Constructions of Simplicial complexes

Žiga Virk

October 20, 2021

Topological methods typically take a simplicial complex as input. However, objects of interest are often not provided in this form. The first step in a topological treatment is thus frequently a creation of simplicial complexes. In this lecture we will present various constructions of complexes from a point cloud, i.e., from a finite collection of points in a metric space. These points may represent a sample of our shape, a collection of numerical data (a subset of $\mathbb{R}^n$), etc.

Our discussion will include two properties we expect from such a construction. The first one is a relationship with the underlying shape, which is often guaranteed by the Nerve Theorem. The second one describes stability to perturbations and proximity of various constructions, as formalized by the interleaving property.

1 Rips complexes

Rips complexes represent perhaps the simplest construction of a complex from a finite collection of points.

☞ Rips complexes are a special case of clique complexes. Suppose G is a graph with vertices V and edges E. The clique complex of G is the abstract simplicial complex with the vertex set V, whose simplices satisfy the following condition: a subset $\sigma \subseteq S$ is a simplex iff each pair of vertices of σ is an edge in G. A Rips complex is the clique complex of its 1-skeleton.

Definition 1.1. *Let X be a metric space and let a sample $S \subset X$ be a finite subset. Choose a scale $r \geq 0$. The **Rips** complex* $\mathrm{Rips}(S, r)$ *is an abstract simplicial complex defined by the following rules:*

1. *The vertex set is S.*
2. *A subset $\sigma \subseteq S$ is a simplex iff* $\mathrm{Diam}(\sigma) \leq r$.

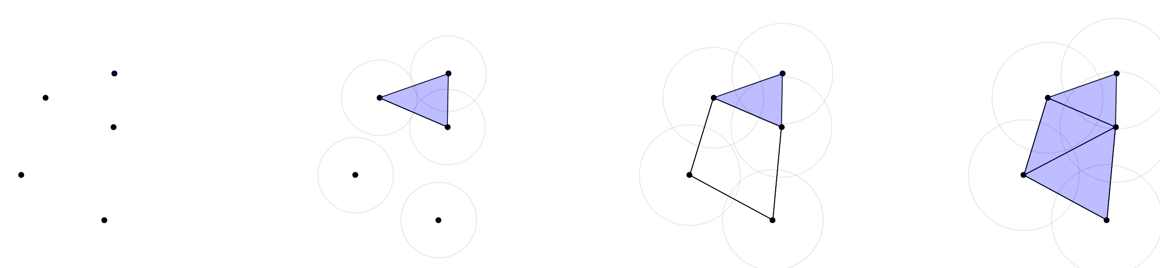

Figure 1: Five points in the plane and three corresponding Rips complexes $\mathrm{Rips}(S, r)$. Visualisation is assisted by circles of radius $r/2$ around each point. For much larger scales the Rips complex is not planar and eventually becomes 4-dimensional.

Remark 1.2. *A few comments about the definition:*

- *Rips complexes are sometimes also called Vietoris-Rips complexes.*
- $\mathrm{Rips}(S, r)$ *represents a combinatorial snapshot of S at scale r.*

- $\mathrm{Diam}(\sigma)$ *is the diameter of* σ. *Condition* $\mathrm{Diam}(\sigma) \leq r$ *means that the distance between each pair of vertices of* σ *is at most* r.
- *It is easy to verify that Rips complexes are indeed abstract simplicial complexes: if* σ *is a simplex then so is each of its subsets.*

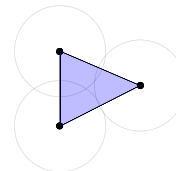

Figure 2: Suppose we are given three points in the plane. These three points span a triangle in the Rips complex for r greater or equal to the maximal pairwise distance between these points. Equivalently, the three balls of radius $r/2$ should pairwise intersect.

Remark 1.3. *Some properties of the Rips complexes:*

1. *Rips complexes are often the preferred construction in TDA due to their simplicity.*
2. $\mathrm{Rips}(S,r)$ *is an abstract simplicial complex, typically not embeddable in* X.
3. *For* r *smaller than the smallest pairwise distance between the points in* S, $\mathrm{Rips}(S,r)$ *is a discrete set, i.e., a complex with no edges or higher-dimensional simplices.*
4. *For* r *at least as large as the largest pairwise distance between the points in* S, $\mathrm{Rips}(S,r)$ *is the* $(|S|-1)$*-simplex, i.e., the simplicial complex on* S *containing all subsets of* S.
5. *If* $r_1 \leq r_2$, *then* $\mathrm{Rips}(S,r_1) \subseteq \mathrm{Rips}(S,r_2)$.

Definition 1.4. *Let* X *be a metric space and let a sample* $S \subset X$ *be a finite subset. The* ***Rips filtration*** *of* S *is a collection of abstract simplicial complexes* $\{\mathrm{Rips}(S,r)\}_{r\geq 0}$ *along with inclusions*

$$i_{r_1,r_2}\colon\ \mathrm{Rips}(S,r_1) \hookrightarrow \mathrm{Rips}(S,r_2) \quad \text{for all } r_1 \leq r_2.$$

A Rips filtration provides a collection of all Rips complexes on S. While a single Rips complex depends on the choice of the scale, the filtration does not. Filtrations will play a fundamental role later in the definition of persistent homology.

2 Čech complexes

Definition 2.1. *Let* X *be a metric space and let a sample* $S \subset X$ *be a finite subset. Choose a scale* $r \geq 0$. *The* ***Čech*** *complex* $\mathrm{Cech}(S,r)$ *is an abstract simplicial complex defined by the following rules:*

1. *The vertex set is* S.
2. *A subset* $\sigma \subseteq S$ *is a simplex iff* $\bigcap_{x\in\sigma} B(x,r) \neq \emptyset$.

☞ Recall $B(x,r) = \{y \in X \mid d(x,y) \leq r\}$ is a closed ball.

☞ Čech complexes are a special case of nerve complexes, a connection that will be explained below.

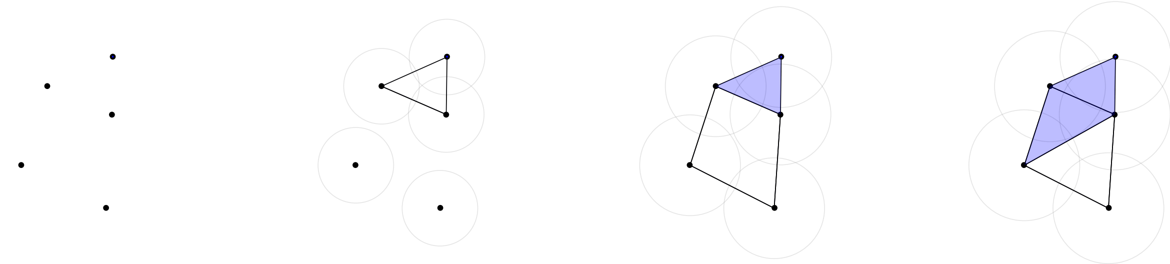

Figure 3: Five points in the plane and three corresponding Čech complexes $\mathsf{Cech}(S, r)$. Visualisation is assisted by circles of radius r around each point. For much larger scales the Čech complex is not planar and eventually becomes 4-dimensional.

Remark 2.2. *A few comments about the definition:*

- *Čech complexes are a classical topological construction used in many contexts throughout topology.*
- $\mathsf{Cech}(S, r)$ *represents a combinatorial snapshot of* S *at scale* r.
- *It is easy to verify that Čech complexes are indeed abstract simplicial complexes.*

Remark 2.3. *Some properties of the Čech complexes:*

1. *While harder to compute*[1], *Čech complexes are attractive due to a well understood geometric interpretation, which will be explained in the next section within the context of nerve complexes.*
2. $\mathsf{Cech}(S, r)$ *is an abstract simplicial complex, typically not embeddable in* X, *although it is often homotopy equivalent*[2] *to a subset of* X.
3. *For* r *smaller than one half of the smallest pairwise distance between the points in* S, $\mathsf{Cech}(S, r)$ *is a discrete set.*
4. *For* r *at least as large as twice the largest pairwise distance between points in* S, $\mathsf{Cech}(S, r)$ *is the* $(|S| - 1)$*-simplex.*
5. *If* $r_1 \leq r_2$, *then* $\mathsf{Cech}(S, r_1) \subseteq \mathsf{Cech}(S, r_2)$.
6. *It is easy to verify*[3] *that* $\mathsf{Cech}(S, r) \subseteq \mathsf{Rips}(S, 2r)$.
7. *In is also easy to see that* $\mathsf{Rips}(S, r) \subseteq \mathsf{Cech}(S, r)$. *A non-trivial inclusion* $\mathsf{Rips}(S, r\sqrt{2}) \subseteq \mathsf{Cech}(S, r)$ *holds in Euclidean spaces by the Jung's Theorem*[4].

[1] See the MiniBall algorithm in the Appendix.

[2] See Nerve Theorem below for details.

[3] If balls of radius r intersect then the pairwise distances between the centers are at most $2r$.

[4]

> **Theorem 2.4** (Jung's Theorem). *If* D *is the diameter of a finite subset* $F \subset \mathbb{R}^n$, *then* F *is contained in a ball of radius at most* $D\sqrt{\frac{n}{2(n+1)}}$.

For $X = \mathbb{R}^n$ we actually obtain

$$\mathsf{Rips}\left(S, r\sqrt{\frac{2(n+1)}{n}}\right) \subseteq \mathsf{Cech}(S, r).$$

The factor $r\sqrt{2}$ in 7 of Remark 2.3 is thus only the smallest upper bound that holds for all n.

> **Definition 2.5.** *Let* X *be a metric space and let a sample* $S \subset X$ *be a finite subset. The* ***Čech filtration*** *of* S *is a collection of abstract*

simplicial complexes $\{\mathsf{Cech}(S, r)\}_{r \geq 0}$ *along with inclusions*

$$i_{r_1, r_2} \colon \mathsf{Cech}(S, r_1) \hookrightarrow \mathsf{Cech}(S, r_2) \quad \text{for all } r_1 \leq r_2.$$

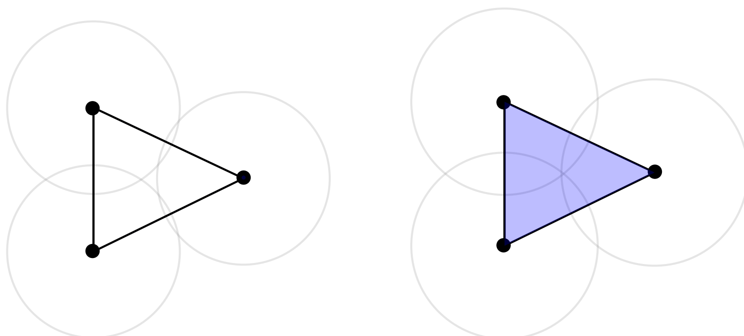

Figure 4: Suppose we are given three points in the plane. These three points span a triangle in the Čech complex iff the three balls of radius r intersect. The left complex consisting of three edges and no triangle does not appear as a Rips complex of any triple of points.

3 Nerve complexes

Cech complexes are a special case of a classical topological construction called the nerve.

Definition 3.1. *For* $k \in \mathbb{N}$ *let* $\mathcal{U} = \{U_1, U_2, \ldots, U_k\}$ *be a collection of subsets of* X*. The* ***nerve*** *of* $\mathcal{U}$ *is the abstract simplicial complex* $\mathcal{N}(\mathcal{U})$ *defined by the following rules:*

1. *The vertex set is* $\mathcal{U} = \{U_1, U_2, \ldots, U_k\}$*, consisting of* k *elements.*
2. *A subset* $\sigma \subseteq \mathcal{U}$ *is a simplex iff* $\bigcap_{i \in \sigma} U_i \neq \emptyset$*.*

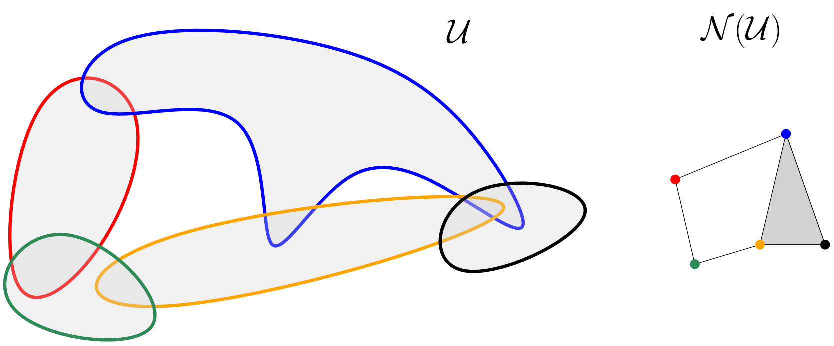

Figure 5: An example of a nerve.

A Čech complex is the nerve of the corresponding collection of r-balls, i.e., $\mathsf{Cech}(S, r) = \mathcal{N}(\{B(s, r)\}_{s \in S})$. Another example is the Delaunay triangulation, which is the nerve of the Voronoi diagram.

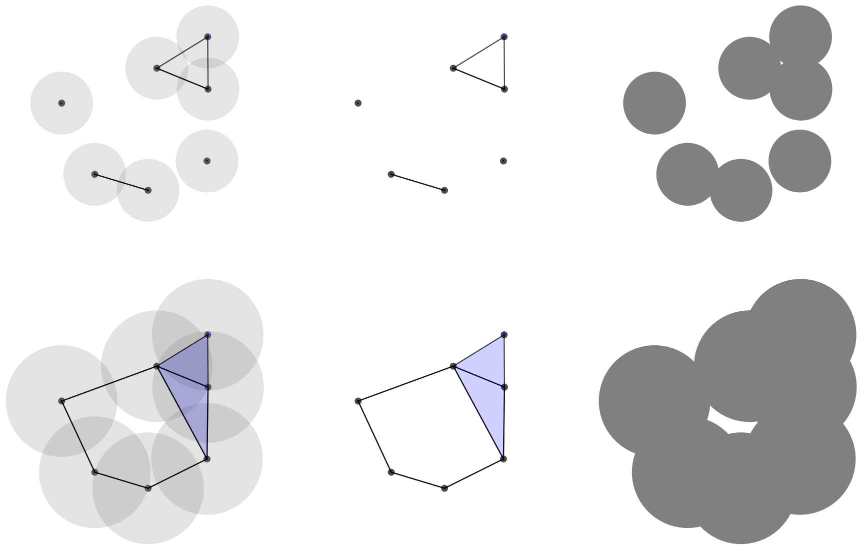

Figure 6: Two examples of Čech complexes: balls and the corresponding complex superimposed (left), complex only (center) and the union of balls homotopy equivalent by the Nerve Theorem(right).

One of the main advantages of nerve complexes is that their homotopy type often represents the union of the elements of $\mathcal{U}$. This is formalized within the context of the Nerve Theorem, of which we now state a special case.

Theorem 3.2. *[Nerve Theorem] Let $n \in \mathbb{N}$ and assume a collection $\mathcal{U} = \{U_1, U_2, \ldots, U_k\}$ consists of closed convex subsets of $\mathbb{R}^n$. Then $U_1 \cup U_2 \cup \ldots \cup U_k \simeq \mathcal{N}(\mathcal{U})$.*

☞ Nerve Theorem actually holds much more generally. For example, assume each finite intersection of sets of $\mathcal{U}$ (including each member $U \in \mathcal{U}$, since it appears as the intersection of $\{U\} \subseteq \mathcal{U}$) is either empty or contractible. If $\mathcal{U}$ is a finite collection of closed subsets in $\mathbb{R}^n$, or an arbitrary collection of open sets in a metric space, then

$$\bigcup_{U \in \mathcal{U}} U \simeq \mathcal{N}(\mathcal{U}).$$

This is a stronger statement than Theorem 3.2 by Lemma 3.3.

An idea of a proof is given in Appendix. The Nerve Theorem does not hold for an arbitrary collection of subsets as Figure 5 demonstrates.

For Delaunay triangulations the Nerve Theorem provides no additional information. As the Voronoi cells are convex, the Nerve Theorem implies that the Delaunay triangulation is contractible, a fact we already know as it triangulates a convex (hence contractible by Lemma 3.3) planar region.

Lemma 3.3. *Let $n \in \mathbb{N}$. Each convex subset of $\mathbb{R}^n$ is contractible.*

Proof. Assume $A \subset \mathbb{R}^n$ is convex and fix $x_0 \in A$. We can slide each $a \in A$ into x_0 along the line segment from a to x_0. This results in a homotopy $H(a,t) = (1-t)a + t\,x_0$ between the identity map on A and the constant map at x_0, hence A is contractible. □

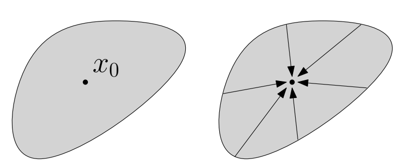

Figure 7: A sketch of Lemma 3.3.

On the other hand, the Nerve Theorem provides a homotopical description of the Čech complex. As Euclidean balls are convex, we obtain

$$Cech(S,r) = \mathcal{N}(\{B(s,r)\}_{s \in S}) \simeq \bigcup_{s \in S} B(s,r) = N(S,r),$$

i.e., the Čech complex $Cech(S,r)$ has the homotopy type of the r-neighborhood of S. This fact is the foremost reason for the computational utilisation of Čech complexes: while they are hader to compute than Rips complexes, we know that the obtained homotopy type represents the r-neighborhood of S. In this spirit we can interpret Figure 6. Furthermore, this observation can be used to prove reconstruction results: given a closed connected surface X in an Euclidean space, for each sufficiently small scale parameter $r \geq 0$ and for each sufficiently dense finite subset $S \subset X$ we have $X \simeq Cech(S, r)$, i.e., the homotopy type of a space X can be reconstructed using Čech complexes (in Euclidean or geodesic metric)[5]. In the Euclidean metric this holds as[6] $X \simeq N(X, r)$ for small r.

[5] The same result holds for more general spaces and under appropriate conditions also for Rips complexes. However, almost all of the proofs are based on the application of the Nerve Theorem to Čech complexes.

[6] Imagine a circle in the plane, a knot in $\mathbb{R}^3$ or a surface in $\mathbb{R}^3$: its small thickening is homotopy equivalent to the space itself.

Alpha complexes

Alpha complexes are a fusion between planar Čech complexes and Delaunay triangulations.

> **Definition 3.4.** *Let $r \geq 0$ and assume $S \subset \mathbb{R}^2$ is a finite collection of points satisfying a general position property: no four points of S lie on the same circle. For each $s \in S$ let V_s denote the corresponding Voronoi cell. The* ***alpha complex*** *of S at scale r is the following nerve: $\mathcal{N}\Big(\{V_s \cap B(s,r)\}_{s \in S}\Big)$.*

Assume S is as in Definition 3.4. While $\mathsf{Cech}(S, r)$ may be of arbitrarily high dimension[7], the Nerve Theorem guarantees it is homotopy equivalent to a planar subset. The alpha complex of S at scale r is a planar[8] complex, which is homotopy equivalent to $\mathsf{Cech}(S, r)$. To see this note that by the Nerve Theorem[9] both are homotopy equivalent to

$$\bigcup_{s \in S} B(s,r) = \bigcup_{s \in S} \big(V_s \cap B(s,r)\big).$$

Thus alpha complexes may be seen as an efficient way of obtaining the homotopy type of a Čech complex in the plane.

[7] This implies, amongst other things, that it may be computationally inefficient.

[8] Note that it is a subcomplex of the Delaunay triangulation on S.

[9] Sets $V_s \cap B(s,r)$ are intersections of closed convex sets thus closed and convex themselves.

☞ To see $\bigcup_{s \in S} B(s,r) = \bigcup_{s \in S} \big(V_s \cap B(s,r)\big)$, take any $x \in \bigcup_{s \in S} B(s,r)$ and note that if $s \in S$ is a closest point to x in S, then $x \in V_s \cap B(x,r)$.

Another way of thinking of alpha complexes is as a model for molecules. Each atom in a molecule has a radius[10] and touches (rather than intersects) other atoms within the range.

[10] Assume all the radii are the same. For different radii there is a well studied concept of a weighted alpha complex.

Mapper

Another example of a construction based on the idea of the nerve is Mapper. In contrast to the constructions above it is typically[11] a one-dimensional simplicial complex, i.e., a graph. Mapper can be thought of as a one-dimensional sketch of a space X as detected through the lens of a single map on X.

[11] By the definition that will be provided, a Mapper is a simplicial complex of arbitrary dimension. However, our discussion and examples will focus on one-dimensional case, as do the practical applications in which Mapper is used.

We first describe a **theoretical setup**. Assume:

- X is a metric space;

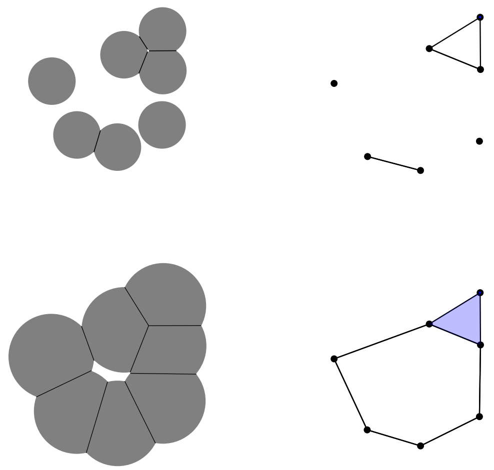

Figure 8: Alpha complexes corresponding to the situation in Figure 6. Note that the alpha complexes are smaller (or equal) yet still homotopy equivalent to the corresponding Čech complexes. For larger r the Čech complexes become higher-dimensional while the alpha complexes maintain the dimensionality bound 2.

A decomposition into regions of the form $V_s \cap B(s,r)$ (on the left) mimics the decomposition of molecules into atoms.

- $f\colon X \to [0,1]$ is a (continuous) map[12];
- $\mathcal{U}$ is a collection[13] of subsets of $[0,1]$, whose union is $[0,1]$.

[12] In accordance with standard practice we restrict ourselves to the cases when the target space is $[0,1]$. However, there is no theoretical reason for doing so and the construction is well defined even if we replace $[0,1]$ by some more complicated space.

[13] Typically we restrict to cases when no three subsets of $\mathcal{U}$ intersect. In such cases Mapper is a one-dimensional simplicial complex.

Definition 3.5. *For each U let $\mathcal{V}_U$ denote the collection of all components of $f^{-1}(U)$ and define $\mathcal{V} = \bigcup_{U\in\mathcal{U}} \mathcal{V}_U$ as the collection of all subsets of X that appear as a component of a preimage $f^{-1}(U)$ for some $U \in \mathcal{U}$.* ***Mapper*** *is defined as $\mathcal{M}(X,f,\mathcal{U}) = \mathcal{N}(\mathcal{V})$.*

An example is provided by Figure 9.

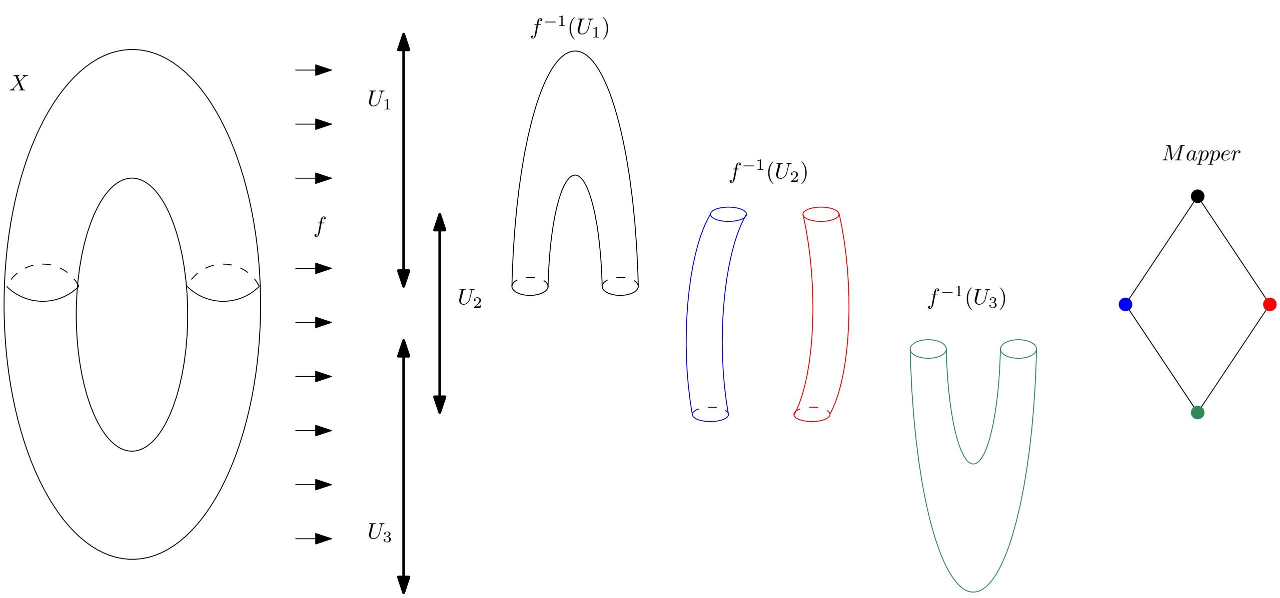

Figure 9: The construction of a Mapper: a space (the torus on the left), a continuous map (projection f onto the vertical axis), cover $\mathcal{U} = \{U_1, U_2, U_3\}$ of the interval $[0,1]$, a decomposition of the preimages into the four components and the resulting graph (right).

In practice data is often given as a finite set of points along with certain measurements. For example, we may have a collection of patients along with their heart rate and blood pressure, or a collection of basketball players along with their statistics, etc. In this case a modified **practical setup** comes into play. Assume:

- X is a finite set;
- $f\colon X \to I$ is a map (measurement);
- $\mathcal{U}$ is a chosen partition of $[0,1]$ into intervals, typically of fixed length $\varepsilon > 0$.
- Π is a chosen clustering scheme[14] on X.

[14] This step possibly includes additional choices of clustering parameters.

Definition 3.6. *For each $U \in \mathcal{U}$ let $\mathcal{V}_U$ denote the collection of all clusters of $f^{-1}(U)$ with respect to Π and define $\mathcal{V} = \bigcup_{U\in\mathcal{U}} \mathcal{V}_U$ as the collection of all subsets of X that appear as a cluster of a preimage $f^{-1}(U)$ for some $U \in \mathcal{U}$.* ***Mapper*** *is defined as $\mathcal{M}(X, f, \mathcal{U}) = \mathcal{N}(\mathcal{V})$.*

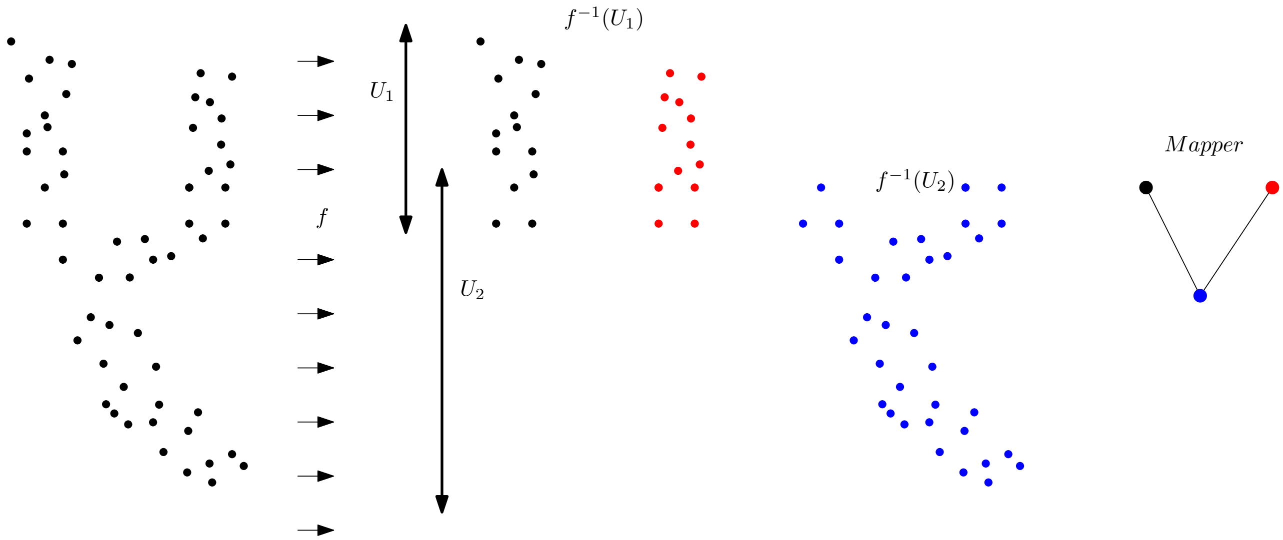

Figure 10: The construction of a Mapper when X is a point cloud.

While a point could and a number of measurements on it are often given, one typically has to construct a single function f and a partition $\mathcal{U}$, and choose other parameters very carefully to extract the desired information. Mapper is usually not analyzed further with topological tools but rather visualized, which is why the one-dimensionality is preferable.

4 Interleaving properties

Given a finite subset of a metric space we have described how to associate various complexes with that set. If we think for a moment about finite abstract complexes we note that these objects are discrete: it would be hard to define an obvious distance between abstract simplicial complexes. On the other hand, we have a continuous selection of inputs and input parameters: scale r is typically positive and there are reasonable notions of a distance between finite subsets of a metric space. As a result any assignment of a single complex is bound to have discontinuities[15] (instabilities) of some sort, see Example 4.1 for a demonstration.

Example 4.1. *Let $X = \{0,1\} \subset \mathbb{R}$. Note that $\mathrm{Rips}(X,r)$ changes discontinuously at $r = 1$: while $\mathrm{Rips}(X,1)$ is a single edge (along with the two boundary points), for each $r < 1$ the complex $\mathrm{Rips}(X,1)$ consists of only two vertices.*

[15] Unless it assigns a constant complex, of course.

However, it turns out we can define a distance on filtrations, for which the assignment of a filtration[16] becomes a continuous function of the input set and the scale parameter. [17]

[16] Of course, this eliminates the dependency on the scale parameter r.

[17] For the sake of simplicity we will restrict ourselves to the mentioned Rips and Čech filtrations although the concept can be defined more generally.

Definition 4.2. *Choose $\varepsilon > 0$. Filtrations $\{A_r\}_{r\geq 0}$ and $\{B_r\}_{r\geq 0}$ (obtained by the Rips or the Čech construction) are ε-**interleaved** if there exist simplicial maps $\varphi_r \colon A_r \to B_{r+\varepsilon}$ and $\psi_r \colon B_r \to A_{r+\varepsilon}$ such that $\varphi_{r+\varepsilon} \circ \psi_r \colon B_r \to B_{r+2\varepsilon}$ and $\psi_{r+\varepsilon} \circ \varphi_r \colon A_r \to A_{r+2\varepsilon}$ are equal to the corresponding inclusions.*

Maps of Definition 4.2 can be visualised by drawing the following commutative[18] "ladder" diagram.

[18] Adjective "commutative" refers to the fact that all maps commute, i.e., going from one complex to another through any viable sequence of maps results in the same inclusion map.

$$\begin{tikzcd}
\cdots \arrow[r] & A_r \arrow[r] \arrow[dr, "\varphi_r"] & A_{r+\varepsilon} \arrow[r] \arrow[dr] & A_{r+2\varepsilon} \arrow[r] & \cdots \\
\cdots \arrow[r] & B_r \arrow[r] \arrow[ur, "\psi_r"] & B_{r+\varepsilon} \arrow[r] \arrow[ur] & B_{r+2\varepsilon} \arrow[r] & \cdots
\end{tikzcd}$$

Definition 4.3. *Given two filtrations their **interleaving distance** is defined as the infimum of all values $\varepsilon > 0$, for which the filtrations are ε-interleaved.*

It turns out that in our context (Rips and Čech filtrations on finite collections of points) the interleaving distance is a metric on the set of filtrations. In contrast, recall that there seems to be no geometrically meaningful metric on the set of single finite simplicial complexes.

The concept of interleaving will play a prominent role later in the context of the stability of persistent homology. At this point we can use it to phrase two proximity results.

Example 4.4. *Let $X = \{0,1\} \subset \mathbb{R}$ and $Y = \{0.1, 1.2\} \subset \mathbb{R}$. $\mathrm{Rips}(X,r)$ consists of:*

- *two points for $r < 1$;*
- *one edge for $r \geq 1$.*

$\mathrm{Rips}(Y,r)$ consists of:

- *two points for $r < 1.1$;*
- *one edge for $r \geq 1.1$.*

The filtrations are 0.1-interleaved.

Theorem 4.5 (Stability with respect to spaces). *Choose $\varepsilon > 0$ and assume $X = \{x_1, x_2, \ldots, x_k\}$ and $Y = \{y, y_2, \ldots, y_k\}$ with $d(x_i, y_i) \leq \varepsilon, \forall i$, i.e., X and Y each consist of k points, such that the corresponding distances are at most ε. Then:*

- *The Rips filtrations of X and Y are 2ε-interleaved.*
- *The Čech filtrations of X and Y are ε-interleaved.*

Proof. It follows directly from the triangle inequality (see Figure 11) that if a subset $\sigma \subset X$ is of diameter r, then the corresponding subset[19] $\tau \subset Y$ is of diameter at most $r + 2\varepsilon$. Hence if σ is a simplex in $\mathsf{Rips}(X, r)$, then τ is a simplex in $\mathsf{Rips}(Y, r + 2\varepsilon)$. Consequently we may deduce that:

- maps $\mathsf{Rips}(X, r) \to \mathsf{Rips}(Y, r + 2\varepsilon)$ defined by $x_i \mapsto y_i$ are simplicial;
- maps $\mathsf{Rips}(Y, r) \to \mathsf{Rips}(X, r + 2\varepsilon)$ defined by $y_i \mapsto x_i$ are simplicial;
- as the above two maps obviously commute with the inclusions we conclude that the Rips filtrations of X and Y are 2ε-interleaved;
- in a similar fashion we may conclude that the Čech filtrations of X and Y are ε-interleaved.

□

[19] Subset τ is formed by taking the points of Y with the same indices as appear in the points of σ.

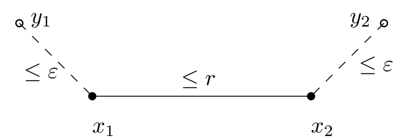

Figure 11: If $d(x_1, x_2) \leq r$ and $d(x_i, y_i) \leq \varepsilon$ then it is apparent that $d(y_1, y_2) \leq r + 2\varepsilon$.

These conclusions tell us that if we perturb our point set slightly, the resulting filtration does not change much in terms of the interleaving distance, i.e., the construction of a filtration is stable. In a similar fashion we can express the relationship between Rips and Čech filtrations.

Rips-Čech correlation

Recall that $\mathsf{Cech}(S, r) \subseteq \mathsf{Rips}(S, 2r)$ and $\mathsf{Rips}(S, r) \subseteq \mathsf{Cech}(S, r) \subseteq \mathsf{Cech}(S, 2r)$. This implies that the Rips and Čech filtrations, when constructed with logarithmic scales[20], are $(\log 2)$-interleaved, i.e.,

$$\{\mathsf{Rips}(S, e^r)\}_{r \geq 0} \quad \text{and} \quad \{\mathsf{Cech}(S, e^r)\}_{r \geq 0}$$

are $(\log 2)$-interleaved.

[20] Note that $\mathsf{Cech}(S, e^r) \subseteq \mathsf{Rips}(S, 2e^r) = \mathsf{Rips}(S, e^{\log 2 + r})$ and similarly $\mathsf{Rips}(S, e^r) \subseteq \mathsf{Cech}(S, e^r) \subseteq \mathsf{Cech}(S, e^{\log 2 + r})$. These inclusions are the interleaving maps.

5 Concluding remarks

Recap (highlights) of this chapter

- Complexes: Rips, Čech, nerve, alpha, Mapper;
- Nerve Theorem;
- Interleaving;

Background and applications

Constructions of simplicial complexes greatly depend on the projected use. Rips and Čech complexes along with the nerve construction are relatively well understood and have been originally introduced for theoretical purposes in the first half of the twentieth century. Their use has recently been extended to the applied setting. They are the complexes most exposed to the curse of dimensionality[21]. Alpha complexes arose decades later within the realm of computational geometry and are intended for computationally intense applications. Mapper is one of the most recent constructions. It is often thought of as a low-dimensional projection method and has turned out to be a commercial success. At about the same time the interleaving distance emerged as a measure of stability of filtrations and persistent homology, although equivalent concepts have been known in pure topology for a long time.

[21] The curse of dimensionality: an annoying fact that the number of simplices typically grows fast with the dimension of a simplicial complex. This presents challenges for their computational applications, which are partially addressed by alternative constructions of complexes.

Appendix: the MiniBall algorithm

Given a finite subset $\sigma \subset X \subset \mathbb{R}^n$ the MiniBall algorithm is a recursive algorithm that returns the miniball of σ, i.e., the minimal[22] ball in $\mathbb{R}^n$ containing σ. As such, the algorithm provides a computational verification of the containment of σ in a Čech complex: $\sigma \in \mathsf{Cech}(X,r)$ iff[23] the radius of the miniball is at most r. As the radius of the ball is also provided, the algorithm actually provides the lower bound for the scales r at which σ is a simplex in $\mathsf{Cech}(X,r)$, hence a single execution of the algorithm suffices for the entire filtration.

[22] Minimality is considered with respect to the radius. Such a ball is unique.

[23] $z \in \bigcap_{x\in\sigma} B(x,r) \Leftrightarrow \sigma \subset B(z,r)$.

- Input: disjoint finite sets $\tau, \nu \subset \mathbb{R}^n$.
- Output: the minimal ball with:
  - τ in the ball;
  - ν on the boundary of the ball.

⚠ Given random finite $\tau, \nu \subset \mathbb{R}^n$, there typically exists no ball containing τ and having ν on the boundary. The algorithm is designed so that only the pairs (τ, ν), for which this condition is satisfied are called.

The algorithm is initiated by calling Miniball$(\sigma, \emptyset)$ [24] and terminates with the miniball B when $\tau = \emptyset$. It inductively scans through the points of τ, either removing[25] a point or putting[26] it into ν. When

[24] I.e., $\tau = \sigma, \nu = \emptyset$

[25] If removing the point from the set does not change the miniball of the set.

[26] If removing the point decreases the miniball.

Algorithm 1: Miniball(τ, ν).

```
if τ = ∅ then
| compute miniball B directly
else
| choose u ∈ τ;
| B =miniball(τ − {u}, ν);
| if u ∉ B then
| | B =miniball(τ − {u}, ν ∪ {u})
return B
```

$\tau = \emptyset$ the set ν consists of at most $n+1$ points that lie on the boundary of the miniball of σ and determine it. In this case we can use the standard[27] circumsphere and circumradius formulas to get the miniball.

[27] The expressions in terms of determinant are quite lengthy and at this point omitted.

Appendix: a sketch of a proof of the Nerve Theorem 3.2

A special case of the proof is illustrated by Figures 12, 13, and 14.

A sketch of a proof of the Nerve Theorem. For the sake of simplicity[28] let us assume the nerve is of dimension 1, i.e., all triple intersections of sets of $\mathcal{U}$ are empty. Define $\mathrm{Z} \subset X \times \mathcal{N}(\mathcal{U})$ as:

[28] A complete general proof is much more technical but broadly follows the same steps as are presented here.

$$\mathrm{Z} = \bigcup_{\sigma \in \mathcal{N}(\mathcal{U})} \Big(\bigcap_{s \in \sigma} U_s \times \sigma \Big).$$

We will prove that $\mathrm{Z} \simeq \mathrm{X}$ and $\mathrm{Z} \simeq \mathcal{N}(\mathcal{U})$.

In order to prove $Z \simeq X$ note that for each $x \in X$ the section $(\{x\} \times \mathcal{N}(\mathcal{U})) \cap \mathrm{Z}$ is a simplex[29] in the nerve spanned by all $s \in S$, for which $x \in U_s$. Contracting each such simplex to a point in a synchronized manner for each $x \in X$ we obtain a deformation of Z to X, hence $\mathrm{Z} \simeq X$.

[29] In our case, either an edge or a vertex.

In order to prove $Z \simeq \mathcal{N}(\mathcal{U})$ note that for each $y \in \mathcal{N}(\mathcal{U})$ the section $(X \times \{y\}) \cap \mathrm{Z}$ is a contractible set by assumptions. Contract first the sections of this form for all non-vertices y, and then conclude by contracting all the sections for vertices. We obtain a deformation of Z to $\mathcal{N}(\mathcal{U})$, hence $\mathrm{Z} \simeq \mathcal{N}(\mathcal{U})$. □

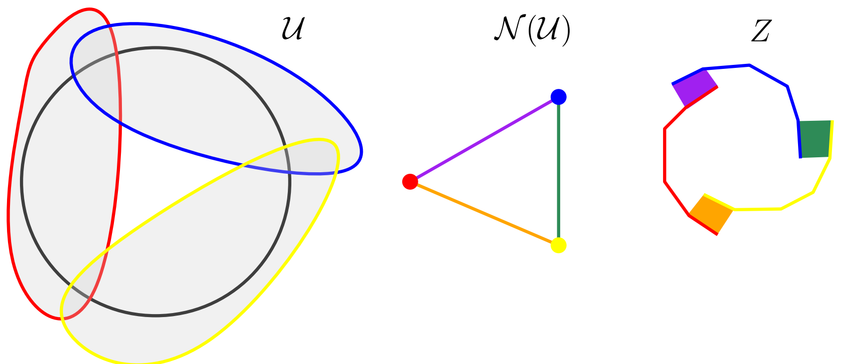

Figure 12: A collection $\mathcal{U}$ of subsets of a circle $X = S^1$ (left), the corresponding nerve (center) and space Z constructed in the proof (right). The sets of $\mathcal{U}$ are illustrated as subsets of the plane for greater clarity, while formally $\mathcal{U}$ consists of their intersections with X.

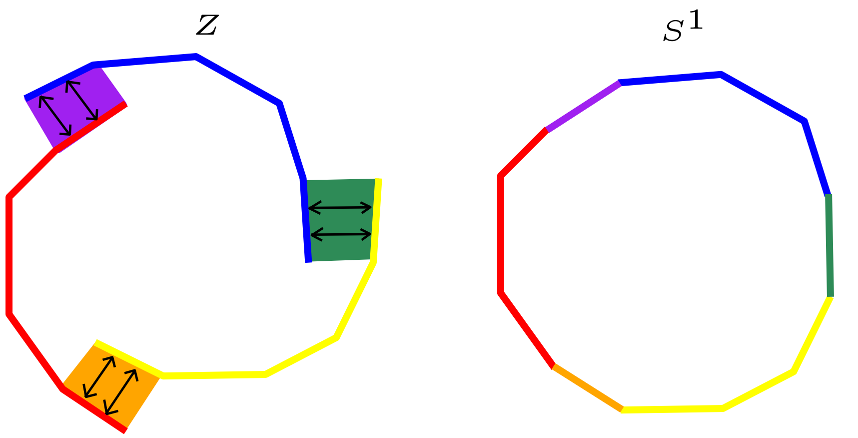

Figure 13: Proving $Z \simeq S^1$ we contract the sections above points of $x \in X$ in Z corresponding to edges in the nerve complex (contract along the indicated arrows on the left) to obtain S^1.

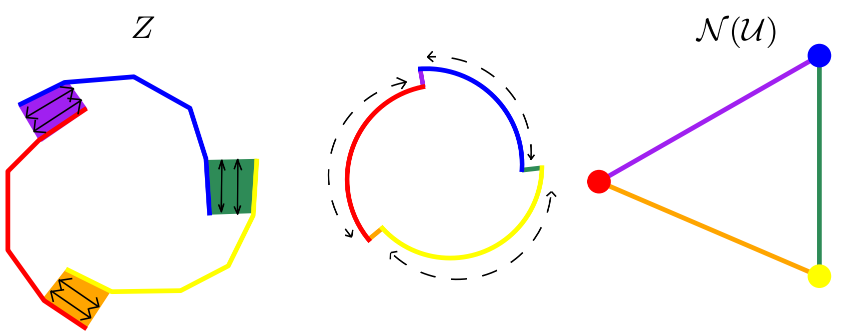

Figure 14: Proving $Z \simeq \mathcal{N}(\mathcal{U})$ we first contract the sections above non-vertices $y \in \mathcal{N}(\mathcal{U})$ in Z (contract along the indicated arrows on the left) to obtain the space in the center. Conclude by contracting the sections above vertices $y \in \mathcal{N}(\mathcal{U})$ in Z (contract along the indicated dashed arrows in the center) to obtain $\mathcal{N}(\mathcal{U})$ on the right.

Appendix: Dowker duality

Nerve complexes are natural complexes arising from a collection of subsets. There is another similar construction, called the Vietoris complexes, that is in a way dual to the nerve construction.

Definition 5.1. *For $k \in \mathbb{N}$ let $\mathcal{U} = \{U_1, U_2, \dots, U_k\}$ be a collection of subsets of a finite space X, whose union is X. The* ***Vietoris complex*** *of $\mathcal{U}$ is the abstract simplicial complex $\mathcal{V}(\mathcal{U})$ defined by the following rules:*

1. *The vertex set is X.*
2. *A set $\sigma \subseteq X$ is a simplex iff there exists $U \in \mathcal{U}$ containing σ.*

We see that maximal simplices of $\mathcal{V}(\mathcal{U})$ are determined by (inclusion-wise) maximal sets of $\mathcal{U}$. There is a surprising connection between the nerve complexes and Vietoris complexes.

☞ Dowker duality actually holds for an arbitrary collection of subsets $\mathcal{U}$ of an arbitrary set X, no additional structure is necessary. Even in such generality it can be proved with ease using a general form of the Nerve Theorem.

Theorem 5.2 (Dowker Duality)**.** *For $k \in \mathbb{N}$ let $\mathcal{U} = \{U_1, U_2, \dots, U_k\}$ be a collection of subsets of a finite space X, whose union is X. Then $\mathcal{N}(\mathcal{U}) \simeq \mathcal{V}(\mathcal{U})$.*

Proof. Consider $\mathcal{V}(\mathcal{U})$ as a subspace of a Euclidean space. Each U_i determines a simplex Δ_{U_i} spanned by all points of U_i. Note that $\{\Delta_{U_1}, \Delta_{U_2}, \dots, \Delta_{U_2}\}$ are closed convex sets whose union is $\mathcal{V}(\mathcal{U})$. By the Nerve Theorem 3.2 $\mathcal{V}(\mathcal{U})$ is homotopy equivalent to the nerve of $\{\Delta_{U_1}, \Delta_{U_2}, \dots, \Delta_{U_2}\}$. This nerve, on the other hand, is actually $\mathcal{N}(\mathcal{U})$ via the correspondence $\Delta_{U_i} \mapsto U_i$, see Figure 15 for a visual sketch of the proof. □

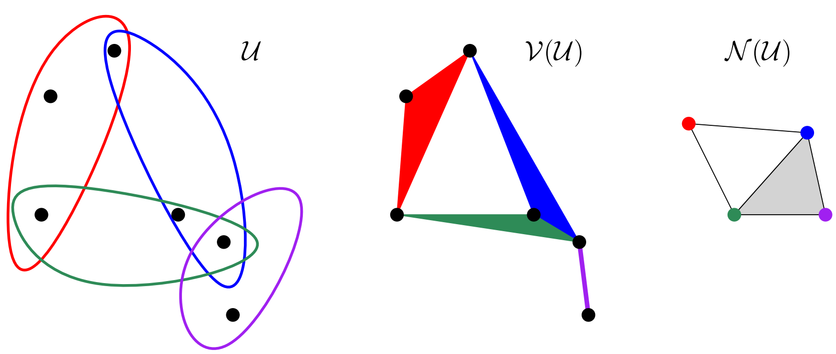

Figure 15: A cover of six points by four colored sets (left), its Vietoris complex (center) and nerve (right). Colored simplices on the central picture provide a collection of subsets satisfying the conditions of the Nerve Theorem. On the other hand, their nerve is actually $\mathcal{N}(\mathcal{U})$ on the right. By the Nerve Theorem we conclude the Dowker Duality.

While Čech complexes are nerves associated to a collection of balls of radius r, Rips complexes are Vietoris complexes associated to a collection of sets of diameter at most r.