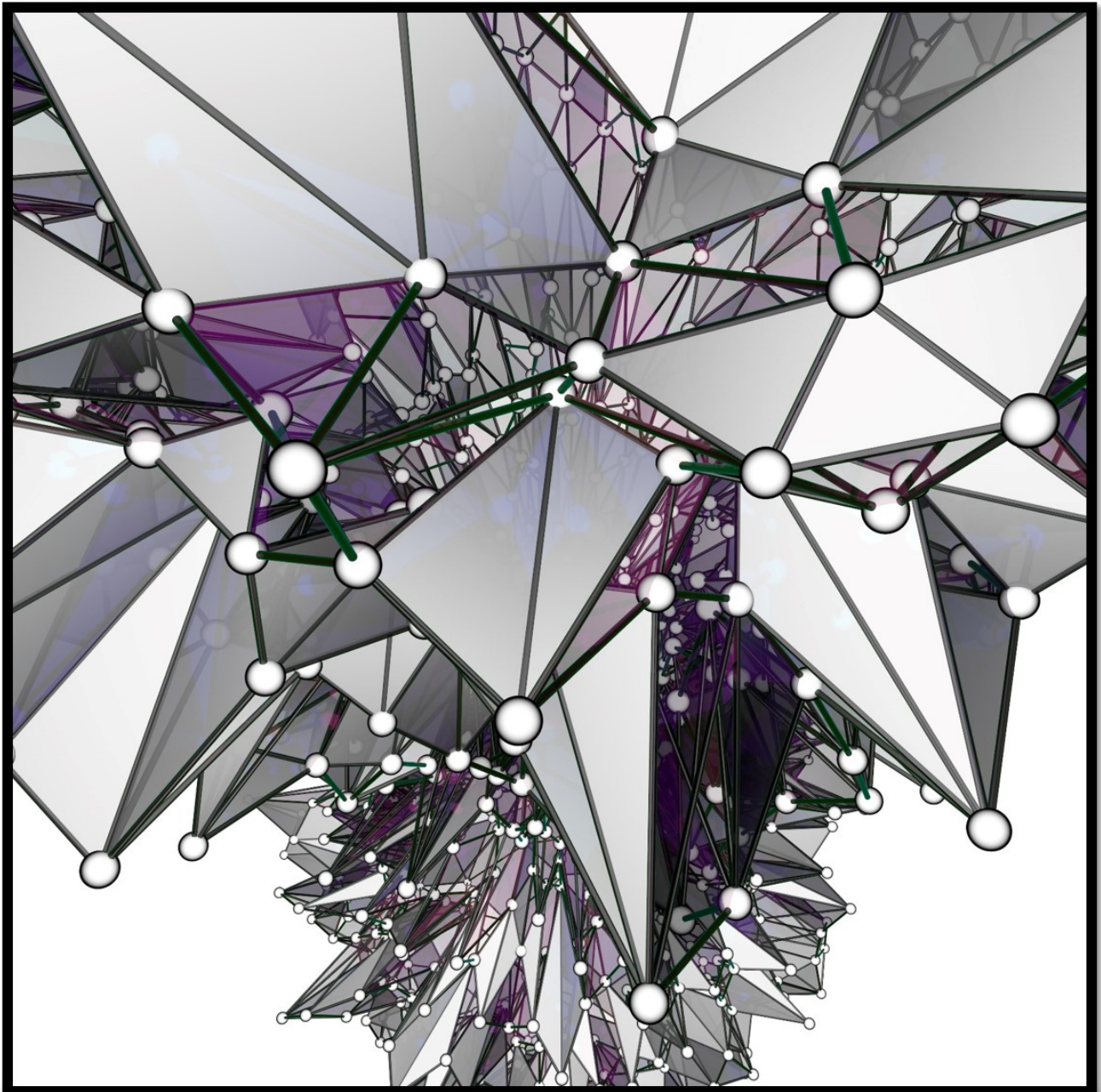


# 1. COMPLEXES



## 1.1 COMBINATORICS

Let  $V$  be a finite nonempty set whose elements we will call *vertices*.

**DEFINITION 1.1.** A **simplicial complex** on  $V$  is a collection  $K$  of nonempty subsets of  $V$  subject to two requirements:

- for each vertex  $v$  in  $V$ , the singleton  $\{v\}$  is in  $K$ , and
- if  $\tau$  is in  $K$  and  $\sigma \subset \tau$  then  $\sigma$  must also be in  $K$ .

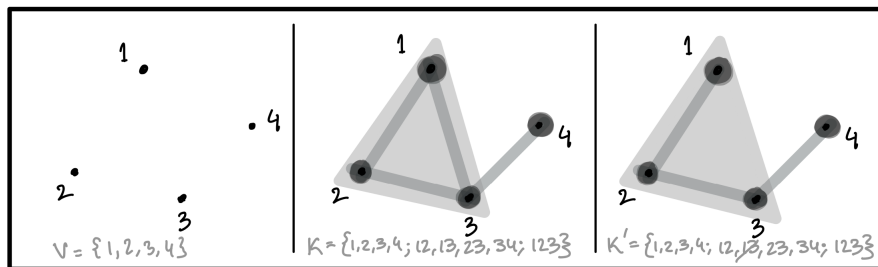
The nonempty subsets which lie in a simplicial complex  $K$  are called the **simplices** of  $K$ . The *dimension* of a simplex  $\sigma$  in  $K$  is defined to be

$$\dim \sigma = \#\sigma - 1,$$

where  $\#\sigma$  denotes the cardinality of (or, the number of vertices contained in)  $\sigma$ . Thus, the singletons  $\{v\}$  all lie in  $K$  and have dimension zero, all pairs  $\{v, v'\}$  which happen to lie in  $K$  have dimension one, and so forth. The dimension of  $K$  itself is given by taking a maximum over constituent simplices, i.e.,

$$\dim K = \max\{\dim \sigma \mid \sigma \in K\}.$$

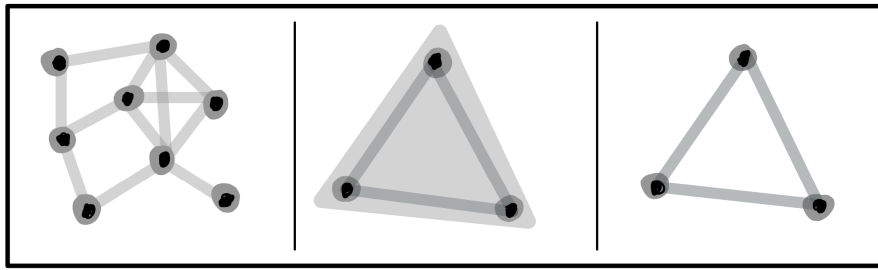
We will write  $K_i$  to denote the set of all  $i$ -dimensional simplices in  $K$ ; the first requirement of Definition 1.1 guarantees that  $K_0$  equals the vertex set  $V$ . The figure below contains cartoon depictions of a vertex set  $V$  with four elements, a simplicial complex  $K$  and a non-simplicial complex  $K'$  — the fact that the set  $\{1, 2, 3\}$  is present in  $K'$  but the subset  $\{1, 3\}$  is not disqualifies  $K'$  from being a simplicial complex.



Here are some more exciting examples of simplicial complexes.

- **Graphs:** a (finite, undirected, simple) graph is a pair  $G = (V, E)$  consisting of a finite set  $V$  (whose elements are called vertices as before) and a set  $E \subset V \times V$  consisting of distinct vertex-pairs, usually called *edges*. Every graph automatically forms a one-dimensional simplicial complex  $K$  with  $V = K_0$  and  $E = K_1$ .
- **Solid Simplices:** for each integer  $k \geq 0$ , the solid  $k$ -simplex is the simplicial complex  $\Delta(k)$  defined on the vertex set  $\{0, 1, \dots, k\}$  whose simplices are *all* possible subsets of vertices.
- **Hollow simplices:** the hollow  $k$ -simplex (for each integer  $k \geq 1$ ) is denoted  $\partial\Delta(k)$  and defined exactly like a solid  $k$ -simplex, *except* that we discard the unique  $(k + 1)$ -dimensional simplex  $\{v_0, \dots, v_k\}$ . Thus,  $\partial\Delta(k)$  has dimension  $k - 1$ .

The figure below illustrates a graph, a solid 2-simplex and a hollow 2-simplex respectively.



So far, the structure of a simplicial complex appears to be purely combinatorial — we are given a universal finite set  $V$  of vertices, and we may select any collection  $K$  of subsets of  $V$  provided that the two constraints of Definition 1.1 are satisfied. The first step towards expanding this perspective beyond combinatorics is to formally relate simplices with their subsets.

**DEFINITION 1.2.** Given two simplices  $\sigma$  and  $\tau$  of a simplicial complex  $K$ , we say that  $\sigma$  is a **face** of  $\tau$ , denoted  $\sigma \leq \tau$ , whenever every vertex of  $\sigma$  is also a vertex of  $\tau$ .

Given a pair  $\sigma \leq \tau$  of simplices of a simplicial complex  $K$ , we call the difference  $\dim \tau - \dim \sigma$  the *codimension* of  $\sigma$  as a face of  $\tau$ ; note that the codimension is always a non-negative integer.

## 1.2 SUBCOMPLEXES, CLOSURES AND FILTRATIONS

Knowledge of face relations between simplices allows us to define subsets of simplicial complexes which are simplicial complexes in their own right.

**DEFINITION 1.3.** Let  $K$  be a simplicial complex. A subset  $L \subset K$  of simplices is called a **subcomplex** of  $K$  if it satisfies the following property: for each simplex  $\tau$  in  $L$ , if  $\sigma$  is a face of  $\tau$  in  $K$ , then  $\sigma$  also belongs to  $L$ .

In general, for a subcomplex  $L \subset K$ , we do *not* require every vertex of  $K$  to be a vertex of  $L$ .

**EXAMPLE 1.4.** Each hollow  $k$ -simplex  $\partial\Delta(k)$  naturally forms a subcomplex of the corresponding solid  $k$ -simplex  $\Delta(k)$ ; each vertex of a given simplicial complex is automatically a subcomplex.

If you are handed a collection  $K'$  of simplices in some simplicial complex  $K$ , it is often desirable to check how far  $K'$  is from being a subcomplex of  $K$ . The following notion is often helpful when performing such checks.

**DEFINITION 1.5.** The **closure** of a collection of simplices  $K'$  in a simplicial complex  $K$  is defined to be the smallest subcomplex  $L \subset K$  satisfying  $K' \subset L$ .

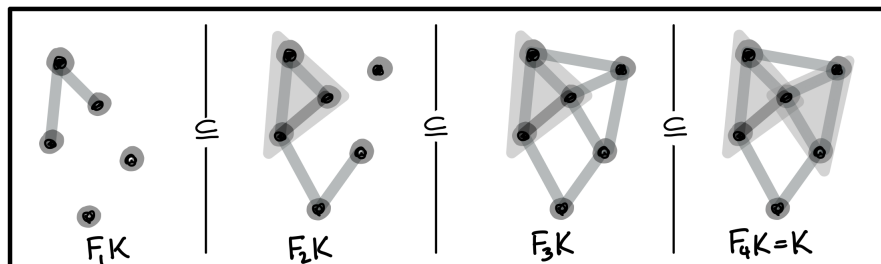
Evidently, a nonempty subcollection  $K' \subset K$  of simplices forms a subcomplex if and only if it equals its own closure. It should be noted that the closure of a given collection  $K'$  of simplices can be *much larger* than  $K'$ . The following exercise is highly recommended: if  $\sigma$  is a single  $k$ -dimensional simplex in a simplicial complex  $K$ , show that the closure of  $\sigma$  in  $K$  contains  $2^k - 1$  simplices. Of particular interest to us here are ascending chains of subcomplexes.

**DEFINITION 1.6.** Let  $K$  be a simplicial complex; a **filtration** of  $K$  (of length  $n$ ) is a nested sequence of subcomplexes of the form

$$F_1K \subset F_2K \subset \cdots \subset F_{n-1}K \subset F_nK = K.$$

In general, the dimensions of the intermediate  $F_i K$  are not constrained by  $i$ . On the other hand, in order to have a well-defined notion of length, we require  $F_i K \neq F_{i+1} K$  for all  $i$ .

The figure below depicts a filtration of length four of the simplicial complex in the right-most panel; the things to check are that each panel contains a genuine simplicial complex, and that these simplicial complexes are getting strictly larger as we scan from left to right.



### 1.3 GEOMETRIC REALIZATION

The **geometric simplex** spanned by a collection of points  $\{x_0, x_1, \dots, x_k\}$  in  $\mathbb{R}^n$  is the closed subset of  $\mathbb{R}^n$  given by

$$\left\{ \sum_{i=0}^k t_i x_i \mid \text{where } t_i \geq 0 \text{ and } \sum_{i=0}^k t_i = 1 \right\}.$$

These points  $\{x_0, \dots, x_k\}$  are said to be *affinely independent* if the collection of vectors

$$\{(x_1 - x_0), (x_2 - x_0), \dots, (x_k - x_0)\}$$

is linearly independent. There can, therefore, be at most  $(n + 1)$  affinely independent points in  $\mathbb{R}^n$ ; the canonical example of such a set has  $x_0$  as the origin while  $x_i$  for  $0 < i \leq n$  is the standard basis vector with 1 in the  $i$ -th coordinate and zeros elsewhere.

**DEFINITION 1.7.** Let  $\phi : K_0 \rightarrow \mathbb{R}^n$  be any function that sends the vertices of  $K$  to points in  $\mathbb{R}^n$ . The **geometric realization** of  $K$  with respect to  $\phi$  is the union

$$|K|_\phi = \bigcup_{\sigma \in K} |\sigma|_\phi,$$

where for each  $\sigma = \{v_0, \dots, v_k\}$  in  $K$ , the set  $|\sigma|_\phi \subset \mathbb{R}^n$  is the geometric simplex spanned by the points  $\{\phi(v_0), \dots, \phi(v_k)\}$ .

If we use a particularly degenerate  $\phi : K_0 \rightarrow \mathbb{R}^n$ , such as the map sending every vertex to the origin, then the topological space  $|K|_\phi \subset \mathbb{R}^n$  might be quite uninteresting and bear no resemblance with  $K$ . We call  $\phi : K_0 \rightarrow \mathbb{R}^n$  an **affine embedding** of  $K$  in  $\mathbb{R}^n$  if  $\phi$  is injective (i.e., it sends different vertices to different points) and if its image  $\phi(K_0)$  is affinely independent. It turns out that the topology of  $|K|_\phi$  is independent of the choice of  $\phi$  provided that we stay within the realm of affine embeddings.

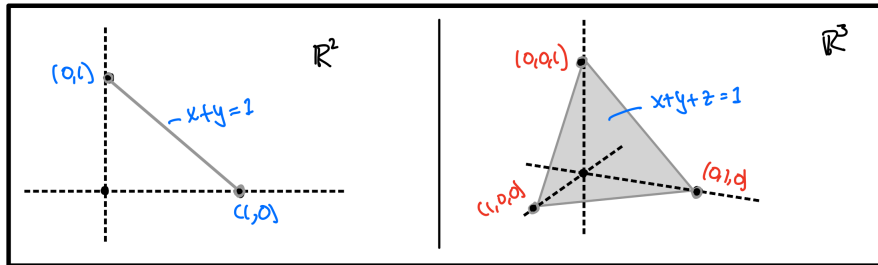
**PROPOSITION 1.8.** For any two affine embeddings  $\phi, \psi : K_0 \rightarrow \mathbb{R}^n$ , there is a homeomorphism  $|K|_\phi \simeq |K|_\psi$  between the corresponding geometric realizations.

**PROOF.** Let  $K_0 = \{v_0, \dots, v_k\}$  be the vertex set of  $K$ ; for each  $i$  in  $\{1, \dots, k\}$  define the following sets of vectors in  $\mathbb{R}^n$

$$x_i = \phi(v_i) - \phi(v_0) \quad \text{and} \quad y_i = \psi(v_i) - \psi(v_0).$$

Since the vectors  $\{x_i\}$  and  $\{y_i\}$  are linearly independent by our assumption on  $\phi$  and  $\psi$ , they each span (possibly distinct)  $k$ -dimensional subspaces of  $\mathbb{R}^n$ . Thus, there is an invertible  $n \times n$  matrix  $M$  sending  $x_i$  to  $y_i$  for each  $i$ , and this  $M$  maps  $|K|_\phi$  to  $|K|_\psi$  homeomorphically.  $\square$

In light of the preceding result, we will usually write the geometric realization of a simplicial complex  $K$  as  $|K|$ , and omit any mention whatsoever of the affine embedding  $\phi$ . It is often convenient to use the endpoints of standard basis vectors in  $\mathbb{R}^n$  as targets of the vertices — this ensures, for instance, that every simplicial complex  $K$  has a geometric realization embeddable in  $\mathbb{R}^n$  for  $n = \#K_0$ . The figure below depicts the geometric realizations of the solid simplices  $\Delta(1)$  and  $\Delta(2)$  with respect to this standard basis embedding.



The geometric realization  $|\Delta(k)|$  is homeomorphic to a  $k$ -dimensional disk while the realization of  $\partial\Delta(k)$  is homeomorphic to the  $(k-1)$ -dimensional sphere. Geometric realizations allow us to look beyond the combinatorial aspects of simplicial complexes and seek structure in the geometry and topology of their realizations. They also provide a rigorous justification for depicting simplices of dimension  $0, 1, 2, 3, \dots$  as points, lines, triangles, tetrahedra, and so forth.

## 1.4 SIMPLICIAL MAPS

Let  $K$  and  $L$  be simplicial complexes.

**DEFINITION 1.9.** A **simplicial map**  $f : K \rightarrow L$  is an assignment  $K_0 \rightarrow L_0$  of vertices to vertices which sends simplices to simplices. So for each simplex  $\sigma = \{v_0, \dots, v_k\}$  of  $K$ , the image  $f(\sigma) = \{f(v_0), \dots, f(v_k)\}$  must be a simplex of  $L$ .

It is important to note that  $f$  as defined above may not be injective, so in general we allow  $f(v_i) = f(v_j)$  even when  $v_i \neq v_j$ . Thus, we only have an inequality  $\dim f(\sigma) \leq \dim \sigma$ .

**EXAMPLE 1.10.** Whenever  $L \subset K$  is a subcomplex, the **inclusion map**  $K \hookrightarrow L$  sends each simplex of  $L$  to the same simplex in  $K$ . In the special case  $L = K$ , this inclusion is called the **identity map** of  $K$ . All such inclusion maps are injective by definition. At the other end of the spectrum, there is a unique surjective simplicial map  $K \twoheadrightarrow \bullet$ , where  $\bullet$  denotes the trivial simplicial complex with only one vertex — so every simplex of  $K$  is sent to this single vertex!

One can compose simplicial maps in a straightforward way — given  $f : K \rightarrow L$  and  $g : L \rightarrow M$ , the composite  $g \circ f : K \rightarrow M$  sends each simplex  $\sigma$  of  $K$  to the simplex  $g(f(\sigma))$  of  $M$ . We call the simplicial map  $f : K \rightarrow L$  an **isomorphism** if there exists an inverse, i.e., a simplicial map  $g : L \rightarrow K$  so that the composites  $g \circ f$  and  $f \circ g$  are the identity maps of  $K$  and  $L$  respectively. Simplicial maps induce honest continuous maps between geometric realizations, which behave as well as one might expect, as described in the following result.

**PROPOSITION 1.11.** For any simplicial map  $f : K \rightarrow L$ ,

- (1) there is an induced continuous function  $|f| : |K| \rightarrow |L|$  between geometric realizations so that for each simplex  $\sigma$  in  $K$ , the geometric simplex  $|f(\sigma)| \subset |L|$  is exactly the image under  $|f|$  of the geometric simplex  $|\sigma| \subset |K|$ ; and moreover,
- (2) if we have a second simplicial map  $g : L \rightarrow M$ , then  $|g \circ f|$  and  $|g| \circ |f|$  coincide as continuous maps  $|K| \rightarrow |M|$ .

The proof of both statements is a reasonable exercise once we explain how to construct  $|f|$  from  $f$ . Let  $\phi : K_0 \rightarrow \mathbb{R}^m$  and  $\psi : L_0 \rightarrow \mathbb{R}^n$  be any affine embeddings. Now each point  $x$  in  $|K| = |K|_\phi$  can be uniquely written as a linear combination  $x = \sum_i t_i \cdot \phi(v_i)$  where  $v_i$  ranges over all the vertices of  $K$  and the  $t_i$  are non-negative real numbers satisfying  $\sum_i t_i = 1$ . The image  $|f|(x)$  of this point in  $|L| = |L|_\psi$  is then given by the formula

$$|f|(x) = \sum_i t_i \cdot \psi \circ f(v_i). \tag{1}$$

If you restrict this map to the realization of a single simplex  $|\sigma|_\phi \subset |K|_\phi$ , you will discover that  $|f|$  is an honest linear map onto the realization of the image simplex  $|f(\sigma)|_\psi \subset |L|_\psi$ . For this reason, such continuous maps are called **piecewise-linear**, and their study forms a rich subject in its own right.

One natural question that you might ask is when two simplicial complexes  $K$  and  $L$  produce homeomorphic geometric realizations  $|K|$  and  $|L|$ . It is a consequence of Proposition 1.8 that any simplicial isomorphism  $f : K \rightarrow L$  induces a homeomorphism  $|f|$  between  $|K|$  and  $|L|$  — but in general  $|K|$  and  $|L|$  can be homeomorphic even if there is no simplicial isomorphism relating  $K$  to  $L$ . We will describe examples of this phenomenon in the next section.

## 1.5 BARYCENTRIC SUBDIVISION

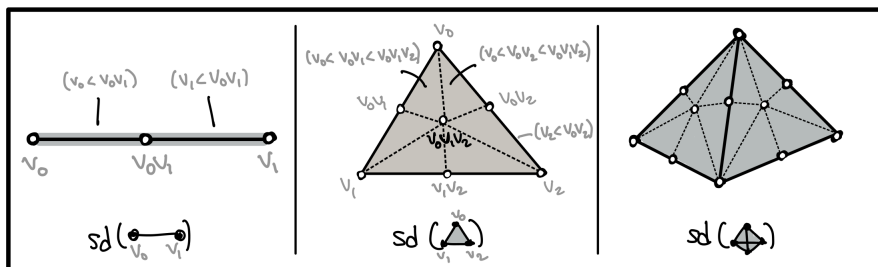
Let  $K$  be a simplicial complex.

**DEFINITION 1.12.** The **barycentric subdivision** of  $K$  is a new simplicial complex  $\mathbf{Sd} K$  defined as follows; for each dimension  $i \geq 0$ , the  $i$ -dimensional simplices are given by all sequences

$$\sigma_0 < \sigma_1 < \dots < \sigma_{i-1} < \sigma_i$$

of (distinct) simplices in  $K$  ordered by the face relation.

This definition is liable to cause confusion until we see what barycentric subdivision looks like geometrically. The figures below depict (some) barycentric simplices within the geometric realizations of the solid simplices  $\Delta(1)$  and  $\Delta(2)$  as well as the hollow 3-simplex  $\partial\Delta(3)$ .



In light of these figures, it is clear that the geometric realizations  $|K|$  and  $|\mathbf{Sd} K|$  agree for every simplicial complex  $K$ ; we record this not-too-surprising fact below.

**PROPOSITION 1.13.** For any simplicial complex  $K$ , there is a homeomorphism between geometric realizations  $|K|$  and  $|\mathbf{Sd} K|$ .

You can check just by counting simplices across various dimensions that for non-trivial  $K$  there can be no simplicial isomorphism  $K \rightarrow \mathbf{Sd} K$ . Since  $\mathbf{Sd} K$  is itself a simplicial complex, it can be further barycentrically subdivided. We refer to this *second* barycentric subdivision as  $\mathbf{Sd}^2 K = \mathbf{Sd}(\mathbf{Sd} K)$ , and similarly define  $\mathbf{Sd}^n K$  for all larger  $n$ . By Proposition 1.13, all the geometric realizations  $|\mathbf{Sd}^n K|$  are homeomorphic regardless of  $n \geq 1$ , even though there are no simplicial isomorphisms which induce these homeomorphisms.

## 1.6 FILTRATIONS FROM DATA

By **data** here we mean a finite set of observations with a well-defined notion of *pairwise distance*, with the typical example being a finite collection of points in  $\mathbb{R}^n$  equipped with the standard Euclidean distance. But in general such observations might not come with any embedding into Euclidean space. One common example is furnished by *dissimilarity matrices* — given a set of observations  $O_1, \dots, O_k$ , one can often build a  $k \times k$  symmetric matrix whose entry in the  $(i, j)$ -th position measures the difference between  $O_i$  and  $O_j$ . Here is a convenient mathematical framework which encompasses all notions of datasets that are relevant to us here.

DEFINITION 1.14. A **metric space**  $(M, d)$  is a pair consisting of a set  $A$  and a function

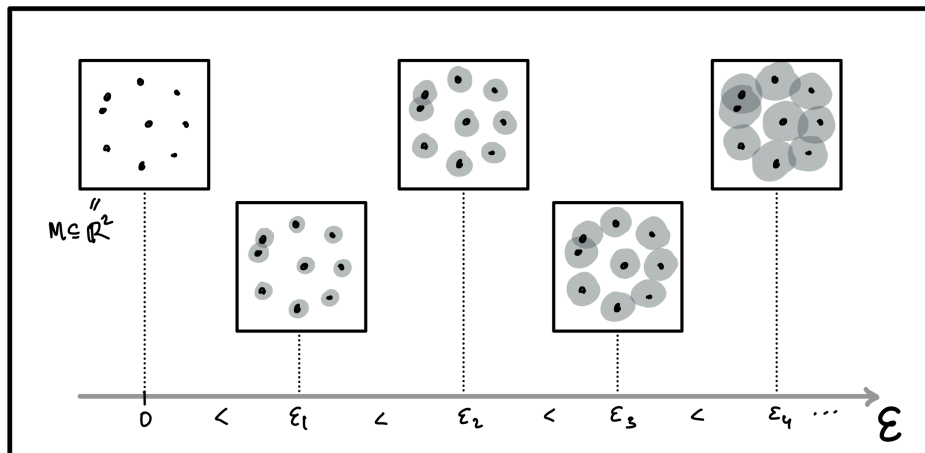
$$d : M \times M \rightarrow \mathbb{R},$$

called the *metric*, satisfying four properties:

- (1) **identity:**  $d(x, x) = 0$  for each  $x$  in  $M$ ,
- (2) **positivity:**  $d(x, y) > 0$  for each  $x \neq y$  in  $M$ ,
- (3) **symmetry:**  $d(x, y) = d(y, x)$  for all  $x, y$  in  $M$ , and most importantly,
- (4) **triangle inequality:**  $d(x, y) + d(y, z) \geq d(x, z)$  for all  $x, y, z$  in  $M$ .

When the metric is clear from context, we will denote the metric space simply by  $M$ ; this happens, for instance, when  $M$  is a subset of some Euclidean space  $\mathbb{R}^n$ . In this case,  $d(x, y)$  is understood to be the Euclidean distance  $\|x - y\|$  for all  $x$  and  $y$  in  $M$ . In fact, any subset  $A \subset M$  of an ambient metric space  $(M, d)$  is automatically given the structure of a metric space in its own right, since we can simply restrict  $d$  to  $A \times A$ .

One fundamental idea behind topological data analysis is best viewed by considering the special case where  $M$  is a finite collection of points in the Euclidean space  $\mathbb{R}^n$ . For such *point clouds*, there is a well-defined notion of *thickening* by any scale  $\epsilon > 0$  — namely,  $M^{+\epsilon}$  is the union of  $\epsilon$ -balls in  $\mathbb{R}^n$  around the points of  $M$ . Various thickenings are illustrated in the figure below.



The hope is to better understand the geometric structure of  $M$  across various scales. One obstacle in this quest is that the union of balls  $M^{+\epsilon}$  is a remarkably inconvenient object from the perspective of designing algorithms — for instance, if you were given a set of points  $M \subset \mathbb{R}^2$  and a scale  $\epsilon > 0$ , how would you program a computer to determine whether or not  $M^{+\epsilon}$  was connected? To address such questions, one replaces unions of balls by filtrations of simplicial complexes (which we encountered in Definition 1.6). There are two common choices of filtrations — **Vietoris-Rips** and **Čech**.<sup>1</sup>

**DEFINITION 1.15.** Let  $(M, d)$  be a finite metric space. The **Vietoris-Rips filtration** of  $M$  is an increasing family of simplicial complexes  $\mathbf{VR}_\epsilon(M)$  indexed by the real numbers  $\epsilon \geq 0$ , defined as follows:

a subset  $\{x_0, x_1, \dots, x_k\} \subset M$  forms a  $k$ -dimensional simplex in  $\mathbf{VR}_\epsilon(M)$  if and only if the pairwise distances satisfy  $d(x_i, x_j) \leq \epsilon$  for all  $i, j$ .

The astute reader may have noticed that we are indexing the simplicial complexes in this filtration by real numbers  $\epsilon \geq 0$  rather than finite subsets of the form  $\{1, 2, \dots, n\}$  as demanded by Definition 1.6. The disparity between the two scenarios is artificial — since we have assumed that  $M$  is finite, there are only finitely many pairwise distances  $d(x, y)$  encountered among the elements of  $M$ , so there are only finitely many  $\epsilon$  values where new simplices are added to  $\mathbf{VR}_\epsilon(M)$ . Those who have not met Vietoris-Rips filtrations before can get better acquainted by verifying the following facts:

- (1) the set  $\mathbf{VR}_\epsilon(M)$  is a simplicial complex for each  $\epsilon > 0$ ,
- (2) the elements of  $M$  are vertices of each such  $\mathbf{VR}_\epsilon(M)$ , and
- (3) for any pair  $0 \leq \epsilon \leq \epsilon'$  of real numbers,  $\mathbf{VR}_\epsilon(M)$  is a subcomplex of  $\mathbf{VR}_{\epsilon'}(M)$ .

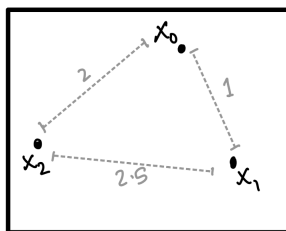
We will see an example of a Vietoris-Rips filtration shortly; first let us examine the Čech alternative.

**DEFINITION 1.16.** Let  $M$  be a finite subset of a metric space  $(Z, d)$ . The **Čech filtration** of  $M$  with respect to  $Z$  is the increasing family of simplicial complexes  $\mathbf{C}_\epsilon$  indexed by  $\epsilon \geq 0$  defined:

a subset  $\{x_0, x_1, \dots, x_k\} \subset M$  forms a  $k$ -dimensional simplex in  $\mathbf{C}_\epsilon(M)$  if and only if there exists some  $z$  in  $Z$  satisfying  $d(z, x_i) \leq \epsilon$  for all  $i$ .

Although the larger metric space  $Z$  plays a starring role in deciding when a simplex lies inside  $\mathbf{C}_\epsilon(M)$ , it is customary to suppress it from the notation (in any case the typical scenario is  $Z = \mathbb{R}^n$  with the Euclidean metric). This blatant dependence on  $Z$  is the biggest immediate difference between Čech filtrations and Vietoris-Rips filtrations — the Vietoris Rips filtration can be defined directly from knowledge of the metric on  $M$  whereas the Čech filtration can not.

To examine the key differences between these two filtrations, consider the three-element metric space  $(M, d)$  illustrated below.



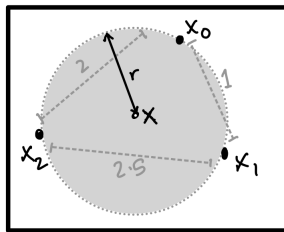
<sup>1</sup>This is pronounced “check”.



The Vietoris Rips filtration of  $M$  at all scales  $\epsilon \geq 0$  is given by the following lists of simplices:

$$\mathbf{VR}_\epsilon(M) = \begin{cases} \{x_0, x_1, x_2\} & 0 \leq \epsilon < 1 \\ \{x_0, x_1, x_2, x_0x_1\} & 1 \leq \epsilon < 2 \\ \{x_0, x_1, x_2, x_0x_1, x_0x_2\} & 2 \leq \epsilon < 2.5 \\ \{x_0, x_1, x_2, x_0x_1, x_0x_2, x_1x_2, x_0x_1x_2\} & \epsilon \geq 2.5 \end{cases}$$

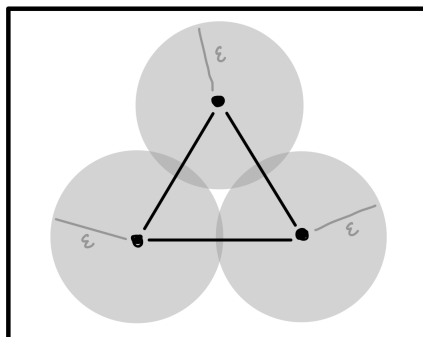
It is crucial to note that the edge  $x_0x_2$  and the 2-simplex  $x_0x_1x_2$  enter the filtration at exactly the same scale, i.e.,  $\epsilon = 2.5$ . Let us now contrast this with the Čech filtration for the same  $M$ , but now viewed as a subset of three points in the Euclidean plane  $\mathbb{R}^2$ . Here, the edge  $x_0x_2$  and the 2-simplex  $x_0x_1x_2$  will not appear simultaneously. Let  $r > 0$  be the radius of the smallest ball which encloses all three points, like so:



The Čech filtration of  $M$  as a subset of  $\mathbb{R}^2$  is given by

$$\mathbf{C}_\epsilon(M) = \begin{cases} \{x_0, x_1, x_2\} & 0 \leq \epsilon < 0.5 \\ \{x_0, x_1, x_2, x_0x_1\} & 0.5 \leq \epsilon < 1 \\ \{x_0, x_1, x_2, x_0x_1, x_0x_2\} & 1 \leq \epsilon < 1.25 \\ \{x_0, x_1, x_2, x_0x_1, x_0x_2, x_1x_2\} & 1.25 \leq \epsilon < r \\ \{x_0, x_1, x_2, x_0x_1, x_0x_2, x_1x_2, x_0x_1x_2\} & \epsilon \geq r \end{cases}$$

Determining the radii of smallest enclosing balls (such as  $r$  above) is quite challenging algorithmically, which is why Vietoris-Rips filtrations are substantially easier to compute. On the other hand, the advantage of the Čech filtration is that it happens to be far more faithful to the underlying geometry of the union of balls  $M^{+\epsilon}$  which we sought to approximate in the first place. For instance, given the union of  $\epsilon$ -balls shown below, the Vietoris-Rips complex at scale  $2\epsilon$  is the solid 2-simplex (which fails to detect the hole) whereas the Čech filtration at scale  $\epsilon$  equals the far more appropriate hollow 2-simplex.



We will study this phenomenon much more carefully in the next Chapter.

## 1.7 BONUS: LOCAL GEOMETRY

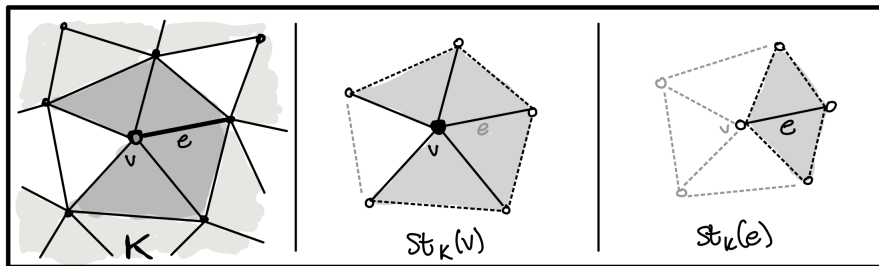
The three notions introduced in this section (stars, links and cones) appear in the exercises of this Chapter and are invoked frequently in subsequent Chapters; but Theorem 1.20 below is not used anywhere else in this text.

Throughout this section, we fix a simplicial complex  $K$  as in Definition 1.1; our goal here is to describe the *neighborhood* of a given simplex  $\sigma$  in (the geometric realization of)  $K$ . The first step is to identify all the simplices which admit  $\sigma$  as a face.

DEFINITION 1.17. The **open star** of  $\sigma$  in  $K$  is the collection of simplices

$$\mathbf{st}_K(\sigma) = \{\tau \text{ in } K \mid \sigma \leq \tau\}.$$

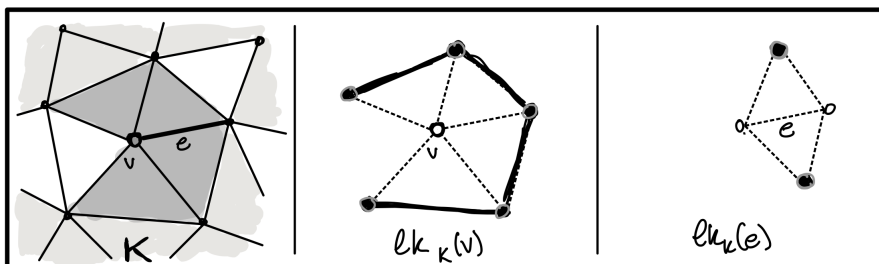
When the ambient simplicial complex  $K$  is clear from context (as it should be here), we simply denote the open star of each simplex  $\sigma$  by  $\mathbf{st}(\sigma)$  rather than dragging  $K$  around in the subscript. The first panel below depicts (a part of) the geometric realization of a 2-dimensional simplicial complex; the open stars of the highlighted vertex  $v$  and edge  $e$  are shown in the next two panels (hollow vertices and dashed edges are *not* included).



Clearly, the open star of  $\sigma$  describes a small simplicial neighborhood of  $\sigma$  in the geometric realization of  $K$ . Since  $\mathbf{st}(\sigma)$  always contains  $\sigma$ , it is guaranteed to be non-empty — but as visible even in the simple examples drawn above, open stars are rarely subcomplexes of  $K$  since they tend to contain simplices without containing all of their faces. Passing to the closure of  $\mathbf{st}(\sigma)$  as described in Definition 1.5 produces a bona fide subcomplex  $\overline{\mathbf{st}}(\sigma) \subset K$ , called the **closed star** of  $\sigma$ . Another useful subset of  $K$  that describes the local geometry of  $\sigma$  is called the **link**.

DEFINITION 1.18. The **link** of  $\sigma$  in  $K$  is the collection  $\mathbf{lk}_K(\sigma)$  of all simplices  $\tau$  in  $K$  which simultaneously satisfy both  $\tau \cup \sigma \in K$  and  $\tau \cap \sigma = \emptyset$ .

Unlike open stars, links of simplices in  $K$  might be empty (for example, the link of a top-dimensional simplex is always empty). But *if* the link of  $\sigma$  is non-empty, then it must be a subcomplex of  $K$ . Here are the links of the vertex  $v$  and edge  $e$  whose open stars we examined in the previous figure.



The final piece of the puzzle is the notion of a **cone** over a simplicial complex.

**DEFINITION 1.19.** The **cone** over  $K$  is a simplicial complex  $\text{Cone}(K)$  defined on the vertex set  $K_0 \cup v_*$ , where  $v_*$  is a new vertex not already present in  $K_0$ . For  $d > 0$ , a  $d$ -simplex of  $\text{Cone}(K)$  is either a  $d$ -simplex of  $K$  itself, or it is  $v_*$  adjoined with a  $(d - 1)$ -simplex of  $K$ .

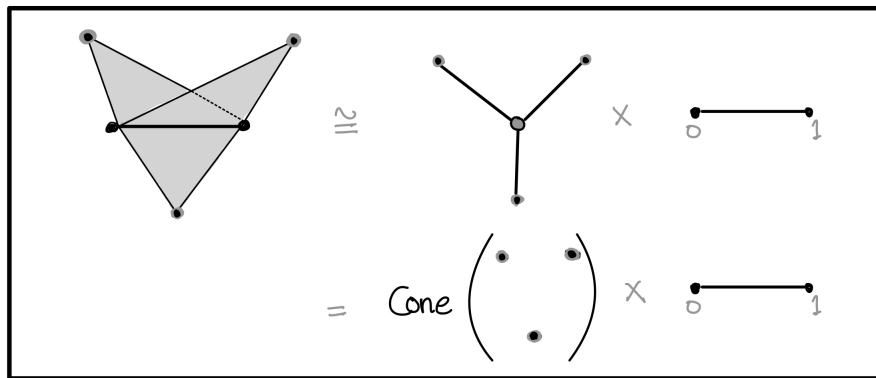
The reason this is called a cone becomes evident if one tries to draw the geometric realization of  $\text{Cone}(K)$  when  $K$ 's geometric realization is homeomorphic to a circle — you are strongly encouraged to try drawing the cone over a hollow 2-simplex. Finally, here is the result which describes neighbourhoods of simplices in every simplicial complex.

**THEOREM 1.20.** For any simplex  $\sigma$  in a simplicial complex  $K$ , there is a homeomorphism

$$|\overline{\text{st}}_K(\sigma)| \simeq |\text{Cone}(\mathbf{lk}_K(\sigma))| \times [0, 1]^{\dim \sigma}.$$

The left side here is the geometric realization of  $\sigma$ 's closed star in  $K$  while the right side is a product of the geometric realization of  $\sigma$ 's link with the closed unit cube in  $\mathbb{R}^{\dim \sigma}$ .

Thus, the smallest simplicial neighborhood around  $\sigma$  in  $K$  (i.e., the closed star of  $\sigma$ ) always decomposes into a product of the cone over the link of  $\sigma$  with Euclidean space of dimension  $\dim \sigma$ . Here is an illustration of this product structure in the special case where  $\sigma$  is a 1-simplex that happens to be a face of three 2-simplices.



## EXERCISES

**EXERCISE 1.1.** For each pair  $i \leq k$  of non-negative integers, how many faces of codimension  $i$  does the solid  $k$ -simplex  $\Delta(k)$  have?

**EXERCISE 1.2.** Show that the face relations between simplices in a finite simplicial complex satisfy the axioms of a *partially ordered set*.

**EXERCISE 1.3.** Show that the set of all subcomplexes of a finite simplicial complex  $K$  satisfy the axioms of a partially ordered set when ordered by containment  $L \subset L'$ .

**EXERCISE 1.4.** Either prove the following, or find a counterexample: if  $K$  is a simplicial complex and  $L \subset K$  a subcomplex with  $L \neq K$ , then the complement  $K - L$  is also a subcomplex of  $K$ .

**EXERCISE 1.5.** Let  $K$  be a  $k$ -dimensional simplicial complex, and for each dimension  $i$  in  $\{0, 1, \dots, k\}$  let  $n_i$  be the number of  $i$ -simplices in  $K$ . How many  $i$ -simplices does the barycentric subdivision  $\text{Sd } K$  have for each dimension  $i$ ?

EXERCISE 1.6. Let  $M$  be a finite metric subspace of an ambient metric space  $(Z, d)$ . Show, for each  $\epsilon > 0$ , that the associated Čech complex  $\mathbf{C}_\epsilon(M)$  is a subcomplex of the Vietoris-Rips complex  $\mathbf{VR}_{2\epsilon}(M)$ . Then, show that – no matter what  $Z$  we had chosen – this  $\mathbf{VR}_{2\epsilon}(M)$  is itself a subcomplex of  $\mathbf{C}_{2\epsilon}(M)$ .

EXERCISE 1.7. Let  $M$  be a finite subset of points in Euclidean space  $\mathbb{R}^n$  (with its standard metric). As a function of  $n$ , can you find the *smallest*  $\delta$  so that  $\mathbf{VR}_\epsilon(M)$  is always a subcomplex of  $\check{\mathbf{C}}_\delta(M)$ ? [Here the Čech complex has been constructed with respect to the ambient Euclidean space  $\mathbb{R}^n$ ]

EXERCISE 1.8. If  $\sigma$  and  $\tau$  are a pair of simplices in a simplicial complex  $K$  satisfying  $\sigma \leq \tau$ , show that  $\mathbf{st}(\sigma) \supset \mathbf{st}(\tau)$  and  $\mathbf{lk}(\sigma) \supset \mathbf{lk}(\tau)$ .

EXERCISE 1.9. Show that if the link  $\mathbf{lk}(\sigma)$  of a simplex  $\sigma$  in a simplicial complex  $K$  is non-empty, then  $\mathbf{lk}(\sigma)$  is a subcomplex of  $K$ .

EXERCISE 1.10. Let  $\sigma$  be a simplex in a simplicial complex  $K$ . Show that a simplex  $\tau$  lies in  $\mathbf{lk}_K(\sigma)$  if and only if the following condition holds: the open stars of  $\sigma$  and  $\tau$  have a non-trivial intersection *and*  $\sigma$  and  $\tau$  have no common faces.