b, a) définissent deux orientations différentes du simplexe $[a, b]$. Le simplexe $[a, b, c]$ a six numérotations différentes qui donnent les deux orientations suivantes : $[a, b, c] = [b, c, a] = [c, a, b]$ et $[b, a, c] = [a, c, b] = [c, b, a]$. Les orientations peuvent être représentées graphiquement à l'aide des flèches, comme le montre la Figure 2.3.

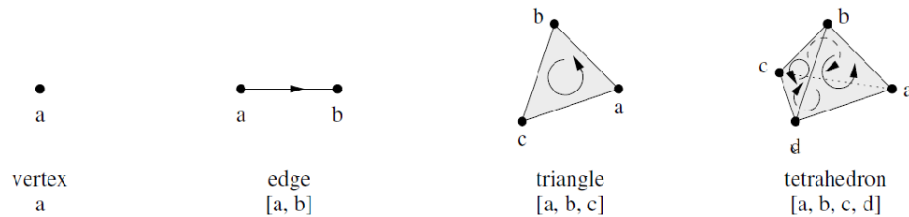


Figure 2.3 – n -simplexes, $0 \leq n \leq 3$. L'orientation sur le tétraèdre est représentée sur ses faces [6].

2.1.2 Groupes d'homologie

Définition 2.9. (*Chaîne élémentaire*) Soit K un complexe simplicial. Une chaîne élémentaire sur K est une fonction c de l'ensemble des simplexes orientés de K dans \mathbb{Z}

telle que :

1. $c(\sigma) = -c(\sigma')$ si σ et σ' sont des orientations opposées du même simplexe.
2. $c(\sigma) = 0$ pour tout autre simplexe

Définition 2.10. (Groupe de chaînes) Le groupe de q -chaînes de complexe simplicial K est $\langle C_q(K), + \rangle$, le groupe abélien libre engendré par les chaînes élémentaires des q -simplexes de K .

Définition 2.11. (Opérateur de bord) On appelle opérateur de bord le morphisme linéaire $\partial_q : C_q(K) \rightarrow C_{q-1}(K)$ défini par :

$$\partial_q(\sigma) = \sum_{i=0}^q (-1)^i [v_0, \dots, \hat{v}_i, \dots, v_q]$$

où \hat{v}_i indique que l'on prend tous les sommets de σ excepté le $i^{\text{ème}}$. L'homomorphisme ∂_q est appelé l'opérateur de bord d'ordre q .

Le bord d'un 0-simplexe est 0. On appellera cycle une chaîne dont le bord est nul.

Exemple 2.12. Prenons le bord des simplexes orientés de la Figure 2.3 :

$$\partial_1[a, b] = b - a.$$

$$\partial_2[a, b, c] = [b, c] - [a, c] + [a, b] = [b, c] + [c, a] + [a, b].$$

$$\partial_3[a, b, c, d] = [b, c, d] - [a, c, d] + [a, b, d] - [a, b, c].$$

Notez que l'opérateur de bord oriente les faces d'un simplexe orienté. Dans le cas du triangle, cette orientation correspond à la marche autour du triangle sur les bords, selon l'orientation du triangle.

Si nous prenons la limite de la limite du triangle, nous obtenons :

$$\partial_1 \partial_2 [a, b, c] = [c] - [b] - [c] + [a] + [b] - [a] = 0$$

Ceci est intuitivement correct : le bord d'un triangle est un cycle, et un cycle n'a pas de frontière. En fait, les applications bords vérifient la propriété suivante.

Propriété 2.13.

$$\forall q \in \mathbb{N}, \partial_q \circ \partial_{q+1} = 0$$

Définition 2.14. (cycle, bord) Le noyau de $\partial_q : C_q(K) \rightarrow C_{q-1}(K)$ est appelé le groupe des q -cycles. On note $\ker \partial_q = Z_q(K)$.

L'image de $\partial_{q+1} : C_{q+1}(K) \rightarrow C_q(K)$ est appelé le groupe des q -bords. On note $\text{Im } \partial_{q+1} = B_q(K)$ (voir Figure 2.4 [5]).

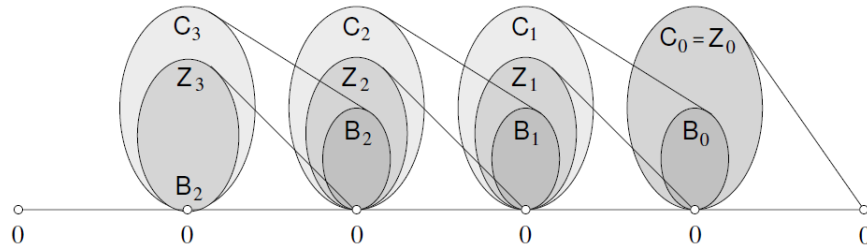


Figure 2.4 – n -simplexes, $0 \leq n \leq 3$. Un complexe de chaîne pour un complexe de dimension trois.

D'après la propriété précédente, on a $B_q(K) \subseteq Z_q(K)$. On dit qu'un complexe est exact si on a égalité entre l'image et le noyau, c-à-d si on a : $B_q(K) = Z_q(K)$.

Définition 2.15. (Groupe d'homologie) On appelle q^{ime} groupe d'homologie le groupe quotient :

$$H_q(K) = Z_q(K)/B_q(K).$$

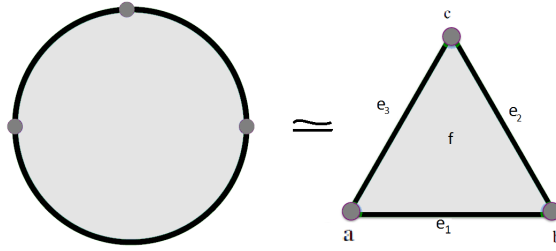


Figure 2.5 – Triangulation d'un disque.

Exemple 2.16. (*Calcul de l'homologie d'un disque*)

Soit l'ensemble simplicial K_1 , le complexe simpliciale qui représente la triangulation du disque (voir Figure 2.5) :

$$K_1 = \{a, b, c, e_1 = a, b, e_2 = b, c, e_3 = c, a, f = a, b, c\}.$$

On a donc le complexe de chaîne suivant :

$$0 \longrightarrow C_2(K_1) \xrightarrow{\partial_2} C_1(K_1) \xrightarrow{\partial_1} C_0(K_1) \xrightarrow{\partial_0} 0,$$

$$\begin{aligned} H_0(K_1) &= \frac{Z_0}{B_0} = \frac{\text{noyau de } \partial_0}{\text{image de } \partial_1} \\ &= \frac{\langle a, b, c \rangle}{\langle a + b, b + c, c + a \rangle}. \end{aligned}$$

Donc $H_0(K_1) = \langle [a] \rangle \simeq \mathbb{Z}$ où $[a] = \{a, b, c\}$,

$$\begin{aligned} H_1(K_1) &= \frac{Z_1}{B_1} = \frac{\text{noyau de } \partial_1}{\text{image de } \partial_2} \\ &= \frac{\langle e_1 + e_2 + e_3 \rangle}{\langle e_1 + e_2 + e_3 \rangle}. \end{aligned}$$

Donc $H_1(K_1) = 0$,

Enfin, $H_1(K_1) = \frac{Z_2}{B_2} = 0$.

Définition 2.17. (Nombre de Betti) On appelle q^{ime} nombre de Betti et on note β_q le rang du groupe $H_q(K)$. Le rang du groupe $H_0(K)$ est égal au nombre de composantes connexes de K (c'est-à-dire le nombre de trous d'ordre 0 dans K , de courbes fermées si $q = 1$, de cavités si $q = 2$, etc.).

Exemple 2.18. (Nombre de Betti d'un tore)

Les nombres de Betti d'un tore T sont $\beta_0(T) = 1$ (il s'agit d'une seule composante connexe) $\beta_1(T) = 2$ (il s'agit de deux trous circulaires, les deux cercles générateurs) et $\beta_2(T) = 1$.

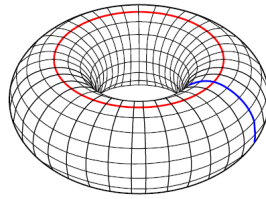


Figure 2.6 – Un tore, avec ses deux cycles d'ordre 1 indiqués en ligne verticale et horizontale.

2.2 La persistance

Dans cette section, on veut introduire le concept de l'homologie persistance [7]. Cette dernière décrit la durée de vie des propriétés topologiques le long d'une famille d'espaces imbriqués indexée par la variable de temps.

2.2.1 Filtration

Définition 2.19. (Sous-complexe) Un sous-complexe d'un complexe simplicial K est un complexe simplicial $L \subseteq K$.

Définition 2.20. (Filtration) Une filtration d'un complexe simplicial K est une séquence imbriquée de sous-complexes,

$$\emptyset = K^0 \subseteq K^1 \subseteq K^2 \subseteq \dots \subseteq K^n = K.$$

On appelle un complexe K avec une filtration un complexe filtré [13].

On note que le complexe $K^{i+1} = K^i \cup \delta^i$, où δ^i est un ensemble de simplexes. Un exemple d'une filtration peut être vu dans la Figure 2.7 :

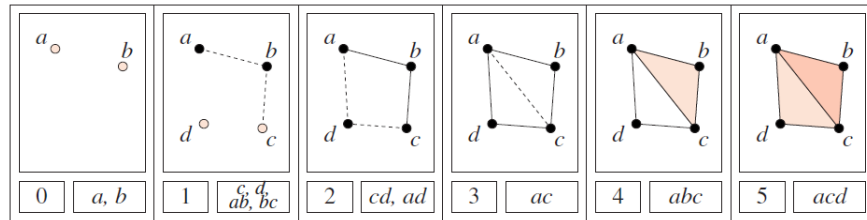


Figure 2.7 – Filtration d'un complexe simplicial [6].

Naturellement, cette filtration donne lieu à une séquence d'homomorphismes pour n'importe quelle dimension de p induite par des inclusions

$$0 \longrightarrow H_q(K^1) \longrightarrow \dots \longrightarrow H_q(K^n) = H_q(K).$$

2.2.2 Méthode de filtration

En pratique, l'espace topologique considéré sera un nuage de point provenant par exemple d'une collecte de données discrètes. Plusieurs méthodes ont été conçues pour construire un complexe et sa filtration à partir d'un échantillon de points (complexes de Čech, complexe de Delaunay, α -complexe,...). [9]

Définition 2.21. (Complexes de Čech) Étant donné un ensemble de points S dans un espace Euclidien \mathbb{E}^n , un nombre $\epsilon > 0$ et $B_\epsilon(s_i) = \{s \in E^n \mid d(s - s_i) \leq \epsilon\}$, le complexe

de Čech C_ϵ est le complexe simplicial abstrait dont les k -simplexes sont déterminés comme suit : pour chaque ensemble S de points, former un $\frac{\epsilon}{2}$ -boule autour de chaque point de S , et inclure S comme un simplexe (de dimension $|S|$) s'il y a un point commun contenu dans toutes les boules dans S .

$$\check{C}ech_\epsilon(S) = \left\{ \{B_\epsilon(s_0), B_\epsilon(s_1), \dots, B_\epsilon(s_k)\} \mid \bigcap_{i=0}^k B_\epsilon(s_i) \neq \emptyset, s_i \in S \right\}.$$

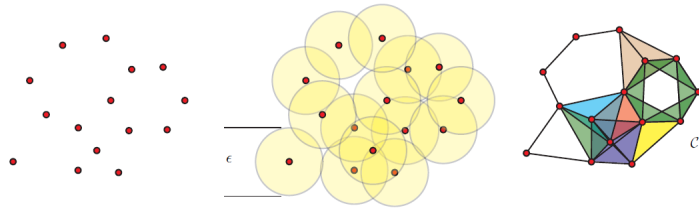


Figure 2.8 – Exemple d'un nuage de points (à gauche) et d'un choix correspondant de boules ($\epsilon/2$). Pour obtenir le complexe de Čech, nous ajoutons un k -simplexe chaque fois que nous voyons un sous-ensemble de k points avec intersection commune. [9]

Cette structure satisfait évidemment la définition d'un complexe simplicial : tout sous-ensemble $S' \subset$ d'un simplexe S sera également un simplexe.

Alpha shape

La forme alpha [11] associée à un ensemble de points est une généralisation du concept d'enveloppe convexe, c'est-à-dire que chaque enveloppe convexe est une forme de alpha shape, mais pas toute forme alpha shape est une enveloppe convexe.

Définition 2.22. (Région de Voronoï) Soit S un ensemble fini de k points de E muni de la distance euclidienne d ; les éléments de S sont appelés germes.

On appelle région de Voronoï ou cellule de Voronoï associée à un élément s de S , l'ensemble des points qui sont plus proches de s que de tout autre point de S :

$$Vor_s(p) = \{x \in E \mid d(x, s) \leq d(x, s') \forall s \neq s', s' \in S\}.$$

On intersecte les boules avec les régions de Voronoï et on trouve l'alpha complexe

$$\text{Alpha}_\epsilon(s) = B_\epsilon(s) \cap Vor(s).$$

On définit alors l'alpha complexe comme :

$$\text{Alpha}_\epsilon(S) = \left\{ \{\text{Alpha}_\epsilon(s_i)\}_{i=0,1,\dots,k} \mid \bigcap_{i=0}^k \text{Alpha}_\epsilon(s_i) \neq \emptyset, s_i \in S \right\}.$$

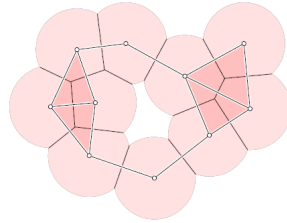


Figure 2.9 – L'union des disques est décomposée en régions convexes par les cellules de Voronoï. Le complexe alpha correspondant est superposé. [7]

Le complexe de Rips

Soient N un nuage de p points de \mathbb{R}^n et d une métrique sur \mathbb{R}^n . Un k -simplexe $\sigma = [a_0, a_1, \dots, a_k]$ appartient au $\text{Rips}(N, R)$ si, et seulement si, on a pour un réel $R > 0$ fixé :

$$\forall 0 \leq i, j \leq k, d(a_i, a_j) \leq R.$$

Pour tout $R' \geq R$, nous avons $\text{Rips}(N, R) \subseteq \text{Rips}(N, R')$. Si en plus on définit $\text{Rips}(N, 0) = \emptyset$, l'ensemble

$$A = \{\text{Rips}(N, R), 0 \leq R \leq R_{max}\}$$

définit une filtration paramétrée par un réel, où $R_{max} \in \mathbb{R}_+$ telle que $\text{Rips}(N, R_{max})$ est le $p - 1$ -simplexe formé par les p points de N .

Un n -simplexe appartient au complexe si, et seulement si, toutes ses faces appartiennent également au complexe. Ceci signifie qu'on peut ne calculer que le 1-squelette du complexe, et qu'on peut ensuite déterminer les simplexes de dimensions supérieures uniquement grâce aux simplexes de dimensions 0 et 1. Les complexes vérifiant cette propriété sont appelés *flag complexes*.

Le Witness Complexe

Cette méthode consiste à considérer un sous-ensemble de points du nuage, qui seront les seuls points à pouvoir entrer dans la filtration [10]. Les autres points nous donneront un critère pour déterminer si une famille de points décrit un simplexe ou non.

Soient N un nuage de p points de \mathbb{R}^n , L un sous-ensemble arbitraire de q points de N et d une métrique sur \mathbb{R}^n . Nous allons associer à N un ensemble de complexes simpliciaux de la manière suivante. Soient $R > 0$ un réel et $v \leq q$ un entier naturel. Chaque sous ensemble $\{a_0, \dots, a_k\} \subset L$ décrit un k -simplexe si, et seulement si, il existe un point $w \in N$ tel que

$$\max \{d(a_0, w), \dots, d(a_k, w)\} \leq R + \delta_v$$

où δ_v est la v^{ieme} plus petite distance entre w et un point de L . Le cas échéant, ce point est appelé témoin du simplexe $[a_0, \dots, a_k]$.

On note l'ensemble de ces simplexes par $W(N, R, v)$ et nous avons $R' \geq R$ implique que

$$W(N, R, v) \subseteq W(N, R', v).$$

Le Lazy Witness Complexe

Le Lazy Witness Complexe est une simple adaptation du Witness Complex afin de vérifier la propriété des flag complexes. Un couple a_1, a_2 de points décrit un segment du complexe si, et seulement si,

$$\max \{d(a_1, w), d(a_2, w)\} \leq R + \delta_v.$$

Ensuite, toute famille de points de cardinalité supérieure à 2 décrit un simplexe si, et seulement si, toutes les faces appartiennent au complexe simplicial.

2.2.3 Formulation algébrique d'homologie persistante

Définition 2.23. (*Groupe d'homologie persistante [6]*) Nous définissons respectivement Z_q^ℓ, B_q^ℓ comme le q^{ime} groupe de cycle et le q^{ime} groupe de bord, du ℓ^{ime} complexe K^ℓ dans la filtration. Le q^{ime} groupe d'homologie persistant de K^ℓ est :

$$H_q^{\ell,p} = Z_q^\ell / (B_q^{\ell+p} \cap Z_q^\ell). \quad (2.1)$$

Cela est bien défini parce que $B_q^{\ell+p} \cap Z_q^\ell$ est l'intersection de deux sous-groupes de $C_q^{\ell+p}$ et est donc lui-même un groupe.

Définition 2.24. (*Invariants de rang [6]*) Le q^{ime} nombre de Betti persistant $\beta_q^{(\ell,p)}$ de K^ℓ est le rang de $H_q^{\ell,p}$.

$$\beta_q^{(\ell,p)} = (\text{rang } H_q^{\ell,p}) = \text{rang } Z_q^\ell - \text{rang } (B_q^{\ell+p} \cap Z_q^\ell).$$

Intuitivement, pour un complexe K de \mathbb{R}^3 , un 0-cycle représente un ensemble de composantes connexes de K . Il en résulte que β_0 est le nombre de composantes connexe de K . Un 1-cycle représente une collection de courbes fermées non contractibles dans K ,

ou bien un ensemble de tunnels formés par K . On peut écrire chaque tunnel en tant que somme des tunnels dans une base, et β_1 est la taille de la base. Un 2-cycle représente une collection de surfaces indépendantes. Enfin, il n'y a pas de 3-cycles parce que K est un complexe dans \mathbb{R}^3 .

2.2.4 Module persistant

Dans cette section, nous prenons une vue différente de l'homologie persistante afin de comprendre sa structure [8]. Nous commençons cette section en combinant l'homologie de tous les complexes de la filtration en une seule structure algébrique. Nous établissons alors une correspondance qui révèle une simple description sur les corps. Plus significativement, nous illustrons que l'homologie persistante d'un complexe filtré est simplement l'homologie standard d'un module gradué particulier sur un anneau polynomial. Enfin, nous illustrons la relation entre nos structures et l'équation de persistance (équation (2.1)). Nous terminons cette section par un algorithme pour calculer l'homologie persistante.

Nous commençons d'abord par présenter des prérequis de l'algèbre en examinant des structures algébriques de plus en plus riches que nous rencontrerons dans notre étude de l'homologie.

Anneaux, corps, anneaux intègres et anneaux principaux

Définition 2.25. (*Anneau unitaire*) Un anneau $(A, +, \cdot)$ est un groupe abélien $(A, +)$ dont l'élément neutre sera noté 0_A , muni d'une opération interne notée « \cdot » vérifiant

1) *Associativité* : $(a \cdot b) \cdot c = a \cdot (b \cdot c)$

2) *Distributivité à gauche et à droite* : $a \cdot (b + c) = a \cdot b + a \cdot c$, $(b + c) \cdot a = b \cdot a + c \cdot a$

Dans la suite on supposera que les anneaux sont commutatifs et unitaires, c'est à dire :

$$3) a \cdot b = b \cdot a$$

4) Il existe un élément unité 1 dans A tel que $1 \cdot a = a \cdot 1 = a$ pour tout a .

Définition 2.26. (corps) Un corps K est un anneau commutatif unitaire tel que, pour chaque $a \in K$, il existe un élément a^{-1} tel que $aa^{-1} = a^{-1}a = 1$.

Définition 2.27. (Anneau intègre) Un anneau commutatif unitaire $(A, +, \cdot)$ est dit intègre s'il est

- différent de l'anneau nul, c'est-à-dire s'il possède au moins deux éléments, et
- sans diviseur de zéro, c'est à dire :

$$\forall (a, b) \in A^2, a \cdot b = 0 \Rightarrow (a = 0 \text{ ou } b = 0).$$

Définition 2.28. (Éléments unités et élément irréductible dans un anneau)

Un élément a d'un anneau commutatif est appelé un inversible (ou une unité) lorsqu'il possède un symétrique pour la multiplication, c'est-à-dire quand il existe un élément de b de l'anneau tel que $ab = 1$. Un élément inversible n'est jamais un diviseur de zéro. Un élément p de A est dit irréductible s'il n'est ni nul, ni inversible, ni produit de deux éléments non inversibles.

Théorème 2.29. Tout corps est un anneau intègre. Tout anneau intègre fini est un corps commutatif.

Exemple 2.30. \mathbb{Z} , \mathbb{Q} , \mathbb{R} , \mathbb{C} sont tous des anneaux sous les opérations d'addition et de multiplication. \mathbb{Z} n'est pas un corps, parce qu'il n'a pas d'inverses multiplicatifs pour ses éléments, mais \mathbb{Z} est un anneau intègre, \mathbb{Q} et \mathbb{R} sont des corps, et donc des anneaux intègres. \mathbb{Z}_p est un anneau intègre si p est premier. Comme \mathbb{Z}_p est fini, le théorème

précédent implique que \mathbb{Z}_p est également un corps. Si p n'est pas un premier, \mathbb{Z}_p n'est pas un anneau intègre, car il possède des éléments non nuls qui divisent zéro. Par exemple, $2 \cdot_6 3 = 0$ dans \mathbb{Z}_6 .

Définition 2.31. (Anneaux polynomiaux) Soit A un anneau commutatif unitaire. On construit l'anneau de polynôme $A[t]$ à indéterminée t comme étant l'ensemble

$$A[t] := \left\{ \sum_{i=0}^n a_i t^i, n_i \in \mathbb{N}, a_i \in A \right\},$$

Définition 2.32. (Anneau principal) Un idéal J est un sous-groupe additif de A stable pour la multiplication par n'importe quel élément a de A . Ainsi si j est élément de J , aj l'est aussi, ou encore aJ est inclus dans J .

Un idéal J de l'anneau A est dit principal s'il est composé des multiples d'un élément donné de l'anneau, autrement dit s'il existe un élément a de A tel que J est égal à aA .

Un anneau intègre D est un anneau principal si chaque idéal dans D est un idéal principal.

Modules et graduation

Définition 2.33. (Module) Un A -module (à gauche) M est un groupe abélien $(M, +)$ muni d'une loi externe :

$$\begin{aligned} A \times M &\rightarrow M \\ (a, m) &\mapsto a \cdot m \end{aligned}$$

vérifiant les trois propriétés suivantes :

1. la bi-additivité, $a \cdot (m_1 + m_2) = am_1 + am_2$, $(a + b) \cdot m_1 = a \cdot m_1 + b \cdot m_1$,
2. l'associativité, $a \cdot (b \cdot m_1) = (ab) \cdot m_1$.
3. l'action triviale de l'unité, $1 \cdot m = m$

où $(a, b) \in A^2$ et $(m_1, m_2) \in M^2$.

On prendra bien soin de ne pas confondre les lois d'addition dans M et dans A , et de même pour la loi de multiplication dans A avec la loi externe de M .

Définition 2.34. (anneau gradué) Un anneau gradué et un anneau $\langle A, +, \otimes \rangle$ muni d'une décomposition de la somme directe des groupes abéliens $A \cong \bigoplus_i A_i$, $i \in \mathbb{Z}$, de sorte que la multiplication est définie par des appariements bilinéaires $A_n \otimes A_m \longrightarrow A_{n+m}$. Les éléments de A_i sont dits homogènes de degré i , $\deg e = i$, $\forall e \in A_i$.

Définition 2.35. (Module gradué) Un module gradué sur un anneau gradué A est un module muni d'une décomposition de la somme directe $M \cong \bigoplus_i M_i$, $i \in \mathbb{Z}$, de sorte que la multiplication est définie par des appariements bilinéaires $A_n \otimes M_m \longrightarrow M_{n+m}$.

Définition 2.36. (graduation non-négative) Un anneau (module) gradué est un à graduation non-négative si $A_i = 0$ ($M_i = 0$, respectivement) pour chaque $i < 0$.

Exemple 2.37. Soit $A[t]$ un anneau polynomial en l'indéterminé t . On a que $A[t]$ est à graduation non-négative avec $(t^n) = t^n \cdot A[t]$, $n \geq 0$. C'est la graduation standard pour $A[t]$.

Théorème de structure

Théorème 2.38. (Théorème de structure) Tout module de type fini sur un anneau principal A est isomorphe à une somme directe de modules de la forme

$$A^r \oplus \left(\bigoplus_{i=1}^m A/d_i A \right),$$

pour $d_i \in A$, $r \in \mathbb{Z}$ et une unique séquence d'éléments non nuls d_1, \dots, d_m de A , telle que d_{i+1} divise d_i .

Le théorème décompose les structures en parties libres (gauche) et torsion (droite).

Module persistant

Définition 2.39. (Module persistant) Un module de persistant \mathbf{M} est une famille de A -modules M^i , ainsi que des homomorphismes $\varphi^i : M^i \longrightarrow M^{i+1}$.

Définition 2.40. (type fini) Un module persistant $\{M^i, \varphi^i\}$ est de type fini si chaque module de la famille est un M -module fini et si les applications φ^i sont des isomorphismes pour $i \geq m$ pour un entier m .

Supposons que nous avons un module persistant $M = \{M^i, \varphi^i\}_{i \geq 0}$ sur un anneau A . Nous munissons maintenant $A[t]$ de la graduation standard et définissons un module gradué sur $A[t]$ par

$$\alpha(M) = \bigoplus_{i=0}^{\infty} M^i,$$

où la structure de A -module est simplement la somme des structures sur les composants individuels, et où l'action de t est donnée par

$$t \cdot (m^0, m^1, m^2, \dots) = (0, \varphi^0(m^0), \varphi^1(m^1), \varphi^2(m^2), \dots).$$

Ainsi, t déplace simplement les éléments du module dans la graduation. Cette correspondance α est la clé pour comprendre la structure de la persistante.

D'autre part, la correspondance nous donne une décomposition simple lorsque l'anneau est un corps K . Ici, l'anneau gradué étant un anneau principal, nous utilisons le théorème de structure pour $A[t]$ -modules de type fini où $A[t]$ est en un anneau principal. On a la décomposition suivante du module persistant :

$$\left(\bigoplus_{i=1}^n \Sigma^{\alpha_i} K[t] \right) \oplus \left(\bigoplus_{j=1}^m \Sigma^{\gamma_j} K[t]/(t^{n_j}) \right).$$

La multiplication par t correspond à l'avancement d'un pas dans le module de persistance. Intuitivement, les parties libres du côté droit correspondent aux générateurs d'homologie qui apparaissent au niveau de filtration α_i et ne disparaissent jamais, tandis que la deuxième partie correspond à ceux qui apparaissent au niveau de filtration γ_j et durent pour n_j filtration (ou de manière équivalente, disparaissent au niveau de filtration $n_j + \gamma_j$).

Exemple 2.41. *L'exemple suivant calcule l'homologie d'ordre 0 de chaîne complexe illustré dans la Figure 2.7*

$$\oplus K_0^i = \langle a, b \rangle \oplus \langle c, d, ta, tb \rangle \oplus \langle tc, td, t^2a, t^2b \rangle \oplus \dots$$

$$a, b : \mathbb{Z}_2[t] = \{n_0 + n_1t + n_2t^2 + \dots : n_i \in \mathbb{Z}_2\} = \{n_0, n_1, n_2, \dots : n_i \in \mathbb{Z}_2\},$$

$$d, c : \Sigma^1 \mathbb{Z}_2[t] = \{n_0t + n_1t^2 + \dots : n_i \in \mathbb{Z}_2\} = \{0, n_0, n_1, \dots : n_i \in \mathbb{Z}_2\}.$$

Maintenant on présente a, b, c, d d'après leur temps de filtration pour calculer l'homologie d'ordre 0.

$$a : \mathbb{Z}_2[t] = \{n_0, n_1, n_2, \dots : n_i \in \mathbb{Z}_2\}.$$

$$b : (\mathbb{Z}_2[t])/t = \{n_0 : n_0 \in \mathbb{Z}_2\}.$$

$$c : \Sigma^1 \mathbb{Z}_2[t]/1 = \emptyset.$$

$$d : \Sigma^1 \mathbb{Z}_2[t]/t = \{n_0 : n_0 \in \mathbb{Z}_2\} = \{(0, n_0, 0, 0, 0, \dots)\}.$$

Enfin on trouve l'homologie d'ordre 0 de notre filtration

$$H_0(K) = \mathbb{Z}_2[t] \oplus (\mathbb{Z}_2[t])/t \oplus \Sigma^1 \mathbb{Z}_2[t]/t.$$

Maintenant, nous nous intéressons à la discussion sur les k -intervalles. Nous souhaitons paramétrer les classes d'isomorphisme de $K[t]$ -modules par des P -intervalles.

Définition 2.42. (P -intervalle) *Un P -intervalle est une paire ordonnée (i, j) avec $0 \leq i < j \in \mathbb{Z}^\infty = \mathbb{Z} \cup \{+\infty\}$.*

Nous associons un $M[t]$ -module gradué à un ensemble S de P -intervalles via une bijection Q . Pour le P -intervalle (i, j) nous définissons

$$Q(i, j) = \Sigma^i M[t]/(t^{j-i}),$$

et pour (i, ∞)

$$Q(i, \infty) = \Sigma^i M[t].$$

Pour l'ensemble de P -intervalles $S = \{(i_1, j_1), (i_2, j_2), \dots, (i_n, j_n)\}$, nous définissons

$$Q(S) = \bigoplus_{l=1}^n Q(i_l, j_l).$$

La correspondance ci-dessus peut maintenant être reformulée comme suit.

Théorème 2.43. *La correspondance $S \rightarrow Q(S)$ définit une bijection entre les ensembles finis de P -intervalles et les modules gradués de type fini sur l'anneau gradué $A[t]$. Par conséquent, les classes d'isomorphisme des modules de persistance de type fini sur M sont en correspondance bijective avec les ensembles finis de P -intervalles.*

Maintenant, en utilisant ce théorème on peut prouver que B est égal au nombre de K -triangles contenant le point (i, p) dans le plan de persistance d'indice.

Lemme 2.44. *Soit \mathcal{T} l'ensemble de triangles définis par P -intervalles pour le k -dimensionnel module persistant. Le rang $\beta_k^{l,p}$ de $H_k^{l,p}$ est le nombre de triangles dans \mathcal{T} contenant le point (l, p) .*

Cela montre aussi que, pour le calcul de l'homologie persistante sur un corps, il est suffisant de trouver l'ensemble correspondant de P -intervalles (voir Figure 2.10).

Calcul de l'homologie persistance avec les matrices

Dans cette section nous calculons l'homologie persistant en utilisant les matrices où les coefficients sont calculés en \mathbb{Z} modulo 2. Nous suivrons l'exemple de filtration de complexe simplicial K (Figure 2.7) pour calculer l'homologie persistance en dimension 0.

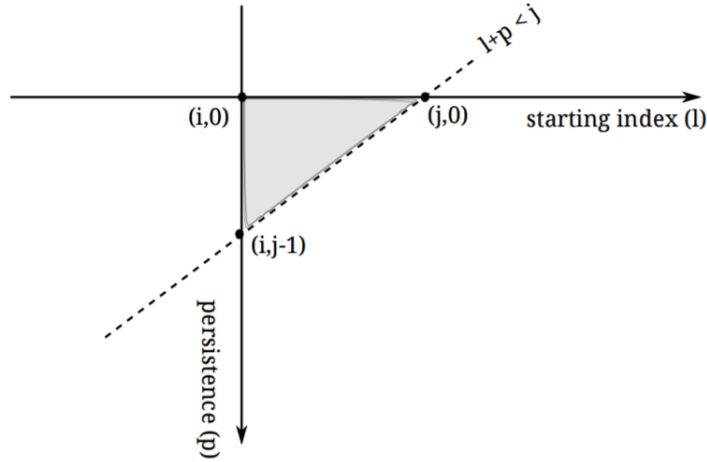


Figure 2.10 – Les inégalités $p \geq 0$, $l \geq i$ et $l + p < j$ définissent une région triangulaire dans le plan d’index et de persistance . Cette région définit quand le cycle est un élément de base pour l’espace vectoriel d’homologie [6].

Tout d’abord, nous représentons l’opérateur de frontière $\partial_k : C_k \rightarrow C_{k-1}$ relatif aux bases standard des groupes de chaînes en tant que matrice M_k qui s’appelle la représentation matricielle standard de ∂_k . Elle a m_k colonnes et m_{k-1} lignes (le nombre de k et $k - 1$ -simplexes, respectivement). Le noyau de la matrice M_k correspond à Z_k et son image à B_{k-1} .

La représentation matricielle standard de ∂_1 de notre exemple est

$$M_1 = \left(\begin{array}{c|ccccc} & ab & bc & cd & ad & ac \\ \hline d & 0 & 0 & 1 & 1 & 0 \\ c & 0 & 1 & 1 & 0 & 1 \\ b & 1 & 1 & 0 & 0 & 0 \\ a & 1 & 0 & 0 & 1 & 1 \end{array} \right)$$

Définition 2.45. Soit σ_i un k -simplexe.

- Le degré de σ_i ($\deg \sigma_i$) est le temps où le simplexe σ_i entre dans la filtration.
- Un polynôme est homogène si tous les termes sont le même degré total. Par exemple dans notre filtration de la Figure 2.7, $t^3a + t^2c + ac$ est en K^3

Nous utilisons $\{e_j\}$ et $\{\hat{e}_i\}$ pour représenter des bases homogènes pour (C_k) et (C_{k-1}) , respectivement. Relativement aux bases homogènes, toute représentation M_k de ∂_k a la propriété fondamentale suivante :

$$\deg \hat{e}_i + \deg M_k(i, j) = \deg e_j,$$

où $M_k(i, j)$ dénote l'élément à l'emplacement (i, j) . Nous obtenons

$$M_1 = \left(\begin{array}{c|ccccc} & ab & bc & cd & ad & ac \\ \hline d & 0 & 0 & t & t & 0 \\ c & 0 & 1 & t & 0 & t^2 \\ b & t & t & 0 & 0 & 0 \\ a & t & 0 & 0 & t^2 & t^3 \end{array} \right), \quad (2.2)$$

pour ∂_1 dans notre exemple. Dans cet exemple $M_1(4, 4) = t^2$. Comme $\deg ad - \deg a = 2$ nous voyons que la propriété précédente est vérifiée.

Évidemment, la base standard pour les groupes de chaînes est homogène. Il faut représenter $\partial_k : C_k \rightarrow C_{k-1}$ par rapport à la base standard pour C_k et à une base homogène pour Z_{k-1} . Nous réduisons ensuite la matrice et lisons la description de H_k . Nous calculons ces représentations de manière inductive par rapport à la dimension. Le cas de base est trivial. Comme $\partial_0 \equiv 0$, $Z_0 = C_0$ et la base standard peut être utilisée pour représenter ∂_1 . Maintenant, supposons que nous avons une représentation matricielle M_k de ∂_k par rapport à la base standard $\{e_j\}$ pour C_k et une base homogène $\{\hat{e}_i\}$ pour Z_{k-1} . Pour la récurrence, nous devons calculer une base homogène pour Z_k et représenter ∂_{k+1} par rapport à C_{k+1} et à la base calculée. Nous commençons par trier la base \hat{e}_i en ordre inverse, comme déjà fait dans la matrice (2.2). Nous transformons ensuite M_k en une matrice échelonnée selon les colonnes \widetilde{M}_k sous forme d'escalier descendant présenté dans la Figure (2.11).

À partir de l'algèbre linéaire, on sait que $M_k = \text{rang} B_{k-1}$ est le nombre de pivots de la forme échelonnée. Les éléments de base correspondant aux colonnes non nulles forment

$$\left[\begin{array}{cccccc} \boxed{*} & 0 & & & & 0 \\ & \boxed{*} & 0 & \dots & & \\ * & * & 0 & & & \vdots \\ & * & \boxed{*} & 0 & \dots & \\ * & & * & 0 & \dots & 0 \end{array} \right]$$

Figure 2.11 – La forme d’une matrice échelonnée selon les colonnes. Une * indique une valeur non nulle et les pivots sont encadrés.

la base souhaitée pour Z_k . Dans notre exemple, nous avons :

$$M_1 = \left(\begin{array}{c|cccccc} & ab & bc & cd & z_1 & z_2 \\ \hline d & t & 0 & 0 & 0 & 0 \\ c & t & 1 & 0 & 0 & 0 \\ b & 0 & t & t & 0 & 0 \\ a & 0 & 0 & t & 0 & 0 \end{array} \right),$$

où $z_1 = ad + cd + t \cdot bc + t \cdot ab$ et $z_2 = ac + t^2 \cdot bc + t^2 \cdot ab$ forme une base homogène Z_1 .

$$B_0 = \text{Im } \partial_1 = \langle tc + td, tb + c, ta + tb \rangle.$$

$$Z_1 = \ker \partial_1 = \langle ad + cd + t \cdot bc + t \cdot ab, ac + t^2 \cdot bc + t^2 \cdot ab \rangle.$$

Nous calculons l’homologie persistante d’ordre 0 de notre exemple :

$$H_0 = \frac{Z_0}{B_0} = \frac{\langle a, b, c, d \rangle}{\langle tc + td, tb + c, ta + tb \rangle}.$$

$$H_0^{i,p} = Z_0^i / (B_0^{i+p} \cap Z_0^i).$$

$$H_0^{0,0} = Z_2^2.$$

$$\text{Pour } p > 0 : H_0^{0,p} = Z_2.$$

Algorithme de calcul de la persistance

Dans cette section, nous expliquons un algorithme pour calculer l’homologie persistante d’un complexe simplicial. L’algorithme total s’exécute dans un temps au plus $O(n^3)$

, où n est le nombre de simplexes qui apparaissent dans le complexe simplicial final. L'algorithme détermine les nombres de Betti et les paires (σ^i, σ^j) , c-à-d les simplexes qui apparaissent en i et disparaissent en j .

Soit K un complexe simplicial composé de n simplexes notés σ_i où $1 \leq i \leq n$ et $\mathcal{F} = \{K_i, i \in \mathbb{R}\}$ la filtration associée à ce complexe.

Soit D_σ la matrice de l'application de bord appliquée à l'ensemble des simplexes de K , telles que les colonnes de D seront les simplexes ordonnés par leur temps de naissance. Avec les coefficients dans \mathbb{Z}_2 , $D_{i,j} = 1$ si σ_i est une face de σ_j et $\dim(\sigma_i) = \dim(\sigma_j) - 1$ et $D_{i,j} = 0$ si non.

On définit l'entier $low(i)$ comme étant le numéro de la ligne où apparait le dernier 1 de la colonne i , ou 0 si la colonne i est nulle. Comme D_σ est une matrice triangulaire strictement supérieure, on a $low(i) < i, \forall 1 \leq i \leq n$. De plus, on dira qu'un simplexe est positif s'il appartient à un cycle, et négatif si non.

L'algorithme va créer des paires de simplexes, le simplexe qui crée une classe d'homologie avec le simplexe qui détruira cette même classe. Si un simplexe crée une classe qui ne disparaîtra pas, on sera dans le cas d'une classe essentielle. Chaque paire sera appelée paire de persistance.

Afin de mettre en évidence les simplexes positifs et les simplexes négatifs, pour déterminer les paires de persistance, nous allons effectuer des opérations sur les simplexes et alors des opérations sur les colonnes de la matrice D_σ . La matrice réduite de D_σ sera notée M dans l'algorithme suivant :

```

M ← Dσ
for j = 1 → m do
while il existe j0 < j tel que low(j0) = low(j) do
  Colonne(j) ← Colonne(j0) + Colonne(j) mod 2
end while
end for

```

Les paires de persistance seront alors les paires $(\sigma_{low(i)}, \sigma_i)$ avec i tel que $low(i) \neq 0$.

Si $low(i) = 0$, alors σ_i est un simplexe positif. En effet, à chaque étape de l'algorithme la colonne i représente le bord de $\sigma_i + \sum_{k < i} \epsilon_k \sigma_k$, où $\epsilon_k \in 0, 1$. On a donc $\partial(\sigma_i + \sum_{k < i} \epsilon_k \sigma_k) = 0$, c'est-à-dire que σ_i fait partie d'un cycle. C'est alors un simplexe positif.

Si $low(i) = j \neq 0$, alors σ_j est un simplexe positif. En effet,

$$\partial\partial(\sigma_i + \sum_{k < i} \epsilon_k \sigma_k) = \partial(\sigma_j + \sum_{p < j} \epsilon_p \sigma_p) = 0,$$

c'est-à-dire que σ_j fait partie d'un cycle. C'est alors un simplexe positif. Comme σ_j est le dernier simplexe de la chaîne $\sigma_j + \sum_{p < j} \epsilon_p \sigma_p$ à être arrivé dans la filtration, c'est aussi lui qui crée ce cycle. Alors σ_i tue la classe d'homologie créée par σ_j . En effet, avant que σ_i n'apparaisse dans la filtration, σ_j n'est associé à aucun autre simplexe, σ_j est apparié, et donc σ_i tue la classe d'homologie qui avait été créée par σ_j .

2.3 Visualisation de l'homologie persistante

Puisque l'homologie persistante représente un invariant algébrique qui détecte la naissance et la mort de chaque caractéristique topologique à mesure que le complexe évolue dans le temps, il est avantageux de coder l'homologie persistante sous la forme d'une version paramétrée du rang du groupe d'homologie (son nombre de Betti).

Visualisation avec code barres : Un code barres est une représentation graphique de l'homologie persistante comme des intervalles dont les extrémités sont des complexes de filtration représentant essentiellement le complexe simplicial filtré aux moments où l'ajout (ou l'élimination) des simplexes a lieu. Durant son existence temporelle, chaque attribut topologique joue un rôle dans la formation d'un certain nombre de Betti et notre intérêt réside dans les propriétés à longue durée de vie (propriétés de la persistance). Les intervalles de paramètres représentent des durées de vie de divers stades de filtration et

peuvent être représentés sur l'axe des indices . Un code barres pour la filtration d'un complexe simplicial est présenté à la Figure 2.12.

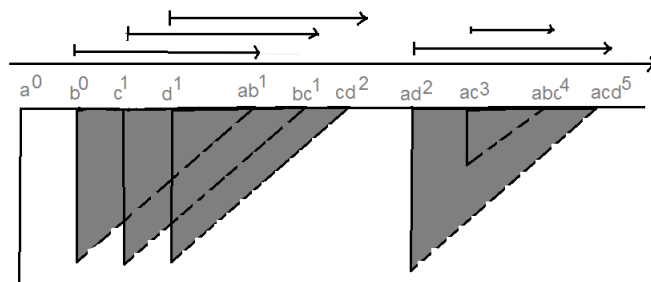


Figure 2.12 – Visualisation par code barres et triangles du complexe simplicial de la Figure 2.7. Les triangles de s et de stw sont non bornés et non dessinés. Les triangles clairs et sombres représentent respectivement les 0-cycles et 1-cycles.

Visualisation avec des triangles : Dans la visualisation avec des code barres, nous représentons chaque paire de simplexes (σ^i, σ^j) sur l'axe des indices par un intervalle à moitié ouvert $[i, j)$ que nous appelons un q -intervalle. Nous étendons cette visualisation à deux dimensions engendrées par les axes des indices et de persistance. Le q -intervalle de simplexes (σ^i, σ^j) se prolonge en un q -triangle engendré par $(i, 0)$, $(j, 0)$, $(i, j - i)$ dans le plan de persistance et des indices. Le q -triangle est fermé le long de ses bords verticaux et horizontaux et s'ouvre le long de la diagonale reliant $(j, 0)$ à $(i, j - i)$, comme le montre la Figure 2.13. Il représente le q -cycle créé par σ^i , qui est détruit par σ^j progressivement à mesure que nous augmentons p .

Toutefois, ces deux visualisations (code barres et triangles) ne peuvent pas être utiles quand le nuage des points contient de centaine de points. Dans ce cas, la visualisation avec un diagramme de persistance qui contient les paires de points (naissance et mort) peut être plus utile (voir Figure 2.14).

L'inconvénient de ce dernier digramme est qu'il y a des parties où on n'a aucune informa-

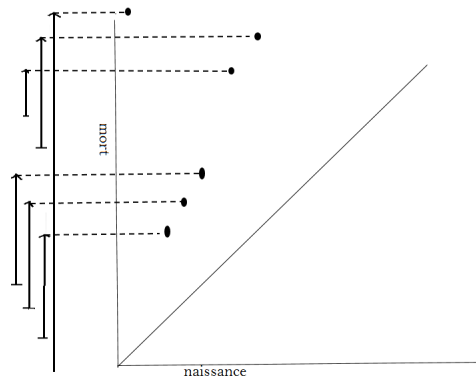


Figure 2.13 – Visualisation par digramme de persistance qui présente les points de pair (naissance, mort).

tion c'est alors inutile de les représenter dans ce diagramme. P. Pranav présente dans [14] un meilleur diagramme où il ne présente que la partie qui contient le maximum d'information, autrement dit la partie qui contient les points persistants qui naissent avant u et meurent après v . [14]

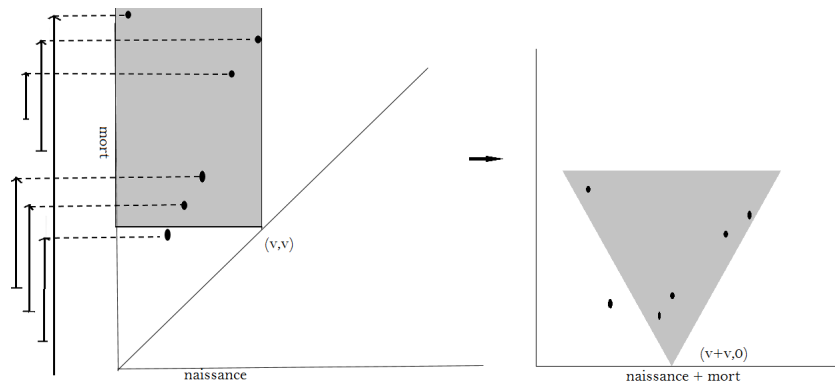


Figure 2.14 – Visualisation par diagramme de Pranav. La région ombrée transforme en panneau à une région en forme de V déterminée par $(v + v, 0)$. Les bras du V ont une pente -1 et 1 respectivement.

CHAPITRE 3

L'homologie persistante multidimensionnelle et la stabilité

La fonction de taille multidimensionnelle est une extension de la fonction unidimensionnelle [19, 20]. Sa filtration de sous-niveau pour une fonction de mesure sur X prend ses valeurs dans \mathbb{R}^k . Celle-ci est choisie par l'utilisateur en fonction de ses besoins. L'idée d'utiliser des fonctions de mesure k -dimensionnelles provient du fait que la forme d'un objet peut être mieux caractérisée par un ensemble de fonctions de mesure, chacune étudiant les caractéristiques spécifiques de la forme étudiée.

Pour illustration, nous prenons l'exemple ci-dessous qui montre deux différents espaces. Si nous utilisons la fonction de mesure unidimensionnelle, nous ne sommes pas en mesure de montrer les différences entre ces deux espaces à l'inverse d'une fonction de mesure bidimensionnelle (voir Figure 3.1).

L'inconvénient de cette filtration multidimensionnelle est l'absence de descripteur pour les nombres de Betti persistants $\beta_{\vec{\varphi}}$ aussi simple que celui des diagrammes de persistance. Il est cependant possible de réduire le calcul de l'homologie persistante multidimensionnelle

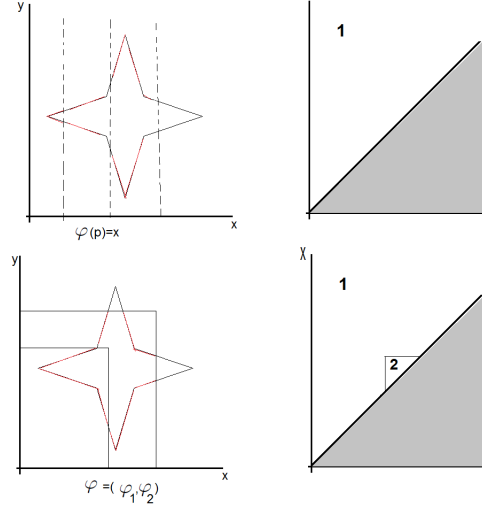


Figure 3.1 – Un espace topologique étudié par une fonction de mesure unidimensionnelle et bidimensionnelle. L’espace topologique filtré par fonction de mesure bidimensionnelle montre plus de détails dans son diagramme de persistance.

au calcul de l’homologie persistante ordinaire par le théorème de réduction en utilisant la méthode de foliation.

Une conséquence importante de cela est la stabilité de ce nouveau descripteur (et donc de la fonction de taille k -dimensionnelle correspondante) par rapport aux perturbations des fonctions de mesure.

Définition 3.1. Soit la paire $(X, \vec{\varphi} = (\varphi_1, \dots, \varphi_k))$, où $\vec{\varphi} : X \rightarrow \mathbf{R}^k$ est une fonction continue. On appelle la fonction $\vec{\varphi}$ fonction de mesure (ou de filtration) k -dimensionnelle. Les relations suivantes \preceq et \prec sont définies dans \mathbf{R}^k : pour $\vec{x} = (x_1, \dots, x_k)$ et $\vec{y} = (y_1, \dots, y_k)$, $\vec{x} \preceq \vec{y}$ ($\vec{x} \prec \vec{y}$) si et seulement si $x_i \leq y_i$ (resp. $x_i < y_i$) pour $i = 1, \dots, k$. De plus, \mathbf{R}^k est muni de la norme infinie usuelle : $\|(x_1, x_2, \dots, x_k)\|_\infty = \max_{1 \leq i \leq k} |x_i|$.

Nous utiliserons les notations suivantes : $\Delta^+ = \{(\vec{x}, \vec{y}) \in \mathbf{R}^k \times \mathbf{R}^k : \vec{x} \prec \vec{y}\}$. Pour chaque k -uplet $\vec{x} = (x_1, \dots, x_k) \in \mathbf{R}^k$, soit $X\langle \vec{\varphi} \preceq \vec{x} \rangle$ l’ensemble $\{P \in X : \varphi_i(P) \leq x_i, i = 1, \dots, k\}$.

Définition 3.2. *L'inclusion $j^{(\vec{x}, \vec{y})} : X\langle \vec{\varphi} \preceq \vec{x} \rangle \hookrightarrow X\langle \vec{\varphi} \preceq \vec{y} \rangle$ pour $\vec{x} \preceq \vec{y}$, induit une application linéaire $H_q(j^{(\vec{x}, \vec{y})})$ entre les homologies $H_q(X\langle \vec{\varphi} \preceq \vec{x} \rangle)$ et $H_q(X\langle \vec{\varphi} \preceq \vec{y} \rangle)$. L'image de $H_q(j^{(\vec{x}, \vec{y})})$ est appelé le q^{ime} groupe d'homologie persistant multidimensionnel de $(X, \vec{\varphi})$ en (\vec{x}, \vec{y}) .*

Autrement dit, l'image de $H_q(j^{(\vec{x}, \vec{y})})$ correspond aux composantes topologiques de dimension q qui existent dans $X_{\vec{x}}$ et existent encore dans $X_{\vec{y}}$. Dans le langage de la persistance, nous dirons que ces composantes sont « nées » au temps \vec{x} ou avant, et « meurent » plus tard que \vec{y} .

Dans ce qui suit, nous travaillerons avec des coefficients dans un corps \mathbb{F} , de sorte que les groupes d'homologie sont des espaces vectoriels. Par conséquent, ils peuvent être complètement décrits par leur dimension, conduisant à la définition suivante.

Définition 3.3. [22] *La fonction $\beta_{\vec{\varphi}} : \Delta^+ \rightarrow \mathbb{N}$ définie par*

$$\beta_{\vec{\varphi}}(\vec{x}, \vec{y}) = \dim \text{Im } H_q(j^{(\vec{x}, \vec{y})}),$$

sera appelée le nombre de Betti persistant (d'un ordre quelconque) pour la fonction $\vec{\varphi}$ (on l'appelle aussi fonction de taille).

3.1 Réduction du cas multidimensionnel au cas unidimensionnel

Dans cette section, nous montrerons qu'il existe une famille paramétrée de demi-plans dans $\mathbb{R}^k \times \mathbb{R}^k$ telle que la restriction des nombres de Betti persistants multidimensionnels pour la fonction de mesure à chacun de ces plans peut être considérée comme des nombres de Betti persistants unidimensionnels pour la fonction à deux variables scalaires.

Définition 3.4. Pour chaque vecteur unitaire $\vec{l} = (l_1, \dots, l_k)$ de \mathbb{R}^k tel que $l_i > 0$ pour $i = 1, \dots, k$, et $\sum_{i=1}^k l_i^2 = 1$ et pour chaque vecteur $\vec{b} = (b_1, \dots, b_k)$ de \mathbb{R}^k tel que $\sum_{i=1}^k b_i = 0$, nous dirons que la paire (\vec{l}, \vec{b}) est admissible. Nous noterons l'ensemble de toutes les paires admissibles dans $\mathbb{R}^k \times \mathbb{R}^k$ par Adm_k . Compte tenu d'une paire admissible (\vec{l}, \vec{b}) , nous définissons le demi-plan $\pi_{(\vec{l}, \vec{b})}$ de $\mathbb{R}^k \times \mathbb{R}^k$ par les équations paramétriques suivantes :

$$\begin{cases} \vec{x} = s \vec{l} + \vec{b} \\ \vec{y} = t \vec{l} + \vec{b} \end{cases}$$

pour $s, t \in \mathbb{R}$, avec $s < t$.

Chaque pair admissible (\vec{l}, \vec{b}) correspond à une droite de \mathbb{R}^k , pour chaque $(\vec{x}, \vec{y}) \in \Delta^+$ il existe une et une seule paire admissible (\vec{l}, \vec{b}) telle que $(\vec{x}, \vec{y}) \in \pi_{(\vec{l}, \vec{b})}$. On observera que \vec{l} et \vec{b} ne dépendent que de (\vec{x}, \vec{y}) .

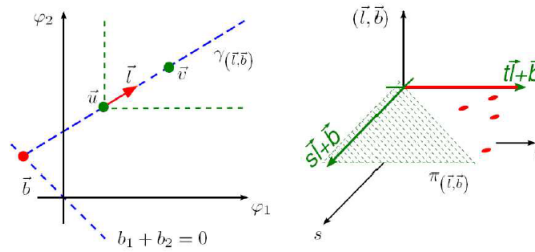


Figure 3.2 – Réduction des nombres de Betti persistants bidimensionnels pour une fonction $\vec{\phi}$ au cas unidimensionnel [22].

Supposons (\vec{l}, \vec{b}) est une paire admissible. On a :

$$\begin{aligned} \vec{\phi}(P) \leq \vec{x} &\Rightarrow \varphi_i(P) \leq x_i = s l_i + b_i \\ &\Rightarrow \frac{\varphi_i(P) - b_i}{l_i} \leq s. \end{aligned}$$

On a aussi

$$\begin{aligned}\vec{\varphi}(P) \leq \vec{y} &\Rightarrow \varphi_i(P) \leq y_i = t l_i + b_i \\ &\Rightarrow \frac{\varphi_i(P) - b_i}{l_i} \leq t,\end{aligned}$$

ce qui nous donne :
$$\begin{cases} \frac{\varphi_i(P) - b_i}{l_i} \leq s \\ \frac{\varphi_i(P) - b_i}{l_i} \leq t. \end{cases}$$

Par conséquent, on pose

$$F_{(\vec{l}, \vec{b})}^{\vec{\varphi}} = \max \left\{ \frac{\varphi_i(P) - b_i}{l_i} \right\}.$$

D'après les arguments précédents on obtient le théorème suivant.

Théorème 3.5. *Soit (\vec{l}, \vec{b}) une paire admissible, et $F_{(\vec{l}, \vec{b})}^{\vec{\varphi}}$ définie comme suit :*

$$F_{(\vec{l}, \vec{b})}^{\vec{\varphi}}(P) = \max_{i=1, \dots, k} \left\{ \frac{\varphi(P) - b_i}{l_i} \right\}.$$

Pour chaque $(\vec{x}, \vec{y}) = (s \vec{l} + \vec{b}, t \vec{l} + \vec{b}) \in \pi_{(\vec{l}, \vec{b})}$ l'égalité suivante est vérifiée :

$$\beta_{(X, \vec{\varphi})}(\vec{x}, \vec{y}) = \beta_{(X, F_{(\vec{l}, \vec{b})}^{\vec{\varphi}})}(s, t).$$

3.2 Distance d'appariement des invariants de rang multidimensionnels

Dans la suite, on note $d_{match}(\beta_{(X, F_{(\vec{l}, \vec{b})}^{\vec{\varphi}})}, \beta_{(Y, F_{(\vec{l}, \vec{b})}^{\vec{\psi}})})$ la distance d'appariement entre les nombres de Betti persistants unidimensionnels $\beta_{(X, F_{(\vec{l}, \vec{b})}^{\vec{\varphi}})}$ et $\beta_{(Y, F_{(\vec{l}, \vec{b})}^{\vec{\psi}})}$.

Corollaire 3.6. *On considère deux fonctions de nombres de Betti persistants $(X, \vec{\varphi}), (Y, \vec{\psi})$.*

Alors, $\beta_{(X, \vec{\varphi})} \equiv \beta_{(Y, \vec{\psi})}$ si et seulement si

$$d_{match}(\beta_{(X, F_{(\vec{l}, \vec{b})}^{\vec{\varphi}})}, \beta_{(Y, F_{(\vec{l}, \vec{b})}^{\vec{\psi}})}) = 0,$$

pour chaque paire admissible (\vec{l}, \vec{b}) .

Le résultat suivant prouve la stabilité de d_{match} par rapport à la fonction de mesure choisie, c'est-à-dire que des changements suffisamment petits en $\vec{\varphi}$ par rapport à la norme du maximum induit de petits changements de $\beta_{(X, F_{(\vec{l}, \vec{b})}^{\vec{\varphi}})}$ par rapport à la distance d'appariement.

Proposition 3.7. *Si $(X, \vec{\varphi}), (X, \vec{\chi})$ sont deux paires de taille*

$\max_{P \in X} \|\vec{\varphi}(P) - \vec{\chi}(P)\|_\infty \leq \epsilon$, alors pour chaque paire admissible (\vec{l}, \vec{b}) , on a que

$$d_{match}(\beta_{(X, F_{(\vec{l}, \vec{b})}^{\vec{\varphi}})}, \beta_{(Y, F_{(\vec{l}, \vec{b})}^{\vec{\chi}})}) \leq \frac{\epsilon}{\min_{i=1, \dots, k} l_i}$$

Preuve. On pose $\eta = \max_{P \in X} |F_{(\vec{l}, \vec{b})}^{\vec{\varphi}}(P) - F_{(\vec{l}, \vec{b})}^{\vec{\chi}}(P)|$ du Théorème 1.15, on obtient

$$d_{match}(\beta_{(X, F_{(\vec{l}, \vec{b})}^{\vec{\varphi}})}, \beta_{(Y, F_{(\vec{l}, \vec{b})}^{\vec{\chi}})}) \leq \max_{P \in X} |F_{(\vec{l}, \vec{b})}^{\vec{\varphi}}(P) - F_{(\vec{l}, \vec{b})}^{\vec{\chi}}(P)|.$$

Fixons maintenant $P \in X$. Alors, on dénote par \hat{i} l'indice pour lequel $\max_i \frac{\phi(P) - b_i}{l_i}$ est atteint, par la définition de $F_{(\vec{l}, \vec{b})}^{\vec{\varphi}}$ et $F_{(\vec{l}, \vec{b})}^{\vec{\chi}}$, nous avons

$$\begin{aligned} F_{(\vec{l}, \vec{b})}^{\vec{\varphi}} - F_{(\vec{l}, \vec{b})}^{\vec{\chi}} &= \max_i \frac{\varphi(P) - b_i}{l_i} - \max_i \frac{\chi_i(P) - b_i}{l_i} \\ &= \frac{\varphi_{\hat{i}}(P) - b_{\hat{i}}}{l_{\hat{i}}} - \max_i \frac{\chi_i(P) - b_i}{l_i} \\ &\leq \frac{\varphi_{\hat{i}}(P) - b_{\hat{i}}}{l_{\hat{i}}} - \frac{\chi_{\hat{i}}(P) - b_{\hat{i}}}{l_{\hat{i}}} \\ &= \frac{\varphi_{\hat{i}}(P) - \chi_{\hat{i}}(P)}{l_{\hat{i}}} \leq \frac{\|\vec{\varphi}(P) - \vec{\chi}(P)\|_\infty}{\min_{i=1, \dots, k} l_i}. \end{aligned}$$

Comme $\max_{P \in X} \|\vec{\varphi}(P) - \vec{\chi}(P)\|_\infty \leq \epsilon$, on trouve alors

$$\begin{aligned} \max_{P \in X} |F_{(\vec{l}, \vec{b})}^{\vec{\varphi}}(P) - F_{(\vec{l}, \vec{b})}^{\vec{\chi}}(P)| &\leq \max_{P \in X} \frac{\|\vec{\varphi}(P) - \vec{\chi}(P)\|_\infty}{\min_{i=1, \dots, k} l_i} \\ &\leq \frac{\epsilon}{\min_{i=1, \dots, k} l_i}. \end{aligned}$$

□

De façon analogue, on peut prouver que d_{match} est stable par rapport au choix des demi-plans dans la foliation. En effet, la proposition suivante indique que des changements suffisamment petits dans (\vec{l}, \vec{b}) par rapport à la norme maximum induisent de petits changements de $\beta_{(X, F_{(\vec{l}, \vec{b})}^{\vec{\varphi}})}$ par rapport à la distance d'appariement.

Proposition 3.8. *Si $(X, \vec{\varphi})$ est une paire de taille, $(\vec{l}, \vec{b}) \in \text{Adm}_k$ et ϵ est un nombre réel tel que $0 \leq \epsilon < \min_{i=1, \dots, k} l_i$, alors pour chaque paire admissible (\vec{l}', \vec{b}') avec $\|(\vec{l}, \vec{b}) - (\vec{l}', \vec{b}')\|_\infty \leq \epsilon$, on trouve que*

$$d_{match}(\beta_{(X, F_{(\vec{l}, \vec{b})}^{\vec{\varphi}})}, \beta_{(Y, F_{(\vec{l}', \vec{b}')}^{\vec{\psi}})}) \leq \epsilon \cdot \frac{\max_{P \in X} \|\vec{\varphi}(P)\|_\infty + \|\vec{l}'\|_\infty + \|\vec{b}'\|_\infty}{\min_{i=1, \dots, k} l_i (l_i - \epsilon)}.$$

Preuve. On pose $\eta = \max_{P \in X} |F_{(\vec{l}, \vec{b})}^{\vec{\varphi}}(P) - F_{(\vec{l}', \vec{b}')}^{\vec{\psi}}(P)|$. Du théorème 1.15, on obtient

$$d_{match}(\beta_{(X, F_{(\vec{l}, \vec{b})}^{\vec{\varphi}})}, \beta_{(Y, F_{(\vec{l}', \vec{b}')}^{\vec{\psi}})}) \leq \max_{P \in X} |F_{(\vec{l}, \vec{b})}^{\vec{\varphi}}(P) - F_{(\vec{l}', \vec{b}')}^{\vec{\psi}}(P)|.$$

Fixons maintenant $P \in X$. Soit \hat{i} l'indice pour lequel $\max_i \frac{\varphi(P) - b_i}{l_i}$ est atteint. D'après

la définition de $F_{(\vec{l}, \vec{b})}^{\vec{\varphi}}$ et $F_{(\vec{l}', \vec{b}')}^{\vec{\varphi}}$, on a

$$\begin{aligned}
F_{(\vec{l}, \vec{b})}^{\vec{\varphi}}(P) - F_{(\vec{l}', \vec{b}')}^{\vec{\varphi}}(P) &= \max_i \frac{\varphi_i(P) - b_i}{l_i} - \max_i \frac{\varphi_i(P) - b'_i}{l'_i} \\
&= \frac{\varphi_i(P) - b_i}{l_i} - \max_i \frac{\varphi_i(P) - b'_i}{l'_i} \\
&\leq \frac{\varphi_i(P) - b_i}{l_i} - \frac{\varphi_i(P) - b'_i}{l'_i} \\
&= \frac{(l'_i - l_i)\varphi_i(P) - l'_i b_i + l_i b'_i}{l_i l'_i} \\
&= \frac{(l'_i - l_i)\varphi_i(P) + l_i(b'_i - b_i) + b_i(l_i - l'_i)}{l_i l'_i} \\
&\leq \frac{|l'_i - l_i|\varphi_i(P) + l_i|b'_i - b_i| + |b_i||l_i - l'_i|}{l_i l'_i} \\
&\leq \frac{\epsilon(\|\vec{\varphi}(P)\|_\infty + \|\vec{l}\|_\infty + \|\vec{b}\|_\infty)}{l_i(l_i - \epsilon)} \\
&\leq \frac{\epsilon(\|\vec{\varphi}(P)\|_\infty + \|\vec{l}\|_\infty + \|\vec{b}\|_\infty)}{\min_{i=1, \dots, k} l_i(l_i - \epsilon)}.
\end{aligned}$$

On trouve alors

$$\max_{P \in X} |F_{(\vec{l}, \vec{b})}^{\vec{\varphi}}(P) - F_{(\vec{l}', \vec{b}')}^{\vec{\varphi}}(P)| \leq \epsilon \cdot \frac{\max_{P \in X} \|\vec{\varphi}(P)\|_\infty + \|\vec{l}\|_\infty + \|\vec{b}\|_\infty}{\min_{i=1, \dots, k} l_i(l_i - \epsilon)}.$$

□

Dans le prochain théorème, on prouve que la borne inférieure de $d((X, \vec{\varphi}), (Y, \vec{\psi}))$ existe et est fournie par les restrictions des fonctions multidimensionnelles aux demi-plans de la foliation.

Théorème 3.9. *Soit $(X, \vec{\varphi})$ et $(Y, \vec{\psi})$ deux paires de taille, avec X, Y homéomorphes. On pose $d((X, \vec{\varphi}), (Y, \vec{\psi})) = \inf_f \max_{P \in X} \|\vec{\varphi}(P) - \vec{\psi}(f(P))\|_\infty$, où f varie entre tous les homéomorphismes entre X et Y . On a alors*

$$\sup_{(\vec{l}, \vec{b}) \in \text{Adm}_k} \left\{ \min_{i=1, \dots, k} l_i \cdot d_{\text{match}}(\beta_{(X, F_{(\vec{l}, \vec{b})}^{\vec{\varphi}})}, \beta_{(Y, F_{(\vec{l}, \vec{b})}^{\vec{\psi}})}) \right\} \leq d((X, \vec{\varphi}), (Y, \vec{\psi})).$$

Définition 3.10. Soit $(X, \vec{\varphi})$ et $(Y, \vec{\psi})$ deux paires de taille. Nous appellerons la distance d'appariement multidimensionnelle la distance étendue définie par :

$$D_{match}(\beta_{(X, \vec{\varphi})}, \beta_{(Y, \vec{\psi})}) = \sup_{(\vec{l}, \vec{b}) \in \text{Adm}_k} \left\{ \min_{i=1, \dots, k} l_i \cdot d_{match}(\beta_{(X, F_{(\vec{l}, \vec{b})} \vec{\varphi})}, \beta_{(Y, F_{(\vec{l}, \vec{b})} \vec{\psi})}) \right\}.$$

Dans la prochaine proposition, on voit que la borne inférieure fournie par $D_{match}(\beta_{(X, \vec{\varphi})}, \beta_{(Y, \vec{\psi})})$ est meilleure que celle obtenue à partir de la distance d'appariement unidimensionnelle $d_{match}(\beta_{(X, \varphi_j)}, \beta_{(Y, \psi_j)})$.

Proposition 3.11. Soit $(X, \vec{\varphi})$ et $(Y, \vec{\psi})$ deux paires de taille, avec $\vec{\varphi} = (\varphi_1, \dots, \varphi_k)$ et $\vec{\psi} = (\psi_1, \dots, \psi_k)$. Alors, pour $j = 1, \dots, k$, on a :

$$d_{match}(\beta_{(X, \varphi_j)}, \beta_{(Y, \psi_j)}) \leq D_{match}(\beta_{(X, \vec{\varphi})}, \beta_{(Y, \vec{\psi})}).$$

Preuve. Soit $\mu = \max_{P \in X} \|\vec{\varphi}(P)\|_\infty$ et $\nu = \max_{Q \in Y} \|\vec{\psi}(Q)\|_\infty$. Pour $j = 1, \dots, k$, on considère une paire admissible (\vec{l}^j, \vec{b}^j) , où $\vec{l}^j = (l_1^j, \dots, l_k^j)$ et $\vec{b}^j = (b_1^j, \dots, b_k^j)$ sont définis par les relations suivantes :

$$l_i^j = \frac{1}{\sqrt{k}}, \quad \text{pour } i = 1, \dots, k,$$

$$b_i^j = \begin{cases} -\frac{2(k-1)}{k} \cdot \max\{\mu, \nu\}, & \text{si } i = j; \\ \frac{2}{k} \cdot \max\{\mu, \nu\}, & \text{si } i \neq j. \end{cases}$$

D'après le théorème 1.20, pour chaque $(\vec{x}, \vec{y}) = (s\vec{l}^j + \vec{b}^j, t\vec{l}^j + \vec{b}^j) \in \pi_{\vec{l}^j, \vec{b}^j}$ il s'en suit que $\beta_{(X, \vec{\varphi})}(\vec{x}, \vec{y}) = \beta_{(X, F_{(\vec{l}^j, \vec{b}^j)} \vec{\varphi})}(s, t)$, $\beta_{(Y, \vec{\psi})}(\vec{x}, \vec{y}) = \beta_{(Y, F_{(\vec{l}^j, \vec{b}^j)} \vec{\psi})}(s, t)$ avec

$$F_{(\vec{l}^j, \vec{b}^j)}^{\vec{\varphi}}(P) = \max_{i=1, \dots, k} \left\{ \frac{\varphi_i(P) - b_i^j}{l_i^j} \right\} = \sqrt{k}(\varphi_j(P) - b_j^j)$$

$$F_{(\vec{l}^j, \vec{b}^j)}^{\vec{\psi}}(Q) = \max_{i=1, \dots, k} \left\{ \frac{\psi_i(Q) - b_i^j}{l_i^j} \right\} = \sqrt{k}(\psi_j(Q) - b_j^j),$$

pour chaque $P \in X$ et $Q \in Y$. Par la définition de d_{match} , on a

$$\begin{aligned} d_{match}(\beta_{(X, F_{(\vec{l}^j, \vec{b}^j)} \vec{\phi})}, \beta_{(Y, F_{(\vec{l}^j, \vec{b}^j)} \vec{\psi})}) &= d_{match}(\beta_{(X, \sqrt{k}(\varphi_j - b_j^j))}, \beta_{(Y, \sqrt{k}(\psi_j - b_j^j))}) \\ &= \sqrt{k} \cdot d_{match}(\beta_{(X, (\varphi_j - b_j^j))}, \beta_{(Y, (\psi_j - b_j^j))}), \end{aligned}$$

et donc

$$\begin{aligned} \min_{i=1, \dots, k} l_i^j \cdot d_{match}(\beta_{(X, F_{(\vec{l}^j, \vec{b}^j)} \vec{\phi})}, \beta_{(Y, F_{(\vec{l}^j, \vec{b}^j)} \vec{\psi})}) &= \frac{1}{\sqrt{k}} \cdot d_{match}(\beta_{(X, F_{(\vec{l}^j, \vec{b}^j)} \vec{\phi})}, \beta_{(Y, F_{(\vec{l}^j, \vec{b}^j)} \vec{\psi})}) \\ &= d_{match}(\beta_{(X, \varphi_j)}, \beta_{(Y, \psi_j)}). \end{aligned}$$

Par la définition de $D_{match}(\beta_{(X, \vec{\phi})}, \beta_{(Y, \vec{\psi})})$, on obtient le résultat. \square

Exemple 3.12. Soit l'ensemble $\Omega = [-1, 1] \times [-1, 1] \times [-1, 1]$ et la sphère S d'équation $x^2 + y^2 + z^2 = 1$. Soit aussi $\vec{\phi} = (\phi_1, \phi_2) : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ une fonction continue définie par $\vec{\phi}(x, y, z) = (|x|, |z|)$. On considère les paires de taille $(X, \vec{\phi})$ et $(Y, \vec{\psi})$ où $X = \partial\Omega$, $Y = S$, et $\vec{\phi}$ et $\vec{\psi}$ sont respectivement la restriction de $\vec{\phi}$ à X et Y .

Afin de comparer les fonctions de taille $\beta_{(X, \vec{\phi})}$ et $\beta_{(Y, \vec{\psi})}$ nous nous intéressons à étudier la

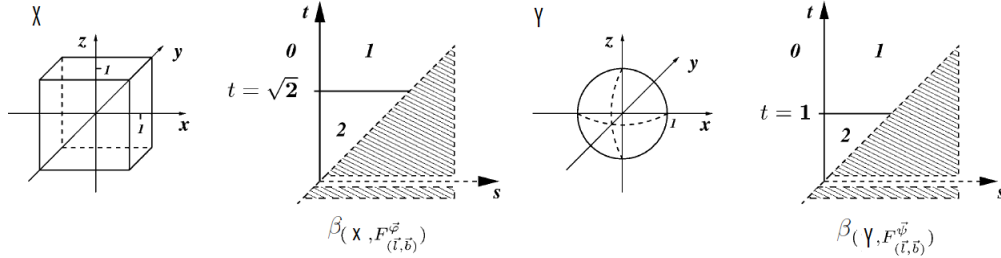


Figure 3.3 – L'espace topologique X et Y et $\beta_{(X, F_{(\vec{l}, \vec{b})} \vec{\phi})}, \beta_{(Y, F_{(\vec{l}, \vec{b})} \vec{\psi})}$ associé au demi-plan $\pi_{(\vec{l}, \vec{b})}$ [19].

foliation dans les demi-plans $\pi_{(\vec{l}, \vec{b})}$, où $\vec{l} = (\cos \theta, \sin \theta)$ avec $\theta \in (0, \frac{\pi}{2})$, et $\vec{b} = (a, -a)$ où $a \in \mathbb{R}$. Chaque demi-plan est représenté par :

$$\begin{cases} x_1 = s \cos \theta + a, \\ x_2 = s \sin \theta - a \\ y_1 = t \cos \theta + a \\ y_2 = t \sin \theta - a \end{cases}$$

avec $s, t \in \mathbb{R}$, $s < t$. Avec $\theta = \frac{\pi}{4}$ et $a = 0$, on a $\vec{l} = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$ et $\vec{b} = (0, 0)$. Avec ce choix, on obtient : $F_{(\vec{l}, \vec{b})}^{\vec{\phi}} = \sqrt{2} \max \{\varphi_1, \varphi_2\} = \sqrt{2} \max \{|x|, |z|\}$ et $F_{(\vec{l}, \vec{b})}^{\vec{\psi}} = \sqrt{2} \max \{\psi_1, \psi_2\} = \sqrt{2} \max \{|x|, |z|\}$. Donc, le Théorème 1.19 implique que, pour chaque $(x_1, x_2, y_1, y_2) \in \pi_{(\vec{l}, \vec{b})}$

$$\begin{aligned} \beta_{(X, \vec{\phi})}(x_1, x_2, y_1, y_2) &= \beta_{(X, \vec{\phi})}\left(\frac{s}{\sqrt{2}}, \frac{s}{\sqrt{2}}, \frac{t}{\sqrt{2}}, \frac{t}{\sqrt{2}}\right) = \beta_{(X, F_{(\vec{l}, \vec{b})}^{\vec{\phi}})}(s, t), \\ \beta_{(Y, \vec{\psi})}(x_1, x_2, y_1, y_2) &= \beta_{(Y, \vec{\psi})}\left(\frac{s}{\sqrt{2}}, \frac{s}{\sqrt{2}}, \frac{t}{\sqrt{2}}, \frac{t}{\sqrt{2}}\right) = \beta_{(Y, F_{(\vec{l}, \vec{b})}^{\vec{\psi}})}(s, t). \end{aligned}$$

La distance d'appariement montre le coût de déplacer le point de coordonnées $(0, \sqrt{2})$ au point de coordonnées $(0, 1)$.

$$\frac{\sqrt{2}}{2} d_{\text{match}}(\beta_{(X, F_{(\vec{l}, \vec{b})}^{\vec{\phi}})}, \beta_{(Y, F_{(\vec{l}, \vec{b})}^{\vec{\psi}})}) = \frac{\sqrt{2}}{2} (\sqrt{2} - 1) = 1 - \frac{\sqrt{2}}{2}.$$

Les fonctions de taille k-dimensionnelles, avec $\vec{\phi}, \vec{\psi}$, sont capables de discriminer le cube et la sphère, alors que les deux fonctions de taille unidimensionnelles, par rapport à $\vec{\phi}_1, \vec{\phi}_2$ et $\vec{\psi}_1, \vec{\psi}_2$, ne peuvent pas le faire. Le pouvoir discriminatoire supérieur des fonctions de taille k-dimensionnelles motive leur définition et leur utilisation.

3.3 Stabilité des nombres de Betti persistants par rapport aux perturbations du domaines

Dans cette section [23], on montre que si l'espace topologique est modifié en raison d'un bruit, les nombres de Betti persistants seront changés aussi.

On propose une approche générale du problème de la stabilité des groupes d'homologie persistants par rapport aux perturbations des ensembles bruités. L'idée principale est de réduire le problème de la stabilité d'espace topologique à celui de la stabilité des fonctions de mesure. Ceci est obtenu en remplaçant l'ensemble d'étude K qui nous intéresse avec une fonction appropriée f_K définie sur un ensemble fixe X contenant K , de sorte que la perturbation de l'ensemble K devient une perturbation de la fonction f_K . En conséquence, la fonction de mesure d'origine $\vec{\varphi}|_K : K \rightarrow \mathbb{R}^n$ est remplacée par une nouvelle fonction de mesure $\vec{\phi} : X \rightarrow \mathbb{R}^{n+1}$, $\vec{\phi} = (f_K, \vec{\varphi})$. Les nombres de Betti persistants de $(X, \vec{\phi})$ peuvent être comparés en utilisant la distance d'appariement multidimensionnelle. De cette façon, nous pouvons obtenir la robustesse des groupes d'homologie persistants sous les perturbations de l'ensemble d'étude .

En particulier, on utilise cette stratégie lorsque les ensembles sont comparés par la distance de Hausdorff δ_H . Dans ce cas, en prenant f_K égale à la fonction de distance de K , nous montrons que la distance d'appariement multidimensionnelle entre les nombres de Betti persistants associés à deux ensembles compacts K_1 et K_2 est toujours majorée par la distance de Hausdorff $\delta_H(K_1, K_2)$ entre K_1 et K_2 (Théorème 3.13). En même temps, nous montrons que, dans notre approche, l'information sur l'ensemble d'étude original K et sa fonction de mesure originale $\vec{\varphi}|_K$ est entièrement maintenue dans les groupes d'homologie persistants de $(X, \vec{\phi})$ (Théorème 3.14).

Théorème 3.13. *Soit K_1, K_2 des sous-ensembles fermés non vides d'un sous-espace triangulable X de \mathbb{R}^m . Soient $d_{K_1}, d_{K_2} : X \rightarrow \mathbb{R}$ leurs fonctions de distance respectives. De plus, soient $\vec{\varphi}_1, \vec{\varphi}_2 : X \rightarrow \mathbb{R}^n$ des fonctions continues de valeurs vectorielles. Alors, si on définit $\vec{\phi}_1, \vec{\phi}_2 : X \rightarrow \mathbb{R}^{n+1}$ par $\vec{\phi}_1 = (d_{K_1}, \vec{\varphi}_1)$ et $\vec{\phi}_2 = (d_{K_2}, \vec{\varphi}_2)$, on a :*

$$D_{match} \left(\beta_{\vec{\phi}_1}, \beta_{\vec{\phi}_2} \right) \leq \max \{ \delta_H(K_1, K_2), \|\vec{\varphi}_1 - \vec{\varphi}_2\|_\infty \},$$

où $\delta_H(K_1, K_2) = \|d_{K_1} - d_{K_2}\|_\infty$.

Jusqu'à maintenant, nous avons présenté que nous pouvons obtenir la stabilité recherchée par rapport à la perturbation de l'espace en considérant des fonctions de filtrage vectorielles appropriées. On pourrait penser que cela se fait au prix de l'oubli des informations sur le problème original. Au contraire, notre méthode permet de récupérer les nombres de Betti persistants de $(K, \varphi|_K)$ à partir des nombres de Betti persistants de $(X, \vec{\phi})$, avec $\vec{\phi} = (d_K, \overline{\phi})$. Ceci est indiqué dans le prochain résultat, montrant que, pour une valeur suffisamment petite de $\eta \in \mathbb{R}$ il existe une valeur suffisamment petite $\varepsilon \in \mathbb{R}$ avec $0 \leq \varepsilon \leq \eta$ telle que $\beta_{\vec{\phi}|_K}(\vec{x}, \vec{y}) = \beta_{\vec{\phi}}((\varepsilon, \vec{x}), (\eta, \vec{y}))$.

Théorème 3.14. *Soit K un sous-ensemble triangulable non vide d'un sous-espace X triangulable de \mathbb{R}^m . De plus, soit $\vec{\phi} : X \rightarrow \mathbb{R}^n$ une fonction continue. On définit $\vec{\phi} : X \rightarrow \mathbb{R}^{n+1}$, où $\vec{\phi} = (d_K, \overline{\phi})$ pour chaque $\vec{x}, \vec{y} \in \mathbb{R}^n$ où $\vec{x} \prec \vec{y}$. Il existe un nombre réel $\hat{\eta} > 0$ tel que, pour tout $\eta \in \mathbb{R}$ où $0 < \eta \leq \hat{\eta}$, il existe $\hat{\varepsilon} = \hat{\varepsilon}(\eta)$ un nombre réel, où $0 < \hat{\varepsilon} < \eta$, pour lequel*

$$\beta_{\vec{\phi}|_K}(\vec{x}, \vec{y}) = \beta_{\vec{\phi}}((\hat{\varepsilon}, \vec{x}), (\eta, \vec{y})),$$

pour tout $\varepsilon \in \mathbb{R}$ avec $0 \leq \varepsilon \leq \hat{\varepsilon}$. En particulier

$$\beta_{\vec{\phi}|_K}(\vec{x}, \vec{y}) = \lim_{\eta \rightarrow 0^+} \beta_{\vec{\phi}}((0, \vec{x}), (\eta, \vec{y})).$$

Exemple 3.15. *Nous travaillons avec l'image numérique binaire représentée dans la Figure 3.4 (à gauche), et la même image mais bruitée cette fois-ci représentée (à droite).*

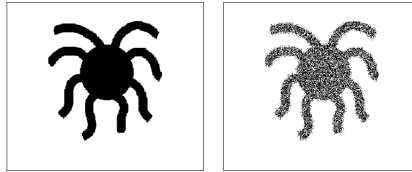


Figure 3.4 – Deux images binaires d'une pieuvre. L'image à droite est une version bruitée de celle de gauche [21].

Les pixels noirs de l'image gauche et droite représentent les ensembles K_1, K_2 à l'étude,

respectivement, alors que dans les deux cas, les rectangle des pixels noirs et blancs constituent l'ensemble X . L'ensemble bruité K_2 ainsi obtenu est proche de l'ensemble d'origine K_1 par rapport à la distance de Hausdorff.

On choisit la fonction de mesure pour les deux cas comme suit : $\varphi : X \rightarrow \mathbb{R}$, $\varphi(p) = -\|p - c\|$ où $c \in X$ correspond au centre de masse de K_1 .

La Figure 3.5 montre le diagramme de persistance du nombre de Betti $B_{\phi|_{K_1}}$ en dimension 0. Il affiche huit points pertinents dans le diagramme de persistance correspondants aux huit tentacules de la pieuvre. Un seul de ces points est à l'infini (et donc représenté par une ligne verticale plutôt que par un cercle) puisque K_1 n'a qu'une seule composante connexes. En ce qui concerne $B_{\phi|_{K_2}}$, en raison de la présence d'une grande quantité de composantes connexes dans l'image bruitée de la pieuvre, son diagramme de persistance présente un très grand nombre de points à l'infini. Ainsi, une figure qui montre tous ces points serait peu lisible. Pour cette raison, la Figure 3.5 (à droite) ne montre qu'un petit sous-ensemble de son diagramme de persistance. Cependant, il suffit de percevoir combien celui-ci est différent de $B_{\phi|_{K_2}}$.

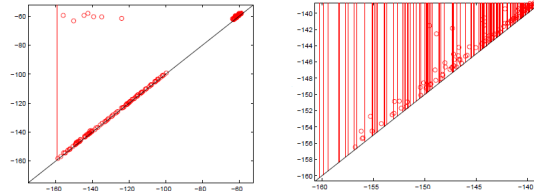


Figure 3.5 – À gauche : le diagramme de persistance correspondant à l'image originale de la pieuvre. À droite : le diagramme de persistance correspondant à l'image de la pieuvre bruitée [21].

D'après le Théorème 3.13, si on compare K_1 et K_2 au moyen de $B_{\vec{\phi}_1}$ et $B_{\vec{\phi}_2}$, où $\vec{\phi}_1 : X \rightarrow \mathbb{R}^2$, $\vec{\phi}_1 = (d_{K_1}, \varphi)$, et $\vec{\phi}_2 : X \rightarrow \mathbb{R}^2$, $\vec{\phi}_2 = (d_{K_2}, \varphi)$, on peut constater la similitude entre K_1 et K_2 .

Puisque le domaine de $B_{\vec{\phi}_1}$ et $B_{\vec{\phi}_2}$ est dans $\mathbb{R}^2 \times \mathbb{R}^2$, nous montrons ces invariants de rang

au moyen de la méthode de la foliation, ce qui permet une réduction unidimensionnelle.

Pour ça on considère la fonction $F_{(\vec{l}, \vec{b})}^{\vec{\phi}_i} : X \rightarrow \mathbb{R}$, pour tout $x \in X$,

$$F_{(\vec{l}, \vec{b})}^{\vec{\phi}_i}(x) = \max \left\{ \frac{d_{K_i}(x) - b_1}{l_1}, \frac{\varphi(x) - b_2}{l_2} \right\},$$

pour $i = 1, 2$, et $\mu = \min \{l_1, l_2\}$.

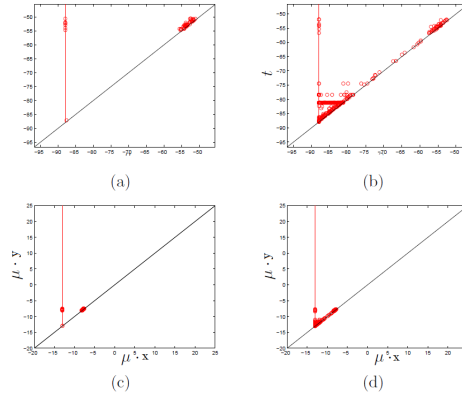


Figure 3.6 – (a) Les nombres de Betti persistants $\beta_{\vec{\phi}_1}^{\vec{l}}$ restreints au demi plan $\pi_{(\vec{l}, \vec{b})}$. (b) Les nombres de Betti persistants $\beta_{\vec{\phi}_2}^{\vec{l}}$ restreints au même demi plan. (c)-(d) Le même (a) et (b) respectivement multiplier par $\mu = \min \{l_1, l_2\}$ [21].

La stabilité de notre approche est illustrée dans la Figure 3.6. La Figure 3.6 (a)-(b), montre les nombres de Betti persistants $B_{\vec{\phi}_1}^{\vec{l}}$ et $B_{\vec{\phi}_2}^{\vec{l}}$ restreints au demi plan $\pi_{(\vec{l}, \vec{b})}$, avec $\vec{l} = (0.1483, 0.9889)$ et $\vec{b} = (13.0434, -13.0434)$. Le demi-plan considéré a été choisi de sorte qu'il contient des points où les invariants de rang prennent des valeurs non triviales. Nous pouvons déjà apprécier leur similitude, même si leur distance n'est pas nécessairement plus petite que la distance de Hausdorff entre K_1 et K_2 . Dans la Figure 3.6 (c)-(d), nous montrons les nombres de Betti persistants de $\mu \cdot F_{(\vec{l}, \vec{b})}^{\vec{\phi}_1}$ et $\mu \cdot F_{(\vec{l}, \vec{b})}^{\vec{\phi}_2}$, qui apparaissent dans la définition de D_{match} . La stabilité de D_{match} garantit que la distance entre ces deux nombres de Betti persistants n'est pas supérieure à la distance Hausdorff entre K_1 et K_2 .

CHAPITRE 4

Applications

Dans ce chapitre, nous allons effectuer des calculs de l'homologie persistante sur différentes formes avec le logiciel Javaplex [15, 16]. Javaplex est un logiciel développé par Vin De Silva et Gunnar Carlsson à l'université de Stanford. Il nous permet de calculer l'homologie persistante de nuages de points avec différentes méthodes de filtration (Figure 4.1).

Nous avons réalisé plusieurs expériences ; dans la première nous avons calculé l'homolo-

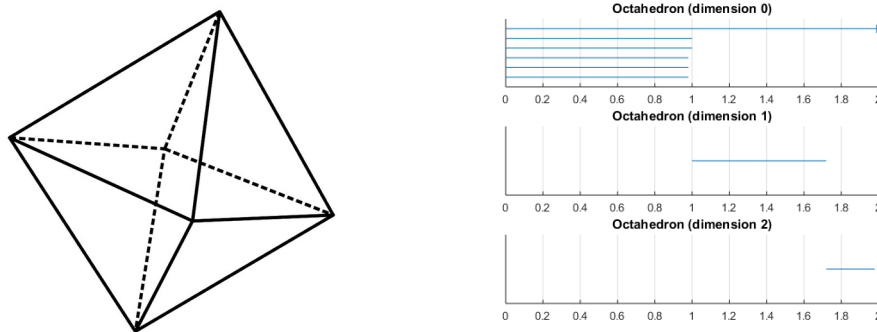


Figure 4.1 – Octaèdre régulier et son diagramme de persistance en dimensions 0,1 et 2.

gie persistante sur un nuage de points de différentes formes plus ou moins proches l'une de l'autre avec la méthode de complexe de Rips [9]. Dans la deuxième expérience, nous avons calculé l'homologie persistante sur des images réelles où la taille des images est en général de 500×500 pixels. En raison d'un très grand nombre de simplexes nous avons changé la méthode de filtration à la méthode de Lazy Witness Complexe [10].

On dit que le diagramme de persistance est stable si un petit changement dans la fonction d'entrée produit un petit changement dans le diagramme de persistance. Nous avons utilisé la distance d'appariement (Bottleneck distance) pour en déduire la stabilité des diagrammes de persistance entre différentes nervures des plantes, différentes plantes et enfin entre un tore et un double tore.

Première expérience :

Exemple 4.1. *Dans cet exemple, nous avons calculé l'homologie persistante et la stabilité du diagramme de persistance. On présente alors sept images différentes : des nervures présentées de différentes façons (opposées, alternées, palmées) et afin de bien étudier la stabilité, nous réalisons des petits changements sur l'image des nervures opposées telles que des nervures opposées moins une seule nervure, des nervures opposées moins une seule nervure autre que la première, des nervures opposées moins deux nervures, des nervures opposées moins trois nervures.*

Afin d'effectuer ces calculs, on applique la filtration avec la méthode du complexe de Rips et le programme Javaplex sur les nuages de points de ces données où la dimension maximale d_{max} choisie est 2, l'homologie persistante du complexe résultant est calculée dans la dimension d_{max-1} , le nombre de divisions est $N = 200$, et enfin la valeur de filtration maximale $t_{max} = 40$. Ces entrées nous permettent de limiter la taille du complexe construit, pour une efficacité de calcul. De plus, au lieu de calculer le complexe simplicial filtré pour tout $t > 0$, on calcule seulement $\{0, t_{max}/(N - 1), \dots, (N - 2)t_{max}/(N - 1), t_{max}\}$.

Les figures suivantes affichent les nervures étudiées, leurs diagrammes de persistance et la visualisation des diagrammes de persistance en dimension 0.

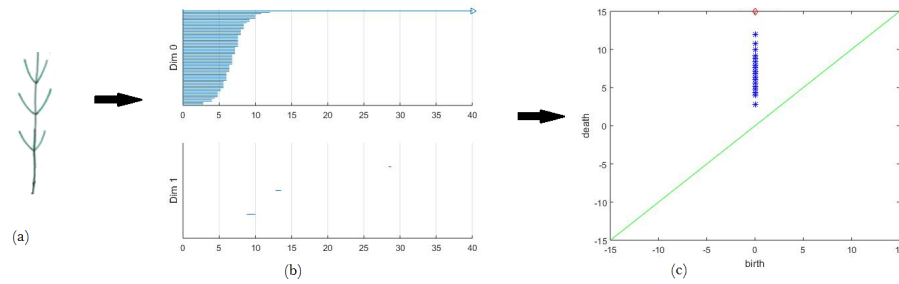


Figure 4.2 – (a) Nervures opposées. (b) Diagramme de persistance. (c) Visualisation du diagramme de persistance.

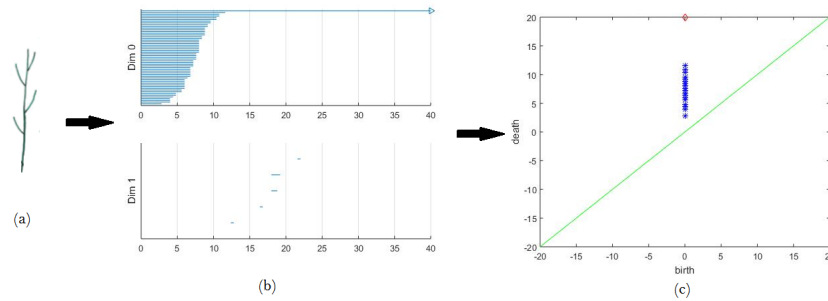


Figure 4.3 – (a) Nervures alternées. (b) Diagramme de persistance. (c) Visualisation du diagramme de persistance.

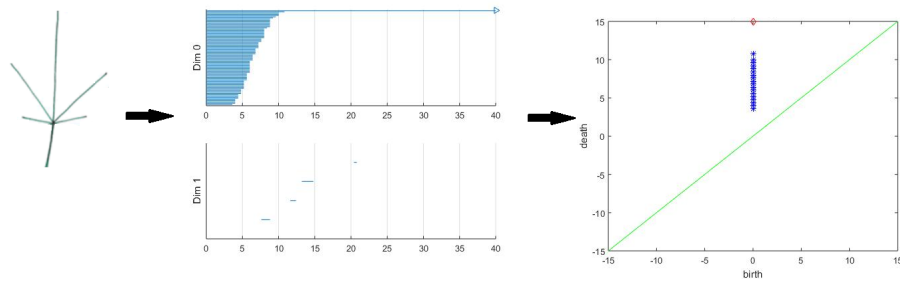


Figure 4.4 – (a) Nervures palmées. (b) Diagramme de persistance. (c) Visualisation du diagramme de persistance.

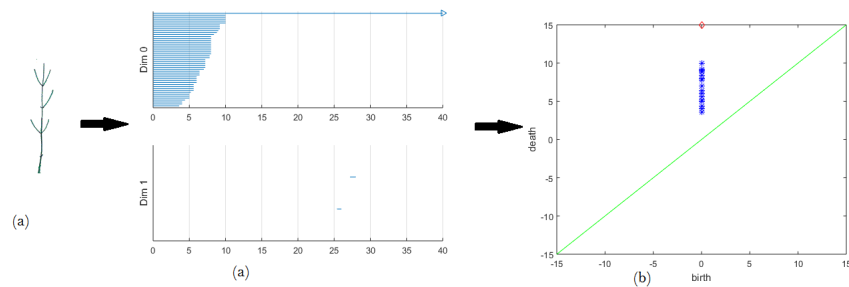


Figure 4.5 – (a) Nervures opposées moins une nervure. (b) Diagramme de persistance. (c) Visualisation du diagramme de persistance.

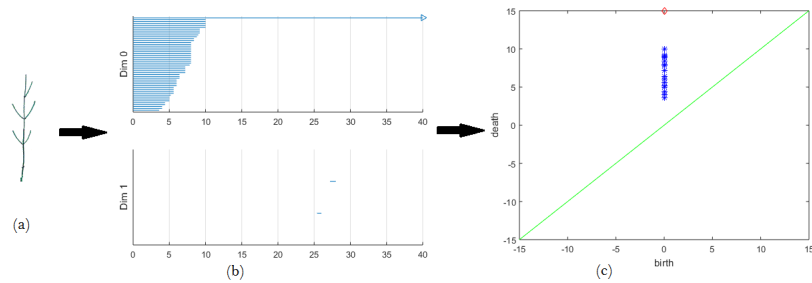


Figure 4.6 – (a) Nervures opposées moins une nervure différente. (b) Diagramme de persistance. (c) Visualisation de diagramme de persistance.

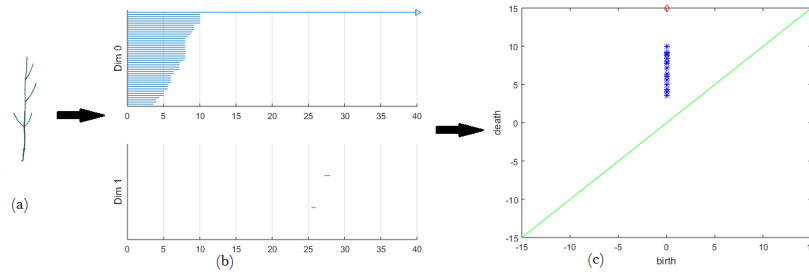


Figure 4.7 – (a) Nervures opposées moins deux nervures. (b) Diagramme de persistance. (c) Visualisation du diagramme de persistance.

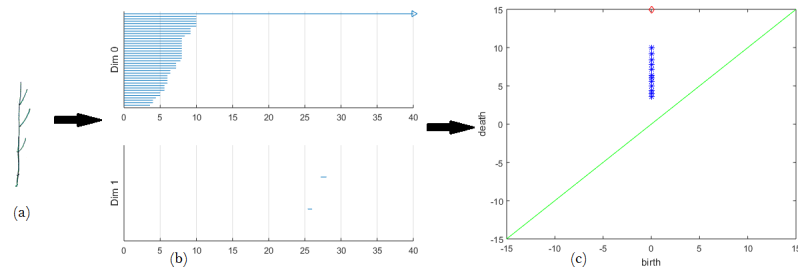










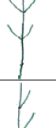






Figure 4.8 – (a) Nervures opposées moins trois nervures. (b) Diagramme de persistance. (c) Visualisation du diagramme de persistance.

Le tableau ci-dessous affiche les distances de Bottleneck en dimension 0 entre toutes les différentes images. Les résultats affichés montrent la stabilité des diagrammes de persistance.

En effet, des changements mineurs dans le nuage de points produisent des petits changements dans le diagramme de persistance. Par exemple, la distance entre les nervures opposées et les nervures opposées moins une seule nervure ($D = 2.4$) augmente à $D = 2.8$ entre les nervures opposées et les nervures opposées moins deux nervures.

							
	0	2.4	2.4	2.8	3	1.8	2.8
	2.4	0	1.8	2.4	2.8	2.4	3
	2.4	1.8	0	2.4	2.8	2.2	3
	2.8	2.4	2.4	0	2.5	3	3.12
	3	2.8	2.8	2.5	0	3.4	3.2
	1.8	2.4	2.2	3	3.4	0	2.8
	2.8	3	3	3.12	3.2	2.8	0

Exemple 4.2. Dans cet exemple, nous avons étudié trois différentes plantes avec 2, 4 et 5 trous dans chaque image respectivement, où les trous représentent des feuilles vides. Le but de cet exemple est d'étudier la stabilité des diagrammes de persistance en dimension un.

Les Figures (4.9), (4.10) et (4.11) affichent en (a) notre image étudiée, (b) les diagrammes de persistance et la visualisation de ces diagrammes en dimension 0 et 1 en (c) et (d) respectivement.

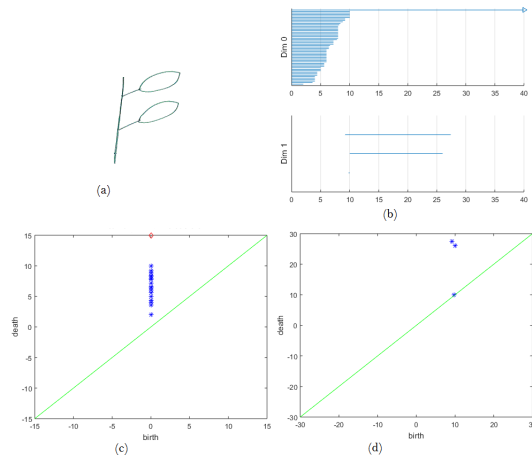


Figure 4.9 – (a) Plante avec deux feuilles. (b) Diagramme de persistance. (c) Visualisation de diagramme de persistance en dimension 0. (d) Visualisation du diagramme de persistance en dimension 1.

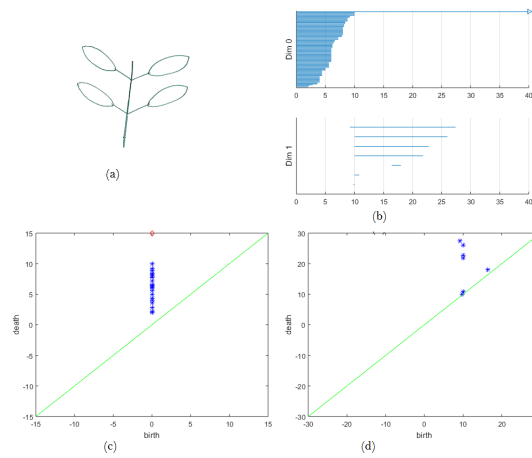


Figure 4.10 – (a) Plante avec quatre feuilles. (b) Diagramme de persistance. (c) Visualisation du diagramme de persistance en dimension 0. (d) Visualisation du diagramme de persistance en dimension 1.

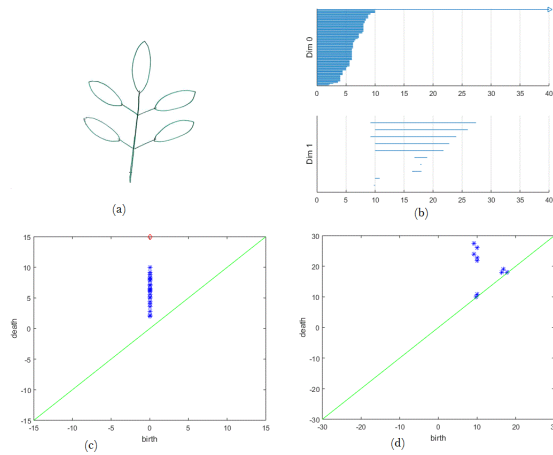


Figure 4.11 – (a) Plante avec cinq feuilles. (b) Diagramme de persistance. (c) Visualisation du diagramme de persistance en dimension 0. (d) Visualisation du diagramme de persistance en dimension 1.

Le tableau ci-dessous affiche les résultats des distances de Bottleneck en dimension 0 et 1 respectivement. On voit bien ici la stabilité des diagrammes de persistance en dimension 1. Par exemple, la distance entre la plante de deux feuilles et celle de quatre feuilles est de 6,4, cette distance augmente à 7,4 avec la plante de cinq feuilles.

			
	0	3	3
	0	6.4	7.4
	3	0	2.2
	6.4	0	5.8
	3	2.2	0
	7.4	5.8	0

Deuxième expérience :

Exemple 4.3. Dans cette expérience, nous avons calculé l'homologie persistante sur des données réelles. Comme la méthode de filtration de complexe de Rips ne fonctionne pas ici en raison du volume très grand des données, nous avons utilisé une nouvelle méthode de filtration plus efficace : la méthode du Lazy witness Complexe, qui a l'avantage d'être un flag complexe (comme le complexe de Rips) et de n'utiliser qu'un nombre réduit de points pour faire les calculs.

La base de données de formes qu'on a utilisée dans cette expérience est la base de données

de silhouettes de formes MPEG-7.

Le nombre de points choisis ici est 300, où la taille de notre image est en général 500×500 pixels, le nombre de filtrations maximales est $t_{max} = 5$ et le nombre de divisions $N = 500$. Enfin, le calcul de l'homologie persistante est fait en dimension maximale $d_{max} = 2$. Les figures suivantes montrent les images, les codes barres et leurs visualisations en dimensions 0 et 1.

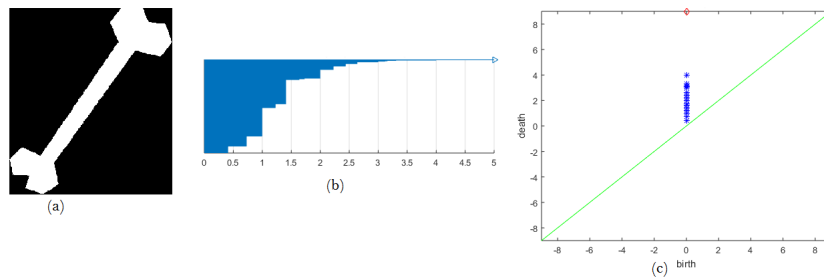


Figure 4.12 – (a) Image d'os. (b) Diagramme de persistance. (c) Visualisation du diagramme de persistance.

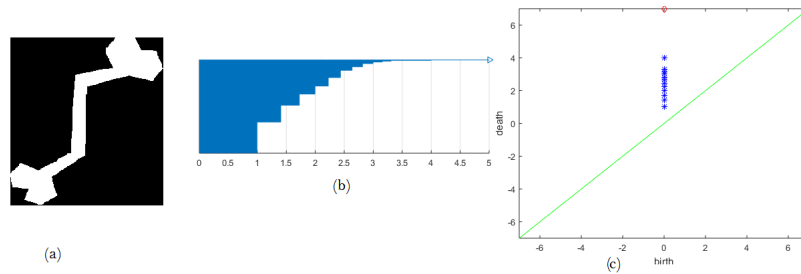


Figure 4.13 – (a) Image d'os différente. (b) Diagramme de persistance. (c) Visualisation du diagramme de persistance.

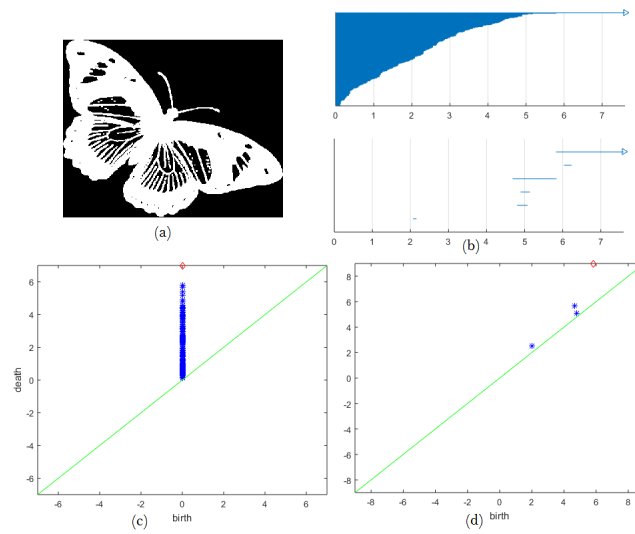


Figure 4.14 – (a) Image de papillon. (b) Diagramme de persistance. (c) Visualisation du diagramme de persistance en dimension 0. (d) Visualisation du diagramme de persistance en dimension 1.

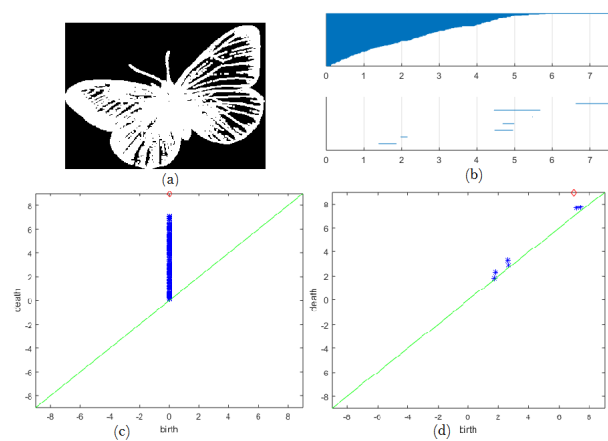


Figure 4.15 – (a) Image de papillon différente. (b) Diagramme de persistance. (c) Visualisation du diagramme de persistance en dimension 0. (d) Visualisation du diagramme de persistance en dimension 1.

Exemple 4.4. Dans cet exemple nous avons la distance entre un tore et un double tore. Les figures suivantes montrent le tore et le double tore, leurs diagrammes de persistance

en dimension 0, 1 et 2, et enfin leurs visualisations.

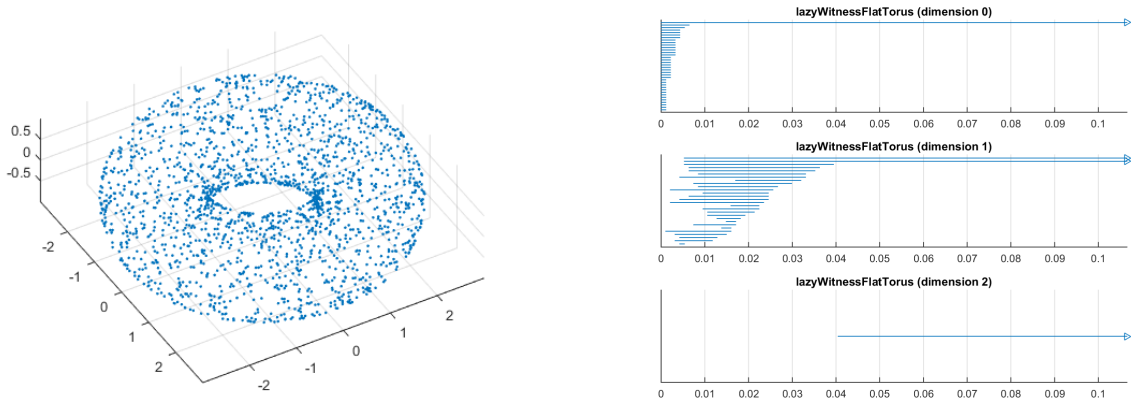


Figure 4.16 – Tore et son diagramme de persistance.

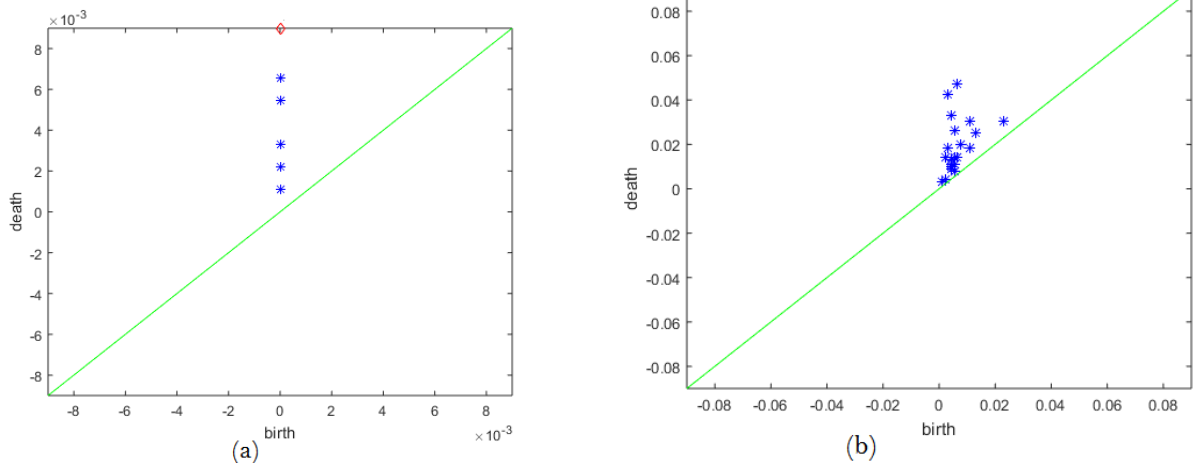


Figure 4.17 – (a) Visualisation du diagramme de persistance en dimension 0. (b) Visualisation du diagramme de persistance en dimension 1.

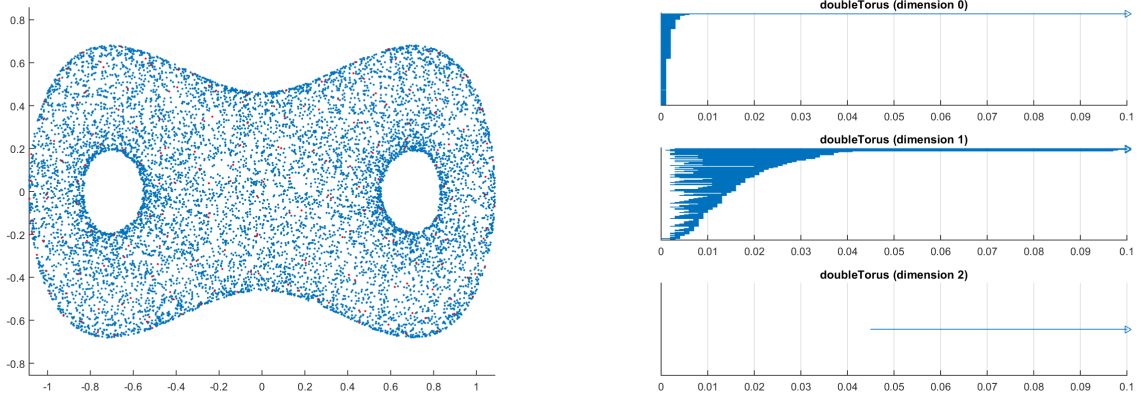


Figure 4.18 – Double tore et son diagramme de persistance.

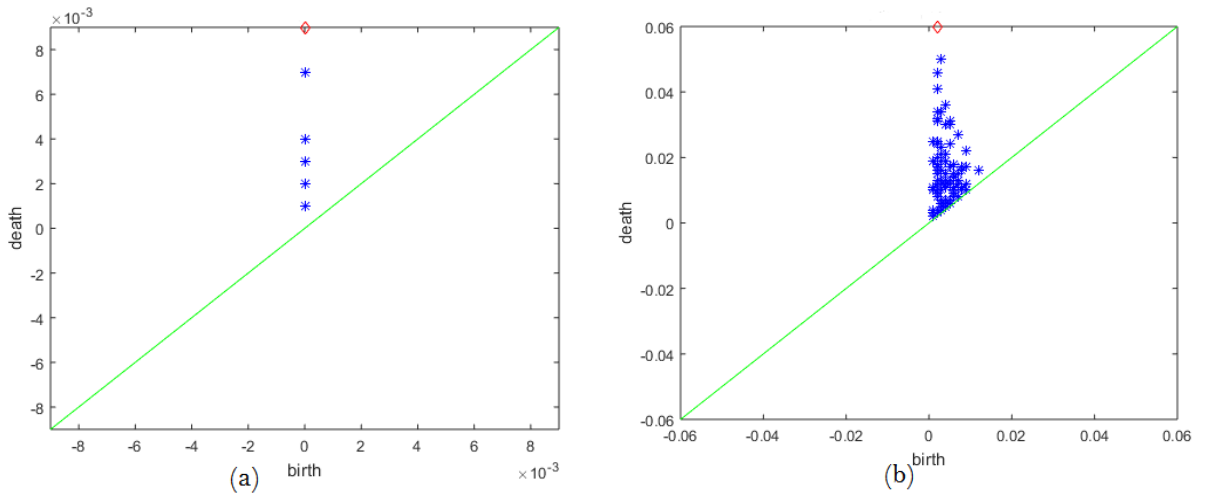


Figure 4.19 – (a) Visualisation du diagramme de persistance en dimension 0. (b) Visualisation du diagramme de persistance en dimension 1.

La distance de Bottleneck du tore et double tore en dimension 1 est de $1.3408 \exp^{154}$. Le nombre de Betti en dimension 1 est de $B_1 = 2$ dans le cas du tore. Par contre, dans le cas du double tore, le nombre de Betti en dimension 1 est de $B_1 = 4$ comme le montrent les diagrammes de persistances avec des lignes continues.

CONCLUSION

L'homologie persistante, qui montre la durée de la persistance des propriétés topologiques le long d'une filtration d'espace, est très appliquée à de nombreuses disciplines, notamment à la reconnaissance et la comparaison des formes dans plusieurs domaines comme l'informatique, la biologie, etc.

Dans le cas de l'homologie persistante unidimensionnelle, on a étudié les formes avec la fonction de taille qui peut être considérée comme une série formelle permettant ainsi une manipulation plus efficace [1]. Cette présentation formelle nous permet de calculer la distance d'appariement entre les formes et de distinguer les différences [3].

Nous avons appliqué cette méthode dans le cas discret sur des formes. Nous avons apporté des changements mineurs à quelques formes pour démontrer la possibilité d'appliquer la théorie de la stabilité unidimensionnelle. D'autres exemples d'applications sont cités dans des articles récents [27–30].

Biasotti et ses collaborateurs [19] ont prouvé dans le cas d'homologie persistante multidimensionnelle, en réduisant la théorie de fonction de taille multidimensionnelle au cas de fonction unidimensionnelle par un changement de variables approprié, que la fonction de taille multidimensionnelle est stable par rapport à la distance d'appariement. D'autre part, [21, 22] ont montré comment se servir de la distance d'appariement multidimensionnelle pour traiter les changements dans l'espace topologique étudié. En effet, en supposant que l'ensemble est triangulable, il suffit d'étudier la fermeture de l'intérieur

de l'ensemble au lieu de l'ensemble lui-même.

Une suite tout à fait envisageable de ce travail serait un calcul rapide de la distance d'appariement multidimensionnelle. En effet, de nouveaux résultats théoriques ont été introduits [31] afin d'estimer l'erreur de calcul dans l'évaluation de la distance d'appariement bidimensionnelle. Ces résultats ont conduit à la définition d'un nouvel algorithme dont l'entrée est un seuil d'erreur arbitraire et la sortie est une approximation de la distance d'appariement bidimensionnelle. Toutefois, des recherches en cours essaient d'étendre l'algorithme à des dimensions plus élevées.

Finalement, pour un travail futur, il serait très intéressant de montrer l'utilité de la théorie de Morse discrète [32] pour le calcul de l'homologie persistante. *Mischaikow* et *Nanda* présentent un algorithme qui utilise la théorie de Morse pour la filtration unidimensionnelle [33]. En effet, ils ont développé un algorithme de prétraitement efficace pour réduire le nombre de cellules dans un complexe cellulaire filtré tout en préservant ses groupes d'homologie persistante. Pour une extension au cas de filtration multidimensionnelle, veuillez consulter [24, 34].

Bibliographie

- [1] P. Frosini, and C. Landi. Size functions and formal series. *Appl. Algebra Eng., Commun. Comput.*, 12(4) : 327-349, 2001.
- [2] A. Verri, C. Uras, P. Frosini, and M. Ferri. On the use of size functions for shape analysis *Biological Cybernetics*, 70(2) : 99-107, 1993.
- [3] C. Landi, and P. Frosini. New pseudo-distances for the size function space *Proc. SPIE, Vision Geometry*, 3168 : 52-60, 1997.
- [4] J. Munkres. *Elements of Algebraic Topology*. Addison-Wesley, 1984.
- [5] F. Croom. *Basic concepts of algebraic topology*. Springer Science & Business Media, 2012.
- [6] A. Zomorodian. *Topology for computing*. Cambridge University Press, 2009.
- [7] H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological persistence and simplification. *Discrete & Comput. Geom.*, 28(4) : 511-533, 2002.
- [8] A. Zomorodian, and G. Carlsson. Computing persistent homology. *Discrete & Comput. Geom.*, 33(2) : 347-356, 2005.
- [9] R. Ghrist. Barcodes : The persistent topology of data. *Bulletin-American Mathematical Society*, 45(1) : 61-75, 2008.

- [10] V. de Silva, G. Carlsson. Topological estimation using witness complexes. *Eurographics Symposium on Point-Based Graphics*, 157-166, 2004.
- [11] H. Edelsbrunner, and J. Harer. *Computational Topology : An Introduction*. American Mathematical Soc., 2010.
- [12] G. Carlsson. Topology and data. *Bulletin of the American Mathematical Society*, 46(2) : 255-308, 2009.
- [13] D. Horak, S. Maletić, and M Rajković. Persistent homology of complex networks. *Journal of Statistical Mechanics : Theory and Experiment*, 2009(3), P03034, 2009.
- [14] P. Pranav. *Persistent holes in the Universe : A hierarchical topology of the cosmic mass distribution*, University of Groningen, 2015.
- [15] A. Tausz, M. Vejdemo-Johansson, and H. Adams. Javaplex : A research software package for persistent (co)homology. In H. Hong and C. Yap, editors, Proceedings of ICMS 2014, Lecture Notes in Computer Science 8592, 129-136, 2014. url : <http://appliedtopology.github.io/javaplex/>.
- [16] H. Adams, and A. Tausz. Javaplex tutorial. 2016.
- [17] F. Iuricich, S. Scaramuccia, C. Landi, and L. De Floriani. A Discrete Morse-based Approach to Multivariate Data Analysis. *SIGGRAPH ASIA 2016 Symposium on Visualization*, 5 :1-5 :8, 2016.
- [18] C. Andrea, and L. Claudia. Hausdorff Stability of Persistence Spaces. *Foundations of Computational Mathematics*, 16(2), 343-367, 2016.
- [19] S. Biasotti, A. Cerri, P. Frosini, D. Giorgi, and C. Landi. Multidimensional size functions for shape comparison. *J. Math. Imaging Vision* , 32(2) : 161-179, 2008.

- [20] F. Cagliari, B. Di Fabio, and M. Ferri. One-dimensional reduction of multidimensional persistent homology. *Proc. Amer. Math. Soc.*, 138(8) : 3003-3017, 2010.
- [21] M. Ferri, P. Frosini, and C. Landi. Stable shape comparison by persistent homology. *Atti Semin. Mat. Fis. Univ. Modena Reggio Emilia.*, 2011.
- [22] A. Cerri, B. Di Fabio, P. Frosini, and C. Landi. Betti numbers in multidimensional persistent homology are stable functions. *Math. Meth. Appl. Sci.*, 2013. DOI : 10.1002/mma.2704.
- [23] P. Frosini, and C. Landi. Stability of multidimensional persistent homology with respect to domain perturbations. arXiv :1001.1078v2 (2010).
- [24] M. Allili, T. Kaczynski, and C. Landi. Reducing complexes in multidimensional persistent homology theory. *J. Symb. Comput.*, 78 : 61-75, 2017.
- [25] A. Cerri, M. Ethier, and P. Frosini. The coherent matching distance in 2D persistent homology. *Computational Topology in Image Context*, 9667 : 216-227, 2016.
- [26] M. Ethier. *De multiples applications de l'homologie à l'imagerie numérique* thèse de doctorat, Université de Sherbrooke, Départ. de math, 2013.
- [27] M. Buchet, F. Chazal, T. K. Dey, F. Fan, S. Y. Oudot, and Y. Wang. Topological analysis of scalar fields with outliers. *CoRR*, abs/1412.1680, 2014.
- [28] M. Li, K. Duncan, C. N. Topp, and D. H. Chitwood. Persistent homology and the branching topologies of plants. *American journal of botany*, 104(3) : 349-353, 2017.
- [29] Z. Cang, and G. -W. Wei. Analysis and prediction of protein folding energy changes upon mutation by element specific persistent homology. *ArXiv e-prints*, arXiv preprint arXiv : 1703.10966, 2017.

- [30] P. Bendich, J.S. Marron, E. Miller, A. Pieloch, and S. Skwerer. Persistent homology analysis of brain artery trees. *The annals of applied statistics*, 10(1) : 198-218, 2016.
- [31] S. Biasotti, A. Cerri, P. Frosini, and D. Giorgi. A new algorithm for computing the 2-dimensional matching distance between size functions. *Pattern recognition letters*, 32(14) : 1735-1746, 2011.
- [32] R. Forman. Morse theory for cell complexes. *Advances in mathematics*, 134(1) : 90-145, 1998.
- [33] K. Mischaikow, and V. Nanda. Morse Theory for filtrations and efficient computation of persistent homology. *Discrete & Comput. Geom.*, 50(2) : 330-353, 2013.
- [34] M. Allili, T. Kaczynski, C. Landi, and F. Masoni. A New Matching algorithm for multidimensional persistence. *CoRR*, abs/1511.05427, 2015.