

# Analysis

Alec Zabel-Mena

**Text**

Principles of Mathematical Analysis ( $3^{rd}$  edition)

Walter Rudin

September 23, 2022



# 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:

- (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.** (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 arbitraritirly 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.** (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:

- (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  $x, y, z \in F$ :*

- (1)  $x + y = x + z$  implies  $y = z$
- (2)  $x + y = x$  implies  $y = 0$
- (3)  $x + y = 0$  implies  $y = -x$
- (4)  $-(-x) = x$ .

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

- (1)  $xy = xz$  implies  $y = z$
- (2)  $xy = x$  implies  $y = 1$
- (3)  $xy = 1$  implies  $y = x^{-1}$
- (4)  $(x^{-1})^{-1} = x$ .

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

- (1)  $0x = 0$
- (2)  $x \neq 0$  and  $y \neq 0$  implies  $xy \neq 0$
- (3)  $(-x)y = -(xy) = x(-y)$
- (4)  $(-x)(-y) = xy$ .

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

- (1)  $x + y < x + z$  whenever  $y < z$ , for  $x, y, z \in F$
- (2)  $xy > 0$  whenever  $x > 0$  and  $y > 0$ , 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:*

- (1)  $x > 0$  implies  $-x < 0$ .
- (2) If  $x > 0$  and  $y < z$ , then  $xy < xz$ .
- (3) If  $x < 0$  and  $y < z$ , then  $xz < xy$ .
- (4) If  $x \neq 0$ , then  $x^2 > 0$ , in particular,  $1 > 0$ .
- (5)  $0 < x < y$  implies that  $0 < y^{-1} < x^{-1}$ .

*Proof.* (1) If  $x > 0$ , then  $0 = x + (-x) > 0 + (-x)$ , 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:

- (1) If  $x \in \mathbb{R}$ , then  $x + \infty = \infty$ ,  $x - \infty = -\infty$ , and  $\frac{x}{\infty} = \frac{x}{-\infty} = 0$ .
- (2) If  $x > 0$ , then  $x(\infty) = \infty$  and  $x(-\infty) = -\infty$ .
- (3) 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  $(a, 0)(b, 0) = (ab, 0)$ .

*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

- (1)  $\overline{z + w} = \bar{z} + \bar{w}$ .
- (2)  $\overline{zw} = \bar{z}\bar{w}$ .
- (3)  $z + \bar{z} = 2 \operatorname{Re} z$  and  $z - \bar{z} = 2i \operatorname{Im} z$ .

(4)  $z\bar{z}$  is a nonnegative real number.

*Proof.* Let  $z = a + ib$ , and let  $w = c + id$ . Then  $z + w = (a + c) + i(b + d)$ , so  $\overline{z + w} = (a + c) - i(b + d) = (a - ib) + (c - id) = \bar{z} + \bar{w}$ ; 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:

(1)  $|z| \geq 0$  and  $|z| = 0$  if and only if  $z = 0$ .

(2)  $|\bar{z}| = |z|$ .

(3)  $|zw| = |z||w|$ .

(4)  $\operatorname{Re} z \leq |z|$ .

(5)  $|z + w| \leq |z| + |w|$ .

*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

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

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  $\langle \cdot, \cdot \rangle : \mathbb{R}^k \mathbb{R}^k \rightarrow \mathbb{R}$  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:

- (1)  $\|x\| \geq 0$  and  $\|x\| = 0$  if and only if  $x_i = 0$  for all  $1 \leq i \leq k$ .
- (2)  $\|\alpha x\| = |\alpha| \|x\|$ .
- (3)  $|\langle x, y \rangle| \leq \|x\| \|y\|$ .
- (4)  $\|x + y\| \leq \|x\| + \|y\|$ , and  $\|x - z\| \leq \|x - y\| + \|y - z\|$

*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. ■

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  $x, y \in X$

- (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  $y, z \in B_r(x)$ , 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  $0 < \lambda < 1$ ,  $a_i \leq \lambda x_i + (1 - \lambda)y_i \leq \lambda b_i + (1 - \lambda)b_i = b_i$ . ■

**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  $d(y, s) < h$ , we have  $d(x, s) \leq d(x, y) + d(y, s) < r - h + h = r$ , thus  $s \in E$ , 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.** (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  $N \cap E = \emptyset$ , 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:*

- (1) If  $\{G_\alpha\}$  is a collection of open sets, then  $\bigcup G_\alpha$  is open.
- (2) If  $\{G_i\}_{i=1}^n$  is a finite collection of open sets, then  $\bigcap_{i=1}^n G_i$  is open.
- (3) If  $\{G_\alpha\}$  is a collection of closed sets, then  $\bigcap G_\alpha$  is closed.
- (4) If  $\{G_i\}_{i=1}^n$  is a finite collection of closed sets, then  $\bigcup_{i=1}^n G_i$  is closed.

*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

- (1)  $\overline{x}$  is closed.
- (2)  $E$  is closed if and only if  $E = \overline{E}$ .
- (3) If  $F \subseteq X$  such that  $E \subseteq F$ , and  $F$  is closed, then  $\overline{E} \subseteq F$ .

*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  $V_p \cap YE$ , thus  $G \cap YE$ , 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  $v_p \in G$ , thus  $V_p \cap Y \subseteq E$ , 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  $K \subseteq \bigcup_{i=1}^n V_i$ , where  $V_i$  is a neighborhood of  $y_i$  for  $1 \leq i \leq n$ . Let  $U = \bigcap_{i=1}^n U_i$ , then  $V \cap W$  is empty, hence  $U \cap V = \emptyset$ , therefore,  $x \in X \setminus K$ , 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 a 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 at least 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:*

- (1)  $E$  is closed and bounded.
- (2)  $E$  is compact.
- (3) Every infinite subset of  $E$  has a limit point in  $E$ .

*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 gave 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  $U_{n+1} \subseteq U_n$ ,  $x_n \notin U_{n+1}$ , and  $U_{n+1} \cap P$  is nonempty. Therefore there is a nonempty  $K_n = U_n \cap P$ . Since  $\hat{U}_n$  is close 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 remove 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  $I_n \subset S$ . 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  $\epsilon > 0$ , there is an  $N \in \mathbb{Z}^+$  such that  $d(x_n, x) < \epsilon$  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  $\mathbb{C}$ .

- (1)  $\{\frac{1}{n}\}$  is bounded, and  $\lim_{n \rightarrow \infty} \frac{1}{n} = 0$ .
- (2) The sequence  $\{n^2\}$  is unbounded and diverges.
- (3)  $1 + \frac{(-1)^n}{n} \rightarrow 1$  as  $n \rightarrow \infty$ , and  $\{1 + \frac{(-1)^n}{n}\}$  is bounded.
- (4)  $\{i^n\}$  is bounded and divergent.
- (5)  $\{1\}$  is bounded and converges to 1.

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

- (1)  $\{x_n\}$  converges to  $x \in X$  if and only if every neighborhood of  $x$  contains  $x_n$  for all but finitely many  $n$ .
- (2) If  $\{x_n\}$  converges to  $x$ , and  $x'$ , then  $x = x'$ .
- (3) If  $\{x_n\}$  converges, then  $x_n$  is bounded.
- (4) 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  $\epsilon > 0$ , there is an  $N \in \mathbb{Z}^+$  for which  $d(x_n, x) < \epsilon$ , 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  $\epsilon > 0$ , we have  $d(x, x_n) < \epsilon$ , hence  $x_n \rightarrow x$ .

Let  $\epsilon > 0$ , then there are  $N_1, N_2 \in \mathbb{Z}^+$  such that  $d(x_n, x) < \frac{\epsilon}{2}$ , and  $d(x_n, x') < \frac{\epsilon}{2}$ . Then choosing  $N = \max\{N_1, N_2\}$ , and letting  $\epsilon$  be arbitrarily small, we have  $d(x, x') \leq d(x, x_n) + d(x_n, x') < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$ ; 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  $N > \frac{1}{\epsilon}$ , then whenever  $n \geq N$ ,  $d(x, x_n) < \epsilon$ ; 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  $n \rightarrow \infty$ :

$$(1) \lim (x_n + y_n) = \lim x_n + \lim y_n = x + y.$$

$$(2) \lim x_n y_n = \lim x_n \lim y_n = xy.$$

$$(3) \lim \frac{x_n}{y_n} = \frac{\lim x_n}{\lim y_n} = \frac{x}{y}; \text{ given that } y_n, y \neq 0.$$

*Proof.* (1) Let  $\epsilon > 0$ , then for  $N_1, N_2 \in \mathbb{Z}^+$ ,  $|x_n - x| < \frac{\epsilon}{2}$  and  $|y_n - y| < \frac{\epsilon}{2}$ . Then choose  $N = \max\{N_1, N_2\}$ , then whenever  $n \geq N$ , we have  $|(x_n + y_n) - (x + y)| \leq |x_n - x| + |y_n - y| < \epsilon$ .

(2) Notice that  $x_n y_n - xy = (x_n - x)(y_n - y) + x(y_n - y) + y(x_n - x)$ , then for  $N_1, N_2 \in \mathbb{Z}^+$ ,  $|x_n - x| < \sqrt{\epsilon}$ , and  $|y_n - y| < \sqrt{\epsilon}$ . Then choosing  $N = \max\{N_1, N_2\}$ , then  $|(x_n - x)(y_n - y)| < \epsilon$ , thus we have  $|x_n y_n - xy| \leq |(x_n - x)(y_n - y)| + |x(y_n - y)| + |y(x_n - x)| < \epsilon$ .

(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  $\epsilon > 0$ , there is an  $N > m$  such that whenever  $n \geq N$ ,  $|y_n - y| < \frac{1}{2}|y|^2 \epsilon$ . Then  $|\frac{1}{y_n} - \frac{1}{y}| \leq \frac{|y_n - y|}{|y_n y|} < \frac{2}{|y|^2} |y_n - y| < \epsilon$ . Then choosing the sequences  $\{x_n\}$  and  $\{\frac{1}{y_n}\}$ , the rest follows. ■

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

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

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  $\epsilon > 0$  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  $|x_n - a| < \epsilon$  and  $|y_n - a| < \epsilon$  for  $n \geq N_1, N_2$ . Now choose  $N = \max\{N_0, N_1, N_2\}$ , if  $n \geq N$ , we have  $-\epsilon < x_n - a < \epsilon$ , and we also have  $x_n - a < w_n - a < y_n - a$ , thus we have that:

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

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

**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  $\epsilon$  there is an  $N \in \mathbb{Z}^+$  such that for  $n \geq N$ ,  $|x_n - 0| < \frac{\epsilon}{M}$ . For  $|x_n y_n - 0| = |x_n y_n| < M|x_n| < M \frac{\epsilon}{M} = \epsilon$ . 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  $\{k\} \subseteq \mathbb{Z}^+$ . Then for  $\epsilon > 0$ , there is a  $K \in \mathbb{Z}^+$  for which  $d(x_{n_k}, x) < \frac{\epsilon}{2}$  for  $k \geq K$ . Let  $N \in \mathbb{Z}^+$ , and choose  $n \geq \max\{N, K\}$ , then  $d(x_n, x) \leq d(x_n, x_{n_k}) + d(x_{n_k}, x) < \epsilon$ . ■

**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  $d(x, x_{n_i}) < \frac{1}{i}$  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  $\epsilon > 0$ , there is an  $N \in \mathbb{Z}^+$  such that  $d(x_n, x_m) < \epsilon$  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  $d(x_n, x_{N_1}) < \frac{\epsilon}{2}$ , and  $d(x_{N_2}, x_m) < \frac{\epsilon}{2}$ . Then we see that  $d(x_{N_1}, x_{N_2}) \leq d(x_{N_1}, x_n) + d(x_n, x_{N_2}) < \epsilon$ , 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  $x_n, x_m \in S$ ,  $d(x_n, 0) < \frac{\epsilon}{2}$  and  $d(0, x_m) < \frac{\epsilon}{2}$  implies that  $d(x_n, x_m) \leq d(x_n, 0) + d(0, x_m) < \epsilon$ , whenever  $n, m > N$ , for  $\epsilon > 0$ . ■

**Theorem 3.3.2.** (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  $\epsilon > 0$ , and choose  $x, y \in \hat{E}$ , then there are points  $x', y' \in E$  such that  $d(x, x') < \frac{\epsilon}{2}$  and  $d(y, y') < \frac{\epsilon}{2}$ . Hence,  $d(x, y) \leq d(x, x') + d(x', y') + d(y', y) < \epsilon + \text{diam } E$ , then choosing  $\epsilon$  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.** (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.* (1) If  $x_n \rightarrow x$ , and  $\epsilon > 0$  such that there is an  $N \in \mathbb{Z}^+$  such that  $d(x_n, x) < \frac{\epsilon}{2}$  for all  $n \geq N$ , then for  $m \geq N$ , we have  $d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \epsilon$ . 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  $\epsilon > 0$ , there is an  $N_0 \in \mathbb{Z}^+$  for which  $\text{diam } \hat{E} < \epsilon$ . Then for all  $x_n \in \hat{E}$ ,  $d(x_n, x) < \epsilon$  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  $\epsilon > 0$ , there is an  $N \in \mathbb{Z}^+$  for which  $x - \epsilon < x_N \leq x$ . Then for  $n \geq N$ ,  $x_n \rightarrow x$ . The converse follows from theorem 3.1.2. ■

## 3.4 Upper and Lower Limits.

Let  $\{x_n\}$  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:*

(1)  $s, s' \in E$ .

- (2) If  $x > s$ , and  $x' > s'$ , there is an  $N \in \mathbb{Z}^+$  such that  $n \geq N$  implies that  $x' < x_n < x$ .

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

- (1) 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  $x_n \rightarrow -\infty$ .
- (2) 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.** (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  $n \rightarrow \infty$ .

- (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.* (1) Let  $n > [p]_{\frac{1}{\epsilon}}$ ; then  $|\frac{1}{n^p}| < \epsilon$ .

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

- (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  $\epsilon > 0$ , there is a  $\delta > 0$  for which  $d_Y(f(x), q) < \epsilon$ , 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.** (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  $\epsilon > 0$ , there is a  $\delta > 0$  such that  $|f(x) - L| < \epsilon$  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  $\epsilon > 0$ , there is a  $\delta > 0$  such that  $|f(x) - L| < \epsilon$  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  $\epsilon > 0$ , there is a  $\delta > 0$  such that  $\|f(x) - L\| < \epsilon$  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  $\epsilon > 0$ , there is a  $\delta > 0$  such that  $d_Y(f(x), q) < \epsilon$  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  $d_Y(f(x_n), q) < \epsilon$  whenever  $d_X(x_n, p) < \delta$ .

Conversely, suppose that  $\lim f \neq q$ , that is for some  $\epsilon > 0$ ,  $d_Y(f(x), q) > \epsilon$  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  $Y$ :

- (1)  $f + g(x) = f(x) + g(x)$ .
- (2)  $fg(x) = f(x)g(x)$ .
- (3)  $(\lambda f)(x) = \lambda f(x)$  for  $\lambda \in X$ .
- (4)  $\frac{f}{g}(x) = \frac{f(x)}{g(x)}$ , provided that  $g(x) \neq 0$ .

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  $x \rightarrow p$ .*

- (1)  $\lim (f + g) = \lim f + \lim g = A + B$ .
- (2)  $\lim fg = \lim f \lim g = AB$ .
- (3)  $\lim \frac{f}{g} = \frac{\lim f}{\lim g} = \frac{A}{B}$ , provided that  $B \neq 0$ .

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

- (1)  $\lim \lambda f = \lambda \lim f = \lambda A$ , and  $\lim (\lambda + f) = \lambda + \lim f = \lambda + A$ .
- (2)  $\lim \frac{1}{f(x)} = \frac{1}{\lim f} = \frac{1}{A}$ , provided that  $A \neq 0$ .

**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  $\epsilon > 0$ , there is a  $\delta > 0$  such that  $d_Y(f(x), f(p)) < \epsilon$  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  $p$ .*

*Proof.* For every  $\epsilon > 0$ , we have  $\delta_1, \delta_2 > 0$  such that  $d_Y(f(x), f(p)) < \epsilon$ , when  $0 < d_X(x, p) < \delta_1$ , and  $d_Z(g(y), g(f(p))) < \epsilon$  whenever  $d_Y(y, f(p)) < \delta_2$ . Then choose  $\delta = \min\{\delta_1, \delta_2\}$ , and we see that  $d_Z(g(f(x)), g(f(p))) < \epsilon$  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  $\epsilon > 0$  such that  $y \in V$  when  $d_Y(y, f(p)) < \epsilon$ . Since  $f$  is continuous, there is a  $\delta > 0$  such that  $d_Y(f(x), f(p)) < \epsilon$ , 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  $\epsilon > 0$ , and let  $V = \{y \in Y : d_Y(y, f(p)) < \epsilon\}$ ;  $V$  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  $d_Y(f(x), f(p)) < \epsilon$ ; 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.** (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  $\epsilon > 0$ , there is a  $\delta > 0$  such that  $d_Y(f(q), f(p)) < \epsilon$ , 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  $\epsilon > 0$ , 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  $d(q, p_m) \leq d_X(q, p) + d_X(p, p_m) < \delta + \frac{1}{2}\phi(p_m) = \phi(p_m)$ , for  $1 \leq m \leq n$ . Therefore,  $d_Y(f(p), f(q)) \leq d_Y(f(p), f(p_m)) + d_Y(f(p_m), f(q)) < \epsilon$ . 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:*

- (1) *There exists a continuous function on  $E$  which is not bounded.*
- (2) *There is a bounded, continuous function on  $E$  which has no maximum.*
- (3) *If  $E$  is bounded, there exists a continuous function on  $E$  that is not uniformly continuous.*

*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  $\epsilon > 0$  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  $|f(x) - f(t)| \geq \epsilon$ , 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  $E$  ( $0 < g \leq 1$ ), 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  $x$  is such that  $a < x \leq b$ , 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.** (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:*

- (1) *If  $f$  is monotonically increasing  $f(x-) \leq f(x) \leq f(x+)$*
- (2) *If  $f$  is monotonically decreasing  $f(x+) \leq f(x) \leq f(x-)$*

*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  $\epsilon > 0$ , then there is a  $\delta > 0$  for which  $a < x - \delta < x$ , and  $A - \epsilon < f(x - \delta) \leq A$ . Then we have  $f(x - \delta) < f(t) \leq A$  for all  $x - \delta < t < x$ , then we get  $|f(t) - A| < \epsilon$ , 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 atmost 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

- (1)  $f$  is monotonically increasing on  $(a, b)$ .
- (2)  $f$  is discontinuous at every point in  $E$  with  $f(x_n+) - f(x_n-) = c_n$ .
- (3)  $f$  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  $t \neq x \in V \cap E$ .

**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  $t \rightarrow x$ .

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

*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  $t \rightarrow x$ :*

$$(1) (f + g)' = f' + g'.$$

$$(2) (fg)' = f'g + fg'.$$

$$(3) \left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}, \text{ provided that } g(x) \neq 0.$$

*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.** (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.** (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  $x \in (a, b)$ :*

- (1) *If  $f' \geq 0$ , then  $f$  is monotonically increasing.*
- (2) *If  $f' = 0$ , then  $f$  is constant.*
- (3) *If  $f' \leq 0$ , then  $f$  is monotonically decreasing.*

*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:

- (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  $g''(x_2) = 0$  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  $f' : (a, b) \rightarrow \mathbb{R}$  such that the values of  $f'$  satisfy equation (5.3)

*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.** (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  $\epsilon > 0$ , there is a partition  $P$  such that:*

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

*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  $\epsilon > 0$ , we assume that  $U(f, P, \alpha) - L(f, P, \alpha) < \epsilon$ , then we get that:

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

implying integrability for small enough  $\epsilon$ .

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

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

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

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

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

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

*Proof.* (1) Theorem 7.1.3 implies this.

- (2) If  $U(f, P, \alpha) - L(f, P, \alpha) < \epsilon$ , 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  $\sum |f(s_i) - f(t_i)| \Delta \alpha_i \leq U(f, P, \alpha) - L(f, P, \alpha) < \epsilon$
- (3) We have  $L(f, P, \alpha) \leq \sum f(t_i) \Delta \alpha_i \leq U(f, P, \alpha)$ , and  $L(f, P, \alpha) \leq \int f d\alpha \leq U(f, P, \alpha)$ , 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  $\epsilon > 0$ , and choose  $\xi > 0$  such that  $(\alpha(b) - \alpha(a))\xi < \epsilon$ . 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  $U(f, P, \alpha) - L(f, P, \alpha) = \sum (M_i - m_i) \Delta \alpha_i \leq \xi(\alpha(b) - \alpha(a)) < \epsilon$ , 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  $\epsilon > 0$ , 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:

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

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  $\epsilon > 0$  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  $\sum \alpha(u_j) - \alpha(v_j) < \epsilon$ .

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  $|f(s) - f(t)| < \epsilon$  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  $M_i - m_i \leq \epsilon$ . Thus for  $\epsilon$  small enough, we have  $U(f, P, \alpha) - L(f, P, \alpha) \leq (\alpha(b) - \alpha(a))\epsilon + 2M\epsilon$ , 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  $\epsilon > 0$ . We have that  $\phi$  is uniformly continuous on  $[m, M]$ , so there is a  $\delta > 0$  for which  $|\phi(s) - \phi(t)| < \epsilon$  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  $M'_i - m'_i \leq \epsilon$ , 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

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

Then for  $\epsilon$  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:*

(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.* (1) 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  $\epsilon > 0$ , there are partitions  $P_1, P_2$ , with  $P = P_1 \cup P_2$  such that  $U(f, P_1, \alpha) - L(f, P_1, \alpha) < \frac{\epsilon}{2}$ , and  $U(g, P_2, \alpha) - L(g, P_2, \alpha) < \frac{\epsilon}{2}$ . Then we get  $U(f + g, P, \alpha) - L(f + g, P, \alpha) < \epsilon$ , so  $f + g$  is intergrable. Moreover, we have that  $U(f, P, \alpha) < \int f + \frac{\epsilon}{2}$  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  $\epsilon > 0$  such that  $U(f, P, \alpha) - L(f, P, \alpha) < \epsilon$ , 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  $U(f, P_1, \alpha) - L(f, P_2, \alpha) < \epsilon$  and  $U(f, P_2, \alpha) - L(f, P, \alpha) < \epsilon$ , thus  $f$  is integrable on  $[a, c]$  and on  $[c, b]$ .

Now we have that  $U(f, P_1, \alpha) < \int_a^c f + \frac{\epsilon}{2}$  and  $\int_c^b f + \frac{\epsilon}{2}$ , 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  $\epsilon > 0$ . Since  $f$  is integrable, we have that

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

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

(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  $U(f, P_1, \alpha_1) - L(f, P_1, \alpha_1) < \frac{\epsilon}{2}$  and  $U(f, P_2, \alpha_2) - L(f, P_2, \alpha_2) < \frac{\epsilon}{2}$ , then we have by the previous inequalities that:

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

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  $\int f(\alpha_1 + \alpha_2) = \int f d\alpha_1 + \int f d\alpha_2$ . ■



**Corollary.**  $\int cf d\alpha = c \int f d\alpha$  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:*

- (1)  $fg$  is integrable with respect to  $\alpha$  on  $[a, b]$ .
- (2)  $|f|$  is integrable with respect to  $\alpha$  on  $[a, b]$  and:

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

*Proof.* (1) 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  $\left| \int f \right| = c \int f = \int cf \leq \int |f|$ .

■

**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  $\epsilon > 0$ , and let  $P$  be a partition of  $[a, b]$  such that  $U(\alpha', P) - L(\alpha', P) < \epsilon$ . 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  $\sum |\alpha'(s_i) - \alpha'(t_i)|\Delta x_i < \epsilon$ , 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

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

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

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

for small enough  $\epsilon$ ,  $\overline{\int} f d\alpha = \overline{\int} f\alpha' dx$ . 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  $U(g, Q, \beta) - L(g, Q, \beta) < \epsilon$  for  $\epsilon > 0$ , 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  $\epsilon > 0$ , we see that  $|F(y) - f(x)| < \epsilon$ , whenever  $|y - x| < \frac{\epsilon}{M}$ . Thus  $f$  is continuous at  $x_0$ . Let  $\epsilon > 0$  and  $\delta > 0$  such that  $|f(t) - f(x_0)| < \epsilon$  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:

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

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  $\epsilon > 0$ , and construct a partition  $P$  of  $[a, b]$  such that  $U(f, P) - L(f, P) < \epsilon$ . 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:

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

■

**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

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

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  $\|\gamma(x_i) - \gamma(x_{i-1})\| = \|\int_{x_{i-1}}^{x_i} \gamma' dt\| \leq \int_{x_{i-1}}^{x_i} \|\gamma'\| dt$ , 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  $\epsilon > 0$ , since  $\gamma'$  is uniformly continuous on  $[a, b]$ , there is a  $\delta > 0$  such that  $\|\gamma'(s) - \gamma'(t)\| < \epsilon$  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 let  $x_{i-1} \leq t \leq x_i$ , we have  $\|\gamma'(t)\| \leq \|\gamma'(x_i)\| + \epsilon$ , hence:

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

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

# Chapter 8

## Lebesgue Theory

### 8.1 $\sigma$ -algebras and Measure.

**Definition.** Let  $X$  be a set and  $\mathcal{A} = \{A_\alpha\}$  a collection of subsets of  $X$ . We call  $\mathcal{A}$  a  $\sigma$ -algebra if:

- (1)  $\emptyset \in \mathcal{A}$ .
- (2)  $X \setminus A_\alpha \in \mathcal{A}$  for any  $\alpha$ .
- (3)  $\mathcal{A}$  is closed under countable unions.

**Theorem 8.1.1.** Let  $X$  be a set and  $\mathcal{A}$  a collection of subsets of  $X$ . Then there exists a smallest  $\sigma$ -algebra containing  $\mathcal{A}$ .

**Definition.** The collection  $\mathcal{B}$  of **Borel sets** of real numbers is the smallest  $\sigma$ -algebra containing all open sets of  $\mathbb{R}$ . We call  $\mathcal{B}$  the **Borel algebra**.

**Definition.** Let  $I$  be a nondegenerate interval on  $\mathbb{R}$ . We define the **length** of  $I$ ,  $l(I)$  to be the difference of its endpoints if  $I$  is bounded and  $l(I) = \infty$  otherwise.

**Definition.** For a set  $A \subseteq \mathbb{R}$  consider a countable collection  $\{I_n\}$  of intervals that covers  $A$ . We define the **outer measure** of  $A$  to be the map  $m^* : 2^{\mathbb{R}} \rightarrow [0, \infty)$  defined by:  $m^*(A) = \inf \{\sum l(I_n)\}$  for all such countable collections.

**Lemma 8.1.2.** Consider the outer measure  $m^* : A \rightarrow m^*(A)$ , then the following are true:

- (1)  $m^*(\emptyset) = 0$ .
- (2) If  $A \subseteq B$ , then  $m^*(A) \leq m^*(B)$ .

*Proof.* Notice that  $\emptyset$  is covered by the interval  $[x, x]$ , whose length is 0. Then by definition  $m^*(\emptyset) = l([x, x]) = 0$ .

Now let  $A \subseteq B$ , if  $\{I_n\}$  covers  $B$ , then  $\{I_n\}$  covers  $A$ . Let  $U = \{l(I_n) : A \subseteq \bigcup I_n\}$  and  $V = \{l(I_n) : B \subseteq \bigcup I_n\}$ . Then  $\inf U \leq \inf V$ , so we have that  $m^*(A) \leq m^*(B)$ . ■

**Example 8.1.** Countable sets have outer measure 0. if  $E$  is countable, then one can list its elements as  $E = \{e_1, \dots, e_n, \dots\}$  such that.  $\{e_i\} \cap \{e_j\} = \emptyset$  if  $i \neq j$ . Notice that  $\{e_i\} \subseteq E$  and that  $E = \bigcup \{e_i\}$ . Moreover, each  $\{e_i\}$  is covered by  $[e_i, e_i]$ , with  $l([e_i, e_i]) = 0$ . So  $m^*(\{e_i\}) = 0$ . Therefore, assuming countable subadditivity of  $m^*$ , we get that  $m^*(E) \leq \sum m^*(\{e_i\}) = 0$ . Since  $0 \leq m^*(E)$ , we get equality.

**Theorem 8.1.3.** *The following is true for outer measure.*

- (1)  $m^*([a, b]) = m^*((a, b)) = b - a$  for all  $a, b \in \mathbb{R}$ .
- (2)  $m^*$  is translation invariant. That is if  $m^*(A + y) = m^*(A)$ .
- (3)  $m^*$  is countably subadditive, that is if  $\{E_k\}$  is a countable collection of subsets of  $E$ , covering  $E$  then

$$m^*(E) \leq \sum m^*(E_k) \quad (8.1)$$

*Proof.* Consider first the interval  $[a, b]$ . For  $\epsilon > 0$  notice  $(a - \epsilon, b + \epsilon)$  contains  $[a, b]$ . Then  $m^*([a, b]) \leq m^*((a - \epsilon, b + \epsilon)) \leq l((a - \epsilon, b + \epsilon)) = b - a + 2\epsilon$ . This shows that  $m^*([a, b]) \leq b - a$ . Now, by the Heine-Borel theorem  $[a, b]$  is compact, so let  $\{I_n\}$  be a open cover of  $[a, b]$  and let  $\{I_k\}_{k=1}^n$  a finite open subcover. Then if  $I_1 = (a_1, b_1)$  contains  $a$ , then we have  $a_1 < a < b_1$ . Now if  $b_1 \geq b$  then we have

$$\sum l(I_1) \leq b_1 - a_1 > b - a$$

If  $a \leq b_1 < b$ , since  $b_1 \notin (a_1, b_1)$ , there is an interval  $I_2 = (a_2, b_2)$  with  $b_1 \in (a_2, b_2)$ , and disjoint from  $I_1$ . Then if  $b_2 \geq b$  then we get

$$\sum l(I_2) \geq (b_1 - a_1) + (b_2 - a_2) > b_2 - a_1 > b - a$$

Proceeding inductively, we get a subcollection  $\{(a_k, b_k)\}_{k=1}^n$  with  $a_1 < a$  and  $a_{k+1} < b_k$  for all  $1 \leq k \leq n - 1$ . Thus we get

$$\sum l(I_k) \geq \sum l((a_k, b_k)) \leq (b_n - a_n) + \cdots + (b_1 - a_1) = b_n - a_1 > b - a$$

Now if  $I$  is any bounded interval for  $\epsilon > 0$  there exist closed bounded intervals  $J_1, J_2$  with  $J_1 \leq I \leq J_2$  with  $l(I) - \epsilon < l(J_1) < l(I) < l(J_2) < l(I) + \epsilon$ . Then we have  $l(I) - \epsilon < m^*(J_1) \leq m^*(I) \leq m^*(J_2) < l(I) + \epsilon$ . This forces  $l(I) = m^*(I)$ .

Now if  $I$  is unbounded there is a subinterval  $J \subseteq I$  with  $m^*(J) = n$ . then  $m^*(J) = n \leq m^*(I)$ . As  $n$  gets large, so does  $l(I)$ , so  $m^*(I) = \infty$ .

Consider now, for any set  $A$ , the set  $A + y$ . Then the collection  $\{I_n\}$  covers  $A$  if, and only if  $\{I_n + y\}$  covers  $A + y$ . Moreover, if  $I_n$  is an interval, then  $l(I_n) = l(I_n + y)$ . This makes  $m^*(A + y) = m^*(A)$ .

Now, let  $\{E_k\}$  be a collection of open subsets of  $E$ , covering  $E$ . If  $m^*(E_k) = \infty$  for some  $k$ , we are done. Suppose then that no  $E_k$  has such measure. Consider then the countable collection  $\{I_{n,k}\}_k$  of open and bounded intervals covering  $E_n$ , i.e.  $E_n \subseteq \bigcup_k I_{n,k}$ , and  $\sum l(I_{n,k}) \leq m^*(E_n) + \frac{\epsilon}{2^n}$ . Then by definition of  $m^*$ , we get  $m^*(\bigcup E_n) \leq \sum l(I_{n,k}) < \sum (m^*(E_k) + \frac{\epsilon}{2^n}) < \sum m^*(E_k) + \epsilon$ . ■

**Definition.** We say that an outermeasure  $m^*$  is **finitely subadditive** if for any finite collection  $\{E_k\}_{k=1}^n$  of open subsets of  $E$ , covering  $E$ , then

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

## 8.2 $\sigma$ -algebras of Lebesgue Measurable Sets

**Definition.** We say a set  $E$  is **measurable** if for any  $A \subseteq \mathbb{R}$ , then:

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

**Lemma 8.2.1.** *A set  $E$  is measurable if, and only if  $m^*(A) \geq m^*(A \cap E) + m^*(A \cap (\mathbb{R} \setminus E))$ .*

**Lemma 8.2.2.** *Any set of outer measure 0 is measurable.*

*Proof.* Let  $E$  be a set of measure 0, and let  $A \subseteq \mathbb{R}$ . Notice that  $A \cap E \subseteq E$  and  $A \cap (\mathbb{R} \setminus E) \subseteq A$ . Then by monotonicity,  $m^*(A \cap E) \leq m^*(E) = 0$  and  $m^*(A \cap (\mathbb{R} \setminus E)) \leq m^*(A)$ . Then we get that  $m^*(A) \geq m^*(A \cap E) + m^*(A \cap (\mathbb{R} \setminus E))$ . ■

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

**Lemma 8.2.3.** *Unions of a finite collection of measurable sets is measurable.*

*Proof.* By induction, for  $n = 2$ , let  $E_1$  and  $E_2$  be measurable. Then for any  $A \subseteq \mathbb{R}$ , we get  $m^*(A) = m^*(A \cap E_1) + m^*(A \cap (\mathbb{R} \setminus E_1)) = m^*(A \cap (\mathbb{R} \setminus E_1) \cap E_2) + m^*(A \cap (\mathbb{R} \setminus E_1) \cap (\mathbb{R} \setminus E_2))$ . Notice that  $A \cap (\mathbb{R} \setminus E_1) \cap (\mathbb{R} \setminus E_2) = A \cap \mathbb{R} \setminus (E_1 \cup E_2)$ , and that  $(A \cap E_1) \cup (A \cap \mathbb{R} \setminus E_1 \cap E_2) = A \cap (E_1 \cup E_2)$ . Then we get that

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

This makes  $E_1 \cup E_2$  measurable.

Now, suppose that for  $n \geq 2$  the collection  $\{E_i\}_{i=1}^n$  of measurable set has measurable union. Consider now the collection  $\{E_i\} \cup E_{n+1}$ , where  $E_{n+1}$  is also measurable. By hypothesis, we have :

$$m^*(A) \geq m^*(A \cap_{i=1}^{n+1} E_i) + m^*(A \cap \mathbb{R} \setminus_{i=1}^{n+1} E_i)$$

therefore the collection  $\{E_i\}_{i=1}^{n+1}$  of measurable sets has measurable union. ■

**Lemma 8.2.4.** *Let  $\{E_i\}_{i=1}^n$  be a collection of mutually disjoint measurable sets. Then for any  $A \subseteq \mathbb{R}$ :*

$$m^*(A \cap \bigcup_{i=1}^n E_i) = \sum_{i=1}^n m^*(A \cap E_i)$$

*Proof.* By induction on  $n$ , we have  $m^*(A \cap E_1) = \sum_{i=1}^1 m^*(A \cap E_i)$ . Now, suppose that for  $n \geq 1$ , that

$$m^*(A \cap \bigcup_{i=1}^n E_i) = \sum_{i=1}^n m^*(A \cap E_i)$$

consider then the collection  $\{E_i\} \cup E_{n+1}$  where  $E_{n+1}$  is measurable and mutually disjoint from all  $E_i$  for  $1 \leq i \leq n$ . Notice that  $(A \cap \bigcup_{i=1}^n E_i) \cap E_{n+1} = A \cap E_{n+1}$  and that  $A \cap \bigcup_{i=1}^{n+1} E_i \cap \mathbb{R} \setminus E_{n+1} = A \cap \bigcup_{i=1}^n E_i$ . Then by hypothesis, we have:

$$m^*(A \cap \bigcup_{i=1}^{n+1} E_i) = m^*(A \cap E_{n+1}) + m^*(A \cap \bigcup_{i=1}^n E_i) = m^*(A \cap E_{n+1}) + \sum_{i=1}^n m^*(A \cap E_i) = \sum_{i=1}^{n+1} m^*(A \cap E_i)$$

■

**Definition.** A collection of subsets of  $\mathbb{R}$  is called an **algebra** if it contains  $\mathbb{R}$ , and it is closed under complements and finite unions.

**Theorem 8.2.5.** *The collection of all measurable sets in  $\mathbb{R}$  forms an algebra.*

*Proof.* By lemma 8.1.3, we have that this collection is closed under finite unions. Also observe that if  $E$  is measurable, then  $m^*(A) = m^*(A \cap E) + m^*(A \cap \mathbb{R} \setminus E)$  implies that  $m^*(A) = m^*(A \cap \mathbb{R} \setminus E) + m^*(A \cap \mathbb{R} \setminus (\mathbb{R} \setminus E))$ . This makes  $\mathbb{R} \setminus E$  measurable. Lastly we see that  $m^*(A) = m^*(A \cap \mathbb{R}) + m^*(A \cap \mathbb{R} \setminus \mathbb{R}) = m^*(A) + m^*(\emptyset)$ , so that  $\mathbb{R}$  is measurable. Therefore the collection of all measurable sets in  $\mathbb{R}$  is measurable. ■

**Corollary.** *The collection of all measurable sets in  $\mathbb{R}$  is closed under finite intersections.*

**Lemma 8.2.6.** *The union of a countable collection of measurable sets is measurable.*

*Proof.* Let  $E = \bigcup E_n$  where  $\{E_n\}_{n \in \mathbb{Z}^+}$  is a countable collection of measurable sets. Let  $A \subseteq \mathbb{R}$ , and for  $n \in \mathbb{Z}^+$ , define  $F_n = \bigcup_{k=1}^n E_k$ . By lemma 8.2.3,  $F_n$  is measurable for all  $n$  and  $\mathbb{R} \setminus E \subseteq \mathbb{R} \setminus F_n$ . Then we have that  $m^*(A \cap F_n) + m^*(A \cap \mathbb{R} \setminus F_n) \geq m^*(A \cap F_n) + m^*(A \cap \mathbb{R} \setminus E)$ . By lemma 8.2.4, we get that  $m^*(A \cap F_n) = \sum_{k=1}^n m^*(E_k)$ , thus we get that  $m^*(A) \geq \sum_{k=1}^n m^*(E_k) + m^*(A \cap \mathbb{R} \setminus E)$ . By countable subadditivity, this gives us  $m^*(A) \geq m^*(A \cap E) + m^*(A \cap \mathbb{R} \setminus E)$ . Therefore  $E$  is measurable. ■

**Lemma 8.2.7.** *Every real interval is measurable.*

*Proof.* Since the collection of all measurable sets form an algebra, in fact they form a  $\sigma$ -algebra, it suffices to show that  $(a, \infty)$  and  $(-\infty, a)$  are measurable for all  $a \in \mathbb{R}$ .

Let  $A \subseteq \mathbb{R}$ , and suppose that  $a \notin A$  without loss of generality. Consider any countable collection  $\{I_n\}$  of open and bounded intervals covering  $A$ ; and define  $I'_n = I_n \cap (-\infty, a)$  and  $I''_n = I_n \cap (a, \infty)$ . Now,  $I'_n$  and  $I''_n$  are open intervals with  $l(I_n) = l(I'_n) + l(I''_n)$  for all  $n$ . Since  $\{I_n\}$  is countable, so are the collections  $\{I'_n\}$  and  $\{I''_n\}$ . They also cover subsets  $A_1, A_2 \subseteq A$ . Then  $m^*(A_1) = \sum l(I'_n)$  and  $m^*(A_2) = \sum l(I''_n)$ , so that  $m^*(A_1) + m^*(A_2) \leq \sum (l(I'_n) + l(I''_n)) \leq \sum l(I_n)$ . This makes  $(a, \infty)$  and  $(-\infty, a)$  measurable. ■

**Corollary.** *Any open set in  $\mathbb{R}$  is measurable.*

**Theorem 8.2.8.** *The collection of all measurable sets contains the Borel algebra.*

**Lemma 8.2.9.** *Measurable sets are translation invariant.*

*Proof.* Let  $E$  be a measurable set and let  $A \subseteq \mathbb{R}$  and  $y \in \mathbb{R}$ . Then  $m^*(A) = m^*(A \setminus y) = m^*((A \setminus y) \cap E) + m^*((A \setminus y) \cap \mathbb{R} \setminus E) = m^*(A \cap E + y) + m^*(A \cap \mathbb{R} \setminus (E + y))$ . ■



# Bibliography

- [1] W. R. Wade, *An introduction to analysis*. Upper Saddle River, NJ: Pearson Education, 2004.
- [2] W. Rudin, *Principles of mathematical analysis*. New York: McGraw-Hill, 1976.