Principals of Analysis

Paul L. Bailey

Department of Mathematics, Southern Arkansas University
E-mail address: plbailey@saumag.edu

Contents

CHAPTER I

Metric Spaces

ABSTRACT. The chapter introduces the initial definitions regarding metric spaces, subspaces, open and closed sets, and product spaces.

1. Metric Spaces

DEFINITION 1. Let X be a set. A *metric* on X is a function

$$\rho : X \times X \to \mathbb{R}$$

satisfying

(M1) $\rho(x, y) \geq 0$ and $\rho(x, y) = 0$ if and only if $x = y$ (Positivity);
(M2) $\rho(x, y) = \rho(y, x)$ (Symmetry);
(M3) $\rho(x, y) + \rho(y, z) \geq \rho(x, z)$ (Triangle Inequality).

The pair (X, ρ) is called a *metric space*.

EXAMPLE 1. Let X be any set and define $\rho : X \times X \to \mathbb{R}$ by

$$\rho(x, y) = \begin{cases} 0 & \text{if } x = y; \\ 1 & \text{otherwise .} \end{cases}$$

Then ρ is a metric on X, called the *discrete metric*, and (X, ρ) is called a *discrete metric space*.

PROOF. In order to prove that (X, ρ) is a metric space, we need to demonstrate **(M1)**, **(M2)**, and **(M3)**. We take the opportunity of this simple example to discuss some general assumptions we can make in the course of such a proof.

(M1) We have two things to verify; that the image of ρ consists only of nonnegative reals, and that $\rho(x, x) = 0$ for every $x \in X$. The first is immediate upon inspection of the definition, and in general won't need to be mentioned unless there is some doubt. The second is directly provided by the definition.

(M2) Let $x, y \in X$. If $x = y$, then $\rho(x, y) = 0 = \rho(y, x)$. Note that if **(M1)** if already verified, then this case need not be considered.

Thus suppose that x and y are distinct. Then $\rho(x, y) = 1 = \rho(y, x)$.

(M3) Let $x, y, z \in X$. If $x = z$, and **(M1)** is already verified, then this condition says that $0 \leq \rho(x, y) + \rho(y, z)$, which is true. If $x = y$ or $y = z$, this statement becomes an immediate equality. So again, we can assume that x, y, and z are distinct. Then

$$\rho(x, z) = 1 < 2 = \rho(x, y) + \rho(y, z).$$

□

EXAMPLE 2. Let $x = \mathbb{R}$ and define $\rho(x, y) = |x - y|$. Then (X, ρ) is a metric space.

PROOF. We address **(M1)**, **(M2)**, and **(M3)**.

(M1) The absolute value is always nonnegative, and the absolute value of zero is zero, so $\rho(x, x) = |x - x| = |0| = 0$. On the other hand, if $x \neq y$, then $x - y \neq 0$, so $\rho(x, y) = |x - y| \neq 0$.

(M2) Let $x, y \in \mathbb{R}$; without loss of generality, assume that $x > y$. Then $\rho(x, y) = |x - y| = x - y = -(y - x) = |y - x| = \rho(y, x)$.

(M3) Let $x, y, z \in \mathbb{R}$; we have seen that $|a + c| \leq |a + b| + |b + c|$ for every $a, b, c \in \mathbb{R}$. Set $a = x - y$, $b = 0$, and $c = y - z$. Then $a + c = x - y$, $a + b = x - y$, and $b + c = y - z$. Thus

$$\rho(x, z) = |x - z| \leq |x - y| + |y - z| = \rho(x, y) + \rho(y, z).$$

□

EXAMPLE 3. Let $X = \mathbb{R}^k$ and define

$$\rho(x, y) = \sqrt{\sum_{i=1}^{k} (x_i - y_i)^2},$$

where $x = (x_1, \ldots, x_k)$ and $y = (y_1, \ldots, y_k)$.

REMARK. The positivity and symmetry of ρ are clear, but the proof of the triangle inequality is involved, and appears in the last section of this chapter, where it is generalized to the product of a finite number of metric spaces. □

EXAMPLE 4. Let $\mathbb{R}^\infty$ denote the set of all sequences of real numbers that are eventually zero, that is, sequences (x_n) such that $x_n = 0$ for all but finitely many n. Let $X = \mathbb{R}^\infty$ and for $x, y \in X$, define

$$\rho(x, y) = \sqrt{\sum_{i=1}^{\infty} (x_i - y_i)^2},$$

where $x = (x_n)$ and $y = (y_n)$. This make sense, since there are only finitely many nonzero summands. Then (X, ρ) is a metric space.

EXAMPLE 5. Let ℓ^2 denote the set of all sequences of real numbers (x_n) that satisfy the converge criterion

$$\sum_{i=1}^{\infty} x_i^2 < 0.$$

Let $X = \ell^2$ and for $x, y \in X$, define

$$\rho(x, y) = \sqrt{\sum_{i=1}^{\infty} (x_i - y_i)^2},$$

where $x = (x_n)$ and $y = (y_n)$. That this series converges follows from the inequality

$$(a \pm b)^2 \leq 2(a^2 + b^2),$$

which the reader is welcome to verify. Then (X, ρ) is a metric space.

EXAMPLE 6. Let $X = \mathbb{Q}$ and let p be a positive prime integer. For each $x \in \mathbb{Q}$, there exists unique $m, n, \alpha \in \mathbb{Z}$ such that $x = p^\alpha \frac{m}{n}$, where $\gcd(m, n) = 1$ and p does not divide m or n. The *p-adic norm* of x is $|x|_p = \frac{1}{p^\alpha}$. Set $\rho(x, y) = |x-y|_p$. Then (X, ρ) is a metric, known as the *p-adic metric* on $\mathbb{Q}$. Here one can show that not only does ρ satisfy the triangle inequality, but also the stronger inequality $|x - y|_p \leq \max\{|x|_p, |y|_p\}$.

EXERCISE 1. Let $\mathcal{F}_{[a,b]}$ denote the set of all bounded functions $f : [a, b] \to \mathbb{R}$. Let $X = \mathcal{F}_{[a,b]}$ and for $f, g \in X$ define

$$\rho(f, g) = \max\{|f(x) - g(x)| \mid x \in [a, b]\}.$$

Show that (X, ρ) is a metric space.

EXERCISE 2. Let $\mathcal{C}_{[a,b]}$ denote the set of all continuous functions $f : [a, b] \to \mathbb{R}$. Let $X = \mathcal{C}_{[a,b]}$ and for $f, g \in X$ define

$$\rho(f, g) = \int_a^b |f - g| \, dx.$$

Show that (X, ρ) is a metric space.

2. Metric Subspaces

DEFINITION 2. Let (X, ρ) be a metric space and let $A \subset X$. ' Let $\rho_A : A \times A \to \mathbb{R}$ be the restriction of ρ to $A \times A \subset X \times X$. Then ρ_A is a metric on A, and (A, ρ_A) is called a *subspace* of (X, ρ).

EXAMPLE 7. Let $X = \mathbb{R}^2$, and define

$$\rho : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R} \quad \text{by} \quad \rho(p_1, p_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2},$$

where $p_i = (x_i, y_i)$. Then ρ is the standard metric on $\mathbb{R}^2$.

Define

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

We call $\mathbb{S}^1$ the *unit circle*. It inherits the metric $\rho_{\mathbb{S}^1}$ from $(\mathbb{R}^2, \rho)$.

Define

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

We call $\mathbb{D}^2$ the *(closed) unit disk*, and $(\mathbb{D}^2, \rho_{\mathbb{D}^2})$ is a metric space.

EXAMPLE 8. Let $\mathbb{S}^1$ be the unit circle, and let ρ be as in Example 7. We may define a metric

$$\alpha : \mathbb{S}^1 \times \mathbb{S}^1 \to \mathbb{R} \quad \text{by} \quad \alpha(p_1, p_2) = 2 \arcsin(\rho(p_1, p_2)).$$

where $p_1, p_2 \in \mathbb{S}^1$. Then $\alpha(p_1, p_2)$ is the angle, measured in radians, from p_1 to the origin and then to p_2; this is the arclength of the shortest path between these two points.

This produces a different metric on $\mathbb{S}^1$. In due course, we will investigate the relationship between these metrics and related consequences for the structure of the metric space.

EXAMPLE 9. Let $X = \mathbb{R}^3$, and define

$$\rho : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R} \quad \text{by} \quad \rho(p_1, p_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2},$$

where $p_i = (x_i, y_i, z_i)$. Then ρ is the standard metric on $\mathbb{R}^3$.

Define

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

We call $\mathbb{S}^2$ the *unit sphere*, and $(\mathbb{S}^2, \rho_{\mathbb{S}^2})$ is a metric space.

Define

$$\mathbb{D}^3 = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 \leq 1\}.$$

We call $\mathbb{D}^2$ the *(closed) unit ball*, and $(\mathbb{D}^3, \rho_{\mathbb{D}^3})$ is a metric space.

3. Product Metric Spaces

The definition of distance in $\mathbb{R}^k$ has been computed using $k-1$ applications of the Pythagorean Theorem; it is clearly the definition that we want. However, in order to apply results to $\mathbb{R}^k$ that we have proven from the metric axioms, we need to first prove that $\mathbb{R}^k$ is indeed a metric space. This involves a demonstration of the triangle inequality; that is, given

$$x = (x_1, \ldots, x_k), y = (y_1, \ldots, y_k), z = (z_1, \ldots, z_k),$$

we need to show that

$$\sqrt{\sum_{j=1}^{k}(x_j - z_j)^2} \leq \sqrt{\sum_{j=1}^{k}(x_j - y_j)^2} + \sqrt{\sum_{j=1}^{k}(y_j - z_j)^2}.$$

Proving this directly would make use of the triangle inequality

$$|a - c| \leq |a - b| + |b - c|$$

in $\mathbb{R}$, and an application of the Cauchy-Schwartz Inequality (below). With approximately the same effort, we can generalize this result to construct the product of a finite number of arbitrary metric spaces. The definition of distance in the product space is motivated by our previous use of the Pythagorean Theorem.

Theorem 1. *Let $(X_1, \rho_1), \ldots, (X_n, \rho_n)$ be a finite collection of metric spaces. Let $X = \times_{k=1}^{n} X_k$, and define $\rho : X \times X \to \mathbb{R}$ by*

$$\rho(x, y) = \sqrt{\sum_{k=1}^{n} \rho_k(x_k, y_k)^2},$$

where $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$, and $x_k, y_k \in X_k$ for $k = 1, \ldots, n$. Then (X, ρ) is a metric space.

We call ρ the *product metric* on X. The difficulty of the proof of this proposition lies in the triangle inequality, a computation which we will break into several intermediate results.

Lemma 1. *Let $a_k, b_k \in \mathbb{R}$ for $k = 1, \ldots, n$. Then*

$$\sum_i \sum_j (a_i b_j - a_j b_i)^2 = 2 \sum_{i \neq j} (a_i^2 b_j^2 - a_i a_j b_i b_j).$$

Proof. Note that

$$(a_i b_j - a_j b_i)^2 = a_i^2 b_j^2 + a_j^2 b_i^2 - 2 a_i a_j b_i b_j.$$

Then

$$\begin{aligned} \sum_i \sum_j (a_i b_j - a_j b_i)^2 &= \sum_i \sum_j (a_i^2 b_j^2 + a_j^2 b_i^2 - 2 a_i a_j b_i b_j) \\ &= 2 \sum_{i \neq j} (a_i^2 b_j^2 - a_i a_j b_i b_j). \end{aligned}$$

$\square$

LEMMA 2. *Let $a_k, b_k \in \mathbb{R}$ for $k = 1, \dots, n$. Then*

$$\Big(\sum_{k=1}^n a_k b_k\Big)^2 = \sum_{k=1}^n a_k^2 \sum_{k=1}^n b_k^2 - \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n (a_i b_j - a_j b_i)^2.$$

PROOF. Compute that

$$\begin{aligned}\sum_k a_k^2 \sum_k b_k^2 &= \sum_k a_k^2 b_k^2 + 2\sum_{i\neq j} a_i^2 b_j^2 \\ &= \Big(\sum_k a_k^2 b_k^2 + 2\sum_{i\neq j} a_i a_j b_i b_j\Big) - 2\sum_{i\neq j} a_i a_j b_i b_j + 2\sum_{i\neq j} a_i^2 b_j^2 \\ &= \Big(\sum_k a_k b_k\Big)^2 + 2\Big(\sum_{i\neq j} a_i^2 b_j^2 - \sum_{i\neq j} a_i a_j b_i b_j\Big).\end{aligned}$$

Subtracting the equation of Lemma 1 to both sides implies the result. $\square$

LEMMA 3 (Cauchy-Schwartz Inequality). *Let $a_k, b_k \in \mathbb{R}$ for $k = 1, \dots, n$. Then*

$$\Big(\sum_{k=1}^n a_k b_k\Big)^2 \leq \sum_{k=1}^n a_k^2 \sum_{k=1}^n b_k^2.$$

PROOF. This follows from Lemma 2 by noting that $\sum_{i=1}^n \sum_{j=1}^n (a_i b_j - a_j b_i)^2$ is always nonnegative. $\square$

LEMMA 4. *Let $a_k, b_k, c_k \in \mathbb{R}$ be positive for $k = 1, \dots, n$. Then*

$$\sqrt{\sum_{k=1}^n a_k^2} \leq \sqrt{\sum_{k=1}^n b_k^2} + \sqrt{\sum_{k=1}^n c_k^2}.$$

PROOF. For $k = 1, \dots, n$, we have $a_k \leq b_k + c_k$, so $a_k^2 \leq b_k^2 + c_k^2 + 2b_k c_k$. Thus

$$(*) \qquad \sum_{k=1}^n a_k^2 \leq \sum_{k=1}^n b_k^2 + \sum_{k=1}^n c_k^2 + 2\sum_{k=1}^m b_k c_k.$$

Now by Lemma 3, we have

$$\Big(\sum_{k=1}^n b_k c_k\Big)^2 \leq \sum_{k=1}^n b_k^2 \sum_{k=1}^n c_k^2.$$

Take the square root of both sides to obtain

$$\Big(\sum_{k=1}^n b_k c_k\Big) \leq \sqrt{\sum_{k=1}^n b_k^2 \sum_{k=1}^n c_k^2}.$$

Combine this with inequality (*) to obtain

$$\sum_{k=1}^n a_k^2 \leq \sum_{k=1}^n b_k^2 + \sum_{k=1}^n c_k^2 + 2\sqrt{\sum_{k=1}^n b_k^2 \sum_{k=1}^n c_k^2}$$

Taking the square root of both sides produces the result. $\square$

PROOF OF THEOREM 1. The positivity of ρ is clear from the use of positive square root in the definition, and the symmetry is given by the symmetry of the metric on the constituent spaces. Thus is suffices to demonstrate the triangle inequality. Let $a_k = \rho(x_k, z_k)$, $b_k = \rho(x_k, y_k)$, and $c_k = \rho(y_k, z_k)$. By the triangle inequality in the constituent spaces, we have $a_k \leq b_k + c_k$ for $i = 1, \ldots, n$. Apply Lemma 4 to obtain the result. $\square$

EXAMPLE 10. Let (X, ρ) be a metric space, and let X^k denote the cartesian product of k copies of X, endowed with the product metric.

EXAMPLE 11. Let $\mathbb{R}^k$ denote the cartesian product of k copies of the real line. The product metric on $\mathbb{R}^k$ as defined in Example 3 is the same as that defined in Theorem 1.

EXAMPLE 12. Let $\mathbb{S}^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2\}$ be the unit circle, considered as a metric subspace of $\mathbb{R}^2$. Let $\mathbb{T} = \mathbb{S}^1 \times \mathbb{S}^1$, endowed with the product metric. Then $\mathbb{T}$ is a torus.

CHAPTER II

Metric Topology

Abstract. The chapter discusses bounded sets, open sets, closed sets, and neighborhoods in a metric space.

1. Bounded Sets

Definition 3. Let (X, ρ) be a metric space, and let $A \subset X$. Define the *diameter* of A with respect to ρ to be

$$\operatorname{diam}(A) = \sup\{\rho(a, b) \mid a, b \in A\};$$

by convention, the diameter of an empty set is zero. Note that $\operatorname{diam}(A)$ is an extended real number which can be ∞.

We say that A is *bounded* if $\operatorname{diam}(A) < \infty$.

Proposition 1. *Let (X, ρ) be a metric space, and let $A, B \subset X$. Then*

(a) $\operatorname{diam}(A) = 0 \Leftrightarrow |A| \leq 1$;
(b) $A \subset B \Rightarrow \operatorname{diam}(A) \leq \operatorname{diam}(B)$;
(c) $A \cap B \neq \varnothing \Rightarrow \operatorname{diam}(A \cup B) \leq \operatorname{diam}(A) + \operatorname{diam}(B)$.

Proof. Recall that $|A|$ is the *cardinality* of A, and is defined to be the number of elements in A. If A contains at least two distinct elements, the distance between them is positive, so the diameter of A is greater than 0. On the other hand, if A contains exactly one element, say $A = \{a\}$, then $\{\rho(a, b) \mid a, b \in A\} = \{\rho(a, a)\} = \{0\}$, and the supremum of this set is zero.

Suppose $A \subset B \subset X$. Set $S_A = \{\rho(a_1, a_2) \mid a_1, a_2 \in A\}$, and $S_B = \{\rho(b_1, b_2) \mid b_1, b_2 \in B\}$. Clearly $S_A \subset S_B$, so $\operatorname{diam}(A) = \sup(S_A) \leq \sup(S_B) = \operatorname{diam}(B)$.

Finally, suppose that $A, B \subset X$ and that $A \cap B \neq \varnothing$. Suppose that $\operatorname{diam}(A \cup B) > \operatorname{diam}(A) + \operatorname{diam}(B)$, and let $\epsilon = \frac{1}{2}(\operatorname{diam}(A \cup B) - (\operatorname{diam}(A) + \operatorname{diam}(B)))$. Then, from the definition of diameter, there exist points $c_1, c_2 \in A \cup B$ such that $\operatorname{diam}(A \cup B) - \rho(c_1, c_2) > \epsilon$. $\square$

Exercise 3. Let (X, ρ) be a metric space, and let $G = \operatorname{diam}(A)$ with respect to ρ. Define a function

$$\widehat{\rho} : X \times X \to \mathbb{R} \quad \text{by} \quad \widehat{\rho}(x, y) = \frac{\rho(x, y)}{1 + \rho(x, y)}.$$

(a) Show that $\widehat{\rho}$ is a metric on X.

Let $H = \operatorname{diam}(X)$ with respect to $\widehat{\rho}$.

(b) Show that $H \leq 1$.
(c) Show that if $G = \infty$, then $H = 1$.
(d) Show that if X is finite, then $H = \frac{G}{1+G}$.

2. Open Sets

Definition 4. Let (X, ρ) be a metric space. Let $x_0 \in X$ and let $\delta > 0$. Set

$$B(x_0, \delta) = \{x \in X \mid \rho(x, x_0) < \delta\};$$

this is known as an *open ball about x_0 of radius δ*.

Let $U \subset X$. We say that U is *open* if

$$\forall u \in U \exists \delta > 0 \mid B(x_0, \delta) \subset U.$$

Proposition 2. *Let (X, ρ) be a metric space, and let $A \subset X$. Then A is open if and only if A can be expressed as a union of open balls.*

Proof. Suppose that A is open; then for every $a \in A$ there exists $\delta_a > 0$ such that $B(a, \delta_a) \subset A$. Then

$$A = \cup_{a \in A} B(a, \delta_a),$$

so A is a union of open balls.

On the other hand, suppose that A is the union of open balls. Let $a \in A$. Then $a \in B(x, \delta)$ for some $x \in X$ and $\delta > 0$, where $B(x, \delta) \subset A$. Then $B(a, \delta - \rho(x, a)) \subset B(x, \delta)$. To see this, let $b \in B(a, \delta - \rho(x, a))$. Then the triangle inequality implies that

$$\rho(x, b) \leq \rho(x, a) + \rho(a, b) \leq \rho(x, a) + (\delta - \rho(x, a)) = \delta.$$

Thus $B(a, \delta) \subset B(x, \delta) \subset A$, and A satisfies the definition of an open set. $\square$

Proposition 3. *Let (X, ρ) be a metric space. Then*

- **(a)** *The sets $\varnothing$ and X are open.*
- **(b)** *The union of any collection of open subsets of X is open.*
- **(c)** *The intersection of any finite collection of open subsets of X is open.*

Proof. The empty set vacuously satisfies the condition for openness; every $x \in \varnothing$ has an open ball contained in $\varnothing$, because there is no $x \in \varnothing$. If $x \in X$, then $B(x, 1) \subset X$ by definition of $B(x, 1)$.

Suppose that $\{U_\alpha \mid \alpha \in I\}$ is a collection of open subsets of X indexed by the indexing set I. Let $U = \cup_{\alpha \in I} U_\alpha$. Let $x \in U$. Then $x \in U_\alpha$ for some $\alpha \in I$. Since U_α is open, $B(x, \delta) \subset U_\alpha$ for some $\delta > 0$. Then $B(x, \delta) \subset U$, since $U_\alpha \subset U$. Thus U is open.

Suppose that $\{U_1, \ldots, U_n\}$ is a finite collection of open subsets of X. Let $U = \cap_{i=1}^n U_i$, and let $x \in U$. Then $x \in U_i$ for $i = 1, \ldots, n$. Since each of these is open, there exist positive real number $\delta_1, \ldots, \delta_n$ such that $x \in B(x, \delta_i)$ for $i = 1, \ldots, n$.

Set $\delta = \min\{\delta_1, \ldots, \delta_n\}$. Then $B(x, \delta) \subset U_i$ for $i = 1, \ldots, n$. Thus $B(x, \delta) \subset \cap_{i=1}^n U_i = U$. In this way, we see that U is open. $\square$

Definition 5. Let X be a set. A *topology* on X is a collection of subsets $\mathcal{T} \subset \mathcal{P}(X)$ satisfying

- **(T1)** $\varnothing \in \mathcal{T}$ and $X \in \mathcal{T}$;
- **(T2)** if $\mathcal{U} \subset \mathcal{T}$, then $\cup \mathcal{U} \in \mathcal{T}$;
- **(T3)** if $\mathcal{U} \subset \mathcal{T}$ and $\mathcal{U}$ is finite, then $\cap \mathcal{U} \in \mathcal{T}$.

The elements of $\mathcal{T}$ are called *open sets*. The pair $(X, \mathcal{T})$ is called a *topological space*.

OBSERVATION 1. If (X, ρ) is a metric space, then the collection of open subsets of X is a topology on X.

3. Closed Sets

DEFINITION 6. Let (X, ρ) be a metric space. Let $F \subset X$. We say that F is *closed* if $X \setminus F$ is open.

WARNING 1. Just because a set is not open does not mean that it is closed. For example, $[1, 2) \subset \mathbb{R}$ is neither.

PROPOSITION 4. *Let (X, ρ) be a metric space. Then*

(a) *The sets $\varnothing$ and X are closed.*
(b) *The intersection of any collection of closed subsets of X is closed.*
(c) *The union of any collection of closed subsets of X is closed.*

PROOF. Recall *DeMorgan's Laws*, which state that the union of complements is the complement of the intersection, and the intersection of complements is the complement of the union. Then Proposition 4 follows from Proposition 3 and DeMorgan's Laws. $\square$

4. Interior, Closure, and Boundary

DEFINITION 7. Let (X, ρ) be a metric space, and let $A \subset X$. The *interior* of A is the union of all open subsets of A:

$$A^\circ = \bigcap_{\substack{U \subset A \\ U \text{ is open}}} U.$$

Since the union of open sets is open, this is clearly the largest open subset of A.

PROPOSITION 5. *Let (X, ρ) be a metric space and let $A \subset X$. Then A is open if and only if $A = A^\circ$.*

DEFINITION 8. Let (X, ρ) be a metric space, and let $A \subset X$. The *closure* of A is the intersection of all closed subsets of X which contain A:

$$\overline{A} = \bigcap_{\substack{A \subset F \\ F \text{ is closed}}} F.$$

Since the intersection of closed sets is closed, this is clearly the smallest closed subset of X which contains A.

PROPOSITION 6. *Let (X, ρ) be a metric space and let $A \subset X$. Then A is closed if and only if $A = \overline{A}$.*

DEFINITION 9. Let (X, ρ) be a metric space and let $A \subset X$. The *boundary* of A is $\partial A = \overline{A} \setminus A^\circ$.

5. Neighborhoods

DEFINITION 10. Let (X, ρ) be a metric space and let $x \in X$. A *basic open neighborhood* of x is an open ball of the form $B(x, \delta)$ for some $\delta > 0$. An *open neighborhood* of x is any open subset of X which contains x. A *neighborhood* of x is any subset of X which contains an open neighborhood of x.

EXERCISE 4. Let (X, ρ) be a metric space. Let $x \in X$ and let $A, B \subset X$ be neighborhoods of x. Show that $A \cap B$ is a neighborhood of x.

DEFINITION 11. If A and B are sets, we say that A *intersects* B if $A \cap B \neq \varnothing$.

Let (X, ρ) be a metric space. Let $A \subset X$ and $x \in X$.

We say that x is an *interior point* of A if there exists a neighborhood of x which is contained in A.

We say that x is a *closure point* of A if for every neighborhood of x intersects A.

We say that x is a *boundary point* of A if for every neighborhood of x intersects both A and $X \setminus A$.

PROPOSITION 7. *Let (X, ρ) be a metric space. Let $A \subset X$ and $x \in X$. Then*

- **(a)** *x is an interior point of A if and only if $x \in A^\circ$;*
- **(b)** *x is a closure point of A if and only if $x \in \overline{A}$;*
- **(c)** *x is a boundary point of A if and only if $x \in \partial A$.*

DEFINITION 12. Let (X, ρ) be a metric space. Let $A \subset X$ and let $x \in X$.

A *deleted neighborhood* of x is a subset $V \subset X$ such that $V = U \setminus \{x\}$ for some neighborhood U of x.

We say that x is an *isolated point* of A if every deleted neighborhood of x is contained in $X \setminus A$.

We say that x is an *accumulation point* of A if every deleted neighborhood of x intersects A.

PROPOSITION 8. *Let (X, ρ) be a metric space. Let $A \subset X$. Set*

$$B = \{x \in X \mid x \textit{ is an isolated point of } A\};$$

$$C = \{x \in X \mid x \textit{ is an accumulation point of } A\}.$$

Then $\overline{A} = B \cup C$.

CHAPTER III

Completeness

ABSTRACT. This chapter discusses sequences, subsequences, bounded sequences, and Cauchy sequences. In the process, the Bolzano-Weierstrass property and the completeness property of metric spaces are discussed. We show that these properties of a metric space carry over to products.

1. Sequences

DEFINITION 13. Let X be a set. A *sequence* in X is a function $a : \mathbb{N} \to X$. We write a_n instead of $a(n)$, and we write $(a_n)_{n\in\mathbb{N}}$ or simply (a_n) to denote the entire sequence.

One can think of a sequence as an ordered tuple with infinity many entries; hence the notation.

DEFINITION 14. Let (X, ρ) be a metric space and let (a_n) be a sequence in X. Let $p \in X$. We say that (a_n) *converges to* p, and write $\lim_{n\to\infty} a_n = p$, if

$$\forall \epsilon > 0\ \exists N \in \mathbb{N} \mid n \geq N \Rightarrow \rho(a_n, p) < \epsilon.$$

If (a_n) converges to p, we call p a *limit point* of (a_n).

DEFINITION 15. Let (X, ρ) be a metric space and let (a_n) be a sequence in X. Let $q \in X$. We say that (a_n) *clusters at* q if

$$\forall \epsilon > 0\ \forall N \in \mathbb{N}\ \exists n \geq N \mid \rho(a_n, q) < \epsilon.$$

If (a_n) clusters at q, we call q a *cluster point* of (a_n).

EXAMPLE 13. Let $X = \mathbb{R}$ and $\rho(x, y) = |x - y|$. Then our new definitions for convergence and clustering become identical to our previous definitions for this particular case.

EXERCISE 5. Let $\mathbb{S}^1$ be the unit circle together with the subspace metric inherited from $\mathbb{R}^2$. Let (a_n) be the sequence in $\mathbb{S}^1$ defined by

$$a_n = \Big(\cos\frac{2\pi n}{6}, \sin\frac{2\pi n}{6}\Big).$$

Find the cluster points of (a_n).

Exercise 6. Let X be a set and define a metric ρ on X by

$$\rho(x,y) = \begin{cases} 0 & \text{if } x = y; \\ 1 & \text{otherwise} . \end{cases}$$

Let (a_n) be a sequence in X.

(a) Show that $p \in X$ is a limit point of (a_n) if and only if

$$\exists N \in \mathbb{N} \mid n \geq N \Rightarrow a_n = p.$$

(b) Show that $q \in X$ is a cluster point of (a_n) if and only if

$$\forall N \in \mathbb{N} \exists n \geq N \mid a_n = q.$$

Definition 16. Let (X, ρ) be a metric space and let (a_n) be a sequence from X.
For each $N \in \mathbb{N}$, the N^{th} *tail* of (a_n) is defined to be the set

$$\{a_n \mid n \geq N\} = \{x \in X \mid x = a_n \text{ for some } n \geq N\}.$$

Proposition 9. *Let (X, ρ) be a metric space, (a_n) a sequence from X, and $p \in X$. Then the following conditions are equivalent:*
(L1) *For every $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that $n \geq N \Rightarrow \rho(a_n, p) < \epsilon$.*
(L2) *For every neighborhood U of p there exists $N \in \mathbb{N}$ such that $n \geq N \Rightarrow a_n \in U$.*
(L3) *Every neighborhood of p contains a tail of (a_n).*
(L4) *Every neighborhood of p contains a_n for all but finitely many $n \in \mathbb{N}$.*

Proof.
(**L1** $\Rightarrow$ **L2**) Suppose that for every $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that $n \geq N \Rightarrow \rho(a_n, p) < \epsilon$. Let U be a neighborhood of p. Then there exists $\epsilon > 0$ such that $B(p, \epsilon) \subset U$. Let N be so large that $\rho(a_n, p) < \epsilon$ whenever $n \geq N$. Then for $n \geq N$, we have $a_n \in B(p, \epsilon) \subset U$.

(**L2** $\Rightarrow$ **L3**) Suppose that for every neighborhood U of p there exists $N \in \mathbb{N}$ such that $n \geq N \Rightarrow a_n \in U$. Let U be a neighborhood of p and let N be so large that $n \geq N \Rightarrow a_n \in U$. Then $\{a_n \mid n \geq N\} \subset U$, so U contains the N^{th} tail of (a_n).

(**L3** $\Rightarrow$ **L4**) Suppose that every neighborhood U of p contains a tail of (a_n). Let U be a neighborhood of p and let $N \in \mathbb{N}$ such that $\{a_n \mid n \geq N\} \subset U$. If $a_n \notin U$ for some $n \in \mathbb{N}$, then $a_n \notin \{a_n \mid n \geq N\}$, so $n < N$. There are only finitely many such n.

(**L4** $\Rightarrow$ **L1**) Suppose that every neighborhood of p contains a_n for all but finitely many n. Let $\epsilon > 0$. Then $B(p, \epsilon)$ is a neighborhood of p, so $a_n \in B(p, \epsilon)$ for all but finitely many $n \in \mathbb{N}$. Let $N = 1 + \max\{n \in \mathbb{N} \mid a_n \notin B(p, \epsilon)\}$. Then for $n > N$, we have $\rho(a_n, p) < \epsilon$. $\square$

Proposition 10. *Let (X, ρ) be a metric space, (a_n) a sequence from X, and $q \in X$. Then the following conditions are equivalent:*
(C1) *For every $\epsilon > 0$ and every $N \in \mathbb{N}$ there exists $n \geq N$ such that $\rho(a_n, q) < \epsilon$.*
(C2) *For every neighborhood U of q and every $N \in \mathbb{N}$ there exists $n \geq N$ such that $a_n \in U$.*
(C3) *Every neighborhood of q intersects every tail of (a_n).*
(C4) *Every neighborhood of q contains a_n for infinitely many $n \in \mathbb{N}$.*

Proof.
(**C1** $\Rightarrow$ **C2**) Suppose that for every $\epsilon > 0$ and every $N \in \mathbb{N}$ there exists $n \geq N$ such that $\rho(a_n, q) < \epsilon$. Let U be a neighborhood of q and let $N \in \mathbb{N}$. Then there exists $\epsilon > 0$ such that $B(p, \epsilon) \subset U$; thus there exists $n \geq N$ such that $\rho(a_n, q) < \epsilon$. But this says that $a_n \in B(q, \epsilon)$, so $a_n \in U$.

(**C2** $\Rightarrow$ **C3**) Suppose that for every neighborhood U of q and every $N \in \mathbb{N}$ there exists $n > N$ such that $a_n \in U$. Let U be a neighborhood of q and let $\{a_n \mid n \geq N\}$ be an arbitrary tail of (a_n). Then for some $n \geq N$, we have $a_n \in U$. But $a_n \in \{a_n \mid n \geq N\}$, so $a_n \in \{a_n \mid n \geq N\} \cap U$, and $\{a_n \mid n \geq N\}$ intersects U.

(**C3** $\Rightarrow$ **C4**) Suppose that every neighborhood of q intersects every tail of (a_n). Let U be a neighborhood of q. Suppose bwoc that U contains a_n for only finitely many $n \in \mathbb{N}$. Let m be the largest natural number such that $a_m \in U$. Then $[a_n : m + 1]$ is a tail of (a_n) which does not intersect U; this is a contradiction.

(**C4** $\Rightarrow$ **C1**) Suppose that every neighborhood of q contains a_n for infinitely many $n \in \mathbb{N}$. Let $\epsilon > 0$ and $N \in \mathbb{N}$. Then $U = B(q, \epsilon)$ is a neighborhood of q, and U contains a_n for infinitely many $n \in \mathbb{N}$. One such n must be larger than N; if $n \in \mathbb{N}$ such that $a_n \in U$, then $\rho(a_n, q) < \epsilon$. □

Proposition 11. *Let (X, ρ) be a metric space, (a_n) a sequence from X, and $p \in X$. If (a_n) converges to p, then (a_n) clusters at p, and p is the only cluster point.*

Proof. Suppose that (a_n) converges to p. Then every neighborhood of p contains a_n for all but finitely many n. Thus there are infinitely many n such that a_n is in the neighborhood. By Proposition 10 **(d)**, (a_n) clusters at p.

To see that p is the only cluster point, let $q \in X$, $q \neq p$; we show that (a_n) does not cluster at q. Let $\epsilon = \frac{\rho(p,q)}{2}$ and let $U = B(p, \epsilon)$ and $V = B(q, \epsilon)$. Then U and V are disjoint neighborhoods of p and q respectively.

Let A be a tail of (a_n) such that $A \subset U$. Since $U \cap V = \varnothing$, we have $A \cap V = \varnothing$, so V is a neighborhood of q which does not intersect A. Thus (a_n) does not cluster at q, by 10 **(c)**. □

Exercise 7. Find an example of a sequence (a_n) of real numbers and a real number $q \in \mathbb{R}$ such that (a_n) clusters at q but does not converge to q.

2. Subsequences

Definition 17. Let (X, ρ) be a metric space and let (a_n) be a sequence in X, where $a : \mathbb{N} \to X$ is the function defining (a_n). A *subsequence* of (a_n) is the composition $a \circ n$ of a with a strictly increasing sequence $n : \mathbb{N} \to \mathbb{N}$ of positive integers. Let $n_k = n(k)$, and denote the subsequence by (a_{n_k}).

Proposition 12. *Let (X, ρ) be a metric space and let (a_n) be a sequence in X, Then $q \in X$ is a cluster point of (a_n) if and only if (a_n) has a subsequence which converges to q.*

3. Bounded Sequences

Definition 18. Let (X, ρ) be a metric space and let (a_n) be a sequence in X. We say that (a_n) is *bounded* if the set $\{a_n \mid n \in \mathbb{N}\}$ is a bounded set.

Exercise 8. Let (X, ρ) be a metric space and let (a_n) be a sequence in X. Then (a_n) is bounded if and only if there exists a point $c \in X$ and a positive real number $R > 0$ such that $\rho(a_n, c) \leq R$ for all $n \in \mathbb{N}$.

Definition 19. Let (X, ρ) be a metric space. We say that X has the *Bolzano-Weierstrass property* if every bounded sequence in X has a convergent subsequence.

Example 14. We have already shown that $\mathbb{R}$ has the Bolzano-Weierstrass property.

Proposition 13. *Let (X, ρ) be a metric space. Then X has the Bolzano-Weierstrass property if and only if every sequence has a cluster point.*

Proof. This follows immediately from Proposition 12. $\square$

Proposition 14. *Let (X, ρ) be a metric space. Then X has the Bolzano-Weierstrass property if and only if every bounded infinite subset of X has an accumulation point.*

Proof. Suppose that X has the Bolzano-Weierstrass property. Then every bounded sequence in X has a cluster point. Let $A \subset X$ be a bounded infinite set. Since A is infinite, there exists an injective function $a : \mathbb{N} \to A$. This produces a sequence (a_n). This sequence is bounded, so it has a cluster point, say $q \in X$.

We claim that q is an accumulation point of A. To see this, let U be a neighborhood of q. Since q is a cluster point, U contains a_n for infinitely many n. Since a is injective, at $a_n = q$ for at most one n. Thus $U \setminus \{q\}$ contains a_n for some n, and $a_n \in A$. Thus U intersects A, and q is a cluster point.

Suppose that every bounded infinite subset of X has an accumulation point. Let (a_n) be a sequence in A. Let $B = \{a_n \mid n \in \mathbb{N}\}$. If B is finite, then there exists $b \in B$ such that $b = a_n$ for infinitely many n. In this case, b is a cluster point of A. On the other hand, if B is infinite, it has an accumulation point, and this accumulation point will be a cluster point of (a_n). $\square$

4. Cauchy Sequences

DEFINITION 20. Let (X, ρ) be a metric space and let (a_n) be a sequence in X. We say that (a_n) is a *Cauchy sequence* if

$$\forall \epsilon > 0 \exists N \in \mathbb{N} \mid m, n \geq N \Rightarrow \rho(a_m, a_n) < \epsilon.$$

DEFINITION 21. Let (X, ρ) be a metric space. We say that X is *complete* if every Cauchy sequence in X converges.

This definition of completeness appears different than the completeness axiom which we use to obtain the reals from the rationals. We now relabel that definition.

DEFINITION 22. Let S be an ordered set. We say that S has the *supremum property* if every subset of S which is bounded above has a least upper bound. We say that S has the *infimum property* if every subset of S which is bounded below has a greatest lower bound.

We have already shown that a sequence in $\mathbb{R}$ converges if and only if it is a Cauchy sequence. We now show that for subsets of $\mathbb{R}$, the supremum and infimum properties are equivalent to the new completeness property; in this way, the new definition is a generalization of the old one.

PROPOSITION 15. *Let $A \subset \mathbb{R}$. Then A is a complete metric subspace of $\mathbb{R}$ if and only if A has the supremum and infimum properties.*

PROOF. Suppose that A is a complete metric subspace of $\mathbb{R}$. Then every Cauchy sequence in A converges to a point in A. Let $B \subset A$ be bounded above; Then B has a supremum in the reals, say $x = \sup B$. Then for each $n \in \mathbb{N}$, there exists $b_n \in B$ such that $x - b_n < \frac{1}{2^n}$. Then for $m < n$, we have $|b_n - b_m| \frac{1}{2^n}$. Therefore (b_n) is a Cauchy sequence, which converges to a point in A. But clearly $\lim b_n = a$, so $\sup B = x \in A$. Similarly, B has the infimum property.

On the other hand, suppose that A has the supremum and infimum properties, and let (a_n) be a Cauchy sequence in A. Then (a_n) converges in $\mathbb{R}$, say to $x \in \mathbb{R}$. Let $u_n = \inf\{a_m \mid m \geq n\}$. Since A has the infimum property, $u_n \in A$ for every $n \in \mathbb{N}$. Also, (u_n) is an increasing sequence which converges to x, so $x = \sup\{u_n \mid n \in \mathbb{N}\}$. Since A has the supremum property, this is also in A. Thus every Cauchy sequence in A converges to a point in A. $\square$

PROPOSITION 16. *Let (X, ρ) be a metric space and let (a_n) be a Cauchy sequence in X. Then (a_n) is bounded.*

PROOF. Since (a_n) is a Cauchy sequence, there exists $N \in \mathbb{N}$ such that $m, n \geq N$ implies $\rho(a_m, a_n) < 1$. Let $M = \max\{\rho(a_i, a_N) \mid i < N\} \cup \{1\}$. Then $\rho(a_n, a_N) < M$ for every $n \in \mathbb{N}$. $\square$

PROPOSITION 17. *Let (X, ρ) be a metric space and let (a_n) be a Cauchy sequence in X. If (a_n) has a subsequence converging to $p \in X$, then (a_n) converges to p.*

PROOF. Suppose that (a_{n_k}) is a subsequence of (a_n) which converges to $p \in X$. Let $\epsilon > 0$, and let K be so large that $k \geq K$ implies that $\rho(a_{n_k}, p) < \frac{\epsilon}{2}$. Let M be so large that $m, n \geq M$ implies $\rho(a_m, a_n) < \frac{\epsilon}{2}$. Let $N = \max\{K, M\}$. Then for $n \geq N$, we have

$$\rho(a_n, p) \leq \rho(a_n, a_N) + \rho(a_N, p) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Therefore (a_n) converges to p. □

PROPOSITION 18. *Let (X, ρ) be a metric space. If X has the Bolzano-Weierstrass property, then X is complete.*

PROOF. Suppose that X has the Bolzano-Weierstrass property, and let (a_n) be a Cauchy sequence. By Proposition 16, (a_n) is bounded, and so has a convergent subsequence. By Proposition 17, (a_n) converges. Thus X is complete. □

We have seen that if a metric space has the Bolzano-Weierstrass property, then it is complete. One may conjecture that these properties are equivalent. The following counterexample shows this is not the case.

EXAMPLE 15. Let X be any set any consider the discrete metric on X such that the distance between distinct points equals 1. In this space, Cauchy sequences are eventually constant, and so they converge. Thus X is complete. However, *every* sequence in X is bounded, so X has the Bolzano-Weierstrass property if and only if X is finite.

Next we would like to show the following propositions.

PROPOSITION 19. *A sequence converges in $\mathbb{R}^k$ if and only if each of the coordinate sequences converges. A sequence is Cauchy in $\mathbb{R}^k$ if and only if each of the coordinate sequences is Cauchy. The metric space $\mathbb{R}^k$ is complete.*

PROPOSITION 20. **Bolzano-Weierstrass Theorem**
Every bounded sequence in $\mathbb{R}^k$ has a convergent subsequence.

DISCUSSION. Proposition 19 is a lemma for Proposition 20, which is a generalization of the Bolzano-Weierstruass Theorem which we have already shown for $\mathbb{R}$ (the case $k = 1$). However, these propositions can be generalized even further, and we postpone the proofs for this more general context, which we take up next. □

5. Product Space Sequences

PROPOSITION 21. *Let $(X_1, \rho_1), \ldots, (X_k, \rho_k)$ be a finite collection of metric spaces. Let $X = \times_{i=1}^{k} X_i$, and let $\rho : X \times X \to \mathbb{R}$ be the product metric on X. Then*

- **(a)** *A sequence is bounded in X if and only if each of the coordinate sequences is bounded.*
- **(b)** *A sequence converges in X if and only if each of the coordinate sequences converges.*

(c) *A sequence is Cauchy in X if and only if each of the coordinate sequences is Cauchy.*
(d) *The metric space X is complete if and only if each of the spaces X_i is complete.*
(e) *The metric space X has the Bolzano-Weierstrass property if and only if each of the spaces X_i has the Bolzano-Weierstrass property.*

PRELIMINARY OBSERVATION. Now suppose that $x = (x_1, \dots, x_k)$ and $y = (y_1, \dots, y_k)$ are points in X, where $x_j, y_j \in X_i$. Observe that, since all metrics are positive, we have

$$\rho_j(x_j, y_j) \leq \sqrt{\sum_{i=1}^{k} \rho(x_i, y_i)} = \rho(x, y) \leq \sqrt{k} \max\{\rho(x_i, y_i) \mid i = 1, \dots, k\}.$$

□

NOTATION. A point in X is an k-tuple with entries for X_1 through X_k. If we denote these entries with subscripts, we must find another place to indicate the position of such an k-tuple in a sequence. Thus let $(x^{(n)})$ denote a sequence in X, where

$$x^{(n)} = (x_1^{(n)}, \dots, x_k^{(n)}),$$

where $x_i^{(n)} \in X_i$. □

PROOF OF **(a)**. This follows from the observation. □

PROOF OF **(b)**. Suppose that $(x_i^{(n)})$ converges for $i = 1, \dots, k$, say to $L_i \in X_i$. Let $L = (L_1, \dots, L_k)$. Let $\epsilon > 0$. Let N be so large that $\rho_i(x_i^{(n)}, L_i^{(n)}) < \frac{1}{k}\epsilon^2$ for $n \geq N$. Then for $n \geq N$ we have

$$\rho(x_n, L) = \sqrt{\sum_{i=1}^{k} \rho(x_i^{(n)}, L_i)} < \sqrt{\sum_{i=1}^{k} \frac{1}{k}\epsilon^2} = \sqrt{k(\frac{1}{k}\epsilon^2)} = \epsilon.$$

Therefore $\lim x^{(n)} = L$, and in particular, $(x^{(n)})$ converges.

Suppose that $(x^{(n)})$ converges, say to $L = (L_1, \dots, L_k)$. Let $\epsilon > 0$ and let n be so large that $\rho(x^{(n)}, L) < \epsilon$ for $n \geq N$. Then for i between 1 and k we have

$$\rho_i(x_i^{(n)}, L_i) \leq \rho(x^{(n)}, L) < \epsilon.$$

Thus $\lim_{n\to\infty} x_i^{(n)} = L_i$, and in particular, the sequence $(x_i^{(n)})$ converges. □

PROOF OF **(c)**. Suppose that $(x_i^{(n)})$ is a Cauchy sequence for $i = 1, \dots, k$. Let $\epsilon > 0$ and let N be so large that $m, n \geq N$ implies

$$\rho_i(x_i^{(m)}, x_i^{(n)}) < \frac{\epsilon}{\sqrt{k}}$$

for all $i = 1, \dots, k$. Then by the observation, we have

$$\rho(x^{(m)}, x^{(n)}) \leq \epsilon.$$

Suppose that $(x^{(n)})$ is a Cauchy sequence. Let $\epsilon > 0$. Let N be so large that $m, n \geq N$ implies $\rho(x^{(m)}, x^{(n)}) < \epsilon$. Then for $m, n \geq N$, we have

$$\rho_i(x_i^{(m)}, x_i^{(n)}) \leq \rho(x^{(m)}, x^{(n)}) < \epsilon,$$

we says that the coordinate sequence $(x_i^{(n)})$ is a Cauchy sequence. $\square$

PROOF OF **(d)**. We know that a metric space is complete if and only if each of its Cauchy sequences converges.

Suppose each space (X_i, ρ_i) is complete, and consider a Cauchy sequence in X. Each of the coordinate sequences are Cauchy by part **(b)**, so each converges since X_i is complete. Then the original sequence converges by part **(a)**, so X is complete.

On the other hand, suppose that (X, ρ) is complete, and let $i \in \{1, \ldots, k\}$. Consider a Cauchy sequence in X_i. Construct a sequence in X by selecting a constant $a_i \in X_i$ in every coordinate other than the i^{th}. These are all Cauchy sequences in the coordinate spaces, so the construct sequence in X converges. Thus the original sequence in X_i converges, and X_i is complete. $\square$

PROOF OF **(e)**. Suppose that X_i has the Bolzano-Weierstrass property for $i = 1, \ldots, k$. Then each bounded sequence in X_i has a convergent subsequence. Given a bounded sequence in X, each of the coordinate sequences is bounded, and has a convergent subsequence. Select a convergent subsequence X_1 for the first coordinate subsequence, and take the corresponding subsequence in X. Now select a convergent subsequence in X_2 for the second coordinate subsequence of the new sequence in X, and again take the corresponding subsequence in X. Continue this process k times, and arrive at a sequence in X such that every subsequence converges. This sequence is a subsequence of the original sequence in X, and it converges. Thus X has the Bolzano-Weierstrass property.

Suppose that X has the Bolzano-Weierstrass property. Let $i \in \{1, \ldots, k\}$ and let consider a bounded sequence in X_i. Construct a sequence in X by selecting a constant $a_i \in X_i$ in every coordinate other than the i^{th}. This is bounded in X, and so has a convergent subsequence. The i^{th} coordinate sequence of this subsequence converges in X_i, and is a subsequence of the original sequence in X_i. Thus X_i has the Bolzano-Weierstrass property. $\square$

COROLLARY 1. *The space $\mathbb{R}^k$ is complete and has the Bolzano-Weierstrass property.*

Example 16. Consider $\mathbb{R}^\infty$, whose points are all infinite tuples of real numbers with all but finitely many entries equal to zero. Construct a sequence $(x^{(n)})$ in $\mathbb{R}^\infty$ by setting

$$x_i^{(n)} = \begin{cases} 1 & \text{if } i = n; \\ 0 & \text{otherwise.} \end{cases} \tag{$\dagger$}$$

Then $(x^{(n)})$ is bounded (it is completely contained inside the closed unit ball), yet has no convergent subsequence. Thus $\mathbb{R}^\infty$ does not have the Bolzano-Weierstrass property. Note that the sequence above is not a Cauchy sequence.

However, consider this example. Construct a sequence $(y^{(n)})$ in $\mathbb{R}^\infty$ by setting

$$y_i^{(n)} = \begin{cases} \frac{1}{2^i} & \text{if } i \leq n; \\ 0 & \text{otherwise.} \end{cases}$$

This is a Cauchy sequence in $\mathbb{R}^\infty$ which does not converge in $\mathbb{R}^\infty$. So this space is not complete.

Example 17. Let ℓ^2 be the space of sequences (x_n) in $\mathbb{R}$ with the convergence criterion $\sum_{i=1}^{\infty} x_i^2 < 0$. Then $\mathbb{R}^\infty$ is a subspace of ℓ^2, and the sequence ($\dagger$) from Example 16 does not have a convergent subsequence in ℓ^2.

However, ℓ^2 is complete. To show this, proceed as follows. Consider a Cauchy sequence $(x_i^{(n)})$ in ℓ^2. Show that the coordinate sequences are Cauchy, and so the converge in $\mathbb{R}$; say that $(x_i^{(n)})$ converges to x_i for each i. Next see that the sequence (x_i) is in ℓ^2.

Clearly there is some relationship between the Bolzano-Weierstrass property and completeness. We need the concept of *compactness* to illuminate this further.

CHAPTER IV

Continuity

1. Continuous Functions

1.1. Continuity at a Point.

DEFINITION 23. Let (X, ρ) and (Y, τ) be metric spaces. Let $f : X \to Y$ and let $a \in X$. We say that f is *continuous at a* if

$$\forall \epsilon > 0 \; \exists \delta > 0 \mid \rho(x, a) < \delta \Rightarrow \tau(f(x), f(a)) < \epsilon.$$

EXAMPLE 18. Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = 3x - 7$ and let $a \in \mathbb{R}$. Show that f is continuous at a.

SOLUTION. Let $\epsilon > 0$ and let $\delta = \frac{\epsilon}{3}$. Then

$$|x-a| < \delta \Rightarrow |x-a| < \frac{\epsilon}{3} \Rightarrow |3x-3a| < \epsilon \Rightarrow |3x-7-(3a-7)| < \epsilon \Rightarrow |f(x)-f(a)| < \epsilon.$$

□

1.2. Sequential Continuity at a Point.

DEFINITION 24. Let (X, ρ) and (Y, τ) be metric spaces. Let $f : X \to Y$ and let $a \in X$. We say that f is *sequentially continuous at a* if for every sequence (x_n) in X converging to a, the sequence $(f(x_n))$ in Y converges to $f(a)$.

PROPOSITION 22. *Let (X, ρ) and (Y, τ) be metric spaces. Let $f : X \to Y$ and let $a \in X$. Then f is continuous at a if and only if f is sequentially continuous at a.*

PROOF. We prove both directions of this implication.

($\Rightarrow$) Suppose that f is continuous at a. Let (x_n) be a sequence in X which converges to a; we wish to show that (x_n) converges to $f(a)$. Let $\epsilon > 0$. Since f is continuous at a, there exists $\delta > 0$ such that $\rho(x, a) < \delta$ implies $\tau(f(x), f(a)) < \epsilon$. Let N be so large that $n \geq N$ implies $\rho(x, a) < \delta$. Then, for $n \geq N$, we have $\tau(f(x), f(a)) < \epsilon$.

($\Leftarrow$) Suppose that f is not continuous at a. Then

$$\exists \epsilon > 0 \mid \forall \delta > 0 \; \exists x \in X, \rho(x, a) < \delta \mid \tau(f(x), f(a)) \geq \epsilon.$$

Let ϵ satisfy the above condition, and for $n \in \mathbb{N}$, let $x_n \in X$ be such that $\rho(x, a) < \frac{1}{n}$, but $\tau(f(x_n), f(a)) \geq \epsilon$. Then (x_n) converges to a, but $f(x_n)$ does not converge to $f(a)$. Therefore, f is not sequentially continuous at a. □

1.3. Topological Continuity at a Point.

Definition 25. Let (X, ρ) and (Y, τ) be metric spaces. Let $f : X \to Y$ and let $a \in X$. We say that f is *topologically continuous at a* if for every open neighborhood V of $f(a)$ there exists an open neighborhood U of a such that $f(U) \subset V$.

Observation 2. Let (X, ρ) and (Y, τ) be metric spaces. Let $f : X \to Y$ and let $a \in X$. The following conditions are equivalent:

(a) $\rho(x, a) < \delta \Rightarrow \tau(f(x), f(a)) < \epsilon$;
(b) $x \in B(a, \delta) \Rightarrow f(x) \in B(f(a), \epsilon)$;
(c) $f(B(a, \delta)) \subset B(f(a), \epsilon)$.

Proposition 23. *Let (X, ρ) and (Y, τ) be metric spaces. Let $f : X \to Y$ and let $a \in X$. Then f is continuous at a if and only if f is topologically continuous at a.*

Proof. We prove both directions.

($\Rightarrow$) Suppose that f is continuous at a, and let V be an open neighborhood of $f(a)$. Then there exists $\epsilon > 0$ such that $B(f(a), \epsilon) \subset V$. Since f is continuous, there exists $\delta > 0$ such that $f(B(a, \delta) \subset B(f(a), \epsilon)$. Let $U = B(a, \delta)$; then $f(U) \subset V$.

($\Leftarrow$) Suppose that f is topologically continuous at a. Let $\epsilon > 0$ and let $V = B(f(a), \epsilon)$. Then V is an open neighborhood of $f(a)$ in Y, so there exists an open neighborhood U of a in X such that $f(U) \subset V$. Since U is open and $a \in U$, there exists $\delta > 0$ such that $B(a, \delta) \subset U$. Then $f(B(a, \delta)) \subset B(f(a), \epsilon)$. □

1.4. Continuity on a Space.

Definition 26. Let (X, ρ) and (Y, τ) be metric spaces. Let $f : X \to Y$ and let $A \in X$. We say that f is *continuous on A* if f is continuous at a for every $a \in A$. If f is continuous on X, we say simply that f is *continuous.*

Example 19. Let $f(x) = \frac{3x-2}{x^2-7x+10}$. The natural real domain of this function is $X = \mathbb{R} \smallsetminus \{2, 5\}$. Thus $f : X \to \mathbb{R}$, and f is continuous. In fact, rational functions are always continuous on their domains.

Definition 27. Let (X, ρ) and (Y, τ) be metric spaces. Let $f : X \to Y$. We say that f is *topologically continuous* if for every open set V in Y, $f^{-1}(V)$ is open in X.

Proposition 24. *Let (X, ρ) and (Y, τ) be metric spaces. Let $f : X \to Y$. Then f is continuous if and only if f is topologically continuous.*

Proof. We prove both directions.

($\Rightarrow$) Suppose that f is continuous, and let $V \subset Y$ be open. Let $U = f^{-1}(V)$, and let $u \in U$. We wish to show that their is an open neighborhood of u contained in U. Since V is open, there exists $\epsilon > 0$ such that $B(f(u), \epsilon) \subset V$. Since f is continuous, there exists $\delta > 0$ such that $\rho(x, u) < \delta$ implies $\tau(f(x), f(u)) < \epsilon$, that is, $x \in B(u, \delta)$ implies that $f(x) \in B(f(u), \epsilon) \subset V$, so that $f(B(u, \delta)) \subset V$. Thus $B(u, \delta) \subset U$, so U is open. Therefore f is topologically continuous.

($\Leftarrow$) Suppose that f is topologically continuous. Let $a \in X$ and let $\epsilon > 0$. Let $V = B(f(a), \epsilon)$; then V is open in Y, so $UY = f^{-1}(V)$ is open in X. Clearly $a \in U$; thus there exists $\delta > 0$ such that $B(a, \delta) \subset U$, and $f(B(a, \delta)) \subset B(f(a), \epsilon)$. This says that $\rho(x, a) < \delta$ implies $\tau(f(x), f(a)) < \epsilon$. □

1.5. Composition of Continuous Functions.

PROPOSITION 25. *Let (X, ρ), (Y, τ), and (Z, χ) be metric spaces. Let $f : X \to Y$ be continuous at $a \in X$ and let $g : Y \to Z$ be continuous at $f(a) \in Y$. Then $g \circ f$ is continuous at a.*

PROOF. Let W be an open neighborhood of $g(f(a))$. Since g is continuous, there exists an open neighborhood V of $f(a)$ such that $g(V) \subset W$. Also, since f is continuous, there exists an open neighborhood U of a such that $f(U) \subset V$. Then $g(f(U)) \subset g(V) \subset W$. □

1.6. Real-Valued Functions.

PROPOSITION 26. *Let (X, ρ) be a metric space, and let $f : X \to \mathbb{R}$ be a real-valued function defined on X. Let $Y = \{x \in X \mid g(x) \neq 0\}$ and let $k \in \mathbb{R}$. Define new functions as follows:*

- $|f| : X \to \mathbb{R}$ *is defined by* $|f|(x) = |f(x)|$.
- $kf : X \to \mathbb{R}$ *is defined by* $(kf)(x) = kf(x)$;
- $\frac{1}{f} : Y \to \mathbb{R}$ *is defined by* $(\frac{1}{f})(x) = \frac{1}{f(x)}$;

If f is continuous at a, then kf and $moduf$ are continuous at a. If $a \in Y$, then $\frac{1}{f}$ is continuous at a.

PROOF. We show, for example, that kf is continuous at $a \in X$. Let (x_n) be a sequence in X that converges to a. Then $\lim kf(x_n) = k \lim f(x_n) = kf(a)$. This shows that kf is continuous at a. □

PROPOSITION 27. *Let (X, ρ) be a metric space, and let $f : X \to \mathbb{R}$ and $g : X \to \mathbb{R}$ be real-valued functions defined on X. Define new functions as follows:*

- $f + g : X \to \mathbb{R}$ *is defined by* $(f + g)(x) = f(x) + g(x)$;
- $fg : X \to \mathbb{R}$ *is defined by* $(fg)(x) = f(x) \cdot g(x)$;
- $\max(f, g) : X \to \mathbb{R}$ *is defined by* $\max(f, g)(x) = \max\{f(x), g(x)\}$;
- $\min(f, g) : X \to \mathbb{R}$ *is defined by* $\min(f, g)(x) = \min\{f(x), g(x)\}$;

If f and g are continuous at a, then the above functions are continuous at a (except when dividing by zero).

PROOF. We show, for example, that $f+g$ is continuous at $a \in X$. Let (x_n) be a sequence in X that converges to a. Then $\lim(f+g)(x_n) = \lim(f(x_n)+g(x_n)) = \lim f(x_n) + \lim g(x_n) = f(a) + g(a) = (f + g)(a)$. This shows that $f + g$ is continuous at a. □

2. Continuity Examples for $f : \mathbb{R} \to \mathbb{R}$

EXAMPLE 20. Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = x^2$. Let $x_0 = 2$. Show that f is continuous at x_0.

PROOF. Let $\epsilon > 0$; we may assume that $\epsilon < 4$. Let $\delta = \sqrt{x_0^2 + \epsilon} - x_0 = \sqrt{4 + \epsilon} - 2$. Thus $(\delta + 2)^2 = 4 + \epsilon$, so $\epsilon = \delta^2 + 4\delta$.

Suppose that $x \in (2 - \delta, 2 + \delta)$. Then $x + 2 < \delta + 4$, and

$$|f(x) - f(x_0)| = |x^2 - 4| = |x - 2|(x + 2) < \delta(4 + \delta) = \epsilon.$$

□

EXAMPLE 21. Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = x^3$. Show that f is continuous.

PROOF. Let $x_0 \in \mathbb{R}$ and let $\epsilon > 0$. We wish to find $\delta > 0$ such that if $|x - x_0| < \delta$, then $|f(x) - f(x_0)| < \epsilon$.

For simplicity, assume that $x_0 > 0$. Let $\delta = \sqrt[3]{x_0^3 + \epsilon} - x_0$. Solving for ϵ yields $\epsilon = (x_0 + \delta)^3 - x_0^3$.

Let $x \in (x_0 - \delta, x_0 + \delta)$. Then $x > 0$, and

$$\begin{aligned}
|f(x) - f(x_0)| &= |x^3 - x_0^3| \\
&= |x - x_0|(x^2 + x_0 x + x_0^2) \\
&< \delta((x_0 + \delta)^2 + x_0(x_0 + \delta) + x_0^2) \\
&= \delta(x_0^2 + 2x_0\delta + \delta^2 + x_0^2 + x_0\delta + x_0^2) \\
&= \delta(3x_0^2 + 3x_0\delta + \delta^2) \\
&= \epsilon.
\end{aligned}$$

□

EXAMPLE 22. Let $f : [0, \infty) \to \mathbb{R}$ be given by $f(x) = \sqrt{x}$. Show that f is continuous.

MOTIVATION. Graph the curve $f(x) = \sqrt{x}$. Select arbitrary $x_0 \in \mathrm{dom}(f)$. Project up and to the right to find the point $\sqrt{x_0}$ on the y-axis. Draw an ϵ-band around this point. Project the intersection of this band with the graph of f onto the x-axis. Notice that the point on the left of this projection is closer to x_0 than is the point on the right. Let δ be one half of the distance between x_0 and the left endpoint of the inverse image of $[f(x_0) - \epsilon, f(x_0) + \epsilon]$. □

PROOF. Let $x_0 \in [0, \infty)$ and let $\epsilon > 0$; wlog assume that $\epsilon^2 \le x_0$. If $x_0 = 0$, let $\delta = \epsilon^2$; clearly this will work. Otherwise set

$$\delta = \frac{1}{2}(x_0 - (\sqrt{x_0} - \epsilon)^2);$$

this is positive. Note that for $x \in \mathbb{R}$, $|x - x_0| = |\sqrt{x} - \sqrt{x_0}|(\sqrt{x} + \sqrt{x_0})$. Then if $|x - x_0| < \delta$, we have

$$\begin{aligned}
|\sqrt{x} - \sqrt{x_0}| &< \frac{\delta}{\sqrt{x} + \sqrt{x_0}} \\
&= \frac{x_0 - (x_0 - 2\sqrt{x_0}\epsilon + \epsilon^2)}{2(\sqrt{x} + \sqrt{x_0})} \\
&= \frac{\epsilon(2\sqrt{x_0} - \epsilon)}{2(\sqrt{x} + \sqrt{x_0})} \\
&< \epsilon\frac{(2\sqrt{x_0} - \epsilon)}{2\sqrt{x_0}} \\
&= \epsilon\Big(1 - \frac{\epsilon}{2\sqrt{x_0}}\Big) \\
&< \epsilon.
\end{aligned}$$

□

EXAMPLE 23. Show that every polynomial function is continuous.

PROOF. This is tedious but obviously important. We build it gradually.

Claim 1: The constant function $f(x) = C$, where $C \in \mathbb{R}$, is continuous.
Let $x_0 \in \mathbb{R}$ and let $\epsilon > 0$. Set $\delta = 1$. Then if $|x - x_0| < \delta$, we have $|f(x) - f(x_0)| = 0 < \epsilon$. Thus f is continuous in this case.

Claim 2: The identity function $f(x) = x$ is continuous.
Let $x_0 \in \mathbb{R}$ and let $\epsilon > 0$. Set $\delta = \epsilon$. Then if $|x - x_0| < \delta$, we have $|f(x) - f(x_0)| = |x - x_0| < \delta = \epsilon$, so f is continuous in this case.

Claim 3: The function $f(x) = x^n$ is continuous.
By induction on n. For $n = 1$, the function $g(x) = x$ is the identity function, and so it is continuous. By induction, $h(x) = x^{n-1}$ is continuous. Then by the Continuous Arithmetic Proposition, $f = gh$ is continuous in this case.

Claim 4: The monomial function $f(x) = a_n x^n$ is continuous, where $a_n \in \mathbb{R}$ is constant.
By Claim 1, $g(x) = a_n$ is continuous, and by Claim 3, $h(x) = x^n$ is continuous, so there product $f = gh$ is continuous.

Claim 5: The polynomial function $f(x) = a_0 + a_1 x + \cdots + a_n x^n$ is continuous.
By induction on n, the degree of the polynomial.

For $n = 0$, $f(x)$ is constant and therefore continuous.

Assume that $g(x) = a_0 + \cdots + a_{n-1}x^{n-1}$ is continuous. By Claim 4, $h(x) = a_n x^n$ is continuous. Then $f = g + h$ is continuous by the Continuous Arithmetic Proposition. □

EXAMPLE 24. Show that every rational function is continuous.

PROOF. Let f be a rational function. Then $f(x) = p(x)/q(x)$, where p and q are polynomial functions. Since p and q are continuous, then f is continuous on its domain by a Proposition from the arithmetic of continuous functions. □

EXAMPLE 25. Let $f : \mathbb{R} \to \mathbb{R}$ be given by

$$f(x) = \begin{cases} 1 & \text{if } x \text{ is rational} \\ 0 & \text{if } x \text{ is irrational} \end{cases}$$

Show that f is discontinuous at every real number.

PROOF. Let $x_0 \in \mathbb{R}$. To show that f is discontinuous at x_0, it suffices to find $\epsilon > 0$ such that for every $\delta > 0$, there exists $x \in (x_0 - \delta, x_0 + \delta)$ with $|f(x) - f(x_0)| \geq \epsilon$.

Let $\epsilon = \frac{1}{2}$ and let $\delta > 0$. Then there exists both a rational and an irrational in $(x_0 - \delta, x_0 + \delta)$. If x_0 is rational, let x_1 be an irrational in this interval, and we have $|f(x_1) - f(x_0)| = 1 > \epsilon$; if x_0 is irrational, let x_2 be a rational in this interval, and we still have $|f(x_2) - f(x_0)| = 1 > \epsilon$. Thus f is not continuous at x_0. □

EXAMPLE 26. Let $f : \mathbb{R} \to \mathbb{R}$ be given by

$$f(x) = \begin{cases} x & \text{if } x \text{ is rational} \\ 0 & \text{if } x \text{ is irrational} \end{cases}$$

Show that f is continuous at $x = 0$ and discontinuous at all nonzero real numbers.

PROOF. Let $x_0 \in \mathbb{R} \setminus \{0\}$; we show that f is discontinuous at x_0. Let $\epsilon = \frac{|x_0|}{2}$ and let $\delta > 0$. Then there exists both a rational and an irrational in $(x_0 - \delta, x_0 + \delta)$. If x_0 is rational, let x_1 be an irrational in this interval, and we have $|f(x_1) - f(x_0)| = |x_0| > \epsilon$. If x_0 is irrational, let x_2 be a rational in this interval such that $|x_2| > |x_0|$ and we still have $|f(x_2) - f(x_0)| = |x_2| > |x_0| > \epsilon$. Thus f is not continuous at x_0.

Now we consider the behavior of f at zero. Let $\epsilon > 0$ and let $\delta = \epsilon$. Then if $|x - 0| < \delta$, we have $|f(x) - f(0)| = 0$ if x is irrational and $|f(x) - f(0)| = |x|$ if x is rational; in either case, $|f(x) - f(0)| \leq |x| < \delta = \epsilon$, so f is continuous at zero. $\square$

EXAMPLE 27. If $r \in \mathbb{Q}$, there exists $p \in \mathbb{Z}$ and $q \in \mathbb{N}$ such that $r = \frac{p}{q}$. Define $q : \mathbb{Q} \to \mathbb{R}$ by

$$q(r) = \min\{q \in \mathbb{N} \mid r = \frac{p}{q} \text{ for some } p \in \mathbb{Z}\}.$$

Define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} 0 & \text{if } x \text{ is irrational} \\ \frac{1}{q(x)} & \text{if } x \text{ is rational} \end{cases}$$

Show that f is discontinuous at every rational and continuous at every irrational.

PROOF. Suppose that x_0 is rational. We wish to show that f is not continuous at x_0. It suffices to find $\epsilon > 0$ such that for every $\delta > 0$ there exists $x_1 \in (x_0 - \delta, x_0 + \delta)$ with $|x_0 - x_1| > \epsilon$.

Since x_0 is rational, we have $x_0 = \frac{p}{q(x_0)}$ for some $p \in \mathbb{Z}$. Let $\epsilon = \frac{1}{2q(x_0)}$ and let $\delta > 0$. Then $(x_0 - \delta, x_0 + \delta)$ contains an irrational number, say x_1; then $|x_0 - x_1| < \delta$ but $|f(x_0) - f(x_1)| = \frac{1}{q(r)} > \epsilon$. Thus f cannot be continuous at x_0.

Suppose that x_0 is irrational. Let $\epsilon > 0$. It suffices to find $\delta > 0$ such that $|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon$.

Let $N \in \mathbb{N}$ be so large that $\frac{1}{N} < \epsilon$. Let a be the greatest integer which is less than x_0 and b be the least integer which is greater than x_0; then $b = a + 1$ and $x_0 \in [a, b]$.

For $q \in \mathbb{Q}$, there exist only finitely many points in the set $[a, b] \cap \{\frac{k}{q} \mid k \in \mathbb{Z}\}$ (in fact, this set contains no more than q points). Thus the set

$$D = [a, b] \cap \{\frac{k}{q} \mid k \in \mathbb{Z}, q \leq N\}$$

is finite (there are no more than $\frac{N(N+1)}{2}$ points in this set). Let

$$\delta = \min\{|x_0 - d| \mid d \in D\};$$

since this set is a finite set of positive real numbers, the minimum exists as a positive real number. Then $(x_0 - \delta, x_0 + \delta) \subset [a, b]$. Let $x \in (x_0 - \delta, x_0 + \delta)$. If x is irrational, we have $|f(x) - f(x_0)| = 0 < \epsilon$, and if x is rational, we have $|f(x) - f(x_0)| = \frac{1}{q(x)} < \frac{1}{N} < \epsilon$. Thus f is continuous at x_0. $\square$

3. Isometries,Contractions, and Homeomorphisms

3.1. Isometries.

DEFINITION 28. Let (X, ρ) and (Y, τ) be metric spaces. Let $f : X \to Y$. We say that f *preserves distance* if

$$\tau(f(a), f(b)) = \rho(a, b).$$

A bijective function which preserves distance is called an *isometry*.

An injective function which preserves distance is called an *isometric embedding*.

PROPOSITION 28. *Let (X, ρ) and (Y, τ) be metric spaces. Let $f : X \to Y$ which preserves distance. Then f is an isometric embedding.*

PROOF. We only have to show that f is injective. Let $a, b \in X$ such that $f(a) = f(b)$. Then $\rho(a, b) = \tau(f(a), f(b)) = 0$, so by property **(M1)** of a metric, $a = b$. Thus f is injective. □

EXAMPLE 28. Let $f : \mathbb{R} \to \mathbb{R}$ be an isometry. Then $f(x) = ux + b$ for some $b \in \mathbb{R}$, where $u = \pm 1$.

PROOF. Let $b = f(0)$. Now let $x \in \mathbb{R}$ and let $y = f(x)$. Now $x = |x - 0| = |f(x) - f(0)| = |y - b|$, so $x = \pm(y - b)$; thus $y = ux + b$, where $u = \pm 1$. It remains to show that u is independent of x.

Thus assume that $f(x_1) = x_1 + b$ and $f(x_2) = -x_2 + b$; it suffices to show that $x_1 = 0$ or $x_2 = 0$. Now $|x_1 - x_2| = |f(x_1) - f(x_2)| = |x_1 + b - (-x_2 + b)| = |x_1 + x_2|$. Squaring both sides and cancelling yields $-x_1x_2 = x_1x_2$, so either $x_1 = 0$ or $x_2 = 0$. □

EXAMPLE 29. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be an isometry. Then exactly one of these conditions hold:

- **(1)** there exists a line $y = mx + b$ such that f is reflection across this line;
- **(2)** there exists a point (x_0, y_0) and an angle α such that f is rotation by α around (x_0, y_0).

EXERCISE 9. Describe the isometries of $\mathbb{R}^3$.

EXERCISE 10. Let $X = \{(\cos\alpha, \sin\alpha) \in \mathbb{R}^2 \mid \alpha = \frac{2\pi}{n} \text{ for some } n \in \mathbb{Z}\}$. Describe the isometries of X.

3.2. Contractions.

DEFINITION 29. Let (X, ρ) and (Y, τ) be metric spaces. Let $f : X \to Y$. We say that f is a *contraction* if there exists $M > 0$ such that

$$\tau(f(a), f(b)) \leq M\rho(a, b)$$

for all $a, b \in X$.

EXAMPLE 30. An isometric embedding is a contraction.

PROPOSITION 29. *Let (X, ρ) and (Y, τ) be metric spaces. Let $f : X \to Y$ be a contraction. Then f is continuous.*

PROOF. Since f is a contraction, there exists $M > 0$ such that $\tau(f(a), f(b)) \leq M\rho(a, b)$ for every $a, b \in X$.

Let $a \in X$ and let $\epsilon > 0$. Let $\delta = \frac{\epsilon}{M}$. Then

$$\rho(a, b) < \delta \Rightarrow \rho(a, b) < \frac{\epsilon}{M} \Rightarrow M\rho(a, b) < \epsilon \Rightarrow \tau(f(a), f(b)) < \epsilon.$$

□

3.3. Homeomorphisms.

DEFINITION 30. Let (X, ρ) and (Y, τ) be metric spaces. Let $f : X \to Y$. We say that f is a *homeomorphism* if f is a continous bijective function whose inverse is also bijective.

It is natural to suppose that a continuous bijective function always has a continuous inverse. This is not the case.

EXAMPLE 31. Let $X = (0, 1) \cup [2, 3)$ and let $Y = (0, 2)$. Define $f : X \to Y$ by

$$f(x) = \begin{cases} x & \text{if } x \in (0, 1); \\ x - 1 & \text{if } x \in [2, 3). \end{cases}$$

This function is clearly bijective and continuous at every point in X; however, its inverse is discontinuous.

EXAMPLE 32. Let $X = \mathbb{R}$. Let ρ be the standard metric on X and let τ be the discrete metric. Then $\mathrm{id} : (X, \tau) \to (X, \rho)$ is bijective and continuous, but the inverse is not continuous.

EXAMPLE 33. Let $X = (-\frac{\pi}{2}, \frac{\pi}{2})$ and $Y = \mathbb{R}$, endowed with the usual metric. Let $f : X \to Y$ be given by $f(x) = \tan x$. Then f is bijective and continuous, and its inverse is $f^{-1}(x) = \arctan x$. Thus a homeomorphism can map a bounded space onto an unbounded space.

EXAMPLE 34. Let (X, ρ) be a metric space, and define

$$\widehat{\rho} : X \times X \to \mathbb{R} \quad \text{by} \quad \widehat{\rho}(x, y) = \frac{\rho(x, y)}{1 + \rho(x, y)}.$$

We have seen that $(X, \widehat{\rho})$ is a metric space. View the identity map $\mathrm{id}_X : X \to X$ as a function from (X, ρ) to $(X, \widehat{\rho})$. Then id_X is a bijective contraction, and its inverse is also continuous. Thus id_X is a homeomorphism.

DEFINITION 31. Let X be a set and let ρ and τ be metrics on X. We say that ρ and τ are *equivalent* if they produce the same open sets.

DEFINITION 32. Let (X, ρ) be a metric space. The *topology induced by ρ on X* is

$$\mathcal{T}_{(X,\rho)} = \{U \subset X \mid U \text{ is open in } X\}.$$

PROPOSITION 30. *Let X be a set and let ρ and τ be metrics on X. The following conditions are equivalent:*

- **(a)** $\mathrm{id} : (X, \rho) \to (X, \tau)$ *is a homeomorphism;*
- **(b)** $\mathcal{T}_{(X,\rho)} = \mathcal{T}_{(X,\tau)}$*;*
- **(c)** *every open ball with respect to one metric contains an open ball with respect to the other metric.*

If any of these conditions hold, we say that the metrics ρ and τ are equivalent.

4. Projections and Injections

4.1. Open Maps.

DEFINITION 33. Let (X, ρ) and (Y, τ) be metric spaces, and let $f : X \to Y$ be a function.

We say that f is an *open map* if for every open set U in X, $f(U)$ is open in Y.

EXAMPLE 35. Let (X, ρ) be any metric space, and let (Y, τ) be a discrete metric space. Then every function $f : (X, \rho) \to (Y, \tau)$ is an open map.

EXAMPLE 36. The function $\sin : \mathbb{R} \to \mathbb{R}$ is a continuous function which sends the open set $(0, 4\pi)$ to the closed interval $[0, 1]$. Thus sin is not an open map.

EXAMPLE 37. Let $\mathbb{S}^1$ denote the unit circle in the Euclidean plane, endowed with the subspace metric. The function $f : \mathbb{R} \to \mathbb{S}^1$ given by $f(\theta) = (\cos\theta, \sin\theta)$ is an open continuous map which is not a homeomorphism.

4.2. Projection.

DEFINITION 34. Let $(X_1, \rho_1), \ldots, (X_k, \rho_k)$ be metric spaces. Let $X = \times_{i=1}^k X_i$, endowed with the product metric. Let $j \in \{1, \ldots, k\}$.

The i^{th} *projection map* is the function

$$p_j : X \to X_j \quad \text{given by } p_j(x_1, \ldots, x_{j-1}, x_j, x_{j+1}, \ldots, x_k) = x_j.$$

PROPOSITION 31. *Let $(X_1, \rho_1), \ldots, (X_k, \rho_k)$ be metric spaces. Let $X = \times_{i=1}^k X_i$, endowed with the product metric ρ. Let $j \in \{1, \ldots, k\}$. Let $a = (a_1, \ldots, a_j, \ldots, a_k) \in X$. Then $p_j(B(a, \delta)) = B(a_j, \delta)$.*

PROOF. We show containment in both directions.
($\subset$) Let $x_j \in p_j(B(a, \delta))$; then $x_j = p_j(x)$ for some $x = (x_1, \ldots, x_{i-1}, x_j, x_{i+1}, \ldots, x_k)$ with $x \in B(a, \delta)$; that is, $\rho(x, a) < \delta$. Now

$$\rho_j(a_j, x_j) \leq \sqrt{\sum_{i=1}^k \rho_i(x_i, a_i)^2} = \rho(a, x) < \delta.$$

So $x_j \in B(a_j, \delta)$.
($\supset$) Let $x_j \in B(a_j, \delta)$, and set $x = (a_1, \ldots, a_{i-1}, x_j, a_{i+1}, \ldots, a_k)$. One sees that

$$\rho(a, x) = \sqrt{\rho_j(a_j, x_j)^2} = \rho_j(a_j, x_j) < \delta.$$

Thus $x \in B(a, \delta)$, so $x_j = p_j(x) \in p_j(B(a, \delta))$. □

PROPOSITION 32. *Let $(X_1, \rho_1), \ldots, (X_k, \rho_k)$ be metric spaces. Let $X = \times_{i=1}^k X_i$, endowed with the product metric ρ. Let $j \in \{1, \ldots, k\}$. Then $p_j : X \to X_j$ is a continuous open map.*

PROOF. Let $a = (a_1, \ldots, a_j, \ldots, a_k) \in X$, so that $p_j(a) = a_j$. Let $\epsilon > 0$, and let $\delta = \epsilon$. Then $p_j(B(a, \delta)) \subset B(a_j, \epsilon) = B(p_j(a), \epsilon)$. Thus p_j is continuous at a.

Now suppose that U is open in X, and let $a_j \in p_j(U)$. Then $a_j = p_j(a)$ for some $a = (a_1, \ldots, a_j, \ldots, a_k) \in U$. Since U is open, there exists $\delta > 0$ such that $B(a, \delta) \subset U$. By the previous proposition, $B(a_j, \delta) = p_j(B(a, \delta)) \subset p_j(U)$. Thus $p_j(U)$ is open. □

4.3. Injection.

Definition 35. Let $(X_1, \rho_1), \ldots, (X_k, \rho_k)$ be metric spaces. Let $X = \times_{i=1}^k X_i$, endowed with the product metric. Let $j \in \{1, \ldots, k\}$, and select a point $a = (a_1, \ldots, a_k) \in X$.

The j^{th} *injection map with respect to* a is the function

$$q_j : X_j \to X \quad \text{given by } q_j(x_j) = (a_1, \ldots, a_{j-1}, x_j, a_{j+1}, \ldots, a_k).$$

Proposition 33. *Let $(X_1, \rho_1), \ldots, (X_k, \rho_k)$ be metric spaces. Let $X = \times_{i=1}^k X_i$, endowed with the product metric. Let $j \in \{1, \ldots, k\}$, and select a point $a = (a_1, \ldots, a_k) \in X$. Let p_j be projection and q_i be injection with respect to a. Then $p_j \circ q_j = \mathrm{id}_{X_j}$.*

Proof. This is clear. □

Proposition 34. *Let $(X_1, \rho_1), \ldots, (X_k, \rho_k)$ be metric spaces. Let $X = \times_{i=1}^k X_i$, endowed with the product metric. Let $i \in \{1, \ldots, k\}$, and select a point $a = (a_1, \ldots, a_k) \in X$. Then $q_i : X_i \to X$ is a continuous function.*

Proof. Let $a = (a_1, \ldots, a_k) \in X$; then $q_j(a_j) = a$. , and $\rho(x, a) = \rho_j(x_j, a_j)$. Let $\epsilon > 0$, and set $\delta = \epsilon$. Let $x_j \in B(a_j, \delta)$, and let $x = \rho_j(x_j)$. Then $\rho(x, a) = \rho_j(x_j, a_j)$, so

$$\rho(x, a) = \rho_j(x_j, a_j) < \delta = \epsilon,$$

which shows that q_j is continuous at a. □

Exercise 11. Show that projection is a contraction.

Exercise 12. Show that injection is an isometric embedding.

Definition 36. Closed map.

Exercise 13. Show that injection is a closed map.

CHAPTER V

Compactness

ABSTRACT. This chapter discusses the concept of compactness. We prove the Heine-Borel Theorem for $\mathbb{R}$, and show that the Heine-Borel property is inherited for finite products of metric spaces.

1. Compactness

DEFINITION 37. Let (X, ρ) be a metric space and let $A \subset X$.

A *cover* of A is a collection of subsets $\mathcal{C} \subset \mathcal{P}(X)$ such that $A \subset \cup\mathcal{C}$.

Let $\mathcal{C}$ be a cover of A. We say that $\mathcal{C}$ is a *finite cover* if A is a finite set. We say that $\mathcal{C}$ is an *open cover* if the elements of $\mathcal{C}$ are open sets. A *subcover* of $\mathcal{C}$ is a subset $\mathcal{D} \subset \mathcal{C}$ such that $\mathcal{D}$ is itself a cover of A.

We say that A is *compact* if every open cover of A has a finite subcover.

REMARK 1. Notice that in the phrase "finite open cover", the word "finite" applies to the cover itself, whereas the word "open" applies to the subsets of X in the cover.

EXAMPLE 38. Let $X = \mathbb{R}$ and $A = \mathbb{Z}$. Let $I_n = (n - \frac{1}{3}, n + \frac{1}{3})$. Let $\mathcal{C} = \{I_n \mid n \in \mathbb{Z}\}$. Then $\mathcal{C}$ is an open cover of $\mathbb{Z}$ with no finite subcover. Thus $\mathbb{Z}$ is not compact.

EXAMPLE 39. Let $X = \mathbb{R}$ and $A = (0, 1)$. Let $I_n = (0, 1 - \frac{1}{n})$. Let $\mathcal{C} = \{I_n \mid n \in \mathbb{N}\}$. Then $\mathcal{C}$ is an open cover of $(0, 1)$ with no finite subcover. Thus $(0, 1)$ is not compact.

PROPOSITION 35. *Let (X, ρ) be a metric space and let $A = \{a_1, \ldots, a_n\} \subset X$ be a finite subset. Then A is compact.*

PROOF. Let $\mathcal{C}$ be an open cover of A. Then for each $a_i \in A$, there exists and open set $U_i \in \mathcal{C}$ such that $a_i \in U_i$. Then $A \subset \cup_{i=1}^{n} U_i$, and $\{U_1, \ldots, U_n\}$ is a finite subcover of $\mathcal{C}$. Thus A is compact. $\square$

Proposition 36. *Let $a, b \in \mathbb{R}$ with $a < b$. Then the closed interval $[a, b] \subset \mathbb{R}$ is compact.*

Proof. Let $\mathcal{C}$ be an open cover of $[a, b]$.

Let $x \in [a, b]$ and let $U_x \in \mathcal{C}$ be an open set which contains x. Then there exists $\epsilon_x > 0$ such that $(x - \epsilon_x, x + \epsilon_x) \subset U_x$. Let

$$B = \{x \in [a, b] \mid [a, x] \text{ can be covered by a finite subcover of } \mathcal{C}\}.$$

Note that B is nonempty, since the closed interval $[a, a + \frac{\epsilon_a}{2}] \subset U_a$, and $\{U_a\}$ is a finite subcover of $\mathcal{C}$, so for example $a + \frac{\epsilon_a}{2} \in B$.

Let $z = \sup B$; clearly $a + \frac{\epsilon_a}{2} \leq z \leq b$. We claim that $z \in B$, and that $z = b$. To see this, let $\epsilon = \min\{\epsilon_z, z - a\}$. Then $z - \frac{\epsilon}{2} \in B$. Let $\mathcal{D}$ be a finite subcover of $\mathcal{C}$ which covers $[a, z - \frac{\epsilon}{2}]$, and let $\mathcal{E} = \mathcal{D} \cup \{U_z\}$. Then $\mathcal{E}$ is finite and covers $[a, z]$, so $z \in B$.

Now suppose that $z < b$, and set $\delta = \min\{\epsilon, z - b\}$. Then $z < z + \frac{\delta}{2} < b$, and $\mathcal{E}$ covers $[a, z + \frac{\delta}{2}]$; since $z + \frac{\delta}{2} \in [a, b]$, this contradicts the definition of z. Thus $z = b$. This completes the proof. $\square$

2. Properties of Compactness

The first proposition says that all compact subsets of a metric space are closed and bounded.

PROPOSITION 37. *Let (X, ρ) be a metric space, and let $K \subset X$ be a compact set. Then K is closed and bounded.*

PROOF. Suppose that K is not bounded, and let $a \in K$. Let

$$\mathcal{C} = \{B(a, n) \mid n \in \mathbb{N}\}.$$

This is a cover of K by open sets (it actually covers all of X). However, since K is unbounded, K is not contained in $B(a, n)$ for any $n \in \mathbb{N}$. Thus $\mathcal{C}$ has no finite subcover, and K is not compact.

Let $\delta > 0$ and let $b \in X$. Let $D(b, \delta) = \{x \in X \mid \rho(x, b) \leq \delta\}$. We claim that this set is closed; to see this, note that if $a \in X \setminus D(b, \delta)$, then $B(a, \rho(a, b) = \delta) \subset X \setminus D(b, \delta)$.

Suppose that K is compact; we wish to show that K is closed. Thus we show that the complement of K is open. Let $b \in X \setminus K$, and set

$$\mathcal{C} = \{X \setminus D(b, \frac{1}{n}) \mid n \in \mathbb{N}\}.$$

Then $\mathcal{C}$ is an open cover of K (in fact, it covers $X \setminus \{b\}$). Thus it has a finite subcover $\mathcal{D}$. Let n be the largest number such that $X \setminus D(b, \frac{1}{n})$ is in $\mathcal{D}$. Then clearly $K \subset X \setminus D(b, \frac{1}{n})$, so $B(b, \frac{1}{n}) \subset X \setminus K$. Thus $X \setminus K$ is open, so K is closed. □

The next proposition says the a closed subset of a compact set is compact.

PROPOSITION 38. *Let (X, ρ) be a metric space, and let $K \subset X$ be a compact set. Let $F \subset K$. If F is closed, then F is compact.*

PROOF. Suppose F is closed. Then $U = X \setminus F$ is open. Let $\mathcal{C}$ be an open cover of F. Then $\mathcal{C} \cup \{U\}$ is an open cover of K, and so it has a finite subcover, say $\mathcal{D}$. Let $\mathcal{E} = \mathcal{D} \setminus \{U\}$. Now $\mathcal{E}$ is a finite subcover of $\mathcal{C}$. □

The next proposition says that the continuous image of a compact set is compact.

PROPOSITION 39. *Let (X, ρ) and (Y, τ) be metric spaces, and let $f : X \to Y$ be a continuous function. If $K \subset X$ is compact, then $f(K)$ is compact.*

PROOF. Let $\mathcal{V}$ be an open cover of $f(K)$, and set

$$\mathcal{U} = \{U \subset X \mid U = f^{-1}(V) \text{ for some } V \in \mathcal{V}\}.$$

Since f is continuous, $\mathcal{U}$ is a collection of open sets which covers K. Thus $\mathcal{U}$ has a finite subcover, say $\{U_1, \ldots, U_n\}$. Now for $i = 1, \ldots, n$, we have $U_i \in \mathcal{U}$, so U_i is the preimage of some set $V_i \in \mathcal{V}$, so that $V_i = f(U_i)$. Then

$$K \subset \cup_{i=1}^n U_i \;\Rightarrow\; f(K) \subset f(\cup_{i=1}^n U_i) = \cup_{i=1}^n f(U_i) = \cup_{i=1}^n V_i.$$

Thus $\{V_1, \ldots, V_n\}$ is a finite subcover of $\mathcal{V}$, and $f(K)$ is compact. □

3. Heine-Borel Theorem

THEOREM 2 (Heine-Borel Theorem for $\mathbb{R}$). *Let $A \subset \mathbb{R}$. Then A is compact if and only if A is closed and bounded.*

PROOF. The forward direction is true in any metric space, as has been stated as Proposition 37. Thus we prove that in $\mathbb{R}$, closed and bounded sets are compact.

Suppose that A is closed and bounded; we wish to show that A is compact. Since A is bounded, there exists $M > 0$ such that $A \subset [-M, M]$. The set $[-M, M]$ is a closed interval, and so it is compact by Proposition 36. Thus A is a closed subset of a compact set, and therefore is compact by Proposition 38. $\square$

PROPOSITION 40. *Let $K \subset \mathbb{R}$ be a compact. Then* $\inf K \in K$ *and* $\sup K \in K$.

PROOF. Since K is bounded, then $\sup K$ exists as a real number, say $b = \sup K$. Suppose $b \notin K$; then $\{(-\infty, b - \frac{1}{n}) \mid n \in \mathbb{N}\}$ is an open cover of K with no finite subcover, contradicting that K is compact. Thus $b \in K$. Similarly, $\inf K \in K$. $\square$

DEFINITION 38. Let (X, ρ) be a metric space. We say that X has the *Heine-Borel property* if every closed and bounded subset of X is compact.

Recall these facts:

(a) Projection is continuous;
(b) Projection is an open map;
(c) Projection is a contraction;
(d) Injection is continuous;
(e) Injection is a closed map;
(f) Injection is an isometric embedding.

PROPOSITION 41. *The finite product of compact sets is compact.*

PROPOSITION 42. *Let $(X_1, \rho_1), \dots, (X_k, \rho_k)$ be a finite collection of metric spaces. Let $X = \times_{i=1}^k X_i$, and let $\rho : X \times X \to \mathbb{R}$ be the product metric on X. Then X has the Heine-Borel property if and only if if and only if each of the spaces X_i has the Heine-Borel property.*

PROOF. Suppose that X has the Heine-Borel property, and let $K_i \subset X_i$ be closed and bounded. Then $\iota_i(K_i)$ is closed and bounded, and so $\iota_i(K_i)$ is compact. Therefore $\pi_i(\iota_i(K_i)) = K$ is compact, so X_i has the Heine-Borel property.

Suppose that X_i has the Heine-Borel property for $i = 1, \dots, k$, and let $K \subset X$ which is closed and bounded.

Then $K_i = \pi_i(K)$ is closed and bounded for each i. Thus K_i is compact, so $\times_{i=1}^k X_i$ is compact by Proposition 41. Since K is a closed subset of $\times_{i=1}^k X_i$, K is compact by Proposition 39. Thus X has the Heine-Borel property. $\square$

CHAPTER VI

Connectedness

Abstract. This chapter discusses the concept of connectedness and its how we may use it to prove the Intermediate Value Theorem.

1. Connectedness

Definition 39. Let (X, ρ) be a metric space and let $A \subset X$. We say that A is *disconnected* if there exist disjoint open sets U_1 and U_2 such that $A \cap U_1 \neq \varnothing$, $A \cap U_2 \neq \varnothing$, and $A \subset U_1 \cup U_2$. We say that A is *connected* if it is not disconnected.

Observation 3. Let $I \subset \mathbb{R}$. Then I is an interval if and only if

- if $x_1, x_2 \in I$ and $x_1 < z < x_2$, then $z \in I$.

Proposition 43. *Let $A \subset \mathbb{R}$. Then A is connected if and only if A is an interval.*

Proof. We prove the contrapositive of each implication.

($\Rightarrow$) Suppose that A is not an interval. Then there exist $x_1, x_2 \in A$ and $z \notin A$ such that $x_1 < z < x_2$. If $U_1 = (-\infty, z)$ and $U_2 = (z, \infty)$, then $x_1 \in U_1$, $x_2 \in U_2$, and $A \subset U_1 \cap U_2$. Thus A is not connected.

($\Leftarrow$) Suppose that A is not connected. Then there exist disjoint open sets $U_1, U_2 \subset \mathbb{R}$ such that $A \cap U_1 \neq \varnothing$, $A \cap U_2 \neq \varnothing$, and $A \subset U_1 \cup U_2$. Let $x_1 \in A \cap U_1$ and $x_2 \in A \cap U_2$. Without loss of generality, assume that $x_1 < x_2$. Let $z = \inf\{u \in U_2 \mid u > x_1\}$; then $x_1 \leq z \leq x_2$.

Suppose $z \in U_1$; then a neighborhood of z is contained in U_1. But since z is the infimum of a subset of U_2, every neighborhood of z intersects U_2. This contradicts that $U_1 \cap U_2 = \varnothing$. Thus $z \notin U_1$. In particular, $x_1 < z$.

Suppose that $z \in U_2$; then a neighborhood of z is contained in U_2, so there exists $y \in U_2$ with $x_1 < y < z$. This contradicts the definition of z. Thus $z \notin U_2$. In particular, $z < x_2$, so $x_1 < z < x_2$.

Since $z \notin U_1 \cup U_2$ but $A \subset U_1 \cup U_2$, $z \notin A$. Thus A is not an interval. $\square$

Proposition 44. *Let (X, ρ) and (Y, τ) be metric spaces, and let $f : X \to Y$ be a continuous function. If $A \subset X$ is connected, then $f(A)$ is connected.*

Proof. Suppose that $f(A)$ is not connected. Then there exist disjoint open subsets V_1 and V_2 of Y such that $f(A) \cap V_1 \neq \varnothing$, $f(A) \cap V_2 \neq \varnothing$, and $f(A) \subset V_1 \cup V_2$.

Let $U_1 = f^{-1}(V_1)$ and $U_2 = f^{-1}(V_2)$. Then $A \cap U_1 \neq \varnothing$, $A \cap U_2 \neq \varnothing$, and $A \subset U_1 \cup U_2$. Thus A is disconnected. $\square$

2. Intermediate Value Theorem

PROPOSITION 45. *Let (X, ρ) be a metric space and let $K \subset X$ be compact and connected. Let $f : X \to \mathbb{R}$ be continuous. Then $f(K)$ is a bounded closed interval.*

PROOF. The image of a compact set is compact, and the compact subsets of $\mathbb{R}$ are closed and bounded.

The image of a connected set is connected, and the connected subsets of $\mathbb{R}$ are intervals.

The result follows. $\square$

THEOREM 3 (Intermediate Value Theorem). *Let $f : [a, b] \to \mathbb{R}$ be a continuous function. If $f(a)f(b) < 0$, then there exists $c \in [a, b]$ such that $f(c) = 0$.*

PROOF. Since f is continuous, the image of $[a, b]$ is a bounded closed interval. Since $f(a)f(b) < 0$, either $f(a) < 0 < f(b)$ or $f(b) < 0 < f(a)$. In either case, 0 is in the image. $\square$

CHAPTER VII

Cardinality

ABSTRACT. We have seen that there are infinitely many rational numbers and infinitely many irrational numbers on the real line. In this chapter we discuss how there are actually more irrational numbers than rational numbers, and attempt to understand exactly how many real numbers there are.

1. Infinite Sets

AXIOM 1 (Axiom of Choice). The cartesian product of a collection of nonempty sets is nonempty.

THEOREM 4 (Zorn's Lemma). *Let A be a partially ordered set. If every chain in A has an upper bound, then A contains a maximal element.*

REMARK. Zorn's Lemma can be proven using the Axiom of Choice, and in fact is logically equivalent to it. It may be used to show many useful set-theoretic propositions, including the following. □

PROPOSITION 46. *Let A and B be sets. Then there exists either a surjective function $A \to B$ or a surjective function $B \to A$.*

PROPOSITION 47. *Let A and B be sets. Then there exists an injective function $A \to B$ if and only if there exists a surjective function $B \to A$.*

THEOREM 5 (Schroeder-Bernstein Theorem). *Let A and B be sets. If there exist injective functions $f : A \to B$ and $g : B \to A$, then there exists a bijective function $h : A \to B$.*

REMARK. We include a proof of this at the end of the chapter. □

DEFINITION 40. Let $\mathbb{N} = \{1, 2, \dots\}$. Let $n \in \mathbb{N}$ and set $\mathbb{N}_n = \{1, \dots, n\}$.

OBSERVATION 4. Let $f : \mathbb{N}_n \to \mathbb{N}_n$. Then f is injective if and only if f is surjective.

OBSERVATION 5. Let $f : \mathbb{N} \to \mathbb{N}_n$ be a function. Then f is not injective.

DEFINITION 41. Let A be a set. We say that A is *finite* there exists a bijective function $\mathbb{N}_n \to A$ for some $n \in \mathbb{N}$. We say that A is *infinite* if there exists an injective function $\mathbb{N} \to A$.

PROPOSITION 48. *Let A be a set. Then A is infinite if and only if it is not finite.*

2. Countable Sets

DEFINITION 42. Let A be a set. We say that A is *countable* if there exists a surjective function $\mathbb{N} \to A$.

PROPOSITION 49. *Every subset of a countable set is countable.*

PROOF. Let A be a countable set and let $B \subset A$. Since A is countable, there exists an injective function $f : A \to \mathbb{N}$. Then $f \restriction_B : B \to \mathbb{N}$ is also injective, so B is countable. $\square$

PROPOSITION 50. *Let A and B be countable sets. Then $A \cup B$ is countable.*

PROOF. Since A and B are countable, there exist surjective functions $g : \mathbb{N} \to A$ and $h : \mathbb{N} \to B$. Define a function

$$f : \mathbb{N} \to A \cup B \quad \text{by} \quad f(n) = \begin{cases} g(\frac{n}{2}) & \text{if } n \text{ is even;} \\ h(\frac{n-1}{2}) & \text{if } n \text{ is odd.} \end{cases}$$

Then f is surjective, so $A \cup B$ is countable. $\square$

PROPOSITION 51. *Let A and B be countable sets. Then $A \times B$ is countable.*

PROOF. Since A and B are countable, there exist injective functions $g : A \to \mathbb{N}$ and $h : B \to \mathbb{N}$. Define a function

$$f : A \times B \to \mathbb{N} \quad \text{by} \quad f(a, b) = 2^{g(a)} \cdot 3^{h(b)}.$$

To see that f is injective, suppose that $f(a_1, b_1) = f(a_2, b_2)$. Then $2^{g(a_1)}3^{h(b_1)} = 2^{g(a_2)}3^{h(b_2)}$. Thus $2^{g(a_1)-g(a_2)} = 3^{h(b_2)-h(b_1)}$, where without loss of generality $g(a_1) \geq g(a_2)$. If $g(a_1) > g(a_2)$, then 2 divides the left side and not the right; this is impossible, so $g(a_1) = g(a_2)$, and since g is injective, we must have $a_1 = a_2$. Similarly, $b_1 = b_2$. $\square$

PROPOSITION 52. *The set $\mathbb{N}$ of natural numbers is countable.*

PROOF. The identity function $\mathrm{id}_{\mathbb{N}} : \mathbb{N} \to \mathbb{N}$ is surjective. $\square$

PROPOSITION 53. *The set $\mathbb{Z}$ of integers is countable.*

PROOF. Define a function

$$f : \mathbb{Z} \to \mathbb{N} \quad \text{by} \quad f(n) = \begin{cases} 1 & \text{if } n = 0; \\ 2n & \text{if } n > 0; \\ 2n + 1 & \text{if } n < 0. \end{cases}$$

Then f is injective, so $\mathbb{Z}$ is countable. $\square$

PROPOSITION 54. *The set $\mathbb{Q}$ of rational numbers is countable.*

PROOF. By Proposition 51, it suffices to find an injective function $\mathbb{Q} \to \mathbb{Z} \times \mathbb{N}$. Every rational number has a unique expression $\frac{p}{q}$ as a ratio of integers, where $\gcd(p, q) = 1$ and $q > 0$. This induces a function $\mathbb{Q} \to \mathbb{Z} \times \mathbb{N}$ given by $\frac{p}{q} \mapsto (p, q)$. This function is bijective; therefore $\mathbb{Q}$ is countable. $\square$

3. Base β Expansions

Let A be a set. A *sequence* in A is a function $a : \mathbb{N} \to A$. We write a_i to mean $a(i)$, and we write $(a_i)_{i=1}^{\infty}$, or simply (a_i), to denote the function a. Let $\mathcal{S}(A)$ denote the set of all sequences in A.

Let β be an integer with $\beta \geq 2$, and let $\mathbb{Z}_\beta = \{0, 1, \dots, \beta - 1\}$. Let $O = (0, 1)$ be the open unit interval in the real line. We are interested in relating the set of sequences in $\mathbb{Z}_\beta$, which we denote by $\mathcal{S}(\mathbb{Z}_\beta)$, to the set O.

Define a function

$$\mu : \mathbb{Z} \to \mathbb{Z}_\beta \quad \text{by} \quad \mu(n) = r,$$

where $n = \beta q + r$ with $q, r \in \mathbb{Z}$ and $0 \leq r < \beta$.

Define a function

$$\zeta : \mathbb{R} \to \mathbb{Z} \quad \text{by} \quad \zeta(x) = \max\{n \in \mathbb{N} \mid n \leq x\}.$$

For each $k \in \mathbb{N}$, define a function

$$\delta_{\beta,k} : \mathbb{R} \to \mathbb{Z}_\beta \quad \text{by} \quad \delta_{\beta,k}(x) = \mu(\zeta(\beta^k x)).$$

This induces a function

$$\delta_\beta : O \to \mathcal{S}(\mathbb{Z}_\beta) \quad \text{by} \quad \delta_\beta(x) = (\delta_{\beta,k}(x))_{k=1}^{\infty}.$$

Then δ_β is an injective function, and we call $\delta_\beta(x)$ the *base β expansion* of x.

Construct a partial inverse to δ_β as follows.

Let $\{a_i\}_{i=1}^{\infty}$ be a sequence in $\mathbb{Z}_\beta$ and set $B = \{\sum_{i=1}^{k} \frac{a_i}{\beta^i} \mid k \in \mathbb{N}\}$. Then $B \subset O$, and in particular, B is a bounded set of real numbers. Set $b = \sup(B)$. For most sequences, $\delta_\beta(b) = (a_i)_{i=1}^{\infty}$.

Call a sequence $(a_i)_{i=1}^{\infty}$ in $\mathbb{Z}_\beta$ a *duplicator* if there exists $N \in \mathbb{N}$ such that $a_i = \beta - 1$ for all $i > N$. These are the only sequences which are not in the image of the function δ_β. If $S = \mathcal{S}(\mathbb{Z}_\beta) \smallsetminus \{\text{ duplicators }\}$, then $\delta_\beta : O \to S$ is bijective.

4. Uncountability

PROPOSITION 55. *The set $\mathbb{R}$ of real numbers is an uncountable set.*

PROOF. Since $O = (0, 1) \subset \mathbb{R}$, it suffices to show that O is uncountable.

Let $\beta = 10$ so that we consider base 10 expansions of the elements in O, and let μ, ζ, and δ_β be as in the previous section.

Let $f : \mathbb{N} \to O$ be any function; we will show that f is not surjective. Define a sequence (a_i) in $\mathbb{Z}_{10}$ by

$$a_i = \begin{cases} 3 & \text{if} \quad \delta_{\beta,i}(f(i)) \neq 3; \\ 6 & \text{if} \quad \delta_{\beta,i}(f(i)) = 3. \end{cases}$$

Define a sequence (s_n) by

$$s_n = \sum_{i=1}^{n} \frac{a_i}{10^i}.$$

Then (s_n) is a bounded increasing sequence; let $b = \lim_{n\to\infty} s_n$. Clearly $\delta_\beta(b) = (a_i)$, and b is not in the image of f. $\square$

5. Measure

In this section, we briefly describe the concepts of measure and probability, and say why the probability of selecting a rational number from the closed unit interval is zero.

Definition 43. Let I be an interval. The *length* of I, denoted $L(I)$, is the distance between its endpoints. Then $L(I)$ is a nonnegative extended real number.

Let $A \subset \mathbb{R}$. An *open interval cover* of A is an open cover of A consisting of open intervals. Let $\mathcal{O}$ be an open interval cover of A. Define the *length* of $\mathcal{O}$, denoted to be the sum of the lengths of the intervals in $\mathcal{O}$:

$$L(\mathcal{O}) = \sum_{I \in \mathcal{O}} L(I).$$

The *outer measure* of A is

$$m(A) = \inf\{L(\mathcal{O}) \mid \mathcal{O} \text{ is an open interval cover of } A\}.$$

We say that A is *measurable* if

$$m(A) = m(A \smallsetminus B) + m(B) \quad \text{for all} \quad B \subset \mathbb{R}.$$

Let $B \subset A$. The *probability* of selected an element of B from A is

$$P(B|A) = \frac{m(B)}{m(A)}.$$

Ponder the following facts:

(a) if I is an interval, then I is measurable, and $m(I) = L(I)$;
(b) if A and B are measurable, then $A \cup B$, $A \cap B$, and $A \smallsetminus B$ are measurable;
(c) if $m(A) = 0$, then A is measurable.

Let $U = [0,1]$; then $m(U) = 1$. Set $Q = U \cap \mathbb{Q}$. We would like to show that $P(Q|U) = 0$. To do this, we only need to show that $m(Q) = 0$.

Proposition 56. *Let $U = [0,1]$ and $Q = U \cap \mathbb{Q}$. Then $m(Q) = 0$.*

Proof. Since $\mathbb{Q}$ is countable, so is Q. Let $q : \mathbb{N} \to Q$ be surjective, and let $q_n = q(n)$; in this way, think of (q_n) as a sequence in U whose image in Q.

Let r be a positive real number. Set $rI_n = (q_n - \frac{r}{2^{n+1}}, q_n + \frac{r}{2^{n+1}})$. Let $r\mathcal{O} = \{rI_n \mid n \in \mathbb{N}\}$. Then

$$L(r\mathcal{O}) = \sum_{n=1}^{\infty} \frac{r}{2^{n+1}} = r.$$

Let $\epsilon > 0$ and let $r = \frac{\epsilon}{2}$. Then $0 < L(r\mathcal{O}) < \epsilon$. Since this is true for every $\epsilon > 0$, we have $m(Q) = 0$. $\square$

6. Cardinality

DEFINITION 44. Let A and B be sets. We say that A and B have the *cardinality* if there exists a bijective function $A \to B$.

If A and B have the same cardinality, we write $|A| = |B|$. If there exists an injective function $A \to B$, we write $|A| \leq |B|$. If there does not exist a surjective function $A \to B$, we write $|A| < |B|$.

DEFINITION 45. Let A be a set. The *power set* of A, denoted $\mathcal{P}(A)$, is the collection of all subsets of A.

PROPOSITION 57. *Let X be a set. Then $|X| < |\mathcal{P}(X)|$.*

PROOF. Let $f : X \to \mathcal{P}(X)$; we wish to show that f is not surjective. Set
$$Y = \{x \in X \mid x \notin f(x)\}.$$
Suppose, by way of contradiction, that $f(x) = Y$ for some $x \in X$. Is $x \in Y$? If it is, then $x \in f(x)$, so by definition of Y, $x \notin Y$. On the other hand, if it is not, then $x \notin f(x)$, so $x \in Y$. Either case is an immediate contradiction. Thus there is no such x satisfying $f(x) = Y$, and Y is not in the image of f. Therefore f is not surjective. $\square$

REMARK 2. This shows that for every set, including infinite sets, there is always a set with a larger cardinality. The cardinality of an infinite set may be regarded as a level of infinity, and in this way, there are infinitely many levels of infinity.

DEFINITION 46. Let A and B be sets. Let $\mathcal{F}(A, B)$ denote the set of all functions from A to B.

PROPOSITION 58. *Let X be any set. Then $|\mathcal{P}(X)| = |\mathcal{F}(X, \mathbb{Z}_2)|$.*

PROOF. Define a function
$$\Phi : \mathcal{F}(X, \mathbb{Z}_2) \to \mathcal{P}(X) \quad \text{by} \quad \Phi(f) = f^{-1}(1).$$
It suffices to show that Φ is bijective.

To see that Φ is injective, suppose that $\Phi(f_1) = \Phi(f_2)$, where $f_1 : X \to T$ and $f_2 : X \to T$. Then $f_1(x) = 1$ if and only if $f_2(x) = 1$. For $x \in X$, $f_i(x)$ is either 1 or 0, so if it is not 1, it is zero. Therefore $f_1(x) = 0$ if and only if $f_2(x) = 0$. So $f_1(x) = f_2(x)$ for every $x \in X$, that is, $f_1 = f_2$.

To see that Φ is surjective, let $A \in \mathcal{P}(X)$. Define a function
$$f : X \to \mathbb{Z}_2 \quad \text{by} \quad f(x) = \begin{cases} 0 & \text{if } x \notin A; \\ 1 & \text{if } x \in A. \end{cases}$$
Then $A = f^{-1}(1)$, so $\Phi(f) = A$. $\square$

7. Intervals

PROPOSITION 59. *Any two intervals have the same cardinality.*

PROOF. We show part of this and leave the remaining details to the reader.

First note that the function $x \mapsto \frac{x-a}{b-a}$ maps (a,b) bijectively onto $(0,1)$. So all intervals of type **(a)** have the same cardinality.

Next consider the function $x \mapsto e^x$, which produces a bijective correspondence between $\mathbb{R}$ and $(0,\infty)$.

Finally consider the function $\arctan : \mathbb{R} \to (-\frac{\pi}{2}, \frac{\pi}{2})$, which is also bijective. This demonstrates how all of the open intervals are equivalent. □

PROPOSITION 60. $|\mathbb{R}| = |\mathcal{F}(\mathbb{N}, \mathbb{Z}_2)|$.

PROOF. Again let $O = (0,1)$. Since $|\mathbb{R}| = |O|$, it suffices to prove that the cardinality of O equals that of $\mathcal{F}(\mathbb{N}, \mathbb{Z}_2)$.

First construct a function

$$f : \mathcal{F}(\mathbb{N}, \mathbb{Z}_2) \to O \quad \text{by} \quad f(a_i) = \sup\Big\{\sum_{i=1}^{k} \frac{a_i}{10^i} \mid k \in \mathbb{N}\Big\}.$$

This function is injective.

Next consider that $\delta_2 : O \to \mathcal{F}(\mathbb{N}, \mathbb{Z}_2)$ is injective.

By the Schoeder-Bernstein theorem, there exists a bijective function $O \to \mathcal{F}(\mathbb{N}, \mathbb{Z}_2)$. □

8. Cardinal Numbers

Let U be a set; we refer to U as a *universal set*, and assume that U contains $\mathbb{R}$.

Let A and B be sets. We say that A and B have the same *cardinality* if there exists a bijective function between them. If A and B have the same cardinality, we write $A \sim B$. Then $\sim$ is a relation on $\mathcal{P}(U)$.

PROPOSITION 61. *The relation $\sim$ is an equivalence relation on $\mathcal{P}(U)$.*

We shall call the equivalence classes of the relation the *cardinal numbers in* U. Let $\beth$ denote the set of cardinal numbers in U. If $A \subset U$, the equivalence class to which it belongs is denoted $|A|$, and is called the *cardinality* of A.

Define a relation $\leq$ on $\beth$ by

$$|A| \leq |B| \Leftrightarrow \exists \text{ injective } f : A \to B;$$

where $A, B \subset U$ are representatives of the cardinal numbers $|A|$ and $|B|$ respectively.

PROPOSITION 62. *The relation $\leq$ on $\beth$ is well defined.*

That is, let $A_1, A_2, B_1, B_2 \subset U$ such that $A_1 \sim A_2$ and $B_1 \sim B_2$, and such that $|A_1| \leq |B_1|$. Show that $|A_2| \leq |B_2|$.

9. Schroeder-Bernstein Theorem

LEMMA 5 (Banach's Lemma). *Let X and Y be sets. and let $f : X \to Y$ and $g : Y \to X$ be injective functions. There exist subsets $A \subset X$ and $B \subset Y$ such that $f(A) = B$ and $g(Y \setminus B) = X \setminus A$.*

PROOF. Fix the following objects:

- Let X and Y be sets.
- Let $f : X \to Y$ and $g : Y \to X$ be injective functions.
- Let $h = g \circ f$.
- Let $C_0 = X \setminus g(Y)$.
- Let $C_n = h(C_{n-1})$, for each $n \in \mathbb{N}$.
- Let $A = \cup_{n=0}^{\infty} C_n$.
- Let $B = f(A)$.

It suffices to show that $g(Y \setminus B) = X \setminus A$.

Claim 1: $h(A) \subset A$.
Let $a_0 \in h(A)$. Then $a_0 = h(a_1)$ for some $a_1 \in A$. By definition of A, $a_1 \in C_n$ for some $n \in \mathbb{N}$. Then $a_0 \in C_{n+1}$. Thus $a_0 \in A$.

Claim 2: $g(Y \setminus B) \subset X \setminus A$.
We want to select an arbitrary $y_0 \in Y \setminus B$ and show that g sends it into $X \setminus A$. Let $x_0 \in g(Y \setminus B)$. Then there exists $y_0 \in Y \setminus B$ such that $g(y_0) = x_0$. Suppose bwoc that $x_0 \in A$. Since $x_0 \in g(Y)$, $x_0 \notin C_0$, so $x_0 \in C_n$ for some $n > 0$. Since $C_n = h(C_{n-1})$, there exists $x_1 \in C_{n-1}$ such that $h(x_1) = x_0$. So $g(f(x_1)) = x_0$. Since g is injective, $f(x_1) = y_0$. But $x_1 \in A$, so $y_0 \in B$. This is a contradiction. Thus $x_0 \notin A$, so $x_0 \in X \setminus A$. Since x_0 was chosen arbitrarily, $g(Y \setminus B) \subset X \setminus A$.

Claim 3: $g(Y \setminus B) \supset X \setminus A$.
We want to select an arbitrary $x_0 \in X \setminus A$ and find $y_0 \in Y \setminus B$ which g sends to it. Let $x_0 \in X \setminus A$. Since $C_0 \subset A$, then $x_0 \in X \setminus C_0$. That is, $x_0 \in g(Y)$, so there exists $y_0 \in Y$ such that $g(y_0) = x_0$. Suppose bwoc that $y_0 \in B$. Then there exists $x_1 \in A$ such that $f(x_1) = y_0$. Thus $h(x_1) = x_0$, so $x_0 \in h(A)$. Since $h(A) \subset A$, $x_0 \in A$, which is a contradiction. Thus $y_0 \notin B$, so $x_0 \in g(Y \setminus B)$. Since x_0 was chosen arbitrarily, $X \setminus A \subset g(Y \setminus B)$. □

THEOREM 6 (The Schroeder-Bernstein Theorem). *Let X and Y be sets. If there exist injective functions $f : X \to Y$ and $g : Y \to X$, then there exists a bijective function $h : X \to Y$.*

PROOF. Let A and B be sets as specified by the lemma. Let $V = X \setminus A$ and $W = Y \setminus B$. Then $f \restriction_A : A \to B$ is bijective, and $g \restriction_W : W \to V$ is bijective. Let $r = (g \restriction_W)^{-1}$. Then $r : V \to W$ is bijective. Thus define $h : X \to Y$ by

$$h(x) = \begin{cases} f(x) & \text{if } x \in A; \\ r(x) & \text{if } x \in V. \end{cases}$$

□

COROLLARY 2. *Let X and Y be sets. If there exist surjective functions $f : X \to Y$ and $g : Y \to X$, then there exists an bijective function $h : X \to Y$.*

PROOF. This follows immediately by combining the Schroeder-Bernstein Theorem with the Axiom of Choice. $\square$

COROLLARY 3. *Let X and Y be sets. The following conditions are equivalent:*

- $|X| = |Y|$;
- $\exists$ *a bijective function* $X \to Y$;
- $\exists$ *injective functions* $X \to Y$ *and* $Y \to X$;
- $\exists$ *surjective functions* $X \to Y$ *and* $Y \to X$.

PROPOSITION 63. *Show that* $(\beth, \leq)$ *is an ordered set.*

PROOF. To show this, one must show that $\leq$ is a total order relation on $\mathcal{P}(U)$. The proof of symmetry uses the Schroeder Bernstein Theorem, and the proof of definiteness requires the Axiom of Choice. $\square$

The total order relation $\leq$ on $\beth$ naturally leads to the following definitions for derived relations on $\beth$:

- $|A| \geq |B| \Leftrightarrow |B| \leq |A|$;
- $|A| < |B| \Leftrightarrow \neg(|A| \geq |B|)$;
- $|A| > |B| \Leftrightarrow \neg(|A| \leq |B|)$.

10. Cardinal Arithmetic

Let A and B be sets. We define the sum, product, and exponentiation of cardinal numbers to match that of finite numbers.

Define

$$|A| + |B| = |(A \times \{0\}) \cup (B \times \{1\})|.$$

Note that even if $A \cap B$ is nonempty, $A \times \{0\}$ and $B \times \{1\}$ are disjoint sets. So if A is any set with m elements and B is any set with n elements, then $(A \times \{0\}) \cup (B \times \{1\})$ is a set with $m + n$ elements.

Define

$$|A| \cdot |B| = |A \times B|.$$

Again, if A and B are finite with m and n elements respectively, then $A \times B$ has mn elements.

Define

$$|A|^{|B|} = |\mathcal{F}(B, A)|,$$

where $\mathcal{F}(B, A)$ denotes the set of all functions from B to A. This again agrees with the finite case.

We have seen that for any set X, there does not exist a surjective function from X to its power set $\mathcal{P}(X)$. Thus $|X| < |\mathcal{P}(X)|$. We have also seen that $|\mathcal{P}(X)| = |\mathcal{F}(X, \mathbb{Z}_2)|$, which can be written as $|\mathcal{P}(X)| = 2^{|X|}$.

Bibliography

[Ga98] Gaughan, Edward D., *Introduction to Analysis*,Brooks/Cole Publishing (1998)

[Ha60] Halmos, Paul R., *Naive Set Theory*, Undergraduate Texts in Mathematics, Springer Verlag (1960,1974)

[KF70] Kolmogorev and Fomin, *Introductory Real Analysis*, Dover (1975)

[Mu75] Munkres, James R., *Topology: A First Course*, Prentice-Hall Inc. (1975)

[Ro80] Ross, Kenneth A., *Elementary Analysis: The Theory of Calculus*, Undergraduate Texts in Mathematics, Springer Verlag (1980)

[SS72] Saxena and Shah, *Introduction to Real Variable Theory*, International Textbook Company 1972