Metric spaces

Žiga Virk

October 4, 2021

Topology and geometry study the shapes of spaces. In the first lecture we will introduce our basic objects: metric spaces. These are sets with a meaningful notion of a distance (metric). The focus will be on an intuitive understanding of three equivalence types of metric spaces: Isometry type, homeomorphic type, and homotopic type of spaces. These types will play a crucial role in later sections.

1 Definition of metric spaces and basic examples

Definition 1.1. *A **metric space** (X, d) is a pair consisting of a set X and a function $d\colon X \times X \to [0, \infty)$, such that for any $x, y, z \in X$ the following hold:*

- $d(x, y) = 0$ *iff* $x = y$,
- *symmetry:* $d(x, y) = d(y, x)$, *and*
- *triangle inequality:* $d(x, z) \leq d(x, y) + d(y, z)$.

*Function d is referred to as a **distance** or a **metric**.*

If X is introduced as a metric space, we implicitly assume d or d_X is the metric on X, unless stated otherwise. Also, $\mathbb{R}^n$ is always considered to be equipped with the Euclidean d_2 metric (see Example 1.2 for definition), unless stated otherwise.

Example 1.2. *The following are some of the metrics d_* on $\mathbb{R}^n$. For $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ we define:*

- $d_1(x, y) = \sum_{i=1}^n |x_i - y_i|$
- $d_2(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$
- $d_p(x, y) = \sqrt[p]{\sum_{i=1}^n (x_i - y_i)^p}$ *for* $p > 1$
- $d_\infty(x, y) = \max_{i \in \{1,2,\ldots,n\}} |x_i - y_i|$

Example 1.3. *Occasionally the underlying space is different that $\mathbb{R}^n$. Here are some examples:*

- *Suppose X is a finite graph with a length associated with each edge. The geodesic distance d_g between two points in X is the length of the shortest path between these points in X.*

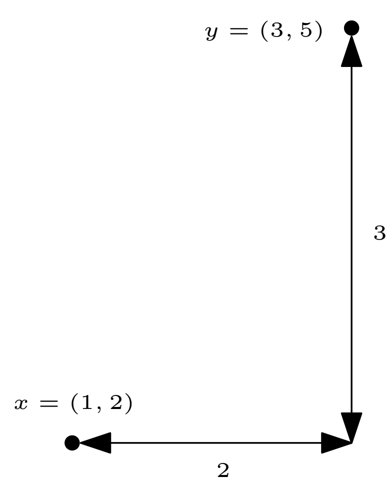

Figure 1: A few distances: $d_1(x, y) = 5$, $d_2(x, y) = \sqrt{13}$, $d_\infty(x, y) = 3$.

- *Suppose X is a surface. We can think of it as a sphere or the surface of the earth. Similarly as above, the geodesic distance* d_g *between two points on X is the length of the shortest path between these points on X. For example, consider the distance between London and Sydney (see Figure 2). The distance usually thought of in this case is the geodesic distance on Earth, that is, the length of the shortest path between the two cities. The actual Euclidean distance in space between the two cities is shorter, but usually not of interest, since the path that realizes it passes fairly close to the center of the Earth.*

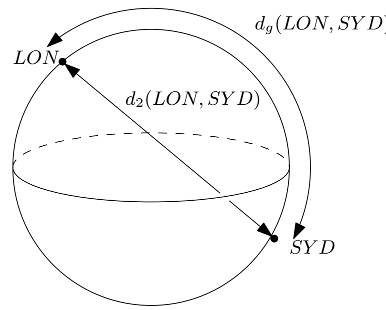

Figure 2: Geodesic vs d_2 distance between London and Sydney.

- *Suppose A is a finite set which we call alphabet. Let X denote a set of finite sequences (words) consisting of the elements of A (letters). The Levenshtein distance between two words in X is defined as the minimum number of edits required to transform one word into another, where the allowed edits are:*
  - *an insertion of a letter at any position;*
  - *a deletion of a letter anywhere;*
  - *a substitution of a letter in any place by another letter.*

  See Figure 3 for example.

$$DOG \quad DOL \quad DOLF \quad WOLF$$

Figure 3: Levenshtein distance between DOG and WOLF is 3 by the following argument. The sequence above demonstrates that the distance is at most 3. As WOLF has three letters that do not appear in DOG, the distance is at least 3.

- *Let X be a finite set and let* 2^X *be the collection of all subsets of X. The Jaccard distance on* 2^X *is defined as*

$$d_J(A,B) = \frac{|A \cup B| - |A \cap B|}{|A \cup B|}.$$

For a metric space (X, d), $x \in X$ and $r > 0$ we define the closed[1] r-**ball** around x as

$$B_d(x,r) = \{y \in X \mid d(x,y) \leq r\}.$$

When the metric is apparent from the context we omit it and use $B(x,r)$.

[1] As we will only consider closed balls, the phrase will be simplified to just "balls".

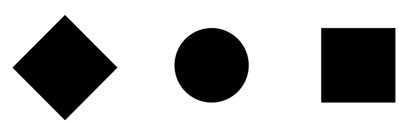

Figure 4: Balls in d_1, d_2 and d_∞ metric in the plane.

2 Maps and equivalence types

When transforming or mapping spaces we will always be using continuous maps.

Definition 2.1. *A map* $f\colon X \to Y$ *between metric spaces is* ***continuous*** *if for each* $x \in X$ *and for each* $\varepsilon > 0$ *there exists* $\delta > 0$ *so* *that the following holds for all $y \in X$:*

$$d_X(x,y) < \delta \implies d_Y(f(x), f(y)) < \varepsilon.$$

The notion of continuity between metric spaces includes the classical continuity from infinitesimal calculus, i.e., all continuous elementary functions $\mathbb{R} \to \mathbb{R}$ are continuous in the sense of Definition 2.1 on $(\mathbb{R}, d_2)$.

An equivalent definition of continuity could be stated in terms of convergent sequences. A sequence of points $\{z_i\}_{i\in\mathbb{N}}$ in Z converges to $w \in Z$ (notation $z_i \overset{i\to\infty}{\longrightarrow} w$) iff $d_Z(z_i, w)$ converges to zero. It turns out that a map $f\colon X \to Y$ between metric spaces is continuous if the following implication holds: If $\{x_i\}_{i\in\mathbb{N}}$ is any sequence in X with $x_i \overset{i\to\infty}{\longrightarrow} u \in X$, then $f(x_i) \overset{i\to\infty}{\longrightarrow} f(u) \in Y$. A practical interpretation of continuity would be the following: if we improve our measurements x_i in the sense that we get a better approximation for the desired state w, then the values over a continuous map $f(x_i)$ also converge to the value $f(w)$. For example, suppose we want to estimate the area of Madagascar from a .bmp image representing a map of the island. We expect that as the resolution increases, we should get a better estimate for the total area.

A continuous map $g\colon [0,1] \to X$ is called a **path** from $g(0)$ to $g(1)$.

Next, we give three different equivalence relations on the class of metric spaces, each of which preserves different level of geometric information. We start with the strictest equivalence, which preserves the most structure.

Definition 2.2. *A map $f\colon X \to Y$ between metric spaces is an **isometry**, if it is bijective and preserves distances, i.e., for each $x_1, x_2 \in X$, $d_X(x_1,x_2) = d_Y(f(x_1), f(x_2))$. Two metric spaces are **isometric**, if there exists an isometry between them.*

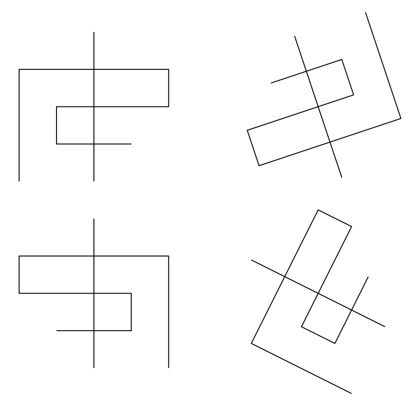

Figure 5: Four isometric planar sets.

Isometries of the plane are combinations of translations, rotations and reflections. In $\mathbb{R}^n$, isometries are combinations of a translation and a linear isometry. Linear isometries in $\mathbb{R}^n$ are represented by orthogonal matrices.

It turns out that no patch of a sphere (equipped with the geodesic metric) is[2] isometric to a subset of a plane. A practical consequence of this fact is that all topographic maps are distorted.

[2] This is a consequence of Gauss' Theorem Egregium.

Although isometries are convenient in many situations, they are essentially a geometric notion that is too rigid for topological treatment. We next introduce a topological counterpart.

> **Definition 2.3.** *A map $f\colon X \to Y$ between metric spaces is a* ***homeomorphism****, if it is bijective, continuous, and f^{-1} is continuous. Two metric spaces are* ***homeomorphic*** *(or of the same topological type; notation: $X \cong Y$), if there exists a homeomorphism between them.*

Obviously, every isometry is a homeomorphism. While homeomorphisms are much more flexible and preserve a number of invariants of a space (later we will mention dimension, number of components and holes, etc.), they do not preserve some of the geometric properties, e.g. diameter (the supremum of pairwise distances in a space), radii of the smallest enclosing balls, etc.

We will often be referring to the following two spaces:

- For $n \in \mathbb{N}$ an n-**sphere** S^n is any space homeomorphic to the n-dimensional sphere

$$\{x \in \mathbb{R}^{n+1} \mid d_2(x,0) = 1\} = \Big\{(x_1,\dots,x_{n+1}) \in \mathbb{R}^{n+1} \mid \sum_{i=1}^{n+1} x_i^2 = 1\Big\},$$

  where 0 is the $(n+1)$-tuple of zeros. Observe that S^0 consists of two points, S^1 is homeomorphic to a circle[3], S^2 to the usual sphere, etc.

- For $n \in \mathbb{N}$ an n-**disc** D^n is any space homeomorphic to

$$\{x \in \mathbb{R}^n \mid d_2(x,0) \leq 1\},$$

  where 0 is the n-tuple of zeros[4]. Observe that D^1 is a closed interval, whose endpoints are S^0. Similarly, D^2 can be thought of as the unit disc in the plane. Note that its boundary in the plane is S^1.

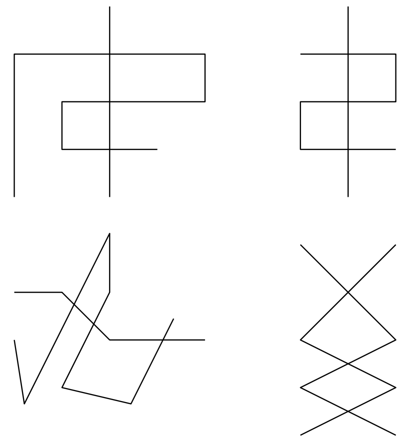

Figure 6: Four homeomorphic sets in the plane.

[3] A circle is a 1-dimensional subset of $\mathbb{R}^2$ defined by $(x-a)^2 + (y-b)^2 = r^2$, i.e., it is "empty inside".

[4] A clarification on terminology: A ball (a metric concept) in a metric space is a particular specific subspace of that metric space. An n-disc (a topological concept) is any space homeomorphic to the standard unit ball in $\mathbb{R}^n$, and thus defined up to homeomorphism. A square in the plane is a 2-disc, but is not a ball in the Euclidean metric. Any unit ball of radius at least 1 on a circle of circumference 1 is the entire circle and so is not a 1-disc.

Example 2.4. *Here we provide some examples of homeomorphisms.*

- *Two finite metric spaces are homeomorphic iff they consist of the same number of points. Each map between finite metric spaces is continuous.*
- *Any two closed intervals are homeomorphic. In particular, a homeomorphism $f\colon [0,1] \to [a,b]$ for $a < b$ is given by $f(t) = a + t(b-a)$.*
- *A square $[-1,1]^2$ in the plane is homeomorphic to the ball $B((0,0),1)$ in the plane. One of the homomorphisms is given by a radial map $B((0,0),1) \to [-1,1]^2$ mapping:*
  - $(0,0) \mapsto (0,0)$ *and*
  - $(x,y) \mapsto \rho(x,y)\cdot(x,y)$ *for*

$$\rho(x,y) = \frac{\sqrt{x^2+y^2}}{\max\{|x|,|y|\}} = \frac{d_2\Big((0,0),(x,y)\Big)}{d_\infty\Big((0,0),(x,y)\Big)}.$$

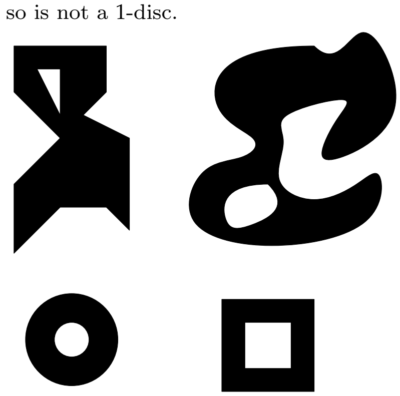

Figure 7: Four homeomorphic sets in the plane.

- *All three balls in Figure 4 are homeomorphic.*
- *For each $n \in \mathbb{N}$, $S^n \cong S^m$ iff $n = m$. We will prove this result using homology in a later chapter.*
- *For each $n \in \mathbb{N}$, $D^n \cong D^m$ iff $n = m$.*
- *No n-disc is homeomorphic to any k-sphere. Each n-sphere can be obtained as a union of two n-discs acting as hemispheres.*

While homeomorphism is the focal equivalence in the field of topology, it turns out that many computable invariants are in fact invariant with respect to a continuous deformation of spaces. These deformations are formalized by the concept of homotopy.

Definition 2.5. *Continuous maps $f, g \colon X \to Y$ between metric spaces are* ***homotopic*** *[$f \simeq g$] if there exists a continuous deformation of f into g, i.e., if there exists a map $H \colon X \times [0,1] \to Y$, such that $H|_{X\times\{0\}} = f$ and $H|_{X\times\{1\}} = g$. Map H is called* ***homotopy****.*

Another way to think about homotopy between f and g would be as a continuous collection of paths from $f(x)$ to $g(x)$ in X.

Example 2.6. *Some examples concerning homotopies:*

1. *For each metric space X, any two maps $f, g \colon X \to \mathbb{R}^n$ are homotopic. A homotopy consists of line segments between $f(x)$ and $g(x)$. In particular,*
$$H(x,t) = (1-t)f(x) + tg(x).$$
2. *Let $w \in S^1$. Then the identity map $id \colon S^1 \to S^1$ is not homotopic to the constant map $c_w \colon S^1 \to S^1$, which maps each point to w. Later we will be able to prove this fact using homology. Note that by the previous example both maps are homotopic in $\mathbb{R}^2$, hence the relation of being homotopic depends on the target space of the maps.*
3. *Consider the two spaces on Figure 12. Space X is a single point, space Y consists of a point, an empty triangle (S^1), a square (D^2) and a disc with a tail. Observe that there are four homotopy classes of maps from X to Y, one for each component of Y.*

We are now ready to introduce homotopy equivalence.

Definition 2.7. *Metric spaces X and Y are* ***homotopy equivalent*** *[$X \simeq Y$] if there exist maps $f \colon X \to Y$ and $g \colon Y \to X$, such that $f \circ g \simeq id_Y$ and $g \circ f \simeq id_X$. Maps f and g are called* ***homotopy***

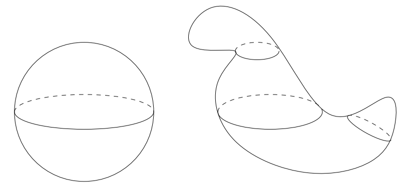

Figure 8: Two homeomorphic surfaces.

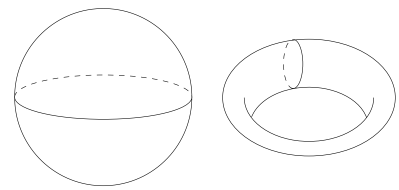

Figure 9: Two non-homeomorphic surfaces.

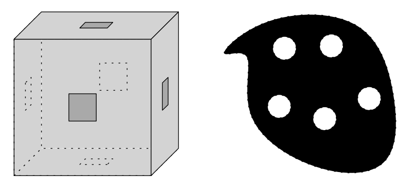

Figure 10: The surface of a cube with a puncture in each of the six sides is homeomorphic to a planar set with five holes.

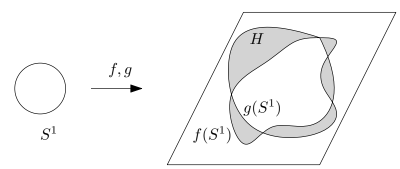

Figure 11: Two maps in the plane are homotopic.

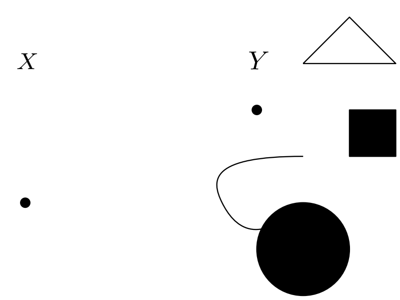

Figure 12: There are four homotopy classes of maps from a single point space X to Y.

equivalences.

Homeomorphic spaces are homotopy equivalent. A metric space X is **contractible**, if it is homotopy equivalent to the one-point space.

Example 2.8. *Some examples concerning homotopy equivalences:*

- *Let* $X = [0,1]$ *and* $Y = \{0\}$. *Then* $X \simeq Y$, *i.e.,* $[0,1]$ *is contractible. Map* $f\colon X \to Y$ *is the constant map and map* $g\colon Y \to X$ *can be chosen to be any map, say* $g(0) = 0$. *Composition* $f \circ g$ *is identity. It remains to show that* $h = g \circ f\colon [0,1] \to [0,1]$, *which is the constant map at* 0, *is homotopic to the identity. Such a homotopy is, for example, the linear homotopy from 1. of Example 2.6. In the same way we can prove that* D^n *is contractible for each* $n \in \mathbb{N}$.
- *Contractible spaces include convex sets and trees.*
- *It turns out that no* S^n *is contractible. The case* $n = 1$ *follows from 2. of Example 2.6.*
- $\mathbb{R}^n \setminus \{(0,0,\ldots,0)\} \simeq S^{n-1}$.

Homotopy equivalence does not preserve all topological properties (for example, dimension), but it does preserve many of those that we can compute: the number of components, holes, etc.

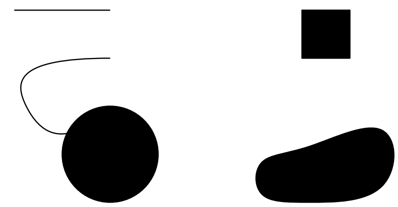

Figure 13: Four contractible spaces.

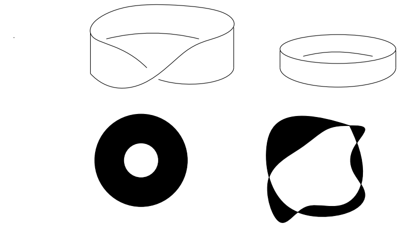

Figure 14: Four spaces homotopy equivalent to S^1: Moebius band (top left), usual band $S^1 \times [0,1]$ (top right) and two planar sets below. Only two of them are homeomorphic.

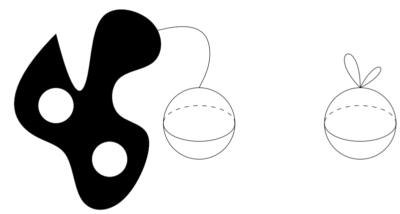

Figure 15: Two more homotopy equivalent spaces.

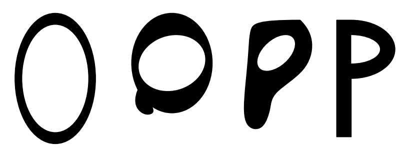

Figure 16: A sequence of steps deforming O to P. While the figure demonstrated a continuous deformation (homotopy equivalence), the spaces presented in this case are actually homeomorphic.

Connectedness

The first homotopy invariant we will mention is connectedness. There are a few versions of it in topology. We will focus on the one generated by paths.

Definition 2.9. *Space* X *is* ***path connected***, *if for each* $x, y \in X$ *there exists a path from* x *to* y *in* X.

Subset $A \subseteq Y$ *of a metric space* Y *is a* ***path component***, *if it is a maximal path connected subset.*

A space is path connected iff it is itself a path component. As was mentioned above, path connectedness is a homotopy invariant: if X is path connected and $Y \simeq X$, then Y is also path connected. Similarly, the number of path components of a metric space is a homotopy invariant. Space Y on Figure 12 has four components.

⚠ From now on we will be dropping adjective “path” and only refer to “connectedness”, and “components”.

3 Concluding remarks

Recap (highlights) of this chapter

- Metric spaces;

- Isometry;
- Homeomorphism;
- Homotopy equivalence;
- Connectedness.

Background and applications

Mathematics is the language of science and scientific concepts are modelled by mathematical objects. These objects can range from simple to sophisticated: a simple Boolean value (0 or 1, i.e., TRUE or FALSE), a numeric value (integer, real, complex, etc.), a collection of numeric values (e.g., a point in $\mathbb{R}^n$), a collection of points in $\mathbb{R}^n$, a function, a vector space, a probability distribution, a graph, a matrix, a metric space, etc. For most of these notions, there is a useful notion of a metric that transfers the possible outputs into a metric space and thus into the realm of geometry and topology, some of which we will explore here.

The notions introduced in this chapter are covered in standard books on topology.