# Part I

# Back Ground Material

# 1

# Introduction / User Guide

Not written as of yet. Topics to mention.

1. A better and more general integral.
   a) Convergence Theorems
   b) Integration over diverse collection of sets. (See probability theory.)
   c) Integration relative to different weights or densities including singular weights.
   d) Characterization of dual spaces.
   e) Completeness.
2. Infinite dimensional Linear algebra.
3. ODE and PDE.
4. Harmonic and Fourier Analysis.
5. Probability Theory

## 1.1 Topology beginnings

Recall the notion of a topology by extrapolating from the open sets on $\mathbb{R}^2$. Also recall what it means to be continuous, namely $f : X \to \mathbb{R}$ is continuous at $x$ if for all $\varepsilon > 0$ there exists $V \in \tau_x$ such that

$$f(V) \subset f(x) + (-\varepsilon, \varepsilon).$$

## 1.2 A Better Integral and an Introduction to Measure Theory

Let $a, b \in \mathbb{R}$ with $a < b$ and let

$$I^0(f) := \int_a^b f(t)dt \text{ for all } f \in C([a, b])$$

denote the Riemann integral. Also let $\mathcal{H}$ denote the smallest **linear subspace** of bounded functions on $[a, b]$ which is closed under bounded convergence and contains $C([a, b])$. Such a space exists since we can take the intersection over all such spaces of functions.

**Theorem 1.1.** *There is an extension $I$ of $I^0$ to $\mathcal{H}$ such that $I$ is still linear and $\lim_{n\to\infty} I(f_n) = I(f)$ for all $f_n \in \mathcal{H}$ with $f_n \to f$ boundedly. Moreover this extension is unique and is **positive** in the sense that $I(f) \geq 0$ if $f \in \mathcal{H}$ and $f \geq 0$.*

*Proof.* We will only prove the uniqueness here. Suppose that $J$ and $I$ are two such extensions and let

$$\mathcal{K} := \{f \in \mathcal{H} : J(f) = I(f)\}.$$

Then $\mathcal{K}$ is a linear subspace closed under bounded convergence which contains $C([a, b])$ and hence $\mathcal{K} = \mathcal{H}$.

The existence of $I$ is the hard part. The positivity of $I$ can be seen from the existence construction. ■

*Example 1.2.* Here are some examples of functions in $\mathcal{H}$ and their integrals:

1. Suppose $[\alpha, \beta] \subset [a, b]$, then $1_{[\alpha,\beta]} \in \mathcal{H}$ and $I\left(1_{[\alpha,\beta]}\right) = \beta - \alpha$.(Draw a picture.)
2. $I\left(1_{\{\alpha\}}\right) = 0$.
3. The space $\mathcal{H}$ is an algebra, i.e. if $f, g \in \mathcal{H}$ then $fg \in \mathcal{H}$. To prove this, first assume that $f \in C([a, b])$ and let
$$\mathcal{H}_f := \{g \in \mathcal{H} : fg \in \mathcal{H}\}.$$
Then $\mathcal{H}_f$ is closed under bounded convergence and contains $C([a, b])$ and hence $\mathcal{H}_f = \mathcal{H}$, i.e. the product of a continuous function and an element in $\mathcal{H}$ is back in $\mathcal{H}$.
Now suppose that $f \in \mathcal{H}$ and again let $\mathcal{H}_f$ be as above. By the same reasoning we may show again that $\mathcal{H}_f = \mathcal{H}$ and this proves the assertion.
4. If $f \in \mathcal{H}$ and $\phi \in C(\mathbb{R})$, then $\phi \circ f \in \mathcal{H}$. This a consequence of the Weierstrass approximation Theorem 22.34. In particular $|f| \in \mathcal{H}$ and $f_\pm := \frac{|f| \pm f}{2} \in \mathcal{H}$ if $f \in \mathcal{H}$.
5. If $f_n \in \mathcal{H}$, $f_n \geq 0$ and $f = \sum_{n=1}^\infty f_n$ is a bounded function, then $f \in \mathcal{H}$ and
$$I(f) = \sum_{n=1}^\infty I(f_n). \tag{1.1}$$
To prove Eq. (1.1) we have
$$\sum_{n=1}^\infty I(f_n) = \lim_{N\to\infty} I\left(\sum_{n=1}^N f_n\right) = I(f).$$

6. As an example of item 4., $1_{\mathbb{Q}\cap[a,b]} = \sum_{n=1}^{\infty} 1_{\{\alpha_n\}} \in \mathcal{H}$ and $I\left(1_{\mathbb{Q}\cap[a,b]}\right) = 0$. Here $\{\alpha_n\}_{n=1}^{\infty}$ is an enumeration of the rational numbers in the interval $[a,b]$.
7. Let $\mathcal{M} := \{A \subset [a,b] : 1_A \in \mathcal{H}\}$ and for $A \in \mathcal{M}$ let $m(A) := I(1_A)$. Then $\mathcal{M}$ and $m$ have the following properties:
   a) $\emptyset, [a,b] \in \mathcal{M}$ and $m(\emptyset) = 0$ and $m([a,b]) = b-a$. Moreover $m(A) \geq 0$ for all $A \in \mathcal{M}$.
   b) If $A \in \mathcal{M}$ then $A^c \in \mathcal{M}$ and $m(A^c) = b - a - m(A)$. This follows from the fact that $1_{A^c} = 1 - 1_A$.
   c) If $A, B \in \mathcal{M}$, then $A \cap B \in \mathcal{M}$ since if $1_{A\cap B} = 1_A \cdot 1_B$ and $\mathcal{H}$ is an algebra.
   **Definition:** a collection of sets $\mathcal{M}$ satisfying a) – c) is called an **algebra** of subsets of $[a,b]$.
   d) More generally if $A_n \in \mathcal{M}$ then $\cap A_n \in \mathcal{M}$ since $1_{\cap A_n} = \lim_{N\to\infty} 1_{A_1} \cdots 1_{A_N}$ and the convergence is bounded.
   **Definition:** a collection of sets $\mathcal{M}$ satisfying a) – d) is called an $\sigma$ – **algebra.**
   e) If $A_n \in \mathcal{M}$, then $\cup A_n \in \mathcal{M}$. Indeed we know $\cup A_n \in \mathcal{M}$ iff $(\cup A_n)^c \in \mathcal{M}$. But
   $$(\cup A_n)^c = \cap A_n^c \in \mathcal{M}$$
   by item d. above.
   f) If $A_n \in \mathcal{M}$ are pairwise disjoint, then
   $$m(\cup A_n) = \sum_{n=1}^{\infty} m(A_n).$$
   To prove this it suffices to observe that $1_{\cup A_n} = \sum_{n=1}^{\infty} 1_{A_n}$.
   g) $\mathcal{M}$ is not $2^{[a,b]}$, i.e. $\mathcal{M}$ is not all subset of $[a,b]$. This is not obvious and it is not possible to really write down an "explicit" subset $[a,b]$ which is not in $\mathcal{M}$. We will prove the existence of such sets later.
8. **Fact:** $\mathcal{M}$ is the smallest $\sigma$ – algebra on $[a,b]$ which contains all subintervals of $[a,b]$.
9. **Fact:** A bounded function $f : [a,b] \to \mathbb{R}$ is in $\mathcal{H}$ iff $\{f > \alpha\} \in \mathcal{M}$ for all $\alpha \in \mathbb{R}$.
10. **Fact:** The integral $I$ may be recovered from the measure $m$ by the formula
$$I(f) = \lim_{\text{mesh}\to 0} \sum_{0<\alpha_1<\alpha_2<\alpha_3<\dots}^{\infty} \alpha_i m(\{x \in [a,b] : \alpha_i < f(x) \leq \alpha_i\}).$$

We will prove items 8. – 10. later in the course. The proof if Items 9. and 10. is not so hard and the energetic reader may wish to give them a try.

**Notation 1.3** *The collection of sets $\mathcal{M}$ is called the Borel $\sigma$ – algebra on $[a,b]$ and the function $m : \mathcal{M} \to \mathbb{R}$ is called Lebesgue measure. We will usually* *write* $I(f)$ *as* $\int_{[a,b]} f dm$ *and* $I(f)$ *will be called the Lebesgue integral of* $f$*. This integral may be extended to all positive functions* $f$ *such that* $f1_{|f|\leq M} \in \mathcal{H}$ *for all* $M$ *by*

$$I(f) = \lim_{M\to\infty} I\left(f1_{|f|\leq M}\right).$$

Again, we will come back to all of this again later.

# 2

# Set Operations

Let $\mathbb{N}$ denote the positive integers, $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ be the non-negative integers and $\mathbb{Z} = \mathbb{N}_0 \cup (-\mathbb{N})$ – the positive and negative integers including 0, $\mathbb{Q}$ the rational numbers, $\mathbb{R}$ the real numbers (see Chapter 3 below), and $\mathbb{C}$ the complex numbers. We will also use $\mathbb{F}$ to stand for either of the fields $\mathbb{R}$ or $\mathbb{C}$.

**Notation 2.1** *Given two sets $X$ and $Y$, let $Y^X$ denote the collection of all functions $f : X \to Y$. If $X = \mathbb{N}$, we will say that $f \in Y^{\mathbb{N}}$ is a sequence with values in $Y$ and often write $f_n$ for $f(n)$ and express $f$ as $\{f_n\}_{n=1}^{\infty}$. If $X = \{1, 2, \ldots, N\}$, we will write $Y^N$ in place of $Y^{\{1,2,\ldots,N\}}$ and denote $f \in Y^N$ by $f = (f_1, f_2, \ldots, f_N)$ where $f_n = f(n)$.*

**Notation 2.2** *More generally if $\{X_\alpha : \alpha \in A\}$ is a collection of non-empty sets, let $X_A = \prod_{\alpha \in A} X_\alpha$ and $\pi_\alpha : X_A \to X_\alpha$ be the canonical projection map defined by $\pi_\alpha(x) = x_\alpha$.*

Recall that an element $x \in X_A$ is a "**choice function**," i.e. an assignment $x_\alpha := x(\alpha) \in X_\alpha$ for each $\alpha \in A$. The **axiom of choice** (See Appendix B.) states that $X_A \neq \emptyset$ provided that $X_\alpha \neq \emptyset$ for each $\alpha \in A$. If $X_\alpha = X$ for some fixed space $X$, then $\prod_{\alpha \in A} X_\alpha = X^A$.

**Notation 2.3** *Given a set $X$, let $2^X$ denote the* ***power set*** *of $X$ – the collection of all subsets of $X$ including the empty set.*

The reason for writing the power set of $X$ as $2^X$ is that if we think of 2 meaning $\{0, 1\}$, then an element of $a \in 2^X = \{0, 1\}^X$ is completely determined by the set

$$A := \{x \in X : a(x) = 1\} \subset X.$$

In this way elements in $\{0, 1\}^X$ are in one to one correspondence with subsets of $X$.

For $A \in 2^X$ let

$$A^c := X \setminus A = \{x \in X : x \notin A\}$$

and more generally if $A, B \subset X$ let

$$B \setminus A := \{x \in B : x \notin A\} = A \cap B^c.$$

We also define the symmetric difference of $A$ and $B$ by

$$A \triangle B := (B \setminus A) \cup (A \setminus B).$$

As usual if $\{A_\alpha\}_{\alpha \in I}$ is an indexed collection of subsets of $X$ we define the union and the intersection of this collection by

$$\cup_{\alpha \in I} A_\alpha := \{x \in X : \exists\, \alpha \in I \ \ni\ x \in A_\alpha\} \text{ and}$$
$$\cap_{\alpha \in I} A_\alpha := \{x \in X : x \in A_\alpha \,\forall\, \alpha \in I\,\}.$$

**Notation 2.4** *We will also write* $\coprod_{\alpha \in I} A_\alpha$ *for* $\cup_{\alpha \in I} A_\alpha$ *in the case that* $\{A_\alpha\}_{\alpha \in I}$ *are pairwise disjoint, i.e.* $A_\alpha \cap A_\beta = \emptyset$ *if* $\alpha \neq \beta$.

Notice that $\cup$ is closely related to $\exists$ and $\cap$ is closely related to $\forall$. For example let $\{A_n\}_{n=1}^\infty$ be a sequence of subsets from $X$ and define

$$\{A_n \text{ i.o.}\} := \{x \in X : \#\{n : x \in A_n\} = \infty\} \text{ and}$$
$$\{A_n \text{ a.a.}\} := \{x \in X : x \in A_n \text{ for all } n \text{ sufficiently large}\}.$$

(One should read $\{A_n \text{ i.o.}\}$ as $A_n$ infinitely often and $\{A_n \text{ a.a.}\}$ as $A_n$ almost always.) Then $x \in \{A_n \text{ i.o.}\}$ iff

$$\forall N \in \mathbb{N}\ \exists\ n \geq N \ni x \in A_n$$

and this may be expressed as

$$\{A_n \text{ i.o.}\} = \cap_{N=1}^\infty \cup_{n \geq N} A_n.$$

Similarly, $x \in \{A_n \text{ a.a.}\}$ iff

$$\exists\ N \in \mathbb{N} \ni \forall\ n \geq N,\ \ x \in A_n$$

which may be written as

$$\{A_n \text{ a.a.}\} = \cup_{N=1}^\infty \cap_{n \geq N} A_n.$$

**Definition 2.5.** *A set $X$ is said to be **countable** if is empty or there is an injective function $f : X \to \mathbb{N}$, otherwise $X$ is said to be **uncountable.***

**Lemma 2.6 (Basic Properties of Countable Sets).**

1. *If $A \subset X$ is a subset of a countable set $X$ then $A$ is countable.*
2. *Any infinite subset $\Lambda \subset \mathbb{N}$ is in one to one correspondence with $\mathbb{N}$.*

3. *A non-empty set $X$ is countable iff there exists a surjective map, $g : \mathbb{N} \to X$.*
4. *If $X$ and $Y$ are countable then $X \times Y$ is countable.*
5. *Suppose for each $m \in \mathbb{N}$ that $A_m$ is a countable subset of a set $X$, then $A = \cup_{m=1}^{\infty} A_m$ is countable. In short, the countable union of countable sets is still countable.*
6. *If $X$ is an infinite set and $Y$ is a set with at least two elements, then $Y^X$ is uncountable. In particular $2^X$ is uncountable for any infinite set $X$.*

*Proof.* 1. If $f : X \to \mathbb{N}$ is an injective map then so is the restriction, $f|_A$, of $f$ to the subset $A$.

2. Let $f(1) = \min \Lambda$ and define $f$ inductively by

$$f(n+1) = \min \Lambda \setminus \{f(1), \dots, f(n)\}.$$

Since $\Lambda$ is infinite the process continues indefinitely. The function $f : \mathbb{N} \to \Lambda$ defined this way is a bijection.

3. If $g : \mathbb{N} \to X$ is a surjective map, let

$$f(x) = \min g^{-1}(\{x\}) = \min \{n \in \mathbb{N} : f(n) = x\}.$$

Then $f : X \to \mathbb{N}$ is injective which combined with item 2. (taking $\Lambda = f(X)$) shows $X$ is countable. Conversely if $f : X \to \mathbb{N}$ is injective let $x_0 \in X$ be a fixed point and define $g : \mathbb{N} \to X$ by $g(n) = f^{-1}(n)$ for $n \in f(X)$ and $g(n) = x_0$ otherwise.

4. Let us first construct a bijection, $h$, from $\mathbb{N}$ to $\mathbb{N} \times \mathbb{N}$. To do this put the elements of $\mathbb{N} \times \mathbb{N}$ into an array of the form

$$\begin{pmatrix} (1,1) & (1,2) & (1,3) & \dots \\ (2,1) & (2,2) & (2,3) & \dots \\ (3,1) & (3,2) & (3,3) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

and then "count" these elements by counting the sets $\{(i,j) : i + j = k\}$ one at a time. For example let $h(1) = (1,1)$, $h(2) = (2,1)$, $h(3) = (1,2)$, $h(4) = (3,1)$, $h(5) = (2,2)$, $h(6) = (1,3)$, etc. etc.

If $f : \mathbb{N} \to X$ and $g : \mathbb{N} \to Y$ are surjective functions, then the function $(f \times g) \circ h : \mathbb{N} \to X \times Y$ is surjective where $(f \times g)(m,n) := (f(m), g(n))$ for all $(m,n) \in \mathbb{N} \times \mathbb{N}$.

5. If $A = \emptyset$ then $A$ is countable by definition so we may assume $A \neq \emptyset$. With out loss of generality we may assume $A_1 \neq \emptyset$ and by replacing $A_m$ by $A_1$ if necessary we may also assume $A_m \neq \emptyset$ for all $m$. For each $m \in \mathbb{N}$ let $a_m : \mathbb{N} \to A_m$ be a surjective function and then define $f : \mathbb{N} \times \mathbb{N} \to \cup_{m=1}^{\infty} A_m$ by $f(m,n) := a_m(n)$. The function $f$ is surjective and hence so is the composition, $f \circ h : \mathbb{N} \to X \times Y$, where $h : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ is the bijection defined above.

6. Let us begin by showing $2^{\mathbb{N}} = \{0,1\}^{\mathbb{N}}$ is uncountable. For sake of contradiction suppose $f : \mathbb{N} \to \{0,1\}^{\mathbb{N}}$ is a surjection and write $f(n)$ as

$(f_1(n), f_2(n), f_3(n), \dots)$. Now define $a \in \{0,1\}^{\mathbb{N}}$ by $a_n := 1 - f_n(n)$. By construction $f_n(n) \neq a_n$ for all $n$ and so $a \notin f(\mathbb{N})$. This contradicts the assumption that $f$ is surjective and shows $2^{\mathbb{N}}$ is uncountable.

For the general case, since $Y_0^X \subset Y^X$ for any subset $Y_0 \subset Y$, if $Y_0^X$ is uncountable then so is $Y^X$. In this way we may assume $Y_0$ is a two point set which may as well be $Y_0 = \{0,1\}$. Moreover, since $X$ is an infinite set we may find an injective map $x : \mathbb{N} \to X$ and use this to set up an injection, $i : 2^{\mathbb{N}} \to 2^X$ by setting $i(a)(x_n) = a_n$ for all $n \in \mathbb{N}$ and $i(a)(x) = 0$ if $x \notin \{x_n : n \in \mathbb{N}\}$. If $2^X$ were countable we could find a surjective map $f : 2^X \to \mathbb{N}$ in which case $f \circ i : 2^{\mathbb{N}} \to \mathbb{N}$ would be surjective as well. However this is impossible since we have already seed that $2^{\mathbb{N}}$ is uncountable. ■

We end this section with some notation which will be used frequently in the sequel.

**Notation 2.7** *If $f : X \to Y$ is a function and $\mathcal{E} \subset 2^Y$ let*

$$f^{-1}\mathcal{E} := f^{-1}(\mathcal{E}) := \{f^{-1}(E) | E \in \mathcal{E}\}.$$

*If $\mathcal{G} \subset 2^X$, let*

$$f_*\mathcal{G} := \{A \in 2^Y | f^{-1}(A) \in \mathcal{G}\}.$$

**Definition 2.8.** *Let $\mathcal{E} \subset 2^X$ be a collection of sets, $A \subset X$, $i_A : A \to X$ be the* ***inclusion map*** *($i_A(x) = x$ for all $x \in A$) and*

$$\mathcal{E}_A = i_A^{-1}(\mathcal{E}) = \{A \cap E : E \in \mathcal{E}\}.$$

## 2.1 Exercises

Let $f : X \to Y$ be a function and $\{A_i\}_{i \in I}$ be an indexed family of subsets of $Y$, verify the following assertions.

**Exercise 2.1.** $(\cap_{i \in I} A_i)^c = \cup_{i \in I} A_i^c$.

**Exercise 2.2.** Suppose that $B \subset Y$, show that $B \setminus (\cup_{i \in I} A_i) = \cap_{i \in I}(B \setminus A_i)$.

**Exercise 2.3.** $f^{-1}(\cup_{i \in I} A_i) = \cup_{i \in I} f^{-1}(A_i)$.

**Exercise 2.4.** $f^{-1}(\cap_{i \in I} A_i) = \cap_{i \in I} f^{-1}(A_i)$.

**Exercise 2.5.** Find a counter example which shows that $f(C \cap D) = f(C) \cap f(D)$ need not hold.

# 3

# The Real and Complex Numbers

Although it is assumed that the reader of this book is familiar with the properties of the real numbers, $\mathbb{R}$, nevertheless I feel it is instructive to define them here and sketch the development of their basic properties. It will most certainly be assumed that the reader is familiar with basic algebraic properties of the natural numbers $\mathbb{N}$ and the ordered field of rational numbers,

$$\mathbb{Q} = \left\{ \frac{m}{n} : m, n \in \mathbb{Z} : n \neq 0 \right\}.$$

As usual, for $q \in \mathbb{Q}$, we define

$$|q| = \begin{cases} q & \text{if } q \geq 0 \\ -q & \text{if } q \leq 0. \end{cases}$$

Notice that if $q \in \mathbb{Q}$ and $|q| \leq \frac{1}{n}$ for all $n$, then $q = 0$. Indeed $q \neq 0$ then $|q| = \frac{m}{n}$ for some $m, n \in \mathbb{N}$ and hence $|q| \geq \frac{1}{n}$. A similar argument shows $q \geq 0$ iff $q \geq -\frac{1}{n}$ for all $n \in \mathbb{N}$. These trivial remarks will be used in the future without further reference.

**Definition 3.1.** *A sequence* $\{q_n\}_{n=1}^{\infty} \subset \mathbb{Q}$ ***converges*** *to* $q \in \mathbb{Q}$ *if* $|q - q_n| \to 0$ *as* $n \to \infty$, *i.e. if for all* $N \in \mathbb{N}$, $|q - q_n| \leq \frac{1}{N}$ *for a.a.* $n$. *As usual if* $\{q_n\}_{n=1}^{\infty}$ *converges to* $q$ *we will write* $q_n \to q$ *as* $n \to \infty$ *or* $q = \lim_{n\to\infty} q_n$.

**Definition 3.2.** *A sequence* $\{q_n\}_{n=1}^{\infty} \subset \mathbb{Q}$ *is* ***Cauchy*** *if* $|q_n - q_m| \to 0$ *as* $m, n \to \infty$. *More precisely we require for each* $N \in \mathbb{N}$ *that* $|q_m - q_n| \leq \frac{1}{N}$ *for a.a. pairs* $(m, n)$.

**Exercise 3.1.** Show that all convergent sequences $\{q_n\}_{n=1}^{\infty} \subset \mathbb{Q}$ are Cauchy and that all Cauchy sequences $\{q_n\}_{n=1}^{\infty}$ are bounded – i.e. there exists $M \in \mathbb{N}$ such that

$$|q_n| \leq M \text{ for all } n \in \mathbb{N}.$$

**Exercise 3.2.** Suppose $\{q_n\}_{n=1}^{\infty}$ and $\{r_n\}_{n=1}^{\infty}$ are Cauchy sequences in $\mathbb{Q}$.

1. Show $\{q_n + r_n\}_{n=1}^{\infty}$ and $\{q_n \cdot r_n\}_{n=1}^{\infty}$ are Cauchy.
   Now assume that $\{q_n\}_{n=1}^{\infty}$ and $\{r_n\}_{n=1}^{\infty}$ are convergent sequences in $\mathbb{Q}$.
2. Show $\{q_n + r_n\}_{n=1}^{\infty}$ $\{q_n \cdot r_n\}_{n=1}^{\infty}$ are convergent in $\mathbb{Q}$ and

$$\lim_{n\to\infty} (q_n + r_n) = \lim_{n\to\infty} q_n + \lim_{n\to\infty} r_n \text{ and}$$
$$\lim_{n\to\infty} (q_n r_n) = \lim_{n\to\infty} q_n \cdot \lim_{n\to\infty} r_n.$$

3. If we further assume $q_n \le r_n$ for all $n$, show $\lim_{n\to\infty} q_n \le \lim_{n\to\infty} r_n$. (It suffices to consider the case where $q_n = 0$ for all $n$.)

The rational numbers $\mathbb{Q}$ suffer from the defect that they are not complete, i.e. not all Cauchy sequences are convergent. In fact, according to Corollary 3.14 below, "most" Cauchy sequences of rational numbers do not converge to a rational number.

**Exercise 3.3.** Use the following outline to construct a Cauchy sequence $\{q_n\}_{n=1}^{\infty} \subset \mathbb{Q}$ which is **not** convergent in $\mathbb{Q}$.

1. Recall that there is no element $q \in \mathbb{Q}$ such that $q^2 = 2$[1]. To each $n \in \mathbb{N}$ let $m_n \in \mathbb{N}$ be chosen so that

$$\frac{m_n^2}{n^2} < 2 < \frac{(m_n + 1)^2}{n^2} \tag{3.1}$$

   and let $q_n := \frac{m_n}{n}$.
2. Verify that $q_n^2 \to 2$ as $n \to \infty$ and that $\{q_n\}_{n=1}^{\infty}$ is a Cauchy sequence in $\mathbb{Q}$.
3. Show $\{q_n\}_{n=1}^{\infty}$ does not have a limit in $\mathbb{Q}$.

## 3.1 The Real Numbers

Let $\mathcal{C}$ denote the collection of Cauchy sequences $a = \{a_n\}_{n=1}^{\infty} \subset \mathbb{Q}$ and say $a, b \in \mathcal{C}$ are equivalent (write $a \sim b$) iff $\lim_{n\to\infty} |a_n - b_n| = 0$. (The reader should check that " $\sim$ " is an equivalence relation.)

**Definition 3.3.** *A* ***real number*** *is an equivalence class, $\bar{a} := \{b \in \mathcal{C} : b \sim a\}$ associated to some element $a \in \mathcal{C}$. The collection of real numbers will be denoted by $\mathbb{R}$. For $q \in \mathbb{Q}$, let $i(q) = \bar{a}$ where $a$ is the constant sequence $a_n = q$ for all $n \in \mathbb{N}$. We will simply write $0$ for $i(0)$ and $1$ for $i(1)$.*

**Exercise 3.4.** Given $\bar{a}, \bar{b} \in \mathbb{R}$ show that the definitions

$$-\bar{a} = \overline{(-a)},\ \bar{a} + \bar{b} := \overline{(a + b)} \text{ and } \bar{a} \cdot \bar{b} := \overline{a \cdot b}$$

[1] This fact also shows that the intermediate value theorem, (See Theorem 10.57 below.) fails when working with continuous functions defined over $\mathbb{Q}$.

are well defined. Here $-a$, $a+b$ and $a \cdot b$ denote the sequences $\{-a_n\}_{n=1}^{\infty}$, $\{a_n + b_n\}_{n=1}^{\infty}$ and $\{a_n \cdot b_n\}_{n=1}^{\infty}$ respectively. Further verify that with these operations, $\mathbb{R}$ becomes a field and the map $i : \mathbb{Q} \to \mathbb{R}$ is injective homomorphism of fields. **Hint:** if $\bar{a} \neq 0$ show that $\bar{a}$ may be represented by a sequence $a \in \mathcal{C}$ with $|a_n| \geq \frac{1}{N}$ for all $n$ and some $N \in \mathbb{N}$. For this representative show the sequence $a^{-1} := \left\{a_n^{-1}\right\}_{n=1}^{\infty} \in \mathcal{C}$. The multiplicative inverse to $\bar{a}$ may now be constructed as: $\frac{1}{\bar{a}} = \bar{a}^{-1} := \overline{\left\{a_n^{-1}\right\}_{n=1}^{\infty}}$.

**Definition 3.4.** *Let $\bar{a}, \bar{b} \in \mathbb{R}$. Then*

1. *$\bar{a} > 0$ if there exists an $N \in \mathbb{N}$ such that $a_n > \frac{1}{N}$ for a.a. $n$.*
2. *$\bar{a} \geq 0$ iff either $\bar{a} > 0$ or $\bar{a} = 0$. Equivalently (as the reader should verify), $\bar{a} \geq 0$ iff for all $N \in \mathbb{N}$, $a_n \geq -\frac{1}{N}$ for a.a. $n$.*
3. *Write $\bar{a} > \bar{b}$ or $\bar{b} < \bar{a}$ if $\bar{a} - \bar{b} > 0$*
4. *Write $\bar{a} \geq \bar{b}$ or $\bar{b} \leq \bar{a}$ if $\bar{a} - \bar{b} \geq 0$.*

**Exercise 3.5.** Show " $\geq$ " make $\mathbb{R}$ into a linearly ordered field and the map $i : \mathbb{Q} \to \mathbb{R}$ preserves order. Namely if $\bar{a}, \bar{b} \in \mathbb{R}$ then

1. exactly one of the following relations hold: $\bar{a} < \bar{b}$ or $\bar{a} > \bar{b}$ or $\bar{a} = \bar{b}$.
2. If $\bar{a} \geq 0$ and $\bar{b} \geq 0$ then $\bar{a} + \bar{b} \geq 0$ and $\bar{a} \cdot \bar{b} \geq 0$.
3. If $q, r \in Q$ then $q \leq r$ iff $i(q) \leq i(r)$.

The **absolute value** of a real number $\bar{a}$ is defined analogously to that of a rational number by

$$|\bar{a}| = \begin{cases} \bar{a} & \text{if } \bar{a} \geq 0 \\ -\bar{a} & \text{if } \bar{a} < 0 \end{cases}.$$

Observe this definition is consistent with our previous definition of the absolute value on $\mathbb{Q}$, namely $i(|q|) = |i(q)|$. Also notice that $\bar{a} = 0$ (i.e. $a \sim 0$ where 0 denotes the constant sequence of all zeros) iff for all $N \in \mathbb{N}$, $|a_n| \leq \frac{1}{N}$ for a.a. $n$. This is equivalent to saying $|\bar{a}| \leq i\left(\frac{1}{N}\right)$ for all $N \in \mathbb{N}$ iff $\bar{a} = 0$.

**Exercise 3.6.** Given $\bar{a}, \bar{b} \in \mathbb{R}$ show

$$\left|\bar{a}\bar{b}\right| = |\bar{a}| \left|\bar{b}\right| \text{ and } \left|\bar{a} + \bar{b}\right| \leq |\bar{a}| + \left|\bar{b}\right|.$$

The latter inequality being referred to as the **triangle inequality**.

By exercise 3.6,

$$|\bar{a}| = \left|\bar{a} - \bar{b} + \bar{b}\right| \leq \left|\bar{a} - \bar{b}\right| + \left|\bar{b}\right|$$

and hence

$$|\bar{a}| - \left|\bar{b}\right| \leq \left|\bar{a} - \bar{b}\right|$$

and by reversing the roles of $\bar{a}$ and $\bar{b}$ we also have

$$-\left(|\bar{a}| - \left|\bar{b}\right|\right) = \left|\bar{b}\right| - |\bar{a}| \leq \left|\bar{b} - \bar{a}\right| = \left|\bar{a} - \bar{b}\right|.$$

Therefore $\left||\bar{a}| - \left|\bar{b}\right|\right| \le \left|\bar{a} - \bar{b}\right|$ and in particular if $\{\bar{a}_n\}_{n=1}^{\infty} \subset \mathbb{R}$ **converges** to $\bar{a} \in \mathbb{R}$ then

$$||\bar{a}_n| - |\bar{a}|| \le |\bar{a}_n - \bar{a}| \to 0 \text{ as } n \to \infty.$$

**Definition 3.5.** *A sequence $\{\bar{a}_n\}_{n=1}^{\infty} \subset \mathbb{R}$ **converges** to $\bar{a} \in \mathbb{R}$ if $|\bar{a} - \bar{a}_n| \to 0$ as $n \to \infty$, i.e. if for all $N \in \mathbb{N}$, $|\bar{a} - \bar{a}_n| \le i\left(\frac{1}{N}\right)$ for a.a. $n$. As before if $\{\bar{a}_n\}_{n=1}^{\infty}$ converges to $\bar{a}$ we will write $\bar{a}_n \to \bar{a}$ as $n \to \infty$ or $\bar{a} = \lim_{n\to\infty} \bar{a}_n$.*

*Remark 3.6.* The field $i(\mathbb{Q})$ is **dense** in $\mathbb{R}$ in the sense that if $\bar{a} \in \mathbb{R}$ there exists $\{q_n\}_{n=1}^{\infty} \subset \mathbb{Q}$ such that $i(q_n) \to \bar{a}$ as $n \to \infty$. Indeed, simply let $q_n = a_n$ where $a$ represents $\bar{a}$. Since $a$ is a Cauchy sequence, to any $N \in \mathbb{N}$ there exits $M \in \mathbb{N}$ such that

$$-\frac{1}{N} \le a_m - a_n \le \frac{1}{N} \text{ for all } m, n \ge M$$

and therefore

$$-i\left(\frac{1}{N}\right) \le i(a_m) - \bar{a} \le i\left(\frac{1}{N}\right) \text{ for all } m \ge M.$$

This shows

$$|i(q_m) - \bar{a}| = |i(a_m) - \bar{a}| \le i\left(\frac{1}{N}\right) \text{ for all } m \ge M$$

and since $N$ is arbitrary that $i(q_m) \to \bar{a}$ as $m \to \infty$.

**Definition 3.7.** *A sequence $\{\bar{a}_n\}_{n=1}^{\infty} \subset \mathbb{R}$ is **Cauchy** if $|\bar{a}_n - \bar{a}_m| \to 0$ as $m, n \to \infty$. More precisely we require for each $N \in \mathbb{N}$ that $|\bar{a}_m - \bar{a}_n| \le i\left(\frac{1}{N}\right)$ for a.a. pairs $(m, n)$.*

**Exercise 3.7.** The analogues of the results in Exercises 3.1 and 3.2 hold with $\mathbb{Q}$ replaced by $\mathbb{R}$. (We now say a subset $\Lambda \subset \mathbb{R}$ is bounded if there exists $M \in \mathbb{N}$ such that $|\lambda| \le i(M)$ for all $\lambda \in \Lambda$.)

For the purposes of real analysis the most important property of $\mathbb{R}$ is that it is "complete."

**Theorem 3.8.** *The ordered field $\mathbb{R}$ is **complete**, i.e. all Cauchy sequences in $\mathbb{R}$ are convergent.*

*Proof.* Suppose that $\{\bar{a}(m)\}_{m=1}^{\infty}$ is a Cauchy sequence in $\mathbb{R}$. By Remark 3.6, we may choose $q_m \in \mathbb{Q}$ such that

$$|\bar{a}(m) - i(q_m)| \le i\left(m^{-1}\right) \text{ for all } m \in \mathbb{N}.$$

Given $N \in \mathbb{N}$, choose $M \in \mathbb{N}$ such that $|\bar{a}(m) - \bar{a}(n)| \le i\left(N^{-1}\right)$ for all $m, n \ge M$. Then

$$\begin{aligned}|i(q_m) - i(q_n)| &\le |i(q_m) - \bar{a}(m)| + |\bar{a}(m) - \bar{a}(n)| + |\bar{a}(n) - i(q_n)| \\ &\le i\left(m^{-1}\right) + i\left(n^{-1}\right) + i\left(N^{-1}\right)\end{aligned}$$

and therefore

$$|q_m - q_n| \le m^{-1} + n^{-1} + N^{-1} \text{ for all } m, n \ge M.$$

It now follows that $q = \{q_m\}_{m=1}^{\infty} \in \mathcal{C}$ and therefore $q$ represents a point $\bar{q} \in \mathbb{R}$. Using Remark 3.6 and the triangle inequality,

$$\begin{aligned}|\bar{a}(m) - \bar{q}| &\le |\bar{a}(m) - i(q_m)| + |i(q_m) - \bar{q}| \\ &\le i\left(m^{-1}\right) + |i(q_m) - \bar{q}| \to 0 \text{ as } m \to \infty\end{aligned}$$

and therefore $\lim_{m\to\infty} \bar{a}(m) = \bar{q}$. ■

**Definition 3.9.** *A number $M \in \mathbb{R}$ is an* ***upper bound*** *for a set $\Lambda \subset \mathbb{R}$ if $\lambda \le M$ for all $\lambda \in \Lambda$ and a number $m \in \mathbb{R}$ is an* ***lower bound*** *for a set $\Lambda \subset \mathbb{R}$ if $\lambda \ge m$ for all $\lambda \in \Lambda$. Upper and lower bounds need not exist. If $\Lambda$ has upper (lower) bound, $\Lambda$ is said to be* ***bounded from above (below)****.*

**Theorem 3.10.** *To each non-empty set $\Lambda \subset \mathbb{R}$ which is bounded from above (below) there is a unique* ***least upper bound*** *denoted by* $\sup \Lambda \in \mathbb{R}$ *(respectively* ***greatest lower bound*** *denoted by* $\inf \Lambda \in \mathbb{R}$).

*Proof.* Suppose $\Lambda$ is bounded from above and for each $n \in \mathbb{N}$, let $m_n \in \mathbb{Z}$ be the smallest integer such that $i\left(\frac{m_n}{2^n}\right)$ is an upper bound for $\Lambda$. The sequence $q_n := \frac{m_n}{2^n}$ is Cauchy because $q_m \in [q_n - 2^{-n}, q_n] \cap \mathbb{Q}$ for all $m \ge n$, i.e.

$$|q_m - q_n| \le 2^{-\min(m,n)} \to 0 \text{ as } m, n \to \infty.$$

Passing to the limit, $n \to \infty$, in the inequality $i(q_n) \ge \lambda$, which is valid for all $\lambda \in \Lambda$ implies

$$\bar{q} = \lim_{n\to\infty} i(q_n) \ge \lambda \text{ for all } \lambda \in \Lambda.$$

Thus $\bar{q}$ is an upper bound for $\Lambda$.

If there were another upper bound $M \in \mathbb{R}$ for $\Lambda$ such that $M < \bar{q}$, it would follow that $M \le i(q_n) < \bar{q}$ for some $n$. But this is a contradiction because $\{q_n\}_{n=1}^{\infty}$ is a decreasing sequence, $i(q_n) \ge i(q_m)$ for all $m \ge n$ and therefore $i(q_n) \ge \bar{q}$ for all $n$. Therefore $\bar{q}$ is the unique least upper bound for $\Lambda$. The existence of lower bounds is proved analogously. ■

**Proposition 3.11.** *If $\{a_n\}_{n=1}^{\infty} \subset \mathbb{R}$ is an increasing (decreasing) sequence which is bounded from above (below), then $\{a_n\}_{n=1}^{\infty}$ is convergent and*

$$\lim_{n\to\infty} a_n = \sup\{a_n : n \in \mathbb{N}\} \ \left(\lim_{n\to\infty} a_n = \inf\{a_n : n \in \mathbb{N}\}\right).$$

*If $\Lambda \subset \mathbb{R}$ is a set bounded from above then there exists $\{\lambda_n\} \subset \Lambda$ such that $\lambda_n \uparrow M := \sup \Lambda$, as $n \to \infty$, i.e. $\{\lambda_n\}$ is increasing and $\lim_{n\to\infty} \lambda_n = M$.*

*Proof.* Let $M := \sup\{a_n : n \in \mathbb{N}\}$, then for each $N \in \mathbb{N}$ there must exist $m \in \mathbb{N}$ such that $M - i\left(N^{-1}\right) < a_m \leq M$. Since $a_n$ is increasing, it follows that

$$M - i\left(N^{-1}\right) < a_n \leq M \text{ for all } n \geq m.$$

From this we conclude that $\lim a_n$ exists and $\lim a_n = M$.

If $M = \sup \Lambda$, for each $n \in \mathbb{N}$ we may choose $\lambda_n \in \Lambda$ such that

$$M - i\left(n^{-1}\right) < \lambda_n \leq M. \tag{3.2}$$

By replacing $\lambda_n$ by $\max\{\lambda_1, \ldots, \lambda_n\}$[2] if necessary we may assume that $\lambda_n$ is increasing in $n$. It now follows easily from Eq. (3.2) that $\lim_{n\to\infty} \lambda_n = M$. ■

### 3.1.1 The Decimal Representation of a Real Number

Let $\alpha \in \mathbb{R}$ or $\alpha \in \mathbb{Q}$, $m, n \in \mathbb{Z}$ and $S := \sum_{k=n}^{m} \alpha^k$. If $\alpha = 1$ then $\sum_{k=n}^{m} \alpha^k = m - n + 1$ while for $\alpha \neq 1$,

$$\alpha S - S = \alpha^{m+1} - \alpha^n$$

and solving for $S$ gives the important geometric summation formula,

$$\sum_{k=n}^{m} \alpha^k = \frac{\alpha^{m+1} - \alpha^n}{\alpha - 1} \text{ if } \alpha \neq 1. \tag{3.3}$$

Taking $\alpha = 10^{-1}$ in Eq. (3.3) implies

$$\sum_{k=n}^{m} 10^{-k} = \frac{10^{-(m+1)} - 10^{-n}}{10^{-1} - 1} = \frac{1}{10^{n-1}} \frac{1 - 10^{-(m-n)}}{9}$$

and in particular, for all $M \geq n$,

$$\lim_{m\to\infty} \sum_{k=n}^{m} 10^{-k} = \frac{1}{9 \cdot 10^{n-1}} \geq \sum_{k=n}^{M} 10^{-k}.$$

Let $\mathbb{D}$ denote those sequences $\alpha \in \{0, 1, 2, \ldots, 9\}^{\mathbb{Z}}$ with the following properties:

1. there exists $N \in \mathbb{N}$ such that $\alpha_{-n} = 0$ for all $n \geq N$ and
2. $\alpha_n \neq 0$ for some $n \in \mathbb{Z}$.

[2] The notation, $\max \Lambda$, denotes $\sup \Lambda$ along with the assertion that $\sup \Lambda \in \Lambda$. Similarly, $\min \Lambda = \inf \Lambda$ along with the assertion that $\inf \Lambda \in \Lambda$.

Associated to each $\alpha \in \mathbb{D}$ is the sequence $a = a(\alpha)$ defined by

$$a_n := \sum_{k=-\infty}^{n} \alpha_k 10^{-k}.$$

Since for $m > n$,

$$|a_m - a_n| = \left| \sum_{k=n+1}^{m} \alpha_k 10^{-k} \right| \le 9 \sum_{k=n+1}^{m} 10^{-k} \le 9 \frac{1}{9 \cdot 10^n} = \frac{1}{10^n},$$

it follows that

$$|a_m - a_n| \le \frac{1}{10^{\min(m,n)}} \to 0 \text{ as } m, n \to \infty.$$

Therefore $a = \underline{a(\alpha)} \in \mathcal{C}$ and we may define a map $D : \{\pm 1\} \times \mathbb{D} \to \mathbb{R}$ defined by $D(\varepsilon, \alpha) = \varepsilon \overline{a(\alpha)}$. As is customary we will denote $D(\varepsilon, \alpha) = \varepsilon \overline{a(\alpha)}$ as

$$\varepsilon \cdot \alpha_m \dots \alpha_0 . \alpha_1 \alpha_2 \dots \alpha_n \dots \tag{3.4}$$

where $m$ is the largest integer in $\mathbb{Z}$ such that $\alpha_k = 0$ for all $k < m$. If $m > 0$ the expression in Eq. (3.4) should be interpreted as

$$\varepsilon \cdot 0.0 \dots 0 \alpha_m \alpha_{m+1} \dots .$$

An element $\alpha \in \mathbb{D}$ has a tail of all 9's starting at $N \in \mathbb{N}$ if $\alpha_n = 9$ and for all $n \ge N$ and $\alpha_{N-1} \ne 9$. If $\alpha$ has a tail of 9's starting at $N \in \mathbb{N}$, then for $n > N$,

$$\begin{aligned} a_n(\alpha) &= \sum_{k=-\infty}^{N-1} \alpha_k 10^{-k} + 9 \sum_{k=N}^{n} 10^{-k} \\ &= \sum_{k=-\infty}^{N-1} \alpha_k 10^{-k} + \frac{9}{10^{N-1}} \cdot \frac{1 - 10^{-(n-N)}}{9} \\ &\to \sum_{k=-\infty}^{N-1} \alpha_k 10^{-k} + 10^{-(N-1)} \text{ as } n \to \infty. \end{aligned}$$

If $\alpha'$ is the digits in the decimal expansion of $\sum_{k=-\infty}^{N-1} \alpha_k 10^{-k} + 10^{-(N-1)}$, then

$$\alpha' \in \mathbb{D}' := \{\alpha \in \mathbb{D} : \alpha \text{ does not have a tail of all 9's}\}.$$

and we have just shown that $D(\varepsilon, \alpha) = D(\varepsilon, \alpha')$. In particular this implies

$$D(\{\pm 1\} \times \mathbb{D}') = D(\{\pm 1\} \times \mathbb{D}). \tag{3.5}$$

**Theorem 3.12 (Decimal Representation).** *The map*

$$D : \{\pm 1\} \times \mathbb{D}' \to \mathbb{R} \setminus \{0\}$$

*is a bijection.*

*Proof.* Suppose $D(\varepsilon, \alpha) = D(\delta, \beta)$ for some $(\varepsilon, \alpha)$ and $(\delta, \beta)$ in $\{\pm 1\} \times \mathbb{D}$. Since $D(\varepsilon, \alpha) > 0$ if $\varepsilon = 1$ and $D(\varepsilon, \alpha) < 0$ if $\varepsilon = -1$ it follows that $\varepsilon = \delta$. Let $a = a(\alpha)$ and $b = a(\beta)$ be the sequences associated to $\alpha$ and $\beta$ respectively. Suppose that $\alpha \neq \beta$ and let $j \in \mathbb{Z}$ be the position where $\alpha$ and $\beta$ first disagree, i.e. $\alpha_n = \beta_n$ for all $n < j$ while $\alpha_j \neq \beta_j$. For sake of definiteness suppose $\beta_j > \alpha_j$. Then for $n > j$ we have

$$\begin{aligned} b_n - a_n &= (\beta_j - \alpha_j) 10^{-j} + \sum_{k=j+1}^{n} (\beta_k - \alpha_k) 10^{-k} \\ &\geq 10^{-j} - 9 \sum_{k=j+1}^{n} 10^{-k} \geq 10^{-j} - 9 \frac{1}{9 \cdot 10^j} = 0. \end{aligned}$$

Therefore $b_n - a_n \geq 0$ for all $n$ and $\lim (b_n - a_n) = 0$ iff $\beta_j = \alpha_j + 1$ and $\beta_k = 9$ and $\alpha_k = 0$ for all $k > j$. In summary, $D(\varepsilon, \alpha) = D(\delta, \beta)$ with $\alpha \neq \beta$ implies either $\alpha$ or $\beta$ has an infinite tail of nines which shows that $D$ is injective when restricted to $\{\pm 1\} \times \mathbb{D}'$.

To see that $D$ is surjective it suffices to show any $\bar{b} \in \mathbb{R}$ with $0 < \bar{b} < 1$ is in the range of $D$. For each $n \in \mathbb{N}$, let $a_n = .\alpha_1 \ldots \alpha_n$ with $\alpha_i \in \{0, 1, 2, \ldots, 9\}$ such that

$$i(a_n) < \bar{b} \leq i(a_n) + i\left(10^{-n}\right). \tag{3.6}$$

Since $a_{n+1} = a_n + \alpha_{n+1} 10^{-(n+1)}$ for some $\alpha_{n+1} \in \{0, 1, 2, \ldots, 9\}$, we see that $a_{n+1} = .\alpha_1 \ldots \alpha_n \alpha_{n+1}$, i.e. the first $n$ digits in the decimal expansion of $a_{n+1}$ are the same as in the decimal expansion of $a_n$. Hence this defines $\alpha_n$ uniquely for all $n \geq 1$. By setting $\alpha_n = 0$ when $n \leq 0$, we have constructed from $\bar{b}$ an element $\alpha \in \mathbb{D}$. Because of Eq. (3.6), $D(1, \alpha) = \bar{b}$. ■

**Notation 3.13** *From now on we will identify $\mathbb{Q}$ with $i(\mathbb{Q}) \subset \mathbb{R}$ and elements in $\mathbb{R}$ with their decimal expansions.*

To summarize, we have constructed a complete ordered field $\mathbb{R}$ "containing" $\mathbb{Q}$ as a dense subset. Moreover every element in $\mathbb{R}$ (modulo those of the form $m10^{-n}$ for some $m \in \mathbb{Z}$ and $n \in \mathbb{N}$) has a unique decimal expansion.

**Corollary 3.14.** *The set $(0,1) := \{a \in \mathbb{R} : 0 < a < 1\}$ is uncountable while $\mathbb{Q} \cap (0,1)$ is countable.*

*Proof.* By Theorem 3.12, the set $\{0, 1, 2 \ldots, 8\}^{\mathbb{N}}$ can be mapped injectively into $(0,1)$ and therefore it follows from Lemma 2.6 that $(0,1)$ is uncountable. For each $m \in \mathbb{N}$, let $A_m := \left\{\frac{n}{m} : n \in \mathbb{N} \text{ with } n < m\right\}$. Since $\mathbb{Q} \cap (0,1) = \cup_{m=1}^{\infty} A_m$ and $\#(A_m) < \infty$ for all $m$, another application of Lemma 2.6 shows $\mathbb{Q} \cap (0,1)$ is countable. ■

## 3.2 The Complex Numbers

**Definition 3.15 (Complex Numbers).** *Let $\mathbb{C} = \mathbb{R}^2$ equipped with multiplication rule*

$$(a, b)(c, d) := (ac - bd, bc + ad) \tag{3.7}$$

*and the usual rule for vector addition. As is standard we will write $0 = (0, 0)$, $1 = (1, 0)$ and $i = (0, 1)$ so that every element $z$ of $\mathbb{C}$ may be written as $z = x1 + yi$ which in the future will be written simply as $z = x + iy$. If $z = x + iy$, let $\operatorname{Re} z = x$ and $\operatorname{Im} z = y$.*

Writing $z = a + ib$ and $w = c + id$, the multiplication rule in Eq. (3.7) becomes

$$(a + ib)(c + id) := (ac - bd) + i(bc + ad) \tag{3.8}$$

and in particular $1^2 = 1$ and $i^2 = -1$.

**Proposition 3.16.** *The complex numbers $\mathbb{C}$ with the above multiplication rule satisfies the usual definitions of a field. For example $wz = zw$ and $z(w_1 + w_2) = zw_1 + zw_2$, etc. Moreover if $z \neq 0$, $z$ has a multiplicative inverse given by*

$$z^{-1} = \frac{a}{a^2 + b^2} - i\frac{b}{a^2 + b^2}. \tag{3.9}$$

*Proof.* The proof is a straightforward verification. Only the last assertion will be verified here. Suppose $z = a + ib \neq 0$, we wish to find $w = c + id$ such that $zw = 1$ and this happens by Eq. (3.8) iff

$$ac - bd = 1 \text{ and} \tag{3.10}$$

$$bc + ad = 0. \tag{3.11}$$

Solving these equations for $c$ and $d$ gives $c = \frac{a}{a^2+b^2}$ and $d = -\frac{b}{a^2+b^2}$ as claimed. ■

**Notation 3.17 (Conjugation and Modulus)** *If $z = a + ib$ with $a, b \in \mathbb{R}$ let $\bar{z} = a - ib$ and*

$$|z| := \sqrt{z\bar{z}} = \sqrt{a^2 + b^2} = \sqrt{|\operatorname{Re} z|^2 + |\operatorname{Im} z|^2}.$$

*See Exercise 3.8 for the existence of the square root as a positive real number.*

Notice that

$$\operatorname{Re} z = \frac{1}{2}(z + \bar{z}) \text{ and } \operatorname{Im} z = \frac{1}{2i}(z - \bar{z}). \tag{3.12}$$

**Proposition 3.18.** *Complex conjugation and the modulus operators satisfy the following properties.*

1. $\bar{\bar{z}} = z$,

*2.* $\overline{zw} = \bar{z}\bar{w}$ *and* $\bar{z} + \bar{w} = \overline{z + w}$.
*3.* $|\bar{z}| = |z|$
*4.* $|zw| = |z|\,|w|$ *and in particular* $|z^n| = |z|^n$ *for all* $n \in \mathbb{N}$.
*5.* $|\mathrm{Re}\, z| \le |z|$ *and* $|\mathrm{Im}\, z| \le |z|$
*6.* $|z + w| \le |z| + |w|$.
*7.* $z = 0$ *iff* $|z| = 0$.
*8. If* $z \neq 0$ *then* $z^{-1} := \frac{\bar{z}}{|z|^2}$ *(also written as* $\frac{1}{z}$*) is the inverse of* $z$.
*9.* $\left|z^{-1}\right| = |z|^{-1}$ *and more generally* $|z^n| = |z|^n$ *for all* $n \in \mathbb{Z}$.

*Proof.* All of these properties are direct computations except for possibly the triangle inequality in item 6 which is verified by the following computation:

$$\begin{aligned}
|z + w|^2 &= (z + w)\left(\overline{z + w}\right) = |z|^2 + |w|^2 + w\bar{z} + \bar{w}z \\
&= |z|^2 + |w|^2 + w\bar{z} + \overline{w\bar{z}} \\
&= |z|^2 + |w|^2 + 2\,\mathrm{Re}\,(w\bar{z}) \le |z|^2 + |w|^2 + 2\,|z|\,|w| \\
&= (|z| + |w|)^2 .
\end{aligned}$$

■

**Definition 3.19.** *A sequence* $\{z_n\}_{n=1}^{\infty} \subset \mathbb{C}$ *is* ***Cauchy*** *if* $|z_n - z_m| \to 0$ *as* $m, n \to \infty$ *and is* ***convergent*** *to* $z \in \mathbb{C}$ *if* $|z - z_n| \to 0$ *as* $n \to \infty$. *As usual if* $\{z_n\}_{n=1}^{\infty}$ *converges to* $z$ *we will write* $z_n \to z$ *as* $n \to \infty$ *or* $z = \lim_{n\to\infty} z_n$.

**Theorem 3.20.** *The complex numbers are complete,i.e. all Cauchy sequences are convergent.*

*Proof.* This follows from the completeness of real numbers and the easily proved observations that if $z_n = a_n + ib_n \in \mathbb{C}$, then

1. $\{z_n\}_{n=1}^{\infty} \subset \mathbb{C}$ is Cauchy iff $\{a_n\}_{n=1}^{\infty} \subset \mathbb{R}$ and $\{b_n\}_{n=1}^{\infty} \subset \mathbb{R}$ are Cauchy and
2. $z_n \to z = a + ib$ as $n \to \infty$ iff $a_n \to a$ and $b_n \to b$ as $n \to \infty$.

■

## 3.3 Exercises

**Exercise 3.8.** Show to every $a \in \mathbb{R}$ with $a \ge 0$ there exists a unique number $b \in \mathbb{R}$ such that $b \ge 0$ and $b^2 = a$. Of course we will call $b = \sqrt{a}$. Also show that $a \to \sqrt{a}$ is an increasing function on $[0, \infty)$. **Hint:** To construct $b = \sqrt{a}$ for $a > 0$, to each $n \in \mathbb{N}$ let $m_n \in \mathbb{N}_0$ be chosen so that

$$\frac{m_n^2}{n^2} < a \le \frac{(m_n + 1)^2}{n^2} \text{ i.e. } i\left(\frac{m_n^2}{n^2}\right) < a \le i\left(\frac{(m_n + 1)^2}{n^2}\right)$$

and let $q_n := \frac{m_n}{n}$. Then show $b = \overline{\{q_n\}_{n=1}^{\infty}} \in \mathbb{R}$ satisfies $b > 0$ and $b^2 = a$.

# 4

# Limits and Sums

## 4.1 Limsups, Liminfs and Extended Limits

**Notation 4.1** *The* ***extended real numbers*** *is the set* $\bar{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$, *i.e. it is* $\mathbb{R}$ *with two new points called* $\infty$ *and* $-\infty$. *We use the following conventions,* $\pm\infty \cdot 0 = 0$, $\pm\infty + a = \pm\infty$ *for any* $a \in \mathbb{R}$, $\infty + \infty = \infty$ *and* $-\infty - \infty = -\infty$ *while* $\infty - \infty$ *is not defined. A sequence* $a_n \in \bar{\mathbb{R}}$ *is said to converge to* $\infty$ $(-\infty)$ *if for all* $M \in \mathbb{R}$ *there exists* $m \in \mathbb{N}$ *such that* $a_n \geq M$ $(a_n \leq M)$ *for all* $n \geq m$.

**Lemma 4.2.** *Suppose* $\{a_n\}_{n=1}^{\infty}$ *and* $\{b_n\}_{n=1}^{\infty}$ *are convergent sequences in* $\bar{\mathbb{R}}$, *then:*

1. *If* $a_n \leq b_n$ *for a.a.* $n$ *then* $\lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n$.
2. *If* $c \in \bar{\mathbb{R}}$, $\lim_{n\to\infty} (ca_n) = c \lim_{n\to\infty} a_n$.
3. *If* $\{a_n + b_n\}_{n=1}^{\infty}$ *is convergent and*
$$\lim_{n\to\infty} (a_n + b_n) = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n \tag{4.1}$$
*provided the right side is not of the form* $\infty - \infty$.
4. $\{a_n b_n\}_{n=1}^{\infty}$ *is convergent and*
$$\lim_{n\to\infty} (a_n b_n) = \lim_{n\to\infty} a_n \cdot \lim_{n\to\infty} b_n \tag{4.2}$$
*provided the right hand side is not of the for* $\infty \cdot 0$.

Before going to the proof consider the simple example where $a_n = n$ and $b_n = -\alpha n$ with $\alpha > 0$. Then

$$\lim (a_n + b_n) = \begin{cases} \infty & \text{if } \alpha < 1 \\ 0 & \text{if } \alpha = 1 \\ -\infty & \text{if } \alpha > 1 \end{cases}$$

while

$$\lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n \text{“}=\text{”} \infty - \infty.$$

This shows that the requirement that the right side of Eq. (4.1) is not of form $\infty - \infty$ is necessary in Lemma 4.2. Similarly by considering the examples $a_n = n$ and $b_n = n^{-\alpha}$ with $\alpha > 0$ shows the necessity for assuming right hand side of Eq. (4.2) is not of the form $\infty \cdot 0$.

*Proof.* The proofs of items 1. and 2. are left to the reader.

**Proof of Eq.** (4.1). Let $a := \lim_{n\to\infty} a_n$ and $b = \lim_{n\to\infty} b_n$. Case 1., suppose $b = \infty$ in which case we must assume $a > -\infty$. In this case, for every $M > 0$, there exists $N$ such that $b_n \geq M$ and $a_n \geq a - 1$ for all $n \geq N$ and this implies

$$a_n + b_n \geq M + a - 1 \text{ for all } n \geq N.$$

Since $M$ is arbitrary it follows that $a_n + b_n \to \infty$ as $n \to b = \infty$. The cases where $b = -\infty$ or $a = \pm\infty$ are handled similarly.

Case 2. If $a, b \in \mathbb{R}$, then for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$|a - a_n| \leq \varepsilon \text{ and } |b - b_n| \leq \varepsilon \text{ for all } n \geq N.$$

Therefore,

$$|a + b - (a_n + b_n)| = |a - a_n + b - b_n| \leq |a - a_n| + |b - b_n| \leq 2\varepsilon$$

for all $n \geq N$. Since $n$ is arbitrary, it follows that $\lim_{n\to\infty} (a_n + b_n) = a + b$.

**Proof of Eq.** (4.2). It will be left to the reader to prove the case where $\lim a_n$ and $\lim b_n$ exist in $\mathbb{R}$. I will only consider the case where $a = \lim_{n\to\infty} a_n \neq 0$ and $\lim_{n\to\infty} b_n = \infty$ here. Let us also suppose that $a > 0$ (the case $a < 0$ is handled similarly) and let $\alpha := \min\left(\frac{a}{2}, 1\right)$. Given any $M < \infty$, there exists $N \in \mathbb{N}$ such that $a_n \geq \alpha$ and $b_n \geq M$ for all $n \geq N$ and for this choice of $N$, $a_n b_n \geq M\alpha$ for all $n \geq N$. Since $\alpha > 0$ is fixed and $M$ is arbitrary it follows that $\lim_{n\to\infty} (a_n b_n) = \infty$ as desired. ■

For any subset $\Lambda \subset \bar{\mathbb{R}}$, let $\sup \Lambda$ and $\inf \Lambda$ denote the least upper bound and greatest lower bound of $\Lambda$ respectively. The convention being that $\sup \Lambda = \infty$ if $\infty \in \Lambda$ or $\Lambda$ is not bounded from above and $\inf \Lambda = -\infty$ if $-\infty \in \Lambda$ or $\Lambda$ is not bounded from below. We will also use the **conventions** that $\sup \emptyset = -\infty$ and $\inf \emptyset = +\infty$.

**Notation 4.3** *Suppose that* $\{x_n\}_{n=1}^{\infty} \subset \bar{\mathbb{R}}$ *is a sequence of numbers. Then*

$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty} \inf\{x_k : k \geq n\} \text{ and} \tag{4.3}$$

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty} \sup\{x_k : k \geq n\}. \tag{4.4}$$

*We will also write* $\underline{\lim}$ *for* $\liminf$ *and* $\overline{\lim}$ *for* $\limsup$.

*Remark 4.4.* Notice that if $a_k := \inf\{x_k : k \ge n\}$ and $b_k := \sup\{x_k : k \ge n\}$, then $\{a_k\}$ is an increasing sequence while $\{b_k\}$ is a decreasing sequence. Therefore the limits in Eq. (4.3) and Eq. (4.4) always exist in $\bar{\mathbb{R}}$ and

$$\begin{aligned}\liminf_{n\to\infty} x_n &= \sup_n \inf\{x_k : k \ge n\} \text{ and} \\ \limsup_{n\to\infty} x_n &= \inf_n \sup\{x_k : k \ge n\}.\end{aligned}$$

The following proposition contains some basic properties of liminfs and limsups.

**Proposition 4.5.** *Let $\{a_n\}_{n=1}^\infty$ and $\{b_n\}_{n=1}^\infty$ be two sequences of real numbers. Then*

1. $\liminf_{n\to\infty} a_n \le \limsup_{n\to\infty} a_n$ *and* $\lim_{n\to\infty} a_n$ *exists in* $\bar{\mathbb{R}}$ *iff* $\liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n \in \bar{\mathbb{R}}$.
2. *There is a subsequence* $\{a_{n_k}\}_{k=1}^\infty$ *of* $\{a_n\}_{n=1}^\infty$ *such that* $\lim_{k\to\infty} a_{n_k} = \limsup_{n\to\infty} a_n$.
3. $$\limsup_{n\to\infty} (a_n + b_n) \le \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n \tag{4.5}$$
   *whenever the right side of this equation is not of the form* $\infty - \infty$.
4. *If* $a_n \ge 0$ *and* $b_n \ge 0$ *for all* $n \in \mathbb{N}$, *then*
   $$\limsup_{n\to\infty} (a_n b_n) \le \limsup_{n\to\infty} a_n \cdot \limsup_{n\to\infty} b_n, \tag{4.6}$$
   *provided the right hand side of (4.6) is not of the form* $0 \cdot \infty$ *or* $\infty \cdot 0$.

*Proof.* Item 1. will be proved here leaving the remaining items as an exercise to the reader. Since

$$\inf\{a_k : k \ge n\} \le \sup\{a_k : k \ge n\} \; \forall\, n,$$

$$\liminf_{n\to\infty} a_n \le \limsup_{n\to\infty} a_n.$$

Now suppose that $\liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n = a \in \mathbb{R}$. Then for all $\varepsilon > 0$, there is an integer $N$ such that

$$a - \varepsilon \le \inf\{a_k : k \ge N\} \le \sup\{a_k : k \ge N\} \le a + \varepsilon,$$

i.e.

$$a - \varepsilon \le a_k \le a + \varepsilon \text{ for all } k \ge N.$$

Hence by the definition of the limit, $\lim_{k\to\infty} a_k = a$.

If $\liminf_{n\to\infty} a_n = \infty$, then we know for all $M \in (0, \infty)$ there is an integer $N$ such that

$$M \le \inf\{a_k : k \ge N\}$$

and hence $\lim_{n\to\infty} a_n = \infty$. The case where $\limsup_{n\to\infty} a_n = -\infty$ is handled similarly.

Conversely, suppose that $\lim_{n\to\infty} a_n = A \in \bar{\mathbb{R}}$ exists. If $A \in \mathbb{R}$, then for every $\varepsilon > 0$ there exists $N(\varepsilon) \in \mathbb{N}$ such that $|A - a_n| \le \varepsilon$ for all $n \ge N(\varepsilon)$, i.e.

$$A - \varepsilon \le a_n \le A + \varepsilon \text{ for all } n \ge N(\varepsilon).$$

From this we learn that

$$A - \varepsilon \le \liminf_{n\to\infty} a_n \le \limsup_{n\to\infty} a_n \le A + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, it follows that

$$A \le \liminf_{n\to\infty} a_n \le \limsup_{n\to\infty} a_n \le A,$$

i.e. that $A = \liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n$.

If $A = \infty$, then for all $M > 0$ there exists $N(M)$ such that $a_n \ge M$ for all $n \ge N(M)$. This show that $\liminf_{n\to\infty} a_n \ge M$ and since $M$ is arbitrary it follows that

$$\infty \le \liminf_{n\to\infty} a_n \le \limsup_{n\to\infty} a_n.$$

The proof for the case $A = -\infty$ is analogous to the $A = \infty$ case. ■

## 4.2 Sums of positive functions

In this and the next few sections, let $X$ and $Y$ be two sets. We will write $\alpha \subset\subset X$ to denote that $\alpha$ is a **finite** subset of $X$ and write $2_f^X$ for those $\alpha \subset\subset X$.

**Definition 4.6.** *Suppose that $a : X \to [0,\infty]$ is a function and $F \subset X$ is a subset, then*

$$\sum_F a = \sum_{x\in F} a(x) := \sup\left\{\sum_{x\in\alpha} a(x) : \alpha \subset\subset F\right\}.$$

*Remark 4.7.* Suppose that $X = \mathbb{N} = \{1, 2, 3, \dots\}$ and $a : X \to [0,\infty]$, then

$$\sum_{\mathbb{N}} a = \sum_{n=1}^{\infty} a(n) := \lim_{N\to\infty} \sum_{n=1}^{N} a(n).$$

Indeed for all $N$, $\sum_{n=1}^N a(n) \le \sum_{\mathbb{N}} a$, and thus passing to the limit we learn that

$$\sum_{n=1}^{\infty} a(n) \le \sum_{\mathbb{N}} a.$$

Conversely, if $\alpha \subset\subset \mathbb{N}$, then for all $N$ large enough so that $\alpha \subset \{1, 2, \ldots, N\}$, we have $\sum_\alpha a \le \sum_{n=1}^N a(n)$ which upon passing to the limit implies that

$$\sum_\alpha a \le \sum_{n=1}^\infty a(n).$$

Taking the supremum over $\alpha$ in the previous equation shows

$$\sum_{\mathbb{N}} a \le \sum_{n=1}^\infty a(n).$$

*Remark 4.8.* Suppose $a : X \to [0, \infty]$ and $\sum_X a < \infty$, then $\{x \in X : a(x) > 0\}$ is at most countable. To see this first notice that for any $\varepsilon > 0$, the set $\{x : a(x) \ge \varepsilon\}$ must be finite for otherwise $\sum_X a = \infty$. Thus

$$\{x \in X : a(x) > 0\} = \bigcup\nolimits_{k=1}^\infty \{x : a(x) \ge 1/k\}$$

which shows that $\{x \in X : a(x) > 0\}$ is a countable union of finite sets and thus countable by Lemma 2.6.

**Lemma 4.9.** *Suppose that $a, b : X \to [0, \infty]$ are two functions, then*

$$\sum_X (a + b) = \sum_X a + \sum_X b \text{ and}$$
$$\sum_X \lambda a = \lambda \sum_X a$$

*for all $\lambda \ge 0$.*

I will only prove the first assertion, the second being easy. Let $\alpha \subset\subset X$ be a finite set, then

$$\sum_\alpha (a + b) = \sum_\alpha a + \sum_\alpha b \le \sum_X a + \sum_X b$$

which after taking sups over $\alpha$ shows that

$$\sum_X (a + b) \le \sum_X a + \sum_X b.$$

Similarly, if $\alpha, \beta \subset\subset X$, then

$$\sum_\alpha a + \sum_\beta b \le \sum_{\alpha\cup\beta} a + \sum_{\alpha\cup\beta} b = \sum_{\alpha\cup\beta} (a + b) \le \sum_X (a + b).$$

Taking sups over $\alpha$ and $\beta$ then shows that

$$\sum_X a + \sum_X b \le \sum_X (a + b).$$

**Lemma 4.10.** *Let $X$ and $Y$ be sets, $R \subset X \times Y$ and suppose that $a : R \to \bar{\mathbb{R}}$ is a function. Let ${}_xR := \{y \in Y : (x,y) \in R\}$ and $R_y := \{x \in X : (x,y) \in R\}$. Then*

$$\sup_{(x,y)\in R} a(x,y) = \sup_{x\in X} \sup_{y\in {}_xR} a(x,y) = \sup_{y\in Y} \sup_{x\in R_y} a(x,y) \text{ and}$$
$$\inf_{(x,y)\in R} a(x,y) = \inf_{x\in X} \inf_{y\in {}_xR} a(x,y) = \inf_{y\in Y} \inf_{x\in R_y} a(x,y).$$

*(Recall the conventions:* $\sup \emptyset = -\infty$ *and* $\inf \emptyset = +\infty$.)

*Proof.* Let $M = \sup_{(x,y)\in R} a(x,y)$, $N_x := \sup_{y\in {}_xR} a(x,y)$. Then $a(x,y) \le M$ for all $(x,y) \in R$ implies $N_x = \sup_{y\in {}_xR} a(x,y) \le M$ and therefore that

$$\sup_{x\in X} \sup_{y\in {}_xR} a(x,y) = \sup_{x\in X} N_x \le M. \tag{4.7}$$

Similarly for any $(x,y) \in R$,

$$a(x,y) \le N_x \le \sup_{x\in X} N_x = \sup_{x\in X} \sup_{y\in {}_xR} a(x,y)$$

and therefore

$$\sup_{(x,y)\in R} a(x,y) \le \sup_{x\in X} \sup_{y\in {}_xR} a(x,y) = M \tag{4.8}$$

Equations (4.7) and (4.8) show that

$$\sup_{(x,y)\in R} a(x,y) = \sup_{x\in X} \sup_{y\in {}_xR} a(x,y).$$

The assertions involving infimums are proved analogously or follow from what we have just proved applied to the function $-a$. ■

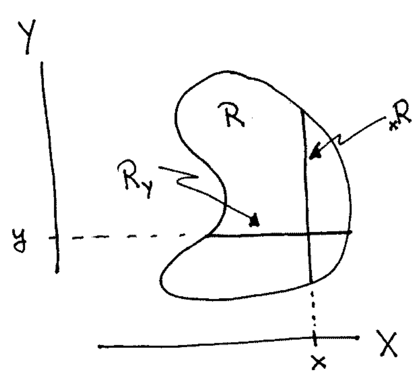


**Fig. 4.1.** The $x$ and $y$ – slices of a set $R \subset X \times Y$.

**Theorem 4.11 (Monotone Convergence Theorem for Sums).** *Suppose that* $f_n : X \to [0,\infty]$ *is an increasing sequence of functions and*

$$f(x) := \lim_{n\to\infty} f_n(x) = \sup_n f_n(x).$$

*Then*

$$\lim_{n\to\infty} \sum_X f_n = \sum_X f$$

*Proof.* We will give two proves. For the first proof, let

$$2_f^X := \{A \subset X : A \subset\subset X\}.$$

Then

$$\begin{aligned}\lim_{n\to\infty} \sum_X f_n &= \sup_n \sum_X f_n = \sup_n \sup_{\alpha\in 2_f^X} \sum_\alpha f_n = \sup_{\alpha\in 2_f^X} \sup_n \sum_\alpha f_n \\ &= \sup_{\alpha\in 2_f^X} \lim_{n\to\infty} \sum_\alpha f_n = \sup_{\alpha\in 2_f^X} \sum_\alpha \lim_{n\to\infty} f_n \\ &= \sup_{\alpha\in 2_f^X} \sum_\alpha f = \sum_X f.\end{aligned}$$

(Second Proof.) Let $S_n = \sum_X f_n$ and $S = \sum_X f$. Since $f_n \le f_m \le f$ for all $n \le m$, it follows that

$$S_n \le S_m \le S$$

which shows that $\lim_{n\to\infty} S_n$ exists and is less that $S$, i.e.

$$A := \lim_{n\to\infty} \sum_X f_n \le \sum_X f. \tag{4.9}$$

Noting that $\sum_\alpha f_n \le \sum_X f_n = S_n \le A$ for all $\alpha \subset\subset X$ and in particular,

$$\sum_\alpha f_n \le A \text{ for all } n \text{ and } \alpha \subset\subset X.$$

Letting $n$ tend to infinity in this equation shows that

$$\sum_\alpha f \le A \text{ for all } \alpha \subset\subset X$$

and then taking the sup over all $\alpha \subset\subset X$ gives

$$\sum_X f \le A = \lim_{n\to\infty} \sum_X f_n \tag{4.10}$$

which combined with Eq. (4.9) proves the theorem. ■

**Lemma 4.12 (Fatou's Lemma for Sums).** *Suppose that $f_n : X \to [0,\infty]$ is a sequence of functions, then*

$$\sum_X \lim \inf_{n\to\infty} f_n \le \lim \inf_{n\to\infty} \sum_X f_n.$$

*Proof.* Define $g_k := \inf_{n\ge k} f_n$ so that $g_k \uparrow \liminf_{n\to\infty} f_n$ as $k \to \infty$. Since $g_k \le f_n$ for all $k \le n$,

$$\sum_X g_k \le \sum_X f_n \text{ for all } n \ge k$$

and therefore

$$\sum_X g_k \le \lim \inf_{n\to\infty} \sum_X f_n \text{ for all } k.$$

We may now use the monotone convergence theorem to let $k \to \infty$ to find

$$\sum_X \lim \inf_{n\to\infty} f_n = \sum_X \lim_{k\to\infty} g_k \overset{\text{MCT}}{=} \lim_{k\to\infty} \sum_X g_k \le \lim \inf_{n\to\infty} \sum_X f_n.$$

■

*Remark 4.13.* If $A = \sum_X a < \infty$, then for all $\varepsilon > 0$ there exists $\alpha_\varepsilon \subset\subset X$ such that

$$A \ge \sum_\alpha a \ge A - \varepsilon$$

for all $\alpha \subset\subset X$ containing $\alpha_\varepsilon$ or equivalently,

$$\left| A - \sum_\alpha a \right| \le \varepsilon \tag{4.11}$$

for all $\alpha \subset\subset X$ containing $\alpha_\varepsilon$. Indeed, choose $\alpha_\varepsilon$ so that $\sum_{\alpha_\varepsilon} a \ge A - \varepsilon$.

## 4.3 Sums of complex functions

**Definition 4.14.** *Suppose that $a : X \to \mathbb{C}$ is a function, we say that*

$$\sum_X a = \sum_{x\in X} a(x)$$

*exists and is equal to $A \in \mathbb{C}$, if for all $\varepsilon > 0$ there is a finite subset $\alpha_\varepsilon \subset X$ such that for all $\alpha \subset\subset X$ containing $\alpha_\varepsilon$ we have*

$$\left| A - \sum_\alpha a \right| \le \varepsilon.$$

The following lemma is left as an exercise to the reader.

**Lemma 4.15.** *Suppose that $a, b : X \to \mathbb{C}$ are two functions such that $\sum_X a$ and $\sum_X b$ exist, then $\sum_X (a + \lambda b)$ exists for all $\lambda \in \mathbb{C}$ and*

$$\sum_X (a + \lambda b) = \sum_X a + \lambda \sum_X b.$$

**Definition 4.16 (Summable).** *We call a function $a : X \to \mathbb{C}$ **summable** if*

$$\sum_X |a| < \infty.$$

**Proposition 4.17.** *Let $a : X \to \mathbb{C}$ be a function, then $\sum_X a$ exists iff $\sum_X |a| < \infty$, i.e. iff $a$ is summable. Moreover if $a$ is summable, then*

$$\left| \sum_X a \right| \le \sum_X |a| \, .$$

*Proof.* If $\sum_X |a| < \infty$, then $\sum_X (\operatorname{Re} a)^{\pm} < \infty$ and $\sum_X (\operatorname{Im} a)^{\pm} < \infty$ and hence by Remark 4.13 these sums exists in the sense of Definition 4.14. Therefore by Lemma 4.15, $\sum_X a$ exists and

$$\sum_X a = \sum_X (\operatorname{Re} a)^+ - \sum_X (\operatorname{Re} a)^- + i \left( \sum_X (\operatorname{Im} a)^+ - \sum_X (\operatorname{Im} a)^- \right).$$

Conversely, if $\sum_X |a| = \infty$ then, because $|a| \le |\operatorname{Re} a| + |\operatorname{Im} a|$, we must have

$$\sum_X |\operatorname{Re} a| = \infty \text{ or } \sum_X |\operatorname{Im} a| = \infty.$$

Thus it suffices to consider the case where $a : X \to \mathbb{R}$ is a real function. Write $a = a^+ - a^-$ where

$$a^+(x) = \max(a(x), 0) \text{ and } a^-(x) = \max(-a(x), 0). \tag{4.12}$$

Then $|a| = a^+ + a^-$ and

$$\infty = \sum_X |a| = \sum_X a^+ + \sum_X a^-$$

which shows that either $\sum_X a^+ = \infty$ or $\sum_X a^- = \infty$. Suppose, with out loss of generality, that $\sum_X a^+ = \infty$. Let $X' := \{x \in X : a(x) \ge 0\}$, then we know that $\sum_{X'} a = \infty$ which means there are finite subsets $\alpha_n \subset X' \subset X$ such that $\sum_{\alpha_n} a \ge n$ for all $n$. Thus if $\alpha \subset\subset X$ is any finite set, it follows that $\lim_{n\to\infty} \sum_{\alpha_n \cup \alpha} a = \infty$, and therefore $\sum_X a$ can not exist as a number in $\mathbb{R}$.

Finally if $a$ is summable, write $\sum_X a = \rho e^{i\theta}$ with $\rho \ge 0$ and $\theta \in \mathbb{R}$, then

$$\begin{aligned}\left|\sum_X a\right| &= \rho = e^{-i\theta}\sum_X a = \sum_X e^{-i\theta} a \\ &= \sum_X \operatorname{Re}\left[e^{-i\theta}a\right] \le \sum_X \left(\operatorname{Re}\left[e^{-i\theta}a\right]\right)^+ \\ &\le \sum_X \left|\operatorname{Re}\left[e^{-i\theta}a\right]\right| \le \sum_X \left|e^{-i\theta}a\right| \le \sum_X |a|.\end{aligned}$$

Alternatively, this may be proved by approximating $\sum_X a$ by a finite sum and then using the triangle inequality of $|\cdot|$. ■

*Remark 4.18.* Suppose that $X = \mathbb{N}$ and $a : \mathbb{N} \to \mathbb{C}$ is a sequence, then it is not necessarily true that

$$\sum_{n=1}^{\infty} a(n) = \sum_{n\in\mathbb{N}} a(n). \tag{4.13}$$

This is because

$$\sum_{n=1}^{\infty} a(n) = \lim_{N\to\infty} \sum_{n=1}^{N} a(n)$$

depends on the ordering of the sequence $a$ where as $\sum_{n\in\mathbb{N}} a(n)$ does not. For example, take $a(n) = (-1)^n/n$ then $\sum_{n\in\mathbb{N}} |a(n)| = \infty$ i.e. $\sum_{n\in\mathbb{N}} a(n)$ does **not** exist while $\sum_{n=1}^{\infty} a(n)$ does exist. On the other hand, if

$$\sum_{n\in\mathbb{N}} |a(n)| = \sum_{n=1}^{\infty} |a(n)| < \infty$$

then Eq. (4.13) is valid.

**Theorem 4.19 (Dominated Convergence Theorem for Sums).** *Suppose that $f_n : X \to \mathbb{C}$ is a sequence of functions on $X$ such that $f(x) = \lim_{n\to\infty} f_n(x) \in \mathbb{C}$ exists for all $x \in X$. Further assume there is a **dominating function** $g : X \to [0,\infty)$ such that*

$$|f_n(x)| \le g(x) \text{ for all } x \in X \text{ and } n \in \mathbb{N} \tag{4.14}$$

*and that $g$ is summable. Then*

$$\lim_{n\to\infty} \sum_{x\in X} f_n(x) = \sum_{x\in X} f(x). \tag{4.15}$$

*Proof.* Notice that $|f| = \lim |f_n| \le g$ so that $f$ is summable. By considering the real and imaginary parts of $f$ separately, it suffices to prove the theorem in the case where $f$ is real. By Fatou's Lemma,

$$\begin{aligned}\sum_X (g \pm f) &= \sum_X \liminf_{n\to\infty} (g \pm f_n) \le \liminf_{n\to\infty} \sum_X (g \pm f_n) \\ &= \sum_X g + \liminf_{n\to\infty} \left( \pm \sum_X f_n \right).\end{aligned}$$

Since $\liminf_{n\to\infty}(-a_n) = -\limsup_{n\to\infty} a_n$, we have shown,

$$\sum_X g \pm \sum_X f \le \sum_X g + \begin{cases} \liminf_{n\to\infty} \sum_X f_n \\ -\limsup_{n\to\infty} \sum_X f_n \end{cases}$$

and therefore

$$\limsup_{n\to\infty} \sum_X f_n \le \sum_X f \le \liminf_{n\to\infty} \sum_X f_n.$$

This shows that $\lim\limits_{n\to\infty} \sum_X f_n$exists and is equal to $\sum_X f$. ■

*Proof.* (Second Proof.) Passing to the limit in Eq. (4.14) shows that $|f| \le g$ and in particular that $f$ is summable. Given $\varepsilon > 0$, let $\alpha \subset\subset X$ such that

$$\sum_{X\setminus\alpha} g \le \varepsilon.$$

Then for $\beta \subset\subset X$ such that $\alpha \subset \beta$,

$$\begin{aligned}\left|\sum_\beta f - \sum_\beta f_n\right| &= \left|\sum_\beta (f - f_n)\right| \\ &\le \sum_\beta |f - f_n| = \sum_\alpha |f - f_n| + \sum_{\beta\setminus\alpha} |f - f_n| \\ &\le \sum_\alpha |f - f_n| + 2\sum_{\beta\setminus\alpha} g \\ &\le \sum_\alpha |f - f_n| + 2\varepsilon.\end{aligned}$$

and hence that

$$\left|\sum_\beta f - \sum_\beta f_n\right| \le \sum_\alpha |f - f_n| + 2\varepsilon.$$

Since this last equation is true for all such $\beta \subset\subset X$, we learn that

$$\left|\sum_X f - \sum_X f_n\right| \le \sum_\alpha |f - f_n| + 2\varepsilon$$

which then implies that

$$\limsup_{n\to\infty}\left|\sum_X f - \sum_X f_n\right| \le \limsup_{n\to\infty}\sum_\alpha |f-f_n| + 2\varepsilon$$
$$= 2\varepsilon.$$

Because $\varepsilon > 0$ is arbitrary we conclude that

$$\limsup_{n\to\infty}\left|\sum_X f - \sum_X f_n\right| = 0.$$

which is the same as Eq. (4.15). ■

*Remark 4.20.* Theorem 4.19 may easily be generalized as follows. Suppose $f_n, g_n, g$ are summable functions on $X$ such that $f_n \to f$ and $g_n \to g$ pointwise, $|f_n| \le g_n$ and $\sum_X g_n \to \sum_X g$ as $n \to \infty$. Then $f$ is summable and Eq. (4.15) still holds. For the proof we use Fatou's Lemma to again conclude

$$\sum_X (g \pm f) = \sum_X \liminf_{n\to\infty} (g_n \pm f_n) \le \liminf_{n\to\infty} \sum_X (g_n \pm f_n)$$
$$= \sum_X g + \liminf_{n\to\infty}\left(\pm \sum_X f_n\right)$$

and then proceed exactly as in the first proof of Theorem 4.19.

## 4.4 Iterated sums and the Fubini and Tonelli Theorems

Let $X$ and $Y$ be two sets. The proof of the following lemma is left to the reader.

**Lemma 4.21.** *Suppose that $a : X \to \mathbb{C}$ is function and $F \subset X$ is a subset such that $a(x) = 0$ for all $x \notin F$. Then $\sum_F a$ exists iff $\sum_X a$ exists and when the sums exists,*

$$\sum_X a = \sum_F a.$$

**Theorem 4.22 (Tonelli's Theorem for Sums).** *Suppose that $a : X \times Y \to [0,\infty]$, then*

$$\sum_{X\times Y} a = \sum_X \sum_Y a = \sum_Y \sum_X a.$$

*Proof.* It suffices to show, by symmetry, that

$$\sum_{X\times Y} a = \sum_X \sum_Y a$$

Let $\Lambda \subset\subset X \times Y$. The for any $\alpha \subset\subset X$ and $\beta \subset\subset Y$ such that $\Lambda \subset \alpha \times \beta$, we have

$$\sum_{\Lambda} a \leq \sum_{\alpha\times\beta} a = \sum_{\alpha}\sum_{\beta} a \leq \sum_{\alpha}\sum_{Y} a \leq \sum_{X}\sum_{Y} a,$$

i.e. $\sum_{\Lambda} a \leq \sum_X \sum_Y a$. Taking the sup over $\Lambda$ in this last equation shows

$$\sum_{X\times Y} a \leq \sum_X \sum_Y a.$$

For the reverse inequality, for each $x \in X$ choose $\beta_n^x \subset\subset X$ such that $\beta_n^x \uparrow$ as $n \uparrow$ and

$$\sum_{y\in Y} a(x,y) = \lim_{n\to\infty} \sum_{y\in\beta_n^x} a(x,y).$$

If $\alpha \subset\subset X$ is a given finite subset of $X$, then

$$\sum_{y\in Y} a(x,y) = \lim_{n\to\infty} \sum_{y\in\beta_n} a(x,y) \text{ for all } x \in \alpha$$

where $\beta_n := \cup_{x\in\alpha}\beta_n^x \subset\subset X$. Hence

$$\begin{aligned}\sum_{x\in\alpha}\sum_{y\in Y} a(x,y) &= \sum_{x\in\alpha}\lim_{n\to\infty}\sum_{y\in\beta_n} a(x,y) = \lim_{n\to\infty}\sum_{x\in\alpha}\sum_{y\in\beta_n} a(x,y) \\ &= \lim_{n\to\infty}\sum_{(x,y)\in\alpha\times\beta_n} a(x,y) \leq \sum_{X\times Y} a.\end{aligned}$$

Since $\alpha$ is arbitrary, it follows that

$$\sum_{x\in X}\sum_{y\in Y} a(x,y) = \sup_{\alpha\subset\subset X}\sum_{x\in\alpha}\sum_{y\in Y} a(x,y) \leq \sum_{X\times Y} a$$

which completes the proof. ■

**Theorem 4.23 (Fubini's Theorem for Sums).** *Now suppose that $a : X \times Y \to \mathbb{C}$ is a summable function, i.e. by Theorem 4.22 any one of the following equivalent conditions hold:*

1. $\sum_{X\times Y} |a| < \infty$,
2. $\sum_X \sum_Y |a| < \infty$ *or*
3. $\sum_Y \sum_X |a| < \infty$.

*Then*

$$\sum_{X\times Y} a = \sum_X\sum_Y a = \sum_Y\sum_X a.$$

*Proof.* If $a : X \to \mathbb{R}$ is real valued the theorem follows by applying Theorem 4.22 to $a^{\pm}$ – the positive and negative parts of $a$. The general result holds for complex valued functions $a$ by applying the real version just proved to the real and imaginary parts of $a$. ■

## 4.5 Exercises

**Exercise 4.1.** Now suppose for each $n \in \mathbb{N} := \{1, 2, \ldots\}$ that $f_n : X \to \mathbb{R}$ is a function. Let

$$D := \{x \in X : \lim_{n\to\infty} f_n(x) = +\infty\}$$

show that

$$D = \cap_{M=1}^{\infty} \cup_{N=1}^{\infty} \cap_{n\geq N}\{x \in X : f_n(x) \geq M\}. \tag{4.16}$$

**Exercise 4.2.** Let $f_n : X \to \mathbb{R}$ be as in the last problem. Let

$$C := \{x \in X : \lim_{n\to\infty} f_n(x) \text{ exists in } \mathbb{R}\}.$$

Find an expression for $C$ similar to the expression for $D$ in (4.16). (Hint: use the Cauchy criteria for convergence.)

### 4.5.1 Limit Problems

**Exercise 4.3.** Prove Lemma 4.15. BRUCE: Move 4.3 and 4.4 after 4.8.

**Exercise 4.4.** Prove Lemma 4.21.

Let $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ be two sequences of real numbers.

**Exercise 4.5.** Show $\liminf_{n\to\infty}(-a_n) = -\limsup_{n\to\infty} a_n$.

**Exercise 4.6.** Suppose that $\limsup_{n\to\infty} a_n = M \in \bar{\mathbb{R}}$, show that there is a subsequence $\{a_{n_k}\}_{k=1}^{\infty}$ of $\{a_n\}_{n=1}^{\infty}$ such that $\lim_{k\to\infty} a_{n_k} = M$.

**Exercise 4.7.** Show that

$$\limsup_{n\to\infty}(a_n + b_n) \leq \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n \tag{4.17}$$

provided that the right side of Eq. (4.17) is well defined, i.e. no $\infty - \infty$ or $-\infty + \infty$ type expressions. (It is OK to have $\infty + \infty = \infty$ or $-\infty - \infty = -\infty$, etc.)

**Exercise 4.8.** Suppose that $a_n \geq 0$ and $b_n \geq 0$ for all $n \in \mathbb{N}$. Show

$$\limsup_{n\to\infty}(a_n b_n) \leq \limsup_{n\to\infty} a_n \cdot \limsup_{n\to\infty} b_n, \tag{4.18}$$

provided the right hand side of (4.18) is not of the form $0 \cdot \infty$ or $\infty \cdot 0$.

### 4.5.2 Dominated Convergence Theorem Problems

**Notation 4.24** *For $u_0 \in \mathbb{R}^n$ and $\delta > 0$, let $B_{u_0}(\delta) := \{x \in \mathbb{R}^n : |x - u_0| < \delta\}$ be the ball in $\mathbb{R}^n$ centered at $u_0$ with radius $\delta$.*

**Exercise 4.9.** Suppose $U \subset \mathbb{R}^n$ is a set and $u_0 \in U$ is a point such that $U \cap (B_{u_0}(\delta) \setminus \{u_0\}) \neq \emptyset$ for all $\delta > 0$. Let $G : U \setminus \{u_0\} \to \mathbb{C}$ be a function on $U \setminus \{u_0\}$. Show that $\lim_{u \to u_0} G(u)$ exists and is equal to $\lambda \in \mathbb{C}$,[1] iff for all sequences $\{u_n\}_{n=1}^{\infty} \subset U \setminus \{u_0\}$ which converge to $u_0$ (i.e. $\lim_{n\to\infty} u_n = u_0$) we have $\lim_{n\to\infty} G(u_n) = \lambda$.

**Exercise 4.10.** Suppose that $Y$ is a set, $U \subset \mathbb{R}^n$ is a set, and $f : U \times Y \to \mathbb{C}$ is a function satisfying:

1. For each $y \in Y$, the function $u \in U \to f(u, y)$ is continuous on $U$.[2]
2. There is a summable function $g : Y \to [0, \infty)$ such that

$$|f(u, y)| \leq g(y) \text{ for all } y \in Y \text{ and } u \in U.$$

Show that

$$F(u) := \sum_{y \in Y} f(u, y) \tag{4.19}$$

is a continuous function for $u \in U$.

**Exercise 4.11.** Suppose that $Y$ is a set, $J = (a, b) \subset \mathbb{R}$ is an interval, and $f : J \times Y \to \mathbb{C}$ is a function satisfying:

1. For each $y \in Y$, the function $u \to f(u, y)$ is differentiable on $J$,
2. There is a summable function $g : Y \to [0, \infty)$ such that

$$\left|\frac{\partial}{\partial u} f(u, y)\right| \leq g(y) \text{ for all } y \in Y \text{ and } u \in J.$$

3. There is a $u_0 \in J$ such that $\sum_{y \in Y} |f(u_0, y)| < \infty$.

Show:

a) for all $u \in J$ that $\sum_{y \in Y} |f(u, y)| < \infty$.

---

[1] More explicitly, $\lim_{u \to u_0} G(u) = \lambda$ means for every every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$|G(u) - \lambda| < \epsilon \text{ whenerver } u \in U \cap (B_{u_0}(\delta) \setminus \{u_0\}).$$

[2] To say $g := f(\cdot, y)$ is continuous on $U$ means that $g : U \to \mathbb{C}$ is continuous relative to the metric on $\mathbb{R}^n$ restricted to $U$.

b) Let $F(u) := \sum_{y \in Y} f(u,y)$, show $F$ is differentiable on $J$ and that

$$\dot{F}(u) = \sum_{y \in Y} \frac{\partial}{\partial u} f(u,y).$$

(Hint: Use the mean value theorem.)

**Exercise 4.12 (Differentiation of Power Series).** Suppose $R > 0$ and $\{a_n\}_{n=0}^{\infty}$ is a sequence of complex numbers such that $\sum_{n=0}^{\infty} |a_n| r^n < \infty$ for all $r \in (0, R)$. Show, using Exercise 4.11, $f(x) := \sum_{n=0}^{\infty} a_n x^n$ is continuously differentiable for $x \in (-R, R)$ and

$$f'(x) = \sum_{n=0}^{\infty} n a_n x^{n-1} = \sum_{n=1}^{\infty} n a_n x^{n-1}.$$

**Exercise 4.13.** Show the functions

$$e^x := \sum_{n=0}^{\infty} \frac{x^n}{n!}, \tag{4.20}$$

$$\sin x := \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \text{ and} \tag{4.21}$$

$$\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} \tag{4.22}$$

are infinitely differentiable and they satisfy

$$\frac{d}{dx} e^x = e^x \text{ with } e^0 = 1$$
$$\frac{d}{dx} \sin x = \cos x \text{ with } \sin(0) = 0$$
$$\frac{d}{dx} \cos x = -\sin x \text{ with } \cos(0) = 1.$$

**Exercise 4.14.** Continue the notation of Exercise 4.13.

1. Use the product and the chain rule to show,

$$\frac{d}{dx} \left[ e^{-x} e^{(x+y)} \right] = 0$$

   and conclude from this, that $e^{-x} e^{(x+y)} = e^y$ for all $x, y \in \mathbb{R}$. In particular taking $y = 0$ this implies that $e^{-x} = 1/e^x$ and hence that $e^{(x+y)} = e^x e^y$. Use this result to show $e^x \uparrow \infty$ as $x \uparrow \infty$ and $e^{-x} \downarrow 0$ as $x \downarrow -\infty$.
2. Use the product rule to show

$$\frac{d}{dx} \left( \cos^2 x + \sin^2 x \right) = 0$$

   and use this to conclude that $\cos^2 x + \sin^2 x = 1$ for all $x \in \mathbb{R}$.

**Exercise 4.15.** Let $\{a_n\}_{n=-\infty}^{\infty}$ be a summable sequence of complex numbers, i.e. $\sum_{n=-\infty}^{\infty} |a_n| < \infty$. For $t \geq 0$ and $x \in \mathbb{R}$, define

$$F(t,x) = \sum_{n=-\infty}^{\infty} a_n e^{-tn^2} e^{inx},$$

where as usual $e^{ix} = \cos(x) + i\sin(x)$, this is motivated by replacing $x$ in Eq. (4.20) by $ix$ and comparing the result to Eqs. (4.21) and (4.22).

1. $F(t,x)$ is continuous for $(t,x) \in [0,\infty)\times\mathbb{R}$. **Hint:** Let $Y = \mathbb{Z}$ and $u = (t,x)$ and use Exercise 4.10.
2. $\partial F(t,x)/\partial t$, $\partial F(t,x)/\partial x$ and $\partial^2 F(t,x)/\partial x^2$ exist for $t > 0$ and $x \in \mathbb{R}$. **Hint:** Let $Y = \mathbb{Z}$ and $u = t$ for computing $\partial F(t,x)/\partial t$ and $u = x$ for computing $\partial F(t,x)/\partial x$ and $\partial^2 F(t,x)/\partial x^2$. See Exercise 4.11.
3. $F$ satisfies the heat equation, namely

$$\partial F(t,x)/\partial t = \partial^2 F(t,x)/\partial x^2 \text{ for } t > 0 \text{ and } x \in \mathbb{R}.$$

# 5

# $\ell^p$ – spaces, Minkowski and Holder Inequalities

In this chapter, let $\mu : X \to (0, \infty)$ be a given function. Let $\mathbb{F}$ denote either $\mathbb{R}$ or $\mathbb{C}$. For $p \in (0, \infty)$ and $f : X \to \mathbb{F}$, let

$$\|f\|_p := (\sum_{x \in X} |f(x)|^p \mu(x))^{1/p}$$

and for $p = \infty$ let

$$\|f\|_\infty = \sup \{|f(x)| : x \in X\}.$$

Also, for $p > 0$, let

$$\ell^p(\mu) = \{f : X \to \mathbb{F} : \|f\|_p < \infty\}.$$

In the case where $\mu(x) = 1$ for all $x \in X$ we will simply write $\ell^p(X)$ for $\ell^p(\mu)$.

**Definition 5.1.** *A* ***norm*** *on a vector space $Z$ is a function $\|\cdot\| : Z \to [0, \infty)$ such that*

1. *(Homogeneity) $\|\lambda f\| = |\lambda| \, \|f\|$ for all $\lambda \in \mathbb{F}$ and $f \in Z$.*
2. *(Triangle inequality) $\|f + g\| \le \|f\| + \|g\|$ for all $f, g \in Z$.*
3. *(Positive definite) $\|f\| = 0$ implies $f = 0$.*

*A pair $(Z, \|\cdot\|)$ where $Z$ is a vector space and $\|\cdot\|$ is a norm on $Z$ is called a* ***normed vector space.***

The rest of this section is devoted to the proof of the following theorem.

**Theorem 5.2.** *For $p \in [1, \infty]$, $(\ell^p(\mu), \| \cdot \|_p)$ is a normed vector space.*

*Proof.* The only difficulty is the proof of the triangle inequality which is the content of Minkowski's Inequality proved in Theorem 5.8 below. ∎

**Proposition 5.3.** *Let $f : [0,\infty) \to [0,\infty)$ be a continuous strictly increasing function such that $f(0) = 0$ (for simplicity) and $\lim_{s\to\infty} f(s) = \infty$. Let $g = f^{-1}$ and for $s,t \geq 0$ let*

$$F(s) = \int_0^s f(s')ds' \text{ and } G(t) = \int_0^t g(t')dt'.$$

*Then for all $s,t \geq 0$,*

$$st \leq F(s) + G(t)$$

*and equality holds iff $t = f(s)$.*

*Proof.* Let

$$A_s := \{(\sigma,\tau) : 0 \leq \tau \leq f(\sigma) \text{ for } 0 \leq \sigma \leq s\} \text{ and}$$
$$B_t := \{(\sigma,\tau) : 0 \leq \sigma \leq g(\tau) \text{ for } 0 \leq \tau \leq t\}$$

then as one sees from Figure 5.1, $[0,s] \times [0,t] \subset A_s \cup B_t$. (In the figure: $s = 3$, $t = 1$, $A_3$ is the region under $t = f(s)$ for $0 \leq s \leq 3$ and $B_1$ is the region to the left of the curve $s = g(t)$ for $0 \leq t \leq 1$.) Hence if $m$ denotes the area of a region in the plane, then

$$st = m\left([0,s] \times [0,t]\right) \leq m(A_s) + m(B_t) = F(s) + G(t).$$

As it stands, this proof is a bit on the intuitive side. However, it will become rigorous if one takes $m$ to be Lebesgue measure on the plane which will be introduced later.

We can also give a calculus proof of this theorem under the additional assumption that $f$ is $C^1$. (This restricted version of the theorem is all we need in this section.) To do this fix $t \geq 0$ and let

$$h(s) = st - F(s) = \int_0^s (t - f(\sigma))d\sigma.$$

If $\sigma > g(t) = f^{-1}(t)$, then $t - f(\sigma) < 0$ and hence if $s > g(t)$, we have

$$\begin{aligned} h(s) &= \int_0^s (t - f(\sigma))d\sigma = \int_0^{g(t)} (t - f(\sigma))d\sigma + \int_{g(t)}^s (t - f(\sigma))d\sigma \\ &\leq \int_0^{g(t)} (t - f(\sigma))d\sigma = h(g(t)). \end{aligned}$$

Combining this with $h(0) = 0$ we see that $h(s)$ takes its maximum at some point $s \in (0,t]$ and hence at a point where $0 = h'(s) = t - f(s)$. The only solution to this equation is $s = g(t)$ and we have thus shown

$$st - F(s) = h(s) \leq \int_0^{g(t)} (t - f(\sigma))d\sigma = h(g(t))$$

with equality when $s = g(t)$. To finish the proof we must show $\int_0^{g(t)} (t - f(\sigma))d\sigma = G(t)$. This is verified by making the change of variables $\sigma = g(\tau)$ and then integrating by parts as follows:

$$\begin{aligned}\int_0^{g(t)} (t - f(\sigma))d\sigma &= \int_0^t (t - f(g(\tau)))g'(\tau)d\tau = \int_0^t (t - \tau)g'(\tau)d\tau \\ &= \int_0^t g(\tau)d\tau = G(t).\end{aligned}$$

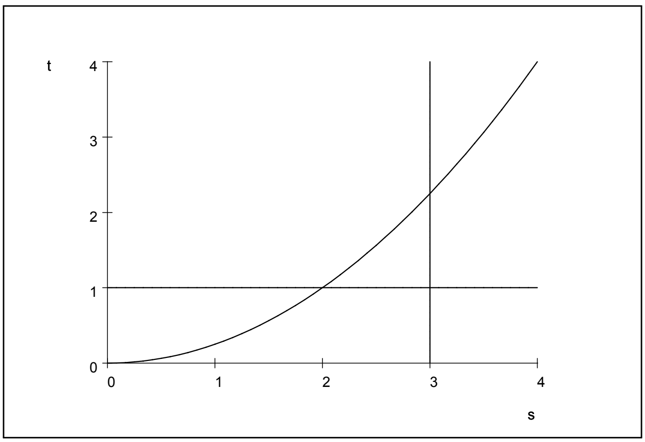


**Fig. 5.1.** A picture proof of Proposition 5.3.

■

**Definition 5.4.** *The conjugate exponent* $q \in [1, \infty]$ *to* $p \in [1, \infty]$ *is* $q := \frac{p}{p-1}$ *with the conventions that* $q = \infty$ *if* $p = 1$ *and* $q = 1$ *if* $p = \infty$. *Notice that* $q$ *is characterized by any of the following identities:*

$$\frac{1}{p} + \frac{1}{q} = 1, \ 1 + \frac{q}{p} = q, \ p - \frac{p}{q} = 1 \text{ and } q(p-1) = p. \tag{5.1}$$

**Lemma 5.5.** *Let* $p \in (1, \infty)$ *and* $q := \frac{p}{p-1} \in (1, \infty)$ *be the conjugate exponent. Then*

$$st \leq \frac{s^p}{p} + \frac{t^q}{q} \text{ for all } s, t \geq 0$$

*with equality if and only if* $t^q = s^p$.

*Proof.* Let $F(s) = \frac{s^p}{p}$ for $p > 1$. Then $f(s) = s^{p-1} = t$ and $g(t) = t^{\frac{1}{p-1}} = t^{q-1}$, wherein we have used $q-1 = p/(p-1) - 1 = 1/(p-1)$. Therefore $G(t) = t^q/q$ and hence by Proposition 5.3,

$$st \leq \frac{s^p}{p} + \frac{t^q}{q}$$

with equality iff $t = s^{p-1}$, i.e. $t^q = s^{q(p-1)} = s^p$.

For those who do not want to use Proposition 5.3, here is a direct calculus proof. Fix $t > 0$ and let

$$h(s) := st - \frac{s^p}{p}.$$

Then $h(0) = 0$, $\lim_{s\to\infty} h(s) = -\infty$ and $h'(s) = t - s^{p-1}$ which equals zero iff $s = t^{\frac{1}{p-1}}$. Since

$$h\left(t^{\frac{1}{p-1}}\right) = t^{\frac{1}{p-1}}t - \frac{t^{\frac{p}{p-1}}}{p} = t^{\frac{p}{p-1}} - \frac{t^{\frac{p}{p-1}}}{p} = t^q\left(1 - \frac{1}{p}\right) = \frac{t^q}{q},$$

it follows from the first derivative test that

$$\max h = \max\left\{h(0), h\left(t^{\frac{1}{p-1}}\right)\right\} = \max\left\{0, \frac{t^q}{q}\right\} = \frac{t^q}{q}.$$

So we have shown

$$st - \frac{s^p}{p} \leq \frac{t^q}{q} \text{ with equality iff } t = s^{p-1}.$$

■

**Theorem 5.6 (Hölder's inequality).** *Let $p, q \in [1, \infty]$ be conjugate exponents. For all $f, g : X \to \mathbb{F}$,*

$$\|fg\|_1 \leq \|f\|_p \cdot \|g\|_q. \tag{5.2}$$

*If $p \in (1, \infty)$ and $f$ and $g$ are not identically zero, then equality holds in Eq. (5.2) iff*

$$\left(\frac{|f|}{\|f\|_p}\right)^p = \left(\frac{|g|}{\|g\|_q}\right)^q. \tag{5.3}$$

*Proof.* The proof of Eq. (5.2) for $p \in \{1, \infty\}$ is easy and will be left to the reader. The cases where $\|f\|_q = 0$ or $\infty$ or $\|g\|_p = 0$ or $\infty$ are easily dealt with and are also left to the reader. So we will assume that $p \in (1, \infty)$ and $0 < \|f\|_q, \|g\|_p < \infty$. Letting $s = |f(x)| / \|f\|_p$ and $t = |g|/\|g\|_q$ in Lemma 5.5 implies

$$\frac{|f(x) g(x)|}{\|f\|_p \|g\|_q} \leq \frac{1}{p} \frac{|f(x)|^p}{\|f\|_p} + \frac{1}{q} \frac{|g(x)|^q}{\|g\|^q}$$

with equality iff

$$\frac{|f(x)|^p}{\|f\|_p} = s^p = t^q = \frac{|g(x)|^q}{\|g\|^q}. \tag{5.4}$$

Multiplying this equation by $\mu(x)$ and then summing on $x$ gives

$$\frac{\|fg\|_1}{\|f\|_p\|g\|_q} \le \frac{1}{p} + \frac{1}{q} = 1$$

with equality iff Eq. (5.4) holds for all $x \in X$, i.e. iff Eq. (5.3) holds. ■

**Definition 5.7.** *For a complex number* $\lambda \in \mathbb{C}$, *let*

$$\operatorname{sgn}(\lambda) = \begin{cases} \frac{\lambda}{|\lambda|} & \text{if } \lambda \neq 0 \\ 0 & \text{if } \lambda = 0. \end{cases}$$

*For* $\lambda, \mu \in \mathbb{C}$ *we will write* $\operatorname{sgn}(\lambda) \overset{\circ}{=} \operatorname{sgn}(\mu)$ *if either* $\lambda\mu = 0$ *or* $\lambda\mu \neq 0$ *and* $\operatorname{sgn}(\lambda) = \operatorname{sgn}(\mu)$.

**Theorem 5.8 (Minkowski's Inequality).** *If* $1 \le p \le \infty$ *and* $f, g \in \ell^p(\mu)$ *then*

$$\|f+g\|_p \le \|f\|_p + \|g\|_p. \tag{5.5}$$

*Moreover, assuming* $f$ *and* $g$ *are not identically zero, equality holds in Eq. (5.5) iff*

$$\operatorname{sgn}(f) \overset{\circ}{=} \operatorname{sgn}(g) \text{ when } p = 1 \text{ and}$$
$$f = cg \text{ for some } c > 0 \text{ when } p \in (1, \infty).$$

*Proof.* For $p = 1$,

$$\|f+g\|_1 = \sum_X |f+g|\mu \le \sum_X (|f|\mu + |g|\mu) = \sum_X |f|\mu + \sum_X |g|\mu$$

with equality iff

$$|f| + |g| = |f+g| \iff \operatorname{sgn}(f) \overset{\circ}{=} \operatorname{sgn}(g).$$

For $p = \infty$,

$$\begin{aligned} \|f+g\|_\infty &= \sup_X |f+g| \le \sup_X (|f| + |g|) \\ &\le \sup_X |f| + \sup_X |g| = \|f\|_\infty + \|g\|_\infty. \end{aligned}$$

Now assume that $p \in (1, \infty)$. Since

$$|f+g|^p \le (2\max(|f|, |g|))^p = 2^p \max(|f|^p, |g|^p) \le 2^p (|f|^p + |g|^p)$$

it follows that

$$\|f+g\|_p^p \le 2^p \left(\|f\|_p^p + \|g\|_p^p\right) < \infty.$$

Eq. (5.5) is easily verified if $\|f+g\|_p = 0$, so we may assume $\|f+g\|_p > 0$. Multiplying the inequality,

$$|f+g|^p = |f+g||f+g|^{p-1} \le (|f| + |g|)|f+g|^{p-1} \tag{5.6}$$

by $\mu$, then summing on $x$ and applying Holder's inequality two times gives

$$\sum_X |f+g|^p \mu \le \sum_X |f|\, |f+g|^{p-1}\mu + \sum_X |g|\, |f+g|^{p-1}\mu$$
$$\le (\|f\|_p + \|g\|_p)\, \|\, |f+g|^{p-1}\, \|_q. \tag{5.7}$$

Since $q(p-1) = p$, as in Eq. (5.1),

$$\||f+g|^{p-1}\|_q^q = \sum_X (|f+g|^{p-1})^q \mu = \sum_X |f+g|^p \mu = \|f+g\|_p^p. \tag{5.8}$$

Combining Eqs. (5.7) and (5.8) shows

$$\|f+g\|_p^p \le (\|f\|_p + \|g\|_p)\, \|f+g\|_p^{p/q} \tag{5.9}$$

and solving this equation for $\|f+g\|_p$ (making use of Eq. (5.1)) implies Eq. (5.5).

Now suppose that $f$ and $g$ are not identically zero and $p \in (1,\infty)$. Equality holds in Eq. (5.5) iff equality holds in Eq. (5.9) iff equality holds in Eq. (5.7) and Eq. (5.6). The latter happens iff

$$\operatorname{sgn}(f) \mathrel{\hat{=}} \operatorname{sgn}(g) \text{ and}$$
$$\left(\frac{|f|}{\|f\|_p}\right)^p = \frac{|f+g|^p}{\|f+g\|_p^p} = \left(\frac{|g|}{\|g\|_p}\right)^p. \tag{5.10}$$

wherein we have used

$$\left(\frac{|f+g|^{p-1}}{\||f+g|^{p-1}\|_q}\right)^q = \frac{|f+g|^p}{\|f+g\|_p^p}.$$

Finally Eq. (5.10) is equivalent $|f| = c\,|g|$ with $c = (\|f\|_p/\|g\|_p) > 0$ and this equality along with $\operatorname{sgn}(f) \mathrel{\hat{=}} \operatorname{sgn}(g)$ implies $f = cg$. ∎

## 5.1 Exercises

**Exercise 5.1.** Generalize Proposition 5.3 as follows. Let $a \in [-\infty, 0]$ and $f : \mathbb{R} \cap [a,\infty) \to [0,\infty)$ be a continuous strictly increasing function such that $\lim_{s\to\infty} f(s) = \infty$, $f(a) = 0$ if $a > -\infty$ or $\lim_{s\to-\infty} f(s) = 0$ if $a = -\infty$. Also let $g = f^{-1}$, $b = f(0) \ge 0$,

$$F(s) = \int_0^s f(s')ds' \text{ and } G(t) = \int_0^t g(t')dt'.$$

Then for all $s, t \ge 0$,

$$st \le F(s) + G(t \vee b) \le F(s) + G(t)$$

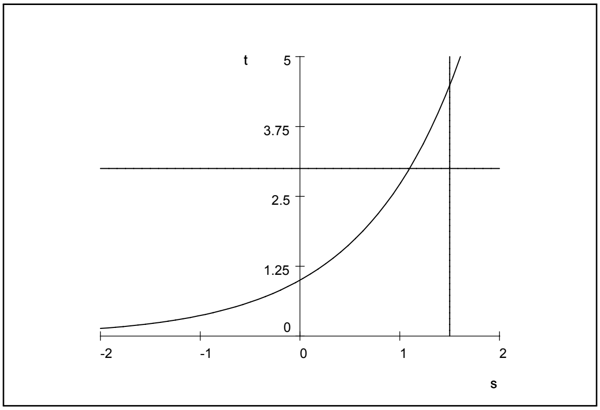


**Fig. 5.2.** Comparing areas when $t \geq b$ goes the same way as in the text.

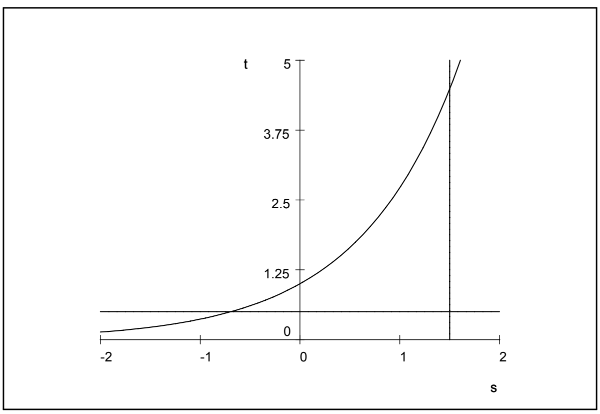


**Fig. 5.3.** When $t \leq b$, notice that $g(t) \leq 0$ but $G(t) \geq 0$. Also notice that $G(t)$ is no longer needed to estimate $st$.

and equality holds iff $t = f(s)$. In particular, taking $f(s) = e^s$, prove Young's inequality stating

$$st \leq e^s + (t \vee 1) \ln (t \vee 1) - (t \vee 1) \leq e^s + t \ln t - t.$$

**Hint:** Refer to the following pictures.