9. Applications of Homology

In this last chapter on the singular homology of topological spaces we now want to study some interesting applications of our theory. Let us start with a few classical applications, some of which are just higher-dimensional generalizations of statements that can be obtained using the fundamental group [G3, Chapter 9], and some are new. As a first very simple statement, we will show that topology is able to distinguish dimensions of real vector spaces in the following sense.

Proposition 9.1 (**Invariance of dimension**, see [G3, Example 8.18])**.** *$\mathbb{R}^m$ is not homeomorphic to $\mathbb{R}^n$ if $m,n \in \mathbb{N}$ with $m \neq n$.*

Proof. Let $f\colon \mathbb{R}^m \to \mathbb{R}^n$ be a homeomorphism with $m,n \in \mathbb{N}$. Removing one point, we obtain from this a restriction $f|_{\mathbb{R}^m\setminus\{0\}}\colon \mathbb{R}^m\setminus\{0\} \to \mathbb{R}^n\setminus\{f(0)\}$ that must be a homeomorphism as well. Hence the source and target must have the same Betti numbers, which by Example 8.2 (c) requires $m = n$. □

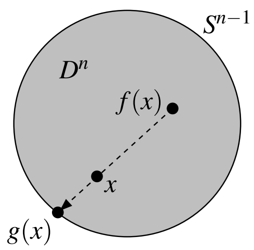
Remark 9.2. Note that Proposition 9.1 is local in the sense that the same statement also holds for open balls: As any open ball in $\mathbb{R}^n$ is homeomorphic to $\mathbb{R}^n$ itself, two open balls in finite-dimensional real vector spaces will be homeomorphic if and only if these vector spaces have the same dimension.

The following fixed point theorem can also be proven in low dimension using the fundamental group already. In fact, the proof given below is almost the same, just replacing the fundamental group with the appropriate homology group.

Proposition 9.3 (**Brouwer's Fixed Point Theorem**, see [G3, Proposition 9.1])**.** *Every continuous map $f\colon D^n \to D^n$ for $n \in \mathbb{N}$ has a fixed point.*

Proof. Assume that $f\colon D^n \to D^n$ is a continuous map without a fixed point.

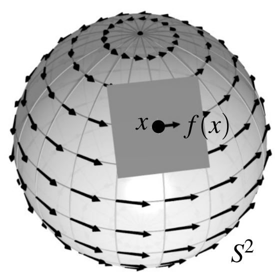
As in the picture on the right, we can then define a map $g\colon D^n \to S^{n-1}$ by sending a point $x \in D^n$ to the unique intersection point of the boundary sphere S^{n-1} with the ray from $f(x)$ that passes through x. This is well-defined since $f(x)$ and x are not the same point, and of course the point $g(x)$ could be computed explicitly to see that g is actually a continuous map. Moreover, as D^n is convex we can move every point x on a straight line to its image point $g(x)$ in D^n, showing since $g|_{S^{n-1}} = \mathrm{id}$ that g is actually a deformation retraction from D^n to S^{n-1}.



But this would imply by Corollary 6.13 that D^n has the same Betti numbers as S^{n-1}, which is not the case by Example 8.2 (b) as D^n is contractible. This contradiction shows that f cannot exist, i. e. that every continuous map from a disk to itself must have a fixed point. □

Example 9.4. Of course, the same fixed point theorem then holds for any topological space homeomorphic to D^n for some n. A common reformulation of Proposition 9.3 in non-mathematical terms is then that any geographic map of a city (which is assumed to be homeomorphic to D^2 here) placed anywhere in this city contains at least one point that lies at the exact point that it represents.

Next, let us consider a construction that we will study in much more detail later on when discussing analysis, namely that of so-called *tangent vector fields* (see Construction 12.13). Currently, we will only look at this situation for a (unit) sphere $S^n \subset \mathbb{R}^{n+1}$ as in the picture on the right, when a tangent vector field is just a continuous map $f\colon S^n \to \mathbb{R}^{n+1}$ such that for all $x \in S^n$ the vector $f(x)$ is orthogonal to x (in the Euclidean norm), i. e. intuitively that $f(x)$ is tangent to the sphere at x We will show now that for even n such a vector field must always have a zero – such as the north and south pole in the picture.



Proposition 9.5 (**Hairy Ball Theorem**). *If $n \in \mathbb{N}$ is even, every continuous map $f\colon S^n \to \mathbb{R}^{n+1}$ with*
13 *$f(x) \perp x$ for all $x \in S^n$ has a zero.*

Proof. Assume that there is a tangent vector field $f\colon S^n \to \mathbb{R}^{n+1}$ without a zero. By normalizing, we may then assume that $\|f(x)\| = 1$ for all $x \in S^n$, i. e. that we have a continuous map $f\colon S^n \to S^n$ with $f(x) \perp x$ for all $x \in S^n$. Now consider the continuous map

$$h\colon S^n \times [0,1] \to S^n,\ x \mapsto \cos(\pi t)\,x + \sin(\pi t)\,f(x)$$

(note that its image actually lies in S^n since $f(x) \perp x$). It is a homotopy between $h(\,\cdot\,,0) = \mathrm{id}$ and $h(\,\cdot\,,1) = -\mathrm{id}$ which moves each point $x \in S^n$ to its opposite point $-x$ along a half-circle in S^n with starting tangent direction $f(x)$. By Proposition 6.12, this means that the map $-\mathrm{id}\colon S^n \to S^n$ induces the identity map $H_n(S^n) \to H_n(S^n)$ on homology, i. e. by Example 8.2 (b) the identity map $K \to K$.

But changing the sign of one coordinate of $S^n \subset \mathbb{R}^{n+1}$ (i. e. a reflection at a hyperplane) induces the negative identity on homology by our simplicial computation of Example 3.13, so the map $-\mathrm{id}\colon S^n \to S^n$ which changes the signs of all $n+1$ coordinates is a multiplication by $(-1)^{n+1}$ in homology. As we have just seen that this map must be the identity on a non-zero vector space, we conclude (by choosing a ground field in which $-1 \neq 1$) that n must be odd. □

Remark 9.6.

(a) The reason for the name “Hairy Ball Theorem“ should readily be visible from the picture above. The statement is also often quoted by saying that a hedgehog cannot be combed without a bald spot.

(b) If $n = 2k+1$ is odd (with $k \in \mathbb{N}$) it is easy to write down a continuous map $f\colon S^n \to \mathbb{R}^{n+1}$ without zeros with $f(x) \perp x$ for all $x \in S^n$, e. g.

$$f\colon S^n \to \mathbb{R}^{n+1},\ (x_1, x_2, \dots, x_{2k+1}, x_{2k+2}) \mapsto (-x_2, x_1, \dots, -x_{2k+2}, x_{2k+1}).$$

Our last classical application of singular homology is a proof of the seemingly obvious statement that a closed non-self-intersecting curve in $\mathbb{R}^2$ divides the plane into exactly two regions (an interior and an exterior). Maybe this assertion seems so clear to you that you would not have expected that it needs a proof at all, but recalling the existence of surface-filling Peano curves might change your mind [G1, Chapter 24.D]. In fact, the proof of this statement is significantly harder than the three previous ones, and clearly depends crucially on what we mean by a “closed non-self-intersecting curve”.

To simplify the following proof a bit, let us actually consider curves in S^2 instead of $\mathbb{R}^2$ – which does not make a difference since we can think of S^2 as $\mathbb{R}^2$ compactified by a “point at infinity“. We will consider a so-called *Jordan curve*, which is just an injective continuous map $f\colon S^1 \to S^2$, and show that the complement of its image $f(S^1)$ has exactly two path-connected components, i. e. that its zeroth Betti number equals 2. Note that this complement is clearly open since $f(S^1)$ is compact and thus closed, and that the requirement of f being injective already rules out many pathological examples: For example, a surjective Peano curve $f\colon S^1 \to S^2$ is no longer possible since it would then be bijective and thus a homeomorphism by [G1, Proposition 24.32], in contradiction to Proposition 9.1 since taking away a point from f would then yield a homeomorphism between $\mathbb{R}$ and $\mathbb{R}^2$.

Proposition 9.7 (**Jordan Curve Theorem**).

(a) *For any injective continuous map $f\colon [0,1] \to S^2$ from the interval $[0,1]$ to the 2-sphere S^2 we have $b_0(S^2 \backslash f([0,1])) = 1$ and $b_1(S^2 \backslash f([0,1])) = 0$.*

(b) *For any injective continuous map $f\colon S^1 \to S^2$ we have $b_0(S^2 \backslash f(S^1)) = 2$.*

Proof.

(a) We have to show that the homology group $H_1(S^2 \backslash f([0,1]))$ is zero, and that the degree map $\deg\colon H_0(S^2 \backslash f([0,1])) \to K$ is an isomorphism. We will prove this by contradiction, so let us assume that for $p = 0$ or $p = 1$ there is a non-zero homology class $c \in H_p(S^2 \backslash f([0,1]))$, with the additional condition $\deg c = 0$ if $p = 0$.

Starting with the interval $I_0 := [0,1]$, we will construct from this a descending sequence of intervals $I_0 \supset I_1 \supset I_2 \supset \cdots$ such that c is non-zero in $H_p(S^2 \backslash f(I_n))$ for all n as well (where we consider c to be a homology class in $S^2 \backslash f(I_n)$ using the pushforward by the inclusion). To do this, assume that I_n is already constructed, and split it into two intervals $I_n = I'_n \cup I''_n$ intersecting at the midpoint P_n. The Mayer-Vietoris sequence of Corollary 7.10 for the open cover of $S^2 \backslash \{f(P_n)\}$ by $S^2 \backslash f(I'_n)$ and $S^2 \backslash f(I''_n)$ then gives as a part the exact sequence

$$H_{p+1}(S^2 \backslash \{f(P_n)\}) \longrightarrow H_p(S^2 \backslash f(I_n)) \longrightarrow H_p(S^2 \backslash f(I'_n)) \times H_p(S^2 \backslash f(I''_n)),$$

of which the left term is zero since $S^2 \backslash \{f(P_n)\}$ is homeomorphic to $\mathbb{R}^2$. Hence the second map in the sequence is injective. As c is non-zero in $H_p(S^2 \backslash f(I_n))$ we can therefore choose $I_{n+1} = I'_n$ or $I_{n+1} = I''_n$ such that c is non-zero in $H_p(S^2 \backslash f(I_{n+1})) > 0$ as well. This completes the construction of the sequence $I_0 \supset I_1 \supset I_2 \supset \cdots$.

But if $P \in [0,1]$ denotes the convergence point of this sequence of intervals, the class c must of course be zero in the homology of $S^2 \backslash \{f(P)\}$ as this space is homeomorphic to $\mathbb{R}^2$ again. In other words, we have $c = \partial c'$ for some $(p+1)$-chain c' in $S^2 \backslash \{f(P)\}$. But the image of all singular simplices in c' in $S^2 \backslash \{f(P)\}$ is a compact set, and this set is covered by the open subsets $S^2 \backslash f(I_n)$ for $n \in \mathbb{N}$. As this is an increasing sequence of open sets, it follows by compactness that c' must be a singular simplex in some $S^2 \backslash f(I_n)$ already. But this means that $c = \partial c'$ is trivial in the homology of $S^2 \backslash f(I_n)$ as well, in contradiction to our construction.

(b) We can obtain the circle S^1 by gluing two intervals I and J at their ends, giving rise to gluing points P and Q. By (a) we have $H_1(S^2 \backslash f(I)) \times H_1(S^2 \backslash f(J)) = 0$, so starting the Mayer-Vietoris sequence for covering $S^2 \backslash f(\{P,Q\})$ with $S^2 \backslash f(I)$ and $S^2 \backslash f(J)$ at this term we obtain the exact sequence

$$0 \to H_1(S^2 \backslash f(\{P,Q\})) \to H_0(S^2 \backslash f(S^1)) \to H_0(S^2 \backslash f(I)) \times H_0(S^2 \backslash f(J)) \to H_0(S^2 \backslash f(\{P,Q\})) \to 0.$$

From this we can now easily obtain the desired result by an Euler characteristic computation: The space $S^2 \backslash f(\{P,Q\})$ is homeomorphic to the punctured plane $\mathbb{R}^2 \backslash \{0\}$ and thus has Betti numbers $b_0(S^2 \backslash f(\{P,Q\})) = b_1(S^2 \backslash f(\{P,Q\})) = 1$ by Example 8.2 (c), and by (a) we have $b_0(S^2 \backslash f(I)) = b_0(S^2 \backslash f(J)) = 1$. Hence, by Remark 4.5 (b) it follows that $1 - b_0(S^2 \backslash f(S^1)) + (1+1) - 1 = 0$, which means that $b_0(S^2 \backslash f(S^1)) = 2$. □

Exercise 9.8. Prove for every $n \in \mathbb{N}_{>0}$:

(a) If $f \colon S^n \to S^n$ is a continuous map with $n \in \mathbb{N}$ even, there is a point $x \in S^n$ with $f(x) = x$ or $f(x) = -x$.

(b) Every real matrix $A \in \mathbb{R}^{n \times n}$ with positive entries has an eigenvector with positive entries.

Exercise 9.9. Let S be a connected graph (i. e. a 1-dimensional simplicial complex) with v vertices and e edges. By an embedding of S into a topological space X we mean a continuous injective map $f \colon |S| \to X$.

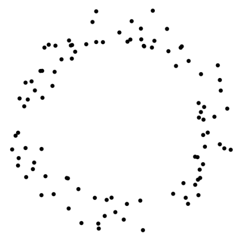
(a) For any embedding of S into the sphere S^2, show that $b_0(S^2 \backslash f(|S|)) = e - v + 2$.

(b) Now let $S = \{\{i,j\} : i,j \in \{1,2,3,4,5\}\}$ be the graph with five vertices and every possible edge between them. Show that f cannot be embedded in S^2, but that it can be embedded in the torus $S^1 \times S^1$.

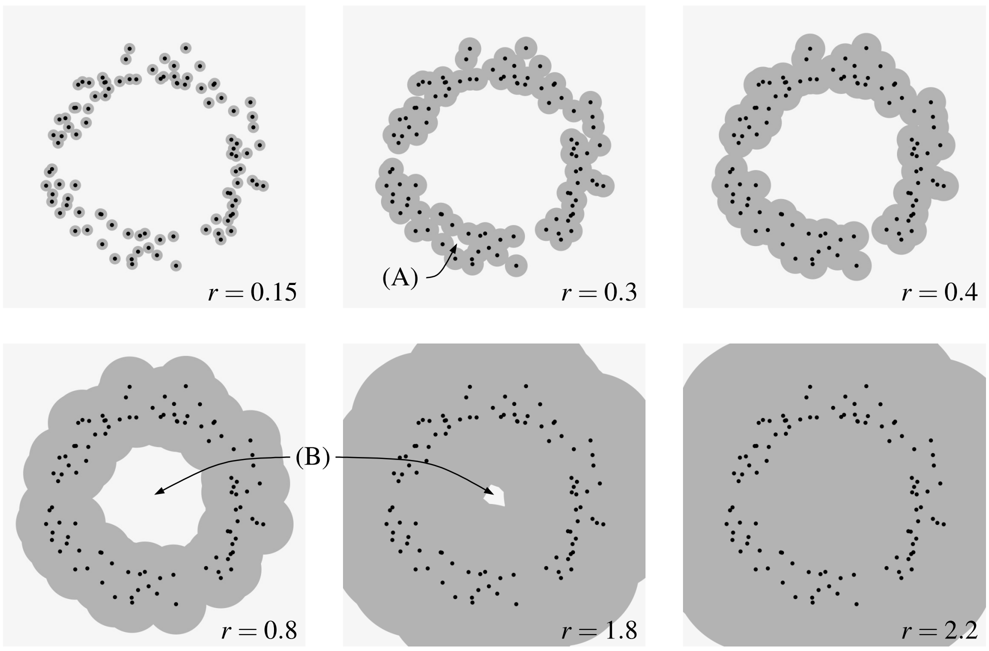
Exercise 9.10. Generalize the proof of Proposition 9.7 to show by induction on $n \in \mathbb{N}_{\geq 2}$ that $b_0(S^n \backslash f(S^{n-1})) = 2$ for any injective continuous map $f \colon S^{n-1} \to S^n$.

To conclude our algebraic study of topology, we now want to discuss a modern application of homology to a branch of mathematics called *topological data analysis*, whose goal is to analyze the geometry of large point cloud data. Let us start with an example explaining the idea behind this technique.

Example 9.11 (Noisy annulus)**.** You would probably agree that the picture on the right shows a noisy image of an annulus. But of course, in reality the picture just consists of a large number of discrete points – in fact, we have generated it by plotting random points in the plane with distance between 2 and 3 from the origin. So what does it mean exactly that the points are arranged in the shape of an annulus, and how can we figure this out? In other words, from an application point of view: If a computer acquired the data of these points, how can it detect the underlying shape?



One possible idea is to replace the points by circles around them with increasing radius r, as shown in the pictures below. Considering the unions of these circles, we obtain subspaces of $\mathbb{R}^2$ of which we can compute and compare the singular homology. As we can see in the pictures, there are sometimes a few spurious holes such as (A) that do not represent an actual feature of the data and vanish again for a slightly bigger radius. In contrast, the main hole (B) persists in a large range of the radius (at least from $r = 0.80$ to $r = 1.80$), indicating that this is an important feature of the image.



To make such statements precise, we have to compare homology groups for different values of the radius. More precisely, as in the picture above we will pick a sequence $r_1 < \cdots < r_k$ of radii, denote by $U^{(i)}$ the union of the circles of radius r_i around the given points (i. e. the gray subspaces of $\mathbb{R}^2$ in the picture), and compute the homology groups $H_p(U^{(i)})$ for all $p \in \mathbb{Z}$. As $U^{(1)} \subset U^{(2)} \subset \cdots \subset U^{(k)}$, the pushforwards by the inclusion maps will induce linear maps

$$H_p(U^{(1)}) \longrightarrow H_p(U^{(2)}) \longrightarrow \cdots \longrightarrow H_p(U^{(k)}),$$

and we want to study how long elements can survive in this sequence. To tackle such questions, let us first consider sequences of vector spaces in general and show that they can always be decomposed uniquely into simple parts, each corresponding to a single element that is created and destroyed at specific stages of the sequence.

Construction 9.12.

(a) For a sequence of vector spaces $C = ((C^{(k)})_{k\in\mathbb{Z}}, (f^{(k)})_{k\in\mathbb{Z}})$, written with ascending upper indices as

$$\cdots \longrightarrow C^{(k-1)} \xrightarrow{f^{(k-1)}} C^{(k)} \xrightarrow{f^{(k)}} C^{(k+1)} \longrightarrow \cdots,$$

we denote by $f^{(k,l)}: C^{(k)} \to C^{(l)}$ for $k,l \in \mathbb{Z}$ with $k \le l$ the composition of the given maps starting at position k and ending at position l, i. e.

$$f^{(k,l)} = f^{(l-1)} \circ \cdots \circ f^{(k+1)} \circ f^{(k)},$$

and set $C^{(k,l)} := \operatorname{im} f^{(k,l)} \subset C^{(l)}$. Note that $f^{(k,k)} = \mathrm{id}_{C^{(k)}}$ and $f^{(k,k+1)} = f^{(k)}$ for all $k \in \mathbb{Z}$; in particular we have $C^{(k,k)} = C^{(k)}$.

(b) For any $i,j \in \mathbb{Z}$ with $i \le j$, the **interval sequence** $I_{i,j}$ is the sequence of vector spaces

$$\cdots \longrightarrow 0 \longrightarrow \underset{\substack{\uparrow \\ i}}{K} \xrightarrow{\mathrm{id}} K \xrightarrow{\mathrm{id}} \cdots \xrightarrow{\mathrm{id}} K \xrightarrow{\mathrm{id}} \underset{\substack{\uparrow \\ j}}{K} \longrightarrow 0 \longrightarrow \cdots$$

corresponding to a single basis element that lives from position i to position j. We call the number $j-i+1$ of its non-zero entries the **length** of the sequence. Clearly, in the notation of (a) we have

$$\dim I_{i,j}^{(k,l)} = \begin{cases} 1 & \text{if } i \le k \le l \le j, \\ 0 & \text{otherwise.} \end{cases}$$

Proposition 9.13 (**Interval decomposition** of sequences)**.** *Let C be a sequence of finite-dimensional vector spaces with only finitely many non-zero entries. Then C is isomorphic to a product of finitely many interval sequences. Moreover, for any $i,j \in \mathbb{Z}$ with $i \le j$ the number of times $I_{i,j}$ occurs in this decomposition is uniquely determined by C.*

Proof. We prove the existence by induction on $\dim C := \sum_{p \in \mathbb{Z}} \dim C^{(p)}$, the case $\dim C = 0$ being trivial. For $\dim C > 0$, let $j \in \mathbb{Z}$ be the minimum such that $f^{(j)}$ is not injective, so that the sequence starts with inclusions $\cdots \subset C^{(j-2)} \subset C^{(j-1)} \subset C^{(j)}$. Pick any non-zero $x \in \ker f^{(j)} \subset C^{(j)}$, and denote by $i \in \mathbb{Z}$ the smallest integer with $x \in C^{(i)}$.

We can then split off $I_{i,j}$ from C: Pick ascending bases $\cdots \subset B^{(j-2)} \subset B^{(j-1)} \subset B^{(j)}$ of the corresponding vector spaces in C, taking x into the bases starting at $B^{(i)}$. If we now set $C'^{(k)}$ to be the vector space with basis $B^{(k)} \backslash \{x\}$ for $k \in \{i,\dots,j\}$ and $C'^{(k)} := C^{(k)}$ otherwise, we obtain a subsequence C' with $C \cong C' \times I_{i,j}$, where all non-zero vector spaces in $I_{i,j}$ are generated by x. As $\dim C' < \dim C$, the existence of a decomposition of C into interval sequences now follows by the induction assumption.

As for uniqueness, assume that we are given a decomposition of C into interval sequences, with $I_{i,j}$ occurring in the decomposition exactly $n_{i,j}$ times for $i \le j$. We will show by descending induction on $j-i$ that $n_{i,j}$ is uniquely determined by C: Clearly, $n_{i,j}$ must be zero for $j-i$ large enough since only finitely many vector spaces in C are non-zero. The induction step follows since by Construction 9.12 (b) we have

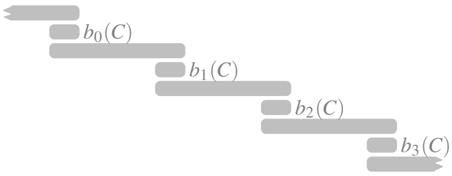
$$\dim C^{(i,j)} = \sum_{k,l:\, k \le i \le j \le l} n_{k,l} = n_{i,j} + \big(\text{terms } n_{k,l} \text{ with } l-k > j-i\big),$$

which determines $n_{i,j}$ in terms of $\dim C^{(i,j)}$ and inductively known terms. □ 14

Example 9.14 (Interval decomposition of chain complexes and exact sequences)**.**

(a) One of the easiest cases of an interval decomposition is obtained for a chain complex. Obviously, an interval sequence is a chain complex if and only if it has length at most 2, and consequently a sequence C of vector spaces satisfying the finiteness conditions of Proposition 9.13 is a chain complex if and only if all interval sequences in its interval decomposition have length at most 2. It is common to draw these interval sequences as actual intervals as indicated in gray in the picture below, resulting in the so-called **barcode** of the sequence. Note that for a general chain complex C each interval in this barcode can occur any number of times (including zero), and that the vertical order of the intervals is chosen arbitrarily.

$$\cdots \longrightarrow C^{(0)} \longrightarrow C^{(1)} \longrightarrow C^{(2)} \longrightarrow C^{(3)} \longrightarrow \cdots$$



(b) Continuing this example, the interval sequences of length 2 are clearly exact, whereas the ones $I_{i,i}$ of length 1 are not and have i-th Betti number equal to 1, with all others being zero. Consequently, the barcode of a chain complex C will contain the interval of length 1 at i exactly $b_i(C)$ times, as shown in the picture above. As a special case, C is an exact sequence if and only if all its Betti numbers are zero, i. e. if and only if its barcode consists only of intervals of length 2. This barcode can then also be interpreted by saying that each vector space $C^{(p)}$ in an exact sequence is made up from parts coming from $C^{(p-1)}$ and $C^{(p+1)}$, i. e. from the immediate left and right – which we had already seen e. g. in Example 4.6.

Remark 9.15 (Decomposition of short exact sequences of chain complexes)**.** In fact, a decomposition similar to Example 9.14 also exists for a short exact sequence $0 \to A \to B \to C \to 0$ of chain complexes, assuming again that all vector spaces in it are finite-dimensional and only finitely many of them are non-zero. This situation is described by a 2-dimensional commutative diagram of vector spaces as in Definition 4.1, with three non-zero columns and an arbitrary number of rows. As in Proposition 9.13, one can show inductively that such a short exact sequence can be decomposed as a product of simple building blocks, in which all non-zero vector spaces are just K and all non-zero maps are the identity. However, these building blocks are now of course no longer intervals, but rather two-dimensional shapes. By the argument of Example 9.14 it is easy to see which shapes are allowed: The rows are exact, hence each allowed shape must have horizontal length exactly 2 in each occurring row. The columns are complexes, hence each allowed shape must have vertical length at most 2 in each column. Moreover, squares as in the picture below on the right are generally forbidden since they violate commutativity. This leaves exactly the five shapes (1) to (5) for the building blocks, which might occur at any vertical position.

$$\begin{array}{ccccc} \vdots & & \vdots & & \vdots \\ K & \to & K & \to & 0 \\ \vdots & & \vdots & & \vdots \end{array} \qquad \begin{array}{ccccc} \vdots & & \vdots & & \vdots \\ 0 & \to & K & \to & K \\ \vdots & & \vdots & & \vdots \end{array} \qquad \begin{array}{ccccc} \vdots & & \vdots & & \vdots \\ K & \to & K & \to & 0 \\ \downarrow & & \downarrow & & \downarrow \\ K & \to & K & \to & 0 \\ \vdots & & \vdots & & \vdots \end{array} \qquad \begin{array}{ccccc} \vdots & & \vdots & & \vdots \\ 0 & \to & K & \to & K \\ \downarrow & & \downarrow & & \downarrow \\ 0 & \to & K & \to & K \\ \vdots & & \vdots & & \vdots \end{array} \qquad \begin{array}{ccccc} \vdots & & \vdots & & \vdots \\ 0 & \to & K & \to & K \\ \downarrow & & \downarrow & & \downarrow \\ K & \to & K & \to & 0 \\ \vdots & & \vdots & & \vdots \end{array} \qquad \begin{array}{ccc} K & \to & K \\ \downarrow & & \downarrow \\ 0 & \to & K \\ \\ K & \to & 0 \\ \downarrow & & \downarrow \\ K & \to & K \end{array}$$

(1) (2) (3) (4) (5)

Note that this gives a visual proof, and thus an intuitive explanation, for the long exact homology sequence of Proposition 4.8: As in Example 9.14, the homology of the complexes sees only the vertical parts of length 1, i. e. the entries marked in dark gray. Hence, the building blocks above give rise to a barcode of the sequence $\cdots \to H_p(A) \to H_p(B) \to H_p(C) \to H_{p-1}(A) \to \cdots$ consisting only of intervals of length 2, i. e. an exact sequence – with shape (5) corresponding to the connecting map $H_p(C) \to H_{p-1}(A)$. The only drawback of this alternative proof besides the finiteness assumption is that the maps in homology constructed in this way are not natural: Whereas the number of times a certain building block occurs in the decomposition is uniquely determined, the actual decomposition is not, corresponding to the choice of bases in the proof of Proposition 9.13.

But let us now apply Proposition 9.13 to sequences of homology groups, as already suggested in Example 9.11. This will lead to longer intervals in our sequences, and it will in fact be the particularly long ones that are important for our applications.

Construction 9.16 (Persistent homology)**.** Let

$$\cdots \longrightarrow C^{(k-1)} \xrightarrow{f^{(k-1)}} C^{(k)} \xrightarrow{f^{(k)}} C^{(k+1)} \longrightarrow \cdots$$

be a sequence of chain complexes as in Definition 4.1. By Lemma 2.17, this gives rise to sequences of vector spaces

$$\cdots \longrightarrow H_p(C^{(k-1)}) \xrightarrow{f_p^{(k-1)}} H_p(C^{(k)}) \xrightarrow{f_p^{(k)}} H_p(C^{(k+1)}) \longrightarrow \cdots \qquad (*)$$

in homology for all $p \in \mathbb{Z}$. As in Construction 9.12 (a), denote by $f_p^{(k,l)} : H_p(C^{(k)}) \to H_p(C^{(l)})$ the composition of these maps from k to l for $k \le l$. Then the images $\operatorname{im} f_p^{(k,l)}$ are called the **persistent homology groups** of the given sequence; their dimensions $b_p^{(k,l)} = \operatorname{rk} f_p^{(k,l)}$ are referred to as its **persistent Betti numbers**. In particular, the number $b_p^{(k,k)} = b_p(C^{(k)})$ is the ordinary p-th Betti number of the chain complex $C^{(k)}$.

Remark 9.17 (Visualization of persistent homology)**.** Consider a sequence of chain complexes as in Construction 9.16, and assume that all sequences $(*)$ of its homology groups for $p \in \mathbb{Z}$ consist of finite-dimensional vector spaces, of which only finitely many are non-zero. According to Proposition 9.13, we can then visualize its persistent homology by the interval decompositions of the homology sequences $(*)$. More precisely, as in the proof of the proposition the persistent Betti number $b_p^{(k,l)}$ (which is by definition the dimension of the space of homology classes surviving at least from position k to position l in the p-th homology sequence) is the same as the number of intervals in the decomposition of the p-th homology sequence containing the interval from k to l. The interval decomposition can then e. g. be drawn as a barcode as in Example 9.14.

In Example 9.11, our idea for persistent homology was to take the singular homology of unions of circles with increasing radius. There is also another way to obtain very similar homology groups; let us introduce both methods now and compare them.

Construction 9.18 (Čech and Vietoris-Rips complex)**.** Let M be a finite subset of $\mathbb{R}^n$ for some $n \in \mathbb{N}$, and let $r \in \mathbb{R}_{>0}$.

(a) The **Čech complex** $\check{C}(M,r)$ is defined to be the nerve of the union $\bigcup_{x \in M} U_r(x) \subset \mathbb{R}^n$ of the open balls with radius r and center points in M. By Definition 8.6, this is just the simplicial complex with vertex set M that contains a subset $\{x_0,\dots,x_p\} \subset M$ as a simplex if and only if $U_r(x_0) \cap \cdots \cap U_r(x_p) \neq \emptyset$.

(b) The **Vietoris-Rips complex** $\mathrm{VR}(M,r)$ is the simplicial complex with vertex set M that contains a subset $\{x_0,\dots,x_p\} \subset M$ as a simplex if and only if $\|x_i - x_j\| < 2r$ for all $i,j \in \{0,\dots,p\}$.

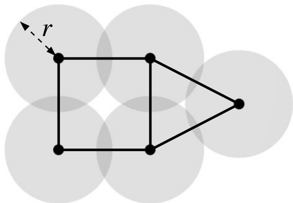
For both constructions, increasing the radius will lead to a bigger simplicial complex, so by picking a finite sequence of increasing radii we can produce a corresponding sequence of simplicial chain complexes (with the morphisms induced by inclusions) and study its persistent homology.

Remark 9.19. Let M be a finite subset of $\mathbb{R}^n$ as above, and let $r \in \mathbb{R}_{>0}$.

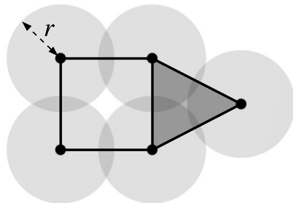
(a) By the Nerve Theorem of Proposition 8.8 (together with Remark 8.9 (a)), the simplicial homology of the Čech complex $\check{C}(M,r)$ as in Construction 9.18 (a) is naturally isomorphic to the singular homology of the union of balls $\bigcup_{x \in M} U_r(x) \subset \mathbb{R}^n$. The picture below on the left shows an example, where M consists of the five black points and it is clearly visible that the simplicial complex and the union of balls have the same Betti numbers (namely 2 in dimension 1 and 1 in dimension 0). Hence, studying the persistent simplicial homology of $\check{C}(M,r)$ for a given sequence of radii is exactly what we have done in Example 9.11.

(b) The construction of the Vietoris-Rips complex follows a similar idea. In fact, the edges of both complexes just agree since the conditions $U_r(x_0) \cap U_r(x_1) \neq \emptyset$ and $\|x_0 - x_1\| < 2r$ are equivalent. Moreover, the Čech complex $\check{C}(M,r)$ is clearly always a subcomplex of $\mathrm{VR}(M,r)$ since $U_r(x_0) \cap \cdots \cap U_r(x_p) \neq \emptyset$ requires $\|x_i - x_j\| < 2r$ for all $i,j \in \{0,\dots,p\}$ by the triangle inequality. However, the Vietoris-Rips complex only checks pairwise distances of points in M and thus cannot detect the full geometry of the union of balls: In the picture below, the three balls on the right have a pairwise non-empty intersection but not a common intersection point, hence the corresponding 2-simplex exists in the Vietoris-Rips complex, but not in the Čech complex. Hence, the simplicial homology of $\mathrm{VR}(M,r)$ does not have an immediate geometric interpretation, which might seem as a disadvantage. However, the Vietoris-Rips complex also has two major advantages:

- It is much faster to compute, as it is easier to check pairwise distances than to compute intersection points of finitely many balls.
- The Vietoris-Rips complex can be defined in exactly the same way for any finite metric space M, without referring to an embedding in $\mathbb{R}^n$.



Čech complex $\check{\mathrm{C}}(M,r)$



Vietoris-Rips complex $\mathrm{VR}(M,r)$

Moreover, from the point of view of persistent homology, it seems in the picture above that the difference between the Čech and Vietoris-Rips complex should not matter much: The hole in the middle of the three balls on the right is very small and will vanish fast when we increase the radius of the balls, hence it will not contribute significantly to the persistent homology. The following lemma and exercise give exact statements along these lines.

Lemma 9.20. *Let M be a finite subset of $\mathbb{R}^n$ for some $n \in \mathbb{N}$, and let $r \in \mathbb{R}_{>0}$. Then*

$$\check{\mathrm{C}}(M,r) \subset \mathrm{VR}(M,r) \subset \check{\mathrm{C}}(M,2r).$$

Proof. We have already observed the first inclusion in Remark 9.19 (b). The second inclusion is equally simple: If $x_0,\dots,x_p \in \mathbb{R}^n$ with $\|x_i - x_j\| < 2r$ for all $i,j \in \{0,\dots,p\}$, every ball $U_{2r}(x_i)$ clearly contains x_0, so $U_{2r}(x_0) \cap \cdots \cap U_{2r}(x_p) \neq \emptyset$. □

Exercise 9.21. Find the smallest constant $c \in \mathbb{R}_{>1}$ such that $\check{\mathrm{C}}(M,r) \subset \mathrm{VR}(M,r) \subset \check{\mathrm{C}}(M,cr)$ for any finite subset M of $\mathbb{R}^2$ and $r \in \mathbb{R}_{>0}$.

Example 9.22 (Noisy annulus revisited)**.** The picture below shows the barcodes for the persistent simplicial homology of the Čech and Vietoris-Rips complex for the noisy annulus of Example 9.11. In contrast to this old example, we have picked many more values of the radius r, ranging from 0 to about 2, to obtain a better resolution. The black bars are the intervals in the decomposition of the homology in dimension 1, and the gray bars correspond to dimension 0. As usual, the vertical order of the intervals in the barcode has been chosen arbitrarily.

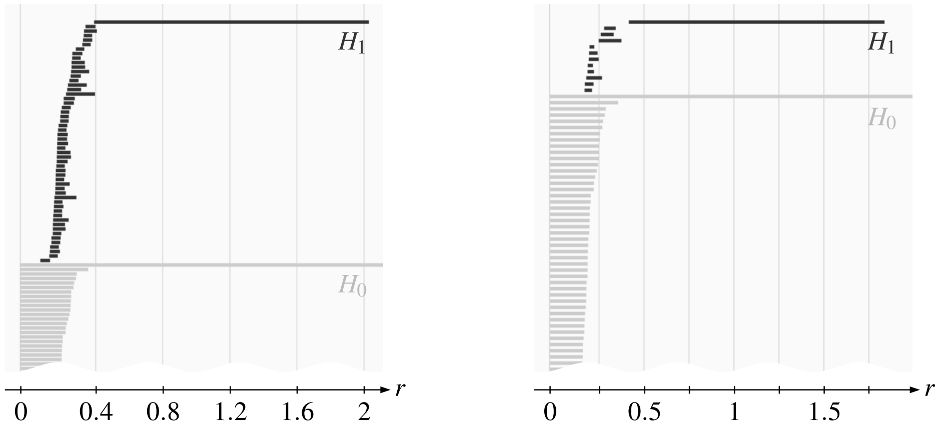
Many features of these barcodes can in fact be explained by the pictures in Example 9.11:

(a) The most important feature of the barcodes is clearly the one very long interval in dimension 0 and 1 each, indicating that the original noisy point cloud looks connected (there is exactly one long-lived homology class in dimension 0) and has one 1-dimensional hole (corresponding to the one long-lived class in dimension 1). This is the main information we are interested in when studying persistent homology.

(b) Looking at the long interval in homology dimension 1 in the Čech barcode more closely, we can see its bounds already from the pictures in Example 9.11 since the 1-dimensional big hole is present from about $r \approx 0.4$ to $r \approx 2$. In homology dimension 0, the $r = 0.3$ picture in Example 9.11 is still disconnected whereas the $r = 0.4$ picture is not – corresponding to the second longest interval in the barcode vanishing between these values.

(c) For small values of r the balls around the given points barely intersect, leading to very many connected components. In the barcodes, this corresponds to many intervals in homology dimension 0 vanishing very fast. In fact, we have cut off several of these intervals in the picture below as they do not contain relevant information.

(d) The Vietoris-Rips barcode has far fewer short-lived intervals in homology dimension 1 than the Čech barcode. This is due to the phenomenon as in the picture of Remark 9.19: Many spurious 1-dimensional holes in the Čech complex are simplicial 1-spheres coming from three balls that intersect pairwise but do not have a common intersection point. By construction, such holes can never appear in the Vietoris-Rips complex as they are always filled in by a simplicial 2-disk. Hence, there are no corresponding intervals in the Vietoris-Rips barcode.



Čech barcode

Vietoris-Rips barcode