Geometric Inference

Frédéric Chazal[1] and David Cohen-Steiner[2]

[1] INRIA Futurs, 4 rue Jacques Monod - 91893 Orsay Cedex France
frederic.chazal@inria.fr
[2] INRIA Sophia-Antipolis, 2004 route des lucioles - BP 93 06902 Sophia Antipolis Cedex France david.cohen-steiner@sophia.inria.fr

1 Introduction

In many practical situations, geometric objects are only known through a finite set of possibly noisy sample points. A natural question is then to recover the geometry and the topology of the unknown object from this information. The most classical example is probably surface reconstruction, where the points are measured on the surface of a real world object. A perhaps more surprising example is the study of the large scale structure formed by the galaxies, which cosmologists believe to be an interconnected network of walls and filaments [35]. In other applications, the shape of interest may be a low dimensional object embedded in a higher-dimensional space, which is the basic assumption in *manifold learning*. This is for example the case in time series analysis, when the shape of study is the attractor of a dynamical system sampled by a sequence of observations [33].

A lot of research was done in this direction, which makes it difficult to give a comprehensive treatment of the problem. We may for example mention the celebrated correlation dimension [27], which is widely used as an estimate of the intrinsic dimension of data sets. Other contributions come from the field of computational geometry, where much effort was done to elaborate provably correct surface reconstruction algorithms, under a suitable sampling condition. We refer to [18] for a thorough review of this approach. However most of this research focused on the case of sampled smooth surfaces in $\mathbb{R}^3$, which is by now fairly well covered. Notable exceptions are [32, 19], which deal with inference of local dimension and topology for higher-dimensional, but still smooth, objects.

The goal of this chapter is to describe particular frameworks that allow to handle more general shapes, *i.e.* (almost) general compact sets in Euclidean space. The first section reviews recent work on the subject based on distance functions, and related concepts such as critical points and medial axes. The second one describes an algebraic-topological tool, persistent homology, which is also highly relevant to geometric inference problems.

2 Distance functions

Topological and geometric features of a shape cannot be directly extracted from an approximating data: for example, the number of connected components of a shape is obviously not the same as the one of a point cloud approximating it. Worse, the occurrence of some features may depend on a "scale" at which both the data and the shape are viewed: for example, viewed with human eyes, the surface of a real world object may look very regular but at a microscopic scale it appears as a much more complicated surface with many holes, tunnels, *etc.*..It has been recently shown that the framework of distance functions allows to investigate these problems in a fairly general setting, assuming a very simple reconstruction process [11, 8].

In the following, all the considered shapes and their approximations (usually point clouds) are represented by compact subsets of Euclidean space $\mathbb{R}^n$, $n \geq 1$. In particular, in the following, the word 'shape' has to be understood as 'compact set'. Given a compact set $K \subset \mathbb{R}^n$, the *distance function* d_K to K is the non-negative function defined by

$$d_K(x) = \inf_{y \in K} d(x, y) \quad \text{for all} \quad x \in \mathbb{R}^n$$

where $d(x, y)$ denotes the usual Euclidean distance. The *offsets* K^r of K are the sublevel sets of the distance function: $K^r = d_K^{-1}([0, r])$. To quantify the notion of approximation it is necessary to introduce a distance measuring the closeness between shapes. The *Hausdorff distance* $d_H(K, K')$ between two compact sets K and K' in $\mathbb{R}^n$ is the minimum number r such that $K \subset K'^r$ and $K' \subset K^r$. The Hausdorff distance defines a distance on the space of compact subsets of $\mathbb{R}^n$ related to distance functions by the following equality:

$$d_H(K, K') = \|d_K - d_{K'}\|_\infty := \sup_{x \in \mathbb{R}^n} |d_K(x) - d_{K'}(x)|$$

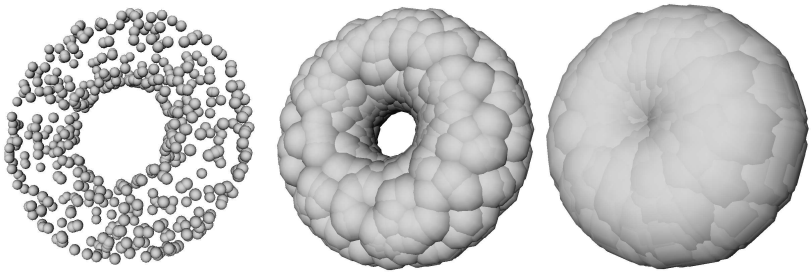

Fig. 1. Three different offsets of a set of points sampled around a torus. The middle one carries the topology of the torus.

The distance function approach for topological and geometric inference aims at comparing the offsets of shapes that are close for the Hausdorff dis-

tance. The underlying intuition is that "at some scales" (i.e. for some range values of the offsets) two close shapes should have the same offsets topology (see Figure 1). Making such an intuition into a formal statement requires to proceed in several steps. First, it is necessary to understand how the topology of the offsets K^r of a single compact K evolves with the offset parameter r. This can be achieved using a critical point theory for distance functions that has been developped in Riemannian geometry [13, 28] (Section 2.1). Second, it is necessary to compare the topology of the offsets of two close compact sets. This leads to stability results and sampling conditions providing theoretical guarantees for a simple reconstruction process (Section 2.2).

2.1 Critical points and offset topology

Although distance functions are continuous and even 1-Lipschitz [3], they are usually not differentiable everywhere. Nevertheless, distance functions have been studied and used for a long time in Riemannian geometry [13, 29, 28] and non-smooth analysis [14] where it has been proved that they behave almost like differentiable functions. In particular, the topology of their level sets can only change at some well-defined critical points. Moreover, it is possible to define a generalized gradient vector field $\nabla_K : \mathbb{R}^n \mapsto \mathbb{R}^n$ of the distance function d_K [31] in the following way.

The gradient of the distance function

Intuitively, the direction of the gradient of d_K at a point x has to be the one along which the directional derivative of d_K at x is the biggest or equivalently the direction in which the "slope"' of the graph $\{(y, d_K(y)) : y \in \mathbb{R}^n\}$ of d_K is the biggest at $(x, d_K(x))$ (see Figure 2). The norm of the gradient has then to be the directional derivative of d_K (or equivalently the "slope" of the graph of d_K) in this direction.

More formally, for any point $x \in \mathbb{R}^n$, we denote by $\Gamma_K(x)$ the set of points in K closest to x (Figure 3):

$$\Gamma_K(x) = \{y \in K \,|\, d(x, y) = d(x, K)\}$$

Note that $\Gamma_K(x)$ is a non-empty compact set. There is a unique smallest closed ball $\sigma_K(x)$ enclosing $\Gamma_K(x)$ (cf. Figure 3). We denote by $\theta_K(x)$ the center of $\sigma_K(x)$ [4] and by $F_K(x)$ its radius. For $x \in \mathbb{R}^n \setminus K$, the generalized gradient $\nabla_K(x)$ is defined by:

$$\nabla_K(x) = \frac{x - \theta_K(x)}{d_K(x)}$$

[3] that is $|d_K(x) - d_K(y)| \leq d(x, y)$ for all $x, y \in \mathbb{R}^n$

[4] Equivalently, one can prove that $\theta_K(x)$ is the projection of x on the convex hull of $\Gamma_K(x)$.

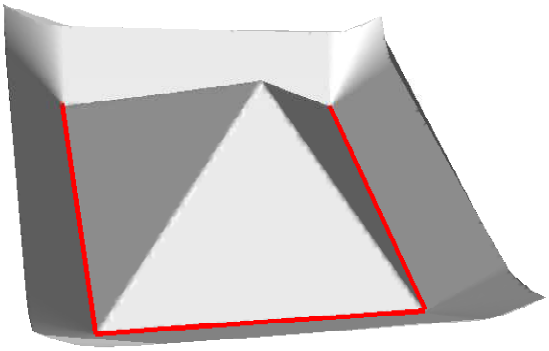

Fig. 2. The graph of the distance function to the boundary of a square in the plane.

It is natural to set $\nabla_K(x) = 0$ for $x \in K$. The norm of $\nabla_K(x)$ is given by

$$\|\nabla_K(x)\|^2 = 1 - \frac{F_K(x)^2}{d_K(x)^2}$$

Equivalently, $||\nabla_K(x)||$ is the cosine of the (half) angle of the smallest cone with apex x that contains $\Gamma_K(x)$, which is intuitively the value of the directional derivative of d_K in the direction of $\nabla_K(x)$.

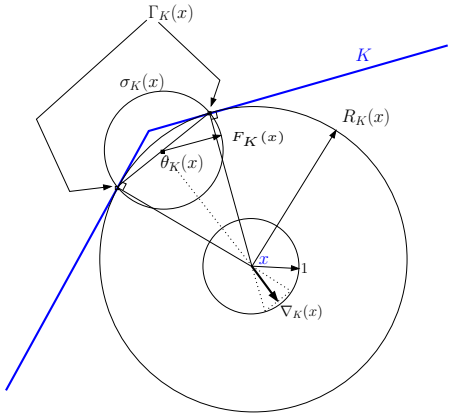

Fig. 3. Definition of the gradient of d_K.

The vector field ∇_K is obviously not continuous as shown on Figure 4 but it can be shown [31] that Euler schemes using ∇_K converge uniformly when the integration step decreases, toward a continuous flow $\mathfrak{C} : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^n$. The curves $t \mapsto \mathfrak{C}(t, x)$ are the trajectories of the gradient vector field and two such trajectories can merge but one trajectory cannot "fork" (see Figure

4). It is proven in [31] that the functions F_K and d_K are increasing along the trajectories of the flow. As we will see in the following, this flow plays a fundamental role in the study of the topological stability properties of the offsets of close shapes. In the particular case where K is a finite set, various notions of flows related to this one have been independently introduced by H. Edelsbrunner [21], J. Giesen and al. [26] and R. Chaine [7] using Voronoi diagrams and Delaunay triangulations.

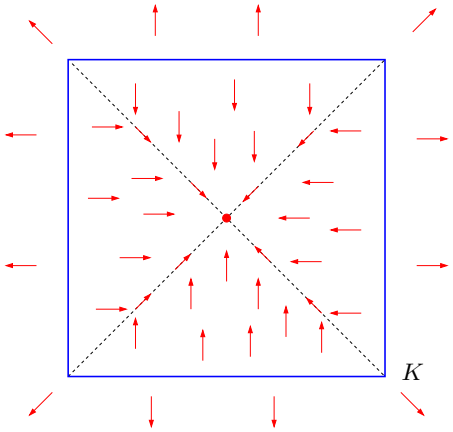

Fig. 4. The gradient vector field of the distance function to a square K in the plane: outside of the diagonals, the norm of ∇_K is equal to 1. On the diagonal it is equal to $\frac{1}{\sqrt{2}}$ except at the center where it is equal to 0 (critical point).

Critical points of d_K and topology of the offsets

The *critical points* of d_K are defined as the points x for which $\nabla_K(x) = 0$. Equivalently, a point x is a critical point if and only if it lies in the convex hull of $\Gamma_K(x)$. When K is a finite point cloud, this last definition means that critical points are precisely the intersections of Delaunay k-dimensional simplices with their dual $(n-k)$-dimensional Voronoi facets [26]. A real $c > 0$ is said to be a *critical value* of d_K is there exists a critical point $x \in \mathbb{R}^n$ such that $d_K(x) = c$. A non critical value of d_K is called a *regular value*. Note that this notion of critical point is the same as the one considered in the setting of non-smooth analysis [14] and Riemannian geometry [13, 28].

The topology of the offsets K^α of a compact set K is closely related to the critical values of d_K.

Theorem 1 (Isotopy lemma [28]). *If $0 < r_1 < r_2$ are such that $K^{r_2} \setminus int(K^{r_1})$ does not contain any critical point of d_K, then all the level sets*

$d_K^{-1}(r)$, $r \in [r_1, r_2]$, are homeomorphic and even isotopic [5] *topological manifolds and*

$$K^{r_2} \setminus int(K^{r_1}) = \{x \in \mathbb{R}^n : r_1 \leq d_K(x) \leq r_2\}$$

is homeomorphic to $d_K^{-1}(r_1) \times [r_1, r_2]$. As a consequence, K^{r_1} and K^{r_2} are isotopic.

In other words, the topology of the offsets of K can only change at critical values. The *weak feature size* of K, or $\text{wfs}(K)$, is defined as the infimum of the positive critical values of d_K. Equivalently it is the minimum distance between K and the set of critical points of d_K. Notice that $\text{wfs}\,(K)$ may be equal to 0. Nevertheless, $\text{wfs}\,(K)$ is non zero for a large class of compact sets including polyhedra and piecewise analytic sets [10, 11]. It follows from the Isotopy Lemma that that $\text{wfs}\,(K)$ may be viewed as the "minimum size of the topological features" of the set K: if $0 < r, s < \text{wfs}\,(K)$ then K^r and K^s are homeomorphic and even isotopic. The same holds for the complements of K^r and K^s and for the boundaries ∂K^r and ∂K^s that are topological $(n-1)$-dimensional manifolds.

Topological stability properties of compact sets with positive wfs

Once we know that the topology of the offsets of a given compact can only change at critical values, it is appealing to compare the topology of two close compact sets with positive weak feature size. Using the flow of the gradient of the distance function, it appears that the homotopy types of two such compacts are the same. Recall that given two topological spaces X and Y, two maps $f : X \to Y$ and $g : X \to Y$ are said *homotopic* if there is a continuous map $H : [0,1] \times X \to Y$, such that $\forall x \in X, H(0,x) = f(x)$ and $H(1,x) = g(x)$. X and Y are said *homotopy equivalent* if there are continuous maps $f : X \to Y$ and $g : Y \to X$ such that $g \circ f$ is homotopic to the identity map of X and $f \circ g$ is homotopic to the identity map of Y. An equivalence class for homotopy equivalence is called a *homotopy type*. If two spaces X and Y are homeomorphic then they are homotopy equivalent. In general, the converse is not true. However, homotopy equivalence between topological spaces implies a one-to-one correspondance between connected components, cycles, holes, tunnels, cavities, or higher-dimensional topological features of the two sets. More precisely, if X and Y have same homotopy type, then their homotopy and homology groups are isomorphic.

Theorem 2 ([11]). *Let K and K' be compact subsets of $\mathbb{R}^n$ and ε such that $\text{wfs}(K) > 2\varepsilon$, $\text{wfs}(K') > 2\varepsilon$ and $d_H(K, K') < \varepsilon$. One has:*

(i) *$\mathbb{R}^n \setminus K$ and $\mathbb{R}^n \setminus K'$ have the same homotopy type.*

[5] Roughly speaking, two subspaces of $\mathbb{R}^n$ are *isotopic* if they can be deformed one into each other without tearing or self-intersection. For example, a circle and a trefoil knot are homeomorphic but not isotopic.

(ii) If $0 < \alpha \leq 2\varepsilon$ then K^α and K'^α have the same homotopy type.

Note that this theorem does not relate the homotopy types of K and K'. Indeed it is possible to construct an example of two non homotopy equivalent compact sets K and K' satisfying the hypothesis of the theorem. Such a construction, out of the scope of this paper, is not provided here but can be easily derived from the example in Section 4.3 of [11].

Theorem 2 provides us with a first stability result but, from a practical point of view, it presents a strong drawback: it may happen that, even if $\mathrm{wfs}\,(K)$ is big enough, the quantity $\mathrm{wfs}\,(K')$ goes to 0 as $d_H(K, K') \to 0$ in such a way that the hypothesis of the theorem is never fulfilled. In particular, this usually happens when K' is a finite set of point approximating K. For example, if $K \subset \mathbb{R}$ is the unit length segment $[0, 1]$ approximated by a finite set of n equally spaced points K', then $d_H(K, K') = 1/(2n-2) = \mathrm{wfs}\,(K')$. It is thus necessary to improve and to extend Theorem 2. One way to extend it is to try to understand what topological properties of K can be recovered from K' when we only assume that K has a positive weak feature size. This approach has been adopted in [11, 15] where it is proven that most of the topological invariants of a compact set K with positive wfs can be recovered from the knowledge of an approximation K' such that $d_H(K, K') < \frac{\mathrm{wfs}\,(K)}{4}$. In particular, when K' is a finite set of points, a simple algorithm allows to recover the Betti numbers of K. This approach is developed in Section 3. The interested reader may refer to [11, 15] for further details. Another way to improve Theorem 2 is to analyze the behavior of the critical points of the distance functions under perturbations of the compact sets. This approach leads to surprisingly strong stability results that are presented in the next section.

2.2 Sampling theory

It appears that the critical points of the distance function to a given compact set $K \subset \mathbb{R}^n$ are not stable under perturbations of K as shown on Figure 5, so that there is no hope to approximate the critical values (and thus the wfs) of a compact set K from an approximation K'. To overcome this unstability problem we introduce the following "parametrized" notion of critical point. Given $0 \leq \mu \leq 1$, a *μ-critical point* of the compact set K is a point at which the norm of the gradient ∇_K does not exceed μ. Notice that the 0-critical points are exactly the critical point of d_K. Unlike the 0-critical points alone, the family of μ-critical points satisfy a stability property.

Theorem 3 (critical point stability theorem [8]). *Let K and K' be two compact subsets of $\mathbb{R}^n$ and $d_H(K, K') \leq \varepsilon$. For any μ-critical point x of K, there is a $(2\sqrt{\varepsilon/d_K(x)} + \mu)$-critical point of K' at distance at most $2\sqrt{\varepsilon d_K(x)}$ from x.*

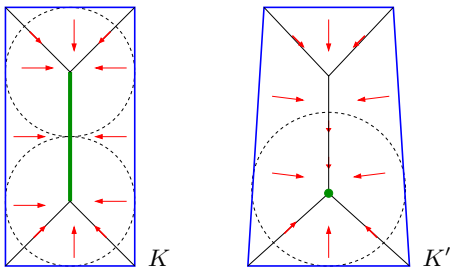

Fig. 5. When K is a rectangle, there is a segment of critical points of d_K along one of the axes of symmetry of K (the bold segment). This segment collapses to one point as soon as one stretches the bottom side of K to obtain K'. Nevertheless, along the previously critical segment, the norm of the gradient of $d_{K'}$ remains small.

Roughly speaking, this theorem states that in a neighborhood of size $O(\sqrt{\varepsilon})$ of a given μ-critical point for K, there is a $(\mu + O(\sqrt{\varepsilon}))$-critical point of K' for any ε-approximation K' of K. In particular, any ε-approximation K' of K has (at least) one $O(\sqrt{\varepsilon})$-critical point in a $O(\sqrt{\varepsilon})$-neighborhood of each critical point of K.

The stability of μ-critical points allows to introduce a general framework for inferring the topology and the geometry of a large class of non-smooth shapes. First, it is possible to "encode" all the μ-critical values of d_K in a real-valued one variable function. Given a compact set $K \subset \mathbb{R}^n$, its *critical function* $\chi_K : (0, +\infty) \to \mathbb{R}_+$ is the real function defined by:

$$\chi_K(d) = \inf_{d_K^{-1}(d)} ||\nabla_K||$$

Notice that from Lemma 1, the zeros of the critical function correspond to the changes in the topology of the offsets of K. As we will see later, whether a compact set is a Hausdorff approximation of a "simple" compact set or not can be directly read from its critical function, which is the main motivation for introducing this concept. Another way to interpret the critical point stability theorem is to notice that if the critical function of K' is greater than some value μ' on a sufficiently large interval of length $O(\sqrt{\varepsilon})$, then the critical function of K in the "middle" of this interval cannot be smaller than some value $\mu = \mu' - O(\sqrt{\varepsilon})$. More precisely, an easy computation leads to the following statement.

Theorem 4 (critical function stability theorem [8]). *Let K and K' be two compact subsets of $\mathbb{R}^n$ and $d_H(K, K') \leq \varepsilon$. For all $d \geq 0$, we have:*

$$\inf\{\chi_{K'}(u) \,|\, u \in I(d, \varepsilon)\} \leq \chi_K(d) + 2\sqrt{\frac{\varepsilon}{d}}$$

where $I(d, \varepsilon) = [d - \varepsilon, d + 2\chi_K(d)\sqrt{\varepsilon d} + 3\varepsilon]$

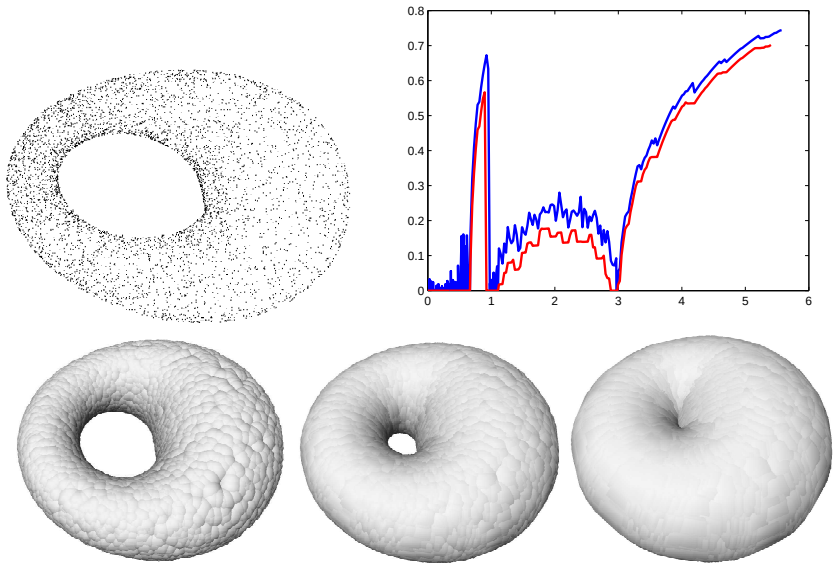

Fig. 6. A 4000 points set (left) sampled around a torus shape in $\mathbb{R}^3$ (which is not a torus of revolution) and its critical function (the upper curve). The lowest curve represents the lower bound for the critical function of any shape K at distance less than some fixed threshold (here 0.001, the diameter of the torus being 10) from the point cloud. We distinguish three intervals with stable topology for K: the first one corresponds to offsets having the topology of a torus (bottom left), the second one corresponds to solid torus with a hole homeomorphic to a ball inside (bottom middle - not visible from outside) and the third one is unbounded and correspond to offsets that have the topology of a ball (bottom right).

Knowing only the critical function of K' and its Hausdorff distance to K, it is thus possible to locate intervals on which the critical function of K does not vanish and the topology of the corresponding offsets of K does not change. Figure 6 illustrates this in the case of a point cloud sampled around a torus shape in $\mathbb{R}^3$. Although we are able to detect, from the critical function, common intervals of stable topology for all the shapes K located at some given distance from K', it remains to relate the topology of the offsets of K' to the ones of K. Fortunately, it happens that if the length of an interval where the critical function of K' does not vanish is sufficiently large with respect to the infimum of $\chi_{K'}$ then the offsets (in this interval) of K and the offsets of K' are homemorphic and even isotopic.

Theorem 5 (Level sets isotopy theorem [9]). *Let $K, K' \subset \mathbb{R}^n$ be two compact sets such that $d_H(K, K') < \varepsilon$ for some $\varepsilon > 0$. If $a > 0$ is such that $\chi_{K'} > 2\sqrt{\frac{2\varepsilon}{a-\varepsilon}}$ on the interval $[a - \varepsilon - 2\sqrt{2\varepsilon(a+\varepsilon)}, a + \varepsilon + 2\sqrt{2\varepsilon(a+\varepsilon)}]$*

then $d_K^{-1}(a)$ *and* $d_{K'}^{-1}(a)$ *are isotopic hypersurfaces and* K^a *and* K'^a *are also isotopic.*

This result has first appeared in a weaker form in [8] where only the homotopy equivalence between the offsets was proven. This weaker form follows directly from the critical point stability theorem and Theorem 2. The isotopy has been proven in [9] and needs a more detailed study of the stability properties of the gradient vector field ∇_K. The isotopy is then obtained by "pushing" the offsets of K' onto the offsets of K using the flow of the vector field ∇_K.

It is important to notice that the previous theorem does not require any knowledge on K (except an upperbound on its distance to K'). This is particularly useful in practical applications where the approximation K' is usually the only information we know about K. In particular, from K' we are able to decide if there exists at some given distance ε some shape with given "intervals of topological stability" for its offsets. From this it is then possible to exhibit sampling conditions insuring that the topology of the offsets of K can be reliably recovered from the offsets of K'. They involve the so-called μ-reach of the considered shape K which is the length of the longest interval starting from 0 on which χ_K is greater than μ. Equivalently, given $\mu \in (0, 1]$, the μ-reach $r_\mu(K)$ of K is the infimum of the values d such that $\chi_K(d) < \mu$. Note that $\mu \to r_\mu(K)$ is a non increasing fonction. For $\mu = 1$, the 1-reach is known as the *reach* and has been introduced by H. Federer [24] in the setting of Geometric Measure Theory. It is well-known that the reach of smooth submanifolds [6] is positive. It has been widely used in smooth surface reconstruction, where it is also known as the minimum feature size, to provide sampling conditions insuring geometrically correct reconstruction (see [1] for example). These conditions require the sampling density to be proportional to the value of the reach. Non-smooth shapes usually having their reach equal to zero, such conditions cannot be fulfilled. So, by introducing the μ-reach it is possible to introduce new sampling conditions that remain relevant for a large class of non-smooth shapes (including smooth shapes, piecewise linear shapes and more generally any small offsets of shapes with positive wfs). Given two non-negative real numbers κ and μ, we say that $K' \subset \mathbb{R}^n$ is a (κ, μ)-*approximation* of $K \subset \mathbb{R}^n$ if the Hausdorff distance between K and K' does not exceed κ times the μ-reach of K [8]. Combining the level sets isotopy theorem with this these sampling conditions leads to the following reconstruction theorem:

Theorem 6 (Isotopic reconstruction theorem). *Let* $K \subset \mathbb{R}^n$ *be a compact set such that* $r_\mu(K) > 0$ *for some* $\mu > 0$. *Let* K' *be a* (κ, μ)*-approximation of* K *where*

$$\kappa < \min\left(\frac{4\sqrt{2}-5}{14}, \frac{\mu^2}{16+2\mu^2}\right)$$

[6] i.e. differentiable submanifolds of class at least $\mathcal{C}^{1,1}$

and let d, d' *be such that* $0 < d < \mathrm{wfs}(K)$ and $\frac{4\kappa r_\mu(K)}{\mu^2} \leq d' < (1-3\kappa) r_\mu(K)$ *Then the level set* $R_{K'}^{-1}(d')$ *is isotopic to the level set* $R_K^{-1}(d)$ *(the same holds for* $K'^{d'}$ *and* K^d*).*

Note that this theorem, unlike the previous one, provides an explicit interval of values close to 0 for which the level set isotopy theorem can be applied. Also, it can be shown that homotopy equivalence of the offsets holds under slightly weaker constraints on κ and μ [8, 9].

2.3 Stable medial axes

Closely related to distance functions is the so-called *medial axis* which appears to be the set of points where d_K is not differentiable. This subset of $\mathbb{R}^n$ finds applications in various domains such as image analysis and mathematical morphology, Solid Modeling, or domain decomposition for CAD (computer aided design) model generation. Intuitively, the medial axis of a shape K is the locus where a wave front starting from K and propagating at constant speed self-intersects. Mathematically, the *medial axis* $\mathcal{M}(K)$ of a compact subset K of $\mathbb{R}^n$ is defined as the set of points x that have at least two nearest neighbors on K (see Figure 7). Equivalently the medial axis can also be defined as the set of points where the gradient of the distance function to K has norm smaller than 1

$$\mathcal{M}(K) = \{x \in \mathbb{R}^n : \|\nabla_K(x)\| < 1\}.$$

It is classically known that $\mathcal{M}(K)$ is exactly the set of points where the distance function to K is not differentiable. The medial axis of a finite set of points is the well-known Voronoi diagram (see Figure 7).

From a topological point of view, the main interest of the medial axis is that it carries topological informations about the topology of the shape. Indeed, using the gradient of the distance function, it has been proven in [31] that the medial axis of a compact shape is homotopy equivalent to its complement [7].

Unfortunately but not surprisingly, the medial axis of a given compact K appears to be very unstable under perturbation or approximation of K as illustrated on Figure 8 where a small perturbation of a circle causes the apparition of a long spike in the medial axis. Avoiding such huge variations of the medial axis under perturbation requires to restrict to perturbations that are not only small in Hausdorff distance but that also have small first and second order derivatives (see [12] for precise statements and details). Unfortunately such restrictions usually do not comply with the constraints of practical applications where Hausdorff closeness is often the only realistic assumption that can be made. Computing and approximating the medial axis of a shape

[7] strictly speaking, for compactness reasons, this is in fact the medial axis of the union of K with a large sphere enclosing it which is homotopy equivalent to the complement of this union (restricted to the ball bounded by the sphere).

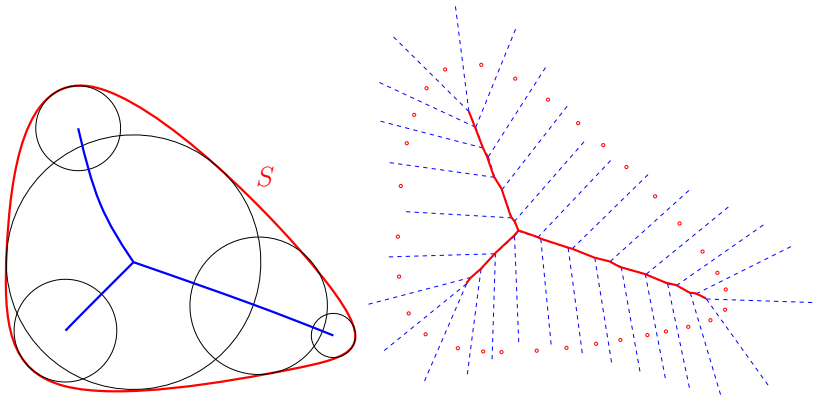

Fig. 7. The medial axis of a closed curve in the plane and of a finite set of points (Voronoi diagram). On the right, the bold part of the Voronoi diagram is a λ-medial axis of the point cloud.

has been a long standing problem that has been adressed in different settings (see [3] for a recent survey). Some of them based upon an exact computation paradigm ensure correctness of the approximation but in practice are subject to efficiency and robustness problems due to the need of an exact knowledge of the input shape [17, 23]. Others are based on heuristics that do not allow to certify the quality of the approximation [1, 20]. More recently, the distance function approach has proposed a solution to overcome the unstability of the medial axis.

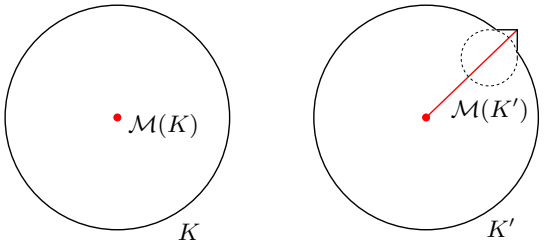

Fig. 8. Unstability of the medial axis.

The λ-medial axis

The idea is to replace the medial axis by one of its subsets that could be proven to be stable. Indeed, looking at the example of Figure 8, one can notice that

along the spike, the function F_K [8] remains very small, its magnitude being approximatively equal to the "size" of the perturbation. It is thus appealing to remove the part of the medial axis where the function F_K is smaller than some user-defined threshold. This has been done in [10] by defining the λ-*medial axis* of a compact K as the set of points where F_K is not smaller than λ:

$$\mathcal{M}_\lambda(K) = \{x \in \mathbb{R}^n : F_K(x) \geq \lambda\}$$

From a topological point of view, the λ-medial axis still carries the topology of the medial axis: if K is a compact set with positive weak feature size then for any $\lambda < \text{wfs}(K)$, $\mathcal{M}_\lambda(K)$ and $\mathcal{M}(K)$ are homotopy equivalent [10]. From a computational point of view, as in the case of μ-critical points, introducing a parametrized notion of medial axis leads to stability properties for the λ-medial axes: if K and K' are two close shapes, then $\mathcal{M}_\lambda(K)$ and $\mathcal{M}_{\lambda'}(K')$ are close to each other for some close values of λ and λ'. The precise statement of this stability result being rather technical the reader is referred to [10] for details. Its proof makes an important use of the properties of the gradient of the distance function. An important consequence of the stability of the λ-medial axis is to make possible to reliably approximate the λ-medial axis of a shape from an approximation by a finite set of points. Noticing that when K is a point cloud the function F_K is constant on the Voronoi facets, one easily deduces an explicit algorithm to compute the λ-medial axes of K from its Voronoi diagram (see Figure 9). Indeed, one only has to compute for each Voronoi facet the radius of the smallest ball containing its Delaunay dual simplex. In particular, the complexity of the computation of all the λ-medial axes of a point cloud is the same as the one of the Voronoi diagram. Nevertheless notice that when λ is not too small, the size of the λ-medial axes may be much smaller than the size of the Voronoi diagram. It is an open question of practical importance to know, given a point cloud and one value $\lambda > 0$, if it is possible to compute the λ-medial axis without computing the whole Voronoi diagram.

The μ-medial axis

To conclude this section, let us quickly mention another construction that leads to stable medial axes. Another way to "filter" the medial axis of a shape K is to remove the parts where the norm of the gradient ∇_K is bigger than some threshold $\mu > 0$ defining a μ-medial axis of K:

$$\mathcal{M}_\mu(K) = \{x \in \mathbb{R}^n : \|\nabla_K(x)\| \leq \mu\}.$$

One thus obtain another filtration of the medial axis which is different from the λ-medial axis. Using the critical point stability theorem for distance functions

[8] recall that for any x, $F_K(x)$ is defined as the radius of the smallest ball containing $\Gamma_K(x)$, the set of points of K nearest to x.

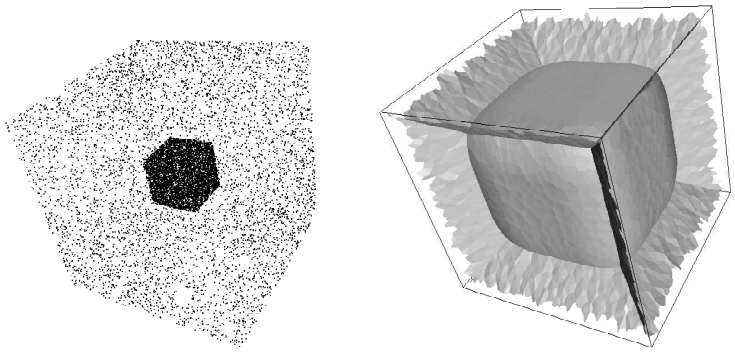

Fig. 9. The λ-medial axis of a set of points sampled around two enclosed cubes.

it is still possible to prove stability properties of this filtration [8]. Nevertheless, the topology of the μ-medial axis is not clearly related to the one of the medial axis. From a computational point of view, the computation of $\mathcal{M}_\mu(K)$ is slightly more awkward than for the computation of the λ-medial axis when K is a point cloud, since $\|\nabla_K\|$ is not constant on the cells of the Voronoi diagram of K.

3 Persistent homology

The previous section showed that it is possible to accurately reconstruct fairly general shapes from samples: under a certain sampling condition, an appropriate offset of the samples provides a topologically correct, and geometrically close approximation of the unknown shape. Now suppose that we are not interested in actually *reconstructing* a topologically correct approximation of the shape, but just in *estimating* its topological invariants, knowing only the sample points. Obviously, a valid approach for solving this problem would be to build a topologically accurate reconstruction and compute its topological invariants. The purpose of this section is to show that for this easier problem, a better solution may actually be found, using the concept of *persistent homology*. This concept, introduced independently by several groups [22, 33, 25], is a rather general tool, which can be applied to other problems than topological inference from samples.

Topological noise

In many practical situations, the object of interest, K, is defined as a particular sublevel set of some "ideal" real function g defined over some topological space. That is, we have $K = g^{-1}(-\infty, x]$ for some number x, where g may be a grey level in a two or three-dimensional image, or a density estimate for

example. Now, if the data is incomplete or corrupted by noise, we may only have access to a noisy approximation f of g. A natural question is then whether the topology of the object of interest K can be recovered knowing only the noisy implicit function f. Perhaps the simplest idea would be to estimate the topology of K by the one of the corresponding sublevel set $f^{-1}(-\infty, x]$ of f. Unfortunately, this naive approach tends to produce *topological noise*, which takes the form of spurious topological features, such as additional connected components as in Figure 10, or additional loops and tunnels as in Figure 11. Persistent homology is the natural way of dealing with this problem.

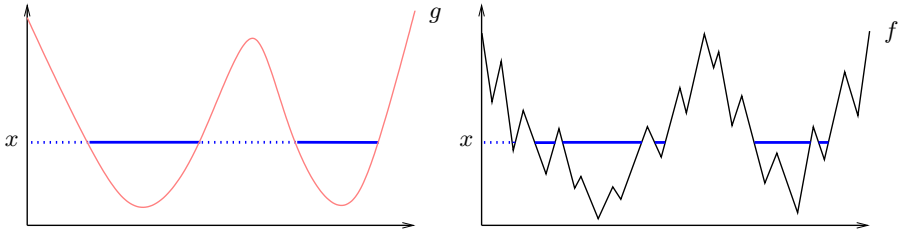

Fig. 10. The x sublevel set of g (left) has two connected components, but the one of its approximation f has four additional components (right).

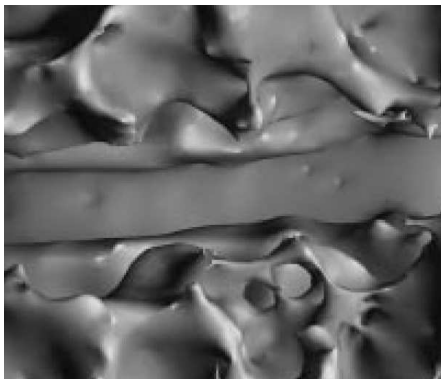

Fig. 11. Estimated interface between grey matter and white matter in the brain contains spurious loops.

3.1 Homology

Homology (see [30] for instance) is a mathematical formulation of the intuitive notions of connected components, loops and voids mentioned in the previous paragraph. Though it can be defined for general topological spaces, we focus in what follows on the simpler case of simplicial complexes, which is also the most relevant from a computational point of view.

Complexes, chains and boundaries

A simplicial complex is a space made out of vertices, edges, triangles, tetrahedra, and their higher-dimensional counterparts. More formally, a *k-simplex* σ in $\mathbb{R}^d$ is the convex hull of $k+1$ points of $\mathbb{R}^d$ in general position, which are called the *vertices* of σ. A *face* of simplex σ is a simplex whose vertices form a subset of the vertices of σ. A *simplicial complex* $\mathbb{X}$ is a union of simplices such that any two simplices intersect along a common face, or are disjoint. The simplices $\mathbb{X}$ is made of, together with their faces, will be referred to as the simplices of $\mathbb{X}$.

A k-chain of a simplicial complex $\mathbb{X}$ is an assignment of either 0 or 1 to each k-simplex of $\mathbb{X}$. It is convenient to think of a chain as the collection of k-simplices with assignment 1. The *sum* of two chains is the chain defined by the modulo 2 sum of the two corresponding assignments. Equivalently, it is the symmetric difference of the two chains. This operation endows the set $C_k(\mathbb{X})$ of k-chains with a structure of a vector space over the field with two elements, $\mathbb{Z}_2$.

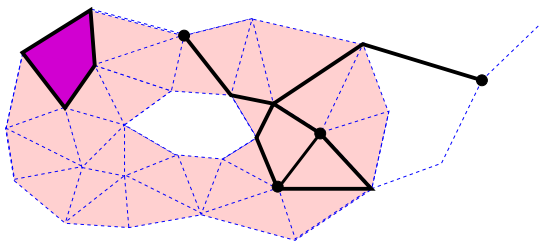

Fig. 12. A 1-chain (in bold on the right) and a 2-chain (the dark quadrilateral on the left) and their respective boundaries (the bold vertices on the right and the edges bounding the quadrilateral on the left).

The boundary operator $\partial_k : C_k(\mathbb{X}) \to C_{k-1}(\mathbb{X})$ is the (unique) linear map that maps each k-simplex (viewed as a chain consisting of a single simplex) to the sum of its $(k-1)$-dimensional faces. Equivalently, the boundary of a k-chain consists of the $(k-1)$-simplices incident to an odd number of k-simplices in the chain (see Figure 12). The kernel of operator ∂_k, that is, the set of chains with an empty boundary, is traditionally denoted by $Z_k(\mathbb{X})$. It

is a subspace of $C_k(\mathbb{X})$, whose elements are called *cycles*. The image of ∂_{k+1}, that is, the set of *boundaries*, is also a subspace of $C_k(\mathbb{X})$, denoted by $B_k(\mathbb{X})$.

Homology groups

It can be easily checked that $B_k(\mathbb{X}) \subseteq Z_k(\mathbb{X})$, meaning that boundaries have empty boundary. Indeed, the boundary of a single simplex clearly has an empty boundary, and the general case follows by linearity. Now homology measures by how much the converse inclusion fails to hold. Formally, the k^{th} *homology group* of $\mathbb{X}$ is defined as the quotient $Z_k(\mathbb{X})/B_k(\mathbb{X})$, and denoted by $H_k(\mathbb{X})$. Elements of $H_k(\mathbb{X})$ are called k-dimensional *homology classes*. The dimension of $H_k(\mathbb{X})$, denoted by $\beta_k(\mathbb{X})$, is called the k^{th} Betti number of $\mathbb{X}$. Intuitively, Betti numbers count the number of k-dimensional topological features in K. For example, $\beta_0(\mathbb{X})$ is the number of connected components of $\mathbb{X}$. Indeed, any two vertices a and b lying in the same component can be joined by a 1-chain c, implying that $a - b = \partial_1(c)$, hence a and b represent the same homology class in $H_0(\mathbb{X}) = Z_0(\mathbb{X})/B_0(\mathbb{X})$. On the other hand, no such construction is possible for vertices lying in different components. Hence vertices belonging to different components are linearly independent. Choosing one vertex in each component thus gives a basis of $H_0(\mathbb{X})$.

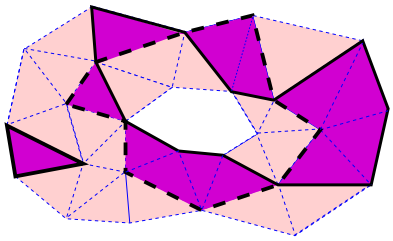

Fig. 13. A simplicial complex with first Betti number equal to 1.

Similarly, $\beta_1(\mathbb{X})$ counts the number of independent loops in $\mathbb{X}$. For example in Figure 13, 1-cycles that do not go around the hole (such as the bold one on the left) are the boundaries of some 2-chain, hence correspond to the zero homology class. On the other hand, any 1-cycle that goes around the hole is not the boundary of any 2-chain, hence corresponds to a non-zero homology class. However, any two such 1-cycles (such as the bold and dashed-bold ones on the right of figure 13) differ by the boundary of some 2-chain, hence correspond to the same homology class. We thus have exactly one non trivial homology class, hence a first Betti number of 1. For a space with several loops, we would get a larger homology group, with one basis element for each loop. For example, the six edges of a tetrahedron form a simplicial complex with first Betti number equal to 3. Indeed, the space of 1-cycles is spanned by the four cycles corresponding to each face of the tetrahedron, however the sum

of these four cycles cancels, implying that there are only three independent classes. Higher-dimensional Betti numbers behave pretty much the same way: $\beta_2(\mathbb{X})$ counts the number of voids, and so on, though common language lacks words for describing higher-dimensional situations.

3.2 Persistence

Topological persistence studies the evolution of the homology of an increasing sequence of spaces. More formally, a *filtration* of a simplicial complex $\mathbb{X}$ is a nested sequence of subcomplexes of $\mathbb{X}$, $\emptyset = \mathbb{X}_0 \subset \mathbb{X}_1 \subset \cdots \subset \mathbb{X}_n = \mathbb{X}$, such that $\mathbb{X}_i$ differs from $\mathbb{X}_{i-1}$ by a single simplex: $\mathbb{X}_i = \mathbb{X}_{i-1} \cup \sigma_i$. An important concept for studying the evolution of homology groups in a filtration is the one of *inclusion maps*, which relate homology groups of different levels in the filtration. Let α be a homology class in $H_k(\mathbb{X}_i)$, represented by some k-cycle a. Cycle a is obviously also a cycle in subsequent levels of the filtration. Given $j \geq i$, the *inclusion map* from $H_k(\mathbb{X}_i)$ to $H_k(\mathbb{X}_j)$ is defined as the (linear) map that sends $\alpha \in H_k(\mathbb{X}_i)$ to the homology class of $H_k(\mathbb{X}_j)$ represented by cycle a. This map is well-defined because $B_k(\mathbb{X}_i)$ is included in $B_k(\mathbb{X}_j)$. Inclusion maps allow to associate to each filtration $\emptyset = \mathbb{X}_0 \subset \mathbb{X}_1 \subset \cdots \subset \mathbb{X}_n = \mathbb{X}$ a *directed system of vector spaces* [6]:

$$\{0\} = H_k(\mathbb{X}_0) \longrightarrow \ldots \longrightarrow H_k(\mathbb{X}_i) \longrightarrow H_k(\mathbb{X}_{i+1}) \longrightarrow \ldots \longrightarrow H_k(\mathbb{X}_i) = H_k(\mathbb{X})$$

which is an algebraic counterpart consisting of the homology groups of different levels of the filtration, joined by inclusion maps. We now describe inclusion maps between consecutive levels.

Birth and death

In a filtration, adding simplex σ_i to $\mathbb{X}_{i-1}$ can have two different effects at the homology level. Let k be the dimension of σ_i. The boundary $\partial_k \sigma_i$ is always a cycle in $\mathbb{X}_{i-1}$. The two cases depend on whether it is also a boundary or not.

If $\partial_k \sigma_i$ is not a boundary in $\mathbb{X}_{i-1}$, then by definition $\partial_k \sigma_i$ corresponds to a non-zero homology class in $H_{k-1}(\mathbb{X}_{i-1})$. However, considered as a chain in $\mathbb{X}_i$, it is clearly a boundary (namely the one of σ_i), hence it corresponds to the zero homology class in $H_{k-1}(\mathbb{X}_i)$. As a consequence, the result of adding simplex σ_i is to destroy the homology class of $H_{k-1}(\mathbb{X}_{i-1})$ represented by $\partial_k \sigma_i$. Simplex σ_i is then said to be a *destructor* (see Figure 14, left). From an algebraic point of view, we have that the inclusion map $H_{k-1}(\mathbb{X}_{i-1}) \to H_{k-1}(\mathbb{X}_i)$ is surjective, and has a 1-dimensional kernel, spanned by the homology class represented by $\partial_k \sigma_i$. The inclusion maps $H_l(\mathbb{X}_{i-1}) \to H_l(\mathbb{X}_i)$ for $l \neq k-1$ all are isomorphisms, since the adjunction of σ_i does not have any impact on homology groups other than the $(k-1)^{th}$.

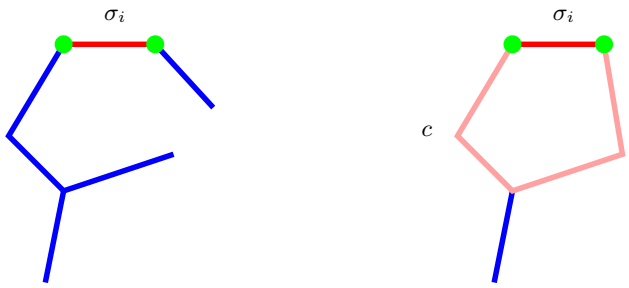

Fig. 14. Edge σ_i is a destructor on the left, and a creator on the right.

Now if $\partial_k \sigma_i$ is a boundary in $\mathbb{X}_{i-1}$, for example $\partial_k \sigma_i = \partial_k c$, where c is a chain in $\mathbb{X}_{i-1}$, the effect of adding σ_i is to create a new homology class in $H_k(\mathbb{X}_i)$. Indeed, the chain $c - \sigma_i$ is a cycle by definition of c. Now c being the last simplex added, it has no coface, implying that the homology class of $c - \sigma_i$ cannot be represented by cycles not containing σ_i. That is, the class of $c - \sigma_i$ does not belong to the image of the inclusion map $H_k(\mathbb{X}_{i-1}) \to H_k(\mathbb{X}_i)$. In this case, we say that σ_i is a *creator* (see figure 14, right). This time, the algebraic situation is that the inclusion map $H_k(\mathbb{X}_{i-1}) \to H_k(\mathbb{X}_i)$ is injective, and its image has codimension 1 in $H_k(\mathbb{X}_i)$, the quotient space $H_k(\mathbb{X}_i)/\mathrm{im}(H_k(\mathbb{X}_{i-1}) \to H_k(\mathbb{X}_i))$ being spanned by the class of $c - \sigma_i$. Again, inclusion maps between other homology groups are isomorphisms.

Pairing

The idea of persistence is to pair destructors with creators. The precise rule is as follows. Let σ_j be a k-dimensional destructor. We know that σ_j destroys the homology class of $\partial_k \sigma_j$ in $H_{k-1}(\mathbb{X}_{i-1})$, which we call α. Now we pair j with the first level in the filtration where α was created. Formally, this is the smallest integer i such that α is in the image of the inclusion map from $H_{k-1}(\mathbb{X}_i)$ to $H_{k-1}(\mathbb{X}_j)$, or, equivalently, such that α can be represented by a cycle in $\mathbb{X}_i$. We will sometimes abuse the terminology and say that the *simplices* σ_i and σ_j are paired. The interval $[i, j]$ is called a k-dimensional *persistence interval*. Intuitively, it represents the life span in the filtration of the k-dimensional feature created by σ_i and destroyed by σ_j. Note that some simplices may remain unpaired by this process. The indices corresponding to such simplices are paired with infinity by convention.

Figure 15 shows an example. First, vertices a, c and b are successively added in the filtration, which creates one new component each time. Then edge ab is added, which destroys the 0-cycle $a - b$. We thus pair edge ab with vertex b, since $a - b$ first appears when b is added. Next we add edge bc, which destroys $b - c$. Now, the homology class of $b - c$ is also represented by the cycle $a - c$, since the difference between the two equals the boundary of the already added edge ab. Hence, the class of $b - c$ was actually born when c was

added, so we pair bc with c. Then, vertex d creates a new component, which is destroyed by edge cd, so we pair cd with d. Finally, the only unpaired creator, a, leads to a half-infinite interval.

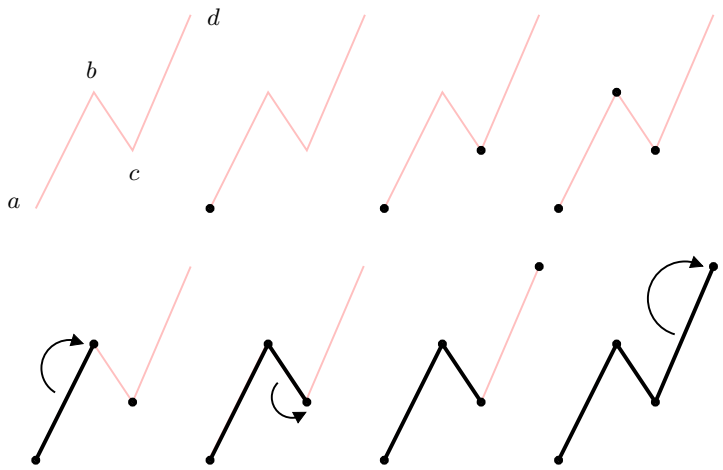

Fig. 15. A filtration of a complex and the associated pairing.

It turns out that the persistence pairing can be computed by a very elegant algorithm in the general case. The first version of this algorithm appears in [22], and was later generalized to arbitrary coefficient fields [6]. The version we present is described in [16]. We define the *depth* of a column vector as the index of the non-zero coefficient of column j that is closest to the bottom of the column. If the column is identically zero, its depth is undefined.

1. Form the $n \times n$ matrix D with $\mathbb{Z}_2$ coefficients defined by $D_{ij} = 1$ if and only if σ_i is a codimension 1 face of σ_j.

2. As long as two columns in D have the same depth, substract the leftmost to the rightmost. Call R the matrix obtained when the reduction process stops.

3. For all j, pair j with the depth of the j^{th} column of R, whenever this column is non-zero.

This algorithm takes cubic time in the number of simplices in the worst case. However, sparse matrix implementations prove to be very effective in practice. The reason why this algorithm works is as follows. First, note that for all j, the space spanned the first j columns of the matrix remains invariant during the reduction process, since we always substract columns from subsequent columns in the filtration. Hence, if the j^{th} column of R is zero, then $\partial\sigma_j$ must be a linear combination of boundaries of simplices in $\mathbb{X}_{j-1}$. That

is, σ_j is a creator. Assume now that the j^{th} column of R is non-zero and let τ_j be the cycle it represents. Calling i the depth of this column, we see that $\partial\sigma_j$ lies in $\tau_j + B_k(\mathbb{X}_{j-1}) \subset C_k(\mathbb{X}_i) + B_k(\mathbb{X}_{j-1})$. We claim that for $i' < i$, $\partial\sigma_j$ cannot lie in $C_k(\mathbb{X}_{i'}) + B_k(\mathbb{X}_{j-1})$. Indeed, if it were the case, τ_j would lie in $C_k(\mathbb{X}_{i'}) + B_k(\mathbb{X}_{j-1})$. This would imply that $C_k(\mathbb{X}_{i'})$ contains the sum of τ_j and a linear combination of $\tau_{j'}$ for $j' < j$. But since no two columns in R share the same depth, the depth of a linear combination of columns of R is the maximum depth of the columns involved in the combination. This shows that the latter sum has depth at least i, hence it cannot belong to $C_k(\mathbb{X}_{i'})$, a contradiction. As a consequence, the class of $\partial\sigma_j$ in $H_k(\mathbb{X}_{j-1})$ may be represented by cycles in $\mathbb{X}_i$ but not by cycles in $\mathbb{X}_{i'}$ for $i' < i$. Hence, σ_j is a destructor and is paired with i.

Persistent Betti numbers

It was shown in [6] that persistence intervals give a complete description of the directed systems of vector spaces associated with a filtration. More precisely, two filtrations have isomorphic directed systems of vector spaces if and only if they have the same persistence intervals. In particular, persistence intervals completely encode an important set of invariants of such systems, called *persistent Betti numbers*. Given two indices i and j in the filtration $0 \leq i \leq j \leq n$, the k^{th} persistent Betti number $\beta_k^{i,j}$ is defined as the rank of the inclusion map $H_k(\mathbb{X}_i) \to H_k(\mathbb{X}_j)$. For $i = j$, persistent Betti numbers $\beta_k^{j,j}$ are just the usual Betti numbers of $\mathbb{X}_j$. For $i < j$, persistent Betti numbers only take into account the homology classes of $\mathbb{X}_j$ that can be represented by cycles in $\mathbb{X}_i$. The connection between persistent Betti numbers and persistence intervals is expressed by the following lemma, which follows easily from the previous discussions:

Theorem 7 (k-triangle lemma[22]). *The persistent Betti number $\beta_k^{i,j}$ equals the number of k-dimensional persistence intervals that contain interval $[i, j]$.*

The intuition behind this lemma is that $\beta_k^{i,j}$ is the number of k-dimensional features that were already born at stage i and not yet destroyed by stage j. It should be noted that conversely, persistence intervals can be deduced from persistent Betti numbers by an inclusion-exclusion formula, as explained in [15, 25]. Hence persistent Betti numbers also describe completely the directed systems of vector spaces associated with a filtration. However, they provide a less compact representation than persistence intervals, since there are n^2 persistent Betti numbers, and only $n/2$ persistence intervals, at most.

3.3 Stability of persistence and geometric inference

The main motivation for persistent homology is that it allows, loosely speaking, to distinguish signal from noise. The goal of this section is make this statement more precise.

Persistence diagrams

Let us first extend persistence from the simplicial framework to the continuous one [15]. Instead of considering filtrations of simplicial complexes, we now assume that $\mathbb{X}$ is a general topological space with the continuous filtration given by the sublevel sets of a given real function f defined on $\mathbb{X}$. We denote the x sublevel set $f^{-1}(-\infty, x]$ by $\mathbb{X}_x$. Such sublevel sets are not simplicial complexes in general, but the concepts of homology groups and inclusion maps can be generalized to handle such sets. This generalization, called *singular homology*, is technically different from the simplicial version described above, but the intuition is the same.

For extending persistence to the continuous case, we need some regularity assumptions on function f. We say that $x \in \mathbb{R}$ is an index k *homological critical value* of f if one can find arbitrarily small values of ε such that the inclusion map $H_k(\mathbb{X}_{x-\varepsilon}) \to H_k(\mathbb{X}_{x+\varepsilon})$ is not an isomorphism. Intuitively, homological critical values are threshold values at which the topology of sublevel sets change. Now function f is called *tame* if it has finitely many homological critical values, and if all its sublevel sets have finite dimensional homology groups. The class of tame functions includes for example piecewise-linear functions defined over simplicial complexes, as well as Morse functions over smooth manifolds. Filtrations of spaces by sublevel sets of tame functions behave pretty much like filtrations of simplicial complexes. One difference is that homological critical values may induce more complex changes in the homology groups: a single critical value may for example destroy several classes and create other classes at the same time. However, it is not difficult to see that there still is a unique set of persistence intervals such that the k-triangle lemma is satisfied. Because several classes may be created by the same value and also killed by the same (different) value, persistence intervals should be counted with a multiplicity possibly higher than 1 in this framework.

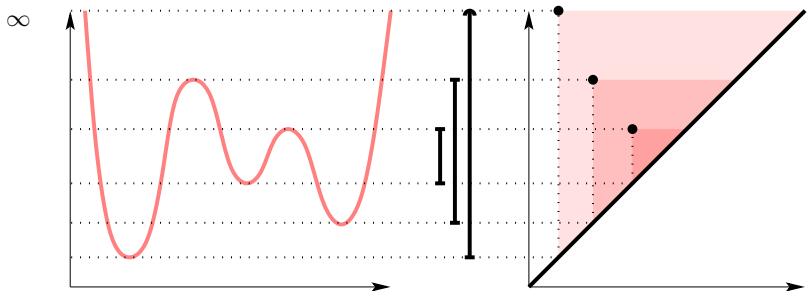

Fig. 16. A function on the line, its 0-dimensional persistence intervals and persistence diagram.

We now introduce a different representation of persistence intervals. To each k-dimensional persistence intervals $[x, y]$ of f, we associate the point with

coordinates (x, y) in the extended plane $[-\infty, +\infty]$, with the same multiplicity as the interval. Also, we include the diagonal $\{(x, x) \mid x \in [-\infty, +\infty]\}$ with infinite multiplicity, for reasons that will become clear in the next paragraph. An example of persistence diagram is shown in Figure 16. The shading in the persistence diagram indicates the values of persistence Betti numbers. Indeed, in this representation, the k-triangle lemma states that $\beta_k^{x,y}$ is the total multiplicity of the k-dimensional persistence diagram falling in the upper left quadrant with corner (x, y), hence we get 3, 2, and 1 as we go from the darkest shaded area to the lightest one.

Stability

Persistence diagrams are a compact encoding of the evolution of the topology of the sublevel sets of a function as the threshold increases. It turns out that they are also stable with respect to perturbation of the function, which is a key property in the context of unprecise data. To make this statement precise, we use the *bottleneck distance* between two multisets in the extended plane. The l_∞ distance between two points (x, y) and (x', y') in the extended plane is the maximum difference between their coordinates, that is, $\max(|x - x'|, |y - y'|)$. Now we say that two multisets have bottleneck distance at most d if there is a one-to-one matching between them such that paired points are at most d away in the l_∞ metric.

Theorem 8 (Stability of persistence diagrams [15, 2]). *Given two continuous tame functions f and g on a triangulable topological space, the bottleneck distance between their k-dimensional persistence diagrams does not exceed* $\sup |f - g|$.

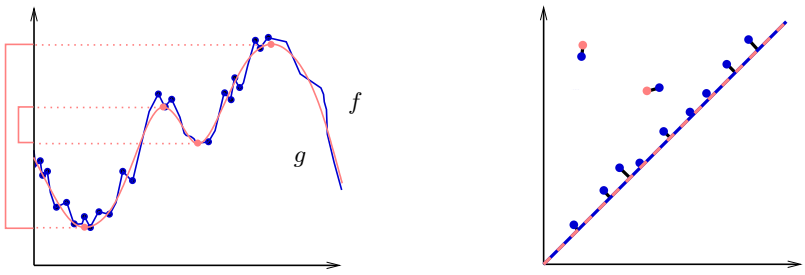

Fig. 17. A real function g, a noisy approximation f, and an optimal matching between their persistence diagrams.

Figure 17 gives an illustration of this theorem. The two points in the persistence diagram of function g are matched with nearby points also present in the diagram of the noisy approximation f. Additional spurious critical

values of f give rise to small intervals, hence points close to the diagonal, which are matched with it. An interpretation of this theorem is that points away from the diagonal (*i.e.* long intervals) correspond to *signal*, whereas points close to the diagonal (*i.e.* short intervals) correspond to *noise*, since they can be created or removed by small perturbations of the function.

One consequence of the stability of persistence diagrams is that Betti numbers of sublevel sets of function g can be inferred from those of a possibly noisy approximation f. The precise statement is as follows:

Theorem 9 (sublevel-set homology inference). *Let f and g two tame functions on a triangulable topological space $\mathbb{X}$, such that $\sup|f-g| < \varepsilon$. If $x, y \in \mathbb{R}$ are such that $|x-y| \geq 2\varepsilon$, and if $[x-\varepsilon, y+\varepsilon]$ contains no homological critical values of g, then with obvious notations:*

$$\beta_k^{x,y}(f) = \beta_k^{x,y}(g) = \beta_k^{z,z}(g)$$

for any $z \in [x, y]$.

The idea behind this theorem is schematically illustrated in Figure 18, which shows a noisy approximation f of an unknown function g with two basins. The persistent Betti number $\beta_0^{x,y}(f)$ is the number of points of the persistence diagram of f contained in the shaded quadrant. Because $|x-y| \geq 2\varepsilon$, the quadrant is at least ε away from the diagonal in the l_∞ metric, hence it cannot contain any spurious point of the persistence diagram. The assumption that $[x, y]$ is far away from the homological critical values of ideal function g implies that the ε-neighborhood of the boundary of the quadrant is void from points. Hence no point in the diagram of g can leave or enter the quadrant during the matching with the diagram of f, which implies that the total number of points within the quadrant is the same for the two functions, that is, 2 in this case.

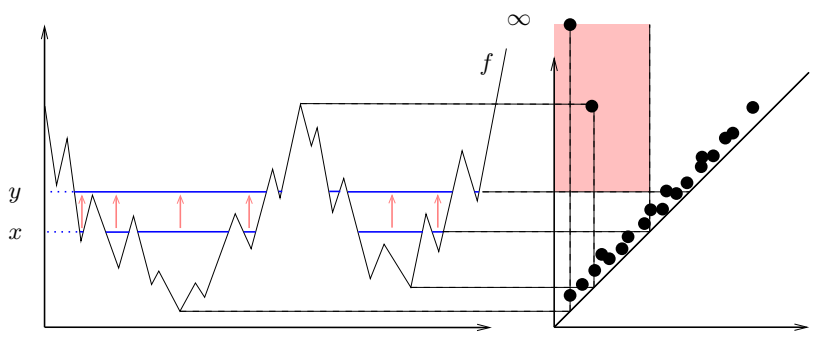

Fig. 18. Persistent Betti numbers allow to recover Betti numbers of the sublevel sets of an "ideal" unknown function.

Homology inference from samples

Let us go back to the problem we started this section with: we want to estimate the Betti numbers of a compact subset, knowing only a possibly noisy point sample. The above theorem directly gives a solution when applied to the distance functions of two nearby compact sets:

Theorem 10 (adapted from [11, 15, 4]). *Let K and K' be two compact subsets of $\mathbb{R}^n$ at Hausdorff distance at most ε. Then if $\varepsilon < \operatorname{wfs}(K)/4$, for all x such that $\varepsilon < x < \operatorname{wfs}(K)/4$ we have:*

$$\beta_k^{x,3x}(d_{K'}) = \beta_k(K^\eta)$$

for arbitrarily small $\eta > 0$.

Indeed we just have to notice that the maximum difference between the distance functions to K and K' is bounded by ε, and that the interval $(0, \operatorname{wfs}(K))$ does not contain any critical value of the distance function to K, by definition. One advantage of this result over the ones stated in the first part of this chapter is that makes much weaker assumptions about the unknown compact K. Indeed, the sampling condition only depends on the weak feature size, as opposed to the μ-reach. For example, Figure 19 shows a sampling of a triangle with a very small angle. Such a triangle has positive μ-reach only for a small value of μ, hence offset based reconstruction is guaranteed to work only for very dense samples, due to the $1/\mu^2$ dependency in the sampling condition (see Theorem 6). However, it is not difficult to find parameters which give a correct estimate of the first Betti number, namely 1. Indeed, there is only one loop in the larger offset that already exists in the smaller one, hence a correct persistent first Betti number. A major drawback of this approach, however, is that it does not say how to reconstruct a topologically correct approximation of the unknown shape from the samples.

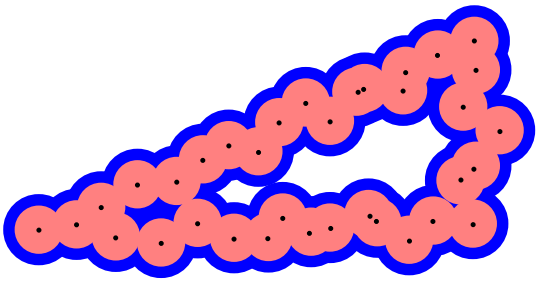

Fig. 19. Persistent Betti numbers allow correct homology inference even on difficult examples.

Persistence diagrams of the distance function to the set of sample points can be computed by applying the persistence algorithm to the α-shape filtration of the Delaunay triangulation of the sample points. However, the cost

of this approach becomes prohibitive in high dimensions, because the size of the Delaunay triangulation grows exponentially with the ambient dimension. An interesting alternative, the *witness complex*, was proposed in [34] to solve this problem. Using this technique, Carlsson et al were able to determine the topology of the set of 3×3 patches in natural images, after suitable normalization. It turns out that this space seems to have the topology of a Klein bottle, that is, a genus 1 non orientable surface. We refer to [5] for an interpretation of this surprising fact.

References

1. N. Amenta, S. Choi and R. Kolluri, *The Power Crust, unions of balls and the Medial Axis Transform* Computational:Theory and Applications, 2001 Vol.19:(2-3), pp.127-153.
2. M. d'Amico, P. Frosini, C. Landi, *Optimal Matching between Reduced Size Functions,* Tech. Report 35, DISMI, Università di Modena e Reggio Emilia (2003).
3. D. Attali, J.-D. Boissonnat and H. Edelsbrunner. *Stability and Computation of the medial axis — a State-of-the-Art Report*, In T. Mller, B. Hamann, and B. Russell, editors, Mathematical Foundations of Scientific Visualization, Computer Graphics, and Massive Data Exploration, Springer-Verlag 2007.
4. P. Bendich, D. Cohen-Steiner, H. Edelsbrunner, J. Harer, D. Morozov, *Inferring local homology from sampled stratified spaces,* 28th Symposium on Foundations of Computer Science (FoCS'07), 536–546.
5. G. Carlsson, T. Ishkanov, V. de Silva, A. Zomorodian, *On the local behavior of spaces of natural images,* International Journal of Computer Vision (2007).
6. G. Carlsson, A. Zomorodian, *Computing Persistent Homology*, Discrete and Computational Geometry, Vol. 33 , 2, p.249 - 274, 2005.
7. R. Chaine, *A Geometric Convection Approach of 3-D Reconstruction*, Proc. 1st Symposium on Geometry Processing, pp. 218-229, 2003.
8. F. Chazal, D. Cohen-Steiner, A. Lieutier, *A Sampling Theory for Compact Sets in Euclidean Space*, Proceedings of the 22nd ACM Symposium on Computational Geometry 2006.
9. F. Chazal, D. Cohen-Steiner, A. Lieutier, *Normal Cone Approximation and Offset Shape Isotopy*, INRIA Research Report RR-6100, January 2007.
10. F. Chazal, A. Lieutier, *The λ-medial axis*, Graphical Models, Volume 67, 4 (2005), Pages 304-331.
11. F. Chazal, A. Lieutier, *Stability and computation of topological invariants of solids in $\mathbb{R}^n$*, Discrete and Computational Geometry, Volume 37, Number 4, pp. 601-617, May 2007.
12. F. Chazal et R. Soufflet, *Stability and finiteness properties of Medial Axis and Skeleton*, Journal of Control and Dynamical Systems vol. 10, No. 2, pp. 149-170, 2004.
13. J. Cheeger, *Critical Points of Distance Functions and Applications to Geometry,* Geometric Topology: recent developments, Montecatini Terme, 1990, Springer Lecture Notes, 1504 (1991), 1-38.
14. F.H. Clarke, *Optimization and NonSmooth Analysis,* Wiley-Interscience, New-York, 1983.

15. D. Cohen-Steiner, H. Edelsbrunner and J. Harer, *Stability of Persistence Diagrams,* Discrete and Computational Geometry, 37:103:120 (2007).
16. D. Cohen-Steiner, H. Edelsbrunner, D. Morozov, *Vines and vineyards by updating persistence in linear time,* 22nd Symposium on Computational Geometry (2006), 119–126.
17. T.Culver, J.Keyser, D.Manocha, *Accurate Computation of the Medial Axis of a Polyhedron,* proceedings of the Fifth ACM Symposium on Solid Modeling and Applications (1999), pp.179-190.
18. T. K. Dey, *Curve and Surface Reconstruction,* Cambridge University Press (2007).
19. T. K. Dey, J. Giesen, S. Goswami, W. Zhao, *Shape Dimension and Approximation from Samples,* Discrete and Computational Geometry, 29:419:434 (2003).
20. T. K. Dey, W. Zhao, *Approximating the medial axis from the Voronoi diagram with a convergence guarantee* Algorithmica, Vol. 38 (2004), 179–200.
21. H. Edelsbrunner, *Surface reconstruction by wrapping finite point sets in space, in* Ricky Pollack and Eli Goodman Festschrift *ed. B. Aronov, S. Basu, J. Pach and M. Sharir.* Springer-Verlag, 379-404.
22. H. Edelsbrunner, D. Letscher, A. Zomorodian, *Topological Persistence and Simplification,* Discrete and Computational Geometry, 28:511:533 (2002).
23. M.Etzion, A.Rappoport, *Computing Voronoi Skeletons of a 3-D Polyhedron by Space Subdivision* Computational Geometry: Theory and Applications, volume 21, number 3, March 2002, pp.87-120.
24. H. Federer, *Curvature measures,* Trans. Amer. Math. Soc. Vol. 93, p 418 (1959).
25. P. Frosini, *Discrete Computation of Size Functions,* Journal of Combinatorics, Information, and System Science, 17:3-4:232:250 (1992).
26. J. Giesen and M. John, *The Flow Complex: A Data Structure for Geometric Modeling,* Proc. 14th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), (2003) 285-294.
27. P. Grassberger, I. Procaccia, Measuring the Strangeness of Strange Attractor, Physica 9D:189:208 (1983).
28. K. Grove, *Critical Point Theory for Distance Functions,* Proc. of Symposia in Pure Mathematics, Vol. 54 (1993), Part 3.
29. K. Grove, K. Shiohama, *A generalized sphere theorem*, Ann. of Math., vol 106, 1977, pp 201-211.
30. A. Hatcher, *Algebraic Topology*, Cambridge University Press (2002).
31. A. Lieutier, *Any open bounded subset of* $\mathbb{R}^n$ *has the same homotopy type as its medial axis*, Computer-Aided Design, 36(2004) 1029-1046, Elsevier.
32. P. Niyogi, S. Smale and S. Weinberger, *Finding the Homology of Submanifolds with High Confidence from Random Samples*, to appear in Discrete and Computational Geometry.
33. V. Robins, *Towards computing homology from approximations,* Topology 24:503:532 (1999).
34. V. de Silva, G. Carlsson, *Topological estimation using witness complexes,* Proc. Symp. Point-based Graphics (2004), 157–166.
35. J.L. Starck, V.J. Martínez, D.L. Donoho, O. Levi, P. Querre and E. Saar, *Analysis of the Spatial Distribution of Galaxies by Multiscale Methods*, arXiv:astro-ph/0406425 v1 18 Jun 04.