

# Real Analysis

Alec Zabel-Mena

June 28, 2023



# Contents

<b>1</b>	<b>The Real and Complex Numbers</b>	<b>5</b>
1.1	Ordered Sets . . . . .	5
1.2	Fields . . . . .	6
1.3	The Field of Real Numbers . . . . .	8
1.4	The Complex Field . . . . .	9
1.5	Euclidean Spaces . . . . .	10
<b>2</b>	<b>Topological Foundations</b>	<b>13</b>
2.1	Finite, Countable, and Uncountable Sets . . . . .	13
2.2	Metric Spaces . . . . .	14
2.3	Compact Sets . . . . .	17
2.4	Perfect Sets . . . . .	20
2.5	Connected Sets . . . . .	20
<b>3</b>	<b>Sequences</b>	<b>23</b>
3.1	Convergent Sequences . . . . .	23
3.2	Subsequences . . . . .	25
3.3	Cauchy Sequences . . . . .	26
3.4	Upper and Lower Limits. . . . .	27
3.5	Special Sequences . . . . .	28
<b>4</b>	<b>Series.</b>	<b>31</b>
4.1	Convergent Series. . . . .	31
<b>5</b>	<b>Continuity</b>	<b>33</b>
5.1	Limits of Functions. . . . .	33
5.2	Continuous Functions. . . . .	34
5.3	Continuity and Compactness. . . . .	36
5.4	Continuity and Connectedness. . . . .	37
5.5	Discontinuities. . . . .	38
5.6	Monotonic Functions. . . . .	38
5.7	Infinite Limits and Limits at Infinity. . . . .	39

<b>6</b>	<b>Differentiation</b>	<b>41</b>
6.1	The Derivative of Real valued Functions. . . . .	41
6.2	Mean Value Theorems. . . . .	42
6.3	The Continuity of Derivatives. . . . .	43
6.4	L'Hospital's Rule. . . . .	43
6.5	Taylor's Theorem. . . . .	44
6.6	Derivatives of vector valued functions. . . . .	44
<b>7</b>	<b>Integration</b>	<b>47</b>
7.1	The Riemann-Stieltjes Integral. . . . .	47
7.2	Properties of the Integral. . . . .	51
7.3	Integration and Differentiation. . . . .	54
7.4	Integration of Vector Valued Functions. . . . .	55
7.5	Rectifiable Curves. . . . .	56
<b>8</b>	<b>Lebesgue Theory</b>	<b>59</b>
8.1	$\sigma$ -algebras . . . . .	59
8.2	The Lebesgue Outer Measure . . . . .	60
8.3	Lebesgue Measurable Sets . . . . .	63
8.4	Outer and Inner Approximations . . . . .	66
8.5	The Borel-Cantelli Lemma . . . . .	68
8.6	Nonmeasurable Sets . . . . .	71
8.7	The Cantor Set . . . . .	72
<b>9</b>	<b>Lebesgue Measurable Functions</b>	<b>75</b>
9.1	Properties of Lebesgue Measurable Functions . . . . .	75
9.2	Simple Approximation . . . . .	78
9.3	Theorems of Littlewood, Egoroff, and Lusin . . . . .	80

# Chapter 1

## The Real and Complex Numbers

### 1.1 Ordered Sets

**Definition.** Let  $S$  be any set. An **order** on  $S$  is a relation  $<$  such that:

[label=(0)]

1. For  $x, y \in S$ , one and only one of the following hold:

$$x < y$$

$$x = y$$

$$y < x$$

We call this property the **trichotomy law**

2.  $<$  is transitive over  $S$ .

We denote the relations  $>$  and  $\leq$  to mean  $x > y$  if and only if  $y < x$ , and  $x \leq y$  if and only if  $x < y$ , or  $x = y$ . We call  $S$  together with  $<$  an **ordered set**.

**Example 1.1.** Define  $<$  on  $\mathbb{Q}$  such that for  $r, s \in \mathbb{Q}$ ,  $r < s$  implies  $< 0s - r$ .

**Definition.** Let  $S$  be an ordered set, and let  $E \subseteq S$ . We say that  $E$  is **bounded above** if there is some  $\beta \in S$  for which  $x \leq \beta$ , for all  $x \in E$ . We say that  $E$  is **bounded below** if  $\beta \leq x$ , for all  $x \in E$ . We say an  $\alpha \in S$  is a **least upperbound** of  $E$ , if  $\alpha$  is an upperbound of  $E$ , and for all other upperbounds,  $\gamma$ , of  $E$ ,  $\alpha \leq \gamma$ . Likewise,  $\alpha$  is a **greatest lowerbound** of  $E$  if  $\alpha$  is a lowerbound of  $E$ , and for all other lowerbounds  $\gamma$  of  $E$ ,  $\gamma \leq \alpha$ . We denote the least upperbound, and greatest lowerbound by  $\sup E$  and  $\inf E$ , respectively.

**Lemma 1.1.1.** *Let  $S$  be an ordered set, and let  $E \subseteq S$ . Then  $E$  has (if they exist) a unique least upperbound, and a unique greatest lowerbound.*

*Proof.* Let  $\alpha, \beta \in S$  be least upperbounds of  $E$ . Then by definition, we have that  $\alpha \leq \beta$ , and  $\beta \leq \alpha$ ; thus by the trichotomy law,  $\alpha = \beta$ . The proof is the same for greatest lowerbounds. ■

**Example 1.2.** [label=(0)]

1. Let  $A = \{p \in \mathbb{Q} : p^2 < 2\}$ , and  $B = \{p \in \mathbb{Q} : p^2 > 2\}$ . Clearly, we have that every element of  $B$  is an upperbound of  $A$ , and every element of  $A$  is a lowerbound of  $B$ . Now take  $p \in \mathbb{Q}$  a positive rational, and take  $q \in \mathbb{Q}$  such that  $q = p - \frac{p^2-2}{p+2}$ . Then  $q^2 - 2 = \frac{2(p^2-2)}{(p+2)^2}$ . Now if  $p \in A$ , then  $p^2 - 2 < 0$ , which implies that  $p < q$ , and  $q^2 < 2$ ; thus  $A$  has no largest element; similarly, if  $p \in B$ , then  $p^2 - 2 > 0$ , which implies that  $q < p$  and  $q^2 > 2$ , which shows that  $B$  has no least element. Thus  $\sup A$  and  $\inf B$  do not exist in  $\mathbb{Q}$ .
2. If  $\alpha = \sup E \in S$ , it may or may not be that  $\alpha \in E$ . Take  $E_1 = \{r \in \mathbb{Q} : r < 0\}$ , and  $E_2 = \{r \in \mathbb{Q} : r \leq 0\}$ . Then  $\sup E_1 = \sup E_2 = 0$ , but  $0 \notin E_1$ , where as  $0 \in E_2$ .
3. Consider the set  $\frac{1}{\mathbb{Z}^+} = \{\frac{1}{n} : n \in \mathbb{Z}^+\}$ . By the well ordering principle, 1 is the least element, and is also an upperbound of all  $\frac{1}{n}$  for  $n > 1$ . Now also notice that as  $n$  gets arbitrarily large, then  $\frac{1}{n}$  gets arbitrarily small; that is to say  $\frac{1}{n}$  “tends” to 0, so  $\sup \frac{1}{\mathbb{Z}^+} = 1 \in \frac{1}{\mathbb{Z}^+}$ , and  $\inf \frac{1}{\mathbb{Z}^+} = 0 \notin \frac{1}{\mathbb{Z}^+}$ .

**Definition.** We say an ordered set  $S$  has the **least upperbound property**, if whenever  $E \subseteq S$ , nonempty, and bounded above, then  $\sup E \in S$  exists; likewise,  $S$  has the **greatest lowerbound property** if whenever  $E$  is nonempty, bounded below then  $\inf E \in S$  exists.

**Example 1.3.** [label=(0)]

1. The set of all rationals  $\mathbb{Q}$  does not have the least upperbound property, nor the greatest lowerbound property, take  $A, B$  as in the previous example. Letting  $E = \{1, \frac{1}{2}, \frac{1}{4}\} \subseteq \frac{1}{\mathbb{Z}^+}$ , we see that  $\frac{1}{\mathbb{Z}^+}$  satisfies both properties, with  $\sup E = 1$ , and  $\inf E = \frac{1}{4}$ .
2. Let  $A \subseteq \mathbb{R}$  be nonempty, and be bounded below. Then by the greatest lowerbound property,  $\alpha = \inf A \in \mathbb{R}$  exists; Then for all  $x \in A$ ,  $\alpha \leq x$ , and for all other lowerbounds  $\gamma$ ,  $\gamma \leq \alpha$ . Then  $-x \leq -\alpha$ , and  $-\alpha \leq -\gamma$ , then we see that  $-\gamma$  and  $-\alpha$  are upperbounds of  $-A$ , and that  $-\alpha$  is the least upperbound of  $-A$ .

**Theorem 1.1.2.** *If  $S$  is an ordered set with the least upperbound property, then  $S$  also inherits the greatest lowerbound property.*

*Proof.* Let  $B \subseteq S$ , and let  $L \subseteq S$  be the set of all lowerbounds of  $B$ . Then we have for any  $y \in L$ ,  $x \in B$ ,  $y \leq x$ . So every element of  $B$  is an upperbound of  $L$ , and  $L$  is nonempty, hence  $\alpha = \sup L \in S$  exists. Now if  $\gamma \leq \alpha$ , then  $\gamma$  is not an upperbound of  $L$ , hence  $\gamma \notin B$ ; thus  $\alpha \leq x$  for all  $x \in B$ , so  $\alpha \in L$ , and by definition of the greatest lowerbound, we get  $\alpha = \inf B$ . ■

## 1.2 Fields

**Definition.** A **field** is a set  $F$ , together with binary operations  $+$  and  $\cdot$  (called **addition** and **multiplication**, respectively) such that:

[label=(0)]

1.  $F$  forms an abelian group under  $+$ .
2.  $F \setminus \{0\}$  forms an abelian group under  $\cdot$  (where  $0$  is the additive identity of  $F$ ).
3.  $\cdot$  distributes over  $+$ .

We now state the following propositions without proof.

**Proposition 1.2.1.** *For all  $x, y, z \in F$ :*

$$[label=(0)] x + y = x + z \text{ implies } y = z \quad x + y = x \text{ implies } y = 0 \quad x + y = 0 \text{ implies } y = -x \quad -(-x) = x.$$

**Proposition 1.2.2.** *For all  $x, y, z \in F \setminus \{0\}$ :*

$$[label=(0)] xy = xz \text{ implies } y = z \quad xy = x \text{ implies } y = 1 \quad xy = 1 \text{ implies } y = x^{-1} \quad (x^{-1})^{-1} = x.$$

**Proposition 1.2.3.** *For all  $x, y, z \in F$ :*

$$[label=(0)] 0x = 0 \quad x \neq 0 \text{ and } y \neq 0 \text{ implies } xy \neq 0 \quad (-x)y = -(xy) = x(-y) \quad (-x)(-y) = xy.$$

**Definition.** An **ordered field** is a field  $F$  that is also an ordered set, such that:

$$[label=(0)] x + y < x + z \text{ whenever } y < z, \text{ for } x, y, z \in F \quad xy > 0 \text{ whenever } x > 0 \text{ and } y > 0, \text{ for } x, y \in F.$$

**Proposition 1.2.4.** *Let  $F$  be an ordered field, then for any  $x, y, z \in F$ , the following hold:*

$$[label=(0)] x > 0 \text{ implies } -x < 0. \text{ If } x > 0 \text{ and } y < z, \text{ then } xy < xz. \text{ If } x < 0 \text{ and } y < z, \text{ then } xz < xy. \text{ If } x \neq 0, \text{ then } x^2 > 0, \text{ in particular, } 1 > 0. \text{ } 0 < x < y \text{ implies that } 0 < y^{-1} < x^{-1}.$$

**4.** *Proof.*  $[label=(0)]$

$$\text{If } x > 0, \text{ then } 0 = x + (-x) > 0 + (-x), \text{ so } -x < 0.$$

2. We have  $0 < z - y$ , so  $0 < x(z - y) = xz - xy$ , so  $xy < xz$ .
3. Do the same as (2), multiplying  $z - y$  by  $-x$ .
4. If  $x > 0$ , we are done. Now suppose that  $x < 0$ , then  $-x > 0$ , so  $(-x)(-x) = xx = x^2 > 0$ ; in particular, we also have that  $1 \neq 0$ , and  $1 = 1^2$ , so  $1 > 0$ .
5. We have  $0 < xy^{-1} < yy^{-1} = 1$ , then  $0 < x^{-1}xy^{-1} = y^{-1} < x^{-1}1 = x^{-1}$

■

### 1.3 The Field of Real Numbers

**Theorem 1.3.1.** *There exists an ordered field  $\mathbb{R}$  with the least upperbound property, such that  $\mathbb{Q} \subseteq \mathbb{R}$ .*

**Definition.** We call the field  $\mathbb{R}$  the **field of real numbers**, and we call the elements of  $\mathbb{R}$  **real numbers**.

**Definition.** Let  $S$  be an ordered field, and let  $E \subseteq S$ . We say that  $E$  is **dense** in  $S$ , if for all  $r, s \in S$ , with  $r < s$ , there is an  $\alpha \in E$  such that  $r < \alpha < s$ .

**Theorem 1.3.2** (The Archimedean Principle). *If  $x, y \in \mathbb{R}$ , and  $x > 0$ , then there is an  $n \in \mathbb{Z}^+$  such that  $nx > y$ .*

*Proof.* Let  $A = \{nx : n \in \mathbb{Z}^+\}$ , and suppose that  $nx \leq y$ . Then  $y$  is an upperbound of  $A$ , and since  $A$  is nonempty,  $\alpha = \sup A \in \mathbb{R}$ , since  $x > 0$ , we have  $\alpha - x < \alpha$ , so  $\alpha - x$  is not an upperbound of  $A$ . Hence  $\alpha - x < mx$  for some  $m \in \mathbb{Z}^+$ . Then  $\alpha < (1 + m)x \in A$ , contradicting that  $\alpha$  is an upperbound of  $A$ . ■

**Theorem 1.3.3** (The density of  $\mathbb{Q}$  in  $\mathbb{R}$ ).  *$\mathbb{Q}$  is dense in  $\mathbb{R}$ .*

*Proof.* Let  $x < y$  be real numbers, then  $y - x > 0$ , so by the Archimedean principle, there is an  $n \in \mathbb{Z}^+$  for which  $n(y - x) > 1$ . By the Archimedean principle again, we have  $m_1, m_2 \in \mathbb{Z}^+$  for which  $m_1 > nx$  and  $m_2 > -nx$ , thus  $-m_2 < nx < m_1$ , and we also have that there is an  $m \in \mathbb{Z}^+$  for which  $-m_2 < m < m_1$ , and  $m - 1 \leq nx < m$ . Thus combining inequalities, we get  $nx < m < ny$ , thus  $x < \frac{m}{n} < y$ . ■

**Theorem 1.3.4** (The existence of  $n^{\text{th}}$  roots of positive reals). *For every real number  $X > 0$ , and for every  $n \in \mathbb{Z}^+$ , there is one, and only one positive real number  $y$  for which  $y^n = x$ .*

*Proof.* Let  $y > 0$  be a real number; then  $y^n > 0$ , so there is at most one such  $y$  for which  $y^n = x$ . Now let  $E = \{t \in \mathbb{R} : t^n < x\}$ , choosing  $t = \frac{x}{1+x}$ , we see that  $0 \leq t < 1$ , hence  $t^n < t < x$ , so  $E$  is nonempty. Now if  $1 + x < t$ , then  $t^n \geq x$ , so  $t \notin E$ , and  $E$  has  $1 + x$  as an upperbound. Therefore,  $\alpha = \sup E \in \mathbb{R}$  exists.

Now suppose that  $y^n < x$ , choose  $0 \leq h < 1$  such that  $h < \frac{x - y^n}{n(y+1)^{n-1}}$ , then  $(y+h)^n - y^n < hn(y+h)^{n-1} < hn(y+1)^{n-1} < x - y^n$ , thus  $(y+h)^n < x$ , so  $y+h \in E$ , contradicting that  $y$  is an upperbound. On the other hand, if  $y^n > x$ , choosing  $k = \frac{y^n - x}{ny^{n-1}}$ , then  $0 \leq k < y$ , and letting  $t \geq y - k$ , we get that  $y^n - t^n \leq y^n + (y - k)^n < kny_{n-1} = y^n - x^n$ , so  $t^n \geq x$ , making  $y - k$  an upperbound of  $E$ , which contradicts  $y = \sup E$ . ■

*Remark.* We denote  $y$  as  $\sqrt[n]{x}$ , or as  $x^{\frac{1}{n}}$ .

**Corollary.** *If  $a, b \in \mathbb{R}$ , with  $a, b > 0$ , and  $n \in \mathbb{Z}^+$ , then  $\sqrt[n]{ab} = \sqrt[n]{a}\sqrt[n]{b}$ .*

*Proof.* Let  $\alpha = \sqrt[n]{a}$ , and  $\beta = \sqrt[n]{b}$ . Then  $\alpha^n = a$ , and  $\beta^n = b$ , so  $ab = \alpha^n \beta^n = (\alpha\beta)^n$ , we are done. ■

**Definition.** We define the **extended real number system** to be the field  $\mathbb{R}$ , together with symbols  $\infty$ , and  $-\infty$ , called **positive infinity** and **negative infinity**, such that  $-\infty < x < \infty$  for all  $x \in \mathbb{R}$ . We call elements of the extended real numbers **infinite**, and every other element not in the extended real numbers **finite**.



**Lemma 1.3.5.**  $\infty$  is an upperbound for every subset  $E$ , of  $\mathbb{R}$ , and  $-\infty$  is a lowerbound for every subset  $E$  of  $\mathbb{R}$ . Moreover, if  $E$  is not bounded above, then  $\sup E = \infty$ , and if  $E$  is not bounded below, then  $\inf E = -\infty$ .

*Remark.* We make the following assumptions for extended real numbers:

[label=(0)]If  $x \in \mathbb{R}$ , then  $x + \infty = \infty$ ,  $x - \infty = -\infty$ , and  $\frac{x}{\infty} = \frac{x}{-\infty} = 0$ . If  $x > 0$ , then  $x(\infty) = \infty$  and  $x(-\infty) = -\infty$ . If  $x < 0$ , then  $x(\infty) = -\infty$  and  $x(-\infty) = \infty$ .

## 1.4 The Complex Field

**Definition.** We define a **complex number** to be a pair of real numbers  $(a, b)$ . We denote the set of all complex numbers by  $\mathbb{C}$ . We define the **addition** and **multiplication** of complex numbers to be the binary operations  $+: \mathbb{C} \rightarrow \mathbb{C}$  and  $\cdot: \mathbb{C} \rightarrow \mathbb{C}$  such that

$$\begin{aligned}(a, b) + (c, d) &= (a + c, b + d) \\ (a, b)(c, d) &= (ac - bd, ad + bc)\end{aligned}$$

Lastly, we define  $i$  to be the complex number such that  $i = (0, 1)$ .

**Theorem 1.4.1.**  $\mathbb{C}$  forms a field together with  $+$  and  $\cdot$ .

**Theorem 1.4.2.** For  $(a, 0), (b, 0) \in \mathbb{C}$ ,  $(a, 0) + (b, 0) = (a + b, 0)$ , and  $(a, 0)(b, 0) = (ab, 0)$ .

**3. Proof.** This is a straightforward application of the addition and multiplication of complex numbers. ■

**Theorem 1.4.3.**  $i^2 = -1$ .

*Proof.*  $i^2 = (0, 1)(0, 1) = (0 - 1, 1 - 1) = (-1, 0) = -1$ . ■

**Theorem 1.4.4.** Let  $(a, b) \in \mathbb{C}$ , then  $(a + b) = a + ib$ .

*Proof.*  $(a, b) = (a, 0) + (0, b) = (a, 0) + (0, 1)(b, 0) = a + ib$ . ■

**Definition.** Let  $a, b \in \mathbb{R}$ , and let  $z \in \mathbb{C}$  such that  $z = a + ib$ . We define the **complex conjugate** of  $z$  to be the complex number  $\bar{z} = a - ib$ . Moreover, we define the **real part** of  $z$  to be  $a$ , and the **imaginary part** of  $z$  to be  $b$ , and we denote them  $a = \operatorname{Re} z$ ,  $b = \operatorname{Im} z$

**Theorem 1.4.5.** Let  $z, w \in \mathbb{C}$ . Then

[label=(0)] $\overline{z + w} = \bar{z} + \bar{w}$ .  $\overline{zw} = \bar{z}\bar{w}$ .  $z + \bar{z} = 2\operatorname{Re} z$  and  $z - \bar{z} = 2i\operatorname{Im} z$ .  $z\bar{z}$  is a nonegative real number.

**4. Proof.** Let  $z = a + ib$ , and let  $w = c + id$ . Then  $z + w = (a + c) + i(b + d)$ , so  $\overline{z + w} = (a + b) - i(b + d) = (a - ib) + (c - id) = \bar{z} + \bar{w}$ ; similarly, we get  $\overline{zw} = \bar{z}\bar{w}$ . Moreover, we have  $(a + ib) + (a - ib) = 2a$ , and  $(a + ib) - (a - ib) = 2ib$ , we also have that  $z\bar{z} = (a + ib)(a - ib) = a^2 + b^2 \geq 0$ , and  $z\bar{z} = 0$  if and only if  $a = b = 0$ . ■

**Definition.** Let  $z \in \mathbb{C}$ . We define the **modulus** of  $z$  to be  $|z| = \sqrt{z\bar{z}}$ .

*Remark.*  $|z|$  exists and is unique.

**Theorem 1.4.6.** Let  $z, w \in \mathbb{C}$ , then:

$$|z| \geq 0 \text{ and } |z| = 0 \text{ if and only if } z = 0. \quad |\bar{z}| = |z|. \quad |zw| = |z||w|. \quad \operatorname{Re} z \leq |z|. \\ |z + w| \leq |z| + |w|.$$

*Proof.* Let  $z = a + ib$ , and  $w = c + id$ . Then  $|z| = \sqrt{a^2 + b^2} \geq 0$ , and  $|z| = 0$  if and only if  $a, b = 0$ . Moreover,  $|\bar{z}| = |a + i(-b)| = \sqrt{a^2 + (-b)^2} = \sqrt{a^2 + b^2} = |z|$ . We also have  $|zw|^2 = (a^2 + b^2)(c^2 + d^2) = |z|^2|w|^2$ , likewise,  $|\operatorname{Re} z| = |a + i0| = \sqrt{a^2} \leq \sqrt{a^2 + b^2}$ . Finally we prove (5).

We have  $|z + w|^2 = (z + w)(\bar{z} + \bar{w}) = z\bar{z} + \bar{z}w + \bar{w}z + w\bar{w} = |z|^2 + w\operatorname{Re} z\bar{w} + |w|^2 \leq |z|^2 + 2|s\bar{w}| + |w|^2 = (|z| + |w|)^2$ . ■

**Theorem 1.4.7** (The Cauchy Schwarz Inequality). Let  $a_i, b_i \in \mathbb{C}$ , for  $1 \leq i \leq n$ . Then:

$$\left| \sum_{i=1}^n a_i \bar{b}_i \right| \leq \sum_{i=1}^n |a_i|^2 \sum_{i=1}^n |b_i|^2 \quad (1.1)$$

*Proof.* Let  $A = \sum |a_j|^2$ ,  $B = \sum |b_i|^2$ , and  $C = \sum a_i \bar{b}_i$ . If  $B = 0$ , then  $b_i = 0$  for  $1 \leq i \leq n$ , and we are done; so suppose that  $B > 0$ . Then

$$\begin{aligned} \sum |Ba_j - Cb_j|^2 &= \sum (Ba_j - Cb_j)(B\bar{a}_j - \bar{C}\bar{b}_j) \\ &= B \sum |a_j|^2 - B\bar{C} \sum a_j \bar{b}_j - BC \sum \bar{a}_j b_j + |C|^2 \sum |b_j|^2 \\ &= (B^2 A - B|C|^2) = B(AB - |C|^2) \geq 0 \end{aligned}$$

Since  $B > 0$ , we get  $|C|^2 \leq AB$  as required. ■

## 1.5 Euclidean Spaces

**Definition.** Let  $k \in \mathbb{Z}^+$ , and let  $\mathbb{R}^k$  be the set of all ordered  $k$ -tuples  $(x_1, x_2, \dots, x_k)$ , with  $x_i \in \mathbb{R}$  for  $1 \leq i \leq k$ . We call  $\mathbb{R}^k$  the **Euclidean space** of **dimension**  $k$ ; more simply the **Euclidean  $k$ -space**. We call elements of  $\mathbb{R}^k$  **vectors** or **points**; and we define **vector addition** and **scalar multiplication** to be:

$$\begin{aligned} (x_1, \dots, x_k) + (y_1, \dots, y_k) &= (x_1 + y_1, \dots, x_k + y_k) \\ \alpha(x_1, \dots, x_k) &= (\alpha x_1, \dots, \alpha x_k) \end{aligned}$$

for  $(x_1, \dots, x_k), (y_1, \dots, y_k) \in \mathbb{R}^k$  and  $\alpha \in \mathbb{R}$ .

**Theorem 1.5.1.**  $\mathbb{R}^k$  forms a vector space together with vector addition and scalar multiplication.

**Definition.** Let  $x, y \in \mathbb{R}^k$ . We define the **inner product** of  $x$  and  $y$  to be the binary operation  $\langle, \rangle : \mathbb{R}^k \times \mathbb{R}^k \rightarrow \mathbb{R}$  such that

$$\langle x, y \rangle = \sum_{i=1}^k x_i y_i$$

We define the **norm** of  $x$  to be  $\|x\| = \sqrt{\langle x, x \rangle} = \sqrt{\sum x_i^2}$ .

**Theorem 1.5.2.** Let  $x, y \in \mathbb{R}^k$ , and  $\alpha \in \mathbb{R}$ . Then:

$$\begin{aligned} & \|x\| \geq 0 \text{ and } \|x\| = 0 \text{ if and only if } x_i = 0 \text{ for all } 1 \leq i \leq k. \quad \|\alpha x\| = |\alpha| \|x\|. \\ & \|\langle x, y \rangle\| \leq \|x\| \|y\|. \quad \|x + y\| \leq \|x\| + \|y\|, \text{ and } \|x - z\| \leq \|x - y\| + \|y - z\| \end{aligned}$$

**Proof.** (1) follows by definition of the norm. We also have that  $\|\alpha x\| = \sqrt{\sum \alpha^2 x_i^2} = \sqrt{\alpha^2} \sqrt{\sum x_i^2} = |\alpha| \|x\|$ .

Now by the Cauchy Schwarz inequality, we have that  $\|\langle x, y \rangle\|^2 = \sum x_i^2 y_i^2 \leq \sum x_i^2 \sum y_i^2 = \|x\| \|y\|$ . Finally we have that  $\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle \leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$ , the last result follows immediately. ■



# Chapter 2

## Topological Foundations

### 2.1 Finite, Countable, and Uncountable Sets

**Definition.** Let  $A$  be a set, and let  $E \subseteq \mathbb{N}$ . We say that  $A$  is **finite** if there exists a 1-1 mapping of  $A$  onto  $E$ , we say  $A$  is **countable** if  $E = \mathbb{N}$ , and we say  $A$  is **atmost countable** if  $A$  is either finite or countable.

**Example 2.1.** The set of all integers  $\mathbb{Z}$  is countable. Take  $f : \mathbb{N} \rightarrow \mathbb{Z}$  such that  $f(n) = 2$  if  $n$  is even, and  $f(n) = -n$  if  $n$  is odd.

**Definition.** Let  $A$  be a set, and let  $E \subseteq \mathbb{N}$ . A **sequence** in  $A$  is a mapping  $f : E \rightarrow A$  such that  $f(n) = x_n$ , for  $x_n \in A$ . We call the values of  $f$  **terms** of the sequence. We denote sequences by  $\{x_n\}_{n=1}^{\infty}$ , and when  $E = \mathbb{N}$ , we denote them simply by  $\{x_n\}$ .

**Theorem 2.1.1.** *Every infinite subset of a countable set is countable.*

*Proof.* Let  $A$  be countable, and let  $E \subseteq A$  be infinite. Arrange the elements of  $A$  into a sequence  $\{x_n\}$ , and construct a sequence  $\{n_k\}$  such that  $n_1$  is the least term for which  $\{x_{n_k}\} \in E$ , and  $n_k$  is the least term greater than  $n_{k-1}$  for which  $x_{n_k} \in E$ . Let  $f(k) = x_{n_k}$ , and we get a 1-1 mapping of  $\mathbb{N}$  onto  $E$ . ■

**Theorem 2.1.2.** *Let  $\{E_n\}$  be a sequence of countable sets. Then  $S = \bigcup E_n$  is also countable.*

*Proof.* Arrange every set  $E_n$  in a sequence  $\{x_{nk}\}$ , and consider the infinite array  $(x_{ij})$ , in which the elements of  $E_n$  form the  $n$ -th row. Then  $(x_{ij})$  contains all the elements of  $S$ , and we can arrange them in a sequence

$$x_{11}, (x_{21}, x_{12}), (x_{31}, x_{22}, x_{13}), \dots$$

Moreover, if  $E_i \cap E_j \neq \emptyset$ , for  $i \neq j$ , then the elements of  $E_i \cap E_j$  appear more than once in the sequence of  $S$ ; so taking  $T \subseteq \mathbb{N}$ , we get a 1-1 mapping of  $T$  onto  $S$ , hence  $S$  is atmost countable, and since  $E_i \subseteq S$  for  $i \in \mathbb{N}$ , is infinite, by theorem 2.1.1,  $S$  is infinite, thus  $S$  is countable. ■

▣Figures/diagonalizationArray.png

Figure 2.1: The infinite array  $(x_{ij})$

**Corollary.** *Let  $A$  be atmost countable, and suppose for all  $\alpha \in A$  that the sets  $B_\alpha$  are atmost countable. Then*

$$T = \bigcup_{\alpha \in A} B_\alpha$$

*is atmost countable.*

**Theorem 2.1.3.** *Let  $A$  be countable, and let  $B_n$  be the set of all  $n$ -tuples  $(a_1, \dots, a_n)$  such that  $a_i \in A$  for  $1 \leq i \leq n$ . Then  $B_n$  is countable.*

*Proof.* By induction on  $n$ , we have that  $B_1 = A$ , which is countable. Now suppose that  $B_n$  is countable, and consider  $B_{n+1}$  whose elements are of the form  $(b, a)$  where  $b \in B_n$  and  $a \in A$ . Fixing  $b$ , we get a 1-1 correspondence between the elements of  $B_{n+1}$  and  $A$ ; therefore  $B$  is countable. ■

**Corollary.**  $\mathbb{Q}$  is countable.

*Proof.* For every rational  $\frac{p}{q} \in \mathbb{Q}$ , represent  $\frac{p}{q}$  as  $(p, q)$ . Then the countability of  $\mathbb{Q}$  follows from theorem 2.1.3. ■

**Theorem 2.1.4.** *Let  $A$  be the set of all sequences of 0 and 1; then  $A$  is uncountable.*

*Proof.* Let  $EA$  be countable, and let  $E$  consist of all the sequences of 0 and 1,  $s_1, s_2, s_3, \dots$ . Construct the sequence  $s$  such that if the  $n$ -th term of the sequence  $s_i$  is 0, then the  $n$ -th term of  $s$  is 1, and vice versa, for  $i \in \mathbb{Z}^+$ . Then the sequence  $s$  differs from the sequence  $s_i$  at atleast one place; thus  $s \notin E$ , but  $s \in A$ . Therefore  $E \subset A$ , which establishes the uncountability of  $A$ . ■

## 2.2 Metric Spaces

**Definition.** A set  $X$ , whose elements we will call **points**, is said to be a **metric space** if there exists a mapping  $d : X \times X \rightarrow \mathbb{R}$ , called a **metric** (or **distance function**) such that for  $x, y \in X$

[label=(0)]

1.  $d(x, y) \geq 0$ , and  $d(x, y) = 0$  if and only if  $x = y$ .
2.  $d(x, y) = d(y, x)$ .
3.  $d(x, y) \leq d(x, z) + d(z, y)$  (The Triangle Inequality).

**Example 2.2.** The absolute value,  $|\cdot|$  for real numbers, the modulus  $|\cdot|$  for complex numbers, and the norm  $\|\cdot\|$  for vectors are all metrics. They turn  $\mathbb{R}$ ,  $\mathbb{C}$ , and  $\mathbb{R}^k$  into metric spaces respectively.

**Definition.** An **open interval** in  $\mathbb{R}$  (or **segment**) is a set of the form  $(a, b) = \{a, b \in \mathbb{R} : a < x < b\}$ , a **closed interval** in  $\mathbb{R}$  is a set of the form  $[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$ ; and **half open intervals** in  $\mathbb{R}$  are sets of the form  $[a, b) = \{x \in \mathbb{R} : a \leq x < b\}$  and  $(a, b] = \{x \in \mathbb{R} : a < x \leq b\}$ .

If  $a_i < b_i$ , for  $1 \leq i \leq k$ , the set of all points  $(x_1, \dots, x_k) \in \mathbb{R}^k$  which satisfy the Inequalities  $a_i \leq x_i \leq b_i$  is called a **k-cell** in  $\mathbb{R}^k$ . If  $x \in \mathbb{R}^k$ , and  $r > 0$ , we call the set  $B_r(x) = \{y \in \mathbb{R}^k : \|x - y\| < r\}$  an **open ball** in  $\mathbb{R}^k$ , and we call the set  $B_r[x] = \{y \in \mathbb{R}^k : \|x - y\| \leq r\}$  a **closed ball** in  $\mathbb{R}^k$ .

**Definition.** We call a set  $E \subseteq \mathbb{R}^k$  **convex**, if whenever  $x, y \in E$ ,  $\lambda x + (1 - \lambda)y \in E$  for  $0 < \lambda < 1$ .

**Lemma 2.2.1.** *Open and closed balls, along with k-cells are convex.*

*Proof.* Let  $B_r(x)$  be an open ball; let  $y, x \in B_r(x)$ , and  $0 < \lambda < 1$ . Then  $\|x - (\lambda y + (1 - \lambda)z)\| = \|\lambda(x - y) - (1 - \lambda)(x - z)\| \leq \lambda\|x - y\| + (1 - \lambda)\|x - z\| < \lambda r + (1 - \lambda)r$ . The proof is analogous for closed ball.

Now let  $K$  be a  $k$ -cell for  $a_i < b_i$ , for  $1 \leq i \leq k$ , let  $x, y \in K$ , then  $a_i \leq x_i, y_i \leq b_i$ , so  $\lambda a_i \leq \lambda x_i \leq \lambda b_i$ , and  $(1 - \lambda)a_i \leq (1 - \lambda)y_i \leq (1 - \lambda)b_i$ , since  $0 < \lambda < 1$ ,  $a_i \leq \lambda a_i + (1 - \lambda)a_i \leq \lambda x_i + (1 - \lambda)y_i \leq \lambda b_i + (1 - \lambda)b_i \leq b_i$ . ■

**Corollary.** *Open and closed intervals, along with half open intervals are convex.*

*Proof.* We just notice that open and closed intervals are open and closed balls in  $\mathbb{R}^1 = \mathbb{R}$ , we also notice that half open intervals  $[a, b)$  and  $(a, b]$  are subsets of the closed interval  $[a, b]$ , and hence inherit convexity. ■

For the following definitions, let  $X$  be a metric space with metric  $d$ .

**Definition.** A **neighborhood** of a point  $x \in X$  is the set  $N_r(x) = \{y \in X : d(x, y) < r\}$  for some  $r > 0$  called the **radius** of the neighborhood. We call  $x$  a **limit point** of a set  $E \subseteq X$  if every neighborhood of  $x$  contains a point  $y \neq x$  such that  $y \in E$ . If  $y \in E$ , and  $y$  is not a limit point, we call  $y$  an **isolated point**.

**Definition.** We call a set  $E \subseteq X$  **closed** if every limit point of  $E$  is in  $E$ . A point  $x \in X$  is an **interior point** of  $E$  if there is a neighborhood  $N$  of  $x$  such that  $N \subseteq E$ . We call  $E$  **open** if every point of  $E$  is an interior point of  $E$ .

**Definition.**  $E \subseteq X$  is called **perfect** if  $E$  is closed, and every point of  $E$  is a limit point of  $E$ . We call  $E$  **dense** if every point of  $X$  is either a limit point of  $E$ , or a point of  $E$ , or both.

**Lemma 2.2.2.** *If  $E \subseteq X$ , then  $E$  is perfect in  $X$  if and only if  $\overline{E} = E$ .*

**Lemma 2.2.3.** *If  $E \subseteq X$  is dense in  $X$ , then either  $E$  is perfect in  $X$ , or  $X = E$ , or both.*

**Definition.** We call  $E \subseteq X$  **bounded** if there is a real number  $M > 0$ , and a point  $y \in X$  such that  $d(x, y) < M$  for all  $x \in E$ .

**Theorem 2.2.4.** *Let  $X$  be a metric space and  $x \in X$ . Every neighborhood of  $x$  is open.*

*Proof.* Consider the neighborhood  $N_r(x)$ , and  $y \in E$ , there is a positive real number  $h$  such that  $d(x, y) = r - h$ , then for  $z \in X$  such that  $d(y, z) < h$ , we have  $d(x, z) \leq d(x, y) + d(y, z) < r - h + h = r$ , thus  $z \in E$ , so  $y$  is an interior point of  $E$ . ■

**Theorem 2.2.5.** *If  $x$  is a limit point of a set  $E$ , then every neighborhood of  $x$  contains infinitely many points of  $E$ .*

*Proof.* Let  $N$  be a neighborhood of  $x$  containing only a finite number of points of  $E$ . Let  $y_1, \dots, y_n$  be points of  $N \cap E$  distinct from  $x$  and let  $r = \min\{d(x, y_i)\}$  for  $1 \leq i \leq n$ , then  $r > 0$ , and the neighborhood  $N_r(x)$  contains no point  $y$  of  $E$  for which  $y \neq x$ , so  $x$  is not a limit point; which is a contradiction. ■

**Corollary.** *A finite point set has no limit points.*

*Proof.* By theorem 2.2.5, if  $x$  is a limit point in the finite point set  $E$ , then every neighborhood of  $x$  contains infinitely many points of  $E$ ; contradicting its finiteness. ■

**Example 2.3.** [label=(0)]

1. The set of all  $z \in \mathbb{C}$  such that  $|z| < 1$  is open, and bounded.
2. The set of all  $z \in \mathbb{C}$  for which  $|z| \leq 1$  is closed, perfect, and bounded.
3. Any nonempty finite set is closed, and bounded.
4.  $\mathbb{Z}$  is closed, but it is not open, perfect, or bounded.
5. The set  $\frac{1}{\mathbb{Z}^+}$  is neither closed, nor open, it is not perfect; but it is bounded.
6.  $\mathbb{C}$  is closed, open, and perfect, but it is not bounded.
7. The open interval in  $(a, b)$  is open (only in  $\mathbb{R}$ ), and bounded.

**Theorem 2.2.6.** *Let  $X$  be a metric space, a set  $E \subseteq X$  is open if and only if  $X \setminus E$  is closed.*

*Proof.* Suppose that  $X \setminus E$  is closed, let  $x \in E$ , then  $x \notin X \setminus E$ , and  $x$  is not a limit point of  $X \setminus E$ . Thus there is a neighborhood  $N$  of  $x$  such that  $N \cap (X \setminus E) = \emptyset$ , thus  $N \subseteq E$ , and so  $x$  is an interior point of  $E$ .

Conversely, suppose that  $E$  is open, and let  $x$  be a limit point of  $X \setminus E$ , then every neighborhood of  $x$  contains a point of  $X \setminus E$ , so  $x$  is not an interior point of  $E$ , since  $E$  is open, it follows that  $x \in X \setminus E$ , thus  $X \setminus E$  is closed. ■

**Corollary.**  *$E$  is closed if and only if  $X \setminus E$  is open.*

*Proof.* This is the converse of theorem 2.2.5. ■

**Theorem 2.2.7.** *Let  $X$  be a metric space. The following are true:*

[label=(0)]If  $\{G_\alpha\}$  is a collection of open sets, then  $\bigcup G_\alpha$  is open. If  $\{G_i\}_{i=1}^n$  is a finite collection of open sets, then  $\bigcap_{i=1}^n G_i$  is open. If  $\{G_\alpha\}$  is a collection of closed sets, then  $\bigcap G_\alpha$  is closed. If  $\{G_i\}_{i=1}^n$  is a finite collection of closed sets, then  $\bigcup_{i=1}^n G_i$  is closed.



**2.** *Proof.* Let  $G = \bigcup G_\alpha$ , then if  $x \in G$ ,  $x \in G_\alpha$  for some  $\alpha$ , then  $x$  is an interior point of  $G_\alpha$ , hence an interior point of  $G$ , so  $G$  is open. Now let  $G = \bigcap_{i=1}^n G_i$ . For  $x \in G$ , there are neighborhoods  $N_i$  of  $x$ , with radii  $r_i$  such that  $N_i \subseteq G_i$  for  $1 \leq i \leq n$ . Then let  $r = \min\{r_1, \dots, r_n\}$ , and let  $N$  be the neighborhood of  $x$  with radius  $r$ , then  $N \subseteq G_i$ , hence  $N \subseteq G$ , so  $G$  is open.

The proofs of (3) and (4) are just the converse of the proofs of (1) and (2). ■

**Definition.** Let  $X$  be a metric space, and let  $E \subseteq X$ , and let  $E'$  be the set of all limit points of  $E$ . We define the **closure** of  $E$  to be the set  $\overline{E} = E \cup E'$ .

**Theorem 2.2.8.** *If  $X$  is a metric space, and  $E \subseteq X$ , then the following hold*

*[label=(0)] $\overline{x}$  is closed.  $E$  is closed if and only if  $E = \overline{E}$ . If  $F \subseteq X$  such that  $E \subseteq F$ , and  $F$  is closed, then  $\overline{E} \subseteq F$ .*

**2.** *Proof.* If  $x \in X$ , and  $x \notin \overline{E}$ , then  $x \notin E$ , nor is it a limit point of  $E$ , thus there is a neighborhood of  $x$  that is disjoint from  $E$ , hence  $X \setminus \overline{E}$  is open.

Now if  $E$  is closed, then  $E' \subseteq E$ , so  $\overline{E} = E$ , conversely, if  $E = \overline{E}$ , then clearly  $E$  is closed. Now if  $F$  is closed and  $E \subseteq F$ , then  $F' \subseteq F$ , and  $E' \subseteq F$ , therefore  $\overline{E} \subseteq F$ . ■

**Theorem 2.2.9.** *Let  $E \subseteq \mathbb{R}$  be nonempty and bounded above, let  $y = \sup E$ , then  $y \in \overline{E}$ , hence  $y \in E$  if  $E$  is closed.*

*Proof.* Suppose that  $y \notin E$ , then for every  $h > 0$ , there exists a point  $x \in E$  such that  $y - h < x < y$ , then  $y$  is a limit point of  $E$ , thus  $y \in \overline{E}$ . ■

**Theorem 2.2.10.** *Let  $Y \subseteq X$ ; a subset  $E$  of  $Y$  is open in  $Y$  if and only if  $E = Y \cap G$  for some open subset  $G$  of  $X$ .*

*Proof.* Suppose  $E$  is open in  $Y$ , then for each  $x \in E$ , there is a  $r_p > 0$  such that  $d(x, y) < r_p$ , if  $y \in Y$ , that implies that  $y \in E$ ; hence let  $V_x$  be the set of all  $y \in X$  such that  $d(x, y) < r_p$ , and define

$$G = \bigcup_{x \in E} V_x$$

Then by theorems 2.2.2 and 2.2.6,  $G$  is open in  $X$ , and  $E \subseteq G \cap Y$ . Now we also have that  $V_x \cap Y \subseteq E$ , thus  $G \cap Y \subseteq E$ , thus  $E = G \cap Y$ . Conversely, if  $G$  is open in  $X$ , and  $E = G \cap Y$ , then every  $x \in E$  has a neighborhood  $V_x \subseteq G$ , thus  $V_x \cap Y \subseteq E$ , hence  $E$  is open in  $Y$ . ■

## 2.3 Compact Sets

**Definition.** Let  $X$  be a metric space, and let  $E \subseteq X$ . An **open cover** of  $E$  is a collection  $\{G_\alpha\}$  of subsets of  $X$  such that  $E \subseteq \bigcup G_\alpha$ . We call a collection  $\{E_\beta\}$  of subsets of  $X$  an **open subcover** of  $E$  if  $\{E_\beta\}$  is a cover of  $E$ , and  $\bigcup E_\beta \subseteq \bigcup G_\alpha$ . We call  $E$  **compact** if every open cover of  $E$  contains a finite open subcover.

**Lemma 2.3.1.** *Every finite set is compact.*

*Proof.* Let  $K$  be finite, and let  $\{G_\alpha\}$  be an open subcover of  $K$ . Since  $K$  is finite, there is a 1-1 mapping of  $K$  onto the set  $\{1, \dots, n\}$ . Let  $\{E_i\}_{i=1}^n$  be the finite collection of all subsets of  $K$ , clearly,  $\{E_i\}$  is an open cover of  $K$ . Moreover, if  $\bigcup E_i \subseteq \bigcup G_\alpha$ , we are done, and if  $\bigcup G_\alpha \subseteq \bigcup E_i$ , then  $\{G_i\}$  is a finite subcollection that covers  $K$ , so in either case,  $K$  is compact. ■

**Theorem 2.3.2.** *Let  $X$  be a metric space, and let  $K \subseteq Y \subseteq X$ . Then  $Y$  is compact in  $X$  if and only if  $K$  is compact in  $Y$ .*

*Proof.* Suppose  $K$  is compact in  $Y$ , and let  $\{G_\alpha\}$  be a collection of subsets of  $Y \setminus X$  that cover  $K$ , and let  $V_\alpha = Y \cap G_\alpha$ , then  $\{V_\alpha\}$  is a collection of subsets of  $X$  covering  $K$ , in which  $V_\alpha \subseteq G_\alpha$  for all  $\alpha$ , therefore  $K$  is compact in  $Y$ .

conversely, suppose that  $K$  is compact in  $X$ , and let  $\{V_\alpha\}$  be a collection of open sets in  $Y$  such that  $K \subseteq \bigcup V_\alpha$ , by theorem 2.2.10, there is a collection  $\{G_\alpha\}$  of open sets in  $Y$  such that  $V_\alpha = Y \cap G_\alpha$ , for all  $\alpha$ . Then  $K \subseteq \bigcup_{i=1}^n G_{\alpha_i}$ ; therefore,  $K$  is compact in  $Y$ . ■

**Theorem 2.3.3.** *Compact subsets of metric spaces are closed.*

*Proof.* Let  $X$  be a metric space, and let  $K$  be compact in  $X$  and let  $x \in X \setminus K$ , if  $y \in K$ , let  $U$  and  $V$  be neighborhoods of  $x$  and  $y$  respectively, each of radius  $r < \frac{1}{2}d(x, y)$ . Since  $K$  is compact, there are finitely many points  $y_1, \dots, y_n$  such that  $K \subseteq \bigcup_{i=1}^n V_i = V$ , where  $V_i$  is a neighborhood of  $y_i$  for  $1 \leq i \leq n$ . Let  $U = \bigcap_{i=1}^n U_i$ , then  $V \cap U$  is empty, hence  $x \notin \overline{K}$ , therefore,  $x \in X \setminus \overline{K}$ , therefore  $K$  is closed. ■

**Theorem 2.3.4.** *Closed subsets of compact sets are compact.*

*Proof.* Let  $X$  be a metric space with  $F \subseteq K \subseteq X$ , with  $F$  closed in  $X$ , and  $K$  compact. Let  $\{V_\alpha\}$  be an open cover of  $F$ . If we append  $X \setminus F$  to  $\{V_\alpha\}$ , we get an open cover  $\Theta$  of  $K$ , and since  $K$  is compact, there is a finite subcollection  $\Phi$  which covers  $K$ , so  $\Phi$  is an open cover of  $F$ ,  $X \setminus F \in \Phi$ , then  $\Phi \setminus (X \setminus F)$  still covers  $F$ , therefore  $F$  is compact. ■

**Theorem 2.3.5.** *Let  $\{K_\alpha\}$  be a collection of compact sets of a metric space  $X$ , such that every finite subcollection of  $\{K_\alpha\}$  is nonempty. Then  $\bigcap K_\alpha$  is nonempty.*

*Proof.* Fix  $K_1 \subseteq \{K_\alpha\}$ , and let  $G_\alpha = X \setminus K_\alpha$ . Suppose no point of  $K_1$  is in  $\bigcap K_\alpha$ , then  $\{G_\alpha\}$  covers  $K_1$ , and since  $K_1$  is compact, we have  $K_1 \subseteq \bigcup_{i=1}^n G_{\alpha_i}$ , for  $1 \leq i \leq n$ , which implies that  $\bigcap K_\alpha$  is empty, a contradiction. ■

**Corollary.** *If  $\{K_\alpha\}$  is a sequence of nonempty compact sets, such that  $K_{n+1} \subseteq K_n$ , then  $\bigcap_{i=1}^\infty K_n$  is nonempty.*

**Theorem 2.3.6.** *If  $E$  is an infinite subset of a compact set  $K$ , then  $E$  has a limit point in  $K$ .*

*Proof.* Suppose no point of  $K$  is a limit point of  $E$ , then for all  $x \in K$ , the neighborhood  $U_x$  contains at most one point in  $E$ . Then no finite subcollection of  $\{U_x\}$  covers  $E$ , which contradicts the compactness on  $K$ . ■

**Theorem 2.3.7** (The Nested Interval Theorem). *if  $\{I_n\}$  is a sequence of intervals in  $\mathbb{R}$  such that  $I_{n+1} \subseteq I_n$ , then  $\bigcap_{j=1}^\infty I_n$  is nonempty.*

*Proof.* We let  $I_n = [a_n, b_n]$ . Letting  $E$  be the set of all  $a_n$ ,  $E$  is nonempty and bounded above by  $b_1$ . Letting  $x = \sup E$ , and  $m \geq n$ , we have  $[a_m, b_m] \subseteq [a_n, b_n]$ , thus  $a_m \leq x \leq b_m$  for all  $m$ , thus  $x \in I_m = \bigcap_{j=i}^n I_j$  ■

**Theorem 2.3.8.** *Let  $k \in \mathbb{Z}^+$ , and  $\{I_n\}$  be a nonempty sequence of  $k$ -cells of  $\mathbb{R}^k$  such that  $I_{n+1} \subseteq I_n$ . Then  $\bigcap_{j=1}^{\infty} I_n$  is nonempty.*

*Proof.* Let  $I_n$  be the set of all points  $x \in \mathbb{R}^k$  such that  $a_{n,j} \leq x_j \leq b_{n,j}$ , and let  $I_{n,j} = [a_{n,j}, b_{n,j}]$ . Then for each  $1 \leq j \leq k$ , by the nested interval theorem,  $\bigcap_{l=1}^{\infty} I_{l,j}$  is nonempty, hence there are real numbers  $x'_j$  such that  $a_{n,j} \leq x'_j \leq b_{n,j}$ . Letting  $x' = (x'_1, \dots, x'_k)$ , we get that  $x' \in I \bigcap_{l=1}^{\infty} I_l$  ■

**Theorem 2.3.9.** *Every  $k$ -cell is compact.*

*Proof.* Let  $I$  be a  $k$ -cell, and let  $\delta = \sqrt{\sum_{j=1}^k a(b_j - a_j)^2}$  we get for  $x, y \in I$ ,  $\|x - y\| \leq \delta$ . Now suppose there is an open cover  $\{G_\alpha\}$  of  $I$  for which no finite subcover is contained. Let  $c_j = \frac{a_j + b_j}{2}$ , then the closed intervals  $[a_j, c_j]$ ,  $[c_j, b_j]$  determine the  $2^k$   $k$ -cells  $Q_i$  such that  $\bigcup Q_i = I$ . Then atleast one  $Q_i$  cannot be covered by any finite subcollection of  $\{G_\alpha\}$ . Subdividing  $Q_1$ , we get a sequence  $\{Q_n\}$  such that  $Q_{n+1} \subseteq Q_n$ ,  $Q_n$  is not covered by any finite subcollection of  $\{G_\alpha\}$ , and  $\|x - y\| \leq \frac{\delta}{2^n}$  for  $x, y \in Q_n$ . Then by theorem 2.3.8, there is a point  $x' \in Q_n$ , and for some  $\alpha$ ,  $x' \in G_\alpha$ ; since  $G_\alpha$  is open, there is an  $r > 0$  for which  $\|x - x'\| < r$  implies  $y \in G_\alpha$ . Then for  $n$  sufficiently large, we have that  $\frac{\delta}{2^n} < r$ , then we get that  $Q_n \subseteq G_\alpha$ , which is a contradiction. ■

**Theorem 2.3.10** (The Heine-Borel Theorem). *If  $E$  is a subset of  $\mathbb{R}^k$ , then the following are equivalent:*

*[label=(0)] $E$  is closed and bounded.  $E$  is compact. Every infinite subset of  $E$  has a limit point in  $E$ .*

**2. Proof.** Suppose that  $E$  is closed and bounded, then  $E \subseteq I$  for some  $k$ -cell  $I$  in  $\mathbb{R}^k$ , and hence it is compact. By theorem 2.3.4,  $E$  is compact. Now suppose that  $E$  is compact, then by theorem 2.3.6, every infinite subset of  $E$  has a limit point in  $E$ .

Now suppose that every infinite subset of  $E$  has a limit point in  $E$ . If  $E$  is not bounded, then  $\|x_n\| > n$  for some  $x_n \in E$  and  $n \in \mathbb{Z}^+$ . Then the set of all such  $x_n$  is infinite, and has no limit point in  $E$ , a contradiction; moreover suppose that  $E$  is not closed. Then there is a point  $x_0 \in \mathbb{R}^k \setminus E$ , which is a limit point of  $E$ . Then there are points  $x_n \in E$  for which  $\|x_n - x_0\| < \frac{1}{n}$ , let  $S$  be the set of all such points. Then  $S$  is infinite and has  $x_0$  as its only limit point; for if  $y \neq x_0 \in \mathbb{R}^k$ , then  $\frac{1}{2}\|x_0 - y\| \leq \|x_0 - y\| - \frac{1}{n} \leq \|x_0 - y\| - \|x_n - x_0\| \leq \|x_n - y\|$  for only some  $n$ . Thus by theorem 2.2.3,  $y$  is not a limit point of  $S$ . Therefore, if every infinite subset of  $E$  has a limit point in  $E$ ,  $E$  must be closed. ■

**Theorem 2.3.11** (The Bolzano-Weierstrass Theorem). *Every bounded infinite subset  $E$  of  $\mathbb{R}^k$  has a limit point in  $\mathbb{R}^k$ .*

*Proof.* We have that  $E \subseteq I$ , for some  $k$ -cell  $I$  in  $\mathbb{R}^k$ . Since  $k$ -cells are compact, by the Heine-Borel theorem,  $E$  is also compact and has a limit point in  $I$ . ■

## 2.4 Perfect Sets

**Theorem 2.4.1.** *If  $P \subseteq \mathbb{R}^k$  is a nonempty perfect set, then  $P$  is uncountable.*

*Proof.* Since every point of  $P$  is a limit point of  $P$ , we have that  $P$  must be infinite. Then suppose that  $P$  is countable. For points  $x_n \in P$ , construct the sequence  $\{U_n\}$  of neighborhoods of  $x_n$ , for  $n \in \mathbb{Z}^+$ ; now by induction, if  $U_1$  is a neighborhood of  $x_1$ , then for  $y \in \hat{U}_1$ ,  $\|x_1 - y\| \leq r$  for some  $r > 0$ . Now suppose the neighborhood  $U_n$  of  $x_n$  has been constructed such that  $U_n \cap P$  is nonempty. Then there is a neighborhood  $U_{n+1}$  of  $x_{n+1}$  such that  $\hat{U}_{n+1} \subseteq U_n$ ,  $x_n \notin \hat{U}_{n+1}$ , and  $U_{n+1} \cap P$  is nonempty. Therefore there is a nonempty  $K_n = U_n \cap P$ . Since  $\hat{U}_n$  is closed and bounded,  $\hat{U}$  is compact, and since  $x_n \notin K_{n+1}$ ,  $x_n \notin \bigcap_{i=1}^{\infty} K_i$ , and since  $K_n \subseteq P$ ,  $\bigcap K_i$  is empty, a contradiction. ■

**Corollary.** *Let  $a < b$  be real numbers. Then the closed interval  $[a, b]$  is uncountable. Moreover,  $\mathbb{R}$  is uncountable.*

*Proof.* We have  $[a, b]$  is closed, and perfect (since  $(a, b)[a, b]$  is perfect), thus  $[a, b]$  is uncountable. Moreover, take  $f : \mathbb{R} \rightarrow [a, b]$ , by  $f(x) = \frac{a+b}{2}x$ ; then  $f$  is a 1-1 mapping of  $\mathbb{R}$  onto  $[a, b]$ , which makes  $\mathbb{R}$  uncountable. ■

**Theorem 2.4.2** (The construction of the Cantor set). *There exists a perfect set in  $\mathbb{R}$  which contains no open interval.*

*Proof.* Let  $E_0 = [0, 1]$ , and remove  $(\frac{1}{3}, \frac{2}{3})$ , and let  $E_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$ . Now remove the open intervals  $(\frac{1}{9}, \frac{2}{9})$ ,  $(\frac{3}{9}, \frac{6}{9})$ ,  $(\frac{7}{9}, \frac{8}{9})$ , and let  $E_2 = [0, \frac{2}{9}] \cup [\frac{2}{9}, \frac{3}{9}] \cup [\frac{6}{9}, \frac{7}{9}] \cup [\frac{7}{9}, \frac{8}{9}]$ . Continuing the removal of the middle third of each interval, we obtain the sequence of compact sets  $\{E_n\}$ , such that  $E_{n+1} \subseteq E_n$ , and  $E_n$  is the union of  $2^n$  closed intervals of length  $\frac{1}{3^n}$ . Then let:

$$P = \bigcap_{i=1}^{\infty} E_i \quad (2.1)$$

Then  $P$  is nonempty, and compact.

Now let  $I$  be the open interval of the form  $(\frac{3k+1}{3^m}, \frac{3k+2}{3^m})$ , with  $k, m \in \mathbb{Z}^+$ . Then by the construction of  $P$ ,  $I$  has no point in  $P$ , we also see that every other open interval contains a subinterval of the form of  $I$ ; then  $P$  contains no open interval.

Now let  $x \in P$ , and let  $S$  be any open interval for which  $x \in S$ . Let  $I_n$  be the closed interval of  $E_n$  such that  $x \in I_n$ . Choose  $n$  sufficiently large such that  $I_n \subseteq S$ . If  $x_n \neq x$  is an endpoint of  $I_n$ , then  $x_n \in P$ , and so  $x$  is a limit point of  $P$ . Therefore  $P$  is perfect. ■

**Definition.** We call the set  $P$  constructed in the proof of theorem 2.4.2 the **Cantor set**.

## 2.5 Connected Sets

**Definition.** Two subsets  $A$  and  $B$  of a metric space  $X$  are **separated** if  $A \cap \hat{B}$  and  $\hat{A} \cap B$  are both empty. We say a subset  $E$  of  $X$  is **connected**, if  $E$  is not the union of two nonempty separated sets.

**Theorem 2.5.1.** *A subset  $E$  of  $\mathbb{R}$  is connected if and only if  $x, y \in E$  and  $x < z < y$  imply  $z \in E$ .*

*Proof.* Let  $x, y \in E$  such that for some  $z \in (x, y)$ ,  $z \notin E$ . Then  $E = A \cup B$ , with  $A = E \cup (-\infty, z)$  and  $B = E \cup (z, \infty)$ . Then  $A$  and  $B$  are separated, which contradicts the connectedness of  $E$ .

Conversely suppose for  $x, y \in E$ , that  $z \in E$  for  $z \in (x, y)$ . Then there are nonempty separated sets  $A$  and  $B$  such that  $A \cup B = E$ . Choose  $x \in A$ ,  $y \in B$  such that  $x < y$ , and let  $z = \sup(A \cap [x, y])$ . Then by theorem 2.2.8,  $z \in \hat{A}$ , so  $z \notin B$ . In particular,  $x \leq z < y$ . Now if  $z \notin A$ , then  $x < z < y$ , with  $z \notin E$ . Now if  $z \in A$ , then  $z \notin \hat{B}$ , hence there is a  $z'$  such that  $z < z' < y$ , and  $z' \notin B$ . Then  $x < z' < y$  and  $z' \notin E$ . ■



# Chapter 3

## Sequences

### 3.1 Convergent Sequences

**Definition.** A sequence  $\{x_n\}$  in a metric space  $X$  is said to **converge** if there is a point  $x \in X$  such that for every  $\varepsilon > 0$ , there is an  $N \in \mathbb{Z}^+$  such that  $d(x_n, x) < \varepsilon$  whenever  $n \geq N$ . We say  $\{x_n\}$  **converges** to  $x$ , and we call  $x$  the **limit** of  $\{x_n\}$  as  $n$  approaches  $\infty$ . We write  $x_n \rightarrow x$  as  $n \rightarrow \infty$ , and  $\lim_{n \rightarrow \infty} x_n = x$  (or  $\lim x_n = x$ ). If  $\{x_n\}$  does not converge, we say the  $\{x_n\}$  **diverges**, or **is divergent**.

**Example 3.1.** Consider the following sequences in  $\mathbb{C}$ .

$\{\frac{1}{n}\}$  is bounded, and  $\lim_{n \rightarrow \infty} \frac{1}{n} = 0$ . The sequence  $\{n^2\}$  is unbounded and diverges.  $1 + \frac{(-1)^n}{n} \rightarrow 1$  as  $n \rightarrow \infty$ , and  $\{1 + \frac{(-1)^n}{n}\}$  is bounded.  $\{i^n\}$  is bounded and divergent.  $\{1\}$  is bounded and converges to 1.

**Theorem 3.1.1.** Let  $\{x_n\}$  be a sequence in a metric space, then:

$\{x_n\}$  converges to  $x \in X$  if and only if every neighborhood of  $x$  contains  $x_n$  for all but finitely many  $n$ . If  $\{x_n\}$  converges to  $x$ , and  $x'$ , then  $x = x'$ . If  $\{x_n\}$  converges, then  $x_n$  is bounded. If  $E \subseteq X$ , and  $x$  is a limit point of  $E$ , then there is a sequence in  $E$  that converges to  $x$ .

**Proof.** Suppose  $x_n \rightarrow x$ , and let  $U$  be a neighborhood of  $x$ . For some  $\varepsilon > 0$ , there is an  $N \in \mathbb{Z}^+$  for which  $d(x_n, x) < \varepsilon$ , whenever  $n \geq N$ , thus  $x_n \in U$  for finitely many  $n$ . Conversely, suppose that  $x_n \in U$  for some  $n \geq N$ , then letting  $\varepsilon > 0$ , we have  $d(x, x_n) < \varepsilon$ , hence  $x_n \rightarrow x$ .

Let  $\varepsilon > 0$ , then there are  $N_1, N_2 \in \mathbb{Z}^+$  such that  $d(x_n, x) < \frac{\varepsilon}{2}$ , and  $d(x_n, x') < \frac{\varepsilon}{2}$ . Then choosing  $N = \max\{N_1, N_2\}$ , and letting  $\varepsilon$  be arbitrarily small, we have  $d(x, x') \leq d(x, x_n) + d(x_n, x') < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ ; and so we get that  $x = x'$ .

Let  $x_n \rightarrow x$ , then there is an  $N \in \mathbb{Z}^+$  for which  $d(x_n, x) < 1$  whenever  $n \geq N$ . Letting  $r = \max\{1, d(x_N, x)\}$ , then  $d(x_n, x) \leq r$ .

Finally, let  $x$  be a limit point of  $E$ , then for each  $n \in \mathbb{Z}^+$ , there is an  $x_n \in E$  such that  $d(x, x_n) < \frac{1}{n}$ , choose  $N > \frac{1}{\varepsilon}$ , then whenever  $n \geq N$ ,  $d(x, x_n) < \varepsilon$ ; hence  $x_n \rightarrow x$ . ■

**Theorem 3.1.2.** Suppose  $\{x_n\}$  and  $\{y_n\}$  are sequences in  $\mathbb{C}$ , and that  $\lim x_n = x$ ,  $\lim y_n = y$  as  $n \rightarrow \infty$ . Then the following hold as  $n \rightarrow \infty$ :

$\lim (x_n + y_n) = \lim x_n + \lim y_n = x + y$ .  $\lim x_n y_n = \lim x_n \lim y_n = xy$ .  
 $\lim \frac{x_n}{y_n} = \frac{\lim x_n}{\lim y_n} = \frac{x}{y}$ ; given that  $y_n, y \neq 0$ .

**2. Proof.** [label=(0)]

Let  $\varepsilon > 0$ , then for  $N_1, N_2 \in \mathbb{Z}^+$ ,  $|x_n - x| < \frac{\varepsilon}{2}$  and  $|y_n - y| < \frac{\varepsilon}{2}$ . Then choose  $N = \max\{N_1, N_2\}$ , then whenever  $n \geq N$ , we have  $|(x_n + y_n) - (x + y)| \leq |x_n - x| + |y_n - y| < \varepsilon$ .

2. Notice that  $x_n y_n - xy = (x_n - x)(y_n - y) + x(y_n - y) + y(x_n - x)$ , then for  $N_1, N_2 \in \mathbb{Z}^+$ ,  $|x_n - x| < \sqrt{\varepsilon}$ , and  $|y_n - y| < \sqrt{\varepsilon}$ . Then choosing  $N = \max\{N_1, N_2\}$ , then  $|(x_n - x)(y_n - y)| < \varepsilon$ , thus we have  $|x_n y_n - xy| \leq |(x_n - x)(y_n - y)| + |x(y_n - y)| + |y(x_n - x)| < \varepsilon$ .
3. We first show that  $\frac{1}{y_n} \rightarrow \frac{1}{y}$ , given that  $y_n, y \neq 0$ . Choose  $m$  such that  $|y_n - y| < \frac{1}{2}|y|$  whenever  $n \geq m$ , then  $|y_n| > \frac{1}{2}|y|$ . Then for  $\varepsilon > 0$ , there is an  $N > m$  such that whenever  $n \geq N$ ,  $|y_n - y| < \frac{1}{2}|y|^2 \varepsilon$ . Then  $|\frac{1}{y_n} - \frac{1}{y}| \leq \frac{|y_n - y|}{|y_n y|} < \frac{2}{|y|^2} |y_n - y| < \varepsilon$ . Then choosing the sequences  $\{x_n\}$  and  $\{\frac{1}{y_n}\}$ , the rest follows. ■

**Corollary.** [label=(0)]

1. For any  $c \in \mathbb{C}$ , and a sequence  $x_n \rightarrow x$ , we have  $\lim c x_n = c \lim x_n = cx$  and  $\lim (c + x_n) = c + \lim x_n = c + x$  as  $n \rightarrow \infty$ .
2. Provided that  $x, x_n \neq 0$ , we have  $\lim \frac{1}{x_n} = \frac{1}{\lim x_n} = \frac{1}{x}$ , as  $n \rightarrow \infty$ .

*Proof.* We choose  $\{x_n\}$  and  $\{y_n\} = \{c\}$  for all  $n$ , then the results follow. ■

**Theorem 3.1.3.** [label=(0)]

1. Let  $x_n = (\alpha_{1n}, \dots, \alpha_{kn}) \in \mathbb{R}^k$ . Then  $\{x_n\}$  converges to  $x$  if and only if  $\lim \alpha_{jn} = \alpha_j$  for  $1 \leq j \leq k$ , as  $n \rightarrow \infty$ .
2. Let  $\{x_n\}, \{y_n\}$  be sequences in  $\mathbb{R}^k$ , and let  $\{\beta_n\}$  be a sequence in  $\mathbb{R}$  such that  $x_n \rightarrow x$ ,  $y_n \rightarrow y$ , and  $\beta_n \rightarrow \beta$ . Then  $\lim (x_n + y_n) = x + y$ ,  $\lim x_n y_n = xy$ , and  $\lim \beta_n x_n = \beta x$ .

*Proof.* If  $x_n \rightarrow x$ , then  $|\alpha_{jn} - \alpha_j| \leq \|x_n - x\| < \varepsilon$ , thus  $\lim \alpha_{jn} = \alpha_j$ . Conversely, suppose that  $\alpha_{jn} \rightarrow \alpha_j$ . Then for  $\varepsilon > 0$  there is an  $N \in \mathbb{Z}^+$  such that  $n \geq N$  implies  $|\alpha_{jn} - \alpha_j| < \frac{\varepsilon}{\sqrt{k}}$ . Then for  $n \geq N$ ,

$$\|x_n - x\| = \sqrt{\sum |\alpha_{jn} - \alpha_j|^2} < \varepsilon$$

To prove (2), we apply part (1) of this theorem together with theorem 3.1.2. ■

**Theorem 3.1.4** (The Sandwich Theorem). Let  $\{x_n\}$ ,  $\{y_n\}$ , and  $\{w_n\}$  be sequences in  $\mathbb{R}$ , and Suppose that  $\lim x_n = \lim y_n = a$  and that there is an  $N \in \mathbb{Z}^+$  such that  $x_n \leq w_n \leq y_n$  for all  $n \geq N$ . Then  $\lim_{n \rightarrow \infty} w_n = a$ .



*Proof.* Let  $\varepsilon > 0$  and let  $\{x_n\}$  and  $\{y_n\}$  both converge to  $a$ . Then by definition there are  $N_1, N_2 \in \mathbb{Z}^+$  such that  $|x_n - a| < \varepsilon$  and  $|y_n - a| < \varepsilon$  for  $n \geq N_1, N_2$ . Now choose  $N = \max\{N_0, N_1, N_2\}$ , if  $n \geq N$ , we have  $-\varepsilon < x_n - a < \varepsilon$ , and we also have  $x_n - a < w_n - a < y_n - a$ , thus we have that:

$$-\varepsilon < x_n - a < w_n - a < y_n - a < \varepsilon$$

Thus we have that  $|w_n - a| < \varepsilon$ . ■

**Corollary.** *If  $x_n \rightarrow \infty$  as  $n \rightarrow \infty$ , and  $\{y_n\}$  is bounded, then  $x_n y_n \rightarrow 0$  as  $n \rightarrow \infty$ .*

*Proof.* We have that  $\{y_n\}$  is bounded, hence, there is  $M > 0$  such that  $|y_n| < M$  for all  $n \in \mathbb{Z}^+$ . And since  $\{x_n\}$  converges to 0 we have that for any  $\varepsilon$  there is an  $N \in \mathbb{Z}^+$  such that for  $n \geq N$ ,  $|x_n - 0| < \frac{\varepsilon}{M}$ . For  $|x_n y_n - 0| = |x_n y_n| < M|x_n| < M \frac{\varepsilon}{M} = \varepsilon$ . Therefore,  $x_n y_n \rightarrow 0$  as  $n \rightarrow \infty$ . ■

**Corollary.** *Let  $\{x_n\}, \{y_n\}$  be sequences such that  $0 \leq x_n \leq y_n$  for  $n \geq N > 0$ . Then if  $y_n \rightarrow 0$ , then  $x_n \rightarrow 0$  as  $n \rightarrow \infty$ .*

*Proof.* This is a special case of the sandwich theorem. ■

## 3.2 Subsequences

**Definition.** Let  $\{x_n\}$  be a sequence, and let  $\{n_k\} \subset \mathbb{Z}^+$  such that  $n_k < n_{k+1}$ . We call the sequence  $\{x_{n_k}\}$  a **Subsequence** of  $\{x_n\}$ . If  $\{x_{n_k}\}$  converges, we call its limit the **subsequential limit** of  $\{x_n\}$ .

**Theorem 3.2.1.** *A sequence  $\{x_n\}$  converges to a point  $x$  if and only if every subsequence  $\{x_{n_k}\}$  converges to  $x$ .*

*Proof.* Clearly if  $x_n \rightarrow x$ , then every subsequence  $x_{n_k} \rightarrow x$ , (since subsequences can be thought of as subsets of their parent sequences). On the other hand, let  $x_{n_k} \rightarrow x$  for  $\{n_k\} \subseteq \mathbb{Z}^+$ . Then for  $\varepsilon > 0$ , there is a  $K \in \mathbb{Z}^+$  for which  $d(x_{n_k}, x) < \frac{\varepsilon}{2}$  for  $k \geq K$ . Let  $N \in \mathbb{Z}^+$ , and choose  $n \geq \max\{N, K\}$ , then  $d(x_n, x) \leq d(x_n, x_{n_k}) + d(x_{n_k}, x) < \varepsilon$ . ■

**Theorem 3.2.2.** *If  $\{x_n\}$  is a sequence in a compact metric space  $X$ , then some subsequence of  $\{x_n\}$  converges to a point  $x$ .*

*Proof.* If  $\{x_n\}$  is finite, then there is an  $x \in \{x_n\}$  and a sequence  $\{n_k\}$  with  $n_k < n_{k+1}$  such that  $x_{n_i} = x$  for  $1 \leq i \leq k$ , then the subsequence converges to  $x$ .

Now if  $\{x_n\}$  is infinite, there is a limit point  $x \in X$  of  $\{x_n\}$ , then choose  $n_i$  such that  $d(x, x_i) < \frac{1}{i}$  for  $1 \leq i \leq k$ . Obtaining  $\{n_k\}$  from this, we see that  $n_k < n_{k+1}$ , and so we get that  $\{x_{n_k}\}$  converges to  $x$ . ■

**Corollary.** *Every bounded sequence in  $\mathbb{R}^k$  contains a convergent subsequence.*

**Theorem 3.2.3.** *The subsequential limits of  $\{x_n\}$  in a metric space  $X$  form a closed subset of  $X$ .*

*Proof.* Let  $E$  be the set of all subsequential limits of  $\{x_n\}$ , and let  $x$  be a limit point of  $E$ . Choose  $n_i$  such that  $x_{n_i} \neq x$  and let  $\delta = d(x, x_{n_i})$ , for  $1 \leq i \leq k$ . Then consider the sequence  $\{n_k\}$ , since  $x$  is a limit point of  $E$ , there is an  $x' \in E$  for which  $d(x, x') < \frac{\delta}{2^i}$ . Thus there is an  $N_I > n_i$  such that  $d(x', x_{n_i}) < \frac{\delta}{2^i}$ , thus  $d(x, x_{n_i}) < \frac{\delta}{2^i}$ . So  $\{x_n\}$  converges to  $x$  and  $x \in E$ . ■

### 3.3 Cauchy Sequences

**Definition.** We call a sequence  $\{x_n\}$  in a metric space  $X$  a **Cauchy sequence** in  $X$ , or more simply, **Cauchy** in  $X$  if for all  $\varepsilon > 0$ , there is an  $N \in \mathbb{Z}^+$  such that  $d(x_n, x_m) < \varepsilon$  whenever  $m, n \geq N$ .

**Definition.** Let  $E$  be a nonempty subset of a metric space  $X$ , and let  $S \subseteq \mathbb{R}$  be the set of all real numbers  $d(x, y)$ , with  $x, y \in E$ . We call  $\sup S$  the **diameter** of  $E$ , and denote it  $\text{diam } E$ .

**Theorem 3.3.1.** Let  $\{x_n\}$  be a sequence, and let  $E_N$  be the set of all points  $p_N$  such that  $N < p_{n+1}$ . Then  $\{x_n\}$  is Cauchy if and only if  $\lim \text{diam } E_N = 0$  as  $N \rightarrow \infty$ .

*Proof.* Let  $\{x_n\}$  be Cauchy. Let  $x_{N_1}, x_{N_2} \in E$  such that  $d(x_n, x_{N_1}) < \frac{\varepsilon}{2}$ , and  $d(x_{N_2}, x_m) < \frac{\varepsilon}{2}$ . Then we see that  $d(x_{N_1}, x_{N_2}) \leq d(x_{N_1}, x_n) + d(x_n, x_{N_2}) < \varepsilon$ , so  $\{x_{N_k}\}$  is Cauchy and we see that  $\lim \text{diam } E_N = 0$ . Now suppose that  $\lim \text{diam } E_N = 0$ , then for any  $x_n, x_m \in S$ ,  $d(x_n, 0) < \frac{\varepsilon}{2}$  and  $d(0, x_m) < \frac{\varepsilon}{2}$  implies that  $d(x_n, x_m) \leq d(x_n, 0) + d(0, x_m) < \varepsilon$ , whenever  $n, m > N$ , for  $\varepsilon > 0$ . ■

**Theorem 3.3.2.** [label=(0)]

1. If  $E \subseteq X$ , then  $\text{diam } \hat{E} = \text{diam } E$ .
2. If  $\{K_n\}$  is a sequence of compact sets in  $X$ , such that  $K_{n+1} \subseteq K_n$ , and if  $\lim \text{diam } K_n = 0$  as  $n \rightarrow \infty$ , then  $\bigcap_{i=1}^{\infty} K_i$  contains exactly one point.

*Proof.* Clearly  $\text{diam } E \leq \text{diam } \hat{E}$ . Now let  $\varepsilon > 0$ , and choose  $x, y \in \hat{E}$ , then there are points  $x', y' \in E$  such that  $d(x, x') < \frac{\varepsilon}{2}$  and  $d(y, y') < \frac{\varepsilon}{2}$ . Hence,  $d(x, y) \leq d(x, x') + d(x', y') + d(y', y) < \varepsilon$ . Then choosing  $\varepsilon$  arbitrarily small,  $\text{diam } \hat{E} \leq \text{diam } E$ .

Now, we also have that by the nested interval theorem that  $K = \bigcap K_i$  is nonempty. Now suppose that  $K$  contains more than one point. then  $\text{diam } K > 0$ , and since  $K \subseteq K_n$  for all  $n$ ,  $\text{diam } K \leq \text{diam } K_n$ , a contradiction. Thus  $K$  contains exactly one element. ■

**Theorem 3.3.3.** [label=(0)]

1. In any metric space  $X$ , every convergent sequence is a Cauchy sequence.
2. If  $X$  is compact, and  $\{x_n\}$  is Cauchy in  $X$ , then  $\{x_n\}$  converges to a point in  $X$ .

*Proof.* [label=(0)]

If  $x_n \rightarrow x$ , and  $\varepsilon > 0$  such that there is an  $N \in \mathbb{Z}^+$  such that  $d(x_n, x) < \frac{\varepsilon}{2}$  for all  $n \geq N$ , then for  $m \geq N$ , we have  $d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \varepsilon$ . Thus  $\{x_n\}$  is Cauchy.

2. Let  $\{x_n\}$  be Cauchy, and let  $E_N$  be the set of all points  $x_N$  for which  $x_N < x_{N+1}$ . Then  $\lim \text{diam } \hat{E} = 0$ , then being closed in  $X$ , each  $\hat{E}_N$  is compact in  $X$ , and  $\hat{E}_{N+1} \subseteq \hat{E}_N$ , so by theorem 3.3.2, there is a unique  $x \in X$  in all of  $\hat{E}_N$ . Now for  $\varepsilon > 0$ , there is an  $N_0 \in \mathbb{Z}^+$  for which  $\text{diam } \hat{E} < \varepsilon$ . Then for all  $x_n \in \hat{E}$ ,  $d(x_n, x) < \varepsilon$  whenever  $n \geq N_0$ . ■

**Corollary** (The Cauchy Criterion). *Every Cauchy sequence in  $\mathbb{R}^k$  converges to a point in  $\mathbb{R}^k$ .*

*Proof.* Let  $\{x_n\}$  be Cauchy in  $\mathbb{R}^k$ , define  $E_N$  as in (2), then for some  $N \in \mathbb{Z}^+$ ,  $\text{diam } E < 1$ , and so  $\{x_n\}$  is the union of all  $E_n$ , and the set of points  $\{x_1, \dots, x_{N-1}\}$ , so  $\{x_n\}$  is bounded, and thus has a compact closure, it follows then that  $x_n \rightarrow x$  for some  $x \in \mathbb{R}^k$ . ■

**Definition.** We call a metric space **complete** if every Cauchy sequence in the space converges.

**Theorem 3.3.4.** *All compact metric spaces, and all Euclidean spaces are complete.*

**Example 3.2.** Consider  $\mathbb{Q}$  together with the metric  $|x - y|$ . The metric space induced on  $\mathbb{Q}$  by  $|\cdot|$  is not complete.

**Definition.** A sequence  $\{x_n\}$  in  $\mathbb{R}$  is said to be **monotonically increasing** if  $x_n \leq x_{n+1}$ ,  $\{x_n\}$  is said to be **monotonically decreasing** if  $x_{n+1} < x_n$ . We call  $\{x_n\}$  **monotonic** if it is either monotonically increasing or monotonically decreasing.

**Theorem 3.3.5.** *A monotonic sequence converges if and only if it is bounded.*

*Proof.* Suppose, without loss of generality, that  $\{x_n\}$  is monotonically increasing. If  $\{x_n\}$  is bounded, then  $x_n \leq x$ , then for all  $\varepsilon > 0$ , there is an  $N \in \mathbb{Z}^+$  for which  $x - \varepsilon < x_N \leq x$ . Then for  $n \geq N$ ,  $x_n \rightarrow x$ . The converse follows from theorem 3.1.2. ■

## 3.4 Upper and Lower Limits.

Let  $\{\xi_n\}$  be a sequence in  $\mathbb{R}$  such that for all  $M > 0$ , there is an  $N \in \mathbb{Z}^+$  for which  $n \geq N$  implies that either  $x_n \geq M$ , or  $x_n \leq -M$ . Then we write  $x_n \rightarrow \infty$  and  $x_n \rightarrow -\infty$ , respectively.

**Definition.** Let  $\{x_n\}$  be a sequence in  $\mathbb{R}$ , and let  $E$  be the set of all extended real numbers  $x$  such that  $x_{n_k} \rightarrow x$  for some subsequence  $\{x_{n_k}\}$ . Then  $E$  contains all subsequential limits of  $\{x_n\}$ , and possible  $\pm\infty$ . We then call  $\sup E$  the **upper limit** of  $E$ , and  $\inf E$  the **lower limit** of  $E$ .

**Theorem 3.4.1.** *Let  $\{x_n\}$  be a sequence in  $\mathbb{R}$ , and let  $E$  be the set of all extended real numbers  $x$ , let  $s = \sup E$  and  $s' = \inf E$ . Then the following hold:*

*[label=(0)]s, s' \in E. If x \notin E, and x' \notin E, there is an N \in \mathbb{Z}^+ such that n \geq N implies that x' < x\_n < x.*

2. *Proof.* We prove the theorem for the case of  $s$ , since it is analogous for  $s'$ .

[label=(0)]If  $s = \infty$ , then  $E$  is not bounded above, so neither is  $\{x_n\}$ , and there is a subsequence for which  $x_n \rightarrow \infty$ . Now if  $s \in \mathbb{R}$ , then  $E$  is bounded above, and has at least one subsequential limit. Then  $s \in E$ . Now if  $s = -\infty$ , then  $E$  contains only  $-\infty$ , and so by definition  $x_n \rightarrow -\infty$ . Suppose there is an  $x > s$ , such that  $x_n \geq x$  for all  $n$ . Then there is a  $y \in E$  such that  $y \geq x \geq s$ , a contradiction of the definition of  $s$ . ■

**Example 3.3.** [label=(0)]

1. Let  $\{x_n\}$  be a sequence in  $\mathbb{Q}$ , then every real number is a subsequential limit, and  $\limsup x_n = \infty$  and  $\liminf x_n = -\infty$ .
2. Let  $\{x_n\} = \{\frac{(-1)^n}{1+\frac{1}{n}}\}$ ; then  $\limsup x_n = 1$  and  $\liminf x_n = -1$  as  $n \rightarrow \infty$ .
3. For a sequence  $\{x_n\}$  in  $\mathbb{R}$ ,  $\lim x_n = x$  if and only if  $\limsup x_n = \liminf x_n = x$  as  $n \rightarrow \infty$ .

**Theorem 3.4.2.** *If  $x_n \leq y_n$ , for  $n \geq N > 0$ , then  $\liminf x_n \leq \liminf y_n$  and  $\limsup x_n \leq \limsup y_n$  as  $n \rightarrow \infty$ .*

## 3.5 Special Sequences

**Theorem 3.5.1.** *Let  $n, p \in \mathbb{Z}^+$ . Then the following hold as  $n \rightarrow \infty$ .*

[label=(0)]

1.  $\lim \frac{1}{n^p} = 0$ .
2.  $\lim \sqrt[p]{n} = 1$ .
3.  $\lim \sqrt[n]{n} = 1$ .
4. If  $\alpha \in \mathbb{R}$ , then  $\lim \frac{n^\alpha}{(1+p)^n} = 0$ .
5. If  $|x| < 1$ , then  $\lim x^n = 0$ .

*Proof.* [label=(0)]

Let  $n > [p] \frac{1}{\varepsilon}$ ; then  $|\frac{1}{n^p}| < \varepsilon$ .

2. If  $p = 1$ , we are done. If  $p > 1$ , let  $x_n = \sqrt[p]{n} - 1$ , then  $x_n > 0$ . By the binomial theorem,  $1 + nx_n \leq (1 + x_n)^p = n$ , hence  $0 \leq x_n \leq \frac{n-1}{p}$ . Now if  $1 > p > 0$ , then  $\frac{1}{p} > 0$ , so we notice that  $0 \leq \frac{1}{x_n} \leq \frac{1}{\frac{1}{n}}$ .
3. Let  $x_n = \sqrt[n]{n} - 1$ , then  $x_n \geq 0$ , then by the binomial theorem again,  $n = (1 + x_n)^n \geq \frac{n(n-1)}{2} x_n^2$ , then  $0 \leq x_n \leq \sqrt{\frac{2}{n-1}}$ .

4. Let  $k \in \mathbb{Z}^+$  such that  $k > \alpha$ . Then  $n > 2k$ , let  $(1+p)^n > \binom{n}{k} p^k > \frac{n^k p^k}{2^k k!}$ . So  $0 < \frac{n^\alpha}{(1+p)^n} < \frac{2^k k!}{p^k} n^{-k}$ , since  $\alpha - k < 0$ ,  $n^{\alpha-k} \rightarrow 0$  and we are done.
5. Take  $\alpha = 0$ , and let  $x = \frac{1}{1+p}$ , then the result follow. ■



# Chapter 4

## Series.

### 4.1 Convergent Series.





# Chapter 5

## Continuity

### 5.1 Limits of Functions.

**Definition.** Let  $X$ , and  $Y$  be metric spaces, and let  $E \subseteq X$ , and let  $f : E \rightarrow Y$  be a function. We say that  $f$  **converges** to a point  $q \in Y$ , as  $x$  **approaches** a limit point  $p \in X$  if for every  $\varepsilon > 0$ , there is a  $\delta > 0$  for which  $d_Y(f(x), q) < \varepsilon$ , whenever  $0 < d_X(x, p) < \delta$ . We say that  $q$  is the **limit** of  $f$  at  $p$  and we write  $f \rightarrow q$  as  $x \rightarrow p$ , and  $\lim_{x \rightarrow p} f(x) = q$ , or more simply,  $\lim f = q$ .

**Example 5.1.** [label=(0)]

1. Let  $X = Y = \mathbb{R}$ , under the absolute value  $|\cdot|$ , and let  $I \subseteq \mathbb{R}$  be an interval, and  $f : I \rightarrow \mathbb{R}$ . Then  $f$  has a limit  $L$  as  $x$  approaches a limit point  $c \in \mathbb{R}$  if for every  $\varepsilon > 0$ , there is a  $\delta > 0$  such that  $|f(x) - L| < \varepsilon$  whenever  $0 < |x - c| < \delta$ . We call functions that map into  $\mathbb{R}$  **real valued**.
2. Let  $X = Y = \mathbb{C}$ , under the modulus  $|\cdot|$ , and let  $D \subseteq \mathbb{R}$  be an domain, and  $f : D \rightarrow \mathbb{R}$ . Then  $f$  has a limit  $L$  as  $z$  approaches a limit point  $w \in \mathbb{R}$  if for every  $\varepsilon > 0$ , there is a  $\delta > 0$  such that  $|f(x) - L| < \varepsilon$  whenever  $0 < |z - w| < \delta$ . We call functions that map into  $\mathbb{C}$  **complex valued**.
3. Let  $X = Y = \mathbb{R}^k$ , under the norm  $\|\cdot\|$ , and let  $D \subseteq \mathbb{R}^k$  be an domain, and  $f : D \rightarrow \mathbb{R}^k$ . Then  $f$  has a limit  $L$  as  $x$  approaches a limit point  $c \in \mathbb{R}^k$  if for every  $\varepsilon > 0$ , there is a  $\delta > 0$  such that  $\|f(x) - L\| < \varepsilon$  whenever  $0 < \|x - c\| < \delta$ . We call functions that map into  $\mathbb{R}^k$  **vector valued**.

**Theorem 5.1.1** (The Sequential Criterion). *Let  $X$  and  $Y$  be metric spaces, and let  $E \subseteq X$ , and  $f : E \rightarrow Y$  be a function, and  $p \in E$  be a limit point. Then  $\lim f(x) = q$  as  $x \rightarrow p$  if and only if  $\lim f(x_n) = q$  as  $n \rightarrow \infty$  for any sequence  $\{x_n\} \in E$ , such that  $x_n \neq p$  and  $\lim x_n = p$ .*

*Proof.* Suppose that  $\lim f(x) = q$  as  $x \rightarrow p$ , and choose  $\{x_n\} \subseteq E$  such that  $x_n \neq p$  and  $\lim x_n = p$  as  $n \rightarrow \infty$ . Then for  $\varepsilon > 0$ , there is a  $\delta > 0$  such that  $d_Y(f(x), q) < \varepsilon$  whenever  $0 < d_X(x, p) < \delta$ , and since  $d_X(x_n, p) < \delta$  whenever  $n \geq N$  for some  $N > 0$ , we have  $d_Y(f(x_n), q) < \varepsilon$  whenever  $d_X(x_n, p) < \delta$ .

Conversely, suppose that  $\lim f \neq q$ , that is for some  $\varepsilon > 0$ ,  $d_Y(f(x), q) \geq \varepsilon$  whenever  $d_X(x, p) < \delta$  for all  $\delta > 0$ . Then choose  $\delta = \frac{1}{n}$ , for  $n \in \mathbb{Z}^+$ , then we have  $\lim x_n = p$ , but  $\lim f(x_n) \neq q$ . ■

The importance of the sequential criterion is that it lets us translate theorems about limits of sequences into theorems about limits of functions.

**Corollary.** *If  $f$  has a limit at  $p$ , then the limit of  $f$  is unique.*

**Definition.** Letting  $f, g : E \rightarrow Y$ , we define the **sum**, **product**, **scalar product** and the **quotient** of  $f$  and  $g$  to be the functions from  $E$  into  $Y$ :

$$\begin{aligned} f + g(x) &= f(x) + g(x). \quad fg(x) = f(x)g(x). \quad (\lambda f)(x) = \lambda f(x) \text{ for } \lambda \in X. \\ \frac{f}{g}(x) &= \frac{f(x)}{g(x)}, \text{ provided that } g(x) \neq 0. \end{aligned}$$

It is well known that the set of all functions from  $E$  into  $Y$  form an algebra under these operations.

**Theorem 5.1.2.** *Let  $E \subseteq X$  a metric space, and let  $p \in E$  be a limit point. Let  $f, g : E \rightarrow Y$  be functions, such that  $\lim f = A$  and  $\lim g = B$  as  $x \rightarrow p$ . Then the following hold as  $x \rightarrow p$ .*

$$\begin{aligned} \lim (f + g) &= \lim f + \lim g = A + B. \quad \lim fg = \lim f \lim g = AB. \quad \lim \frac{f}{g} = \\ &= \frac{\lim f}{\lim g} = \frac{A}{B}, \text{ provided that } B \neq 0. \end{aligned}$$

**Corollary.** *The following hold:*

$$\begin{aligned} \lim \lambda f &= \lambda \lim f = \lambda A, \text{ and } \lim (\lambda + f) = \lambda + \lim f = \lambda + A. \quad \lim \frac{1}{f} = \\ &= \frac{1}{\lim f} = \frac{1}{A}, \text{ provided that } A \neq 0. \end{aligned}$$

**Theorem 5.1.3** (The Sandwich Theorem). *Let  $f, g$ , and  $h$  be real valued functions defined on  $\mathbb{R}$  such that  $\lim f = \lim g = A$  as  $x \rightarrow p$ , and suppose that  $f(x) \leq h(x) \leq g(x)$  for all  $x \in \mathbb{R}$ . Then  $\lim h = A$  as  $x \rightarrow p$ .*

**Corollary.** *Let  $f, g$  be real valued functions defined on  $\mathbb{R}$  such that  $0 \leq f(x) \leq g(x)$  for all  $x \in \mathbb{R}$ . Then if  $g \rightarrow 0$  as  $x \rightarrow p$ , then  $f \rightarrow 0$ .*

The proofs of all these are the result of applying the sequential criterion.

## 5.2 Continuous Functions.

**Definition.** Let  $X$  and  $Y$  be metric spaces and let  $p \in E \subseteq X$ , and  $f : E \rightarrow Y$  be a function. We say that  $f$  is **continuous** at  $p$  if for every  $\varepsilon > 0$ , there is a  $\delta > 0$  such that  $d_Y(f(x), f(p)) < \varepsilon$  whenever  $0 < d_X(x, p) < \delta$ . If  $f$  is continuous at every point in  $X$ , we say that  $f$  is **continuous on  $X$** .

**Theorem 5.2.1.** *If  $E \subseteq X$  a metric space, and if  $f$  is a function defined on  $X$ , and  $p \in E$  is a limit point, then  $f$  is continuous if and only if  $\lim f(x) = f(p)$  as  $x \rightarrow p$ .*

**Theorem 5.2.2.** *Suppose  $X, Y$ , and  $Z$  are metric spaces, and that  $f : E \rightarrow Y$ ,  $g : Y \rightarrow Z$ , are functions (with  $E \subseteq X$ ) such that  $f$  is continuous at  $p$  and  $g$  is continuous at  $f(p)$ . Then  $g \circ f$  is continuous at  $p$ .*

2. *Proof.* For every  $\varepsilon > 0$ , we have  $\delta_1, \delta_2 > 0$  such that  $d_Y(f(x), f(p)) < \varepsilon$ , when  $0 < d_X(x, p) < \delta_1$ , and  $d_Z(g(y), g(f(p))) < \varepsilon$  whenever  $d_Y(y, f(p)) < \delta_2$ . Then choose  $\delta = \min\{\delta_1, \delta_2\}$ , and we see that  $d_Z(g(f(x)), g(f(p))) < \varepsilon$  whenever  $0 < d_X(x, p) < \delta$ . ■

**Theorem 5.2.3.** *A mapping  $f$  of a metric space  $X$  into a metric space  $Y$  is continuous if and only if for every open set  $V \subseteq Y$ ,  $f^{-1}(V)$  is open in  $X$ .*

*Proof.* Let  $f$  be continuous on  $X$ , and let  $V$  be open in  $Y$ . For  $p \in X$ ,  $f(p) \in V$ , and since  $V$  is open, there is an  $\varepsilon > 0$  such that  $y \in V$  when  $d_Y(y, f(p)) < \varepsilon$ . Since  $f$  is continuous, there is a  $\delta > 0$  such that  $d_Y(f(x), f(p)) < \varepsilon$ , whenever  $0 < d_X(x, p) < \delta$ . Thus  $f^{-1}(V)$  is open in  $X$ .

Conversely, suppose that  $f^{-1}(V)$  is open in  $X$  for  $V$  open in  $Y$ . Let  $p \in X$  and  $\varepsilon > 0$ , and let  $V = \{y \in Y : d_Y(y, f(p)) < \varepsilon\}$ ;  $V$  is open in  $Y$ , so  $f^{-1}(V)$  is open in  $X$ , thus there is a  $\delta > 0$  such that  $x \in f^{-1}(V)$  when  $0 < d_X(x, p) < \delta$ , then  $f(x) \in V$ , so  $d_Y(f(x), f(p)) < \varepsilon$ ; therefore,  $f$  is continuous at  $p$ . ■

**Corollary.** *A mapping  $f$  from  $X$  into  $Y$  is continuous if and only if  $f^{-1}(C)$  is closed in  $X$ , whenever  $C$  is closed in  $Y$ .*

*Proof.* This is the converse of the previous theorem. ■

**Theorem 5.2.4.** *Let  $f, g : X \rightarrow \mathbb{C}$  be continuous complex valued functions defined on a metric space  $X$ , then  $f + g$ ,  $fg$ , and  $\frac{f}{g}$  are continuous.*

*Proof.* This follows from theorem 5.1.2 and the sequential criterion. ■

**Theorem 5.2.5.** *Let  $f_1, \dots, f_k$  be realvalued functions defined on a metric space  $X$ , and define  $f : X \rightarrow \mathbb{R}^k$  by  $f(x) = (f_1(x), \dots, f_k(x))$  for all  $x \in X$ . Then  $f$  is continuous if and only if  $f_i$  is continuous for  $1 \leq i \leq k$ . Moreover, if  $g : X \rightarrow \mathbb{R}^k$  and  $f$  are continuous, then so is  $f + g$  and  $fg$ .*

*Proof.* Notice that  $|f_i(x) - f_i(y)| \leq \|f(x) - f(y)\| = \sqrt{\sum |f_i(x) - f_i(y)|^2}$  for  $1 \leq i \leq k$ . It follows then that  $f$  is continuous if and only if  $f_i$  is. Moreover, if  $g : X \rightarrow \mathbb{R}^k$  is also continuous, then by the previous theorem, so is  $f + g$  and  $fg$ . ■

**Example 5.2.** [label=(0)]

1. Let  $x \in \mathbb{R}^k$ , define the functions  $\phi_i : \mathbb{R}^k \rightarrow \mathbb{R}$  by  $\phi_i(x) = x_i$  for all  $1 \leq i \leq k$ , then  $\phi_i$  is continuous on  $\mathbb{R}^k$
2. The monomials  $x_1^{n_1} x_2^{n_2} \dots x_k^{n_k}$ , with  $n_i \in \mathbb{Z}^+$  for  $1 \leq i \leq k$  are continuous on  $\mathbb{R}^k$ . So are all constant multiples, thus the polynomial  $\sum c_{n_1, \dots, n_k} x_1^{n_1} x_2^{n_2} \dots x_k^{n_k}$  is also continuous on  $\mathbb{R}^k$ .
3. We have  $|||x|| - ||y||| \leq ||x - y||$  for all  $x, y \in \mathbb{R}^k$ , thus the mapping  $x \rightarrow ||x||$  is continuous on  $\mathbb{R}^k$ .

### 5.3 Continuity and Compactness.

**Definition.** A mapping  $f : E \rightarrow \mathbb{R}^k$  is said to be **bounded** if there is a real number  $M > 0$  such that  $\|f\| \leq M$  for all  $x \in E$ .

**Theorem 5.3.1.** *Let  $f$  be a continuous mapping of a compact metric space  $X$  into a metric space  $Y$ . Then  $f(X)$  is compact in  $Y$ .*

*Proof.* Let  $\{V_\alpha\}$  be an open cover of  $f(X)$ , since  $f$  is continuous, then  $f^{-1}(V_\alpha)$  is open in  $X$ , and since  $X$  is compact,  $X \subseteq \bigcup_{i=1}^n V_{\alpha_i}$ , and  $f(f^{-1}(E)) \subseteq E$ , we have that  $f(X) \subseteq \bigcup_{i=1}^n V_{\alpha_i}$ . ■

**Theorem 5.3.2.** *If  $f : X \rightarrow \mathbb{R}^k$  is continuous, where  $X$  is a compact metric space, then  $f(X)$  is closed and bounded; in particular,  $f$  is bounded.*

*Proof.* From theorem 5.3.1, we have that  $f(X)$  is compact in  $\mathbb{R}^k$ , therefore, it is closed and bounded. ■

**Theorem 5.3.3** (The Extreme Value Theorem). *Suppose  $f$  is a continuous, realvalued function on a metric space  $X$ , and that  $M = \sup f$ , and  $m = \inf f$ . Then there exist points  $p, q \in X$  such that  $f(p) = M$  and  $f(q) = m$ .*

*Proof.* By theorem 5.3.2,  $f(X)$  is closed and bounded, thus  $M, m \in f(X)$ . ■

**Theorem 5.3.4.** *Suppose  $f$  is a continuous 1-1 mapping of a compact metric space  $X$  onto a metric space  $Y$ . Then the inverse mapping  $f^{-1} : Y \rightarrow X$  is a Continuous mapping of  $Y$  onto  $X$ .*

*Proof.* By theorem 5.2.3, it suffices to show that  $f(V)$  is open in  $Y$  whenever  $V$  is open in  $X$ . We have that  $X \setminus V$  is closed in  $X$ , and compact, thus  $f(X \setminus V)$  is closed and compact in  $Y$ , thus  $f(V) = Y \setminus f(X \setminus V)$  is open in  $Y$ . ■

**Definition.** Let  $f$  be a mapping of a metric space  $X$  into a metric space  $Y$ . We say that  $f$  is **uniformly continuous** on  $X$  if for every  $\varepsilon > 0$ , there is a  $\delta > 0$  such that  $d_Y(f(q), f(p)) < \varepsilon$ , for all  $p, q \in X$  for which  $d_X(p, q) < \delta$ .

**Lemma 5.3.5.** *If  $f$  is uniformly continuous, then  $f$  is continuous.*

**Theorem 5.3.6.** *Let  $f$  be a continuous mapping of a compact metric space  $X$  into a metric space  $Y$ . Then  $f$  is uniformly continuous on  $X$ .*

*Proof.* Let  $\varepsilon > 0$ , by the continuity of  $f$ , we can associate for each  $p \in X$  a number  $\phi(p) > 0$  such that for  $q \in X$ ,  $d_X(p, q) < \phi(p)$  implies  $d_Y(f(p), f(q)) < \frac{1}{2}\phi(p)$ . Now let  $J(p) = \{q \in X : d_X(p, q) < \phi(p)\}$ . Clearly,  $p \in J(p)$ , so  $J(p)$  is an open cover of  $X$ , and since  $X$  is compact, there are  $p_1, \dots, p_n$  for which  $X \subseteq \bigcup_{i=1}^n J(p_i)$ , then take  $\delta = \min\{\phi(p_1), \dots, \phi(p_n)\}$ ; we have  $\delta > 0$ . Now let  $p, q \in X$  such that  $d_X(p, q) < \delta$ . Then there is an  $m \in \mathbb{Z}^+$  with  $1 \leq m \leq n$  such that  $p \in J(p_m)$ , thus  $d_X(p, q) < \frac{1}{2}\phi(p_m)$ , by the triangle inequality, we get  $d_Y(q, p_m) \leq d_Y(q, p) + d_Y(p, p_m) < \delta + \frac{1}{2}\phi(p_m) = \phi(p_m)$ , for  $1 \leq m \leq n$ . Therefore,  $d_Y(f(p), f(q)) \leq d_Y(f(p), f(p_m)) + d_Y(f(p_m), f(q)) < \varepsilon$ . Thus,  $f$  is uniformly continuous. ■

*Remark.* What this theorem says, is that in any compact metric space, continuity and uniform continuity are equivalent.

**Theorem 5.3.7.** *Let  $E \subseteq \mathbb{R}$  be noncompact, then:*

*[label=(0)]There exists a continuous function on  $E$  which is not bounded. There is a bounded, continuous function on  $E$  which has no maximum. If  $E$  is bounded, there exists a continuous function on  $E$  that is not uniformly continuous.*

**2. Proof.** Suppose first that  $E$  is bounded. Then there is a limit point  $x_0 \notin E$  of  $E$ . Consider the function

$$f(x) = \frac{1}{x - x_0} \text{ for all } x \in E$$

Then  $f$  is continuous on  $E$ , but not bounded. Then let  $\varepsilon > 0$  and  $\delta > 0$ , and choose  $x \in E$  such that  $|x - x_0| < \delta$ , then taking  $t$  arbitrarily close to  $x_0$ , we can get  $|f(x) - f(t)| \geq \varepsilon$ , even though  $|x - t| < \delta$ . Thus  $f$  is not uniformly continuous.

Now choose

$$g(x) = \frac{1}{1 + (x - x_0)^2} \text{ for all } x \in E$$

$g$  is continuous, and bounded on  $E$  ( $0 \leq g \leq 1$ ), then  $\sup g = 1$ , and since  $g(x) < 1$  for all  $x$ , we see that  $g$  attains no maximum.

Lastly, suppose that  $E$  is unbounded, then the functions  $f(x) = x$  and  $h(x) = \frac{x^2}{1+x^2}$  for all  $x \in E$  establish (1) and (2). ■

**Example 5.3.** Let  $f$  be the mapping of the interval  $[0, 2\pi)$  onto the unit circle. That is  $f(t) = (\cos t, \sin t)$  for  $0 \leq t < 2\pi$ . Then  $f$  is a continuous 1-1 mapping of  $[0, 2\pi)$  onto the unit circle, however, the inverse mapping,  $f^{-1}$  fails to be continuous at the point  $f(0) = (1, 0)$ .

## 5.4 Continuity and Connectedness.

**Theorem 5.4.1.** *If  $f$  is a continuous mapping of a metric space  $X$  into a metric space  $Y$ , and if  $E \subseteq X$  is Connected, then so is  $f(E)$ .*

*Proof.* Suppose that  $f(E) = A \cup B$  with  $A, B \subseteq Y$  nonempty and separated. Let  $G = E \cap f^{-1}(A)$  and  $H = E \cap f^{-1}(B)$ , then  $E = G \cup H$ , and  $G$  and  $H$  are both nonempty. Then since  $A \subseteq \overline{A}$ ,  $G \subseteq f^{-1}(\overline{A})$ , and since  $f$  is continuous,  $f^{-1}(\overline{A})$  is closed, so  $\overline{G} \subseteq f^{-1}(\overline{A})$ , thus  $f(\overline{G}) \subseteq \overline{A}$ . Since  $f(H) = B$ , and  $\overline{A} \cap B$  is empty,  $\overline{G} \cap H$ , and  $H \cap \overline{H}$  are also empty, which contradicts the connectedness of  $E$ . ■

**Theorem 5.4.2** (The Intermediate Value Theorem). *Let  $f[a, b] \rightarrow \mathbb{R}$  be a realvalued function. If  $f(a) < f(b)$ , and  $c \in \mathbb{R}$  such that  $f(a) < c < f(b)$ , then there is an  $x \in (a, b)$  such that  $f(x) = c$ .*

*Proof.* We have that  $[a, b]$  is connected in  $\mathbb{R}$ , thus by theorem 5.4.1,  $f([a, b])$  is connected in  $\mathbb{R}$ , hence there is an  $x \in (a, b)$  for which  $f(x) = c$ . ■

**Corollary.** *If  $f : [a, b] \rightarrow \mathbb{R}$  is a realvalued function such that  $f(a) < 0 < f(b)$ , then there is an  $x \in (a, b)$  such that  $f(x) = 0$ .*

## 5.5 Discontinuities.

**Definition.** Let  $X$  and  $Y$  be metric spaces, and let  $f : E \rightarrow Y$  for  $E \subseteq X$ . If there is a point  $x \in E$  for which  $f$  is not continuous, we say that  $f$  is discontinuous at  $x$ , and we say that  $f$  has a **discontinuity** at  $x$ .

**Definition.** Let  $f$  be defined on  $(a, b)$ , and let  $x$  be such that  $a \leq x < b$ . We write  $f(x+) = q$  if  $f(t_n) \rightarrow q$  for all sequences  $\{t_n\}$  in  $(x, b)$  such that  $t_n \rightarrow x$ . Similarly, if  $a < x \leq b$ , we write  $f(x-) = q$  if  $f(t_n) \rightarrow q$  for all sequences  $\{t_n\}$  in  $(a, x)$  such that  $t_n \rightarrow x$ . We call  $f(x+)$  and  $f(x-)$  the **right handed limit** and **left handed limit** of  $f$  at  $x$  respectively, and write  $\lim_{t \rightarrow x+} f = f(x+)$  and  $\lim_{t \rightarrow x-} f = f(x-)$ .

**Theorem 5.5.1.** If  $x \in (a, b)$ , then  $\lim f$  exists as  $t \rightarrow x$  if and only if,  $f(x+) = f(x-) = \lim f$ .

*Proof.* Suppose that  $\lim f$  exists, by the uniqueness of the limit, and the sequential criterion, we get that  $f(x+) = f(x-) = \lim f$ . Conversely, suppose that  $f(x+) = f(x-) = q$ . Then  $f(t_n) \rightarrow q$  for all sequences  $\{t_n\}$  in  $(x, b)$  and  $(a, x)$ , then  $f(t_n) \rightarrow q$  for all sequences  $\{t_n\}$  in  $(a, b)$ , thus by the sequential criterion again,  $\lim f$  exists, and  $\lim f = q$ . ■

**Definition.** Let  $f$  be defined on  $(a, b)$ . If  $f$  is discontinuous at a point  $x$ , and  $f(x+)$  and  $f(x-)$  exists, we say that  $f$  has a **removable discontinuity** at  $x$ , otherwise, we say the  $f$  has an **infinite discontinuity**.

**Example 5.4.** [label=(0)]

1. The function  $f(x) = 1$  for  $x \in \mathbb{Q}$  and  $f(x) = 0$  for  $x \in \mathbb{R} \setminus \mathbb{Q}$  has an infinite discontinuity at every point  $x$ .
2. The function  $f(x) = x$  for  $x \in \mathbb{Q}$  and  $f(x) = 0$  for  $x \in \mathbb{R} \setminus \mathbb{Q}$  is continuous at  $x = 0$ , and has an infinite discontinuity at every other point  $x$ .
3. The function  $f(x) = \sin \frac{1}{x}$  for  $x \neq 0$  and  $f(x) = 0$  for  $x = 0$ , has an infinite discontinuity at  $x = 0$ .
4. The function  $f(x) = x + 2$  for  $-3 < x < -2$  and  $0 \leq x < 1$  and  $f(x) = -x - 2$  for  $-2 \leq x < 0$  has a removable discontinuity at  $x = 0$ , and is continuous everywhere else.

*Remark.* The discontinuities in examples (1) and (2) are the result of  $\mathbb{Q}$  and  $\mathbb{R} \setminus \mathbb{Q}$  being dense in  $\mathbb{R}$ .

## 5.6 Monotonic Functions.

**Definition.** Let  $f$  be a realvalued function on an interval  $(a, b)$ . We say that  $f$  is **monotonically increasing** on  $(a, b)$  if  $a < x < y < b$  implies  $f(x) \leq f(y)$ . We say that  $f$  is **monotonically decreasing** on  $(a, b)$  if  $a < x < y < b$  implies  $f(y) \leq f(x)$ . We say  $f$  is **monotonic** if it is either monotonically increasing or monotonically decreasing.

**Theorem 5.6.1.** *Let  $f$  be monotonic on  $(a, b)$  then  $f(x+)$  and  $f(x-)$  exist at every point of  $(a, b)$  and  $\sup f = f(x-)$  and  $\inf f = f(x+)$ , and the following hold:*

*[label=(0)]If  $f$  is monotonically increasing  $f(x-) \leq f(x) \leq f(x+)$  If  $f$  is monotonically decreasing  $f(x+) \leq f(x) \leq f(x-)$*

(b) *Proof.* We prove only (1), since (2) is analogous. Suppose that  $f$  is monotonically increasing, clearly,  $f$  has an upperbound  $A$  for which  $A \leq f$ . Now let  $\varepsilon > 0$ , then there is a  $\delta > 0$  for which  $a < x - \delta < x$ , and  $A - \varepsilon < f(x - \delta) \leq A$ . Then we have  $f(x - \delta) < f(t) \leq A$  for all  $x - \delta < t < x$ , then we get  $|f(t) - A| < \varepsilon$ , hence  $f(x-) = A$ . Similarly, we get  $f(+) = -\inf f$ . Now since  $\sup f \leq f \leq \inf f$ , we get the desired result. ■

**Corollary.** *Monotonic functions have no infinite discontinuities.*

**Theorem 5.6.2.** *Let  $f$  be monotonic on  $(a, b)$ , then the set of all points of  $(a, b)$  for which  $f$  is discontinuous is at most countable.*

*Proof.* Suppose, without loss of generality that  $g$  is monotonically increasing, and let  $E$  be the set of all points of  $(a, b)$  for which  $f$  is discontinuous. By the density of  $\mathbb{Q}$  in  $\mathbb{R}$ , for each  $x \in E$  associate  $r(x) \in \mathbb{Q}$  such that  $f(x+) < f(x) < f(x-)$ . Since  $x_1 < x_2$  implies  $f(x_1+) \leq f(x_2-)$ , then  $r(x_1) \neq r(x_2)$ , thus  $x_1 \neq x_2$ , and so  $r$  is a 1-1 mapping of  $E$  into  $\mathbb{Q}$ . ■

Now, given a countable  $E$  in an interval  $(a, b)$ , we can construct a monotonic function  $f$  that is discontinuous at every point in  $E$  and continuous everywhere else. Arrange the points of  $E$  into a sequence  $\{x_n\}$  and let  $\{c_n\}$  be a sequence such that  $c_n > 0$  for all  $n \in \mathbb{Z}^+$ , such that  $\sum c_n$  converges. Define  $f(x) = \sum_{x_n < x} c_n$ , for  $x \in (a, b)$ . Then we have that

*[label=(0)] $f$  is monotonically increasing on  $(a, b)$ .  $f$  is discontinuous at every point in  $E$  with  $f(x_n+) - f(x_n-) = c_n$ .  $f$  is continuous at every point in  $(a, b) \setminus E$ .*

**Definition.** Let  $f$  be a realvalued function defined on an interval  $(a, b)$ . We say that  $f$  is **continuous from the right** if  $f(x+) = f(x)$ , and we say  $f$  is **continuous from the left** if  $f(x-) = f(x)$ .

## 5.7 Infinite Limits and Limits at Infinity.

**Definition.** For any  $c \in \mathbb{R}$ , the set of all real numbers  $x$  such that  $x > c$  is called the **neighborhood of  $\infty$** , and denoted  $(c, \infty)$ . The set of all real numbers  $x$  such that  $x < c$  is called the **neighborhood of  $-\infty$** , and denoted  $(-\infty, c)$ .

**Definition.** Let  $f : E \rightarrow \mathbb{R}$  be a realvalued function. We say that  $f(t) \rightarrow A$  as  $t \rightarrow x$ , with  $A$ , and  $x$  extended real numbers if for every neighborhood of  $U$   $A$ , there is a neighborhood  $V$  of  $x$  such that  $V \cap E$  is nonempty, and  $f(t) \in U$  for all  $t \neq x \in V \cup E$ .

**Theorem 5.7.1.** *Let  $f, g : E \rightarrow \mathbb{R}$  be realvalued functions such that  $f \rightarrow A$ , and  $g \rightarrow B$  as  $t \rightarrow x$ , for extended real numbers  $A, B$ , and  $x$ . Then the following hold as  $t \rightarrow x$ .*

*[label=(0)] $f \rightarrow A'$  implies  $A = A'$ .  $f + g \rightarrow A + B$ .  $fg \rightarrow AB$ .  $\frac{f}{g} \rightarrow \frac{A}{B}$ . Provided that (1), (2), and (3) are not of the forms  $\infty - \infty$ ,  $0 \cdot \infty$ ,  $\frac{\infty}{\infty}$ , and  $A_0$ , respectively.*

**2.** *Proof.* This is a direct application of the sequential criterion using the appropriate definition. ■



# Chapter 6

## Differentiation

### 6.1 The Derivative of Real valued Functions.

**Definition.** Let  $f : [a, b] \rightarrow \mathbb{R}$  be a realvalued function defined on  $[a, b]$ . The **derivative** of  $f$  at a point  $x \in (a, b)$  is the function  $f' : (a, b) \rightarrow \mathbb{R}$  defined by

$$f'(x) = \lim_{t \rightarrow x} \frac{f(t) - f(x)}{t - x} \quad (6.1)$$

If  $f'$  is defined at  $x \in [a, b]$ , then we say that  $f$  is **differentiable** at  $x$ , and if  $f'$  is defined for all  $x \in (a, b)$ , we say that  $f$  is **differentiable** on  $(a, b)$ .

**Theorem 6.1.1.** Let  $f : [a, b] \rightarrow \mathbb{R}$  be a realvalued function. If  $f$  is differentiable at a point  $x \in (a, b)$ , then  $f$  is continuous.

*Proof.* As  $t \rightarrow x$ , we get  $|f(t) - f(x)| = \left| \frac{f(t) - f(x)}{t - x} \right| |t - x| \rightarrow f'(x) 0 = 0$ , thus  $f(t) \rightarrow f(x)$ . ■

**Theorem 6.1.2.** Suppose  $f, g : [a, b] \rightarrow \mathbb{R}$  are realvalued functiond differentiable at a point  $x \in (a, b)$ . Then  $f + g$ ,  $fg$ , and  $\frac{f}{g}$  are differentiable at  $x$ , and as  $t \rightarrow x$ :

$$[label=(0)](f + g)' = f' + g'. \quad (fg)' = f'g + fg'. \quad \left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}, \text{ provided that } g(x) \neq 0.$$

**2. Proof.** (1) follows directly from the definiton. Now notice that  $fg(t) - fg(x) = f(t)(g(t) - g(x)) + g(t)(f(t) - f(x))$ , then dividing by  $t - x$ , the result follows by definition.

Now also notice that  $\frac{\frac{f}{g}(t) - \frac{f}{g}(x)}{t - x} = \frac{1}{g(t)g(x)}(g(x)\frac{f(t) - f(x)}{t - x} - f(x)\frac{g(t) - g(x)}{t - x})$ , and the result again follows by definition. ■

**Example 6.1.** [label=(0)]

1. The derivative of constant functions are always 0, and the derivative of the identity function is always 1.
2. Let  $f(x) = x^n$ , for  $n \in \mathbb{Z}$ , and  $x \neq 0$  for  $n < 0$ , then  $f$  is differentiable and  $f'(x) = nx^{n-1}$ .
3. Polynomial functions are differentiable, and so are rational functions  $\frac{p}{q}$ , provided that  $q \neq 0$ .

**Theorem 6.1.3** (Caratheodory's Theorem). *Let  $f : [a, b] \rightarrow \mathbb{R}$  be a continuous realvalued function. Then  $f$  is differentiable at a point  $x \in (a, b)$  if and only if there is a continuous function  $\phi : (a, b) \rightarrow \mathbb{R}$  such that  $f(t) - f(x) = \phi(t)(t - x)$ ; moreover,  $\phi = f'$ .*

*Proof.* Suppose  $f'$  exists at  $x$ , and define  $\phi : (a, b) \rightarrow \mathbb{R}$  by  $\phi(t) = \frac{f(t)-f(x)}{t-x}$  when  $t \neq x$ , and  $\phi(t) = f'(x)$  at  $t = x$ . Then by the continuity of  $f$ ,  $\phi$  is continuous at  $x$ , moreover, at  $t \neq x$  we see that  $f(t) - f(x) = \phi(t)(t - x)$ .

Conveersesly, sup[ose there is a  $\phi$ , continuous at  $x$  such that  $f(t) - f(x) = \phi(t)(t - x)$ , then clearly,  $\lim \phi = f'(x)$  as  $t \rightarrow x$ , and since  $\phi$  is continuous,  $\phi(x) = f'(x)$ . ■

**Theorem 6.1.4** (The Chain Rule). *Suppose that  $f : [a, b] \rightarrow \mathbb{R}$  and  $g : I \rightarrow \mathbb{R}$  are continuous, where  $f([a, b]) \subseteq I \subseteq [a, b]$ , and suppose that  $f$  is differentiable at  $x$ , and that  $g$  is differentiable at  $f(x)$ . Then  $g \circ f$  is differentiable at  $x$ , and  $(g \circ f)' = (g' \circ f)f'$ .*

*Proof.* We have by Caratheodory's theorem that  $f(t) - f(x) = (t - x)(f'(x) - u(t))$ , and  $g(s) - g(y) = (s - y)(g'(y) - v(s))$ . Then letting  $y = f(x)$ , and  $s \rightarrow y$  as  $t \rightarrow x$ , we see that  $u, v \rightarrow 0$ , and we get that  $g(f(t)) - g(f(x)) = g'(f(x))f(t) - g'(f(x))f(x)$ , dividing by  $t - x$  give the desired result. ■

**Example 6.2.** [label=(0)]

1. Let  $f(x) = \sin \frac{1}{x}$  at  $x \neq 0$ , and  $f(x) = 0$  at  $x = 0$ . We have at  $x \neq 0$ , that  $f'(x) = \sin \frac{1}{x} - \frac{1}{x} \cos \frac{1}{x}$ , but at  $x = 0$ , we must appeal to the definition, and we get  $f(t) = \sin \frac{1}{t}$ , which diverges at  $t \rightarrow 0$ , thus  $f'(0)$  does not exist.
2. Let  $f(x) = x^2 \sin \frac{1}{x}$  at  $x \neq 0$ , and  $f(x) = 0$  at  $x = 0$ . For  $x \neq 0$ , we get  $f'(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x}$ , and at  $x = 0$ , we notice that  $|t \sin \frac{1}{t}| \leq |t|$ , so by the sandwich theorem,  $f'(0) = 0$  as  $t \rightarrow 0$ .

## 6.2 Mean Value Theorems.

**Definition.** Let  $f : X \rightarrow \mathbb{R}$  be defined on a metric space  $X$ . We say that  $f$  has a **local maximum** at a point  $p \in X$ , if there is a  $\delta > 0$  for which  $f(q) \leq f(p)$  whenever  $d(q, p) < \delta$ . Likewise  $f$  has a **local minimum** at a point  $p \in X$ , if there is a  $\delta > 0$  for which  $f(q) \geq f(p)$  whenever  $d(q, p) < \delta$ . We call local maxima and local minumums **local extrema**.

**Theorem 6.2.1.** *Let  $f : [a, b] \rightarrow \mathbb{R}$  be a realvalued function, and suppose that  $f$  has a local extremum at  $x \in (a, b)$ . If  $f'$  exists, then  $f'(x) = 0$ .*

*Proof.* Suppose, without loss of generality that  $f$  has a local maximum at  $x$ . Choosse  $\delta > 0$  such that  $a < x - \delta < x < x + \delta < b$ . Then if  $x - \delta < t < x$ , we have  $|t - x + \delta| < \delta$ , so  $f(t) \leq f(x)$ , thus  $\frac{f(t)-f(x)}{t-x} \leq 0$ . Similarly, for  $x < t < x + \delta$ , we get  $\frac{f(t)-f(x)}{t-x} \geq 0$ , hence, as  $t \rightarrow x$ , we get  $0 \leq f'(0) \leq 0$ , thus  $f'(x) = 0$ . ■

**Theorem 6.2.2** (The Generalized Mean Value Theorem). *If  $f, g : [a, b] \rightarrow \mathbb{R}$  are continuous on  $[a, b]$ , and differentiable on  $(a, b)$ , then there is a point  $x \in (a, b)$  such that  $(f(b) - f(a))g'(x) = (g(b) - g(a))f'(x)$ .*

*Proof.* Let  $h(t) = (f(b) - f(a))g'(x) - (g(b) - g(a))f'(x)$ , for  $t \in [a, b]$ , then  $h$  is continuous on  $[a, b]$ , and differentiable on  $(a, b)$ , moreover, we have  $h(b) = f(b)g(a) - f(a)g(b) = h(a)$ . Now if  $h$  is constant, then  $h' = 0$  for all  $t$  and we are done. Now suppose that  $h(a) < h(b)$ , and let  $x \in (a, b)$ , be a local minimum of  $h$ , then  $h'(x) = 0$ , and we are done; the same result follows for local maxima of  $h$ . ■

**Corollary** (The Mean Value Theorem). *Let  $f : [a, b] \rightarrow \mathbb{R}$  be continuous on  $[a, b]$ , and differentiable on  $(a, b)$ . Then there is an  $x \in (a, b)$  such that  $f(b) - f(a) = (b - a)f'(x)$ .*

*Proof.* Take  $g(t) = t$ . ■

**Theorem 6.2.3.** *Suppose that  $f : [a, b] \rightarrow \mathbb{R}$  is differentiable on  $(a, b)$ . Then the following hold for all  $x \in (a, b)$ :*

*[label=(0)]If  $f' \geq 0$ , then  $f$  is monotonically increasing. If  $f' = 0$ , then  $f$  is constant. If  $f' \leq 0$ , then  $f$  is monotonically decreasing.*

**2.** *Proof.* Let  $x_1, x_2 \in (a, b)$ , then by the mean value theorem, there is an  $x \in (x_1, x_2)$  such that  $f(x_2) - f(x_1) = (x_2 - x_1)f'(x)$ . Then if  $f'(x) = 0$ , we get  $f(x_2) = f(x_1)$ , and that  $f$  is constant. If  $f'(x) \geq 0$ , we get  $f(x_2) \geq f(x_1)$ , making  $f$  monotonically increasing, similarly, if  $f'(x) \leq 0$ , we get  $f$  monotonically decreasing. ■

## 6.3 The Continuity of Derivatives.

**Theorem 6.3.1.** *Let  $f : [a, b] \rightarrow \mathbb{R}$  be differentiable on all of  $[a, b]$ , and suppose that  $f'(a) < \lambda < f'(b)$ . Then there is an  $x \in (a, b)$  such that  $f'(x) = \lambda$ .*

*Proof.* Let  $g(t) = f(t) - \lambda t$ , then  $g'(a) < 0$  and  $g'(b) > 0$ . Then for  $t_1, t_2 \in (a, b)$ ,  $g(t_1) < g(a)$ , and  $g(b) < g(t_2)$ . Then by the extreme value theorem,  $g$  attains a maximum at a point  $x \in (t_1, t_2)$ , hence  $g'(x) = 0$ , hence  $f'(x) = \lambda$ . ■

**Corollary.** *If  $f : [a, b] \rightarrow \mathbb{R}$  is differentiable, then  $f$  cannot have any removable discontinuities, nor jump discontinuities.*

*Remark.*  $f'$  may have infinite discontinuities.

## 6.4 L'Hospital's Rule.

**Theorem 6.4.1** (L'Hospital's Rule). *Suppose  $f$  and  $g$  are realvalued functions differentiable on  $(a, b)$ , and that  $g' \neq 0$  for all  $x \in (a, b)$ , where  $-\infty \leq a < b \leq \infty$ , and suppose that  $\frac{f'}{g'} \rightarrow A$  as  $x \rightarrow a$ . If  $f, g \rightarrow 0$ , or if  $g \rightarrow \pm\infty$ , as  $x \rightarrow a$ , then  $\frac{f}{g} \rightarrow A$  as  $x \rightarrow a$ .*

*Proof.* Suppose first that  $-\infty \leq A < \infty$ , and choose  $q, r \in \mathbb{R}$  such that  $A < r < q$ . By hypothesis, there is a  $c \in (a, b)$  for which  $a, x < c$  implies  $\frac{f}{g} < r$ . If  $a < x < y < c$ , then by the generalized mean value theorem,  $\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(t)}{g'(t)} < r$ , thus letting  $x \rightarrow a$ , we see that  $\frac{f(y)}{g(y)} \leq r < q$ . Now suppose, without loss of generality, that  $g \rightarrow \infty$ . Fixing  $y$ , and choosing

$c_1 \in (a, y)$  such that  $g(x) > g(y)$ , and  $g(x) > 0$ , if  $a < x < c_1$ , then  $\frac{f(x)}{g(x)} < r - r \frac{g(y)+f(y)}{g(x)}$ , then as  $x \rightarrow a$ , there is a  $c_2 \in (x, c_1)$  such that  $\frac{f}{g} < q$ .

Likewise, if we suppose that  $-\infty < A \leq \infty$ , by the same reasoning, we can choose a  $p < A$  and  $c_3 \in (a, b)$  such that  $p < \frac{f}{g}$  as  $x \rightarrow a$ . Since  $p < A < q$ , by the sandwich theorem, we get  $\frac{f}{g} = A$  as  $x \rightarrow a$ . ■

## 6.5 Taylor's Theorem.

**Definition.** If  $f$  has a derivative  $f'$  on an interval, and  $f'$  is differentiable, we denote  $f''$  to be  $(f')'$  and call it the **second derivative** of  $f$ ; likewise, if  $f''$  is differentiable, we denote the **third derivative** by  $f^{(3)} = (f'')'$ . More generally, for  $n \in \mathbb{Z}^+$ , we define recursively the  **$n$ th derivative** to be:

[label=(0)]

1.  $f^{(0)} = f$  and  $f^{(1)} = f'$ .
2.  $f^{(n+1)} = (f^{(n)})'$ , given that  $f^{(n)}$  is differentiable.

We call  $f$   **$n$ th differentiable** if  $f^{(n)}$  exists.

**Theorem 6.5.1** (Taylor's Theorem). *Suppose  $f : [a, b] \rightarrow \mathbb{R}$  is a realvalued function, that is  $n$ th differentiable, and let  $n \in \mathbb{Z}^+$  be such that  $f^{(n-1)}$  is continuous on  $[a, b]$ , and that  $f^{(n)}$  exists on  $(a, b)$ . Let  $\alpha, \beta \in [a, b]$  be distinct, and define:*

$$p(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!} (t - \alpha)^k \quad (6.2)$$

*Then there exists a point  $x \in (\alpha, \beta)$  such that  $f(\beta) = p(\beta) + \frac{f^{(n)}(x)}{n!} (\beta - \alpha)^n$ .*

*Proof.* For  $n = 1$ , this reduces to the mean value theorem, so suppose that  $n > 1$ . Let  $M \in \mathbb{R}$  be such that  $f(\beta) = p(\beta) + M(\beta - \alpha)^n$ , and let  $g(t) = f(t) - p(t) + M(\beta - \alpha)^n$ , for  $t \in [a, b]$ . Then  $g$  is  $n$ th differentiable, and we get  $g^{(n)} = f^{(n)} - n!M$  for  $t \in (a, b)$ . We wish to show that  $f^{(n)} = n!M$ .

We have that  $p^{(k)} = f^{(k)}(\alpha)$  for  $0 \leq k \leq n-1$ , then  $g(\alpha) = g'(\alpha) = \dots = g^{(n-1)}(\alpha) = 0$ , and our choice of  $M$  shows that  $g(\beta) = 0$ . So  $g'(x_1) = 0$  for  $x_1 \in (\alpha, \beta)$ , so by the mean value theorem, since  $g'(\alpha) = 0$ , then  $g''(\xi_2) = 0$  for  $x_2 \in (\alpha, x_1)$ . Proceeding inductively, we then get that  $g^{(n)}(x_n) = 0$  for  $x_n \in (\alpha, x_{n-1})$ , hence we get that  $n!M = f^{(n)}(x)$ . ■

**Definition.** We call the series in equation (5.2) the **Taylor series** (or **Taylor polynomial**) of  $f$  about  $\alpha$ . We call the realnumber  $M$  such that  $n!M = f^{(n)}(x)$  the **tail**, (or **error**) of the Taylor series.

## 6.6 Derivatives of vector valued functions.

**Definition.** Let  $f : [a, b] \rightarrow \mathbb{C}$  be a complex valued function, such that  $f(t) = f_1(t) + if_2(t)$ . We say that  $f$  is **differentiable** at a point  $x$  if and only if  $f_1$  and  $f_2$  are differentiable, and we denote the **derivative** of  $f$  to be the function  $f : (a, b) \rightarrow \mathbb{C}$  such that  $f' = f'_1 + if'_2$

**Definition.** Let  $f : [a, b] \rightarrow \mathbb{R}^k$  be a vectorvalued function for  $k \in \mathbb{Z}^+$ .  $f$  is said to be **differentiable** at  $x \in (a, b)$  if there is some point  $f'(x) \in \mathbb{R}^k$  such that:

$$\lim_{t \rightarrow x} \left\| \frac{f(t) - f(x)}{t - x} - f'(x) \right\| = 0 \quad (6.3)$$

We define the **derivative** of  $f$  at  $x$  to be the function  $f' : (a, b) \rightarrow \mathbb{R}^k$  such that the values of  $f'$  satisfy equation (6.3)

*Remark.* If  $f : [a, b] \rightarrow \mathbb{R}^k$  is defined by  $f = (f_1, \dots, f_k)$ , then  $f$  is differentiable at a point  $x \in (a, b)$  if and only if  $f_i$  is differentiable at  $x$  for  $1 \leq i \leq k$ , and we have that  $f' = (f'_1, \dots, f'_k)$ .

Theorem 6.1.1 follows naturally, and so does theorem 6.1.2(a) and (2), where we define  $fg$  as  $\langle f, g \rangle$ , however, the mean value theorem in general does not hold.

**Example 6.3.** [label=(0)]

1. Define  $f : \mathbb{R} \rightarrow \mathbb{C}$  by  $f(x) = e^{ix} = \cos x + i \sin x$ . Then  $f(2\pi) - f(0) = 0$ , however,  $f'(x) = ie^{ix} \neq 0$  for all  $x$  (moreover,  $|f'| = 1$ ), so the generalized mean value theorem fails here.
2. Define  $f, g : (0, 1) \rightarrow \mathbb{C}$  by  $f(x) = x$  and  $g(x) = x + x^2 e^{\frac{i}{x^2}}$  for all  $x$ . Since  $|e^{it}| = 1$ , we have that  $\lim_{x \rightarrow 0} \frac{f}{g} = 1$  as  $x \rightarrow 0$ . Now  $g'(x) = 1 + (2x - i\frac{2}{x})e^{\frac{i}{x^2}}$  on  $(0, 1)$ , hence  $|g'| = |2x - i\frac{2}{x}| - 1 \geq \frac{2}{x} - 1$ , so  $|\frac{f'}{g'}| \leq \frac{x}{2-x} \rightarrow 0$  as  $x \rightarrow 0$ , so L'Hospital's rule fails in  $\mathbb{C}$  as well, and hence in  $\mathbb{R}^2$  (as  $\mathbb{C}$  is isomorphic to  $\mathbb{R}^2$ ).

**Theorem 6.6.1.** Suppose  $f : [a, b] \rightarrow \mathbb{R}^k$ , for  $k \in \mathbb{Z}^+$  is continuous, and that  $f$  is differentiable on  $(a, b)$ . Then there is an  $x \in (a, b)$  for which  $\|f(b) - f(a)\| \leq (b - a)\|f'(x)\|$ .

*Proof.* Let  $z = f(b) - f(a)$ , and define  $\phi = \langle f, g \rangle$  for all  $t \in [a, b]$ , then  $\phi$  is a realvalued function continuous on  $[a, b]$ , moreover it is differentiable on  $(a, b)$ ; therefore, by the mean value theorem,  $\phi(b) - \phi(a) = (b - a)\phi'(a) = (b - a)\langle z, f'(x) \rangle$  for  $x \in (a, b)$ . On the other hand, we have that  $\phi(b) - \phi(a) = \langle z, z \rangle = \|z\|^2$ , hence, by the Cauchy Schwarz inequality, we have that  $\|z\|^2 = (b - a)\langle z, f' \rangle \leq \|z\|\|f'\|$ , which gives the desired result. ■



# Chapter 7

## Integration

### 7.1 The Riemann-Stieltjes Integral.

**Definition.** Let  $[a, b]$  be an interval. A **partition** of  $[a, b]$  is a set of points  $P = \{x_0, x_1, \dots, x_n\}$  such that  $a = x_0 < x_1 < \dots < x_n = b$ , and we write  $\Delta x_i = x_i - x_{i-1}$ . Now let  $f : [a, b] \rightarrow \mathbb{R}$  be a bounded realvalued function, and for each partition  $P$  of  $[a, b]$  let  $M_i = \sup f$  and  $m_i = \inf f$  for all  $x_{i-1} \leq x \leq x_i$ . We define the **upper Riemann sum** and the **lower Riemann sum** to of  $f$  with respect to be:

$$U(f, P) = \sum_{i=1}^n M_i \Delta x_i \quad (7.1)$$

$$L(f, P) = \sum_{i=1}^n m_i \Delta x_i \quad (7.2)$$

respectively. We also define the **upper Riemann integral** and the **lower Riemann integral** of  $f$  over  $[a, b]$  to be:

$$\overline{\int_a^b} f(x) dx = \inf U(f, P) \quad (7.3)$$

$$\underline{\int_a^b} f(x) dx = \sup L(f, P) \quad (7.4)$$

Respectively.

If  $\overline{\int_a^b} f = \underline{\int_a^b} f$ , then we say that  $f$  is **Riemann integrable** on  $[a, b]$ , and we its value the **Riemann integral**, and denote it to be:

$$\int_a^b f(x) dx = \overline{\int_a^b} f(x) dx = \underline{\int_a^b} f(x) dx \quad (7.5)$$

**Lemma 7.1.1.**  $\overline{\int_a^b} f$ , and  $\underline{\int_a^b} f$  are defined for every bounded realvalued function  $f$  over  $[a, b]$ .

*Proof.* Let  $f$  be bounded on  $[a, b]$ , then there are  $m$  and  $M$  such that  $m \leq f \leq M$  for all  $a \leq x \leq b$ . Now let  $P$  be a partition of  $[a, b]$ . Since  $\inf f \leq \sup f$ , we have that  $m \leq m_i = \inf f \leq M_i = \sup f \leq M$ , thus  $m(b-a) \leq L(f, P) \leq U(f, P) \leq M(b-a)$ , hence  $L$  and  $U$  form a bounded set, and we are done. ■

**Corollary.**  $L(f, P) \leq U(f, P)$  for every bounded function  $f$ .

Now the question of the integrability of  $f$  is a very delicate matter, and requires a closer scrutiny on the concepts of upper and lower sums. Infact, it turns out that the Riemann integral is a consequence of a more general class of integrals. Developng this more general situation will allow us to discern facts about the Riemann integral.

**Definition.** Let  $\alpha$  be a bounded monotonically increasing function on  $[a, b]$ , and let  $P$  be a partition of  $[a, b]$  and let  $\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1})$ . For any realvalued, bounded function on  $[a, b]$ , defined the **upper sum** and the **lower sum** of  $f$  with respect to  $P$  and  $\alpha$  to be:

$$U(f, P, \alpha) = \sum_{i=1}^n M_i \Delta\alpha_i \quad (7.6)$$

$$L(f, P, \alpha) = \sum_{i=1}^n m_i \Delta\alpha_i \quad (7.7)$$

Where  $M_i = \sup f$  and  $m_i = \inf f$  for all  $x_{i-1} \leq x \leq x_i$ , and again, define the **upper integral** and **lower integral** of  $f$  with respect to  $\alpha$  on  $[a, b]$  to be:

$$\overline{\int_a^b} f(x) d\alpha = \inf U(f, P, \alpha) \quad (7.8)$$

$$\underline{\int_a^b} f(x) d\alpha = \sup L(f, P, \alpha) \quad (7.9)$$

If  $\overline{\int_a^b} f d\alpha = \underline{\int_a^b} f d\alpha$ , we call the value:

$$\int_a^b f(x) d\alpha = \overline{\int_a^b} f(x) d\alpha = \underline{\int_a^b} f(x) d\alpha \quad (7.10)$$

the **Riemann-Stieltjes integral** of  $f$  with respect to  $\alpha$  on  $[a, b]$ . If such an integra exists, we say that  $f$  is **integrable** with respect to  $\alpha$  on  $[a, b]$ .

**Example 7.1.** Let  $\alpha(x) = x$ , be defined over  $[a, b]$ . Then  $\alpha$  is monotonically increasing, and our definitions reduces to those for the Riemann integral. Here  $U(f, P, x) = U(f, P)$  and  $L(f, P, x) = L(f, P)$ .

We are now in a position to investigate the properties of integrability, in the Riemann-Stieltjes sense.

**Definition.** Let  $[a, b]$  be an interval, and let  $P$  and  $Q$  be partitions of  $[a, b]$ . We say that  $Q$  is a **refinement** of  $P$  if  $PQ$ , and we also say that  $Q$  is **finer** than  $P$ . Now if neither  $P$  nor  $Q$  is a refinement of the other, we say that the two partitions are **noncomparable**.



**Lemma 7.1.2.** *Let  $P$  and  $Q$  be partitions of an interval  $[a, b]$ , then  $P \cup Q$  is a partition of  $[a, b]$ , and is a refinement of both  $P$  and  $Q$ .*

*Proof.* If  $P$  is a refinement of  $Q$ , or viceversa, then we are done; so suppose that  $P$  and  $Q$  are noncomparable. Let  $P = \{x_0, x_1, \dots, x_n\}$  and  $Q = \{y_0, y_1, \dots, y_m\}$  with  $a = x_0 < x_1 < \dots < x_n = b$  and  $a = y_0 < y_1 < \dots < y_m = b$ . Then  $P \cup Q = \{x_0, y_0, x_1, y_1, \dots, x_n, y_m\}$  and  $a = x_0 = y_0 < x_1, y_1 < \dots < x_n = y_m = b$ , thus  $P \cup Q$  is a partition of  $[a, b]$ , that it is a refinement of  $P$  and  $Q$  follows trivially. ■

**Theorem 7.1.3.** *Let  $\alpha$  be monotonically increasing, and bounded on  $[a, b]$ , and let  $P$  and  $Q$  be partitions of  $[a, b]$ . If  $Q$  is a refinement of  $P$ , then  $L(f, P, \alpha) \leq L(f, Q, \alpha)$  and  $U(f, Q, \alpha) \leq U(f, P, \alpha)$ .*

*Proof.* Let  $Q = P \cup \{x'\}$  and suppose that  $x_{i-1} \leq x' \leq x_i$ . Let  $w_1 = \inf f$  for  $x_{i-1} \leq x \leq x'$  and let  $w_2 = \inf f$  for  $x' \leq x \leq x_i$ . Then  $m_i \leq w_1, w_2$ , thus  $L(f, Q, \alpha) - L(f, P, \alpha) = (w_1 - m_i)(\alpha(b) - \alpha(a)) - (w_2 - m_i)(\alpha(b) - \alpha(a)) \geq 0$ , we are done. The proof is analogous for  $U$ . ■

**Corollary.**  *$L(f, P, \alpha)$  is monotonically increasing and  $U(f, P, \alpha)$  is monotonically decreasing.*

*Proof.* We note that if  $Q$  is a refinement of  $P$ , then  $|P| \leq |Q|$ , the result follows by direct application. ■

*Remark.* If  $Q$  contains  $k$  more points than  $P$ , we can repeat the proof inductively.

**Theorem 7.1.4.**  $\int f d\alpha \leq \bar{\int} f d\alpha$ .

*Proof.* Let  $P = P_1 \cup P_2$  for partitions  $P_1$  and  $P_2$  of  $[a, b]$ . By theorem 7.1.3 and lemma 7.1.1, we have:

$$L(f, P_1, \alpha) \leq L(f, P_2, \alpha) \leq U(f, P_2, \alpha) \leq U(f, P_1, \alpha) \quad (7.11)$$

Fixing  $P_2$  and taking the supremum over all  $P_1$ , we get  $\int f \leq U(f, P_2, \alpha)$ , the infimum over  $P_2$  we get  $\int f \leq \bar{\int} f$  ■

**Theorem 7.1.5.** *A realvalued function  $f$  is integrable over an interval  $[a, b]$  if and only if for  $\varepsilon > 0$ , there is a partition  $P$  such that:*

$$U(f, P, \alpha) - L(f, P, \alpha) < \varepsilon \quad (7.12)$$

*Proof.* For every  $P$ , we have that  $L(f, P, \alpha) \leq \int f \leq \bar{\int} f \leq U(f, P, \alpha)$ , so if for  $\varepsilon > 0$ , we assume that  $U(f, P, \alpha) - L(f, P, \alpha) < \varepsilon$ , then we get that:

$$0 \leq \bar{\int}_a^b f d\alpha - \int_a^b f d\alpha < \varepsilon \quad (7.13)$$

implying integrability for small enough  $\varepsilon$ .

Conversely, suppose that  $\bar{\int} f = \int f$ . Let  $\varepsilon > 0$ . Then there are partitions  $P_1$  and  $P_2$  such that  $U(f, P_2, \alpha) - \int f < \frac{\varepsilon}{2}$  and  $\int f - L(f, P_1, \alpha) < \frac{\varepsilon}{2}$ . Take  $P = P_1 \cup P_2$ , then by theorem 7.1.4, adding the inequalities we get  $U(f, P, \alpha) - L(f, P, \alpha) < \varepsilon$ . ■

**Theorem 7.1.6.** *The following hold:*

*[label=(0)]If  $U(f, P, \alpha) - L(f, P, \alpha) < \varepsilon$  for some  $\varepsilon > 0$  and some  $P$ , then it holds, with the same  $\varepsilon$ , for every refinement of  $P$ . If  $U(f, P, \alpha) - L(f, P, \alpha) < \varepsilon$  for  $P = \{x_0, \dots, x_n\}$ , and if  $s_i, t_i \in [x_{i-1}, x_i]$ , then:*

$$\sum_{i=1}^n |f(s_i) - f(t_i)| \Delta \alpha_i < \varepsilon \quad (7.14)$$

*If  $f$  is integrable with respect to  $\alpha$  on  $[a, b]$ , and (2) holds, then:*

$$\left| \sum_{i=1}^n f(t_i) \Delta \alpha_i - \int_a^b f d\alpha \right| < \varepsilon \quad (7.15)$$

**2. Proof.** [label=(0)]

Theorem 7.1.3 implies this.

2. If  $U(f, P, \alpha) - L(f, P, \alpha) < \varepsilon$ , we have that  $f(s_i), f(t_i) \in [m_i, M_i]$ , hence  $|f(s_i) - f(t_i)| \leq M_i - m_i$ , thus we get that  $\sum |f(s_i) - f(t_i)| \Delta \alpha_i \leq U(f, P, \alpha) - L(f, P, \alpha) < \varepsilon$
3. We have  $L(f, P, \alpha) \leq \sum f(t_i) \Delta \alpha_i \leq U(f, P, \alpha)$ , and  $L(f, P, \alpha) \leq \int f d\alpha_i \leq U(f, P, \alpha)$ , taking differences, we get the following. ■

**Theorem 7.1.7.** *If  $f$  is continuous on  $[a, b]$ , then  $f$  is integrable with respect to  $\alpha$  on  $[a, b]$ .*

*Proof.* Let  $\varepsilon > 0$ , and choose  $\xi > 0$  such that  $(\alpha(b) - \alpha(a))\xi < \varepsilon$ . Now we have that  $f$  is uniformly continuous on  $[a, b]$ , so there is a  $\delta > 0$  such that  $|f(x) - f(t)| < \xi$  whenever  $|x - t| < \delta$  for  $x, t \in [a, b]$ . Now since  $P$  is a partition of  $[a, b]$ , with  $\Delta x_i < \delta$  for all  $i$ , then we have  $M_i - m_i \leq \xi$  for  $i = 1, \dots, n$  thus  $U(f, P, \alpha) - L(f, P, \alpha) = \sum (M_i - m_i) \Delta \alpha_i \leq \xi(\alpha(b) - \alpha(a)) < \varepsilon$ , therefore,  $f$  is integrable. ■

**Theorem 7.1.8.** *If  $f$  is monotonic on  $[a, b]$ , and  $\alpha$  is monotonic and continuous on  $[a, b]$ , then  $f$  is integrable with respect to  $\alpha$  on  $[a, b]$ .*

*Proof.* Let  $\varepsilon > 0$ , and for  $n > 0$ , construct a partition  $P$  such that  $\Delta \alpha_i = \frac{\alpha(b) - \alpha(a)}{n}$  for  $1 \leq i \leq n$ , which is guaranteed by the intermediate value theorem.

Now suppose without loss of generality that  $f$  is monoton increasing, then  $M_f(x_i)$  and  $m_i = f(x_{i-1})$  by the extreme value theorem, Then:

$$U(f, P, \alpha) - L(f, P, \alpha) = \frac{\alpha(b) - \alpha(a)}{n} \sum f(x_i) - f(x_{i-1}) = \frac{\alpha(b) - \alpha(a)}{n} (f(b) - f(a)) < \varepsilon$$

Taking  $n$  large enough implies integrability. ■

**Theorem 7.1.9.** *Suppose that  $f$  is bounded on  $[a, b]$ , with only finitely many discontinuities on  $[a, b]$ . Suppose that  $\alpha$  is continuous at every discontinuity of  $f$ , then  $f$  is integrable with respect to  $\alpha$  on  $[a, b]$ .*

*Proof.* Let  $\varepsilon > 0$  and  $M = \sup |f|$  and let  $E$  be the set of all discontinuities of  $f$ . Since  $E$  is finite, and  $\alpha$  is continuous on  $E$ , then  $E$  can be covered by finitely many disjoint intervals  $[u_j, v_j] \subseteq [a, b]$  such that  $\sum \alpha(u_j) - \alpha(v_j) < \varepsilon$ .

Now suppose that we construct these intervals in such a way that every point of  $E \cap [a, b]$  lies in  $(u_j, v_j)$ , for some  $j$ . Taking  $K = [a, b] \setminus (u_j, v_j)$ ,  $K$  is compact, so  $f$  is uniformly continuous on  $K$ . Thus there is a  $\delta > 0$  such that  $|f(s) - f(t)| < \varepsilon$  whenever  $|s - t| < \delta$  for  $s, t \in K$ .

Now construct the partition  $P$  of  $[a, b]$  such that  $u_j, v_j \in P$ , but  $(u_j, v_j) \not\subseteq P$  for some  $j$ . If  $x_i \neq u_j$ , then  $\Delta x_i < \delta$ . Notice that  $M_i - m_i \leq 2M$  for all  $i$  and that  $M_i - m_i \leq \varepsilon$ . Thus for  $\varepsilon$  small enough, we have  $U(f, P, \alpha) - L(f, P, \alpha) \leq (\alpha(b) - \alpha(a))\varepsilon + 2M\varepsilon$ , thus  $f$  is integrable. ■

**Theorem 7.1.10.** *Suppose that  $f$  is integrable with respect to  $\alpha$  on an interval  $[a, b]$ , and that  $m \leq f \leq M$ , and let  $\phi$  be continuous on  $[m, M]$ , and let  $h = \phi \circ f$  be defined on  $[a, b]$ . Then  $h$  is integrable with respect to  $\alpha$  on  $[a, b]$ .*

*Proof.* Let  $\varepsilon > 0$ . We have that  $\phi$  is uniformly continuous on  $[m, M]$ , so there is a  $\delta > 0$  for which  $|\phi(s) - \phi(t)| < \varepsilon$  whenever  $|s - t| < \delta$ , for  $s, t \in [m, M]$ .

Now since  $f$  is integrable, there is a partition  $P$  of  $[a, b]$  such that  $U(f, P, \alpha) - L(f, P, \alpha) < \delta^2$ . Let  $M_i = \sup f$ ,  $m_i = \inf f$  and  $M'_i = \sup h$ ,  $m'_i = \inf h$  for all  $x_{i-1} \leq x \leq x_i$ . Now divide  $\{1, \dots, n\}$  into two sets  $A$  and  $B$  such that  $i \in A$  if  $M_i - m_i < \delta$  and  $i \in B$  if  $M_i - m_i > \delta$  (then  $A$  and  $B$  are disjoint). Then for  $i \in A$ , we get  $M'_i - m'_i \leq \varepsilon$ , and for  $i \in B$ , we get  $M'_i - m'_i \leq 2K$  with  $K = \sup |\phi|$  for  $m \leq t \leq M$ .

$$\delta \sum_{i \in B} \Delta \alpha_i \leq \sum_{i \in B} (M_i - m_i) \Delta \alpha_i < \delta^2$$

so  $\sum \Delta \alpha_i < \delta$ . Then

$$\begin{aligned} U(h, P, \alpha) - L(h, P, \alpha) &= \sum_{i \in A} (M'_i - m'_i) \Delta \alpha_i + \sum_{i \in B} (M'_i - m'_i) \Delta \alpha_i \\ &\leq \varepsilon(\alpha(b) - \alpha(a)) + 2K\delta \\ &< \varepsilon(\alpha(b) - \alpha(a) + 2K) \end{aligned}$$

Then for  $\varepsilon$  small enough, we get that  $h$  is integrable. ■

## 7.2 Properties of the Integral.

**Theorem 7.2.1.** *Let  $f$  and  $g$  be realvalued functions integrable with respect to  $\alpha$  over an interval  $[a, b]$ . Then:*

[label=(0)]

1.  $f + g$  and  $cf$

$$\int_a^b f + g d\alpha = \int_a^b f d\alpha + \int_a^b g d\alpha \quad (7.16)$$

2. If  $f \leq g$  on  $[a, b]$ , then:

$$\int_a^b f d\alpha \leq \int_a^b g d\alpha \quad (7.17)$$

3. If  $c \in (a, b)$  and  $f$  then:

$$\int_a^c f d\alpha + \int_c^b f d\alpha = \int_a^b f d\alpha \quad (7.18)$$

4. If  $|f| \leq M$  on  $[a, b]$ , then:

$$\left| \int_a^c f d\alpha \right| \leq M(\alpha(b) - \alpha(a)) \quad (7.19)$$

5. If  $f$  is intergrable with respect to  $\alpha_1$ , and  $\alpha_2$  on  $[a, b]$ , and  $c > 0$ , then:

$$\int_a^b f d(\alpha_1 + \alpha_2) = \int_a^b f d\alpha_1 + \int_a^b f d\alpha_2 \quad (7.20)$$

*Proof.* [label=(0)]

Let  $P$  be a partition of  $[a, b]$ , then  $L(f, P, \alpha) + L(f, P, \alpha) \leq L(f + g, P, \alpha) \leq U(f + g, P, \alpha) \leq U(f, P, \alpha) + U(g, P, \alpha)$ . Now let  $\varepsilon > 0$ , there are partitions  $P_1, P_2$ , with  $P = P_1 \cup P_2$  such that  $U(f, P_1, \alpha) - L(f, P_1, \alpha) < \frac{\varepsilon}{2}$ , and  $U(g, P_2, \alpha) - L(g, P_2, \alpha) < \frac{\varepsilon}{2}$ . Then we get  $U(f + g, P, \alpha) - L(f + g, P, \alpha) < \varepsilon$ , so  $f + g$  is intergrable. Moreover, we have that  $U(f, P, \alpha) < \int f + \frac{\varepsilon}{2}$  which guarantees  $\int f + g \leq \int f + \int g$ . Doing the same with  $L$ , we get  $\int f + \int g \leq \int f + g$ , and we are done.

2. If  $f \leq g$ , then  $0 \leq g - f$ , and we have by part (1) that  $g - f$  is intergrable, and that  $0 \leq \int g - f = \int g - \int f$ , wich gives us the established inequality.

3. Let  $P_1$  be a ppartition of  $[a, c]$ , and let  $P_2$  be a partition of  $[c, b]$ , and let  $P = P_1 \cup P_2$  be a partition of  $[a, b]$ . Since  $f$  is intergrable on  $[a, b]$ , there is an  $\varepsilon > 0$  such that  $U(f, P, \alpha) - L(f, P, \alpha) < \varepsilon$ , by theorem 7.1.3, we have that  $L(f, P_1, \alpha) \leq L(f, P, \alpha) \leq U(f, P, \alpha) \leq U(f, P_1, \alpha)$  and  $L(f, P_2, \alpha) \leq L(f, P, \alpha) \leq U(f, P, \alpha) \leq U(f, P_2, \alpha)$ , thus we get  $U(f, P_1, \alpha) - L(f, P_2, \alpha) < \varepsilon$  and  $U(f, P_2, \alpha) - L(f, P, \alpha) < \varepsilon$ , thus  $f$  is integrable on  $[a, c]$  and on  $[c, b]$ .

Now we have that  $U(f, P_1, \alpha) < \int_a^c f + \frac{\varepsilon}{2}$  and  $\int_c^b f + \frac{\varepsilon}{2}$ , so  $\int_a^b f \leq \int_a^c f + \int_c^b f$ . With  $L$  we get the reverse inequality, and we are done.

4. We have that  $-M \leq f \leq M$ , now let  $P$  be a partition of  $[a, b]$ , and let  $\varepsilon > 0$ . Since  $f$  is integrable, we have that

$$-M(\alpha(b) - \alpha(a)) \leq U(f, P, \alpha) - L(f, P, \alpha) < \varepsilon \leq M(\alpha(b) - \alpha(a)) \quad (7.21)$$

which goves us for arbitrarily small  $\varepsilon$ ,  $|\int f| \leq M(\alpha(b) - \alpha(a))$ .

5. Let  $P_1$  and  $P_2$  be partitions of  $[a, b]$ , respective to  $\alpha_1$  and  $\alpha_2$ , and construct  $P = P_1 \cup P_2$  respective to  $\alpha_1 + \alpha_2$ . We have that  $U(f, P_1, \alpha_1) - L(f, P_1, \alpha_1) < \frac{\varepsilon}{2}$  and  $U(f, P_2, \alpha_2) - L(f, P_2, \alpha_2) < \frac{\varepsilon}{2}$ , then we have by the previous inequalities that:

$$U(f, P, \alpha_1 + \alpha_2) - L(f, P, \alpha_1 + \alpha_2) \leq (U(f, P_1, \alpha_1) + U(f, P_2, \alpha_2)) - (L(f, P_1, \alpha_1) + L(f, P_2, \alpha_2)) < \varepsilon$$

Thus,  $f$  is integrable with respect to  $\alpha_1 + \alpha_2$ . Now again, by the similar reasoning used in parts (1) and (3), we get that  $\int f d(\alpha_1 + \alpha_2) = \int f d\alpha_1 + \int f d\alpha_2$ . ■

**Corollary.**  $\int c f d\alpha = c \int f d\alpha$  and  $\int f d(c\alpha) = c \int f d\alpha$

**Theorem 7.2.2.** *If  $f$  and  $g$  are integrable with respect to  $\alpha$  on an interval  $[a, b]$ , then:*

*[label=(0)]  $fg$  is integrable with respect to  $\alpha$  on  $[a, b]$ .  $|f|$  is integrable with respect to  $\alpha$  on  $[a, b]$  and:*

$$\left| \int_a^b f d\alpha \right| \leq \int_a^b |f| d\alpha \quad (7.22)$$

2. *Proof.* [label=(0)]

Take  $\phi(t) = t^2$ , then by theorem 7.1.10,  $\phi \circ f = f^2$  is integrable; then notice that  $4fg = (f+g)^2 - (f-g)^2$  is integrable, hence, so is  $fg$ .

2. Take  $\phi(t) = |t|$ , then again by theorem 7.1.10,  $\phi \circ f = |f|$  is integrable. Furthermore, choose  $c = \pm 1$  such that  $0 \leq c \int f$ , since  $cf \leq |f|$ , we have that  $|\int f| = c \int f = \int cf \leq \int |f|$ . ■

**Definition.** We define the **unit step function**  $I : \mathbb{R} \rightarrow \{0, 1\}$  to be  $I(x) = 1$  if  $x > 0$  and  $I(x) = 0$  if  $x \leq 0$ .

**Theorem 7.2.3.** *If  $s \in (a, b)$ , and  $f$  is bounded on  $[a, b]$  and continuous at  $s$ , and if  $\alpha(x) = I(x - s)$ , then  $\int f = f(s)$ .*

*Proof.* Let  $P = \{x_0, x_1, x_2, x_3\}$  be a partition of  $[a, b]$ , such that  $a = x_0 < s = x_1 < x_2 < x_3 = b$ . Then  $U(f, P, \alpha) = M_2$  and  $L(f, P, \alpha) = m_2$ . Now since  $f$  is continuous,  $M_2, m_2 \rightarrow f(s)$  as  $x_2 \rightarrow s$ , and we are done. ■

**Theorem 7.2.4.** *Suppose that  $\alpha$  is monotonically increasing, and differentiable on an interval  $[a, b]$ , and suppose that  $\alpha'$  is Riemann integrable on  $[a, b]$ . Let  $f$  be a bounded realvalued function on  $[a, b]$ , then  $f$  is integrable with respect to  $\alpha$  on  $[a, b]$  if and only if  $f\alpha'$  is Riemann integrable.*

*Proof.* Let  $\varepsilon > 0$ , and let  $P$  be a partition of  $[a, b]$  such that  $U(\alpha', P) - L(\alpha', P) < \varepsilon$ . By the mean value theorem, there is a  $t_i \in [x_{i-1}, x_i]$  for which  $\Delta\alpha_i = \alpha'(t_i)\Delta x_i$ , for  $1 \leq i \leq n$ . Now if  $s_i \in [x_{i-1}, x_i]$ , then  $\sum |\alpha'(s_i) - \alpha'(t_i)|\Delta x_i < \varepsilon$ , and by theorem 7.1.6, with  $M = \sup |f|$ , since  $\sum f(s_i)\Delta\alpha_i = \sum f(s_i)\alpha'(t_i)\Delta x_i$ , we get

$$\left| \sum f(s_i)\Delta\alpha_i - \sum f(s_i)\alpha'(t_i)\Delta x_i \right| \leq M\varepsilon$$

in particular,  $\sum f(s_i)\Delta\alpha_i \leq U(f, \alpha', P) + M\varepsilon$ , hence we get  $U(f, P, \alpha) \leq U(f\alpha', P) + M\varepsilon$ , hence we have:

$$\left| \overline{\int} f d\alpha - \overline{\int} f\alpha' dx \right| \leq M\varepsilon$$

for small enough  $\varepsilon$ ,  $\overline{\int} f d\alpha = \overline{\int} f\alpha' dx$ . We get an analogous result for lower sums. ■

*Remark.* Taking  $\alpha$  to be a step function allows us to reduce the integral to be a finite or infinite series; similarly, if the derivative of  $\alpha$  (if it exists) is integrable, then the Riemann-Stielhes integral reduces to just the Riemann integral.

**Theorem 7.2.5** (Change of Variables). *Of  $\phi$  is a strictly increasing continuous function, mapping the interval  $[A, B]$  onto the interval  $[a, b]$ , and if  $\alpha$  is monotonically increasing on  $[a, b]$ , and if  $f$  is integrable with respect to  $\alpha$  on  $[a, b]$ , define  $\beta$  and  $g$  on  $[a, b]$  by  $\beta = \alpha \circ \phi$  and  $g = f \circ \phi$ , then  $g$  is integrable with respect to  $\beta$  on  $[A, B]$ , and*

$$\int_A^B g d\beta = \int_a^b f d\alpha \quad (7.23)$$

*Proof.* For each partition  $P$  of  $[a, b]$ , construct a partition  $Q$  of  $[A, B]$  such that  $x_i = \phi(y_i)$ . Then the values of  $f$  on  $[x_{i-1}, x_i]$  are the same as the values of  $g$  on  $[y_{i-1}, y_i]$ . Then  $U(g, Q, \beta) = U(f, P, \alpha)$  and  $L(g, Q, \beta) = L(f, P, \alpha)$ . It follows that since  $f$  is integrable, then  $U(g, Q, \beta) - L(g, Q, \beta) < \varepsilon$  for  $\varepsilon > 0$ , thus  $g$  is integrable; and the equality is established by the equality of their sums. ■

**Corollary.** *Let  $\alpha(x) = x$  and  $\beta = \phi$ , and suppose that  $\phi'$  is Riemann integrable on  $[A, B]$ . Then:*

$$\int_a^b f dx = \int_A^B (f \circ \phi) \phi' dy \quad (7.24)$$

## 7.3 Integration and Differentiation.

**Theorem 7.3.1.** *Let  $f$  be Riemann integrable on an interval  $[a, b]$ , and for  $a \leq x \leq b$ , let  $F(x) = \int_a^x f(t) dt$ , then  $F$  is continuous on  $[a, b]$ . Furthermore, if  $f$  is continuous at  $x_0 \in [a, b]$ , then  $F$  is differentiable at  $x_0$  and  $F'(x_0) = f(x_0)$ .*

*Proof.* We have that  $f$  is bounded, so for some  $a \leq t \leq b$ , we have  $|f| \leq M$ . Now if  $a \leq x < y \leq b$ , then  $|F(y) - F(x)| = |\int_x^y f(t) dt| \leq M(y - x)$ .

Now given  $\varepsilon > 0$ , we see that  $|F(y) - f(x)| < \varepsilon$ , whenever  $|y - x| < \frac{\varepsilon}{M}$ . Thus  $f$  is continuous at  $x_0$ . Let  $\varepsilon > 0$  and  $\delta > 0$  such that  $|f(t) - f(x_0)| < \varepsilon$  whenever  $|t - x_0| < \delta$  whenever  $a \leq t \leq b$ . Thus if  $x_0 - \delta < s \leq x_0 \leq t < x_0 + \delta$ , and  $a \leq s < t \leq b$ , then:

$$\left| \frac{F(s) - F(t)}{s - t} \right| = \left| \frac{\int_x^y f(u) - f(x_0) du}{s - t} \right| < \varepsilon$$

thus, we get  $F'(x_0) = f(x_0)$ . ■

**Theorem 7.3.2** (The Fundamental Theorem of Calculus). *If  $f$  is Riemann integrable on  $[a, b]$ , and if there is a differentiable function  $F$  on  $[a, b]$  such that  $F' = f$ , then:*

$$\int_a^b f(x) dx = F(b) - F(a) \quad (7.25)$$

*Proof.* Let  $\varepsilon > 0$ , and construct a partition  $P$  of  $[a, b]$  such that  $U(f, P) - L(f, P) < \varepsilon$ . Then by the mean value theorem, there is a  $t_i \in [x_{i-1}, x_i]$  such that  $F(x_i) - F(x_{i-1}) = f(t_i)\Delta x_i$ , for  $1 \leq i \leq n$ . Taking sums, we have

$$\sum f(t_i)\Delta x_i = F(b) - F(a) \quad (7.26)$$

thus, by theorem 7.1.6, we have:

$$\left| \int_a^b f(x)dx - (F(b) - F(a)) \right| < \varepsilon.$$

■

**Theorem 7.3.3** (Integration by Parts). *Let  $F$  and  $G$  be differentiable functions on  $[a, b]$ , and let  $F' = f$  and  $G' = g$  be Riemann integrable on  $[a, b]$ . Then:*

$$\int_a^b Fg(x)dx = FG(b) - FG(a) - \int_a^b fG(x)dx \quad (7.27)$$

*Proof.* By theorem 7.2.2,  $FG$  is Riemann integrable, and also notice that  $(FG)' = Fg + fG$ , thus by the fundamental theorem of calculus:

$$\int_a^b FGdx = \int_a^b Fgdx + \int_a^b fGdx = FG(b) - FG(a)$$

■

*Remark.* Integration by parts is what will allow us to find integrals of products of functions.

## 7.4 Integration of Vector Valued Functions.

We extend the concept of integrability to general Euclidean spaces. Let  $f_1, \dots, f_k$  be realvalued functions defined on an interval  $[a, b]$ , and let  $f = (f_1, \dots, f_k)$  be a vectorvalued function of  $[a, b]$  onto  $\mathbb{R}^k$ . If  $\alpha$  is a monotonically increasing function on  $[a, b]$ , we say that  $f$  is **integrable** with respect to  $\alpha$  of  $[a, b]$  if and only if  $f_i$  is integrable with respect to  $\alpha$  for  $1 \leq i \leq k$ . If this integrable exists, we define it to be:

$$\int_a^b f d\alpha = \left( \int_a^b f_1 d\alpha, \dots, \int_a^b f_k d\alpha \right). \quad (7.28)$$

The usual theorems for integrability extend to  $\mathbb{R}^k$ .

**Theorem 7.4.1** (The Fundamental Theorem of Calculus for Vector valued functions). *If  $f : [a, b] \rightarrow \mathbb{R}^k$  is a vector valued function that is Riemann integrable, and if  $F$  is a differentiable function such that  $F' = f$ , then:*

$$\int_a^b f dx = F(b) - f(a) \quad (7.29)$$

*Proof.* We apply the fundamental theorem of calculus to each of the components of  $f$ . ■

**Theorem 7.4.2.** *If  $f : [a, b] \rightarrow \mathbb{R}^k$  is integrable with respect to  $\alpha$  on an interval  $[a, b]$ , then  $\|f\|$  is integrable, with respect to  $\alpha$  on  $[a, b]$ , and:*

$$\left\| \int_a^b f d\alpha \right\| \leq \int_a^b \|f\| d\alpha$$

*Proof.* Let  $f = (f_1, \dots, f_k)$ , then  $\|f\| = \sqrt{\sum f_i^2}$  for  $1 \leq i \leq k$ . Now  $f$  is integrable, hence so is  $\sum f_i^2$ , and furthermore, since  $\|f\|$  is continuous on  $[0, M]$  for  $M > 0$ , it is integrable by theorem 7.1.10.

Now let  $y = (y_1, \dots, y_k)$ , with  $y_i = \int_a^b f_i d\alpha$  for  $1 \leq i \leq k$ . Then  $y = \int_a^b f d\alpha$ , and

$$\|y\|^2 = \sum y_i^2 = \sum y_i \int_a^b f_i d\alpha = \int_a^b \sum y_i f_i d\alpha$$

hence, by the Cauchy-Schwarz inequality,  $\sum y_i f_i \leq \|y\| \|f\|$  for  $a \leq t \leq b$ , hence  $\|y\|^2 \leq \|y\| \int_a^b \|f\| d\alpha$ . ■

## 7.5 Rectifiable Curves.

**Definition.** We call a continuous mapping  $\gamma : [a, b] \rightarrow \mathbb{R}^k$  a **curve** in  $\mathbb{R}^k$ , on an interval  $[a, b]$ . If  $\gamma$  is 1-1, we call the curve an **arc**, and we call  $\gamma$  a **closed curve** if  $\gamma(a) = \gamma(b)$ .

Now, for each partition  $P$  of  $[a, b]$ , and for  $\gamma$  a curve on  $[a, b]$ , and let:

$$\Lambda(\gamma, P) = \sum_{i=1}^n \|\gamma(x_i) - \gamma(x_{i-1})\| \quad (7.30)$$

We define the **length** of  $\gamma$  to be :

$$\Lambda(\gamma) = \sup_P \Lambda(\gamma, P) \quad (7.31)$$

If  $\gamma$  is of finite length (i.e.  $\Lambda(\gamma) < \infty$ ), then we call  $\gamma$  a **rectifiable** curve.

**Theorem 7.5.1.** *Let  $\gamma$  be a differentiable curve on  $[a, b]$ , if  $\gamma'$  is continuous, then  $\gamma$  is rectifiable and*

$$\Lambda(\gamma) = \int_a^b \|\gamma'\| dt \quad (7.32)$$

*Proof.* If  $a \leq x_{i-1} < x_i \leq b$ , then  $\|\gamma(x_i) - \gamma(x_{i-1})\| = \left\| \int_{x_{i-1}}^{x_i} \gamma' dt \right\| \leq \int_{x_{i-1}}^{x_i} \|\gamma'\| dt$ , by the Fundamental theorem of calculus, thus  $\Lambda(\gamma, P) \leq \int_a^b \|\gamma'\| dt$  for every partition  $P$  of  $[a, b]$ , then consequently, we have  $\Lambda(\gamma) = \int_a^b \|\gamma'\| dt$ . Consequently,  $\gamma$  is rectifiable.

Now let  $\varepsilon > 0$ , since  $\gamma'$  is uniformly continuous on  $[a, b]$ , there is a  $\delta > 0$  such that  $\|\gamma'(s) - \gamma'(t)\| < \varepsilon$  whenever  $|s - t| < \delta$  for  $s, t \in [a, b]$ . Let  $P$  be a partition of  $[a, b]$ , with



$\Delta x_i < \delta$  for all  $i$ , and if  $x_{i-1} \leq t \leq x_i$ , we have  $\|\gamma'(t)\| \leq \|\gamma'(x_i)\| + \varepsilon$ , hence:

$$\begin{aligned}
 \int_{x_{i-1}}^{x_i} \|\gamma'\| \Delta x_i &\leq \|\gamma'(x_i)\| \Delta x_i + \varepsilon \Delta x_i \\
 &= \left\| \int_{x_{i-1}}^{x_i} (\gamma'(t) + \gamma'(x_i) - \gamma'(t)) dt \right\| + \varepsilon \Delta x_i \\
 &\leq \left\| \int_{x_{i-1}}^{x_i} \gamma' dt \right\| + \left\| \int_{x_{i-1}}^{x_i} (\gamma'(x_i) - \gamma'(t)) dt \right\| + \varepsilon \Delta x_i \\
 &\leq \|\gamma(x_i) - \gamma(x_{i-1})\| + 2\varepsilon \Delta x_i
 \end{aligned}$$

Thus,  $\int_a^b \|\gamma'\| dt \leq \Lambda(\gamma, P) + 2\varepsilon(b-a)$ , so for  $\varepsilon$  small enough, we get that  $\int \|\gamma'\| \leq \Lambda(\gamma)$ . ■



# Chapter 8

## Lebesgue Theory

### 8.1 $\sigma$ -algebras

**Definition.** Let  $X$  be a nonempty set. An **algebra** of sets on  $X$  is a nonempty collection  $\mathcal{A}$  of subsets of  $X$  which are closed under finite unions and complements in  $X$ . We call  $\mathcal{A}$  a  **$\sigma$ -algebra** if it is closed under countable unions.

**Lemma 8.1.1.** *Let  $X$  be a set and  $\mathcal{A}$  an algebra on  $X$ . Then  $\mathcal{A}$  is closed under finite intersections.*

*Proof.* Let  $\{E_\lambda\}$  be a collection of sets of  $\mathcal{A}$ . Then by finite union  $E = \bigcup E_\lambda \in \mathcal{A}$ . Then by complements,  $X \setminus E = \bigcap X \setminus E_\lambda \in \mathcal{A}$ . ■

**Corollary.**  *$\sigma$ -algebras are closed under countable unions.*

**Lemma 8.1.2.** *Let  $X$  be a set, and  $\mathcal{A}$  an algebra on  $X$ . Then  $\emptyset \in \mathcal{A}$  and  $X \in \mathcal{A}$ .*

*Proof.* By closure of finite unions, notice that if  $E \in \mathcal{A}$ , then  $E \cup X \setminus E = X \in \mathcal{A}$  lemma 8.1.1 gives us that  $E \cap X \setminus E = \emptyset \in \mathcal{A}$ . ■

**Example 8.1.** (1) The collections  $\{\emptyset, X\}$  and  $2^X$  are  $\sigma$ -algebras on any set  $X$ .

(2) Let  $X$  be an uncountable set. Then the collection

$$\mathcal{C} = \{E \subseteq X : E \text{ is countable or } X \setminus E \text{ is countable}\}$$

defines a  $\sigma$ -algebra of sets on  $X$ , since countable unions of countable sets are countable, and  $\mathcal{C}$  is closed under complements. We call  $\mathcal{C}$  the  **$\sigma$ -algebra of countable or cocountable sets**.

**Lemma 8.1.3.** *Let  $\{A_\lambda\}$  be a collection of  $\sigma$ -algebras on a set  $X$ . Then the intersection*

$$\mathcal{A} = \bigcap \mathcal{A}_\lambda$$

*is a  $\sigma$ -algebra on  $X$ . Moreover, if  $F \subseteq X$ , then there exists a unique smallest  $\sigma$ -algebra containing  $F$ ; in particular, it is the intersection of all  $\sigma$ -algebras containing  $F$ .*

*Proof.* Notice that since each  $\mathcal{A}_\lambda$  is a  $\sigma$ -algebra, they are closed under countable unions and complements. Hence by definition,  $\mathcal{A}$  must also be closed under countable unions and complements.

Now, let  $F \subseteq X$  and let  $\{\mathcal{A}_\lambda\}$  be the collection of all  $\sigma$ -algebras containing  $F$ . Then the intersection  $\mathcal{A} = \bigcap \mathcal{A}_\lambda$  is also a  $\sigma$ -algebra containing  $F$ ; by above. Now, suppose that there is a smallest  $\sigma$ -algebra  $\mathcal{B}$  containing  $F$ . Then we have that  $\mathcal{B} \subseteq \mathcal{A}$ . Now, by definition of  $\mathcal{A}$  as the intersection of all  $\sigma$ -algebras containing  $F$ , we get that  $\mathcal{A} \subseteq \mathcal{B}$ ; so that  $\mathcal{B} = \mathcal{A}$ . ■

**Definition.** Let  $X$  be a nonempty set and  $F \subseteq X$ . We define the  $\sigma$ -algebra **generated** by  $F$  to be the smallest such  $\sigma$ -algebra  $\mathcal{M}(F)$  containing  $F$ .

**Lemma 8.1.4.** Let  $X$  be a set and let  $E, F \subseteq X$ . Then if  $E \subseteq \mathcal{M}(F)$ , then  $\mathcal{M}(E) \subseteq \mathcal{M}(F)$ .

*Proof.* We have that since  $E \subseteq \mathcal{M}(F)$ , and  $\mathcal{M}(E)$  is the intersection of all  $\sigma$ -algebras containing  $E$ , then  $\mathcal{M}(E) \subseteq \mathcal{M}(F)$ . ■

**Definition.** Let  $X$  be a topological space. We define the **Borel  $\sigma$ -algebra** on  $X$  to be the  $\sigma$ -algebra  $\mathcal{B}(X)$  generated by all open sets of  $X$ ; that is

$$\mathcal{B}(X) = \mathcal{M}(\mathcal{T})$$

where  $\mathcal{T}$  is the topology on  $X$ . We call the elements of  $\mathcal{B}(X)$  **Borel-sets**

**Definition.** Let  $X$  be a topological space. We call a countable intersection of open sets of  $X$  a  $G_\delta$ -**set** of  $X$ . We call a countable union of closed sets of  $X$  an  $F_\sigma$ -**set** of  $X$ .

**Theorem 8.1.5.** The Borel  $\sigma$ -algebra on  $\mathbb{R}$ ,  $\mathcal{B}(\mathbb{R})$ , is generated by the following.

- (1) All open intervals of  $\mathbb{R}$ .
- (2) All closed intervals of  $\mathbb{R}$ .
- (3) All half-open intervals of  $\mathbb{R}$ .
- (4) All open rays of  $\mathbb{R}$ .
- (5) All closed rays of  $\mathbb{R}$ .

## 8.2 The Lebesgue Outer Measure

**Definition.** Let  $I \subseteq \mathbb{R}$  be an interval. We define the **length**  $l(I)$ , of  $I$  to be  $\infty$  if  $I$  is unbounded, and the difference of its endpoints otherwise.

**Definition.** Let  $\{I_k\}$  be a countable collection of open bounded intervals covering a set  $A \subseteq \mathbb{R}$ . We define the **Lebesgue outer measure** of  $A$  to be

$$m^*(A) = \inf \left\{ \sum_{k=1}^{\infty} l(I_k) : A \subseteq \bigcup_{k=1}^{\infty} I_k \right\}$$

We now go over some basic properties of the Lebesgue outer measure.

**Lemma 8.2.1.** *For any set  $A \subseteq \mathbb{R}$ ,  $m^*(A) \geq 0$ ; in particular,  $m^*(\emptyset) = 0$ .*

*Proof.* By definition, since  $l(I_k) \geq 0$ , each  $\sum l(I_k) \geq 0$ . This makes  $m^*(A) \geq 0$  for any  $\{I_k\}$  a countable cover of  $A$  by bounded open intervals.

Notice that  $\emptyset \subseteq (-\varepsilon, \varepsilon)$  for some  $\varepsilon > 0$ , and that this covers  $\emptyset$ , so that  $m^*(\emptyset) = l((-\varepsilon, \varepsilon)) = 2\varepsilon$ . Then choosing  $\varepsilon$  small enough gives us  $m^*(\emptyset) = 0$ . ■

**Lemma 8.2.2** (Monotonicity). *The Lebesgue Outer Measure is monoton; that is, if  $A \subseteq B$ , then*

$$m^*(A) \leq m^*(B)$$

*Proof.* Let  $\{I_k\}$  be a countable cover of  $B$  by bounded open intervals. Then notice that  $\{I_k\}$  covers  $A$  as well. Now, let

$$E = \left\{ \sum_{k=1}^{\infty} l(I_k) : A \subseteq \bigcup_{k=1}^{\infty} I_k \right\}$$

$$F = \left\{ \sum_{k=1}^{\infty} l(I_k) : B \subseteq \bigcup_{k=1}^{\infty} I_k \right\}$$

Then since  $A \subseteq B$ ,  $F \subseteq E$ . Therefore, we get that

$$\inf E = m^*(A) \leq m^*(B) = \inf F$$

■

**Lemma 8.2.3.** *Countable sets have Lebesgue outer measure 0.*

*Proof.* Let  $C$  be a countable set, and  $C = \{c_k\}_{k \in \mathbb{Z}^+}$  an enumeration for  $C$ . Let  $\varepsilon > 0$ , then for every  $k \in \mathbb{Z}^+$ , define the interval

$$I_k = (c_k - \frac{\varepsilon}{2^{k+1}}, c_k + \frac{\varepsilon}{2^{k+1}})$$

Then  $\{I_k\}$  is a countable cover of  $C$  by bounded open intervals. Thus we get

$$0 \leq m^*(C) \leq \sum_{k=1}^{\infty} l(I_k) = \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = \varepsilon$$

Therefore taking  $\varepsilon$  small enough, we get  $0 \leq m^*(C) \leq 0$ , and equality is established. ■

**Lemma 8.2.4.** *Let  $I \subseteq \mathbb{R}$  be an interval. Then  $m^*(I) = l(I)$ .*

*Proof.* Consider first that  $I = [a, b]$ , a closed bounded interval, where  $a < b$ . Let  $\varepsilon > 0$ , then we have that  $[a, b] \subseteq (a - \varepsilon, b + \varepsilon)$ , so that by monotonicity

$$m^*([a, b]) \leq l((a - \varepsilon, b + \varepsilon)) = b - a + 2\varepsilon$$

Then for  $\varepsilon$  small enough, this shows that

$$m^*([a, b]) \leq b - a$$

Now, let  $\{I_k\}$  be a countable cover of  $[a, b]$  by bounded open intervals. Then since  $[a, b]$  is compact (by Heine-Borel), there is a finite subcover  $\{I_k\}_{k=1}^n$  of  $[a, b]$ . Now, since  $a \in \bigcup_{k=1}^n I_k$ , there is at least one interval  $I_k$  containing  $a$ ; denote it  $(a_1, b_1)$ . Now,  $a_1 < a < b_1$ . If  $b_1 \geq b$ , then we are done as

$$\sum_{k=1}^n l(I_k) \geq b_1 - a_1 > b - a$$

Otherwise,  $b_1 \in [a, b)$ , and since  $b_1 \notin (a_1, b_1)$ , there is an interval  $(a_2, b_2)$ , distinct from  $(a_1, b_1)$ , containing  $b_1$ . Now, if  $b_2 \geq b$ , we are done. Otherwise, proceeding recursively, we obtain a subcollection  $\{(a_k, b_k)\}_{k=1}^N$  of  $\{I_k\}_{k=1}^n$  for which

$$a_1 < a \text{ and } a_{k+1} < b_k \text{ for all } 1 \leq k \leq N-1$$

The process of selecting such a subcollection must terminate, which leaves us with  $b_N > b$ , so that

$$\sum_{k=1}^n l(I_k) \geq \sum_{k=1}^N l((a_k, b_k)) = (b_N - a_N) + \cdots + (b_1 - a_1) \geq b_N - a_1 > b - a$$

Indeed, we get

$$\sum_{k=1}^n l(I_k) \geq b - a$$

so that  $m^*([a, b]) = b - a$ .

Now, suppose that  $I$  is any bounded interval. Then for  $\varepsilon > 0$ , there exist closed bounded interval  $J_1$  and  $J_2$  such that

$$J_1 \subseteq I \subseteq J_2$$

and

$$l(I) - \varepsilon < l(J_1) \leq l(I) \leq l(J_2) < l(I) + \varepsilon$$

By the monotonicity, and the above discussion of closed bounded intervals, we get

$$l(I) - \varepsilon < m^*(J_1) \leq m^*(I) \leq m^*(J_2) < l(I) + \varepsilon$$

Therefore, for  $\varepsilon$  small enough, we get that  $m^*(I) = l(I)$ .

Finally, suppose that  $I$  is an unbounded interval. Then for every  $n \in \mathbb{Z}^+$ , there is an interval  $J$  with  $l(J) = n$ . So that

$$n = m^*(J) \leq m^*(I)$$

This makes  $m^*(I) = l(I) = \infty$ , by definition of  $l(I)$ . ■

**Lemma 8.2.5** (Translation Invariance). *The Lebesgue outer measure is translation invariant; that is, if  $A \subseteq \mathbb{R}$ , and  $y \in \mathbb{R}$ ,*

$$m^*(A + y) = m^*(A)$$

*Proof.* Let  $\{I_k\}$  be a countable cover of  $A$  by open bounded intervals. Then the collection  $\{I_k + y\}$  is a countable cover of the set  $A + y$  by open bounded intervals. Moreover, notice that  $l(I_k) = l(I_k + y)$ . This gives us

$$\sum l(I_k) = \sum l(I_k + y)$$

and we are done. ■

**Lemma 8.2.6** (Countable Subadditivity). *The Lebesgue outer measure is countable subadditive; that is, if  $\{E_k\}$  is a countable collection of subsets of  $\mathbb{R}$ , then*

$$m^*\left(\bigcup E_k\right) \leq \sum m^*(E_k)$$

*Proof.* Let  $\{E_k\}$  be a countable collection of subsets of  $\mathbb{R}$ , and let

$$E = \bigcup E_k$$

If at least one of the  $E_k$  has  $m^*(E_k) = \infty$ , then we are done. Suppose then that  $m^*(E_k)$  is finite for all  $k \in \mathbb{Z}^+$ . Then for each  $E_k$ , there is a countable cover  $\{I_{k,i}\}_{i \in \mathbb{Z}^+}$  by bounded open intervals for which

$$\bigcup l(I_{k,i}) < m^*(E_k) + \frac{\varepsilon}{2^k}$$

Now, consider the countable collection

$$\{I_{k,i}\}_{k,i \in \mathbb{Z}^+} = \bigcup_{k \in \mathbb{Z}^+} \{I_{k,i}\}_{i \in \mathbb{Z}^+}$$

Then  $\{I_{k,i}\}_{k,i \in \mathbb{Z}^+}$  is a countable cover of  $E$  by bounded open intervals. We get

$$m^*(E) \leq \sum_{k \in \mathbb{Z}^+} \left( \sum_{i \in \mathbb{Z}^+} l(I_{k,i}) \right) < \sum_{k \in \mathbb{Z}^+} \left( m^*(E_k) + \frac{\varepsilon}{2^k} \right) = \left( \sum m^*(E_k) \right) + \varepsilon$$

Taking  $\varepsilon > 0$  small enough, gives us the required subadditivity. ■

**Corollary.** *The Lebesgue outer measure is finitely subadditive.*

*Proof.* Recall that finite collections of sets are countable. ■

## 8.3 Lebesgue Measurable Sets

**Definition.** A set  $E$  of  $\mathbb{R}$  is said to be **measurable** if for any set  $A \subseteq \mathbb{R}$ , we have

$$m^*(A) = m^*(A \cap E) + m^*(A \cap (\mathbb{R} \setminus E))$$

**Lemma 8.3.1.** *A set  $E$  is measurable if for any  $A \subseteq \mathbb{R}$ , we have*

$$m^*(A) \geq m^*(A \cap E) + m^*(A \cap (\mathbb{R} \setminus E))$$

*Proof.* This follows by finite subadditivity, notice that both  $A \cap E \subseteq A$  and  $A \cap (\mathbb{R} \setminus E) \subseteq A$ , and that  $A = (A \cap E) \cup (A \cap (\mathbb{R} \setminus E))$ . Thus

$$m^*(A) \leq m^*(A \cap E) + m^*(A \cap (\mathbb{R} \setminus E))$$

■

**Corollary.** *A set  $E$  is measurable if, and only if  $\mathbb{R} \setminus E$  is measurable.*

**Lemma 8.3.2.** *Any set of Lebesgue outer measure 0 is measurable.*

*Proof.* Let  $E$  be a set with  $m^*(E) = 0$ , and let  $A \subseteq \mathbb{R}$ . Then we have  $A \cap E \subseteq E$  and  $A \cap (\mathbb{R} \setminus E) \subseteq A$ , so that by monotonicity

$$m^*(A \cap E) \leq m^*(E) = 0 \text{ and } m^*(A \cap (\mathbb{R} \setminus E)) \leq m^*(A)$$

Adding both relations gives us the result. ■

**Corollary.** *Countable sets are measurable.*

**Lemma 8.3.3.** *The union of a finite collection of measurable sets is measurable.*

*Proof.* Let  $E_1$  and  $E_2$  be measurable and  $A \subseteq \mathbb{R}$ . Then

$$\begin{aligned} m^*(A) &= m^*(A \cap E_1) + m^*(A \cap (\mathbb{R} \setminus E_1)) \\ &= m^*(A \cap E_1) + m^*((A \cap E_1) \cap E_2) + m^*((A \cap (\mathbb{R} \setminus E_1)) \cap (\mathbb{R} \setminus E_2)) \end{aligned}$$

Now, we have that

$$(A \cap (\mathbb{R} \setminus E_1)) \cap (\mathbb{R} \setminus E_2) = A \cap \mathbb{R} \setminus (E_1 \cup E_2) \text{ and } (A \cap E_1) \cup (A \cap (\mathbb{R} \setminus E_1) \cap E_2) = A \cap (E_1 \cup E_2)$$

which gives us

$$m^*(A) \geq m^*(A \cap (E_1 \cup E_2)) + m^*(A \cap \mathbb{R} \setminus (E_1 \cup E_2))$$

which makes the union  $E_1 \cup E_2$  measurable. Extending the argument for arbitrary  $n \in \mathbb{Z}^+$  by induction gives us the result. ■

**Lemma 8.3.4.** *Let  $A$  be a set and  $\{E_k\}_{k=1}^n$  a disjoint collection of measurable sets. Then*

$$m^*\left(A \cap \left(\bigcup_{k=1}^n E_k\right)\right) = \sum_{k=1}^n m^*(A \cap E_k)$$

*Proof.* By induction on  $n$ , for  $n = 1$ , we are done. Now, suppose that our assertion holds for all  $n \geq 1$ , and consider the finite disjoint collection  $\{E_k\}_{k=1}^{n+1}$ . The since this collection is disjoint, we have

$$A \cap \left(\bigcup_{k=1}^{n+1} E_k\right) \cap E_{n+1} = A \cap E_{n+1}$$



and

$$A \cap \left( \bigcup_{k=1}^{n+1} E_k \right) \cap (\mathbb{R} \setminus E_{n+1}) = A \cap \left( \bigcup_{k=1}^n E_k \right)$$

By the measurability of  $E_{n+1}$ , we get that

$$m^* \left( A \cap \left( \bigcup_{k=1}^{n+1} E_k \right) \right) = m^*(A \cap E_{n+1}) + \sum_{k=1}^n m^*(A \cap E_k) = \sum_{k=1}^{n+1} m^*(A \cap E_k)$$

■

**Lemma 8.3.5.** *Countable unions of measurable sets are measurable.*

*Proof.* Let  $\{E_k\}$  be a countable collection of measurable sets, and  $E = \bigcup E_k$  the countable union; suppose further without loss of generality that the collection  $\{E_k\}$  is a disjoint collection. Let  $A \subseteq \mathbb{R}$ , and  $n \in \mathbb{Z}^+$  and define

$$F_n = \bigcup_{k=1}^n E_k$$

Then  $F$  is measurable since it is a finite union of measurable sets and  $\mathbb{R} \setminus E \subseteq \mathbb{R} \setminus F$ , moreover, we have

$$m^*(A) = m^*(A \cap E_n) + m^*(A \cap (\mathbb{R} \setminus F)) \geq m^*(A \cap F_n) + m^*(A \cap (\mathbb{R} \setminus E))$$

Now, by the lemma 8.3.4, we have that

$$m^*(A \cap F) = \sum_{k=1}^n m^*(A \cap E_k)$$

so that

$$m^*(A) \geq \sum_{k=1}^n m^*(A \cap E_k) + m^*(A \cap (\mathbb{R} \setminus E))$$

Now, choosing  $n \in \mathbb{Z}^+$  large enough, we get

$$m^*(A) \geq \sum m^*(A \cap E_k) + m^*(A \cap (\mathbb{R} \setminus E))$$

so that  $m^*(A) \geq m^*(A \cap E) + m^*(A \cap (\mathbb{R} \setminus E))$

■

**Theorem 8.3.6.** *The measurable sets of  $\mathbb{R}$  form a  $\sigma$ -algebra on  $\mathbb{R}$ .*

**Lemma 8.3.7.** *Every interval is measurable.*

*Proof.* It suffices to show that the intervals of the form  $(a, \infty)$  are measurable. Let  $A \subseteq \mathbb{R}$ , and suppose that  $a \notin A$ . Now, let  $\{I_k\}$  be a countable collection of open bounded intervals covering  $A$ . Now, define

$$A' = A \cap (-\infty, a) \text{ and } A'' = A \cap (a, \infty)$$

and

$$I'_k = I_k \cap (-\infty, a) \text{ and } I''_k = I_k \cap (a, \infty)$$

Then the  $I'_k$  and  $I''_k$  are intervals with  $l(I_k) = l(I'_k) + l(I''_k)$ . Moreover, the countable collections  $\{I'_k\}$  and  $\{I''_k\}$  form open covers of  $A'$  and  $A''$  respectively. Therefore, by the definition of outer measure, we get

$$m^*(A') \leq \sum l(I'_k) \text{ and } m^*(A'') \leq \sum l(I''_k)$$

so that

$$m^*(A') + m^*(A'') \leq \sum l(I_k)$$

which makes  $(a, \infty)$  measurable. ■

**Corollary.** *Every Borel set of  $\mathbb{R}$  is measurable.*

**Corollary.**  *$G_\delta$  sets and  $F_\sigma$  sets are measurable.*

**Lemma 8.3.8.** *Translates of measurable sets are measurable.*

*Proof.* Let  $E$  be measurable, and  $y \in \mathbb{R}$ , and let  $A \subseteq \mathbb{R}$ . We have by the measurability of  $E$ , and by translation invariance of outer measure that

$$\begin{aligned} m^*(A) &= m^*(A \setminus y) \\ &= m^*((A \setminus y) \cap E) + m^*((A \setminus y) \cap (\mathbb{R} \setminus E)) \\ &= m^*(A \cap (E + y)) + m^*(A \cap \mathbb{R} \setminus (E + y)) \end{aligned}$$
■

**Lemma 8.3.9** (The Excision Property). *If  $A$  is a measurable set of finite outer measure contained in a set  $B$  of  $\mathbb{R}$ , then*

$$m^*(B \setminus A) = m^*(B) - m^*(A)$$

*Proof.* We have that

$$m^*(B) = m^*(B \cap A) + m^*(B \cap (\mathbb{R} \setminus A)) = m^*(A) + m^*(B \setminus A)$$
■

## 8.4 Outer and Inner Approximations

**Theorem 8.4.1** (The Outer Approximation Theorem). *Let  $E \subseteq \mathbb{R}$ . Then the following are true.*

- (1)  *$E$  is measurable.*
- (2) *For every  $\varepsilon > 0$ , there is an open set  $U$ , containing  $E$  for which  $m^*(U \setminus E) < \varepsilon$*
- (3) *There exists a  $G_\delta$  set  $G$ , containing  $E$  for which  $m^*(G \setminus E) = 0$ .*

*Proof.* Suppose that  $E$  is measurable, and let  $\varepsilon > 0$ . If  $m^*(E)$  is finite, then by definition of outer measure, there exists a countable collection of open bounded intervals  $\{I_k\}$  covering  $E$ , for which

$$\sum l(I_k) < m^*(E) + \varepsilon$$

Define now,  $U = \bigcup I_k$ , then  $U$  is open in  $\mathbb{R}$ , and  $E \subseteq U$ . Again, by the definition of outer measure, we have

$$m^*(U) \leq \sum l(I_k) < m^*(E) + \varepsilon$$

so that

$$m^*(U) - m^*(E) < \varepsilon$$

Now, since  $E$  is measurable, and  $E \subseteq U$ , by the excision property we get  $m^*(U \setminus E) < \varepsilon$ .

Now, if  $m^*(E)$  is infinite, then there exists a disjoint countable collection  $\{E_k\}$  of measurable sets, each of finite outer measure, for which  $E = \bigcup E_k$ . Then for any  $k \in \mathbb{Z}^+$ , there exists an open set  $U_k$  with  $E_k \subseteq U_k$  for which

$$m^*(U_k \setminus E_k) < \frac{\varepsilon}{2^k}$$

Now, let  $U = \bigcup U_k$ , so that  $U$  is open and  $E \subseteq U$ . Moreover, we get  $U \setminus E \subseteq \bigcup U_k \setminus E_k$ ; therefore, by monotonicity, we get

$$m^*(U \setminus E) \leq \sum m^*(U_k \setminus E_k) < \sum \frac{\varepsilon}{2^k} = \varepsilon$$

Now, suppose that the second assertion holds for  $E$ . Then for any  $k \in \mathbb{Z}^+$ , choose an open set  $U_k$  containing  $E$  for which

$$m^*(U_k \setminus E) < \frac{1}{k}$$

Define  $G = \bigcap U_k$ . Then  $G$  is a  $G_\delta$  set and  $E \subseteq G$ . Moreover, for each  $k$ , we have  $G \setminus E \subseteq U_k \setminus E$  so by monotonicity

$$m^*(G \setminus E) \leq m^*(U_k \setminus E) < \frac{1}{k}$$

Then as  $k \rightarrow \infty$ ,  $m^*(G \setminus E) \rightarrow 0$ .

Finally, suppose that the third assertion holds for  $E$ . Now, since sets of measure 0 are measurable, and  $G_\delta$  are measurable, we get

$$E = G \cap \mathbb{R} \setminus (G \setminus E)$$

is measurable, by theorem 8.3.6. ■

**Theorem 8.4.2** (The Inner Approximation Theorem). *Let  $E \subseteq \mathbb{R}$ . Then the following are true.*

- (1)  $E$  is measurable.
- (2) For every  $\varepsilon > 0$ , there is a closed set  $F$ , contained in  $E$  for which  $m^*(E \setminus F) < \varepsilon$
- (3) There exists a  $F_\sigma$  set  $F$ , contained in  $E$  for which  $m^*(E \setminus F) = 0$ .

*Proof.* This follows by DeMorgan's laws on the outer approximation theorem. ■

**Theorem 8.4.3.** *Let  $E$  be measurable of finite outer measure. Then for every  $\varepsilon > 0$ , there exists a finite disjoint collection of open intervals  $\{I_k\}_{k=1}^n$  for which if  $U = \bigcup_{k=1}^n I_k$ , then*

$$m^*(E \setminus U) + m^*(U \setminus E) < \varepsilon$$

*Proof.* Since  $E$  is measurable, there exists an open set  $V$ , containing  $E$  and for which

$$m^*(V \setminus E) < \frac{\varepsilon}{2}$$

Then by the excision property, we observe that since  $E$  has finite outer measure, then  $V$  must also have finite outer measure. Now, let  $V = \bigcup I_k$  where  $\{I_k\}$  is a countable disjoint collection of open intervals. Therefore, by lemma 8.3.3, and monotonicity we get

$$\sum_{k=1}^n l(I_k) = m^*\left(\bigcup_{k=1}^n I_k\right) \leq m^*(V)$$

Then as  $n \rightarrow \infty$ , we see the sum  $\sum l(I_k)$  is finite. Then choosing  $n \in \mathbb{Z}^+$  for which

$$\sum_{k=n+1}^{\infty} l(I_k) < \frac{\varepsilon}{2}$$

and defining

$$U = \bigcup_{k=1}^n I_k$$

we get  $U \setminus E \subseteq V \setminus E$  so by monotonicity,

$$m^*(U \setminus E) \leq m^*(V \setminus E) < \frac{\varepsilon}{2}$$

on the other hand, since  $E \subseteq V$ , we get  $E \setminus U \subseteq V \setminus U = \bigcup_{k=n+1}^{\infty} I_k$ . Therefore by definition of outer measure

$$m^*(E \setminus U) \leq \sum_{k=n+1}^{\infty} l(I_k) < \frac{\varepsilon}{2}$$

This gives us

$$m^*(U \setminus E) + m^*(E \setminus U) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

■

## 8.5 The Borel-Cantelli Lemma

**Definition.** We define the **Lebesgue measure**  $m$  to be the restriction of the Lebesgue outer measure to the  $\sigma$ -algebra of measurable sets of  $\mathbb{R}$ . That is, if  $E \subseteq \mathbb{R}$  is measurable, then

$$m(E) = m^*(E)$$

**Lemma 8.5.1.** *The Lebesgue measure is countable additive. That is, if  $\{E_k\}$  is a countable disjoint collection of measurable sets, then*

$$m\left(\bigcup E_k\right) = \sum m(E_k)$$

*Proof.* Let  $\{E_k\}$  be a countable disjoint collection of measurable sets, and let  $E = \bigcup E_k$ . Then  $E$  is also measurable, and we have by countable subadditivity

$$m(E) \leq \sum m(E_k)$$

Now, by lemma 8.3.4, for each  $n \in \mathbb{Z}^+$ , we have

$$m\left(\bigcup_{k=1}^n E_k\right) = \sum_{k=1}^n m(E_k)$$

and that

$$\bigcup_{k=1}^n E_k \subseteq E$$

Thus, by monotonicity, we have

$$\sum_{k=1}^n m(E_k) \leq m(E)$$

Now, since  $n$  is arbitrary, choosing  $n$  large enough, we get

$$\sum m(E_k) \leq m(E)$$

and so equality is established. ■

**Theorem 8.5.2** (Continuity of Measure). *The following is true for the Lebesgue measure.*

(1) *If  $\{A_k\}$  is an ascending collection of measurable sets, then*

$$m\left(\bigcup A_k\right) = \lim_{k \rightarrow \infty} m(A_k)$$

(2) *If  $\{B_k\}$  is a descending collection of measurable sets, and  $m(B_1)$  is finite, then*

$$m\left(\bigcap B_k\right) = \lim_{k \rightarrow \infty} m(B_k)$$

*Proof.* Let  $A = \bigcup A_k$ . First, suppose that there is a  $k_0 \in \mathbb{Z}^+$  for which  $m(A_{k_0}) = \infty$ . Then  $m(A) = \infty$ , and so  $m(A_k) = \infty$  for all  $k \geq k_0$ , and we are done. Now, suppose that  $m(A_k)$  is finite for all  $k \in \mathbb{Z}^+$ . Define

$$A_0 = \emptyset \text{ and } C_{k+1} = A_{k+1} \setminus A_k \text{ for all } k \geq 1$$

Since  $\{A_k\}$  is an ascending collection, the collection  $\{C_k\}$  is disjoint, and we have

$$A = \bigcup A_k = \bigcup C_k$$

Then, by the countable additivity of the Lebesgue measure, we get

$$m(A) = \sum m(A_{k+1} \setminus A_k)$$

and by excision, we get

$$\begin{aligned} \sum m(A_{k+1} \setminus A_k) &= \sum (m(A_{k+1}) - m(A_k)) \\ &= \lim_{n \rightarrow \infty} \sum_{k=0}^{n-1} (m(A_{k+1}) - m(A_k)) \\ &= \lim_{n \rightarrow \infty} (m(A_n) - m(A_0)) \end{aligned}$$

Now, since  $m(A_0) = m(\emptyset) = 0$ , we get that

$$m(A) = \lim_{k \rightarrow \infty} m(A_k)$$

Now, let  $B = \bigcap B_k$ , and suppose that  $m(B_1)$  is finite. Define

$$D_k = B_1 \setminus B_k \text{ for all } k \in \mathbb{Z}^+$$

Since  $\{B_k\}$  is a descending collection,  $\{D_k\}$  is an ascending collection. By the above assertion, we get

$$m\left(\bigcup D_k\right) = \lim_{k \rightarrow \infty} m(D_k)$$

Now, by DeMorgan's laws, we have

$$\bigcup D_k = B_1 \setminus \left(\bigcap B_k\right)$$

so that by the excision property

$$m\left(B_1 \setminus \left(\bigcap B_k\right)\right) = \lim_{k \rightarrow \infty} (m(B_1) - m(B_k))$$

by the excision property on the left hand side of the equation, we get

$$m(B_1) - m(B) = m(B_1) - \lim_{k \rightarrow \infty} m(B_k)$$

which gives us the result. ■

**Definition.** We say that a property on a Lebesgue measurable set  $E$  of  $\mathbb{R}$  holds **almost everywhere** provided there exists a subset  $E_0 \subseteq E$  of  $m(E_0) = 0$  for which the property holds for all  $x \in E \setminus E_0$ .

**Lemma 8.5.3** (The Borel-Cantelli Lemma). *Let  $\{E_k\}$  be a countable collection of measurable sets for which  $\sum m(E_k)$  is finite. Then almost all  $x \in \mathbb{R}$  belong to at most finitely many  $E_k$ .*

*Proof.* Let  $E = \bigcup_{k=n}^{\infty} E_k$ , then  $E$  is measurable. By subadditivity

$$m(E) \leq \sum_{k=n}^{\infty} m(E_k) < \infty$$

Therefore, by the continuity of measure, we have

$$m\left(\bigcap_{n=1}^{\infty} E\right) = \lim_{n \rightarrow \infty} m(E) \leq \lim_{n \rightarrow \infty} \sum_{k=1}^n m(E_k) = 0$$

Therefore, almost all  $x \in \mathbb{R}$  fail to belong to  $\bigcap_{n=1}^{\infty} E$ , and so must belong to at most finitely many  $E_k$ . ■

## 8.6 Nonmeasurable Sets

**Definition.** We call a set  $E$  of  $\mathbb{R}$  **nonmeasurable** if it is not a measurable set.

**Lemma 8.6.1.** *Let  $E$  be a bounded measurable set, and suppose there exists a bounded countably infinite set of reals  $\Lambda$  for which the collection of translates  $\{\lambda + E\}$  is disjoint. Then  $m(E) = 0$ .*

*Proof.* Since translates of measurable sets are measurable, by countable additivity

$$m\left(\bigcup_{\lambda \in \Lambda} (\lambda + E)\right) = \sum_{\lambda \in \Lambda} m(\lambda + E) = \sum_{\lambda \in \Lambda} m(E)$$

Now, since both  $E$  and  $\Lambda$  are bounded, then the union

$$\bigcup_{\lambda \in \Lambda} (\lambda + E)$$

is also bounded, and has finite measure. Thus, for each  $\lambda \in \Lambda$ ,  $m(\lambda + E) = m(E)$  is finite. Additionally, by translation invariance,  $m(\lambda + E) = m(E) > 0$ , and since  $\Lambda$  is countably infinite, we must have  $m(E) = 0$ . ■

**Definition.** Let  $E \subseteq \mathbb{R}$  be an arbitrary set. We call two points  $x, y \in E$  **rationally equivalent** in  $E$  provided  $x - y \in \mathbb{Q}$ . We write  $x \sim y$ .

**Lemma 8.6.2.** *For any set  $E \subseteq \mathbb{R}$ , rational equivalence defines an equivalence relation on  $E$ .*

*Proof.* We have that for any  $x \in E$ ,  $x - x = 0 \in \mathbb{Q}$ , so that  $x \sim x$ . Now, let  $x, y \in E$  with  $x \sim y$ . Then  $x - y \in \mathbb{Q}$ , and since  $\mathbb{Q}$  is a field,  $-(x - y) = y - x \in \mathbb{Q}$  so that  $y \sim x$ . Finally, let  $x, y, z \in E$  with  $x \sim y$  and  $y \sim z$ . Then  $x - y, y - z \in \mathbb{Q}$ , and again since  $\mathbb{Q}$  is a field,  $(x - y) + (y - z) = x - z \in \mathbb{Q}$  which makes  $x \sim z$ . ■

**Definition.** Let  $E$  be a nonempty set of  $\mathbb{R}$ . We define a **choice set** for rational equivalence on  $E$  to be a set  $\mathcal{C}_E$  consisting of exactly one member of each equivalence class in  $E/\sim$ ; such that

- (1) For all  $x, y \in \mathcal{C}_E$ ,  $x \not\sim y$ .
- (2) For every  $x \in E$ , there exists a  $c \in \mathcal{C}_E$  such that  $x = c + q$  for some  $q \in \mathbb{Q}$ .

**Theorem 8.6.3** (Vitali's Theorem). *Any set  $E$  with positive outer measure contains a non-measurable set.*

*Proof.* By countable subadditivity, suppose that  $E$  is bounded. Now, let  $\mathcal{C}_E$  be a choice set of rational equivalence on  $E$  (which exists by the axiom of choice). Then we claim that  $\mathcal{C}_E$  is nonmeasurable.

Indeed, suppose that  $\mathcal{C}_E$  is measurable, and let  $\Lambda$  be any countably infinite set of rationals. Then the collection  $\{\lambda + \mathcal{C}_E\}$  is disjoint and  $m(\lambda + \mathcal{C}_E) = m(\mathcal{C}_E) = 0$ , by lemma 8.6.1. Then, by translatio invariance, and countable additivity, we have

$$m\left(\bigcup_{\lambda \in \Lambda} \lambda + \mathcal{C}_E\right) = \sum_{\lambda \in \Lambda} m(\lambda + \mathcal{C}_E) = \sum_{\lambda \in \Lambda} m(\mathcal{C}_E) = 0$$

Now, since  $E$  is bounded,  $E \subseteq [-b, b]$  for some  $b \in \mathbb{R}$ . Choose, then a set  $\Lambda_0 = [-2b, 2b] \cap \mathbb{Q}$ . Then  $\Lambda_0$  is bounded and countably infinite. Now, if  $x \in E$ , then there is a  $c \in \mathcal{C}_E$  for which  $x = c + q$  for some  $q \in \mathbb{Q}$ . But  $x, c \in [-b, b]$ , so that  $x, c \in [-2b, 2b]$ . Thus

$$E \subseteq \bigcup_{\lambda \in \Lambda_0} (\lambda + \mathcal{C}_E)$$

but  $m^*(E) > 0$ , and  $m(\mathcal{C}_E) = 0$ , which is a contradiction! Therefore,  $\mathcal{C}_E$  cannot be measurable. ■

**Theorem 8.6.4.** *There exist nonmeasurable disjoint sets  $A, B \subseteq \mathbb{R}$  for which*

$$m^*(A \cup B) < m^*(A) \cup m^*(B).$$

## 8.7 The Cantor Set

**Definition.** We define the **Cantor set**  $\mathcal{C}$  to be the intersection

$$\mathcal{C} = \bigcap C_k$$

where  $\{C_k\}$  is a descending collection of closed sets in which for each  $k \in \mathbb{Z}^+$ ,  $C_k$  is the disjoint union of  $2^k - 1$  closed intervals each of length  $\frac{1}{3^k}$ .

**Theorem 8.7.1.** *The Cantor set is a closed and uncountable set of measure  $m(\mathcal{C}) = 0$ .*

*Proof.* Since  $\mathcal{C}$  is an arbitrary intersection of closed sets in  $\mathbb{R}$ ,  $\mathcal{C}$  is closed in  $\mathbb{R}$ . Moreover, by the definition of each  $C_k$ ,  $C_k$  is measurable and

$$m(C_k) = \left(\frac{2}{3}\right)^k \text{ for each } k \in \mathbb{Z}^+$$



Then  $\mathcal{C}$  is also measurable, and we have by monotonicity of the Lebesgue measure

$$m(\mathcal{C}) \leq m(C_k) = \left(\frac{2}{3}\right)^k$$

Taking  $k \rightarrow \infty$ , gives us that  $m(\mathcal{C}) = 0$ .

Now, suppose that  $\mathcal{C}$  is countable, and let  $\mathcal{C} = \{c_k\}$  be an enumeration of  $\mathcal{C}$ . Now, one of the disjoint intervals whose union is  $C_1$  fails to contain  $c_1$ ; call it  $F_1$ . Proceeding, one of the disjoint intervals whose union is  $F_1$  fails to contain  $c_2$ ; call it  $F_2$ . Proceeding inductively, we get a descending collection  $\{F_k\}$  of closed sets, for which  $c_k \notin F_k$  for each  $k \in \mathbb{Z}^+$ . Now, by the nested set theorem, the intersection

$$F = \bigcap F_k$$

is nonempty. Moreover,  $F \subseteq \mathcal{C}$ . This implies that there is an element  $x \in \mathcal{C}$  which is not equal to any  $c_k$ ; but  $\{c_k\}$  is an enumeration of  $\mathcal{C}$ , which is absurd! Therefore,  $\mathcal{C}$  fails to be countable. ■

**Definition.** Let  $\mathcal{C}$  be the Cantor set, and define  $U_k$  such that  $C_k = [0, 1] \setminus U_k$ , and define  $\mathcal{U} = \bigcup U_k$ ; so that  $[0, 1] = \mathcal{C} \cup \mathcal{U}$ . Fix  $k \in \mathbb{Z}^+$  and define

$$\phi : U_k \rightarrow \mathbb{R}$$

to be the increasing function which is constant on each of the  $2^k - 1$  intervals of  $U_k$ , and whose image is

$$\phi([0, 1]) = \left\{ \frac{1}{2^k}, \frac{2}{2^k}, \dots, \frac{2^k - 1}{2^k} \right\}$$

We define the **Cantor-Lebesgue function** to be the extension  $\Phi$  of  $\phi$  to  $[0, 1]$ ; defined in terms of  $\mathcal{C}$  to be

$$\Phi(0) = \phi(0) = 0 \text{ and } \Phi(x) = \sup \{ \phi(t) : t \in \mathcal{U} \cap [0, x] \} \text{ for all } x \in \mathcal{C} \setminus \{0\}$$

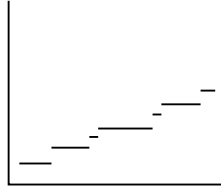


Figure 8.1: The Cantor Lebesgue Function on  $U_3$ .

**Theorem 8.7.2.** *The Cantor-Lebesgue function is an increasing continuous function mapping  $[0, 1]$  onto  $[0, 1]$ , and differentiable on  $\mathcal{U}$  with*

$$\Phi'(x) = 0 \text{ on } \mathcal{U} \text{ where } m(\mathcal{U}) = 1$$

*Proof.* Since  $\phi$  is an increasing function, and  $\Phi$  is the extension of  $\phi$  to  $[0, 1]$ , then  $\Phi$  must also be increasing. Now,  $\phi$  is continuous at each point of  $\mathcal{U}$ , since it is constant on each  $U_k$ . Now, consider  $x_0 \in \mathcal{C}$  with  $x_0 \neq 0, 1$ . Then  $x_0 \notin U_k$  for all  $k \in \mathbb{Z}^+$ . So for  $k$  large enough,  $x_0$  lies in between consecutive intervals of  $U_k$ . Choose  $a_k$  to be the lowerbound of the lower interval, and  $b_k$  the upperbound of the upper interval. Then by definition of  $\phi$ , we get

$$a_k < x_0 < b_k \text{ and } \phi(b_k) - \phi(a_k) = \frac{1}{2^k}$$

Since  $k$  is arbitrarily large,  $\Phi$  fails to have a jump discontinuity at  $x_0$ , which is the only possible discontinuity it can have. Therefore  $\Phi$  is continuous at  $x_0$ . A similar argument shows that  $\Phi$  is continuous at  $x_0 = 0, 1$ .

Now, since  $\phi$  is constant on each  $U_k$ ,  $\phi$  is differentiable on  $U_k$  and has  $\phi' = 0$  on each  $U_k$ , hence  $\phi' = 0$  on  $\mathcal{U}$ . Since  $m(\mathcal{C}) = 0$  and  $[0, 1] = \mathcal{C} \cup \mathcal{U}$ , then  $m(\mathcal{U}) = 1$ . Since  $\Phi$  is an extension, we get that

$$\Phi' = 0 \text{ on } \mathcal{U} \text{ where } m(\mathcal{U}) = 1$$

Finally, since  $\Phi(0) = 0$  and  $\Phi(1) = 1$ , and  $\Phi$  is increasing, by the intermediate value theorem  $\Phi([0, 1]) = [0, 1]$ . ■

**Lemma 8.7.3.** *Let  $\Phi$  be the Cantor-Lebesgue function, and define  $\Psi : [0, 1] \rightarrow \mathbb{R}$  by*

$$\Psi(x) = \Phi(x) + x \text{ for all } x \in [0, 1]$$

*Then  $\Psi$  is strictly increasing, and maps  $[0, 1]$  onto  $[0, 2]$ . Moreover,  $\Psi$  maps the Cantor set onto a set of positive measure, and maps a measurable subset of the Cantor set onto a nonmeasurable set.*

*Proof.* Since  $\Psi$  is the sum of the strictly increasing function  $f(x) = x$  and the increasing function  $\Phi(x)$ ,  $\Psi$  is strictly increasing. Moreover,  $\Psi$  is continuous since it is the sum of continuous functions, and  $\Psi(0) = 0$ , and  $\Psi(1) = 2$ , so by the intermediate value theorem,  $\Psi([0, 1]) = [0, 2]$ . Finally, notice that since  $\Phi$  and  $f(x) = x$  are 1-1, then  $\Psi$  is also 1-1. Therefore  $\Psi$  has a continuous inverse  $\Psi^{-1}$ . This makes  $\Psi(\mathcal{C})$  closed, and  $\Psi(\mathcal{U})$  open. Therefore both  $\Psi(\mathcal{C})$  and  $\Psi(\mathcal{U})$  are measurable.

Now, let  $\{I_k\}$  be the collection of intervals removed from  $[0, 1]$  to form  $\mathcal{C}$ ; i.e.

$$\mathcal{U} = \bigcup I_k$$

Since  $\Phi$  is constant on each of these intervals, we get that  $\Psi$  maps  $I_k$  onto the translate  $I_k + x$ . Since  $\Psi$  is 1-1, we have that  $\{\Psi(I_k)\}$  is a disjoint collection of measurable sets. Therefore, by countable additivity

$$m\left(\bigcup \Psi(I_k)\right) = \sum m(\Psi(I_k)) = \sum m(I_k + x) = \sum m(I_k) = m(\mathcal{U}) = 1$$

Since  $[0, 2] = \Psi(\mathcal{C}) \cup \Psi(\mathcal{U})$ , we get  $m(\Psi(\mathcal{C})) = 1$ .

Now, by Vitali's theorem, there exists a nonmeasurable subset  $W \subseteq \Psi(\mathcal{C})$ . Notice then that  $\Psi^{-1}(W) \subseteq \mathcal{C}$  is measurable of measure  $m(\Psi^{-1}(W)) = 0$ , since  $m(\mathcal{C}) = 0$ . That is, we have mapped a measurable subset of  $\mathcal{C}$  to a nonmeasurable set. This concludes the proof. ■

**Theorem 8.7.4.** *There exists a measurable subset of the Cantor set which is not Borel.*

# Chapter 9

## Lebesgue Measurable Functions

### 9.1 Properties of Lebesgue Measurable Functions

**Lemma 9.1.1.** *Let  $f$  a realvalued function on a measurable domain  $E$ . Then the following statements are equivalent for any  $c \in \mathbb{R}$ .*

- (1)  $f^{-1}((c, \infty)) = \{x \in E : f(x) > c\}$  is measurable.
- (2)  $f^{-1}([c, \infty)) = \{x \in E : f(x) \geq c\}$  is measurable.
- (3)  $f^{-1}((-\infty, c)) = \{x \in E : f(x) < c\}$  is measurable.
- (4)  $f^{-1}((-\infty, c]) = \{x \in E : f(x) \leq c\}$  is measurable.

*Proof.* It suffices to show that statements (1) and (2) are equivalent. Suppose that (1) holds. Notice that

$$f^{-1}([c, \infty)) = \bigcap_{k=1}^{\infty} \{x \in E : f(x) > c - \frac{1}{k}\}$$

which is a countable intersection of measurable sets, which makes  $f^{-1}([c, \infty))$  measurable. Conversely, if (2) holds, then we have

$$f^{-1}((c, \infty)) = \bigcup_{k=1}^{\infty} \{x \in E : f(x) \geq c - \frac{1}{k}\}$$

which is a countable union of measurable sets, which makes  $f^{-1}((c, \infty))$  measurable. ■

**Corollary.** *If  $f$  is a realvalued function on a measurable domain  $E$ , then for any  $c \in \mathbb{R}$*

$$f^{-1}(c) = \{x \in E : f(x) = c\}$$

*is measurable.*

*Proof.* Take  $c \in \mathbb{R}$ . Then  $f^{-1}([c, \infty))$  and  $f^{-1}((-\infty, c])$  are measurable, hence so is their intersection, which is  $f^{-1}(c)$ . ■

**Corollary.** *If  $f$  is an extended realvalued function, then  $f^{-1}(\infty)$  is measurable.*

*Proof.* Notice that

$$f^{-1}(\infty) = \{x \in E : f(x) > k\}$$

so that  $f^{-1}(\infty)$  is measurable by the above lemma. ■

**Definition.** We call a real-valued function  $f$  on a measurable domain  $E$  **Lebesgue measurable** if for any  $c \in \mathbb{R}$  and  $I$  a bounded interval of the form  $(c, \infty)$ ,  $(-\infty, c)$ ,  $[c, \infty)$ , or  $(-\infty, c]$ , then  $f^{-1}(I)$  is measurable.

**Lemma 9.1.2.** *Let  $f$  be a realvalued function defined on a measurable domain  $E$ . Then  $f$  is Lebesgue measurable if, and only if for any  $U$  open in  $\mathbb{R}$ ,  $f^{-1}(U)$  is measurable.*

*Proof.* Let  $U$  be open in  $\mathbb{R}$ , and  $f^{-1}(U)$  be measurable. Since  $(c, \infty)$  is open, then  $f^{-1}((c, \infty))$  is measurable by hypothesis, which makes  $f$  measurable.

Conversely, suppose that  $f$  is measurable, and let  $U$  be open in  $\mathbb{R}$ . Then

$$U = \bigcup I_k$$

where  $\{I_k\}$  is a countable collection of open bounded intervals. Moreover, we have each  $I_k = A_k \cap B_k$  where  $A_k = (a_k, \infty)$  and  $B_k = (-\infty, b_k)$ . Since  $f$  is measurable, then  $f^{-1}(A_k)$  and  $f^{-1}(B_k)$  are measurable. Hence

$$f^{-1}(U) = f^{-1}\left(\bigcup (A_k \cap B_k)\right) = \bigcup f^{-1}(A_k) \cap f^{-1}(B_k)$$

which is a countable union of measurable sets, which makes  $f^{-1}(U)$  measurable. ■

**Corollary.** *If  $f$  is continuous on its domain, then it is measurable.*

*Proof.* Let  $f$  be continuous on  $E$ , and let  $U$  be open. Then  $f^{-1}(U)$  is open and  $f^{-1}(U) = E \cap V$  with some  $V$  open in  $\mathbb{R}$ . Now,  $E$  and  $V$  are measurable sets, which makes  $f^{-1}(U)$  measurable. ■

**Lemma 9.1.3.** *Monotone functions defined on an interval are measurable.*

**Lemma 9.1.4.** *Let  $f$  and  $g$  be extended realvalued functions on a measurable domain  $E$ . Then the following are true.*

- (1) *If  $f$  is Lebesgue measurable, and  $f = g$  almost everywhere on  $E$ , then  $g$  is Lebesgue measurable.*
- (2) *If  $D$  is a measurable subset of  $E$ , then  $f$  is Lebesgue measurable if, and only if  $f$  is Lebesgue measurable when restricted to  $D$  and to  $E \setminus D$ ; i.e. both  $f|_D$  and  $f|_{E \setminus D}$  are Lebesgue measurable.*

*Proof.* Suppose that  $f$  is measurable and that  $f = g$  almost everywhere on  $E$ . Let  $A = \{x \in E : f(x) \neq g(x)\}$ . Then notice that

$$g^{-1}((c, \infty)) = \{x \in A : g(x) > c\} \cup ((\{x \in E : f(x) > c\}) \cap (E \setminus A))$$

Now, since  $f = g$  almost everywhere on  $E$ , then  $m(A) = 0$ , so that the set  $\{x \in A : g(x) > c\}$  is measurable. Now,  $f^{-1}((c, \infty))$  is measurable since  $f$  is measurable, which makes  $g^{-1}((c, \infty))$  measurable.

Now, let  $D \subseteq E$  a measurable subset. Observe that for every  $c \in \mathbb{R}$

$$f^{-1}((c, \infty)) = \{x \in D : f(x) > c\} \cup \{x \in E \setminus D : f(x) > c\}$$

hence the equivalence is proved. ■

**Theorem 9.1.5.** *Let  $f$  and  $g$  be Lebesgue measurable functions defined on a measurable domain  $E$ , and finite almost everywhere on  $E$ . Then for all  $\alpha, \beta \in \mathbb{R}$*

(1)  $\alpha f + \beta g$  is Lebesgue measurable.

(2)  $fg$  is Lebesgue measurable.

*Proof.* Suppose, without loss of generality, that  $f$  and  $g$  are finite on all of  $E$ . Now, for  $\alpha = 0$ , we have that  $\alpha f = 0$  is measurable. Now, if  $\alpha \neq 0$ , observe that

$$(\alpha f)^{-1}((c, \infty)) = \begin{cases} \{x \in E : f(x) > \frac{c}{\alpha}\}, & \alpha > 0 \\ \{x \in E : f(x) < \frac{c}{\alpha}\}, & \alpha < 0 \end{cases}$$

Now, since  $f$  is measurable, both these sets are measurable. This makes  $\alpha f$  measurable. Now, consider the function  $f + g$ . If  $f(x) + g(x) < c$  for all  $x \in E$ , then  $f(x) < c - g(x)$ . Now, by the density of  $\mathbb{Q}$  in  $\mathbb{R}$ , there is a rational  $q \in \mathbb{Q}$  for which  $f(x) < q < c - g(x)$ . Thus

$$(f + g)^{-1}((c, \infty)) = \bigcup_{q \in \mathbb{Q}} (g^{-1}((-\infty, c - q)) \cap f^{-1}((-\infty, q)))$$

By the countability of  $\mathbb{Q}$  we have a countable union of measurable sets, which makes  $(f + g)^{-1}((c, \infty))$  measurable. Hence  $f + g$  is measurable.

Lastly, notice that  $fg = \frac{1}{2}((f + g)^2 - f^2 - g^2)$ . It suffices to show then that  $f^2$  is measurable. For  $c \geq 0$ , we have that

$$(f^2)^{-1}((c, \infty)) = \{x \in E : f(x) > \sqrt{c}\} \cup \{x \in E : f(x) < -\sqrt{c}\}$$

and for  $c < 0$ , we have  $(f^2)^{-1}((c, \infty)) = \{x \in E : f^2(x) > c\} = E$ . In either case we get measurable sets, which makes  $f^2$  measurable. ■

**Example 9.1.** Consider the strictly increasing function  $\psi$  on  $[0, 1]$  defined by  $\psi(x) = \phi(x) + x$ , where  $\phi$  is the Cantor-Lebesgue function. Then there exists a measurable subset  $A$  of  $[0, 1]$  for which  $\psi(A)$  is nonmeasurable. Now, extend  $\psi$  to the continuous function  $\Psi : \mathbb{R} \rightarrow \mathbb{R}$ . Then  $\Psi^{-1}$  is continuous, and hence measurable. Now, consider the function

$$\chi_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A \end{cases}$$

and take  $f(x) = \chi_A \circ \Psi^{-1}(x)$ . We claim that  $f$  is not a measurable function. Take  $I$  an open interval with  $1 \in I$  and  $0 \notin I$ . Then  $f^{-1}(I) = \Psi(\chi_A^{-1}(I)) = \Psi(A) = \psi(A)$  which is not measurable. Hence  $f$  fails to be a measurable function. In general, compositions of Lebesgue measurable functions need not be Lebesgue measurable.

**Theorem 9.1.6.** *Let  $g$  be a Lebesgue measurable function on a domain  $E$ , and let  $f$  be a continuous realvalued function on all  $\mathbb{R}$ . Then  $f \circ g$  is Lebesgue measurable on  $E$ .*

*Proof.* Let  $U$  be open in  $\mathbb{R}$ . Then  $(f \circ g)^{-1}(U) = g^{-1}(f^{-1}(U))$ . Now, since  $f$  is continuous, we get  $f^{-1}(U)$  is open in  $\mathbb{R}$ , and since  $g$  is measurable, this makes  $g^{-1}(f^{-1}(U))$  measurable by lemma 9.1.2. Therefore  $f \circ g$  is Lebesgue measurable. ■

**Corollary.**  $|f|^p$  is Lebesgue measurable for any  $p > 0$ .

**Lemma 9.1.7.** *Let  $\{f_k\}_{k=1}^n$  be a finite sequence of Lebesgue measurable functions on a common measurable domain  $E$ . Then the functions*

$$\begin{aligned}\bar{f}(x) &= \max \{f_1(x), \dots, f_n(x)\} \\ \underline{f}(x) &= \min \{f_1(x), \dots, f_n(x)\}\end{aligned}$$

*are Lebesgue measurable.*

## 9.2 Simple Approximation

**Definition.** Let  $\{f_n\}$  be a sequence of realvalued functions on a measurable domain  $E$ , and let  $f$  be a function on  $E$ , and let  $A \subseteq E$ . We say that  $\{f_n\}$  **converges pointwise** to  $f$  on  $A$  if

$$\lim_{n \rightarrow \infty} f_n(x) = f(x) \text{ for all } x \in A$$

We say that  $\{f_n\}$  **converges uniformly** to  $f$  on  $A$  provided for all  $\varepsilon > 0$ , there is an  $N \in \mathbb{Z}^+$  for which

$$|f(x) - f_n(x)| < \varepsilon \text{ for all } n \geq N$$

**Lemma 9.2.1.** *Let  $\{f_n\}$  be a sequence of measurable functions on a domain  $E$ , converging pointwise almost everywhere to a realvalued function  $f$  on  $E$ . Then  $f$  is measurable.*

*Proof.* Let  $E_0 \subseteq E$  have measure  $m(E_0) = 0$ , and let  $\{f_n\}$  converge pointwise to  $f$  on  $E \setminus E_0$ . Then  $f$  is measurable if, and only if  $f|_{E_0}$  and  $f|_{E \setminus E_0}$  are measurable. Suppose, then, that  $\{f_n\}$  converges pointwise to  $f$  on all  $E$ .

Now, fix  $c \in \mathbb{R}$  and observe that for every  $x \in E$ , that  $\lim_{n \rightarrow \infty} f_n(x) = f(x)$  as  $n \rightarrow \infty$ . Thus  $f(x) < c$  if, and only if there exists  $n, k \in \mathbb{Z}^+$  for which

$$f_j(x) < c - \frac{1}{n} \text{ for all } j \geq k$$

Now, since  $f_j$  is measurable, we get  $f_j^{-1}((-\infty, c - \frac{1}{n}))$  is measurable, hence the intersection

$$F_k = \bigcap_{j \geq k} f_j^{-1}((-\infty, c - \frac{1}{n}))$$

is measurable. Notice then that  $f^{-1}((-\infty, c)) = \bigcup_k F_k$ . ■

**Definition.** Let  $A$  be any set of  $\mathbb{R}$ . We define the **characteristic function** of  $A$  to be the realvalued function  $\chi_A : \mathbb{R} \rightarrow \{0, 1\}$  defined by

$$\chi_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A \end{cases}$$

**Lemma 9.2.2.** *If  $A$  is a measurable set of  $\mathbb{R}$ , then  $\chi_A$  is Lebesgue measurable.*

**Definition.** We call a realvalued function  $\phi$  defined on a measurable set  $E$  **simple** if it takes only finitely many values. If  $\phi$  takes the values  $c_1, \dots, c_n$ ; i.e.  $\phi(E) = \{c_1, \dots, c_n\}$ , then we define

$$\phi(x) = \sum_{k=1}^n c_k \chi_{E_k}(x) \text{ where } E_k = \phi^{-1}(c_k)$$

the **canonical representation** of  $\phi$ .

**Lemma 9.2.3.** *Simple functions are Lebesgue measurable.*

**Lemma 9.2.4** (The Simple Approximation Lemma). *Let  $f$  be a measurable function bounded on a measurable  $E$ . Then for every  $\varepsilon > 0$ , there exists simple function  $\phi_\varepsilon$  and  $\psi_\varepsilon$  on  $E$  such that*

$$\phi_\varepsilon(x) \leq f(x) \leq \psi_\varepsilon(x) \text{ and } 0 \leq \psi_\varepsilon(x) - \phi_\varepsilon(x) < \varepsilon$$

*Proof.* Let  $(c, d)$  be an open bounded interval containing  $f(E)$ , and let

$$P : \{c = y_0 < \dots < y_n = d\}$$

a partition of the closed bounded interval  $[c, d]$  such that

$$y_{k+1} - y_k < \varepsilon \text{ for all } 0 \leq k \leq n-1$$

Define  $I_k = [y_k, y_{k+1})$  and  $E_k = f^{-1}(I_k)$  for all  $0 \leq k \leq n-1$ . Since each  $I_k$  is an interval, and  $f$  is measurable, then each  $E_k$  is measurable. Now, define

$$\phi_\varepsilon(x) = \sum_{k=1}^n y_k \chi_{E_k} \text{ and } \psi_\varepsilon(x) = \sum_{k=1}^n y_{k+1} \chi_{E_k}$$

Then  $\phi_\varepsilon$  and  $\psi_\varepsilon$  are simple functions on  $E$ . Now, for each  $x \in E$ , we have that  $f(x) \in (c, d)$ , so that there exists a unique  $0 \leq k \leq n-1$  for which  $y_k \leq f(x) < y_{k+1}$ , so that

$$y_k = \phi_\varepsilon(x) \leq f(x) < y_{k+1} = \psi_\varepsilon(x)$$

Moreover, we have that  $y_{k+1} - y_k = \psi_\varepsilon(x) - \phi_\varepsilon(x) < \varepsilon$ . ■

**Theorem 9.2.5** (The Simple Approximation Theorem). *An extended realvalued function  $f$  on a measurable domain  $E$  is measurable if, and only if there exists a sequence  $\{\phi_n\}$  of simple functions on  $E$ , converging pointwise to  $f$  such that*

$$|\phi_n(x)| \leq |f(x)| \text{ for all } n \in \mathbb{Z}^+$$

*Proof.* Since simple functions are measurable, if  $\{\phi_n\}$  converges pointwise to  $f$  on  $E$ , then  $f$  must also be measurable by lemma 9.2.1.

Now, suppose that  $f$  is measurable, and that  $f(x) \geq 0$  for all  $x \in E$ . Let  $n \in \mathbb{Z}^+$ , and define  $E_n = f^{-1}((-\infty, n])$ . Then  $E_n$  is measurable, and  $f|_{E_n}$  is a nonnegative bounded measurable function. By the simple approximation lemma, choosing  $\varepsilon = \frac{1}{n}$ , select simple functions  $\phi_n$  and  $\psi_n$  on  $E$  such that

$$\phi_n(x) \leq f(x) \leq \psi_n(x) \text{ and } 0 \leq \psi_n(x) - \phi_n(x) < \frac{1}{n} \text{ on } E_n$$

Now, extend  $\phi_n$  to a function  $\Phi_n$  on all of  $E$  such that  $\Phi_n(x) = n$  if  $f(x) > n$ . Then  $\Phi_n$  is simple, on  $E$  and

$$0 \leq \Phi_n(x) \leq f(x) \text{ on } E$$

Now, we claim that the sequence  $\{\psi_n\}$  converges to  $f$  pointwise on  $E$ . Indded, let  $x \in E$ , if  $f(x)$  is finite, choose an  $N \in \mathbb{Z}^+$  such that  $f(x) < N$ , then

$$0 \leq f(x) - \Phi_n(x) < \frac{1}{n} \text{ for all } n \geq N$$

so that  $\{\psi_n\} \rightarrow f$  as  $n \rightarrow \infty$ . Now, if  $f(x) = \infty$ , then  $\Phi_n(x) = n$  for all  $n \in \mathbb{Z}^+$  so that  $\{\phi_n\} \rightarrow f$  as  $n \rightarrow \infty$ . ■

### 9.3 Theorems of Littlewood, Egoroff, and Lusin

**Lemma 9.3.1.** *Let  $E$  be a measurable set having finite measure, and let  $\{f_n\}$  a sequence of measurable functions on  $E$  converging pointwise to a realvalued function  $f$  on  $E$ . Then for all  $\eta > 0$  and  $\delta > 0$ , there exists a measurable subset  $A$  of  $E$  and an  $N \in \mathbb{Z}^+$  for which*

$$|f(x) - f_n(x)| < \eta \text{ on } A \text{ for all } n \geq N \text{ and } m(E \setminus A) < \delta$$

*Proof.* Since  $f$  is realvalued and measurable (by lemma 9.2.1), the set

$$\{x \in E : |f(x) - f_n(x)| < \eta\}$$

is measurable. Hence, the union

$$E_n = \{x \in E : |f(x) - f_n(x)| < \eta \text{ for all } k \geq n\}$$

is measurable. Then  $\{E_n\}$  is an ascending collection of measurable sets with

$$E = \bigcup E_n$$

Now, since  $\{f_n\}$  converges to  $f$  pointwise, by the continuity of measure, we conclude that

$$m(E) = \lim_{n \rightarrow \infty} m(E_n)$$

Since  $m(E)$  is finite, choose an  $N \in \mathbb{Z}^+$  for which  $m(E_N) > m(E) - \varepsilon$ , and define  $A = E_N$ . Then by excision, we get  $m(E \setminus A) = m(E) - m(A) < \varepsilon$ . ■



**Theorem 9.3.2** (Egoroff's Theorem). *Let  $E$  be a measurable set with finite measure, and let  $\{f_n\}$  a sequence of measurable functions on  $E$  converging pointwise to a realvalued function  $f$  on  $E$ . Then for each  $\varepsilon > 0$ , there exists a closed subset  $F$  of  $E$  for which  $\{f_n\}$  converges uniformly to  $f$  on  $F$ , and  $m(E \setminus F) < \varepsilon$ .*

*Proof.* For every  $n \in \mathbb{Z}^+$ , let  $A_n$  be a measurable subset of  $E$ , and let  $N(n) \in \mathbb{Z}^+$  such that

$$|f(x) - f_k(x)| < \frac{1}{n} \text{ on } A_n \text{ for all } k \geq N(n) \text{ and where } m(E \setminus A_n) < \frac{\varepsilon}{2^{n+1}}$$

Define

$$A = \bigcap_{n=1}^{\infty} A_n$$

Then by DeMorgan's laws and countable subadditivity

$$m(E \setminus A) = \sum_{n=1}^{\infty} m(E \setminus A_n) < \sum_{n=0}^{\infty} \frac{\varepsilon}{2^{n+1}} < \frac{\varepsilon}{2}$$

Now, let  $\varepsilon > 0$  and choose an  $n_0 \in \mathbb{Z}^+$  such that  $\frac{1}{n_0} < \varepsilon$ . Then

$$|f(x) - f_k(x)| < \frac{1}{n_0} \text{ on } A \text{ for all } k \geq N(n_0)$$

Now, since  $A \subseteq A_{n_0}$ , and our hypothesis on  $n_0$ , we get

$$|f(x) - f_k(x)| < \varepsilon \text{ on } A \text{ for all } k \geq N(n_0)$$

so that  $\{f_n\}$  converges uniformly to  $f$  on  $A$  with  $m(E \setminus A) < \frac{\varepsilon}{2}$ . Now, by the inner approximation theorem, choose an  $F \subseteq A$ , closed, for which  $m(A \setminus F) < \frac{\varepsilon}{2}$ . Then  $m(E \setminus F) < m(E \setminus A) + m(A \setminus F) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ . So that  $\{f_n\}$  converges uniformly to  $f$  on  $F$ . ■

**Theorem 9.3.3** (Littlewood's Theorem). *Let  $f$  be a simple function on a measurable set  $E$ . Then for every  $\varepsilon > 0$ , there exists a continuous function  $g$  on  $\mathbb{R}$  and a closed subset  $F$  of  $E$  such that  $f = g$  on  $F$  and  $m(E \setminus F) < \varepsilon$ .*

*Proof.* Let  $\{a_1, \dots, a_n\}$  the distinct values taken on by  $f$ , taken on the sets  $E_1, \dots, E_n$ , respectively. Then  $\{E_k\}_{k=1}^n$  is a finite disjoint collection of measurable sets whose union is  $E$ , since each  $a_i$  is disjoint. Now, choose closed sets  $F_1, \dots, F_n$  such that for all  $1 \leq k \leq n$ ,  $F_k \subseteq E_k$  and  $m(E_k \setminus F_k) < \frac{\varepsilon}{n}$ . Then

$$F = \bigcup_{k=1}^n F_k$$

is closed and since  $\{E_k\}_{k=1}^n$  is a disjoint collection, we get

$$m(E \setminus F) = \sum_{k=1}^n m(E_k \setminus F_k)$$

Now, define  $g : F \rightarrow \mathbb{R}$  to take  $g(x) = a_k$  for all  $x \in F_k$ , for all  $1 \leq k \leq n$ . Since  $\{F_k\}$  is a disjoint collection,  $g$  is well defined. Moreover,  $g$  is continuous on  $F$ . Then for every  $x \in F_i$ , there exists an interval  $I$ , containing  $x$ , and disjoint from the union  $\bigcup_{k \neq i} F_k$ . Thus  $g$  is constant on  $F \cap I$ . Now, extend  $g$  to a continuous function  $G : \mathbb{R} \rightarrow \mathbb{R}$ . Since  $G(x) = a_k$  on each  $F_k$ , and  $m(E_k \setminus F_k) < \frac{\varepsilon}{n}$ , we get  $f = G$  on  $F$  where  $m(E \setminus F) < \varepsilon$ . ■

**Theorem 9.3.4** (Lusin's Theorem). *Let  $f$  be a measurable function on a measurable set  $E$ . Then for every  $\varepsilon > 0$ , there exists a continuous function  $g$  on  $\mathbb{R}$  and a closed subset  $F$  of  $E$  such that  $f = g$  on  $F$  and  $m(E \setminus F) < \varepsilon$ .*

*Proof.* Consider first when  $E$  has finite measure. By the simple approximation theorem (theorem 9.2.5), there exists a sequence  $\{f_n\}$  of simple functions on  $E$ , converging pointwise to  $f$  on  $E$ . Let  $n \in \mathbb{Z}^+$ , then by Littlewood's theorem, choose a  $g_n$  continuous on  $\mathbb{R}$  and a closed subset  $F_n$  of  $E$  for which  $f = g_n$  on  $F_n$  and  $m(E \setminus F_n) < \frac{\varepsilon}{2^{n+1}}$ . By Egoroff's theorem, there exists a closed set  $F_0 \subseteq E$  such that  $\{f_n\}$  converges uniformly to  $f$  on  $F_0$ , with  $m(E \setminus F_0) < \frac{\varepsilon}{22}$ . Define then

$$F = \bigcap_{n=0}^{\infty} F_n$$

Then  $F$  is closed and by DeMorgan's laws,

$$m(E \setminus F) \leq \frac{\varepsilon}{2} + \sum_{n=0}^{\infty} \frac{\varepsilon}{2^{n+1}} = \varepsilon$$

Now, each  $f_n$  is continuous on  $F$ , and since  $F \subseteq F_n$ , and  $f = g_n$  on  $F_n$ , we get  $\{f_n\}$  converges uniformly to  $f$  on  $F$ , since  $F \subseteq F_0$ . Moreover,  $f|_F$  is continuous on  $F$ , thus choose a continuous function  $g : \mathbb{R} \rightarrow \mathbb{R}$  such that  $f = g$  on  $F$ , and we are done. ■

# Bibliography

- [1] D. Dummit, *Abstract algebra*. Hoboken, NJ: John Wiley & Sons, Inc, 2004.
- [2] I. N. Herstein, *Topics in algebra*. New York: Wiley, 1975.