Topology
A chapter for the Mathematics++ Lecture Notes

Jiří Matoušek

Rev. 2/XII/2014 JM

Topology has spectacular applications in discrete mathematics and computer science, such as in lower bounds for the chromatic number of graphs (which will be discussed later to some extent), in results about the behavior of distributed computing systems (see Herlihy, Kozlov, and Rajsbaum [HKR13]), or in methods for reconstructing 3-dimensional shapes from point samples, whose importance increases with the advent of ubiquitous 3D printing.

Yet the entrance barriers of topology are relatively high, according to the author's experience. This has to do with the extent, maturity, and technical sophistication of the field. At the very beginning of serious study, a newcomer is confronted with new language and conventions, such as commutative diagrams, exact sequences, and categorical concepts. At the same time, in order to honestly reach the first real results, one also has to work through a number of technicalities such as approximations of continuous maps. These things can be experienced once and then more or less forgotten, yet skipped they should not be. Last but not least, some of the fundamental concepts are truly sophisticated.

The notion of homology seems to be a particularly high stumbling block. Many computer scientists with some topological background switch off when a homology or cohomology group appears on the board. In this chapter we thus aim at an introduction with as few technicalities as possible reaching all the way to (simplicial) homology groups, including their independence of the triangulation. The latter is technical, but we do not see any other way of getting used to the machinery without actually working through a number of details.

The chapter does not get one very far in topology, but it may make a systematic study of full-fledged textbooks easier for those wishing to get deeper.

We fix notation for two sets in $\mathbb{R}^n$, which are used all the time in topology. The n-dimensional **ball** is

$$B^n = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$$

(some sources prefer the word *disk* and the notation D^n), and the $(n-1)$-dimensional **sphere** is the boundary of B^n, i.e.,

$$S^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$$

(note that S^2 lives in $\mathbb{R}^3$). Both are considered with the Euclidean metric.

1 Topological spaces and continuous maps

A topological space is a mathematical structure for capturing the notion of continuity, one of the most basic concepts of all mathematics, on a very general level.

The usual definition of continuity of a mapping from introductory courses uses the notion of distance: a mapping is continuous if the images of sufficiently close points are again close.

This can be formalized for mappings between metric spaces. We recall that a **metric space** is a pair (X, d_X), where X is a set and $d_X\colon X \times X \to \mathbb{R}$ is a **metric** satisfying several natural axioms (x, y, z are arbitrary points of X): $d_X(x, y) \geq 0$, $d_X(x, x) = 0$, $d_X(x, y) > 0$ for $x \neq y$, $d_X(y, x) = d_X(x, y)$, and $d_X(x, y) + d_X(y, z) \geq d_X(x, z)$ (the *triangle inequality*). The most important example of a metric space is $\mathbb{R}^n$ with the Euclidean metric, and another, of particular interest in computer science, is a graph with the shortest-path metric.

Formally, a mapping $f\colon X \to Y$ between metric spaces is continuous if for every $x \in X$ and every $\varepsilon > 0$ there exists $\delta > 0$ such that whenever $y \in X$ and $d_X(x, y) < \delta$, we have $d_Y(f(x), f(y)) < \varepsilon$.

One can think of a topological space as starting with a metric space and forgetting the metric, remembering only which sets are open. (We recall that a set $U \subseteq X$ in a metric space is open if for every $x \in U$ there is $\varepsilon > 0$ such that U contains the ε-ball around x.) This is not quite precise since topological spaces are much more general than metric spaces and there are many interesting specimens which cannot be obtained from any metric space, but in applications of topology we mostly encounter topological spaces coming from metric ones.

Topological space. Here is the general definition.

Definition 1.1. *A* **topological space** *is a pair* $(X, \mathcal{O})$*, where* X *is a (typically infinite) ground set and* $\mathcal{O} \subseteq 2^X$ *is a set system, whose members are called the* **open sets***, such that* $\emptyset \in \mathcal{O}$*,* $X \in \mathcal{O}$*, the intersection of finitely many open sets is an open set, and so is the union of an arbitrary collection of open sets.*

The system $\mathcal{O}$ as in the definition is sometimes called a **topology** on X.

In this chapter, we will often say just space instead of topological space.

Two topological spaces $(X, \mathcal{O}_X)$ and $(Y, \mathcal{O}_Y)$ are considered "the same" from the point of view of topology if there is a bijective map $f\colon X \to Y$ that preserves open sets in both directions; that is, $V \in \mathcal{O}_Y$ implies $f^{-1}(V) \in \mathcal{O}_X$ and $U \in \mathcal{O}_X$ implies $f(U) \in \mathcal{O}_Y$. For most mathematical structures, such as groups or graphs, an f with analogous structure-preserving properties is called an isomorphism, but in topology an f as above is called a **homeomorphism**. Topological spaces X and Y are said to be **homeomorphic**, written $X \cong Y$, if there is a homeomorphism between them. (Strictly speaking, we should write that the topological spaces $(X, \mathcal{O}_X)$ and $(Y, \mathcal{O}_Y)$ are homeomorphic, but in agreement with a common practice we mostly use the same letter for the topological space and for the underlying set.)

Here we see a substantial difference between metric and topological spaces: two spaces which are metrically quite different can be homeomorphic and thus topologically the same.

Exercise 1.2. *Verify the following homeomorphisms (the topology is always given by the Euclidean metric):*

(a) $\mathbb{R}$*, the open interval* $(0,1)$*, and* $S^1 \setminus \{(0,1)\}$ *(the unit circle in the plane minus one point).*

(b) S^1 *and the boundary of the unit square* $[0,1]^2$*.*

Similarly, different metrics on X may induce the same topology: this is the case for all ℓ_p metrics on $\mathbb{R}^n$ (n fixed), for example. For readers familiar with Banach spaces we also mention that all infinite-dimensional separable Banach spaces are homeomorphic as topological spaces—this is a nontrivial theorem of Kadets; in this case, from the point of view of functional analysis, the topology carries too little information.

Subspaces. The topological spaces encountered most often in applications, as well as in a substantial part of topology itself, are subspaces of some $\mathbb{R}^n$ with the standard topology (i.e., the one induced by the Euclidean metric), or are at least homeomorphic to such subspaces.

In general, for a topological space $(X, \mathcal{O})$, every subset $Y \subseteq X$ induces a **subspace** of $(X, \mathcal{O})$, namely, the topological space $(Y, \{U \cap Y : U \in \mathcal{O}\})$. (This is quite different, e.g., from groups, where only quite special subsets correspond to subgroups.) Note that the open sets of the subspace need not be open as subsets of X: for instance, let X be the Euclidean plane and Y a segment in it; then Y is open in Y but, of course, not in the plane.

Neighborhoods, bases, closure, boundary, interior. A set N in a topological space X is called a **neighborhood** of a point $x \in X$ if there is an open set U such that $x \in U \subseteq N$.

The system $\mathcal{O}$ of all open sets in a topological space can often be described more economically by specifying a **base** of $\mathcal{O}$, which is a collection $\mathcal{B} \subseteq \mathcal{O}$ such that every $U \in \mathcal{O}$ is a union of some of the sets in $\mathcal{B}$. For example, the system of all open intervals is a base of the standard topology of $\mathbb{R}$, and so is the system of all open intervals with rational endpoints.

Exercise 1.3. *Check that the system of all open balls of radius* $\frac{1}{n}$*,* $n = 1, 2, \ldots$*, constitutes a base of the topology of a metric space.*

A possibly still more compact specification of a topology $\mathcal{O}$ is a *subbase*, which is a system $\mathcal{S}$ such that the system of all finite intersections of sets from $\mathcal{S}$ forms a base of $\mathcal{O}$. An example is the system of all intervals $(-\infty, a)$ and (a, ∞), $a \in \mathbb{R}$, for $\mathbb{R}$.

A set $F \subseteq X$ is **closed** if $X \setminus F$ is open. Traditionally one uses letters U, V, W for open sets and F, G, H for closed sets, and in sketches, open sets are drawn as smooth ovals and closed sets as polygons.

The **closure** $\operatorname{cl} Y$ of a set in a topological space X is the intersection of all closed sets containing Y (an alternative notation is $\overline{Y}$). In the metric case, the closure consists of all points with zero distance to Y (where $d_X(x, Y) =$

$\inf_{y\in Y} d_X(x,y))$. The **boundary** of Y is $\partial Y := \mathrm{cl}(Y) \cap \mathrm{cl}(X \setminus Y)$, and the **interior** $\mathrm{int}\, Y := Y \setminus \partial Y$.

We note that these last three notions depend not only on Y, but also on the space X in which they are considered: for example, if $X = \mathbb{R}$ and Y is the closed interval $[0,1]$, then $\partial Y = \{0,1\}$ and $\mathrm{int}\, Y = (0,1)$, but if we consider the segment Y' connecting the points $(0,0)$ and $(1,0)$ as a subspace of $\mathbb{R}^2$, then $Y' \cong Y$ but $\partial Y' = Y'$ and $\mathrm{int}\, Y' = \emptyset$. To avoid ambiguities one sometimes writes $\mathrm{cl}_X Y$, $\partial_X Y$, $\mathrm{int}_X Y$.

Continuous maps. Now we return to continuity, whose topological definition is strikingly simple.

Definition 1.4. *A* **continuous mapping** *of a topological space $(X, \mathcal{O}_X)$ into a topological space $(Y, \mathcal{O}_Y)$ is a mapping $f\colon X \to Y$ of the underlying sets such that $f^{-1}(U) \in \mathcal{O}_X$ for all $U \in \mathcal{O}_Y$. In words, a mapping is continuous if the preimages of all open sets are open.*

In topological texts, *all mappings between topological spaces are usually assumed to be continuous* unless stated otherwise. We will also sometimes use this convention.

The next exercise is definitely worth doing.

Exercise 1.5. *Show that for mappings $\mathbb{R} \to \mathbb{R}$ (where $\mathbb{R}$ has the standard topology), or more generally for mappings between metric spaces, this definition of continuity is equivalent to the ε-δ definition recalled earlier.*

Exercise 1.6. *A curious reader might ask why the definition of continuity requires preimages, rather than images, of open sets to be open. We define a mapping $f\colon X \to Y$ between topological spaces to be an* open mapping *if $f(U)$ is open for every open set U. Find examples, involving mappings between subspaces of $\mathbb{R}$, of a continuous map that is not open, as well as of an open map that is not continuous.*

Exercise 1.7. (a) *Check that a homeomorphism of topological spaces can equivalently be defined as a bijective continuous mapping with continuous inverse.* (b) *Find an example of a bijective continuous mapping between suitable subspaces of $\mathbb{R}$ that is not a homeomorphism.*

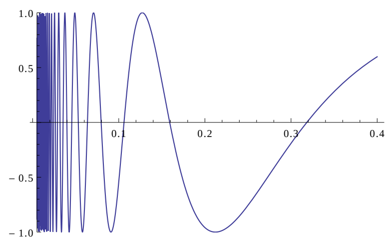
Exercise 1.8. *Let X, Y be a topological spaces, let $f\colon X \to Y$ be a mapping, and let $A_1, \ldots, A_n \subseteq X$ be closed sets that together cover all of X. Let us assume that the restriction of f to the subspace of X induced by A_i is continuous, for every $i = 1, 2, \ldots, n$ (while we do not apriori assume f continuous). Prove that f is continuous.*

2 Bits of general topology

There is a sizeable list of properties a topological space may or may not have. (These properties are all invariant under homeomorphism.) Here we present a brief selection.

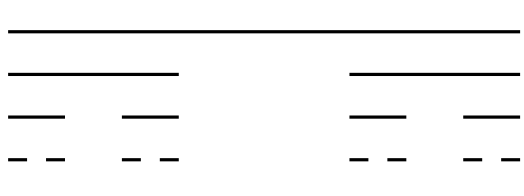
Connectedness. There are two different definitions capturing the intuitive idea that a topological space "has just one piece." A topological space X is **connected** if X cannot be written as a union of two disjoint nonempty open sets.[1] And X is **path-connected** if every two points x, y are connected by a path, where in the topological setting, a **path** from x to y is a continuous map $f\colon [0,1] \to X$ of the unit interval with $f(0) = x$ and $f(1) = y$.

Connectedness and path-connectedness are not equivalent: the latter implies the former, but a famous example of a connected space that is not path-connected is the *topologist's sine curve*, the subspace of $\mathbb{R}^2$ consisting of the vertical segment from $(0,-1)$ to $(0,1)$ and the graph of the function $x \mapsto \sin \frac{1}{x}$ for $x > 0$:



For applications, path-connectedness seems to be more important.

One can define *connected components* of a space X as inclusion-maximal subsets that, considered as topological subspaces of X, are connected, and analogously for path-connected components. Among the wilder examples we have the famous **Cantor set** $C \subset \mathbb{R}$, given by $C = \bigcap_{i=1}^{\infty} C_i$, where $C_0 = [0,1]$ and $C_i = \frac{1}{3}C_{i-1} \cup (\frac{1}{3}C_{i-1} + \frac{2}{3})$:



All of its connected (or path-connected) components are singletons, and there are uncountably many.

For every topological property we can hope that it allows us to distinguish some pairs of non-homeomorphic spaces. In the case of (path-)connectedness, we can prove that no two of the spaces S^1, $(0,1)$ (open interval) and $[0,1]$ (closed interval) are homeomorphic: indeed, if we remove a single point, then S^1 *always* stays connected, $(0,1)$ *never*, and $[0,1]$ *sometimes* stays connected and sometimes not. (Can you see other ways of proving any of these non-homeomorphisms?)

Bizarre spaces and general topology. So far we may have made the impression that all topological spaces look more or less like subspaces of Euclidean

[1] The literature is not quite unified concerning the question of whether the empty topological space is connected. It should be according to the general definition, but for many purposes it is better to define that it is not.

spaces, but this is very far from the truth—they need not even look like metric spaces.

A topological space X whose topology can be obtained from some metric is called **metrizable**. A traditional subfield of topology called **general topology** or *point-set topology* studies mainly various properties of topological spaces more general than metrizability, relations among them, conditions making a space metrizable, etc.

Let us list several examples taken from the vast supply built in general topology over the years. We will not prove any properties for them, except possibly in exercises—the intention is to give the reader some feeling for the possible pathologies occurring in arbitrary topological spaces, as well as a supply of candidate counterexamples for refuting too general claims. The reader may at least want to check in passing that these are indeed topological spaces.

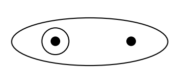
All of the examples except for (A) are non-metrizable, which in several cases is nontrivial to prove.

(A) Any set X, such as the real numbers, can be given the **discrete topology**, in which all subsets are open. Note that the integers $\mathbb{Z}$ inherit such a topology as a subspace of $\mathbb{R}$ with the standard topology, but discrete topology becomes more exotic if the ground set is uncountable.

(B) Let X be an infinite set. The *topology of finite complements* has open sets $\emptyset$ and $X \setminus B$ for all $B \subseteq X$ finite. Similarly one can define the topology of *countable complements* on an uncountable set.

(C) We recall that an *algebraic variety* in $\mathbb{R}^n$ (or, for that matter, in $\mathbb{K}^n$ for any field $\mathbb{K}$) is the set of common zeros of a set of n-variate polynomials. The open sets of the *Zariski topology* on $\mathbb{R}^n$ has all complements of algebraic varieties as open sets. The reader may want to check that for $n = 1$ we get the topology of finite complements. This is a (somewhat rare) example of an exotic topology used heavily outside the field of general topology, namely, in algebraic geometry.

(D) The two-point space $\{1, 2\}$ in which open sets are $\emptyset$, $\{1\}$, and $\{1, 2\}$:



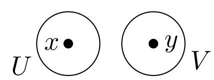
Here the closure of the singleton set $\{1\}$ is $\{1, 2\}$, while 1 is not in the closure of $\{2\}$, which probably cannot be considered good manners.

(E) The *Sorgenfrey line* is $\mathbb{R}$ with the topology whose base are all half-open intervals $[a, b)$. The *Sorgenfrey plane* is the product of the Sorgenfrey line with itself (products will be introduced soon); explicitly, this is $\mathbb{R}^2$ with the topology whose base are half-open rectangles $[a, b) \times [c, d)$.

(F) Let ω_1 be the first uncountable ordinal (assuming that the reader knows or looks up what ordinal numbers are). The set $L = \omega_1 \times [0, 1)$ is ordered lexicographically, and then given the topology whose base are all open intervals in this linear ordering. The resulting topological space is called the *long ray*; locally it looks like $\mathbb{R}$ with the standard topology, but globally it is "too long" to be metrizable.

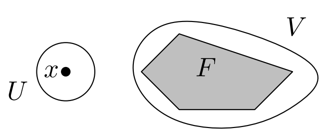
Separation axioms. One class of properties intended to measure how close a given space is to metrizability are traditionally called the *separation axioms*. The most popular ones are called T_0, T_1, T_2, T_3, $T_{3\frac{1}{2}}$, T_4 in the order of increasing strength (T abbreviates the German *Trennungsaxiom*, i.e., separation axiom), and one can also find $T_{2\frac{1}{2}}$, T_5, and T_6 in the literature, plus a number of others not quite fitting the T_i scale. Metrizable spaces have all of these properties.

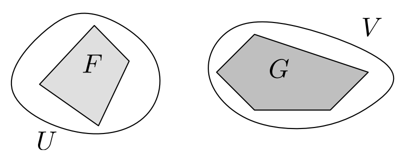
Probably the most important to remember is T_2: a space X is T_2 or **Hausdorff** if for every two distinct points $x, y \in X$ there are open sets $U \ni x$ and $V \ni y$ with $U \cap V = \emptyset$. Briefly, distinct points can be separated by open sets:



Decent topological spaces are at least Hausdorff (possibly with the honorable exception of the Zariski topology); examples (B)–(D) above are not.

For illustration, we also mention that a T_3 or *regular* space is one that is T_2 and in which every closed set F can be separated from every point $x \notin F$ by open sets, while a T_4 or *normal* space is T_2 and disjoint closed sets can be separated by open sets:





There are examples showing that all of the hierarchy is strict, i.e., T_i does not imply T_j for $i < j$. Sometimes these are quite sophisticated, the hardest being one showing $T_3 \not\Rightarrow T_{3\frac{1}{2}}$. As far as our examples above are concerned, the Sorgenfrey plane is $T_{3\frac{1}{2}}$ but not T_4.

We conclude this brief mention of the separation axioms by a warning: The literature is far from unified concerning terminology. The main difference is in whether, for the higher separation axioms like T_3 or T_4, one automatically assumes T_1 (or, equivalently, T_2) or not. Indeed, the modern usage seems to prefer "normal" to mean "disjoint closed sets separable by open sets" while T_4 means "normal+T_1." So it is advisable to check the definitions carefully.

Cardinality restrictions. A very important notion is that of a dense subset: a set $D \subseteq X$ is **dense** in a topological space X if $\operatorname{cl} D = X$.

A space X is **separable** if it has a countable dense set. The space $\mathbb{R}^n$ with the standard topology is separable because the set $\mathbb{Q}^n$ of all rational points is dense in it, and so is every subspace.

Exercise 2.1. (a) *Show that the Sorgenfrey plane in (E) above is separable, but it has a non-separable subspace.*

(b) *Prove that a subspace of a separable metric space is separable.*

A notion with less importance outside topology is a space with countable base (meaning a base for its topology as introduced earlier), which for historical reasons is often called a *second-countable space*. This is a property much stronger than separability.

Polish spaces. In many fields of mathematics, when one wants to work only with "sufficiently nice" topological spaces, one makes assumptions even stronger than metrizability. The most frequent such concept is perhaps a **Polish space**, which is a separable completely metrizable space.

Here one needs to know that a *complete metric space* is one in which every Cauchy sequence[2] converges to a limit. For example, the Euclidean metric on $\mathbb{R}$ is complete, but on $(0,1)$ it is not. The definition of Polish space requires the existence of at least one complete metric inducing the topology; so, for example, $(0,1)$ is a Polish space.

Let us conclude this section with two examples of nice basic theorems of general topology. The first one we state without proof:

Theorem 2.2 (Tietze extension theorem)**.** *Let X be a metric space, or more generally, a T_4 topological space, let $A \subseteq X$ be closed, and let $f\colon A \to \mathbb{R}$ be a continuous map. Then there exists a continuous extension $\overline{f}\colon X \to \mathbb{R}$ of f, for which we may moreover assume* $\sup_{x\in X} |\overline{f}(x)| \le \sup_{a\in A} |f(x)|$.

Theorem 2.3 (Urysohn metrization theorem)**.** *Every T_3 topological space with a countable base is metrizable.*

We present a proof, assuming for convenience T_4 instead of just T_3.

Exercise 2.4. *Prove that a T_3 space with a countable base is also T_4.*

The proof of Theorem 2.3 contains a very useful and general trick (appearing, e.g., in the theory on low-distortion embeddings of finite metric spaces, a recent hot topic in computer science, all the time).

The countable base assumption, as well as Tietze's extension theorem, are used in the next lemma.

Lemma 2.5. *For every T_4 space X with a countable base there exists a countable sequence $(f_1, f_2, \ldots)$ of continuous functions $X \to [0,1]$ such that for every point $x \in X$ and every open set U with $x \in U$ there is an f_i that is 0 outside U and 1 in x.*

Proof. For every pair (B, B') of the assumed countable base $\mathcal{B}$ of X with $\operatorname{cl} B' \subset B$, we use the Tietze extension theorem to get a function $X \to [0,1]$ that equals 1 on $\operatorname{cl} B'$ and equals 0 on $X \setminus B$. These are the desired f_i.

To check that this works, we consider $x \in U$ as in the lemma. We find $B \in \mathcal{B}$ with $x \in B \subseteq U$, and then we use the T_3 property to separate x from $X \setminus B$ by disjoint open sets $V \ni x$ and $W \supseteq X \setminus B$. It follows that $\operatorname{cl} V \subseteq X \setminus W \subseteq B$. Finally we shrink V to some $B' \in \mathcal{B}$ still containing x.

We now have $x \in B' \subseteq \operatorname{cl} B' \subseteq B \subseteq U$, and it is clear that the separating function made above for (B, B') is 1 at x and 0 outside U. $\square$

[2]A sequence $(x_1, x_2, \ldots)$ is *Cauchy* if for every $\varepsilon > 0$ there is n such that for all $i, j \ge n$ we have $d_X(x_i, x_j) < \varepsilon$.

Proof of Theorem 2.3 under the T_4 assumption. Let H, the **Hilbert cube**, be the metric space of all infinite sequences $x = (x_1, x_2, \ldots)$, $x_i \in [0, \frac{1}{i}]$, $i = 1, 2, \ldots$, with the ℓ_2 metric, meaning that the distance of x and y is $\left(\sum_{i=1}^{\infty}(x_i - y_i)^2\right)^{1/2}$.

We will show that the space X as in the theorem is homeomorphic to a subspace of H. Then the metrizability of X will be clear.

We define a mapping $f\colon X \to H$ by

$$f(x) := \left(\tfrac{1}{1}f_1(x), \tfrac{1}{2}f_2(x), \tfrac{1}{3}f_3(x), \ldots\right)$$

where the f_i are as in the lemma (this definition is the main trick!).

Exercise 2.6. *Check that f is continuous (this uses nothing but the continuity of the f_i) and injective.*

It remains to verify that the inverse mapping $f^{-1}\colon f(X) \to X$ is continuous. To this end it suffices to check that for every $U \subseteq X$ open and every $x \in U$, there is an $\varepsilon > 0$ such that $f(U)$ contains the ε-ball around $f(x)$ (ball in $f(X)$, not in all of H, that is).

As expected, we fix i with $f_i(x) = 1$ and f_i zero outside U, and we let $\varepsilon := \frac{1}{2i}$. Now we suppose that $y \in X$ is such that $f(x)$ and $f(y)$ have distance at most ε in H; we want to conclude $y \in U$. We have, in particular, $\frac{1}{i}|f_i(x) - f_i(y)| \le \varepsilon$, so $f_i(y) \ge \frac{1}{2}$, and thus $f_i(y) \ne 0$. Hence $y \in U$ as needed. □

3 Compactness

One of the most important and most applied topological properties is *compactness*. Intuitively, a compact space is one that does not have too much room inside. The topological definition is quite simple:

Definition 3.1. *A topological space X is* **compact** *if for every collection $\mathcal{U}$ of open sets in X whose union is all of X, there exists a finite $\mathcal{U}_0 \subseteq \mathcal{U}$ whose union also covers all of X. In brief, every open cover of X has a finite subcover.*

A set $C \subseteq X$ is a **compact set** in X if C with the subspace topology is a compact space.

The notion of compactness was first developed in the metric setting, with a different definition, which is still presented in many introductory courses. Namely, a metric space X is compact if every infinite sequence $(x_1, x_2, \ldots)$ contains a subsequence $(x_{i_1}, x_{i_2}, \ldots)$, $i_1 < i_2 < \cdots$, that is convergent.

Exercise 3.2. *Prove that if X is a metric space that is compact according to Definition 3.1, then every infinite sequence has a convergent subsequence. Hint: construct an open cover by balls "witnessing" that there is no convergent subsequence.*

Diligent readers may also do the opposite implication for metric spaces, but this is more difficult.

While one can naturally define convergent sequences in a topological space, and thus transfer the definition with sequences to topological spaces, one obtains a different, and much less well behaved, notion of *sequential compactness.* From this point of view, the topological approach, as opposed to the metric one, greatly clarified the essence of the notion.

Mainly in order to show typical proofs in general topology, we will now develop some properties of compactness, culminating in two extremely useful results concerning compact sets.

Lemma 3.3.

(i) *A closed subset of a compact space is compact.*

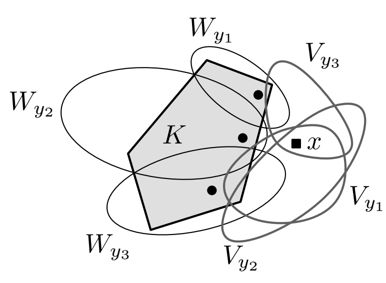
(ii) *A compact subset in a Hausdorff space is closed.*

(iii) *If $f\colon X \to Y$ is continuous and $K \subseteq X$ is compact, then $f(K)$ is compact (and hence closed if Y is Hausdorff).*

To appreciate (iii), one should realize that continuous maps need not map closed sets to closed sets in general.

Proof. In (i), let X be compact and $F \subseteq X$ be closed. Consider an open cover $\mathcal{U}$ of F, and for every $U \in \mathcal{U}$, fix an open set $\tilde{U}$ in X with $\tilde{U} \cap K = U$. Then $\tilde{\mathcal{U}} := \{\tilde{U} : U \in \mathcal{U}\} \cup \{X \setminus F\}$ is an open cover of X. From a finite subcover of $\tilde{\mathcal{U}}$ we obtain a finite subcover of $\mathcal{U}$ by restricting everything back to F.

For (ii), let X be Hausdorff and $K \subseteq X$ be compact. It suffices to show that for every $x \notin K$ there is an open U_x such that $U_x \cap K = \emptyset$. For every $y \in K$ we can fix, by the Hausdorff property, disjoint open sets $V_y \ni x$ and $W_y \ni y$. The W_y for all $y \in K$ form an open cover of K, so we select a finite subcover, say $W_{y_1}, \ldots, W_{y_n}$, and we set $U_x := \bigcap_{i=1}^n V_{y_i}$.



Finally, (iii) is easy based on the observation that if $\mathcal{U}$ is an open cover of $f(K)$, then $\{f^{-1}(U) : U \in \mathcal{U}\}$ is an open cover of K. $\square$

Here is the first often-applied result.

Theorem 3.4. *Let K be compact, and let $f\colon K \to \mathbb{R}$ be a continuous function. Then f attains its minimum: there exists $x_0 \in K$ with $f(x_0) = \inf_{x\in K} f(x)$. In particular, a continuous function on a compact set is bounded, and a function on a compact set that is never zero is bounded away from 0; that is, there is $\varepsilon > 0$ such that $|f(x)| \geq \varepsilon$ for all $x \in K$.*

Proof. By Lemma 3.3(iii), $Y := f(K) \subseteq \mathbb{R}$ is compact. Set $m := \inf Y$, choose a sequence $(y_1, y_2, \ldots)$, $y_i \in Y$, converging to m, and set $U_i := (y_i, \infty)$.

If the U_i do not cover Y, then this can be only because they all avoid m, and in particular, $m \in Y$. So we suppose that $\{U_i\}$ is an open cover of Y, and we select a finite subcover $U_{i_1}, \ldots, U_{i_n}$. Let $y^* := \min\{y_{i_1}, \ldots, y_{i_n}\}$. Then $Y \subseteq \bigcup_{j=1}^{n} U_{i_j} = (y^*, \infty)$, but this is a contradiction since $y^* \in Y$. $\square$

Products. The product of two topological spaces $(X, \mathcal{O}_X)$ and $(Y, \mathcal{O}_Y)$ is defined in an expected way, with the ground set $X \times Y$ and the collection $\{U \times V : U \in \mathcal{O}_X, V \in \mathcal{O}_Y\}$ of *open rectangles* as a base of the topology.

The definition of a product of *infinitely many* spaces is trickier (but often needed): we do not take all open rectangles, but only those having only finitely many coordinates in which the open set is not the whole space. Thus, if $(X_i, \mathcal{O}_i)_{i \in I}$ is a collection of spaces indexed by an arbitrarily large set I, then the product space $\prod_{i \in I}(X_i, \mathcal{O}_i)$ has ground set $\prod_{i \in I} X_i$, and a base of the topology is

$$\Big\{\prod_{i \in I} U_i : U_i \in \mathcal{O}_i, |\{i \in I : U_i \neq X_i\}| < \infty\Big\}.$$

For example, the product of countably many copies of the two-point discrete space $\{0, 1\}$ turns out to be homeomorphic to the Cantor set C, and the product of countably many copies of $\{0, 1, 2, \ldots\}$, again with the discrete topology, is homeomorphic to the set of all irrational numbers with the standard topology inherited from $\mathbb{R}$ (ambitious readers may want to prove these).

Exercise 3.5. *Prove that a product of Hausdorff spaces is Hausdorff.*

Theorem 3.6 (Tychonoff's theorem)**.** *The product of an arbitrary collection of compact topological spaces is compact.*

Exercise 3.7. (a) *Prove that if $X \times Y$ is a product of two topological spaces such that every open cover of $X \times Y$ by open rectangles (i.e., sets of the form $U \times V$, U open in X, V open in Y) has a finite subcover, then $X \times Y$ is compact.*

(b) *Prove Tychonoff's theorem for products of two spaces.*

The proof of Tychonoff's theorem for infinitely many factors needs more work, and more significantly, it relies on the *axiom of choice*—Tychonoff's theorem is actually one of the important theorems *equivalent* to the axiom of choice.

Instead of a proof, we will demonstrate a typical combinatorial application (similar considerations underlie compactness principles in logic and elsewhere). We recall that a graph $G = (V, E)$ is *k-chromatic* if there is a mapping (coloring) $c\colon V \to [k] := \{1, 2, \ldots, k\}$ such that $f(u) \neq f(v)$ whenever $\{u, v\}$ is an edge of G.

Proposition 3.8. *Let G be an infinite graph. If every finite subgraph of G is k-chromatic, then G is k-chromatic.*

For countable graphs there is an elementary inductive proof. Tychonoff's theorem provides a quick proof in general.

Proof. For every vertex $v \in V$, let X_v be a copy of the discrete topological space $[k]$, and let $X := \prod_{v \in V} X_v$. Since the X_v are (trivially) compact, X is compact.

A point of X can be identified with a mapping $f\colon V \to [k]$. For every edge $e = \{u, v\} \in E$, let $F_e \subseteq X$ consist of those mappings $f\colon V \to [k]$ for which $f(u) \neq f(v)$. We want to prove that $\bigcap_{e \in E} F_e \neq \emptyset$.

What we know is that whenever $E_0 \subseteq E$ is a finite set of edges, we have $\bigcap_{e \in E_0} F_e \neq \emptyset$. This is because the finite graph consisting of the edges of E_0 and their vertices is assumed to be k-chromatic.

By the definition of the product topology, it is easy to see that every F_e is closed. So it suffices to verify the following claim: *If $\mathcal{F}$ is a collection of closed sets in a compact space X such that every finite subcollection has a nonempty intersection, then $\mathcal{F}$ has a nonempty intersection.* But this is a reformulation of the definition of compactness—just consider $\mathcal{U} := \{X \setminus F : F \in \mathcal{F}\}$. $\square$

Compact subsets of $\mathbb{R}^n$. Now we can easily establish the following well-known characterization.

Theorem 3.9. *A subset $A \subseteq \mathbb{R}^n$ with the standard topology is compact if and only if it is both closed and bounded.*

Proof. First we assume A compact. Then A is closed by Lemma 3.3(ii), and boundedness follows by considering the open cover by balls $B(0, n)$, $n = 1, 2, \ldots$.

For the other direction, it suffices to prove that the cube $[-m, m]^n$ is compact for every m, n, since then the case of a general A follows by Lemma 3.3(i).

The crucial part is in proving the interval $[0, 1]$ compact; the rest follows by re-scaling and by Tychonoff's theorem. The compactness of closed intervals is built deeply in the construction of the reals, and it is more or less a rephrasing of the fact that every subset of $\mathbb{R}$ has a supremum.

So let $\mathcal{U}$ be an open cover of $[0, 1]$, and let s be the supremum of those $a \leq 1$ for which $[0, a]$ can be covered by finitely many members of $\mathcal{U}$.

Clearly $s > 0$. If $0 < s < 1$, then there is $\varepsilon > 0$ such that $[s - \varepsilon, s + \varepsilon]$ is covered by some $U \in \mathcal{U}$. Together with the assumed finite cover of $[0, s - \varepsilon]$, this U forms a finite cover of $[0, s + \varepsilon]$—a contradiction. $\square$

Exercise 3.10. *The previous result shows, in particular, that the Euclidean unit ball in $\mathbb{R}^n$ is compact.*

(a) *Consider the (infinite-dimensional Hilbert) space ℓ_2 consisting of all infinite sequences $x = (x_1, x_2, \ldots)$ of real numbers such that $\|x\| := \left(\sum_{i=1}^{\infty} x_i^2\right)^{1/2}$ is finite. Regard it as a topological space with topology induced by $\|.\|$, i.e., by the metric given by $d(x, y) = \|x - y\|$. Show that the unit ball $\{x \in \ell_2 : \|x\| \leq 1\}$ is not compact.*

(b) *Explain where the proof above, showing that B^n is compact, fails for the unit ball in ℓ_2.*

Paracompactness. There are many variations on compactness, most of them weaker than compactness, and none as significant. We mention just one notion, paracompactness, which often occurs among assumptions in other fields of mathematics.

We do not give the standard definition but an equivalent property which is most often used in applications. So let us assume that X is a Hausdorff space; then X is **paracompact** if every open cover $\mathcal{U}$ of X admits a partition of unity subordinated to $\mathcal{U}$. Here a *partition of unity subordinated to* $\mathcal{U}$ is a collection, finite or infinite, $(f_i)_{i \in I}$ of continuous functions $f_i \colon X \to [0,1]$ such that, first, for every $x \in X$, the sum $\sum_{i \in I} f(x)$ has only finitely many nonzero terms and equals 1, and second, for every $i \in I$ there is $U \in \mathcal{U}$ such that f_i is zero everywhere outside U.

Partitions of unity are a useful technical tool for gluing "locally defined" objects on X into a global object. Paracompactness is a relatively weak property: in particular, every compact space is paracompact, and all metric spaces are paracompact (which is a hard result). A non-paracompact example is the long ray introduced in (F) above.

4 Homotopy and homotopy equivalence

So far we have considered two topological spaces equivalent (the same) if they are homeomorphic. But finding out whether two given spaces are homeomorphic is a very ambitious and generally hopeless task, since it is known that the algorithmic problem, given two spaces X and Y, decide whether $X \cong Y$, is *algorithmically unsolvable*. (At the same time, homeomorphism can be decided in many specific settings, and topology is full of remarkable results of this kind. For example, later we will see that $\mathbb{R}^m \not\cong \mathbb{R}^n$ for $m \neq n$, which is well known but quite nontrivial.)

Even stronger undecidability claims hold; for example, it is undecidable whether a given space X is homeomorphic to the 5-dimensional sphere S^5, a very simple-looking space.

An attentive reader might wonder how a topological space, a highly infinite object in general, is given to an algorithm that can accept only finite inputs. This question will be discussed later, but for the moment, one may think of the input X to the question of homeomorphism with S^5 as a space living in some $\mathbb{R}^n$ and built of finitely many 5-dimensional Lego cubes, for example.

Algebraic topology, a branch which we are now slowly entering, considers topological spaces with a coarser equivalence, called *homotopy equivalence*. For example, as we will see, all of the spaces $\mathbb{R}^n$, $n = 1, 2, \ldots$, are homotopy equivalent, and actually homotopy equivalent to a one-point space.

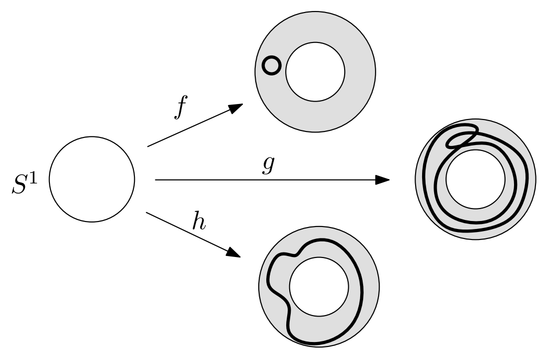
While deciding homotopy equivalence is still undecidable in general, chances of success in concrete cases are much better than for homeomorphism. The reason is that there are many wonderful tools (the reader may have heard keywords like fundamental group, homotopy groups, homology and cohomology groups, etc.) that cannot distinguish between two homotopy equivalent spaces, but they can often prove homotopy non-equivalence.

Homotopy of maps. Homotopy equivalence is a somewhat sophisticated concept, which needs some time to be digested. We begin with an analogous but simpler notion for maps.

Definition 4.1. *Two (continuous) maps $f, g\colon X \to Y$ between the same spaces are called* **homotopic**, *written $f \sim g$, if there exists a continuous map $H\colon X \times [0,1] \to Y$, a* homotopy *between f and g, satisfying $H(.,0) = f$ and $H(.,1) = g$.*

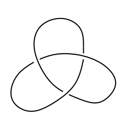
Intuitively, f and g are homotopic if f can be continuously deformed into g. The homotopy H specifies such a deformation: we can think of the second coordinate t as time, and for every point $x \in X$, the mapping $h_x(t) = H(x,t)$ specifies the trajectory of the image of x during the deformation: it starts in $f(x)$ at time $t = 0$, moves continuously, and reaches $g(x)$ at time $t = 1$. The continuity of H implies that this trajectory is continuous for every x, and also that close points must have close trajectories.

The next picture shows three maps of S^1 into the annulus (a part of the plane with a hole).



We have $f \sim g$ (imagine an appropriate deformation). But h is not homotopic to either of f, g—this is quite intuitive, since h goes once around the hole, while f and g do not go around, in a suitably defined sense, but proving it rigorously is nontrivial, and we will leave it without proof for now.

Exercise 4.2. (a) *Is the mapping $f\colon S^1 \to \mathbb{R}^3$ that maps S^1 to a geometric circle homotopic to a mapping $g\colon S^1 \to \mathbb{R}^3$ sending the circle to a knot, such as the trefoil? Answer before reading further!*



(b) *Let X be a space. Prove that every two maps $X \to B^n$ are homotopic.*
(c) *Prove that every two maps $B^n \to X$ are homotopic, provided X is path-connected.*

It is not difficult to show that being homotopic is an equivalence relation (writing down the proof of transitivity may take some work, but the idea is absolutely straightforward). We write $[X, Y]$ for the set of all homotopy classes of continuous maps $X \to Y$.

While there are usually uncountably many maps $X \to Y$, $[X, Y]$ is countable for spaces normally encountered in applications, sometimes even finite, and in many cases of interest it is well understood.

As a simple example we mention, again without proof, that the homotopy classes of maps of S^1 into the annulus are in a bijective correspondence with $\mathbb{Z}$, where each mapping is assigned the number of times the image winds around the hole, in positive (counterclockwise) or negative (clockwise) direction.

A map homotopic to a constant map $X \to Y$ (i.e., mapping all of X to a single point) is called, with a bit illogical-looking terminology, **nullhomotopic**.

Homotopy equivalence. Now we come to spaces. The usual definition of homotopy equivalence is not very intuitive but good to work with.

Definition 4.3. *Two spaces X and Y are* **homotopy equivalent***, written $X \simeq Y$, if there are continuous maps $f\colon X \to Y$ and $g\colon Y \to X$ such that the composition $fg\colon Y \to Y$ is homotopic to the identity map* id_Y *and* $gf \sim \mathrm{id}_X$.

The map g as in the definition is called a *homotopy inverse* to f (and vice versa).

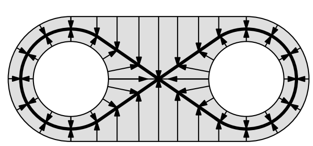
Similar to homotopy of maps, it is a simple exercise to show that homotopy equivalence is transitive. A class of homotopy equivalence of spaces is called a **homotopy type**.

Exercise 4.4. (a) *Show that the dumbbell ⚯ and the letter θ are homotopy equivalent.*

(b) *(This is a very basic fact.) Check that* $\mathbb{R}^n \setminus \{0\} \simeq S^{n-1}$.

A way of visualizing homotopy equivalence uses the notion of *deformation retract*. Let X be a space and Y a subspace of X (this is important). A *deformation retraction* of X onto Y is a continuous map $R\colon X \times [0,1] \to X$ such that $R(.,0)$ is the identity map id_X, $R(t,y) = y$ for all $y \in Y$ and all $t \in [0,1]$ (Y remains pointwise fixed), and $R(x,1) \in Y$ for all $x \in X$. We say that Y is a **deformation retract** of X if there is a deformation retraction as above.

The deformation retraction R describes a continuous motion of points of X within X such that every point ends up in Y and Y remains fixed all the time. Here is an example, with X a thick figure 8 and Y a thin one:

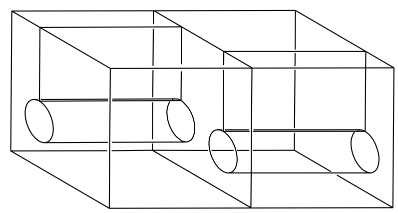


Now it is a theorem that two spaces X, Y are homotopy equivalent if and only if there exists a space Z such that both X and Y are deformation retracts of Z. The direction which helps us with visualization, i.e., being deformation retracts of the same space implies homotopy equivalence, is exercise-level, and the other, with a right idea, is simple as well.

Exercise 4.5. *Take an S^2 in $\mathbb{R}^3$ and connect the north and south poles by a segment, obtaining a space X. Take another copy of S^2 and attach a circle S^1 to the north pole by a single point, which yields Y. Show that $X \simeq Y$ (you may use deformation retracts).*

A space that is homotopy equivalent to a single point is called **contractible**.

Some spaces are "obviously" contractible, such as the ball B^n, but for others, contractibility is not easy to visualize. An example is **Bing's house**, one of the puzzling and beautiful objects of topology:



Bing's house is a hollow box with a wall inside separating it into two rooms, left and right. Each room has its own entrance, but by the architect's caprice, the entrance to the right room goes through a tunnel inside the left room (but is not accessible from the left room), and vice versa. Each of the tunnels is also attached to the ceiling by a vertical wall, which assures contractibility.

To check contractibility, one can visualize a deformation retraction of a solid cube onto Bing's house. If the cube is made of clay, one can push in a hole from the left and hollow out the right room through the hole, and similarly for the left room.

5 The Borsuk–Ulam theorem

Here we interrupt our gradual introduction of basic topological notions and ideas, and we present the Borsuk–Ulam theorem, which is arguably one of the most useful tools topology has to offer to non-topologists. (Another theorem of comparable fame and usefulness is Brouwer's, which we will treat later.)

We begin by stating three versions, easily seen to be equivalent. The following notion will be useful: Let $X \subseteq \mathbb{R}^m$ and $Y \subseteq \mathbb{R}^n$ be antipodally symmetric sets; that is, $x \in X$ implies $-x \in X$. We call a continuous mapping $f\colon X \to Y$ an **antipodal map** if $f(-x) = -f(x)$ for all $x \in X$ (so an antipodal map is automatically assumed continuous).

Theorem 5.1 (Borsuk–Ulam)**.** (i) *For every continuous mapping $f\colon S^n \to \mathbb{R}^n$ there is a point $x \in S^n$ with $f(x) = f(-x)$.*

(ii) *Every antipodal map $g\colon S^n \to \mathbb{R}^n$ maps some point $x \in S^n$ to 0, the origin in $\mathbb{R}^n$.*

(iii) *There is no antipodal mapping $S^n \to S^{n-1}$.*

Exercise 5.2. *Prove the equivalence (i)$\Leftrightarrow$(ii)$\Leftrightarrow$(iii).*

Exercise 5.3. *(Harder) Derive the following from Theorem 5.1: An antipodal map $S^n \to S^n$ cannot be nullhomotopic.*

The Borsuk–Ulam theorem comes from the 1930s and many different proofs are known. Unfortunately, conceptual proofs providing deeper insight require topological machinery beyond our scope, and the more elementary proofs we are aware of are often nice and clever, but one needs to spend considerable time with inessential technicalities. So we refer to the literature for a proof (e.g., [Mat03] or references therein), and instead we derive yet another, different-looking version.

Theorem 5.4 (Lyusternik–Schnirel'man). *Let $A_1, \dots, A_{n+1} \subseteq S^n$ be $n+1$ sets that together cover S^n, and let us assume that, for each i, A_i is either open or closed. Then some A_i contains a pair of antipodal points, x and $-x$.*

This theorem is traditionally presented either with all A_i closed or all A_i open, but allowing for a mixture can be useful, as we will see.

Exercise 5.5. (a) *Construct a covering of S^n with $n+2$ closed sets, none containing an antipodal pair.*

(b) *Cover S^n with two sets, neither containing an antipodal pair.*

Proof of Lyusternik–Schnirel'man from Borsuk–Ulam. First we assume that all the A_i are closed, and we define a continuous map $f\colon S^n \to \mathbb{R}^n$ by $f(x)_i = \mathrm{dist}(x, A_i)$, the Euclidean distance of x from A_i. By the Borsuk–Ulam theorem there is $x \in S^n$ with $f(x) = f(-x)$. If $f(x)_i = 0$ for some i, then $x \in A_i$ (here we use the closedness), as well as $-x \in A_i$, and we are done. If, on the other hand, $f(x)_i > 0$ for all i, then x and $-x$ do not belong to any of $A_1, \dots, A_n$, and so they both lie in A_{n+1}, the set which was seemingly neglected in the definition of f.

Next, let the A_i be all open. It suffices to show that there are closed $F_1 \subset A_1, \dots, F_{n+1} \subset A_{n+1}$ that together still cover S^n, since then we can use the version with the A_i closed.

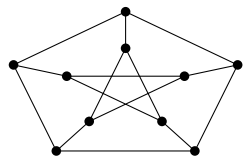
The proof of the last claim is a typical application of compactness. For every $x \in S^n$ we choose $i = i(x)$ such that $x \in A_i$, and an open neighborhood U_x of x whose closure is contained in $A_{i(x)}$. The U_x form an open cover of S^n, so we can choose a finite subcover, say $U_{x_1}, \dots, U_{x_m}$. Then we set $F_i := \bigcup_{j:i(x_j)=i} \mathrm{cl}\, U_{x_j}$.

Finally, let $A_1, \dots, A_k$ be open and $A_{k+1}, \dots, A_{n+1}$ closed. We proceed by contradiction, supposing that no A_i contains an antipodal pair. Then, for each $i \geq k+1$, A_i has some positive distance $\varepsilon_i > 0$ from $-A_i$, and we let A_i' be the open $(\varepsilon_i/3)$-neighborhood of A_i. We still have $A_i' \cap (-A_i') = \emptyset$, and hence the open sets $A_1, \dots, A_k, A_{k+1}', \dots, A_{m+1}'$ contradict the version of the theorem for open sets proved above. $\square$

Exercise 5.6. *Derive the Borsuk–Ulam theorem from the Lyusternik–Schnirel'man theorem. Hint: use Exercise 5.5(a).*

Kneser graphs. For integers n and k, the **Kneser graph** $\mathrm{KG}_{n,k}$ has all k-element subsets of some fixed n-element set X as vertices. Two such subsets F_1, F_2 are connected by an edge in $\mathrm{KG}_{n,k}$ if they are disjoint.

A Kneser graph is typically quite large; it has $\binom{n}{k}$ vertices. As a small example, we note that $\mathrm{KG}_{5,2}$ is isomorphic to the famous *Petersen graph*:



There are several reasons why Kneser graphs constitute an extremely interesting class of graph-theoretic examples (recently they have also been used in computer science in connection with the PCP theorem). Perhaps the most remarkable property is that they have a significantly large chromatic number, but their chromatic number is not explained by any of the "usual" reasons, as we will indicate below.

We have already mentioned k-chromatic graphs in connection with Proposition 3.8; here we just add that the **chromatic number** $\chi(G)$ of a graph G is the smallest k such that G is k-chromatic.

The following celebrated result was conjectured by Kneser and proved by Lovász:

Theorem 5.7 (Lovász–Kneser). *For $n \geq 2k$, we have $\chi(\mathrm{KG}_{n,k}) \geq n - 2k + 2$.*

The chromatic number of $\mathrm{KG}_{n,k}$ actually equals $n-2k+2$; finding a coloring is an elementary but nice exercise.

The perhaps most common general lower bound for $\chi(G)$ is $\chi(G) \geq |V(G)|/\alpha(G)$, where $\alpha(G)$, the *independence number* of G, is the size of a maximum independent set in G. This lower bound has a simple reason, since an equivalent definition of a k-chromatic graph is that the vertex set can be covered by k independent sets.

Now $\mathrm{KG}_{n,k}$ has quite large independent sets, of size $\binom{n-1}{k-1}$, corresponding to the collection of all k-element sets containing a given point of the ground set. Setting $n = 3k - 2$, for example, we see that $\chi(\mathrm{KG}_{3k-2,k}) = k$, while the $|V(G)|/\alpha(G)$ lower bound yields less than 3.

Even more strongly, $\mathrm{KG}_{3k-2,k}$ also has the *fractional chromatic number* less than 3, where the fractional chromatic number $\chi_f(G)$ can be compactly defined as the infimum of fractions $\frac{a}{b}$ such that $V(G)$ can be covered by a independent sets so that every vertex is covered at least b times. The fractional chromatic number is an important graph parameter, and examples with a large gap between χ_f and χ are *very* rare.

Many proofs of the Lovász–Kneser theorem are known, but all of them are topological, or at least strongly inspired by the topological proofs. We present a particularly short and neat one.

Proof of the Lovász–Kneser theorem. The Kneser graph $\mathrm{KG}_{n,k}$ needs an n-element ground set X; we choose X as an n-point set in $\mathbb{R}^{d+1}$ in general position, where $d = n - 2k + 1$, and where general position means that no $d+1$ points of X lie on a common hyperplane passing through the origin.

For contradiction, we suppose that there is a proper coloring of $\mathrm{KG}_{n,k}$ by at most $n-2k+1 = d$ colors. We fix one such proper coloring and we define sets $A_1, \ldots, A_d \subseteq S^d$: For a point $x \in S^d$, we have $x \in A_i$ if there is at least one k-tuple $F \subset X$ of color i contained in the open halfspace $H(x) := \{y \in \mathbb{R}^d :$

$\langle x, y \rangle > 0\}$ (i.e., x is a unit normal of the boundary of $H(x)$ and points into $H(x)$). Finally, we put $A_{d+1} = S^d \setminus (A_1 \cup \cdots \cup A_d)$.

Clearly, A_1 through A_d are open sets, while A_{d+1} is closed. By our version of the Lyusternik–Schnirel'man theorem, there exist $i \in [d+1]$ and $x \in S^d$ such that $x, -x \in A_i$.

If $i \leq d$, we get two disjoint k-tuples colored by color i, one in the open halfspace $H(x)$ and one in the opposite open halfspace $H(-x)$. This means that the considered coloring is not a proper coloring of the Kneser graph.

If $i = d+1$, then $H(x)$ contains at most $k-1$ points of X, and so does $H(-x)$. Therefore, the common boundary hyperplane of $H(x)$ and $H(-x)$ contains at least $n-2k+2 = d+1$ points of X, and this contradicts the choice of X. $\square$

6 Operations on topological spaces

We have seen the *product* of topological spaces as an operation creating new spaces from old ones. Here we introduce some more operations.

Quotient. Given a topological space X and a subset $A \subset X$, we can form a new space by "shrinking A to a point." Two spaces can be "glued together" to form another space. A space can be factored using a group acting on it. Here is a general definition capturing all of these cases.

Definition 6.1. *Let X be a topological space and let $\approx$ be an equivalence relation on the set X. The points of the* **quotient space** *$X/\!\approx$ are the classes of the equivalence $\approx$, and a set $U \subseteq X/\!\approx$ is open if $q^{-1}(U)$ is open in X, where $q\colon X \to X/\!\approx$ is the* quotient map *that maps each $x \in X$ to the equivalence class $[x]_\approx$ containing it.*

If A is a subspace of X, one writes X/A for the quotient space $X/\approx$, where the classes of $\approx$ are A and the singletons $\{x\}$ for all $x \in X \setminus A$. This formalizes the "shrinking of A to a single point" mentioned above.

More generally, if $(A_i)_{i \in I}$ is a collection of disjoint subspaces, the notation $X/(A_i)_{i \in I}$ is used, with the expected meaning (each A_i is shrunk to a point).

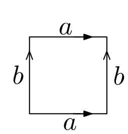
It is not hard to see, even rigorously, that $[0,1]/\{0,1\} \cong S^1$. Here are examples requiring more of mental gymnastics:

Exercise 6.2. *Substantiate, at least on an intuitive level, the following homeomorphisms:*

(a) $(S^n \times [0,1])/(S^n \times \{0\}) \cong B^{n+1}$.

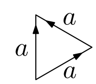
(b) $B^n/S^{n-1} \cong S^n$.

(c) $[0,1]^2/\!\approx\, \cong S^1 \times S^1$, *where $\approx$ is given by the following identification of the sides of the square:*



The picture means that each point of an arrow labeled a is to be identified with the corresponding point of the other a-arrow, and similarly for the b-arrows *(so, in particular, all four corners are glued together). This is a well-known construction of the* **torus***.*

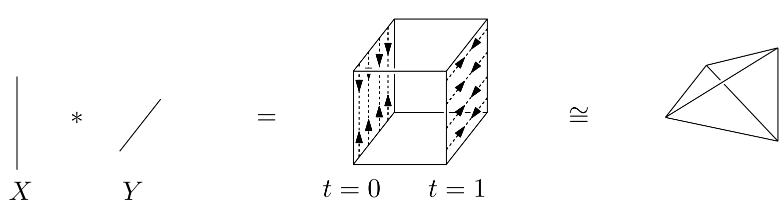
The following identification of the sides of a triangle leads to a mind-boggling space called the **dunce hat**, with properties similar to those of Bing's house. The dunce hat can be made in $\mathbb{R}^3$, even from cloth, for example, but it is quite hard to picture mentally.



We should warn that if a quotient space is made in an irresponsible manner, we can obtain a badly-behaved topology even if we start with a nice space. For example, the quotient $\mathbb{R}^2/B^2$ can be shown to be homeomorphic to $\mathbb{R}^2$, but $\mathbb{R}^2/(\operatorname{int} B^2)$ is not even Hausdorff. Generally speaking, under normal circumstances, only *closed* subspaces should be shrunk to a point, but even that does not always guarantee good behavior.

If A is a closed subspace of X that is *contractible*, examples suggest that X/A should be homotopy equivalent to X (why not homeomorphic?). This, unfortunately, is *not* true in general, but it works for cases one is likely to encounter. Technically, an assumption guaranteeing that $X/A \simeq X$ for contractible A is called the *homotopy extension property* of the pair (X, A). We will not define it here; it suffices to say, with a forward reference to the next section, that if X is a simplicial or CW complex and A is a contractible subcomplex, then $X/A \simeq X$ holds.

Join. While various products and quotients are encountered in many mathematical structures, joins appear more specific to topology (joins in lattices or in database theory are similar to joins in topology only by name). The join $X * Y$ of spaces X and Y is obtained by taking the Cartesian product $X \times Y$, "fattening" it by another product with $[0,1]$, and finally, collapsing the initial and final slices $X \times Y \times \{0\}$ and $X \times Y \times \{1\}$: in the former, each copy $X \times \{y\} \times \{0\}$ of X is collapsed to a point, while in the latter, the copies $\{x\} \times Y \times \{1\}$ of Y are collapsed. After these collapses, $X \times Y \times \{0\}$ becomes homeomorphic to Y, and $X \times Y \times \{1\}$ to X. Here is an illustration with X and Y segments:



The formal definition goes as follows.

Definition 6.3. *The join $X * Y$ of spaces X and Y is the quotient space $(X \times Y \times [0,1])/{\approx}$, where $\approx$ is given by $(x, y, 0) \approx (x', y, 0)$ for all $x, x' \in X$ and all $y \in Y$ ("for $t = 0$, x does not matter") and $(x, y, 1) \approx (x, y', 1)$ for all $x \in X$ and all $y, y' \in Y$ ("for $t = 1$, y does not matter").*

We observe that $X * Y$ contains the product $X \times Y$, e.g., as the "middle slice" $X \times Y \times \{\frac{1}{2}\}$. The join may look more complicated than the product, but in many respects it is better behaved; some of the advantages will be mentioned later.

There is a nice geometric interpretation of the join. Namely, suppose that X is represented as a bounded subspace of some $\mathbb{R}^m$, and Y of some $\mathbb{R}^n$. We then further insert $\mathbb{R}^m$ and $\mathbb{R}^n$ into $\mathbb{R}^{m+n+1}$ as skew affine subspaces, concretely $\{x \in \mathbb{R}^{m+n+1} : x_{n+1} = \cdots = x_{n+m+1} = 0\}$ and $\{y \in \mathbb{R}^{m+n+1} : x_1 = \cdots = x_n = 0, x_{n+1} = 1\}$ (so for $m = n = 1$ we have two skew lines in $\mathbb{R}^3$). With this placement of X and Y in $\mathbb{R}^{m+n+1}$ it can be verified that $X*Y$ is homeomorphic to the subspace $\bigcup_{x \in X, y \in Y} xy$ of $\mathbb{R}^{m+n+1}$, where xy is the segment connecting x and y. The point of placing X and Y into skew affine subspaces is to guarantee that two segments xy and $x'y'$, $x, x' \in X$, $y, y' \in Y$ never intersect, except possibly at one of the endpoints.

The join is commutative up to homeomorphism, but unfortunately not associative in general (although some of the literature claims so). For our purposes, though, it is amply sufficient that it is associative (up to homeomorphism of course) on the class of all compact Hausdorff spaces.

Cone and suspension. These are two popular special case of the join. The **cone** of a space X is $\mathrm{C}X := X * \{p\}$, the join with a one-point space. Geometrically, the cone is the union of all segments connecting the points of X to a new point. We can also write $\mathrm{C}X$ as another quotient space, simpler than the one for a general join: $(X \times [0,1])/(X \times \{1\})$.

One of the simple ways of proving contractibility of a space Y is to show that Y is the cone of another space.

The join with a two-point space, $X * S^0$, is called the **suspension** of X and denoted by $\mathrm{S}X$. It can be interpreted as erecting a double cone over X. (Readers who find S^0 as two-point space puzzling may want to think it over—S^0 is used quite frequently.)

Exercise 6.4. (a) *Show* $\mathrm{S}S^n \cong S^{n+1}$.

(b) *Prove* $S^k * S^\ell \cong S^{k+\ell+1}$. *Hint: use (a) and associativity of the join.*

While the cone operation makes every space homotopically trivial, i.e., contractible, the suspension more or less preserves the topological complexity, only pushing it one dimension higher. Very roughly speaking, it converts "k-dimensional holes" in X into "$(k+1)$-dimensional holes" in $\mathrm{S}X$.

6.1 Note on categorical definitions

The topology of the quotient $X/{\approx}$ can also be defined as the finest one for which the quotient map $q\colon X \to X/{\approx}$ is continuous. Here a topology $\mathcal{O}'$ is finer than $\mathcal{O}$ if $\mathcal{O} \subseteq \mathcal{O}'$. In the definition earlier we described explicitly what the open sets are, but the formulation just given is equivalent.

The definition of the product topology on the Cartesian product $X := \prod_{i \in I} X_i$ in Section 3 can be rephrased similarly using the *projection maps* $p_i\colon X \to X_i$, where p_i maps an $|I|$-tuple $(x_i)_{i \in I} \in X$ to its ith component

x_i. Namely, *the product topology is the coarsest topology on X that makes all of the p_i continuous* (a topology $\mathcal{O}$ is coarser than $\mathcal{O}'$ if $\mathcal{O} \subseteq \mathcal{O}'$) of open sets is inclusion-minimal among all topologies making the p_i continuous).

This is not only equivalent to the definition of Section 3, but it also explains one possibly ad-hoc looking aspect of that definition, namely, why we admit only finitely many nontrivial factors in the open rectangles.

Exercise 6.5. *Check the equivalence of both of the definitions of the product topology.*

Disjoint union. There is another, rather simple operation, which can be defined in a similar way. Namely, given a collection, finite or infinite, $(X_i)_{i\in I}$ of topological spaces, their **disjoint union** (or sometimes *disjoint sum*) $\coprod_{i\in I} X_i$ corresponds to the intuitive notion of putting disjoint copies of the X_i "side by side."

The ground set of $\coprod_{i\in I} X_i$ is the disjoint union of the sets X_i. Concretely, we may take $\bigcup_{i\in I} X_i \times \{i\}$, so that the elements of X_i are marked with i. This time we have the *inclusion maps* $\iota_i \colon X_i \to \coprod_{i\in I} X_i$, and the topology of the disjoint union is the finest one making all the ι_i continuous. Of course, it is not hard to describe the open sets explicitly as well: a set in $\coprod_{i\in I} X_i$ is open exactly if its intersection with each X_i is open.

The categorical approach. Here "categorical" is not related to Immanuel Kant but rather to the mathematical field of *category theory*, which studies general abstract structures in all mathematics.

Why do we feel obliged to say something about categories in an introductory text on topology? First, category theory was invented by algebraic topologists, it has greatly helped cleaning up some unmanageably complicated, and thus potentially wrong, proofs in topology, facilitated much progress in the field, and it is heavily used in topology both as a language and as a tool.

Second, even if one does not intend to learn much about category theory, there are several basic principles definitely worth knowing about. In almost any field of mathematics or computer science, even a little bit of category-theory thinking can prevent one from re-inventing the wheel, or from riding on octagonal wheels where round ones are available.

Objects and morphisms. One of the starting points of category theory is that *mappings* between mathematical objects deserve at least equal status as the objects. Moreover, knowing all mappings into an object and from it often gives enough information about the object, so that we need not consider the object's internal structure at all.

For example, in the category Top of topological spaces, we take all topological spaces as objects. We do not consider just any old mappings between spaces, but the "right" structural maps, namely, all *continuous maps.*

In category theory, the maps of the "right kind" for a given type of objects are called **morphisms**. When studying some type of mathematical objects, what the morphisms are is not God-given, but it is to be user-defined. But for many standard cases the morphisms are clear. For the category Set of sets they are arbitrary mappings, for the category Grp of groups they are group homomorphisms, and for the category Gra of (simple, undirected) graphs they are graph homomorphisms.

Exercise 6.6. *Recall as many mathematical structures as you can, and think what morphisms between them should be.*

The next conceptual step in creating the category Top of topological spaces is to forget what are the ground set and open sets of each space, and where individual points are sent by the various maps. What is left? Well, a (tremendously infinite) *directed multigraph.* The spaces are the vertices, and each morphism (continuous map) $f\colon X \to Y$ gives rise to one arrow from X to Y. Importantly, information about *composition of morphisms* is also retained: given two arrows $f\colon X \to Y$ and $g\colon Y \to Z$, we know which of the arrows $X \to Z$ corresponds to the composition gf.

In general, a category is just that, a directed multigraph with an associative composition rule (or, if you prefer an algebraic language, a partial monoid). In more detail, a **category** C consists of the following data:

- A class[3] $\mathrm{Ob}(\mathsf{C})$ of objects.
- For every two objects $X, Y \in \mathrm{Ob}(\mathsf{C})$, a class $\mathrm{Hom}(X,Y)$ of morphisms from X to Y (with $\mathrm{Hom}(X,Y) \cap \mathrm{Hom}(U,V) = \emptyset$ whenever $(X,Y) \neq (U,V)$).
- For every $X \in \mathrm{Ob}(\mathsf{C})$, a unique *identity morphism* $\mathrm{id}_X \in \mathrm{Hom}(X,X)$.
- A *composition law* assigning to every $f \in \mathrm{Hom}(X,Y)$ and $g \in \mathrm{Hom}(Y,Z)$ an $h \in \mathrm{Hom}(X,Z)$, written as $h = gf$.

The composition is required to be associative, $f(gh) = (fg)h$, and satisfies $f\,\mathrm{id}_X = \mathrm{id}_Y\, f = f$ for every $f \in \mathrm{Hom}(X,Y)$.

Surprisingly many properties and constructions can be expressed solely in terms of objects and morphisms. Take the concepts of injectivity, surjectivity, and isomorphism. In category theory, the counterparts are:

- A **monomorphism**, which is a left-cancellable morphism $f\colon X \to Y$, in the sense that $fg_1 = fg_2$ implies $g_1 = g_2$ for any two morphisms into X.
- An **epimorphism** is a right-cancellable morphism $f\colon X \to Y$, with $g_1 f = g_2 f$ implying $g_1 = g_2$.
- An **isomorphism** is a morphism $f\colon X \to Y$ that has a two-sided inverse; i.e., $g\colon Y \to X$ with $fg = \mathrm{id}_Y$ and $gf = \mathrm{id}_X$. An isomorphism is is both a monomorphism and an epimorphism, but these conditions are not sufficient in general.

[3] We cannot say set because of Russell's paradox. For example, if every set is an object of C, we cannot form the set of all sets, as Russell tells us. This is why the word *class* is used. Informally, a class is "like a set but possibly bigger"; for a mathematical foundation for working with classes see, e.g., [AHS06]. Categories whose class of objects is a set are called *small*.

Exercise 6.7. (a) *Check that in the category* Set*, monomorphisms and epimorphisms correspond to injective and surjective maps, respectively.*

(b) *Consider the category* Haus *of all* Hausdorff *topological spaces with continuous maps as morphisms. Let us consider the rationals* $\mathbb{Q}$ *as a subspace of* $\mathbb{R}$ *with the standard topology, and let* $f\colon \mathbb{Q} \to \mathbb{R}$ *be the standard inclusion. Is* f *an epimorphism? Can you characterize what epimorphisms are in this category?*

Products revisited. Products, for example, have a general categorical definition. Given objects X and Y in a category C, this definition identifies the product of X and Y, if one exists, up to isomorphism.

Namely, the **product** $X \times Y$ in C is an object P *plus* morphisms $p_X\colon P \to X$ and $p_Y\colon P \to Y$ with the following *universal property*: whenever P' is an object and $p'_X\colon P' \to P$ and $p'_Y\colon P' \to Y$ are morphisms, there is a unique morphism $f\colon P' \to P$ with $p'_X = p_X f$ and $p'_Y = p_Y f$. Or, expressed in a way category theorists and topologist prefer, there is a unique f making the following diagram commutative:

$$
\begin{tikzcd}
& P' \arrow[dl, "p'_X"'] \arrow[d, dashed, "f"] \arrow[dr, "p'_Y"] & \\
X & P \arrow[l, "p_X"'] \arrow[r, "p_Y"] & Y
\end{tikzcd}
$$

It is easy to see that such a P, if it exists, is unique up to isomorphism. The definition for the product of arbitrarily many objects is entirely analogous. As we have already indicated, not every category has products, but many do.

This definition may very well look nonintuitive and difficult to work with, and certainly it takes time and training to get used to that kind of reasoning. For specific categories, it may take some work to figure out what the product "looks like." On the other hand, the categorical approach maintains that once we know that a product exist, the defining property above is the only one we really need for working with it, and that we may never *need* to figure out the specific structure, especially if we are working in some less common category.

Exercise 6.8. (a) *Check that the product of topological spaces satisfies the categorical definition.*

(b) *Take* Gra*, graphs with graph homomorphisms. Describe the categorical product (for two graphs) concretely, in terms of vertices and edges.*

Limits. The product construction is a special case of categorical **limit**. That definition tells us what is the limit of a given (commutative) diagram in a given category C. Since we do not want to define diagrams in general, let us give just an example.

We consider three objects A, X, Y with morphisms $f\colon X \to A$ and $g\colon Y \to A$. The limit of the diagram

$$
\begin{tikzcd}
& X \arrow[d, "f"] \\
Y \arrow[r, "g"] & A
\end{tikzcd}
$$

is an object T plus morphisms $p_X : T \to X$ and $p_Y : T \to Y$ making the following digram commutative

$$\begin{tikzcd} T \arrow[r, "p_X"] \arrow[d, "p_Y"'] \arrow[dr, "p_A"] & X \arrow[d, "f"] \\ Y \arrow[r, "g"] & A \end{tikzcd}$$

and satisfying the universality property: whenever T' and p'_A, p'_X, p'_Y is another completion to a commutative diagram, there is a unique morphism $u: T' \to T$ such that $p'_A = p_A u$, $p'_X = p_X u$, and $p'_Y = p_Y u$.

For this particular diagram, the limit is called the *pullback*.

The same definition of a limit works for any commutative diagram in C; the morphisms p_X go from the limit object to every object in the diagram. The product is the special case of a limit where the diagram has just objects and no morphisms.

Exercise 6.9. *Work out what the pullback looks like in* Set.

Opposite category and conotions. For every category C we can immediately form a new category C^{op} by reversing all arrows. This, of course, would be highly problematic for actual mappings, since how should one invert a mapping that is not bijective, but it is no problem for a category theorist, who regards morphisms as abstract arrows.

For every categorical notion, we can form a "dual" notion by reversing all arrows. From product we get **coproduct**, which for topological spaces turns out to be just the disjoint union. (Here and in many other categories, the coproduct is rather dull, but for example, in the groups category Grp it is the *free product* of groups.) From limit we get *colimit*, etc., the prefix co- expressing the dual nature of the notion. (This terminology has some common sense exceptions, such as epimorphism instead of comonomorphism and *pushout* instead of copullback. But physicists may have missed an opportunity here with their *bra* and *ket* terminology.)

Category theory has a number of general constructions and theorems, and many concrete constructions get simplified by observing that they are but special realizations of these general abstract results. In topological and other proofs, references to such general categorical considerations are often (proudly) prefixed by the phrase "by abstract nonsense it follows that...."

7 Simplicial complexes and relatives

7.1 Simplicial complexes and simplicial maps

We have already touched upon the question, how can interesting topological spaces be described in a finite way? Simplicial complexes provide the simplest systematic way. Real topologists often frown on them and consider them old-fashioned as a theoretical tool and not economical enough compared to other tools. These are perfectly valid concerns, but for computer-science and combinatorial uses, simplicial complexes may often be the winners because of their combinatorial simplicity.

As a combinatorial object, a simplicial complex is simply a hereditary system of finite sets:

Definition 7.1. *A* **simplicial complex** *is a system K of finite subsets of a (possibly infinite) set V, with the property that if $F \in K$ and $F' \subset F$, then $F' \in K$ as well. The set V, called the* **vertex set** *of K and denoted by $V(K)$, is the union of all sets of K.*

In rare cases, it may be useful to also admit, unlike in the definition above, points of V that do not belong to any $F \in K$.

The definition implies, in particular, that $\emptyset \in K$ whenever $K \neq \emptyset$; in some of the literature, though, the empty set is not regarded as a member of K.

The sets in K are called the **simplices** of K. The vertex set is sometimes also called the **ground set**.

There is some formal ambiguity in using the term **vertex** of a simplicial complex: it may mean a point v of the vertex set V or a singleton set $\{v\}$, which is a simplex of K. But in practice this does not lead to confusion.

A **subcomplex** of a simplicial complex K is a simplicial complex $L \subseteq K$. We say that L is an **induced subcomplex** of K if $L = \{F \in K : F \subseteq V(L)\}$, i.e., every simplex of K living on the vertex set of L also belongs to L.

The **dimension** of a simplicial complex K is $\dim K := \sup_{F \in K}(|F| - 1)$. The "$-1$" in this definition is logical, of course, since, e.g., a three-point $F \in K$ will correspond to a geometric triangle, which is 2-dimensional, but it is an eternal source of potential confusion.

A useful example to keep in mind are 1-dimensional simplicial complexes, which can be regarded as simple graphs: the 0-dimensional simplices correspond to vertices and 1-dimensional ones to edges. Historically, the study of graphs has for some time been regarded as a part of topology.

Finite and infinite simplicial complexes. A simplicial complex is **finite** if it has a finite ground set. By definition, a simplicial complex can also be infinite, for a good reason: as we will see, finite simplicial complexes can describe only compact subspaces of some $\mathbb{R}^n$, which excludes spaces like $(0, 1)$ or $\mathbb{R}^n$ itself.

On the other hand, only finite simplicial complexes can naturally serve as inputs to algorithms, which was one of our main motivations for considering simplicial complexes. Moreover, for many purposes, including most of computer-science related applications, finite simplicial complexes suffice. Infinite simplicial complexes originally served as a theoretical tool for building algebraic topology, but in that role they have been replaced by other, more modern tools.

We will restrict ourselves to finite simplicial complexes, except for a couple of remarks.

Simplicial maps. By now the reader may be impatient to see what is the topological space described by a simplicial complex, but before explaining that, we will still want to say what are the appropriate maps (*morphisms* in the categorical jargon newly introduced above) between simplicial complexes.

Definition 7.2. *A* **simplicial map** *of a simplicial complex K into a simplicial complex L is a map $s\colon V(K) \to V(L)$ of the vertex sets that maps simplices to simplices, i.e., $s(F) \in L$ for every $F \in K$. An* **isomorphism** *of simplicial complexes is a bijective simplicial map with simplicial inverse.*

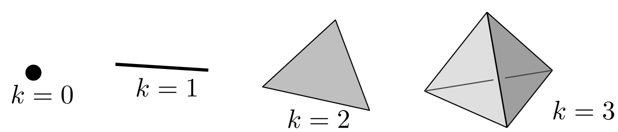
Isomorphism, similar to many other mathematical structures, means that the simplicial complexes have identical structure and differ only by renaming vertices.

We note that simplicial maps for 1-dimensional simplicial complexes are not the same as graph homomorphisms, since unlike homomorphisms, they allow for edges to be collapsed to vertices. But isomorphism is the same notion for graphs and 1-dimensional simplicial complexes.

7.2 Geometric realization and polyhedra

Now we want to say what the topological space described by a (finite) simplicial complex K is.

First we recall that a (geometric) **simplex** is the convex hull of a set of affinely independent points[4] in some $\mathbb{R}^n$; simplices of dimension 0, 1, 2, 3 are points, segments, triangles, and tetrahedra, respectively.

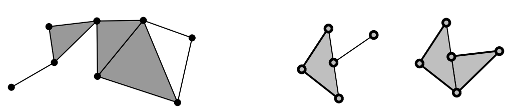


The *faces* of a simplex σ are the convex hulls of subsets of the vertex set. For example, a tetrahedron has 16 faces: itself, 4 triangles, 6 edges, 4 vertices, and the empty set. The faces of dimension one lower than σ are called the *facets* of σ; a k-dimensional simplex has $k+1$ facets.

Definition 7.3. *A* **geometric simplicial complex** *is a collection Δ of geometric simplices of various dimensions satisfying the following two conditions:*

(i) *(Hereditary) If $\sigma \in \Delta$ and σ' is a face of σ, then $\sigma' \in \Delta$.*
(ii) *(Intersecting in faces) For every $\sigma, \sigma' \in \Delta$, $\sigma \cap \sigma'$ is a face of both σ and σ'.*

Somewhat informally, the simplices in a geometric simplicial complex may be glued only along common faces:



GOOD BAD

[4] Points $p_0, p_1, \ldots, p_k \in \mathbb{R}^n$ ($k{+}1$ of them) are called **affinely independent** if the k vectors $p_1 - p_0, \ldots, p_k - p_0$ are linearly independent.

A geometric simplicial complex Δ defines a simplicial complex $K = K(\Delta)$ in the sense of Definition 7.1 in an obvious way: we set $V(K) = V(\Delta)$, the latter denoting the set of all vertices of the simplices in Δ, and the simplices of K are vertex sets of the simplices in Δ.

Now the geometric simplicial complex Δ is called a **geometric realization** of this K, and also of any simplicial complex K' isomorphic to K.

Proposition 7.4. *Every finite simplicial complex K has a geometric realization; if $k = \dim K$ then the realization can be taken in $\mathbb{R}^{2k+1}$.*

Sketch of proof. A geometric realization of K in some $\mathbb{R}^n$ is fully specified by the placement of the vertex set. Thus, we seek an (injective) mapping $\rho\colon V(K) \to \mathbb{R}^{2k+1}$.

The condition we need is that, for every two simplices $F, G \in K$, $\mathrm{conv}(\rho(F)) \cap \mathrm{conv}(\rho(G)) = \mathrm{conv}\,\rho(F \cap G)$, where $\mathrm{conv}(.)$ denotes the convex hull. A sufficient condition for this is that $\rho(F \cup G)$ be affinely independent, since then $\mathrm{conv}\,\rho(F \cup G)$ is a geometric simplex, both $\mathrm{conv}\,\rho(F)$ and $\mathrm{conv}\,\rho(G)$ are faces of it, and they intersect in the (possibly empty) face $\mathrm{conv}\,\rho(F \cap G)$ as they should.[5]

So it suffices to show that for every n there is an n-point set in $\mathbb{R}^{2k+1}$ in which every $2k + 2$ points are affinely independent (because $2k + 2$ is the maximum possible size of $F \cup G$). This we leave as an exercise for the readers not familiar with the trick. □

Exercise 7.5. *Verify that every $d + 1$ distinct points on the* moment curve *$\{(t, t^2, \ldots, t^d) : t \in \mathbb{R}\} \subset \mathbb{R}^d$ are affinely independent. Hint: a polynomial of degree at most d has at most d roots.*

Now, finally, we define the space associated with a simplicial complex.

Definition 7.6. *Let Δ be a geometric simplicial complex, and suppose that all simplices of Δ are contained in $\mathbb{R}^n$. The* **polyhedron** *of Δ is the topological subspace of $\mathbb{R}^n$ induced by the union of all simplices of Δ. A* **polyhedron** *of a finite simplicial complex K is the polyhedron of a geometric realization of K.*

The polyhedron of K is not defined uniquely, but as we will soon see, all polyhedra of K are homeomorphic. The polyhedron of K is usually denoted by $|K|$, but often one writes K for the polyhedron as well, and one has to distinguish from the context whether the combinatorial object or the geometric one is meant.

Remark on infinite simplicial complexes. As we have mentioned above, defining the polyhedron of an infinite simplicial complex is somewhat more demanding. An immediate trouble is that all of the geometric simplices may not fit in the same $\mathbb{R}^n$, for example if the dimension is unbounded.

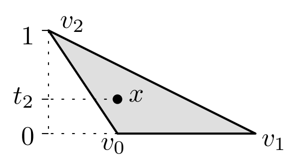
The solution uses quotient spaces. First we assign a k-dimensional geometric simplex $\rho(F)$ to every k-dimensional $F \in K$, possibly each $\rho(F)$ in a different

[5] Obvious as it may seem, this fact still needs a little proof, which we allow ourselves to omit. Here we are basically asserting that the set of all faces of a geometric simplex constitutes a geometric simplicial complex.

Euclidean space. Then we introduce a suitable equivalence relation $\approx$ on the disjoint union of these simplices, which amounts to identifying, for every $G \subset F$, the simplex $\rho(G)$ with the appropriate face of the simplex $\rho(F)$ (some care is needed in saying how exactly these identifications are performed; it is helpful to fix a linear ordering of the vertices of K first). Finally, $|K|$ is defined as the quotient of the disjoint union by $\approx$.

How simplicial maps yield continuous maps. Let K and L be a simplicial complexes, and let $s\colon V(K) \to V(L)$ be a simplicial map. There is a canonical continuous map $|s|\colon |K| \to |L|$ of the polyhedra associated to s.

One often says that $|s|$ is a linear extension of s on the simplices of $|K|$ (although, strictly speaking, it is an affine extension). To define $|s|$ precisely, we need to recall that if σ is a geometric simplex with vertices $v_0, \ldots, v_k$, then every point $x \in \sigma$ can be uniquely written as $\mathbf{x} = \sum_{i=0}^{k} t_i v_i$, where $t_0, \ldots, t_k \geq 0$ and $\sum_{i=0}^{k} t_i = 1$. Here $(t_0, \ldots, t_k)$ is called the *barycentric coordinates* of x; t_i is the height of x above the facet of σ not containing v_i, scaled so that v_i has height 1:



So let us fix geometric realizations Δ and Δ' of K and L, respectively, and regard s as a map $V(\Delta) \to V(\Delta')$. For a point x in the polyhedron of Δ we choose a lowest-dimensional simplex σ containing x (such a σ is called the **support** of x in Δ and it is determined uniquely). We have $x = \sum_{i=0}^{k} t_i v_i$, where $v_0, \ldots, v_k$ are the vertices of σ, and we set

$$|s|(x) := \sum_{i=0}^{k} t_i s(v_i).$$

The sum is well defined because, by the definition of a simplicial map, $\{s(v_i) : i = 0, 1, \ldots, k\}$ is the vertex set of some simplex in Δ'.

Note that, since simplicial maps are allowed to map higher-dimensional simplices to lower-dimensional ones, the image of a k-dimensional simplex under $|s|$ may have any dimension $\ell \leq k$.

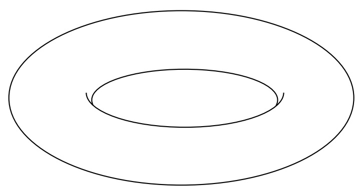
One needs to check that $|s|$ is continuous when we go from the interior of some simplex towards a point of a facet, but this is straightforward.

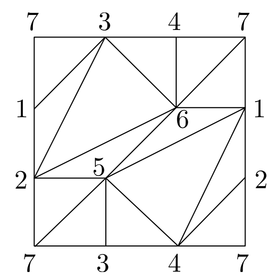
It is also not hard to see that if s is injective, then so is $|s|$, and if s is an isomorphism, then $|s|$ is a homeomorphism. From this we immediately get that isomorphic simplicial complexes have homeomorphic polyhedra. In particular, the polyhedron of a simplicial complex is uniquely defined, upto a homeomorphism.

Triangulations. A simplicial complex K is called a **triangulation** of a space X if $X \cong |K|$. Naturally not all topological spaces possess a triangulation: some for reasons of local pathology, such as not being Hausdorff, but some others are non-triangulable in spite of being locally very nice. The perhaps most striking example is a 4-dimensional compact manifold (the *Freedman E8 manifold*; manifolds will be introduced later).

The simplest triangulation of the sphere S^{n-1} is the boundary of an n-dimensional simplex, with n simplices of dimension $n-1$ but $2^n - 1$ simplices in total. Combinatorially, denoting the vertices by $1, 2, \ldots, n$, the simplicial complex is $\{F \subseteq [n] : F \neq [n]\}$. Another, more symmetric triangulation will be mentioned soon.

It can be shown that every triangulation of the torus must have at least 7 vertices and at least 14 triangles, and here is one attaining these minimal numbers:





The triangulation is drawn as a square, but the sides of the square should be identified as in Exercise 6.2(c)—this is also indicated by the numbering of the vertices.

It may be worthwhile if the reader draws her own triangulation of a torus, trying to get a small number of triangles, and notes the pitfalls in such an enterprise.

The study of triangulations is a major and fast-growing area, but here we leave it aside, referring to [DLRS10].

Simplicial joins. The join operation can also be done on the level of simplicial complexes in a straightforward way.

Let K, L be simplicial complexes, and first assume $V(K) \cap V(L) = \emptyset$. Then the simplicial join $K * L$ is $\{F \cup G : F \in K, G \in L\}$, on the vertex set $V(K) \cup V(L)$. If the vertex sets are not disjoint, we must first replace L, say, with an isomorphic simplicial complex whose vertex set is disjoint from $V(K)$.

It is not hard to show that $|K * L| \cong |K| * |L|$. The main step is in checking that the join of a geometric k-simplex and ℓ-simplex is a $(k + \ell + 1)$-simplex, which is easy using the interpretation of join with skew affine subspaces.

We saw (Exercise 6.4) that $S^n \cong (S^0)^{*(n+1)}$, the $(n+1)$-fold join of the 0-dimensional sphere, or two-point space. If we do this join simplicially, we obtain the following triangulation of S^n: the vertex set is $\{a_1, b_1, \ldots, a_{n+1}, b_{n+1}\}$, and a set of vertices forms a simplex exactly if it does not contain a pair $\{a_i, b_i\}$ for any i.

The geometric realization is the boundary of the *crosspolytope*, a regular octahedron for $n = 2$ (just identify a_i with e_i, the ith vector of the standard basis of $\mathbb{R}^{n+1}$, and b_i with $-e_i$). One often speaks of the *octahedral sphere.*

7.3 Combinatorial examples

A great feature of simplicial complexes is that they give a way of assigning a topological space to all kinds of combinatorial objects: whenever we have a system of finite sets, we can close it under taking subsets, if it is not yet hereditary by itself, and we have a simplicial complex. Sometimes this topology connection is fruitful, sometimes not so much, but definitely there is something to study. We list several cases where this approach has been used with great success; many others can be found in the literature.

Clique (or flag) complexes of graphs. Given a (simple, undirected) graph G, we define a simplicial complex $C = C(G)$ on the vertex set $V(G)$ whose simplices are sets of vertices forming a clique (every two vertices connected by an edge). This C has several common names: **clique complex**, **flag complex**, *Whitney complex*, and probably others. A similar complex $I(G)$ whose simplices are independent sets in G is the **independence complex** of G.

Clique complexes, besides constituting an interesting subclass of simplicial complexes, carry lot of information about the underlying graph in their topology. They feature, for example, in a proof by Meshulam [Mes01] of a generalization of a lovely theorem of Aharoni and Haxell [AH00], a Hall-type theorem for hypergraphs. The Aharoni–Haxell theorem was later used for proving a tantalizing combinatorial conjecture known as the *tripartite Ryser conjecture* [Aha01]. In computational geometry one finds a special case of the clique complex as the *Vietoris–Rips complex*.

Order complex. Let $(X, \preceq)$ be a partially ordered set. Its **order complex** lives on the vertex set X and the simplices correspond to *chains*, i.e., subsets of X linearly ordered by $\preceq$. (Equivalently, this is the clique complex of the comparability graph of $(X, \preceq)$.)

There is an extensive topological theory of order complexes; see, e.g., [Bjö95, Wac07]. For example, there is a famous fixed-point theorem for posets of Baclawski and Björner, which has a combinatorial statement but only topological proofs. Deep connections were found to questions in algebra and Lie theory.

Nerve. Let $\mathcal{F} = \{F_1, F_2, \ldots, F_n\}$ be a family of sets, so far arbitrary. The **nerve** $\mathcal{N}(\mathcal{F})$ of $\mathcal{F}$ is the simplicial complex with vertex set $[n]$, the set indices, and simplices corresponding to nonempty intersections:

$$\mathcal{N}(\mathcal{F}) = \Big\{ I \subseteq [n] : \bigcap_{i \in I} F_i \neq \emptyset \Big\}.$$

Here is a remarkable theorem.

Theorem 7.7 (Nerve theorem)**.** *Let $K_1, K_2, \ldots, K_n$ be subcomplexes of a finite simplicial complex K that together cover K (each simplex of K is in at least one K_i). Suppose that the intersection $\bigcap_{i \in J} |K_i|$ is empty or contractible for each nonempty $J \subseteq [n]$. Then the (polyhedron of the) nerve of $\{|K_1|, \ldots, |K_n|\}$ is homotopy equivalent to $|K|$.*

This often allows one to simplify a simplicial complex drastically while keeping the homotopy type.

There are many variations of the nerve theorem in the literature. For example, a useful geometric setting where a nerve theorem holds is when $F_1, \ldots, F_n$ are closed convex sets in a Euclidean space. This result, and various generalizations, are closely related to Helly-type theorems in geometry (see, e.g., [CdVGG12] for recent progress).

The usual proofs of nerve theorems use somewhat more machinery than we are going to develop here, so we refer, e.g., to Björner [Bjö03] for a relatively simple and elementary proof which also yields a powerful generalization.

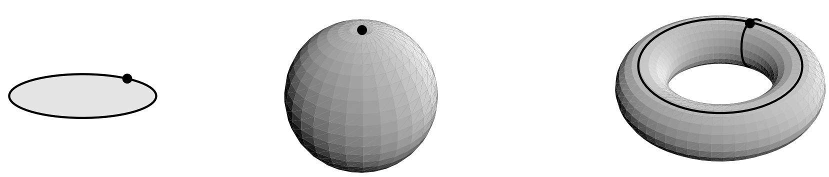
7.4 Simplicial sets and cell complexes

In geometric simplicial complexes, the simplices can be glued only in a somewhat rigid face-to-face manner. For example, we are not allowed to glue two triangles by two sides without actually making them identical, although geometrically, such a construction makes perfect sense. Other descriptions of spaces used in topology allow for more flexible gluing, and some of them use building blocks other than simplices.

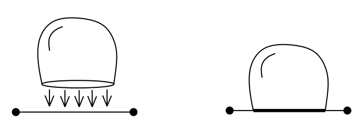
Cell (or CW) complexes. This is perhaps the most popular way in topology, but it does not provide a finite description of a space, so we mention it only briefly.

The building blocks here are topological balls of various dimensions, called *cells*, which can be thought of as being completely "flexible" and which can be glued together in an almost arbitrary continuous fashion. Essentially the only condition is that each n-dimensional cell has to be attached along its boundary to the $(n-1)$-*skeleton* of the space, i.e., to the part that has already been built, inductively, from lower-dimensional cells.

Here are some pictorial examples:



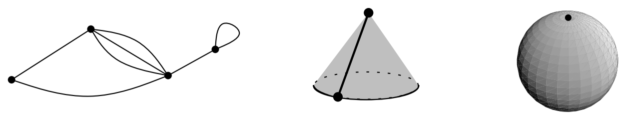
The left one is a B^2 made of a 0-cell (point), 1-cell, and 2-cell. The middle shows an S^2 made from one 0-cell and one 2-cell (whose boundary is shrunk to a point), and the right picture is the torus from a 0-cell, two 1-cells, and a single 2-cell. It is also perfectly legal to glue something to the middle of an edge, for example:



We refer to standard textbooks for a formal definition of a cell complex.

Simplicial sets. Although the name may suggest the opposite, simplicial sets are more complicated than simplicial complexes.

Intuitively, a simplicial set, similar to simplicial complexes, is made by gluing simplices of various dimensions together. The gluing is still face-to-face but much more permissive than for simplicial complexes. For example, one may have several 1-dimensional simplices connecting the same pair of vertices, a 1-simplex forming a loop, two edges of a 2-simplex identified to create a cone, or the boundary of a 2-simplex all contracted to a single vertex, forming an S^2.



The features mentioned so far are also shared by *Delta-complexes*, as used, e.g., in Hatcher [Hat01]. They ensure both the possibility of a finite combinatorial description (assuming finitely many simplices), and also considerable economy of description, at least in small cases: the torus can be described as a simplicial set with two triangles, as opposed to the minimum of 14 for simplicial complexes, or S^n needs only one 0-dimensional and one n-dimensional simplex, as opposed to at least 2^n simplices for a simplicial complex.

There is still another feature of simplicial sets, which looks peculiar at first sight: *degenerate simplices.* If σ is a simplex in a simplicial set X, say two-dimensional with vertices v_0, v_1, v_2, then X also contains the degenerate simplex, denoted by $s_0\sigma$, which we can think of as simplex that geometrically coincides with σ but in which the vertex v_0 has been duplicated (and it is actually considered 3-dimensional). Not only that, X also has to contain $s_1\sigma$, $s_2\sigma$, $s_0 s_0\sigma$, and so on, degeneracies of σ of all possible orders.

We thus see that every nonempty simplicial set must have infinitely many simplices, which seems to ruin the purpose of finite description. Fear not: the degenerate simplices can be represented implicitly, since, as it turns out, every degenerate simplex can be specified by some nondegenerate starting simplex and a canonical sequence of the degeneracy ("vertex-duplicating" operators) s_i. So the degenerate simplices need not be stored explicitly. If a simplicial set has finitely many nondegenerate simplices, then it has a (quite efficient) finite encoding; this holds, in particular, when a finite simplicial complex is converted to a simplicial set.

After this misty introduction, what about a formal definition of a simplicial set? The preferred one in modern texts goes as follows: *A simplicial set is a contravariant functor $\Delta \to \mathsf{Set}$, where Δ denotes the category of nonempty finite linearly ordered sets, with (non-strictly) monotone maps as morphisms.*

This kind of definition may be one of the reasons why people find modern algebraic topology inaccessible. Even if one knew what a contravariant functor is (which will be mentioned here later on), it seems rather hard without long training to make any intuitive sense of this definition, and see how it may correspond to the things discussed informally above (of course, there may be exceptions among our readers).

Fortunately, there is a friendly pictorial treatment of simplicial sets by Friedman [Fri12], where one can also find a still rigorous but more descriptive reformulation of the above concise definition.

This section can be taken as an advertisement of simplicial sets, which, unlike simplicial complexes, seem to remain almost unknown in computer science. One of their remarkable uses is theoretical.

According to a theory worked out mainly by Kan, using sufficiently rich (necessarily infinite) simplicial sets, one can capture all homotopy classes of continuous maps between spaces by simplicial maps among the corresponding simplicial sets, similarly for homotopy classes of homotopies, etc. In this way, "continuous" homotopy theory can be imitated purely discretely in the category of simplicial sets with simplicial maps, and these ideas also allow one to do homotopy theory in, say, algebraic categories with no notion of continuity.

Simplicial sets do not appear outside topology with the spontaneity of simplicial complexes, yet they have also found impressive algorithmic uses. They constitute the main data structure in algorithms for computing homotopy groups or homotopy classes of continuous maps. We refer to [RS12] for nice lecture notes on this subject and to [ČKM$^+$14] for a sample of such algorithms.

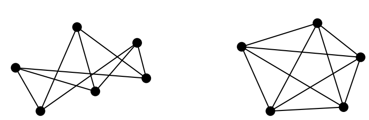
8 Non-embeddability

One of the first questions addressed in basic graph theory is, which graphs are planar? A planar graph is one that can be drawn in the plane without edge crossings. In topological terms, we have a 1-dimensional simplicial complex G and we ask whether its polyhedron $|G|$ can be embedded in $\mathbb{R}^2$.

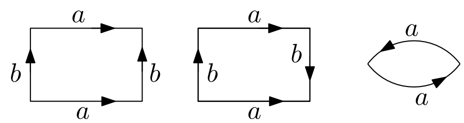
In general, if X and Y are topological spaces, an **embedding** of X in Y is a mapping $f : X \to Y$ that is a homeomorphism of X with the image $f(X)$. In other words, we are looking for a subspace of Y homeomorphic to X.

Let us remark that if X is compact Hausdorff, as is the case for polyhedra of finite simplicial complexes and for finite graphs in particular, then we only need to look for an injective continuous map $X \to Y$; the inverse is continuous automatically.

Graph theory has a number of very satisfactory answers to the question above. One of them is Kuratowski's theorem, asserting that a graph is planar if and only if it does not contain a subdivision of $K_{3,3}$ or K_5, which are thus the basic nonplanar graphs:



The planarity question has obvious generalizations, where we consider higher-dimensional simplicial complexes K and we ask whether they embed in $\mathbb{R}^d$ for some given d. (Of course, one can be even much broader and go beyond Euclidean spaces as targets, etc.) One of the classical facts very often mentioned when discussing 2-dimensional surfaces is that while the torus, whose quotient-space construction is recalled below on the left, obviously embeds in $\mathbb{R}^3$, neither the *Klein bottle* (middle picture) nor the *projective plane* (right) do.



Various cases of the embeddability question has been an important topic in topology for many decades, and lots of very interesting partial results are known. However, it is clear that Kuratowski's theorem is a great positive exception—there cannot be any comparably conclusive results about embeddability of k-dimensional simplicial complexes in $\mathbb{R}^d$ for any $k \geq 2$. In particular, it is known that the algorithmic question, does a given finite 4-dimensional simplicial complex embed in $\mathbb{R}^5$, is *undecidable*, and so there is no hope at reasonable characterization theorems.

There is another point which should be stressed before we leave this general introduction. For planar graphs, a remarkable theorem of Fáry tells us that every planar graph has a straight-edge planar drawing, or in other words, that every 1-dimensional simplicial complex embeddable in $\mathbb{R}^2$ also has a geometric realization in $\mathbb{R}^2$, in which the simplices are straight.

For simplicial complexes of dimension 2 and higher, these two notions (embeddability in $\mathbb{R}^d$ and having a geometric realization in $\mathbb{R}^d$), are completely different in general. Here we will talk exclusively about (topological) embeddability.

In the rest of the section we want to demonstrate a single nonembeddability result, where we show an interesting general method, as well as some of the notions and tools introduced earlier in action.

Van Kampen–Flores complexes. We saw in Proposition 7.4 that every k-dimensional finite simplicial complex K embeds in $\mathbb{R}^{2k+1}$. Here we will prove a complementary and classical result.

Theorem 8.1 (Van Kampen; Flores)**.** *For every $k = 1, 2, \ldots$, there are k-dimensional finite simplicial complexes that cannot be embedded in $\mathbb{R}^{2k}$.*

Thus, the bound of $2k+1$ of Proposition 7.4 cannot be improved in general. We can also see that the question of embeddability of a k-dimensional simplicial complex in $\mathbb{R}^d$ is nontrivial in general for $k \leq d \leq 2k$.

Let D_3 denote the simplicial complex with three isolated vertices. The complex for which we prove non-embeddability in $\mathbb{R}^{2k}$ is the $(k+1)$-fold join $D_3^{*(k+1)}$. More graphically, we think of the vertex set of $D_3^{*(k+1)}$ as $k+1$ rows with 3 vertices each, and the simplices are the subsets that use at most one vertex per row. In particular, for $k = 1$ we get $K_{3,3}$, one of the two Kuratowski graphs.

An abstract version of antipodality. In order to prove non-embeddability of a simplicial complex K in $\mathbb{R}^d$, we consider a hypothetical continuous injective map $f\colon |K| \to \mathbb{R}^d$, and we want to derive a contradiction. But the injectivity condition, $f(x) \neq f(y)$ for $x \neq y$, does not appear very suitable to work with directly.

Instead, we are going to use a surprising trick of general importance: from $|K|$ and from $\mathbb{R}^d$ we construct new, more complicated-looking spaces, and the given f yields a mapping between these new spaces. The key advantage is that this new mapping satisfies a more global and convenient condition: it is antipodal in a suitable sense, and for our particular simplicial complex $D_3^{*(k+1)}$ and for $d = 2k$ we will be able to use the Borsuk–Ulam theorem to conclude that the new map cannot exist.

The first thing to do is generalizing the definition of antipodal maps suitably. The definition of antipodal maps $\mathbb{R}^m \to \mathbb{R}^n$ given earlier relies on the particular map, one for each $\mathbb{R}^n$, sending x to $-x$. The important properties are that this map is a self-homeomorphism of $\mathbb{R}^n$, and that applying it twice gives the identity.

We thus define a $\mathbb{Z}_2$**-space** as a pair (X, ν), where X is a topological space and $\nu\colon X \to X$ is a homeomorphism with $\nu\nu = \mathrm{id}_X$. (Here $\mathbb{Z}_2$ refers to the (only) two-element group $\{0, 1\}$ with addition modulo 2, indicating that our considerations could be generalized to G-spaces with G a finite group or a topological group, but we will not pursue this direction in this introductory treatment.) The homeomorphism ν is often called the $\mathbb{Z}_2$*-action* of the considered $\mathbb{Z}_2$-space.

With $\mathbb{Z}_2$-spaces as objects, we also want to define the corresponding maps (morphisms). A $\mathbb{Z}_2$**-map** between $\mathbb{Z}_2$-spaces (X, ν) and (Y, ω) is a continuous map $f\colon X \to Y$ with $f\nu = \omega f$; this is the analog of $f(-x) = -f(x)$ for antipodality.

Deleted product and the Gauss map. We return to the earlier setting with an injective continuous $f\colon |K| \to \mathbb{R}^d$. From the space $|K|$ we construct the **deleted product** $|K|^2_\Delta$, where for an arbitrary space X we have

$$X^2_\Delta := \{(x, y) : x, y \in X, x \neq y\}.$$

This is a subspace of the product $X \times X$, and the subscript Δ should suggest that we delete the *diagonal* $\Delta := \{(x, x) : x \in X\}$. Moreover, we can naturally consider X^2_Δ as $\mathbb{Z}_2$-space, with the $\mathbb{Z}_2$-action $(x, y) \mapsto (y, x)$.

Based on f, we define a $\mathbb{Z}_2$-map $\tilde{f}\colon |K|^2_\Delta \to S^{d-1}$, where the sphere on the right is considered with the usual antipodality $x \mapsto -x$, as follows:

$$\tilde{f}(x, y) := \frac{f(x) - f(y)}{\|f(x) - f(y)\|}.$$

This is sometimes called the **Gauss map**; note that it is well defined only because the deleted product contains only pairs with $x \neq y$ and because f is assumed to be injective..

We have thus arrived at the following sufficient condition for non-embeddability:

Proposition 8.2. *Let K be a finite simplicial complex, and suppose that there is no $\mathbb{Z}_2$-map $|K|^2_\Delta \to S^{d-1}$. Then $|K|$ cannot be embedded in $\mathbb{R}^d$.*

There is no obvious reason to expect the condition to be also necessary; after all, it is hard to imagine why *every* possible $\mathbb{Z}_2$-map should look like the Gauss map for some embedding f. The condition indeed may not always be necessary, but remarkably enough, for a wide range of parameters it turns out to be necessary. This is the statement of a celebrated theorem of Haefliger an Weber, which asserts that, for $\dim K \le \frac{2}{3}d - 1$, the proposition actually holds as equivalence: $|K|$ embeds in $\mathbb{R}^d$ if and only if a $\mathbb{Z}_2$-map $|K|^2_\Delta \to S^{d-1}$ exists. The proof, unlike for Proposition 8.2, is difficult and we will not say anything more about it.

Deleted joins. The reader might reasonably expect that we will now use Proposition 8.2 to establish the Van Kampen–Flores theorem, but the problem with this is that the deleted product of our specific complexes is not so easy to work with, at least by elementary means. Instead, we first derive a variant of the proposition in which deleted *products* are replaced by deleted *joins*.

We need an analogy of the Gauss map for joins: a $\mathbb{Z}_2$-map from the twofold join X^{*2}, with something like the diagonal deleted, into a sphere, a reasobable guess being S^d instead of S^{d-1} as before, since the join has dimension one larger than the product. We recall that points of X^{*2} are triples (x, y, t), $x, y \in X$, $t \in [0, 1]$, with appropriate identifications for $t = 0$ and $t = 1$. After some experimenting one can arrive at the following Gauss-like map formula:

$$\tilde{f}(x, y, t) := \frac{(tf(x) - (1-t)f(y), 2t - 1)}{\|(tf(x) - (1-t)f(y), 2t - 1)\|} \in S^d.$$

On the right-hand side, $(tf(x) - (1-t)f(y), 2t - 1)$ is a $(d+1)$-component vector with the first d components given by $tf(x) - (1-t)f(y)$ and the last one by $2t - 1$. The formula is well-defined unless both $tf(x) = (1-t)f(y)$ and $2t - 1 = 0$, or in other words, unless $f(x) = f(y)$ and $t = \frac{1}{2}$. So it is sufficient to delete all triples $(x, x, \frac{1}{2})$ from X^{*2}, which gives us a (somewhat ad-hoc) notion of deleted join.

Some care is needed to check continuity of $\tilde{f}$; we need to see that it respects the identifications in the definition of the join as a quotient space. Actually, the definition of $\tilde{f}$ has been reverse-engineered to obey these identifications. Finally, we can see that $\tilde{f}$ is a $\mathbb{Z}_2$-map if the $\mathbb{Z}_2$-action on X^{*2} is given by $(x, y, t) \mapsto (y, x, 1 - t)$, a natural choice.

Simplicial deleted join. In our setting, where $X = |K|$ is the polyhedron of a simplicial complex, the join $|K|^{*2}$ is, as we recall, the polyhedron of the simplicial complex K^{*2}. However, deleting the points $(x, x, \frac{1}{2})$ as above destroys this structure; look, for example, what happens if K is a segment (1-dimensional simplex).

We will thus define a simplicial version of the deleted join, in which we delete more points but keep a nice simplicial structure. Let K' and K'' be two vertex-disjoint copies of K, and for a simplex $F \in K$, let us write F' and F'' for the corresponding simplices in K' and K'', respectively. Then the simplicial deleted join of K is

$$K^{*2}_\Delta := \{F' \cup G'' : F, G \in K, F \cap G = \emptyset\}.$$

We thus delete all joins of intersecting pairs of simplices.

The polyhedron $|K^{*2}_\Delta|$ is again a $\mathbb{Z}_2$-space, with the same $\mathbb{Z}_2$-action as the one for the join. Moreover, it is contained, usually strictly, in the deleted join of $|K|$ as a space, as announced. Instead of a general proof, we invite the reader to work out an example or two, to see what is going on.

Exercise 8.3. *Describe/visualize the deleted join of the* 3*-cycle and of the* 4*-cycle (understood as* 1*-dimensional complexes). Check that it does not contain any points of the usual join of the form* $(x, x, \frac{1}{2})$*.*

By the above considerations, we thus have the following analog of Proposition 8.2 with deleted joins.

Proposition 8.4. *Let* K *be a finite simplicial complex, and suppose that there is no* $\mathbb{Z}_2$*-map* $|K^{*2}_{\Delta}| \to S^d$*. Then* $|K|$ *cannot be embedded in* $\mathbb{R}^d$*.*

Proof of Theorem 8.1. We need to understand the deleted join $(D_3^{*(k+1)})^{*2}_{\Delta}$.

Exercise 8.5. (a) *Work out what the simplicial deleted join of* D_3 *is (the* $k = 0$ *case).*

(b) *Check that the join and deleted join commute:* $(K * L)^{*2}_{\Delta} \cong K^{*2}_{\Delta} * L^{*2}_{\Delta}$*. Using this, check that* $(D_3^{*(k+1)})^{*2}_{\Delta}$ *is isomorphic to the join of* $k+1$ *copies of* S^1 *(represented as* 6*-cycles), and hence its polyhedron is an* S^{2k+1}*; see Exercise 6.4.*

Now we know that the deleted join is homeomorphic to S^{2k+1}. With some more care, it can be verified that the homeomorphism obtained in this way is also a $\mathbb{Z}_2$-map; we omit this part here, since it is not very instructive (moreover, it can be shown that every two $\mathbb{Z}_2$-actions on S^n without fixed points, i.e., with $\nu(x)$ never equal to x, are equivalent, in the sense that there is a $\mathbb{Z}_2$-map between the resulting $\mathbb{Z}_2$-spaces in both directions).

Thus, a $\mathbb{Z}_2$-map of the deleted join into S^{2k} yields an antipodal map $S^{2k+1} \to S^{2k}$, which contradicts the Borsuk–Ulam theorem (Theorem 5.1(iii)). $\square$

9 Homotopy groups

We start introducing two fundamental concepts of algebraic topology: *homotopy groups* of a space on the one hand, and *homology groups* and the closely related *cohomology groups* on the other hand.

Rudiments of these notions go back to a theory of integration along curves in the complex plane in the 19th century. Homotopy and homology groups were explicitly introduced by Poincaré at the beginning of the 20th century, in a form somewhat different from the modern one. They were the key concepts that have made a large part of topology algebraic, in the sense of associating algebraic structures with topological spaces.

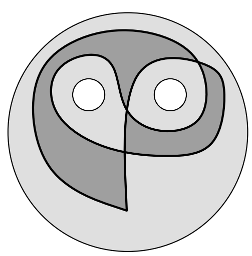
For both homotopy and homology, we can take as a starting point the question, what makes a disk topologically different from a disk with a hole, i.e. an annulus A? The first impulse may be to say that A has a hole and a disk does not, but then, what is a hole? Apparently in the annulus case it is something outside A, and to think of such a hole, one must imagine the annulus as a subspace of something else, say the plane. But what happens to the hole if we think of A as the surface of a cylinder in $\mathbb{R}^3$, for example? We need an intrinsic notion, talking only about the space itself.

Both in homotopy and homology, we look at loops in the considered space X. From the homotopy point of view, we consider a loop as a mapping $S^1 \to X$. Every loop in the disk is nullhomotopic, i.e., can be continuously shrunk to a point, but the annulus has nontrivial loops, ones that are not nullhomotopic (left picture):

nullhomotopic   not nullhomotopic   a boundary   not a boundary

On the other hand, thinking homologically, we consider a loop just as the image of the map $S^1 \to X$, i.e., a point set,[6] and we ask, is the considered loop the boundary of something 2-dimensional[7] in X? Again, every loop is a boundary in the disk, but not so in the annulus—see the right picture (these claims should be quite intuitive, but we have not proved them).

Here is an example showing that these two notions of triviality of a loop, being nullhomotopic and being a boundary, are different. We believe that experiencing this early and hands-on is important for developing some intuition about homology later on. So now we consider a disk with two holes, and the self-intersecting loop as in the picture:



The loop is the boundary of the indicated region. But it is not nullhomotopic; again we do not prove this but we can perhaps recommend the reader a physical experience with a string and a bar with two sticks or something similar.

Pointed everything. We postpone further discussion of homology and focus on homotopy groups.

The idea of looking at homotopy classes of loops in the considered space still needs a refinement: we want only loops that begin and end in some distinguished point, called the *basepoint*, $x_0 \in X$, since such loops can be composed. The

[6] This applies to "baby" homology over a 2-element field, which we will mostly deal with in this text. In more "grown-up" homology with integer coefficients, normally used in textbooks and in many applications, we would need to consider loops also with orientation and possibly with multiplicity.

[7] Here the meaning of "boundary" is somewhat different from the definition in general topology; we want the boundary of a 2-dimensional disk in $\mathbb{R}^3$ to be a circle.

composition of loops a and b is the loop $a.b$ obtained by first traversing a and then b. More formally, if a is represented by a map $a\colon [0,1] \to X$ with $a(0) = a(1) = x_0$, and similarly for b, then $c = a.b$ is given by

$$c(t) = \begin{cases} a(2t) & 0 \le t \le \frac{1}{2} \\ b(2t-1) & \frac{1}{2} \le t \le 1. \end{cases}$$

In order to do everything consistently, it is good to work in the category of **pointed spaces**. A pointed space is a pair (X, x_0), where X is a topological space and $x_0 \in X$ is a basepoint. The appropriate morphisms are **pointed maps**, i.e., continuous maps sending the basepoint to the basepoint. Here are some other "pointed" notions:

- A **pointed homotopy** H of pointed maps $f, g\colon (X, x_0) \to (Y, y_0)$ is a homotopy that, moreover, fixes the basepoint all the time, i.e., $H(x_0, t) = y_0$ for all $t \in [0,1]$. Let $[f]_*$ denote the pointed homotopy class of a pointed map f (the star usually refers to the pointed setting), and let $[(X, x_0), (Y, y_0)]_*$ be the set of all such classes for given (X, x_0) and (Y, y_0).
- The **wedge** $X \vee Y$ of pointed spaces X and Y is obtained by taking the disjoint union and identifying the basepoints (and similarly for any number of spaces). This is actually the coproduct in the category of pointed spaces (while the product has an obvious unique basepoint, and so no change is needed).
- More as an interesting illustration than something we would actually need, we mention that the pointed analog of the suspension $\mathrm{S}X$ is the *reduced suspension* ΣX, obtained from $\mathrm{S}X$ by collapsing the two segments above the basepoint of X to a new basepoint (we are thinking of $\mathrm{S}X$ as a double cone over X).

Let us remark that it is very convenient and technically useful to extend pointed spaces to **pairs** (X, A), where X is a space and A is a subspace of X (usually assumed to sit nicely in X, say as a subcomplex of a finite simplicial complex). A *map of pairs* $f\colon (X, A) \to (Y, B)$ is a continuous map $X \to Y$ with $f(A) \subseteq B$—a very simple concept but with large expressive power, and a basis of notions such as *relative homotopy or homology groups*, which are important tools for working with the usual homotopy and homology groups. For example, now we can also think of a pointed map $S^1 \to X$ as a map of pairs $([0,1], \{0,1\}) \to (X, \{x_0\})$.

The fundamental group. We are ready to introduce the first homotopy group $\pi_1(X)$ of a pointed space, one of the great inventions of Poincaré, also called the *fundamental group*.

Definition 9.1. *Let (X, x_0) be a pointed space. The* **fundamental group** *$\pi_1(X, x_0)$ has the ground set $[(S^1, s_0), (X, x_0)]_*$ of pointed homotopy classes of pointed loops in X, and the group operation is given by composition of loops, i.e., $[a]_*[b]_* = [a.b]_*$.*

Exercise 9.2. (a) *Check that the operation is well defined.*

(b) *Show that it is associative (a proper picture makes this quite obvious).*

(c) *Take a disk with two holes and find an example of loops a, b witnessing noncommutativity of the fundamental group. The example should be just informal since we have not built the tools for showing nontriviality of any loop.*

The basepoints in the notation are annoying and one would like to get rid of them as soon as possible. Unfortunately, for different basepoints x_0, x_0' the groups $\pi_1(X, x_0)$ and $\pi_1(X, x_0')$ are certainly not equal (they are disjoint as sets). But, under reasonable circumstances, they are at least isomorphic, which gives us a good enough reason to ignore the basepoint.

Exercise 9.3. *Let X be a space, and let x_0, x_0' be two points connected by a path γ (a map $[0,1] \to X$ with $\gamma(0) = x_0$, $\gamma(1) = x_0'$). Exhibit an isomorphism $\pi_1(X, x_0) \cong \pi_1(X, x_0')$. Hint:* The Hobbit*, full title.*

A space X with a trivial fundamental group (i.e. with no nontrivial loops), which is usually written as $\pi_1(X) = 0$, is called **simply connected**.[8]

Functors. From every pointed map $f\colon (X, x_0) \to (Y, y_0)$, we obtain a map $f_*\colon \pi_1(X, x_0) \to \pi_1(Y, y_0)$ of the fundamental groups: if you think of it, there is only one possible definition of f_*, namely, $f_*([a]_*) = [fa]_*$—just look at the loop's image under f. Routine check shows that this is well defined, and moreover, one finds easily that f_* is a group homomorphism. Finally, the construction respects composition of maps: $(fg)_* = f_* g_*$.

Because of these wonderful properties, π_1 is called a **functor**, more precisely, a functor from the category of pointed spaces to the category of groups. A functor is yet another key notion of category theory we wanted to mention, but we have been waiting for a good enough example.

A functor F in general is something like a morphism of categories (but it must not be called that way). If F goes from a category C to a category D, then it assigns an object $F(X) \in \mathrm{Ob}(\mathsf{D})$ to every $X \in \mathrm{Ob}(\mathsf{C})$ and a morphism $F(f) \in \mathrm{Hom}(F(X), F(Y))$ to every morphism $f \in \mathrm{Hom}(X, Y)$ in C, so that identity morphisms are mapped to identity morphisms and F respects composition, $F(fg) = F(f)F(g)$.

Another example of a functor we have met is the polyhedron of a finite simplicial complex, from the category of finite simplicial complexes with simplicial maps to topological space with continuous maps (or, strictly speaking, equivalence classes of homeomorphic topological spaces).

If we take a commutative diagram of pointed spaces with pointed maps and apply a functor, such as π_1, we automatically obtain another commutative diagram, in this case with groups and homomorphisms. Since many things can be expressed by commutativity of suitable diagrams, some steps in proofs become almost mechanical, once one gets used to this approach. When considering some construction taking objects of some kind and producing objects

[8] Often this also includes the assumption that X is path-connected, but not in all sources. On the other hand, the fundamental group can "see" only the path-connected component containing the basepoint.

of another kind, say simplicial complexes from graphs, it may be worth asking whether it is a functor, or if it can be adjusted to behave functorially.

The preservation of commutative diagrams by functors might perhaps suggest that functors should preserve limits, such as products, but this is not the case in general—the problem is with the uniqueness requirement in the definition of limit.

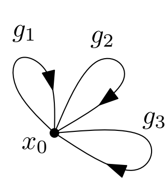
On the uncomputability of the fundamental group. Most of the elementary texts on algebraic topology cover basic properties of the fundamental group and tools for working with it, such as covering spaces, as well as rigorous computations of the fundamental group of S^1 and of compact 2-dimensional surfaces. Here we do not go in this direction, mainly because in many computer-science uses of topology, spaces with a nontrivial fundamental group are difficult or impossible to deal with, while interesting things can be done once we assume a simply connected space.

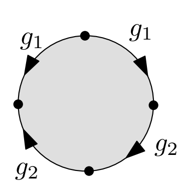
The basic difficulty with $\pi_1(X)$ is that almost nothing about it is algorithmically computable. These uncomputability results all go back to uncomputability results for groups.

A group G is said to be *finitely presented by generators and relations* if we are given a list $g_1, \ldots, g_n$ of elements of G that together generate G, plus a finite list of relations such as $g_3 g_5 g_4^7 g_2^{-3} = e$, e the unit element of G. For example, the free group on 2 generators may have the generator list a, b and no relations, while adding the relation $aba^{-1}b^{-1} = e$ yields the Abelian group $\mathbb{Z}^2$, etc.

Practically everything about groups presented in this way is uncomputable in general, such as *nontriviality* (does G have an element distinct from e?), or the *word problem* (do the defining relations of G imply a given relation?).

Given a group G finitely presented by generators and relations, one can algorithmically construct a 2-dimensional simplicial complex having G as the fundamental group. The idea is extremely simple: given the generators $g_1, \ldots, g_n$, we make a wedge of n circles, each corresponding to one generator. Then, for every relation, say $g_1 g_2^2 g_1^{-1} = e$, we take a new disk and glue its boundary to the loops, as is illustrated below, so that it does not intersect anything else from the space constructed so far:





This usually cannot be pictured in $\mathbb{R}^3$, but the resulting space is well defined. It should also be clear that it has a finite triangulation.

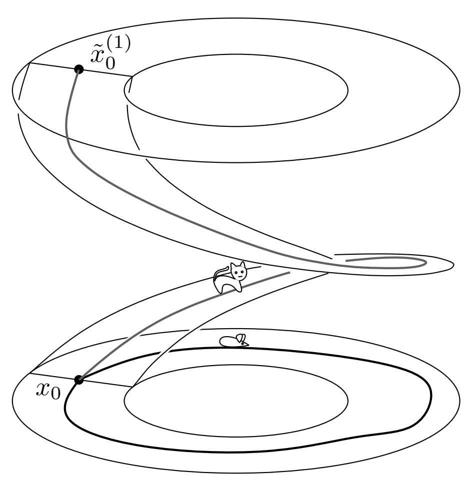
Basic results about the fundamental group (the Seifert–Van Kampen theorem) immediately imply that the fundamental group of the resulting space is isomorphic to G, since it has the same generators and relations. So the nontriviality of $\pi_1(K)$, already for a 2-dimensional finite simplicial complex K, is algorithmically undecidable!

There are numerous other, more difficult uncomputability results, such as the impossibility of recognizing an S^5 mentioned earlier. These usually require refinements of the basic results above both on the group side (proving uncomputability for groups of some restricted type) and on the topological side (implanting such groups as fundamental groups in restricted classes of spaces, such as 4-dimensional manifolds).

Fortunately, while testing simple connectedness is hopeless in general, in many concrete cases it may have simple reasons and be easy to check (or be known).

So how can we prove the nontriviality of a loop? We have said that we will not develop the theory of the fundamental group, but we cannot resist sketching one of the basic tricks at least informally. We consider the annulus A with basepoint x_0 and a pointed loop a going around once, and we would like to prove that a cannot be nullhomotopic.

We think of A as a corridor in the ground (0th) floor, and we consider a spiral staircase $\tilde{A}$ that winds above A and also below it, in infinitely many loops (the geometric shape of A is called a *helicoid*; we ignore the stairs and regard $\tilde{A}$ as smooth surface). The staircase $\tilde{A}$ reaches the ground floor A exactly at the basepoint x_0, and above x_0, we have exactly one point $\tilde{x}_0^{(k)} \in \tilde{A}$ in each floor, $k \in \mathbb{Z}$, with $\tilde{x}_0^{(0)} = x_0$.



There is a mouse following the loop a in A: at time $t \in [0, 1]$ it is in $a(t)$, starting and ending in x_0. A cat starts at x_0 at time 0 as well, but moves along $\tilde{A}$, not along A (we do not address a practical solution of how the mouse can freely cross $\tilde{A}$ and the cat A, but we do not let this issue distract us from the essence). The cat always stays precisely vertically above the mouse (or below it).

It seems plausible that the cat's path $\tilde{a}$ is determined uniquely (although proving this rigorously is one of the technical parts we wanted to avoid), and for the particular loop a, the cat ends up at $\tilde{x}_0^{(1)}$, one floor above the mouse. Moreover, by technically very similar considerations one can prove that if a and b are pointed-homotopic loops, then the corresponding cat's paths $\tilde{a}$ and $\tilde{b}$ are homotopic too, with homotopy fixing each of the $\tilde{x}_0^{(k)}$. In particular, the cat ends up in the same floor for both $\tilde{a}$ and $\tilde{b}$.

We can already see that our particular loop a cannot be nullhomotopic, since for the constant (unmoving) loop the cat keeps sitting in x_0.

The space $\tilde{A}$ is an example of a *covering space*, and the vertical projection $\tilde{A} \to A$ is called a *covering map*. The method of covering spaces allows one to compute, e.g., the fundamental groups of all compact 2-dimensional surfaces.

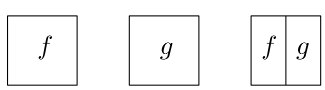
Higher homotopy groups. Higher homotopy groups $\pi_k(X)$, $k = 2, 3, \ldots$, were first introduced by Čech, long after the fundamental group $\pi_1(X)$. Čech actually withdrew his paper on the advice of senior colleagues, who believed that, unlike the groups in his definition, the true higher homotopy groups should not be commutative in general.

The definition is now the accepted one, though, and in spite of being "only" Abelian, higher homotopy groups belong among the most challenging objects in mathematics. As we will see, the reason why $\pi_1(X)$ need not be commutative but all higher homotopy groups always are is that the 1-dimensional sphere S^1 is like a rail track where moving points cannot pass one another without colliding, while in S^k, $k \geq 2$, there is enough room for points to move around without collisions.

To define π_k, we must again consider a pointed space (X, x_0). The *elements* of $\pi_k(X)$ are easy to define with the notation we already have: they are pointed homotopy classes of pointed maps $(S^k, s_0) \to (X, x_0)$ (where s_0 is a basepoint in S^k, say the north pole).

Understanding the group operation is a bit more challenging. A good way is to regard S^k as the quotient $I^k/\partial I^k$ of the k-dimensional cube I^k, $I = [0, 1]$, by its boundary ∂I^k (we saw this representation of the sphere, with B^k instead of I^k, in Exercise 6.2). Then a pointed map $f\colon (S^k, s_0) \to (X, x_0)$ can also be regarded as a map of pairs $(I^k, \partial I^k) \to (X, x_0)$, i.e., a map from the cube that sends all of the boundary to x_0.

Let us now consider two elements $[f]_*$ and $[g]_*$ of $\pi_k(X, x_0)$, and let us think of the representing maps f, g as maps of the cube as above. Then the map h representing $[f]_* + [g]_*$ (the operation in $\pi_k(X)$, $k \geq 2$, is usually written as addition) is constructed as follows: We split I^k into the left and right half along the x_1 coordinate, we squeeze the cube on which f is defined twice and put it over the left half, and similarly the cube with g is squeezed twice and identified with the right half. Then h equals the squeezed f on the left half and the squeezed g on the right half.



Note that this directly generalizes the way we have introduced the operation in the fundamental group.

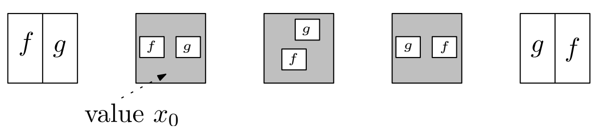
Definition 9.4. *For a pointed space (X, x_0), the* **kth homotopy group** *$\pi_k(X, x_0)$ is the set $[(S^k, s_0), (X, x_0)]_*$ of pointed homotopy equivalence classes of pointed maps of the k-sphere into (X, x_0), with the addition operation described above, "putting f on the left half-cube and g on the right half-cube."*

Exercise 9.5. *Describe the addition in π_2 directly using pointed maps, rather than maps of the cube.*

Exercise 9.6. *How does one get the inverse of $[f]_*$?*

One again has to verify, routinely, that the operation is well defined (the result does not depend on the choice of representatives), associative, and has inverses. It is also easy to show that for a path-connected X, all choices of the basepoint x_0 give isomorphic $\pi_k(X, x_0)$. Moreover, for every $k \geq 2$, π_k is a functor from pointed spaces to Abelian groups—this is not really different from the $k = 1$ case, with the exception of commutativity, which we now explain.

We want to show that the representative of $[f]_* + [g]_*$ constructed as above is homotopic to the representative of $[g]_* + [f]_*$. The homotopy, in the setting of maps from the cube, is illustrated below:



Homotopy groups of spheres. Contractible spaces, such as balls, have all homotopy groups zero. One would think that the next simplest example should be the spheres. So what is $\pi_k(S^n)$?

First, for $k < n$, it is 0. This looks quite plausible if you think of the image of S^1 in S^2, say; one should be able to pick a point of S^2 not in the image, take it as the north pole and let northern wind contract the image continuously to the south pole. There is a technical difficulty with this, since the image could be a space-filling curve covering all of S^2, but this can be dealt with, as we will see later (Corollary 11.3).

Good; now for $k = n$ we have $\pi_n(S^n) = \mathbb{Z}$ for all $n \geq 1$. This hides a nice theorem of Hopf, stating that two maps $S^n \to S^n$ are homotopic if and only if they have the same *degree*. A rigorous definition of degree can be given using homology. Here is an informal explanation: We think of S^2 in $\mathbb{R}^3$, and color it green from inside and red from outside. Then we map it to another S^2 by a map f. If f is locally sufficiently nice and we look at a generic point x of the target S^2, there are locally several sheets of the red-green S^2 over x, some of them are green when we look from inside the target S^2 and others red. The degree of f is the number of green ones minus the number of red ones. For $k = 1$ this is the winding number – number of times f "goes around" the target S^1.

How about the case $k > n$? Historically it came as a great surprise that $\pi_k(S^n)$ can be nonzero in this situation. The first instance was discovered by

Hopf, who found a nontrivial map $\eta : S^3 \to S^2$ (i.e., not nullhomotopic). It is now called the **Hopf map** and it belongs among a handful of key examples in topology (and also with uses in quantum physics and elsewhere). The Hopf map has a one-line definition using complex numbers, but a longer explanation and pictures are needed to see what is going on—not speaking of a proof of nontriviality, which needs considerable apparatus. So we refer to the literature.

A concise answer to the question of what the $\pi_k(S^n)$ are, after many decades of research, is—nobody knows. Many deep and interesting facts have been proved. For example, for $k > n$, the $\pi_k(S^n)$ are finite, with the sole exception of the cases $\pi_{4n-1}(S^{2n})$. One of the perhaps most remarkable phenomena is *stability*: for all $n > i+1$ the homotopy groups $\pi_{n+i}(S^n)$ depend only on i.

But these stable homotopy groups of spheres, for instance, are known only up to $i = 64$, and known part of the table looks fairly chaotic; just to give a taste, we have $\pi_{n+35}(S^n) = \mathbb{Z}_8 \oplus \mathbb{Z}_2^2 \oplus \mathbb{Z}_{27} \oplus \mathbb{Z}_7 \oplus \mathbb{Z}_{19}$ for all $n \geq 37$.

Let us stress that these mysterious homotopy groups are, in principle, *computable*. There are algorithms that, given k and a simply connected finite simplicial complex K, outputs a description of $\pi_k(K)$; some of the more recent ones might even be practically usable for k very small.

But so far such general algorithms have been useless for the problem of homotopy groups of spheres. Researchers have developed a number of very sophisticated methods tailor-made for the case of spheres; this problem has helped to keep algebraic topology progressing for several decades.

Perhaps surprisingly, most of the known $\pi_k(S^n)$ were actually computed by hand. This illustrates that, given a sufficiently structured problem and enough interest, mathematical theory can usually beat supercomputers.

10 Homology of simplicial complexes

Here we begin considering homology groups. They are more difficult to introduce than homotopy groups, and in some respects they carry less information, but from a computer-science perspective, they offer an immense advantage: They are efficiently computable, both in theory (polynomial-time algorithms) and in practice (reasonably large instances can be dealt with). With a bit of exaggeration one can say that in topology, whatever can be inferred from homology is efficiently computable, and almost nothing else is.

Coefficients for homology. For every integer $k \geq 0$, we want to assign an Abelian group $H_k(X)$ to a topological space X, the *kth homology group* of X. More precisely, the construction has a parameter R, which is a commutative ring (ring as in algebra); then we speak of *homology with coefficients in* R and write $H_k(X; R)$.

The most standard choice is $R = \mathbb{Z}$, the ring of integers—the notation $H_k(X)$ refers to this default case. Other common choices are $R = \mathbb{Q}$, the field of the rationals, and $R = \mathbb{Z}_p$, the finite field with p elements, p a prime. It is known that the homology groups with integer coefficients collectively encode all the information contained in homology with coefficients in any other ring R, so in a sense they are the best one can get, but other choices of coefficients may sometimes be easier to work with.

Here, in the interest of simplicity, we are going to work with coefficients in $\mathbb{Z}_2$, the two-element field. This brings two distinct kinds of simplification compared to integer coefficients.

First, we have $-1 = 1$ in $\mathbb{Z}_2$, and this relieves us from having to keep track of *orientations* of simplices and signs in formulas. Handling these issues is not conceptually difficult at all, although error-prone, but it is pleasant to ignore them at first encounter.

Second, $\mathbb{Z}_2$ is a field, and working with homology over a field is basically linear algebra—in particular, the homology groups $H_k(X;\mathbb{Z}_2)$ are really vector spaces. On the other hand, if R is not a field, then vector spaces have to be replaced with R*-modules*, a probably much less familiar concept (in particular, $\mathbb{Z}$-modules can be regarded as Abelian groups).

Simplicial homology. We will follow more or less the historical route of developing homology: first we define homology groups for a finite simplicial complex, and then we will say that the homology of a triangulable space X is the homology of a triangulation of X. For this to make sense, we must show that the result is the same for all possible triangulation, and this takes quite a bit of work.

Modern topology textbooks often prefer a different way, the so-called *singular homology*.[9] This is defined very quickly and compactly, if not too intuitively, directly for a space without a detour through triangulations, and some of the general properties are immediately obvious. However, it is difficult to compute the singular homology of almost anything, and if one wants a computational tool, some version of simplicial homology is needed anyway.

Chains, cycles, boundaries. Let K be a finite simplicial complex, and let $k \geq 0$ be an integer. Let us write K_k for the set of all simplices of dimension precisely k in K.

First we define a vector space $C_k(K;\mathbb{Z}_2)$. Its vectors are all formal linear combinations of simplices in K_k. In more detail, we fix a vector space over $\mathbb{Z}_2$ of dimension $|K_k|$, we fix a basis in it, and we identify the basis elements bijectively with the k-simplices of K. The vectors of $C_k(K;\mathbb{Z}_2)$ are called the ***k*-chains** of K, and typically they are denoted by c, or c_k if we want to stress which k we have in mind.

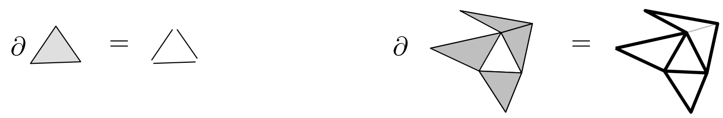
In the very special case of $\mathbb{Z}_2$, we can also think of k-chains as subsets of K_k: for a k-chain c, the subset consists of those simplices in K_k that have coefficient 1 in c.

Next, we define a linear map $\partial = \partial_k \colon C_k(K;\mathbb{Z}_2) \to C_{k-1}(K;\mathbb{Z}_2)$ called the **boundary operator**. It suffices to define the values of a linear map on basis vectors, and in our case we just specify the value of ∂_k on each k-simplex $F \in K_k$. Namely, $\partial_k F$ is the sum of all $(k-1)$-simplices that are facets of F; formally, and recalling that we regard simplices of a simplicial complex as finite

[9] There are actually many ways of defining homology groups, which may differ in generality of the considered spaces or in various fine points, so many that already in 1945 Eilenberg and Steenrod found it desirable to put all the various homology theories on an axiomatic footing, in order to isolate essential properties common to all of them.

sets, we can thus write $\partial_k F = \sum_{v \in F} F \setminus v$.

The left picture illustrates this definition for $k = 2$, while the right picture shows the boundary of a more complicated 2-chain.



The boundary operator in the $\mathbb{Z}_2$ case can also easily be described combinatorially: a $(k-1)$-dimensional simplex G belongs to $\partial_k c$ for a given k-chain c exactly if it is contained in an odd number of the simplices of c.

We also make the convention that $\partial_0 = 0$; i.e., vertices have no boundary. (Sometimes it is technically convenient to change this convention slightly, which leads to the so-called *reduced homology groups* $\tilde{H}_k$, but we do not consider these here—the difference is minor anyway, and only in the 0th homology group.)

Exercise 10.1. *(Trivial) To see that there is no mystery in the boundary operator, write down the matrix of* ∂_2 *and of* ∂_1 *for the simplicial complex from the right picture above.*

Starting from the boundary operators, we now obtain the homology groups of K by pure linear algebra. We define two vector subspaces of $C_k(K; \mathbb{Z}_2)$:

- $Z_k(K; \mathbb{Z}_2)$ consists of all k-chains whose boundary is zero (empty); concisely, $Z_k(K; \mathbb{Z}_2) = \ker \partial_k$. The vectors of $Z_k(K; \mathbb{Z}_2)$ are called ***k*-cycles** (Z is for the German word *Zyklus*).

- $B_k(K; \mathbb{Z}_2)$ consists of all boundaries of $(k+1)$-chains; i.e., $B_k(K; \mathbb{Z}_2) = \operatorname{im} \partial_{k+1}$. Its vectors are ***k*-boundaries**.

A key observation in homology is this.

Observation 10.2. *The composition* $\partial_k \partial_{k+1}$ *is zero (every boundary has zero boundary). Thus,* $B_k(K; Z_2) \subseteq Z_k(K; \mathbb{Z}_2)$ *(all boundaries are cycles).*

Proof. It suffices to prove $\partial_k \partial_{k+1} F = 0$ for every $(k+1)$-simplex F. This is immediate since every $(k-1)$-dimensional face G of F is contained in exactly two k-dimensional faces of F, namely, $F \setminus \{a\}$ and $F \setminus \{b\}$, where $F \setminus G = \{a, b\}$. □

Definition 10.3. *The kth homology group* $H_k(K; \mathbb{Z}_2)$ *of a simplicial complex* K *is the quotient vector space* $Z_k(K; \mathbb{Z}_2)/B_k(K; \mathbb{Z}_2)$*, "cycles modulo boundaries."*

If we are interested only in the isomorphism type of $H_k(K; \mathbb{Z}_2)$, then a single number suffices: the dimension, which equals $\dim Z_k(K; \mathbb{Z}_2) - \dim B_k(K; \mathbb{Z}_2)$. We trust that the reader can imagine how the computation of these dimensions, with K as input, could be programmed using basic subroutines for linear algebra (Gaussian elimination, say)—if not, we recommend to spend a couple of minutes on that.

For some applications, though, the isomorphism type may not suffice; we may also be interested in seeing the k-cycles representing elements of some basis of $H_k(K;\mathbb{Z}_2)$. These are also easily computed.

The elements of the quotient vector space $H_k(K;\mathbb{Z}_2)$ are equivalence classes of k-cycles, of the form $z + B_k(K;\mathbb{Z}_2)$. They are called *homology classes*. In general, two k-chains c and c' are called *homologous* if $c - c'$ is a boundary.

Here we come back to the geometric intuition sketched when we mentioned homology for the first time: the nonzero elements of the k-th homology group correspond to the k-dimensional cycles (e.g., in the 1-dimensional case, loops and linear combinations of loops) that cannot be expressed as boundaries of $(k+1)$-dimensional "things."

Enough definitions—it is time to practice a little.

Exercise 10.4. *Regard a simple graph as a* 1*-dimensional simplicial complex. Describe k-cycles and k-boundaries in graph-theoretic terms, $k = 0, 1$. What is the meaning of H_0 and H_1?*

The next, very basic exercise should show that, unlike homotopy groups, the homology groups of spheres are very simple and predictable.

Exercise 10.5. (a) *Consider the n-dimensional simplex as a simplicial complex (all subsets of $[n+1]$). Compute all of the homology groups with $\mathbb{Z}_2$ coefficients.*

(b) *Now remove the simplex itself from the simplicial complex, leaving only its boundary (this is a triangulation of S^{n-1}). What changes in the homology groups?*

Functors. We would like to see that, for every k, the assignment of the kth homology group to a finite simplicial complex behaves as a functor (we still work with $\mathbb{Z}_2$ coefficients but this applies to any coefficient ring, and later on we will of course want an analog for spaces instead of simplicial complexes). Given simplicial complexes K, L and a simplicial map $f\colon K \to L$ (these are the morphisms), we want a linear map $f_{*k}\colon H_k(K;\mathbb{Z}_2) \to H_k(L;\mathbb{Z}_2)$—what should it be?

Short reflection shows that there is not much choice. First we define a linear map $f_{\#k}\colon C_k(K;\mathbb{Z}_2) \to C_k(L;\mathbb{Z}_2)$ by specifying the values on the usual basis of k-simplices. Namely, the image $f(F)$ of a simplex $F \in K$ is a simplex of L, which can be of dimension $\dim F$ or smaller. We set

$$f_{\#k}(F) := \begin{cases} f(F) & \text{if } \dim f(F) = \dim F, \\ 0 & \text{otherwise.} \end{cases}$$

Now the map f_{*k} of the homology group is defined by taking a representative k-cycle z of a homology class $h = z + B_k(K;\mathbb{Z}_2)$ and defining $f_{*k}(h)$ as the homology class $f_{\#k}(z) + B_k(L;\mathbb{Z}_2)$. Two things could possibly go wrong with this definition; the reader is invited to check that they do not.

Exercise 10.6. (a) *Check that $f_{\#(k-1)}\partial_k = \partial_k f_{\#k}$, and consequently, that the image of a cycle under $f_{\#k}$ is a cycle and the image of a boundary is a boundary.*

(b) *Verify that $(gf)_{*k} = g_{*k} f_{*k}$ for simplicial maps $f\colon K \to L$ and $g\colon L \to M$.*

Chain complexes. It is useful to isolate an intermediate algebraic object between a simplicial complex (or space) and its homology groups. A **chain complex** $\mathcal{C}$ is an infinite sequence of vector spaces (in our case of field coefficients) connected by linear maps, which are also called boundary operators:

$$\cdots \xrightarrow{\partial_3} C_2 \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\partial_0} C_{-1} \xrightarrow{\partial_{-1}} \cdots$$

The sole axiom is $\partial_{k-1}\partial_k = 0$ for all $k \in \mathbb{Z}$, exactly the property of the boundary operators that allowed us to define homology groups. So we can define homology groups for an arbitrary chain complex $\mathcal{C}$ in the same way.

Of course, a primary example of a chain complex is one obtained from a simplicial complex, with $C_k = C_k(K;\mathbb{Z}_2)$ (where $C_k = 0$ for $k < 0$ by convention). But chain complexes proved useful in many other contexts, and they are the object of study of *homological algebra*.

A morphism $f_\#\colon \mathcal{C} \to \mathcal{D}$ of chain complexes, called a **chain map**, is modeled after the maps $f_{\#k}$ constructed above; we just postulate the property we needed so that $f_\#$ induce a map of the homology groups. Namely, $f_\# = (f_{\#k})_{k\in\mathbb{Z}}$ is a sequence of linear maps $f_{\#k}\colon C_k \to D_k$ satisfying $\partial_k f_{\#k} = f_{\#(k-1)}\partial_k$, where ∂_k on the left is the boundary operator in $\mathcal{C}$, while ∂_k on the right comes from $\mathcal{D}$.

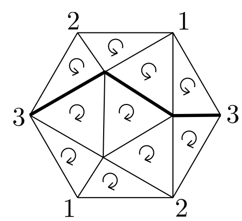
What is the point of this exercise with defining chain complexes? Besides general applicability already mentioned, there is a specific advantage for our later considerations: even between chain complexes derived from simplicial complexes, we will have useful and fairly obvious chain maps that do not come from any simplicial map, though.

What changes for $\mathbb{Z}$ instead of $\mathbb{Z}_2$? For integer coefficients, first of all, we must assign every simplex of K some *orientation*. A simple way is to number the vertices of K from 1 to n and then orient every simplex "from left to right". Now the definition of the boundary operator involves signs, so the boundary of a triangle is no longer the sum of its edges, but something like "the first edge minus the second edge plus the third one."

The oriented simplices appear with arbitrary integer coefficients in k-chains. The k-chains, k-cycles, and k-boundaries are no longer vector spaces, but $\mathbb{Z}$-modules or Abelian groups. For computing with them we no longer suffice with linear algebra, but we need to manipulate integer matrices. Most notably, instead of Gaussian elimination, we use algorithms for computing the *Smith normal form*; these are considerably more sophisticated but reasonably well understood and fast.

The resulting homology group $H_k(K;\mathbb{Z})$ can be an arbitrary Abelian group (finitely generated for a finite simplicial complex), such as $\mathbb{Z}^{17} \oplus \mathbb{Z}_2^6 \oplus \mathbb{Z}_3^2 \oplus \mathbb{Z}_6 \oplus \mathbb{Z}_{293}$. Here 17, the exponent of the infinite cyclic component, is called the *kth Betti number* of K (warning: it need not equal $\dim H_k(K;\mathbb{Z}_2)$, the kth $\mathbb{Z}_2$ Betti number).

The finite cyclic summands are called the **torsion part** of $H_k(K;\mathbb{Z})$. They have no analog for field coefficients, and their geometric meaning is not so easy to visualize. The simplest case can perhaps be shown for the projective plane, which can be obtained by taking a disk and identifying every pair of diametrically opposite points.



The picture shows a triangulation of the projective plane (the smallest one, actually; identifications of vertices are indicated by the numbers 1,2,3) with some chosen orientations of the triangles. The curve drawn thick is a closed loop, a homological 1-cycle (zero boundary). The curve itself is not a boundary (with $\mathbb{Z}$ coefficients), but if we take it *twice*, it becomes a boundary of the sum of all 2-simplices of the projective plane with the given orientations; note that the orientations of adjacent triangles flip locally exactly along our curve.

Cohomology. First we recall from linear algebra that if V is a vector space over a field $\mathbb{K}$, then the set of all linear functions $f\colon V \to \mathbb{K}$ also forms a vector space V^*, called the *dual* of V. If V is finite-dimensional, then $\dim V^* = \dim V$; given a basis $b_1, \dots, b_n$ of V, we can form a basis $b_1^*, \dots, b_n^*$ of V^*, where b_i^* attains value 1 on b_i and value 0 on all the other b_j.

Every linear map $L\colon V \to W$ determines a linear map $L^*\colon W^* \to V^*$ (note the change of direction; this will be happening all the time here), called the *adjoint* of L. There is only one reasonable way of defining L^*, namely, $L(f) := fL\colon V \to \mathbb{K}$, $f \in W^*$.

To introduce cohomology with $\mathbb{Z}_2$ coefficients (or with any field coefficients, there is no difference), we apply this kind of duality to the vector spaces $C_k(K;\mathbb{Z}_2)$ of k-chains and the boundary operators ∂_k. So $C^k(K;\mathbb{Z}_2)$, the k-*cochains*, is the vector space dual to $C_k(K;\mathbb{Z}_2)$.

How should one think of k-cochains? It is enough to specify the values of a k-cochain on a basis, in this case the standard basis of the k-chains consisting of k-simplices—so a k-cochain "looks" precisely like a k-chain.[10]

Next, the **coboundary operator** $\delta_k\colon C^k(K;\mathbb{Z}_2) \to C^{k+1}(K;\mathbb{Z}_2)$ now *increases* dimension by 1, and it is the adjoint of ∂_k. The k-cocycles $Z^k(K;\mathbb{Z}_2) := \ker \delta_k$ and the k-coboundaries $B^k(K;\mathbb{Z}_2) := \operatorname{im} \delta_{k-1}$ are defined analogously to the case of homology, and the $\boldsymbol{k}$**th cohomology group** is then

$$H^k(K;\mathbb{Z}_2) := Z^k(K;\mathbb{Z}_2)/B^k(K;\mathbb{Z}_2).$$

Exercise 10.7. *Similar to homology, interpret the coboundary operator, cocycles, and coboundaries in graph-theoretic terms (for* 1*-dimensional simplicial complexes).*

It turns out that cohomology groups are similar to homology groups in many respects, and in particular, all cohomology groups of a space X can be deduced from the knowledge of all homology groups of X. So why bother with cohomology?

[10]In the finite-dimensional case, that is. For infinite simplicial complexes, say, a k-chain may attain only finitely many nonzero values, while a k-cochain may have any number of nonzeros.

Perhaps the key reason is that if we look at a simplicial maps $f\colon K \to L$ (or, later, a continuous map), the induced map in cohomology goes in the *opposite direction*, $f^{*k}\colon H^k(L;\mathbb{Z}_2) \to H^k(K;\mathbb{Z}_2)$. This is because of the dual nature of cohomology, and the reader may want to contemplate the definition of f^{*k} for a while.

Category theorist say that cohomology groups are **contravariant functors**; all of the functors mentioned so far were **covariant**. This is often useful, since cohomology can be pulled in the direction opposite to homology (and homotopy groups).

One of the most important manifestations of this general phenomenon is the existence of a product structure on the union of all cohomology groups of a given space. Namely, one can define the **cup product** in cohomology (we will not do this here, because this would take us too far), which assigns to two cohomology classes $c^p \in H^p(X)$ and $c^q \in H^q(X)$ a cohomology class $c^p \smile c^q \in H^{p+q}(X)$. This operation makes the union $\bigcup_{k=0}^{\infty} H^k(X)$ into a ring, the *cohomology ring* of X. For homology this does not work in general, since one of the key maps needed for defining a product goes in the wrong direction.

The cohomology ring in general carries strictly more information about X than just all homology groups or all cohomology groups. The cup product is also used for formalizing intuitive geometric notions such as *linking number* (given two images of S^1 in $\mathbb{R}^3$, how many times are they "linked"?, or similarly for an image of S^p and an image of S^q in $\mathbb{R}^{p+q+1}$).

11 Simplicial approximation

We have assigned a sequence of homology groups to a finite simplicial complex. To make the construction topologically useful, we still need to show that the homology groups depend only on the considered topological space (and actually, only on its homotopy type) and not on the chosen triangulation.

We will present a substantial part of a proof. Along the way, we will encounter a surprising phenomenon concerning triangulations. As we have mentioned earlier, the usual contemporary proof would go via singular homology. However, the considerations needed to connect singular homology to simplicial homology are also nontrivial.

In the present section we will not talk about homology, but we prepare a tool for approximating continuous maps by simplicial maps—the simplicial approximation theorem. Historically, this kind of statement made possible the first rigorous (and correct or almost correct) proofs of foundational results such as that $\mathbb{R}^m$ is not homeomorphic to $\mathbb{R}^n$ for $m \neq n$.

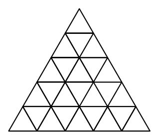
Subdivisions, especially barycentric. A geometric simplicial complex Δ' is called a **subdivision** of a geometric simplicial complex Δ if their polyhedra are the same, $|\Delta'| = |\Delta|$, and every simplex of Δ' is fully contained in some simplex of Δ.

The idea of subdivision is that each simplex of Δ is further sliced into smaller simplices, but this must be done so that the slicings are compatible across lower-dimensional faces.

The main thing about subdivisions one usually needs is that every finite geometric simplicial complex has arbitrarily fine subdivisions:

Lemma 11.1. *Let Δ be a (finite) geometric simplicial complex, considered as a subspace of $\mathbb{R}^n$ with the Euclidean metric, and let $\varepsilon > 0$ be given. Then there is a subdivision Δ' of Δ in which no simplex has diameter larger than ε.*

A two-dimensional picture

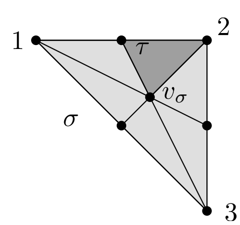


might mislead one into considering the lemma trivial. But the "parallel layer slicing" idea does not work in dimensions 3 and higher. The usual proof of the lemma, shown below, uses a subdivision which is not very economical for the purpose of reducing the diameter of the simplices, but very simple to describe.

Sketch of proof. First we define the **barycentric subdivision** of a simplicial complex Δ. We formulate the definition recursively.

If $\dim \Delta = k$, we first produce the barycentric subdivision, call it Γ, of the subcomplex of Δ consisting of all simplices of dimension at most $k-1$ (by induction on k; for $k = 0$ we do nothing). Then, for every k-dimensional simplex σ of Δ, we consider the barycenter (center of gravity) v_σ of σ, and for every simplex $\tau \in \Gamma$ lying on the boundary of σ, we construct the cone with base τ and apex v_σ (which is a simplex). All of these simplices together with those of Γ form the barycentric subdivision of Δ.

The picture illustrates the procedure for a single 2-dimensional simplex:



Let us mention in passing that there is also a direct, combinatorial description: simplices of the barycentric subdivision of Δ are in one-to-one correspondence with chains of nonempty simplices of Δ under inclusion. For example, the darker 2-simplex in the picture above corresponds to the chain $\{2\} \subset \{1,2\} \subset \{1,2,3\}$. This correspondence has several nice uses, but we will consider barycentric subdivisions only for the purposes of the present proof.

For us it is important that if we apply the barycentric subdivision to a simplex of diameter at most 1, then each simplex in the subdivision has diameter at most $1 - \frac{1}{n+1}$ (this is not so hard to verify, but we will accept it as a fact without proof). Hence if we start with a given complex Δ and iterate the barycentric subdivision sufficiently many times, we are guaranteed to get the diameter of all simplices as small as needed. $\square$

Now we state the main theorem of this section. Exceptionally, it contains an undefined term “simplicial approximation”; we prefer to develop the definition later. Even ignoring this term the theorem still contains a highly interesting statement.

Theorem 11.2 (Simplicial approximation theorem)**.** *Let $f\colon |K| \to |L|$ be an arbitrary continuous map of polyhedra of two finite simplicial complexes. Then there is a subdivision K' of K and a simplicial approximation $s\colon V(K') \to V(L)$ of f. In particular, s is a simplicial map such that the affine extension $|s|\colon |K'| \to |L|$ is homotopic to f.*

Note that only the “source” complex K is subdivided, while L stays as is. The mapping $|s|$ need not be close to f in metric sense; we are only guaranteed that they are homotopic.

Here is a nontrivial consequence:

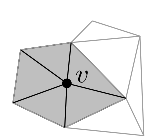
Corollary 11.3. *Every continuous map $f\colon S^{n-1} \to S^n$ is nullhomotopic.*

Proof. If we have a point y not in the image of f, the argument is obvious: we continuously push the image of f away from y until all of it ends up in the opposite point $-y$.

But the catch is that f may be surjective, like a space-filling curve. Then the simplicial approximation theorem comes to rescue: We consider S^{n-1} and S^n as polyhedra of K and L, respectively, and obtain a simplicial map $|s|$ from a subdivision of K into L that is homotopic to f. Such an $|s|$ cannot be surjective, because its image in L consists of simplices of dimension at most $n-1$. $\square$

Elementary homotopies and stars. We still need to define the meaning of “simplicial approximation.” First we formulate a simple condition for two continuous maps $f, g\colon |K| \to |L|$ to be homotopic. Namely, if we assume “for every $x \in K$, $f(x)$ and $g(x)$ share a simplex of L,” then surely $f \sim g$, since then the desired homotopy can be obtained by moving $f(x)$ towards $g(x)$ along a segment within the simplex containing both of these points. Let us say that f and g satisfy the *elementary homotopy condition*.

Next, we state a rather different-looking condition. Let K be a simplicial complex and let v be a vertex of it. The **open star** $\mathrm{st}^\circ v$ in K is defined as $|K|$ minus the union of all simplices of K that do not contain v.[11].



Let $f\colon |K| \to |L|$ be a continuous map of polyhedra and let $s\colon V(K) \to V(L)$ be a map of the vertex sets of the underlying simplicial complexes; at this

[11]The standard definition in the literature looks different, but from our version it is clear that the open star is an open set, which is far from clear for the standard formulation.

moment we (exceptionally) do not apriori assume s simplicial. We say that f and s satisfy the *star condition* if for every vertex $v \in V(K)$ we have

$$f(\mathrm{st}^\circ v) \subseteq \mathrm{st}^\circ(s(v)).$$

Interestingly, the elementary homotopy condition and the star condition turn out to be equivalent, and if one of them holds (and hence both) for a continuous map $f\colon |K| \to |L|$ and a simplicial map $s\colon V(K') \to V(L)$, we say that s is a **simplicial approximation** of f.

Exercise 11.4. (a) *Assuming that f and s satisfy the star condition, prove that s is necessarily simplicial.*

(b) *Show that if f and s satisfy the star condition, then f and $|s|$ satisfy the elementary homotopy condition.*

(c) *(Optional) Prove that, conversely, if f and $|s|$ satisfy the elementary homotopy condition, where s is simplicial, then f and s satisfy the star condition.*

A simple but very useful observation is that simplicial approximations respect composition: if $f\colon |K| \to |L|$ and $g\colon |L| \to |M|$ are continuous, $s\colon V(K) \to V(L)$ is a simplicial approximation of f, and $t\colon V(L) \to V(M)$ is a simplicial approximation of g, then ts is a simplicial approximation of gf.

Exercise 11.5 (Lebesgue covering lemma)**.** *Let X be a compact metric space, and let $\mathcal{U}$ be an open cover of X. Prove that there exists $\delta > 0$ (the* Lebesgue number *of the covering $\mathcal{U}$) such that for every $x \in X$, there is $U \in \mathcal{U}$ that contains x together with its open δ-neighborhood.*

Proof of Theorem 11.2. Given a map $f\colon |K| \to |L|$ as in the theorem, we consider the open cover $\mathcal{U}$ of $|K|$ with preimages of open stars in L: $\mathcal{U} := \{U_w = f^{-1}(\mathrm{st}^\circ w) : w \in V(L)\}$. We let $\delta > 0$ be a Lebesgue number of this cover, and let K' be a refinement of K in which simplices have diameter below δ.

We want to construct s so that f and s satisfy the star condition. For a vertex $v \in V(K)$, the open star $\mathrm{st}^\circ v$ is contained in the δ-neighborhood of v, and hence in some $U = U_w \in \mathcal{U}$. So $f(\mathrm{st}^\circ v) \subseteq \mathrm{st}^\circ w$. It suffices to put $s(v) := w$ (choosing arbitrarily if there are several possibilities); then f and s satisfy the star condition. $\square$

12 Homology does not depend on triangulation

In the subsequent considerations we stick to the $\mathbb{Z}_2$ coefficients in the notation, but the arguments work for any coefficients.

The case of a subdivision. First we consider the setting where K is a finite simplicial complex and K' is a subdivision of it. As expected, there is an isomorphism $H_k(K;\mathbb{Z}_2) \cong H_k(K';\mathbb{Z}_2)$ (for every k), but this is not completely easy to prove.

Let us describe the mappings inducing this isomorphism and its inverse. In the direction $K \to K'$, there seems to be no reasonable simplicial map to use, but there is a very natural chain map. Namely, for every k, we let

$\lambda_{\#k}\colon C_k(K;\mathbb{Z}_2) \to C_k(K';\mathbb{Z}_2)$ be the linear map that assigns to every k-simplex σ of K the sum of all simplices $\tau \in K'$ that are contained in σ. It is easy to check that this is indeed a chain map, and so it induces homomorphisms λ_{*k} in homology.

For the other direction, we use a simplicial map $\gamma\colon V(K') \to V(K)$. Namely, as in the proof of the simplicial approximation theorem (Theorem 11.2), we obtain γ as a simplicial approximation to the identity $\mathrm{id}\colon |K'| \to |K|$. For this, it suffices to observe that whenever $v \in V(K')$ is a vertex, we can find a vertex $w \in V(K)$ with $\mathrm{st}^\circ v \subseteq \mathrm{st}^\circ w$; then we set $\gamma(v) := w$ and the star condition holds.

We leave the following result without proof (referring, e.g., to Munkres [Mun84]).

Fact 12.1. *The homomorphisms $\lambda_{*k}\colon H_k(K;\mathbb{Z}_2) \to H_k(K';\mathbb{Z}_2)$ and $\gamma_{*k}\colon H_k(K';\mathbb{Z}_2) \to H_k(K;\mathbb{Z}_2)$ are mutually inverse, and thus isomorphisms.*

The proof uses two additional tools, *chain homotopy* and the method of *acyclic carriers*, which we prefer not to discuss here for space reasons. Interestingly, the proof is easier if K' is the barycentric subdivision of K. This case is actually sufficient for the proof that homology does not depend on triangulation, since it is enough to have arbitrarily fine subdivisions, but their structure does not matter.

Later we will need the following (easier) statement, which is proved using chain homotopy as well, and whose proof we also omit.

Fact 12.2. *Let $f\colon |K| \to |L|$ be a continuous map of polyhedra and let $s, t\colon V(K) \to V(L)$ be two simplicial approximations of f. Then they induce the same homomorphism in homology: $s_* = t_*$.*

(In accordance with our earlier notation, we should write s_{*k} and t_{*k}, but since all of the considerations are valid for every k, from now on we will mostly use the simpler notation.)

The Hauptvermutung. Once topologists proved that a simplicial complex and a subdivision of it have isomorphic homology groups, they hoped to obtain the independence of homology of the triangulation by establishing the following conjecture, which became known as the **Hauptvermutung** (main conjecture, from 1908): *Every two triangulations of a triangulable topological space have a common refinement.*

However, the conjecture was much more resistant to attacks than expected, and after some time researchers found a way of bypassing it (which will be presented below). Much later it turned out that the difficulty in proving the conjecture has a good reason: as shown by Milnor in 1961, the Hauptvermutung is actually *false*.

By now there are many examples and a great related theory concerned mainly with triangulations of manifolds (the Kirby–Siebenmann classification of manifolds and related developments), but the failure of the Hauptvermutung remains one of the truly mind-boggling phenomena in geometric topology. (See, e.g., Rudyak [Rud01] for a somewhat advanced exposition; unfortunately, we are not aware of any treatment easily accessible to beginners.)

Functorial properties. In order to bypass the Hauptvermutung, we will consider the homomorphisms induced by continuous maps in homology.

For a continuous map $f\colon |K| \to |L|$, we can define the induced homomorphism f_* in homology using the simplicial approximation theorem. Namely, we choose a subdivision K' of K such that there is a simplicial approximation $s\colon V(K') \to V(L)$ of f, and we set

$$f_* := s_*\lambda_* = s_*\gamma_*^{-1},$$

where λ_* and γ_* are the mutually inverse isomorphisms of $H_k(K;\mathbb{Z}_2)$ and $H_k(K';\mathbb{Z}_2)$ as in Fact 12.1.

The next complicated-looking statement tells us that this definition of f_* does not depend on K'.

Lemma 12.3. *Let K' and K'' be subdivisions of K, let $\gamma'\colon V(K') \to V(K)$ and $\gamma''\colon V(K'') \to V(K)$ be simplicial approximations of the identity, and let s', s'' be simplicial approximations of $f\colon |K| \to |L|$, where s' is defined on K' and s'' on K''. Finally, let $f'_* = s'_*(\gamma'_*)^{-1}$ and $f''_* = s''_*(\gamma''_*)^{-1}$ be the homomorphisms in homology as above. Then $f'_* = f''_*$.*

Proof. First we assume that there is a simplicial approximation $\gamma\colon V(K'') \to V(K')$ of the identity map $\mathrm{id}\colon |K''| \to |K'|$ (this does not mean that K'' is a refinement of K'!); later we will see how to arrange this. Here is a diagram summarizing the situation:

$$\begin{tikzcd}
K'' \arrow[r, "\gamma"'] \arrow[rr, bend left, "\gamma''"] \arrow[dr, "s''"'] & K' \arrow[r, "\gamma'"'] \arrow[d, "s'"] & K \arrow[dl, dotted, "f"] \\
& L &
\end{tikzcd}$$

We use Fact 12.2 twice: first, $s'\gamma$ and s'' are both simplicial approximations of f defined on K'', and so $s'_*\gamma_* = s''_*$, and second, $\gamma'\gamma$ and γ'' are both simplicial approximations of the identity defined on K'', and thus $\gamma'_*\gamma_* = \gamma''_*$.

Then we calculate quite mechanically

$$\begin{aligned}
f''_* &= s''_*(\gamma''_*)^{-1} = s'_*\gamma_*(\gamma''_*)^{-1} = s'_*\gamma_*(\gamma'_*\gamma_*)^{-1} \\
&= s'_*\gamma_*(\gamma_*)^{-1}(\gamma'_*)^{-1} = s'_*(\gamma'_*)^{-1} = f'_*.
\end{aligned}$$

Now in the general case, with K' and K'' arbitrary, we fix a subdivision K''' such that the identity has simplicial approximations $\gamma_1\colon V(K''') \to V(K')$ and $\gamma_2\colon V(K''') \to V(K'')$. Then we use the result of the first part twice, once on $K''' \to K' \to K$ and once on $K''' \to K'' \to K$. $\square$

Exercise 12.4. *Do the second step (the general case) in the above proof carefully in detail. How exactly do we get K'''?*

Theorem 12.5 (Functoriality of homology)**.**

(i) *Let* $\mathrm{id}\colon |K| \to |K|$ *be the identity map. Then the induced homomorphism in homology* id_* *is the identity.*

(ii) *Let* $f\colon |K| \to |L|$ *and* $g\colon |L| \to |M|$ *be continuous maps of polyhedra. Then* $(gf)_* = g_* f_*$.

Part (i) is immediate if we use the identity simplicial map $V(K) \to V(K)$ in the definition of id_*. Part (ii) is proved in a way very similar to the proof of the previous lemma, again using Fact 12.2 twice.

Exercise 12.6. *Prove the theorem. Draw a suitable diagram of the maps involved.*

The theorem immediately implies that two triangulations of the same space give the same homology groups, or in other words, that a homeomorphism of triangulable spaces induces isomorphism in homology. Indeed, if $h\colon |K| \to |L|$ is a homeomorphism, then we have $h_*(h^{-1})_* = (hh^{-1})_* = \mathrm{id}$, and so h_* has an inverse.

With some more work, one can show that homotopic maps induce the same homomorphism in homology: $f \sim g$ implies $f_* = g_*$ (the main idea is to use a simplicial approximation of the homotopy between f and g). It follows that not only homeomorphic, but also homotopy equivalent spaces have the same homology.

13 A quick harvest and two more results about homology

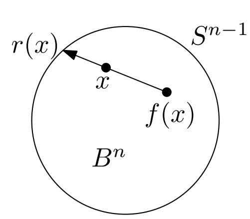
The sphere S^n has nonzero homology groups exactly in dimensions 0 and n. Hence S^n is not homotopy equivalent to S^m for $m \neq n$.

Corollary 13.1. $\mathbb{R}^m \not\cong \mathbb{R}^n$ *for* $m \neq n$.

Proof. A homeomorphism $\mathbb{R}^m \cong \mathbb{R}^n$ would yield a homeomorphism $\mathbb{R}^m \setminus \{0\} \cong \mathbb{R}^n \setminus \{0\}$. But $\mathbb{R}^m \setminus \{0\}$ is homotopy equivalent to S^{m-1} and $\mathbb{R}^n \setminus \{0\}$ to S^{n-1}. □

Theorem 13.2 (Brouwer's fixed-point theorem)**.** *Every continuous mapping* $f\colon B^n \to B^n$ *has a fixed point; i.e., there exists* $x \in B^n$ *with* $f(x) = x$.

Proof. The first step appears in almost all proofs of Brouwer's theorem. Assuming for contradiction that there is an f with $f(x) \neq x$ for all x, we construct a mapping $r\colon B^n \to S^{n-1}$ as in the picture:



That is, we send a ray from $f(x)$ to x, and $r(x)$ is the point where it hits the boundary of the ball.

The (hypothetical) mapping r is a *retraction* of the ball onto the sphere, meaning that $r(B^n) = S^{n-1}$ and on S^{n-1}, r acts as the identity. Using homology, we will show that r cannot exist.

Let $i\colon S^{n-1} \to B^n$ stand for the inclusion map. The composition ri is the identity $\mathrm{id}_{S^{n-1}}$, so we have the following diagram:

$$\begin{tikzcd} S^{n-1} \arrow[r, "i"'] \arrow[rr, bend left, "\mathrm{id}"] & B^n \arrow[r, "r"'] & S^{n-1} \end{tikzcd}$$

Now we apply the $(n-1)$st homology group functor $H_{n-1}(.;\mathbb{Z}_2)$ and obtain

$$\begin{tikzcd} H_{n-1}(S^{n-1};\mathbb{Z}_2) = \mathbb{Z}_2 \arrow[r, "i*"'] \arrow[rr, bend left, "\mathrm{id}"] & H_{n-1}(B^n;\mathbb{Z}_2) = 0 \arrow[r, "r_*"'] & H_{n-1}(S^{n-1};\mathbb{Z}_2) = \mathbb{Z}_2. \end{tikzcd}$$

But this is impossible, since an isomorphism $\mathbb{Z}_2 \cong \mathbb{Z}_2$ cannot factor through the trivial group. $\square$

Two nice theorems. We conclude our discussion of homology by presenting two more advanced results, which may be useful to know about.

The first one is a theorem of Hurewicz relating homology and homotopy groups. We state a somewhat special case of the theorem, which may be easier to grasp and remember than the usual general statement.

Theorem 13.3 (Hurewicz)**.** *If X is a simply connected space (i.e., $\pi_1(X) = 0$), then the first nonzero homotopy group and the first nonzero integral homology group occur in the same dimension and they are isomorphic. That is, for some $k \geq 2$ (possibly $k = \infty$) we have $\pi_j(X) = H_j(X;\mathbb{Z}) = 0$ for all $j < k$, and $\pi_k(X) \cong H_k(X;\mathbb{Z}) \neq 0$.*

Moreover, for every space X, $H_1(X;\mathbb{Z})$ is the Abelianization of $\pi_1(X)$, i.e., the quotient of $\pi_1(X)$ by its commutator subgroup.

The second result we want to highlight is concerned with a situation where we have the sphere S^n and a nonempty proper closed subspace $X \subset S^n$, and we would like to say something about the topological properties of the complement $S^n \setminus X$.

The unit circle S^1 can be embedded in S^3 in the "standard" way, say as a great circle, but also in various knotted ways, for example as the trefoil knot. It can be shown that the complements of the unknotted circle and of the trefoil are not homeomorphic (for instance, they have nonisomorphic fundamental groups). Therefore, the topology of X does not determine the topology of the complement.

Surprisingly, though, all *homology groups* of $S^n \setminus X$ are determined by X, and actually by the cohomology groups of X—this is the claim of a theorem known as the Alexander duality.

We should stress that our definition of homology using finite simplicial complexes is not adequate for this setting, since even if X is triangulable by a finite simplicial complex, its complement is an open set, which does not admit a finite triangulation.

Thus, one has to use other, more general definitions of homology. With the most usual one, *singular (co)homology*, we must make an extra assumption on X, namely, that X be *locally contractible* (every point has a contractible neighborhood). To get a general statement, one uses yet another definition, the *Čech cohomology*.[12]

The last remark before the statement of the Alexander duality is that we need to use *reduced* homology and cohomology, which is marked by a tilde above H. It influences only the 0th (co)homology group: while $H_0(X;\mathbb{Z})$ is $\mathbb{Z}^c$, where c is the number of path-connected components of X, $\tilde{H}_0(X;\mathbb{Z})$ is $\mathbb{Z}^{c-1}$.

Theorem 13.4 (Alexander duality). *Let $X \subset S^n$ be a nonempty proper closed subset of the sphere. Then we have, for $i = 0, 1, \ldots, n-1$,*

$$\tilde{H}^{n-i-1}(X;\mathbb{Z}) \cong \tilde{H}_i(S^n \setminus X;\mathbb{Z}),$$

where we need to use Čech cohomology on the left-hand side in general, and singular homology on the right-hand side.

For example, the Alexander duality implies that the complement of S^1 embedded in S^2 has two connected components, which is the contents of the famous *Jordan curve theorem.*

14 Manifolds

Manifolds constitute the most studied and most often applied class of topological spaces. Motivation and examples, besides pure mathematics, come mainly from physics, where certain kinds of manifolds play essential role.

> A seccond-countable Hausdorff topological space M is a **manifold** if for every point $x \in M$ there exists an open neighborhood U_x of x in M that is homeomorphic to some $\mathbb{R}^n$.

Simple examples of manifolds are R^n, S^n, the torus, or the projective plane. A more challenging type of example is $\mathrm{SO}(\mathbb{R}, n)$, the group of all rotations in $\mathbb{R}^n$ around the origin, which can be represented by the set of all orthogonal $n \times n$ real matrices with unit determinant. Here the condition in the definition of a manifold is non-obvious.

The number n in the definition above has to be constant for all points in every path-connected component of M, and it is called the **dimension** of that component. Some authors insist that all components have the same dimension as well (in this case we speak of an *n-manifold* and sometimes write M^n), while others admit combining components of different dimensions, but this is just a matter of convention.

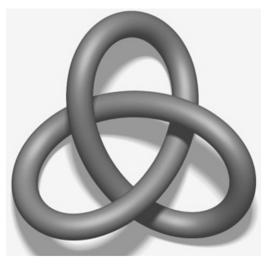
[12] All of these definitions agree with the simplicial one in the case of triangulable spaces, and they differ for quite pathological sets.

The assumption "second-countable and Hausdorff" is made to exclude pathologies such as the long ray of Example (F) in Section 2. Actually, most texts focus on *compact* manifolds, where a number of annoying technical difficulties disappear, and we will follow the suit. As examples of non-compact manifolds, one may think of an infinite cylinder or an infinite string of tori.

A **manifold with boundary** is defined in almost the same way as a manifold, only the neighborhood U_x is allowed to be homeomorphic *either* to $\mathbb{R}^n$ *or* to a closed halfspace in $\mathbb{R}^n$.

The points with a neighborhood homeomorphic to $\mathbb{R}^n$ are called *interior points*, and the remaining points form the *boundary*.

An obvious example of a manifold with boundary is the ball B^n. For a more interesting one, let us consider a knot in S^3. Mathematically, a knot is an S^1 embedded in S^3, but fancier drawings of knots show a thickened S^1: we imagine that we drill a thin non-self-intersecting tunnel in S^3 along the embedded S^1.



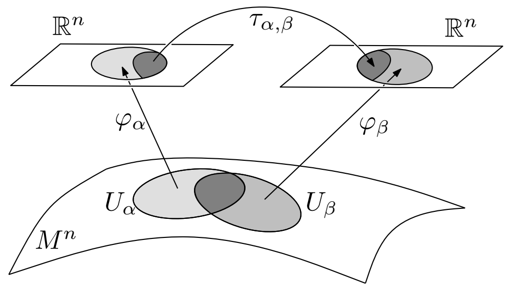
If we consider the tunnel T as an open set, so that $S^3 \setminus T$ is closed, then $S^3 \setminus T$ is a 3-manifold with boundary called a *knot manifold*.

Atlases and additional structures on manifolds. The manifolds we have defined above are **topological manifolds**, i.e., topological spaces satisfying an additional condition. But very often one wants a manifold to carry more structure, in order to calculate derivatives of functions in given directions, tangent spaces, curvature, and similar quantities, to integrate real functions or more complicated quantities, to set up and solve differential equations, or to do all kinds of other mathematics. Indeed, historically manifolds have emerged as an abstraction of such situations in various areas, most notably in the geometry of surfaces, theory of elliptic integrals, complex analysis, and analytical mechanics.

A device for introducing such additional structures on a manifold is an **atlas**. An atlas for a manifold M^n is a collection of pairs $(U_\alpha, \varphi_\alpha)_{\alpha\in\Lambda}$, where each U_α is an open subset of M^n, and $\varphi_\alpha\colon U_\alpha \to \mathbb{R}^n$ is a homeomorphism of the set U_α with an open subset of $\mathbb{R}^n$. Unlike in atlases of Earth, the pairs $(U_\alpha, \varphi_\alpha)$ are called *charts*, not maps,[13] but otherwise, an atlas of Earth gives a good example. All charts together are required to cover all of M^n.

An important concept is a **transition map**. It is a mapping that answers the question, If I am now here on my map, am I also on your map, and where? More formally, if $U_\alpha \cap U_\beta \neq \emptyset$, the transition map $\tau_{\alpha,\beta}$ is the composition $\varphi_\beta\varphi_\alpha^{-1}$, which goes from $\varphi_\alpha(U_\alpha \cap U_\beta)$ to $\varphi_\beta(U_\alpha \cap U_\beta)$, both open sets in $\mathbb{R}^n$.

[13]Or actually, sometimes the φ_α are also called *coordinate maps*.



An atlas is called *differentiable* (or *smooth*) if all of the transition maps are of the class C^∞, i.e., if they have continuous partial derivatives of all orders (this is well defined since the transition maps are mappings between open subsets of $\mathbb{R}^n$). A **differentiable** or **smooth manifold** is a manifold equipped with a differentiable atlas. (One also says that the manifold has a *differentiable structure.*)

Informally, one can think of the differentiablity of an atlas as follows: a path that is drawn as a smooth curve on one of the charts must be smooth on all other charts.

The point of this definition is that on such a manifold, we can define which real functions are differentiable, and develop a (coordinate-free) calculus for them (the latter belongs to differential geometry and we will not discuss it here). Namely, a function $f\colon M^n \to \mathbb{R}$ is differentiable if each of the functions $f\varphi_\alpha^{-1}$ is differentiable (C^∞) in the usual sense. The assumed differentiability of the transition maps guarantees that this definition is globally consistent.

In a similar way, one could define a C^k-atlas for given k; we recall that C^k is the class of functions with continuous partial derivatives up to order k. However, these are seldom mentioned in the literature.

The reason is a theorem of Whitney, stating that once a manifold has a C^k structure S, there is a unique C^∞ structure S' on it that is equivalent, as a C^k-structure, to S. So in this sense, there is only one notion of differentiable manifold.

What differentiable structures are possible on the most basic manifolds like $\mathbb{R}^n$ or S^n? It seems hard to imagine that there could be any others but the standard ones, but people studying this question discovered quite strange phenomena.

As for $\mathbb{R}^n$, the differentiable structure is unique for all n *except* for $n = 4$, where there are uncountably many mutually nonequivalent differentiable structures.

The first *exotic sphere*, i.e., a sphere with differentiable structure not equivalent to the standard one, was discovered by Milnor in dimension 7, and later it was found that S^7 admits 28 nonequivalent structures. The general picture is quite complicated and it is, among others, related to the stable homotopy groups of spheres; in some dimensions there is just one structure, and typically there are finitely many.

The recipe for defining differentiable manifold is quite general. Once we require that the transition maps of an atlas belong to some particular class of mappings, we can say that the corresponding manifold is of that class. The literature abounds with such classes; e.g., *analytic manifolds*, *complex manifolds*, *symplectic manifolds*, or *contact manifolds*.

Another important class, whose definition does not fit the pattern above, though, are *Riemannian manifolds*—see, e.g., the book [Ber03] for interesting views of Riemannian geometry, which we neglect here completely.

Two, three, four, many. The difficulty of studying manifolds depends very much on their dimension. And surprisingly, the dependence is not "the higher dimension, the harder" as one might perhaps expect—it is four-dimensional manifolds that pose the most tantalizing questions, while dimension 5 and above are again easier.

Two-dimensional manifolds. This is a classical area presented in many textbooks. There is a well known complete classification of compact 2-dimensional manifolds, also known as 2-dimensional closed surfaces.

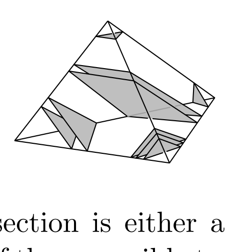
Up to homeomorphism, they fall into two groups: the orientable ones, consisting of the spere S^2, the torus, the double torus, etc. (all are obtained by attaching handles to the sphere), and the unorientable ones, which are the projective plane, the Klein bottle, and others that can be constructed by attaching handles to one of these two. The compact 2-manifolds with boundary are obtained from the ones without boundary by cutting out finitely many disjoint disks.

Three-dimensional manifolds. Lots of knowledge has been accumulated about these, although outstanding problems still remain, many of them of algorithmic nature. The techniques are mostly very specific to this area, which is almost separate from the rest of topology.

A basic tool for three-dimensional manifolds is the *theory of normal surfaces*, which was originally developed by Haken in order to get an algorithm for testing whether a given (piecewise linear) embedding $S^1 \to S^3$ is knotted.

Let us make a digression and sketch the main ideas of this beautiful algorithm. It is well known that an embedding of S^1 in S^3 is unknotted exactly if there is a disk D embedded in S^3 such that the embedded S^1 is its boundary. Given an embedded S^1, the algorithm considers the corresponding knot manifold M (i.e., S^3 minus the slightly thickened knot) and searches for a suitable embedded disk D in M (such that the boundary of D is contained in the torus bounding M and goes "once around" it). We assume that M is triangulated, and that the triangulation has t tetrahedra.

Assuming that such a D exist, we pull it taut within M and then perturb it slightly. Then the intersection of D with each tetrahedron is of the type shown in the picture:



Each component of the intersection is either a triangle, one of four possible types, or a quadrilateral, one of three possible types (where two types of quadrilaterals cannot coexist in a single tetrahedron). (A *normal surface* in general is a surface in M whose intersections with all tetrahedra have the just explained form.)

For every tetrahedron τ, we write down 7 nonnegative integers, corresponding to the number of components of the intersection $D \cap \tau$ of each of the 7 possible types, where the types are enumerated in some fixed order. Doing this for every tetrahedron, we obtain a vector $v_D \in \mathbb{Z}^{7t}_{\geq 0}$, the *coordinate vector* of D.

The coordinate vector describes D uniquely up to a continuous deformation within M. Moreover, given a vector $v \in \mathbb{Z}^{7t}_{\geq 0}$, we can check whether it actually describes the desired disk D witnessing unknotedness.

Of course, this does not yet give an algorithm, since there are infinitely many nonnegative integer vectors. We observe that by far not all such vectors actually describe a surface in M; one necessary condition is that the vectors be compatible on the boundaries of the tetrahedra. That is, if T is a triangle bounding a tetrahedron τ, then the components of the coordinate vector v_D corresponding to τ determine the number and type of segments of $D \cap T$. Since each triangle, except for those on the boundary of M, bounds two tetrahedra, it imposes a condition on v_D, which can be expressed as several *homogeneous* linear equations for the components of v_D.

Thus, it suffices to look for the desired coordinate vector v_D only among the (nonnegative) solutions of the linear system expressing the compatibility conditions.

Now, crucially, it can be shown (nontrivially) that we can restrict the search to *fundamental solutions*, where a solution v is fundamental if it cannot be written as $v = v_1 + v_2$, where both v_1 and v_2 are nonzero solutions.

A basic result in the theory of integral cones asserts that a system of m linear equations (and possibly inequalities) with n unknowns and with integer coefficients has only a finite number of fundamental nonnegative solutions, and this number can be bounded by a function of m, n, and of the size of the coefficients. The fundamental solutions can also be enumerated algorithmically.

The algorithm thus consists in generating all fundamental solutions and testing each of them if it provides the desired disk D. We refer to Hass et al. [HLP99] for a more detailed presentation.

Algorithms obtained by the method of normal surfaces are typically at least exponential, and the existence of polynomial-time algorithms for some of the problems (e.g., recognizing whether a given simplicial complex is a triangulation of S^3, or detecting knotedness) presents fascinating questions.

Another, more recent tool for studying 3-mainfolds, comes from *Thurston's geometrization conjecture*. Roughly speaking, the conjecture asserts that every compact orientable 3-manifold (without boundary) can be cut into finitely many pieces so that each of the pieces can be endowed with one of 8 very special geometries.

Here "geometry" basically means metric; more precisely, a Riemannian metric. To give a simple example, let us consider 2-dimensional manifolds first. The sphere S^2 with the geodesic metric (measuring shortest distance along the surface) has constant positive curvature. It turns out that S^2 is the *only* orientable 2-manifold that can be given a geometry of constant positive curvature. The torus, for example, admits a flat (zero-curvature) Euclidean geometry, and the double torus a hyperbolic, constant negative curvature geometry. The *uniformization theorem* implies that every compact 2-manifold can be endowed with one of these three types of geometry.

The geometrization conjecture is a similar kind of statement, except that the 3-manifold must in general be cut into "geometrizable" pieces, and that there are 8 possible geometries, some of them quite exotic-looking. But knowing the geometry provides very good understanding of the considered manifold.

The geometrization conjecture was proved by Perel'man in 2003, building on the work of Thurston and Hamilton, by an approach through a nonlinear partial differential equation (Perel'man's techniques have also had great impact in PDE's).

The most celebrated result of Perel'man, which is a consequence of the geometrization conjecture but does not need the full strength of it, was a proof of the **Poincaré conjecture**.

Poincaré was initially wondering whether the homology groups characterize the 3-sphere among compact 3-manifolds. Soon he found a counterexample, known as the *Poincaré homology sphere*, one of the important and elegant examples in topology, which has a nontrivial fundamental group and thus cannot be homeomorphic to S^3. So next he asked whether every compact simply connected 3-manifold is homeomorphic to S^3, and this is what became known as the Poincaré conjecture (and one of the Clay Institute's "Millennium Problems").

A (nontrivial) reformulation of the Poincaré conjecture is whether every 3-manifold homotopy equivalent to S^3 must also be homeomorphic to S^3, and in this form, it makes sense for every dimension n. This is the *generalized Poincaré conjecture.*

Interestingly, the $n = 3$ was the last case to be solved: the generalized Poincaré conjecture for $n \geq 5$ was proved in the 1960s, and for $n = 4$ in 1982.

Dimensions 5 and more. Five- and higher-dimensional manifolds were originally thought to be even more difficult than four-dimensional ones, but after a breakthrough of Smale in the 1960s, they are now much better understood, especially in the differentiable case. Smale proved a celebrated result known as the *h-cobordism* theorem, claiming that certain kind of equivalence of simply connected manifolds in dimension at least 5 actually implies homeomorphism. An immediate consequence was the generalized Poincaré conjecture for $n \geq 5$.

Later on, a *surgery theory* was developed, which provides an algebraic classification of in these dimensions manifolds. Since homeomorphism of two given triangulated manifolds of every dimension $n \geq 4$ is algorithmically undecidable, the classification is "inefficient" in a sense, but it has been used with success to solve various concrete problems. For example, these methods have been used in the classification of exotic spheres (nonequivalent differential structures in S^n), as well as in results related to the Hauptvermutung and triangulability of manifolds mentioned earlier.

Four-dimensional manifolds. These are the most problematic ones. Some of the higher-dimensional theory was extended, mainly by Freedman, to work with 4-dimensional *topological* manifolds (and in this way the generalized Poincaré conjecture was also proved for $n = 4$), but not at all for differentiable manifolds, and basic questions about them remain unresolved.

Scorpan's book [Sco05] is devoted to 4-manifolds, has many nice pictures, and in the first part it also explains material around the h-cobordism theorem, how things work for dimension 5 and more and why they fail in dimension 4.

15 Literature

A usual source for general topology is Kelley [Kel75]; our favorite book is Engelking [Eng89]. For an accessible introduction to combinatorial and algebraic topology the best recommendation we can provide are two books of Prasolov [Pra06, Pra07] (in [Pra95] the same author explains quite sophisticated topological examples almost entirely by intuitive pictorial problems and challenges [Pra95]). A standard, on-line accessible, and quite readable (not to be confused with easy) textbook of algebraic topology is Hatcher [Hat01].

For manifolds, differential topology and such the literature is vast, and some references have already been given earlier. For the beginning it is advisable to read Milnor; all of his books and lecture notes, although quite old by now, seem hard to beat in quality and accessibility.

A good, if a bit dated, introduction to knots and 3-dimensional topology is Rolfsen [Rol90]. Hempel [Hem76] is a book on 3-manifolds. Category theory is presented in [ML98] by its inventor; another more recent book is Adámek et al. [AHS06].

References

[AH00] R. Aharoni and P. Haxell. Hall's theorem for hypergraphs. *J. Graph Theory*, 35(2):83–88, 2000.

[Aha01] R. Aharoni. Ryser's conjecture for tripartite 3-graphs. *Combinatorica*, 21(1):1–4, 2001.

[AHS06] J. Adámek, H. Herrlich, and G. E. Strecker. Abstract and concrete categories: the joy of cats. *Repr. Theory Appl. Categ.*, 17, 2006. Reprint of the 1990 original [Wiley, New York], also available online.

[Ber03] M. Berger. *A panoramic view of Riemannian geometry.* Springer-Verlag, Berlin, 2003.

[Bjö95] A. Björner. Topological methods. In R. Graham, M. Grötschel, and L. Lovász, editors, *Handbook of Combinatorics*, volume II, chapter 34, pages 1819–1872. North-Holland, Amsterdam, 1995.

[Bjö03] A. Björner. Nerves, fibers and homotopy groups. *J. Combin. Theory Ser. A*, 103:88–93, 2003.

[CdVGG12] É. Colin de Verdière, G. Ginot, and X. Goaoc. Multinerves and Helly numbers of acyclic families. In *Proc. ACM Sympos. Comput. Geometry (SoCG'12)*, pages 209–217. ACM, New York, 2012.

[ČKM+14] M. Čadek, M. Krčál, J. Matoušek, F. Sergeraert, L. Vokřínek, and U. Wagner. Computing all maps into a sphere. *J. ACM*, 61(3), 2014. Article No. 17. Preprint in arXiv:1105.6257.

[DLRS10] J. A. De Loera, J. Rambau, and F. Santos. *Triangulations*, volume 25 of *Algorithms and Computation in Mathematics*. Springer-Verlag, Berlin, 2010.

[Eng89] R. Engelking. *General topology*. Heldermann Verlag, Berlin, second edition, 1989.

[Fri12] G. Friedman. An elementary illustrated introduction to simplicial sets. *Rocky Mountain J. Math.*, 42(2):353–423, 2012.

[Hat01] A. Hatcher. *Algebraic Topology*. Cambridge University Press, Cambridge, 2001. Electronic version available at `http://math.cornell.edu/~hatcher#AT1`.

[Hem76] J. Hempel. 3-*Manifolds*. Princeton University Press, Princeton, N. J., 1976. Ann. of Math. Studies, No. 86.

[HKR13] M. Herlihy, D. Kozlov, and S. Rajsbaum. *Distributed Computing Through Combinatorial Topology*. Morgan Kaufmann, 2013.

[HLP99] J. Hass, J. C. Lagarias, and N. Pippenger. The computational complexity of knot and link problems. *J. ACM*, 46(2):185–211, 1999.

[Kel75] J. L. Kelley. *General topology*. Springer-Verlag, New York-Berlin, 1975.

[Mat03] J. Matoušek. *Using the Borsuk–Ulam theorem*. Springer, Berlin etc., 2003.

[Mes01] R. Meshulam. The clique complex and hypergraph matching. *Combinatorica*, 21(1):89–94, 2001.

[ML98] S. Mac Lane. *Categories for the working mathematician*. Springer-Verlag, New York, second edition, 1998.

[Mun84] J. R. Munkres. *Elements of Algebraic Topology*. Addison-Wesley, Reading, MA, 1984.

[Pra95] V. V. Prasolov. *Intuitive topology*. American Mathematical Society, Providence, RI, 1995.

[Pra06] V. V. Prasolov. *Elements of combinatorial and differential topology*. American Mathematical Society, Providence, RI, 2006.

[Pra07] V. V. Prasolov. *Elements of homology theory.* American Mathematical Society, Providence, RI, 2007.

[Rol90] D. Rolfsen. *Knots and links.* Publish or Perish, Inc., Houston, TX, 1990. Corrected reprint of the 1976 original.

[RS12] J. Rubio and F. Sergeraert. Constructive homological algebra and applications. Preprint, arXiv:1208.3816, 2012. Written in 2006 for a MAP Summer School at the University of Genova.

[Rud01] Yu. B. Rudyak. Piecewise linear structures on topological manifolds. Preprint arXiv:math/0105047, 2001.

[Sco05] A. Scorpan. *The wild world of 4-manifolds.* American Mathematical Society, Providence, RI, 2005.

[Wac07] M. L. Wachs. Poset topology: tools and applications. In *Geometric combinatorics*, pages 497–615. Amer. Math. Soc., Providence, RI, 2007.