Lecture Notes on Topology for MAT3500/4500 following J. R. Munkres' textbook

John Rognes

November 21st 2018

Contents

Introduction

Topology (from Greek *topos* [place/location] and *logos* [discourse/reason/logic]) can be viewed as the study of continuous functions, also known as maps. Let X and Y be sets, and $f\colon X \to Y$ a function from X to Y. In order to make sense of the assertion that f is a continuous function, we need to specify some extra data. After all, continuity roughly asserts that if x and y are elements of X that are "close together" or "nearby", then the function values $f(x)$ and $f(y)$ are elements of Y that are also close together. Hence we need to give some sense to a notion of closeness for elements in X, and similarly for elements in Y.

In many cases this can be done by specifying a real number $d(x, y)$ for each pair of elements $x, y \in X$, called the distance between x and y, and saying at x and y are close together if $d(x, y)$ is sufficiently small. This leads to the notion of a *metric space* (X, d), when the distance function (or metric) d satisfies some reasonable properties.

The only information available about two elements x and y of a general set X is whether they are equal or not. Thus a set X appears as an unorganized collection of its elements, with no further structure. When (X, d) is equipped with a metric, however, it acquires a shape or form, which is why we call it a *space*, rather than just a set. Similarly, when (X, d) is a metric space we refer to the $x \in X$ as *points*, rather than just as elements.

However, metric spaces are somewhat special among all shapes that appear in Mathematics, and there are cases where one can usefully make sense of a notion of closeness, even if there does not exist a metric function that expresses this notion. An example of this is given by the notion of pointwise convergence for real functions. Recall that a sequence of functions f_n for $n = 1, 2, \dots$ converges pointwise to a function g if for each point t in the domain the sequence $f_n(t)$ of real numbers converges to the number $g(t)$. There is no metric d on the set of real functions that expresses this notion of convergence.

To handle this, and many other more general examples, one can use a more general concept than that of metric spaces, namely *topological spaces.* Rather than specifying the distance between any two elements x and y of a set X, we shall instead give a meaning to which subsets $U \subset X$ are "open". Open sets will encode closeness as follows:

> If U is open and $x \in U$, then all $y \in X$ that are "sufficiently close" to x also satisfy $y \in U$.

The shape of X is thus defined not by a notion of distance, but by the specification of which subsets U of X are open. When this specification satisfies some reasonable conditions, we call X together with the collection of all its open subsets a "topological space". The collection of all open subsets will be called the *topology* on X, and is usually denoted $\mathscr{T}$.

As you can see, this approach to the study of shapes involves not just elements and functions, like the theory of metric spaces, but also subsets and even collections of subsets. In order to argue effectively about topological spaces, it is therefore necessary to have some familiarity with the basic notions of set theory. We shall therefore start the course with a summary of the fundamental concepts concerning sets and functions.

Having done this, we can reap some awards. For instance, the definition of what it means for a function $f: X \to Y$, from a topological space X to a topological space Y, to be continuous, is simply:

> For each open subset V in Y the preimage $f^{-1}(V)$ is open in X.

This may be compared with the (ϵ, δ)-definition for a function $f: X \to Y$, from a metric space (X, d) to another metric space (Y, d), to be continuous:

> For each point x in X and each $\epsilon > 0$ there exists a $\delta > 0$ such that for each point y in X with $d(x, y) < \delta$ we have $d(f(x), f(y)) < \epsilon$.

It may be worth commenting that the definition of a topological space may seem more abstract and difficult to fully comprehend than the subsequent definition on a continuous map. The situation is analogous to that in linear algebra, where we say that a function $f: V \to W$ between real vector spaces is linear if it satisfies

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$$

for all vectors $x, y \in V$ and all real numbers λ and μ (the Greek letters 'lambda' and 'mu').

However, to make sense of this, we must first give the abstract definition of a real vector space, as a set V of vectors, with a vector sum operation $+: V \times V \to V$ and a scalar multiplication $\cdot: \mathbb{R} \times V \to V$ satisfying a list of properties. That list in turn presupposes that the set of real numbers $\mathbb{R}$ is a *field*, which also involves the two operations addition and multiplication and about nine axioms, expressing associativity, commutativity, existence of neutral elements and inverses, for both sum and product, plus the distributive law relating the two operations.

The moral is that the axiomatization of the most fundamental objects, such as topological spaces and real vector spaces, may be so general as to make it difficult to immediately grasp their scope. However, it is often the *relations* between these objects that we are most interested in, such as the properties of continuous functions or linear transformations, and these will then often appear to be relatively concrete.

Once we have established the working definitions of topological spaces and continuous functions, or maps, we shall turn to some of the most useful properties that such topological spaces may satisfy, including being *connected* (not being a disjoint union of open subsets), *compact* (not having too many open subsets globally) or *Hausdorff* (having enough open subsets locally). Then we discuss consequences of these properties, such as general forms of the intermediate value theorem, existence of maximal values, or uniqueness of limits, and many more. In the case of real functions of one real variable, these are familiar from first-year Calculus, but the generalized results apply to a vastly wider range of shapes, including the plane and higher-dimensional Euclidean spaces, infinite-dimensional function spaces, and finite partially ordered sets.

These lecture notes are intended for the course MAT4500 at the University of Oslo, following James R. Munkres' textbook "Topology". The §-signs refer to the sections in that book.

Once the foundations of Topology have been set, as in this course, one may proceed to its proper study and its applications. A well-known example of a topological result is the classification of surfaces, or more precisely, of connected compact 2-dimensional manifolds. The answer is that two facts about a surface suffice to determine it up to topological equivalence, namely, whether the surface "can be oriented", and "how many handles it has". The number of handles is also known as the *genus*. A sphere has genus 0, while a torus has genus 1, and the surface of a mug with two handles has genus 2.

An interesting result about the relation between the global topological type of a surface and its local geometry is the Gauss–Bonnet theorem. For a surface F equipped with a so-called

Riemannian metric, this is a formula

$$\int_F K \, dA = 2\pi \cdot \chi$$

expressing the integral of the locally defined curvature K of the surface in terms of the globally defined genus g, or more precisely in terms of the Euler characteristic $\chi = 2 - 2g$. (That's the Greek letter 'chi'.) For example, a sphere of radius r has curvature $1/r^2$ everywhere, and surface area $4\pi r^2$. The integral of the curvature over the whole surface is the product of these two quantities, i.e., 4π. This equals 2π times the Euler characteristic of the sphere, which is 2. These results will be covered in the course MAT4510 Geometric Structures, followed by MAT4520 on Manifolds.

Let $[0,1] \subset \mathbb{R}$ be the unit interval. One form of the intermediate value theorem tells us that any continuous function $f\colon [0,1] \to [0,1]$ has a fixed-point, i.e., an element $x \in [0,1]$ such that $f(x) = x$. We will prove this as a consequence of the fact that $[0,1]$ is connected, while $\{0,1\}$ is disconnected. Moving up one dimension, let $[0,1]^2 \subset \mathbb{R}^2$ be the unit square. We will also prove that any continuous function $g\colon [0,1]^2 \to [0,1]^2$ has a fixed-point, i.e., a point $(x,y) \in [0,1]^2$ such that $g(x,y) = (x,y)$. This will be a consequence of the fact that the square $[0,1]^2$ is simply-connected, but the boundary of this square is not simply-connected. Going on to more complicated spaces, the subject of algebraic topology provides tools for analyzing the higher-dimensional analogues of connectedness and simple-connectedness. For instance, the n-dimensional sphere S^n, consisting of the unit vectors in $\mathbb{R}^{n+1}$, admits a continuous vector field of nonzero vectors if and only if n is odd. Thus you "cannot comb a hairy ball flat" in any even dimension. This theory is developed in the course MAT4530 Algebraic Topology I.

When considering surfaces given by algebraic equations among complex numbers, such as

$$x^2 + y^2 = 1$$

or

$$x^5 + y^2 = 1$$

(with $x, y \in \mathbb{C}$) there is also a subtle relationship between the topological type of the solution surface (known as an algebraic curve, since it has real dimension 2 but complex dimension 1) and the number of rational solutions to the equation (with $x, y \in \mathbb{Q}$). The first equation describes a sphere, and has infinitely many rational solutions, while the second equation describes a curve of genus 2, and has only finitely many rational solutions. It was conjectured by Mordell, and proved by Gerd Faltings in 1983, that any rationally defined algebraic curve of genus greater than one has only finitely many rational points. In other words, if the defining equation has rational coefficients, and a topological condition is satisfied, then there are only finitely many rational solutions. For more on algebraic curves, see the courses in Algebraic Geometry.

Chapter 1

Set Theory and Logic

1.1 (§1) Fundamental Concepts

1.1.1 Membership

In naive set theory, a *set* (norsk: mengde) is any collection of Mathematical objects, called its *elements*. We often use uppercase letters, like A or B, to denote sets, and lowercase letters, like x or y, to denote its elements.

If A is a set, and x is one of its elements, we write $x \in A$ and say that "x is an element of A". Otherwise, if x is not an element of A, we write $x \notin A$. The symbol "$\in$" thus denotes *membership* in a collection.

We can specify sets by listing its elements, as in the set of decimal digits:

$$D = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\} ,$$

or by selecting the elements of some previously given set that satisfy some well-defined condition:

$$P = \{n \in \mathbb{N} \mid n \text{ is a prime}\} ,$$

read as "the set of $n \in \mathbb{N}$ such that n is a prime". Here

$$\mathbb{N} = \{1, 2, 3, \dots\}$$

is the set of natural numbers (= positive integers). In this case $691 \in P$, while $693 \notin P$, since 691 is a prime, while $693 = 3 \cdot 3 \cdot 7 \cdot 11$ is not a prime. When the condition begins with a symbol similar to '$|$', we may use a colon instead, as in:

$$B(x, \epsilon) = \{y \in \mathbb{R} : |y - x| < \epsilon\} .$$

We may also specify the list of elements by means of an expression, as in the set

$$S = \{n^2 \mid n \in \mathbb{N}\}$$

of squares. We shall sometimes use informal notations, like $P = \{2, 3, 5, \dots\}$ and $S = \{1, 4, 9, \dots\}$, when it should be clear from the context what we really mean.

Note that for any object x, the *singleton set* $\{x\}$ is different from x itself, so $x \in \{x\}$ but $x \neq \{x\}$. Be careful with the braces!

The *empty set* $\varnothing = \{\}$ has no elements, so $x \notin \varnothing$ for all objects x. If a set A has one or more elements, so that $x \in A$ for some object x, then we say that A is *nonempty*.

1.1.2 Inclusion and equality

Let A and B be sets. We say that A is a *subset* of B, and write $A \subset B$, if each element of A is also an element of B. In logical terms, the condition is that $x \in A$ only if $x \in B$, so $(x \in A) \Longrightarrow (x \in B)$. We might also say that A is contained in B. For example, $\{1\} \subset \{1, 2\}$.

Less commonly, we might say that A is a *superset* of B, or that A contains B, and write $A \supset B$, if each element of B is also an element of A. In logical terms, the condition is that $x \in A$ if $x \in B$, so $(x \in A) \Longleftarrow (x \in B)$. This is of course equivalent to $B \subset A$. For example, $\{1, 2\} \supset \{2\}$.

We say that A is *equal* to B, written $A = B$, if $A \subset B$ and $B \subset A$. This means that $x \in A$ if and only if $x \in B$, so $(x \in A) \iff (x \in B)$. For example, $\{1, 2\} = \{1, 1, 2\}$, since the notion of a set only captures whether an element is an element of a set, not how often it is listed.

If A is not a subset of B we might write $A \not\subset B$, and if A is not equal to B we write $A \neq B$. If $A \subset B$ but $A \neq B$, so that A is a *proper subset* of B, we write $A \subsetneq B$.

(In other texts you may find the alternate notations $A \subseteq B$ and $A \subset B$ for $A \subset B$ and $A \subsetneq B$, respectively.)

1.1.3 Intersection and union

Let A and B be sets. The *intersection* $A \cap B$ is the set of objects that are elements in A and in B:

$$A \cap B = \{x \mid x \in A \text{ and } x \in B\} .$$

Clearly $A \cap B \subset A$ and $A \cap B \subset B$. We say that A *meets* B if $A \cap B \neq \varnothing$ is nonempty, so that there exists an x with $x \in A \cap B$, or equivalently, with $x \in A$ and $x \in B$.

The *union* $A \cup B$ is the set of objects that are elements in A or in B:

$$A \cup B = \{x \mid x \in A \text{ or } x \in B\} .$$

Note that "or" in the mathematical sense does not exclude the possibility that both $x \in A$ and $x \in B$. Hence $A \subset A \cup B$ and $B \subset A \cup B$.

In addition to the commutative and associative laws, these operations satisfy the following two *distributive laws*, for all sets A, B and C:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

1.1.4 Difference and complement

Let A and B be sets. The *difference* $A - B$ is the set of objects of A that are not elements in B:

$$A - B = \{x \in A \mid x \notin B\} .$$

Note that $A - B \subset A$, $(A - B) \cap B = \varnothing$ and $(A - B) \cup B = A \cup B$.

We shall also call $A - B$ the *complement* of B in A. (Some texts denote the difference set by $A \setminus B$, or introduce a notation like $\complement B$ for the complement of B in A, if $B \subset A$ and the containing set A is implicitly understood.)

The complement of a union is the intersection of the complements, and the complement of an intersection is the union of the complements. These rules are known as *De Morgan's laws*:

$$A - (B \cup C) = (A - B) \cap (A - C)$$
$$A - (B \cap C) = (A - B) \cup (A - C)$$

1.1.5 Collections of sets, the power set

A set is again a mathematical object, and may therefore be viewed as an element of another set. When considering a set whose elements are sets, we shall usually refer to it as a *collection* of sets (norsk: samling), and denote it with a script letter like $\mathscr{A}$ or $\mathscr{B}$. (Sometimes sets of sets are called *families*.)

For example, each student at the university may be viewed as a mathematical object. We may consider the set of all students:

$$S = \{s \mid s \text{ is a student at UiO}\}\,.$$

Similarly, each course offered at the university may be viewed as another mathematical object. There is a set of courses

$$C = \{c \mid c \text{ is a course at UiO}\}\,.$$

For each course $c \in C$, we may consider the set E_c of students enrolled in that course:

$$E_c = \{s \in S \mid s \text{ is enrolled in } c\}\,.$$

Now we may consider the collection $\mathscr{E}$ of these sets of enrolled students:

$$\mathscr{E} = \{E_c \mid c \in C\}\,.$$

This $\mathscr{E}$ is a set of sets. Its elements are the sets of the form E_c, for some course $c \in C$. These sets in turn have elements, which are students at the university.

It may happen that no students are enrolled for a specific course c. In that case, $E_c = \varnothing$. If this is the case for two different courses, c and d, then both $E_c = \varnothing$ and $E_d = \varnothing$. Hence it may happen that $E_c = E_d$ in $\mathscr{E}$, even if $c \neq d$ in C.

For a given set A, the collection of all subsets $B \subset A$ is called the *power set* of A, and is denoted $\mathscr{P}(A)$:

$$\mathscr{P}(A) = \{B \mid B \subset A\}\,.$$

For example, the power set $\mathscr{P}(\{a,b\}) = \{\varnothing, \{a\}, \{b\}, \{a,b\}\}$ of the two-element set $\{a,b\}$ has four elements.

1.1.6 Arbitrary intersections and unions

Given a collection $\mathscr{A}$ of sets, the *intersection* of the elements of $\mathscr{A}$ is

$$\bigcap_{A \in \mathscr{A}} A = \{x \mid \text{for every } A \in \mathscr{A} \text{ we have } x \in A\}$$

and the *union* of the elements of $\mathscr{A}$ is

$$\bigcup_{A \in \mathscr{A}} A = \{x \mid \text{there exists an } A \in \mathscr{A} \text{ such that } x \in A\}\,.$$

When $\mathscr{A} = \{A, B\}$, these are the same as the previously defined sets $A \cap B$ and $A \cup B$, respectively. When $\mathscr{A} = \{A\}$ is a singleton set, both are equal to A.

When $\mathscr{A} = \varnothing$ is empty, the intersection $\bigcap_{A \in \mathscr{A}} A$ could be interpreted as the "set of all x", but this leads to set-theoretic difficulties. We shall therefore only use that notation in the context of subsets A of a fixed "universal set" X, in which case $\bigcap_{A \in \varnothing} A = X$. There is no difficulty with the empty union: $\bigcup_{A \in \varnothing} A = \varnothing$.

Returning to the example of students and courses, the intersection

$$A = \bigcap_{E_c \in \mathscr{E}} E_c$$

is the set of those students that are enrolled to every single course at the university, which most likely is empty. The union

$$B = \bigcup_{E_c \in \mathscr{E}} E_c$$

is the subset of S consisting of those students that are enrolled to one or more courses. They might be referred to as the active students. The complement, $S - B$, is the set of students that are not enrolled to any courses. They might be referred to as the inactive students.

1.1.7 Cartesian products

Given two sets A and B, the *cartesian product* $A \times B$ consists of the ordered pairs of elements (x, y) with $x \in A$ and $y \in B$:

$$A \times B = \{(x, y) \mid x \in A \text{ and } y \in B\} .$$

The notion of an ordered pair (x, y) is different from that of the set $\{x, y\}$. For example, $(x, y) = (x', y')$ if and only if $x = x'$ and $y = y'$. (If desired, one can define ordered pairs in terms of sets by letting $(x, y) = \{\{x\}, \{x, y\}\}$.)

The cartesian product of two copies of the real numbers, $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$, is the set consisting of pairs (x, y) of real numbers. Thinking of these two numbers as the (horizontal and vertical) coordinates of a point in the plane, one is led to René Descartes' formulation of analytic geometry, as opposed to Euclid's synthetic approach to geometry.

1.2 (§2) Functions

1.2.1 Domain, range and graph

A *function* f from a set A to a set B is a rule that to each element $x \in A$ associates a unique element $f(x) \in B$. We call A the *domain* (or *source*) of f, and B the *range* (or *codomain*, or *target*) of f. We use notations like $f \colon A \to B$ or

$$A \xrightarrow{f} B$$

to indicate that f is a function with domain A and range B. Note that the sets A and B are part of the data in the definition of the function f, even though we usually emphasize the rule taking x to $f(x)$. To show both aspects of the function, we might write

$$\begin{aligned} f \colon A &\longrightarrow B \\ x &\longmapsto f(x) . \end{aligned}$$

Sometimes that rule is defined by some explicit procedure or algorithm for computing $f(x)$ from x, such as in

$$f(x) = \begin{cases} 3x + 1 & \text{if } x \text{ is odd} \\ x/2 & \text{if } x \text{ is even} \end{cases}$$

for natural numbers x, but we make no such assumption in general.

In set theoretic terms, we can define a function $f\colon A \to B$ to be the subset of $A \times B$ given by its *graph*, i.e., the subset

$$\Gamma_f = \{(x, f(x)) \in A \times B \mid x \in A\}\,.$$

(Here Γ is the upper-case Greek letter 'Gamma'.) The subsets $\Gamma \subset A \times B$ that arise in this way are characterized by the property that for each $x \in A$ there exists one and only one $y \in B$ such that $(x, y) \in \Gamma$.

Hence we can define a function f to be a triple of sets (A, B, Γ), with $\Gamma \subset A \times B$ having the property that for each element $x \in A$ there exists a unique $y \in B$ with $(x, y) \in \Gamma$. In this case we let $f(x) = y$ and call y the *value* of f at the *argument* x. We call A the domain and B the range of f.

1.2.2 Image, restriction, corestriction

Let $f\colon A \to B$ be a function. The *image* of f is the subset

$$f(A) = \{f(x) \in B \mid x \in A\}$$

of the range B, whose elements are all the values of f. The image may, or may not, be equal to the range. (Other texts may refer to the image of f as its range, in which case they usually have no name for the range/codomain/target.)

If $S \subset A$ is a subset of the domain, we define the *restriction* of f to S to be the function $f|S\colon S \to B$ given by $(f|S)(x) = f(x)$ for all $x \in S$. In terms of graphs, $f|S$ corresponds to the subset

$$\Gamma_f \cap (S \times B)$$

of $S \times B$, where $\Gamma_f \subset A \times B$ is the graph of f.

If $T \subset B$ is a subset of the range with the property that $f(A) \subset T$, then there is also a well-defined function $g\colon A \to T$ given by $g(x) = f(x)$ for all $x \in T$. In terms of graphs, g corresponds to the subset

$$\Gamma_f \cap (A \times T)$$

of $A \times T$. There does not seem to be a standard notation for this "corestriction" of f. Note that this construction only makes sense when the new range T contains the image of f.

1.2.3 Injective, surjective, bijective

Let $f\colon A \to B$. We say that f is *injective* (or *one-to-one*) if $f(x) = f(y)$ only if $x = y$, for $x, y \in A$. We say that f is *surjective* (or *onto*) if for each $y \in B$ there exists a $x \in A$ with $f(x) = y$. Note that f is surjective if and only if its image equals its range, $f(A) = B$.

We say that f is *bijective* (or a *one-to-one correspondence*) if it is both injective and surjective, so that for each $y \in B$ there exists one and only one $x \in A$ with $f(x) = y$. A bijective function is also called a *bijection*.

When $f\colon A \to B$ is bijective, we can define a new function $f^{-1}\colon B \to A$, with the domain and range interchanged, by the rule that takes $y \in B$ to the unique $x \in A$ such that $f(x) = y$. Hence $f^{-1}(y) = x$ precisely when $y = f(x)$. We call f^{-1} the *inverse function* of f.

Note that the graph of f^{-1} is obtained from the graph of f by interchanging the two factors in the cartesian product $A \times B$. It is the subset

$$\Gamma_{f^{-1}} = \{(y, x) \in B \times A \mid (x, y) \in \Gamma_f\}$$

of $B \times A$.

The inverse function is not defined when f is not bijective. If the restriction $f|S$ of f to a subset $S \subset A$ is injective, and we let $T = f(S)$ be the image of the restricted function, then the resulting function $g \colon S \to T$ is bijective, and has an inverse function $g^{-1} \colon T \to S$. (Various abuses of notations are common here.)

1.2.4 Composition

Let $f \colon A \to B$ and $g \colon B \to C$ be functions, such that the range of f equals the domain of g. The *composite function* $g \circ f \colon A \to C$ is then defined by the rule

$$(g \circ f)(x) = g(f(x))$$

for all $x \in A$. We often abbreviate $g \circ f$ to gf. The graph Γ_{gf} of gf is the subset

$$\{(x, z) \in A \times C \mid \text{there exists a } y \in B \text{ with } (x, y) \in \Gamma_f \text{ and } (y, z) \in \Gamma_g\}$$

of $A \times C$.

When $f \colon A \to B$ is bijective, with inverse function $f^{-1} \colon B \to A$, the composite $f^{-1} \circ f \colon A \to A$ is defined and equals the identity function $id_A \colon A \to A$ taking $x \in A$ to $x \in A$. Furthermore, the composite $f \circ f^{-1} \colon B \to B$ is defined, and equals the identity function $id_B \colon B \to B$ taking $y \in B$ to $y \in B$.

Composition of functions is unital and associative, so that $f \circ id_A = f = id_B \circ f$ and $(h \circ g) \circ f = h \circ (g \circ f)$ (for $h \colon C \to D$), but hardly ever commutative. Even if $A = B = C$, so that both $g \circ f$ and $f \circ g$ are defined and have the same domains and ranges, it is usually not the case that $g(f(x)) = f(g(x))$ for all $x \in A$, so usually $g \circ f \neq f \circ g$. One exception is the case when $f = g$.

The composite $g \circ f \colon A \to C$ of two bijections $f \colon A \to B$ and $g \colon B \to C$ is again bijective, with inverse $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Let $f \colon A \to B$ and $S \subset A$. Define the *inclusion function* $i \colon S \to A$ by the rule $i(x) = x$ for all $x \in S$. This is not the identity function, unless $S = A$. The composite $f \circ i \colon S \to B$ equals the restriction $f|S \colon S \to B$, since these functions have the same sources and ranges, and both map $x \in S$ to $f(x) \in B$.

Let $T \subset B$ and assume that $f(A) \subset T$. Let $j \colon T \to B$ be the inclusion function given by the rule $j(y) = y$ for all $y \in T$. Then the corestriction $g \colon A \to T$ of f is characterized by the property that $j \circ g = f$.

1.2.5 Images of subsets

Let $f \colon A \to B$ be a function. For each subset $S \subset A$ we let the *image* $f(S) \subset B$ of S under f be the set of values

$$f(S) = \{f(x) \in B \mid x \in S\}$$

of f, as the argument x ranges over S. When $S = A$, this agrees with the image of f.

The rule taking S to $f(S)$ defines a function

$$\begin{aligned} f \colon \mathscr{P}(A) &\longrightarrow \mathscr{P}(B) \\ S &\longmapsto f(S) . \end{aligned}$$

Using the same symbol for this function $\mathscr{P}(A) \to \mathscr{P}(B)$ and the original function $f \colon A \to B$ is an abuse of notation. Since the two functions are defined on disjoint sets, namely for $x \in A$ and $S \in \mathscr{P}(A)$, respectively, one can usually avoid confusion by inspecting the argument of f, but some care is certainly appropriate.

The image function f respects inclusions and unions: If $S \subset T \subset A$, then

$$f(S) \subset f(T) .$$

Similarly, if $S, T \subset A$, meaning that $S \subset A$ and $T \subset A$, then

$$f(S \cup T) = f(S) \cup f(T) .$$

For intersections, we only have the inclusion

$$f(S \cap T) \subset f(S) \cap f(T)$$

in general. The inclusions $S \cap T \subset S$ and $S \cap T \subset T$ imply inclusions $f(S \cap T) \subset f(S)$ and $f(S \cap T) \subset f(T)$, and these imply the displayed inclusion. If f is injective, we have equality, but in general this can be a proper inclusion.

For complements, we have the inclusion

$$f(T) - f(S) \subset f(T - S)$$

for $S, T \subset A$, with equality if f is injective, but no equality in general.

1.2.6 Preimages of subsets

Let $f\colon A \to B$ as before. For each subset $T \subset B$ we let the *preimage* $f^{-1}(T) \subset A$ of T under f be the set of arguments

$$f^{-1}(T) = \{x \in A \mid f(x) \in T\}$$

for which f takes values in T. When $T = B$ this is all of A. (The preimage is also called the *inverse image*.)

The rule taking T to $f^{-1}(T)$ defines a function

$$\begin{aligned} f^{-1}\colon \mathscr{P}(B) &\longrightarrow \mathscr{P}(A) \\ T &\longmapsto f^{-1}(T) . \end{aligned}$$

Note that we use this notation also in the cases where f is not bijective, i.e., even if the inverse function $f^{-1}\colon B \to A$ is not defined. So the use of the notation $f^{-1}(T)$ for the preimage of T under f does not imply that f is invertible.

In the special case when $f\colon A \to B$ is bijective, so that the inverse function $f^{-1}\colon B \to A$ is defined, we have the equality of sets

$$\{x \in A \mid f(x) \in T\} = \{f^{-1}(y) \in A \mid y \in T\}$$

so that the preimage $f^{-1}(T)$ of T under f is equal to the image $f^{-1}(T)$ of T under f^{-1}. Hence the potential conflict of notations does not lead to any difficulty.

The preimage function f^{-1} respects inclusions, unions, intersections and complements: If $S \subset T \subset B$, then

$$f^{-1}(S) \subset f^{-1}(T) .$$

If $S, T \subset B$, then

$$\begin{aligned} f^{-1}(S \cup T) &= f^{-1}(S) \cup f^{-1}(T) , \\ f^{-1}(S \cap T) &= f^{-1}(S) \cap f^{-1}(T) \end{aligned}$$

and

$$f^{-1}(S - T) = f^{-1}(S) - f^{-1}(T) .$$

The image and preimage constructions satisfy the relations

$$S \subset f^{-1}(f(S)) \quad \text{and} \quad f(f^{-1}(T)) \subset T$$

for $f\colon A \to B$, $S \subset A$ and $T \subset B$.

1.3 (§5) Cartesian Products

1.3.1 Indexed families

Let $\mathscr{A}$ be a collection of sets. An *indexing function* for $\mathscr{A}$ is a surjective function $f\colon J \to \mathscr{A}$ from some set J to $\mathscr{A}$. We call J the *index set*, and we call the collection $\mathscr{A}$ together with the indexing function f an *indexed family of sets*.

If

$$\mathscr{A} = \{A_1, A_2, \dots, A_n\} = \{A_i\}_{i=1}^n$$

is a finite collection, with $n \geq 0$, we may let $J = \{1, 2, \dots, n\}$ and let $f(i) = A_i$ for $1 \leq i \leq n$. If

$$\mathscr{A} = \{A_1, A_2, \dots\} = \{A_i\}_{i=1}^\infty$$

is a countably infinite sequence of sets, we may let $J = \mathbb{N}$ and let $f(i) = A_i$ for $i \in \mathbb{N}$. In general, we often use the notation $A_\alpha = f(\alpha)$, so that f is the rule taking α to A_α, and the indexed family is denoted

$$\{A_\alpha\}_{\alpha \in J} .$$

The surjectivity of f ensures that each set $A \in \mathscr{A}$ occurs as $A_\alpha = f(\alpha)$ for some $\alpha \in J$. We do not require that f is injective, so we may have $A_\alpha = A_\beta$ even if $\alpha \neq \beta$ in J.

1.3.2 General intersections and unions

We use the following alternate notations for general intersections and unions of sets. Let $\mathscr{A}$ be a collection of sets, with indexing function $f\colon J \to \mathscr{A}$ as above.

If $\mathscr{A}$ is nonempty (so J is nonempty, too), we define

$$\bigcap_{\alpha \in J} A_\alpha = \bigcap_{A \in \mathscr{A}} A$$

to be the set of x such that $x \in A_\alpha$ for all $\alpha \in J$, which is the same as the set of x such that $x \in A$ for all $A \in \mathscr{A}$.

In general, we define

$$\bigcup_{\alpha \in J} A_\alpha = \bigcup_{A \in \mathscr{A}} A$$

to be the set of x such that $x \in A_\alpha$ for some $\alpha \in J$, or equivalently, such that $x \in A$ for some $A \in \mathscr{A}$.

If $\mathscr{A} = \{A_1, A_2, \dots, A_n\}$ with $n \geq 1$ we also write

$$\bigcap_{i=1}^n A_i = A_1 \cap A_2 \cap \dots \cap A_n$$

for the intersection of the sets in $\mathscr{A}$, and

$$\bigcup_{i=1}^n A_i = A_1 \cup A_2 \cup \dots \cup A_n$$

for the union, and similarly in the countably infinite case.

1.3.3 Finite cartesian products

Let $n \geq 0$. Given a set X, an *n-tuple* of elements in X is a function

$$x\colon \{1, 2, \dots, n\} \to X\,.$$

Writing $x_i = x(i)$ for its value at $1 \leq i \leq n$, the function is determined by its list of values, which is the ordered n-tuple

$$(x_i)_{i=1}^n = (x_1, x_2, \dots, x_n)\,.$$

A family of sets $\mathscr{A} = \{A_1, A_2, \dots, A_n\}$ indexed by the set $J = \{1, 2, \dots, n\}$ is equivalent to an ordered n-tuple of sets $(A_1, A_2, \dots, A_n)$. Let

$$X = A_1 \cup A_2 \cup \dots \cup A_n$$

be the union of the n sets. We define the *cartesian product* of this indexed family, denoted

$$\prod_{i=1}^n A_i = A_1 \times A_2 \times \dots \times A_n\,,$$

to be the set of all n-tuples $(x_1, x_2, \dots, x_n)$ in X where $x_i \in A_i$ for each $1 \leq i \leq n$.

If all of the sets A_i are equal, so that each $A_i = X$, we write

$$X^n = X \times X \times \dots \times X$$

(n copies of X) for the n-fold cartesian product of X. It is the set of all n-tuples in X.

1.3.4 Countable cartesian products

We use similar notation for *sequences* in X, which are functions

$$x\colon \mathbb{N} \to X$$

or ordered sequences

$$(x_i)_{i=1}^\infty = (x_1, x_2, \dots)\,.$$

(These are also called *ω-tuples*, where ω is the Greek letter 'omega'. Here ω denotes the smallest infinite ordinal, corresponding to the well-ordered set underlying $\mathbb{N}$.) Given a sequence of sets $\mathscr{A} = (A_i)_{i=1}^\infty = (A_1, A_2, \dots)$ we let $X = \bigcup_{i=1}^\infty A_i$ be the union, and define the cartesian product

$$\prod_{i=1}^\infty = A_1 \times A_2 \times \dots$$

to be the set of all sequences $(x_1, x_2, \dots)$ in X such that $x_i \in A_i$ for each $i \in \mathbb{N}$.

When all $A_i = X$, we write

$$X^\omega = \prod_{i=1}^\infty X$$

for the countably infinite product of copies of X. It is the set of sequences in X.

1.3.5 General cartesian products

Let J be any indexing set. Given a set X, a *J-tuple* of elements in X is a function

$$x\colon J \to X\,.$$

Writing $x_\alpha = x(\alpha)$ for its value at $\alpha \in J$, the α-th coordinate of x, we can also denote the J-tuple x by its values $(x_\alpha)_{\alpha\in J}$.

Let $\{A_\alpha\}_{\alpha\in J}$ be an indexed family of sets, with union $X = \bigcup_{\alpha\in J} A_\alpha$. Its *cartesian product*, denoted

$$\prod_{\alpha\in J} A_\alpha$$

is the set of all J-tuples $(x_\alpha)_{\alpha\in J}$ of elements in X such that $x_\alpha \in A_\alpha$ for all $\alpha \in J$. In other words, it is the set of functions

$$x\colon J \to \bigcup_{\alpha\in J} A_\alpha$$

such that $x(\alpha) \in A_\alpha$ for all $\alpha \in J$.

When all $A_\alpha = X$, we write

$$X^J = \prod_{\alpha\in J} X$$

for the J-fold product of copies of X. It is the set of functions $x\colon J \to X$.

1.4 (§6) Finite Sets

1.4.1 Cardinality

For each nonnegative integer n, the set

$$\{1, 2, \ldots, n\}$$

of natural numbers less than or equal to n is called a *section* of the natural numbers $\mathbb{N}$. For $n = 0$ this is the empty set $\varnothing$, for $n = 1$ it is the singleton set $\{1\}$.

Lemma 1.4.1. *If there exists an injective function*

$$f\colon \{1, 2, \ldots, m\} \to \{1, 2, \ldots, n\}$$

then $m \le n$.

Proof. We prove this by induction on $n \ge 0$. For $n = 0$ this is clear, since there only exists a function $\{1, 2, \ldots, m\} \to \varnothing$ if $m = 0$. For the inductive step, let $n \ge 1$ and suppose that the lemma holds for $n - 1$. Let $f\colon \{1, 2, \ldots, m\} \to \{1, 2, \ldots, n\}$ be injective. Let $f(m) = k$. Then f restricts to an injective function $g\colon \{1, 2, \ldots, m-1\} \to \{1, 2, \ldots, n\} - \{k\}$. Define a bijection $h\colon \{1, 2, \ldots, n\} - \{k\} \to \{1, 2, \ldots, n-1\}$ by $h(x) = x$ for $x \ne n$, and $h(n) = k$. (If $k = n$ we let h be the identity.) The composite $h \circ g\colon \{1, 2, \ldots, m-1\} \to \{1, 2, \ldots, n-1\}$ is injective, so by the inductive hypothesis we know that $m - 1 \le n - 1$. It follows that $m \le n$, as desired. □

Corollary 1.4.2. *There does not exist an injective function $f\colon \{1, 2, \ldots, m\} \to \{1, 2, \ldots, n\}$ if $m > n$.*

Proposition 1.4.3. *If there exists a bijection $f\colon \{1, 2, \ldots, m\} \to \{1, 2, \ldots, n\}$ then $m = n$.*

Proof. Both f and its inverse f^{-1} are injective, so by the previous lemma we have $m \le n$ and $n \le m$. Thus $m = n$. $\square$

Corollary 1.4.4. *There does not exist a bijective function* $f\colon \{1, 2, \dots, m\} \to \{1, 2, \dots, n\}$ *if* $m \ne n$.

Definition 1.4.5. A set A is *finite* if there exists a bijective function $f\colon A \to \{1, 2, \dots, n\}$ for some $n \ge 0$. In this case, we say that A has *cardinality* n.

For example, the empty set is finite with cardinality 0, and each singleton set is finite with cardinality 1.

Lemma 1.4.6. *The cardinality of a finite set* A *is well-defined. That is, if there exists bijections* $f\colon A \to \{1, 2, \dots, n\}$ *and* $g\colon A \to \{1, 2, \dots, m\}$ *for some* $m, n \ge 0$*, then* $m = n$.

Proof. The composite

$$f \circ g^{-1}\colon \{1, 2, \dots, m\} \to \{1, 2, \dots, n\}$$

is a bijection, with inverse $g \circ f^{-1}$, so $m = n$ by the proposition above. $\square$

1.4.2 Subsets

Lemma 1.4.7. *Let* $A \subset \{1, 2, \dots, n\}$ *be a subset. There exists a bijection* $f\colon A \to \{1, 2, \dots, m\}$ *for some* $m \le n$*. Hence* A *is a finite set, of cardinality* $\le n$.

Proof. We prove this by induction on $n \ge 0$. For $n = 0$ it is clear, since the only subset of $\varnothing$ is $\varnothing$. For the inductive step, let $n \ge 1$ and suppose that the lemma holds for $n - 1$. Let $A \subset \{1, 2, \dots, n\}$ be a subset. If $n \notin A$, then $A \subset \{1, 2, \dots, n - 1\}$, so there exists a bijection $f\colon A \to \{1, 2, \dots, m\}$ for some $m \le n-1$ by the inductive hypothesis. Hence $m \le n$. Otherwise, we have $n \in A$. Let $B = A - \{n\}$. Then $B \subset \{1, 2, \dots, n - 1\}$, and by the inductive hypothesis there exists a bijection $g\colon B \to \{1, 2, \dots, k\}$ for some $k \le n-1$. Let $m = k+1$, so $m \le n$. Define the bijection $f\colon A \to \{1, 2, \dots, m\}$ by $f(x) = g(x)$ if $x \in B$ and $f(n) = m$. This completes the inductive step. $\square$

Proposition 1.4.8. *Let* $A \subsetneq \{1, 2, \dots, n\}$ *be a proper subset. Then the cardinality of* A *is strictly less than* n*, so there does not exist a bijection* $f\colon A \to \{1, 2, \dots, n\}$.

Proof. Since A is a proper subset, we can choose a $k \in \{1, 2, \dots, n\}$ with $k \notin A$. Define a bijection $h\colon \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$ by $h(k) = n$, $h(n) = k$ and $h(x) = x$ for the remaining x. Let $B = h(A)$, so that $h|A\colon A \to B$ is a bijection. Then $B \subset \{1, 2, \dots, n - 1\}$, so by the lemma above there is a bijection $g\colon B \to \{1, 2, \dots, m\}$ with $m \le n - 1$. The composite $g \circ (h|A)\colon A \to \{1, 2, \dots, m\}$ is then a bijection. Hence the cardinality of A is $m \le n-1 < n$. $\square$

Theorem 1.4.9. *If* A *is a finite set, then there is no bijection of* A *with a proper subset of itself.*

Proof. Suppose that $B \subsetneq A$ is a proper subset, and that there exists a bijection $f\colon A \to B$. Since A is finite there is a bijection $g\colon A \to \{1, 2, \dots n\}$ for some $n \ge 0$, where n is the cardinality of A. Let $C = g(B) \subset \{1, 2, \dots, n\}$. Then $g|B\colon B \to C$ is a bijection, and C is a proper subset of $\{1, 2, \dots, n\}$. Hence there is a bijection $h\colon C \to \{1, 2, \dots, m\}$ for some $m < n$. If there exists a bijection $f\colon A \to B$, then the composite bijection $h \circ (g|B) \circ f\colon A \to \{1, 2, \dots, m\}$ would say that the cardinality of A is m, and not equal to n. This contradicts the fact that the cardinality is well-defined, hence no such bijection f exists. $\square$

Corollary 1.4.10. *The set* $\mathbb{N}$ *of natural numbers is not finite.*

Proof. The function $f\colon \mathbb{N} \to \mathbb{N} - \{1\}$ defined by $f(x) = x + 1$ is a bijection of $\mathbb{N}$ with a proper subset of itself. □

Corollary 1.4.11. *Any subset B of a finite set A is finite. If B is a proper subset of A, then its cardinality is strictly less than the cardinality of A.*

1.4.3 Injections and surjections

Proposition 1.4.12. *Let A be a set. The following are equivalent:*

(1) A is finite.

(2) There exists a surjective function $\{1, 2, \dots, n\} \to A$ for some integer $n \geq 0$.

(3) There exists an injective function $A \to \{1, 2, \dots, n\}$ for some integer $n \geq 0$.

Proof. (1) $\implies$ (2): Since A is finite there exists a bijection $f\colon A \to \{1, 2, \dots, n\}$ for some integer $n \geq 0$. Then f^{-1} is surjective, as required.

(2) $\implies$ (3): Let $g\colon \{1, 2, \dots, n\} \to A$ be surjective. Define a function $h\colon A \to \{1, 2, \dots, n\}$ by letting $h(x)$ be the smallest element of the subset

$$g^{-1}(x) = \{i \mid g(i) = x\}\,.$$

of $\{1, 2, \dots, n\}$. This subset is nonempty since g is surjective. Then h is injective, since if $x \neq y$ then $g^{-1}(x)$ and $g^{-1}(y)$ are disjoint, so their smallest elements must be different.

(3) $\implies$ (1): If $f\colon A \to \{1, 2, \dots, n\}$ is injective, let $B = f(A) \subset \{1, 2, \dots, n\}$. The corestriction of f is then a bijection $g\colon A \to B$. By a previous lemma there is a bijective function $h\colon B \to \{1, 2, \dots, m\}$ for some $m \leq n$. The composite bijection $h \circ g\colon A \to \{1, 2, \dots, m\}$ shows that A is finite. □

Proposition 1.4.13. *Finite unions and finite products of finite sets are finite.*

Proof. Let A and B be finite. Choose bijections $f\colon \{1, 2, \dots, m\} \to A$ and $g\colon \{1, 2, \dots, n\} \to B$ for suitable $m, n \geq 0$. We define a surjection

$$h\colon \{1, 2, \dots, m + n\} \to A \cup B$$

by

$$h(x) = \begin{cases} f(x) & \text{if } 1 \leq x \leq m \\ g(x - m) & \text{if } m + 1 \leq x \leq m + n. \end{cases}$$

Here we regard A and B as subsets of $A \cup B$. Hence $A \cup B$ is finite by the proposition above.

By the case $n = 2$, the formula

$$A_1 \cup \dots \cup A_n = (A_1 \cup \dots \cup A_{n-1}) \cup A_n$$

and induction on n it follows that if $A_1, \dots, A_n$ are finite then $A_1 \cup \dots \cup A_n$ is finite, for all $n \geq 0$.

The cartesian product $A \times B$ is the union of the subsets $A \times \{y\}$ for all $y \in B$. Hence if A and B are finite this is a finite union of finite sets, so $A \times B$ is finite.

By the case $n = 2$, the formula

$$A_1 \times \dots \times A_n = (A_1 \times \dots \times A_{n-1}) \times A_n$$

and induction on n it follows that if $A_1, \dots, A_n$ are finite then $A_1 \times \dots \times A_n$ is finite, for all $n \geq 0$. □

Chapter 2

Topological Spaces and Continuous Functions

2.1 (§12) Topological Spaces

2.1.1 Open sets

Definition 2.1.1. Let X be a set. A *topology* on X is a collection $\mathscr{T}$ of subsets of X, such that:

(1) $\varnothing$ and X in $\mathscr{T}$.

(2) For any subcollection $\{U_\alpha\}_{\alpha\in J}$ of $\mathscr{T}$, the union $\bigcup_{\alpha\in J} U_\alpha$ is in $\mathscr{T}$.

(3) For any finite subcollection $\{U_1, \dots, U_n\}$ of $\mathscr{T}$ the intersection $U_1 \cap \dots \cap U_n$ is in $\mathscr{T}$.

A *topological space* $(X, \mathscr{T})$ is a set X with a chosen topology $\mathscr{T}$.

The subsets $U \subset X$ with $U \in \mathscr{T}$ are said to be *open*. Note that this *defines* the property of being open. With this terminology, the axioms above assert that:

(1) $\varnothing$ and X are open (as subsets of X).

(2) The union of any collection of open subsets of X is open.

(3) The intersection of any finite collection of open subsets of X is open.

With the convention that $\varnothing$ is the union of the empty collection of subsets of X, and X is the intersection of the empty collection of subsets of X, one may agree that (1) follows from (2) and (3), but condition (1) is usually included for clarity. We express (2) by saying that $\mathscr{T}$ is closed under (arbitrary) unions, and express (3) by saying that $\mathscr{T}$ is closed under finite intersections.

To check that $\mathscr{T}$ is closed under finite intersections, it suffices to prove that if U_1, $U_2 \in \mathscr{T}$ then $U_1 \cap U_2 \in \mathscr{T}$. This follows by induction on n from the formula

$$U_1 \cap \dots \cap U_n = (U_1 \cap \dots \cap U_{n-1}) \cap U_n \,.$$

2.1.2 Discrete and trivial topologies

Let X be any set. Here are two extreme examples of topologies on X.

Definition 2.1.2. The *discrete topology* on X is the topology $\mathscr{T}_{\text{disc}}$ where *all* subsets $U \subset X$ are defined to be open. Hence the collection of open subsets equals the power set of X: $\mathscr{T}_{\text{disc}} = \mathscr{P}(X)$. It is clear that the axioms of a topology are satisfied, since it is so easy to be open in this topology. We call $(X, \mathscr{T}_{\text{disc}})$ a *discrete topological space.*

The terminology can be explained as follows. Note that for each point $x \in X$, the singleton set $\{x\}$ is a subset of X, hence is a open in the discrete topology. Thus all other points $y \neq x$ of X are separated away from x by this open set $\{x\}$. We therefore think of X with the discrete topology as a space of separate, isolated points, with no close interaction between different points. In this sense, the space is discrete.

Definition 2.1.3. The *trivial topology* on X is the topology $\mathscr{T}_{\text{triv}}$ where only the subsets $\varnothing$ and X are defined to be open. Hence $\mathscr{T}_{\text{triv}} = \{\varnothing, X\}$. It is clear that the axioms of a topology are satisfied, since there are so few collections of open subsets. We call $(X, \mathscr{T}_{\text{triv}})$ a *trivial topological space.* (Some texts call this the *indiscrete topology.*)

This terminology probably refers to the fact that the trivial topology is the minimal example of a topology on X, in the sense that only those subsets of X that axiom (1) demand to be open are open, and no others.

2.1.3 Finite topological spaces

Definition 2.1.4. If X is a finite set, and $\mathscr{T}$ is a topology on X, we call $(X, \mathscr{T})$ a *finite topological space.*

When X is finite, the power set $\mathscr{P}(X)$ and any topology $\mathscr{T} \subset \mathscr{P}(X)$ is finite, so the distinction between finite and arbitrary unions plays no role. Hence to check conditions (2) and (3) for a topology, it suffices to check that if $U_1, U_2 \in \mathscr{T}$ then $U_1 \cup U_2 \in \mathscr{T}$ and $U_1 \cap U_2 \in \mathscr{T}$.

In the case when X is empty, or a singleton set, the discrete topology on X is equal to the trivial topology on X, and these are the only possible topologies on X.

Example 2.1.5. Let $X = \{a, b\}$ be a 2-element set. There are four different possible topologies on X:

(1) The minimal possibility is the trivial topology $\mathscr{T}_{\text{triv}} = \{\varnothing, X\}$.

(2) An intermediate possibility is $\mathscr{T}_a = \{\varnothing, \{a\}, X\}$.

(3) Another intermediate possibility is $\mathscr{T}_b = \{\varnothing, \{b\}, X\}$.

(4) The maximal possibility is the discrete topology $\mathscr{T}_{\text{disc}} = \{\varnothing, \{a\}, \{b\}, X\}$.

We already explained why cases (1) and (4) are topologies. Examples (2) and (3) are known as *Sierpinski spaces.* To see that $\mathscr{T}_a$ is a topology on $\{a, b\}$, note that $\{b\}$ does not occur as the union or the intersection of any collection of sets in $\mathscr{T}_a$. Interchanging the role of a and b we also see that $\mathscr{T}_b$ is a topology.

In the Sierpinski space $(X = \{a, b\}, \mathscr{T}_a)$, the element a is separated away from the other point by the open set $\{a\}$, while the element b not separated away from the other point by any open set. For the only open set containing b is $\{a, b\}$, which also contains a. This means that a is "arbitrarily close" to b, even if b is not arbitrarily close to a. This kind of asymmetry of "closeness" in topological spaces is not seen in metric spaces.

In these examples each collection of subsets containing the trivial topology defined a topology. When X has cardinality 3 this is no longer true.

Example 2.1.6. Let $X = \{a, b, c\}$. There are 29 different topologies on X. Here are nine of them:

(1) The trivial topology $\mathscr{T}_1 = \mathscr{T}_{\text{triv}} = \{\varnothing, X\}$.

(2) $\mathscr{T}_2 = \{\varnothing, \{a\}, X\}$.

(3) $\mathscr{T}_3 = \{\varnothing, \{a, b\}, X\}$.

(4) $\mathscr{T}_4 = \{\varnothing, \{a\}, \{a, b\}, X\}$.

(5) $\mathscr{T}_5 = \{\varnothing, \{a, b\}, \{c\}, X\}$.

(6) $\mathscr{T}_6 = \{\varnothing, \{a\}, \{b\}, \{a, b\}, X\}$.

(7) $\mathscr{T}_7 = \{\varnothing, \{a\}, \{a, b\}, \{a, c\}, X\}$.

(8) $\mathscr{T}_8 = \{\varnothing, \{a\}, \{c\}, \{a, b\}, \{a, c\}, X\}$.

(9) The discrete topology $\mathscr{T}_9 = \mathscr{T}_{\text{disc}} = \mathscr{P}(X)$ (with 8 elements).

The reader should check that each of these is closed under unions and intersections. The remaining topologies on X arise by permuting the elements a, b and c.

Example 2.1.7. Let $X = \{a, b, c\}$. Here are some collections of subsets of X that are not topologies:

(1) $\{\{a\}, \{c\}, \{a, b\}, \{a, c\}\}$ does not contain $\varnothing$ and X.

(2) $\{\varnothing, \{a\}, \{b\}, X\}$ is not closed under unions.

(3) $\{\varnothing, \{a, b\}, \{a, c\}, X\}$ is not closed under intersections.

2.1.4 The cofinite topology

Definition 2.1.8. Let X be a set. Let the *cofinite topology* $\mathscr{T}_{\text{cof}}$ be the collection of subsets $U \subset X$ whose complement $X - U$ is finite, together with the empty set $U = \varnothing$.

The word "cofinite" refers to the fact that complements of finite sets are open, since if $F \subset X$ is finite, then $U = X - F$ has complement $X - U = X - (X - F) = F$, which is finite. Calling $\mathscr{T}_{\text{cof}}$ a "topology" requires justification:

Lemma 2.1.9. *The collection $\mathscr{T}_{\text{cof}}$ is a topology on X.*

Proof. We check the three conditions for a topology.

(1): The subset $\varnothing$ is in $\mathscr{T}_{\text{cof}}$ by definition. The subset X is in $\mathscr{T}_{\text{cof}}$ since its complement $X - X = \varnothing$ is finite.

(2): Let $\{U_\alpha\}_{\alpha \in J}$ be a subcollection of $\mathscr{T}_{\text{cof}}$, so for each $\alpha \in J$ we have that $X - U_\alpha$ is finite, or $U_\alpha = \varnothing$. Let $V = \bigcup_{\alpha \in J} U_\alpha$. We must prove that $X - V$ is finite, or that $V = \varnothing$.

If each U_α is empty, then V is empty. Otherwise, there is a $\beta \in J$ such that $X - U_\beta$ is finite. Since $U_\beta \subset V$, the complements satisfy $X - V \subset X - U_\beta$. Hence $X - V$ is a subset of a finite set, and is therefore finite, as desired.

(3): Let $\{U_1, \dots, U_n\}$ be a finite subcollection of $\mathscr{T}_{\text{cof}}$, so for each $1 \le i \le n$ we have that $X - U_i$ is finite, or $U_i = \varnothing$. Let $W = U_1 \cap \dots \cap U_n$. We must prove that $X - W$ is finite, or $W = \varnothing$.

If some U_i is empty, then $W \subset U_i$ is empty. Otherwise, $X - U_i$ is finite for each $1 \le i \le n$. By De Morgan's law,

$$X - W = X - (U_1 \cap \cdots \cap U_n) = (X - U_1) \cup \cdots \cup (X - U_n) .$$

The right hand side is a finite union of finite sets, hence is again finite. Thus $X - W$ is finite, as desired. □

When X is a finite set, the condition that $X - U$ is finite is always satisfied, so in this case the cofinite topology equals the discrete topology: $\mathscr{T}_{\mathrm{cof}} = \mathscr{T}_{\mathrm{disc}}$.

2.1.5 Coarser and finer topologies

Definition 2.1.10. Let $\mathscr{T}$ and $\mathscr{T}'$ be two topologies on the same set X. We say that $\mathscr{T}$ is *coarser* than $\mathscr{T}'$, or equivalently that $\mathscr{T}'$ is *finer* than $\mathscr{T}$, if $\mathscr{T} \subset \mathscr{T}'$. This means that each subset $U \subset X$ that is open in $(X, \mathscr{T})$ is also open in $(X, \mathscr{T}')$.

Lemma 2.1.11. *The trivial topology is coarser than any other topology, and the discrete topology is finer than any other topology.*

Proof. For any topology $\mathscr{T}$ on X we have

$$\mathscr{T}_{\mathrm{triv}} = \{\varnothing, X\} \subset \mathscr{T} \subset \mathscr{P}(X) = \mathscr{T}_{\mathrm{disc}} .$$

□

The set of topologies on X becomes partially ordered by the "coarser than"-relation. Note that two topologies need not be comparable under this relation. For example, neither one of the two Sierpinski topologies $\mathscr{T}_a$ and $\mathscr{T}_b$ on $\{a, b\}$ is coarser (or finer) than the other.

When X is an infinite set, the cofinite topology is strictly coarser than the discrete topology: $\mathscr{T}_{\mathrm{cof}} \subsetneq \mathscr{T}_{\mathrm{disc}}$. For example, in this case each finite, nonempty subset $F \subset X$ is open in the discrete topology, but not open in the cofinite topology. To see this, note that for F to be open in $\mathscr{T}_{\mathrm{cof}}$ its complement $X - F$ would have to be finite. Then $X = F \cup (X - F)$ would be the union of two finite sets, and therefore would be finite. This contradicts the assumption that X is infinite. Such finite, nonempty subsets $F \subset X$ exist. For example, each singleton set $F = \{x\}$ for $x \in X$ will do. Hence $\mathscr{T}_{\mathrm{cof}} \ne \mathscr{T}_{\mathrm{disc}}$ for infinite X.

Example 2.1.12. Let $X = \mathbb{N}$ be the set of natural numbers. The discrete topology $\mathscr{T}_{\mathrm{disc}}$ on $\mathbb{N}$ is strictly finer than the cofinite topology $\mathscr{T}_{\mathrm{cof}}$ on $\mathbb{N}$, which is strictly finer than the trivial topology $\mathscr{T}_{\mathrm{triv}}$ on $\mathbb{N}$.

Remark 2.1.13. Given two topological spaces X and Y, a function $f \colon X \to Y$ will be said to be *continuous* if:

for each open $V \subset Y$ the preimage $f^{-1}(V)$ is open in X.

Suppose given two topologies $\mathscr{T}$ and $\mathscr{T}'$ on X, with $\mathscr{T} \subset \mathscr{T}'$, so that $\mathscr{T}'$ is finer than $\mathscr{T}$, and $\mathscr{T}$ is coarser than $\mathscr{T}'$. To an analyst considering real-valued functions

$$f \colon X \longrightarrow \mathbb{R}$$

(where $\mathbb{R}$ has a fixed, metric, topology), it is harder for f to be continuous with respect to $\mathscr{T}$ than with respect to $\mathscr{T}'$. The analyst would therefore say that $\mathscr{T}$ is the stronger topology, and $\mathscr{T}'$ is the weaker topology.

On the other hand, to a topologist considering paths

$$g\colon [0,1] \longrightarrow X$$

(where $[0,1] \subset \mathbb{R}$ has a fixed topology), it is easier for g to be continuous with respect to $\mathscr{T}$ than with respect to $\mathscr{T}'$. The topologist would therefore say that $\mathscr{T}$ is the weaker topology, and $\mathscr{T}'$ is the stronger topology.

These (conflicting) terminologies presume that the condition to be continuous is some sort of obstacle or barrier to be overcome, and that the stronger the barrier is, the harder it is to satisfy the condition.

2.1.6 Metric spaces

Definition 2.1.14. A *metric* on a set X is a function $d\colon X \times X \to \mathbb{R}$ such that:

(1) $d(x,y) \geq 0$ for all $x,y \in X$, and $d(x,y) = 0$ if and only if $x = y$.

(2) $d(x,y) = d(y,x)$ for all $x,y \in X$.

(3) $d(x,z) \leq d(x,y) + d(y,z)$ for all $x,y,z \in X$ (the triangle inequality).

A *metric space* (X,d) is a set X with a chosen metric d.

Example 2.1.15. The real line $X = \mathbb{R}$ is a metric space, with distance function $d(x,y) = |y-x|$. More generally, $X = \mathbb{R}^n$ is a metric space with the Euclidean distance

$$d(x,y) = \|y-x\| = \sqrt{(y_1-x_1)^2 + \cdots + (y_n-x_n)^2}$$

for $x = (x_1,\ldots,x_n)$ and $y = (y_1,\ldots,y_n)$. We call $\mathbb{R}^n$ with this metric Euclidean n-space.

Definition 2.1.16. Let (X,d) be a metric space. For each point $x \in X$ and each positive real number $\epsilon > 0$, let

$$B_d(x,\epsilon) = \{y \in X \mid d(x,y) < \epsilon\}$$

be the *ϵ-ball* around x in (X,d).

Definition 2.1.17. Let (X,d) be a metric space. The *metric topology* $\mathscr{T}_d$ on X is the collection of subsets $U \subset X$ satisfying the property: for each $x \in U$ there exists an $\epsilon > 0$ such that $B_d(x,\epsilon) \subset U$.

Lemma 2.1.18. *The collection $\mathscr{T}_d$ is a topology on X.*

Proof. We check the three conditions for a topology.

(1): The subset $\varnothing$ is in $\mathscr{T}_d$ since there are no $x \in \varnothing$ for which anything needs to be checked. The subset X is in $\mathscr{T}$ since for each $x \in U$ we can take $\epsilon = 1$, since $B_d(x,1) \subset X$.

(2): Let $\{U_\alpha\}_{\alpha\in J}$ be a subcollection of $\mathscr{T}_d$. Let $V = \bigcup_{\alpha\in J} U_\alpha$ and consider any $x \in V$. By the definition of the union there exists an $\alpha \in J$ with $x \in U_\alpha$. By the property satisfied by the U_α in $\mathscr{T}_d$, there exists an $\epsilon > 0$ such that $B_d(x,\epsilon) \subset U_\alpha$. Since $U_\alpha \subset V$ it follows that $B_d(x,\epsilon) \subset V$. Hence $V \in \mathscr{T}_d$.

(3): Let $\{U_1,\ldots,U_n\}$ be a finite subcollection of $\mathscr{T}_d$. Let $W = U_1 \cap \cdots \cap U_n$ and consider any $x \in W$. For each $1 \leq i \leq n$ we have $W \subset U_i$ so $x \in U_i$. By the defining property of $\mathscr{T}_d$ there exists an $\epsilon_i > 0$ such that $B_d(x,\epsilon_i) \subset U_i$. Let $\epsilon = \min\{\epsilon_1,\ldots,\epsilon_n\}$. This makes sense since n is finite, and $\epsilon > 0$. Then $B_d(x,\epsilon) \subset B_d(x,\epsilon_i) \subset U_i$ for each $1 \leq i \leq n$, which implies that $B_d(x,\epsilon) \subset W$. Hence $W \in \mathscr{T}_d$. □

Definition 2.1.19. Let (X, d) be a metric space and let $A \subset X$ be any subset. Let $d_A \colon A \times A \to \mathbb{R}$ be the restriction of d to $A \times A \subset X \times X$. Then d_A is a metric on A, so (A, d_A) is a metric space. We call (A, d_A) a metric subspace of (X, d).

Example 2.1.20. Let the n-sphere

$$S^n = \{x \in \mathbb{R}^{n+1} \colon \|x\| = 1\}$$

be the unit sphere in Euclidean $(n+1)$-space. It is a metric subspace, with the restricted distance function $d_r = d_{S^n}$ given by $d_r(x, y) = \|y - x\|$ for $x, y \in S^n \subset \mathbb{R}^{n+1}$. When $n = 1$ we call $S^1 \subset \mathbb{R}^2$ the unit circle. By definition, $S^0 = \{+1, -1\}$ consists of two points, and $S^{-1} = \varnothing$ is empty. Note that the restricted Euclidean metric d_r is different from the intrinsic metric d_i given by minimizing curve length within S^n. Nonetheless, we these two metrics give the same underlying topology on S^n. In fact, the metrics are related by

$$d_r(x, y) = 2 \sin(d_i(x, y)/2)$$

so the collection of ϵ-balls $B_{d_r}(x, \epsilon) \subset S^n$ for $\epsilon > 0$ is equal to the collection of ϵ-balls $B_{d_i}(x, \epsilon) \subset S^n$ for $\epsilon > 0$, just with a different parametrization in terms of ϵ. Hence the collections of open sets are equal: $\mathscr{T}_{d_r} = \mathscr{T}_{d_i}$.

2.2 (§13) Basis for a Topology

2.2.1 Bases

Remark 2.2.1. Recall that given two topological spaces Z and X, a function $g \colon Z \to X$ is continuous if for each open $U \subset X$ the preimage $g^{-1}(U)$ is open in Z. Suppose that

$$U = \bigcup_{\alpha \in J} B_\alpha$$

can be written as the union of a collection $\{B_\alpha\}_{\alpha \in J}$ of open subsets $B_\alpha \subset X$. Then

$$g^{-1}(U) = \bigcup_{\alpha \in J} g^{-1}(B_\alpha) .$$

If each $g^{-1}(B_\alpha)$ is open in Z, then so is their union, by condition (2) for a topology. Hence this will imply that $g^{-1}(U)$ is open. Thus, to verify that g is continuous, it is enough to find a subcollection $\mathscr{B}$ of the topology $\mathscr{T}$ on X such that:

(1) Each open subset of X can be written as a union $\bigcup_{\alpha \in J} B_\alpha$ of sets $B_\alpha \in B$.

(2) $g^{-1}(B)$ is open in Z, for each $B \in \mathscr{B}$.

A collection $\mathscr{B} \subset \mathscr{T}$ satisfying condition (1) will be called a basis for the topology $\mathscr{T}$ on X. In this situation, to prove that a given function $g \colon Z \to X$ is continuous it suffices to verify condition (2), with $B \in \mathscr{B}$ replacing $U \in \mathscr{T}$.

It will often be convenient to define a topology $\mathscr{T}$ by only specifying a basis $\mathscr{B}$ for that topology. The open subsets of X will then be precisely the unions of subcollections of $\mathscr{B}$. In this way the basis $\mathscr{B}$ determines, or generates, $\mathscr{T}$. However, different bases $\mathscr{B}$ and $\mathscr{B}'$ may well generate the same topology. (Compare with the role of bases for vector spaces.)

We now define what it means for a collection $\mathscr{B}$ of subsets of X to be a basis. Then we define the topology $\mathscr{T}$ generated by $\mathscr{B}$, and prove that it satisfied the axioms for a topology. Thereafter we show that $\mathscr{T}$ precisely consists of the unions of subcollections of $\mathscr{B}$, i.e., that the open sets in the topology $\mathscr{T}$ are precisely the sets

$$U = \bigcup_{\alpha \in J} B_\alpha$$

where $\{B_\alpha\}_{\alpha \in J}$ ranges through the subcollections of $\mathscr{B}$.

Definition 2.2.2. Let X be a set. A collection $\mathscr{B}$ of subsets of X is a *basis* (for a topology) if

(1) For each $x \in X$ there exists a $B \in \mathscr{B}$ with $x \in B$.

(2) If $B_1, B_2 \in \mathscr{B}$ and $x \in B_1 \cap B_2$ then there exists a $B_3 \in \mathscr{B}$ with $x \in B_3 \subset B_1 \cap B_2$.

The sets $B \in \mathscr{B}$ are called *basis elements*. They are elements in $\mathscr{B}$ and subsets of X.

Example 2.2.3. Let $X = \mathbb{R}^2$ be the plane, and let $\mathscr{B}$ be the set of all open circular regions in the plane. This is the set of all ϵ-balls (or ϵ-disks)

$$B(x, \epsilon) = \{y \in \mathbb{R}^2 : \|y - x\| < \epsilon\}$$

with respect to the Euclidean metric $d(x, y) = \|y - x\|$ in the plane. We verify the two conditions for a basis:

(1): For each $x \in \mathbb{R}^2$ the 1-ball $B(x, 1) = \{y \in \mathbb{R}^2 : \|y - x\| < 1\}$ lies in $\mathscr{B}$ and $x \in B(x, 1)$.

(2): For $B_1 = B(x_1, \epsilon_1)$ and $B_2 = B(x_2, \epsilon_2)$ in $\mathscr{B}$, consider any $x \in B_1 \cap B_2$. Then $x \in \mathbb{R}^2$ and $\|x - x_1\| < \epsilon_1$ and $\|x - x_2\| < \epsilon_2$. Let

$$\epsilon = \min\{\epsilon_1 - \|x - x_1\|, \epsilon_2 - \|x - x_2\|\}\,.$$

Then $\epsilon > 0$, and we claim that $B_3 = B(x, \epsilon) \subset B_1 \cap B_2$. Since $x \in B_3$, this will finish the proof.

We prove that $B_3 \subset B_1$, by means of the triangle inequality. A similar proof shows that $B_3 \subset B_2$, so that $B_3 \subset B_1 \cap B_2$. Consider any $y \in B_3$, with

$$\|y - x\| < \epsilon \le \epsilon_1 - \|x - x_1\|\,.$$

Then

$$\|y - x_1\| \le \|y - x\| + \|x - x_1\| < \epsilon_1 - \|x - x_1\| + \|x - x_1\| = \epsilon_1$$

(using the triangle inequality for y, x and x_1). Hence $y \in B_1 = B(x_1, \epsilon_1)$. Since $y \in B_3$ was arbitrarily chosen, we have proved that $B_3 \subset B_1$, as required.

Example 2.2.4. Let $X = \mathbb{R}^2$ be the xy-plane, and let $\mathscr{B}'$ be the set of all open rectangular regions

$$(a, b) \times (c, d) \subset \mathbb{R} \times \mathbb{R}$$

with $a < b$ and $c < d$. This is the rectangle bounded by the vertical lines $x = a$ and $x = b$, and the horizontal lines $y = c$ and $y = d$. We verify the two conditions for a basis:

(1): For each $(x, y) \in \mathbb{R}^2$, the open rectangle $(x - 1, x + 1) \times (y - 1, y + 1)$ lies in $\mathscr{B}'$ and contains (x, y).

(2): Consider $B_1 = (a_1, b_1) \times (c_1, d_1)$, $B_2 = (a_2, b_2) \times (c_2, d_2)$ and $(x, y) \in B_1 \cap B_2$, so that $a_1 < x < b_1$, $c_1 < y < d_1$, $a_2 < x < b_2$ and $c_2 < y < d_2$. Let $a_3 = \max\{a_1, a_2\}$, $b_3 = \min\{b_1, b_2\}$, $c_3 = \max\{c_1, c_2\}$ and $d_3 = \min\{d_1, d_2\}$. Then $B_3 = (a_3, b_3) \times (c_3, d_3)$ satisfies

$$(x, y) \in B_3 = B_1 \cap B_2$$

so (a stronger form of) condition (2) is satisfied.

Example 2.2.5. Let X be a set, and let $\mathscr{B}''$ be the collection of singleton sets $\{x\}$ for $x \in X$. It is a basis for the discrete topology $\mathscr{T}_{\text{disc}}$ on X.

Definition 2.2.6. Let $\mathscr{B}$ be a basis for a topology on X. The topology $\mathscr{T}$ *generated* by $\mathscr{B}$ is the collection of subsets $U \subset X$ such that

for each $x \in U$ there exists a $B \in \mathscr{B}$ with $x \in B \subset U$.

In other words, a subset $U \subset X$ is defined to be open in this topology if for each $x \in U$ there exists a basis element $B \subset U$ with $x \in B$.

Lemma 2.2.7. *The collection $\mathscr{T}$ generated by a basis $\mathscr{B}$ is a topology on X.*

Proof. We check the three conditions for a topology.

(1): The subset $U = \varnothing$ is in $\mathscr{T}$, since there are no $x \in U$ for which a condition must be satisfied. The subset $U = X$ is in $\mathscr{T}$, since for each $x \in X$ there exists a $B \in \mathscr{B}$ with $x \in B$, by condition (1) for a basis, and then $B \subset U = X$.

(2): Let $\{U_\alpha\}_{\alpha \in J}$ be a subcollection of $\mathscr{T}$. Let $V = \bigcup_{\alpha \in J} U_\alpha$ be its union. We must show that $V \in \mathscr{T}$. Consider any $x \in V$. By the definition of the union, there exists an $\alpha \in J$ with $x \in U_\alpha$. Since $U_\alpha \in \mathscr{T}$, this means that there exists a basis element $B \in \mathscr{B}$ with $x \in B \subset U_\alpha$. Here $U_\alpha \subset V$, so $x \in B \subset V$. Since this holds for each $x \in V$, it follows that $V \in \mathscr{T}$.

(3): Let $\{U_i\}_{i=1}^n$ be a finite subcollection of $\mathscr{T}$. Let $W = U_1 \cap \cdots \cap U_n$ be its intersection. We must show that $W \in \mathscr{T}$. By induction on n, it suffices to do this in the case $n = 2$. Hence assume that $W = U_1 \cap U_2$. Consider any $x \in W$. Since $W \subset U_1$ we have $x \in U_1$, so there exists a basis element $B_1 \in \mathscr{B}$ with $x \in B_1 \subset U_1$. Furthermore, since $W \subset U_2$ we have $x \in U_2$, so there exists a basis element $B_2 \in \mathscr{B}$ with $x \in B_2 \subset U_2$. Hence $x \in B_1 \cap B_2 \subset U_1 \cap U_2 = W$. By consition (2) for a basis, these exists a basis element $B_3 \in \mathscr{B}$ with $x \in B_3 \subset B_1 \cap B_2$. It follows that $x \in B_3 \subset W$. Since x was arbitrarily chosen in W, we have verified that $W \in \mathscr{T}$. □

Proposition 2.2.8. *Let $\mathscr{B}$ be a basis for a topology $\mathscr{T}$ on X, i.e., let $\mathscr{T}$ be the topology generated by $\mathscr{B}$.*

(1) Each $B \in \mathscr{B}$ is open in X. Hence each union of basis elements is also open in X.

(2) Conversely, each open $U \subset X$ is a union of basis elements.

Proof. (1): Let $B \in \mathscr{B}$ be any basis element. For each $x \in B$ we obviously have $x \in B$ and $B \subset B$. Hence $B \in \mathscr{T}$ is open. It follows that any union of basis elements is a union of opens sets, hence is open.

(2): Let $U \in \mathscr{T}$ be open. For each $x \in U$ there exists a $B_x \in \mathscr{B}$ with $x \in B_x$ and $B_x \subset U$. Then

$$U = \bigcup_{x \in U} B_x$$

is the union of the collection of basis elements $\{B_x \mid x \in U\}$. To check the displayed equality, note that for each $x \in U$ we have $x \in B_x$, so $x \in \bigcup_{x \in U} B_x$. Hence $U \subset \bigcup_{x \in U} B_x$. On the other hand, each $B_x \subset U$, so $\bigcup_{x \in U} B_x \subset U$. □

Example 2.2.9. In any metric space (X, d), the collections of ϵ-balls

$$\mathscr{B} = \{B_d(x, \epsilon) \mid x \in X, \epsilon > 0\}$$

is a basis. The proof is the same as for $\mathbb{R}^2$ with the Euclidean metric, writing $d(x, y)$ in place of $\|y - x\|$. The topology generated by $\mathscr{B}$ is equal to the metric topology, $\mathscr{T}_d$. Hence a subset $U \subset X$ is open if and only if it is a union of ϵ-balls.

Example 2.2.10. Let $\mathbb{R}$ be the real line. With the usual metric $d(x, y) = |y - x|$, the ϵ-neighborhoods

$$B(x, \epsilon) = (x - \epsilon, x + \epsilon)$$

are open intervals, and each open interval

$$(a, b) = \{x \in \mathbb{R} \mid a < x < b\}$$

with $a < b$ has this form for $x = (a+b)/2$, $\epsilon = (b-a)/2$. Let $\mathscr{B}$ be the collection of all intervals $(a, b) \subset \mathbb{R}$ for $a < b$. It is a basis for the *standard topology* on $\mathbb{R}$, that is, the metric topology $\mathscr{T}_d$.

In this topology, open intervals are open subsets. Furthermore, the open subsets are precisely the unions of (arbitrary collections of) open intervals. By the general theory, finite intersections of open subsets are open, and arbitrary unions of opens subsets are open. To see that infinite intersections of open subsets need not be open, consider the example

$$[0, 1] = \bigcap_{n=1}^{\infty} (-1/n, 1 + 1/n) .$$

Here $(-1, 3), (-1/2, 3/2), (-1/3, 4/3), \dots$ are open, but their intersection is not.

Here is a 'recognition principle' for when a collection $\mathscr{C}$ is a basis for a given topology $\mathscr{T}$.

Lemma 2.2.11. *Let $(X, \mathscr{T})$ be a topological space. Suppose $\mathscr{C} \subset \mathscr{T}$ is a subcollection such that for each open $U \in \mathscr{T}$ and point $x \in U$ there exists an element $C \in \mathscr{C}$ with $x \in C$ and $C \subset U$. Then $\mathscr{C}$ is a basis for the topology $\mathscr{T}$.*

Proof. We first check that $\mathscr{C}$ is a basis.

(1): The set X is open in itself, so for each $x \in X$ there exists a $C \in \mathscr{C}$ with $x \in C$.

(2): Let $C_1, C_2 \in \mathscr{C}$. Since C_1 and C_2 are open, so is the intersection $C_1 \cap C_2$. Hence, for each $x \in C_1 \cap C_2$ there exists a $C_3 \in \mathscr{C}$ with $x \in C_3$ and $C_3 \subset C_1 \cap C_2$.

Next we check that the topology $\mathscr{T}'$ generated by $\mathscr{C}$ equals $\mathscr{T}$.

$\mathscr{T} \subset \mathscr{T}'$: If $U \in \mathscr{T}$ and $x \in U$ there exists a $C \in \mathscr{C}$ with $x \in C$ and $C \subset U$, by hypothesis, so $U \in \mathscr{T}'$ by definition.

$\mathscr{T}' \subset \mathscr{T}$: If $U \in \mathscr{T}'$ then U is a union of elements of $\mathscr{C}$ by the proposition above. Each element of $\mathscr{C}$ is in $\mathscr{T}$, hence so is the union U. □

2.2.2 Comparing topologies using bases

Lemma 2.2.12. *Let $\mathscr{B}$ and $\mathscr{B}'$ be bases for the topologies $\mathscr{T}$ and $\mathscr{T}'$ on X, respectively. The following are equivalent:*

(1) $\mathscr{T}'$ is finer than $\mathscr{T}$, so $\mathscr{T} \subset \mathscr{T}'$.

(2) For each basis element $B \in \mathscr{B}$ and each point $x \in B$ there is a basis element $B' \in \mathscr{B}'$ with $x \in B'$ and $B' \subset B$.

Proof. (1) $\implies$ (2): Let $B \in \mathscr{B}$ and $x \in B$. Since $B \in \mathscr{T}$ and $\mathscr{T} \subset \mathscr{T}'$ we have $B \in \mathscr{T}'$. Since $\mathscr{T}'$ is the topology generated by $\mathscr{B}'$ there exists a $B' \in \mathscr{B}'$ with $x \in B'$ and $B' \subset B$.

(2) $\implies$ (1): Let $U \in \mathscr{T}$. For each $x \in U$ there exists a $B \in \mathscr{B}$ with $x \in B$ and $B \subset U$, since $\mathscr{B}$ generates $\mathscr{T}$. By hypothesis there exists a $B' \in \mathscr{B}'$ with $x \in B'$ and $B' \subset B$. Hence $B' \subset U$. Since this holds for each $x \in U$, it follows that $U \in \mathscr{T}'$. □

Note that in order to have $\mathscr{T} \subset \mathscr{T}'$ it is not necessary to have $\mathscr{B} \subset \mathscr{B}'$ (each basis element $B \in \mathscr{B}$ does not need to be a basis element in $\mathscr{B}'$), but for each $x \in B$ there should be some potentially smaller basis element $B' \in \mathscr{B}'$ with $x \in B' \subset B$.

Corollary 2.2.13. *Two bases $\mathscr{B}$ and $\mathscr{B}'$ for topologies on X generate the same topology if and only if (1) for each $x \in B \in \mathscr{B}$ there is a basis element $B' \in \mathscr{B}'$ with $x \in B' \subset B$, and furthermore, (2) for each $x \in B' \in \mathscr{B}'$ there is a basis element $B \in \mathscr{B}$ with $x \in B \subset B'$.*

Example 2.2.14. The basis $\mathscr{B}$ of open circular regions in the plane and the basis $\mathscr{B}'$ of open rectangular regions generate the same topology on $\mathbb{R}^2$, namely the metric topology.

2.2.3 Subbases

Starting with any collection $\mathscr{S}$ of subsets of a set X, we can form a basis $\mathscr{B}$ for a topology by taking all finite intersections

$$B = S_1 \cap \cdots \cap S_n$$

of elements in $\mathscr{S}$. The open sets in the topology $\mathscr{T}$ generated by $\mathscr{B}$ are then all unions of such basis elements B, which are all unions of all finite intersections of sets in $\mathscr{S}$. Such a collection $\mathscr{S}$ is called a subbasis for the topology $\mathscr{T}$. To avoid ambiguity about the intersection of an empty collection, we mildly restrict the collection $\mathscr{S}$ as follows:

Definition 2.2.15. A *subbasis* for a topology on X is a collection $\mathscr{S}$ of subsets of X, with union equal to X. The *basis* associated to $\mathscr{S}$ is the collection $\mathscr{B}$ consisting of all finite intersections

$$B = S_1 \cap \cdots \cap S_n$$

of elements $S_1, \dots, S_n \in \mathscr{S}$, for $n \geq 1$. By the topology $\mathscr{T}$ generated by the subbasis $\mathscr{S}$ we mean the topology generated by the associated basis $\mathscr{B}$.

Clearly $\mathscr{S} \subset \mathscr{B} \subset \mathscr{T}$.

Lemma 2.2.16. *Let $\mathscr{S}$ be a subbasis on X. The associated collection $\mathscr{B}$ is a basis for a topology.*

Proof. (1): Each $x \in X$ lies in some $S \in \mathscr{S}$, hence is an element of the basis element $B = S \in \mathscr{B}$.

(2): Suppose that $B_1 = S_1 \cap \cdots \cap S_n$ and $B_2 = S_{n+1} \cap \cdots \cap S_{n+m}$ are basis elements, and $x \in B_1 \cap B_2$. Let $B_3 = B_1 \cap B_2 = S_1 \cap \cdots \cap S_{m+n}$. Then B_3 is a basis element, and $x \in B_3 = B_1 \cap B_2$. □

2.3 (§15) The Product Topology on $X \times Y$

2.3.1 A basis for the product topology

Recall (from §5) that for two sets A and B, the cartesian product $A \times B$ is the set of ordered pairs (x, y), with $x \in A$ and $y \in B$.

Definition 2.3.1. Let X and Y be topological spaces. The *product topology* on $X \times Y$ is the topology generated by the basis

$$\mathscr{B} = \{U \times V \mid U \text{ open in } X \text{ and } V \text{ open in } Y\}$$

consisting of all sets $U \times V \subset X \times Y$, where U ranges over all open subsets of X and V ranges over all open subsets of Y.

Lemma 2.3.2. *The collection $\mathscr{B}$ (as above) is a basis for a topology on $X \times Y$.*

Proof. (1): $X \times Y$ is itself a basis element.

(2): Let $B_1 = U_1 \times V_1$ and $B_2 = U_2 \times V_2$ be two basis elements. In view of the identity

$$(U_1 \times V_1) \cap (U_2 \times V_2) = (U_1 \cap U_2) \times (V_1 \cap V_2)$$

we have that $B_1 \cap B_2 = B_3$, where $B_3 = U_3 \times V_3$ is the basis element given by the product of the two open sets $U_3 = U_1 \cap U_2$ and $V_3 = V_1 \cap V_2$. □

The union of two basis elements

$$(U_1 \times V_1) \cup (U_2 \times V_2)$$

is usually not a basis element. The open sets in the product topology on $X \times Y$ are the unions

$$\bigcup_{\alpha \in J} (U_\alpha \times V_\alpha)$$

of arbitrary collections $\{B_\alpha = U_\alpha \times V_\alpha\}_{\alpha \in J}$ of basis elements.

Theorem 2.3.3. *Let X have the topology generated by a basis $\mathscr{B}$ and let Y have the topology generated by a basis $\mathscr{C}$. Then the collection*

$$\mathscr{D} = \{B \times C \mid B \in \mathscr{B} \text{ and } C \in \mathscr{C}\}$$

is a basis for the product topology on $X \times Y$.

Proof. We apply Lemma 2.2.11. The elements $B \times C$ of the collection $\mathscr{D}$ are open in the product topology, since each $B \in \mathscr{B}$ is open in X and each $C \in \mathscr{C}$ is open in Y, so $B \times C$ is one of the basis elements for the product topology.

Let $(x, y) \in W \subset X \times Y$ where W is open in the product topology. By definition of the topology generated by a basis, there exists a basis element $U \times V$ for the product topology, such that $(x, y) \in U \times V \subset W$. Since $x \in U$, U is open in X and $\mathscr{B}$ is a basis for the topology on X, there exists a basis element $B \in \mathscr{B}$ such that $x \in B \subset U$. Similarly, there exists a basis element $C \in \mathscr{C}$ such that $y \in C \subset V$. Then $B \times C$ is in the collection $\mathscr{D}$, and $(x, y) \in B \times C \subset U \times V \subset W$. □

Example 2.3.4. The collection $\mathscr{B}'$ of open rectangular regions

$$(a, b) \times (c, d)$$

for $a < b$ and $c < d$ is a basis for the product topology on $\mathbb{R} \times \mathbb{R}$, since the collection of open intervals (a, b) for $a < b$ is a basis for the standard topology on $\mathbb{R}$. As previously noted, this product topology is the same as the metric topology.

2.3.2 A subbasis for the product topology

Definition 2.3.5. Let $\pi_1 \colon X \times Y \to X$ denote the (first) projection $\pi_1(x, y) = x$, and let $\pi_2 \colon X \times Y \to Y$ denote the (second) projection $\pi_2(x, y) = y$, for $x \in X$ and $y \in Y$.

Lemma 2.3.6. *The preimage of $U \subset X$ under $\pi_1 \colon X \times Y \to X$ equals*

$$\pi_1^{-1}(U) = U \times Y\,.$$

Similarly, the preimage of $V \subset Y$ under $\pi_2 \colon X \times Y \to Y$ equals

$$\pi_2^{-1}(V) = X \times V\,.$$

Proof. An element (x, y) lies in $\pi_1^{-1}(U)$ if and only if $x = \pi_1(x, y)$ lies in U, which for $y \in Y$ is equivalent to asking that (x, y) lies in $U \times Y$. The second case is similar. □

Note the identity

$$(U \times Y) \cap (X \times V) = U \times V$$

of subsets of $X \times Y$. Hence each basis element $B = U \times V$ for the product topology on $X \times Y$ is the intersection of two subsets of the form $S_1 = \pi_1^{-1}(U) = U \times Y$ and $S_2 = \pi_2^{-1}(V) = X \times V$. It follows that the basis for the product topology is generated by a smaller subbasis:

Definition 2.3.7. Let

$$\begin{aligned}\mathscr{S} &= \{U \times Y \mid U \subset X \text{ open}\} \cup \{X \times V \mid V \subset Y \text{ open}\} \\ &= \{\pi_1^{-1}(U) \mid U \subset X \text{ open}\} \cup \{\pi_2^{-1}(V) \mid V \subset Y \text{ open}\}\,.\end{aligned}$$

Lemma 2.3.8. *The collection $\mathscr{S}$ (as above) is a subbasis for the product topology on $X \times Y$.*

Proof. The finite intersections of elements in the subbasis are all of the form

$$\pi_1^{-1}(U) \cap \pi_2^{-1}(V) = U \times V$$

for U open in X and V open in Y, hence the subbasis generates the usual basis for the product topology. □

2.4 (§16) The Subspace Topology

2.4.1 Subspaces

Definition 2.4.1. Let $(X, \mathscr{T})$ be a topological space, and let $A \subset X$ be a subset. The collection

$$\mathscr{T}_A = \{A \cap U \mid U \in \mathscr{T}\}$$

of subsets of A is called the *subspace topology* on A. With this topology, $(A, \mathscr{T}_A)$ is called a *subspace* of X.

Lemma 2.4.2. *The collection $\mathscr{T}_A$ (as above) is a topology on A.*

Proof. (1): $\varnothing = A \cap \varnothing$ and $A = A \cap X$ are in $\mathscr{T}_A$, since $\varnothing$ and X are in $\mathscr{T}$.

(2): Each subcollection of $\mathscr{T}_A$ can be indexed as $\{A \cap U_\alpha\}_{\alpha \in J}$ for some subcollection $\{U_\alpha\}_{\alpha \in J}$ of $\mathscr{T}$. Then

$$\bigcup_{\alpha \in J} (A \cap U_\alpha) = A \cap \bigcup_{\alpha \in J} U_\alpha$$

by the distributive law, and $\bigcup_{\alpha \in J} U_\alpha$ is in $\mathscr{T}$, hence this union is in $\mathscr{T}_A$.

(3): Each finite subcollection of $\mathscr{T}_A$ can be indexed as $\{A \cap U_1, \dots, A \cap U_n\}$ for some finite subcollection $\{U_1, \dots, U_n\}$ of $\mathscr{T}$. Then

$$(A \cap U_1) \cap \cdots \cap (A \cap U_n) = A \cap (U_1 \cap \cdots \cap U_n)$$

and $U_1 \cap \cdots \cap U_n$ is in $\mathscr{T}$, hence this intersection is in $\mathscr{T}_A$. □

When $(A, \mathscr{T}_A)$ is a subspace of $(X, \mathscr{T})$, and $V \subset A \subset X$, there are two possible meanings of the assertion "V is open", namely $V \in \mathscr{T}$ or $V \in \mathscr{T}_A$. In general, these two meanings are different.

Definition 2.4.3. We say that "V is open in X", or that "V is an open subset of X", to indicate that $V \in \mathscr{T}$, while we say that "V is open in A", or that "V is an open subset of A", to indicate that $V \in \mathscr{T}_A$. The latter means that $V = A \cap U$ for some U that is open in X.

Lemma 2.4.4. *If $\mathscr{B}$ is a basis for a topology $\mathscr{T}$ on X, and $A \subset X$, then the collection*

$$\mathscr{B}_A = \{A \cap B \mid B \in \mathscr{B}\}$$

is a basis for the subspace topology $\mathscr{T}_A$ on A.

Proof. We apply Lemma 2.2.11 for the topological space $(A, \mathscr{T}_A)$ and the collection $\mathscr{B}_A$. Each subset $A \cap B$ in $\mathscr{B}_A$ is open in A, since each basis element $B \in \mathscr{B}$ is open in X. Furthermore, each open subset of A has the form $A \cap U$ for some open subset U of X. If $x \in A \cap U$ is any point, then $x \in U$, so since $\mathscr{B}$ is a basis for the topology $\mathscr{T}$ there exists a $B \in \mathscr{B}$ with $x \in B$ and $B \subset U$. Then $A \cap B \in \mathscr{B}_A$, $x \in A \cap B$, and $A \cap B \subset A \cap U$. By the cited lemma, $\mathscr{B}_A$ is a basis for the topology $\mathscr{T}_A$. □

Example 2.4.5. Give $X = \mathbb{R}$ the standard topology generated by the open intervals (a, b), and let $A = [0, 1)$. The subspace topology on A has a basis consisting of the intersections $[0, 1) \cap (a, b)$, i.e., the subsets $[0, b)$ and (a, b) for $0 < a < b \le 1$. For instance, $[0, 1/2) = [0, 1) \cap (-1/2, 1/2)$ and $(0, 1/2) = [0, 1) \cap (0, 1/2)$ are both open subsets of $[0, 1)$ in the subspace topology.

Example 2.4.6. Let (X, d) be a metric space, with basis $\mathscr{B} = \{B_d(x, \epsilon) \mid x \in X, \epsilon > 0\}$ for the metric topology $\mathscr{T} = \mathscr{T}_d$. Here

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

Let $A \subset X$ be any subset, with metric $d_A = d|A \times A$ given by $d_A(x, y) = d(x, y)$ for all $x, y \in A$. The metric space (A, d_A) has basis $\mathscr{B}' = \{B_{d_A}(x, \epsilon) \mid x \in A, \epsilon > 0\}$ for the metric topology $\mathscr{T}' = \mathscr{T}_{d_A}$, where

$$B_{d_A}(x, \epsilon) = \{y \in A \mid d_A(x, y) < \epsilon\} .$$

Note that

$$B_{d_A}(x, \epsilon) = A \cap B_d(x, \epsilon)$$

for all $x \in A$. Hence $\mathscr{B}' \subset \mathscr{B}_A$ and $\mathscr{T}' \subset \mathscr{T}_A$, where $\mathscr{T}_A$ is the subspace topology on A.

To prove that the two topologies are equal, so that $\mathscr{T}_A = \mathscr{T}'$, we use Lemma 2.2.12 to check that $\mathscr{T}_A \subset \mathscr{T}'$. Thus consider any basis element $A \cap B_d(x, \epsilon)$ in $\mathscr{T}_A$ and any element $y \in A \cap B_d(x, \epsilon)$. Then $\delta = \epsilon - d(x, y)$ is positive, and $B_{d_A}(y, \delta)$ is a basis element in $\mathscr{B}'$, $y \in B_{d_A}(y, \delta)$ and $B_{d_A}(y, \delta) \subset A \cap B_d(x, \epsilon)$.

Definition 2.4.7. By an *open subspace* of X we mean an open subset $A \subset X$ with the subspace topology.

Lemma 2.4.8. *Let A be an open subspace of X. Then a subset $V \subset A$ is open in A if and only if it is open in X.*

Proof. Suppose first that V is open in A, in the subspace topology. Then $V = A \cap U$ for some U that is open in X. Since A is open in X, it follows that the intersection, $V = A \cap U$ is open in X.

Conversely, suppose that $V \subset A$ is open in X. Then $V = A \cap V$, so V is also open in A. □

2.4.2 Products vs. subspaces

Lemma 2.4.9. *Let X and Y be topological spaces, with subspaces A and B, respectively. Then the product topology on $A \times B$ is the same as the subspace topology on $A \times B$ as a subset of $X \times Y$.*

Proof. The subspace topology on A is generated by the basis with elements $A \cap U$, where U ranges over all open subsets of X. Likewise, the subspace topology on B is generated by the intersections $B \cap V$, where V ranges over all open subsets of Y. Hence the product topology on $A \times B$ is generated by the basis with elements

$$(A \cap U) \times (B \cap V)$$

where U and V range over the open subsets of X and Y, respectively.

On the other hand, the collection of products $U \times V$ is a basis for the product topology on $X \times Y$, so the collection of intersections

$$(A \times B) \cap (U \times V)$$

is a basis for the subspace topology on $A \times B$, where U and V still range over the open subsets of X and Y. In view of the identity

$$(A \cap U) \times (B \cap V) = (A \times B) \cap (U \times V)$$

these two bases are in fact equal, hence they generate the same topology. □

Example 2.4.10. With $A = S^1 \subset \mathbb{R}^2$ and $B = [0, 2\pi] \subset \mathbb{R}$, the product space

$$A \times B = S^1 \times [0, 2\pi] \subset \mathbb{R}^2 \times \mathbb{R} \cong \mathbb{R}^3$$

is a cylinder. If we view $\mathbb{R}$ as $\mathbb{R} \times \{0\} \subset \mathbb{R}^2$, there is room to bend the interval $B \cong [0, 2\pi] \times \{0\} \subset \mathbb{R}^2$ to bring its ends closer together, eventually forming the loop $B' = S^1 \subset \mathbb{R}^2$. The product space

$$A \times B' = S^1 \times S^1 \subset \mathbb{R}^2 \times \mathbb{R}^2 \cong \mathbb{R}^4$$

is obtained from the cylinder $S^1 \times [0, 2\pi]$ by bringing the two end circles together. This can be realized within $\mathbb{R}^3$ by the torus surface. Hence $S^1 \times S^1$ is topologically "the same" as the torus surface.

2.5 (§17) Closed Sets and Limit Points

2.5.1 Closed subsets

Definition 2.5.1. A subset K of a topological space X is said to be *closed* if (and only if) the complement $X - K$ is open. In other words, the closed subsets of X are the subsets of the form $X - U$ where U is open.

Example 2.5.2. The interval $[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$ is closed in $\mathbb{R}$ (with the standard topology), since the complement $\mathbb{R} - [a, b] = (-\infty, a) \cup (b, \infty)$ is open.

Example 2.5.3. In the discrete topology $\mathscr{T}_{\text{disc}}$ on a set X, every subset is closed. In the trivial topology $\mathscr{T}_{\text{triv}}$, only the subsets $\varnothing$ and X are closed.

Example 2.5.4. In the cofinite topology $\mathscr{T}_{\text{cof}}$ on a set X, the closed subsets are the finite subsets $F \subset X$, together with X itself.

Theorem 2.5.5. *Let X be a topological space.*

(1) $\varnothing$ and X are closed (as subsets of X).

(2) The intersection of any collection of closed subsets of X is closed.

(3) The union of any finite collection of closed subsets of X is closed.

Proof. (1): $\varnothing = X - X$ and $X = X - \varnothing$ are closed.

(2): If $\{K_\alpha\}_{\alpha \in J}$ is any collection of closed subsets of X, then the complements $U_\alpha = X - K_\alpha$ are all open, so that $\{U_\alpha\}_{\alpha\in J}$ is a collection of open subsets of X. To prove that the intersection $\bigcap_{\alpha\in J} K_\alpha$ is closed, we must check that its complement is open. By De Morgan's law

$$X - \bigcap_{\alpha\in J} K_\alpha = \bigcup_{\alpha\in J}(X - K_\alpha) = \bigcup_{\alpha\in J} U_\alpha$$

is a union of open sets, hence is open, as desired.

(3): If $\{K_1, \dots, K_n\}$ is a finite collection of closed subsets of X, then the complements $U_i = X - K_i$ are all open, so that $\{U_1, \dots, U_n\}$ is a finite collection of open subsets of X. To prove that the union $K_1 \cup \dots \cup K_n$ is closed, we must check that its complement is open. By De Morgan's law

$$X - (K_1 \cup \dots \cup K_n) = (X - K_1) \cap \dots \cap (X - K_n) = U_1 \cap \dots \cap U_n$$

is a finite intersection of open sets, hence is open, as desired. □

Clearly the collection $\mathscr{C} = \{X - U \mid U \in \mathscr{T}\}$ of closed subsets of a topological space $(X, \mathscr{T})$ uniquely determine the topology $\mathscr{T}$, and any collection $\mathscr{C}$ of subsets, called closed subsets, satisfying the three conditions of the theorem above, will determine a topology $\mathscr{T} = \{X - K \mid K \in \mathscr{C}\}$ in this way.

When $(A, \mathscr{T}_A)$ is a subspace of $(X, \mathscr{T})$, and $K \subset A \subset X$, there are two possible meanings of the assertion "K is closed", namely $X - K \in \mathscr{T}$ or $A - K \in \mathscr{T}_A$. In general, these two meanings are different.

Definition 2.5.6. We say that "K is closed in A", or that "K is a closed subset of A", if K is a subset of A and K is closed in the subspace topology on A, so that $A - K$ is open in the subspace topology on A.

Theorem 2.5.7. *Let A be a subspace of X. A subset $K \subset A$ is closed in A if and only if there exists a closed subset $L \subset X$ with $K = A \cap L$.*

Proof. If K is closed in A, then $V = A - K$ is open in A, so there exists an open $U \subset X$ with $V = A \cap U$. Then $L = X - U$ is closed in X, and

$$A \cap L = A \cap (X - U) = A - (A \cap U) = A - V = K\,.$$

Conversely, if L is closed in X and $K = A \cap L$, then $U = X - L$ is open in X so $V = A \cap U$ is open in A. Now

$$A - K = A - A \cap L = A \cap (X - L) = A \cap U = V\,,$$

so K is closed in A. □

Definition 2.5.8. By a *closed subspace* of X we mean a closed subset $A \subset X$ with the subspace topology.

Lemma 2.5.9. *Let A be a closed subspace of X. Then a subset $K \subset A$ is closed in A if and only if it is closed in X.*

Proof. Suppose first that K is closed in A, in the subspace topology. Then $K = A \cap L$ for some L that is closed in X. Since A is closed in X, it follows that the intersection, $K = A \cap L$ is closed in X. Conversely, suppose that $K \subset A$ is closed in X. Then $K = A \cap K$, so K is also closed in A. □

2.5.2 Closure and interior

Definition 2.5.10. Let X be a topological space and $A \subset X$ a subset. The *closure* $\mathrm{Cl}\, A = \bar{A}$ of A is the intersection of all the closed subsets of X that contain A. The *interior* $\mathrm{Int}\, A$ of A is the union of all the open subsets of X that are contained in A.

Since unions of open sets are open, and intersections of closed sets are closed, the following lemmas are clear.

Lemma 2.5.11. *(1) The closure $\mathrm{Cl}\, A$ is a closed subset of X.*

(2) $A \subset \mathrm{Cl}\, A$.

(3) If $A \subset K \subset X$ with K closed, then $\mathrm{Cl}\, A \subset K$.

Lemma 2.5.12. *(1) The interior $\mathrm{Int}\, A$ is an open subset of X.*

(2) $\mathrm{Int}\, A \subset A$.

(3) If $U \subset A \subset X$ with U open, then $U \subset \mathrm{Int}\, A$.

Example 2.5.13. Let $X = \mathbb{R}$ and $A = [a, b)$ with $a < b$. The closure of A is the closed interval $[a, b]$, and the interior of A is the open interval (a, b). The closure cannot be smaller, since $[a, b)$ is not closed, and the interior cannot be larger, since $[a, b)$ is not open.

Example 2.5.14. If X has the discrete topology, $\mathrm{Int}\, A = A = \mathrm{Cl}\, A$ for each $A \subset X$, since each A is both open and closed.

If X has the indiscrete topology, and $A \subset X$ is a proper, non-empty subset, then $\mathrm{Int}\, A = \varnothing$ and $\mathrm{Cl}\, A = X$.

If $X = \{a, b\}$ has the Sierpinski topology $\mathscr{T}_a = \{\varnothing, \{a\}, X\}$, where $\{a\}$ is open and $\{b\}$ is closed, then $\mathrm{Cl}\{a\} = X$ but $\mathrm{Cl}\{b\} = \{b\}$.

Lemma 2.5.15. *The complement of the closure is the interior of the complement, and the complement of the interior is the closure of the complement:*

$$X - \mathrm{Cl}\, A = \mathrm{Int}(X - A)$$
$$X - \mathrm{Int}\, A = \mathrm{Cl}(X - A)$$

Proof. Since $\mathrm{Cl}\, A = \bigcap\{K \mid A \subset K,\ K \text{ closed in } X\}$,

$$\begin{aligned} X - \mathrm{Cl}\, A &= \bigcup\{X - K \mid A \subset K,\ K \text{ closed in } X\} \\ &= \bigcup\{U \mid U \subset X - A,\ U \text{ open in } X\} = \mathrm{Int}(X - A)\,, \end{aligned}$$

since $A \subset K$ is equivalent to $X - K \subset X - A$, where we substitute U for $X - K$.

The other assertion follows by considering $X - A$ in place of A. □

Example 2.5.16. Let $X = \mathbb{R}$ and $A = \mathbb{Q}$, the set of rational numbers. The closure of A equals $\mathbb{R}$, while the interior of A is empty. For every ϵ-ball $(x - \epsilon, x + \epsilon)$ in $\mathbb{R}$ contains both rational and irrational numbers, so the interiors of A and $X - A$ are both empty. We say that $\mathbb{Q}$ is dense in $\mathbb{R}$.

Definition 2.5.17. A subset $A \subset X$ is *dense* if $\operatorname{Cl} A = X$.

2.5.3 Closure in subspaces

To emphasize the role of the ambient space, we might write $\operatorname{Cl}_X(A)$ for the closure of A in X.

Theorem 2.5.18. *Let X be a topological space, $Y \subset X$ a subspace, and $A \subset Y$ a subset. Let $\bar{A}$ denote the closure of A in X. Then the closure of A in Y equals $Y \cap \bar{A}$.*

$$\operatorname{Cl}_Y(A) = Y \cap \operatorname{Cl}_X(A)$$

Proof. Let B denote the closure of A in Y.

To see that $B \subset Y \cap \bar{A}$, note that $\bar{A}$ is closed in X, so $Y \cap \bar{A}$ is closed in Y and contains A. Hence it contains the closure B of A in Y.

To prove the opposite inclusion, note that B is closed in Y, hence has the form $B = Y \cap C$ for some C that is closed in X. Then $A \subset B \subset C$, so C is closed in X and contains A. Hence $\bar{A} \subset C$ and $Y \cap \bar{A} \subset Y \cap C = B$. □

Example 2.5.19. Let $X = \mathbb{R}$, $Y = \mathbb{Q}$ and $A = \mathbb{Q} \cap [0, \pi)$. The closure of A in X is $[0, \pi]$, and the closure of A in Y is $\mathbb{Q} \cap [0, \pi] = \mathbb{Q} \cap [0, \pi)$ (since π is not rational). Hence A is closed in Y.

2.5.4 Neighborhoods

Definition 2.5.20. Let X be a topological space, $U \subset X$ a subset and $x \in X$ a point. We say that *U is a neighborhood of x* (norsk: "*U er en omegn om x*") if $x \in U$ and U is open in X.

One way to formalize that we consider "all points y sufficiently close to x", for a given point $x \in X$, is to consider " all points $y \in U$ for some neighborhood U of x".

We say that a set A *meets*, or *intersects*, a set B if $A \cap B$ is not empty. Here is a criterion for detecting which points $x \in X$ lie in the closure of A:

Theorem 2.5.21. *Let A be a subset of a topological space X. A point $x \in X$ lies in the closure $\bar{A}$ if and only if A meets every open set U in X that contains x. Equivalently, $x \in \bar{A}$ if and only if A meets U for each neighborhood U of x.*

Proof. Consider the complement $X - A$ and its interior. We have $x \in \operatorname{Int}(X - A)$ if and only if there exists an open U with $x \in U$ such that $U \subset X - A$. The negation of $x \in \operatorname{Int}(X - A)$ is $x \in X - \operatorname{Int}(X - A) = \bar{A}$. The negation of $U \subset X - A$ is $A \cap U \neq \varnothing$. The negation of "there exists an open U with $x \in U$ such that $U \subset X - A$" is therefore "for each open U with $x \in U$ we have $A \cap U \neq \varnothing$". Hence x is in the closure of A if and only if A meets each neighborhood U of x. □

It suffices to check this for neighborhoods in a basis:

Theorem 2.5.22. *Let $\mathscr{B}$ be a basis for a topology on X, and let $A \subset X$. A point $x \in X$ lies in $\bar{A}$ if and only if A meets each basis element $B \in \mathscr{B}$ with $x \in B$.*

Proof. If there exists an open U with $x \in U$ such that $U \subset X - A$ then there exists a basis element $B \in \mathscr{B}$ with $x \in B$ such that $B \subset X - A$, and conversely. Hence we may replace "an open U" by "a basis element B" in the previous proof. □

Example 2.5.23. Let $A = \{1/n \mid n \in \mathbb{N}\} \subset \mathbb{R}$. Then $0 \in \bar{A}$, since each basis element (a, b) for the standard topology on $\mathbb{R}$ with $0 \in (a, b)$ contains $(-\epsilon, \epsilon)$ for some $\epsilon > 0$, hence also contains $1/n \in A$ for each $n > 1/\epsilon$. The closure of A is $\bar{A} = \{0\} \cup A$. This is a closed subset of $\mathbb{R}$, since the complement is the union of the open sets $(-\infty, 0)$, $(1/(n + 1), 1/n)$ for $n \in \mathbb{N}$, and $(1, \infty)$.

2.5.5 Limit points

Definition 2.5.24. Let A be a subset of a topological space X. A point $x \in X$ is a *limit point* of A if each neighborhood U of x contains a point of A other than x, or equivalently, of x belongs to the closure of $A - \{x\}$. (If $x \notin A$, recall that $A - \{x\} = A$.)

The set of limit points of A is often denoted A', and is called the *derived set* of A in X.

Example 2.5.25. Let $A = \{1/n \mid n \in \mathbb{N}\} \subset \mathbb{R}$. Then $0 \in \mathbb{R}$ is a limit point of A. In this case there are no other limit points of A.

Theorem 2.5.26. *Let A be a subset of a topological space X, with closure $\bar{A}$ and set of limit points A'. Then*

$$\bar{A} = A \cup A' .$$

Proof. $A \cup A' \subset \bar{A}$: Clearly $A \subset \bar{A}$. If $x \in A'$ then every neighborhood U of x meets $A - \{x\}$, hence it also meets A, so $x \in \bar{A}$.

$\bar{A} \subset A \cup A'$: Let $x \in \bar{A}$. If $x \in A$, then $x \in A \cup A'$. Otherwise, $x \notin A$, so $A - \{x\} = A$. Since $x \in \bar{A}$, every neighborhood U of x meets $A - \{x\} = A$, so $x \in A'$ is a limit point of A. □

Corollary 2.5.27. *A subset A of a topological space is closed if and only if it contains all its limits points.*

Proof. We have $A = \bar{A}$ if and only if $A = A \cup A'$, which holds if and only if $A \supset A'$. □

2.5.6 Convergence to a limit

Definition 2.5.28. Let $(x_1, x_2, \dots) = (x_n)_{n=1}^{\infty}$ be a sequence of points in a topological space X, so $x_n \in X$ for each $n \in \mathbb{N}$. We say that $(x_n)_{n=1}^{\infty}$ *converges* to a point $y \in X$ if for each neighborhood U of y there is an $N \in \mathbb{N}$ such that $x_n \in U$ for all $n \geq N$. In this case we call y *a limit* of the sequence $(x_n)_{n=1}^{\infty}$, and may write $x_n \to y$ as $n \to \infty$.

A sequence $(x_n)_{n=1}^{\infty}$ in X can also be viewed as a function $f \colon \mathbb{N} \to X$, with $f(n) = x_n$ for each $n \in \mathbb{N}$.

Example 2.5.29. Consider the Sierpinski space $X = \{a, b\}$ with the topology $\mathscr{T}_a = \{\varnothing, \{a\}, X\}$. The constant sequence $(x_n)_{n=1}^{\infty}$ with $x_n = a$ for all $n \in \mathbb{N}$ converges to a, since the only neighborhoods of a are $\{a\}$ and $\{a, b\}$, both of which contain x_n for all n. Hence a is a limit of $(a, a, \dots)$.

However, the same sequence also converges to b, since the only neighborhood of b is $\{a, b\}$, which also contains x_n for all n. Hence b is also a limit for $(a, a, \dots)$.

On the other hand, the constant sequence $(y_n)_{n=1}^{\infty}$ with $y_n = b$ for all $n \in \mathbb{N}$ converges to b, since the only neighborhood $\{a, b\}$ of b contains y_n for all n.

This constant sequence does not converge to a, since the neighborhood $\{a\}$ of a does not contain y_n for any n, hence there is no $N \in \mathbb{N}$ such that $y_n \in \{a\}$ for all $n \geq N$.

2.5.7 Hausdorff spaces

To obtain unique limits for convergent sequences, and be able to talk about *the limit* of a sequence, we must assume that the topology is sufficiently fine to separate the individual points. Such additional hypotheses are called *separation axioms* (German: Trennungsaxiome). The most common separation axiom is known as the Hausdorff property.

Definition 2.5.30. A topological space X is called a *Hausdorff space* if for each pair of points $x, y \in X$, with $x \neq y$, there exist open sets $U, V \subset X$ with $x \in U$, $y \in V$ and $U \cap V = \varnothing$. In other words, there exist neighborhood U and V of x and y, respectively, that are disjoint.

Example 2.5.31. The set $X = \{a, b\}$ with the discrete topology is a Hausdorff space, since the only pair of distinct points is a and b, and the open subsets $\{a\}$ and $\{b\}$ are neighborhoods of a and b, respectively, with empty intersection.

The set $X = \{a, b\}$ with the Sierpinski topology $\mathscr{T}_a = \{\varnothing, \{a\}, X\}$ is not a Hausdorff space, since the only neighborhood of b is $V = X$, and no neighborhood U of a can be disjoint from X.

Lemma 2.5.32. *Each metric space (X, d) is Hausdorff.*

Proof. Let $x, y \in X$ be two distinct points. Then $\delta = d(x, y) > 0$. Consider the neighborhoods $U = B_d(x, \delta/2)$ and $V = B_d(y, \delta/2)$ of x and y, respectively. Then $U \cap V = \varnothing$ by the triangle inequality. □

Remark 2.5.33. If $(X, \mathscr{T})$ is Hausdorff, clearly $(X, \mathscr{T}')$ is also Hausdorff if $\mathscr{T}'$ is a finer topology than $\mathscr{T}$. In rough terms, the Hausdorff property asserts that there are "enough" open sets, locally in X.

2.5.8 Uniqueness of limits in Hausdorff spaces

Theorem 2.5.34. *If X is a Hausdorff space, then a sequence $(x_n)_{n=1}^\infty$ of points in X converges to at most one point in X.*

Proof. Suppose that $(x_n)_{n=1}^\infty$ converges to y and z. We must prove that $y = z$.

Suppose, to achieve a contradiction, that $y \neq z$. Then there exist neighborhoods U of y and V of z with $U \cap V = \varnothing$. Since $(x_n)_{n=1}^\infty$ converges to y there exists an $N \in \mathbb{N}$ such that $x_n \in U$ for all $n \geq N$. Since $(x_n)_{n=1}^\infty$ converges to z there exists an $M \in \mathbb{N}$ such that $x_n \in V$ for all $n \geq M$. Hence, for $n \geq \max\{N, M\}$ we have $x_n \in U \cap V$, which is impossible, since this intersection is empty. □

Definition 2.5.35. If X is a Hausdorff space, and a sequence $(x_n)_{n=1}^\infty$ of points in X converges to a point $y \in X$, we say that y is *the limit* of $(x_n)_{n=1}^\infty$, and write

$$y = \lim_{n\to\infty} x_n \,.$$

2.5.9 Closed sets and limits points in Hausdorff spaces

Theorem 2.5.36. *Each finite subset $A \subset X$ in a Hausdorff space is closed.*

Proof. The set A is the union of a finite collection of singleton sets $\{x\}$, so it suffices to prove that each singleton set $\{x\}$ is closed in X.

Consider any other point $y \in X$, with $x \neq y$. By the Hausdorff property there are open subsets $U, V \subset X$ with $x \in U$ and $y \in V$, such that $U \cap V = \varnothing$. Then $x \notin V$, so $X - V$ is a closed set that contains $\{x\}$. Hence $\mathrm{Cl}\{x\} \subset X - V$, so $y \notin \mathrm{Cl}\{x\}$. Since $\mathrm{Cl}\{x\}$ cannot contain any other points than x, it follows that $\{x\} = \mathrm{Cl}\{x\}$ and $\{x\}$ is closed. □

[[We omit the T_1 version of the following.]]

Theorem 2.5.37. *Let A be a subset of a Hausdorff space X. A point $x \in X$ is a limit point of A if and only if each neighborhood U of x meets A in infinitely many points.*

Proof. If $U \cap A$ consists of infinitely many points, then it certainly contains other points than x, so U meets $A - \{x\}$.

Conversely, if $U \cap A$ is finite, then then $U \cap (A - \{x\}) = \{x_1, \dots, x_n\}$ is closed, so

$$V = U - \{x_1, \dots, x_n\} = U \cap (X - \{x_1, \dots, x_n\})$$

is open. Then $x \in V$, V is open, and $V \cap (A - \{x\}) = \varnothing$, so x is not a limit point of A. □

2.5.10 Products of Hausdorff spaces

Lemma 2.5.38. *If X and Y are Hausdorff spaces, then so is $X \times Y$.*

Proof. Let (x, y) and (x', y') be distinct points in $X \times Y$. Then $x \neq x'$ or $y \neq y'$. If $x \neq x'$ there are open subsets $U, V \subset X$ with $x \in U$, $x' \in V$ and $U \cap V = \varnothing$. Then $U \times Y, V \times Y$ are open subsets of $X \times Y$, with $(x, y) \in U \times Y$, $(x', y') \in V \times Y$ and $(U \times Y) \cap (V \times Y) = \varnothing$. The argument if $y \neq y'$ is very similar, obtained by interchanging the roles of X and Y. □

2.6 (§18) Continuous Functions

2.6.1 Continuity in terms of preimages

Definition 2.6.1. Let X and Y be topological spaces. A function $f\colon X \to Y$ is said to be *continuous* if for each open set V in Y the preimage

$$f^{-1}(V) = \{x \in X \mid f(x) \in V\}$$

is open in X. A continuous function is also called a *map*.

Lemma 2.6.2. *Let X, Y and Z be topological spaces. If $f\colon X \to Y$ and $g\colon Y \to Z$ are continuous functions, then the composite $g \circ f\colon X \to Z$ is continuous.*

Proof. Let $W \subset Z$ be open. Then $g^{-1}(W) \subset Y$ is open since g is continuous, and $f^{-1}(g^{-1}(W)) \subset X$ is open since f is continuous. But this set equals $(g \circ f)^{-1}(W)$, so $g \circ f$ is continuous. □

Lemma 2.6.3. *Let X and Y be topological spaces, and suppose that $\mathscr{B}$ is a basis for the topology on Y. Then a function $f\colon X \to Y$ is continuous if and only if for each basis element $B \in \mathscr{B}$ the preimage $f^{-1}(B)$ is open in X.*

Proof. Each basis element B is open in Y, so if f is continuous then $f^{-1}(B)$ is open in X.

Conversely, each open $V \subset Y$ is a union $V = \bigcup_{\alpha \in J} B_\alpha$ of basis elements, and

$$f^{-1}(V) = \bigcup_{\alpha \in J} f^{-1}(B_\alpha)\,,$$

so if each $f^{-1}(B_\alpha)$ is open in X, so is $f^{-1}(V)$. □

Lemma 2.6.4. *Let X and Y be topological spaces, and suppose that $\mathscr{S}$ is a subbasis for the topology on Y. Then a function $f\colon X \to Y$ is continuous if and only if for each basis element $S \in \mathscr{S}$ the preimage $f^{-1}(S)$ is open in X.*

Proof. We build on the previous lemma.

Each subbasis element S is a basis element in the associated basis $\mathscr{B}$, so if f is continuous then $f^{-1}(S)$ is open in X.

Conversely, each basis element $B \in \mathscr{B}$ is a finite intersection $B = S_1 \cap \cdots \cap S_n$ of subbasis elements, and

$$f^{-1}(B) = f^{-1}(S_1) \cap \cdots \cap f^{-1}(S_n)\,,$$

so if each $f^{-1}(S_i)$ is open in X, so is $f^{-1}(B)$. □

Example 2.6.5. If (X, d) and (Y, d') are metric spaces, then $f: X \to Y$ is continuous if and only if for each $x \in X$ and $\epsilon > 0$ there is a $\delta > 0$ such that $B_d(x, \delta) \subset f^{-1}(B_{d'}(f(x), \epsilon))$, i.e., such that for all $y \in X$ with $d(x, y) < \delta$ we have $d'(f(x), f(y)) < \epsilon$.

Example 2.6.6. Let $\mathbb{R}_d$ and $\mathbb{R}_{\text{cof}}$ be the real numbers with the standard (metric) topology and the cofinite topology, respectively. The identity function

$$id\colon \mathbb{R}_{\text{cof}} \to \mathbb{R}_d$$

(given by $id(x) = x$) is not continuous, since (a, b) is not open in the cofinite topology, for $a < b$. However, the identity function

$$id\colon \mathbb{R}_d \to \mathbb{R}_{\text{cof}}$$

(still given by $id(x) = x$) is continuous, since $\mathbb{R} - F$ is open in the standard topology, for F finite.

2.6.2 Continuity at a point

Theorem 2.6.7. *Let X and Y be topological spaces, and $f: X \to Y$ a function. Then f is continuous if and only if for each $x \in X$ and each neighborhood V of $f(x)$ there is a neighborhood U of x with $f(U) \subset V$.*

Definition 2.6.8. We say that f is *continuous at x* if for each neighborhood V of $f(x)$ there is a neighborhood U of x with $f(U) \subset V$. Hence $f: X \to Y$ is continuous if and only if it is continuous at each $x \in X$.

Proof. If f is continuous, $x \in X$ and V is a neighborhood of $f(x)$, then $U = f^{-1}(V)$ is a neighborhood of x with $f(U) \subset V$.

Conversely, if V is open in Y and $x \in f^{-1}(V)$ then V is a neighborhood of $f(x)$, so by hypothesis there is a neighborhood U_x of x with $f(U_x) \subset V$. Then $x \in U_x \subset f^{-1}(V)$. Taking the union over all $x \in f^{-1}(V)$ we find that $\bigcup_{x \in f^{-1}(V)} U_x = f^{-1}(V)$ is a union of open sets, hence is open. □

2.6.3 Continuity in terms of closed sets and the closure

Theorem 2.6.9. *Let X and Y be topological spaces, and $f: X \to Y$ a function. The following are equivalent:*

(1) f is continuous.

(2) For every subset $A \subset X$ we have $f(\bar{A}) \subset \overline{f(A)}$.

(3) For every closed L in Y the preimage $f^{-1}(L)$ is closed.

Proof. (1) $\implies$ (2): Assume that f is continuous and $A \subset X$. Each point in $f(\bar{A})$ has the form $f(x)$ for some $x \in \bar{A}$. We must show that $f(x) \in \overline{f(A)}$. Let V be a neighborhood of $f(x)$. By continuity, $f^{-1}(V)$ is a neighborhood of x. Since $x \in \bar{A}$, the intersection $A \cap f^{-1}(V)$ is nonempty. Choose a $y \in A \cap f^{-1}(V)$. Then $f(y) \in f(A) \cap V$, since $y \in A$ implies $f(y) \in f(A)$ and $y \in f^{-1}(V)$ implies $f(y) \in V$. In particular, $f(A)$ meets V. Since V was an arbitrary neighborhood of $f(x)$ we have $f(x) \in \overline{f(A)}$.

(2) $\implies$ (3): Let $L \subset Y$ be closed, and let $A = f^{-1}(L)$. We will show that $A = \bar{A}$, so that A is closed. Now $f(A) \subset L$, so $\overline{f(A)} \subset L$, since L is closed. By hypothesis $f(\bar{A}) \subset \overline{f(A)}$, so $f(\bar{A}) \subset L$, hence $\bar{A} \subset f^{-1}(L) = A$. This implies $A = \bar{A}$.

(3) $\implies$ (1): Let $V \subset Y$ be open, then $L = Y - V$ is closed. By hypothesis

$$f^{-1}(L) = f^{-1}(Y - V) = X - f^{-1}(V)$$

is closed, so $f^{-1}(V)$ is open. Hence f is continuous. □

2.6.4 Homeomorphism = topological equivalence

Definition 2.6.10. A bijective function $f\colon X \to Y$ between topological spaces with the property that both f and $f^{-1}\colon Y \to X$ are continuous, is called a *homeomorphism*. If there exists a homeomorphism $f\colon X \to Y$ we say that X and Y are *homeomorphic spaces*, or that they are *topologically equivalent*, and write $X \cong Y$. (Another common notation for homeomorphism is $X \approx Y$.)

Lemma 2.6.11. *Being homeomorphic is an equivalence relation on any set of topological spaces:*

(1) For each space X the identity function $id\colon X \to X$, with $id(x) = x$ for all $x \in X$, is a homeomorphism.

(2) If $f\colon X \to Y$ is a homeomorphism, then so is the inverse map $f^{-1}\colon Y \to X$.

(3) If $f\colon X \to Y$ and $g\colon Y \to Z$ are homeomorphisms, then so is the composite map $gf\colon X \to Z$.

Lemma 2.6.12. *Let $f\colon X \to Y$ be a bijective function between topological spaces. The following are equivalent:*

(1) f is a homeomorphism.

(2) A set $U \subset X$ is open in X if and only if the image $f(U) \subset Y$ is open in Y.

(3) A set $V \subset Y$ is open in Y if and only if the preimage $f^{-1}(V) \subset X$ is open in X.

Proof. To say that f is continuous means that $V \subset Y$ open implies $f^{-1}(V) \subset X$ open. To say that f^{-1} is continuous means that $U \subset X$ open implies $f(U) \subset Y$ open. Now each $U \subset X$ has the form $U = f^{-1}(V)$ for a unique $V \subset Y$, with $f(U) = f(f^{-1}(V)) = V$. Hence to say that f^{-1} is continuous also means that $f^{-1}(V) \subset X$ open implies $V \subset Y$ open.

This proves that (1) and (3) are equivalent. Replacing f by f^{-1} proves that (1) and (2) are equivalent. □

In other words, f is a homeomorphism if and only the image function

$$f\colon \mathscr{P}(X) \to \mathscr{P}(Y)$$

induces a bijection from the topology on X to the topology on Y, or equivalently, if the preimage function

$$f^{-1}\colon \mathscr{P}(Y) \to \mathscr{P}(X)$$

induces a bijection from the topology on Y to the topology on X.

Remark 2.6.13. Any property of X that can be expressed in terms of its elements and its open subsets is called a *topological property* of X. If $f\colon X \to Y$ is a homeomorphism, than any topological property of X is logically equivalent to the corresponding topological property of Y obtained by replacing each element $x \in X$ by its image $f(x) \in Y$, and each open subset $U \subset X$ by its image $f(U) \subset Y$. Such topological properties are thus preserved by homeomorphisms.

Example 2.6.14. For instance, being a finite topological space, having the discrete, trivial or cofinite topology, or being a Hausdorff space, are all examples of topological properties. So if X is a Hausdorff space and $X \cong Y$ then Y is a Hausdorff space. We shall study other topological properties, like compactness and connectedness, in later sections.

2.6.5 Examples

Example 2.6.15. The two closed intervals $X = [0, 1]$ and $Y = [a, b]$ with $a < b$, each with the subspace topology from $\mathbb{R}$, are homeomorphic:

$$[0, 1] \cong [a, b] \,.$$

One example of a homeomorphism $f\colon [0, 1] \to [a, b]$ is given by the linear function $f(x) = a + (b - a)x = (1 - x)a + xb$. The inverse $f^{-1}\colon [a, b] \to [0, 1]$ is given by the linear function $f^{-1}(y) = (y - a)/(b - a)$. It is well known that both f and f^{-1} are continuous. Hence any two closed intervals $[a, b]$ and $[c, d]$ are homeomorphic, for $a < b$ and $c < d$.

Example 2.6.16. Any two open intervals of the form (a, b), (a, ∞), $(-\infty, b)$ or $\mathbb{R} = (-\infty, \infty)$ are homeomorphic, for $a < b$. First, $(0, 1)$ and (a, b) are homeomorphic:

$$(0, 1) \cong (a, b) \,.$$

The linear map $f\colon (0, 1) \to (a, b)$ given by $f(x) = a+(b-a)x = (1-x)a+xb$ is a homeomorphism, with inverse $f^{-1}\colon (a, b) \to (0, 1)$ given by $f^{-1}(y) = (y - a)/(b - a)$. Next $(0, 1)$ and $(1, \infty)$ are homeomorphic:

$$(0, 1) \cong (1, \infty) \,.$$

The function $f\colon (0, 1) \to (1, \infty)$ given by $f(x) = 1/x$ is a homeomorphism, with inverse $f^{-1}\colon (1, \infty) \to (0, 1)$ given by $f^{-1}(y) = 1/y$. Using $f(x) = x + c$ or $f(x) = -x$ it is easy to see that $(1, \infty) \cong (a, \infty)$ and $(-\infty, b) \cong (-b, \infty)$ for any a and b. Finally, function $f\colon (-1, 1) \to \mathbb{R}$ given by $f(x) = x/(1 - x^2)$ is a homeomorphism:

$$(-1, 1) \cong \mathbb{R} \,.$$

To find the inverse we rewrite the equation $x/(1 - x^2) = y$ as $yx^2 + x - y = 0$ and solve for $x \in (-1, 1)$ as a function of $y \in \mathbb{R}$, namely

$$f^{-1}(y) = \frac{-1 + \sqrt{1 + 4y^2}}{2y} = \frac{2y}{1 + \sqrt{1 + 4y^2}} \,.$$

It is well known that both f and f^{-1} are continuous, hence $\mathbb{R}$ is homeomorphic to $(-1, 1)$, and therefore to any open interval (a, b). Alternatively, $\tan\colon (-\pi/2, \pi/2) \to \mathbb{R}$ and $\arctan\colon \mathbb{R} \to (-\pi/2, \pi/2)$ are both known to be continuous, hence mutually inverse homeomorphisms.

Example 2.6.17. Let $X = \{a, b\}$, and let $\mathscr{T}_a = \{\varnothing, \{a\}, X\}$ and $\mathscr{T}_b = \{\varnothing, \{b\}, X\}$ be the two Sierpinski topologies. These are homeomorphic

$$(X, \mathscr{T}_a) \cong (X, \mathscr{T}_b)$$

by the function $f\colon X \to X$ given by $f(a) = b$ and $f(b) = a$, which is equal to its own inverse. Note that the image function $f\colon \mathscr{P}(X) \to \mathscr{P}(X)$ takes $\mathscr{T}_a$ bijectively to $\mathscr{T}_b$, with inverse given by the preimage function $f^{-1}\colon \mathscr{P}(X) \to \mathscr{P}(X)$. In other words, any statement S about the elements and open sets of $(X, \mathscr{T}_a)$ corresponds to a logically equivalent statement S' about the elements and open sets of $(X, \mathscr{T}_b)$, given by interchanging the roles of a and b. Note also that the identity function $id\colon (X, \mathscr{T}_a) \to (X, \mathscr{T}_b)$ is not a homeomorphism—it is not even continuous.

Example 2.6.18. Let $S^1 \subset \mathbb{R}^2$ be the (unit) circle, and let $N = (0, 1) \in S^1$ be the uppermost point. There is a homeomorphism

$$f\colon S^1 - \{N\} \xrightarrow{\cong} \mathbb{R}$$

called *stereographic projection*, mapping a point $(x, y) \in S^1 - \{N\}$ to $t \in \mathbb{R}$ where $(t, 0)$ is the intersection of the line though N and (x, y) with the x-axis. Hence

$$f(x, y) = t = \frac{x}{1 - y}$$

for all $(x, y) \in S^1 - \{N\}$. This is the restriction of a continuous function from $\mathbb{R} \times (\mathbb{R} - \{1\}) \subset \mathbb{R}^2$ to $\mathbb{R}$ to the subspace $S^1 - \{N\}$, hence f is continuous. The inverse function $f^{-1}\colon \mathbb{R} \to S^1 - \{N\}$ maps t to $f^{-1}(t) = (x, y)$ with $t = x/(1 - y)$ and $x^2 + y^2 = 1$, which we can solve for $y \neq 1$ and then x to obtain

$$f^{-1}(t) = (x, y) = \Big(\frac{2t}{t^2 + 1}, \frac{t^2 - 1}{t^2 + 1}\Big) .$$

Clearly f and f^{-1} are continuous, hence define mutually inverse homeomorphisms.

Note that when (x, y) approaches N from the right $(x > 0)$, the value of t approaches $+\infty$, while when (x, y) approaches N from the left $(x < 0)$, the value of t approaches $-\infty$. If we were to extend f to a homeomorphism from S^1 to a space containing $\mathbb{R}$, the latter space would have to contain exactly one point corresponding to N, which would have to be the limit of sequences $(t_n)_{n=1}^{\infty}$ both for $t_n \to +\infty$ as $n \to \infty$, and for $t_n \to -\infty$ as $n \to \infty$. Hence the additional point, which we might call ∞, should be the limit of all sequences $(t_n)_{n=1}^{\infty}$ with $|t_n| \to \infty$ as $n \to \infty$. There is indeed a topology on the set

$$\mathbb{R} \cup \{\infty\}$$

with this propery, called the *one-point compactification* of $\mathbb{R}$. We will generalize this construction from $X = \mathbb{R}$ to arbitrary locally compact Hausdorff spaces in §29 (Local Compactness).

Example 2.6.19. Let $S^n \subset \mathbb{R}^{n+1}$ be the (unit) n-sphere, and let $N = (0, \ldots, 0, 1) \in S^n$ be the 'north pole'. There is a homeomorphism

$$f\colon S^n - \{N\} \xrightarrow{\cong} \mathbb{R}^n$$

mapping $x = (x_1, \ldots, x_n, y) \in S^n - \{N\} \subset \mathbb{R}^{n+1}$ to $t = (t_1, \ldots, t_n) \in \mathbb{R}^n$ where $(t_1, \ldots, t_n, 0)$ is the intersection of the line though N and x with the linear subspace $\mathbb{R}^n \times \{0\} \subset \mathbb{R}^{n+1}$. Hence

$$f(x) = t = \Big(\frac{x_1}{1 - y}, \ldots, \frac{x_n}{1 - y}\Big) .$$

The inverse function $f^{-1}\colon \mathbb{R}^n \to S^n - \{N\}$ maps $t = (t_1, \ldots, t_n) \in \mathbb{R}^n$ to

$$f^{-1}(t) = x = \Big(\frac{2t_1}{|t|^2 + 1}, \ldots, \frac{2t_n}{|t|^2 + 1}, \frac{|t|^2 - 1}{|t|^2 + 1}\Big)$$

where $|t|^2 = t_1^2 + \cdots + t_n^2$. Clearly f and f^{-1} are continuous, hence define mutually inverse homeomorphisms.

The one-point compactification of $\mathbb{R}^n$ is a space $\mathbb{R}^n \cup \{\infty\}$, such that the stereographic projection f extends to a homeomorphism

$$S^n \cong \mathbb{R}^n \cup \{\infty\}\,.$$

[[Other compactifications, such as the real projective spaces $\mathbb{R}P^n = \mathbb{R}^n \cup \mathbb{R}P^{n-1}$, can be constructed as quotient spaces. We consider these in §22 (The Quotient Topology).]]

2.6.6 Nonexamples

Example 2.6.20. The identity function $id\colon \mathbb{R}_d \to \mathbb{R}_{\text{cof}}$ is a bijection that is continuous but not a homeomorphism, since the inverse function ($id\colon \mathbb{R}_{\text{cof}} \to \mathbb{R}_d$) is not continuous.

Example 2.6.21. Let $X = [0, 1)$ and $Y = S^1$, where

$$S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$$

is the unit circle in the xy-plane, with the subspace topology. Let $f\colon X \to Y$ be given by $f(t) = (\cos(2\pi t), \sin(2\pi t))$ for $t \in [0, 1)$. It is a continuous bijection, but the inverse function $g = f^{-1}\colon S^1 \to [0, 1)$ is not continuous at $f(0) = (1, 0)$. For $U = [0, 1/2)$ is open in $[0, 1)$, while the preimage $g^{-1}(U)$ is not open in S^1. This preimage equals the image $f(U)$, which is the part of S^1 that lies strictly in the upper half-plane (where $y > 0$), together with the point $(1, 0)$. No neighborhood of $(1, 0)$ in S^1 is contained in $f(U)$, so $f(U)$ is not open.

2.6.7 Constructing maps

Theorem 2.6.22. *Let A be a subset of a topological space X. The subspace topology on A is the coarsest topology for which the inclusion $i\colon A \to X$ is continuous, where $i(a) = a$ for all $a \in A$.*

Proof. For i to be continuous in a topology $\mathscr{T}'$ on A, the inverse image $i^{-1}(U) = A \cap U$ must be open in A for each open $U \subset X$, and conversely. This just means that $\mathscr{T}'$ must contain the subspace topology on A. □

Corollary 2.6.23. *The restriction $f|A\colon A \to Y$ of any continuous function $f\colon X \to Y$ to a subspace $A \subset X$ is continuous.*

Lemma 2.6.24. *The corestriction $g\colon X \to B$ of any continuous function $f\colon X \to Y$ is continuous, where $B \subset Y$ is a subspace containing $f(X)$.*

Proof. Each open subset $V \subset B$ has the form $B \cap U$, where $U \subset Y$ is open, hence $g^{-1}(V) = \{x \in X \mid g(x) \in V\} = \{x \in X \mid f(x) \in U\} = f^{-1}(U)$ is open. □

Definition 2.6.25. A map $f\colon X \to Y$ is called an *embedding* (also called an *imbedding*) if the corestriction $g\colon X \to f(X)$ is a homeomorphism, where the image $f(X) \subset Y$ has the subspace topology.

Lemma 2.6.26. *A map $f\colon X \to Y$ is an embedding if and only if it factors as the composite of a homeomorphism $h\colon X \to B$ and the inclusion $j\colon B \to Y$ of a subspace. In particular, any embedding is an injective map.*

Proof. It is clear that an embedding f factors in this way, with $B = f(X)$. Conversely, if $f = j \circ h$ with $j\colon B \to Y$ the inclusion, then $f(X) = B$ and $h\colon X \to B$ equals the corestriction of f. □

Example 2.6.27. The map $f\colon [0,1) \to \mathbb{R}^2$ given by $f(t) = (\cos(2\pi t), \sin(2\pi t))$ is an example of an injective continuous function that is not an embedding, since the corestriction $g\colon [0,1) \to S^1$ to its image is not a homeomorphism.

Theorem 2.6.28. *Let $f\colon X \to Y$ be a function, and suppose that $X = A_1 \cup \cdots \cup A_n$ is covered by a finite collection of closed subsets $A_i \subset X$. Then f is continuous if (and only if) each restriction $f|A_i\colon A_i \to Y$ is continuous.*

Proof. Let $L \subset Y$ be closed. For each $1 \le i \le n$, the preimage $(f|A_i)^{-1}(L)$ is closed in A_i, since $f|A_i$ is continuous, hence is closed in X, since A_i is closed. Then

$$f^{-1}(L) = (f|A_1)^{-1}(L) \cup \cdots \cup (f|A_n)^{-1}(L)$$

is a finite union of closed subsets of X, hence is closed. □

Theorem 2.6.29. *Let $f\colon X \to Y$ be a function, and suppose that $X = \bigcup_{\alpha\in J} A_\alpha$ is covered by a collection of open subsets $A_\alpha \subset X$. Then f is continuous if (and only if) each restriction $f|A_\alpha\colon A_\alpha \to Y$ is continuous.*

Proof. Let $V \subset Y$ be open. For each $\alpha \in J$, the preimage $(f|A_\alpha)^{-1}(V)$ is open in A_α, since $f|A_\alpha$ is continuous, hence is open in X, since A_α is open. Then

$$f^{-1}(V) = \bigcup_{\alpha\in J} (f|A_\alpha)^{-1}(V)$$

is a union of open subsets of X, hence is open. □

[[See Munkres p. 107/109 for examples.]]

Lemma 2.6.30. *Let $P = \{p\}$ be a singleton set, with the unique topology. For each topological space X the unique map $f\colon X \to P$ is continuous.*

Proof. The only open subsets of P are $\varnothing$ and P, with preimages $\varnothing$ and X, respectively, and these are open. □

Corollary 2.6.31. *Each constant function $c\colon X \to Y$ to a point $p \in Y$ is continuous, where $c(x) = p$ for all $x \in X$.*

2.6.8 Maps into products

Let X and Y be topological spaces, and give $X \times Y$ the product topology. Recall, or introduce, the notations $\pi_1\colon X \times Y \to X$ for the (first) projection $\pi_1(x,y) = x$ and $\pi_2\colon X \times Y \to Y$ for the (second) projection $\pi_2(x,y) = y$.

Lemma 2.6.32. *$\pi_1\colon X \times Y \to X$ and $\pi_2\colon X \times Y \to Y$ are continuous.*

Proof. For each open subset $U \subset X$ the preimage $\pi_1^{-1}(U) = U \times Y$ is in the basis for the product topology on $X \times Y$. In particular, it is open. Hence π_1 is continuous. The case for π_2 is very similar. □

Theorem 2.6.33. *Let W be any topological space. A function $f\colon W \to X \times Y$ is continuous if and only if both of its components $f_1 = \pi_1 \circ f\colon W \to X$ and $f_2 = \pi_2 \circ f\colon W \to Y$ are continuous.*

Proof. The 'only if' part follows from the continuity of π_1 and π_2, and the fact that the composite of two maps is a map.

For the 'if' part, note that for each open $U \subset X$ we have

$$f_1^{-1}(U) = f^{-1}(U \times Y)$$

and for each open subset $V \subset Y$ we have

$$f_2^{-1}(V) = f^{-1}(X \times V) \,.$$

If f_1 and f_2 are continuous, then $f_1^{-1}(U)$ and $f_2^{-1}(V)$ are open. Hence so is their intersection

$$f_1^{-1}(U) \cap f_2^{-1}(V) = f^{-1}(U \times Y) \cap f^{-1}(X \times V) = f^{-1}(U \times Y \cap X \times V) = f^{-1}(U \times V) \,.$$

Letting U and V vary, $U \times V$ ranges over the 'standard' basis for the product topology on $X \times Y$, and we have just seen that the preimage $f^{-1}(U \times V)$ of each of these basis elements is open in W. This suffices to prove that f is continuous. □

Example 2.6.34. Let $W = (a, b) \subset \mathbb{R}$ and $X = Y = \mathbb{R}$. A function $f\colon (a, b) \to \mathbb{R}^2$ can be written $f(t) = (f_1(t), f_2(t))$. Then f is continuous if and only if both of the component functions f_1 and f_2 are continuous.

Corollary 2.6.35. *The product topology on $X \times Y$ is the coarsest topology for which both of the projection maps $\pi_1\colon X \times Y \to X$ and $\pi_2\colon X \times Y \to Y$ are continuous.*

Proof. The projection maps are continuous for the product topology. Conversely, for π_1 and π_2 to be continuous with respect to a topology $\mathscr{T}'$ on $X \times Y$ is equivalent to asking that $U \times Y$ and $X \times V$ lie in $\mathscr{T}'$ for all open $U \subset X$ and $V \subset Y$. This is in turn equivalent to asking that

$$U \times V = (U \times Y) \cap (X \times V)$$

lies in $\mathscr{T}'$, for all open $U \subset X$ and $V \subset Y$, i.e., that $\mathscr{T}'$ is finer than the product topology. □

2.6.9 Maps out of products

Let Z be any topological space. The analysis of continuous functions $X \times Y \to Z$ is not as simple as the case of maps $W \to X \times Y$, but for reasonable spaces X (or Y) there are reasonable answers.

First, ignoring topology, there is a bijective correspondence between functions

$$f\colon X \times Y \longrightarrow Z$$

(of two variables, $x \in X$ and $y \in Y$) and functions

$$g\colon X \longrightarrow \mathrm{Func}(Y, Z)$$

where $\mathrm{Func}(Y, Z) = \{h\colon Y \to Z\}$ is the set of all functions from Y to Z. The correspondence sends f to $g\colon x \mapsto g(x)$ where $g(x)\colon Y \to Z$ is given by

$$f(x, y) = g(x)(y) \,.$$

Letting f vary, we have a bijection (of sets)

$$\mathrm{Func}(X \times Y, Z) \cong \mathrm{Func}(X, \mathrm{Func}(Y, Z)) \,.$$

Taking topologies into account, we might ask for a similar description of the subset

$$\mathrm{Cont}(X \times Y, Z) \subset \mathrm{Func}(X \times Y, Z)$$

of continuous functions $X \times Y \to Z$. Here $\mathrm{Cont}(Y, Z) = \{\text{continuous } h\colon Y \to Z\}$. (We might later abbreviate this to $\mathscr{C}(Y, Z)$.)

Suppose hereafter that $f\colon X \times Y \to Z$ is continuous. For each $x \in X$, the inclusion

$$\begin{aligned} i_x\colon Y &\longrightarrow X \times Y \\ y &\mapsto (x, y) \end{aligned}$$

is an embedding, topologically identifying Y with its image $\{x\} \times Y \subset X \times Y$ in the subspace topology. In particular, i_x is continuous, so for each $x \in X$ the function

$$g(x) = f \circ i_x\colon Y \longrightarrow Z$$

is the composite of two maps, hence is itself a map. Thus, $g\colon X \to \mathrm{Func}(Y, Z)$ in fact takes values in $\mathrm{Cont}(Y, Z)$, and we get an injective function

$$\begin{aligned} \mathrm{Cont}(X \times Y, Z) &\longrightarrow \mathrm{Func}(X, \mathrm{Cont}(Y, Z)) \\ f &\longmapsto (g\colon x \mapsto g(x) = f \circ i_x) \end{aligned}$$

However, this function is not surjective. In other words, a function $f\colon (x, y) \mapsto f(x, y)$ may be continuous in the second variable (y) for each value of the first variable (x), without being continuous.

This is perhaps not so surprising. A little less obvious is that f can be continuous in each variable separately, and still not be continuous. A standard example is $f\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ given by

$$f(x, y) = \begin{cases} \dfrac{2xy}{x^2 + y^2} & \text{for } (x, y) \neq (0, 0), \\ 0 & \text{for } (x, y) = (0, 0). \end{cases}$$

Here $y \mapsto f(x, y)$ is continuous as a function $\mathbb{R} \to \mathbb{R}$ for each y, and $x \mapsto f(x, y)$ is continuous for each x, but f is not continuous at $(0, 0)$ as a function $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$. This is clear from the formula $f(x, y) = \sin(2\theta)$ where $(x, y) = (r \cos\theta, r \sin\theta) \neq (0, 0)$.

To recognize the maps $f\colon X \times Y \to Z$ in terms of the maps $g(x)\colon Y \to Z$ for $x \in X$, we might impose further conditions on the function

$$\begin{aligned} g\colon X &\longrightarrow \mathrm{Cont}(Y, Z) \\ x &\longmapsto g(x) \end{aligned}$$

In particular, we might ask that the function g is itself continuous. However, we can only make sense of this if the set $\mathrm{Cont}(Y, Z)$ is turned into a topological space. In other words, we have to specify a topology on the set of maps $Y \to Z$. In §46 (Pointwise and Compact Convergence) we shall introduce a topology on $\mathrm{Cont}(Y, Z)$, called the *compact–open topology* (S. Lefschetz, ca. 1942(?)). The functions $g\colon X \to \mathrm{Cont}(Y, Z)$ that arise from maps $f\colon X \times Y \to Z$ will then be continuous, with respect to the given topology on X and the compact–open topology on $\mathrm{Cont}(Y, Z)$. Hence we get an injective function

$$\begin{aligned} \mathrm{Cont}(X \times Y, Z) &\longrightarrow \mathrm{Cont}(X, \mathrm{Cont}(Y, Z)) \\ f &\longmapsto g \end{aligned}$$

and for locally compact Hausdorff spaces Y, this function is also surjective, hence bijective (R. H. Fox, 1945). Viewing $\mathrm{Cont}(X \times Y, Z)$ and $\mathrm{Cont}(X, \mathrm{Cont}(Y, Z))$ as topological spaces, with the compact–open topologies, it also follows that this bijection is in fact a homeomorphism (J. R. Jackson, 1952).

In general there are many possible topologies that can be imposed on sets of functions, such as $\mathrm{Func}(Y, Z)$ and $\mathrm{Cont}(Y, Z)$, and the resulting spaces are called *function spaces*. In the first case, $\mathrm{Func}(Y, Z)$ can also be viewed as a product of one copy of Z for each element of Y:

$$\begin{aligned} \mathrm{Func}(Y, Z) &\xrightarrow{\cong} \prod_{y \in Y} Z \\ h &\longmapsto (h(y))_{y \in Y} \end{aligned}$$

In the case when $Y = \{y_1, y_2\}$ has only two elements, we have already discussed the product topology on $\prod_{y \in Y} Z = Z \times Z$. We now extend this to the case of arbitrary indexing sets Y. We can also relax that condition that all factors Z in the product are the same. This leads us to general products, which is the subject of the next section.

2.7 (§19) The Product Topology

Definition 2.7.1. Let $\{X_\alpha\}_{\alpha \in J}$ be an indexed collection of sets. The *cartesian product*

$$\prod_{\alpha \in J} X_\alpha$$

is the set of J-indexed sequences $(x_\alpha)_{\alpha \in J}$ with $x_\alpha \in X_\alpha$, for each $\alpha \in J$.

Definition 2.7.2. For each $\beta \in J$ there is a *projection function*

$$\pi_\beta \colon \prod_{\alpha \in J} X_\alpha \to X_\beta$$

taking $(x_\alpha)_{\alpha \in J}$ to x_β.

Now suppose that each X_α is a topological space. We wish to equip $\prod_{\alpha \in J} X_\alpha$ with the coarsest possible topology $\mathscr{T}$ such that each projection π_β is continuous. In other words, for each open subset $U_\beta \subset X_\beta$, the preimage $\pi_\beta^{-1}(U_\beta)$ must be open in $\prod_{\alpha \in J} X_\alpha$. Here

$$\pi_\beta^{-1}(U_\beta) = \prod_{\alpha \in J} A_\alpha$$

where

$$A_\alpha = \begin{cases} U_\alpha & \text{for } \alpha = \beta, \\ X_\alpha & \text{otherwise.} \end{cases}$$

The collection

$$\mathscr{S} = \{\pi_\beta^{-1}(U_\beta) \mid \beta \in J,\, U_\beta \subset X_\beta \text{ open}\}$$

must therefore be contained in the topology $\mathscr{T}$.

Furthermore, each finite intersection

$$B = \pi_{\beta_1}^{-1}(U_{\beta_1}) \cap \cdots \cap \pi_{\beta_n}^{-1}(U_{\beta_n})$$

of such preimages must be open in $\prod_{\alpha \in J} X_\alpha$, since any finite intersection of open sets is open. Here $\beta_i \in J$ and U_{β_i} is an open subset of X_{β_i}, for each $1 \le i \le n$. We may assume that $\beta_i \ne \beta_j$ for $i \ne j$, and in this case the intersection can be written as

$$B = \bigcap_{i=1}^{n} \pi_{\beta_i}^{-1}(U_{\beta_i}) = \prod_{\alpha \in J} A_\alpha$$

where

$$A_\alpha = \begin{cases} U_\alpha & \text{for } \alpha = \beta_i,\ 1 \le i \le n, \\ X_\alpha & \text{otherwise.} \end{cases}$$

The collection

$$\begin{aligned} \mathscr{B} &= \{S_1 \cap \cdots \cap S_n \mid S_i \in \mathscr{S}, 1 \le i \le n\} \\ &= \{\pi_{\beta_1}^{-1}(U_{\beta_1}) \cap \cdots \cap \pi_{\beta_n}^{-1}(U_{\beta_n}) \mid \beta_i \in J,\ U_{\beta_i} \text{ open in } X_{\beta_i},\ 1 \le i \le n\} \end{aligned}$$

of finite intersections of the elements in $\mathscr{S}$ is evidently a basis. For $\prod_{\alpha \in J} X_\alpha$ is an element in the collection, and given any two elements B_1 and B_2 in $\mathscr{B}$ the intersection $B_1 \cap B_2$ is again an element in $\mathscr{B}$.

Definition 2.7.3. The *product topology* on $\prod_{\alpha \in J} X_\alpha$ is the topology $\mathscr{T}$ generated by the basis

$$\mathscr{B} = \{\pi_{\beta_1}^{-1}(U_{\beta_1}) \cap \cdots \cap \pi_{\beta_n}^{-1}(U_{\beta_n}) \mid \beta_i \in J,\ U_{\beta_i} \text{ open in } X_{\beta_i},\ 1 \le i \le n\}$$

consisting of all finite intersections of preimages

$$\pi_\beta^{-1}(U_\beta)\,,$$

where β ranges over the indexing set J and U_β ranges over the open subsets of X_β.

In this situation we call the collection $\mathscr{S}$ a *subbasis* for the topology $\mathscr{T}$. The collection $\mathscr{B}$ of all finite intersections of sets in $\mathscr{S}$ is a basis, and the associated topology $\mathscr{T}$ consists of all unions of sets in $\mathscr{B}$. See §13 (Basis for a Topology) for more details.

Lemma 2.7.4. *The elements of the subbasis $\mathscr{S}$ generating the product topology on $\prod_{\alpha \in J} X_\alpha$ are precisely the products*

$$S = \prod_{\alpha \in J} U_\alpha$$

where each $U_\alpha \subset X_\alpha$ is open, and $U_\alpha \ne X_\alpha$ for at most one $\alpha \in J$.

The elements of the associated basis $\mathscr{B}$ for the product topology on $\prod_{\alpha \in J} X_\alpha$ are precisely the products

$$B = \prod_{\alpha \in J} U_\alpha$$

where each $U_\alpha \subset X_\alpha$ is open, and $U_\alpha \ne X_\alpha$ for only finitely many $\alpha \in J$.

Proof. For each $\beta \in J$ and $U_\beta \subset X_\beta$ open we have

$$\pi_\beta^{-1}(U_\beta) = \prod_{\alpha \in J} U_\alpha$$

where $U_\alpha = X_\alpha$ for all $\alpha \ne \beta$.

The intersection of two subbasis elements $\pi_\beta^{-1}(U)$ and $\pi_\beta^{-1}(V)$, with $U, V \subset X_\beta$ for the same $\beta \in J$, is equal to the subbasis element $\pi_\beta^{-1}(U \cap V)$. Hence the basis elements all have the form

$$\pi_{\beta_1}^{-1}(U_{\beta_1}) \cap \cdots \cap \pi_{\beta_n}^{-1}(U_{\beta_n})$$

for some finite subset $\{\beta_1, \ldots, \beta_n\} \subset J$ and open subsets $U_{\beta_i} \subset X_{\beta_i}$ for $1 \leq i \leq n$, where all of the β_i are distinct. This finite intersection equals

$$\prod_{\alpha \in J} U_\alpha$$

where $U_\alpha = X_\alpha$ for all $\alpha \notin \{\beta_1, \ldots, \beta_n\}$. □

As always, each open subset of $\prod_{\alpha \in J} X_\alpha$ is the union of a collection of basis elements, and each such union is open.

2.7.1 Pointwise convergence

In the special case of a J-indexed sequence $\{X_\alpha\}_{\alpha \in J}$ where all the spaces X_α are equal (to one space X), the product

$$\prod_{\alpha \in J} X = X^J = \mathrm{Func}(J, X)$$

is the set of functions $f \colon J \to X$, with f corresponding to $(f(\alpha))_{\alpha \in J}$. Note that $\pi_\beta(f) = f(\beta)$, for each $\beta \in J$. The product topology on this set is often called the topology of *pointwise convergence*, due to the following property:

Proposition 2.7.5. *Let $(f_n)_{n=1}^\infty$ be a sequence of functions $f_n \colon J \to X$, and let $f \colon J \to X$ be another such function. Then*

$$f_n \to f \text{ as } n \to \infty$$

in $\mathrm{Func}(J, X)$ *if and only if the functions f_n converge pointwise to f, i.e., if and only if for each $\beta \in J$ we have*

$$f_n(\beta) \to f(\beta) \text{ as } n \to \infty$$

in X.

Proof. Suppose that $f_n \to f$ as $n \to \infty$ in the product topology on $\prod_J X = X^J = \mathrm{Func}(J, X)$. Consider any $\beta \in J$. To show that $f_n(\beta) \to f(\beta)$ as $n \to \infty$ we must show that for any neighborhood U of $f(\beta)$ in X there is an N such that $f_n(\beta) \in U$ for all $n \geq N$. Note that $\pi_\beta^{-1}(U)$ is a neighborhood of f in $\mathrm{Func}(J, X)$. Hence there is an N such that $f_n \in \pi_\beta^{-1}(U)$ for all $n \geq N$. This is equivalent to the required property, that $f_n(\beta) \in U$ for all $n \geq N$.

Conversely, suppose that $f_n(\beta) \to f(\beta)$ for each $\beta \in J$. To show that $f_n \to f$ as $n \to \infty$ we must show that for any neighborhood V of f in $\mathrm{Func}(J, X)$ there is an N such that $f_n \in V$ for all $n \geq N$. Given f and V there exists a basis element

$$B = \pi_{\beta_1}^{-1}(U_{\beta_1}) \cap \cdots \cap \pi_{\beta_m}^{-1}(U_{\beta_m})$$

for the product topology, with $f \in B \subset V$. Here $\beta_1, \ldots, \beta_m \in J$ and $U_{\beta_1}, \ldots, U_{\beta_m}$ are open in X. Since $f \in B \subset \pi_{\beta_1}^{-1}(U_{\beta_1})$ we have $f(\beta_1) \in U_{\beta_1}$, and similarly $f(\beta_2) \in U_{\beta_2}, \ldots, f(\beta_m) \in U_{\beta_m}$. Since $f_n(\beta_1) \to f(\beta_1)$ as $n \to \infty$, there exists an N_1 such that $f_n(\beta_1) \in U_{\beta_1}$ for all $n \geq N_1$. Similarly, there exists N_2 such that $f_n(\beta_2) \in U_{\beta_2}$ for all $n \geq N_2$, ..., and N_m such that $f_n(\beta_m) \in U_{\beta_m}$ for all $n \geq N_m$. Let $N = \max\{N_1, N_2, \ldots, N_m\}$. Then $f(\beta_i) \in U_{\beta_i}$ for all $1 \leq i \leq m$ and $n \geq N$. Hence $f \in B \subset V$ for all $n \geq N$. Since the neighborhood V of f was arbitrarily chosen, this means that $f_n \to f$ as $n \to \infty$. □

2.7.2 Properties of general product spaces

Theorem 2.7.6. *Let $\{X_\alpha\}_{\alpha\in J}$ be a collection of topological spaces. A function $f\colon A \to \prod_{\alpha\in J} X_\alpha$ is continuous if and only if all of its components $f_\beta = \pi_\beta \circ f\colon A \to X_\alpha$ are continuous.*

Corollary 2.7.7. *The product topology on $\prod_{\alpha\in J} X_\alpha$ is the coarsest topology for which all of the projection maps $\pi_\beta\colon \prod_{\alpha\in J} X_\alpha \to X_\beta$ are continuous.*

Theorem 2.7.8. *Let A_α be a subspace of X_α for each $\alpha \in J$. The product topology on $\prod_{\alpha\in J} A_\alpha$ equals the subspace topology from $\prod_{\alpha\in J} X_\alpha$.*

Theorem 2.7.9. *If X_α is a Hausdorff space, for each $\alpha \in J$, then $\prod_{\alpha\in J} X_\alpha$ is Hausdorff.*

Theorem 2.7.10. *Let A_α be a subset of X_α for each $\alpha \in J$. The closure of the product $\prod_{\alpha\in J} A_\alpha$ equals the product of the closures:*

$$\overline{\prod_{\alpha\in J} A_\alpha} = \prod_{\alpha\in J} \bar{A}_\alpha \,.$$

[[See Munkres page 116.]]

2.8 (§20) The Metric Topology

2.8.1 Bounded metrics

Let (X, d) be a metric space. Recall the associated metric topology $\mathscr{T}_d$ on X.

Definition 2.8.1. A topological space $(X, \mathscr{T})$ is *metrizable* if there exists a metric d on X so that $\mathscr{T}$ is the topology associated to d.

Definition 2.8.2. A metric space (X, d) is *bounded* if there exists a number M such that $d(x, y) \le M$ for all $x, y \in X$. If (X, d) is bounded the *diameter* of X is the least upper bound

$$\operatorname{diam}(X) = \sup\{d(x, y) \mid x, y \in X\} \,.$$

Being bounded is a metric, not a topological property:

Theorem 2.8.3. *Let (X, d) be a metric space. Define the* standard bounded metric $\bar{d}\colon X\times X \to \mathbb{R}$ *by*

$$\bar{d}(x, y) = \min\{d(x, y), 1\} \,.$$

Then $\bar{d}$ is a bounded metric on X that defines the same topology as d.

Proof. Checking that $\bar{d}(x, y) = 0$ if and only if $x = y$, and $\bar{d}(y, x) = \bar{d}(x, y)$, is trivial. To prove the triangle inequality

$$\bar{d}(x, z) \le \bar{d}(x, y) + \bar{d}(y, z)$$

we divide into two cases. If $d(x, y) \ge 1$ or $d(y, z) \ge 1$, then $\bar{d}(x, y) = 1$ or $\bar{d}(y, z) = 1$, so

$$\bar{d}(x, z) \le 1 \le \bar{d}(x, y) + \bar{d}(y, z) \,.$$

Otherwise, $\bar{d}(x, y) = d(x, y)$ and $\bar{d}(y, z) = d(y, z)$, so

$$\bar{d}(x, z) \le d(x, z) \le d(x, y) + d(y, z) = \bar{d}(x, y) + \bar{d}(y, z) \,.$$

It is clear that $\bar{d}(x, y) \le 1$ for all $x, y \in X$, so $(X, \bar{d})$ is bounded.

In any metric space, the collection of ϵ-balls with $\epsilon < 1$ is a basis for the associated topology. These collections are the same for d and $\bar{d}$. □

2.8.2 Euclidean n-space

Definition 2.8.4. Let $X = \mathbb{R}^n$ be the set of real n-tuples $x = (x_1, \dots, x_n)$. The *Euclidean norm* on $\mathbb{R}^n$ is given by

$$\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$$

and the *Euclidean metric* is defined by

$$d(x, y) = \|y - x\| .$$

The *sup norm* = *max norm* on $\mathbb{R}^n$ is given by

$$\|x\|_\infty = \max\{|x_1|, \dots, |x_n|\}$$

and the *square metric* is defined by

$$\rho(x, y) = \|y - x\|_\infty .$$

Here ρ is the Greek letter 'rho'.

Theorem 2.8.5. *The Euclidean metric d, the square metric ρ, and the product metric of n copies of $\mathbb{R}$, all define the same topology on $\mathbb{R}^n$.*

Proof. For each $x \in \mathbb{R}^n$, any neighborhood of x in one of these topologies contains neighborhoods of x in the two other topologies. □

2.8.3 Infinite dimensional Euclidean space

For any indexing set J, consider the set $\mathbb{R}^J$ of real J-tuples $x = (x_\alpha)_{\alpha \in J}$, or equivalently, of functions

$$x \colon J \to \mathbb{R} .$$

For example, when $J = \{1, 2, \dots, n\}$ we can identify $\mathbb{R}^{\{1,2,\dots,n\}}$ with $\mathbb{R}^n$.

When $J = \{1, 2, \dots\} = \mathbb{N}$ we write $\mathbb{R}^\omega$ for the set of real sequences $x = (x_n)_{n=1}^\infty$. (Recall that ω denotes the first infinite ordinal, represented by $\mathbb{N}$ with the usual linear ordering. The ordering plays no role in the definition of the topology on $\mathbb{R}^\omega$.)

The formulas

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \dots}$$

and

$$\|x\|_\infty = \sup\{|x_1|, |x_2|, \dots\}$$

are not well-defined for all $x \in \mathbb{R}^\omega$.

However, $\|x\|_\infty$ does make sense for all bounded sequences in $\mathbb{R}$. Replacing the usual metric on $\mathbb{R}$ with the standard bounded metric will therefore allow us to generalize the square metric to infinite dimensions.

Definition 2.8.6. Let (X, d) be any metric space, with associated standard bounded metric $\bar{d}$. Let J be any set, and let X^J be the set of J-tuples $x \colon J \to X$ in X. The function

$$\bar{\rho}(x, y) = \sup\{\bar{d}(x_\alpha, y_\alpha) \mid \alpha \in J\}$$

defines a metric on X^J, called the *uniform metric* on X. The associated topology is called the *uniform topology.*

Remark 2.8.7. The sets

$$B(x,\epsilon) = \{y \in X^J \mid \ d(x_\alpha, y_\alpha) < \epsilon \text{ for each } \alpha \in J\ \}$$

with $x \in X$ and $0 < \epsilon < 1$ also form a basis for the uniform topology. This follows from the inclusions

$$B_{\bar\rho}(x, \epsilon/2) \subset B(x, \epsilon) \subset B_{\bar\rho}(x, \epsilon)\,.$$

Theorem 2.8.8. *The uniform topology on $\mathbb{R}^J$ is finer than the product topology.*

Proof. Consider a point $x = (x_\alpha)_{\alpha \in J}$ in $\mathbb{R}^J$ and a product topology basis neighborhood

$$B = \prod_{\alpha \in J} U_\alpha$$

of x, where each U_α is open and $U_\alpha \neq \mathbb{R}$ only for $\alpha \in \{\alpha_1, \dots, \alpha_n\} \subset J$. For each i choose an $\epsilon_i > 0$ so that $B_{\bar d}(x_{\alpha_i}, \epsilon_i) \subset U_i$. Let $\epsilon = \min\{\epsilon_1, \dots, \epsilon_n\}$. Then $B_{\bar\rho}(x, \epsilon) \subset B$, since if $y \in B_{\bar\rho}(x, \epsilon)$ then

$$\bar\rho(x, y) = \sup\{\bar d(x_\alpha, y_\alpha) \mid \alpha \in J\} < \epsilon$$

so that $\bar d(x_\alpha, y_\alpha) < \epsilon$ for all $\alpha \in J$. In particular, $y_\alpha \in U_\alpha$ for all $\alpha \in J$. □

[[Exercise: The uniform topology is strictly finer than the product topology for J infinite.]]

The product topology is metrizable when J is countable. It suffices to consider the case $J = \mathbb{N}$.

Theorem 2.8.9. *Let $\bar d(x, y) = \min\{|y - x|, 1\}$ be the standard bounded metric on $\mathbb{R}$. For points $x = (x_n)_{n=1}^\infty$, $y = (y_n)_{n=1}^\infty \in \mathbb{R}^\omega$ define*

$$D(x, y) = \sup\Big\{\frac{\bar d(x_n, y_n)}{n} \mid n \in \mathbb{N}\Big\}\,.$$

Then D is a metric that induces the product topology on $\mathbb{R}^\omega$.

Proof. ((See Munkres p. 125 for the proof that D is a metric.))

Let us show that the product topology on $\mathbb{R}^\omega$ is finer than the metric topology from D. Suppose that $x \in V$ where $V \subset \mathbb{R}^\omega$ is open in the metric topology. Then there is an $\epsilon > 0$ with $B_D(x, \epsilon) \subset V$. We seek an open set U in the product topology with $x \in U \subset B_D(x, \epsilon)$.

Let

$$U_n = B_{\bar d}(x_n, n\epsilon/2)$$

for all $n \in \mathbb{N}$. More explicitly, $U_n = (x_n - n\epsilon/2, x_n + n\epsilon/2)$ for all $1/n \geq \epsilon/2$, and $U_n = \mathbb{R}$ for all $1/n < \epsilon/2$. Then each U_n is open, and $U_n \neq \mathbb{R}$ only for finitely many n, so

$$U = \prod_{n=1}^\infty U_n$$

is open in the product topology, with $x \in U$. If $y \in U$ then $y_n \in U_n$, so $\bar d(x_n, y_n) < n\epsilon/2$ and $\bar d(x_n, y_n)/n < \epsilon/2$, for each $n \in \mathbb{N}$. Hence $D(x, y) \leq \epsilon/2 < \epsilon$ and $U \subset B_D(x, \epsilon)$, as required.

Conversely, we show that the metric topology from D is finer than the product topology on $\mathbb{R}^\omega$. Let $x \in U$ where $U \subset \mathbb{R}^\omega$ is open in the product topology. Then there is a basis element B for the product topology, with $x \in B \subset U$. We seek an $\epsilon > 0$ such that $B_D(x, \epsilon) \subset B$.

The basis element has the form

$$B = \prod_{n=1}^\infty U_n$$

where $x_n \in U_n$ is open for each n, and $U_n \neq \mathbb{R}$ only for finitely many n. If $U_n = \mathbb{R}$ let $\delta_n = n$, otherwise choose $0 < \delta_n < 1$ such that $(x_n - \delta_n, x_n + \delta_n) \subset U_n$. Let

$$\epsilon = \min\Big\{\frac{\delta_n}{n} \mid n \in \mathbb{N}\Big\}.$$

The minimum is well-defined since only finitely many of these numbers are different from 1. If $y \in B_D(x, \epsilon)$ then

$$\bar{d}(x_n, y_n)/n < \epsilon \leq \delta_n/n$$

for all n. It follows that $y_n \in U_n$, since there is only something to check if $U_n \neq \mathbb{R}$, in which case $\delta_n < 1$, and $\bar{d}(x_n, y_n) < \delta_n$ implies $y_n \in (x_n - \delta_n, x_n + \delta_n) \subset U_n$. Hence $B_D(x, \epsilon) \subset B$, as required. □

2.9 (§21) The Metric Topology (continued)

Definition 2.9.1. A topological space X has a *countable basis at a point* $x \in X$ if there is a countable collection $\{B_n\}_{n=1}^\infty$ of neighborhoods of x in X, such that each neighborhood U of x contains (at least) one of the B_n. A space with a countable basis at each of its points is said to satisfy the *first countability axiom*. Replacing each B_n by $B_1 \cap \dots \cap B_n$ we may assume that the neighborhoods are nested:

$$B_1 \supset B_2 \supset \dots \supset B_n \supset \dots.$$

Lemma 2.9.2. *Each metric space (X, d) satisfies the first countability axiom.*

Proof. A countable basis at $x \in X$ is given by the neighborhoods $B_d(x, 1/n)$ for $n \in \mathbb{N}$. □

Lemma 2.9.3 (The sequence lemma)**.** *Let $A \subset X$.*

(a) If there is a sequence $(x_n)_{n=1}^\infty$ of points in A that converges to x, then $x \in \bar{A}$.

(b) If X is metrizable, and $x \in \bar{A}$, then there is a sequence $(x_n)_{n=1}^\infty$ of points in A that converges to x.

Proof. (a) Suppose $x_n \to x$ as $n \to \infty$. For any neighborhood U of x then there is an $N \in \mathbb{N}$ such that $x_n \in U$ for all $n \geq N$. In particular, $A \cap U \neq 0$. Since U was arbitrary, $x \in \bar{A}$.

(b) Let $x \in \bar{A}$. For each $n \in \mathbb{N}$ the neighborhood $B_d(x, 1/n)$ of x meets A, so we can choose an $x_n \in A \cap B_d(x, 1/n)$. Then $x_n \to x$ as $n \to \infty$, as each neighborhood U of x contains some $B_d(x, 1/N)$, hence also each $B_d(x, 1/n)$ for $n \geq N$, so $x_n \in U$ for all $n \geq N$. □

Theorem 2.9.4. *Let $f\colon X \to Y$ be a function.*

(a) If f is continuous, then for every convergent sequence $x_n \to x$ in X the image sequence $f(x_n)$ converges to $f(x)$ in Y.

(b) If X is metrizable, and for every convergent sequence $x_n \to x$ in X the image sequence $f(x_n)$ converges to $f(x)$ in Y, then f is continuous.

Proof. (a) Consider any neighborhood V of $f(x)$ in Y. Since f is continuous, the preimage $f^{-1}(V)$ is a neighborhood of x in X. If $x_n \to x$ as $n \to \infty$ then there is an $N \in \mathbb{N}$ such that $x_n \in f^{-1}(V)$ for all $n \geq N$. Then $f(x_n) \in V$ for all $n \geq N$. Since V was arbitrarily chosen, $f(x_n) \to f(x)$ as $n \to \infty$.

(b) Let $A \subset X$ be any subset. We prove that $f(\bar{A}) \subset \overline{f(A)}$, which implies that f is continuous. Any point in $f(\bar{A})$ has the form $f(x)$ with $x \in \bar{A}$. By the lemma above, there exists a sequence $(x_n)_{n=1}^\infty$ in A with $x_n \to x$ as $n \to \infty$. By hypothesis, $f(x_n) \to f(x)$ as $n \to \infty$, where $(f(x_n))_{n=1}^\infty$ is a sequence of points in $f(A) \subset Y$. Hence $f(x)$ is in the closure of $f(A)$. □

Lemma 2.9.5. *An uncountable product* $\mathbb{R}^J$ *of copies of* $\mathbb{R}$ *is not metrizable.*

Proof. Let J be uncountable. We show that $X = \mathbb{R}^J$ does not satisfy the converse claim in the sequence lemma. Let

$$A = \{(x_\alpha)_{\alpha \in J} \mid x_\alpha \neq 0 \text{ for only finitely many } \alpha \in J\}$$

and $x = (1)_{\alpha \in J}$. We shall prove (a) that $x \in \bar{A}$, but (b) there is no sequence $(x_n)_{n=1}^\infty$ of points in A that converges to x. Here each x_n is a J-tuple $(x_{n,\alpha})_{\alpha \in J}$.

Claim (a): Consider any basis element $B = \prod_{\alpha \in J} U_\alpha$ containing $x = (1)_{\alpha \in J}$. Here $U_\alpha \subset \mathbb{R}$ is open, $1 \in U_\alpha$, and $U_\alpha \neq \mathbb{R}$ only for finitely many $\alpha \in J$. Let

$$y_\alpha = \begin{cases} 0 & \text{if } U_\alpha = \mathbb{R}, \\ 1 & \text{if } U_\alpha \neq \mathbb{R}. \end{cases}$$

Then $y = (y_\alpha)_{\alpha \in J}$ is in A and in B, hence $A \cap B \neq \varnothing$. So $x \in \bar{A}$.

Claim (b): Let $(x_n)_{n=1}^\infty$ be any sequence of points in A. For each $n \in \mathbb{N}$, let J_n be the finite set of $\alpha \in J$ such that $x_{n,\alpha} \neq 0$. Then $\bigcup_{n=1}^\infty J_n$ is a countable union of finite sets, hence is countable. Since J is uncountable there exists a $\beta \in J - \bigcup_{n=1}^\infty J_n$. Then $x_{n,\beta} = 0$ for all n, and

$$U = \pi_\beta^{-1}(1/2, 3/2)$$

is a neighborhood of $x = (1)_{\alpha \in J}$ that does not contain any x_n. Hence $(x_n)_{n=1}^\infty$ cannot converge to x. □

2.10 (§22) The Quotient Topology

Each injective function $f: X \to Y$ factors as a bijection $g: X \to B$ followed by an inclusion $i: B \subset Y$, where $B = f(X)$ is the image of f. When Y is a topological space, we defined the subspace topology on B to be the coarsest topology making $i: B \to Y$ continuous. We may also give X the unique topology making $g: X \to B$ a homeomorphism, and then $f: X \to Y$ is an embedding.

In this section we consider the more-or-less dual situation of a surjective function $f: X \to Y$. When X is a topological space we shall explain how to give Y the finest topology making f continuous, called the *quotient topology*.

2.10.1 Equivalence relations

Definition 2.10.1. A *relation* $\sim$ (read: "tilde") on a set X is a subset $R \subset X \times X$, where we write $x \sim y$ if and only of $(x, y) \in R$. An *equivalence relation* on X is a relation $\sim$ satisfying the three properties:

(1) $x \sim x$ for each $x \in X$.

(2) $x \sim y$ implies $y \sim x$ for $x, y \in X$.

(3) $x \sim y$ and $y \sim z$ implies $x \sim z$ for $x, y, z \in X$.

Definition 2.10.2. If $\sim$ is an equivalence relation on X, let

$$[x] = \{y \in X \mid x \sim y\}$$

be the *equivalence class* of $x \in X$. (This is a subset of X.) Note that $x \in [x]$ for each $x \in X$, and $[x] = [y]$ if and only if $x \sim y$. Hence the equivalence classes are nonempty subsets of X that cover X, and which are mutually disjoint. Let

$$X/{\sim} = \{[x] \mid x \in X\}$$

(read: "X mod tilde") be the set of equivalence classes. (This is a set of subsets of X.) The *canonical surjection*

$$\pi\colon X \to X/{\sim}$$

is given by $\pi(x) = [x]$.

Lemma 2.10.3. *Let $f\colon X \to Y$ be a surjective function. Define an equivalence relation $\sim$ on X by $x \sim y$ if and only if $f(x) = f(y)$. There is an induced bijection*

$$h\colon X/{\sim} \to Y$$

given by $h([x]) = f(x)$. Its inverse h^{-1} takes $y \in Y$ to the preimage $f^{-1}(y)$, which equals $[x]$ for any choice of $x \in f^{-1}(y)$. The surjection $f\colon X \to Y$ thus factors as the composite of the canonical surjection $\pi\colon X \to X/{\sim}$ and the bijection $h\colon X/{\sim} \to Y$.

Proof. The function h is well-defined since $[x] = [y]$ only if $x \sim y$, in which case $f(x) = f(y)$ by assumption. It is surjective since each element of Y has the form $f(x) = h([x])$ for some $x \in X$. It is injective since $h([x]) = h([y])$ implies $f(x) = f(y)$ so $x \sim y$ and $[x] = [y]$. □

In this way we can go back and forth between equivalence relations on X and surjective functions $X \to Y$, at least up to a bijection.

2.10.2 Quotient maps

Definition 2.10.4. Let $f\colon X \to Y$ be a surjective function, where X is a topological space. The *quotient topology* on Y (induced from X) is the collection of subsets $V \subset Y$ such that $f^{-1}(V)$ is open in X.

Lemma 2.10.5. *The quotient topology is a topology on Y.*

Proof. (1): $f^{-1}(\varnothing) = \varnothing$ and $f^{-1}(Y) = X$ are both open in X, so $\varnothing$ and Y are open in the quotient topology on Y.

(2): If $\{V_\alpha\}_{\alpha\in J}$ is a collection of open subsets of Y then each $f^{-1}(V_\alpha)$ is open in X, so

$$f^{-1}\Big(\bigcup_{\alpha\in J} V_\alpha\Big) = \bigcup_{\alpha\in J} f^{-1}(V_\alpha)$$

is a union of open sets in X hence is open in X, so $\bigcup_{\alpha\in J} V_\alpha$ is open in the quotient topology on Y.

(3): If $\{V_1, \dots, V_n\}$ is a finite collection of open subsets of Y then each $f^{-1}(V_i)$ is open in X, so

$$f^{-1}(V_1 \cap \dots \cap V_n) = f^{-1}(V_1) \cap \dots \cap f^{-1}(V_n)$$

is a finite intersection of open sets in X hence is open in X, so $V_1 \cap \dots \cap V_n$ is open in the quotient topology on Y. □

Definition 2.10.6. A surjective function $f\colon X \to Y$ between topological spaces is called a *quotient map* if Y has the quotient topology from X, i.e., if $V \subset Y$ is open if and only if $f^{-1}(V)$ is open in X.

Lemma 2.10.7. *A quotient map $f\colon X \to Y$ is continuous.*

Proof. When Y has the quotient topology from X, $V \subset Y$ is open (if and) only if $f^{-1}(V)$ is open in X. In particular f is continuous. □

Lemma 2.10.8. *Let $f\colon X \to Y$ be a surjective function, where X is a topological space. The quotient topology is the finest topology on Y such that $f\colon X \to Y$ is continuous.*

Proof. Let $\mathscr{T}$ be a topology on Y such that $f\colon X \to Y$ is continuous. Then for each $V \in \mathscr{T}$ we have that $f^{-1}(V)$ is open in X, so V is in the quotient topology. Hence $\mathscr{T}$ is coarser than the quotient topology. □

Lemma 2.10.9. *A surjective function $f\colon X \to Y$ is a quotient map if and only if the following condition holds: a subset $L \subset Y$ is closed if and only if $f^{-1}(L)$ is closed in X.*

Proof. Let $V = Y - L$. Then L is closed if and only if V is open, and $f^{-1}(L)$ is closed if and only if

$$X - f^{-1}(L) = f^{-1}(Y - L) = f^{-1}(V)$$

is open. □

Lemma 2.10.10. *A bijective quotient map $f\colon X \to Y$ is a homeomorphism, and conversely.*

2.10.3 Open and closed maps

Definition 2.10.11. Let $f\colon X \to Y$ be a function.

We say that f is an *open function* if for each open $U \subset X$ the image $f(U)$ is open in Y. If f is also continuous, we say that f is an *open map*.

We say that f is a *closed function* if for each closed $K \subset X$ the image $f(K)$ is closed in Y. If f is also continuous, we say that f is a *closed map*.

Lemma 2.10.12. *Let $\mathscr{B}$ be a basis for a topology on X. A map $f\colon X \to Y$ is open if and only if $f(B)$ is open in Y for each basis element $B \in \mathscr{B}$.*

Proof. Each basis element is open in X, so if f is an open map then $f(B)$ will be open in Y.

Conversely, any open subset $U \subset X$ is a union of basis elements $U = \bigcup_{\alpha \in J} B_\alpha$, so if each $f(B_\alpha)$ is open in Y then

$$f(U) = f(\bigcup_{\alpha \in J} B_\alpha) = \bigcup_{\alpha \in J} f(B_\alpha)$$

is open. □

Lemma 2.10.13. *(1) Each surjective, open map $f\colon X \to Y$ is a quotient map.*

(2) Each surjective, closed map $f\colon X \to Y$ is a quotient map.

Proof. (1): Let $V \subset Y$. If V is open then $f^{-1}(V)$ is open since f is continuous. Conversely, if $f^{-1}(V)$ is open in X then

$$V = f(f^{-1}(V))$$

because f is surjective, and this is an open subset of Y since f is an open map.

(2): Let $L \subset Y$. If L is closed then $f^{-1}(L)$ is closed since f is continuous. Conversely, if $f^{-1}(L)$ is closed in X then

$$L = f(f^{-1}(L))$$

because f is surjective, and this is a closed subset of Y since f is a closed map. □

Example 2.10.14. Let $X = [0,1]$ in the subspace topology from $\mathbb{R}$, and let $Y = S^1$ be the *circle* in the subspace topology from $\mathbb{R}^2$. Let

$$f\colon [0,1] \to S^1$$

be given by $f(t) = (\cos(2\pi t), \sin(2\pi t))$. It is clearly continuous and surjective. It can also be shown to be closed (most easily using compactness, later), hence is a quotient map. It is not open, since $U = [0, 1/2)$ is open in $[0,1]$, but the image $f(U)$ is not open in S^1.

Define an equivalence relation $\sim$ on $[0,1]$ by $s \sim t$ if and only if $f(s) = f(t)$. The equivalence classes for this relation are $[0] = [1] = \{0,1\}$ and $[t] = \{t\}$ for $0 < t < 1$. We get an induced bijection

$$h\colon [0,1]/{\sim} \to S^1$$

and $f = h \circ \pi$. (We might also write $0 \sim 1$ for this equivalence relation.) Then h is a homeomorphism from $[0,1]/{\sim}$ with the quotient topology from $[0,1]$ to S^1 with the quotient topology from $[0,1]$, which equals the subspace topology from $\mathbb{R}^2$.

Example 2.10.15. Let $X = [0,1] \times [0,1]$ be a square in the subspace topology from $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$, and let $Y = S^1 \times S^1$ in the product topology, which equals the subspace topology from $\mathbb{R}^2 \times \mathbb{R}^2 = \mathbb{R}^4$. We call Y a *torus*. Let

$$g\colon [0,1] \times [0,1] \to S^1 \times S^1$$

be given by $g(s,t) = (\cos(2\pi s), \sin(2\pi s), \cos(2\pi t), \sin(2\pi t))$. This is the product $f \times f$ of two copies of the quotient map $f\colon [0,1] \to S^1$ discussed above.

The function g is continuous and surjective. It is also closed, hence a quotient map, and induces a homeomorphism

$$h\colon ([0,1] \times [0,1])/{\sim} \to S^1 \times S^1$$

where $\sim$ is the equivalence relation given by $(s,t) \sim (s',t')$ precisely if $g(s,t) = g(s',t')$. The equivalence classes of this relation are the 4-element set

$$\{(0,0), (0,1), (1,0), (1,1)\},$$

the 2-element sets

$$\{(s,0), (s,1)\} \quad \text{and} \quad \{(0,t), (1,t)\},$$

and the singleton sets (s,t), for $0 < s,t < 1$

In this way the torus $S^1 \times S^1$ is realized, up to homeomorphism, as a quotient space of the square $[0,1] \times [0,1]$, with respect to the equivalence relation $\sim$ generated by the relations $(s,0) \sim (s,1)$ and $(0,t) \sim (1,t)$ for all $s,t \in [0,1]$.

Example 2.10.16. Let

$$D^n = \{x \in \mathbb{R}^n \mid \|x\| \le 1\}$$

the *unit disc* in $\mathbb{R}^n$. Its boundary $\partial D^n = \mathrm{Cl}(D^n) - \mathrm{Int}(D^n) = S^{n-1}$ is the unit sphere. The stereographic projection $f\colon S^n - \{N\} \to \mathbb{R}^n$ restricts to a homeomorphism

$$f_-\colon D^n_- \xrightarrow{\cong} D^n$$

where

$$D^n_- = \{x = (x_1, \dots, x_n, y) \in S^n \mid y \le 0\}$$

is the *lower hemisphere* of S^n. Similarly,

$$D^n_+ = \{x = (x_1, \dots, x_n, y) \in S^n \mid y \ge 0\}$$

is the *upper hemisphere*. We can map

$$q\colon D^n_- \longrightarrow S^n$$

by

$$q(x_1, \dots, x_n, y) = (sx_1, \dots, sx_n, 2y + 1)\,,$$

with $s \ge 0$ characterized by $s^2(x_1^2 + \cdots + x_n^2) + (2y+1)^2 = 1$. Then q is continuous, identifies $S^{n-1} \subset D^n_-$ to $N \in S^n$, and is otherwise bijective. Using compactness we can show that q is closed, hence a quotient map. We obtain homeomorphisms

$$D^n/S^{n-1} \xleftarrow{\cong} D^n_-/S^{n-1} \xrightarrow{\cong} S^n\,.$$

Example 2.10.17. Let $X = \mathbb{R}$ and $Y = \{n, z, p\}$. Define $f\colon X \to Y$ by

$$f(x) = \begin{cases} n & \text{if } x < 0, \\ z & \text{if } x = 0, \\ p & \text{if } x > 0. \end{cases}$$

The quotient topology on Y from $\mathbb{R}$ is the collection

$$\{\varnothing, \{n\}, \{p\}, \{n, p\}, Y\}\,.$$

It is a non-Hausdorff topology, where z is the only closed point.

Example 2.10.18. Let $X = \mathbb{R} \times \mathbb{R}$ and $Y = \mathbb{R}$. Let $\pi_1\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be the projection $\pi_1(x, y) = x$. Then π_1 is continuous and surjective. It is an open map, since for each basis element $B = U \times V \subset \mathbb{R} \times \mathbb{R}$ for the product topology, with $U, V \subset \mathbb{R}$ open, the image $\pi_1(B) = U$ is open in $\mathbb{R}$. Hence π_1 is a quotient map.

It is not a closed map, since the hyperbola

$$C = \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid xy = 1\}$$

is a closed subset of $\mathbb{R} \times \mathbb{R}$, but $\pi_1(C) = \mathbb{R} - \{0\}$ is not closed in $\mathbb{R}$. To see that C is closed, use the continuous function $m\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ taking (x, y) to xy, and note that $C = m^{-1}(1)$ is the preimage of the closed point $\{1\}$.

Example 2.10.19. Let $A = C \cup \{(0, 0)\}$ be a subspace of $X = \mathbb{R}^2$, and consider the restricted map $f = \pi_1|A\colon A \to \mathbb{R}$. It is continuous and surjective (even bijective), but not a quotient map. For $\{0\}$ is not open in $\mathbb{R}$, but its preimage $f^{-1}(0) = \{(0, 0)\}$ is open in the subspace topology on A.

Theorem 2.10.20. *Let $f\colon X \to Y$ be a quotient map, let $B \subset Y$ and $A = f^{-1}(B) \subset X$ be subspaces, and let $g\colon A \to B$ be the restricted map.*

(1) If A is open (or closed) in X, then g is a quotient map.

(2) If f is an open map (or a closed map), then g is a quotient map.

Proof. Consider the diagram

$$\begin{array}{ccccccc} g^{-1}(V) & \rightarrowtail & A & \rightarrowtail & X & \leftarrowtail & U \\ & & \downarrow{\scriptstyle g} & & \downarrow{\scriptstyle f} & & \\ V & \rightarrowtail & B & \rightarrowtail & Y & \leftarrowtail & f(U) \end{array}$$

We first prove the identities

$$g^{-1}(V) = f^{-1}(V)$$

for $V \subset B$, and

$$g(A \cap U) = B \cap f(U)$$

for $U \subset X$. It is clear that $g^{-1}(V) \subset f^{-1}(V)$. Conversely, if $x \in f^{-1}(V)$ then $f(x) \in V \subset B$, so $x \in f^{-1}(B) = A$, hence $x \in g^{-1}(V)$. It is also clear that $g(A \cap U) = f(A \cap U) \subset B \cap f(U)$. Conversely, if $y \in B \cap f(U)$ then there is an $x \in U$ with $f(x) = y$. Since $f(x) = y \in B$ we get $x \in f^{-1}(B) = A$, so $x \in A \cap U$. Hence $y \in f(A \cap U)$.

Next, suppose that A is open. Given $V \subset B$ with $g^{-1}(V)$ open in A we want to prove that V is open in B. Since A is assumed to be open in X we know that $g^{-1}(V) = f^{-1}(V)$ is open in X. Since f is a quotient map, V is open in Y. Hence $V = B \cap V$ is open in B. (Same argument for A closed.)

Finally, suppose that f is an open map. Given $V \subset B$ with $g^{-1}(V)$ open in A we want to prove that V is open in B. There is an open subset $U \subset X$ with $g^{-1}(V) = A \cap U$. Then

$$g(g^{-1}(V)) = g(A \cap U) = B \cap f(U)\,.$$

Here $V = g(g^{-1}(V))$ since g is surjective, and $f(U)$ is open in Y since f is assumed to be an open map. Hence $V = B \cap f(U)$ is open in B. (Same argument for f a closed map.) □

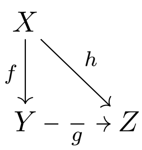
Remark 2.10.21. The composite of two quotient maps $f\colon X \to Y$ and $g\colon Y \to Z$ is a quotient map $gf\colon X \to Z$.

The product of two quotient maps is in general not a quotient map. Some condition like local compactness is usually needed.

The image of a Hausdorff space X under a quotient map $f\colon X \to Y$ needs not be a Hausdorff space.

The quotient topology has a universal mapping property, somewhat dual to that of the subspace and product topologies.

Theorem 2.10.22. *Let $f\colon X \to Y$ be a quotient map, and Z any topological space. Let $h\colon X \to Z$ be a function such that $h(x) = h(y)$ whenever $f(x) = f(y)$. Then h induces a unique function $g\colon Y \to Z$ with $h = g \circ f$.*



The induced function g is continuous if and only if h is continuous. Furthermore, g is a quotient map if and only if h is a quotient map.

Proof. Each element $y \in Y$ has the form $y = f(x)$ for $x \in X$, since f is surjective, so we can (and must) define $g(y) = h(x)$.

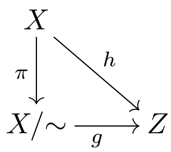
If g is continuous, then so is the composite $h = g \circ f$. Conversely, suppose that h is continuous. To prove that g is continuous, let $V \subset Z$ be open. To prove that $g^{-1}(V)$ is open in Y it suffices to show that $f^{-1}(g^{-1}(V))$ is open in X, since f is a quotient map. But $f^{-1}(g^{-1}(V)) = h^{-1}(V)$, and $h^{-1}(V)$ is open in X by the assumption that h is continuous.

If g is a quotient map, then so is the composite $h = g \circ f$. Conversely, suppose that h is a quotient map. Then g is surjective, since any $z \in Z$ has the form $z = h(x) = g(f(x))$ for some $x \in X$. Lastly, let $V \subset Z$ and suppose that $g^{-1}(V)$ is open. We must show that V is open. Now $f^{-1}(g^{-1}(V)) = h^{-1}(V)$ is open, since f is continuous. Hence V is open, by the assumption that h is a quotient map. □

Corollary 2.10.23. *Let $h\colon X \to Z$ be a surjective, continuous map. Let $\sim$ be the equivalence relation on X given by $x \sim y$ if and only if $h(x) = h(y)$, and let $X/{\sim}$ be the set of equivalence classes:*

$$X/{\sim} = \{h^{-1}(z) \mid z \in Z\}$$

Give $X/{\sim}$ the quotient topology from the canonical surjection $\pi\colon X \to X/{\sim}$.



(1) The map h induces a bijective, continuous map $g\colon X/{\sim} \to Z$.

(2) The map $g\colon X/{\sim} \to Z$ is a homeomorphism if and only if h is a quotient map.

(3) If Z is Hausdorff, then so is $X/{\sim}$.

Chapter 3

Connectedness and Compactness

3.1 (§23) Connected Spaces

3.1.1 Sums of spaces

Definition 3.1.1. Let C and D be topological spaces. Assume that $C \cap D = \varnothing$. We then write $C \sqcup D$ for the *disjoint union* $C \cup D$. There are canonical inclusions

$$i_C \colon C \to C \sqcup D \qquad \text{and} \qquad i_D \colon D \to C \sqcup D\,.$$

The *sum topology* on $C \sqcup D$ is the collection of subsets $W \subset C \sqcup D$ such that $C \cap W$ is open in C and $D \cap W$ is open in D. It is the finest topology on $C \sqcup D$ for which both i_C and i_D are continuous.

Remark 3.1.2. The disjoint union $C \sqcup D$ is also known as the *coproduct* of the two spaces C and D. It has a universal property dual to that of the product $C \times D$.

Each space X is homeomorphic to a disjoint union $C \sqcup D$ in some trivial ways, if $C = \varnothing$ or $D = \varnothing$. If $X \cong C \sqcup D$ in a non-trivial way, with both C and D nonempty, then we say that X is *disconnected.* Otherwise, X is a *connected* space.

3.1.2 Separations

Definition 3.1.3. Let X be a topological space. A *separation* of X is a pair U, V of disjoint, nonempty open subsets of X whose union is X. The space X is said to be *connected* (norsk: sammenhengende) if there does not exist a separation of X. Otherwise it is *disconnected.*

Remark 3.1.4. Being connected is a topological property. The empty space $X = \varnothing$ may require special care. Some authors make an exception, and say that it is not connected. The situation is similar to that of prime factorization, where the unit 1 has no proper factors in natural numbers, but is still not counted as a prime.

Lemma 3.1.5. *A space X is connected if and only if the only subsets of X that are both open and closed are $\varnothing$ and X itself.*

Proof. In a separation U, V of X, since $U \cap V = \varnothing$ and $U \cup V = X$ we must have $V = X - U$. Asking that U and V are open is equivalent to asking that U is both open and closed. Asking that both U and V are nonempty is equivalent to asking that U is different from $\varnothing$ and X. □

Lemma 3.1.6. *If U, V is a separation of X, then $X \cong U \sqcup V$ is homeomorphic to the disjoint union of the subspaces U and V.*

Proof. If $W \subset X$ is open then $U \cap W$ and $V \cap W$ are open in the subspace topologies on U and V, respectively, so W is open in $U \sqcup V$. Conversely, if W is open in $U \sqcup V$, then $U \cap W$ and $V \cap W$ are open in U and V, respectively, hence are open in X, since U and V are open in X. Thus

$$W = (U \cap W) \cup (V \cap W)$$

is a union of open sets, hence is open in X. □

Lemma 3.1.7. *If $X = C \sqcup D$ with C and D nonempty, then C, D is a separation of X.*

Proof. It is clear that C and D are disjoint, nonempty subsets of X whose union equals X. To see that C is open in the disjoint union topology, not that $C \cap C = C$ is open in C and $D \cap C = \varnothing$ is open in D. Similarly, D is open in the disjoint union topology. □

Lemma 3.1.8. *A separation of a topological space X is a pair of disjoint, nonempty subsets A and B whose union is X, neither of which contains any limit points of the other. (In symbols, $A \cap B' = \varnothing$ and $A' \cap B = \varnothing$.)*

Proof. If A and B form a separation of X, then A is closed, so $A' \subset \bar{A} = A$ does not meet B. Similarly, B' does not meet A.

Conversely, suppose that A and B are disjoint, nonempty sets with union X, $A' \cap B = \varnothing$ and $A \cap B' = \varnothing$. Then $\bar{A} \cap B = \varnothing$, since $\bar{A} = A \cup A'$, so $\bar{A} \subset X - B = A$, hence $A = \bar{A}$ is closed and B is open. Likewise, B is closed and A is open. Hence A and B form a separation of X. □

Example 3.1.9. Each 1-point space $X = \{a\}$ is connected, since there are no proper, nonempty subsets.

Example 3.1.10. Let $X = \{a, b\}$ with the Sierpinski topology $\mathscr{T}_a = \{\varnothing, \{a\}, X\}$. The proper, nonempty subsets of X are $\{a\}$ and $\{b\}$, where the first is open and the second is closed, but neither is both open and closed. Hence X is connected.

Example 3.1.11. Let $X = [-1, 0) \cup (0, 1]$ be a subspace of $\mathbb{R}$. Then $U = [-1, 0)$ and $V = (0, 1]$ is a separation of X, so X is disconnected.

Remark 3.1.12. We shall prove in the next section that $\mathbb{R}$ is connected, as is each interval $[a, b]$, $[a, b)$, $(a, b]$ and (a, b) for $-\infty \le a \le b \le \infty$.

Example 3.1.13. Each subspace $X \subset \mathbb{Q}$ with at least 2 elements is disconnected: If $p < q \in X$ choose an irrational $a \in (p, q)$. Then $U = X \cap (-\infty, a)$ and $V = X \cap (a, \infty)$ is a separation of X. We say that $\mathbb{Q}$ is *totally disconnected.*

3.1.3 Constructions with connected spaces

Lemma 3.1.14. *If U and V form a separation of X, and A is a connected subspace, then $A \subset U$ or $A \subset V$.*

Proof. The intersection $A \cap U$ is open and closed in A. Since A is connected, $A \cap U$ is empty or all of A. In the first case, $A \subset V$. In the second case, $A \subset U$. □

Theorem 3.1.15. *The union of a collection of connected subspaces of X, that all have a point in common, is connected.*

Proof. Let $\{A_\alpha\}_{\alpha\in J}$ be a collection of connected subspaces of X, let $p \in \bigcap_{\alpha\in J} A_\alpha$, and let $Y = \bigcup_{\alpha\in J} A_\alpha$. To show that Y is connected, suppose that $Y = U \cup V$ is a separation of Y. Then $p \in U$ or $p \in V$. Suppose, without loss of generality, that $p \in U$. For each $\alpha \in J$ the connected space A_α is contained in U or in V. Since $p \in A_\alpha$ and $p \notin V$ we must have $A_\alpha \subset U$. This holds for each α, hence $Y \subset U$. This contradicts the assumption that V is nonempty. □

Theorem 3.1.16. *The product of two connected spaces is connected.*

Proof. Let X and Y be connected. By hypothesis $X \times \{y\} \cong X$ and $\{x\} \times Y \cong Y$ are connected, for all $x \in X$ and $y \in Y$. Choose $a \in X$ and $b \in Y$. For each $y \in Y$ the space

$$C_y = X \times \{y\} \cup \{a\} \times Y$$

is a union of two connected spaces with a common point, namely (a, y), hence is connected. It follows that

$$X \times Y = \bigcup_y C_y$$

is also a union of connected spaces with a common point, namely (a, b), hence is connected. □

Corollary 3.1.17. *Any finite product $X_1 \times \cdots \times X_n$ of connected spaces is connected.*

Proof. This is clear for $n = 1$, and we have just proved it for $n = 2$. The general case follows by induction on n, using the homeomorphism

$$X_1 \times \cdots \times X_n \cong (X_1 \times \cdots \times X_{n-1}) \times X_n\,.$$

□

Theorem 3.1.18. *Let $A \subset B \subset \bar{A}$ be subspaces of X. If A is connected then B is connected.*

Proof. Suppose that $B = U \cup V$ is a separation. Since A is connected, we have $A \subset U$ or $A \subset V$. Without loss of generality assume that $A \subset U$. Then $B \subset \bar{A} \subset \bar{U}$. Since U is closed in B, it equals its closure $B \cap \bar{U}$ in B. Combining $B \subset \bar{U}$ and $U = B \cap \bar{U}$ we deduce that $B \subset U$. This contradicts the assumption that V is nonempty. □

Theorem 3.1.19. *The continuous image of a connected space is connected.*

Proof. Let $f\colon X \to Y$ be continuous, with X connected. We prove that the image space $Z = f(X)$ is connected. Suppose that $Z = U \cup V$ is a separation. Then $f^{-1}(U)$ and $f^{-1}(V)$ are disjoint, nonempty open subsets of X whose union equals X. This contradicts the assumption that X is connected. □

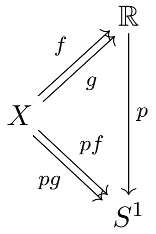
Lemma 3.1.20. *Any map $f\colon X \to Y$ from a connected space X to a totally disconnected space Y (for example, a discrete space) is constant.*

Proof. Since X is connected and f is continuous, its image $f(X)$ is a connected subspace of Y, hence consists of at most one point. This means that f is constant. □

Example 3.1.21. Consider the *covering map*

$$\begin{aligned} p\colon \mathbb{R} &\longrightarrow S^1 \\ t &\longmapsto (\cos(2\pi t), \sin(2\pi t))\,. \end{aligned}$$

If X is connected and $f, g\colon X \to \mathbb{R}$ are two maps such that $p \circ f, p \circ g\colon X \to S^1$ are equal, then either $f(x) = g(x)$ for all $x \in X$, or $f(x) \neq g(x)$ for all $x \in X$.



To see this, note that $h(x) = g(x) - f(x)$ defines a map $h\colon X \to \mathbb{Z}$, where $\mathbb{Z} \subset \mathbb{R}$ is discrete. Hence h is constant. If $f(x) = g(x)$ for some $x \in X$, then $h = 0$ everywhere and $f(x) = g(x)$ for all $x \in X$. Otherwise $f(x) \neq g(x)$ for all $x \in X$. We will return to covering spaces/maps in §53.

3.2 (§24) Connected Subspaces of the Real Line

We now use the existence of least upper bounds for nonempty, bounded subsets of $\mathbb{R}$ to prove that $\mathbb{R}$ is connected.

Definition 3.2.1. A subset $C \subset \mathbb{R}$ is *convex* if for any two points $a < b$ in C the closed interval $[a, b]$ is a subset of Y.

Example 3.2.2. The convex subsets of $\mathbb{R}$ are $\varnothing$, the intervals (a, b), $[a, b)$, $(a, b]$ and $[a, b]$, the rays $(-\infty, b)$, $(-\infty, b]$, (a, ∞) and $[a, \infty)$, and $\mathbb{R}$ itself.

Theorem 3.2.3. *Each convex subset $C \subset \mathbb{R}$ is connected.*

Proof. Suppose that U, V is a separation of C. Choose $a \in U$ and $b \in V$. By symmetry we may assume that $a < b$. Let $A = [a, b] \cap U$ and $B = [a, b] \cap V$. Then A, B is a separation of $[a, b]$, with $a \in A$ and $b \in B$. Let $c = \sup A$ be the least upper bound of the elements in A. Clearly $a \leq c \leq b$, since $a \in A$ and b is an upper bound for A.

If $c = a$ then $A = \{a\}$, contradicting the assumption that A is open in $[a, b]$.

If $c = b$ then there are points $x \in A$ arbitrarily close to b. Since A is closed we must have $b \in A$, contradicting the assertion that $A \cap B = \varnothing$.

Otherwise $a < c < b$. Then $(c, b] \subset B$, since c is an upper bound for A. Hence there are points $y \in (c, b] \subset B$ arbitrarily close to c. Since B is closed, we must have $c \in B$. Since B is open, there is an $\epsilon > 0$ with $(c - \epsilon, c + \epsilon) \subset B$. Then any $x \in (c - \epsilon, c)$ is also an upper bound for A, contradicting the definition of c as the least upper bound.

Each of the three cases leads to a contradiction. Hence no separation of C exists, and C is connected. □

Theorem 3.2.4 (Intermediate value theorem)**.** *Let $f\colon X \to \mathbb{R}$ be a continuous map, where X is a connected space. If $a, b \in X$ are points, and $r \in \mathbb{R}$ lies between $f(a)$ and $f(b)$, then there exists a point $c \in X$ with $f(c) = r$.*

Proof. Suppose that $f(X) \subset \mathbb{R} - \{r\} = (-\infty, r) \cup (r, \infty)$. Then X is the union of the disjoint, nonempty subsets $U = f^{-1}((-\infty, r))$ and $V = f^{-1}((r, \infty))$, each of which is open in X. This contradicts the assumption that X is connected. □

3.2.1 Path connected spaces

Definition 3.2.5. Given points $x, y \in X$ a *path* in X from x to y is a map $f\colon [a,b] \to X$ with $f(a) = x$ and $f(b) = y$, where $[a,b] \subset \mathbb{R}$.

A space X is *path connected* (norsk: veisammenhengende) if for any two points x and y of X there exists a path in X from x to y.

Lemma 3.2.6. *A path connected space is connected.*

Proof. Let X be path connected, and suppose that U and V separate X. Choose points $x \in U$ and $y \in V$, and a path $f\colon [a,b] \to X$ in X from x to y. Then $f^{-1}(U)$ and $f^{-1}(V)$ form a separation of the connected space $[a,b]$, which is impossible. □

Lemma 3.2.7. *The continuous image of a path connected space is path connected.*

Proof. If $g\colon X \to Y$ is a map, any two points in $f(X)$ can be written as $g(x)$ and $g(y)$ for $x, y \in X$. Since X is path connected, there is a path $f\colon [a,b] \to X$ in X from x to y. Then $g \circ f\colon [a,b] \to Y$ is a path in $f(X)$ from $f(x)$ to $f(y)$. Hence $f(X)$ is path connected. □

Definition 3.2.8. A subset C of a real vector space V is *convex* if for each pair of points $x, y \in C$ the straight-line path $f\colon [0,1] \to V$ defined by

$$f(t) = (1-t)x + ty$$

takes all of its values in C.

Example 3.2.9. Any convex subset of $\mathbb{R}^n$ is path connected, since the path f is continuous. For example, the n-dimensional unit ball

$$B^n = \{x \in \mathbb{R}^n \mid \|x\| \le 1\}$$

is convex for $n \ge 0$, hence path connected.

Example 3.2.10. The punctured Euclidean space $\mathbb{R}^n - \{0\}$ is path connected for $n \ge 2$. For $n = 1$, the space $\mathbb{R}^1 - \{0\}$ is not (path) connected. For $n = 0$ the space $\mathbb{R}^0 - \{0\}$ is empty, hence is path connected by convention.

Example 3.2.11. The $(n-1)$-dimensional unit sphere

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

is the continuous image of $g\colon \mathbb{R}^n - \{0\} \to S^{n-1}$ given by $g(x) = x/\|x\|$, hence is path connected for $n \ge 2$. For $n = 1$, the 0-sphere $S^0 = \{+1, -1\}$ is not path connected. For $n = 0$, the (-1)-sphere S^{-1} is empty, hence is also path connected, by convention.

Example 3.2.12. Let

$$S = \{(x, \sin(1/x)) \mid 0 < x \le 1\}$$

be a subset of $\mathbb{R}^2$. It is the image of the connected space $(0,1]$ under the continuous map $g(t) = (t, \sin(1/t))$, hence is connected. Therefore its closure $\bar{S}$ in $\mathbb{R}^2$ is connected. This closure

$$\bar{S} = S \cup V$$

is the union of S with the vertical interval $V = \{0\} \times [-1,1]$. The space $\bar{S}$ is called the *topologist's sine curve*.

We show that $\bar{S}$ is not path connected. Suppose that $f\colon [a,b] \to \bar{S}$ is a map with $f(a) \in V$ and $f(b) \in S$. The set of $t \in [a,b]$ with $f(t) \in V$ is closed, since V is closed in $\bar{S}$, hence has a greatest element c. The restricted function $f|[c,b]\colon [c,b] \to \bar{S}$ is then a path in $\bar{S}$ starting in V and ending in S. By reparametrizing, we may replace $[c,b]$ by $[0,1]$. We then have a map $f\colon [0,1] \to \bar{S}$ with $f(0) \in V$ and $f(t) \in S$ for all $t \in (0,1]$.

Write $f(t) = (x(t), y(t))$. Then $x(0) = 0$ and $x(t) > 0$ for all $t \in (0,1]$. We show that there is a sequence of points $t_n \in (0,1]$ with $0 < t_n < 1/n$ and $y(t_n) = (-1)^n$. Then $t_n \to 0$ as $n \to \infty$, but the sequence $(y(t_n))_{n=1}^{\infty} = ((-1)^n)_{n=1}^{\infty}$ does not converge. This contradicts the continuity of y.

For each n we choose a $v > 1/x(1/n)$ such that $\sin(v) = (-1)^n$. Let $u = 1/v$, then $0 = x(0) < u < x(1/n)$ and $\sin(1/u) = (-1)^n$. By the intermediate value theorem for x, there is a $t_n \in (0, 1/n)$ with $x(t_n) = u$. Then $y(t_n) = (-1)^n$, as desired.

3.3 (§25) Components and Local Connectedness

Definition 3.3.1. Define an equivalence relation $\sim$ for points in a topological space X by $x \sim y$ if there is a connected subset $C \subset X$ with $x, y \in C$. The equivalence classes for $\sim$ are called the *(connected) components* of X.

Lemma 3.3.2. *$\sim$ is an equivalence relation on X.*

Theorem 3.3.3. *The components of X are connected, disjoint subspaces whose union is X. Each nonempty, connected subset of X is contained in precisely one component.*

Proof. It is clear that the components are disjoint, nonempty subspaces whose union is X, since the components are the equivalence classes for an equivalence relation. If $A \subset X$ is connected and meets two components C_1 and C_2, in points x_1 and x_2, say, then $x_1 \sim x_2$, so $C_1 = C_2$.

To show that each component C is connected, let $x_0 \in C$. For each $x \in C$ we have $x_0 \sim x$, so there exists a connected subset A_x with $x_0, x \in A_x$. Then $A_x \subset C$, so $\bigcup_{x\in C} A_x = C$. Since all A_x are connected, and all contain x_0, it follows that the union is connected. □

Using the notion of connected components we can simplify the proof of the following theorem (which we already proved as Theorem 3.1.16).

Theorem 3.3.4. *A finite product of connected spaces is connected.*

Proof. Let X and Y be connected spaces. We show that any two points (x,y) and (x',y') in $X \times Y$ are in the same component. First, $X \times \{y\}$ is homeomorphic to X, hence is connected. So (x,y) and (x',y) are in the same component. Next, $\{x'\} \times Y$ is homeomorphic to Y, hence is connected. So (x',y) and (x',y') are in the same component. The claim follows by associativity of $\sim$.

Given n connected spaces $X_1, \dots, X_n$, the homeomorphism

$$X_1 \times \cdots \times X_n \cong (X_1 \times \cdots \times X_{n-1}) \times X_n$$

and induction on n shows that $X_1 \times \cdots \times X_n$ is connected. □

Remark 3.3.5. In fact, an arbitrary product of connected spaces is connected, in the product topology.

3.3.1 Path components

Definition 3.3.6. Define another equivalence relation $\simeq$ for points in a topological space X by $x \simeq y$ if there is a path in X from x to y. The equivalence classes for $\simeq$ are called the *path components* of X.

Lemma 3.3.7. *$\simeq$ is an equivalence relation on X.*

Theorem 3.3.8. *The path components of X are path connected, disjoint subspaces whose union is X. Each nonempty, path connected subset of X is contained in precisely one path component.*

Example 3.3.9. The topologist's sine curve $\bar{S} = S \cup V$ is connected, hence consist of only one component. It has two path components, S and V. Note that S is open in $\bar{S}$, but not closed, and V is closed in $\bar{S}$, but not open.

3.3.2 Locally connected spaces

Definition 3.3.10. A space X is *locally connected at a point $x \in X$* if for each neighborhood U of x there is a connected neighborhood V of x contained in U:

$$x \in V \subset U \subset X$$

We say that X is *locally connected* if it is locally connected at each of its points.

Definition 3.3.11. A space X is *locally path connected at a point $x \in X$* if for each neighborhood U of x there is a path connected neighborhood V of x contained in U:

$$x \in V \subset U \subset X$$

We say that X is *locally path connected* if it is locally path connected at each of its points.

Example 3.3.12. The real line is locally (path) connected, since each neighborhood U of any point x contains a (path) connected basis neighborhood $(x - \epsilon, x + \epsilon)$ for some $\epsilon > 0$.

Example 3.3.13. Any open subset $\Omega \subset \mathbb{R}^n$ is locally (path) connected, since each neighborhood U of any point $x \in \Omega$ contains a (path) connected basis neighborhood $B_d(x, \epsilon)$ for some $\epsilon > 0$.

Example 3.3.14. The topologist's sine curve is not locally (path) connected, since small neighborhoods of points in V are not connected.

Theorem 3.3.15. *If X is locally connected, then each component C of X is open in X.*

Proof. Let $x \in C$. Then $U = X$ is a neighborhood of x, so by local connectivity there is a connected neighborhood V of x with $V \subset U = X$. Since V is connected, $V \subset C$. Hence C contains a neighborhood around each of its points, and must be open. □

Theorem 3.3.16. *If X is locally path connected, then each path component P of X is open in X.*

Theorem 3.3.17. *If X is a topological space, each path component of X lies in a unique component of X. If X is locally path connected, then the components and the path components of X are the same.*

Proof. Each path component P is nonempty and connected, hence lies in a unique component C. If X is locally path connected we show that $P = C$.

Let U be the union of the path components Q of X that are different from P and meet C. Since each such path component Q is connected, it lies in C, so that

$$C = P \cup U$$

is a disjoint union. Because X is locally path connected, each path component P or Q is open in X, hence so is the union U. Hence P is a nonempty, open and closed subset of C. Since C is connected, it follows that $P = C$. □

3.3.3 The Jordan curve theorem

The line segment

$$[0, 1] \times \{0\} \subset \mathbb{R}^2$$

can be parametrized as an arc = simple curve. The complement $\mathbb{R}^2 - [0, 1] \times \{0\}$ is connected. The circle

$$S^1 \subset \mathbb{R}^2$$

can be parametrized as a simple closed curve. The complement $\mathbb{R}^2 - S^1$ has two connected components. The following is proved in §61.

Theorem 3.3.18 (Jordan curve theorem)**.** *Let $C \subset \mathbb{R}^2$ be a simple closed curve, i.e., a subspace homeomorphic to S^1. Then $\mathbb{R}^2 - C$ has exactly two connected (path) components.*

Here C is closed in $\mathbb{R}^2$ (this will follow from compactness), so $\Omega = \mathbb{R}^2 - C$ is open, hence locally path connected. Thus the components and path components of Ω are the same, and the theorem asserts that there are precisely two components: one bounded and one unbounded (containing a 'neighborhood of infinity').

Theorem 3.3.19 (Arc theorem)**.** *Let $A \subset \mathbb{R}^2$ be an arc, i.e., a subspace homeomorphic to $[0, 1]$. Then $\mathbb{R}^2 - A$ is (path) connected.*

This is proved in §63. Note that it is not sufficient that A admits a simple (injective) parametrization by a half-open interval. The 'closed topologist's sine curve' $C \subset \mathbb{R}^2$ is obtained from $\bar{S}$ by connecting $(1, \sin(1)) \in S$ to $(0, -1) \in V$. There is a continuous bijection $[0, \infty) \to C$. However, $\mathbb{R}^2 - C$ is not (path) connected.

3.3.4 Gaussian elimination

Let

$$M_n(\mathbb{R}) = \{A = (a_{i,j})_{i,j=1}^n \mid a_{i,j} \in \mathbb{R}\}$$

be the space of $n \times n$ real matrices. Viewing a matrix $A = (\alpha_1, \dots, \alpha_n)$ as an order n-tuple of column vectors in $\mathbb{R}^n$, we can identify $M_n(\mathbb{R})$ with the Euclidean space

$$(\mathbb{R}^n)^n = \mathbb{R}^n \times \cdots \times \mathbb{R}^n \cong \mathbb{R}^{n^2}$$

with the standard topology. The determinant function $\det\colon M_n(\mathbb{R}) \to \mathbb{R}$ is given by a (degree n) polynomial in the entries $(a_{i,j})_{i,j=1}^n$ of a matrix $A \in M_n(\mathbb{R})$, hence is continuous.

Let

$$GL_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) \mid \det(A) \neq 0\}$$

denote the space of $n \times n$ invertible real matrices. It is the preimage of the open subset $\mathbb{R} - \{0\}$ of $\mathbb{R}$ for the map det, hence is an open subspace of $M_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$. Since the image of $\det\colon GL_n(\mathbb{R}) \to \mathbb{R}$ is $\mathbb{R} - \{0\}$ (for $n \geq 1$), which is disconnected, it follows that $GL_n(\mathbb{R})$ is not connected.

Let $GL_n^+(\mathbb{R}) \subset GL_n(\mathbb{R})$ and $SL_n(\mathbb{R}) \subset GL_n(\mathbb{R})$ be the subspaces of matrices A with $\det(A) > 0$ and $\det(A) = 1$, respectively.

Proposition 3.3.20. *$GL_n^+(\mathbb{R})$ and $SL_n(\mathbb{R})$ are (path) connected.*

Proof. We use the method of row reduction to connect any given matrix A with $\det(A) > 0$ to the identity matrix I_n. For $1 \leq i \neq j \leq n$ and $r \in \mathbb{R}$ let

$$E_{i,j}(r) = \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & r & \dots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix}$$

be the $n \times n$ *elementary matrix*, which differs from I_n only in the (i,j)-th entry, which is equal to r. Note that left multiplication by $E_{i,j}(r)$, sends A to the matrix $E_{i,j}(r)A$ obtained by adding r times the j-th row of A to the i-th row of A. This is one of the three standard *row operations* used in Gaussian elimination. Note that

$$t \mapsto E_{i,j}(tr)A$$

for $t \in [0,1]$ defines a path in $GL_n(\mathbb{R})$ from A to $E_{i,j}(r)A$. Hence A and $E_{i,j}(r)A$ lie in the same path component of $GL_n(\mathbb{R})$.

Another standard row operation is given by interchanging two rows. Instead, we will use the operation given by interchanging two rows, and reversing the sign of one of the rows. In the case $n = 2$, this amounts to left multiplication by the matrix

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} .$$

However, in view of the factorization

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

this left multiplication can be achieved by a series of three row operations of the previous kind. A similar argument, acting only on two of n rows, shows that if B is obtained from A by interchanging two rows are reversing the sign of one of them, then A and B lie in the same path component of $GL_n(\mathbb{R})$.

Using these operations we can create a path that connects any invertible matrix A to a diagonal matrix of the form

$$D = \mathrm{diag}(d_1, d_2, \dots, d_n) = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{pmatrix}$$

where $d_1, \dots, d_n$ are all nonzero. Furthermore, if d_i and d_j are negative, for $i \neq j$, we can interchange the i-th and j-th rows twice, changing the signs each time, so as to replace d_i by $-d_i$ and d_j by $-d_j$. We may therefore arrange that at most one of the diagonal entries is negative. Here

$$\det(D) = d_1 d_2 \cdots d_n$$

so if $\det(D) > 0$ we have arranged that all of the diagonal entries are positive. In fact $\det(A) = \det(D)$, since the row operations considered so far do not alter the determinant. So the assumption that $\det(A) > 0$ precisely ensures that $\det(D) > 0$, and we can conclude that $d_1, \dots, d_n > 0$.

The final standard row operation is given by multiplying one row by a nonzero scalar $r \in \mathbb{R} - \{0\}$. Since $\mathbb{R} - \{0\}$ is not connected, we instead only allow the operation of multiplying one row by a positive scalar, $r \in (0, \infty)$. When acting on the i-th row, this operation is given by left multiplication by the diagonal matrix

$$\mathrm{diag}(1, \dots, 1, r, 1, \dots, 1)$$

that only differs from I_n in the (i,i)-th entry, which is equal to r. For $r > 0$ the path

$$t \mapsto \mathrm{diag}(1, \dots, 1, (1-t) + tr, 1, \dots, 1)A$$

for $t \in [0,1]$ connects A, within $GL_n(\mathbb{R})$, to the result of this row operation. Applying this to the i-th row with $r = 1/d_i$, for each $1 \leq i \leq n$, we have finally found a path in $GL_n(\mathbb{R})$ connecting A with $\det(A) > 0$ to the identity matrix.

This proves that $GL_n^+(\mathbb{R})$ is connected. Finally, there is a retraction

$$g \colon GL_n^+(\mathbb{R}) \longrightarrow SL_n(\mathbb{R})$$
$$A \longmapsto \frac{1}{\sqrt[n]{\det(A)}} A$$

from $GL_n^+(\mathbb{R})$ onto the subspace $SL_n(\mathbb{R})$, hence the latter is the continuous image of a path connected space, and is itself therefore path connected. □

It follows that any map from $GL_n^+(\mathbb{R})$ or $SL_n(\mathbb{R})$ to a totally disconnected space (e.g. a discrete space) is constant.

The proof of the complex case is easier, since $\mathbb{C} - \{0\}$ is path connected, and will be omitted.

Proposition 3.3.21. *$GL_n(\mathbb{C})$ and $SL_n(\mathbb{C})$ are (path) connected.*

3.4 (§26) Compact Spaces

3.4.1 Open covers and finite subcovers

Definition 3.4.1. A collection $\mathscr{C} = \{U_\alpha\}_{\alpha \in J}$ of subsets of X is said to *cover* X, or to be a *covering* of X, if the union of its elements is equal to X, so $X = \bigcup_{\alpha \in J} U_\alpha$. If each element in the collection is an open subset of X, then we say that $\mathscr{C}$ is an *open cover*.

A subcollection $\mathscr{D} \subset \mathscr{C}$ that also covers X is called a *subcover* of $\mathscr{C}$. If $\mathscr{F} \subset \mathscr{C}$ is a subcover with finitely many elements, then we call $\mathscr{F}$ a *finite subcover* of $\mathscr{C}$. In other words, this means that there is a finite subset $\{\alpha_1, \dots, \alpha_n\} \subset J$ with $\mathscr{F} = \{U_{\alpha_1}, \dots, U_{\alpha_n}\}$ and $X = U_{\alpha_1} \cup \dots \cup U_{\alpha_n}$.

Definition 3.4.2. A space X is said to be *compact* if for each open cover $\mathscr{C} = \{U_\alpha\}_{\alpha \in J}$ of X there exists a finite subcollection $\mathscr{F} = \{U_{\alpha_1}, \dots, U_{\alpha_n}\}$ that also covers X. In other words, X is compact if each open cover of X contains a finite subcover.

Remark 3.4.3. Note that the word 'finite' in 'finite subcover' refers to the collection $\mathscr{F}$ being finite, not that it elements are finite. On the other hand, the word 'open' in 'open cover' refers to the elements in $\mathscr{C}$ being open, not that $\mathscr{C}$ itself is open in some topology.

Example 3.4.4. A finite topological space X is compact, since there are only finitely many different open subsets $U \subset X$, so any collection covering X is finite.

Example 3.4.5. The real line $\mathbb{R}$ is not compact, since the open cover $\mathscr{C} = \{(n-1, n+1) \mid n \in \mathbb{N}\}$ does not admit a finite subcover.

Remark 3.4.6. From S. G. Krantz' "Mathematical apocrypha".

> There is a story about Sir Michael Atiyah (1929–) and Graeme Segal (1941–) giving an oral exam to a student at Cambridge. Evidently the poor student was a nervous wreck, and it got to a point where he could hardly answer any questions as all.
>
> At one point, Atiyah (endeavoring to be kind) asked the student to give an example of a compact set. The student said: "The real line." Trying to play along, Segal said: "In what topology?"

Example 3.4.7. The real line $\mathbb{R}$ in the trivial topology is compact, since the only open covers are the collections $\{\mathbb{R}\}$ and $\{\varnothing, \mathbb{R}\}$, which are finite.

Example 3.4.8. The subspace

$$X = \{0\} \cup \{1/n \mid n \in \mathbb{N}\}$$

of $\mathbb{R}$ is compact. Given an open covering $\mathscr{C}$ of X, choose $U \in \mathscr{C}$ with $0 \in U$. Since U is open in the subspace topology, there is an $N \in \mathbb{N}$ such that $1/n \in U$ for all $n > N$. For each $1 \le n \le N$ choose $U_n \in \mathscr{C}$ with $1/n \in U_n$. Then $\mathscr{F} = \{U, U_1, \dots, U_N\}$ is a finite subcover of $\mathscr{C}$.

Example 3.4.9. Any space X in the cofinite topology is compact. If X is finite, this is trivially true. Otherwise, if $\mathscr{C} = \{U_\alpha\}_{\alpha \in J}$ is an open cover, not all U_α can be empty. Choose a $\beta \in J$ so that $U_\beta \subset X$ is nonempty. Then $X - U_\beta = \{x_1, \dots, x_n\}$ is a finite set. For each i choose an $\alpha_i \in J$ so that $x_i \in U_{\alpha_i}$. Then $\{U_\beta, U_{\alpha_1}, \dots, U_{\alpha_n}\}$ is a finite subcover of $\mathscr{C}$. Since $\mathscr{C}$ was an arbitrary open cover of X, it follows that X is compact.

Definition 3.4.10. If A is a subspace of X, a collection $\mathscr{D}$ of subsets of X *covers* A if the union of the elements of $\mathscr{D}$ contains A.

Lemma 3.4.11. *Let A be a subspace of X. Then A is compact if and only if each covering of A by open subsets of X contains a finite subcollection covering A.*

Proof. Suppose that A is compact, and that $\mathscr{D} = \{U_\alpha\}_{\alpha \in J}$ is a covering of A by open subsets of X. Then $\mathscr{C} = \{A \cap U_\alpha\}_{\alpha \in J}$ is an open cover of A, hence contains a finite subcover $\mathscr{F} = \{A \cap U_{\alpha_1}, \dots, A \cap U_{\alpha_n}\}$. Then $\mathscr{G} = \{U_{\alpha_1}, \dots, U_{\alpha_n}\}$ is a finite subcollection of $\mathscr{D}$ that covers A.

Conversely, suppose that each covering of A by open subsets of X contains a finite subcollection covering A, and let $\mathscr{C} = \{V_\alpha\}_{\alpha \in J}$ be an open covering of A. For each $\alpha \in J$ we can write $V_\alpha = A \cap U_\alpha$ for some open subset $U_\alpha \subset X$. Then the collection $\mathscr{D} = \{U_\alpha\}_{\alpha \in J}$ is a covering of A by open subsets of X, which we have assumed contains a finite subcollection $\mathscr{G} = \{U_{\alpha_1}, \dots, U_{\alpha_n}\}$ covering A. Then $\mathscr{F} = \{V_{\alpha_1}, \dots, V_{\alpha_n}\}$ is a finite subcollection of $\mathscr{C}$ that covers A. Since $\mathscr{C}$ was an arbitrary open cover, it follows that A is compact. □

3.4.2 Compact subspaces of Hausdorff spaces

Theorem 3.4.12. *Every closed subspace of a compact space is compact.*

Proof. Let $A \subset X$ with A closed and X compact. Let $\mathscr{D} = \{U_\alpha\}_{\alpha \in J}$ be any covering A by open subsets in X. The complement $X - A$ is also open in X, so

$$\mathscr{C} = \mathscr{D} \cup \{X - A\}$$

is an open cover of X. There exists a finite subcover $\mathscr{F} \subset \mathscr{C}$ of X, which we may assume contains $X - A$, hence is of the form

$$\mathscr{F} = \{U_{\alpha_1}, \dots, U_{\alpha_n}, X - A\}$$

for some finite set of indices $\alpha_1, \dots, \alpha_n \in J$. Then

$$\mathscr{G} = \{U_{\alpha_1}, \dots, U_{\alpha_n}\}$$

is a finite subcollection of $\mathscr{D}$, and it covers A. Hence A is compact. □

Recall that finite subsets of Hausdorff spaces are closed. Compact subspaces generalize finite sets in this respect.

Theorem 3.4.13. *Every compact subspace of a Hausdorff space is closed.*

Proof. Let X be a Hausdorff space and let $K \subset X$ be a compact subspace. Let $p \in X - K$ be any point. We prove that $p \notin \bar{K}$, so that $K = \bar{K}$ is closed.

For each point $q \in K$ we have $p \neq q$, so there exist neighborhoods U_q and V_q of p and q, respectively, with $U_q \cap V_q = \varnothing$. The collection $\{V_q \mid q \in K\}$ of open subsets in X covers K, since

$$K \subset \bigcup_{q \in K} V_q \, .$$

By compactness of K, there is a finite subcollection $\{V_{q_1}, \dots, V_{q_n}\}$ that also covers K:

$$K \subset V = V_{q_1} \cup \dots \cup V_{q_n} \, .$$

Let $U = U_{q_1} \cap \dots \cap U_{q_n}$. Then U is neighborhood of p. We claim that $U \cap K = \varnothing$, so p is not in $\bar{K}$. In fact $U \cap V = \varnothing$, for if $x \in V$ then $x \in V_{q_i}$ for some i, but then $x \notin U_{q_i}$ so $x \notin U$. □

The following lemma was established in the course of the previous proof. It serves as inspiration for the notions of 'regular' and 'normal' spaces, to be considered in §31 (The Separation Axioms).

Lemma 3.4.14. *If X is a Hausdorff space, $K \subset X$ a compact subspace, and $p \in X - K$, then there exist disjoint open subsets U and V of X with $p \in U$ and $K \subset V$.*

Example 3.4.15. The intervals $(a, b]$, $[a, b)$ and (a, b) are not closed in $\mathbb{R}$, hence cannot be compact. We shall prove in the next section that each closed interval $[a, b]$ in $\mathbb{R}$ is compact.

Example 3.4.16. If X is an infinite set with the cofinite topology, then any subspace has the cofinite topology, hence is compact, but not every subspace is closed.

Theorem 3.4.17. *The continuous image of a compact space is compact.*

Proof. Let $f\colon X \to Y$ be continuous, and assume that X is compact. We prove that $f(X)$ is a compact subspace of Y. Let $\mathscr{C} = \{U_\alpha\}_{\alpha \in J}$ be a covering of $f(X)$ by open subsets of Y. Then $\{f^{-1}(U_\alpha)\}_{\alpha \in J}$ is an open cover of X. By compactness there exists a finite subcover $\{f^{-1}(U_{\alpha_1}), \dots, f^{-1}(U_{\alpha_n})\}$. Then the sets $\mathscr{F} = \{U_{\alpha_1}, \dots, U_{\alpha_n}\}$ cover $f(X)$. □

Theorem 3.4.18. *Let $f\colon X \to Y$ be a map from a compact space X to a Hausdorff space Y.*

(1) f is a closed map.

(2) If f is surjective, then f is a quotient map.

(3) If f is bijective, then f is a homeomorphism.

(4) If f is injective, then f is an embedding.

Proof. (1): Let $A \subset X$ be a closed subset. Since X is compact, A is compact. Since f is continuous, $f(A)$ is compact. Since Y is Hausdorff, $f(A) \subset Y$ is closed.

(2): Any closed, surjective map is a quotient map.

(3): If f is bijective, the inverse function $h = f^{-1}\colon Y \to X$ is continuous, since for each closed subset $A \subset X$ the preimage $h^{-1}(A) = f(A)$ is closed in Y.

(4): If f is injective, the corestriction $g\colon X \to f(X)$ is bijective, and $f(X)$ is Hausdorff, so g is a homeomorphism. □

Example 3.4.19. The map $\mathbb{R} \to \{0\}$ shows that the continuous preimage of a compact space needs not be compact.

Definition 3.4.20. A map $f\colon X \to Y$ is said to be *proper* if for each compact subspace $L \subset Y$ the preimage $f^{-1}(L)$ is compact.

3.4.3 Finite products of compact spaces

Theorem 3.4.21. *Let X and Y be compact spaces. Then $X \times Y$ is compact.*

Corollary 3.4.22. *Any finite product of compact spaces is compact.*

Lemma 3.4.23 (The tube lemma)**.** *Consider the product $X \times Y$, let $p \in X$, and assume that Y is compact. If $N \subset X \times Y$ is open, with $\{p\} \times Y \subset N$, then there exists a neighborhood $U \subset X$ of p with $U \times Y \subset N$.*

Proof. For each $q \in Y$ we have $(p, q) \in \{p\} \times Y \subset N$. Since N is open there is a basis element $U_q \times V_q \subset N$ for the product topology on $X \times Y$, with $p \in U_q$ open in X and $q \in V_q$ open in Y. The collection $\{V_q\}_{q \in Y}$ is an open cover of Y. By compactness of Y, there exists a finite subcover $\{V_{q_1}, \dots, V_{q_n}\}$. Let $U = U_{q_1} \cap \dots \cap U_{q_n}$. Then $p \in U$ is open in X. We claim that $U \times Y \subset N$. For any $(x, y) \in U \times Y$ there is an $1 \le i \le n$ with $y \in V_{q_i}$. Then $x \in U \subset U_{q_i}$, so $(x, y) \in U_{q_i} \times V_{q_i} \subset N$. □

Example 3.4.24. The tube lemma fails if Y is not compact. Consider the neighborhood

$$N = \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid |xy| \le 1\}$$

of $\{0\} \times \mathbb{R}$.

Proof of theorem. Let $\mathscr{C} = \{W_\alpha\}_{\alpha \in J}$ be an open cover of $X \times Y$. For each point $p \in X$, the subspace $\{p\} \times Y$ is compact, and is therefore covered by a finite subcollection $\mathscr{F}_p = \{W_{\alpha_1}, \dots, W_{\alpha_n}\}$ of $\mathscr{C}$. Let $N = W_{\alpha_1} \cup \dots \cup W_{\alpha_n}$. Then $N \subset X \times Y$ is an open subset containing $\{p\} \times Y$. By the tube lemma, there is a neighborhood $U_p \subset X$ of p with $U_p \times Y \subset N$. Note that $U_p \times Y$ is covered by the finite subcollection $\mathscr{F}_p$ of $\mathscr{C}$.

Now let $p \in X$ vary. The collection $\{U_p\}_{p \in X}$ is an open cover of X, hence admits a finite subcover $\{U_{p_1}, \dots, U_{p_m}\}$. For each $1 \le j \le m$ the subspace $U_{p_j} \times Y$ is covered by the finite subcollection $\mathscr{F}_{p_j}$ of $\mathscr{C}$. Hence the union

$$X \times Y = \bigcup_{j=1}^{m} U_{p_j} \times Y$$

is covered by the subcollection

$$\mathscr{F}_{p_1} \cup \dots \cup \mathscr{F}_{p_m}$$

of $\mathscr{C}$. This is a finite union of finite collections, hence is a finite subcollection of $\mathscr{C}$. Since $\mathscr{C}$ was an arbitrary open cover, it follows that $X \times Y$ is compact. □

Example 3.4.25. In the next section we show that $X = [0, 1]$ is compact in the subspace topology from $\mathbb{R}$. Hence for each $n \ge 0$ the n-fold product

$$[0,1]^n = [0,1] \times \dots \times [0,1]$$

is compact. Note that we can embed $[0, n]$ as a subspace of $[0,1]^n$

$$f_n \colon [0, n] \longrightarrow [0,1]^n$$

by sending $[i, i+1]$ to the edge from $(1, \dots, 1, 0, 0, \dots, 0)$ (with i copies of 1) to $(1, \dots, 1, 1, 0, \dots, 0)$ (with $i + 1$ copies of 1), for each $0 \le i < n$. In §37 we discuss the Tychonoff theorem, implying that also the infinite product

$$[0,1]^\omega = \prod_{n \ge 1} [0,1]$$

is compact in the product topology. Including $[0,1]^n$ in $[0,1]^\omega$, the paths above combine to an injective map

$$f_\infty \colon [0, \infty) \longrightarrow [0,1]^\omega$$

with closed, hence compact, image $K = f_\infty([0, \infty))$. However, $[0, \infty)$ is not compact, so f_∞ is a continuous bijection from a non-compact space to a compact Hausdorff space. The resolution to this apparent paradox is that the product topology on $[0,1]^\omega$ restricts to a coarser topology on the subspace A than the topology that would make f_∞ a homeomorphism.

3.4.4 The finite intersection property

Let $\mathscr{C}$ be a collection of open subsets of a space X. Let $\mathscr{E} = \{X - U \mid U \in \mathscr{C}\}$ be the collection of closed complements. To say that $\mathscr{C}$ is a cover of X is equivalent to saying that $\mathscr{E}$ has empty intersection:

$$\bigcap_{C \in \mathscr{E}} C = \bigcap_{U \in \mathscr{C}} (X - U) = X - \bigcup_{U \in \mathscr{C}} U$$

is empty if and only if $X = \bigcup_{U \in \mathscr{C}} U$.

Definition 3.4.26. A collection $\mathscr{E}$ of subsets of X has the *finite intersection property* if for each finite subcollection $\{C_1, \dots, C_n\} \subset \mathscr{E}$ the intersection

$$C_1 \cap \cdots \cap C_n$$

is nonempty.

Theorem 3.4.27. *A topological space X is compact if and only if for each collection $\mathscr{E}$ of closed subsets of X, having the finite intersection property, the intersection $\bigcap_{C \in \mathscr{E}} C$ is nonempty.*

Proof. "X is compact" is equivalent to the assertion:

> For any collection $\mathscr{C}$ of open subsets of X, if $\mathscr{C}$ covers X then some finite subcollection of $\mathscr{C}$ covers X.

This is logically equivalent to the contrapositive statement:

> For any collection $\mathscr{C}$ of open subsets of X, if no finite subcollection of $\mathscr{C}$ covers X, then $\mathscr{C}$ does not cover X.

Translated to a statement about the complementary collection of closed subsets $\mathscr{E}$, this is equivalent to:

> For any collection $\mathscr{E}$ of closed subsets of X, if no finite subcollection of $\mathscr{E}$ has empty intersection, then $\mathscr{E}$ does not have empty intersection.

In other words:

> For any collection $\mathscr{E}$ of closed subsets of X, if each finite subcollection of $\mathscr{E}$ has nonempty intersection, then $\mathscr{E}$ has nonempty intersection.

□

3.5 (§27) Compact Subspaces of the Real Line

Theorem 3.5.1. *Each closed interval $[a, b] \subset \mathbb{R}$ of the real line is compact.*

(Here $a, b \in \mathbb{R}$, we are not considering infinite intervals.)

Proof. Let $\mathscr{C} = \{U_\alpha\}_{\alpha \in J}$ be a covering of $[a, b]$ by open subsets of $\mathbb{R}$. Consider the set S of all $x \in [a, b]$ such that $[a, x]$ can be covered by a finite subcollection of $\mathscr{C}$. Then $a \in S$, since $a \in U_\alpha$ for some $\alpha \in J$, and then $\{U_\alpha\}$ is a finite subcollection of $\mathscr{C}$ that covers $[a, a] = \{a\}$. Hence S is nonempty and bounded above. Let $c = \sup S$ be the least upper bound of S. Clearly $c \in [a, b]$.

We claim that $c \in S$. Choose $\beta \in J$ with $c \in U_\beta$. Since U_β is open there exists an $\epsilon > 0$ with $(c - \epsilon, c + \epsilon) \subset U_\beta$. Since the supremum c is in the closure of S, there is some point $x \in S \cap (c - \epsilon, c + \epsilon)$. Since c is an upper bound for S, $x \le c$. Then $[a, x]$ can be covered by a finite subcollection $\{U_{\alpha_1}, \dots, U_{\alpha_n}\}$ of $\mathscr{C}$, and $[x, c]$ is contained in U_β. This implies that $[a, c] = [a, x] \cup [x, c]$ is covered by the finite subcollection $\{U_{\alpha_1}, \dots, U_{\alpha_n}, U_\beta\}$ of $\mathscr{C}$, so that $x \in S$.

We also claim that $c = b$. Suppose that $c < b$, to achieve a contradiction. Then there is a $y \in [a, b] \cap (c - \epsilon, c + \epsilon)$ with $c < y$ such that $[a, y]$ is covered by the same finite subcollection $\{U_{\alpha_1}, \dots, U_{\alpha_n}, U_\beta\}$. Hence $y \in S$, contradicting the assumption that c is an upper bound. □

Example 3.5.2. The surjective map $f \colon [0, 1] \to S^1$ given by $f(t) = (\cos(2\pi t), \sin(2\pi t))$ is a quotient map, since $[0, 1]$ is compact and $S^1 \subset \mathbb{R}$ is Hausdorff. Similarly for $f \times f \colon [0, 1] \times [0, 1] \to S^1 \times S^1$.

Theorem 3.5.3 (Heine-Borel theorem). *A subspace A of $\mathbb{R}^n$ is compact if and only if it is closed and bounded (in any of the equivalent metrics coming from a norm).*

Proof. Since $[a, b] \subset \mathbb{R}$ is compact, any finite product

$$[a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$$

is compact, as is any closed subset A of such a finite product. (These are the bounded subsets in the square metric.)

Conversely, the collection of open subsets

$$U_M = (-M, M) \times \cdots \times (-M, M) \subset \mathbb{R}^n$$

for $M \in \mathbb{N}$ has union $\mathbb{R}^n$, so if A is compact then there is a finite subcollection $\{U_{M_1}, \ldots, U_{M_k}\}$ of these that covers A. Let $M = \max\{M_1, \ldots, M_k\}$. Then $A \subset U_M$ is bounded. Since $\mathbb{R}^n$ is Hausdorff we must also have that A is closed. □

Theorem 3.5.4 (Extreme value theorem). *Let $f\colon X \to \mathbb{R}$ be continuous, with X compact. Then there exist points $c, d \in X$ with*

$$f(c) \leq f(x) \leq f(d)$$

for all $x \in X$.

Proof. The continuous image $f(X) \subset \mathbb{R}$ is compact, hence closed, so contains both its infimum and its supremum. Writing these values as $f(c)$ and $f(d)$, we get the conclusion. □

3.5.1 The Lebesgue number

Definition 3.5.5. Let (X, d) be a metric space and let $A \subset X$ be a nonempty subset. For each $x \in X$ the *distance from x to A* is

$$d(x, A) = \inf\{d(x, a) \mid a \in A\}\,.$$

The *diameter of A* is

$$\operatorname{diam}(A) = \sup\{d(a, b) \mid a, b \in A\}\,.$$

Lemma 3.5.6. *The function $x \mapsto d(x, A)$ is continuous.*

Proof. Let $x, y \in X$. By the triangle inequality,

$$d(x, A) \leq d(x, a) \leq d(x, y) + d(y, a)$$

for all $a \in A$, so

$$d(x, A) - d(x, y) \leq \inf\{d(y, a) \mid a \in A\} = d(y, A)$$

and

$$|d(x, A) - d(y, A)| \leq d(x, y)$$

by symmetry in x and y. Hence $x \mapsto d(x, A)$ is continuous. □

Lemma 3.5.7. *Let $\mathscr{C}$ be an open cover of a compact metric space (X, d). There exists a $\delta > 0$ such that for each subset $B \subset X$ of diameter $< \delta$ there exists an element $U \in \mathscr{C}$ with $B \subset U$. The number δ is called a* Lebesgue number *of $\mathscr{C}$.*

Proof. If $X \in \mathscr{C}$ then any positive number is a Lebesgue number for $\mathscr{C}$. Otherwise, by compactness there is a finite subcollection $\{U_1, \dots, U_n\}$ of $\mathscr{C}$ that covers X. Let $C_i = X - U_i$ be the closed complement; each C_i is nonempty. To say that $x \in U_i$ is equivalent to saying $d(x, C_i) > 0$, since $d(x, C_i) = 0$ if and only if $x \in C_i$.

Define $f \colon X \to \mathbb{R}$ as the average

$$f(x) = \frac{1}{n} \sum_{i=1}^{n} d(x, C_i)\,.$$

Claim: $f(x) > 0$ for all $x \in X$. Proof: For any given $x \in X$ there is an i with $x \in U_i$. Since U_i is open, there exists an $\epsilon > 0$ with $B_d(x, \epsilon) \subset U_i$. Then $d(x, C_i) \geq \epsilon$. Hence $f(x) \geq \epsilon/n > 0$.

Since f is continuous, it has a positive minimum value δ. Claim: δ is a Lebesgue number for $\{U_1, \dots, U_n\}$, hence for $\mathscr{C}$. Proof: Let $B \subset X$ have diameter $< \delta$. There is only something to prove for B nonempty; choose a point $p \in B$. Then $B \subset B_d(p, \delta)$. Consider the numbers $d(p, C_i)$ for $1 \leq i \leq n$. Choose m so that $d(p, C_m)$ is the largest of these numbers. Then

$$\delta \leq f(p) \leq d(p, C_m)$$

so $B_d(p, \delta) \cap C_m = \varnothing$. Hence $B \subset B_d(p, \delta) \subset U_m$. □

3.5.2 Uniform continuity

Definition 3.5.8. Let $f \colon (X, d) \to (Y, d)$ be a function between metric spaces. We say that f is *uniformly continuous* if given $\epsilon > 0$ there exists a $\delta > 0$ such that for any two points $x, x' \in X$ with $d(x, x') < \delta$ we have $d(f(x), f(x')) < \epsilon$, or equivalently, if for any $x \in X$ we have $f(B_d(x, \delta)) \subset B_d(f(x), \epsilon)$.

Theorem 3.5.9. *Let $f \colon (X, d) \to (Y, d)$ be a continuous map om metric spaces. If X is compact then f is uniformly continuous.*

Proof. Given $\epsilon > 0$ cover Y by the balls $B_d(y, \epsilon/2)$ and let

$$\mathscr{C} = \{f^{-1}(B_d(y, \epsilon/2))\}_{y \in Y}$$

be the open covering of X by the preimages of these balls. Choose a Lebesgue number $\delta > 0$ for this open covering. If $x, x' \in X$ with $d(x, x') < \delta$ then $\{x, x'\} \subset f^{-1}(B_d(y, \epsilon/2))$ for some $y \in Y$, hence $\{f(x), f(x')\} \subset B_d(y, \epsilon/2)$. Thus $d(f(x), f(x')) \leq d(f(x), y) + d(y, f(x')) < \epsilon$. □

3.5.3 The Gram–Schmidt process

Let $A^T = (a_{j,i})_{i,j=1}^n$ denote the transpose of a matrix $A = (a_{i,j})_{i,j=1}^n$, and let

$$O_n = \{A \in M_n(\mathbb{R}) \mid A^T A = I\}$$

be the space of $n \times n$ orthogonal matrices. (The notation $O(n)$ is also commonly used.) Here $(A^T A)_{i,j} = \alpha_i \cdot \alpha_j$ is the dot product of the i-th and the j-th column vectors of $A = (\alpha_1, \dots, \alpha_n)$, so the condition $A^T A = I$ asserts that the α_i are pairwise orthogonal, each of unit length. Note that $A^T A = I$ implies $\det(A)^2 = 1$, so each orthogonal matrix is invertible and $O_n \subset GL_n(\mathbb{R})$. Let

$$SO_n = \{A \in O_n \mid \det(A) = 1\}$$

(also denoted $SO(n)$) be the space of special orthogonal matrices. For example, SO_2 is the space of rotations of the plane, and SO_3 is the space of rotations of $\mathbb{R}^3$. The determinant restricts to a map det: $O_n \to \{\pm 1\}$, where $\{\pm 1\}$ is discrete, so SO_n is the preimage of the open and closed point $1 \in \{\pm 1\}$. Hence SO_n is open and closed in O_n. For $n \geq 1$ the image of det on O_n is disconnected, hence O_n is not connected.

The rule $A \mapsto A^T A$ defines a continuous function $M_n(\mathbb{R}) \to M_n(\mathbb{R})$, since each component $\alpha_i \cdot \alpha_j$ is a (2nd order) polynomial in the entries of A. Hence $O_n \subset M_n(\mathbb{R})$ is the preimage of the closed point $I \in M_n(\mathbb{R})$. It follows that O_n is a closed subspace of $M_n(\mathbb{R})$, as well as of $GL_n(\mathbb{R})$. Hence SO_n is also closed in $GL_n(\mathbb{R})$ and $M_n(\mathbb{R})$.

Let $S^{n-1} \subset \mathbb{R}^n$ be the subspace of unit vectors. It is clearly closed and bounded, hence compact. Since each column vector of A has unit length, we have an inclusion

$$\begin{aligned} O_n &\subset S^{n-1} \times \cdots \times S^{n-1} \\ A &\mapsto (\alpha_1, \dots, \alpha_n) \end{aligned}$$

Here $S^{n-1} \times \cdots \times S^{n-1}$ is a product of n compact spaces, hence is compact. It follows that the space O_n of orthogonal matrices is a closed subspace of a compact space and is therefore itself compact. Furthermore, SO_n is a closed subspace of O_n, and is therefore also compact.

The Gram–Schmidt orthogonalization process defines a retraction

$$gs\colon GL_n(\mathbb{R}) \longrightarrow O_n\,.$$

Here $A = (\alpha_1, \dots, \alpha_n) \in GL_n(\mathbb{R})$ is sent to $gs(A) = B = (\beta_1, \dots, \beta_2) \in O_n$, where $\beta_1 = \alpha_1/\|\alpha_1\|$, is the normalization of α_1, β_2 is the normalization of $\alpha_2 - (\alpha_2 \cdot \beta_1)\beta_1$, etc. The matrix B depends continuously on A, and if A was orthogonal to start with, then $B = A$. This process does not change the sign of the determinant, so gs restricts to a retraction

$$gs\colon SL_n(\mathbb{R}) \longrightarrow SO_n\,.$$

We showed earlier that $SL_n(\mathbb{R})$ is (path) connected. From this it follows that SO_n is connected.

The spaces SO_n are fundamental examples of connected compact Lie groups. As spaces, we can identify SO_2 with the circle S^1, while SO_3 is homeomorphic to the projective space $\mathbb{R}P^3$ of lines through the origin in $\mathbb{R}^4$.

A similar discussion applies for matrices with complex entries, leading to the spaces $U_n \subset GL_n(\mathbb{C})$ of unitary matrices, and their subspaces $SU_n \subset U_n$ of special unitary matrices, all of which are examples of connected compact Lie groups. Here $U_1 \cong S^1$, while $SU_2 \cong S^3$.

3.6 (§28) Limit Point Compactness

Definition 3.6.1. Let $(x_n)_{n=1}^\infty$ be a sequence of points in X. If

$$n_1 < n_2 < \cdots < n_k < \dots$$

is a strictly increasing sequence of natural numbers, the sequence

$$x_{n_1}, x_{n_2}, \dots, x_{n_k}, \dots$$

is called a *subsequence* of $(x_n)_{n=1}^\infty$. It is a *convergent subsequence* if $x_{n_k} \to p$ as $k \to \infty$, for some $p \in X$.

Definition 3.6.2. A space X is *sequentially compact* if every sequence $(x_n)_{n=1}^\infty$ in X has a convergent subsequence $(x_{n_k})_{k=1}^\infty$.

Theorem 3.6.3. *Let X be a metrizable space. Then X is compact if and only if X is sequentially compact.*

Proof. (Munkres first proves that compact spaces are 'limit point compact', and then proves that metrizable limit point compact spaces are sequentially compact. We give a direct proof of the composite implication, compact $\implies$ sequentially compact for metrizable spaces.) Consider any sequence $(x_n)_{n=1}^\infty$ in X. For each $n \ge 1$, let

$$A_n = \{x_m \mid m \ge n\}$$

be the set of points in the subsequence obtained by omitting $x_1, \dots, x_{n-1}$ from the sequence, and let

$$C_n = \bar{A}_n$$

be the closure (in X) of this set. Here

$$C_1 \supset C_2 \supset \dots \supset C_n \supset \dots$$

and each $C_n \ne \varnothing$, so no finite intersection of the collection $\mathscr{E} = \{C_n\}_{n=1}^\infty$ is empty. By the finite intersection property of the compact space X we can choose a point

$$p \in \bigcap_{n=1}^{\infty} C_n \,,$$

since this intersection is nonempty. For each k we have $p \in C_k = \bar{A}_k$, so $B_d(p, 1/k) \cap A_k \ne \varnothing$ and we can choose an index $n_k \ge k$ with $x_{n_k} \in B_d(p, 1/k)$. We then have

$$x_{n_k} \to p \qquad \text{as} \qquad k \to \infty \,.$$

As chosen, the sequence

$$n_1 \le n_2 \le \dots \le n_k \le \dots$$

may not be strictly increasing, but since $n_k \ge k$ for each k there are at most finitely many repetitions of each value, so we can delete any repetitions and renumber, so as to obtain a genuine convergent subsequence $(x_{n_k})_{k=1}^\infty$ of $(x_n)_{n=1}^\infty$.

The proof of the implication sequentially compact $\implies$ compact remains. First we show that if X is sequentially compact then the Lebesgue number lemma holds for X:

Lemma 3.6.4. *Let $\mathscr{C}$ be an open cover of a sequentially compact metric space (X, d). Then there exists a $\delta > 0$ such that for each subset $B \subset X$ of diameter $< \delta$ there is an element $U \in \mathscr{C}$ with $B \subset U$.*

We assume that no such δ exists, and achieve a contradiction. For each $n \in \mathbb{N}$ there is a set B_n of diameter $< 1/n$ that is not contained in any element of $\mathscr{C}$. Choose $x_n \in B_n$. By the assumed sequential compactness, the sequence $(x_n)_{n=1}^\infty$ has a convergent subsequence $(x_{n_k})_{k=1}^\infty$, with $(n_k)_{k=1}^\infty$ a strictly increasing sequence. Let $p \in X$ be its limit: $x_{n_k} \to p$ as $k \to \infty$. There is an $U \in \mathscr{C}$ with $p \in U$. Since U is open, there is an $\epsilon > 0$ with $B_d(p, \epsilon) \subset U$. For k sufficiently large we have $1/n_k < \epsilon/2$ and $d(x_{n_k}, p) < \epsilon/2$. Then B_{n_k} lies in the $\epsilon/2$-neighborhood of x_{n_k}, hence in the ϵ-neighborhood of p:

$$B_{n_k} \subset B_d(x_{n_k}, \epsilon/2) \subset B_d(p, \epsilon) \subset U$$

This contradicts the choice of B_{n_k}, not being contained in any element of $\mathscr{C}$.

Next we show that if X is sequentially compact then it is *totally bounded*:

Lemma 3.6.5. *Let (X, d) be sequentially compact. For each $\epsilon > 0$ there exists a finite covering of X by ϵ-balls.*

Assume that for some $\epsilon > 0$ there is no finite covering of X by ϵ-balls, to reach a contradiction. Construct a sequence $(x_n)_{n=1}^{\infty}$ as follows. Choose any point $x_1 \in X$. Having chosen $x_1, \dots, x_n$, note that the finite union

$$B_d(x_1, \epsilon) \cup \cdots \cup B_d(x_n, \epsilon)$$

is not all of X, so we can choose x_{n+1} in its complement, and continue. By construction, $d(x_m, x_n) \geq \epsilon$ for all $m \neq n$, so $(x_n)_{n=1}^{\infty}$ does not contain any convergent subsequence. This contradicts sequential compactness of X.

We can now finish the proof of the theorem. Let (X, d) be sequentially compact. To prove that X is compact, consider any open cover $\mathscr{C}$ of X. It has a Lebesgue number $\delta > 0$. Let $\epsilon = \delta/3$. Choose a finite covering of X by ϵ-balls. Each ϵ-ball has diameter $\leq 2\epsilon < \delta$, hence is contained in an element of $\mathscr{C}$. Hence X is covered by finitely many of the elements of $\mathscr{C}$, so $\mathscr{C}$ has a finite subcover. □

3.7 (§29) Local Compactness

We have seen that the closed subspaces of a compact Hausdorff space are the same as the compact subspaces. We shall now consider a condition satisfied by the open subspaces of compact Hausdorff spaces.

Definition 3.7.1. A space X is *locally compact at x* if there is a compact subspace C of X that contains a neighborhood V of x:

$$x \in V \subset C \subset X$$

It is *locally compact* if it is locally compact at each of its points.

Example 3.7.2. Any compact space is locally compact. (Take $V = C = X$.)

Example 3.7.3. The real line $\mathbb{R}$ is locally compact. Each point $x \in \mathbb{R}$ is contained in the compact subspace $C = [x-1, x+1]$, which contains the neighborhood $V = (x-1, x+1)$.

Example 3.7.4. The set of rational numbers $\mathbb{Q}$, in the subspace topology from $\mathbb{R}$, is not locally compact. Any subset $C \subset \mathbb{Q}$ containing a basis neighborhood $\mathbb{Q} \cap (x-\epsilon, x+\epsilon)$ cannot be compact. For instance, choosing an irrational number $a \in (x-\epsilon, x+\epsilon)$ the real function $f(t) = (t-a)^2$ is continuous on C and takes arbitrarily small positive values, but is never zero. By the extreme value theorem, C cannot be compact.

Example 3.7.5. Euclidean n-space $\mathbb{R}^n$ is locally compact. Each point $(x_1, \dots, x_n) \in \mathbb{R}^n$ is contained in the compact subspace

$$C = [x_1 - 1, x_1 + 1] \times \cdots \times [x_n - 1, x_n + 1]\,,$$

which contains the neighborhood

$$V = (x_1 - 1, x_1 + 1) \times \cdots \times (x_n - 1, x_n + 1)\,.$$

Example 3.7.6. The countably infinite product $\mathbb{R}^\omega$ is not locally compact. Any neighborhood V of $0 = (0)_{n=1}^{\infty}$ contains a basis neighborhood

$$(-\epsilon, \epsilon) \times \cdots \times (-\epsilon, \epsilon) \times \mathbb{R} \times \cdots \times \mathbb{R} \times \dots$$

for some $\epsilon > 0$. If V were contained in a compact subspace C, then the closure

$$[-\epsilon, \epsilon] \times \cdots \times [-\epsilon, \epsilon] \times \mathbb{R} \times \cdots \times \mathbb{R} \times \dots$$

of the basis neighborhood would be compact, which it is not.

3.7.1 The one-point compactification

The *Gauss sphere* is the model S^2 for $\mathbb{C}\cup\{\infty\}$. We now discuss how to give $\mathbb{C}\cup\{\infty\}$ a topology homeomorphic to S^2, in a way that generalizes from $\mathbb{C}$ to all locally compact Hausdorff spaces.

Definition 3.7.7. Let X be a locally compact Hausdorff space. Let $Y = X \cup \{\infty\}$ where $\infty \notin X$. Give Y the topology $\mathscr{T}_\infty$ consisting of

(1) the open subsets $U \subset X$, and

(2) the complements $Y - C$ of compact subsets $C \subset X$.

We call Y the *one-point compactification* of X.

Theorem 3.7.8. *Let X be a locally compact Hausdorff space. The one-point compactification $Y = X \cup \{\infty\}$ is a compact Hausdorff space, $X \subset Y$ is a subspace, and $Y - X$ consists of a single point.*

Proof. We first prove that the given collection $\mathscr{T}_\infty$ is a topology on Y. The empty set is of type (1) and Y is of type (2). To check that the intersection of two open sets is open, there are three cases:

$$U_1 \cap U_2 \subset X$$
$$(Y - C_1) \cap (Y - C_2) = Y - (C_1 \cup C_2)$$
$$U_1 \cap (Y - C_2) = U_1 \cap (X - C_2)$$

These are of type (1), (2) and (1), respectively, since $C_1 \cup C_2$ is compact (Exercise!) and $X - C_2$ is open, since X is assumed to be Hausdorff.

To check that the union of a (nonempty) collection of open sets is open, there are three cases:

$$\bigcup_{\alpha\in J} U_\alpha = U \subset X$$
$$\bigcup_{\beta\in K} (Y - C_\beta) = Y - \bigcap_{\beta\in K} C_\beta = Y - C$$
$$U \cup (Y - C) = Y - (C - U)$$

These are of type (1), (2) and (2), respectively, since $C = \bigcap_{\beta\in K} C_\beta \subset X$ is a closed subspace of some compact space C_β, hence is compact, and $C - U$ is a closed subspace of C, hence is compact.

Next we show that $X \subset Y$ is a subspace: The open sets in the subspace topology are of the form $X \cap V$ where V is open in Y. If $V = U \subset X$ is of type (1), then $X \cap V = U$ is open in X. If $V = Y - C$ is of type (2), then $X \cap V = X - C$ is open in X since $C \subset X$ is compact, hence closed, in the Hausdorff space X. Conversely, if $U \subset X$ is open, then U is open of type (1) in Y.

To show that Y is compact, let $\mathscr{C}$ be an open cover of Y. Some element $V \in \mathscr{C}$ must contain $\infty \notin X$, hence be of the form $V = Y - C$. The collection $\mathscr{C}$ of open subsets of Y covers the compact space C, so there is a finite subcollection $\{U_1, \dots, U_n\} \subset \mathscr{C}$ that covers C. Then $\mathscr{F} = \{V, U_1, \dots, U_n\}$ is a finite subcover of $\mathscr{C}$.

To show that Y is Hausdorff, let $x, y \in Y$. If both lie in X, then there are open subsets $U, V \subset X$ with $x \in U$, $y \in V$, $U \cap V = \varnothing$. Then U and V are also open and disjoint in Y. Otherwise, we may assume that $x \in X$ and $y = \infty$. Since X is locally compact at x there exists a compact $C \subset X$ containing a neighborhood U of x. Let $V = Y - C$. Then $x \in U$, $\infty \in V$, U and V are open in Y and $U \cap V = \varnothing$. □

Here is a converse.

Proposition 3.7.9. *Let $X \subset Y$ be a subspace of a compact Hausdorff space, such that $Y - X$ consists of a single point. Then X is locally compact and Hausdorff.*

Proof. As a subspace of a Hausdorff space, it is clear that X is Hausdorff. We prove that it is locally compact. Let $x \in X$ and let y be the single point of $Y - X$. Since Y is Hausdorff, there are open sets $U, V \subset Y$ with $x \in U$, $y \in V$, $U \cap V = \varnothing$. Let $C = Y - V$. It is a closed subset of a compact space, hence compact. Thus $x \in U \subset C \subset X$, as required for local compactness at x. □

There is also the following uniqueness statement, which justifies why we say "the one-point compactification", not just "a one-point compactification".

Proposition 3.7.10. *Let X be locally compact Hausdorff, with one-point compactification $Y = X \cup \{\infty\}$, and suppose that Y' is a compact Hausdorff space such that $X \subset Y'$ is a subspace and $Y' - X$ is a single point. Then the unique bijection $Y' \to Y$ that is the identity on X is a homeomorphism.*

Proof. It suffices to prove that the bijection $f\colon Y' \to Y$ is continuous, since Y' is compact and Y is Hausdorff. An open subset of Y is of the form U or $Y - C$, with $U \subset X$ open and $C \subset X$ compact. The preimage $f^{-1}(U) = U$ is then open in X, hence also in Y', since X must be open in the Hausdorff space Y' because its complement is a single point. The preimage $f^{-1}(Y - C) = Y' - C$ will also be open in Y', because C is compact and Y' is Hausdorff, so $C \subset Y'$ is closed. □

Example 3.7.11. The one-point compactification of the open interval $(0, 1)$ is homeomorphic to the circle S^1. This follows from the uniqueness statement above, and the homeomorphism

$$f\colon (0,1) \to S^1 - \{(1,0)\}$$

given by $f(t) = (\cos(2\pi t), \sin(2\pi t))$. The closed interval $[0, 1]$ is a different compactification of $(0, 1)$, with $[0, 1] - (0, 1) = \{0, 1\}$ consisting of two points.

Since $(0, 1) \cong \mathbb{R}$, the one-point compactification of $\mathbb{R}$ is also homeomorphic to the circle:

$$\mathbb{R} \cup \{\infty\} \cong S^1$$

Example 3.7.12. The one-point compactification of the open unit n-ball

$$B(0,1) = \{x \in \mathbb{R}^n \mid \|x\| < 1\}$$

is homeomorphic to the n-sphere S^n. The closed n-ball

$$D^n = \bar{B}(0,1) = \{x \in \mathbb{R}^n \mid \|x\| \le 1\}$$

is a different compactification, with $\bar{B}(0, 1) - B(0, 1) = S^{n-1}$ consisting of the $(n - 1)$-sphere.

Since $B(0, 1) \cong \mathbb{R}^n$, the one-point compactification of $\mathbb{R}^n$ is also homeomorphic to the n-sphere:

$$\mathbb{R}^n \cup \{\infty\} \cong S^n$$

3.7.2 The local nature of local compactness

For Hausdorff spaces, the property of being locally compact is a local property in the following sense.

Theorem 3.7.13. *Let X be a Hausdorff space. Then X is locally compact if and only if for each point $x \in X$ and each neighborhood U of x there is a neighborhood V of x with compact closure $\bar{V}$ contained in U:*

$$x \in V \subset \bar{V} \subset U$$

Proof. The stated property implies local compactness at x, by taking $U = X$ and $C = \bar{V}$.

For the converse, suppose that X is locally compact (and Hausdorff), and let $x \in U \subset X$ be a neighborhood. Let $Y = X \cup \{\infty\}$ be the one-point compactification, and let $K = Y - U$. Then $K \subset Y$ is closed, hence compact. Since Y is Hausdorff and $x \notin K$ we can find open subsets $V, W \subset Y$ with $x \in V$, $K \subset W$ and $V \cap W = \varnothing$. Then $\infty \in K \subset W$, so $W = Y - C$ for some compact $C \subset X$. Hence $x \in V \subset C \subset U$ with V open and C compact. Hence $\bar{V}$ is also compact. □

Corollary 3.7.14. *Let X be locally compact Hausdorff. If $A \subset X$ is an open or closed subspace, then A is locally compact.*

Proof. Suppose that A is open in X. Let $x \in A$. By the previous theorem there is a neighborhood V of x with $\bar{V}$ compact and $\bar{V} \subset A$. This shows that A is locally compact at x.

Suppose instead that A is closed in X. Let $x \in A$. Since X is locally compact there is a compact subspace $C \subset X$ that contains a neighborhood V of x. Then $A \cap C$ is closed in C, hence compact, and contains the neighborhood $A \cap V$ of x in the subspace topology on A. □

Corollary 3.7.15. *A space X is homeomorphic to an open subspace of a compact Hausdorff space if and only if X is locally compact and Hausdorff.*

Proof. If X is an open subspace of a compact Hausdorff space, then X is locally compact by the corollary above, and obviously Hausdorff. The same applies if X is homeomorphic to such an open subspace.

Conversely, if X is locally compact and Hausdorff then X is an open subspace of its one-point compactification $Y = X \cup \{\infty\}$, which is compact Hausdorff. □

Chapter 4

Countability and Separation Axioms

We have seen the first countability axiom (each point has a countable neighborhood basis) and the Hausdorff separation axiom (two points can be separated by disjoint neighborhoods).

Our aim is to prove the Urysohn metrization theorem, saying that if a topological space X satisfies a countability axiom (it is second countable) and a separation axiom (is it regular), then we can construct enough continuous functions $X \to \mathbb{R}$ to embed X into a metric space, so that X is metrizable.

The Hausdorff property, the first and second countability axioms, and the term 'metric space' were introduced in 'Grundzüge der Mengenlehre' by Felix Hausdorff, published in 1914. Hausdorff, together with his wife and her sister, committed suicide in January 1942, when they and other Jews in Bonn were ordered to move to the Endenich camp.

Paul Urysohn proved his lemma and metrization theorem in 1923 or 1924. The results were published posthumously, after Urysohn drowned when swimming off the coast of Brittany in August 1924.

4.1 (§30) The Countability Axioms

It may be a good idea to look through §7 on Countable and Uncountable Sets, if this material is unfamiliar.

Definition 4.1.1. A set C is *countable* if there is a surjection $f\colon \mathbb{N} \to C$. Otherwise C is *uncountable*. Finite sets are countable. A set is *countably infinite* if it is countable but not finite, or equivalently, if there is a bijection $g\colon \mathbb{N} \to C$.

Theorem 4.1.2. *(a) A quotient set of a countable set is countable.*

(b) A subset of a countable set is countable.

(c) A finite or countable union of countable sets is countable.

(d) A finite product of countable sets is countable.

(e) The set $\mathbb{Q}$ of rational numbers is countable.

Theorem 4.1.3 (Cantor's diagonal argument)**.** *(a) A countable product of sets with 2 or more elements each is uncountable.*

(b) The set $\mathbb{R}$ of real numbers is uncountable.

Proof. (a) Let $2 = \{0, 1\}$. We show that $X = 2^\omega = \{0, 1\}^\omega$ is uncountable. Write an element of X as a sequence

$$x = (x_1, x_2, \dots, x_n, \dots)$$

with each $x_n \in \{0,1\}$. Consider any function $f\colon \mathbb{N} \to X$. For each $m \in \mathbb{N}$ we have

$$f(m) = (f(m)_1, f(m)_2, \dots, f(m)_n, \dots)$$

with $f(m)_n \in \{0,1\}$. Let $y \in \{0,1\}^\omega$ be the sequence given by

$$y_n = \begin{cases} 0 & \text{if } f(n)_n = 1, \\ 1 & \text{if } f(n)_n = 0. \end{cases}$$

Then y is not in the image of f, because if $y = f(m)$ then $y_m = f(m)_m$ and we chose y so that $y_m \neq f(m)_m$. In particular, f is not surjective.

(b) This follows from (a), e.g. by consideration of binary expansions of real numbers. □

4.1.1 First-countable spaces

Definition 4.1.4. A space X has a *countable basis at x* if there is a countable collection $\mathscr{B}_x = \{B_n \mid n \in \mathbb{N}\}$ of neighborhoods of x such that each neighborhood U of x contains at least one of the elements of $\mathscr{B}_x$:

$$x \in B_n \subset U\,.$$

A space having a countable basis at each of its points is said to satisfy the *first countability axiom*, or to be *first-countable*.

Every metric space (X, d) is first-countable, since the collection $\{B_d(x, 1/n) \mid n \in \mathbb{N}\}$ is a countable basis at x, for each $x \in X$. The fact that $B_d(x, 1/n) \supset B_d(x, 1/n + 1)$ is typical of the general case.

Lemma 4.1.5. *If X has a countable basis at x then it has a nested countable basis at x, i.e., a descending sequence*

$$B'_1 \supset B'_2 \supset \dots \supset B'_n \supset \dots$$

of neighborhoods, such that each neighborhood U of x contains B'_n for all sufficiently large n.

Proof. Given a countable basis $\{B_n\}_n$ at x, let

$$B'_n = B_1 \cap B_2 \cap \dots \cap B_n$$

for each $n \geq 1$. □

The following results, which we proved in §21 for metric spaces, also hold for all first-countable spaces.

Theorem 4.1.6. *Let X be a first-countable space.*

(a) Let $A \subset X$. If $x \in \bar{A}$ then there is a sequence $(x_n)_{n=1}^\infty$ of points in A converging to x.

(b) Let $f\colon X \to Y$. If for each sequence $(x_n)_{n=1}^\infty$ in X converging to x the sequence $(f(x_n))_{n=1}^\infty$ in Y converges to $f(x)$, then f is continuous.

Theorem 4.1.7. *(a) Any subspace A of a first-countable space X is first-countable.*

(b) Any countable product $\prod_{k=1}^\infty X_k$ of first-countable spaces X_k is first-countable.

Proof. (a) If $x \in A$ and $\mathscr{B} = \{B_n \mid n \in \mathbb{N}\}$ is a countable basis at x in X, then $\{A \cap B_n \mid n \in \mathbb{N}\}$ is a countable basis at x in A.

(b) If $x = (x_k)_{k=1}^\infty \in X = \prod_{k=1}^\infty X_k$ and $\mathscr{B}_k$ is a countable basis at x_k in X_k, then the collection of products

$$\prod_{k=1}^\infty U_k\,,$$

where $U_k \in \mathscr{B}_k$ for finitely many values of k and $U_k = X_k$ for the remaining values of k, is a countable basis at x in X. (Exercise: Check that this collection is countable. It is a countable union of finite products of countable sets.) □

4.1.2 Second-countable spaces

Definition 4.1.8. A space X has a countable basis (for its topology) if there is a countable collection $\mathscr{B}$ of subsets of X that is a basis for the topology on X. In this case X is said to satisfy the *second countability axiom*, or to be *second-countable.*

Lemma 4.1.9. *Second-countability implies first-countability.*

Proof. Let $\mathscr{B} = \{U_n\}_{n=1}^{\infty}$ be a countable basis for the topology of a space X. For each point $x \in X$ the subcollection $\mathscr{B}_x = \{U \in \mathscr{B} \mid x \in U\}$ is a countable neighborhood basis at x. □

Example 4.1.10. The real line is second-countable. A countable basis for the topology is given by the open intervals (a, b) with $a < b \in \mathbb{Q}$ both rational.

Euclidean n-space $\mathbb{R}^n$ is second-countable. A countable basis is given by the products

$$(a_1, b_1) \times \cdots \times (a_n, b_n)$$

where all $a_k, b_k \in \mathbb{Q}$ are rational.

Even the infinite product $\mathbb{R}^\omega$ is second-countable. A countable basis is given by the products

$$\prod_{k=1}^{\infty} U_k \subset \prod_{k=1}^{\infty} \mathbb{R} = \mathbb{R}^\omega$$

where $U_k = (a_k, b_k)$ with rational endpoints, for finitely many k, and $U_k = \mathbb{R}$ for all other k.

Not every metric space is second-countable. A counterexample is $\mathbb{R}^\omega$ in the uniform topology. Another counterexample is $\mathbb{R}$ with the discrete topology, corresponding to the metric d with $d(x, y) = 1$ for all $x \neq y$ in $\mathbb{R}$.

Theorem 4.1.11. *(a) Any subspace A of a second-countable space X is second-countable.*

(b) Any countable product $\prod_{k=1}^{\infty} X_k$ of second-countable spaces X_k is second-countable.

Proof. (a) If $\mathscr{B} = \{B_n\}_{n=1}^{\infty}$ is a countable basis for X, and $A \subset X$, then $\{A \cap B_n\}_{n=1}^{\infty}$ is a countable basis for the subspace topology on A.

(b) If $\mathscr{B}_k$ is a countable basis for X_k, for $k \in \mathbb{N}$, then the collection of products

$$\prod_{k=1}^{\infty} U_k$$

where $U_k \in \mathscr{B}_k$ for finitely many k, and $U_k = X_k$ for the remaining k, is a countable basis for the product topology on $\prod_{k=1}^{\infty} X_k$. □

4.1.3 Countable dense subsets

Recall that $A \subset X$ is dense if $\bar{A} = X$, i.e., if each nonempty open subset of X meets A.

Example 4.1.12. The rational numbers $\mathbb{Q} \subset \mathbb{R}$ are dense in the real line.

Theorem 4.1.13. *Suppose that X is second-countable. Then there exists a countable dense subset of X.*

A space with a countable dense subset is sometimes called a *separable space.*

Proof. Let $\{B_n\}_{n=1}^{\infty}$ be a countable basis for the topology on X. We may assume that each B_n is nonempty. Choose a point $x_n \in B_n$ for each $n \geq 1$, and let $D = \{x_n \mid n \geq 1\}$. We claim that D is dense in X. Consider any nonempty open subset $U \subset X$ and choose a point $y \in U$. Then there is a basis neighborhood B_n with $y \in B_n \subset U$. Hence $x_n \in D \cap B_n \subset D \cap U$, so D meets U. This implies that $\bar{D} = X$. □

$$\boxed{\text{first-countable} \Longleftarrow \text{second-countable} \Longrightarrow \text{separable}}$$

4.2 (§31) The Separation Axioms

We can strengthen the Hausdorff property (T2) by demanding to be able to separate not only pairs of points, but pairs of points and closed sets, or pairs of closed sets. This leads to regular (T3) and normal (T4) spaces.

$$\boxed{\text{normal} \Longrightarrow \text{regular} \Longrightarrow \text{Hausdorff}}$$

Here "regular" comes from the Latin "regula", originally meaning a straight piece of wood, as in a ruler. Similarly, "normal" comes from "norma", a carpenter's square with four right angles. Its edges are normal, or perpendicular, to one another. The pendulum (of "perpendicular") is another tool for the recognition of vertical lines.

Definition 4.2.1. A topological space X is *regular* if

1. the singleton set $\{x\}$ is closed in X for each $x \in X$, and
2. for each point $x \in X$ and each closed subset $B \subset X$, with $x \notin B$, there exist disjoint open subsets $U, V \subset X$ with $x \in U$ and $B \subset V$.

We then say that U and V *separate* x and B.

Lemma 4.2.2. *Regular spaces are Hausdorff.*

Proof. Given x and y consider $B = \{y\}$. □

We could therefore replace condition (1) by asking that X is Hausdorff. We cannot, however, omit condition (1), since a space with the trivial (indiscrete) topology satisfies (2).

Lemma 4.2.3. *Let X be a space with closed points. Then X is regular if and only if for each point $x \in X$ and neighborhood W of x there is a neighborhood U of x with $\bar{U} \subset W$.*

$$x \in U \subset \bar{U} \subset W \subset X$$

Proof. If X is regular and $x \in W \subset X$, consider $B = X - W$. Let $x \in U$ and $B \subset V$ with U and V open and disjoint. Then $X - V$ is closed and contains U, so $\bar{U} \subset X - V \subset X - B = W$.

Conversely, if $x \in X$ and B closed in X are given, with $x \notin B$, consider $W = X - B$. Then W is a neighborhood of x. If U is a neighborhood of x with $\bar{U} \subset W$ then U and $V = X - \bar{U}$ separate x and B, as required for X to be regular. □

Theorem 4.2.4. *(a) Any subspace of a Hausdorff space is Hausdorff.*
(b) Any product of Hausdorff spaces is Hausdorff.

We have discussed this in exercises. Here is the analogous result for regular spaces.

Theorem 4.2.5. *(a) Any subspace of a regular space is regular.*

(b) Any product of regular spaces is regular.

Proof. (a) Let X be regular and $A \subset X$ a subspace. We already know that A is Hausdorff, so we must show how to separate points and closed subspaces in A. Consider $x \in A$ and $B \subset A$ closed, with $x \notin B$. Since B is closed in A, there is a closed subspace $K \subset X$ with $B = A \cap K$. Then $x \notin K$ (Exercise: Why?), so there exist open and disjoint $U, V \subset X$ with $x \in U$, $K \subset V$. Then $A \cap U$ and $A \cap V$ are open and disjoint subsets of A with $x \in A \cap U$ and $B \subset A \cap V$.

(b) Let $(X_\alpha)_{\alpha \in J}$ be any collection of regular spaces. We already know that $\prod_\alpha X_\alpha$ is Hausdorff. We use the previous lemma to show that $X = \prod_\alpha X_\alpha$ is regular. Let $x = (x_\alpha)_{\alpha \in J} \in X$ and consider any neighborhood W of x. Then there is a basis neighborhood $\prod_\alpha W_\alpha$ of x contained in W, where W_α is a neighborhood of x_α for each $\alpha \in J$, with $W_\alpha = X_\alpha$ for all but finitely many indices α. By regularity of X_α we can choose a neighborhood U_α of x_α with $\bar{U}_\alpha \subset W_\alpha$, for each $\alpha \in J$. When $W_\alpha = X_\alpha$ we can and will choose $U_\alpha = X_\alpha$. Then $U = \prod_\alpha U_\alpha$ is a (basis) neighborhood of x, with

$$\bar{U} \overset{!}{=} \prod_\alpha \bar{U}_\alpha \subset \prod_\alpha W_\alpha \subset W.$$

Here the first equality uses the theorem below. Hence $x \in U \subset \bar{U} \subset W$. This shows that X is regular. □

The previous proof used the following theorem from §19, which we omitted at the time.

Theorem 4.2.6. *Consider subspaces $A_\alpha \subset X_\alpha$ for each $\alpha \in J$. The closure of their product equals the product of their closures:*

$$\overline{\prod_{\alpha \in J} A_\alpha} = \prod_{\alpha \in J} \bar{A}_\alpha \,.$$

Proof. ($\subset$): Let $A = \prod_\alpha A_\alpha$ and $X = \prod_\alpha X_\alpha$. Let $x = (x_\alpha)_\alpha$ be a point in $\bar{A}$. We show that $x \in \prod_\alpha \bar{A}_\alpha$, by showing that $x_\beta \in \bar{A}_\beta$, for each $\beta \in J$. Let V be a neighborhood of x_β in X_β. Then $U = \pi_\beta^{-1}(V)$ is a neighborhood of x in X, hence meets A. Choose $y = (y_\alpha)_\alpha \in A \cap U$. Then $y_\beta \in A_\beta \cap V$. Since V was arbitrary, $x_\beta \in \bar{A}_\beta$, as required.

For the opposite inclusion, see Munkres (page 114/116). □

Definition 4.2.7. A topological space X is *normal* if

1. the singleton set $\{x\}$ is closed in X for each $x \in X$, and
2. for each pair of disjoint closed subsets $A, B \subset X$ there exist disjoint open subsets $U, V \subset X$ with $A \subset U$ and $B \subset V$.

We then say that U and V *separate* A and B.

Lemma 4.2.8. *Normal spaces are regular.*

Proof. Given x and B consider $A = \{x\}$. □

We could replace condition (1) by asking that X is Hausdorff (or regular). We cannot, however, omit condition (1), since a space with the trivial (indiscrete) topology satisfies (2).

Lemma 4.2.9. *Let X be a space with closed points. Then X is normal if and only if for each closed subset $A \subset X$ and open set W containing A there is an open set U containing A with $\bar{U} \subset W$.*

$$A \subset U \subset \bar{U} \subset W \subset X$$

Proof. If X is normal and $A \subset W \subset X$, consider $B = X - W$. Let $A \subset U$ and $B \subset V$ with U and V open and disjoint. Then $X - V$ is closed and contains U, so $\bar{U} \subset X - V \subset X - B = W$.

Conversely, if disjoint and closed A and B in X are given, consider $W = X - B$. Then W is an open set containing A. If U is an open set containing A with $\bar{U} \subset W$ then U and $V = X - \bar{U}$ separate A and B, as required for X to be normal. □

Remark 4.2.10. Counterexamples exist to show that a subspace of a normal space needs not be normal, and a product of (two or more) normal spaces needs not be normal. Hence there is no analogue of Theorem 4.2.5 for normal spaces.

4.3 (§32) (More About) Normal Spaces

We show that for second-countable spaces, regularity and normality are equivalent, while for compact spaces, the Hausdorff property, regularity and normality are all equivalent. Metric spaces satisfy all three of these separation axioms.

Theorem 4.3.1 (Tychonoff (1926)). *Every second-countable regular space is normal.*

Proof. Let X be a regular space with a countable basis $\mathscr{B}$ for its topology, and let A and B be disjoint closed subsets of X. We seek disjoint open subsets U and V of X with $A \subset U$ and $B \subset V$.

For each point $x \in A$, regularity for x and B gives us a neighborhood $U(x)$ of x with closure disjoint from B:

$$x \in U(x) \subset \overline{U(x)} \subset X - B\,.$$

Any smaller neighborhood of x still has closure disjoint from B, so we may assume that $U(x)$ is chosen from the basis $\mathscr{B}$ for the topology. Then the covering $\{U(x) \mid x \in A\}$ of A is a subcollection of $\mathscr{B}$, hence is countable. Choosing an enumeration $n \mapsto U_n$ of this countable collection, we have

$$\{U(x) \mid x \in A\} = \{U_n \mid n \in \mathbb{N}\} \subset \mathscr{B}\,.$$

We now have a countable covering $\{U_n \mid n \in \mathbb{N}\}$ of A, such that each U_n is chosen from $\mathscr{B}$ and $\overline{U_n}$ is disjoint from B.

Similarly, we can find a countable covering $\{V_n \mid n \in \mathbb{N}\}$ of B, such that each V_n is chosen from $\mathscr{B}$ and $\overline{V_n}$ is disjoint from A.

The union $\bigcup_n U_n$ is now an open set containing A and disjoint from B. Similarly, the union $\bigcup_n V_n$ is an open set containing B and disjoint from A. However, $\bigcup_n U_n$ will typically not be disjoint from $\bigcup_n V_n$. To achieve this, we replace U_n and V_n with the following open subsets:

$$U'_n = U_n - \bigcup_{i=1}^{n} \overline{V_i}$$
$$V'_n = V_n - \bigcup_{i=1}^{n} \overline{U_i}\,.$$

Since each $\bar{V}_i$ is closed, and disjoint from A, the set U'_n is open and covers the same part of A as U_n did ($A \cap U_n = A \cap U'_n$). Hence

$$U = \bigcup_n U'_n$$

is an open set containing A. Similarly,

$$V = \bigcup_n V'_n$$

is an open set containing B. To show that $U \cap V = \varnothing$, suppose that $x \in U \cap V$. Then $x \in U'_j$ and $x \in V'_k$ for some $j, k \geq 1$. By symmetry, we may assume that $j \leq k$. Then $x \in U'_j \subset U_j$, but $x \in V'_k \subset V_k - \overline{U_j}$ implies $x \notin \overline{U_j}$. This contradiction shows that U cannot meet V. □

Theorem 4.3.2. *Every metrizable space is normal.*

Proof. Let (X, d) be a metric space. Any metric space has closed points, so condition (1) in the definition of a normal space is satisfied. To establish condition (2), let $A, B \subset X$ be disjoint, closed subsets. If A is empty the condition is trivially satisfied (with $U = \varnothing$ and $V = X$), and similarly if B is empty, so we may assume that A and B are nonempty.

Recall the definition

$$d(x, B) = \inf\{d(x, b) \mid b \in B\}$$

of the distance from x to B. It is continuous as a function of x, and $d(x, B) = 0$ if and only if $x \in B$, since B is closed. Let

$$U = \{x \in X \mid d(x, A) < d(x, B)\}$$
$$V = \{x \in X \mid d(x, A) > d(x, B)\}.$$

Then U and V are disjoint open subsets of X, with $A \subset U$ and $B \subset V$. □

Theorem 4.3.3. *Every compact Hausdorff space is normal.*

Proof. Let X be compact Hausdorff. Points in X are closed, so we must show that disjoint closed subsets $A, B \subset X$ can be separated by disjoint open subsets. Being closed subsets of a compact space, A and B are themselves compact.

We have already considered the case $A = \{x\}$. Let us review the argument: For each $y \in B$ we can separate x from y, i.e., find open sets $U_y \ni x$ and $V_y \ni y$ with $U_y \cap V_y = \varnothing$. The collection $\{V_y \mid y \in B\}$ covers B, which is compact, so there is a finite subcollection

$$\{V_{y_1}, \dots, V_{y_n}\}$$

that also covers B. Then $U = U_{y_1} \cap \dots \cap U_{y_n}$ and $V = V_{y_1} \cup \dots \cup V_{y_n}$ are disjoint open sets with $x \in U$ and $B \subset V$.

Now consider the case of a general, compact A. For each $x \in A$ we can choose disjoint, open sets $U_x \ni x$ and $V_x \supset B$. The collection $\{U_x \mid x \in A\}$ covers A, which is compact, so there is a finite subcollection

$$\{U_{x_1}, \dots, U_{x_m}\}$$

that also covers A. Then $U = U_{x_1} \cup \dots \cup U_{x_m}$ and $V = V_{x_1} \cap \dots \cap V_{x_m}$ are disjoint open sets with $A \subset U$ and $B \subset V$, as required. □

4.4 (§33) The Urysohn Lemma

For normal spaces we can separate disjoint closed subsets by real-valued functions, in the following sense:

Theorem 4.4.1 (Urysohn's lemma)**.** *Let A and B be disjoint, closed subsets of a normal space X. There exists a map*

$$f \colon X \to [0, 1]$$

such that $f(x) = 0$ for all $x \in A$ and $f(x) = 1$ for all $x \in B$.

Proof. A *dyadic number* is a rational number of the form $r = a/2^n$, where a and n integers with $n \geq 0$. The dyadic numbers are dense in $\mathbb{R}$.

For each dyadic number $0 \leq r \leq 1$ we shall construct an open subset $U_r \subset X$, with $A \subset U_r \subset X - B$, so that for each pair of dyadic numbers $0 \leq p < q \leq 1$ we have $\bar{U}_p \subset U_q$.

Let $U_1 = X - B$. Then $A \subset U_1$, so by normality there exists an open U_0 with $A \subset U_0 \subset \bar{U}_0 \subset U_1$. By normality again, there exists an open $U_{1/2}$ with $\bar{U}_0 \subset U_{1/2} \subset \bar{U}_{1/2} \subset U_1$.

Let $n \geq 2$ and assume inductively that we have constructed the U_r for all $0 \leq r \leq 1$ of the form $b/2^{n-1} = 2b/2^n$. We must construct the U_r for r of the form $a/2^n$ with $a = 2b+1$ odd. By induction we have constructed $U_{2b/2^n}$ and $U_{(2b+2)/2^n}$ with $\bar{U}_{2b/2^n} \subset U_{(2b+2)/2^n}$. Using normality we can choose an open $U_{(2b+1)/2^n}$ with

$$\bar{U}_{2b/2^n} \subset U_{(2b+1)/2^n} \subset \bar{U}_{(2b+1)/2^n} \subset U_{(2b+2)/2^n} .$$

Continuing for all natural numbers n, we are done.

Extend the definition of the U_r to all dyadic numbers r, by letting $U_r = \varnothing$ for $r < 0$, and $U_r = X$ for $r > 1$. We still have the key property that $\bar{U}_p \subset U_q$ for all dyadic numbers $p < q$.

Let $x \in X$ and consider the set

$$D(x) = \{r \text{ dyadic} \mid x \in U_r\} .$$

Since $x \notin U_r$ for $r < 0$, the displayed set is bounded below by 0. Since $x \in U_r$ for all $r > 1$, the displayed set contains all dyadic $r > 1$, and is nonempty. Hence the greatest lower bound

$$f(x) = \inf D(x)$$

exists as a real number, and lies in the interval $[0, 1]$.

Claim (1): If $x \in \bar{U}_r$ then $f(x) \leq r$.

If $x \in \bar{U}_r$ then $x \in U_q$ for all $r < q$, so $D(x)$ contains all dyadic numbers greater than r. The dyadic numbers are dense in the reals, so $f(x) \leq r$.

Claim (2): If $x \notin U_r$ then $r \leq f(x)$.

If $x \notin U_r$ then $x \notin U_p$ for all $p < r$, so $D(x)$ contains no dyadic numbers less than r. Hence r is a lower bound for $D(x)$, and $r \leq f(x)$.

Claim (3): f is continuous.

Let $x \in X$ and consider any neighborhood (c, d) in $\mathbb{R}$ of $f(x)$. We shall find a neighborhood U of x with $f(U) \subset (c, d)$.

Choose dyadic numbers p and q with $c < p < f(x) < q < d$. Then $x \notin \bar{U}_p$ by (1), and $x \in U_q$ by (2), so $U = U_q - \bar{U}_p$ is a neighborhood of x.

If $y \in U$ then $y \notin U_p \subset \bar{U}_p$, so $c < p \leq f(y)$. Also $y \in U_q \subset \bar{U}_q$, so $f(y) \leq q < d$. Hence $f(U) \subset (c, d)$. □

4.5 The Hilbert Cube

We review material covered in Exercise 8 of §20.

Definition 4.5.1. The *Hilbert cube* is the product

$$H = \prod_{n=1}^{\infty} [0, \frac{1}{n}]$$

in the product topology.

A point $x \in H$ can be viewed as a sequence $(x_n)_{n=1}^{\infty}$ with

$$0 \le x_n \le \frac{1}{n}$$

for each $n \ge 1$. We consider two different metrics on H. The *uniform metric* $\rho = d_\infty$ is given by

$$\rho(x, y) = \sup_{n \ge 1} |y_n - x_n| \,.$$

The *ℓ^2-metric* d_2 is given by

$$d_2(x, y) = \sqrt{\sum_{n=1}^{\infty} (y_n - x_n)^2} \,.$$

Here the infinite series is bounded by

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} < \infty \,,$$

hence converges. (This ζ-series is easily seen to be bounded by $1+\int_1^\infty 1/x^2\, dx = 1+[-1/x]_1^\infty = 2$. The exact value $\zeta(2) = \pi^2/6$ is due to Euler (1734). Here ζ is the Greek letter 'zeta', and $\zeta(s) = \sum_n 1/n^s$ defines Riemann's zeta-function.) The ℓ^2-metric on H is restricted from the inner product

$$x \cdot y = \sum_n x_n y_n$$

on the complete vector space $\ell^2(\mathbb{R})$ of square-summable sequences. Such complete inner-product spaces are known as *Hilbert spaces.*

Proposition 4.5.2. *The uniform metric ρ, and the ℓ^2-metric d_2 define the same topology on the Hilbert cube H as the product topology. In particular, H is metrizable.*

Proof. Consider any point $x \in H$. Any neighborhood of x in the product topology contains a basis neighborhood of the form

$$\prod_{n=1}^{k} [0, 1/n] \cap (x_n - \epsilon, x_n + \epsilon) \times \prod_{n=k+1}^{\infty} [0, 1/n]$$

for some $\epsilon > 0$ and finite k. (Exercise: Why?) This, in turn, contains the uniform metric neighborhood

$$B_\rho(x, \epsilon) = \{y \in H \mid \sup_n |y_n - x_n| < \epsilon\} \,.$$

Hence the uniform metric topology is equal to or finer than the product topology.

Furthermore,

$$\rho(x, y) \le d_2(x, y)$$

since

$$\sup_n \,(y_n - x_n)^2 \le \sum_n (y_n - x_n)^2 \,.$$

Hence $B_{d_2}(x, \epsilon) \subset B_\rho(x, \epsilon)$, so any uniform metric neighborhood of x contains an ℓ^2-metric neighborhood. Thus the ℓ^2-metric topology is equal to or finer than the uniform metric topology.

It remains to prove that any ℓ^2-neighborhood $B_{d_2}(x, \epsilon)$ of x contains a neighborhood in the product topology. Here $\epsilon > 0$ is given. Since the series $\sum_n 1/n^2$ converges, there is a finite $k \geq 1$ such that

$$\sum_{n=k+1}^{\infty} 1/n^2 < \epsilon^2/2\,.$$

Let $\delta = \epsilon/\sqrt{2k} > 0$. We claim that the basis neighborhood

$$B = \prod_{n=1}^{k} [0, 1/n] \cap (x_n - \delta, x_n + \delta) \times \prod_{n=k+1}^{\infty} [0, 1/n]$$

of x, for the product topology, is contained in $B_{d_2}(x, \epsilon)$. To check this, consider $y \in B$. Then $|y_n - x_n| < \delta$ for $1 \leq n \leq k$, and $0 \leq y_n \leq 1/n$ for $n \geq k+1$. Hence $(y_n - x_n)^2 < \delta^2 = \epsilon^2/2k$ for $1 \leq n \leq k$ and $(y_n - x_n)^2 \leq 1/n^2$ for $n \geq k+1$. Thus

$$d_2(x, y)^2 = \sum_{n=1}^{k} (y_n - x_n)^2 + \sum_{n=k+1}^{\infty} (y_n - x_n)^2 < \sum_{n=1}^{k} \epsilon^2/2k + \sum_{n=k+1}^{\infty} 1/n^2 < \epsilon^2/2 + \epsilon^2/2 = \epsilon^2\,.$$

This implies $d_2(x, y) < \epsilon$, so $B \subset B_{d_2}(x, \epsilon)$. Hence the product topology on H is equal to or finer than the ℓ^2-metric topology. □

Corollary 4.5.3. *The countably infinite product*

$$[0, 1]^\omega = \prod_{n=1}^{\infty} [0, 1]$$

is metrizable.

Proof. We have evident homeomorphisms $[0, 1] \cong [0, 1/n]$, taking x to x/n. Their product defines a homeomorphism

$$[0, 1]^\omega = \prod_n [0, 1] \cong \prod_n [0, 1/n] = H\,.$$

The product topology on the right hand side comes from various metrics, including the uniform metric and the ℓ^2-metric. Hence the product topology on the left hand side is also metrizable. □

For example, the uniform metric ρ on H corresponds to the metric

$$D(x, y) = \sup_n \frac{|y_n - x_n|}{n}$$

on $[0, 1]^\omega$ considered in §20.

We will use this corollary to show that certain spaces are metrizable by embedding them in $[0, 1]^\omega$.

4.6 (§34) The Urysohn Metrization Theorem

Theorem 4.6.1 (Urysohn's metrization theorem)**.** *Every second-countable regular space is metrizable.*

Proof. Let X be a second-countable regular space. By Theorem 4.3.1 this is the same as a second-countable normal space. Hence X is a normal space with a countable basis $\mathscr{B} = \{B_k\}_{k=1}^{\infty}$ for its topology. We shall prove that X is metrizable by embedding it into the metrizable space $[0,1]^{\omega} = \prod_{n=1}^{\infty}[0,1]$, with the product topology.

Claim 1: There is a countable collection $\{f_n\}_{n=1}^{\infty}$ of maps $f_n \colon X \to [0,1]$, such that for any

$$p \in U \subset X$$

with U open there is an f_n in the collection with $f_n(p) = 1$ and $f_n(X - U) \subset \{0\}$.

Consider $\mathscr{B} = \{B_k\}_{k=1}^{\infty}$. For each pair (i,j) of indices with $\bar{B}_i \subset B_j$ use Urysohn's lemma to choose a map $g_{i,j} \colon X \to [0,1]$ with $g_{i,j}(\bar{B}_i) \subset \{1\}$ and $g_{i,j}(X - \bar{B}_j) \subset \{0\}$. Then the collection $\{g_{i,j}\}$ satisfies the claim. To see this, consider $p \in U$ open in X. Since $\mathscr{B}$ is a basis, there is a basis element B_j with $p \in B_j \subset U$. By regularity, there is an open V with $p \in V \subset \bar{V} \subset B_j$, and by the basis property there is a basis element B_i with $p \in B_i \subset V$. Then $\bar{B}_i \subset \bar{V} \subset B_j$, so $p \in \bar{B}_i \subset B_j \subset U$. Then $g_{i,j}$ is defined, and satisfies $g_{i,j}(p) = 1$ and $g_{i,j}(X - U) \subset \{0\}$. We reindex the countable collection $\{g_{i,j}\}_{i,j}$ as $\{f_n\}_{n=1}^{\infty}$.

Define a map

$$F \colon X \to \prod_{n=1}^{\infty}[0,1] = [0,1]^{\omega}$$

by the rule

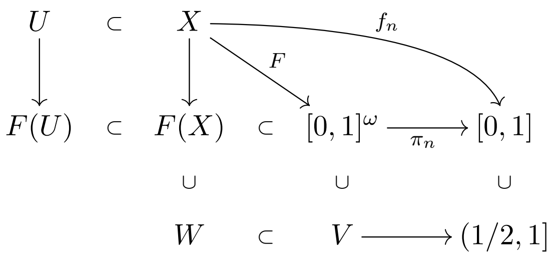
$$F(x) = (f_1(x), f_2(x), \dots)\,.$$

In other words, $\pi_n \circ F = f_n \colon X \to [0,1]$ for each $n \geq 1$.

Claim 2: F is an embedding of X into $[0,1]^{\omega}$.

It is clear that F is continuous, since each component f_n is continuous and $[0,1]^{\omega}$ has the product topology. It is also clear that F is injective, since for $x \neq y$ in X the complement $U = X - \{y\}$ is a neighborhood of x, so there is an index n with $f_n(x) = 1$ and $f_n(X - U) \subset \{0\}$, so $f_n(y) = 0$. Hence the n-th coordinates of $F(x)$ and $F(y)$ are different, so $F(x) \neq F(y)$.

Let $F(X) \subset [0,1]^{\omega}$ be the image of F in the subspace topology. We have proved that $F \colon X \to [0,1]^{\infty}$ is continuous and injective. In order to show that it is an embedding, it remains to show that its corestriction $X \to F(X)$ is an open map. Let $U \subset X$ be open. We must show that $F(U)$ is open in $F(X)$.



Let $q \in F(U)$ be any point. We shall find an open $W \subset F(X)$ with $q \in W \subset F(U)$. Let $p \in U$ be the (unique) point with $F(p) = q$. Choose an index n such that $f_n(p) = 1$ and $f_n(X - U) \subset \{0\}$. Let

$$V = \pi_n^{-1}((1/2, 1])$$

be the open set of sequences $(x_1, x_2, \dots,)$ in $[0,1]$ with $x_n > 1/2$, and let $W = F(X) \cap V$. Then W is open in the subspace topology on $F(X)$.

Claim 3: $q \in W$ and $W \subset F(U)$.

We have $\pi_n(q) = \pi_n(F(p)) = f_n(p) = 1$, so $q \in V$. Since $q \in F(X)$ we get $q \in F(X) \cap V = W$. Let $y \in W$ be any point. Then $y = F(x)$ for some (unique) $x \in X$, and $\pi_n(y) = \pi_n(F(x)) = f_n(x) \in (1/2, 1]$, which implies $x \in U$, since $f_n(X - U) \subset \{0\}$. Hence $y = F(x) \in F(U)$, so $W \subset F(U)$. □

4.7 (§35) The Tietze Extension Theorem

Theorem 4.7.1. *Let A be a closed subspace of a normal space X. Any map $f\colon A \to [0,1]$ (resp. $f\colon A \to \mathbb{R}$) may be extended to a map $g\colon X \to [0,1]$ (resp. $g\colon X \to \mathbb{R}$) with $g|A = f$.*

This is an application of Urysohn's lemma. One may of course replace $[0,1]$ by $[a,b]$, and replace $\mathbb{R}$ by $(0,1)$ or (a,b), for any $a < b$.

4.8 (§36) Embeddings of Manifolds

Definition 4.8.1. An *m-dimensional manifold* is a second-countable Hausdorff space X such that each point $p \in X$ has a neighborhood U that is homeomorphic to an open subset V of $\mathbb{R}^m$.

The composite of such a homeomorphism $U \xrightarrow{\cong} V$ and the inclusion $V \subset \mathbb{R}^m$ defines an embedding

$$g\colon U \longrightarrow \mathbb{R}^m$$

We call g a *coordinate map* and U a *coordinate domain.* The main assumption on X is that it is covered by coordinate domains.

Any neighborhood V of a point q in $\mathbb{R}^m$ contains a neighborhood that is homeomorphic to $\mathbb{R}^m$. Hence we may just as well ask that each point $p \in X$ has a neighborhood that is homeomorphic to $\mathbb{R}^m$. We say that X is *locally homeomorphic to* $\mathbb{R}^m$. The assumptions that X is Hausdorff and second-countable are made to avoid pathological examples.

A 1-dimensional manifold is called a *curve.* A 2-dimensional manifold is called a *surface.*

Lemma 4.8.2. *(a) An open subspace of an m-manifold is an m-manifold.*

(b) The product of an m-manifold and an n-manifold is an $(m+n)$-manifold.

Proof. Clear. □

Example 4.8.3. (a) *Euclidean m-space* $\mathbb{R}^m$,

(b) The *m-sphere*

$$S^m = \{x \in \mathbb{R}^{m+1} \mid \|x\| = 1\}$$

(c) Any *hypersurface*

$$X = f^{-1}(r) = \{x \in \mathbb{R}^{m+1} \mid f(x) = r\}$$

where $f\colon \mathbb{R}^{m+1} \to \mathbb{R}$ is a continuously differentiable (C^1) function and r is a regular value of f. This follows from the implicit function theorem.

Example 4.8.4. Real *projective m-space* is the quotient space

$$\mathbb{R}P^m = S^m/\!\sim,$$

where $\sim$ is the equivalence relation with equivalence classes $[x] = \{x, -x\}$, for $x \in S^m$. Here x and $-x$ are *antipodal* points. There is a quotient map

$$f\colon S^m \longrightarrow \mathbb{R}P^m$$

taking x to its equivalence class $f(x) = [x]$, and defining the topology on $\mathbb{R}P^m$. There is a bijection

$$\mathbb{R}P^m \xrightarrow{\cong} \{\text{lines } L \text{ through } 0 \text{ in } \mathbb{R}^{m+1}\}$$

taking $[x]$ to the line $L = \{rx \in \mathbb{R}^{m+1} \mid r \in \mathbb{R}\}$. We can give the right hand side the topology making this a homeomorphism. This makes $\mathbb{R}P^m$ a compact m-manifold. For $m = 2$ we obtain the projective plane $\mathbb{R}P^2$. This admits an incidence geometry of "points" and "lines", and is "better" than the Euclidean plane $\mathbb{R}^2$ in that any two distinct lines in $\mathbb{R}P^2$ meet in a unique point. (This fails for parallel lines in $\mathbb{R}^2$.)

Example 4.8.5. Let $0 \le k \le n$. The *Grassmann manifold* of k-planes in $\mathbb{R}^n$ is the set

$$Gr_k(\mathbb{R}^n) = \{V \subset \mathbb{R}^n \mid \dim(V) = k\}$$

of k-dimensional vector subspaces of $\mathbb{R}^n$. For example, $\mathbb{R}P^m = Gr_1(\mathbb{R}^{m+1})$, and $Gr_2(\mathbb{R}^4)$ is the space of 2-dimensional planes through 0 in $\mathbb{R}^4$. There is a surjection

$$f\colon GL_n(\mathbb{R}) \longrightarrow Gr_k(\mathbb{R}^n)$$

sending an invertible $n \times n$ matrix $A = (\alpha_1, \ldots, \alpha_k)$ to the subspace of $\mathbb{R}^n$ spanned by the first k column vectors:

$$f(A) = \operatorname{span}\{\alpha_1, \ldots, \alpha_k\} \subset \mathbb{R}^n\,.$$

We can give $Gr_k(\mathbb{R}^n)$ the quotient topology from $GL_n(\mathbb{R})$ (which has the subspace topology from $\mathbb{R}^{n^2}$), and this makes $Gr_k(\mathbb{R}^n)$ a manifold. Its dimension is $m = k(n-k)$. Replacing $GL_n(\mathbb{R})$ with the compact space O_n of orthogonal matrices, once can see that $Gr_k(\mathbb{R}^n)$ is compact.

Example 4.8.6. Let $X \subset \mathbb{R}^n$ be a k-manifold differentiably embedded in $\mathbb{R}^n$. For each point $p \in X$ the tangent space T_pX is then a k-dimensional subspace of $\mathbb{R}^n$. We can think of T_pX as a point $g(x) \in Gr_k(\mathbb{R}^n)$. This defines the *Gauss map*

$$\begin{aligned} g\colon X &\longrightarrow Gr_k(\mathbb{R}^n) \\ p &\longmapsto T_pX \end{aligned}$$

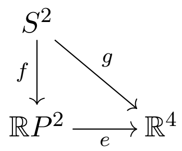
of the embedded manifold X, which is an important tool in differential geometry.

Combining these examples with the lemma above gives rise to a number of examples of manifolds. For example, any open subspace

$$X \subset \mathbb{R}P^{m_1} \times \cdots \times \mathbb{R}P^{m_k}$$

is an $(m_1 + \cdots + m_k)$-manifold.

Example 4.8.7. The real projective plane $\mathbb{R}P^2$ cannot be embedded in $\mathbb{R}^3$, but does admit an embedding in $\mathbb{R}^4$. An explicit embedding $e\colon \mathbb{R}P^2 \to \mathbb{R}^4$ is obtained by factoring $g\colon S^2 \to \mathbb{R}^4$ given by $g(x, y, z) = (yz, xz, xy, x^2 - y^2)$ over the quotient map $f\colon S^2 \to \mathbb{R}P^2$:



Note that $g(x, y, z) = g(x', y', z')$ if and only if $(x', y', z') = \pm(x, y, z)$, so e is a well-defined injective map from a compact space to a Hausdorff space, hence is an embedding.

It is a difficult problem, much studied in the 1950s and 60s, for a given m to determine the minimal N such that $\mathbb{R}P^m$ embeds in $\mathbb{R}^N$. I believe it is not known whether $\mathbb{R}P^6$ can be embedded in $\mathbb{R}^{10}$.

Theorem 4.8.8 (Embedding theorem for compact manifolds)**.** *If X is a compact m-dimensional manifold, then X can be embedded in $\mathbb{R}^N$ for some $N \in \mathbb{N}$.*

For the proof, we will use the existence of partitions of unity.

Definition 4.8.9. The *support* of a function $\phi\colon X \to \mathbb{R}$ is

$$\operatorname{supp}(\phi) = \overline{\{x \in X \mid \phi(x) \neq 0\}}\,.$$

Hence $\phi = 0$ on the open set $X - \operatorname{supp}(\phi)$.

Here ϕ is the Greek letter 'phi', sometimes written φ. The capital letter is Φ.

Definition 4.8.10. Let $\mathscr{C} = \{U_i\}_{i=1}^n$ be a finite, indexed, open covering of X. A *partition of unity* dominated by (or subordinate to) $\mathscr{C}$ is an indexed family $\Phi = \{\phi_i\}_{i=1}^n$ of maps

$$\phi_i\colon X \longrightarrow [0,1]$$

such that $\operatorname{supp}(\phi_i) \subset U_i$ for each $1 \le i \le n$, and $\sum_{i=1}^n \phi(x) = 1$ for each $x \in X$.

At each $x \in X$, the number 1 is written as a sum of non-negative numbers $\phi_1(x), \dots, \phi_n(x)$, where $\phi_i(x) > 0$ only for x in (a set with closure in) U_i.

Theorem 4.8.11 (Existence of finite partitions of unity)**.** *Let X be a normal space, and let $\mathscr{C} = \{U_i\}_{i=1}^n$ be a finite open covering of X. Then there exists a partition of unity $\Phi = \{\phi_i\}_{i=1}^n$ dominated by $\mathscr{C}$.*

The first step of the proof gives the following lemma.

Lemma 4.8.12 (Shrinking lemma)**.** *Let X be a normal space, and let $\mathscr{C} = \{U_i\}_{i=1}^n$ be a finite open covering of X. Then there exists a finite open covering $\mathscr{D} = \{V_i\}_{i=1}^n$ of X such that*

$$\bar{V}_i \subset U_i$$

for each i.

Proof. We define V_k by induction. Let $1 \le k \le n$, and suppose that $V_1, \dots, V_{k-1}$ have been defined so that $\bar{V}_i \subset U_i$ for $1 \le i < k$, and

$$\{V_1, \dots, V_{k-1}, U_k, \dots, U_n\}$$

covers X. This is trivially satisfied for $k = 1$. Consider

$$A = X - (V_1 \cup \dots \cup V_{k-1} \cup U_{k+1} \cup \dots \cup U_n)\,.$$

Then $A \subset U_k$ with A closed and U_k open in X. By normality, we can choose V_k open with $A \subset V_k \subset \bar{V}_k \subset U_k$. This completes the inductive step. □

Proof of existence of partitions of unity. Given the finite open covering $\mathscr{C} = \{U_i\}_{i=1}^n$ apply the shrinking lemma to obtain a finite open covering $\mathscr{D} = \{V_i\}_{i=1}^n$ with $\bar{V}_i \subset U_i$. Apply the shrinking lemma again, to obtain another finite open covering $\mathscr{E} = \{W_i\}_{i=1}^n$ with $\bar{W}_i \subset V_i$.

For each $1 \le i \le n$, use Urysohn's lemma for $\bar{W}_i$ and $X - V_i$ to choose a map $\psi_i\colon X \to [0,1]$ with $\psi_i|\bar{W}_i = 1$ and $\psi_i|X - V_i = 0$. Then $\{x \in X \mid \psi_i(x) \neq 0\} \subset V_i$, so

$$\operatorname{supp}(\psi_i) = \overline{\{x \in X \mid \psi_i(x) \neq 0\}} \subset \bar{V}_i \subset U_i$$

for each i. The sum $\Psi(x) = \sum_{i=1}^{n} \psi_i(x)$ is eveywhere greater than or equal to 1, since any $x \in X$ lies is some W_i, and then $\psi_i(x) = 1$.

To obtain a partition of unity we use Ψ to normalize the functions ψ_i:

$$\phi_i(x) = \frac{\psi_i(x)}{\Psi(x)}$$

for $1 \le i \le n$. Then each ϕ_i is continuous,

$$\text{supp}(\phi_i) = \text{supp}(\psi_i) \subset U_i$$

and $\sum_{i=1}^{n} \phi_i(x) = 1$ for each $x \in X$. □

Here ψ is the Greek letter 'psi', and Ψ is the upper-case form.

Proof of embedding theorem. Let X be a compact m-manifold. Each point $p \in X$ has a neighborhood U_p that is homeomorphic to an open subset $V_p \subset \mathbb{R}^m$. The collection $\{U_p \mid p \in X\}$ is an open cover of X, so by compactness there exists a finite subcover $\{U_{p_1}, \dots, U_{p_n}\}$. Choose embeddings

$$g_i \colon U_{p_i} \longrightarrow \mathbb{R}^m$$

with image V_{p_i}, for each $1 \le i \le n$.

Let $\{\phi_1, \dots, \phi_n\}$ be a partition of unity dominated by $\{U_{p_1}, \dots, U_{p_n}\}$. Recall that $\text{supp}(\phi_i) \subset U_{p_i}$. The function

$$\begin{aligned} U_{p_i} &\longrightarrow \mathbb{R}^m \\ x &\longmapsto \phi_i(x) \cdot g_i(x) \end{aligned}$$

is continuous, and equal to 0 outside $\text{supp}(\phi_i)$. We can extend this to a map

$$h_i \colon X \longrightarrow \mathbb{R}^m$$

by setting $h_i(x) = 0$ if $x \notin \text{supp}(\phi_i)$. Then h_i is well-defined, and continuous on the open sets U_{p_i} and $X - \text{supp}(\phi_i)$ that cover X, hence is continuous on all of X.

An embedding of X is now given by

$$F \colon X \longrightarrow (\mathbb{R} \times \mathbb{R}^m)^n \cong \mathbb{R}^N$$

with

$$F(x) = (\phi_1(x), h_1(x), \dots, \phi_n(x), h_n(x))$$

and $N = n(m+1)$. Since each component of F is continuous, so is F. By assumption X is compact, and $\mathbb{R}^N$ is Hausdorff, so in order to show that F is an embedding it suffices to prove that it is injective.

Consider $x, y \in X$ with $F(x) = F(y)$. Then $\sum_i \phi_i(x) = 1$ so $\phi_i(x) > 0$ for some i. Then $\phi_i(y) = \phi_i(x) > 0$, and $x, y \in \text{supp}(\phi_i) \subset U_{p_i}$. From

$$\phi_i(x) \cdot g_i(x) = h_i(x) = h_i(y) = \phi_i(y) \cdot g_i(y)$$

we deduce $g_i(x) = g_i(y)$. Here g_i is injective, so $x = y$. □

This proof does not give the optimal (minimal) N, but given an embedding $F \colon X \to \mathbb{R}^N$ one can look for projections $\pi \colon \mathbb{R}^N \to \mathbb{R}^{N-1}$ and ask if $\pi \circ F \colon X \to \mathbb{R}^{N-1}$ remains an embedding. For differentiable manifolds X this works to bring N down to $2m+1$. Hassler Whitney (1944) proved that $N = 2m$ is possible (for $m \ge 1$), but to get lower than that, (much) more care is needed.

Chapter 5

The Tychonoff Theorem

5.1 (§37) The Tychonoff Theorem

Theorem 5.1.1 (Tychonoff theorem). *Any product of compact spaces is compact.*

In other words, for any set J and any collection $\{X_\alpha\}_{\alpha \in J}$ of compact topological spaces X_α, the product space

$$\prod_{\alpha \in J} X_\alpha$$

is compact in the product topology.

We proved this for finite J, but the result is also true for infinite J. Proofs typically involve well-ordering, nets or transfinite induction.

5.1.1 The Hilbert cube is compact

A special case of Tychonoff's theorem is that the Hilbert cube

$$H = \prod_{m=1}^{\infty} [0, 1/m]$$

is compact. Here $J = \mathbb{N}$ and $X_m = [0, 1/m]$ for $m \geq 1$.

This can be given an elementary proof (Morris, 1984) using the Cantor set. It suffices to prove that the homeomorphic space

$$[0, 1]^\omega = \prod_{m=1}^{\infty} [0, 1]$$

is compact in the product topology. The Cantor set $C = \bigcap_{n=1}^{\infty} C_n$ is the closed subspace of $[0, 1]$ obtained by iteratively removing middle thirds:

$$C_1 = [0, 1/3] \cup [2/3, 1]$$
$$C_2 = [0, 1/9] \cup [2/9, 3/9] \cup [6/9, 7/9] \cup [8/9, 1]$$
$$\dots$$

We know that C is compact, because it is a closed subset of $[0, 1]$.

The countably infinite product

$$\{0, 1\}^\omega = \prod_{n=1}^{\infty} \{0, 1\}$$

of the discrete two-point space $\{0,1\}$ maps (continuously) onto the unit interval $[0,1]$ by

$$f\colon (a_n)_n \longmapsto \sum_{n=1}^{\infty} \frac{a_n}{2^n}\,.$$

This corresponds to the presentation of real numbers $0 \le x \le 1$ by binary expansions

$$x = 0.b_1b_2b_3\dots$$

where $b_n = a_n$. It also maps onto the Cantor set C by

$$h\colon (a_n)_n \longmapsto \sum_{n=1}^{\infty} \frac{2a_n}{3^n}\,.$$

This corresponds to the (unique) presentation of elements $y \in C$ by ternary (= base 3) expansions

$$y = 0.t_1t_2t_3\dots$$

where $t_n = 2a_n$. In fact $h\colon \{0,1\}^\omega \to C$ is a homeomorphism. The n-th component of the inverse h^{-1} takes a number $y \in C \subset [0,1]$ to $a_n = t_n/2$ where $t_n \in \{0,2\}$ is the n-th term in the ternary expansion of y, which is locally constant on C.

Choosing a bijection $\mathbb{N} \times \mathbb{N} \cong \mathbb{N}$ taking (m,n) to k, we obtain a homeomorphism

$$\prod_{k=1}^{\infty}\{0,1\} \cong \prod_{m=1}^{\infty}\prod_{n=1}^{\infty}\{0,1\}\,.$$

The left hand side is homeomorphism to C, hence compact. For each m we have a surjective map $f\colon \prod_{n=1}^{\infty}\{0,1\} \to [0,1]$. Taking their product we obtain a surjecitve map

$$\prod_{m=1}^{\infty} f\colon \prod_{m=1}^{\infty}\prod_{n=1}^{\infty}\{0,1\} \longrightarrow \prod_{m=1}^{\infty}[0,1]\,.$$

Hence $\prod_{m=1}^{\infty}[0,1]$ is the continuous image of a compact space, and is therefore compact.

5.1.2 The profinite integers

Let $n \in \mathbb{N}$ be a natural number. We say that two integers a and b are *congruent modulo* n, and write $a \equiv b \mod n$, if $b - a$ is a multiple of n, i.e., if $n \mid b - a$. Congruence modulo n is an equivalence relation on $\mathbb{Z}$, and the equivalence class of an integer a is

$$[a]_n = a + n\mathbb{Z} = \{a + kn \mid k \in \mathbb{Z}\}\,.$$

The n integers

$$\{0, 1, 2, \dots, n-1\}$$

are commonly chosen representatives for the n different equivalence classes for this equivalence relation. The set of equivalence classes for congruence modulo n is denoted $\mathbb{Z}/n\mathbb{Z}$, $\mathbb{Z}/(n)$ or $\mathbb{Z}/n$. It is a commutative ring, called the *ring of integers modulo* n, with sum

$$[a]_n + [b]_n = [a+b]_n$$

and product

$$[a]_n \cdot [b]_n = [ab]_n$$

defined by choosing representatives. There is a surjective ring homomorphism

$$\phi_n\colon \mathbb{Z} \to \mathbb{Z}/n$$

taking a to $[a]_n$.

Lemma 5.1.2. *View $\mathbb{Z}$ and $\mathbb{Z}/n$ as discrete topological spaces. The product space*

$$\prod_{n=1}^{\infty} \mathbb{Z}/n = \mathbb{Z}/1 \times \mathbb{Z}/2 \times \mathbb{Z}/3 \times \dots$$

is a compact Hausdorff space. The function

$$\Phi \colon \mathbb{Z} \to \prod_{n=1}^{\infty} \mathbb{Z}/n$$

with components $(\phi_1, \phi_2, \phi_3, \dots)$, taking a to $([a]_1, [a]_2, [a]_3, \dots)$, is an injective, continuous function

Proof. Each $\mathbb{Z}/n$ has only finitely many open subsets, hence any open cover is finite, so $\mathbb{Z}/n$ is compact. By Tychonoff's theorem, the product space $\prod_{n=1}^{\infty} \mathbb{Z}/n$ is also compact. (This is not too hard to prove directly)

Each $\mathbb{Z}/n$ is discrete, hence Hausdorff, so also the product space $\prod_{n=1}^{\infty} \mathbb{Z}/n$ is Hausdorff.

The function Φ is injective, since if $\Phi(a) = \Phi(b)$ then $\phi_n(a) = \phi_n(b)$ for all $n \in \mathbb{N}$, so $b - a$ is divisible by each natural number n. Taking $n > |b - a|$ it follows that $b - a = 0$, so $a = b$.

Each function ϕ_n is continuous, since $\mathbb{Z}$ has the discrete topology. Hence Φ is continuous, since each of its components is continuous and $\prod_{n=1}^{\infty} \mathbb{Z}/n$ has the product topology. □

Let $m, d \in \mathbb{N}$ be natural numbers, and suppose that m is a multiple of d, so that $d \mid m$. The function

$$\rho_{m,d} \colon \mathbb{Z}/m \to \mathbb{Z}/d$$

taking $[a]_m$ to $[a]_n$ is then a well-defined ring homomorphism. We call $\rho_{m,d}(x)$ the *reduction modulo d* of $x \in \mathbb{Z}/m$. Notice that $\rho_{m,d} \circ \phi_m = \phi_d$

$$\begin{array}{ccc} \mathbb{Z} & \xrightarrow{\phi_m} & \mathbb{Z}/m \\ & {\scriptstyle \phi_d} \searrow & \downarrow {\scriptstyle \rho_{m,d}} \\ & & \mathbb{Z}/d \end{array}$$

as functions $\mathbb{Z} \to \mathbb{Z}/d$. Hence the image $\Phi(a) \in \prod_{n=1}^{\infty} \mathbb{Z}/n$ of an integer $a \in \mathbb{Z}$ is a sequence $(x_n)_{n=1}^{\infty}$ with the property that

$$\rho_{m,d}(x_m) = x_d$$

for all $d \mid m$. Let

$$\hat{\mathbb{Z}} = \{(x_n)_{n=1}^{\infty} \in \prod_{n=1}^{\infty} \mathbb{Z}/n \mid \rho_{m,d}(x_m) = x_d \text{ for all } d \mid m\}$$

be the subspace defined by this property. In other words, an element of $\hat{\mathbb{Z}}$ is a sequence $(x_n)_{n=1}^{\infty}$ with x_n an integer modulo n, such that x_d is the reduction modulo d of x_m, for each $d \mid m$.

The set $\hat{\mathbb{Z}}$ is a commutative ring, called the *ring of profinite integers*, with sum

$$(x_n)_{n=1}^{\infty} + (y_n)_{n=1}^{\infty} = (x_n + y_n)_{n=1}^{\infty}$$

and product

$$(x_n)_{n=1}^{\infty} \cdot (y_n)_{n=1}^{\infty} = (x_n \cdot y_n)_{n=1}^{\infty}$$

defined termwise. It is a topological ring, in the sense that the ring operations are continuous.

Lemma 5.1.3. *The ring of profinite integers $\hat{\mathbb{Z}}$ is a closed subspace of the product space $\prod_{n=1}^{\infty} \mathbb{Z}/n$, hence is a compact Hausdorff space.*

Proof. For each $d \mid m$ consider the continuous function

$$f_{m,d} \colon \prod_{n=1}^{\infty} \mathbb{Z}/n \to \mathbb{Z}/d$$

taking $(x_n)_{n=1}^{\infty}$ to $\rho_{m,d}(x_m) - x_d$. The preimage

$$C_{m,d} = f_{m,d}^{-1}([0]_d)$$

of $[0]_d$ is closed, since $\mathbb{Z}/d$ is discrete. Hence the intersection

$$\hat{\mathbb{Z}} = \bigcap_{d \mid m} C_{m,d}$$

is also closed in $\prod_{n=1}^{\infty} \mathbb{Z}/n$. □

Lemma 5.1.4. *The function Φ corestricts to an injective, continuous ring homomorphism*

$$\Psi \colon \mathbb{Z} \to \hat{\mathbb{Z}}$$

taking a to $([a]_n)_{n=1}^{\infty}$, from the discrete space $\mathbb{Z}$ to the compact Hausdorff space $\hat{\mathbb{Z}}$. The image of Ψ is dense in $\hat{\mathbb{Z}}$.

Proof. Let $p = (x_n)_{n=1}^{\infty}$ be any point in $\hat{\mathbb{Z}}$. Any neighborhood V of p contains a basis element $\hat{\mathbb{Z}} \cap U$ for the subspace topology, where

$$U = \prod_{n=1}^{\infty} U_n \subset \prod_{n=1}^{\infty} \mathbb{Z}/n$$

is a basis element for the product topology. Here $x_n \in U_n$ for all n, and $U_n = \mathbb{Z}/n$ for all but finitely many n. Let $N \in \mathbb{N}$ be a common multiple of all the n with $U_n \neq \mathbb{Z}/n$. Choose any integer a with $[a]_N = x_N$. Claim:

$$\Psi(a) \in \Psi(\mathbb{Z}) \cap U \,,$$

so that $\Psi(\mathbb{Z}) \cap V \neq \varnothing$, and $\Psi(\mathbb{Z})$ is dense in $\hat{\mathbb{Z}}$.

To prove the claim, it is enough to prove that $\Phi(a) \in U$, or equivalently, that $[a]_n \in U_n$ for all n. This is clear when $U_n = \mathbb{Z}/n$. When $U_n \neq \mathbb{Z}/n$ we have $n \mid N$, and then

$$[a]_n = \rho_{N,n}([a]_N) = \rho_{N,n}(x_N) = x_n \in U_n \,.$$

□

The map $\Psi \colon \mathbb{Z} \to \hat{\mathbb{Z}}$ is not an embedding of the discrete space $\mathbb{Z}$ in the topological sense. The topology on $\mathbb{Z}$ that makes Ψ an embedding, i.e., the subspace topology from $\hat{\mathbb{Z}}$, may be called the *Fürstenberg topology*.

Theorem 5.1.5 (Euclid, ca. 300 BC)**.** *There are infinitely many prime numbers.*

Proof. Here is Fürstenberg's proof from 1955. The open sets of the Fürstenberg topology on $\mathbb{Z}$ are of the form $\Psi^{-1}(V)$ with V open in $\hat{\mathbb{Z}}$, or equivalently, of the form $\Phi^{-1}(U)$ with U open in $\prod_{n=1}^{\infty} \mathbb{Z}/n$. It follows that each open set in the Fürstenberg topology on $\mathbb{Z}$ is either empty or infinite.

For each prime number p, the subset

$$p\mathbb{Z} = \{kp \mid k \in \mathbb{Z}\}$$

of $\mathbb{Z}$ is closed in the Fürstenberg topology. Consider the subset

$$A = \bigcup_{p \text{ prime}} p\mathbb{Z}$$

of $\mathbb{Z}$. Its complement is

$$X - A = \{\pm 1\},$$

since the only integers not divisible by any primes as the units 1 and -1.

If there is only a finite set of primes, then A is a finite union of closed subsets, hence is closed in $\mathbb{Z}$, so that $X - A = \{\pm 1\}$ is open. This contradicts the fact that the open subsets of $\mathbb{Z}$ are either empty or infinite. □

Chapter 6

Complete Metric Spaces and Function Spaces

6.1 (§43) Complete Metric Spaces

Definition 6.1.1. Let (X, d) be a metric space. A sequence $(x_n)_{n=1}^{\infty}$ of points in X is a *Cauchy sequence* if for each $\epsilon > 0$ there is an $N \in \mathbb{N}$ such that

$$d(x_m, x_n) < \epsilon$$

for all $m, n \geq N$.

Each convergent sequence is a Cauchy sequence.

Definition 6.1.2. A metric space (X, d) is *complete* if each Cauchy sequence in X is convergent.

Lemma 6.1.3. *Any closed metric subspace of a complete metric space is complete.*

Proof. Let (X, d) be complete and $A \subset X$ closed. If $(x_n)_n$ is Cauchy in A with the restricted metric then it is Cauchy in (X, d), hence has a limit in X, which must also be a limit in A, since A is closed. □

Lemma 6.1.4. *A metric space (X, d) is complete if every Cauchy sequence in X has a convergent subsequence.*

Proof. If $(x_n)_n$ is Cauchy and $x_{n_k} \to y$ as $k \to \infty$ then $x_n \to y$ as $n \to \infty$. □

Proposition 6.1.5. *Every compact metric space is complete.*

Proof. Every Cauchy sequence in a compact metric space contains a convergent subsequence, by (sequential) compactness, so by the lemma above it is convergent. □

Theorem 6.1.6. *Euclidean space $\mathbb{R}^n$ is complete (in any of the equivalent metrics coming from a norm).*

Proof. Each Cauchy sequence in $(\mathbb{R}^n, d)$ is bounded, hence lies is a closed and bounded subspace of $\mathbb{R}^n$, which is compact. It therefore has a convergent subsequence, and is itself convergent. □

Remark 6.1.7. A vector space V equipped with a norm $\| - \|$ can be viewed as a metric space, with metric $d(x, y) = \|y - x\|$. The normed vector spaces $(V, \| - \|)$ such that (V, d) is complete are called *Banach spaces* and play a key role in linear analysis (MAT3400/4400), which also serves a foundation for the existence and uniqueness theory for differential equations. A starting point for the theory is the fact that each closed subspace of a Banach space is again complete. In the special case that the complete norm comes from an inner product $\langle -, - \rangle$, with $\|x\| = \sqrt{\langle x, x \rangle}$, we refer to $(V, \langle -, - \rangle)$ as a *Hilbert space*.

Given a metric space (X, d), there exists a universal complete metric space $(\hat{X}, \hat{d})$ containing (X, d) as a dense metric subspace.

Theorem 6.1.8. *Let (X, d) be a metric space. Let* $\text{Cauchy}(X)$ *be the set of Cauchy sequences $(x_n)_n$ in (X, d), and let*

$$\hat{X} = \text{Cauchy}(X)/\sim$$

be the set of equivalence classes of Cauchy sequences, where $(x_n)_n \sim (y_n)_n$ if $d(x_n, y_n) \to 0$ as $n \to \infty$. Define $\hat{d} \colon \hat{X} \times \hat{X} \to \mathbb{R}$ by

$$\hat{d}(\hat{x}, \hat{y}) = \lim_n d(x_n, y_n) .$$

where $\hat{x} = [(x_n)_n]$ and $\hat{y} = [(y_n)_n]$. Then $(\hat{X}, \hat{d})$ is a complete metric space. The map

$$\begin{aligned} e \colon X &\longrightarrow \hat{X} \\ x &\longmapsto [(x, x, x, \dots)] \end{aligned}$$

sending $x \in X$ to the class of the constant sequence at x is an isometric embedding (meaning that $d(x, y) = \hat{d}(e(x), e(y))$ for all $x, y \in X$) with dense image $e(X) \subset \hat{X}$.

Proof. We outline the proof. See Exercise 43.9 in Munkres' book.

(a) $\sim$ is an equivalence relation and $\hat{d}$ is well-defined, by the axioms for the metric d, including the triangle inequality.

(b) $e \colon X \to \hat{X}$ is an isometry, hence identifies (X, d) with the metric subspace $(e(X), \hat{d})$ of $(\hat{X}, \hat{d})$ obtained by restricting the metric $\hat{d}$ to $e(X) \subset \hat{X}$.

(c) $e(X)$ is dense in $\hat{X}$, since for any $\hat{x} = [(x_n)_n]$ in $\hat{X}$ the sequence $(e(x_n))_n$ in $e(X)$ converges to $\hat{x}$.

(d) Every Cauchy sequence in the dense subset $e(X)$ of $\hat{X}$ converges in $\hat{X}$.

(e) This implies that $\hat{X}$ is complete. □

Definition 6.1.9. The complete metric space $(\hat{X}, \hat{d})$ is called the *completion* of (X, d).

The completion is well-defined up to unique isometry.

Theorem 6.1.10. *Let $g \colon (X, d) \to (Y, d_Y)$ be any isometric embedding from a metric space (X, d) to a complete metric space (Y, d_Y). There is a unique factorization of g as a composite*

$$g = f \circ e \colon (X, d) \xrightarrow{e} (\hat{X}, \hat{d}) \xrightarrow{f} (Y, d_Y)$$

where e is the canonical isometric embedding of (X, d) in $(\hat{X}, \hat{d})$, and f is also an isometric embedding.

Proof. We outline the proof. See Exercise 43.10 in Munkres' book.

For $\hat{x} = [(x_n)_n]$ in $\hat{X}$, note that $(g(x_n))_n$ is a Cauchy sequence in the complete metric space (Y, d_Y). Let $f(\hat{x}) = \lim_n g(x_n)$. □

Example 6.1.11. We have used the real numbers in our definition of a metric, and hence of completeness, so the following example may appear a bit circular.

Let $X = \mathbb{Q}$ be the set of rational numbers, with metric

$$d_\infty(x, y) = |y - x|$$

given by the absolute value. The usual inclusion $g\colon (\mathbb{Q}, d_\infty) \to (\mathbb{R}, d)$, where $Y = \mathbb{R}$ has the usual metric, factors as

$$(\mathbb{Q}, d_\infty) \xrightarrow{e} (\hat{\mathbb{Q}}, \hat{d}_\infty) \xrightarrow{f} (\mathbb{R}, d)\,.$$

Since $\mathbb{Q}$ is dense in $(\mathbb{R}, d)$, it follows that the isometric embedding f is surjective, hence an isometric homeomorphism. We can therefore recover the real numbers $\mathbb{R}$, with the usual metric, as the completion $\hat{\mathbb{Q}}$ of the rational numbers with respect to the metric d_∞.

Example 6.1.12. Let p be any prime number. A rational number $x = A/B$ can be written in the form ap^n/b by collecting any factors p in A or B in the power p^n, where n may be any integer. The *p-adic norm*

$$| - |_p \colon \mathbb{Q} \longrightarrow \mathbb{R}$$

is defined by

$$\Big|\frac{ap^n}{b}\Big|_p = \frac{1}{p^n}$$

for a and b integers not divisible by p, and n any integer. We also set $|0|_p = 0$. Then $|x|_p \geq 0$ for all $x \in \mathbb{Q}$, $|-x|_p = |x|_p$, and the triangle inequality

$$|x + y|_p \leq |x|_p + |y|_p$$

can be proved to hold. We can therefore define a metric

$$d_p \colon \mathbb{Q} \times \mathbb{Q} \longrightarrow \mathbb{R}$$

on $\mathbb{Q}$, called the *p-adic metric*, by $d_p(x) = |y - x|_p$.

The metric space $(\mathbb{Q}, d_p)$ has a very different topology than the metric subspace $(\mathbb{Q}, d_\infty)$ of $\mathbb{R}$. For example, the sequence

$$1, p, p^2, p^3, \dots$$

converges to 0 in $(\mathbb{Q}, d_p)$, while it is unbounded and diverges in $(\mathbb{Q}, d_\infty) \subset (\mathbb{R}, d)$.

Definition 6.1.13. The completion $(\hat{\mathbb{Q}}, \hat{d}_p)$ of the rational numbers with respect to the p-adic metric is denoted $\mathbb{Q}_p$, and is called the *p-adic numbers*. It is a complete metric space, containing $(\mathbb{Q}, d_p)$ as a dense metric subspace. A concrete model for $\mathbb{Q}_p$ is given by the formal p-adic expansions

$$\sum_n a_n p^n$$

where $n \in \mathbb{Z}$, $a_n \in \{0, 1, \dots, p-1\}$, and $a_n = 0$ for all sufficiently negative n. The subset of such formal sums where $a_n = 0$ for $n < 0$ is called the set of *p-adic integers*, denoted $\mathbb{Z}_p$.

One can prove that $\mathbb{Q}_p$ is a field of characteristic zero, containing $\mathbb{Q}$ as a subfield, and $\mathbb{Z}_p$ is a subring of $\mathbb{Q}_p$, contaning $\mathbb{Z}$ as a subring. We therefore call $\mathbb{Q}_p$ the field of p-adic numbers, and $\mathbb{Z}_p$ the ring of p-adic integers. The closure of $\mathbb{Z}$ in $\mathbb{Q}_p$ is equal to $\mathbb{Z}_p$. There is also a description of $\mathbb{Z}_p$ as the inverse limit

$$\mathbb{Z}_p = \lim_n \mathbb{Z}/(p^n)$$

of the rings of integers modulo p^n, where the limit is formed in the algebraic/categorical sense.

Remark 6.1.14. An important aspect of modern number theory is to treat the completions $\mathbb{R}$ and $\mathbb{Q}_p$ of $\mathbb{Q}$, with respect to the metrics d_∞ and d_p for all primes p, on an equal footing. In other words, anything arithmetic that can be done with real coefficients should also be done with p-adic coefficients. This leads to important 'reciprocity' results, the first glimpse of which is the formula

$$|x|_\infty \cdot \prod_p |x|_p = 1$$

for any $x \in \mathbb{Q} - \{0\}$. The Langlands program takes the point of view that Galois groups should not be studied only through their real representations (actions on finite-dimensional real vector spaces), but also through their p-adic representations for all primes p.

6.2 (§45) Compactness in Metric Spaces

Definition 6.2.1. A metric space (X, d) is *totally bounded* if for every $\epsilon > 0$ there is a finite covering of X by ϵ-balls.

Proposition 6.2.2. *Every compact metric space is totally bounded.*

Proof. Let (X, d) be a compact metric space and consider any $\epsilon > 0$. The collection of all ϵ-balls is an open covering of X. By compactness there exists a finite subcover, which is a finite covering of X by ϵ-balls. □

Theorem 6.2.3. *A metric space (X, d) is compact if and only if it is complete and totally bounded.*

Proof. We have proved that a compact metric space is complete and totally bounded. Conversely, we will prove that a complete and totally bounded is sequentially compact. This implies that it is compact, as we proved in §28 (Theorem 3.6.3).

Let $(x_n)_{n=1}^\infty$ be any sequence of points in X. We shall construct a Cauchy subsequence $(x_{n_k})_{k=1}^\infty$. By the assumed completeness of X, this will be a convergent subsequence.

First cover X by finitely many balls of radius 1. At least one of these, call it B_1, will contain x_n for infinitely many n. Let

$$J_1 = \{n \in \mathbb{N} \mid x_n \in B_1\}$$

be this infinite set of indices.

Inductively suppose, for some $k \geq 1$, that we have chosen a ball B_k of radius $1/k$ that contains x_n for all n in an infinite set $J_k \subset \mathbb{N}$. Cover X by finitely many balls of radius $1/(k+1)$. At least one of these, call it B_{k+1}, will contain x_n for infinitely many $n \in J_k$. Let

$$J_{k+1} = \{n \in J_k \mid x_n \in B_{k+1}\}$$

be this infinite set. Continue for all k, to get an infinite descending sequence of infinite sets:

$$J_1 \supset \cdots \supset J_k \supset J_{k+1} \supset \ldots .$$

Choose $n_1 \in J_1$. Inductively suppose, for some $k \geq 1$, that we have chosen $n_k \in J_k$. Since J_{k+1} is infinite, we can choose an $n_{k+1} \in J_{k+1}$ with $n_{k+1} > n_k$. Continue for all k. The sequence

$$n_1 < \cdots < n_k < n_{k+1} < \ldots$$

is strictly increasing, so $(x_{n_k})_{k=1}^\infty$ is a subsequence of $(x_n)_{n=1}^\infty$.

We claim that it is a Cauchy sequence. Let $\epsilon > 0$ and choose k with $1/k < \epsilon/2$. For all $i, j \geq k$ we have $n_i \in J_i \subset J_k$ and $n_j \in J_j \subset J_k$, so $x_{n_i}, x_{n_j} \in B_k$. Since B_k has diameter $\leq 2/k < \epsilon$, we get that $d(x_{n_i}, x_{n_j}) < \epsilon$. Hence $(x_{n_k})_{k=1}^\infty$ is Cauchy. □

6.3 (§46) Pointwise and Compact Convergence

6.3.1 Pointwise and uniform convergence

Let X and Y be topological spaces, and consider the set

$$\mathrm{Func}(X,Y) = Y^X \cong \prod_X Y$$

of functions from X to Y. Here $f\colon X \to Y$ corresponds to the X-indexed sequence $(f(x))_x$ of points in Y. Evaluation of a function at x corresponds to projection to the x-th factor, for each $x \in X$.

Definition 6.3.1. For $x \in X$ and $U \subset Y$ open let

$$S(x,U) = \{f\colon X \to Y \mid f(x) \in U\}\,.$$

This corresponds to $\pi_x^{-1}(U)$, where $\pi_x\colon \prod_X Y \to Y$ projects to the x-th factor. The collection of these sets $S(x,U)$ is a subbasis for the *topology of pointwise convergence* on $\mathrm{Func}(X,Y)$, corresponding to the product topology on $\prod_X Y$. The associated basis consists of the finite intersections

$$S(x_1,U_1) \cap \dots \cap S(x_n,U_n)$$

for $x_1,\dots,x_n \in X$ and $U_1,\dots,U_n \subset Y$ open.

Lemma 6.3.2. *A sequence $(f_n)_{n=1}^\infty$ of functions $f_n\colon X \to Y$ converges to $f\colon X \to Y$ in the topology of pointwise convergence if and only if $f_n(x) \to f(x)$ as $n \to \infty$ for each $x \in X$.*

Now suppose that (Y,d) is a metric space.

Definition 6.3.3. For $f\colon X \to Y$ and $\epsilon > 0$ let

$$B(f,\epsilon) = \{g\colon X \to Y \mid \sup_{x\in X} d(f(x),g(x)) < \epsilon\}$$

be the set of functions $g\colon X \to Y$ such that the distances $d(f(x),g(x))$ for $x \in X$ are bounded above by a number less than ϵ. For $\epsilon \le 1$ this is the same set as the ϵ-ball

$$B_{\bar\rho}(f,\epsilon) = \{g\colon X \to Y \mid \bar\rho(f,g) < \epsilon\}$$

for the uniform metric

$$\bar\rho(f,g) = \sup_{x\in X} \bar d(f(x),g(x))\,,$$

where $\bar d(y_1,y_2) = \min\{d(y_1,y_2),1\}$ is the standard bounded metric on Y associated to d. The collection of subsets $B(f,\epsilon) \subset \mathrm{Func}(X,Y)$ is a basis for the *topology of uniform convergence*, which equals the topology associated to the uniform metric $\bar\rho$.

Whenever we discuss the uniform topology on $\mathrm{Func}(X,Y)$ we assume that Y is equipped with a metric.

Lemma 6.3.4. *A sequence $(f_n)_{n=1}^\infty$ of functions $f_n\colon X \to Y$ converges to $f\colon X \to Y$ in the topology of uniform convergence if and only if the sequence f_n converges uniformly to f as $n \to \infty$.*

To say that $f_n \to f$ pointwise as $n \to \infty$ is equivalent to for each $\epsilon > 0$ and $x \in X$ there is an $N \in \mathbb{N}$ such that for each $n \geq N$ we have $d(f_n(x), f(x)) < \epsilon$,

while to say that $f_n \to f$ uniformly as $n \to \infty$ is equivalent to

for each $\epsilon > 0$ there is an $N \in \mathbb{N}$ such that for each $x \in X$ and $n \geq N$ we have $d(f_n(x), f(x)) < \epsilon$.

The latter condition is (strictly) stronger, since N can only depend on ϵ, while in the former condition N may also depend on x.

Example 6.3.5. Let $X = [0, 1]$ and $Y = \mathbb{R}$. Then sequence $f_n(x) = x^n$ converges pointwise, but not uniformly, to the function f given by $f(x) = 0$ for $x \in [0, 1)$ and $f(1) = 1$.

Theorem 6.3.6 (Uniform limit theorem). *The uniform limit of a sequence of continuous functions is continuous. In other words, if the sequence $(f_n)_{n=1}^{\infty}$ of functions $f_n \colon X \to Y$ converges to $f \colon X \to Y$ in the topology of uniform convergence, and each f_n is continuous, then f is continuous.*

Proof. We show that f is continuous at each point $p \in X$. Let $\epsilon > 0$. By the assumption of uniform convergence, there is an $N \in \mathbb{N}$ such that $\sup_{x \in X} d(f(x), f(x_n)) < \epsilon/3$ for all $n \geq N$. In particular this holds for $n = N$. The function f_N is continuous at p, so there is a neighborhood U of p such that $f_N(U) \subset B_d(f_N(p), \epsilon/3)$. Then for all $q \in U$,

$$d(f(p), f(q)) \leq d(f(p), f_N(p)) + d(f_N(p), f_N(q)) + d(f_N(q), f(q)) < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon\,.$$

Hence f is continuous at p. □

Definition 6.3.7. Let $\mathscr{C}(X, Y) = \mathrm{Map}(X, Y)$ be the set of maps (= continuous functions) $f \colon X \to Y$.

Corollary 6.3.8. *The subset* $\mathrm{Map}(X, Y)$ *is closed in* $\mathrm{Func}(X, Y)$ *with the uniform topology, but not (in general) with the pointwise topology.*

The topology of uniform convergence is finer than the topology of pointwise convergence. There is an intermediate *topology of compact convergence*, or uniform convergence on compact sets, defined so that $f_n \to f$ uniformly on compact sets as $n \to \infty$ if any only if for each compact $C \subset X$ we have $f_n|C \to f|C$ uniformly as $n \to \infty$.

$$\text{(pointwise convergence)} \subset \text{(compact convergence)} \subset \text{(uniform convergence)}$$

Example 6.3.9. Consider functions $\mathbb{R} \to \mathbb{R}$. The Taylor polynomials

$$f_n(x) = \sum_{k=0}^{n} \frac{x^k}{k!}$$

converge pointwise, and uniformly on compact subsets, but not uniformly, to the exponential function $f(x) = e^x$.

If X is compact, then the topology of compact convergence equals the topology of uniform convergence, while if X is discrete then the topology of compact convergence equals the topology of pointwise convergence. For a large class of topological spaces X, called *compactly generated* spaces, the subset $\mathrm{Map}(X, Y)$ is closed in $\mathrm{Func}(X, Y)$ with the compact convergence topology. Hence for compactly generated X and metric Y, if $f_n \to f$ uniformly on compact sets and each f_n is continuous then f is continuous. The class of compactly generated spaces plays a significant role in modern algebraic topology. We omit to discuss it only due to a lack of time.

The pointwise, compact convergence and uniform topologies restrict to give topologies on the subspace

$$\mathscr{C}(X, Y) = \mathrm{Map}(X, Y) \subset \mathrm{Func}(X, Y) = Y^X$$

of maps $f\colon X \to Y$. It turns out that the topology of compact convergence on $\mathscr{C}(X, Y)$ will not depend on the choice of metric on Y, and can be extended to the case of arbitrary topological spaces Y. This construction, called the compact-open topology, is due to Ralph Fox (1945). (Fox is known for his work in knot theory, and was the PhD advisor of John Milnor.)

6.3.2 The compact-open topology

Let X and Y be (arbitrary) topological spaces.

Definition 6.3.10. For $C \subset X$ compact and $U \subset Y$ open, let

$$S(C, U) = \{f \in \mathscr{C}(X, Y) \mid f(C) \subset U\}.$$

The collection of these subsets $S(C, U) \subset \mathscr{C}(X, Y)$ is a subbasis for the *compact-open topology.* The associated basis consists of the finite intersections

$$S(C_1, U_1) \cap \cdots \cap S(C_n, U_n)$$

where $C_1, \ldots, C_n \subset X$ are compact and $U_1, \ldots, U_n \subset Y$ are open.

The compact-open topology on $\mathscr{C}(X, Y)$ is finer than the topology of pointwise convergence, since one-point sets are compact.

Theorem 6.3.11. *Let X be a topological space and (Y, d) a metric space. On $\mathscr{C}(X, Y)$ the compact-open topology equals the topology of compact convergence.*

We omit the proof.

Corollary 6.3.12. *Let Y be a metric space. The compact convergence topology on $\mathscr{C}(X, Y)$ only depends on the underlying topology on Y. If X is compact, the uniform topology on $\mathscr{C}(X, Y)$ only depends on the underlying topology on Y.*

6.3.3 Joint continuity

A function

$$f\colon X \times Y \to Z$$

corresponds to a function

$$g\colon X \to \mathrm{Func}(Y, Z)$$

defined by the equation $g(x)(y) = f(x, y)$, and conversely. In this situation, we may call f the *left adjoint* and g the *right adjoint*

Example 6.3.13. Let $X = Y = Z = \mathbb{R}$, with

$$f(x, y) = \begin{cases} \dfrac{2xy}{x^2 + y^2} & \text{if } (x, y) \neq 0, \\ 0 & \text{if } (x, y) = 0. \end{cases}$$

If $(x, y) = (r\cos\theta, r\sin\theta)$ with $r > 0$ then $f(x, y) = \sin(2\theta)$ is independent of r, so if $(x, y) \to (0, 0)$ along a line of slope θ, then $f(x, y) \to \sin(2\theta)$. Thus f is not continuous at 0. However, the adjoint function g sends $x \neq 0$ to

$$g(x)\colon y \longmapsto \frac{2xy}{x^2 + y^2}$$

and $x = 0$ to

$$g(0)\colon 0 \longmapsto 0\,,$$

so that $g(x)$ is continuous for each $x \in \mathbb{R}$.

If f is continuous, each function $g(x)\colon Y \to Z$ is continuous, so that g is a function

$$g\colon X \to \mathscr{C}(Y, Z)\,.$$

We shall show that g itself is continuous if we give $\mathscr{C}(Y, Z)$ the compact-open topology, and that the converse holds e.g. if X is locally compact Hausdorff.

Theorem 6.3.14. *Let X, Y and Z be spaces, and give $\mathscr{C}(Y, Z)$ the compact-open topology. If $f\colon X \times Y \to Z$ is continuous, then so is the right adjoint function $g\colon X \to \mathscr{C}(Y, Z)$.*

Proof. Consider a point $p \in X$ and a subbasis element $S(C, U) \subset \mathscr{C}(Y, Z)$ that contains $g(p)$, so that $g(p)(C) \subset U$, or equivalently, $f(\{p\} \times C) \subset U$. We wish to find a neighborhood V of p such that $g(V) \subset S(C, U)$.

By continuity of f, the preimage $f^{-1}(U)$ is an open subset of $X \times Y$ that contains $\{p\} \times C$. Then $f^{-1}(U) \cap X \times C$ is an open subset in $X \times C$ that contains $\{p\} \times C$. By compactness of C, using the "tube lemma" of §26 (Lemma 3.4.23), there is a neighborhood V of $\{p\}$ such that $V \times C \subset f^{-1}(U) \cap X \times C$. Then for $q \in V$ we have $f(\{q\} \times C) \subset U$, so $g(q)(C) \subset U$ and $g(q) \in S(C, U)$. Hence $g(V) \subset S(C, U)$. □

Corollary 6.3.15. *The* inclusion map

$$\eta\colon X \to \mathscr{C}(Y, X \times Y)$$

given by $\eta(x)(y) = (x, y)$ is continuous.

Proposition 6.3.16. *Let Y be a locally compact Hausdorff space, and give $\mathscr{C}(Y, Z)$ the compact-open topology. The* evaluation map

$$\epsilon\colon \mathscr{C}(Y, Z) \times Y \to Z$$

given by $\epsilon(g, y) = g(y)$ is continuous.

Proof. Let $(g, p) \in \mathscr{C}(Y, Z) \times Y$ and consider any neighborhood U of $\epsilon(g, p) = g(p) \in Z$. We wish to find a neighborhood W of (g, p) such that $\epsilon(W) \subset U$.

By the continuity of g, the preimage $g^{-1}(U)$ is a neighborhood of $p \in Y$. By the assumption that Y locally compact and Hausdorff (recall the local nature of local compactness, see §29, Theorem 3.7.13), there is a neighborhood V of p with compact closure $\bar{V} \subset g^{-1}(U)$. Then

$$W = S(\bar{V}, U) \times V \subset \mathscr{C}(Y, Z) \times Y$$

is an open subset containing (g, p), since $g(\bar{V}) \subset U$ and $p \in V$. Furthermore, if $(h, q) \in S(\bar{V}, U) \times V$ then $\epsilon(h, q) = h(q) \in U$, so $\epsilon(W) \subset U$. □

Theorem 6.3.17. *Let X, Y and Z be spaces, with Y locally compact and Hausdorff, and give $\mathscr{C}(Y,Z)$ the compact-open topology. If $g\colon X \to \mathscr{C}(Y,Z)$ is continuous, then so is the left adjoint function $f\colon X \times Y \to Z$.*

Proof. This follows from the proposition, since f is the composite

$$X \times Y \xrightarrow{g \times id} \mathscr{C}(Y,Z) \times Y \xrightarrow{\epsilon} Z\,.$$

□

Corollary 6.3.18. *Let Y be locally compact, and give $\mathscr{C}(Y,Z)$ the compact-open topology. There are one-to-one correspondences*

$$\mathscr{C}(X \times Y, Z) \cong \mathscr{C}(X, \mathscr{C}(Y,Z)) \cong \mathscr{C}(Y \times X, Z) \tag{6.1}$$

$$f \leftrightarrow g \qquad\qquad \leftrightarrow h \tag{6.2}$$

between the maps

$$f\colon X \times Y \longrightarrow Z$$
$$g\colon X \longrightarrow \mathscr{C}(Y,Z)$$
$$h\colon Y \times X \longrightarrow Z$$

specified by

$$f(x,y) = g(x)(y) = h(y,x)$$

in Z for $x \in X$ and $y \in Y$. If also X is locally compact Hausdorff, then there is also a one-to-one correspondence

$$\mathscr{C}(Y \times X, Z) \cong \mathscr{C}(Y, \mathscr{C}(X,Z))$$
$$h \leftrightarrow k$$

where $\mathscr{C}(X,Z)$ has the compact-open topology and the map

$$k\colon Y \longrightarrow \mathscr{C}(X,Z)$$

is specified by $k(y)(x) = h(y,x)$.

The *unit interval* $I = [0,1]$ in the subspace topology from $\mathbb{R}$ is compact and Hausdorff, and in particular is locally compact Hausdorff.

Definition 6.3.19. Let $\alpha, \beta\colon X \to Z$ be maps. A *homotopy* from α to β is a map $f\colon X \times I \to Z$ such that $f(x,0) = \alpha(x)$ and $f(x,1) = \beta(x)$ for all $x \in X$. We then write $f\colon \alpha \simeq \beta$. We say that α and β are *homotopic* if there exists a homotopy from α to β.

Equivalently, a homotopy from α to β is a map $g\colon X \to \mathscr{C}(I,Z)$ such that $g(x)(0) = \alpha(x)$ and $g(x)(1) = \beta(x)$ for all $x \in X$, or a map $h\colon I \times X \to Z$ such that $h(0,x) = \alpha(x)$ and $h(1,x) = \beta(x)$ for all $x \in X$. If X is locally compact Hausdorff, a homotopy from α to β is equivalent to a map $k\colon I \to \mathscr{C}(X,Z)$ with $k(0) = \alpha$ and $k(1) = \beta$, i.e., a path from α to β in the space $\mathscr{C}(X,Z)$.

Homotopy theory is the part of (algebraic) topology that views homotopic maps α and β as being identical, i.e., only considers homotopy classes $[\alpha] = \{\beta\colon \alpha \simeq \beta\}$ of maps. This leads to a weaker notion of isomorphism than homeomorphism (topological equivalence), namely homotopy equivalence: Two spaces X and Y are homotopy equivalent if there exist maps

$\phi\colon X \to Y$ and $\psi\colon Y \to X$ such that the composite $\psi\phi\colon X \to X$ is homotopic to the identity $id\colon X \to X$, and the composite $\phi\psi\colon Y \to Y$ is homotopic to the identity $id\colon Y \to Y$.

For general topological spaces, homotopy equivalence is a much weaker condition than homeomorphism. However, for some classes of spaces, such as manifolds, the relationship is surprisingly much closer. For instance, the Poincaré conjecture asserts that any 3-manifold that is homotopy equivalent to S^3 is in fact homeomorphic to S^3. This was famously proved by Grigori Perelman in 2003. The generalized Poincaré conjecture, that any m-manifold that is homotopy equivalent to S^m is homeomorphic to S^m is easy for $m = 1$, classical for $m = 2$, was proved by Stephen Smale (1961) for $m \geq 5$, and by Michael Freedman (1982) for $m = 4$.

Chapter 7

The Fundamental Group

7.1 (§51) Homotopy of Paths

7.1.1 Line integrals

Let $X \subset \mathbb{R}^2$ be an open region in the plane, and consider a continuous vector field v defined on X. Here $v\colon X \to \mathbb{R}^2$ is a map giving a vector $v(x) \in \mathbb{R}^2$ at each point $x \in X$. Consider also a continuously differentiable path $f\colon [0,1] \to X$, parametrizing a curve C in X from $f(0) = x_0$ to $f(1) = x_1$. The line integral of v along C is then defined to be

$$\int_C v = \int_0^1 v(f(s)) \cdot f'(s)\, ds$$

If we write $v(x) = (v_1(x), v_2(x))$ and $f(s) = (f_1(s), f_2(s))$, then $v(f(s)) \cdot f'(s) = v_1(f(s))f_1'(s) + v_2(f(s))f_2'(s)$.

If $v = \nabla p$ is the gradient field of a potential, i.e., a continuously differentiable function $p\colon X \to \mathbb{R}$ with $\partial p/\partial x_1 = v_1$ and $\partial p/\partial x_2 = v_2$, then $v(f(s)) \cdot f'(s) = (p \circ f)'(s)$ by the chain rule, and

$$\int_0^1 v(f(s)) \cdot f'(s)\, ds = \int_0^1 (p \circ f)'(s)\, ds = \big[p \circ f\big]_0^1 = p(x_1) - p(x_0)$$

only depends on p and the end-points x_0 and x_1 of C. In particular, it does not depend on the choice of curve leading from x_0 to x_1 (other than that such a curve exists). In this case we say that v is exact, following Poincaré (1899).

If v is itself continuously differentiable, it may happen that $\partial v_1/\partial x_2 = \partial v_2/\partial x_1$, even if v is not actually the gradient of a potential defined over all of X. We call such a vector field closed. Exact vector fields are closed, since $\partial v_1/\partial x_2 = \partial^2 p/\partial x_1 \partial x_2 = \partial v_2/\partial x_1$. The Poincaré lemma states that any closed vector field is locally the gradient of a two times continuously differentiable function, but these local potentials may not fit together to a global potential.

Example 7.1.1. For the vector field

$$v(x_1, x_2) = \frac{(-x_2, x_1)}{x_1^2 + x_2^2}$$

on $X = \mathbb{R}^2 - \{(0,0)\}$ we have

$$\frac{\partial v_1}{\partial x_2} = \frac{x_2^2 - x_1^2}{(x_1^2 + x_2^2)^2} = \frac{\partial v_2}{\partial x_1}\,,$$

so v is closed. However, for the loop $f(s) = (\cos(2\pi s), \sin(2\pi s))$ parametrizing the unit circle C,

$$\int_C v = \int_0^1 v(f(s)) \cdot f'(s)\, ds = \int_0^1 2\pi\, ds = 2\pi \neq 0$$

cannot be $p(x_1) - p(x_0)$ for any potential p, since $x_0 = f(0) = (1,0)$ is equal to $x_1 = f(1) = (1,0)$. Hence v is not exact.

However, for closed vector fields v the line integral $\int_C v$ is still somewhat independent of the curve C leading from x_0 to x_1. More precisely, the line integral is invariant under suitable deformations of C. If C_0 and C_1 are curves from x_0 to x_1 that together form a simple closed curve bounding a region D in the plane, with $D \subset X$, then Green's theorem tells us that

$$\int_{C_0} v - \int_{C_1} v = \iint_D \left(\frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2}\right) dx_1 dx_2 \,.$$

If v is closed, then the integrand over D is zero, so $\int_{C_0} v = \int_{C_1} v$. The key role of the geometric hypothesis on C_0, C_1 and the region D can be expressed in terms of parametrizations $f_0 \colon [0,1] \to X$ of C_0 and $f_1 \colon [0,1] \to X$ of C_1. We can deform C_0 through D, keeping the endpoints fixed, to end up at C_1. In terms of the parametrizations, there is a continuous family of functions

$$f_t \colon [0,1] \longrightarrow X$$

for $0 \le t \le 1$, starting at f_0 and ending at f_1, such that $f_t(0) = x_0$ and $f_t(1) = x_1$ for all t. The continuity of the family $(f_t)_t$ is more conveniently expressed in terms of the adjoint map

$$F \colon [0,1] \times [0,1] \longrightarrow X$$
$$F(s,t) = f_t(s) \,,$$

subject to $F(s,0) = f_0(s)$, $F(s,1) = f_1(s)$ for all $s \in [0,1]$ and $F(0,t) = x_0$, $F(1,t) = x_1$ for all $t \in [0,1]$. We call F a path homotopy from f_0 to f_1. If F is continuously differentiable, we can pull the vector field v on X back to a vector field $w = (w_1, w_2)$ on $[0,1] \times [0,1]$, given by

$$(w_1(s,t), w_2(s,t)) = (v_1(F(s,t)), v_2(F(s,t)) \begin{pmatrix} \partial F_1/\partial s & \partial F_1/\partial t \\ \partial F_2/\partial s & \partial F_2/\partial t \end{pmatrix} .$$

If v is closed, then w is closed and by Green's theorem applied to the parametrizing region $[0,1]^2$,

$$\int_{\partial[0,1]^2} w = \int_{C_0} v + 0 - \int_{C_1} v - 0$$

is equal to

$$\iint_{[0,1]^2} \left(\frac{\partial w_2}{\partial s} - \frac{\partial w_1}{\partial t}\right) ds\, dt = 0 \,.$$

Hence the existence of the path homotopy F from f_0 to f_1 as paths from x_0 to x_1 in X is sufficient to ensure that

$$\int_{C_0} v = \int_{C_1} v$$

for closed vector fields v, more so than the hypothesis that C_0 (going forward) and C_1 (in reverse) form a simple closed curve bounding a region contained in X.

7.1.2 Homotopy

Let Y and Z be topological spaces.

Definition 7.1.2. Two maps $f_0, f_1 \colon Y \to Z$ are *homotopic* if there exists a map

$$F \colon Y \times [0,1] \to Z$$

with $F(y,0) = f_0(y)$ and $F(y,1) = f_1(y)$ for all $y \in Y$. In this case we write $f_0 \simeq f_1$, and call F a *homotopy* from f_0 to f_1.

Remark 7.1.3. We often use $t \in [0,1]$ to indicate the homotopy parameter. For each $t \in [0,1]$ let $f_t \colon Y \to Z$ be given by $f_t(y) = F(y,t)$. Then each f_t is continuous, and the rule $t \mapsto f_t$ defines a path $[0,1] \to \mathscr{C}(Y,Z)$ connecting f_0 to f_1, in the space of maps $Y \to Z$ with the compact-open topology. If Y is locally compact and Hausdorff we can conversely recover the homotopy F from this path in the mapping space.

Lemma 7.1.4. *The homotopy relation $\simeq$ is an equivalence relation on the set of maps $Y \to Z$.*

Proof. For each $f \colon Y \to Z$ there is a constant homotopy $F \colon Y \times [0,1] \to Z$ from f to f given by $F(y,t) = f(y)$.

If $F \colon Y \times [0,1] \to Z$ is a homotopy from $f_0 \colon Y \to Z$ to $f_1 \colon Y \to Z$ then $\bar{F} \colon Y \times [0,1] \to Z$ given by $\bar{F}(y,t) = F(y, 1-t)$ is a homotopy from f_1 to f_0.

If $F \colon Y \times [0,1] \to Z$ is a homotopy from $f_0 \colon Y \to Z$ to $f_1 \colon Y \to Z$ and $G \colon Y \times [0,1] \to Z$ is a homotopy from f_1 to $f_2 \colon Y \to Z$, then there is a homotopy $F \star G \colon Y \times [0,1] \to Z$ from f_0 to f_2 given by

$$(F \star G)(y,t) = \begin{cases} F(y,2t) & \text{for } 0 \le t \le 1/2, \\ G(y, 2t-1) & \text{for } 1/2 \le t \le 1. \end{cases}$$

This defines a continuous map by the pasting lemma (Theorem 2.6.28) from §18, since $Y \times [0,1/2]$ and $Y \times [1/2,1]$ are closed subsets that cover $Y \times [0,1]$, and the functions $F(y,2t)$ and $G(y,2t-1)$ agree on the overlap $Y \times \{1/2\}$. □

Definition 7.1.5. The equivalence class of $f \colon Y \to Z$ under $\simeq$ is denoted

$$[f] = \{f' \colon Y \to Z \mid f \simeq f'\}$$

and is called the *homotopy class* of f. It is common to write

$$[Y,Z] = \mathscr{C}(Y,Z)/\simeq$$

for the set of homotopy classes of maps $Y \to Z$.

Example 7.1.6. Let $f, f' \colon Y \to \mathbb{R}^2$ be any two maps to the plane. Then $f \simeq f'$ via the *straight-line homotopy*

$$F(y,t) = (1-t)f(y) + tf'(y)\,.$$

For each y, the rule $t \mapsto F(y,t)$ parametrizes the straight line segment in $\mathbb{R}^2$ from $F(y,0) = f(y)$ to $F(y,1) = f'(y)$.

More generally, if $Z \subset \mathbb{R}^n$ is any convex subspace of $\mathbb{R}^n$, so that $(1-t)z + tz' \in Z$ for all $z, z' \in Z$ and $t \in [0,1]$, then any two maps $f, f' \colon Y \to Z$ are homotopic.

Example 7.1.7. If $Y = \{y_0\}$ is a single point, then the maps $f \colon Y \to Z$ correspond to the points $z = f(y_0) \in Z$. Two maps $f, f' \colon Y \to Z$ are homotopic if and only if there is a path from $z = f(y_0)$ to $z' = f(y_0)$. Hence the homotopy classes of maps $Y \to Z$ correspond to the path components of Z, and the set of homotopy classes $[Y,Z]$ corresponds to the set of path components of Z, often denoted $\pi_0(Z)$.

7.1.3 Path homotopy

Let X be any topological space.

We are particularly interested in the case when $Y = [0,1]$ and $Z = X$, so that $f_0, f_1 \colon [0,1] \to X$ are paths in X. We often use $s \in [0,1]$ as the path parameter. (In differential geometry, s is often used more specifically to indicate parametrization by path length, but we will not conform to this usage here.)

Definition 7.1.8. Let $x_0, x_1 \in X$. A *path in X from x_0 to x_1* is a map $f \colon [0,1] \to X$ with $f(0) = x_0$ and $f(1) = x_1$.

Given x_0 and x_1, a path in X from x_0 to x_1 is restricted in that the end points are already prescribed. The corresponding notion of path homotopy is similarly restricted.

Definition 7.1.9. Two paths $f_0, f_1 \colon [0,1] \to X$ in X from x_0 to x_1, are *path homotopic* if there is a map

$$F \colon [0,1] \times [0,1] \to X$$

with $F(s,0) = f_0(s)$ and $F(s,1) = f_1(s)$ for all $s \in [0,1]$, and $F(0,t) = x_0$ and $F(1,t) = x_1$ for all $t \in [0,1]$. In this case we write $f \simeq_p g$, and call F a *path homotopy* from f to g.

For a path homotopy $F \colon f_0 \simeq_p f_1$ the corresponding paths $f_t \colon [0,1] \to X$, with $f_t(s) = F(s,t)$, are all paths from x_0 to x_1.

Lemma 7.1.10. *The path homotopy relation $\simeq_p$ is an equivalence relation.*

Proof. The following three claims are easy to check by inspection of the definitions.

The constant homotopy from f to f is a path homotopy.

If $F \colon f_0 \simeq f_1$ is a path homotopy, then $\bar{F} \colon f_1 \simeq f_0$ is a path homotopy.

If $F \colon f_0 \simeq f_1$ and $G \colon f_1 \simeq f_2$ are path homotopies, then $F \star G \colon f_0 \simeq f_2$ is a path homotopy. □

Definition 7.1.11. If $f \colon [0,1] \to X$ is a path in X from x_0 to x_1 we let

$$[f] = \{f' \mid f \simeq_p f'\}$$

denote its path homotopy class. We then write

$$\pi_1(X, x_0, x_1) = \{[f] \mid f \text{ is a path in } X \text{ from } x_0 \text{ to } x_1\}$$

for the set of such path homotopy classes.

Example 7.1.12. If $x_0, x_1 \in X = \mathbb{R}^2$ is the plane, and $f, f' \colon [0,1] \to \mathbb{R}^2$ are paths from x_0 to x_1, then the straight-line homotopy

$$F(s,t) = (1-t)f(s) + tf'(s)$$

is a path homotopy, since $F(0,t) = (1-t)x_0 + tx_0 = x_0$ and $F(1,t) = (1-t)x_1 + tx_1 = x_1$ for all $t \in [0,1]$.

More generally, if $x_0, x_1 \in X \subset \mathbb{R}^n$ is any convex subspace of $\mathbb{R}^n$, so that $(1-t)x + tx' \in X$ for all $x, x' \in X$ and $t \in [0,1]$, then any two paths $f, f' \colon [0,1] \to X$ from x_0 to x_1 are path homotopic.

Example 7.1.13. Let $X = \mathbb{R}^2 - \{(0,0)\}$ be the punctured plane. The paths

$$f(s) = (\cos(\pi s), \sin(\pi s))$$
$$g(s) = (\cos(\pi s), 2\sin(\pi s))$$

from $(1,0)$ to $(-1,0)$ are path homotopic: the straight-line homotopy in $\mathbb{R}^2$ factors through X. However, the straight-line homotopy in $\mathbb{R}^2$ from f to the path

$$h(s) = (\cos(\pi s), -\sin(\pi s))$$

from $(1,0)$ to $(-1,0)$ passes through the origin at $s = t = 1/2$, hence does not give a (path) homotopy from f to h as paths in X. We shall prove later that f and h are not path homotopic in X. (They are homotopic as maps, ignoring the endpoint conditions, which is one reason why it is path homotopy that is the interesting relation.)

7.1.4 Sums of line integrals

If A is a curve in X from x_0 to x_1, parametrized by $f\colon [0,1] \to X$, and B is a curve in X from x_1 to x_2, parametrized by $g\colon [0,1] \to X$, then $C = A \cup B$ is a curve from x_0 to x_2, which can be parametrized by

$$f * g\colon [0,1] \longrightarrow X$$

given by

$$(f * g)(s) = \begin{cases} f(2s) & \text{for } 0 \le s \le 1/2, \\ g(2s-1) & \text{for } 1/2 \le s \le 1. \end{cases}$$

If f and g are continuously differentiable, then $f * g$ will at least be piecewise continuously differentiable, which is good enough to make sense of line integrals over C. In this case

$$\int_C v = \int_A v + \int_B v\,.$$

Hence the deformation-invariant values taking by the integrals $\int_C v$ for closed vector fields v also have an algebraic structure, with addition of integrals corresponding to a composition of curves.

7.1.5 Composition of paths

Consider points $x_0, x_1, x_2, \dots$ in X.

Definition 7.1.14. If $f\colon [0,1] \to X$ is a path from x_0 to x_1, and $g\colon [0,1] \to X$ is a path from x_1 to x_2, the *composition* $f * g\colon [0,1] \to X$ is defined to be the path

$$(f * g)(s) = \begin{cases} f(2s) & \text{for } 0 \le s \le 1/2, \\ g(2s-1) & \text{for } 1/2 \le s \le 1 \end{cases}$$

from x_0 to x_2. Note that $f(1) = x_1 = g(0)$, so $f * g$ is continuous by the pasting lemma.

Lemma 7.1.15. *If $f_0 \simeq_p f_1$ are path homotopic paths from x_0 to x_1, and $g_0 \simeq_p g_1$ are path homotopic paths from x_1 to x_2, then the compositions $f_0 * g_0 \simeq_p f_1 * g_1$ are path homotopic paths from x_0 to x_2. Hence there is a well-defined pairing of path homotopy classes*

$$\pi_1(X, x_0, x_1) \times \pi_1(X, x_1, x_2) \longmapsto \pi_1(X, x_0, x_2)$$
$$([f],[g]) \longmapsto [f] * [g] = [f * g]$$

where f is a path from x_0 to x_1 and g is a path from x_1 to x_2.

Proof. Let $F\colon [0,1]\times[0,1]\to X$ be a path homotopy from f_0 to f_1, and let $G\colon [0,1]\times[0,1]\to X$ be a path homotopy from g_0 to g_1. Then $F*G\colon [0,1]\times[0,1]\to X$ defined by

$$(F*G)(s,t)=\begin{cases}F(2s,t) & \text{for } 0\le s\le 1/2,\\ G(2s-1,t) & \text{for } 1/2\le s\le 1\end{cases}$$

is a path homotopy from f_0*g_0 to f_1*g_1. □

Theorem 7.1.16. *The operation $*$ on path homotopy classes in a space X has the following properties:*

1. *(Associativity) If f is a path from x_0 to x_1, g a path from x_1 to x_2, and h a path from x_2 to x_3, then*
$$([f]*[g])*[h]=[f]*([g]*[h])\,.$$
2. *(Left and right units) For $x\in X$ let $c_x\colon[0,1]\to X$ denote the* constant path *at x, with $c_x(s)=x$ for all $s\in[0,1]$. If f is a path from x_0 to x_1 then*
$$[c_{x_0}]*[f]=[f]=[f]*[c_{x_1}]\,.$$
3. *(Inverse) For $f\colon[0,1]\to X$ a path from x_0 to x_1 let $\bar f$ be the* reverse path *from x_1 to x_0, with $\bar f(s)=f(1-s)$ for all $s\in[0,1]$. Then*
$$[f]*[\bar f]=[c_{x_0}]\qquad\text{and}\qquad[\bar f]*[f]=[c_{x_1}]\,.$$

Proof. (1) The composition $([f]*[g])*[h]$ is the path homotopy class of the path $(f*g)*h$ given by

$$((f*g)*h)(s)=\begin{cases}f(4s) & \text{for } 0\le s\le 1/4,\\ g(4s-1) & \text{for } 1/4\le s\le 1/2,\\ h(2s-1) & \text{for } 1/2\le s\le 1\end{cases}$$

while $[f]*([g]*[h])$ is the path homotopy class of the path $f*(g*h)$ given by

$$(f*(g*h))(s)=\begin{cases}f(2s) & \text{for } 0\le s\le 1/2,\\ g(4s-2) & \text{for } 1/2\le s\le 3/4,\\ h(4s-3) & \text{for } 3/4\le s\le 1.\end{cases}$$

These are not equal, but they are path homotopic. One choice of path homotopy $F\colon (f*g)*h\simeq_p f*(g*h)$ is given by the family of paths from x_0 to x_3 that at time $t\in[0,1]$ traverses f for path parameter values

$$0\le s\le(1-t)(1/4)+t(1/2)=(1+t)/4\,,$$

then traverses g for

$$(1+t)/4\le s\le(1-t)(1/2)+t(3/4)=(2+t)/4\,,$$

and finally traverses h for $(2+t)/4\le s\le 1$. This leads to the following explicit formula:

$$F(s,t)=\begin{cases}f(4s/(1+t)) & \text{for } 0\le s\le(1+t)/4,\\ g(4s-1-t) & \text{for } (1+t)/4\le s\le(2+t)/4,\\ h((4s-2-t)/(2+t)) & \text{for } (2+t)/4\le s\le 1\,.\end{cases}$$

It is continuous as a function of $(s,t) \in [0,1]^2$, by the pasting lemma.

(2) The composition $[c_{x_0}] * [f]$ is the path homotopy class of the path $c_{x_0} * f$ given by

$$(c_{x_0} * f)(s) = \begin{cases} x_0 & \text{for } 0 \le s \le 1/2, \\ f(2s-1) & \text{for } 1/2 \le s \le 1. \end{cases}$$

This is not equal to f, but they are path homotopic. One path homotopy $F\colon c_{x_0} * f \simeq_p f$ is given by the family of paths that at time $t \in [0,1]$ remains at x_0 for parameter values

$$0 \le s \le (1-t)(1/2) + t(0) = (1-t)/2$$

and then traverses f for

$$(1-t)/2 \le s \le 1\,.$$

This corresponds to the formula

$$F(s,t) = \begin{cases} x_0 & \text{for } 0 \le s \le (1-t)/2, \\ f((2s-1+t)/(1+t)) & \text{for } (1-t)/2 \le s \le 1. \end{cases}$$

Similarly, $[f] * [c_{x_1}]$ is the path homotopy class of

$$(f * c_{x_1})(s) = \begin{cases} f(2s) & \text{for } 0 \le s \le 1/2, \\ x_1 & \text{for } 1/2 \le s \le 1. \end{cases}$$

A path homotopy $F\colon f \simeq_p f * c_{x_1}$ is given by the family of paths that at time t traverses f for

$$0 \le s \le (1-t)(1) + t(1/2) = 1 - t/2$$

and then remains at x_1 for

$$1 - t/2 \le s \le 1\,.$$

This corresponds to the formula

$$F(s,t) = \begin{cases} f(2s/(2-t)) & \text{for } 0 \le s \le 1 - t/2, \\ x_1 & \text{for } 1 - t/2 \le s \le 1. \end{cases}$$

(3) The composition $[f] * [\bar{f}]$ is the path homotopy class of

$$f * \bar{f} = \begin{cases} f(2s) & \text{for } 0 \le s \le 1/2, \\ f(2-2s) & \text{for } 1/2 \le s \le 1, \end{cases}$$

since $\bar{f}(2s-1) = f(1-(2s-1)) = f(2-2s)$. A path homotopy $c_{x_0} \simeq_p f * \bar{f}$ is given by the family of paths that at time t follows $s \mapsto f(2s)$ for $0 \le s \le t/2$, remains at $f(t)$ for $t/2 \le s \le 1 - t/2$, and then returns with $s \mapsto f(2-2s)$ for $1 - t/2 \le s \le 1$.

$$F(s,t) = \begin{cases} f(2s) & \text{for } 0 \le s \le t/2, \\ f(t) & \text{for } t/2 \le s \le 1 - t/2, \\ f(2-2s) & \text{for } 1 - t/2 \le s \le 1. \end{cases}$$

Again, this is continuous in (s,t) by the pasting lemma.

A path homotopy $c_{x_1} \simeq_p \bar{f} * f$ is obtained by replacing f with $\bar{f}$ in the path homotopy $f * \bar{f} \simeq_p c_{x_0}$. □

Remark 7.1.17. The sets of path homotopy classes

$$\pi_1(X, x_0, x_1)$$

for all pairs of points $x_0, x_1 \in X$, together with the pairings

$$\pi_1(X, x_0, x_1) \times \pi_1(X, x_1, x_2) \longrightarrow \pi_1(X, x_0, x_2)\,,$$

subject to the properties of the theorem, form a categorical structure called a *groupoid*. For many purposes, little is lost by selecting a single point $x_0 \in X$, and focussing on the part of this structure where $x_0 = x_1 = x_2$. The resulting structure is a *group*, which is a far more classical mathematical notion.

7.2 (§52) The Fundamental Group

7.2.1 The fundamental group

Let $x_0 \in X$ be a fixed *base point*. We refer to the pair (X, x_0) as a *based space*.

Definition 7.2.1. A path in X from x_0 to x_0 is called a *loop in X based at x_0*, or *a loop in (X, x_0)*. Let

$$\pi_1(X, x_0) = \{[f] \mid f \text{ is a loop in } (X, x_0)\}$$

be the set of path homotopy classes of loops in X based at x_0.

$\pi_1(X, x_0)$ is the same set as $\pi_1(X, x_0, x_0)$. Note that the composition $f * g$ of two loops in (X, x_0) is again a loop in (X, x_0).

Lemma 7.2.2. *The composition of path homotopy classes specializes to a pairing*

$$\begin{aligned} \pi_1(X, x_0) \times \pi_1(X, x_0) &\longrightarrow \pi_1(X, x_0) \\ ([f], [g]) &\longmapsto [f] * [g] = [f * g] \end{aligned}$$

where f and g are loops in X based at x_0.

Theorem 7.2.3. *The set $\pi_1(X, x_0)$ with the composition operation $*$ is a group, with neutral element $e = [c_{x_0}]$ and group inverse $[f]^{-1} = [\bar{f}]$ for each loop f in (X, x_0).*

Proof. The composition operation defines a group structure if it is (1) associative, (2) has a left and right unit, and (3) each element has a left and right inverse. All three conditions follow by specializing the previous theorem to the case where all paths are loops in X based at x_0. □

Definition 7.2.4. $\pi_1(X, x_0)$, with this group structure, is called the *fundamental group* of X based at x_0.

Example 7.2.5. If $X \subset \mathbb{R}^n$ is convex, and $x_0 \in X$, then $\pi_1(X, x_0) = \{e\}$ is the trivial group.

Theorem 7.2.6. *If X is path connected and $x_0, x_1 \in X$ then $\pi_1(X, x_0) \cong \pi_1(X, x_1)$.*

We omit the proof. (See Theorem 52.1 in Munkres.)

Definition 7.2.7. A path connected space X is said to be *simply connected* if $\pi_1(X, x_0)$ is the trivial group for some, hence any, base point $x_0 \in X$.

A path connected space X is simply connected if and only if any two paths f and f' in X from x_0 to x_1 are path homotopic.

Remark 7.2.8. For the fundamental group to be useful, we must of course have examples of spaces with non-trivial fundamental group. We shall see that for $X = S^1$ the circle, based at $s_0 = (1, 0)$ (or any other point), the fundamental group $\pi_1(S^1, s_0)$ is an infinite cyclic group, isomorphic to the additive group $\mathbb{Z}$ of the integers. In fact, any discrete group G can be realized as the fundamental group of some based space.

7.2.2 Periods

If v is a closed vector field on $X \subset \mathbb{R}^2$, the rule

$$[f] \longmapsto \int_C v$$

where f is a ((piecewise continuously differentiable)) loop in X based at x_0, and $C = f([0,1])$, defines a group homomorphism

$$\int_? v \colon \pi_1(X, x_0) \longrightarrow \mathbb{R}\,.$$

Up to some issues about whether continuous loops and path homotopies can be replaced with continuously differentiable loops and path homotopies, the example of the closed vector field

$$v(x_1, x_2) = \frac{(-x_2, x_1)}{x_1^2 + x_2^2}$$

on $X = \mathbb{R}^2 - \{(0,0)\}$ with $\int_C v = 2\pi$, where $C = f([0,1])$ is parametrized as a loop based at $(1,0)$, going 'once around' the origin, shows that $\int_? v$ is nonzero. In particular, this shows that $\pi_1(X, x_0)$ is an infinite group when $X = \mathbb{R}^2 - \{(0,0)\}$, except for the abovementioned issues about replacing continuous functions with differentiable ones.

7.2.3 Functoriality

Definition 7.2.9. Let (X, x_0) and (Y, y_0) be based spaces. A *based map* $h \colon (X, x_0) \to (Y, y_0)$ is a map $h \colon X \to Y$ such that $h(x_0) = y_0$. If f is a loop in X based at x_0, then $h \circ f$ is a loop in Y based at y_0. Let

$$h_* \colon \pi_1(X, x_0) \to \pi_1(Y, y_0)$$

be defined by $h_*([f]) = [h \circ f]$.

Lemma 7.2.10. *h_* is a group homomorphism.*

We call h_* the homomorphism *induced* by the based map h.

Proof. This means that for closed loops f and g in (X, x_0), the identity

$$h_*([f] * [g]) = h_*([f]) * h_*([g])$$

holds in $\pi_1(Y, y_0)$. The two sides are the path homotopy classes of the loops $h \circ (f * g)$ and $(h \circ f) * (h \circ g)$, respectively. These are in fact the same loop, given by

$$(h \circ (f * g))(s) = ((h \circ f) * (h \circ g))(s) = \begin{cases} h(f(2s)) & \text{for } 0 \le s \le 1/2, \\ h(g(2s-1)) & \text{for } 1/2 \le s \le 1. \end{cases}$$

□

Theorem 7.2.11 (Functoriality)**.** *If $h \colon (X, x_0) \to (Y, y_0)$ and $k \colon (Y, y_0) \to (Z, z_0)$, then*

$$(k \circ h)_* = k_* \circ h_* \colon \pi_1(X, x_0) \longrightarrow \pi_1(Z, z_0)\,.$$

If $id_X \colon (X, x_0) \to (X, x_0)$ is the identity map, then

$$(id_X)_* \colon \pi_1(X, x_0) \longrightarrow \pi_1(X, x_0)$$

is the identity homomorphism.

Proof. If $[f]$ is a loop in (X, x_0) then

$$(k \circ h)_*([f]) = [(k \circ h) \circ f]$$

and

$$k_*(h_*([f])) = [k \circ (h \circ f)]$$

in $\pi_1(Z, z_0)$, and these are equal by the associativity of composition of maps:

$$[0,1] \xrightarrow{f} X \xrightarrow{h} Y \xrightarrow{k} Z\,.$$

Similarly, $(id_X)_*([f]) = [id_X \circ f]$ is equal to $[f]$. □

Remark 7.2.12. We say that the combined rule sending each based space (X, x_0) to its fundamental group $\pi_1(X, x_0)$ and sending each based map $h\colon (X, x_0) \to (Y, y_0)$ to its induced homomorphism h_* is a *functor* from the category of based spaces and based maps to the category of groups and homomorphisms. For now this is mostly fancy language to express the identity $(k \circ h)_* = k_* \circ h_*$.

Corollary 7.2.13. *If* $h\colon (X, x_0) \to (Y, y_0)$ *is a homeomorphism then*

$$h_*\colon \pi_1(X, x_0) \to \pi_1(Y, y_0)$$

is a group isomorphism.

Proof. If $k = h^{-1}\colon (Y, y_0) \to (X, x_0)$ is the inverse homeomorphism to k, then $k_* = (h^{-1})_*$ is an inverse isomorphism to h_*, because $k_* \circ h_* = (k \circ h)_* = (id_X)_*$ is the identity on $\pi_1(X, x_0)$, and $h_* \circ k_* = (h \circ k)_* = (id_Y)_*$ is the identity on $\pi_1(Y, y_0)$. □

Hence homeomorphic based spaces have isomorphic fundamental groups. In this sense the fundamental group is a topological invariant of based spaces.

Once we show that $\pi_1(S^1, s_0)$ is a nontrivial group, we can deduce the following 2-dimensional analogue of the intermediate value theorem. In fact, we will see that the inclusion $j\colon S^1 \to \mathbb{R}^2 - \{(0,0)\}$ induces an isomorphism

$$j_*\colon \pi_1(S^1, s_0) \xrightarrow{\cong} \pi_1(\mathbb{R}^2 - \{(0,0)\}, s_0)\,,$$

so determining the fundamental group of the circle is equivalent to determining the fundamental group of the punctured plane.

Theorem 7.2.14 (Brouwer's fixed point theorem)**.** *Let* $D^2 = \{x \in \mathbb{R}^2 : \|x\| \le 1\}$ *be the unit disc in the plane. Each map* $f\colon D^2 \to D^2$ *has a fixed point, i.e., a point* $p \in D^2$ *such that* $f(p) = p$*.*

Proof. Suppose that $f\colon D^2 \to D^2$ is a map with no fixed point, so that $f(p) \ne p$ for each $p \in D^2$. Define a map $r\colon D^2 \to S^1$, where $S^1 = \{x \in \mathbb{R}^2 : \|x\| = 1\}$ is the unit circle, as follows: Draw the ray

$$s \mapsto (1-s)f(p) + sp$$

in the plane, starting at $f(p)$ for $s = 0$, passing through p for $s = 1$, and leaving D^2 at a point $r(p)$ when $s \ge 1$ is such that $\|(1-s)f(p) + sp\| = 1$. Here s and $r(p)$ depend continuously on p, so we get a map

$$r\colon D^2 \longrightarrow S^1\,.$$

This would be a retraction of D^2 on S^1, in the sense that $r|S^1$ is the identity. To check this, note that for $p \in S^1$ the ray from $f(p)$ through p reaches S^1 at the point p, for $s = 1$, so $r(p) = p$. This is the assertion that $r|S^1 = id_{S^1}$.

The functoriality of the fundamental group and the non-triviality of the fundamental group of the circle now leads to a contradiction. Let $i\colon S^1 \subset D^2$ denote the inclusion map. The composite map

$$S^1 \xrightarrow{i} D^2 \xrightarrow{r} S^1$$

is the identity on S^1. Hence the composite $r_* \circ i_*$ of the induced homomorphisms

$$\pi_1(S^1, s_0) \xrightarrow{i_*} \pi_1(D^2, s_0) \xrightarrow{r_*} \pi_1(S^1, s_0)$$

is the identity homomorphism. However, $D^2 \subset \mathbb{R}^2$ is convex, so $\pi_1(D^2, s_0) \cong \{e\}$ is the trivial group and the image of $r_* \circ i_*$ consists only of the unit element. Thus $\pi_1(S^1, s_0)$ must be a trivial group (if f has no fixed point). □

7.3 (§53) Covering Spaces

Let $S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2+y^2 = 1\}$ be the unit circle, based at $s_0 = (1, 0)$. In order to calculate the fundamental group

$$\pi_1(S^1, s_0)$$

of path homotopy classes of loops in S^1 at s_0, we will use the map

$$p\colon \mathbb{R} \longrightarrow S^1$$
$$p(x) = (\cos(2\pi x), \sin(2\pi x))\,.$$

This map is particularly well-behaved, exhibiting $\mathbb{R}$ as a 'covering space' of S^1. This will allow us to relate loops in S^1 to paths in $\mathbb{R}$, in a way that is compatible with path homotopy. From the fact that $\mathbb{R}$ is simply connected, we will then obtain a bijection

$$\phi\colon \pi_1(S^1, s_0) \xrightarrow{\cong} p^{-1}(s_0) = \mathbb{Z}$$

between the elements of $\pi_1(S^1, s_0)$ and the points in the preimage $p^{-1}(s_0) \subset \mathbb{R}$. That preimage is the set of integers, and a further analysis shows that the bijection ϕ is a group isomorphism: the pairing in the fundamental group corresponds to the addition of integers.

Definition 7.3.1. A map $p\colon E \to B$ is called a *covering map* if each point in B has a neighborhood $U \subset B$ that is *evenly covered*. This means that

$$p^{-1}(U) = \coprod_{\alpha \in J} V_\alpha$$

is the disjoint union of open subsets $V_\alpha \subset E$ such that the restriction

$$p|_{V_\alpha}\colon V_\alpha \xrightarrow{\cong} U$$

is a homeomorphism, for each $\alpha \in J$.

We call p the *projection*, E the *total space* and B the *base space* of the covering map. For each $b \in B$ the preimage $p^{-1}(b) = \{e \in E \mid p(e) = b\}$ is called the *fiber* over b.

Example 7.3.2. The map $p\colon \mathbb{R} \to S^1$ given by $p(x) = (\cos(2\pi x), \sin(2\pi x))$ is a covering map. The open subsets $U = S^1 - \{(-1,0)\}$ and $U' = S^1 - \{(1,0)\}$ are evenly covered. To verify this for U, note that

$$p^{-1}(U) = \coprod_{\alpha\in\mathbb{Z}} (\alpha - 1/2, \alpha + 1/2)$$

is the disjoint union of open subsets $V_\alpha = (\alpha - 1/2, \alpha + 1/2) \subset \mathbb{R}$, and the restricted map

$$p|_{V_\alpha} \colon V_\alpha \xrightarrow{\cong} U$$

is a homeomorphism for each integer $\alpha \in \mathbb{Z}$. Similarly, for U' note that

$$p^{-1}(U') = \coprod_{\alpha\in\mathbb{Z}} (\alpha, \alpha + 1)$$

is the disjoint union of open subsets $V'_\alpha = (\alpha, \alpha + 1) \subset \mathbb{R}$, and the restricted map

$$p|_{V'_\alpha} \colon V'_\alpha \xrightarrow{\cong} U'$$

is a homeomorphism for each integer $\alpha \in \mathbb{Z}$.

Example 7.3.3. Let $n \in \mathbb{N}$. The n-th power map

$$p\colon S^1 \longrightarrow S^1$$
$$p(\cos(\theta), \sin(\theta)) = (\cos(n\theta), \sin(n\theta))$$

is a covering map. Again the open subsets $U = S^1 - \{(-1,0)\}$ and $U' = S^1 - \{(1,0)\}$ in the base are evenly covered. For example,

$$p^{-1}(U) = \{(\cos(\theta), \sin(\theta) \mid n\theta \neq \pi \mod 2\pi\}$$

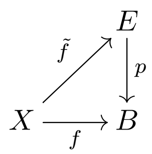
is the disjoint union of n open arcs in the total space.

Remark 7.3.4. Note the difference between a cover, or covering, $\mathscr{C} = \{U_\alpha\}_{\alpha\in J}$ of a space $X = \bigcup_{\alpha\in J} U_\alpha$ and a covering map $p\colon E \to B$. The cover also gives rise to a map $\coprod_{\alpha\in J} U_\alpha \to X$, but the notions are different. They are, however, both special cases of a more general notion. In the context of algebraic geometry, the common notion is that of a Grothendieck (pre-)topology, with the covers corresponding to the Zariski topology and the covering maps corresponding to the étale topology.

7.4 (§54) The Fundamental Group of the Circle

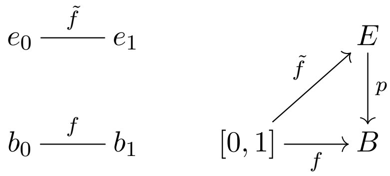
Let $p\colon E \to B$ be a covering spaces. We now relate paths in E to paths in B.

Definition 7.4.1. Let $f\colon X \to B$ be any map. A *lifting* of f is a map $\tilde{f}\colon X \to E$ such that $p \circ \tilde{f} = f$:



Hence, for each $x \in X$ the point $\tilde{f}(x)$ lies in the fiber of E above $f(x)$. If $X = [0,1]$, so that f is a path in B, a lifting of f is a path in E.

Proposition 7.4.2 (Unique path lifting property). *Let $p\colon E \to B$ be a covering map, let $e_0 \in E$ and $b_0 = p(e_0)$, and let f be a path in B from b_0. Then there exists a unique path $\tilde{f}$ in E from e_0 that lifts f.*



Proof. Cover B by evenly covered open subsets U. The preimages $f^{-1}(U)$ form an open cover of $[0,1]$. By compactness, there is a Lebesgue number $\epsilon > 0$ for this open cover. Choose $n \in \mathbb{N}$ such that $1/n < \epsilon$, and let $s_k = k/n$ for each $0 \le k \le n$.

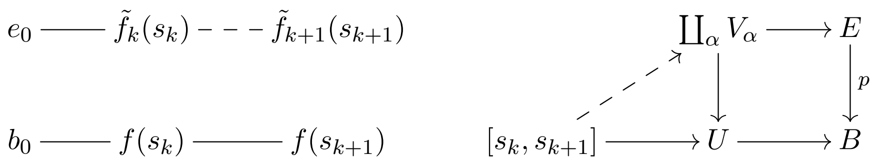
$$0 = s_0 < s_1 < \cdots < s_k < s_{k+1} < \cdots < s_n = 1\,.$$

Then for each $0 \le k < n$ the interval $[s_k, s_{k+1}]$ lies in one of the preimages $f^{-1}(U)$, so

$$f([s_k, s_{k+1}]) \subset U$$

lies in an evenly covered subset of B.

We first show the existence of a lift $\tilde{f}$ of f. Let $f_k = f|_{[0,s_k]}$ denote the restriction of f to the first k of these intervals of length $1/n$. By (finite) induction on k, assume that f_k admits a lift $\tilde{f}_k\colon [0, s_k] \to E$, subject to the initial condition $\tilde{f}_k(0) = e_0$. This is clear for $k = 0$. We will show that $\tilde{f}_k$ can be extended to a lift $\tilde{f}_{k+1}\colon [0, s_{k+1}] \to E$ of f_{k+1}.



As already noted, $f([s_k, s_{k+1}])$ lies in an evenly covered $U \subset B$. Let $p^{-1}(U) = \coprod_{\alpha \in J} V_\alpha$. The end point $\tilde{f}_k(s_k)$ of the lift $\tilde{f}_k$ lies in the fiber over $f_k(s_k) = f(s_k) \in U$, hence lies in V_β for a unique $\beta \in J$. Since

$$p|_{V_\beta}\colon V_\beta \xrightarrow{\cong} U$$

is a homeomorphism, we can use its inverse to define the extended lift:

$$\tilde{f}_{k+1}(x) = \begin{cases} \tilde{f}_k(x) & \text{for } 0 \le x \le s_k, \\ p|_{V_\beta}^{-1} \circ f_{k+1}(x) & \text{for } s_k \le x \le s_{k+1}. \end{cases}$$

This is well-defined at $x = s_k$ by the choice of β, hence is continuous by the pasting lemma. Clearly, $p \circ \tilde{f}_{k+1} = f_{k+1}$, so $\tilde{f}_{k+1}$ lifts f_{k+1} over $[0, s_{k+1}]$. After finitely many steps we obtain a lift $\tilde{f} = \tilde{f}_n$ of f, starting at e_0.

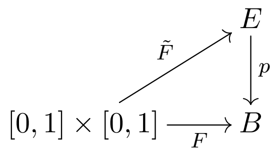
Next we show that $\tilde{f}\colon [0,1] \to E$ is uniquely determined by $p \circ \tilde{f} = f$ and $\tilde{f}(0) = e_0$. Suppose that $\hat{f}\colon [0,1] \to E$ is a second such lift. We show that $\tilde{f}|_{[0,s_k]} = \hat{f}|_{[0,s_k]}$ for $0 \le k \le n$ by induction on k. The case $k = 0$ asserts that $\tilde{f}(0) = \hat{f}(0)$, which is clear since both lifts start at e_0. For $0 \le k < n$ we can assume that $\tilde{f}(s_k) = \hat{f}(s_k)$, and want to prove that $\tilde{f}|_{[s_k,s_{k+1}]} = \hat{f}|_{[s_k,s_{k+1}]}$. Since f maps $[s_k, s_{k+1}]$ into an evenly covered $U \subset B$, with $p^{-1}(U) = \coprod_{\alpha \in J} V_\alpha$ a disjoint union of open subsets of E, both $\tilde{f}$ and $\hat{f}$ map $[s_k, s_{k+1}]$ into $\coprod_{\alpha \in J} V_\alpha$. Let $\beta \in J$ be the index such that $\tilde{f}(s_k) = \hat{f}(s_k)$ lies in V_β. Since $[s_k, s_{k+1}]$ is connected and the V_α are open, $\tilde{f}|_{[s_k,s_{k+1}]}$ must be contained in only one of the sets V_α, and since $\tilde{f}(s_k) \in V_\beta$, that set must in fact be V_β.

Likewise, $\hat{f}|_{[s_k,s_{k+1}]}$ must be contained in V_β. Now $p|_{V_\beta}: V_\beta \to U$ is a homeomorphism, and in particular it is injective. Hence

$$p|_{V_\beta} \circ \tilde{f}|_{[s_k,s_{k+1}]} = f|_{[s_k,s_{k+1}]} = p|_{V_\beta} \circ \hat{f}|_{[s_k,s_{k+1}]}$$

implies that $\tilde{f}|_{[s_k,s_{k+1}]} = \hat{f}|_{[s_k,s_{k+1}]}$, as required. □

Proposition 7.4.3 (Homotopy lifting property)**.** *Let $p\colon E \to B$ be a covering map, let $e_0 \in E$ and $b_0 = p(e_0)$, and let $F\colon [0,1] \times [0,1] \to B$ be a map with $F(0,0) = b_0$. Then there exists a unique lifting $\tilde{F}\colon [0,1] \times [0,1] \to E$ with $\tilde{F}(0,0) = e_0$.*



If F is a path homotopy from f_0 to f_1, then $\tilde{F}$ is a path homotopy from a lift $\tilde{f}_0$ to a lift $\tilde{f}_1$.

Proof. The proof is similar to that of the path lifting property, breaking $[0,1]\times[0,1]$ into smaller squares $[s_k, s_{k+1}] \times [t_k, t_{k+1}]$ that are mapped by F into evenly covered subsets of B, and using $p|_{V_\beta}^{-1}$ to inductively extend a lift of F from $(0,0)$ in a unique way to all of $[0,1] \times [0,1]$

If F is a path homotopy from f_0 to f_1, where these are paths from b_0 to b_1, then $t \mapsto F(0,t)$ is the constant path at b_0. Hence $t \mapsto \tilde{F}(0,t)$ is a lift of this path, starting at $\tilde{F}(0,0) = e_0$. The constant path at e_0 is also such a lift, so by the uniqueness of path lifting we must have $\tilde{F}(0,t) = e_0$ for all $t \in [0,1]$. Similarly, $t \mapsto F(1,t)$ is the constant path at b_1. Let $e_1 = \tilde{F}(1,0)$, with $p(e_1) = b_1$. The lift $t \mapsto \tilde{F}(1,t)$ must then be the constant path at e_1, by the same uniqueness argument as before. This shows that $\tilde{F}$ is a path homotopy, connecting two paths in E from e_0 to e_1. □

We now come to the construction of the *lifting correspondence*

$$\phi\colon \pi_1(B, b_0) \longrightarrow p^{-1}(b_0)$$

that we will prove is a bijection in certain cases.

Definition 7.4.4. Let $p\colon E \to B$ be a covering map, choose $e_0 \in E$ and let $b_0 = p(e_0)$. Given an element $[f] \in \pi_1(B, b_0)$, equal to the path homotopy class of a loop $f\colon [0,1] \to B$ in B at b_0, let $\tilde{f}\colon [0,1] \to E$ be the unique lift of f to a path in E starting at e_0. Let

$$\phi([f]) = \tilde{f}(1)$$

be the end-point of that lifted path. Then $p(\phi([f])) = f(1) = b_0$, so $\phi([f]) \in p^{-1}(b_0)$ lies in the fiber of p over b_0.

If $f_0 \simeq_p f_1$ are path homotopic loops at b_0, and $\tilde{f}_0$ and $\tilde{f}_1$ are lifts in E starting at e_0, then the homotopy lifting property tells us that $\tilde{f}_0 \simeq_p \tilde{f}_1$. In particular, $\tilde{f}_0(1) = \tilde{f}_1(1)$. This means that ϕ is well-defined.

Theorem 7.4.5. *Let $p\colon E \to B$ be a covering map, choose $e_0 \in E$ and let $b_0 = p(e_0)$. If E is path connected then the lifting correspondence*

$$\phi\colon \pi_1(B, b_0) \longrightarrow p^{-1}(b_0)$$

is surjective. If E is simply-connected, then ϕ is a bijection.

Proof. To prove that ϕ is surjective, when E is path connected, consider any point $e \in p^{-1}(b_0)$. By hypothesis there exists a path $h\colon [0,1] \to E$ in E from e_0 to e. Let $f = p \circ h\colon [0,1] \to B$ be its projection to B. Then $f(0) = p(e_0) = b_0$ and $f(1) = p(e) = b_0$, so f is a loop in B at b_0, and $h = \tilde{f}$ is the unique lift of f to a path in E starting at e_0. Hence ϕ maps the path homotopy class $[f]$ in $\pi_1(B, b_0)$ to $\phi([f]) = \tilde{f}(1) = h(1) = e$. Thus ϕ is surjective.

To prove that ϕ is injective, consider two loops $f, g\colon [0,1] \to B$ in B at b_0 such that ϕ maps both $[f]$ and $[g]$ to the same point $e \in E$. Let $\tilde{f}$ and $\tilde{g}$ be the unique lifts of f and g to paths in E starting at e_0, respectively. By assumption, $\tilde{f}(1) = e = \tilde{g}(1)$, so we can form the loop $\tilde{f} * \overline{\tilde{g}}$ in E at e_0. This is the composition of $\tilde{f}$ with the reverse of $\tilde{g}$. By the hypothesis of simple-connectivity, there is a path homotopy

$$H\colon \tilde{f} * \overline{\tilde{g}} \simeq_p c_{e_0}$$

from that loop to the constant loop at e_0. Composing with p we obtain a path homotopy

$$pH\colon f * \overline{g} \simeq_p c_{b_0}$$

from $f * \overline{g}$ to the constant loop at b_0. This means that $[f] * [g]^{-1} = e$ in $\pi_1(B, b_0)$, so $[f] = [g]$. Hence ϕ is injective. □

Theorem 7.4.6. *There is a group isomorphism* $\pi_1(S^1, s_0) \cong \mathbb{Z}$.

Proof. The covering map $p\colon \mathbb{R} \to S^1$ given by $p(x) = (\cos(2\pi x), \sin(2\pi x))$ maps $e_0 = 0$ to $b_0 = s_0 = (1, 0)$, and $E = \mathbb{R}$ is path connected and simply-connected, so the lifting correspondence

$$\phi\colon \pi_1(S^1, s_0) \xrightarrow{\cong} p^{-1}(s_0)$$

is a bijection. Here $p^{-1}(s_0) = \mathbb{Z} \subset \mathbb{R}$. It remains to prove that this bijection respects the group structures, sending the composition $*$ in $\pi_1(S^1, s_0)$ to the addition in $\mathbb{Z}$.

Let $f, g\colon [0,1] \to S^1$ be loops in S^1 at s_0, with unique lifts $\tilde{f}, \tilde{g}\colon [0,1] \to \mathbb{R}$ to paths in $\mathbb{R}$ starting at 0. Let $\tilde{f}(1) = \phi([f]) = m$ and $\tilde{g}(1) = \phi([g]) = n$. The composition $[f] * [g]$ is the path homotopy class of $f * g$. We must show that $\phi([f * g]) = m + n$. To evaluate $\phi([f * g])$ we need to lift $f * g$. The first half of such a lift is given by $\tilde{f}$, traversed at double speed, and ending at $\tilde{f}(1) = m$. To continue the lift we may therefore not use $\tilde{g}$, which starts at 0. Instead we use $\hat{g} = m + g$, given by

$$\hat{g}(s) = m + g(s)$$

for $s \in [0,1]$. Then $\hat{g}$ is a lift of g starting at m, so the composition $\tilde{f} * \hat{g}$ is well-defined. This is a lift of $p(\tilde{f} * \hat{g}) = f * g$ starting at 0, i.e.,

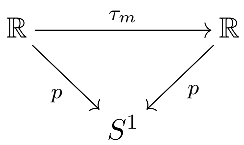
$$\widetilde{f * g} = \tilde{f} * \hat{g}\,.$$

Thus

$$\phi([f * g]) = \widetilde{f * g}(1)) = \hat{g}(1) = m + \tilde{g}(1) = m + n\,.$$

This proves that $\phi([f] * [g]) = m + n = \phi([m]) + \phi([n])$, so that ϕ is a group homomorphism. Since it is bijective, it is a group isomorphism. □

Remark 7.4.7. This proof used the existence of maps $\tau_m\colon x \mapsto m + x\colon \mathbb{R} \to \mathbb{R}$ with $p \circ \tau_m = p$, for each $m \in \mathbb{Z}$. These maps τ_m are called the *deck transformations* of the covering map $p\colon \mathbb{R} \to S^1$.



7.5 (§55) Retractions and Fixed Points

Definition 7.5.1. Let $i\colon A \subset X$. A *retraction* of X onto A is a map $r\colon X \to A$ such that $r \circ i = id_A$.

$$\begin{array}{ccc} A & \overset{i}{\rightarrowtail} & X \\ & {\scriptstyle id_A} \searrow & \downarrow {\scriptstyle r} \\ & & A \end{array}$$

If such a map r exists, we say that A is a *retract* of X.

Lemma 7.5.2. *If $x_0 \in A \subset X$ and A is a retract of X, then $i_*\colon \pi_1(A, x_0) \to \pi_1(X, x_0)$ is injective.*

Proof.

$$\begin{array}{ccc} \pi_1(A, x_0) & \overset{i_*}{\longrightarrow} & \pi_1(X, x_0) \\ & {\scriptstyle id} \searrow & \downarrow {\scriptstyle r_*} \\ & & \pi_1(A, x_0) \end{array}$$

□

Theorem 7.5.3. *There is no retraction of D^2 onto S^1.*

Proof. $\pi_1(S^1, s_0) \to \pi_1(D^2, s_0)$ is not injective. □

Theorem 7.5.4 (Brouwer's fixed point theorem for the disc)**.** *If $f\colon D^2 \to D^2$ is continuous, then there is a $p \in D^2$ with $f(p) = p$.*

We already proved this: if $f(x) \neq x$ for all $x \in D^2$ then there is a retraction $r\colon D^2 \to S^1$, specified by asking that $f(x)$, x and $r(x)$ lie on a line, in that order.

Corollary 7.5.5. *If $A \in M_3(\mathbb{R})$ is a 3×3 matrix with positive entries, then A has a positive eigenvalue $\lambda > 0$.*

Proof. Let $D = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1; x, y, z \geq 0\}$. There is a homeomorphism $h\colon D \cong D^2$. Define $f\colon D \to D$ by $f(v) = Av/\|Av\|$. Then $hfh^{-1}\colon D^2 \to D^2$ is continuous, hence has a fixed point p. Let $v = h^{-1}(p)$. Then $f(v) = v$, so $Av = \lambda v$ with $\lambda = \|Av\| > 0$. □

7.6 (§56) The Fundamental Theorem of Algebra

Theorem 7.6.1. *Let $n \geq 1$ and $a_{n-1}, \dots, a_0 \in \mathbb{C}$. The polynomial*

$$P(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$$

has at least one root $z \in \mathbb{C}$.

Proof. Choose $R \geq 1$ so that

$$|a_{n-1}| + \cdots + |a_1| + |a_0| < R\,.$$

For $w \in \mathbb{C}$ with $|w| = R$ we have

$$\begin{aligned} |a_{n-1}w^{n-1} + \cdots + a_1 w + a_0| &\leq |a_{n-1}|R^{n-1} + \cdots + |a_1|R + a_0 \\ &\leq (|a_{n-1}| + \cdots + |a_1| + |a_0|)R^{n-1} \\ &< R^n = |w|^n \end{aligned}$$

so the straight-line homotopy

$$
\begin{aligned}
F(w,t) &= (1-t)w^n + t(w^n + a_{n-1}w^{n-1} + \cdots + a_1 w + a_0) \\
&= w^n + t(a_{n-1}w^{n-1} + \cdots + a_1 w + a_0)
\end{aligned}
$$

is never zero for $|w| = R$ and $t \in [0,1]$.

Assume, in order to derive a contradiction, that $P(z) \neq 0$ for all $z \in \mathbb{C}$. Consider $S^1 \subset D^2$ as subspaces of $\mathbb{C}$, based at $s_0 = 1$. Let

$$X = S^1 \times [0,1] \cup D^2 \times \{1\}$$

be based at $x_0 = (1,0)$, and define a map $G\colon X \to S^1$ by

$$G(z,t) = F(Rz,t)/|F(Rz,t)|\,.$$

This is well-defined, because for $z \in D^2$ we have $F(Rz,1) = P(Rz) \neq 0$, by our assumption, and for $z \in S^1$ we have $|Rz| = R$, so $F(Rz,t) \neq 0$ for $t \in [0,1]$. In particular,

$$G(z,0) = F(Rz,0)/|F(Rz,0)| = (Rz)^n/|(Rz)^n| = z^n$$

for $z \in S^1$. Hence we have a commutative diagram of based spaces and based maps

$$
\begin{array}{ccc}
(S^1, s_0) & \xrightarrow{\;i\;} & (X, x_0) \\
& \searrow^{g} & \downarrow{\scriptstyle G} \\
& & (S^1, s_0)
\end{array}
$$

where $i(z) = (z,0)$ and $g(z) = z^n$. Note that there is a homeomorphism $h\colon (X,x_0) \cong (D^2, s_0)$, given by $h(z,t) = (1-t/2)z$ for $(z,t) \in S^1 \times [0,1]$ and $h(z,1) = z/2$ for $(z,1) \in D^2 \times \{1\}$.

Applying the fundamental group, we obtain a commutative diagram of groups and homomorphisms

$$
\begin{array}{ccc}
\pi_1(S^1, s_0) & \xrightarrow{\;i_*\;} & \pi_1(X, x_0) \\
& \searrow^{g_*} & \downarrow{\scriptstyle G_*} \\
& & \pi_1(S^1, s_0)\,.
\end{array}
$$

On one hand, $g_*\colon \pi_1(S^1,s_0) \to \pi_1(S^1,s_0)$ maps the generator of $\pi_1(S^1,s_0) \cong \mathbb{Z}$, given by the path homotopy class $[f]$ of the loop $f(s) = e^{2\pi i \cdot s} \leftrightarrow (\cos(2\pi s), \sin(2\pi s))$ to the path homotopy class of the loop

$$gf(s) = (e^{2\pi i \cdot s})^n = e^{2\pi i \cdot ns} \leftrightarrow (\cos(2\pi ns), \sin(2\pi ns))\,.$$

Under the lifting correspondence $\phi\colon \pi_1(S^1,s_0) \cong \mathbb{Z}$, we have $\phi([f]) = 1$, since $\tilde{f}(s) = s$ lifts f and $\tilde{f}(1) = 1$. Similarly, $\phi([gf]) = n$, since $\widetilde{gf}(s) = ns$ lifts gf and $\widetilde{gf}(1) = n$. Hence $g_*\colon \pi_1(S^1,s_0) \to \pi_1(S^1,s_0)$ maps the generator $[f]$ of $\pi_1(S^1,s_0)$ to

$$g_*([f]) = [f] * \cdots * [f]$$

(with n copies of $[f]$), corresponding to multiplication by n in $\mathbb{Z}$. Since $n \geq 1$, this is not the trivial homomorphism.

On the other hand, $h_*\colon \pi_1(X,x_0) \cong \pi_1(D^2,s_0) = \{e\}$ is the trivial group, so the composite $G_* \circ i_*$ is the trivial homomorphism. This contradicts the identity $g_* = G_* \circ i_*$. $\square$

7.7 3-manifolds

Any compact 3-manifold is the disjoint union of a finite set of compact, connected 3-manifolds. A compact, connected 3-manifolds is said to be *irreducible* if each embedded sphere (S^2) bounds an embedded ball (B^3). Otherwise the 3-manifold is reducible and can be simplified by cutting it open along the embedded 2-sphere and gluing in a ball on each side. This process stops after finitely many steps.

Here is a post on irreducible 3-manifolds and their fundamental groups, by Bruno Martelli on mathoverflow.net:

> Perelman has proved Thurston's geometrization conjecture, which says that every irreducible 3-manifold decomposes along its canonical decomposition along tori into pieces, each admitting a geometric structure. A "geometric structure" is a nice riemannian metric, which is in particular complete and of finite volume.
>
> There are eight geometric structures for 3 manifolds: three structures are the constant curvature ones (spherical, flat, hyperbolic), while the other 5 structures are some kind of mixing of low-dimensional structures (for instance, a surface Σ of genus 2 has a hyperbolic metric, and the three-manifold $\Sigma \times S^1$ has a mixed hyperbolic $\times S^1$ structure).
>
> The funny thing is that geometrization conjecture was already proved by Thurston when the canonical decomposition is non-trivial, i.e., when there is at least one torus in it. In that case the manifold is a Haken manifold because it contains a surface whose fundamental group injects in the 3-manifold. Haken manifolds have been studied by Haken himself (of course) and by Waldhausen, who proved in 1968 that two Haken manifolds with isomorphic fundamental groups are in fact homeomorphic.
>
> If the canonical decomposition of our irreducible manifold M is empty, now we can state by Perelman's work that M admits one of these 8 nice geometries. The manifolds belonging to 7 of these geometries are well-known and have been classified some decades ago (six of these geometries actually coincide with the well-known Seifert manifolds, classified by Seifert already in 1933). From the classification one can see that the only distinct manifolds with isomorphic fundamental groups are lens spaces (which belong to the elliptic geometry, since they have finite fundamental group).
>
> The only un-classified geometry is the hyperbolic one. However, Mostow rigidity theorem says that two hyperbolic manifolds with isomorphic fundamental group are isometric, hence we are done. Some simple considerations also show that two manifolds belonging to distinct geometries have non-isomorphic fundamental groups.
>
> Therefore now we know that the fundamental group is a complete invariant for irreducible 3-manifolds, except lens spaces.

— — — END OF NOTES — — —