Notes on Topology

David K. Zhang

Last updated January 16, 2022

1 Introduction

Topology is the mathematical study of "stretchy geometry"; that is, the study of geometric properties that do not change when a space is continuously deformed. This means that topology does not deal with concrete, measurable quantities like lengths and angles, since these can change under continuous deformations. Instead, topology studies much more basic, abstract properties, like whether a space is connected or disconnected, and how many dimensions a given space has.

Topology draws much coarser, more flexible distinctions than the rigid plane geometry that we all learned in grade school. For example, from the perspective of topology, triangles, squares, and circles are all the same shape. If you drew a circle on a rubber sheet, you could imagine that with enough pulling and stretching (but not tearing — that's discontinuous!), you could make that circle look like a triangle or square.

Now, you might ask why topology is worth studying if it can't even tell the difference between a circle and a square. Indeed, topology is not very useful for studying geometry in a 2D plane or in 3D space, where we, as humans, have a highly-developed sense of spatial intuition. We don't need an abstract definition of "connectedness" when we can tell whether a 2D shape is connected or not just by looking.

The true utility of topology lies in studying higher-dimensional spaces, or spaces defined in other abstract ways, where our visual and spatial intuition no longer serves us. Even though we could never concretely visualize lengths and angles in such a space, topology gives us the tools to conceptualize and understand them in a precise, logical fashion. This will allow us to reason about abstract spaces and answer mathematical questions that would otherwise be impossible.

2 Continuity

Before we can begin studying topology, we first need to define what we mean by the terms "geometric space" and "continuous deformation." While our discussion will be guided by familiar examples, such as the real number line $\mathbb{R}$ and the plane $\mathbb{R}^2$, we ultimately want to construct

definitions that are as general as possible. As a starting point, if you've ever taken a class on calculus or real analysis, you may have learned how to define the notion of **continuity** for a real-valued function of a real variable.

Definition: Continuity of a Function $\mathbb{R} \to \mathbb{R}$ at a Point

A function $f : \mathbb{R} \to \mathbb{R}$ is **continuous at a point** $x_0 \in \mathbb{R}$ if, for every $\epsilon > 0$, there exists $\delta > 0$ such that for all $x \in \mathbb{R}$, if $|x - x_0| < \delta$, then $|f(x) - f(x_0)| < \epsilon$.

Definition: Continuity of a Function $\mathbb{R} \to \mathbb{R}$

A function $f : \mathbb{R} \to \mathbb{R}$ is **continuous** if f is continuous at every point $x_0 \in \mathbb{R}$.

This definition formalizes the idea that a function is continuous if wiggling the input x by a small amount only causes a small wiggle in the output $f(x)$. It introduces the notion of an "input window" of size δ, defined by $|x - x_0| < \delta$, and an "output window" of size ϵ, defined by $|f(x) - f(x_0)| < \epsilon$. It then requires that, as δ shrinks to zero, so too should ϵ.

Now, the only reason that f needs to map real numbers to real numbers in this definition is to allow us to define the input window $|x - x_0| < \delta$ and the output window $|f(x) - f(x_0)| < \epsilon$ using the operations of subtraction and absolute value. The purpose of the expression $|x - x_0|$ is to measure the **distance** between the numbers x and x_0. This suggests that if we had another means of measuring distance, then we could define a more general notion of continuity for functions between sets other than $\mathbb{R}$.

Definition: Metric, Distance Function

Let X be a set. A **metric** on X, also known as a **distance function** on X, is a function $d : X \times X \to \mathbb{R}$ that satisfies the following requirements:

- **Positive-definiteness**: $d(x, y) \geq 0$ for all $x, y \in X$, and $d(x, y) = 0$ if and only if $x = y$.
- **Symmetry**: $d(x, y) = d(y, x)$ for all $x, y \in X$.
- **Triangle inequality**: $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in X$.

For example, the function $d(x, y) := |x - y|$ is a metric on the real number line $\mathbb{R}$, and the Euclidean distance function $d(\mathbf{x}, \mathbf{y}) := \sqrt{(\mathbf{x} - \mathbf{y})^T (\mathbf{x} - \mathbf{y})}$ is a metric on $\mathbb{R}^n$. These metrics are so commonly used that we call them the **usual metrics** on $\mathbb{R}$ and $\mathbb{R}^n$, respectively.

Definition: Metric Space, Underlying Set, Point

A **metric space** is an ordered pair (X, d) consisting of a set X, called the **underlying set** of the metric space, and a metric d on X. The elements of X are called the **points** of the metric space (X, d).

We also generalize the notion of a "window" to an arbitrary metric space. In the context of the real numbers, the "window" $|x - x_0| < \delta$ is nothing more than the open interval $(x_0 - \delta, x_0 + \delta)$. In a general metric space, we call a set of this type an **open ball**.

Definition: Open Ball, $B_r(x_0)$

Let (X, d) be a metric space, $x_0 \in X$, and $r \in \mathbb{R}$. The **open ball** of radius r centered at x_0, denoted by $B_r(x_0)$, is the set defined by

$$B_r(x_0) := \{x \in X : d(x, x_0) < r\}.$$

This is called an *open* ball in analogy with open intervals in $\mathbb{R}$, because it does not contain the points on the boundary of distance exactly r away from x_0. Of course, we can also define a ball that *does* contain these boundary points, which is called a **closed ball**.

Definition: Closed Ball, $B_r[x_0]$

Let (X, d) be a metric space, $x_0 \in X$, and $r \in \mathbb{R}$. The **closed ball** of radius r centered at x_0, denoted by $B_r[x_0]$, is the set defined by

$$B_r[x_0] := \{x \in X : d(x, x_0) \le r\}.$$

The notation $B_r(x_0)$ and $B_r[x_0]$ is inspired by the notation for open intervals (a, b) and closed intervals $[a, b]$ in $\mathbb{R}$. With these definitions in hand, we are now ready to generalize the definition of continuity to a function between any two metric spaces.

Definition: Continuity at a Point in a Metric Space

Let (X, d_X) and (Y, d_Y) be metric spaces, and let $x_0 \in X$. A function $f : X \to Y$ is **continuous at x_0** if for every $\epsilon > 0$, there exists $\delta > 0$ such that for all $x \in B_\delta(x_0)$, we have $f(x) \in B_\epsilon(f(x_0))$.

Definition: Continuity in a Metric Space

Let (X, d_X) and (Y, d_Y) be metric spaces. A function $f : X \to Y$ is **continuous** if it is continuous at every point $x_0 \in X$.

Definition: Open Set

Let (X, d) be a metric space. A subset $U \subseteq X$ is called an **open set** if, for every point $x \in U$, there exists a radius $r > 0$ such that $B_r(x) \subseteq U$.

Definition: Closed Set

Let (X, d) be a metric space. A subset $F \subseteq X$ is called a **closed set** if its complement $X \setminus F$ is an open set.

Note that in topology, the words "open" and "closed" are not antonyms! It is possible for a subset of a metric space to be open, closed, neither, or both.

Definition: Image

Let X and Y be sets. Given a function $f : X \to Y$, the **image** of a subset $A \subseteq X$ under f,

denoted by $f[A]$, is the set defined by

$$f[A] := \{f(a) : a \in A\}.$$

In other words, $f[A]$ is the subset of Y containing all elements that f maps to from A. For example, consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) := x^2$. The image of the interval $(-1, 1)$ under f is the interval $f[(-1, 1)] = [0, 1)$.

Definition: Inverse Image, Preimage

Let X and Y be sets. Given a function $f : X \to Y$, the **inverse image** or **preimage** of a subset $B \subseteq Y$ under f, denoted by $f^{-1}[B]$, is the set defined by

$$f^{-1}[B] := \{x \in X : f(x) \in B\}.$$

In other words, $f^{-1}[B]$ is the subset of X containing all elements that map into B under f. For example, consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) := x^2$. The inverse image of the interval $(1, 4)$ is the union of the two intervals $f^{-1}[(1, 4)] = (-2, -1) \cup (1, 2)$.

Continuity $\iff$ Preimages of Open Sets are Open

Theorem: Let (X, d_X) and (Y, d_Y) be metric spaces. A function $f : X \to Y$ is continuous if and only if, for every open set $U \subseteq Y$, the inverse image $f^{-1}[U]$ is open in X.

- -

Proof: Suppose f is continuous, and let an open set $U \subseteq Y$ be given. We must show that the set $f^{-1}[U]$ is open, i.e., for every point $x_0 \in f^{-1}[U]$, there exists a radius $r > 0$ such that $B_r(x_0) \subseteq f^{-1}[U]$. By definition, since $x_0 \in f^{-1}[U]$, we know that $f(x_0) \in U$. Moreover, we know that U is an open set, so there exists a radius $\epsilon > 0$ such that $B_\epsilon(f(x_0)) \subseteq U$. Using the definition of continuity, it follows that there exists $\delta > 0$ such that $f[B_\delta(x_0)] \subseteq B_\epsilon(f(x_0))$. It follows that $B_\delta(x_0) \subseteq f^{-1}[U]$, which is the desired result.

Conversely, suppose that the preimage of every open set in Y under f is open in X. Let $x_0 \in X$ and $\epsilon > 0$ be given. We must show that there exists $\delta > 0$ such that $f[B_\delta(x_0)] \subseteq B_\epsilon(f(x_0))$. Since $B_\epsilon(f(x_0))$ is an open set, it follows that $f^{-1}[B_\epsilon(f(x_0))]$ is also open. Now, $x_0 \in f^{-1}[B_\epsilon(f(x_0))]$, so by definition, there exists a radius $r > 0$ such that $B_r(x_0) \subseteq f^{-1}[B_\epsilon(f(x_0))]$. It follows that $f[B_r(x_0)] \subseteq B_\epsilon(f(x_0))$, which is the desired result.

List of Definitions

List of Theorems