Bruce K. Driver

Undergraduate Analysis Tools

January 10, 2013  File:unanal.tex
## Contents

### Part I Numbers

1. Natural, integer, and rational Numbers
   1.1 Limits in $\mathbb{Q}$
   1.2 The Problem with $\mathbb{Q}$
   1.3 Peano’s arithmetic (Highly Optional)

2. Fields
   2.1 Basic Properties of Fields
   2.2 Ordered Fields

3. Real Numbers
   3.1 Extended real numbers
   3.2 Limsups and Liminf
   3.3 Partitioning the Real Numbers
   3.4 The Decimal Representation of a Real Number
   3.5 Summary of Key Facts about Real Numbers
   3.6 (Optional) Proofs of Theorem 3.6 and Theorem 3.3
   3.7 Supremum and Infimums of sets

4. Complex Numbers
   4.1 A Matrix Perspective (Optional)

5. Set Operations, Functions, and Counting
   5.1 Set Operations and Functions
   5.1.1 Exercises
   5.2 Cardinality
   5.3 Finite Sets
   5.4 Countable and Uncountable Sets
10.2.3 Compactness Problems

11 Connectedness in Metric Spaces
11.0.4 Connectedness Problems

Part III Appendices

A Appendix: Notation and Logic
B Appendix: More Set Theoretic Properties (highly optional)
B.1 Appendix: Zorn’s Lemma and the Hausdorff Maximal Principle (optional)
B.2 Cardinality
B.3 Formalities of Counting
C Appendix: Math 140A Topics
C.1 Summary of Key Facts about Real Numbers
C.2 Test 2 Review Topics
C.3 After Test 2 Review Topics

References

Index
Natural, integer, and rational Numbers

Notation 1.1 Let \( \mathbb{N} = \{1, 2, \ldots \} \) denote the natural numbers, \( \mathbb{N}_0 = \{0\} \cup \mathbb{N} \), \( \mathbb{Z} := \{0, \pm 1, \pm 2, \ldots \} = \{ \pm n : n \in \mathbb{N}_0 \} \) be the integers, and \( \mathbb{Q} := \left\{ \frac{m}{n} : m \in \mathbb{Z} \text{ and } n \in \mathbb{N} \right\} \) be the rationale numbers.

I am going to assume that the reader is familiar with all the standard arithmetic operations (addition, multiplication, inverses, etc.) on \( \mathbb{N}_0, \mathbb{Z}, \) and \( \mathbb{Q} \). However let us review the important induction axiom of the natural numbers.

Induction Axiom If \( S \subset \mathbb{N} \) is a subset such that \( 1 \in S \) and \( n + 1 \in S \) whenever \( n \in S \), then \( S = \mathbb{N} \).

This axiom leads takes on two other useful forms which we describe in the next Propositions.

Proposition 1.2 (Strong form of Induction). Suppose \( S \subset \mathbb{N} \) is a subset such that \( 1 \in S \) and \( n + 1 \in S \) whenever \( \{1, 2, \ldots, n\} \subset S \), then \( S = \mathbb{N} \).

Proof. Let \( T := \{n \in \mathbb{N} : \{1, 2, \ldots, n\} \subset S \} \). Then \( 1 \in T \) and if \( n \in T \) then \( n + 1 \in T \) by assumption. Therefore by the induction axiom, \( T = \mathbb{N} \) so that \( \{1, 2, \ldots, n\} \subset S \) in for all \( n \in \mathbb{N} \). This suffices to show \( S = \mathbb{N} \).

Proposition 1.3 (Well ordering principle). Suppose \( S \subset \mathbb{N} \) is a non-empty subset, then there exists a smallest element \( m \) of \( S \).

Proof. Let \( S \) be a subset of \( \mathbb{N} \) for which there is no smallest element, \( m \in S \). Let \( T = \{n \in \mathbb{N} : n < s \text{ for all } s \in S \} \).

If \( 1 \not\in T \), then \( 1 \in S \) and \( 1 \) would be a smallest element of \( S \). Hence we must have \( 1 \in T \). Now suppose that \( n \in T \) so that \( n < s \) for all \( s \in S \). If \( n + 1 \not\in T \) then there exists \( s \in S \) such that \( n < s \leq n + 1 \) which would force \( s = n + 1 \in S \). But we would then have \( n + 1 \) is the minimal element of \( S \) which is assumed not to exist. So we have shown if \( n \in T \) then \( n + 1 \in T \). So by the induction axiom of \( \mathbb{N} \) it follows that \( T = \mathbb{N} \) and therefore \( n \not\in S \) for all \( n \in \mathbb{N} \), i.e. \( S = \emptyset \).

Remark 1.4. Let us further observe that the well ordering principle implies the induction axiom. Indeed, suppose that \( S \subset \mathbb{N} \) is a subset such that \( 1 \in S \) and \( n + 1 \in S \) whenever \( n \in S \). For sake of contradiction suppose that \( S \neq \mathbb{N} \) so that \( T := \mathbb{N} \setminus S \) is not empty. By the well ordering principle there \( T \) has a unique minimal element \( m \) and in particular \( T \subset \{m, m+1, \ldots \} \). This implies that \( \{1,2,\ldots,m-1\} \subset S \) and then by assumption that \( \{1,2,\ldots,m\} \subset S \). But this then implies \( T \subset \{m+1, \ldots \} \) and therefore \( m \not\in T \) which violates \( m \) being the minimal element of \( T \). We have arrived at the desired contradiction and therefore conclude that \( S = \mathbb{N} \).

Remark 1.5. Recall that, for \( q \in \mathbb{Q} \), we define

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

Recall that, for all \( a, b \in \mathbb{Q} \),

\[
|a + b| \leq |a| + |b|, \quad |ab| = |a| |b|, \quad \text{and} \quad \left| \frac{1}{a} \right| = \frac{1}{|a|} \text{ when } a \neq 0.
\]

It is also often useful to keep in mind the following statements are equivalent for \( a, b \in \mathbb{Q} \) with \( b \geq 0 \):

1. \( |a| \leq b \),
2. \( -b \leq a \leq b \), and
3. \( \pm a \leq b \), i.e. \( a \leq b \) and \( -a \leq b \).

Lemma 1.6. If \( a, b \in \mathbb{Q} \), then

\[
||b| - |a|| \leq |b - a|.
\]  \hspace{1cm} (1.1)

Proof. Since both sides of Eq. (1.1) are symmetric in \( a \) and \( b \), we may assume that \( |b| \geq |a| \) so that \( ||b| - |a|| \leq |b - |a|| \).

Since

\[
|b| = |b - a + a| \leq |b - a| + |a|,
\]

it follows that

\[
||b| - |a|| = |b - |a|| \leq |b - a|.
\]
Theorem 1.8. The rational numbers have the following properties:

1. For any \( p \in \mathbb{Q} \) there exists \( N \in \mathbb{N} \) such that \( p < N \).
2. For all \( \varepsilon \in \mathbb{Q} \) with \( \varepsilon > 0 \) there exists an \( N \in \mathbb{N} \) such that \( 0 < \frac{1}{N} < \varepsilon \).
3. If \( a, b \in \mathbb{Q} \) and \( a \leq b + \varepsilon \) for all \( \varepsilon > 0 \), then \( a \leq b \).

Proof. 1. If \( p \leq 0 \) we may take \( N = 1 \). So suppose that \( p = \frac{m}{n} \) with \( m, n \in \mathbb{N} \). In this case let \( N = m \).

2. Write \( \varepsilon = \frac{m}{n} \) with \( m, n \in \mathbb{N} \) and then take \( N = 2n \).

3. If \( a \leq b \) is false happens iff \( a > b \) which is equivalent to \( a - b > 0 \). If we now let \( \varepsilon := \frac{a - b}{2} > 0 \), then

\[
a = b + (b - a) > b + \varepsilon
\]

which would violate the assumption that \( a \leq b + \varepsilon \) for all \( \varepsilon > 0 \).

Remark 1.7 (Some basic philosophies of real analysis). Let \( a, b, \varepsilon \) be numbers (i.e. in \( \mathbb{Q} \) or later real numbers). We will often prove:

1. \( a \leq b \) by showing that \( a \leq b + \varepsilon \) for all \( \varepsilon > 0 \). (See the next theorem.)
2. \( a = b \) by proving \( a \leq b \) and \( b \leq a \) or
3. \( a = b \) by showing \( |b - a| \leq \varepsilon \) for all \( \varepsilon > 0 \).

Theorem 1.8. The rational numbers have the following properties:

1. For any \( p \in \mathbb{Q} \) there exists \( N \in \mathbb{N} \) such that \( p < N \).
2. For any \( \varepsilon \in \mathbb{Q} \) with \( \varepsilon > 0 \) there exists an \( N \in \mathbb{N} \) such that \( 0 < \frac{1}{N} < \varepsilon \).
3. If \( a, b \in \mathbb{Q} \) and \( a \leq b + \varepsilon \) for all \( \varepsilon > 0 \), then \( a \leq b \).

Proof. 1. If \( p \leq 0 \) we may take \( N = 1 \). So suppose that \( p = \frac{m}{n} \) with \( m, n \in \mathbb{N} \). In this case let \( N = m \).

2. Write \( \varepsilon = \frac{m}{n} \) with \( m, n \in \mathbb{N} \) and then take \( N = 2n \).

3. If \( a \leq b \) is false happens iff \( a > b \) which is equivalent to \( a - b > 0 \). If we now let \( \varepsilon := \frac{a - b}{2} > 0 \), then

\[
a = b + (b - a) > b + \varepsilon
\]

which would violate the assumption that \( a \leq b + \varepsilon \) for all \( \varepsilon > 0 \).
Proof. Since $a \neq 0$ we know that $|a| > 0$. Hence, there exists $M := M_{|a|} \in \mathbb{N}$ such that $|a_n - a| < \frac{|a|}{2}$ for all $n \geq M$. Therefore for $n \geq M$

$$|a| = |a - a_n + a_n| \leq |a - a_n| + |a_n| < \frac{|a|}{2} + |a_n|$$

from which it follows that $|a_n| > \frac{|a|}{2}$ for all $n \geq M$. If $\varepsilon > 0$ is given arbitrarily, we may choose $N \geq M$ such that $|a - a_n| < \varepsilon$ for all $n \geq M$. Then for $n \geq N$ we have,

$$\left| \frac{1}{a_n} - \frac{1}{a} \right| = \frac{|a - a_n|}{a_n a} = \frac{|a - a_n|}{|a_n|} \leq \frac{\varepsilon}{2} < \frac{|a|}{2}.$$

As $\varepsilon > 0$ is arbitrary it follows that $\frac{|a|}{2} > 0$ is arbitrarily small as well (replace $\varepsilon$ by $\varepsilon |a|^2/2$ if you feel it is necessary), and hence we may conclude that Eq. 1.2 holds.

Variation on the method. In order to make these arguments more routine, it is often a good idea to write $a_n = a + \delta_n$, where $\delta_n := a_n - a$ is the error between $a_n$ and $a$. By assumption, $\lim_{n \to \infty} \delta_n = 0$ and so for any $\delta > 0$ given there exists $N(\delta) \in \mathbb{N}$ such that $|\delta_n| \leq \delta$. With this notation we have,

$$\left| \frac{1}{a_n} - \frac{1}{a} \right| = \left| \frac{1}{a + \delta_n} - \frac{1}{a} \right| = \left| \frac{-\delta_n}{a + \delta_n} \frac{1}{a} \right| \leq \frac{\delta}{|a|}.$$ 

So if we assume that $\delta \leq |a|/2$ we find that

$$\left| \frac{1}{a_n} - \frac{1}{a} \right| \leq \frac{2}{|a|^2} \delta$$

Taking $\delta = \delta(\varepsilon) = \min \left( |\varepsilon|/2, |\varepsilon|^2/2 \right)$ in Eq. 1.3 shows for $n \geq N(\delta(\varepsilon))$ that $\left| \frac{1}{a_n} - \frac{1}{a} \right| \leq \varepsilon$. Since $\varepsilon > 0$ is arbitrary we may conclude that $\lim_{n \to \infty} \frac{1}{a_n} = \frac{1}{a}$.

End of Lecture 1, 9/28/2012

Definition 1.14. A sequence $\{a_n\}_{n=1}^\infty \subset \mathbb{Q}$ is Cauchy if $|a_n - a_m| \to 0$ as $m, n \to \infty$. More precisely we require for each $\varepsilon > 0$ in $\mathbb{Q}$ that $|a_m - a_n| \leq \varepsilon$ for a.a. pairs $(m, n)$, i.e. there should exist $N \in \mathbb{N}$ such that $|a_m - a_n| \leq \varepsilon$ for all $m, n \geq N$.

Exercise 1.1. Show that all convergent sequences $\{a_n\}_{n=1}^\infty \subset \mathbb{Q}$ are Cauchy.

Exercise 1.2. Show all Cauchy sequences $\{a_n\}_{n=1}^\infty$ are bounded – i.e. there exists $M \in \mathbb{N}$ such that

$$|a_n| \leq M$$

for all $n \in \mathbb{N}$.

Exercise 1.3. Suppose $\{a_n\}_{n=1}^\infty$ and $\{b_n\}_{n=1}^\infty$ are Cauchy sequences in $\mathbb{Q}$. Show $\{a_n + b_n\}_{n=1}^\infty$ and $\{a_n \cdot b_n\}_{n=1}^\infty$ are Cauchy.

Exercise 1.4. Assume that $\{a_n\}_{n=1}^\infty$ and $\{b_n\}_{n=1}^\infty$ are convergent sequences in $\mathbb{Q}$. Show $\{a_n + b_n\}_{n=1}^\infty$ and $\{a_n \cdot b_n\}_{n=1}^\infty$ are convergent in $\mathbb{Q}$ and

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

Exercise 1.5. Assume that $\{a_n\}_{n=1}^\infty$ and $\{b_n\}_{n=1}^\infty$ are convergent sequences in $\mathbb{Q}$ such that $a_n \leq b_n$ for all $n$. Show $A := \lim_{n \to \infty} a_n \leq \lim_{n \to \infty} b_n := B$.

Exercise 1.6 (Sandwich Theorem). Assume that $\{a_n\}_{n=1}^\infty$ and $\{b_n\}_{n=1}^\infty$ are convergent sequences in $\mathbb{Q}$ such that $\lim_{n \to \infty} a_n = \lim_{n \to \infty} b_n$. If $\{x_n\}_{n=1}^\infty$ is another sequence in $\mathbb{Q}$ which satisfies $a_n \leq x_n \leq b_n$ for all $n$, then

$$\lim_{n \to \infty} x_n := a := \lim_{n \to \infty} a_n = \lim_{n \to \infty} b_n.$$ 

Please note that that main part of the problem is to show that $\lim_{n \to \infty} x_n$ exists in $\mathbb{Q}$. Hint: start by showing; if $a \leq x \leq b$ then $|x| \leq \max(|a|, |b|)$.

Definition 1.15 (Subsequence). We say a sequence, $\{y_k\}_{k=1}^\infty$ is a subsequence of another sequence, $\{x_n\}_{n=1}^\infty$, provided there exists a strictly increasing function, $\mathbb{N} \ni k \to n_k \in \mathbb{N}$ such that $y_k = x_{n_k}$ for all $k \in \mathbb{N}$. Example, $n_k = k^2 + 3$, and $\{y_k := x_{k^2 + 3}\}_{k=1}^\infty$ would be a subsequence of $\{x_n\}_{n=1}^\infty$.

Exercise 1.7. Suppose that $\{x_n\}_{n=1}^\infty$ is a Cauchy sequence in $\mathbb{Q}$ (or $\mathbb{R}$) which has a convergent subsequence, $\{y_k = x_{n_k}\}_{k=1}^\infty$ in $\mathbb{Q}$ (or $\mathbb{R}$). Show that $\lim_{n \to \infty} x_n$ exists and is equal to $\lim_{k \to \infty} y_k$.

1.2 The Problem with $\mathbb{Q}$

The problem with $\mathbb{Q}$ is that it is full of “holes.” To be more precise, $\mathbb{Q}$ is not “complete,” i.e. not all Cauchy sequences are convergent. In fact, according to Corollary 5.31 below, “most” Cauchy sequences of rational numbers do not converge to a rational number. Let us demonstrate some examples pointing out this flaw. We first pause to recall how to sum geometric series.
Lemma 1.16 (Geometric Series). Let \( \alpha \in \mathbb{Q}, m, n \in \mathbb{Z} \) with \( n \leq m \), and \( S := \sum_{k=n}^m \alpha^k \). Then
\[
S = \begin{cases} 
  m - n + 1 & \text{if } \alpha = 1 \\
  \frac{\alpha^{m+1} - \alpha^n}{\alpha - 1} & \text{if } \alpha \neq 1.
\end{cases}
\]
Moreover if \( 0 \leq \alpha < 1 \), then
\[
\sum_{k=n}^m \alpha^k = \frac{\alpha^n (1 - \alpha^{m-n+1})}{1 - \alpha} \leq \frac{\alpha^n}{1 - \alpha}.
\]  
(1.4)

**Proof.** When \( \alpha = 1 \),
\[
S = \sum_{k=n}^m 1^k = m - n + 1.
\]
If \( \alpha \neq 1 \), then
\[
\alpha S - S = \alpha^{m+1} - \alpha^n.
\]
Solving for \( S \) gives
\[
S = \sum_{k=n}^m \alpha^k = \frac{\alpha^{m+1} - \alpha^n}{\alpha - 1} \quad \text{if } \alpha \neq 1.
\]  
(1.5)

**Example 1.17.** Let \( S_n := \sum_{k=0}^n \frac{1}{k!} \in \mathbb{Q} \) for all \( n \in \mathbb{N} \). For \( n > m \) in \( \mathbb{N} \) we have,
\[
0 \leq S_n - S_m = \sum_{k=m+1}^n \frac{1}{k!} = \sum_{j=1}^{n-m} \frac{1}{(m+j)!} \\
= \frac{1}{(m+1)!} + \cdots + \frac{1}{m!} \\
\leq \frac{1}{m!} \left[ \frac{1}{m+1} + \left( \frac{1}{m+1} \right)^2 + \cdots + \left( \frac{1}{m+1} \right)^{n-m} \right] \\
\leq \frac{1}{m!} \frac{1}{m+1} \frac{1}{m+2} \cdots \frac{1}{m+n} = \frac{1}{m \cdot m!}.
\]  
(1.6)

wherein we have used Eq. (1.4) for the last inequality. From this inequality it follows that \( \{S_n\}_{n=0}^\infty \) is a Cauchy sequence and we also have,
\[
\frac{1}{(m+1)!} \leq S_n - S_m \leq \frac{1}{m \cdot m!} \quad \text{for all } n > m.
\]  
(1.7)

Suppose that \( e := \lim_{n \to \infty} S_n \) were to exist in \( \mathbb{Q} \). Then letting \( n \to \infty \) in Eq. (1.7) would show,
\[
0 < \frac{1}{(m+1)!} \leq e - S_m \leq \frac{1}{m \cdot m!}.
\]

Multiplying this inequality by \( m! \) then implies,
\[
0 < m! e - m! S_m \leq \frac{1}{m}.
\]

However for \( m \) sufficiently large \( m! e \in \mathbb{N} \) (as \( e \) is assumed to be rational) and \( m! S_m \) is always in \( \mathbb{N} \) and therefore \( k := m! e - m! S_m \in \mathbb{N} \). But there is no element \( k \in \mathbb{N} \) such that \( 0 < k < \frac{1}{m} \) and hence we must conclude \( \lim_{n \to \infty} S_n \) can not exist in \( \mathbb{Q} \). **Moral:** the number \( e = \sum_{n=0}^\infty \frac{1}{n!} = \lim_{n \to \infty} (1 + \frac{1}{n})^n \) that you learned about in calculus is not in \( \mathbb{Q} \)!

**Example 1.18 (Square roots need not exist).** The square root, \( \sqrt{2} \), of 2 does not exist in \( \mathbb{Q} \). Indeed, if \( \sqrt{2} = \frac{m}{n} \) where \( m \) and \( n \) have no common factors (in particular no common factors of 2 so that either \( m \) or \( n \) is odd), then
\[
\frac{m^2}{n^2} = 2 \implies m^2 = 2n^2.
\]

This shows that \( m^2 \) is even which would then imply that \( m = 2k \) is even (since odd-odd=odd). However this implies \( 4k^2 = 2n^2 \) from which it follows that \( n^2 = 2k^2 \) is even and hence \( n \) is even. But this contradicts the assumption that \( m \) and \( n \) had no common factors (of 2).
Exercise 1.8. Use the following outline to construct another Cauchy sequence \( \{q_n\}_{n=1}^\infty \subseteq \mathbb{Q} \) which is not convergent in \( \mathbb{Q} \).

1. Recall that there is no element \( q \in \mathbb{Q} \) such that \( q^2 = 2 \). To each \( n \in \mathbb{N} \) let \( m_n \in \mathbb{N} \) be chosen so that
\[
\frac{m_n^2}{n^2} < 2 < \frac{(m_n + 1)^2}{n^2} \tag{1.8}
\]
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 \( \mathbb{Q} \).

Example 1.19. It is also a fact that \( \pi \notin \mathbb{Q} \) where
\[
\pi = 2 \int_0^\infty \frac{1}{1 + x^2} \, dx = 2 \lim_{N \to \infty} \int_0^N \frac{1}{1 + x^2} \, dx
= 2 \lim_{N \to \infty} \sum_{k=1}^{N^2} \frac{1}{1 + \left( \frac{k}{N} \right)^2} \cdot \frac{1}{N}
= \lim_{N \to \infty} \sum_{k=1}^{N^2} \frac{2N}{N^2 + k^2}.
\]

The point is that the basic operations from calculus tend to produce “real numbers” which are not rational even though we start with only rational numbers.

- End of Lecture 2, 10/1/2012

1.3 Peano’s arithmetic (Highly Optional)

This section is for those who want to understand \( \mathbb{N} \) at a more fundamental level. Here we start with Peano’s rather minimalist axioms for \( \mathbb{N} \) and show how they lead to all the standard properties you are used to using for \( \mathbb{N} \). Here are the axioms;

non-empty \( \mathbb{N}_0 \) is a non-empty set which contains a distinguished element, 0.

We let \( \mathbb{N} := \mathbb{N}_0 \setminus \{0\} \) and call these the natural numbers.

Successor Function There is an injective\(^4\) function, \( s : \mathbb{N}_0 \to \mathbb{N} \) and we let \( 1 := s(0) \in \mathbb{N} \).

Induction hypothesis If \( S \subseteq \mathbb{N}_0 \) is a set such that \( 0 \in S \) and \( s(n) \in S \) whenever \( n \in S \), then \( S = \mathbb{N}_0 \).

Assuming these axioms one may develop all of the properties or \( \mathbb{N}_0 \) that you are accustomed to seeing. I will develop the basic properties of addition, multiplication, and the ordering on \( \mathbb{N}_0 \) in this section. For more on this point and then the further construction of \( \mathbb{Z} \) and \( \mathbb{Q} \) from \( \mathbb{N}_0 \), the reader is referred to the notes; “Numbers” by M. Taylor. You may also consult E. Landau’s book \cite{Landau} for a very detailed (but perhaps too long winded) exposition of these topics.

Lemma 1.20. The map \( s : \mathbb{N}_0 \to \mathbb{N} \) is a bijection.

Proof. Let \( S := s(\mathbb{N}_0) \cup \{0\} \subset \mathbb{N}_0 \). Then \( 0 \in S \) and \( s(0) \in s(\mathbb{N}_0) \subset S \). Moreover if \( x \in \mathbb{N} \cap S \) then \( s(x) \in s(\mathbb{N}_0) \subset S \) so that \( x \in S \implies s(x) \in S \) and hence \( S = \mathbb{N}_0 \) and therefore \( s(\mathbb{N}_0) = \mathbb{N} \).

Theorem 1.21 (Addition). There exists a function \( p : \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}_0 \) such that \( p(x,0) = x \) for all \( x \in \mathbb{N}_0 \) and \( p(x,s(y)) = s(p(x,y)) \) for all \( x,y \in \mathbb{N}_0 \). Moreover, we may construct \( p \) so that \( p(s(x),y) = p(x,s(y)) \) for all \( x,y \in \mathbb{N}_0 \).

This function \( p \) satisfies the following properties;

1. \( p(x,0) = x = p(0,x) \) for all \( x \in \mathbb{N}_0 \),
2. \( p(x,1) = p(1,x) = s(x) \) for all \( x \in \mathbb{N}_0 \),
3. \( p(x,y) = p(y,x) \) for all \( x,y \in \mathbb{N}_0 \),
4. \( p(x,p(y,z)) = p(p(x,y),z) \) for all \( x,y,z \in \mathbb{N}_0 \).

Proof. We will construct \( p \) inductively. Let
\[
S := \{ x \in \mathbb{N} : \exists p_x : \mathbb{N}_0 \to \mathbb{N}_0 \ni p_x(0) = x \text{ and } p_x(s(y)) = s(p_x(y)) \forall y \in \mathbb{N}_0 \}.
\]
Taking \( p_0(y) = y \) shows \( 0 \in S \). Moreover if \( x \in S \) we define
\[
p_x(y) := s(p_x(y)) \text{ for all } y \in \mathbb{N}_0.
\]
We then have \( p_{s(x)}(0) = s(p_x(0)) = s(x) \) and
\[
p_{s(x)}(s(y)) := s(p_x(s(y))) = s(s(p_x(y))) = s(p_{s(x)}(y))
\]
which shows \( s(x) \in S \). Thus we may conclude \( S = \mathbb{N}_0 \) and we may now define \( p(x,y) := p_x(y) \) for all \( x, y \in \mathbb{N}_0 \). By construction this function satisfies,
\[
p(s(x),y) = s(p(x,y)) = p(x,s(y)).
\]
We now verify the properties in items 1. – 4.
1. By construction $p(x, 0) = x$ for all $x \in \mathbb{N}_0$. Let $S = \{ x \in \mathbb{N} : p(0, x) = x \}$, then $0 \in S$ and if $x \in S$ we have $p(0, s(x)) = s(p(0, x)) = s(x)$ so that $s(x) \in S$. Therefore $S = \mathbb{N}_0$ and the first item holds.

2. $p(x, 1) = p(x, s(0)) = s(p(x, 0)) = s(x)$ and $p(1, x) = p(s(0), x) = s(p(0, x)) = s(x)$ so that item 2. is proved.

3. Let $S = \{ x \in \mathbb{N}_0 : p(x, \cdot) = p(\cdot, x) \}$. Then by items 1 and 2. it follows that $0, 1 \in S$. Moreover if $x \in S$, then for all $y \in \mathbb{N}_0$ we find,

$$p(s(x), y) = s(p(x, y)) = s(p(y, x)) = p(y, s(x))$$

which shows $s(x) \in S$. Therefore $S = \mathbb{N}_0$ and item 3. is proved.

4. Let

$$S := \{ x \in \mathbb{N}_0 : p(x, p(y, z)) = p(p(x, y), z) \ \forall y, z \in \mathbb{N}_0 \}.$$ 

Then $0 \in S$ and if $x \in S$ we find,

$$p(s(x), p(y, z)) = s(p(x, p(y, z))) = s(p(p(x, y), z)) = p(s(p(x, y)), z) = p(p(s(x), y), z)$$

which shows that $s(x) \in S$ and therefore $S = \mathbb{N}_0$ and item 4. is proved.

\[ \blacksquare \]

**Notation 1.22** We now write $x + y$ for $p(x, y)$ and refer to the symmetric binary operator, $+$, as addition.

To summarize we have now shown addition satisfies for all $x, y, z \in \mathbb{N}_0$:

1. $x + 0 = 0 + x = x$,
2. $s(x) + 1 = x + 1$,
3. $x + y = y + x$,
4. $(x + y) + z = x + (y + z)$.
5. The induction hypothesis may now be written as; if $S \subset \mathbb{N}_0$ is a subset such that $0 \in S$ and $n + 1 \in S$ whenever $n \in S$, then $S = \mathbb{N}_0$.

**Proposition 1.23 (Additive Cancellation).** If $x, y, z \in \mathbb{N}_0$ and $x + z = y + z$, then $x = y$.

**Proof.** Let $S$ be those $z \in \mathbb{N}_0$ for which the statement $x + z = y + z$ implies $x = y$ holds. It is clear that $0 \in S$. Moreover if $z \in S$ and $x + (z + 1) = y + (z + 1)$ then $(x + 1) + z = (y + 1) + z$ and so by the inductive hypothesis $s(x) = x + 1 = y + 1 = s(y)$. Recall that $s$ is one to one by assumption and therefore we may conclude $x = y$ and we have shown $s(z) \in S$. Therefore $S = \mathbb{N}_0$ and the proposition is proved.

\[ \blacksquare \]

**Definition 1.24.** Given $x, y \in \mathbb{N}_0$, we say $x < y$ iff $y = x + n$ for some $n \in \mathbb{N}$ and $x \leq y$ iff $y = x + n$ for some $n \in \mathbb{N}_0$. We further let $R_x := \{ x + n : n \in \mathbb{N}_0 \}$ so that $y \geq x$ iff $y \in R_x$.

**Proposition 1.25.** If $x, y \in \mathbb{N}_0$ and $x \leq y$ and $y \leq x$ then $x = y$. Moreover if $x \leq y$ then either $x < y$ or $x = y$.

**Proof.** By assumption there exists $m, n \in \mathbb{N}_0$ such that $y = x + m$ and $x = y + n$ and therefore $y = y + (m + n)$. Hence by cancellation it follows that $m + n = 0$. If $n \neq 0$ then $n = s(x)$ for some $x \in \mathbb{N}_0$ and we have $m + n = m + s(x) = s(m + x) \in \mathbb{N}$ which would imply $m + n \neq 0$. Thus we conclude that $m = 0 = n$ and therefore $x = y$.

If $x \leq y$ and $x \neq y$ then $y = x + n$ for some $n \in \mathbb{N}_0$ with $n \neq 0$, i.e. $x < y$.

**Proposition 1.26.** If $x, y \in \mathbb{N}_0$ then precisely one of the following three choices must hold, 1) $x < y$, 2) $x = y$, 3) $y < x$.

**Proof.** Suppose that $x \leq y$ does not hold, i.e. $y \notin R_x$. We wish to show that $y < x$, i.e. that $x = y + n$ for some $n \in \mathbb{N}$. We do this by induction on $y$. That is let $S$ be the the set of $y \in \mathbb{N}$ such that the statement $y \notin R_x$ implies $y < x$ holds. If $y = 0 \notin R_x$ implies $n := x \neq 0$ so that $y = x + n$, i.e. $y < x$. This shows $0 \in S$. Now suppose that $y \in S$ and that $y + 1 \notin R_x = \{ x + m : m \in \mathbb{N}_0 \}$. It follows that $y + 1 \neq x + m + 1$ for all $m \in \mathbb{N}_0$ and hence that $y \neq x + m$ for all $m \in \mathbb{N}_0$, i.e. $y \notin R_x$. So by induction $y < x$ and therefore $x = y + k$ for some $k \in \mathbb{N}$. Since $k \in \mathbb{N}$ we know there exists $k' \in \mathbb{N}_0$ such that $k = k'$ and it follows that $x = y + 1 + k'$, i.e. $y + 1 \leq x$. Since $y + 1 \notin R_x$ we may conclude that in fact $y + 1 < x$ and therefore $y + 1 \in S$. So by induction $S = \mathbb{N}_0$ and we have shown if $x < y$ does not hold iff $y \leq x$.Combining this statement with the Proposition completes the proof.

We have now set up a satisfactory addition operations and ordering on $\mathbb{N}_0$. Our next goal is to define multiplication on $\mathbb{N}_0$.

**Theorem 1.27.** There exists a function $M : \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}_0$ such that $M(x, 0) = 0$ for all $x \in \mathbb{N}_0$ and $M(x, y + 1) = M(x, y) + x$ for all $x, y \in \mathbb{N}_0$. This function $M$ satisfies the following properties:

1. $M(x, 0) = 0 = M(0, x)$ for all $x \in \mathbb{N}_0$,
2. $M(x, 1) = M(1, x) = x$ for all $x \in \mathbb{N}_0$,
3. $M(x, y) = M(y, x)$ for all $x, y \in \mathbb{N}_0$,
4. $M(x, y + z) = M(x, y) + M(x, z)$ for all $x, y, z \in \mathbb{N}_0$.
5. $M(x, M(y, z)) = M(M(x, y), z)$ for all $x, y, z \in \mathbb{N}_0$. 

\[ \blacksquare \]
Proof. Let $S$ denote those $x \in \mathbb{N}_0$ such that there exists a function $M_x : \mathbb{N}_0 \to \mathbb{N}_0$ satisfying $M_x(0) = 0$ and $M_x(y + 1) = M_x(y) + x$ for all $y \in \mathbb{N}_0$. Taking $M_0(y) := 0$ shows $0 \in S$. Moreover if $x \in S$ we define $M_{x+1}(y) := M_x(y) + y$. Then $M_{x+1}(0) = 0$ and
\[ M_{x+1}(y + 1) = M_x(y + 1) + y + 1 = M_x(y) + x + y + 1 \]
while
\[ M_{x+1}(y) + (x + 1) = M_x(y) + y + x + 1 = M_{x+1}(y + 1) . \]
This shows that $x + 1 \in S$ and so by induction $S = \mathbb{N}_0$ and we may now define $M(x,y) := M_x(y)$ for all $x,y \in \mathbb{N}_0$. We now prove the properties of $M$ stated above.

1. By construction $M(x,0) = 0$ for all $x$. Let $S := \{ x \in \mathbb{N}_0 : M(0,x) = 0 \}$. Then $0 \in S$ and if $x \in S$ we have
\[ M(0,x+1) = M(0,x) + 0 = 0 + 0 = 0 \]
which shows $x + 1 \in S$. Therefore by induction $S = \mathbb{N}_0$ and $M(0,x) = 0$ for all $x \in \mathbb{N}_0$.

2. Let $S := \{ x \in \mathbb{N}_0 : M(1,x) = x \}$. Then $0 \in S$ and if $x \in S$ we have
\[ M(1,x+1) = M(1,x) + 1 = x + 1 \]
which shows $x + 1 \in S$. Therefore $S = \mathbb{N}_0$ and $M(1,x) = x$ for all $x \in \mathbb{N}_0$.

3. Let $S := \{ x \in \mathbb{N}_0 : M(x,\cdot) = M(\cdot,\cdot) \}$. Then by items 1. and 2. we know that $0,1 \in S$. Now suppose that $x \in S$, then by construction,
\[ M(x+1,y) = M_{x+1}(y) = M(x,y) + y \]
while
\[ M(y,x+1) = M(y,x) + y. \]
The last two displayed equations along with the induction hypothesis shows $x + 1 \in S$ and therefore $S = \mathbb{N}_0$ and item 3. is proved.

4. Let $S$ denotes those $x \in S$ such that $M(x,y+z) = M(x,y) + M(x,z)$ for all $y,z \in \mathbb{N}_0$. Then $0,1 \in S$ and if $x \in S$ we have,
\[
M(x+1,y+z) = M(x,y+z) + y + z \\
= M(x,y) + M(x,z) + y + z \\
= M(x,y) + y + M(x,z) + z \\
= M(x+1,y) + M(x+1,z)
\]
which shows $x + 1 \in S$. Therefore $S = \mathbb{N}_0$ and we have proved item 4.

5. Let
\[ S := \{ x \in \mathbb{N}_0 : M(x,M(y,z)) = M(M(x,y),z) \quad \forall y,z \in \mathbb{N}_0 \} . \]
Then $0 \in S$ and if $x \in S$ we find,
\[ M(x+1,M(y,z)) = M(x,M(y,z)) + M(y,z) \]
while
\[ M(M(x+1,y),z) = M(M(x,y) + y,z) = M(M(x,y),z) + M(y,z) . \]
The last two equations along with the induction hypothesis shows $x + 1 \in S$ and therefore $S = \mathbb{N}_0$ and item 5. is proved.

\[ \blacksquare \]

Notation 1.28 We now write $x \cdot y$ for $M(x,y)$ and refer to the symmetric binary operator $\cdot$ as multiplication.

To summarize Theorem 1.27 we have shown multiplication satisfies for all $x,y,z \in \mathbb{N}_0$:

1. $x \cdot 0 = 0 \cdot x$,
2. $x \cdot 1 = x = 1 \cdot x$,
3. $x \cdot y = y \cdot x$,
4. $x \cdot (y+z) = x \cdot y + x \cdot z$,
5. $(x \cdot y) \cdot z = x \cdot (y \cdot z)$.

Proposition 1.29 (Multiplicative Cancellation). If $x,y \in \mathbb{N}_0$ and $z \in \mathbb{N}$ such that $x \cdot z = y \cdot z$, then $x = y$.

Proof. If $x \neq y$, say $x < y$, then $y = x + n$ for some $n \in \mathbb{N}$ and therefore
\[ y \cdot z = (x + n) \cdot z = x \cdot z + n \cdot z. \]
Hence if $x \cdot z = y \cdot z$, then by additive cancellation we must have $n \cdot z = 0$. As $n,x \in \mathbb{N}$ we may write $n = n' + 1$ and $z = z' + 1$ with $n',z' \in \mathbb{N}_0$ and therefore,
\[ 0 = n \cdot z = (n' + 1) \cdot (z' + 1) = n' \cdot z' + n' + z' + 1 \neq 0 \]
which is a contradiction. \[ \blacksquare \]

Remark 1.30 (Base 10 counting). The typical method of counting is to use base 10 enumeration of $\mathbb{N}_0$. The rules are;
\[
0 := 0, \quad 1 := 1, \quad 2 := 1 + 1, \quad 3 := 2 + 1, \quad 4 := 3 + 1, \quad 5 := 4 + 1 \quad 6 := 5 + 1, \quad 7 := 6 + 1, \quad 8 := 7 + 1, \quad 9 := 8 + 1, \quad 10 := 9 + 1.
\]
Once these elements of $\mathbb{N}_0$ have been defined, then given $a_0, \ldots, a_n \in \{0, 1, \ldots, 9\}$ with $a_n \neq 0$, we let

$$a_n a_{n-1} \ldots a_0 := \sum_{k=0}^{n} a_k 10^k.$$  

For example, $35 = 3 \cdot 10 + 5 = 34 + 1$, etc.

As mentioned above one can formalize $\mathbb{Z}$ and $\mathbb{Q}$ using $\mathbb{N}_0$ constructed above. I will omit the details here and refer the reader to the references already mentioned.
Fields

The basic question we want to eventually address is: What are the real numbers? Our answer is going to be; the real numbers is the essentially unique complete ordered field, see Theorem 3.3 below. In order to make sense of this answer we need to explain the terms, “complete,” “ordered,” and “field.” We will start with the notion of a field which loosely stated means something that can reasonably be interpreted a “numbers.”

Definition 2.1 (Fields, i.e.“ numbers”). A field is a non-empty set $\mathbb{F}$ equipped with two operations called addition and multiplication, and denoted by $+$ and $\cdot$, respectively, such that the following axioms hold;

1. **Closure** of $\mathbb{F}$ under addition and multiplication. For all $a, b \in \mathbb{F}$, both $a + b$ and $a \cdot b$ are in $\mathbb{F}$ (or more formally, $+$ and $\cdot$ are binary operations on $\mathbb{F}$).
2. **Associativity of addition and multiplication.** For all $a, b,$ and $c$ in $\mathbb{F}$, the following equalities hold: $a + (b + c) = (a + b) + c$ and $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.
3. **Commutativity of addition and multiplication.** For all $a$ and $b$ in $\mathbb{F}$, the following equalities hold: $a + b = b + a$ and $a \cdot b = b \cdot a$.
4. **Additive and multiplicative identity.** There exists an element of $\mathbb{F}$, called the additive identity element and denoted by $0 = 0_\mathbb{F}$, such that for all $a$ in $\mathbb{F}$, $a + 0 = a$. Likewise, there is an element, called the multiplicative identity element and denoted by $1 = 1_\mathbb{F}$, such that for all $a$ in $\mathbb{F}$, $a \cdot 1 = a$.
   It is assumed that $0_\mathbb{F} \neq 1_\mathbb{F}$.
5. **Additive and multiplicative inverses.** For every $a$ in $\mathbb{F}$, there exists an element $-a$ in $\mathbb{F}$, such that $a + (-a) = 0$. Similarly, for any $a$ in $\mathbb{F}$ other than $0$, there exists an element $a^{-1}$ in $\mathbb{F}$, such that $a \cdot a^{-1} = 1$. (The elements $a + (-b)$ and $a \cdot b^{-1}$ are also denoted $a - b$ and $a/b$, respectively.) In other words, subtraction and division operations exist.
6. **Distributivity of multiplication over addition.** For all $a, b$ and $c$ in $\mathbb{F}$, the following equality holds: $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$.

(Note that all but the last axiom are exactly the axioms for a commutative group, while the last axiom is a compatibility condition between the two operations.)

2.1 Basic Properties of Fields

Here are some sample properties about fields. For more information about Fields see 5-8 of Rudin.

**Lemma 2.2.** Let $\mathbb{F}$ be a field, then;

1. There is only one additive and multiplicative inverses.
2. If $x, y, z \in \mathbb{F}$ with $x \neq 0$ and $xy = xz$ then $y = z$.
3. $0 \cdot x = 0$ for all $x \in \mathbb{F}$.
4. If $x, y \in \mathbb{F}$ such that $xy = 0$ then $x = 0$ or $y = 0$.
5. $(−x) y = −(xy)$.
6. $−(−x) = x$ for all $x \in \mathbb{F}$.
7. $(−x)(−y) = xy$ or all $x, y \in \mathbb{F}$.

**Proof.** We take each item in turn.

1. Suppose that $x + y = 0 = x + y′$, then adding $−x$ to both sides of this equation shows $y = y′$. Taking $y = −x$ then shows $y = −x = y′$, i.e. additive inverses are unique. Similarly if $x \neq 0$ and $xy = 1$ then multiplying this equation by $x^{-1}$ shows $y = x^{-1}$ and so there is only one multiplicative inverse.
2. If $xy = xz$ then multiplying this equation by $x^{-1}$ shows $y = z$.
3. $0 \cdot x + z = 0 \cdot x + 1 \cdot x = (0 + 1) \cdot x = 1 \cdot x = x$.

Adding $−x$ to both side of this equation using associativity and commutativity of addition then implies $0 \cdot x = 0$.
4. If $x \in \mathbb{F} \setminus \{0\}$ and $y \in \mathbb{F}$ such that $xy = 0$, then

$$0 = x^{-1} \cdot 0 = x^{-1}(xy) = (x^{-1}x)y = 1y = y.$$

5. $(−x)y + xy = (−x + x)y = 0 \cdot y = 0 \implies (−x)y = −(xy)$.
6. Since $(−x) + x = 0$ we have $−(−x) = x$.
7. $(−x)(−y) = −(x \cdot (−y)) = −(−(xy)) = xy$ by 6.

**Example 2.3.** Here are a few examples of Fields:
1. \(F_2 = \{0, 1\}\) with \(0 + 0 = 0 = 1 + 1,\) and \(0 + 1 = 1 + 0 = 0 \cdot 1 = 1 \cdot 0 = 0 \cdot 0 = 0\) and \(1 - 1 = 1\). In this case \(-1 = 1, 1^{-1} = 1\) and \(-0 = 0\).

2. \(\mathbb{Q}\) - the rational numbers with the usual addition and multiplication of fractions. \((\frac{m}{n})^{-1} = \frac{1}{n}\) if \(m \neq 0\) and \(-\frac{m}{n} = \frac{-m}{n}\).

3. \(F = \mathbb{Q}(t)\) where
\[
\mathbb{Q}(t) = \left\{ \frac{p(t)}{q(t)} : p(t) and q(t) are polynomials over \mathbb{Q} \ni q(t) \neq 0 \right\}.
\]

Again the multiplication and addition are as usual.

**Example 2.4.** \(\mathbb{Z}\) is not a field. For example, 2 has no multiplicative inverse in \(\mathbb{Z}\).

The inverse to 2, \(2^{-1}\), should be \(\frac{1}{2}\) but this is not in \(\mathbb{Z}\).

**Definition 2.5.** We say a map \(\varphi : \mathbb{Z} \rightarrow F\) is a (ring) homomorphism iff \(\varphi(1) = 1_F, \varphi(0) = 0_F,\) and for all \(x, y \in \mathbb{Z};\)
\[
\varphi(x + y) = \varphi(x) + \varphi(y) \text{ and } \varphi(xy) = \varphi(x) \varphi(y).
\]

[The assumption that \(\varphi(0) = 0_F\) is redundant since \(\varphi(0) = \varphi(0 + 0) = \varphi(0) + \varphi(0)\) and therefore \(\varphi(0) = 0_F.\)]

**Lemma 2.6.** For every field \(F\) there a unique (ring) homomorphism, \(\varphi : \mathbb{Z} \rightarrow F.\) In fact, \(\varphi(n) = n1_F\) for all \(n \in \mathbb{Z}\) where \(0 \cdot 1_F = 0_F.\)
\[
\text{n times}
\]
\[
n1_F := \overbrace{1 + \cdots + 1}^{\text{n times}} \text{ if } n \in \mathbb{N} \text{ and}
\]
\[
(n)1_F := -(n1_F) \text{ if } n \in \mathbb{N}.
\]

[The map \(\varphi\) need not be injective as is seen by taking \(F = F_2.\)]

**Proof.** Let us first work on \(\mathbb{N}_0 \subset \mathbb{Z}.\) We must define \(\varphi(0) = 0\) and \(\varphi(1) = 1\) and then \(\varphi\) inductively by \(\varphi(n + 1) = \varphi(n) + \varphi(1) = \varphi(n) + 1_F\) so that
\[
\text{n times}
\]
\[
\varphi(n) = \overbrace{1 + \cdots + 1}^{\text{n times}}.
\]

We now write \(n1_F\) for \(\varphi(n)\) with the convention that \(01_F = 0_F.\) For \(n \in \mathbb{N}\) we must set \(\varphi(-n) = -\varphi(n) = -(n1_F).\) Thus we have \(\varphi(n) = n1_F\) for all \(n \in \mathbb{Z}.\)

We now must show \(\varphi\) is a homomorphism.

**Additive homomorphism:** First suppose that \(m, n \in \mathbb{N}_0\) and let
\[
S := \{m \in \mathbb{N}_0 : \varphi(mn) = \varphi(m) \varphi(n) \text{ for all } n \in \mathbb{N}_0\}.
\]

One easily sees that \(0 \in S\) and that \(1 \in S\) by construction. Moreover if \(m \in S,\) then
\[
\varphi((m + 1) + n) = \varphi(m + n) = \varphi(n) + \varphi(1) + \varphi(n)
\]
\[
= \varphi(m) \varphi(n) + \varphi(n) = \varphi(m + 1) \varphi(n)
\]
\[
= \varphi(m + 1) \varphi(n) = \varphi(m + 1 + \varphi(n) = \varphi(m + 1) + \varphi(n)
\]
which shows \(m + 1 \in S.\) Therefore by induction, \(S = \mathbb{N}_0\) and \(\varphi(m + n) = \varphi(m) + \varphi(n)\) for all \(m, n \in \mathbb{N}_0.\)

If \(m \in \mathbb{N}_0\) we have \(\varphi(-m) = -\varphi(m)\) by construction. If \(n > m \in \mathbb{N}_0,\) then
\[
\varphi(n + (m)) = \varphi(n) + (-\varphi(m)) = \varphi(n) + \varphi(-m).
\]

If \(n < m \in \mathbb{N}_0,\) then
\[
\varphi(n + (m)) = \varphi(n) - \varphi(m) = \varphi(n) + \varphi(-m).
\]

Putting all of this together shows \(\varphi\) is an additive homomorphism.

**Multiplicative homomorphism:** First suppose that \(m, n \in \mathbb{N}_0\) and let
\[
S := \{m \in \mathbb{N}_0 : \varphi(mn) = \varphi(m) \varphi(n) \text{ for all } n \in \mathbb{N}_0\}.
\]

It is easily seen that \(0, 1 \in S.\) Moreover if \(m \in S\) and \(n \in \mathbb{N}_0,\) then
\[
\varphi((m + 1) \varphi(n) = \varphi(mn + n) = \varphi(mn) + \varphi(n)
\]
\[
= \varphi(m) \varphi(n) + \varphi(n) = (\varphi(m) + 1_F) \varphi(n)
\]
\[
= \varphi(m + 1) \varphi(n),
\]
which shows \(m + 1 \in S.\) Therefore by induction, \(S = \mathbb{N}_0\) and \(\varphi(mn) = \varphi(m) \varphi(n)\) for all \(m, n \in \mathbb{N}_0.\)

If \(m, n \in \mathbb{N}_0,\) then
\[
\varphi((-m) n) = \varphi(-mn) = -\varphi(mn) = -[\varphi(m) \varphi(n)] = -\varphi(m) \varphi(n) = \varphi(-m) \varphi(n)
\]
and
\[
\varphi((-m) (-n)) = \varphi(mn) = \varphi(mn) = (\varphi(m) \varphi(n) = \varphi(m) \varphi(n)
\]
which completes the verification that \(\varphi\) is a multiplicative homomorphism. ■
2.2 Ordered Fields

Definition 2.7 (Ordered Field). We say $\mathbb{F}$ is an ordered field if there exists, $P \subset \mathbb{F}$, called the positive elements, such that

Ord 1. $\mathbb{F}$ is the disjoint union of $P$, $\{0\}$, and $-P$, i.e. if $x \in \mathbb{F}$ then precisely one of following happens; $x \in P$, $x = 0$, or $-x \in P$.

Ord 2. $P + P \subset P$ and $P \cdot P \subset P$.

Lemma 2.8. Let $(\mathbb{F}, P)$ be an ordered field, then;

1. For all $x \in \mathbb{F} \setminus \{0\}$, $x^2 \in P$. In particular $1 = 1^2 \in P$.
2. If $x \in P$ and $y \in -P$ then $xy \in -P$.
3. If $x \in P$ then $x^{-1} \in P$.

Proof. If $x \in P$ then $x^2 \in P \cdot P \subset P$ while if $x \in -P$ then $-x \in P$ and $x^2 = (-x)^2 \in P$. For item 3. we have $x \cdot x^{-1} = 1$.

Example 2.9. The field $\mathbb{F} = \{0, 1\}$ is not ordered. The only possible choice for $P$ is $P = \{1\}$ which does not work since $1 + 1 = 0 \notin P$.

Example 2.10. Take $\mathbb{F} = \mathbb{Q}$ and $P = \left\{ \frac{m}{n} : m, n > 0 \right\}$. This is in fact the unique choice we can make for $P$ in this case. Indeed suppose that $P$ is any order on $\mathbb{Q}$. By Lemma 2.8 we know $1 \in P$ and then by induction it follows that $\mathbb{N} \subset P$. Then again by Lemma 2.8 we must have $m \cdot n^{-1} \in P$ for all $m, n \in \mathbb{Q}$.

Example 2.11. Take $\mathbb{F} = \mathbb{Q}(t)$ and

$$P = \left\{ \frac{p(t)}{q(t)} \in \mathbb{F} : \frac{p(t)}{q(t)} > 0 \text{ for } t > 0 \text{ large} \right\},$$

i.e. $\frac{p(t)}{q(t)} \in P$ iff the highest order coefficients of $p(t)$ and $q(t)$ have the same sign. For example $\frac{t^2 - 25t + 7}{t^4 - 100t^2} \in P$ while $\frac{t^2 + 25t + 7}{t^4 - 100t^2} \in -P$.

Notice that $t > n$ for all $n \in \mathbb{N}$ and $\frac{1}{2} \leq \frac{1}{n}$ for all $n \in \mathbb{N}$. This is kind of strange and explains why you have to prove the “obvious” in this course!!

Moral: obvious statements are often false.

Notation 2.12 (Max and Min) We will often use the following notation in the sequel. If $a, b$ are elements of an ordered field, let

$$a \wedge b := \min (a, b) = \begin{cases} a & \text{if } a \leq b \\ b & \text{if } b \leq a \end{cases}$$

and

$$a \lor b := \max (a, b) = \begin{cases} b & \text{if } a \leq b \\ a & \text{if } b \leq a \end{cases}.$$

More generally if $\{a_i\}_{i=1}^n \subset \mathbb{F}$ we let

$$a_1 \wedge \cdots \wedge a_n := \min (a_1, \ldots, a_n)$$

and

$$a_1 \lor \cdots \lor a_n := \max (a_1, \ldots, a_n)$$

be the smallest and largest element in the finite list $(a_1, \ldots, a_n)$.

Definition 2.13. Suppose that $\mathbb{F}$ and $\mathbb{G}$ are fields. A map, $\varphi : \mathbb{F} \rightarrow \mathbb{G}$ is a (field) homomorphism iff $\varphi (1_{\mathbb{F}}) = 1_{\mathbb{G}}$, $\varphi (0_{\mathbb{F}}) = 0_{\mathbb{G}}$, and for all $x, y \in \mathbb{F}$;

$$\varphi (x + y) = \varphi (x) + \varphi (y) \text{ and } \varphi (xy) = \varphi (x) \varphi (y).$$

Lemma 2.14 ($\mathbb{Q}$ embeds into an ordered field). For every ordered field $(\mathbb{F}, P)$, there a unique field homomorphism, $\varphi : \mathbb{Q} \rightarrow \mathbb{F}$. In fact,

$$\varphi \left( \frac{m}{n} \right) = \frac{m}{n} \cdot 1_{\mathbb{F}} := m_{1_{\mathbb{F}}} \cdot (n_{1_{\mathbb{F}}})^{-1} \quad (2.1)$$

times $n$

where $n_{1_{\mathbb{F}} := 1_{\mathbb{F}} + \cdots + 1_{\mathbb{F}}}$ and $(-n)_{1_{\mathbb{F}}} := -(n_{1_{\mathbb{F}}})$ for $n \in \mathbb{N}$ and $0 \cdot 1_{\mathbb{F}} = 0_{\mathbb{F}}$.

Moreover;

1. $\varphi (x) \in P$ whenever $x > 0$,
2. and $\varphi$ is injective. Thus we may identify $\mathbb{Q}$ with $\varphi (\mathbb{Q})$ and consider $\mathbb{Q}$ as a sub-field of $\mathbb{F}$.

[In particular, ordered fields must be fields with an infinite number of elements in it.]

Proof. From Lemma 2.6 we know there is a unique ring homomorphism, $\varphi : \mathbb{Z} \rightarrow \mathbb{F}$, given by $\varphi (m) = m \cdot 1_{\mathbb{F}}$. So for $m \in \mathbb{Z}$ and $n \in \mathbb{N}$ we must have

$$\varphi \left( \frac{m}{n} \right) \cdot n_{1_{\mathbb{F}}} = \varphi \left( \frac{m}{n} \right) \cdot \varphi (n) = \varphi \left( \frac{m}{n} \cdot n \right) = \varphi (m) = m_{1_{\mathbb{F}}}$$

which forces us to define $\varphi$ as in Eq. (2.1). Notice that is easy to verify by induction that $n_{1_{\mathbb{F}}} = \varphi (n) \in P$ for all $n \in \mathbb{N}$ and in particular $n_{1_{\mathbb{F}}} \neq 0$ for $n \in \mathbb{N}$. In particular if $x = m/n > 0$ then $\varphi \left( \frac{m}{n} \right) = m_{1_{\mathbb{F}}} \cdot (n_{1_{\mathbb{F}}})^{-1} \in P$ by Lemma 2.8. We must still check that $\varphi$ is well defined homomorphism.

Well defined. Suppose that $k \in \mathbb{N}$, we must show

$$(km)_{1_{\mathbb{F}}} \cdot ((kn)_{1_{\mathbb{F}}})^{-1} = m_{1_{\mathbb{F}}} \cdot (n_{1_{\mathbb{F}}})^{-1}.$$ 

By cross multiplying, this will happen iff

$$(km)_{1_{\mathbb{F}}} \cdot (n_{1_{\mathbb{F}}}) = ((kn)_{1_{\mathbb{F}}} \cdot m_{1_{\mathbb{F}}}.$$
which is the case as \( \varphi : \mathbb{Z} \to \mathbb{F} \) is a ring homomorphism.

**Homomorphism property.** We have

\[
\varphi \left( \frac{m}{n} + \frac{p}{n} \right) = \varphi \left( \frac{m + p}{n} \right) = \varphi (m + p) \cdot \varphi (n)^{-1} \\
= [\varphi (m) + \varphi (p)] \cdot \varphi (n)^{-1} \\
= \varphi (m) \cdot \varphi (n)^{-1} + \varphi (p) \cdot \varphi (n)^{-1} \\
= \varphi \left( \frac{m}{n} \right) + \varphi \left( \frac{p}{n} \right)
\]

and

\[
\varphi \left( \frac{m}{n} \right) \varphi \left( \frac{q}{p} \right) = \varphi (m) \varphi (n)^{-1} \varphi (q) \varphi (p)^{-1} \\
= \varphi (m) \varphi (q) [\varphi (n) \varphi (p)]^{-1} \\
= \varphi (mq) [\varphi (np)]^{-1} = \varphi \left( \frac{mq}{np} \right).
\]

**Injectivity.** If \( 0 = \varphi \left( \frac{m}{n} \right) \) then

\[
0 = \varphi (m) \cdot \varphi (n)^{-1}
\]

which implies \( \varphi (m) = 0 \) which happens iff \( m = 0 \), i.e. \( m/n = 0 \).

---

**Exercise 2.1.** Let \((\mathbb{F}, P)\) be an ordered field and \( x, y \in \mathbb{F} \) with \( y > x \). Show:

1. \( y + a > x + a \) for all \( a \in \mathbb{F} \),
2. \( -x > -y \),
3. if we further suppose \( x > 0 \), show \( \frac{1}{x} > \frac{1}{y} \).

**Definition 2.17.** Given \( x \in \mathbb{F} \), we say that \( y \in \mathbb{F} \) is a square root of \( x \) if \( y^2 = x \). [From Lemma 2.8, it follows that if \( x \in \mathbb{F} \) has a square root then \( x \geq 0 \).]

**Lemma 2.18.** Suppose \( x, y \in \mathbb{F} \) with \( x^2 = y^2 \), then either \( x = y \) or \( x = -y \). In particular, there are at most 2 square roots of any number \( x \geq 0 \) in \( \mathbb{F} \).

**Proof.** Observe that

\[
(x - y)(x + y) = (x - y)x + (x - y)y = x^2 - xy + xy - y^2 = x^2 - y^2 = 0.
\]

Thus it follows that either \( x - y = 0 \) or \( x + y = 0 \), i.e. \( x = y \) or \( x = -y \).

**Definition 2.19.** If \( x > 0 \) admits a square root we let \( \sqrt{x} \) be the unique positive root. We also define \( \sqrt{0} = 0 \).

**Lemma 2.20.** Suppose that \( 0 < x < y \), i.e. \( x, y \in P \), then \( x^2 < y^2 \).

**Proof.** By Lemma 2.16 we know \( x \cdot y < x \cdot y \) and \( x \cdot y < y \cdot y \) and therefore \( x^2 < y^2 \).

**Corollary 2.21.** If \( 0 \leq x < y \) and \( \sqrt{x} \) and \( \sqrt{y} \) exists, then \( 0 \leq \sqrt{x} < \sqrt{y} \).

**Proof.** If \( \sqrt{x} = \sqrt{y} \) then \( x = (\sqrt{x})^2 = (\sqrt{y})^2 = y \) which is impossible. Similarly if \( \sqrt{x} > \sqrt{y} \) then

\[
x = (\sqrt{x})^2 > (\sqrt{y})^2 = y
\]

which is again false.

Alternatively: starting with \( y^2 - x^2 = (y - x)(y + x) \) and then replacing \( y \) and \( x \) by \( \sqrt{y} \) and \( \sqrt{x} \) respectively (assuming they exist) shows,

\[
y - x = (\sqrt{y} - \sqrt{x}) (\sqrt{y} + \sqrt{x}) \implies \sqrt{y} - \sqrt{x} = (y - x)(\sqrt{y} + \sqrt{x})^{-1}
\]

from which it follows that \( \sqrt{y} - \sqrt{x} \in \mathbb{P} \) if \( (y - x) \in \mathbb{P} \). More importantly this shows \( \sqrt{y} \) depends “continuously” in on \( y \).

**Definition 2.22.** The absolute value, \( |x| \), of \( x \) in ordered field \( \mathbb{F} \) is defined by

\[
|x| = \begin{cases} 
  x & \text{if } x \geq 0 \\
  -x & \text{if } x < 0
\end{cases}
\]

Alternatively we may define

\[
|x| = \sqrt{x^2}.
\]
Proposition 2.23. For all \( x, y \in \mathbb{F} \), then
\begin{enumerate}
\item \( |x| \geq 0 \)
\item \( |xy| = |x| |y| \)
\item \( |x + y| \leq |x| + |y| \).
\end{enumerate}

Proof. 1. holds by definition since \(-x > 0\) if \( x < 0 \).
2. As \( |x| |y| \geq 0 \) and \( (|x| |y|)^2 = |x|^2 |y|^2 = x^2 y^2 = (xy)^2 \), we have
\[ |x| |y| = \sqrt{(xy)^2} = |xy|. \]
3. It suffices to show \( x + y \leq (|x| + |y|)^2 \). However,
\[ x + y \geq (x + y)^2 = x^2 + y^2 + 2xy \]
\[ \leq x^2 + y^2 + 2|x|y \quad (x \leq |xy|) \]
\[ = |x|^2 + |y|^2 + 2|x| |y| \]
\[ = (|x| + |y|)^2. \]

Definition 2.24. Let \((\mathbb{F}, P)\) be an ordered field and \( S \) be a subset of \( \mathbb{F} \).
\begin{enumerate}
\item We say that \( S \subseteq \mathbb{F} \) is bounded from above (below) if there exists \( x \in \mathbb{F} \) such that \( x \geq s \) (\( x \leq s \)) for all \( s \in S \). Any such \( x \) is called an upper (lower) bound of \( S \).
\item If \( S \) is bounded from above (below), we say that \( y \in \mathbb{F} \) is a least upper bound (greatest lower bound) for \( S \) if \( y \) is an upper (lower) bound for \( S \) and \( y \leq x \) (\( y \geq x \)) for any other upper (lower) bound, \( x \), of \( S \).
\end{enumerate}

Notice that least upper bounds and greatest lower bounds are unique if they exist. We will write and
\[ y = \text{l.u.b.} \,(S) = \sup \,(S) \]
if \( y \) is the least upper bound for \( S \) and
\[ y = \text{g.l.b.} \,(S) = \inf \,(S) \]
if \( y \) is the greatest lower bound for \( S \).

Example 2.25. Let \( \mathbb{F} = \mathbb{Q} \), then;
1. \( \max \,(a, b) \) and \( \min \,(a, b) \) are least upper respectively greatest lower bounds respectively for \( S = \{a, b\} \). More generally, if \( S = \{a_1, \ldots, a_n\} \), then
\[ \sup \,(S) = a_1 \vee \cdots \vee a_n := \max \,(a_1, \ldots, a_n) \quad \text{and} \]
\[ \inf \,(S) = a_1 \wedge \cdots \wedge a_n := \min \,(a_1, \ldots, a_n). \]
2. \( S = \mathbb{N} \) is not bounded from above while \( \inf \,(S) = 1 \).
3. \( S = \{1 - \frac{1}{n} : n \in \mathbb{N}\} \) is bounded from above and \( 1 = \sup \,(S) \) while \( \inf \,(S) = 1/2 \).
4. Let \( S = \{1, 1.4, 1.41, 1.414, 1.4142, 1.41421, 1.414213, 1.4142135, \ldots\} \)
where I am getting these digits from the decimal expansion of \( \sqrt{2} \);\n\[ \sqrt{2} \approx 1.41421356237309504880168872420968907856967187537694807317667973799. \]

In this case \( S \) is bounded above by 2, or 1.42, or 1.415, etc. Nevertheless \( \sqrt{2} = \sup \,(S) \) does not exists in \( \mathbb{Q} \).

Example 2.26. Now let \( \mathbb{F} = \mathbb{Q} \,(t) \) be the field of rational functions described in Example 2.11 then; \( S = \mathbb{N} \) is bounded from above. For example \( t \) is an upper bound but there is not least upper bound. For example \( \frac{1}{n} t \) is also an upper bound for \( S \).

End of Lecture 4, 10/5/2012

Definition 2.27 (Dedekind Cuts). A subset \( \alpha \subset \mathbb{Q} \) is called a cut (see [2, p. 17]) if;
\begin{enumerate}
\item \( \alpha \) is a proper subset of \( \mathbb{Q} \), i.e. \( \alpha \neq \emptyset \) and \( \alpha \neq \mathbb{Q} \),
\item if \( p \in \alpha \) and \( q \in \mathbb{Q} \) and \( q < p \), then \( q \in \alpha \),
\item if \( p \in \alpha \), then there exists \( r \in \alpha \) with \( r > p \).
\end{enumerate}

Example 2.28. To each \( a \in \mathbb{Q} \), let \( \alpha_a := \{ q \in \mathbb{Q} : q < a \} \). Then \( \alpha_a \) is a cut and \( a \) is the least upper bound of \( \alpha_a \) in \( \mathbb{Q} \).

Example 2.29. Let \( \{S_n\}_{n=0}^{\infty} \subset \mathbb{Q} \) be any bounded sequence such that \( S_n \leq S_{n+1} \) for all \( n \). Then \( \alpha := \bigcup_{n=0}^{\infty} \alpha_{S_n} = \{ q \in \mathbb{Q} : q < S_n \text{ a.a. } n \} \) is a cut as the reader should verify. Let us further suppose that \( \lim_{n \to \infty} S_n \) does not exist in \( \mathbb{Q} \). [For example from Example 1.17 we may take \( S_n := \sum_{k=0}^{\infty} \frac{1}{k^2} \in \mathbb{Q} \).] If \( m \in \mathbb{Q} \) is an upper bound for \( \alpha \), then \( m \geq S_n \) for all \( n \) since if \( m < S_n \) for some \( n \) then \( q := \frac{1}{2} (m + S_n) \in \alpha \) with \( q > m \). Since \( \lim_{n \to \infty} S_n \neq m \) as \( m \in \mathbb{Q} \) there must exists \( \varepsilon > 0 \) such that
\[ m - S_n = |m - S_n| \geq \varepsilon \text{ i.o. } n. \]
As \( m - S_n \) is decreasing we may conclude that \( m - S_n \geq \varepsilon \) for all \( n \), i.e. \( S_n \leq m - \varepsilon \) for all \( n \). From this it now follows that \( m - \varepsilon \) is an upper bound for \( \alpha \) which is strictly smaller that \( m \). So there can be no least upper bound.
Real Numbers

As we saw in Section 1.2, \( \mathbb{Q} \) is full of holes and calculus tends to produce
answers which live in these holes. So it is imperative that we fill the holes. Doing
so will lead to the real numbers provided we fill in the holes without adding too
much extra filler along the way. One good answer to the question, What are
the real numbers?, is contained in the statement of Theorem 3.3.

**Definition 3.1.** An order preserving field isomorphism between two ordered
fields, \((\mathbb{F}_1, P_1)\) and \((\mathbb{F}_2, P_2)\), is a bijection, \( f : \mathbb{F}_1 \to \mathbb{F}_2 \) such that
\( f(0) = 0, f(1) = 1, f(P_1) = P_2, \) and
\[
f(x + y) = f(x) + f(y) \quad \text{and} \quad f(xy) = f(x)f(y) \quad \text{for all } x, y \in \mathbb{F}_1.
\]

**Definition 3.2.** An ordered field \((\mathbb{F}, P)\) is has the least upper bound property (or is
complete) if every non-empty subset, \( S \subseteq \mathbb{F} \), which is bounded from
above possesses a least upper bound in \( \mathbb{F} \). [As we have seen in examples above, \( \mathbb{Q} \) does not have the least upper bound property.]

**Theorem 3.3 (The real numbers).** Up to order preserving field isomorphism
(see Definition 3.7), there is exactly one complete ordered field. It is this field
that we refer to as the **real numbers** and denote by \( \mathbb{R} \).

**Definition 3.4.** We say two Cauchy sequences \( \{a_n\}_{n=1}^{\infty} \) and \( \{b_n\}_{n=1}^{\infty} \) of rational
numbers are equivalent and write \( \{a_n\}_{n=1}^{\infty} \sim \{b_n\}_{n=1}^{\infty} \) iff
\[
\lim_{n \to \infty} |a_n - b_n| = 0.
\]

We then define \( \alpha := \{[a_n]_{n=1}^{\infty}\} \) to be the equivalence class of the Cauchy sequence
\( \{a_n\}_{n=1}^{\infty} \) and refer to the collection of these equivalence classes as the **real numbers**.
The set of real numbers will be denoted by \( \mathbb{R} \).

**Notation 3.5** Let \( i : \mathbb{Q} \to \mathbb{R} \) be defined by \( i(a) := \{(a, a, a, \ldots)\} \), i.e. \( i(a) \) is
the equivalence class of the constant sequence \( a \).

Notice that if \( i(a) = i(b) \) iff \( a = \lim_{n \to \infty} a = \lim_{n \to \infty} b = b \). Thus the map,
\( i : \mathbb{Q} \to \mathbb{R} \) is injective and we will often simply identify \( a \) with \( i(a) \) and in this
way consider \( \mathbb{Q} \) as a subset of \( \mathbb{R} \).

**Theorem 3.6.** Let \( \mathbb{R} \) be as in Theorem 3.3. For \( \alpha := \{[a_n]_{n=1}^{\infty}\} \) and \( \beta := \{[b_n]_{n=1}^{\infty}\} \in \mathbb{R} \) we define
\[
\alpha + \beta = \{[a_n + b_n]_{n=1}^{\infty}\} \quad \text{and} \quad \alpha \cdot \beta = \{[a_n \cdot b_n]_{n=1}^{\infty}\}.
\]

1. With these definitions, \( \mathbb{R} \), satisfies the axioms of a field.
2. Moreover, we can make this into an ordered field by setting \( P := \{\alpha \in \mathbb{R} : \alpha > 0\} \) where we say \( \alpha > 0 \) iff there exists an \( N \in \mathbb{N} \) such that \( a_n > \frac{1}{N} \) for a.a. \( n \).\footnote{Roughly speaking here, you should think of \( \alpha = \lim_{n \to \infty} a_n \) and so \( \alpha > 0 \) should happen iff \( \alpha > \frac{1}{N} \) for some \( N \in \mathbb{N} \) which then implies \( a_n \geq \frac{1}{N} \) for a.a. \( n \).}
3. The ordered field \((\mathbb{R}, P)\) is complete, i.e. has the least upper bound property.

The proof of Theorem 3.6 and Theorem 3.3 will be relegated to Section
3.6 at the end of this chapter. For an alternative existence proof of \( \mathbb{R} \) using
Dedekind cuts as the elements of \( \mathbb{R} \) is covered in Rudin [2] pages 17-21.]. One
may also construct the Real numbers using decimal expansions, see [T. Gower’s
notes on real numbers as decimals]. We will prove the uniqueness assertion of
Theorem 3.3 in Section at the end of this section. From now on we are going
to take Theorem 3.3 for granted and derive from this the “familiar” properties
of the real numbers.

Observe that \( \mathbb{Q}, \mathbb{Q}(t), \mathbb{R}(t) \) are not complete and hence are not the real
numbers, \( \mathbb{R} \). For example \( \mathbb{N} \subset \mathbb{Q}(t) \) (or \( \mathbb{N} \subset \mathbb{R}(t) \)) is bounded by \( t \) say but
has no least upper bound. However, we do know that \( \mathbb{Q} \subset \mathbb{R} \) by Lemma 2.14.
We will soon see that \( \mathbb{Q} \) is “dense” in \( \mathbb{R} \). We now pause to discuss some of the
basic properties of \( \mathbb{R} \).

**Theorem 3.7.** Suppose that \( \mathbb{R} \) is a complete ordered field which we assume we have
already embedded \( \mathbb{Q} \) into \( \mathbb{R} \) as in Lemma 2.14. Then;

1. For all \( x \geq 0 \) there exists \( n \in \mathbb{N} \) such that \( n \geq x \).
2. For all \( \varepsilon > 0 \) in \( \mathbb{R} \) there exists \( n \in \mathbb{N} \) such that \( 0 < \frac{1}{n} \leq \varepsilon \).
3. If \( \varepsilon \geq 0 \) satisfies \( \varepsilon \leq 1/n \) for all \( n \in \mathbb{N} \) then \( \varepsilon = 0 \).
4. If \( a, b \in \mathbb{R} \) and \( a \leq b + \frac{1}{n} \) for all \( n \in \mathbb{N} \) or \( a \leq b + \varepsilon \) for all \( \varepsilon > 0 \), then
   \( a \leq b \).

**Proof.** We take each item in turn.

1. If \( n < x \) for all \( n \in \mathbb{N} \), then \( \mathbb{N} \) is bounded from above and so \( a := \sup(\mathbb{N}) \)
exists in \( \mathbb{R} \) by the completeness axiom. As \( a \) is the least upper bound for \( \mathbb{N} 
there must be an \( n \in \mathbb{N} \) such that \( n > a - 1 \). However this implies \( n + 1 > a \)
which violates \( a \) be an upper bound for \( \mathbb{N} \).
2. If \( \varepsilon > 0 \) in \( \mathbb{R} \) and \( \frac{1}{n} > \varepsilon \) for all \( n \in \mathbb{N} \), then \( n < \frac{1}{\varepsilon} \) for all \( n \in \mathbb{N} \) which is impossible by item 1.

3. If there exists \( \varepsilon > 0 \) such that \( \varepsilon \leq \frac{1}{n} \) for all \( n \) then \( n \leq 1/\varepsilon \) for all \( n \) which is again impossible by item 1.

4. It suffices to prove the first assertion. We may assume \( a \geq b \) for otherwise we are done. If \( a \leq b + \frac{1}{n} \) for all \( n \), then \( 0 \leq a - b \leq \frac{1}{n} \) for all \( n \in \mathbb{N} \) and hence \( a = b \) and in particular \( a \leq b \).

\[ \text{Proposition 3.8. If } \mathbb{R} \text{ is a complete ordered field, then every subset } S \subset \mathbb{R} \text{ which is bounded from below has a greatest lower bound, } \text{glb}(S) = \inf(S). \text{ In fact,} \]

\[ \inf(S) = -\sup(-S). \]

**Proof.** We let \( m := -\sup(-S) \). Then we have \(-s \leq -m \) for all \( s \in S \), i.e. \( s \geq m \) for all \( s \in S \) so that \( m \) is a lower bound for \( S \). Moreover if \( \varepsilon > 0 \) is given there exists \( s \in S \) such that \(-s \geq -m - \varepsilon \), i.e. \( s \leq m + \varepsilon \). This shows that any lower bound, \( m \) of \( S \) must satisfy, \( k \leq m + \varepsilon \) for all \( \varepsilon > 0 \), i.e. \( k \leq m \). This shows that \( m \) is the greatest lower bound for \( S \).

Let me sketch one way to construct \( R \) based on Cauchy sequences of rational numbers.

\[ \text{Definition 3.9. A sequence } \{q_n\}_{n=1}^{\infty} \subset \mathbb{Q} \text{ converges to } 0 \text{ in } \mathbb{R} \text{ if for all } \varepsilon > 0 \text{ there exists } N \in \mathbb{N} \text{ such that } |q_n| \leq \varepsilon \text{ for all } n \geq N. \text{ Alternatively put, for all } M \in \mathbb{N} \text{ we have } |q_n| \leq \frac{1}{M} \text{ for a.a. } n. \]

\[ \text{Definition 3.10. A sequence } \{q_n\}_{n=1}^{\infty} \subset \mathbb{Q} \text{ converges to } q \text{ in } \mathbb{R} \text{ if } |q - q_n| \to 0 \text{ as } n \to \infty, \text{ i.e. if for all } N \in \mathbb{N}, |q - q_n| \leq \frac{1}{N} \text{ for a.a. } n. \text{ As usual if } \{q_n\}_{n=1}^{\infty} \text{ converges to } q \text{ we will write } q_n \to q \text{ as } n \to \infty \text{ or } q = \lim_{n \to \infty} q_n. \]

\[ \text{Definition 3.11. A sequence } \{q_n\}_{n=1}^{\infty} \subset \mathbb{Q} \text{ is Cauchy if } |q_n - q_m| \to 0 \text{ as } m, n \to \infty. \text{ More precisely we require for each } \varepsilon > 0 \text{ in } \mathbb{R} \text{ that } |q_m - q_n| \leq \varepsilon \text{ for a.a. pairs } (m, n), \text{ i.e. there should exists } N \in \mathbb{N} \text{ such that } |q_m - q_n| \leq \varepsilon \text{ for all } m, n \geq N. \]

The next few results are analogous to what you have already shown in the case \( S \) is replaced by \( Q \). As the proofs are essentially identical to those of Theorem 1.13 and Exercise 1.16.

\[ \text{Proposition 3.12. If } \{a_n\}_{n=1}^{\infty} \subset \mathbb{R} \text{ is a convergent sequence then it is Cauchy.} \]

\[ \text{If } \{a_n\}_{n=1}^{\infty} \text{ is Cauchy sequence then it is bounded.} \]

\[ \text{Theorem 3.13 (Basic Limit Results). Suppose that } \{a_n\} \text{ and } \{b_n\} \text{ are sequences of real numbers such that } A := \lim_{n \to \infty} a_n \text{ and } B := \lim_{n \to \infty} b_n \text{ exists in } \mathbb{R}. \text{ Then; } \]

1. \( \lim_{n \to \infty} |a_n| = |A| \).
2. If \( A \neq 0 \) then \( \lim_{n \to \infty} \frac{1}{a_n} = \frac{1}{A} \).
3. \( \lim_{n \to \infty} (a_n + b_n) = A + B \).
4. \( \lim_{n \to \infty} (a_n b_n) = A \cdot B \).
5. If \( a_n \leq b_n \) for all \( n \), then \( A \leq B \).
6. If \( \{x_n\} \subset \mathbb{R} \) is another sequence such that \( a_n \leq x_n \leq b_n \) and \( A = B \), then \( \lim_{n \to \infty} x_n = A = B \).

\[ \text{Theorem 3.14. If } S \subset \mathbb{R} \text{ is a non-empty set which is bounded from above, then there exists } \{x_n\}_{n=1}^{\infty} \subset S \text{ such that } x_n \uparrow \sup S \text{ as } n \to \infty, \text{ i.e. } x_n \leq x_{n+1} \text{ for all } n \text{ and } \lim_{n \to \infty} x_n = \sup S. \]

**Proof.** Let \( M := \sup S \). For each \( n \in \mathbb{N} \), there exists \( y_n \in S \) such that \( M \geq y_n \geq M - \frac{1}{n} \). We now let \( x_n := \max \{\{y_1, \ldots, y_n\}\} \) in which case \( M \geq x_n \geq M - \frac{1}{n} \) and \( x_n \) is increasing. By the Sandwich theorem it follows that \( \lim_{n \to \infty} x_n = M \).

- End of Lecture 5, 10/8/2012

\[ \text{Theorem 3.15. If } \{x_n\}_{n=1}^{\infty} \subset \mathbb{R} \text{ is bounded from above and } x_n \text{ is non-decreasing, then } \lim_{n \to \infty} x_n = \sup_{n \to \infty} x_n. \text{ Similarly if } \{x_n\}_{n=1}^{\infty} \subset \mathbb{R} \text{ is bounded from below and } x_n \text{ is non-increasing, then } \lim_{n \to \infty} x_n = \inf_{n \to \infty} x_n. \]

**Proof.** Let \( M := \sup_{n \to \infty} x_n \), then for all \( \varepsilon > 0 \), there exists \( N_\varepsilon \in \mathbb{N} \) such that \( M \geq x_{N_\varepsilon} \geq M - \varepsilon \). As \( x_n \) is non-decreasing it follows that \( M \geq x_n \geq M - \varepsilon \) for all \( n \geq N_\varepsilon \), i.e.

\[ |M - x_n| \leq \varepsilon \text{ for all } n \geq N_\varepsilon. \]
As $\varepsilon > 0$ was arbitrary, we may conclude that $\lim_{n \to \infty} x_n = M$. If $x_n$ is decreasing instead, then $-x_n \uparrow$ and we have $\lim_{n \to \infty} x_n = \sup_n (-x_n) = -\inf_n x_n$.

**Theorem 3.16.** Suppose that $\mathbb{R}$ is a complete ordered field which we assume we have already embedded $\mathbb{Q}$ into $\mathbb{R}$ as in Lemma 2.14 Then:

1. For all $m \in \mathbb{R}$, if 
   \[ \alpha_m := \{ y \in \mathbb{Q} : y < m \}, \]
   then $\sup \alpha_m = m$.

2. If $a, b \in \mathbb{R}$ with $a < b$, then there exists $q \in \mathbb{Q}$ such that $a < q < b$.

3. If $\alpha \subseteq \mathbb{Q}$ is a cut and $m := \sup \alpha$, then $\alpha = \alpha_m$.

**Proof.** We take each item in turn.

1. **Proof 1.** Let $\alpha_m := \{ y \in \mathbb{Q} : y < m \}$ and $M := \sup \alpha_m \in \mathbb{R}$. Then $M \leq m$. If $M \neq m$ then $M < m$. To see this last case is not possible $\varepsilon := m - M > 0$ and choose $n \in \mathbb{N}$ such that $0 < \frac{1}{n} \leq \varepsilon$. Then choose $y \in \mathbb{Q}$ such that 
   \[ M - \frac{1}{2n} < y < M. \]
   From this it follows that 
   \[ M < y + \frac{1}{2n} < M + \frac{1}{2n} < m \]
   which shows $y + \frac{1}{2n} \in \alpha_m$ is greater than $M$ violating the assumption that $M$ is an upper bound for $\alpha_m$.

2. **Proof 2.** [Here is a slight rewriting of the above argument.] Choose $y_m \in \alpha_m$ such that $y_m \uparrow M$ as $m \to \infty$. Choose $n \in \mathbb{N}$ so that $m - M > \frac{1}{n}$. Then 
   \[ y_m + \frac{1}{n} \uparrow M + \frac{1}{n} < m \text{ as } m \to \infty. \]
   So for large $m$, $y_m + \frac{1}{n} < m$ while $y_m + \frac{1}{n} > M$, i.e. $y_m + \frac{1}{n} \in \alpha_m$ yet $y_m + \frac{1}{n} > M$. This violates the assumption that $M$ is an upper bound for $\alpha_m$.

2. By item 1. and Theorem 3.14 we can choose $q \in \alpha_b$ to be as close to $b$ as we choose and in particular $q$ can be chosen to be in $\alpha_b$ with $q > a$.

3. You are asked to prove this in Exercise 3.1 below.

**Exercise 3.1.** Suppose that $\alpha \subseteq \mathbb{Q}$ is a cut as in Definition 2.27. Show $\alpha$ is bounded from above. Then let $m := \sup \alpha$ and show that $\alpha = \alpha_m$, where $\alpha_m$ 
\[ \alpha_m := \{ y \in \mathbb{Q} : y < m \}. \]
Also verify that $\alpha_m$ is a cut for all $m \in \mathbb{R}$. [In this way we see that we may identify $\mathbb{R}$ with the cuts of $\mathbb{Q}$. This should motivate Dedekind’s construction of the real numbers as described in Rudin.]

**Proposition 3.17 (Rationals are dense in the reals).** For all $b \in \mathbb{R}$, there exists $q_n \in \mathbb{Q}$ such that $q_n \uparrow b$. Similarly there exists $p_n \in \mathbb{Q}$ such that $p_n \downarrow b$.

**Proof.** Given $b \in \mathbb{Q}$ we know that $b = \sup \alpha_b$ by Theorem 3.16 Then by Theorem 3.14 there exists $q_n \in \alpha_b$ such that $q_n \uparrow b$ as $n \to \infty$. The second assertion can be proved in much the same way as the first. Alternatively, let $q_n \in \mathbb{Q}$ such that $q_n \uparrow -b$ and set $p_n := -q_n \in \mathbb{Q}$. Then $p_n \downarrow b$.

**Definition 3.18.** The real numbers which are not rational are called **irrational** so the irrational numbers are $\mathbb{R} \setminus \mathbb{Q}$.

**Example 3.19 (Euler’s number).** Let $S_n := \sum_{k=0}^{n} \frac{1}{k!}$ for all $n \in \mathbb{N}_0$. We define Euler’s number to be,
\[ e := \lim_{n \to \infty} S_n = \sup \{ S_n : n \in \mathbb{N}_0 \} \in \mathbb{R}. \]
From Example 1.17 we have seen that $e \in \mathbb{R} \setminus \mathbb{Q}$.

**Theorem 3.20 (nth roots).** Let $n \in \mathbb{N}$ and $x > 0$ in $\mathbb{R}$, then there exists a unique $y \in \mathbb{R}_{+}$ such that $y^n = x$. We of course denote $y$ by $x^{1/n}$ for $\sqrt[n]{x}$. The function $x \to x^{1/n}$ is increasing. [See Rudin for more properties of $x^{1/n}$ and $x^{m/n}$ where $m \in \mathbb{Z}$ and $n \in \mathbb{N}$.]

**Proof.** **Uniqueness.** First of $t > s \geq 0$ then $t^n > s^n \geq 0$ as can be proved by induction.\(^3\) Thus if $x, y \geq 0$ and $x^n = y^n$ then $x = y$ for otherwise $x > y$ or $x < y$.

\(^3\) For what it is worth, as dictionary definition of irrational is “not consistent with or using reason.” Let’s try to use irrational numbers in a rational way!

\(^4\) The statement holds for $n = 1$ by assumption and if $t^n > s^n$, then $t^{n+1} > ts^n > s^{n+1}$. For the last equality we used $t > s$ implies $ts^n > s \cdot s^n$.
y > x in which case \( x^n > y^n \) or \( y^n > x^n \) respectively. This shows that there is at most one \( n^{th} \) root if it exists. I also claim that \( x^{1/n} < y^{1/n} \) if \( x < y \). If not then \( x^{1/n} \geq y^{1/n} \) and this would then imply \( x = (x^{1/n})^n \geq (y^{1/n})^n = y \) which contradicts \( x < y \).

**Existence.** Let \( \Lambda := \{ t \in \mathbb{R}^+ : t^n \leq x \} \). If \( t = \frac{1}{1+x} \in (0,1) \), then \( t^n \leq t \leq x \) so that \( t \in \Lambda \) and \( \Lambda \neq \emptyset \). If \( t = 1 + x \), then \( t^n = (1 + x)^n \geq 1 + nx > x \) and therefore \( \Lambda \) is bounded from above. Hence we may define \( y := \sup \Lambda \). We will now show that \( y^n = x \).

By Theorem 3.14 there exists \( t_k \in \Lambda \) such that \( t_k \uparrow y \) as \( k \to \infty \). By definition of \( \Lambda \), \( \frac{1}{t_k} \leq x \) for all \( k \). Passing to the limit as \( k \to \infty \) in this inequality implies \( y^n = \lim_{k \to \infty} t_k^n \leq x \).

If \( y^n < x \) then (using the Binomial theorem) and properties of limits,

\[
y^n = \left( y + \frac{1}{m} \right)^n = y^n + \sum_{k=1}^{n} \binom{n}{k} y^{n-k} \left( \frac{1}{m} \right)^k \rightarrow y^n < x \text{ as } m \to \infty.
\]

Hence for sufficiently large \( m \) we will have \( \left( y + \frac{1}{m} \right)^n < x \). But this shows that \( y + \frac{1}{m} \in \Lambda \) which violates \( y \) being an upper bound for \( \Lambda \). Therefore we conclude that \( y^n = x \).

- End of Lecture 6, 10/10/2012.

### 3.1 Extended real numbers

**Notation 3.21** 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, \( \pm \infty \cdot a = \mp \infty \) if \( a \in \mathbb{R} \) with \( a > 0 \), \( \pm \infty \cdot a = \mp \infty \) if \( a \in \mathbb{R} \) with \( a < 0 \), \( \pm \infty + a = \pm \infty \) for any \( a \in \mathbb{R} \), \( \infty + \infty = \infty \) and \( -\infty - \infty = -\infty \) while the following expressions are not defined:

\[
-\infty - \infty, \quad -\infty + \infty, \quad \infty / 0, \quad 0 \cdot \infty, \quad \text{and} \quad \infty \cdot 0.
\]

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 \( n \geq m \). In these case we write \( \lim_{n \to \infty} a_n = \pm \infty \) or \( a_n \to \pm \infty \) as \( n \to \infty \).

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 \( \inf \emptyset = +\infty \). The next theorem is a fairly simple but often useful result about computing least upper bounds.

**Theorem 3.22 (Sup Sup Theorem).** Suppose that \( \Lambda \) is a subset of \( \mathbb{R} \) such that \( \Lambda = \bigcup_{\alpha \in I} \Lambda_{\alpha} \) where \( \Lambda_{\alpha} \subset \Lambda \) and \( I \) is some index set. Then

\[
\sup \Lambda = \sup \alpha \in I \sup \Lambda_{\alpha}.
\]

The convention here is that the supremum of a set which is not bounded from above is \( \infty \) and the sup\( \emptyset = -\infty \).

**Proof.** Let \( M := \sup \Lambda \) and \( M_{\alpha} := \sup \Lambda_{\alpha} \) for all \( \alpha \in I \). As \( \Lambda_{\alpha} \subset \Lambda \) we have \( M_{\alpha} \leq M \) for all \( \alpha \in I \) and therefore \( \sup_{\alpha \in I} M_{\alpha} \leq M \). Conversely, if \( \lambda \in \Lambda \), then \( \lambda \in \Lambda_{\alpha} \) for some \( \alpha \in I \) and therefore \( \lambda \leq \sup_{\alpha \in I} M_{\alpha} \). From this it follows that \( \Lambda \leq \sup_{\alpha \in I} \Lambda_{\alpha} \) and as \( \lambda \in \Lambda \) is arbitrary we may conclude that

\[
M = \sup \Lambda = \sup_{\alpha \in I} \sup \Lambda_{\alpha}.
\]

The next corollary records a typical way the Sup Sup theorem is used.

**Corollary 3.23.** Suppose that \( X \) and \( Y \) are sets and \( S : X \times Y \to \mathbb{R} \) is a function. Then

\[
\sup_{x \in X, y \in Y} S(x, y) = \sup_{(x,y) \in X \times Y} \sup_{y \in Y} S(x, y).
\]

In particular, if \( S_{m,n} \in \mathbb{R} \) for all \( m, n \in \mathbb{N} \), then

\[
\sup_{m,n} S_{m,n} = \sup_{m} S_{m,n} = \sup_{n} S_{m,n}.
\]

**Proof.** Let \( \Lambda := \{ S(x, y) : (x, y) \in X \times Y \} \), and for \( x \in X \) let \( \Lambda_x := \{ S(x, y) : y \in Y \} \). Then \( \Lambda = \bigcup_{x \in X} \Lambda_x \) and therefore,

\[
\sup_{(x,y) \in X \times Y} S(x, y) = \sup_{x \in X} \sup_{y \in Y} S(x, y).
\]

The same reasoning also shows,

\[
\sup_{(x,y) \in X \times Y} S(x, y) = \sup_{y \in Y} \sup_{x \in X} S(x, y).
\]

The next Lemma records some basic limit theorems involving the extended real numbers.

**Lemma 3.24.** 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 \( c \in \mathbb{R} \), \( \lim_{n \to \infty} (ca_n) = c \lim_{n \to \infty} a_n \).

5 The only sequences that do not converge in \( \bar{\mathbb{R}} \) are those which oscillate too much.
3. If \( \{a_n + b_n\}_{n=1}^{\infty} \) is convergent and
\[
\lim_{n \to \infty} (a_n + b_n) = \lim_{n \to \infty} a_n + \lim_{n \to \infty} b_n
\]
provided the right side is not of the form \( \infty - \infty \).

4. \( \{a_nb_n\}_{n=1}^{\infty} \) is convergent and
\[
\lim_{n \to \infty} (a_nb_n) = \lim_{n \to \infty} a_n \cdot \lim_{n \to \infty} b_n
\]
provided the right hand side is not of the form \( \pm \infty \cdot 0 \) of \( 0 \cdot (\pm \infty) \).

Before going to the proof consider the simple example where \( a_n = n \) and \( b_n = -an + c \) with \( a > 0 \) and \( c \in \mathbb{R} \). Then\(^6\)
\[
\lim (a_n + b_n) = \begin{cases} 
\infty & \text{if } a < 1 \\
c & \text{if } a = 1 \\
-\infty & \text{if } a > 1
\end{cases}
\]
while
\[
\lim a_n + \lim b_n = "\infty - \infty".
\]
This shows that the requirement that the right side of Eq. 3.1 is not of form \( \infty - \infty \) is necessary in Lemma 3.24. 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. 3.2 is not of the form \( \infty \cdot 0 \).

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

**Proof of Eq. 3.1.** 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 \( n \to \infty \). 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,
\[
|a + b - (a_n + b_n)| = |a - a_n + b - b_n| \leq |a - a_n| + |b - b_n| \leq 2\varepsilon
\]
\(\text{for all } n \geq N. \) Since \( n \) is arbitrary, it follows that \( \lim_{n \to \infty} (a_n + b_n) = a + b \).

**Proof of Eq. 3.2.** It will be left to the reader to prove the case where \( a_n \) and \( b_n \) 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 \( x := \min \left( \frac{a}{2}, 1 \right) \). Given any \( M < \infty \), there exists \( N \in \mathbb{N} \) such that \( a_n \geq x \) and \( b_n \geq M \) for all \( n \geq N \) and for this choice of \( N \), \( a_nb_n \geq M \alpha \) for all \( n \geq N \). Since \( \alpha > 0 \) is fixed and \( M \) is arbitrary it follows that \( \lim_{n \to \infty} (a_nb_n) = \infty \) as desired. \[\blacksquare\]

**Exercise 3.2.** Show \( \lim_{n \to \infty} a^n = \infty \) and \( \lim_{n \to \infty} \frac{1}{c^n} = 0 \) whenever \( \alpha > 1 \).

**Exercise 3.3.** Suppose \( \alpha > 1 \) and \( k \in \mathbb{N} \), show there is a constant \( c = c(\alpha,k) > 0 \) such that \( a^n \geq cn^k \) for all \( n \in \mathbb{N} \). [In words, \( \alpha^n \) grows in \( n \) faster than any polynomial in \( n \).]

**Lemma 3.25.** Suppose that \( \{a_n\}_{n=1}^{\infty} \subset \mathbb{R} \) and \( \lim_{n \to \infty} a_n = A \in \mathbb{R} \). Then every subsequence, \( \{b_k := a_{n_k}\}_{k=1}^{\infty} \), also converges to \( A \).

**Exercise 3.4.** Prove Lemma 3.25.

### 3.2 Limsups and Liminfs

**Notation 3.26.** Suppose that \( \{x_n\}_{n=1}^{\infty} \subset \overline{\mathbb{R}} \) is a sequence of numbers. Then
\[
\liminf_{n \to \infty} x_n = \liminf_{n \to \infty} \{x_k : k \geq n\} \quad \text{and} \quad \limsup_{n \to \infty} x_n = \limsup_{n \to \infty} \{x_k : k \geq n\}.
\]
We will also write \( \underline{\lim} \) for \( \liminf \) and \( \overline{\lim} \) for \( \limsup \).

**Remark 3.27.** Notice that if \( a_k := \inf \{x_k : k \geq n\} \) and \( b_k := \sup \{x_k : k \geq n\} \), then \( \{a_k\} \) is an increasing sequence while \( \{b_k\} \) is a decreasing sequence. Therefore the limits in Eq. 3.3 and Eq. 3.4 always exist in \( \overline{\mathbb{R}} \) (see Theorem 3.15) and
\[
\liminf_{n \to \infty} x_n = \inf \{x_k : k \geq n\} \quad \text{and} \quad \limsup_{n \to \infty} x_n = \sup \{x_k : k \geq n\}.
\]
Owing to the following exercise, one may reduce properties of the \( \liminf \) to those of the \( \limsup \).

**Exercise 3.5.** Show \( \liminf_{n \to \infty} (-a_n) = -\limsup_{n \to \infty} a_n \).
Proposition 3.28. Let \( \{a_n\}_{n=1}^{\infty} \) and \( \{b_n\}_{n=1}^{\infty} \) be two sequences of real numbers. Then

1. \( \liminf_{n \to \infty} a_n \leq \limsup_{n \to \infty} a_n \).
2. \( \lim_{n \to \infty} a_n \) exists in \( \mathbb{R} \) iff
   \[
   \lim_{n \to \infty} a_n = \limsup_{n \to \infty} a_n \in \mathbb{R}.
   \]
3. \( \limsup_{n \to \infty} (a_n + b_n) \leq \limsup_{n \to \infty} a_n + \limsup_{n \to \infty} b_n \)
   whenever the right side of this equation is not of the form \( \infty - \infty \).
4. If \( a_n \geq 0 \) and \( b_n \geq 0 \) for all \( n \in \mathbb{N} \), then
   \[
   \lim_{n \to \infty} (a_n b_n) \leq \lim_{n \to \infty} a_n \cdot \lim_{n \to \infty} b_n,
   \]
   provided the right hand side of (3.7) is not of the form \( 0 \cdot \infty \) or \( \infty \cdot 0 \).

Proof. Items 1. and 2. will be proved here leaving the remaining items as an exercise to the reader. For item 1. we have

\[
\inf \{a_k : k \geq n\} \leq \sup \{a_k : k \geq n\} \quad \forall n,
\]
and therefore by the Sandwich theorem, \( \liminf_{n \to \infty} a_n \leq \limsup_{n \to \infty} a_n \).

2. \((\iff)\) Let \( A := \liminf_{n \to \infty} a_n = \limsup_{n \to \infty} a_n \in \mathbb{R} \). Since
   \[
   \inf_{k \geq n} a_k \leq a_n \leq \sup_{k \geq n} a_k,
   \]
if \( A \in \mathbb{R} \) then it follows by the sandwich theorem that \( \lim_{n \to \infty} a_n = A \). If \( A = \infty \), then for all \( M \in \mathbb{N} \) we have \( M \leq \inf_{k \geq n} a_k \) for a.a. \( n \). Therefore \( a_k \geq M \) for a.a. \( k \) and we have shown \( \lim_{k \to \infty} a_k = \infty \). If \( A = -\infty \) then for all

\[ M \in \mathbb{N} \] we have \( \sup_{k \geq n} a_k \leq -M \) for a.a. \( n \). Therefore \( a_k \leq -M \) for a.a. \( k \) and we have shown \( \lim_{k \to \infty} a_k = -\infty \).

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

\[ A - \varepsilon \leq a_n \leq A + \varepsilon \quad \forall n \geq N(\varepsilon). \]

From this we learn that

\[ A - \varepsilon \leq \inf_{k \geq n} a_k \leq \sup_{k \geq n} a_k \leq A + \varepsilon \quad \text{for a.a.} \quad n \]

and so passing to the limit as \( n \to \infty \) implies

\[ A - \varepsilon \leq \lim_{n \to \infty} a_n \leq \lim_{n \to \infty} \sup_{n \to \infty} a_n \leq A + \varepsilon. \]

Since \( \varepsilon > 0 \) is arbitrary, it follows that

\[ A \leq \lim_{n \to \infty} a_n \leq \lim_{n \to \infty} \sup_{n \to \infty} a_n, \]
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 \( N = N(M) \) such that \( a_n \geq M \) for all \( n \geq N \). This show that \( \liminf_{n \to \infty} a_n \geq M \) and since \( M \) is arbitrary it follows that

\[ \infty \leq \liminf_{n \to \infty} a_n \leq \limsup_{n \to \infty} a_n. \]

The proof for the case \( A = -\infty \) is analogous to the \( A = \infty \) case.

Exercise 3.8. Show that

\[ \limsup_{n \to \infty} (a_n + b_n) \leq \limsup_{n \to \infty} a_n + \limsup_{n \to \infty} b_n \]
provided that the right side of Eq. (3.8) is well defined, i.e. no \( \infty - \infty \) or \( -\infty + \infty \) type expressions. (It is OK to have \( \infty + \infty = \infty \) or \( \infty - \infty = -\infty = \infty \), etc.)

Exercise 3.9. Suppose that \( a_n \geq 0 \) and \( b_n \geq 0 \) for all \( n \in \mathbb{N} \). Show

\[ \limsup_{n \to \infty} (a_n b_n) \leq \limsup_{n \to \infty} a_n \cdot \limsup_{n \to \infty} b_n, \]
provided the right hand side of (3.9) is not of the form \( 0 \cdot \infty \) or \( \infty \cdot 0 \).

Exercise 3.10. Suppose that \( \{a_n\}_{n=1}^{\infty} \) and \( \{b_n\}_{n=1}^{\infty} \) are two non-negative sequences and assume \( A = \lim_{n \to \infty} a_n \) exists in \( (0, \infty) \). Show

\[ \lim_{n \to \infty} a_n b_n = A \cdot \limsup_{n \to \infty} b_n. \]
Definition 3.29. A sequence, \(\{a_n\}_{n=1}^{\infty}\), of positive numbers is said to have **sub-geometric growth** if for all \(\alpha > 1\) there exists \(c = c(\alpha) < \infty\) such that
\[
a_n \leq cn^\alpha \text{ for a.a. } n \in \mathbb{N}.
\]

Lemma 3.30. Suppose that \(\{a_n\}_{n=1}^{\infty}\) is a sequence of non-negative numbers.

1. If \(\{a_n\}_{n=1}^{\infty}\) has sub-geometric growth, then \(\limsup_{n \to \infty} a_n^{1/n} \leq 1\).
2. If \(a_n > 0\) for a.a. \(n\) and \(\{a_n^{-1}\}_{n=1}^{\infty}\) has sub-geometric growth (i.e. for all \(\alpha > 1\) there exists \(\varepsilon = \varepsilon(\alpha) < \infty\) such that \(a_n \geq \varepsilon a_n^{-\alpha}\) for a.a. \(n\)), then \(\liminf_{n \to \infty} a_n^{1/n} \geq 1\).
3. In particular if \(a_n > 0\) for a.a. \(n\) and both \(\{a_n\}_{n=1}^{\infty}\) and \(\{a_n^{-1}\}_{n=1}^{\infty}\) have sub-geometric growth, then \(\lim_{n \to \infty} a_n^{1/n} = 1\).

**Proof.** 1. For any \(\beta > 1\) choose \(\alpha \in (1, \beta)\) and let \(c = c(\alpha)\). Since \(\lim_{n \to \infty} c \left(\frac{2}{\beta}\right)^n = 0 < 1\), it follows that \(a_n \leq c n^\alpha = c \left(\frac{2}{\beta}\right)^n \beta n \leq \beta n\) for a.a. \(n\) and therefore
\[
a_n^{1/n} \leq \beta \text{ for a.a. } n \implies \limsup_{n \to \infty} a_n^{1/n} \leq \beta.
\]

As \(\beta > 1\) is arbitrary we may conclude that \(\limsup_{n \to \infty} a_n^{1/n} \leq 1\).

2. This may be proved similarly to item 1, or it can be reduced to item 1 as we show here. By item 1, applies to \(\{a_n^{-1}\}\),
\[
\limsup_{n \to \infty} \frac{1}{a_n^{1/n}} = \limsup_{n \to \infty} a_n^{-1/n} \leq 1.
\]

The proof is completed because \(\limsup_{n \to \infty} \frac{1}{a_n^{1/n}} = \frac{1}{\liminf_{n \to \infty} a_n^{1/n}}\) as the reader should prove.

3. From items 1. and 2. we learn that
\[
1 \leq \liminf_{n \to \infty} a_n^{1/n} \leq \limsup_{n \to \infty} a_n^{1/n} \leq 1
\]
from which the result follows.

**Example 3.31.** If \(e^{-1/n} - p \leq a_n \leq cn^p\) for some \(p \in \mathbb{N}\) and \(c < \infty\), then \(\lim_{n \to \infty} a_n^{1/n} = 1\). Indeed using Exercise 3.3, we may easily verify the hypothesis of Lemma 3.30.

**Exercise 3.11.** Suppose that \(p(t)\) is a non-zero polynomial of \(t \in \mathbb{R}\) with (possibly) complex coefficients. Show
\[
\lim_{n \to \infty} |p(n)|^{1/n} = 1.
\]

- End of Lecture 7, 10/12/2012.

**Exercise 3.12.** If \(a_n \geq 0\), then \(\lim_{n \to \infty} a_n = 0\) iff \(\sup_{n \to \infty} a_n = 0\).

**Proposition 3.32.** Suppose that \(\{a_n\}_{n=1}^{\infty}\) is a sequence of real numbers and let
\[
B := \{y \in \mathbb{R} : a_n \geq y \text{ for i.o. } n\}.
\]

Then \(\sup B = \limsup_{n \to \infty} a_n\) with the convention that \(\sup B = -\infty\) if \(B = \emptyset\).

**Proof.** If \(\{a_n\}_{n=1}^{\infty}\) is not bounded from above, then \(B = \mathbb{R}\), and hence \(\sup B = -\infty = \limsup_{n \to \infty} a_n\).

If \(y > \beta\), then \(a_n < y\) for a.a. \(n\). But then, let \(B := \{a_n \geq 1\} = \limsup_{n \to \infty} a_n\).

If \(y > \beta\), then \(a_n < y\) for a.a. \(n\). Hence \(a_n \leq \beta\) for a.a. \(n\). Thus we have shown \(\beta = 1\).

Theorem 3.33. There is a subsequence \(\{a_{n_k}\}_{k=1}^{\infty}\) of \(\{a_n\}_{n=1}^{\infty}\) such that \(\lim_{k \to \infty} a_{n_k} = \limsup_{n \to \infty} a_n\). Similarly, there is a sequence \(\{a_{n_k}\}_{k=1}^{\infty}\) of \(\{a_n^{-1}\}_{n=1}^{\infty}\) such that \(\lim_{k \to \infty} a_{n_k} = \liminf_{n \to \infty} a_n\). Moreover, every convergent subsequence, \(\{b_k := a_{n_k}\}_{k=1}^{\infty}\) of \(\{a_n\}_{n=1}^{\infty}\) satisfies
\[
\liminf_{n \to \infty} a_n \leq \lim_{k \to \infty} b_k \leq \limsup_{n \to \infty} a_n.
\]

**Proof.** Let me prove the last assertion first. Suppose that \(b_k := a_{n_k}\) is some convergent subsequence of \(\{a_n\}_{n=1}^{\infty}\). Then we have,
\[
\inf_{n \geq n_k} a_n \leq b_k \leq \sup_{n \geq n_k} a_n \text{ for all } k \in \mathbb{N}.
\]

Passing to the limit in this equation then implies,
\[
\liminf_{n \to \infty} a_n = 1 \leq \lim_{k \to \infty} b_k \leq \limsup_{n \to \infty} a_n = 1 \leq \limsup_{k \to \infty} a_{n_k} = \limsup_{n \to \infty} a_n.
\]

We have used, \(\{\inf_{n \geq n_k} a_n\}_{k=1}^{\infty}\) and \(\{\sup_{n \geq n_k} a_n\}_{k=1}^{\infty}\) are subsequence of the convergent sequences \(\{\inf_{n \geq k} a_n\}_{k=1}^{\infty}\) and \(\{\sup_{n \geq k} a_n\}_{k=1}^{\infty}\) respectively and therefore converge to the same limits respectively, see Lemma 3.30.

Now let us prove the first assertions. I will cover the \(\sup\) case here as the \(\inf\) case is similar or can be deduced from the \(\sup\) case with the aid of Exercise 3.5. Let \(A := \limsup_{n \to \infty} a_n\). We will need to consider three case, \(A \in \mathbb{R}\), \(A = \infty\), and \(A = -\infty\).

\[\text{This can be done more formally by choosing a sequence } \{y_k\}_{k=1}^{\infty}\text{ such that } y_k \downarrow \alpha \text{ so that } a^* \leq y_k. \text{ Now let } k \to \infty \text{ to conclude } a^* \leq \limsup_{k \to \infty} y_k = \alpha.\]
i) If \( A = \limsup_{n \to \infty} a_n = -\infty \) so that for all \( M \in \mathbb{N} \) there exists \( N \in \mathbb{N} \) such that \( \sup_{k \geq n} a_k \leq -M \) for all \( n \geq N \), i.e. \( a_n \leq -M \) for all \( n \geq N \). In this case it follows that in fact \( \lim_{n \to \infty} a_n = -\infty \) and we do not have to even choose as subsequence.

**Corollary 3.34 (Bolzano–Weierstrass Property / Compactness).** Every bounded sequence of real numbers, \( \{a_n\}_{n=1}^{\infty} \), has a convergent in \( \mathbb{R} \) subsequence, \( \{a_{n_k}\}_{k=1}^{\infty} \). If we drop the bounded assumption then we may only assert that there is a subsequence which is convergent in \( \mathbb{R} \).

**Proof.** Let \( M < \infty \) such that \( |a_n| \leq M \) for all \( n \in \mathbb{N} \), i.e. \( -M \leq a_n \leq M \) for all \( n \). We may then conclude from Exercise 3.7 that,

\[
-\infty \leq \limsup_{n \to \infty} a_n \leq \infty.
\]

It now follows from Theorem 3.33 that there exists a subsequence, \( \{a_{n_k}\}_{k=1}^{\infty} \) of \( \{a_n\}_{n=1}^{\infty} \) such that

\[
\lim_{k \to \infty} a_{n_k} = \limsup_{n \to \infty} a_n \in [-M, M] \subset \mathbb{R}.
\]

**Theorem 3.35 (\( \mathbb{R} \) is Cauchy complete).** If \( \{a_n\}_{n=1}^{\infty} \subset \mathbb{R} \) is a Cauchy sequence, then \( \lim_{n \to \infty} a_n \) exists in \( \mathbb{R} \) and in fact,

\[
\lim_{n \to \infty} a_n = \limsup_{n \to \infty} a_n = \liminf_{n \to \infty} a_n.
\]

**Proof.** We will give two proofs of this important theorem. Each proof uses the fact that \( \{a_n\}_{n=1}^{\infty} \subset \mathbb{R} \) is Cauchy implies \( \{a_n\}_{n=1}^{\infty} \) is bounded. This is proved exactly in the same way as the solution to Exercise 1.2.

First proof. By Corollary 3.34 there is a subsequence, \( \{a_{n_k}\}_{k=1}^{\infty} \), such that \( \lim_{k \to \infty} a_{n_k} = L \in \mathbb{R} \). As in the proof of Exercise 1.7 it follows that \( \lim_{n \to \infty} a_n \) exists and is equal to \( L \).

Second proof. Let \( a := \liminf_{n \to \infty} a_n \) and \( b := \limsup_{n \to \infty} a_n \). It suffices to show \( a = b \). As we always know that \( a \leq b \) it will suffice to show \( b \leq a \).

Given \( \varepsilon > 0 \), there exists \( N \in \mathbb{N} \) such that

\[
|a_m - a_n| \leq \varepsilon \text{ for all } m, n \geq N.
\]

In particular, for \( m, n \geq k \geq N \) we have \( a_m \leq a_n + \varepsilon \) and hence

\[
b \leq \sup_{m \geq k} a_m \leq a_n + \varepsilon \text{ for all } n \geq k.
\]

From this inequality we may further conclude,

\[
b \leq \inf_{n \geq k} a_n + \varepsilon \leq a + \varepsilon.
\]

As \( \varepsilon > 0 \) is arbitrary, we have indeed shown \( b \leq a \).
3.3 Partitioning the Real Numbers

Notation 3.36 (Intervals) For \( a, b \in \mathbb{R} \) with \( a < b \) we define,
\[
(a, b) := \{ x \in \mathbb{R} : a < x < b \}, \\
[a, b) := \{ x \in \mathbb{R} : a \leq x < b \}, \\
(a, b] := \{ x \in \mathbb{R} : a < x \leq b \}, \\
[a, b] := \{ x \in \mathbb{R} : a \leq x \leq b \}.
\]
We also also \( a = -\infty \) in the intervals, \((a, b)\) and \((a, b]\) and allows \( b = +\infty \) in the intervals \((a, b)\) and \([a, b].\)

Notation 3.37 (Pairwise disjoint unions) If \( X \) is a set and \( A_\alpha \subseteq X \) for \( \alpha \in I \), we write \( X = \sum_{\alpha \in I} A_\alpha \) to mean; \( X = \bigcup_{\alpha \in I} A_\alpha \) and \( A_\alpha \cap A_\beta \) for all \( \alpha \neq \beta \).

Exercise 3.16. Suppose that \( a, b, c, d \in \mathbb{R} \) such that \( a < b \leq c < d \). Show \((a, b) \cap (c, d) = \emptyset\) and \([a, b] \cap [c, d] = \emptyset\).

Lemma 3.38 (Well Ordering II). Suppose that \( S \) is a non-empty subset of \( \mathbb{Z} \) which is bounded from below, then \( \inf(S) \in S \), i.e. \( S \) has a (unique) minimizer.

Proof. As \( S \) is bounded from below, there exists \( k \in \mathbb{Z} \) such that \( k \leq s \) for all \( s \in S \). Therefore \( S := \{ s - k + 1 : s \in S \} \subseteq \mathbb{N} \) and hence by the Well ordering principle, \( \min(S) := m \in \mathbb{N} \) exists. That is \( m \leq s - k + 1 \) for all \( s \in S \) and there exists \( s_0 \in S \) such that \( m = s_0 - k + 1 \). These last statements are equivalent to saying,
\[
s_0 = m + k - 1 \leq s \quad \text{for all} \quad s \in S,
\]
which is to say \( s_0 = \min(S) \).

Proposition 3.39. Suppose that \( \{S_\alpha\}^{\infty}_{\alpha=-\infty} \subseteq \mathbb{R} \) such that \( S_\alpha < S_{\alpha+1} \) for all \( n \in \mathbb{Z} \), \( \lim_{n \to \infty} S_n = \infty \) and \( \lim_{n \to -\infty} S_n = -\infty \). Then
\[
\sum_{n \in \mathbb{Z}} (S_{n-1}, S_n] = \mathbb{R} = \sum_{n \in \mathbb{Z}} [S_n, S_{n+1}).
\]

Proof. The fact that \((S_n, S_{n+1}] \cap (S_m, S_{m+1}] = \emptyset\) follows from Exercise 3.16. For \( x \in \mathbb{R} \), let \( n_0 := \min\{ \{ n \in \mathbb{Z} : x \leq S_n \} \} \) which exists since \( \{ n \in \mathbb{Z} : x \leq S_n \} \) is non-empty as \( S_n \to \infty \) as \( n \to \infty \) and is bounded from below since \( S_n \to -\infty \) as \( n \to -\infty \). It then follows that \( x \leq S_{n_0} \) while \( x \neq S_{n_0-1} \), i.e. \( S_{n_0-1} < x \leq S_{n_0} \) and we have shown \( x \in (S_{n_0-1}, S_{n_0}] \) which completes the proof of the first equality in Eq. (3.12). The proof of the second equality is similar and so will be omitted.

Proposition 3.40. Suppose that \(-\infty < a < b < \infty \) and \( \{S_n\}^{N}_{n=0} \subseteq [a, b] \) such that \( a = S_0 < S_1 < \cdots < S_{N-1} < S_N = b \), then
\[
[a, b) = \bigcup_{n=1}^{N} (S_{n-1}, S_n).
\]
This result also holds if \( N = \infty \) provided we now assume \( S_n < S_{n+1} \) for all \( n \), \( a = S_0 \), and \( S_n \uparrow b \) as \( n \to \infty \).

Proof. This proof is very similar to the proof of Proposition 3.39 and so will be omitted.

3.4 The Decimal Representation of a Real Number

Lemma 3.41 (Geometric Series). Let \( \alpha \in \mathbb{R} \) or \( \alpha \in \mathbb{Q} \), \( m, n \in \mathbb{Z} \) and \( S := \sum_{k=n}^{m} \alpha^k \). Then
\[
S = \begin{cases} 
  m - n + 1 & \text{if } \alpha = 1 \\
  \frac{\alpha^{m+1} - \alpha^n}{\alpha - 1} & \text{if } \alpha \neq 1.
\end{cases}
\]

Proof. When \( \alpha = 1 \),
\[
S = \sum_{k=n}^{m} 1^k = m - n + 1.
\]
If \( \alpha \neq 1 \), then
\[
\alpha S - S = \alpha^{m+1} - \alpha^n.
\]
Solving for \( S \) gives
\[
S = \sum_{k=n}^{m} \alpha^k = \frac{\alpha^{m+1} - \alpha^n}{\alpha - 1} \quad \text{if } \alpha \neq 1.
\] (3.12)

Taking \( \alpha = 10^{-1} \) in Eq. (3.13) implies
\[
\sum_{k=n}^{m} 10^{-k} = \frac{10^{-(m+1)} - 10^{-n}}{10^{-1} - 1} = \frac{10^{n-1} - 10^{-(m-n+1)}}{9}
\]
and in particular, for all \( M \geq n \),
\[
\lim_{m \to \infty} \sum_{k=n}^{m} 10^{-k} = \frac{1}{9} \cdot \frac{1}{10^{n-1}} \geq \sum_{k=n}^{M} 10^{-k}.
\]
**Definition 3.42 (Decimal Numbers).** Let $\mathbb{D}$ denote those sequences $\alpha \in \{0, 1, 2, \ldots, 9\}^\infty$ 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 $n \in \mathbb{Z}$.

A decimal number is then an expression of the form

$$\alpha = -N \alpha_{-N+1} \ldots \alpha_0 \alpha_1 \alpha_2 \alpha_3 \ldots.$$ 

For example

$$52 + \sqrt{2} \approx 53.41421356237309504880168872420969807856967187537694807 \ldots.$$ 

To every decimal number $\alpha \in \mathbb{D}$ is the sequence $a_n = a_n(\alpha)$ defined for $n \in \mathbb{N}$ by

$$a_n := \sum_{k=-\infty}^{n} \alpha_k 10^{-k}. \quad \text{(a finite sum)}.$$ 

Since for $m > n$,

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

it follows that

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

which shows $\{a_n\}_{n=1}^{\infty}$ is a Cauchy sequence. Thus to every decimal number we may associate the real number

$$a(\alpha) := \lim_{n \to \infty} a_n.$$ 

**Theorem 3.43.** If $x \geq 0$ is a real number, there exists $\alpha \in \mathbb{D}$ such that $x = a(\alpha)$, i.e. all real numbers can be represented in decimal form.

**Proof.** If $x = 0$, we can take $\alpha_n = 0$ for all $n$ so that $0 = a(\alpha)$. So suppose that $x > 0$ and let $p := \min(\{n \in \mathbb{N} : x < n\})$. Set $m = p - 1$, then $m \leq x < m + 1$. We then define $\alpha_k$ for $k \leq 0$ so that $m = \alpha_{-N} \ldots \alpha_0$. We now construct $\alpha_k$ for $k \geq 1$. For $k = 1$ we write

$$[m, m+1) = \sum_{l=0}^{9} \left( m + \frac{l}{10^l} \right)$$

and then choose $\alpha_1 = l$ if $x \in \left[ \frac{m}{10}, \frac{m+1}{10} \right)$. We then construct $\alpha_2$ using,

$$[m + \frac{\alpha_1}{10}, m + \frac{\alpha_1 + 1}{10}) \sum_{l=0}^{9} \left( m + \frac{\alpha_1}{10} + \frac{l+1}{10} \right)$$

and set $\alpha_2 = l$ for $x \in \left[ \frac{m + \frac{\alpha_1}{10}}{10}, \frac{m + \frac{\alpha_1 + 1}{10}}{10} \right)$. Continuing this way inductively we construct $\{\alpha_k\}_{k=1}^{\infty}$ such that

$$x \in \left[ \frac{m}{10^k}, m + \frac{\alpha_1}{10^k} + \frac{\alpha_2}{10^k} + \ldots + \frac{\alpha_k}{10^k} \right).$$

It is now easy to see that $x = a(\alpha)$.

**Remark 3.44.** The representation of $x \geq 0$ as a decimal number may not be unique. For example,

$$0.9999 = \sum_{k=1}^{\infty} \frac{9}{10^k} = \frac{9}{10} \cdot \frac{1}{1 - \frac{1}{10}} = 1.000.$$ 

[Or note that

$$\frac{1}{10} \cdot \frac{9}{10} \cdot \frac{9}{10} \cdot \ldots = 0.1 \cdot 0.0 \cdot 0.1 = 0.01 \cdot 0.01 \cdot 0.01 = 10^{-n} \rightarrow 0 \text{ as } n \rightarrow \infty.]$$

On the other hand if we agree to not allow a tail of repeated 9’s as an element of $\mathbb{D}$, then the representation would be unique.

### 3.5 Summary of Key Facts about Real Numbers

1. The real numbers, $\mathbb{R}$, is the unique (up to order preserving field isomorphism) ordered field with the least upper bound property or equivalently which is Cauchy complete.
2. Informally the real numbers are the rational numbers with the (irrational) hole filled in.
3. Monotone bounded sequence always converge in $\mathbb{R}$.
4. A sequence converges in $\mathbb{R}$ iff it is Cauchy.
5. Cauchy sequences are bounded.
6. $\mathbb{N}$ is unbounded from above in $\mathbb{R}$.
7. For all $\varepsilon > 0$ in $\mathbb{R}$ there exists $n \in \mathbb{N}$ such that $\frac{1}{n} < \varepsilon$.
8. $\mathbb{Q}$ and $\mathbb{R} \setminus \mathbb{Q}$ is dense in $\mathbb{R}$. In particular, between any two real numbers $a < b$, there are infinitely many rational and irrational numbers.
9. Decimal numbers map (almost 1-1) into the real numbers by taking the limit of the truncated decimal number.
10. If $a, b, \varepsilon \in \mathbb{R}$, then

$$\left[ \frac{a}{10^n}, \frac{b}{10^n} \right) \cap \left[ \frac{m}{10^n}, \frac{m+1}{10^n} \right)$$
11. A number of standard limit theorems hold, see Theorem 3.13.

12. Unlike limits, lim sup and lim inf always exist. Moreover we have: \[ \liminf_{n \to \infty} a_n \leq \limsup_{n \to \infty} a_n \]
with equality iff \( \lim_{n \to \infty} a_n \) exists in which case
\[ \liminf_{n \to \infty} a_n = \lim_{n \to \infty} a_n = \limsup_{n \to \infty} a_n. \]

We may allow the values of \( \pm \infty \) in these statements.

13. If \( b_k := \{ a_{nk} \}_{k=1}^{\infty} \) is a convergent subsequence of \( \{ a_n \} \), then
\[ \liminf_{n \to \infty} a_n \leq \liminf_{k \to \infty} b_k \leq \limsup_{k \to \infty} b_k \leq \limsup_{n \to \infty} a_n \]

and we may choose \( \{ b_k \} \) so that \( \lim_{k \to \infty} b_k = \limsup_{n \to \infty} a_n \) or \( \lim_{k \to \infty} b_k = \liminf_{n \to \infty} a_n \).

14. Bounded sequences of real numbers always have convergence subsequences.

15. If \( S \subset \mathbb{R} \) and \( A := \sup (S) \), then there exists \( \{ a_n \}_{n=1}^{\infty} \subset S \) such that \( a_n \leq a_{n+1} \) for all \( n \) and \( \lim_{n \to \infty} a_n = \sup (S) \).

16. If \( S \subset \mathbb{R} \) and \( A := \inf (S) \), then there exists \( \{ a_n \}_{n=1}^{\infty} \subset S \) such that \( a_{n+1} \leq a_n \) for all \( n \) and \( \lim_{n \to \infty} a_n = \inf (S) \).

### 3.6 (Optional) Proofs of Theorem 3.6 and Theorem 3.3

In this section, we assume that \( \mathbb{R} \) is as describe in Theorem 3.6. The next exercise is relatively straightforward.

**Exercise 3.17.** Prove the following properties of \( \mathbb{R} \).

1. Show addition and multiplication in Theorem 3.6 are well defined.
2. Show \( (\mathbb{R}, +, \cdot) \) satisfies the axioms of a field. **Hint:** for constructing multiplicative inverses, make use of Proposition 3.45 below to conclude if \( \alpha := \{ a_n \}_{n=1}^{\infty} \in \mathbb{R} \) and \( a \neq 0 = i(0) \), then there exists \( N \in \mathbb{N} \) such that \( |a_n| \geq \frac{1}{N} \) for all \( n \). By redefining the first few terms of \( a_n \) if necessary, you may assume that \( |a_n| \geq \frac{1}{N} \) for all \( n \) and then take
\[ \alpha^{-1} = \{ (a_n)^{-1} \}_{n=1}^{\infty}. \]
3. Show \( i : \mathbb{Q} \to \mathbb{R} \) is an injective homomorphism of fields.

To finish the proof of Theorem 3.6 we must show that \( P \) is an ordering on \( \mathbb{R} \) with the least upper bound property. This will be carried out in the remainder of this section.

**Proposition 3.45.** Suppose that \( \alpha := \{ a_n \}_{n=1}^{\infty} \) and \( \beta := \{ b_n \}_{n=1}^{\infty} \) are real numbers. Then precisely one of the following three cases can happen:

1. \( \lim_{n \to \infty} (a_n - b_n) = 0 \), i.e. \( \alpha = \beta \),
2. there exists \( \varepsilon = \frac{1}{N} > 0 \) such that \( a_n - b_n \geq \varepsilon \) for a.a. \( n \) in which case \( \alpha > \beta \), or
3. there exists \( \varepsilon = \frac{1}{N} > 0 \) such that \( b_n - a_n \geq \varepsilon \) for a.a. \( n \) in which case \( \beta > \alpha \).

**Proof.** If case 1. does not hold then there exists \( \varepsilon > 0 \) such that \( |a_n - b_n| \geq \delta \) for infinitely many \( n \). There are now two possibilities (which will turn out to me mutually exclusive):

i) \( a_n - b_n \geq \delta \) i.o. \( n \),

ii) \( b_n - a_n \geq \delta \) i.o. \( n \).

Since \( \{ a_n \} \) and \( \{ b_n \} \) are Cauchy sequences, there exists \( N \in \mathbb{N} \) such that
\[ |a_n - a_m| \geq \delta/3 \quad \text{and} \quad |b_n - b_m| \geq \delta/3 \quad \text{for all} \quad m,n \geq N. \]

If case i) holds, we may choose an \( m \geq N \) such that \( a_m - b_m \geq \delta \) and so for \( n \geq N \) we find,
\[
\delta \leq a_m - b_m = a_m - a_n + a_n - b_n + b_n - b_m \\
\leq |a_m - a_n| + |a_n - b_n| + |b_n - b_m| \\
= \delta/3 + a_n - b_n + \delta/3
\]

from which it follows that \( a_n - b_n \geq \varepsilon := \delta/3 \) for all \( n \geq N \) and we are in case 2. Similarly if case ii) holds then we are in fact in case 3. of the proposition. ■

**Corollary 3.46.** Suppose that \( \alpha := \{ a_n \}_{n=1}^{\infty} \) and \( \beta := \{ b_n \}_{n=1}^{\infty} \) are real numbers, then \( \alpha \geq \beta \) iff for all \( N \in \mathbb{N} \),
\[
a_n - b_n \geq -\frac{1}{N} \quad \text{for a.a.} \quad n.
\] (3.13)

Alternatively put, \( \alpha \geq \beta \) iff for all \( N \in \mathbb{N} \),
\[
b_n \leq a_n + \frac{1}{N} \quad \text{for a.a.} \quad n.
\]

**Proof.** If \( \alpha = \beta \), then \( \lim_{n \to \infty} (a_n - b_n) = 0 \) and therefore Eq. (3.14) holds.
If \( \alpha > \beta \), then in fact \( a_n - b_n \geq \varepsilon > 0 > -1/N \) for a.a. \( n \).

Conversely, if \( \alpha < \beta \), then there exists \( \varepsilon > 0 \) such that \( b_n \geq a_n + \varepsilon \) for a.a. \( n \). Thus if Eq. (3.14) were to also hold we could conclude for each \( N \in \mathbb{N} \) that
\[
a_n \geq b_n - \frac{1}{N} \geq a_n + \varepsilon - \frac{1}{N} \quad \text{for a.a.} \quad n.
\]

This leads to a contradiction as soon as we choose \( N \) so large as to make \( 1/N < \varepsilon \). Thus if Eq. (3.14) holds we must have \( \alpha \geq \beta \). ■
Proposition 3.47. Suppose that \( \lambda \in \mathbb{R} \), \( \{a_n\}_{n=1}^\infty \) be a Cauchy sequence in \( \mathbb{Q} \), and \( \alpha := \{\lim_{n \to \infty} a_n\} \). If \( \lambda \leq i(a_k) \) for all \( k \) then \( \lambda \leq \alpha \). Similarly if \( i(a_k) \leq \lambda \) for all \( k \) then \( \alpha \leq \lambda \).

**Proof.** Let \( \lambda = \{\lambda_n\}_{n=1}^\infty \) and suppose that \( \lambda \leq i(a_n) \) for all \( n \). For sake of contradiction, suppose that \( \lambda > \alpha \), i.e. there exists an \( N \in \mathbb{N} \) such that \( \lambda_n \geq a_n + \frac{1}{N} \) for a.a. \( n \). The assumption that \( \lambda \leq i(a_k) \) implies that \( \lambda_n \leq a_k + \frac{1}{N} \) for all \( k \). Because \( \{a_k\} \) is Cauchy, we may conclude there exists \( M \in \mathbb{N} \) such that

\[
\lambda_n \leq a_k + \frac{1}{2N} \quad \text{for all} \quad k \geq M.
\]

By making \( M \) even larger if necessary, we may assume that \( \lambda_n \geq a_n + \frac{1}{N} \) for all \( n \geq M \) as well. From these two inequalities with \( k = n \geq M \) we learn

\[
a_n + \frac{1}{N} \leq \lambda_n \leq a_n + \frac{1}{2N} \quad \implies \quad \frac{1}{2N} \geq \frac{1}{N}
\]

and we have reached the desired contradiction. The fact that \( i(a_k) \leq \lambda \) for all \( k \) implies \( \alpha \leq \lambda \) is proved similarly. Alternatively if \( i(a_k) \leq \lambda \) then \( -\lambda \leq i(-a_k) \) which implies \( -\lambda \leq -\alpha \), i.e. \( \alpha \leq \lambda \).

With these results in hand, let us now show that \( \mathbb{R} \) as defined in Theorem 3.6 has the least upper bound property.

**Proof of the least upper bound property.** So suppose that \( A \subset \mathbb{R} \) is a non empty set which is bounded from above. For each \( m \in \mathbb{N} \), let \( k_m \in \mathbb{Z} \) be the smallest integer such that \( i(a_m) := i\left(\frac{k_m}{2^m}\right) \) is an upper bound for \( A \). Since, for all \( n \geq m \), \( a_m - 2^{-m} \leq a_n \leq a_m \), we may conclude that

\[
|a_n - a_m| \leq 2^{-\min(n,m)} \to 0 \quad \text{as} \quad n, m \to \infty.
\]

This shows \( \{a_n\}_{n=1}^\infty \) is Cauchy and hence we defined an element \( \alpha := \{\lim_{n \to \infty} a_n\} \in \mathbb{R} \). We now will show \( \alpha = \sup A \).

If \( \lambda \in \mathbb{R} \), then \( \lambda \leq i(a_n) \) for all \( n \) and so by Proposition 3.47 we conclude that \( \lambda \leq \alpha \), i.e. \( \alpha \) is an upper bound for \( A \). Now suppose that \( \beta \) is another upper bound for \( A \). As \( i(a_n - 2^{-n}) \) is not an upper bound for \( A \) there exists \( \lambda \in A \) such that

\[
i(a_n - 2^{-n}) < \lambda \leq \beta.
\]

So by another application of Proposition 3.47 we learn that

\[
\alpha = \{\lim_{n \to \infty} a_n\} = \{\lim_{n \to \infty} a_n - 2^{-n}\} = \sup A.
\]

This shows that \( \alpha \) is in fact the least upper bound for \( A \).

**Theorem 3.48 (Real numbers are unique).** Suppose that \( \mathbb{F} \) and \( \mathbb{G} \) are two complete ordered fields. Then there is a unique order preserving isomorphism, \( \varphi: \mathbb{F} \to \mathbb{G} \).

**(Sketch).** Suppose that \( \varphi: \mathbb{F} \to \mathbb{G} \) is an order preserving homomorphism. The usual arguments show that any homomorphism, \( \varphi: \mathbb{F} \to \mathbb{G} \) must satisfy \( \varphi(q1_p) = q1_G \). We know that \( \{q \cdot 1_p : q \in \mathbb{Q}\} \) and \( \{q \cdot 1_G : q \in \mathbb{Q}\} \) are dense copies of \( \mathbb{Q} \) inside of \( \mathbb{F} \) and \( \mathbb{G} \) respectively. Now for general \( \varphi \in \mathbb{F} \) choose \( q_n, p_n \in \mathbb{Q} \) that \( q_n1_p \uparrow a \) and \( p_n1_p \downarrow a \). Since \( \varphi \) is order preserving we must have \( q_n1_G = \varphi(q_n1_p) \) is increasing and \( p_n1_G = \varphi(p_n1_p) \) is decreasing. Moreover, since \( p_n - q_n \to 0 \) we must have \( \lim_{n \to \infty} \varphi(q_n1_p) = \lim_{n \to \infty} \varphi(p_n1_p) \). Since \( \varphi(q_n1_p) \leq \varphi(a) \leq \varphi(p_n1_p) \) for all \( n \) it then follows that \( \varphi(a) = \lim_{n \to \infty} q_n1_G = \lim_{n \to \infty} p_n1_G \) and we have shown \( \varphi \) is uniquely determined.

For the converse, if \( q_n \in \mathbb{Q} \) we know that

\[
|q_n1_p - q_m1_p| = |q_n - q_m| \quad \text{and} \quad |q_n1_G - q_m1_G| = |q_n - q_m| \quad \text{G}.
\]

Thus if \( \{q_n1_G\}_{n=1}^\infty \) is convergent in \( \mathbb{F} \) iff \( \{q_n1_G\}_{n=1}^\infty \) is convergent in \( \mathbb{G} \). Thus for any \( \alpha \in \mathbb{F} \) we choose \( q_n \in \mathbb{Q} \) such that \( q_n1_p \to a \) and then define \( \varphi(a) := \lim_{n \to \infty} q_n1_G \). One now checks that this formula is well defined (independent of the choice of \( \{q_n\} \subset \mathbb{Q} \) such that \( q_n1_p \to a \)) and defines an order preserving isomorphism. For example, if \( a \leq b \) we may choose \( \{q_n\} \subset \mathbb{Q} \) and \( \{p_n\} \subset \mathbb{Q} \) such that \( q_n1_p \uparrow a \) and \( p_n1_p \downarrow b \). Then \( q_n1_G \leq p_n1_G \) for all \( n \) and letting \( n \to \infty \) shows,

\[
\varphi(a) = \lim_{n \to \infty} q_n1_G \leq \lim_{n \to \infty} p_n1_G = \varphi(b) \quad \text{G}.
\]

The other properties of \( \varphi \) are proved similarly.

3.7 Supremum and Infimums of sets

**Definition 3.49.** Given a set \( A \subset X \), let

\[
1_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}
\]

be the **indicator function** of \( A \).

**Lemma 3.50 (Properties of inf and sup).** We have:

1. \( (\cup_n A_n)^c = \cap_n A_n^c \)
2. \( \{A_n \text{ i.o.}\}^c = \{A_n^c \text{ a.a.}\} \)
3. \( \limsup_{n \to \infty} A_n = \{x \in X : \sum_{n=1}^\infty 1_{A_n}(x) = \infty\} \)
4. \( \liminf_{n \to \infty} A_n = \{x \in X : \sum_{n=1}^\infty 1_{A_n}(x) < \infty\} \)
5. \( \sup_{k \geq n} 1_{A_k}(x) = 1_{\cup_{k \geq n} A_k} = 1_{\sup_{k \geq n} A_k} \)
6. \( \inf_{k \geq n} 1_{A_k}(x) = 1_{\cap_{k \geq n} A_k} = 1_{\inf_{k \geq n} A_k} \)
7. \( \limsup_{n \to \infty} A_n = \limsup_{n \to \infty} 1_{A_n} \), and
8. \( \liminf_{n \to \infty} A_n = \liminf_{n \to \infty} 1_{A_n} \).

**Proof.** These results follow fairly directly from the definitions and so the proof is left to the reader. (The reader should definitely provide a proof for herself.) □
Complex Numbers

Definition 4.1 (Complex Numbers). Let $\mathbb{C} = \mathbb{R}^2$ equipped with multiplication rule
\[(a, b)(c, d) \equiv (ac - bd, bc + ad)\]
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 $z = x + yi$ which in the future will be written simply as $z = x + iy$. If $z = x + iy$, let $Re z = x$ and $Im z = y$.

Writing $z = a + ib$ and $w = c + id$, the multiplication rule in Eq. (4.1) becomes
\[(a + ib)(c + id) \equiv (ac - bd) + i(bc + ad)\]
and in particular $1^2 = 1$ and $i^2 = -1$.

Proposition 4.2. The complex numbers $\mathbb{C}$ with the above multiplication rule satisfies the usual definitions of a field – see Definition 2.1. For example $1 + 1 = (1 + 1)i$ and the usual rule for vector addition. As is standard we will write $\mathbb{C}$ with the above multiplication rule.

Proof. Suppose $z = a + ib \neq 0$, we wish to find $w = c + id$ such that $zw = 1$ and this happens by Eq. (4.2) iff
\[\frac{a}{a^2 + b^2} - i\frac{b}{a^2 + b^2}.\]

Moreover $\mathbb{C}$ contains $\mathbb{R}$ as sub-field under the identification
\[\mathbb{R} \ni a \rightarrow a1 + 0i = (a, 0) \in \mathbb{C}.\]

Notation 4.3 We will write $1/z$ for $z^{-1}$ and $w/z$ to mean $z^{-1} \cdot w$.

Notation 4.4 (Conjugation and Modulus) If $z = a + ib$ with $a, b \in \mathbb{R}$ let $\bar{z} = a - ib$ and
\[|z|^2 \equiv \bar{z} = a^2 + b^2.\]

Notice that
\[Re z = \frac{1}{2}(z + \bar{z})\text{ and } Im z = \frac{1}{2i}(z - \bar{z}).\]

Proposition 4.5. Complex conjugation and the modulus operators satisfy:
1. $\bar{\bar{z}} = z$.
2. $\bar{zw} = \bar{z}\bar{w}$ and $\bar{z} + \bar{w} = \bar{z + w}$.
3. $|z| = |ar{z}|$.
2. Say \( z = a + ib \) and \( w = c + id \), then \( \bar{z}w \) is the same as \( zw \) with \( b \) replaced by \(-b\) and \( d \) replaced by \(-d\), and looking at Eq. (4.2) we see that

\[
\bar{z}w = (ac - bd) - i(bc + ad) = zw.
\]

4. \(|zw|^2 = zw\bar{z}w = |z|^2 |w|^2\) as real numbers and hence \(|zw| = |z||w|\).

5. Geometrically obvious or also follows from

\[
|z| = \sqrt{|\text{Re } z|^2 + |\text{Im } z|^2}.
\]

6. This is the triangle inequality which may be understood geometrically or by the computation

\[
|z + w|^2 = (z + w)(\bar{z} + \bar{w}) = |z|^2 + |w|^2 + \bar{z} \bar{w} + w \bar{z}
\]

\[
= |z|^2 + |w|^2 + w \bar{z} + \bar{w}z
\]

\[
= |z|^2 + |w|^2 + 2 \text{Re}(w \bar{z}) \leq |z|^2 + |w|^2 + 2|z||w|
\]

\[
= (|z| + |w|)^2.
\]

7. Obvious.

8. Follows from Eq. (4.3). Alternatively if \( \rho = \rho + i0 > 0 \) is a real number then \( \rho^{-1} = \rho^{-1} + i0 \) as is easily verified since \( \mathbb{R} \) is a sub-field of \( \mathbb{C} \). Thus since \( \bar{z}z = |z|^2 \) we find

\[
\frac{1}{|z|^2} \bar{z}z = \frac{1}{|z|^2} |z|^2 = 1 \quad \Rightarrow 
\]

\[
z^{-1} = \frac{\text{Re } z}{|z|^2} - i\frac{\text{Im } z}{|z|^2}.
\]

9. \(|z^{-1}| = \frac{|z|^{-1}}{|z|^2} \) \(= \frac{1}{|z|^2} \) \(|z| = \frac{1}{|z|} \).

**Corollary 4.6.** If \( w, z \in \mathbb{C} \), then

\[
||z| - |w|| \leq |z - w|.
\]

**Proof.** Just copy the proof of Lemma \[1.6\] ■

**Lemma 4.7.** For complex numbers \( u, v, w, z \in \mathbb{C} \) with \( v \neq 0 \neq z \), we have

\[
\frac{u}{v} \frac{w}{z} = \frac{uw}{vz} \text{ and } \frac{u}{v} + \frac{w}{z} = \frac{uz + vw}{vz}.
\]

[These statements hold in any field.]
Proof. For the first item, it suffices to check that
\[(uv)(u^{-1}v^{-1}) = u^{-1}ww^{-1} = 1 \cdot 1 = 1.\]
The rest follow using
\[
\frac{uw}{vz} = uw^{-1}wz^{-1} = uww^{-1}z^{-1} = uw(vz)^{-1} = \frac{uw}{vz}.
\]
Proof. By assumption there exists \(M < \) such that \(|z_n| \leq M\) for all \(n \in \mathbb{N}\).
Writing \(z_n = a_n + ib_n\) with \(a_n, b_n \in \mathbb{R}\) we may conclude that \(|a_n|, |b_n| \leq M\).
According to Corollary 3.34 there exists an increasing function \(N \ni k \to n_k \in \mathbb{N}\) such that \(\lim_{k \to \infty} a_{n_k} = A\) exists. Similarly, we can apply Corollary 3.34 again to find an increasing function \(N \ni l \to k_l \in \mathbb{N}\) such that \(\lim_{l \to \infty} b_{n_{k_l}} = B\) exists. We now let \(w_l := z_{n_{k_l}}\) for \(l \in \mathbb{N}\). Then \(\{w_l\}_{l=1}^\infty\) is a subsequence of \(\{z_n\}_{n=1}^\infty\) which is convergent to \(A + iB \in \mathbb{C}\). Indeed,
\[
|w_l - (A + iB)| = \left|a_{n_{k_l}} - A + i\left(b_{n_{k_l}} - B\right)\right| \\
\leq |a_{n_{k_l}} - A| + |b_{n_{k_l}} - B| \to 0 \text{ as } l \to \infty.
\]
Definition 4.8. 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 \(|z_n| \to 0\) as \(n \to \infty\). As usual if \(\{z_n\}_{n=1}^\infty\) converges to \(z\) as \(n \to \infty\) or \(z = \lim_{n \to \infty} z_n\).

Theorem 4.9. 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 \(n \to \infty\).

The complex numbers satisfy all the same limit theorems as the real numbers.

Theorem 4.10. If \(\{w_n\}_{n=1}^\infty\) and \(\{z_n\}_{n=1}^\infty\) are convergent sequences of complex numbers, then
1. \(\{w_n\}_{n=1}^\infty\) and \(\{z_n\}_{n=1}^\infty\) are Cauchy sequences.
2. \(\lim_{n \to \infty} (w_n + z_n) = \lim_{n \to \infty} w_n + \lim_{n \to \infty} z_n\).
3. \(\lim_{n \to \infty} (w_n \cdot z_n) = \lim_{n \to \infty} w_n \cdot \lim_{n \to \infty} z_n\).
4. if we further assume that \(\lim_{n \to \infty} z_n \neq 0\), then
\[
\lim_{n \to \infty} \left(\frac{w_n}{z_n}\right) = \frac{\lim_{n \to \infty} w_n}{\lim_{n \to \infty} z_n}.
\]

End of Lecture 11, 10/24/2012.

Lemma 4.11 (Bolzano–Weierstrass property). Every bounded sequence, \(\{z_n\}_{n=1}^\infty \subset \mathbb{C}\), has a convergent subsequence.
\[ M_z w = \begin{pmatrix} ac - bd \\
bc + ad \end{pmatrix} = \begin{pmatrix} a - b \\
 -b 

\begin{pmatrix} c \\
a 
\end{pmatrix} \] so that

\[ M_z = \begin{pmatrix} a - b \\
b 

\begin{pmatrix} b 
\end{pmatrix} \] = aI + bJ

where

\[ J := \begin{pmatrix} 0 & -1 \\
1 & 0 \end{pmatrix}, \]

We now have the following simple observations:

1. \( J^2 = -I \) and \( J^* = -J \),
2. \( M_z M_w = M_w M_z \) because \( J \) and \( I \) commute,
3. we have

\[ M_z M_w = (aI + bJ)(cI + dJ) = (ac - bd)I + (ad + bc)J = M_z w, \]

4. the associativity of complex multiplication follows from the associativity of matrix multiplication,
5. \( M_z^* = aI - bJ = M_z \) and in particular
6. \( M_z^* M_z = (M_z M_w)^* = M_w^* M_z^* = M_w M_z = M_{wz}, \)
7. \( M_z^* M_z = M_{|z|^2} = \det(M_z), \)
8. \( |wz| = \det(M_w M_z) = \det(M_w) \det(M_z) = |w| |z|, \)
9. \( M_z \) is invertible iff \( \det(M_z) \neq 0 \) which happens iff \( |z|^2 \neq 0 \) and in this case we know from basic linear algebra that

\[ M_z^{-1} = \frac{1}{a^2 + b^2} \begin{pmatrix} a & b \\
 -b & a \end{pmatrix} = \frac{1}{|z|^2} M_z^* = M_{\frac{1}{|z|^2} z}, \]

10. With this notation we have \( M_z M_w = M_{zw} \) and since \( I \) and \( J \) commute it follows that \( zw = wz \). Moreover, since matrix multiplication is associative so is complex multiplication. Also notice that \( M_z \) is invertible iff \( \det(M_z) = a^2 + b^2 = |z|^2 \neq 0 \) in which case

\[ M_z^{-1} = \frac{1}{|z|^2} \begin{pmatrix} a & b \\
 -b & a \end{pmatrix} = M_{\frac{1}{|z|^2} z}, \]

as we have already seen above.
Set Operations, Functions, and Counting

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, and $\mathbb{C}$ the complex numbers. We will also use $\mathbb{F}$ to stand for either of the fields $\mathbb{R}$ or $\mathbb{C}$.

### 5.1 Set Operations and Functions

**Notation 5.1** Given two sets $X$ and $Y$, let $Y^X$ denote the collection of all functions $f : X \rightarrow Y$. If $X = \mathbb{N}$, we will say that $f \in Y^X$ 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 5.2** 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 $\pi_\alpha : X_A \rightarrow X_\alpha$ be the canonical projection map defined by $\pi_\alpha(x) = x_\alpha$. If $X_\alpha = X$ for some fixed space $X$, then we will write $\prod_{\alpha \in A} X_\alpha$ as $X^A$ rather than $X_A$.

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 states that $X_A \neq \emptyset$ provided that $X_\alpha \neq \emptyset$ for each $\alpha \in A$.

**Notation 5.3** 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

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

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

$$\cup_{\alpha \in I} A_\alpha := \{x \in X : \exists \alpha \in I \quad \exists x \in A_\alpha\}$$

and

$$\cap_{\alpha \in I} A_\alpha := \{x \in X : x \in A_\alpha \forall \alpha \in I\}.$$ 

**Example 5.4.** Let $A, B,$ and $C$ be subsets of $X$. Then

$$A \cap (B \cup C) = [A \cap B] \cup [A \cap C].$$

Indeed, $x \in A \cap (B \cup C) \iff x \in A$ and $x \in B \cup C \iff x \in A$ and $x \in B$ or $x \in A$ and $x \in C \iff x \in A \cap B$ or $x \in A \cap C \iff x \in [A \cap B] \cup [A \cap C]$.

**Notation 5.5** We will also write $\bigcup_{\alpha \in I} A_\alpha$ 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

$$\{A_n \text{ i.o.}\} := \{x \in X : \# \{n : x \in A_n\} = \infty\}$$

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

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

and this may be expressed as

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

Similarly, $x \in \{A_n \text{ a.a.}\}$ iff
\[ \exists N \in \mathbb{N} \exists \forall n \geq N, \; x \in A_n \]

which may be written as
\[ \{ A_n \text{ a.a.} \} = \bigcup_{N=1}^{\infty} \cap_{n \geq N} A_n. \]

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

**Definition 5.6.** If \( f : X \to Y \) is a function and \( B \subset Y \), then
\[ f^{-1}(B) := \{ x \in X : f(x) \in B \}. \]

If \( A \subset X \) we also write,
\[ f(A) := \{ f(x) : x \in A \} \subset Y. \]

**Example 5.7.** If \( f : X \to Y \) is a function and \( B \subset Y \), then \( f^{-1}(B^c) = [f^{-1}(B)]^c \) or to be more precise,
\[ f^{-1}(Y \setminus B) = X \setminus f^{-1}(B). \]

To prove this notice that
\[ x \in f^{-1}(B^c) \iff f(x) \in B^c \iff f(x) \notin B \iff x \notin f^{-1}(B) \iff x \in [f^{-1}(B)]^c. \]

On the other hand, if \( A \subset X \) it is **not** necessarily true that \( f(A^c) = [f(A)]^c \).

For example, suppose that \( f : \{1, 2\} \to \{1, 2\} \) is the defined by \( f(1) = f(2) = 1 \)
and \( A = \{1\} \). Then \( f(A) = f(A^c) = \{1\} \) where \( [f(A)]^c = \{1\}^c = \{2\} \).

**Notation 5.8** If \( f : X \to Y \) is a function and \( E \subset 2^Y \) let
\[ f^{-1}E := f^{-1}(E) := \{ f^{-1}(E) \mid E \in E \}. \]

If \( G \subset 2^X \), let
\[ f_\ast G := \{ A \in 2^Y \mid f^{-1}(A) \in G \}. \]

**Definition 5.9.** Let \( E \subset 2^X \) be a collection of sets, \( A \subset X \), \( i_A : A \to X \) be the **inclusion map** \( (i_A(x) = x \text{ for all } x \in A) \) and
\[ E_A = i_A^{-1}(E) = \{ A \cap E : E \in E \} \]
Proposition 5.11. If $X$ and $Y$ are sets, then $\text{card}(X) \leq \text{card}(Y)$ iff $\text{card}(Y) \geq \text{card}(X)$.

Proof. If $f : X \to Y$ is an injective map, define $g : Y \to X$ by $g|_{f(X)} = f^{-1}$ and $g|_{Y \setminus f(X)} = x_0 \in X$ chosen arbitrarily. Then $g : Y \to X$ is surjective.

If $g : Y \to X$ is a surjective map, then $Y_x := g^{-1} \{\{x\}\} \neq \emptyset$ for all $x \in X$ and so by the axiom of choice there exists $f \in \prod_{x \in X} Y_x$. Thus $f : X \to Y$ such that $f(x) \in Y_x$ for all $x$. As the $\{Y_x\}_{x \in X}$ are pairwise disjoint, it follows that $f$ is injective.

Theorem 5.12 (Schröder-Bernstein Theorem). If $X$ and $Y$ are sets then either $\text{card}(X) \leq \text{card}(Y)$ or $\text{card}(Y) \leq \text{card}(X)$. Moreover, if $\text{card}(X) \leq \text{card}(Y)$ and $\text{card}(Y) \leq \text{card}(X)$, then $\text{card}(X) = \text{card}(Y)$. [Stated more explicitly: if there exists injective maps $f : X \to Y$ and $g : Y \to X$, then there exists a bijective map, $h : X \to Y$.]

Proof. These results are proved in the appendices. For the first assertion see [13.8] and for the second see Theorem 5.11.

Exercise 5.6. If $X = X_1 \cup X_2$ with $X_1 \cap X_2 = \emptyset$, $Y = Y_1 \cup Y_2$ with $Y_1 \cap Y_2 = \emptyset$, and $X_i \sim Y_i$ for $i = 1, 2$, then $X \sim Y$. This exercise generalizes to an arbitrary number of factors.

5.3 Finite Sets

Notation 5.13 (Integer Intervals) For $n \in \mathbb{N}$ we let

$$J_n := \{1, 2, \ldots, n\} := \{k \in \mathbb{N} : k \leq n\}.$$ 

Definition 5.14. We say a non-empty set, $X$, is finite if $\text{card}(X) = \text{card}(J_n)$ for some $n \in \mathbb{N}$. We will also write $\#(X) = n$ to indicate that $\text{card}(X) = \text{card}(J_n)$. [It is shown in Theorem 5.17 below that $\#(X)$ is well defined, i.e. it is not possible for $\text{card}(X) = \text{card}(J_n)$ and $\text{card}(X) = \text{card}(J_m)$ unless $m = n$.]

Lemma 5.15. Suppose $n \in \mathbb{N}$ and $k \in J_{n+1}$, then $\text{card}(J_{n+1} \setminus \{k\}) = \text{card}(J_n)$.

Proof. Let $f : J_n \to J_{n+1} \setminus \{k\}$ be defined by

$$f(x) = \begin{cases} x & \text{if } x \leq k \\ x + 1 & \text{if } x > k \end{cases}$$

Then $f$ is the desired bijection.

Alternatively. If $n = 1$, then $J_2 = \{1, 2\}$ and either $J_2 \setminus \{k\} = J_2$ or $J_2 \setminus \{k\} = \{2\}$, either way $\text{card}(J_2 \setminus \{k\}) = \text{card}(J_1)$. Now suppose that result holds for a given $n \in \mathbb{N}$ and $k \in J_{n+2}$. If $k = (n+2)$ we have $J_{n+2} \setminus \{k\} = J_{n+1}$ so $\text{card}(J_{n+2} \setminus \{k\}) = \text{card}(J_{n+1})$ while if $k \in J_{n+1} \subset J_{n+2}$, then $J_{n+2} \setminus \{k\} = (J_{n+1} \setminus \{k\}) \cup \{n+2\} \sim J_n \cup \{n \sim J_n \cup \{n+1\} = J_{n+1}$.

Lemma 5.16. If $m, n \in \mathbb{N}$ with $n > m$, then every map, $f : J_n \to J_m$, is not injective.

Proof. If $f : J_n \to J_m$ were injective, then $f|_{J_{m+1}} : J_{m+1} \to J_m$ would be injective as well. Therefore it suffices to show there is no injective map, $f : J_{m+1} \to J_m$ for all $m \in \mathbb{N}$. We prove this last assertion by induction on $m$.

The case $m = 1$ is trivial as $J_1 = \{1\}$ so the only function, $f : J_2 \to J_1$ is the function, $f(1) = 1 = f(2)$ which is not injective.

Now suppose $m \geq 1$ and there were an injective map, $f : J_{m+2} \to J_m$. Letting $k := f(m + 2)$ we would have, $f|_{J_{m+1}} : J_{m+1} \to J_{m+1} \setminus \{k\}$ $\sim J_m$, which would produce an injective map from $J_{m+1}$ to $J_m$. However this contradicts the induction hypothesis and thus completes the proof.

Theorem 5.17. If $m, n \in \mathbb{N}$, then $\text{card}(J_m) \leq \text{card}(J_n)$ iff $m \leq n$. Moreover, $\text{card}(J_m) = \text{card}(J_n)$ iff $m = n$ and hence $\text{card}(J_m) < \text{card}(J_n)$ iff $m < n$.

Proof. As $J_m \subseteq J_m$ if $m \leq n$ and $J_m = J_m$ if $m = n$, it is only the forward implications that have any real content. If $\text{card}(J_m) \leq \text{card}(J_n)$, there exists an injective map, $g : J_m \to J_n$. According to Lemma 5.16 this can only happen if $m \leq n$. If $\text{card}(J_n) = \text{card}(J_m)$, then $\text{card}(J_n) \leq \text{card}(J_m) + \text{card}(J_m) \leq \text{card}(J_n) = \text{card}(J_m) + \text{card}(J_m)$ and hence $m \leq n$ and $n \leq m$, i.e. $m = n$.

Proposition 5.18. If $X$ is a finite set with $\#(X) = n$ and $S$ is a non-empty subset of $X$, then $S$ is a finite set and $\#(S) \leq n$. Moreover if $\#(S) = n$, then $S = X$.

Proof. It suffices to assume that $X = J_n$ and $S \subseteq J_n$. We now give two proofs of the result.

Proof 1. Let $S_1 = S$ and $f(1) := \min S \geq 1$. If $S_2 := S_1 \setminus \{f(1)\}$ is not empty, let $f(2) := \min S_2 \geq 2$. We then continue this construction inductively. So if $f(k) = \min S_k \geq k$ has been constructed, then we define $S_{k+1} := S_k \setminus \{f(k)\}$. If $S_{k+1} \neq \emptyset$ we define $f(k+1) := \min S_{k+1} \geq k + 1$. As $f(k) \geq k$.

1 You should read $\#(X) = n$, as $X$ is a set with $n$ elements.
for all $k$ that $f$ is defined, this process has to stop after at most $n$ steps. Say it stops at $k$ so that $S_{k+1} = \emptyset$. Then $f: J_k \to S$ is a bijection and therefore $S$ is finite and $\#(S) = k \leq n$. Moreover, the only way that $k = n$ is if $f(k) = k$ at each step of the construction so that $f: J_n \to S$ is the identity map in this case, i.e. $S = J_n$.

**Proof.** We prove this by induction on $n$. When $n = 1$ the only no-empty subset of $S$ of $J_1$ is $J_1$ itself. Thus $\#(S) = 1$ and $S = J_1$. Now suppose that the result hold for some $n \in \mathbb{N}$ and let $S \subset J_{n+1}$. If $n+1 \notin S$, then $S \subset J_n$ and by the induction hypothesis we know $\#(S) = k \leq n < n+1$. So now suppose that $n+1 \in S$ and let $S' := S \setminus \{n+1\} \subset J_n$. Then by the induction hypothesis, $S'$ is a finite set and $\#(S') = k \leq n$, i.e. there exists a bijection, $f': J_k \to S'$ and $S' = J_k$ is $k = n$. Therefore $f: J_{k+1} \to S$ given by $f = f'$ on $J_k$ and $f(k+1) = n+1$ is a bijections from $J_{k+1}$ to $S$. Therefore $\#(S) = k+1 \leq n+1$ with equality iff $S' = J_n$ which happens iff $S = J_{n+1}$.

**Proposition 5.19.** If $f: J_n \to J_n$ is a map, then the following are equivalent,

1. $f$ is injective,
2. $f$ is surjective,
3. $f$ is bijective.

**Proof.** If $n = 1$, the only map $f: J_1 \to J_1$ is $f(1) = 1$. So in this case there is nothing to prove. Now suppose the proposition holds for level $n$ and $f: J_{n+1} \to J_{n+1}$ is a given map.

If $f: J_{n+1} \to J_{n+1}$ is an injective map and $f(J_{n+1})$ is a proper subset of $J_{n+1}$, then card $(J_{n+1}) < card (f(J_{n+1})) = card (J_{n+1})$ which is absurd. Thus $f$ is injective implies $f$ is surjective.

Conversely suppose that $f: J_{n+1} \to J_{n+1}$ is surjective. Let $g: J_{n+1} \to J_{n+1}$ be a right inverse, i.e. $f \circ g = id$, which is necessarily injective, see the proof of Proposition 5.11. By the previous paragraph we know that $g$ is necessarily surjective and therefore $f = g^{-1}$ is a bijection.

**Theorem 5.20.** A subset $S \subset \mathbb{N}$ is finite iff $S$ is bounded. Moreover if $\#(S) = n \in \mathbb{N}$ then the sup $(S) \geq n$ with equality iff $S = J_n$.

**Proof.** If $S$ is bounded then $S \subset J_n$ for some $n \in \mathbb{N}$ and hence $S$ is a finite set by Proposition 5.18. Also observe that if $\#(S) = n = sup (S)$, then $S \subset J_n$ and $\#(S) = n = \#(J_n)$. Thus it follows from Proposition 5.18 that $S = J_n$.

Conversely suppose that $S \subset \mathbb{N}$ is a finite set and let $n = \#(S)$. We will now complete the proof by induction. If $n = 1$ we have $S \sim J_1$ and therefore $S = \{k\}$ for some $k \in \mathbb{N}$. In particular sup $S = k \geq 1$ with equality iff $S = J_1$.

Suppose the truth of the statement for some $n \in \mathbb{N}$ and let $S \subset \mathbb{N}$ be a set with $\#(S) = n+1$. If we choose a point, $k \in S$, we have by Lemma 5.15 that $\#(S \setminus \{k\}) = n$. Hence by the induction hypothesis, sup$(S \setminus \{k\}) \geq n$ with equality iff $S \setminus \{k\} = J_n$. If sup$(S \setminus \{k\}) > n$ then sup$(S) \geq sup(S \setminus \{k\}) \geq n+1$ as desired. If sup$(S \setminus \{k\}) = n$ then $S \setminus \{k\} = J_n$ therefore $S \ni k > n$. Hence it follows that sup$(S) = k \geq n+1$.

**Corollary 5.21.** Suppose $S$ is a non-empty subset of $\mathbb{N}$. Then $S$ is an unbounded subset of $\mathbb{N}$ iff card$(J_n) \leq$ card$(S)$ for all $n \in \mathbb{N}$.

**Proof.** If $S$ is bounded we know card$(S) = card(J_k)$ for some $k \in \mathbb{N}$ which would violate the hypothesis that card$(J_n) \leq$ card$(S)$ for all $n \in \mathbb{N}$. Conversely if card$(S) \leq$ card$(J_n)$ for some $n \in \mathbb{N}$, then there exists an injective map, $f: S \to J_n$. Therefore $card(S) = card(f(S)) = card(J_k)$ for some $k \leq n$. So $S$ is finite and hence bounded in $\mathbb{N}$ by Theorem 5.20.

**Exercise 5.7.** Suppose that $m,n \in \mathbb{N}$, show $J_{m+n} = J_m \cup (m+J_n)$ and $(m+J_n) \cap J_m = \emptyset$. Use this to conclude if $X$ is a disjoint union of two non-empty finite sets, $X_1$ and $X_2$, then $\#(X) = \#(X_1) + \#(X_2)$.

**Exercise 5.8.** Suppose that $m,n \in \mathbb{N}$, show $J_m \times J_n \sim J_{mn}$. Use this to conclude if $X$ and $Y$ are two non-empty sets, then $\#(X 	imes Y) = \#(X) \cdot \#(Y)$.

### 5.4 Countable and Uncountable Sets

**Definition 5.22 (Countability).** A set $X$ is said to be countable if $X = \emptyset$ or if there exists a surjective map, $f: \mathbb{N} \to X$. Otherwise $X$ is said to be uncountable.

**Remark 5.23.** From Proposition 5.11 it follows that $X$ is countable iff there exists an injective map, $g: X \to \mathbb{N}$. This may be succinctly stated as: $X$ is countable iff card$(X) \leq$ card$(\mathbb{N})$. From a practical point of view as set $X$ is countable iff the elements of $X$ may be arranged into a linear list,

$$X = \{x_1, x_2, x_3, \ldots\}.$$ 

**Example 5.24.** The integers, $\mathbb{Z}$, are countable. In fact $\mathbb{N} \sim \mathbb{Z}$, for example define $f: \mathbb{N} \to \mathbb{Z}$ by

$$(f(1), f(2), f(3), f(4), f(5), f(6), f(7), \ldots) := (0, 1, -1, 2, -2, 3, -3, \ldots).$$

**Remark 5.25 (Countability in a nutshell).** If $f: \mathbb{N} \to X$ is surjective, then $g(x) := \min f^{-1}(\{x\})$ defines an injective map, $g: X \to \mathbb{N}$. If $g: X \to \mathbb{N}$ is injective, then $f(n) := g^{-1}(n)$ for $n \in g(X) := S$ and $f(n) = x_0 \in X$ for $n \notin S$ defines a surjective map, $f: X \to \mathbb{N}$. Moreover, if $S$ is a subset of $\mathbb{N}$ we may list its elements in increasing order so that either
In more detail, let
\[ S = \{ n_1 < n_2 < \cdots < n_k \} \text{ for some } k \in \mathbb{N} \text{ or} \]
\[ S = \{ n_1 < n_2 < \cdots < n_k < \ldots \}. \]

In the first case card \((X) = \text{card} (J_k)\) while in the second case \((S) = \text{card} (\mathbb{N})\).

[Define \(f(j) := n_j\) to set up the bijections between \(J_k\) or \(\mathbb{N}\) and \(S\).]

The above arguments demonstrate that the following statements are equivalent:

1. \(X\) is countable, i.e. there exists a surjective map \(f : \mathbb{N} \rightarrow X\).
2. \(\text{card} (X) \leq \text{card} (\mathbb{N})\), i.e. there exists an injective map, \(g : X \rightarrow \mathbb{N}\).
3. There exists \(S \subset \mathbb{N}\) such that \(\text{card} (X) = \text{card} (S)\). Furthermore \(\text{card} (S) = \text{card} (J_k)\) for some \(k \in \mathbb{N}\) iff \(S\) is bounded and \(\text{card} (S) = \text{card} (\mathbb{N})\) iff \(S\) is unbounded.
4. Either \(\text{card} (X) = \text{card} (\mathbb{N})\) for \(\text{card} (X) = \text{card} (J_k)\) for some \(k \in \mathbb{N}\).

Formal proofs of these observations are given above and below.

**Lemma 5.26.** If \(S \subset \mathbb{N}\) is an unbounded set, then \(\text{card} (S) = \text{card} (\mathbb{N})\).

**Proof.** The main idea is that any subset, \(S \subset \mathbb{N}\), may be given as an infinite or infinite list written in increasing order, i.e.

\[ S = \{ n_1, n_2, n_3, \ldots \} \text{ with } n_1 < n_2 < n_3 < \ldots . \]

If the list is finite, say \(S = \{ n_1, \ldots, n_k \}\), then \(n_k\) is an upper bound for \(S\). So \(S\) will be unbounded iff only if the list is infinite in which case \(f : \mathbb{N} \rightarrow S\) defined by \(f(k) = n_k\) defines a bijection.

**Formal proof.** Define \(f : \mathbb{N} \rightarrow S\) via, let

\[
S_1 := S \text{ and } f(1) := \min S_1, \\
S_2 := S_1 \setminus \{ f(1) \} \text{ and } f(2) := \min S_2, \\
S_3 := S_2 \setminus \{ f(2) \} \text{ and } f(3) := \min S_3, \\
\vdots
\]

In more detail, let \(T\) denote those \(n \in \mathbb{N}\) such that there exists \(f : J_n \rightarrow S\) and \(\{ S_k \subset S \}_{k=1}^n\) satisfying, \(S_1 = S\), \(f(k) = \min S_k\) and \(S_{k+1} = S_{k} \setminus \{ f(k) \}\) for \(1 \leq k < n\). If \(n \in T\), we may define \(S_{n+1} := S_n \setminus \{ f(n) \}\) and \(f(n+1) := \min S_{n+1}\) in order to show \(n+1 \in T\). Thus \(T \subset \mathbb{N}\) and we have constructed an injective map, \(f : \mathbb{N} \rightarrow S\). Moreover \(\bigcap_{k \in \mathbb{N}} S_k \subset \mathbb{N} \setminus J_n\) for all \(n\) and therefore \(\bigcap_{k \in \mathbb{N}} S_k = \emptyset\). Thus it follows that \(f\) is a bijection. \(\blacksquare\)

- End of Lecture 14, 10/31/2012.

The following theorem summarizes most of what we need to know about counting and countability.

**Theorem 5.27.** The following properties hold:
1. \(N \times N\) is countable and in fact \(N \times N \sim N\), i.e. there exists a bijective map, \(h, \text{ from } N \text{ to } N \times N\).
2. If \(X\) and \(Y\) are countable, then \(X \times Y\) is countable.
3. If \(\{ X_n \}_{n \in \mathbb{N}}\) are countable sets then \(X := \bigcup_{n=1}^{\infty} X_n\) is a countable set.
4. If \(X\) is countable, then either there exists \(n \in \mathbb{N}\) such that \(X \sim J_n\) or \(X \sim N\).
5. If \(S \subset \mathbb{N}\) and \(S \sim J_n\) for some \(n \in \mathbb{N}\) then \(S\) is bounded.
6. If \(X\) is a set and \(J_n \leq \text{card} (X)\) for all \(n \in \mathbb{N}\) then \(\text{card} N \leq \text{card} X\).
7. If \(A \subset X\) is a subset of a countable set \(X\) then \(A\) is countable.

**Proof.** We take each item in turn.

1. Put the elements of \(N \times N\) into an array of the form

\[
\begin{pmatrix}
(1,1) & (1,2) & (1,3) & \ldots \\
(2,1) & (2,2) & (2,3) & \ldots \\
(3,1) & (3,2) & (3,3) & \ldots \\
\vdots & \vdots & \vdots & \ddots
\end{pmatrix}
\]

and then “count” these elements by counting the sets \(\{(i,j) : i + j = k\}\) one at a time. For example let \(h(1) = (1, 1), h(2) = (2, 1), h(3) = (1, 2), h(4) = (3, 1), h(5) = (2, 2), h(6) = (1, 3)\) and so on. In other words we put \(N \times N\) into the following list form,

\[
N \times N = \{(1,1), (1,2), (1,3), (2,1), (2,2), (2,3), (3,1), (3,2), (3,3), \ldots \}.
\]

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

3. By assumption there exists surjective maps, \(f_n : N \rightarrow X_n\), for each \(n \in \mathbb{N}\). Let \(h(n) := (a(n), b(n))\) be the bijection constructed for item 1. Then \(f : N \rightarrow X\) defined by \(f(n) := f_{h(n)}(b(n))\) is a surjective map.

4. To see this let \(f : N \rightarrow X\) be a surjective map and let \(g(x) := \min f^{-1}\{\{x\}\}\) for all \(x \in X\). Then \(g : X \rightarrow N\) is an injective map. Let \(S := g(X)\), then \(g : X \rightarrow S \subset N\) is a bijection. So it remains to show \(S \sim N\) or \(S \sim J_n\) for some \(n \in \mathbb{N}\). If \(S\) is unbounded, then \(S \sim N\) as we have already seen. So it suffices to consider the case where \(S\) is bounded. If \(S\) is bounded by \(1\) then \(S = \{1\} = J_1\) and we are done. Now assume the result is true if \(S\) is bounded by \(n \in \mathbb{N}\) and now suppose that \(S\) is bounded by \(n + 1\). If \(n + 1 \notin S\), then \(S\) is bounded by \(n\) and by induction, \(S \sim J_k\) for some \(k \leq n < n + 1\). If \(n + 1 \in S\), then from above, \(S \setminus \{n + 1\} \sim J_k\) for some \(k \leq n\). Therefore there exists a bijection, \(f : J_k \rightarrow S \setminus \{n + 1\}\). We then extend \(f\) to \(J_{k+1}\) by setting \(f(k + 1) := n + 1\) which shows \(J_{k+1} \sim S\).
5. We again prove this by induction on \( n \). If \( n = 1 \), then \( S = \{m\} \) for some \( m \in \mathbb{N} \) which is bounded. So suppose for some \( n \in \mathbb{N} \), every subset \( S \subset \mathbb{N} \) with \( S \sim J_n \) is bounded. Now suppose that \( S \subset \mathbb{N} \) with \( S \sim J_{n+1} \). Then \( f(J_n) \sim J_n \) and hence \( f(J_n) \) is bounded in \( \mathbb{N} \). Then \( \max f(J_n) \cup \{f(n+1)\} \) is an upper bound for \( S \). This completes the inductive argument.

6. For each \( n \in \mathbb{N} \) there exists injection, \( f_n : J_n \to X \). By replacing \( X \) by \( X_n := \cup_{n \in \mathbb{N}} f_n(J_n) \) we may assume that \( X = \cup_{n \in \mathbb{N}} f_n(J_n) \). Thus there exists a surjective map, \( f : \mathbb{N} \to X \) by item 3. Let \( g : \mathbb{N} \to X \) be defined by \( g(n) := \min f^{-1}(\{x\}) \) for all \( x \in X \) and let \( S := g(X) \). To finish the proof we need only show that \( S \) is unbounded. If \( S \) were bounded, then we would find \( k \in \mathbb{N} \) such that \( J_k \sim S \). However this is impossible since \( \text{card}(J_n) \leq \text{card}(g(X)) \) for all \( n \) which implies \( g(X) \) is unbounded by Corollary 5.21. Therefore \( X \sim g(X) \sim \mathbb{N} \) by Lemma 5.26.

Lemma 5.28. If \( X \) is a countable set which contains \( Y \subset X \) with \( Y \sim \mathbb{N} \), then \( X \sim \mathbb{N} \).

Proof. By assumption there is an injective map, \( g : X \to \mathbb{N} \) and a bijective map, \( f : \mathbb{N} \to Y \). It then follows that \( g \circ f : \mathbb{N} \to \mathbb{N} \) is injective from which it follows that \( g(X) \) is unbounded. Indeed, \( (g \circ f)(J_n) \subset g(X) \) for all \( n \) implies \( \text{card}(J_n) \leq \text{card}(g(X)) \) for all \( n \) which implies \( g(X) \) is unbounded by Corollary 5.21. Therefore \( X \sim g(X) \sim \mathbb{N} \) by Lemma 5.26.

Corollary 5.29. We have \( \text{card}(\mathbb{Q}) = \text{card}(\mathbb{N}) \) and in fact, for any \( a < b \) in \( \mathbb{R} \), \( \text{card}(\mathbb{Q} \cap (a,b)) = \text{card}(\mathbb{N}) \).

Proof. First off \( \mathbb{Q} \) is a countable since \( \mathbb{Q} \) may be expressed as a countable union of countable sets;

\[
\mathbb{Q} = \bigcup_{m \in \mathbb{N}} \left\{ \frac{n}{m} : n \in \mathbb{Z} \right\}.
\]

From this it follows that \( \mathbb{Q} \cap (a,b) \) is countable for all \( a < b \) in \( \mathbb{R} \). As these sets are not finite, they must have the cardinality of \( \mathbb{N} \).

Theorem 5.30 (Uncountability results). 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. 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

\[
(f_1(n),f_2(n),f_3(n), \ldots).\]

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 \( Y_0 = \{0,1\} \). 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 \( i(A) := \{x_n : n \in \mathbb{N}\} \subset X \) for all \( A \subset \mathbb{N} \). 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 seen that \( 2^\mathbb{N} \) is uncountable.

Corollary 5.31. The set \( \{0,1\} : a < b \) in \( \mathbb{R} \), card \( (\mathbb{Q} \cap (a,b)) = \text{card}(\mathbb{N}) \) while card \( (\mathbb{Q} \cap (a,b)) > \text{card}(\mathbb{N}) \).

Proof. From Section 3.4 the set \( \{0,1,2, \ldots, 8\}^\mathbb{N} \) can be mapped injectively into \( (0,1) \) and therefore it follows from Theorem 5.30 that \( (0,1) \) is uncountable. For each \( m \in \mathbb{N} \), let \( A_m := \{\frac{n}{m} : n \in \mathbb{N} \text{ with } n < m\} \). Since \( \mathbb{Q} \cap (0,1) = \bigcup_{n=1}^\infty X_m \) and \( \#(X_m) < \infty \) for all \( m \), another application of Theorem 5.27 shows \( \mathbb{Q} \cap (0,1) \) is countable.

The fact that these results hold for any other finite interval follows from the fact that \( f(0,1) \to (a,b) \) defined by \( f(t) := a + t(b-a) \) is a bijection.

Definition 5.32. We say a non-empty set \( X \) is infinite if \( X \) is not a finite set.

Example 5.33. Any unbounded subset, \( S \subset \mathbb{N} \), is an infinite set according to Theorem 5.20.

Theorem 5.34. Let \( X \) be a non-empty set. The following are equivalent;

1. \( X \) is an infinite set,
2. \( \text{card}(J_n) \leq \text{card}(X) \) for all \( n \in \mathbb{N} \),
3. \( \text{card}(\mathbb{N}) \leq \text{card}(X) \),
4. \( \text{card}(X \setminus \{x\}) = \text{card}(X) \) for some (or all) \( x \in X \).

Proof. 1. \( \implies \) 2. Suppose that \( X \) is an infinite set. We show by induction that \( \text{card}(J_n) \leq \text{card}(X) \) for all \( n \in \mathbb{N} \). Since \( X \) is not empty, there exists \( x \in X \) and we may define \( f : J_1 \to X \) by \( f(1) = x \) in order to learn \( \text{card}(J_1) \leq \text{card}(X) \). Suppose we have shown \( \text{card}(J_n) \leq \text{card}(X) \) for some \( n \in \mathbb{N} \), i.e. there exists and injective map, \( f : J_n \to X \). If \( f(J_n) = X \) it would follow that \( \text{card}(X) = \text{card}(J_n) \) and would violated the assumption that \( X \) is not a finite set. Thus there exists \( x \in X \setminus f(J_n) \) and we may define \( f' : J_{n+1} \to X \).
Given a non-empty
Theorem 5.35 (Cardinality/Counting Summary I).
been proven above.

Exercise 5.9. Show that

5.4.1 Exercises

Exercise 5.10. Let \( \mathbb{Q}[t] \) be the set of polynomial functions, \( p \), such that \( p \) has rationale coefficients. That is \( p \in \mathbb{Q}[t] \) iff there exists \( n \in \mathbb{N}_0 \) and \( a_k \in \mathbb{Q} \) for \( 0 \leq k \leq n \) such that

\[
p(t) = \sum_{k=0}^{n} a_k t^k \quad \text{for all } t \in \mathbb{R}.
\]

Show \( \mathbb{Q}[t] \) is a countable set.

Definition 5.37 (Algebraic Numbers). A real number, \( x \in \mathbb{R} \), is called algebraic number, if there is a non-zero polynomial \( p \in \mathbb{Q}[t] \) such that \( p(x) = 0 \). [That is to say, \( x \in \mathbb{R} \) is algebraic if it is the root of a non-zero polynomial with coefficients from \( \mathbb{Q} \).]

Note that for all \( q \in \mathbb{Q} \), \( p(t) := t - q \) satisfies \( p(q) = 0 \). Hence all rational numbers are algebraic. But there are many more algebraic numbers, for example \( y^{1/n} \) is algebraic for all \( y \geq 0 \) and \( n \in \mathbb{N} \).

Exercise 5.11. Show that the set of algebraic numbers is countable. [Hint: any polynomial of degree \( n \) has at most \( n \) – real roots.] In particular, “most” irrational numbers are not algebraic numbers, i.e. there is still an uncountable number of non-algebraic numbers.
Normed and Metric Spaces
Metric Spaces

Definition 6.1. A function \( d : X \times X \to [0, \infty) \) is called a metric if
1. (Symmetry) \( d(x, y) = d(y, x) \) for all \( x, y \in X \).
2. (Non-degenerate) \( d(x, y) = 0 \) if and only if \( x = y \in X \), and
3. (Triangle inequality) \( d(x, z) \leq d(x, y) + d(y, z) \) for all \( x, y, z \in X \).

Example 6.2. Here are a few immediate examples of metric spaces;
1. Let \( X \) be any set and then define,
   \[
   d(x, y) = \begin{cases} 
   0 & \text{if } x = y \\
   1 & \text{if } x \neq y 
   \end{cases}
   \]
2. Let \( X = \mathbb{R} \) with \( d(x, y) := |y - x| \). Notice that
   \[
   d(x, z) = |x - y + y - z| 
   \leq |x - y| + |y - z| = d(x, y) + d(y, z).
   \]
3. Let \( X \) be any subset of \( \mathbb{C} \) and define \( d(w, z) := |z - w| \).

In general our typical example of a metric space will often be a generalization of the last example above, see Example 6.12.

6.1 Normed Spaces [Linear Algebra Meets Analysis]

6.1.1 Review of Vector Spaces and Subspaces

Definition 6.3 (Vector Space). A vector space is a non-empty set \( Z \) of objects, called vectors, equipped with an addition operation “+” and scalar (\( = \mathbb{R} \) or maybe \( \mathbb{C} \)) multiplication “\( \cdot \)” satisfying all of the properties above: i.e. For all \( u, v, w \in Z \) and \( a, b \in \mathbb{R} \):
1. Associativity of addition: \( u + (v + w) = (u + v) + w \).
2. Commutativity of addition: \( v + w = w + v \).
3. Identity element of addition: \( 0 + v = v \) for all \( v \).
4. Inverse elements of addition: \(-v + v = 0 \) for all \( v \in Z \). (In fact \(-v = (-1) \cdot v\).)
5. Distributivity of scalar multiplication with vector addition: \( a \cdot (v + w) = a \cdot v + a \cdot w \).
6. Distributivity of scalar multiplication with respect to field addition \( (a + b) \cdot v = a \cdot v + b \cdot v \).
7. Compatibility of scalar multiplication with the multiplication on \( \mathbb{R} \): \( a \cdot (b \cdot v) = (ab) \cdot v \).
8. Identity element of scalar multiplication \( 1 \cdot v = v \) for all \( v \in Z \).

Example 6.4. Here are two fundamental examples of vector spaces.
1. \( \mathbb{R} \) with usual vector addition and scalar multiplication is vector space over \( \mathbb{R} \).
2. \( \mathbb{C}^n \) with usual vector addition and scalar multiplication is vector space over \( \mathbb{C} \).

Notation 6.5 If \( T \) and \( X \) are sets, let \( X^T \) denote the collection of functions, \( f : T \to X \).

Example 6.6 (The Main Umbrella Example). Let \( T \) be a non-empty set and let \( Z := \mathbb{R}^T \). For \( f, g \in Z \) and \( \lambda \in \mathbb{R} \) we define \( f + g \) and \( \lambda \cdot f \) by
   \[
   (f + g)(t) = f(t) + g(t) \quad \text{(addition in } \mathbb{R}) \text{ for all } t \in T \\
   (\lambda \cdot f)(t) = \lambda f(t) \quad \text{(multiplication in } \mathbb{R}) \text{ for all } t \in T.
   \]

It can now be checked that \( Z \) is a vector space so that functions have now become vectors! Essentially all other examples of vector spaces we give will be related to an example of this form. The same observations show \( \mathbb{C}^T \) is a complex vector space.

Example 6.7. For example; \( \mathbb{R}^3 = \mathbb{R}^{\{1,2,3\}} \) and more generally \( \mathbb{R}^n = \mathbb{R}^{J_n} \) where \( J_n := \{1,2,\ldots,n\} \). In this setting we usually specify \( x \in \mathbb{R}^{J_n} \) by listing its values \( (x(1), \ldots, x(n)) \). To abbreviate notation a bit more we will usually write \( x(i) \) as \( x_i \) so that \( (x(1), \ldots, x(n)) \) becomes \( (x_1, \ldots, x_n) \).
Example 6.8. The vector space of $2 \times 2$ matrices:

$$M_{2 \times 2} = \{ A : A \text{ is a } 2 \times 2 \text{ matrix } \} = \{ A : \{(1,1), (1,2), (2,1), (2,2)\} \to \mathbb{R}\}.$$  

This can be generalized.

Definition 6.9 (Subspace). Let $Z$ be a vector space. A non-empty subset, $H \subset Z$, is a subspace of $Z$ if $H$ is closed under addition and scalar multiplication. Note, if $H$ is a subspace and $v \in H$, then $0 = 0 \cdot v \in H$.

The vector space $\mathbb{R}^T$ and $\mathbb{C}^T$ are typically the “largest” vector spaces we will consider in this course.

Example 6.10. Here are three common subspaces of $\mathbb{R}^n$:

1. $H = \{ f \in Z : f \text{ is continuous} \}$.
2. $H = \{ f \in Z : f \text{ is continuously differentiable} \}$.
3. $H = \{ f \in Z : f \text{ is differentiable at } \pi \}$.

6.1.2 Normed Spaces

Definition 6.11. 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) $\| f + g \| \leq \| f \| + \| g \|$ 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 or normed space for short.

Example 6.12. If $(Z, \| \cdot \|)$ is a normed space, then $d(x, y) := \| x - y \|$ is a metric on $Z$ and restricts to a metric on any subset of $Z$.

Example 6.13 (Normed Spaces). The following are normed spaces;

1. $Z = \mathbb{R}$ with $\| x \| = |x|$.
2. $Z = \mathbb{C}$ with $\| z \| = |z|$.
3. $Z = \mathbb{C}^n$ with

$$\| z \|_1 := \sum_{i=1}^{n} |z_i| \quad \text{for } z = (z_1, \ldots, z_n) \in Z.$$  

The triangle inequality is easily verified here since,

$$\| z + w \|_1 = \sum_{i=1}^{n} |z_i + w_i| \leq \sum_{i=1}^{n} (|z_i| + |w_i|) = \sum_{i=1}^{n} |z_i| + \sum_{i=1}^{n} |w_i| = \| z \|_1 + \| w \|_1.$$  

4. Let $X$ be a set and for any function $f : X \to \mathbb{C}$, let

$$\| f \|_u := \sup_{x \in X} |f(x)|.$$  

Then $Z := \{ f : X \to \mathbb{C} : \| f \|_u < \infty \}$ is a vector space and $\| \cdot \|_u$ is a norm on $Z$.

Exercise 6.1. Verify the last item of Example 6.13. That is let $X$ be a set and for any function $f : X \to \mathbb{C}$, let

$$\| f \|_u := \sum_{x \in X} |f(x)|.$$  

Show $Z := \{ f : X \to \mathbb{C} : \| f \|_u < \infty \}$ is a vector space and $\| \cdot \|_u$ is a norm on $Z$.

Our next goal is to show that $\| \cdot \|$ defined in Eq. (4.7) defines a norm on $\mathbb{R}^n$ and $\mathbb{C}^n$. We will begin by proving the important Cauchy-Schwarz inequality.

Lemma 6.14. If $x, y \geq 0$ and $\rho > 0$, then

$$xy \leq \frac{1}{2} \left( \rho x^2 + \frac{1}{\rho} y^2 \right)$$  

(6.1)

with equality when $\rho = y/x$ in the case $x > 0$.

Proof. Since

$$0 \leq \left( \sqrt{\rho x} - \frac{y}{\sqrt{\rho}} \right)^2 = \rho x^2 + \frac{1}{\rho} y^2 - 2xy,$$

with equality iff $\sqrt{\rho x} = \frac{y}{\sqrt{\rho}}$, i.e. iff $\rho = y/x$, we see that

$$xy \leq \frac{1}{2} \left( \rho x^2 + \frac{1}{\rho} y^2 \right)$$  

with equality iff $\rho = y/x$.  

\close
Definition 6.15. The two norm, \( \| \|_2 \) on \( \mathbb{C}^n \) is the function defined by

\[
\| z \|_2 = \sqrt{\sum_{i=1}^{n} |z_i|^2} \text{ for all } z \in \mathbb{C}^n.
\]

Theorem 6.16 (Cauchy-Schwarz Inequality). For \( a, b \in \mathbb{C}^n \), \( \langle a \cdot b \rangle \leq \| a \|_2 \cdot \| b \|_2 \).

Proof. The inequality holds true if \( a = 0 \) so we may now assume \( a \neq 0 \). Using Lemma 6.14 with \( x = |a_i| \) and \( y = |b_i| \) we find for any \( \rho > 0 \) that

\[
\| a \cdot b \| \leq \sum_{i=1}^{n} |a_i b_i| = \sum_{i=1}^{n} |a_i| \cdot |b_i| = \frac{1}{2} \left( \rho |a_i|^2 + \frac{1}{\rho} |b_i|^2 \right) \leq \frac{1}{2} \left( \rho \| a \|_2^2 + \frac{1}{\rho} \| b \|_2^2 \right).
\]

Taking \( \rho = \| b \|_2 / \| a \|_2 \) then completes the proof. ■

Theorem 6.17 (Triangle Inequality). The function, \( \| \|_2 \), on \( \mathbb{C}^n \) is a norm and in particular for \( a, b \in \mathbb{C}^n \),

\[
\| a + b \|_2 \leq \| a \|_2 + \| b \|_2.
\]

Proof. The main point is to prove the triangle inequality. The proof is as follows;

\[
\| a + b \|_2^2 = \langle a + b \rangle \cdot \overline{a + b} = \langle a + b \rangle \cdot (\overline{a} + \overline{b})
\]

\[
= \langle a + b \rangle \cdot \overline{a} + (a + b) \cdot \overline{b} = \| a \|_2^2 + b \cdot \overline{a} + a \cdot \overline{b} + \| b \|_2^2
\]

\[
= \| a \|_2^2 + \| b \|_2^2 + 2 \Re (a \cdot \overline{b}) \leq \| a \|_2^2 + \| b \|_2^2 + 2 \left| \Re (a \cdot \overline{b}) \right|
\]

\[
\leq \| a \|_2^2 + \| b \|_2^2 + 2 \| a \|_2 \cdot \| b \|_2 \quad \text{(Theorem 6.16)}
\]

\[
= \| a \|_2^2 + \| b \|_2^2 + 2 \| a \|_2 \cdot \| b \|_2 = (\| a \|_2 + \| b \|_2)^2.
\]

The remaining properties of a norm are easily checked. For example, if \( \lambda \in \mathbb{C} \) and \( a \in \mathbb{C}^n \), then

\[
\| \lambda a \|_2 = \sqrt{\sum_{i=1}^{n} |\lambda a_i|^2} = \sqrt{\sum_{i=1}^{n} |\lambda|^2 |a_i|^2}
\]

\[
= \sqrt{|\lambda|^2 \sum_{i=1}^{n} |a_i|^2} = |\lambda| \sqrt{\sum_{i=1}^{n} |a_i|^2} = |\lambda| \| a \|_2.
\]

Fact 6.18 For \( 0 < p < \infty \), let \( \| \|_p : \mathbb{C}^n \to [0, \infty) \) be defined by

\[
\| z \|_p = \sqrt[p]{\sum_{i=1}^{n} |z_i|^p} \text{ for all } z \in \mathbb{C}^n.
\]

Then \( \| \|_p \) is a norm on \( \mathbb{C}^n \) for all \( 1 \leq p < \infty \) but is not a norm (the triangle inequality fails) for \( 0 < p < 1 \).

Definition 6.19 (Balls). Let \( (X, d) \) be a metric space. For \( x \in X \) and \( r \geq 0 \) let

\[
B_x (r) := \{ y \in X : d (x, y) < r \} \quad \text{and} \quad C_x (r) := \{ y \in X : d (x, y) \leq r \}.
\]

We refer to \( B_x (r) \) and \( C_x (r) \) as the open and closed ball respectively about \( x \) with radius \( r \).

Example 6.20. Let us consider \( \| \|_p \) to \( \mathbb{R}^2 \). Figure 6.1 shows the boundary of the balls of radius 1 centered at 0 for some of the different \( p - \) norms.

![Fig. 6.1. Balls in \( \mathbb{R}^2 \) corresponding to \( p \in \{ \frac{1}{2}, 1, 2, 5, 20 \} \). The case \( p = \frac{1}{2} \) is not convex like the other balls which is an indication that the triangle inequality fails.](image)

Example 6.21. The next figures explain how to understand balls in \( C ([0, 1], \mathbb{R}) \) equipped with the uniform norm.
48  6 Metric Spaces

Fig. 6.2. The ball in $C([0, 1], \mathbb{R})$ of radius $1/4$ centered at $f(x) = \sin(x^2)$ are all the continuous functions whose graphs lie between the green envelope.

Exercise 6.2 (Weighted $2 -$ norms). Suppose that $\rho_i \in (0, \infty)$ for $1 \leq i \leq n$ and for $a, b \in \mathbb{C}^n$ let

$$a \ast b := \sum_{i=1}^{n} a_i b_i \rho_i \quad \text{and} \quad \|a\| := \sqrt{a \ast a} = \sqrt{\sum_{i=1}^{n} |a_i|^2 \rho_i}.$$  

Show, for all $a, b \in \mathbb{C}^n$ that $|a \ast b| \leq \|a\| \cdot \|b\|$ and that $\|\cdot\|$ is a norm on $\mathbb{C}^n$. [Hint: reduce to the case where $\rho_i = 1$ for all $i$.]

For the next two exercise you will be using some concepts from calculus which we will developed in detail next quarter. For now, I assume you know what the Riemann integral is for continuous functions on $[0, 1]$ with values in $\mathbb{R}$. Let $Z$ denote the continuous functions on $[0, 1]$ with values in $\mathbb{R}$. The only properties that you need to know about the Riemann integral are:

1. The integral is linear, namely for all $f, g \in Z$ and $\lambda \in \mathbb{R}$,

$$\int_{0}^{1} (f(t) + \lambda g(t)) dt = \int_{0}^{1} f(t) dt + \lambda \int_{0}^{1} g(t) dt.$$  

2. If $f, g \in Z$ and $f(t) \leq g(t)$ for all $t \in [0, 1]$, then

$$\int_{0}^{1} f(t) dt \leq \int_{0}^{1} g(t) dt.$$  

3. For all $f \in Z$,

$$\left| \int_{0}^{1} f(t) dt \right| \leq \int_{0}^{1} |f(t)| dt.$$  

In fact this item follows from items 1. and 2. Indeed, since $\pm f(t) \leq |f(t)|$ for all $t$, we find

$$\pm \int_{0}^{1} f(t) dt = \int_{0}^{1} \pm f(t) dt \leq \int_{0}^{1} |f(t)| dt \iff \left| \int_{0}^{1} f(t) dt \right| \leq \int_{0}^{1} |f(t)| dt.$$  

$^1$ The notion of continuity will be formally developed shortly.
Exercise 6.3. Let $Z$ denote the continuous function on $[0,1]$ with values in $\mathbb{R}$ and for $f \in Z$ let

$$\|f\|_1 := \int_0^1 |f(t)| \, dt.$$  

Show $\|\cdot\|_1$ satisfies,

1. (Homogeneity) $\|\lambda f\| = |\lambda| \|f\|$ for all $\lambda \in \mathbb{R}$ and $f \in Z$.
2. (Triangle inequality) $\|f + g\| \leq \|f\| + \|g\|$ for all $f, g \in Z$.

Remark 6.22 (An interpretation of $\|\cdot\|_1$). If we interpret $f(t)$ as the speed of a particle on the real line at time $t$, then $\|f\|_1$ represents the total distance (including retracing of its path) the particle travels over the time interval $[0,1]$.

Exercise 6.4. Let $Z$ denote the continuous function on $[0,1]$ with values in $\mathbb{R}$ and for $f \in Z$ let

$$\|f\|_2 := \sqrt{\int_0^1 |f(t)|^2 \, dt}.$$  

Show;

1. for $f, g \in Z$ that

$$\left| \int_0^1 f(t) g(t) \, dt \right| \leq \|f\|_2 \cdot \|g\|_2,$$

2. Homogeneity ($\|\lambda f\|_2 = |\lambda| \|f\|_2$ for all $\lambda \in \mathbb{R}$ and $f \in Z$), and
3. (Triangle inequality) $\|f + g\|_2 \leq \|f\|_2 + \|g\|_2$ for all $f, g \in Z$.

Remark 6.23 (An interpretation of $\|\cdot\|_2$). Let us suppose that $f(t)$ is the voltage across a 1 Ohm resistor. By Ohm’s law the current through this resistor is $f(t)/1 = f(t)$ and the power dissipated by the resistor at time $t$ is \textit{(Voltage-Current)} is $f(t)^2$. The work done over the time interval, $[0,1]$ is then

$$\int_0^1 \text{Power}(t) \, dt = \int_0^1 f^2(t) \, dt = \|f\|_2^2.$$  

On the other hand if we had a constant voltage of $\|f\|_2$ across the resistor over the time interval $[0,1]$, the work done over this period would again be $\|f\|_2^2$. Thus $\|f\|_2$ is often referred to as the RMS voltage (root mean squared voltage) and represents the equivalent DC (Direct Current, i.e. constant) voltage necessary to produce the same amount of work over the time interval $[0,1]$.

Exercise 6.5. Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$ such that $\|a\| \leq b$ whenever $0 \leq a_i \leq b$ for $1 \leq i \leq n$. Further suppose that $(X_i, d_i)$ for $i = 1, \ldots, n$ is a finite collection of metric spaces and for $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ in $X := \prod_{i=1}^n X_i$, let

$$d(x, y) = \| (d_1(x_1, y_1), d_2(x_2, y_2), \ldots, d_n(x_n, y_n)) \|.$$  

Show $(X, d)$ is a metric space.

- End of Lecture 13, 10/28/2012. [We started Section 5.1 above as well this day.]

6.2 Sequences in Metric Spaces

Definition 6.24. A sequence $\{x_n\}_{n=1}^\infty$ in a metric space $(X, d)$ is said to be convergent if there exists a point $x \in X$ such that $\lim_{n \to \infty} d(x, x_n) = 0$. In this case we write $\lim_{n \to \infty} x_n = x$ or $x_n \to x$ as $n \to \infty$.

Exercise 6.6. Show that $x$ in Definition 6.24 is necessarily unique.

Definition 6.25 (Cauchy sequences). A sequence $\{x_n\}_{n=1}^\infty$ in a metric space $(X, d)$ is Cauchy provided that $\lim_{m,n \to \infty} d(x_n, x_m) = 0$, i.e. for all $\varepsilon > 0$ there exists $N(\varepsilon) \in \mathbb{N}$ such that

$$d(x_n, x_m) \leq \varepsilon$$

when $m, n \geq N(\varepsilon)$.

- End of Lecture 15, 11/2/2012.

Exercise 6.7. Show that convergent sequences in metric spaces are always Cauchy sequences. The converse is not always true. For example, let $X = \mathbb{Q}$ be the set of rational numbers and $d(x, y) = |x - y|$. Choose a sequence $\{x_n\}_{n=1}^\infty \subset \mathbb{Q}$ which converges to $\sqrt{2} \in \mathbb{R}$, then $\{x_n\}_{n=1}^\infty$ is (Q, d) – Cauchy but not (Q, d) – convergent. Of course the sequence is convergent in $\mathbb{R}$.

Exercise 6.8. If $\{x_n\}_{n=1}^\infty$ is a Cauchy sequence in a metric space $(X, d)$, $\lim_{n \to \infty} d(x_n, y)$ exists in $\mathbb{R}$ for all $y \in X$. In particular, $\{d(x_n, y)\}_{n=1}^\infty$ is a bounded sequence in $\mathbb{R}$ for all $y \in X$.

Definition 6.26. A metric space $(X, d)$ (or normed space $(X, \|\cdot\|)$) is complete if all Cauchy sequences are convergent sequences. A complete normed space is called a Banach space.

Lemma 6.27. Let $X$ be a non-empty set and

$$\|f\|_u := \sup_{x \in X} |f(x)|$$

for all $f \in \mathbb{C}^X$.

Then the subspace, $Z := \{f \in \mathbb{C}^X : \|f\|_u < \infty\}$ is a Banach space, i.e. $(Z, \|\cdot\|_u)$ is a complete normed space.
Proof. Let \( \{ f_n \}_{n=1}^{\infty} \subset Z \) be a Cauchy sequence. Since for any \( x \in X \), we have
\[
|f_n(x) - f_m(x)| \leq \|f_n - f_m\|_u \tag{6.2}
\]
which shows that \( \{ f_n(x) \}_{n=1}^{\infty} \subset \mathbb{C} \) is a Cauchy sequence of complex numbers. Because \( \mathbb{C} \) is complete, \( f(x) := \lim_{n \to \infty} f_n(x) \) exists for all \( x \in X \). Passing to the limit \( n \to \infty \) in Eq. (6.3) implies
\[
|f(x) - f_m(x)| \leq \lim_{n \to \infty} \|f_n - f_m\|_u
\]
and taking the supremum over \( x \in X \) of this inequality implies
\[
\|f - f_m\|_u \leq \lim_{n \to \infty} \|f_n - f_m\|_u \to 0 \quad \text{as} \quad m \to \infty
\]
showing \( f_m \to f \) in \( Z \).
\[
\text{Definition 6.28. We say that two norms, } \|\cdot\|_a \text{ and } \|\cdot\|_b \text{, on a vector space } X \text{ are equivalent if there are constants } C_1, C_2 \in (0, \infty) \text{ such that}
\]
\[
\|x\|_a \leq C_1 \|x\|_b \text{ and } \|x\|_b \leq C_2 \|x\|_a \text{ for all } x \in X.
\]

Similarly two metrics, \( d_a \) and \( d_b \) on a set \( X \) are said to be equivalent if there are constants \( C_1, C_2 \in (0, \infty) \) such that
\[
d_a(x, y) \leq C_1 d_b(x, y) \quad \text{and} \quad d_b(x, y) \leq C_2 d_a(x, y) \text{ for all } x, y \in X.
\]

Exercise 6.9. Show that two norms, \( \|\cdot\|_a \) and \( \|\cdot\|_b \) on a vector space \( X \) are equivalent iff the corresponding metrics, \( d_a(x, y) := \|y - x\|_a \) and \( d_b(x, y) := \|y - x\|_b \), on \( X \) are equivalent metrics.

Corollary 6.29. If \( d_a \) and \( d_b \) are two equivalent metrics on a set \( X \) then \( (X, d_a) \) is a complete metric space iff \( (X, d_b) \) is a complete metric space.

Proof. Suppose that \( (X, d_b) \) is complete. If \( \{x_n\}_{n=1}^{\infty} \) is \( d_a - \) Cauchy implies
\[
d_b(x_n, x_m) \leq C_2 d_a(x_n, x_m) \to 0 \quad \text{as} \quad m, n \to \infty
\]
which shows that \( \{x_n\}_{n=1}^{\infty} \) is \( d_b - \) Cauchy. As \( (X, d_b) \) is complete, there exists \( x \in X \) such that \( d_b(x_n, x) \to 0 \) as \( n \to \infty \). Since
\[
d_a(x_n, x_n) \leq C_1 d_b(x_n, x_n) \to 0 \quad \text{as} \quad n \to \infty
\]
we see that \( x_n \to x \) in the \( d_a \) – metric as well. This shows \( (X, d_a) \) is complete. The reverse implication is proved the same way.

Exercise 6.10 (Equivalence of 3 norms on \( \mathbb{C}^n \)). Let \( \|\cdot\|_1, \|\cdot\|_u \), and \( \|\cdot\|_2 \) be the three norms on \( \mathbb{C}^n \) given above. Show for all \( z \in \mathbb{C}^n \) that
\[
\|z\|_2 \leq \sqrt{n} \|z\|_1, \tag{Hint: use Cauchy Schwarz.}
\]
\[
\|z\|_2 \leq \sqrt{n} \|z\|_u.
\]
It follows from these inequalities that \( \|\cdot\|_1, \|\cdot\|_u \), and \( \|\cdot\|_2 \) are equivalent norms on \( \mathbb{C}^n \).

Theorem 6.30 (Completeness of \( \mathbb{C}^n \)). Let \( n \in \mathbb{N} \) and \( \|\cdot\| \) denote any one of the norms, \( \|\cdot\|_1, \|\cdot\|_2 \), or \( \|\cdot\|_u \) on \( \mathbb{C}^n \). Then \( \mathbb{C}^n, \|\cdot\| \) is complete.

Proof. By Exercise 6.10 all of these norms are equivalent to \( \|\cdot\|_u \) and hence it suffices to show that \( \|\cdot\|_u \) is a complete norm on \( \mathbb{C}^n \). This is a special case of Lemma 6.27 with \( X = \{1, 2, \ldots, n\} \).

Exercise 6.11. Let \( X \) be a set and \( (Y, \rho) \) be a complete metric space. Suppose that \( f_n : X \to Y \) are functions such that
\[
\delta_{m,n} := \sup_{x \in X} d(f_n(x), f_m(x)) \to 0 \quad \text{as} \quad m, n \to \infty.
\]
Show there exists a (unique) functions, \( f : X \to Y \) such that
\[
\lim_{n \to \infty} \sup_{x \in X} d(f_n(x), f(x)) = 0.
\]

Hint: mimic the proof of Lemma 6.27.

Exercise 6.12. Let \( Z \) denote the continuous functions on \( [0, 1] \) with values in \( \mathbb{R} \) and as above let
\[
\|f\|_1 := \int_0^1 |f(t)| \, dt, \quad \|f\|_2 := \sqrt{\int_0^1 |f(t)|^2 \, dt}, \quad \text{and} \quad \|f\|_u = \sup_{0 \leq t \leq 1} |f(t)|.
\]

Show for all \( f \in Z \) that;
\[
\|f\|_1 \leq \|f\|_2 \quad \text{and} \quad \|f\|_2 \leq \|f\|_u.
\]

[Hint: for the first inequality use Cauchy Schwarz.] Also show there is no constant \( C < \infty \) such that
\[
\|f\|_u \leq C \|f\|_2 \quad \text{for all} \quad f \in Z.
\]

[Hint: consider the sequence, \( f_n(t) = t^n \).]
Example 6.31. Let $Z = C([0,1],\mathbb{R})$ be the vector space of continuous functions on $[0,1]$ with values in $\mathbb{R}$ and for $f \in Z$ let

$$\|f\|_1 := \int_0^1 |f(t)| \, dt.$$ 

Let us show that $(Z,\|\cdot\|_1)$ is not complete. To this end let

$$g(t) := \begin{cases} 2t & \text{if } t \leq 1/2 \\ 1 & \text{if } 1/2 \leq t \leq 1 \end{cases}$$

and then set $f_n(t) = g(t)^n$ or all $t \in [0,1]$. Then

$$\lim_{n \to \infty} f_n(t) = h(t) = \begin{cases} 0 & \text{if } t < \frac{1}{2} \\ 1 & \text{if } t \geq \frac{1}{2} \end{cases}$$

which is discontinuous at $1/2$. Let us now observe that $\|h\|_1 = \frac{1}{2}$ and

$$\|f_n - h\|_1 = \int_0^{1/2} (2t)^n \, dt = \frac{1}{2} \left( \frac{1}{n+1} \right)$$

so that $f_n \to h$ in $\|\cdot\|_1$. If there were some $f \in Z$ so that $\|f - f_n\|_1 \to 0$ we would have to have $\|f - h\|_1 = 0$. If $\varepsilon := |f(t_0) - h(t_0)|$ for some $t_0 \neq \frac{1}{2}$, by continuity we would have $|f(t) - h(t)| \geq \varepsilon/2$ for $t$ near $t_0$ from which it would follow that $\|f - h\|_1 > 0$. Therefore we must have $f(t) = h(t)$ for all $t \neq t_0$. Since

$$\lim_{t \uparrow \frac{1}{2}} f(t) = \lim_{t \downarrow \frac{1}{2}} f(t) = \lim_{t \to \frac{1}{2}} h(t) = 0$$

there is not choice for $f(1/2)$ for which $f$ would be continuous at $1/2$. Hence we have shown that $f_n$ can not converge to any element, $f \in Z$.

In fact, $(Z,\|\cdot\|_1)$ is full of uncountably many “holes” and $h$ is the location of just one of these holes. In the third quarter of this course we will fill these holes.

6.3 General Limits and Continuity in Metric Spaces

Suppose now that $(X,\rho)$ and $(Y,d)$ are two metric spaces and $f : X \to Y$ is a function.

Definition 6.32 (Limits of functions). If $x_0 \in X$ and $f : X \setminus \{x_0\} \to Y$ is a function, then we say $\lim_{x \to x_0} f(x) = y_0 \in Y$ if for all $\varepsilon > 0$ there is a $\delta = \delta(\varepsilon,x_0) > 0$ such that

$$d(f(x),y_0) \leq \varepsilon$$

provided that $0 < \rho(x,x_0) \leq \delta(\varepsilon,x_0)$.

[In generally when we write $\lim_{x \to x_0} f(x)$ we do not need to assume that $f(x_0)$ is defined.]

Theorem 6.33 (Computing Limits Using Sequences). If $x_0 \in X$ and $f : X \setminus \{x_0\} \to Y$ is a function as above, then $\lim_{x \to x_0} f(x) = y_0 \in Y$ iff $\lim_{n \to \infty} f(x_n) = y_0$ for all sequences $(x_n)_{n=1}^{\infty} \subset X \setminus \{x_0\}$ such that $\lim_{n \to \infty} x_n = x_0$. 

We say \( f \) is continuous at \( x \) if for all \( \varepsilon > 0 \) there is a \( \delta > 0 \) such that \( d(x, x') < \delta \) implies \( d(f(x), f(x')) < \varepsilon \) provided that \( \rho(x, x') < \frac{\varepsilon}{\delta} \). We then have \( \lim_{n \to \infty} x_n = x \) while \( \lim_{n \to \infty} f(x_n) \neq y_0 \).

**Definition 6.34 (Continuity).** A function \( f : X \to Y \) is continuous at \( x \in X \) if for all \( \varepsilon > 0 \) there is a \( \delta > 0 \) such that
\[
d(f(x), f(x')) < \varepsilon \quad \text{provided that} \quad \rho(x, x') < \delta.
\] (6.5)
The function \( f \) is said to be *continuous* if \( f \) is continuous at all points \( x \in X \). We will write \( C(X,Y) \) for the collection of continuous functions from \( X \) to \( Y \).

**Definition 6.35 (Sequential Continuity).** A function \( f : X \to Y \) is continuous at \( x \in X \) if \( \lim_{n \to \infty} f(x_n) = f(x) \) for all \( \{x_n\}_{n=1}^{\infty} \subset X \) with \( \lim_{n \to \infty} x_n = x \).

We say \( f \) is sequentially continuous on \( X \) if it is continuous at all points in \( X \).

**Corollary 6.36.** Continuity and sequential continuity are the same notions.

**Proof.** This follows rather directly from Theorem 6.33. \( \blacksquare \)

**Example 6.37.** The functions \( f : \mathbb{C} \setminus \{0\} \to \mathbb{C} \) defined by \( f(z) = 1/z \) is continuous. Indeed, if \( \{z_n\}_{n=1}^{\infty} \subset \mathbb{C} \setminus \{0\} \) and \( \lim_{n \to \infty} z_n = z \in \mathbb{C} \setminus \{0\} \), then
\[
\lim_{n \to \infty} f(z_n) = \lim_{n \to \infty} \frac{1}{z_n} = \frac{1}{z} = f(z).
\]

**Example 6.38.** Let \( f : \mathbb{R} \to \mathbb{R} \) be the function defined by
\[
f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin \mathbb{Q}. \end{cases}
\]
The function \( f \) is discontinuous at all points in \( \mathbb{R} \). For example, if \( x_0 \in \mathbb{Q} \) we may choose \( x_n \in \mathbb{R} \setminus \mathbb{Q} \) such that \( \lim_{n \to \infty} x_n = x_0 \) while
\[
\lim_{n \to \infty} f(x_n) = \lim_{n \to \infty} 0 = 0 \neq 1 = f(x_0).
\]
Similarly, if \( x_0 \in \mathbb{R} \setminus \mathbb{Q} \) we may choose \( x_n \in \mathbb{Q} \) such that \( \lim_{n \to \infty} x_n = x_0 \) while
\[
\lim_{n \to \infty} f(x_n) = \lim_{n \to \infty} 1 = 1 \neq 0 = f(x_0).
\]

**Exercise 6.13.** Consider \( \mathbb{N} \) as a metric space with \( d(m,n) := |m-n| \) and suppose that \( (Y,d) \) is a metric space. Show that every function, \( f : \mathbb{N} \to Y \) is continuous.

**Exercise 6.14.** Suppose that \( (X,d) \) is a metric space and \( f,g : X \to \mathbb{C} \) are two continuous functions on \( X \). Show:
1. \( f + g \) is continuous,
2. \( f \cdot g \) is continuous,
3. \( f/g \) is continuous provided \( g(x) \neq 0 \) for all \( x \in X \).

**Exercise 6.15.** Show the following functions from \( \mathbb{C} \) to \( \mathbb{C} \) are continuous.
1. \( f(z) = c \) for all \( z \in \mathbb{C} \) where \( c \in \mathbb{C} \) is a constant.
2. \( f(z) = |z| \).
3. \( f(z) = z \) and \( f(z) = \bar{z} \).
4. \( f(z) = \text{Re} z \) and \( f(z) = \text{Im} z \).
5. \( f(z) = \sum_{m,n=0}^{N} a_{m,n} z^m \bar{z}^n \) where \( a_{m,n} \in \mathbb{C} \).

**Exercise 6.16.** Suppose now that \( (X,\rho) \), \( (Y,d) \), and \( (Z,\delta) \) are three metric spaces and \( f : X \to Y \) and \( g : Y \to Z \). Let \( x \in X \) and \( y = f(x) \in Y \), show \( g \circ f : X \to Z \) is continuous at \( x \) if \( f \) is continuous at \( x \) and \( g \) is continuous at \( y \). Recall that \( (g \circ f)(x) := g(f(x)) \) for all \( x \in X \). In particular this implies that if \( f \) is continuous on \( X \) and \( g \) is continuous on \( Y \) then \( f \circ g \) is continuous on \( X \).
Example 6.39. If $f : X \to \mathbb{C}$ is a continuous function then $|f|$ is continuous and

$$F := \sum_{m,n=0}^{N} a_{mn} f^m \cdot \bar{f}^n$$

is continuous.

- End of Lecture 16, 11/5/2012.

Definition 6.40 (One sided limits). Suppose $(Y,d)$ is a metric space, $-\infty < a < b < \infty$, and $f : (a,b) \to Y$ is a function. For $x_0 \in (a,b)$ we say

$$\lim_{x \to x_0^+} f(x) = y_0 \iff \forall \varepsilon > 0 \exists \delta = \delta(\varepsilon) > 0 \exists d(f(x),y_0) \leq \varepsilon \text{ if } 0 < x - x_0 \leq \delta$$

and

$$\lim_{x \to x_0^-} f(x) = y_0 \iff \forall \varepsilon > 0 \exists \delta = \delta(\varepsilon) > 0 \exists d(f(x),y_0) \leq \varepsilon \text{ if } 0 < x_0 - x \leq \delta.$$

Theorem 6.41 (One sided limit criteria). Suppose that $(Y,d)$ is a metric space, $(a,b) \subset \mathbb{R}$, $f : (a,b) \to Y$ is a function, and $x_0 \in (a,b)$. Then the following are equivalent:

1. $\lim_{x \to x_0^+} f(x) = y_0$
2. $\lim_{n \to \infty} f(x_n) = y_0$ for all $\{x_n\}_{n=1}^{\infty} \subset (a,b)$ such that $\lim_{n \to \infty} x_n = x_0$.
3. $\lim_{n \to \infty} f(x_n) = y_0$ for all $\{x_n\}_{n=1}^{\infty} \subset (a,b)$ such that $x_n \downarrow x_0$, i.e. $x_{n+1} \leq x_n$ and $x_0 < x_n$ for all $n$ and $\lim_{n \to \infty} x_n = x_0$.

We also have the following equivalent statements:

- $a. \lim_{x \to x_0^+} f(x) = y_0$
- $b. \lim_{n \to \infty} f(x_n) = y_0$ for all $\{x_n\}_{n=1}^{\infty} \subset (a,b)$ such that $\lim_{n \to \infty} x_n = x_0$.
- $c. \lim_{n \to \infty} f(x_n) = y_0$ for all $\{x_n\}_{n=1}^{\infty} \subset (a,b)$ such that $x_n \uparrow x_0$, i.e. $x_{n+1} \geq x_n$ and $x_0 < x_n$ for all $n$ and $\lim_{n \to \infty} x_n = x_0$.

Moreover, $\lim_{x \to x_0} f(x) = y_0$ iff $\lim_{x \to x_0^+} f(x) = y_0 = \lim_{x \to x_0^-} f(x)$.

Proof. (1. $\Rightarrow$ 2.) If $\lim_{x \to x_0^+} f(x) = y_0$ then for all $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that $d(f(x),y_0) \leq \varepsilon$ if $0 < x - x_0 \leq \delta$. Hence if $\{x_n\}_{n=1}^{\infty} \subset (a,b)$ such that $x_n \to x_0$ then $0 < x_n - x_0 \leq \delta$ for a.a. $n$ and hence $d(f(x_n),y_0) \leq \varepsilon$ for a.a. $n$. Since $\varepsilon > 0$ was arbitrary, it follows that $\lim_{n \to \infty} d(f(x_n),y_0) = 0$, i.e. $\lim_{n \to \infty} f(x_n) = y_0$. It is trivial that (2. $\Rightarrow$ 3.).

(3. $\Rightarrow$ 1.) If $\lim_{x \to x_0^-} f(x) \neq y_0$ then there exists $\varepsilon > 0$ such that for all $\delta = \frac{1}{n} > 0$ there exists $x_n \in (x_0,b)$ such that $0 \leq x_n - x_0 < \frac{1}{n}$ while $d(f(x_n),y_0) \geq \varepsilon$. Then take $x_n = \min(x_n^1,\ldots,x_n^n)$. Then $x_n \downarrow x$ while $d(f(x_n),y_0) \geq \varepsilon$. The limit $\lim_{n \to \infty} f(x_n) = y_0$ is not continuous.

For the last statement it is clear that $\lim_{x \to x_0} f(x) = y_0$ implies $\lim_{x \to x_0^-} f(x) = y_0 = \lim_{x \to x_0^+} f(x)$. For the converse assertion, suppose that $\varepsilon > 0$ is given, then choose $\delta = \delta(\varepsilon) > 0$ such that

$$d(f(x),y_0) \leq \varepsilon \text{ when } 0 < x - x_0 \leq \delta^+ \varepsilon \text{ or } 0 < x_0 - x \leq \delta^- \varepsilon.$$

If we take $\delta(\varepsilon) = \min(\delta^+,\delta^-)(\varepsilon)$, then we will have

$$d(f(x),y_0) \leq \varepsilon \text{ when } 0 < |x - x_0| \leq \delta(\varepsilon)$$

and since $\varepsilon > 0$ was arbitrary, it follows that $\lim_{x \to x_0} f(x) = y_0$.

Corollary 6.42 (A monotone continuity criteria). Suppose that $(Y,d)$ is a metric space, $X = (a,b) \subset \mathbb{R}$, and $f : X \to Y$ is a function. Then $f$ is continuous at $x \in X$ iff $\lim_{n \to \infty} f(x_n) = f(x)$ whenever $\{x_n\}_{n=1}^{\infty} \subset X$ converges monotonically to $x$ as $n \to \infty$. In other words, it is sufficient to check sequential continuity along sequences which are either increasing or decreasing.

Proof. This is a direct consequence of Theorem 6.41.

Exercise 6.17 (Continuity of $x^{1/m}$). Show for each $m \in \mathbb{N}$ that the function $f(x) := x^{1/m}$ is continuous on $[0,\infty)$.

Exercise 6.18 (Differentiability of $x^{1/m}$). Show for each $m \in \mathbb{N}$ that the function $f(x) := x^{1/m}$ is differentiable on $(0,\infty)$ and that

$$\frac{d}{dx} x^{1/m} := \lim_{y \to x} \frac{y^{1/m} - x^{1/m}}{y - x} = \frac{1}{m} x^{1/m-1}.$$

Exercise 6.19 (Intermediate value theorem). Suppose that $-\infty < a < b < \infty$ and $f : [a,b] \to \mathbb{R}$ is a continuous function such that $f(a) \leq f(b)$. Show for any $y \in [f(a),f(b)]$, there exists a $c \in [a,b]$ such that $f(c) = y$.

Hint: Let $S := \{t \in [a,b] : f(t) \leq y\}$ and let $c := \sup(S)$.

Exercise 6.20 (Inverse Function Theorem I). Let $f : [a,b] \to [c,d]$ be a strictly increasing (i.e. $f(x_1) < f(x_2)$ whenever $x_1 < x_2$) continuous function such that $f(a) = c$ and $f(b) = d$. Then $f$ is bijective and the inverse function, $g := f^{-1} : [c,d] \to [a,b]$, is strictly increasing and is continuous.

Notations 6.43 Let $(X,\rho)$ and $(Y,d)$ be metric spaces and $f : X \to Y$ be a function.

3 The same result holds for $y \in [f(b),f(a)]$ if $f(b) \leq f(a)$, just replace $f$ by $-f$ in this case.
1. We say $f$ is uniformly continuous, iff for all $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that

$$\forall x, x' \in X \text{ with } \rho(x, x') \leq \delta \implies d(f(x), f(x')) \leq \varepsilon.$$ 

2. A function, $f : X \to Y$, is said to be Lipschitz if there is a constant $C < \infty$ such that

$$d(f(x), f(x')) \leq C \rho(x, x') \text{ for all } x, x' \in X.$$

Recall that a function $f : X \to Y$ is continuous at $x_0 \in X$ if for all $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon, x_0) > 0$ such that

$$\forall x \in X \text{ with } \rho(x, x_0) \leq \delta \implies d(f(x), f(x_0)) \leq \varepsilon.$$ 

Thus we see that a function is uniformly provided we can take $\delta(\varepsilon, x_0) > 0$ to be independent of $x_0$. If $f$ is Lipschitz and $\varepsilon > 0$, we may take $\delta := \varepsilon/C$ in order to see that if

$$\rho(x, x') \leq \delta \implies d(f(x), f(x')) \leq C \rho(x, x') \leq C \delta = \varepsilon$$ 

which shows $f$ is uniformly continuous.

**Example 6.44**. Any function, $f : \mathbb{R} \to \mathbb{R}$ which is everywhere differentiable is Lipschitz iff $K := \sup_{t \in \mathbb{R}} |f'(t)| < \infty$. Indeed if

$$|f(y) - f(x)| \leq K |y - x| \text{ for all } x, y \in \mathbb{R}$$

then

$$|f'(x)| = \lim_{y \to x} \frac{|f(y) - f(x)|}{|y - x|} \leq K \text{ for all } x \in \mathbb{R}.$$ 

Conversely, if $K := \sup_{t \in \mathbb{R}} |f'(t)| < \infty$, then by the mean value theorem, for all $y > x$ there exists $c \in (x, y)$ such that

$$\frac{|f(y) - f(x)|}{|y - x|} = |f'(c)| \leq K.$$

It turns out that every metric spaces with an infinite number of elements comes equipped with a large collection of Lipschitz functions.

**Lemma 6.45 (Distance to a Set)**. For any non empty subset $A \subset X$, let

$$d_A(x) := \inf\{d(x, a) | a \in A\},$$

then

$$|d_A(x) - d_A(y)| \leq d(x, y) \forall x, y \in X. \tag{6.6}$$

In particular, $d_A : X \to [0, \infty)$ is continuous.

**Proof.** Let $a \in A$ and $x, y \in X$, then

$$d_A(x) \leq d(x, a) \leq d(x, y) + d(y, a).$$

Take the infimum over $a$ in the above equation shows that

$$d_A(x) \leq d(x, y) + d_A(y) \forall x, y \in X.$$ 

Therefore, $d_A(x) - d_A(y) \leq d(x, y)$ and by interchanging $x$ and $y$ we also have that $d_A(y) - d_A(x) \leq d(x, y)$ which implies Eq. (6.6).

**Corollary 6.46.** The function $d$ satisfies,

$$|d(x, y) - d(x', y')| \leq d(y, y') + d(x, x').$$

Therefore $d : X \times X \to [0, \infty)$ is continuous in the sense that $d(x, y)$ is close to $d(x', y')$ if $x$ is close to $x'$ and $y$ is close to $y'$. In particular, if $x_n \to x$ and $y_n \to y$ then

$$\lim_{n \to \infty} d(x_n, y_n) = d(x, y) = d \left( \lim_{n \to \infty} x_n, \lim_{n \to \infty} y_n \right).$$

**Proof. First Proof.** By Lemma 6.45 for single point sets and the triangle inequality for the absolute value of real numbers,

$$|d(x, y) - d(x', y')| \leq |d(x, y) - d(x, y')| + |d(x', y) - d(x', y')| \leq d(y, y') + d(x, x').$$

**Second Proof.** By the triangle inequality,

$$d(x, y) \leq d(x, x') + d(x', y) \leq d(x, x') + d(x', y') + d(y', y)$$

from which it follows that

$$d(x, y) - d(x', y') \leq d(x, x') + d(y', y).$$

Interchanging $x$ with $x'$ and $y$ with $y'$ in this inequality shows,

$$d(x', y') - d(x, y) \leq d(x, x') + d(y', y)$$

and the result follows from the last two inequalities.

**Exercise 6.21 (Continuity of integration).** Let $Z = C([0,1], \mathbb{R})$ be the continuous functions from $[0,1]$ to $\mathbb{R}$ and $\|\cdot\|_\infty$ be the uniform norm, $\|f\|_\infty := \sup_{0 \leq t \leq 1} |f(t)|$. Define $K : Z \to Z$ by

$$K(f)(x) := \int_0^x f(t) \ dt \text{ for all } x \in [0,1].$$
Show that $K$ is a Lipschitz function. In more detail, show
\[
\|K(f) - K(g)\| \leq \|f - g\| \quad \text{for all } f, g \in Z.
\]

In this problem please take for granted the standard properties of the integral including

1. The function $x \to K(f)(x)$ is indeed continuous (in fact differentiable by the fundamental theorem of calculus).
2. $K : Z \to Z$ is a linear transformation.
3. If $f(t) \leq g(t)$ for all $t \in [0, 1]$, then $\int_0^x f(t) \, dt \leq \int_0^x g(t) \, dt$ for all $x \in X$.
4. From 3. it follows that $\int_0^x |f(t)| \, dt \leq \int_0^x |g(t)| \, dt$.

**Exercise 6.22 (Discontinuity of differentiation).** Let $Z$ be the polynomial functions in $C([0,1], \mathbb{R})$, i.e. functions of the form $p(t) = \sum_{k=0}^n a_k t^k$ with $a_k \in \mathbb{R}$. As above we let $\|p\|_u := \sup_{0 \leq t < 1} |p(t)|$. Define $D : Z \to Z$ by $D(p) = p'$, i.e. if $p(t) = \sum_{k=0}^n a_k t^k$ then
\[
D(p)(t) = \sum_{k=1}^n k a_k t^{k-1}.
\]

1. Show $D$ is discontinuous at $0$ – where $0$ represents the zero polynomial.
2. Show $D$ is discontinuous at all points $p \in Z$.

**Exercise 6.23 (Continuity of integration II).** Let $Z = C([0,1], \mathbb{R})$ and $K$ be as in Exercise 6.21. Further let
\[
\|f\|_2 := \sqrt{\int_0^1 |f(t)|^2 \, dt}.
\]

Show
\[
\|K(f) - K(g)\|_2 \leq \frac{1}{\sqrt{2}} \|f - g\|_2 \quad \text{for all } f, g \in Z.
\]

**Definition 6.47 (Pointwise Convergence).** Let $(X, d)$ and $(Y, \rho)$ be metric spaces and $f_n : X \to Y$ be functions for each $n \in \mathbb{N}$. We say that $f_n$ converges pointwise to $f : X \to Y$ provided\n\[
\lim_{n \to \infty} f_n(x) = f(x) \quad \text{for all } x \in X,
\]

i.e. provided
\[
\lim_{n \to \infty} d(f(x), f_n(x)) = 0 \quad \text{for each } x \in X.
\]

**Definition 6.48 (Uniform Convergence).** Let $(X, d)$ and $(Y, \rho)$ be metric spaces and $f_n : X \to Y$ be functions for each $n \in \mathbb{N}$. We say that $f_n$ converges uniformly to $f : X \to Y$ provided
\[
\delta_n := \sup_{x \in X} d(f(x), f_n(x)) \to 0 \quad \text{as } n \to \infty.
\]

**Theorem 6.49 (Uniform Convergence Preserves Continuity).** Suppose that $(f_n)_{n=1}^\infty$ are continuous functions from $X$ to $Y$ and $f_n$ converges uniformly to $f : X \to Y$. If $f_n$ is continuous at $x \in X$ for all $n$ then $f$ is continuous at $x$ as well. In particular if $f_n$ is continuous on $X$ for all $n$ then $f$ is continuous on $X$ as well.

**Proof.** We will give three proofs of this important theorem. In these proofs we will let
\[
\delta_n := \sup_{x \in X} d(f(x), f_n(x)).
\]

**First Proof.** Suppose that $f$ were discontinuous at some point $x_0 \in X$. Then there would exist $\varepsilon > 0$ and $x_k \in X \setminus \{x_0\}$ such that $\lim_{k \to \infty} x_k = x_0$ while $\rho(f(x_k), f(x_0)) \geq \varepsilon$ for all $\varepsilon > 0$. Let $n \in \mathbb{N}$ and set $g := f_n$, then
\[
\varepsilon \leq \rho(f(x_k), f(x_0)) \leq \rho(f(x_k), g(x_k)) + \rho(g(x_k), g(x_0)) + \rho(g(x_0), f(x_0)) \leq \delta_n + \rho(g(x_k), g(x_0)) + \delta_n = 2\delta_n + \rho(g(x_k), g(x_0)).
\]

Letting $k \to \infty$ in this inequality implies, $\varepsilon \leq 2\delta_n$ and then letting $n \to \infty$ implies $\varepsilon = 0$ and we have reached the desired contradiction, see Figure 6.6.
Second Proof. We must show $\lim_{k \to \infty} f(x_k) = f(x)$ whenever $\{x_k\}_{k=1}^\infty \subset X$ is a convergent sequence such that $x := \lim_{k} x_k \in X$. So assume we are given such a sequence $\{x_k\}_{k=1}^\infty$. Then for any $n \in \mathbb{N}$ we have,

$$
\rho(f(x), f(x_k)) \leq \rho(f(x), f_n(x)) + \rho(f_n(x), f(x_k)) \\
\leq \rho(f(x), f_n(x)) + \rho(f_n(x), f_n(x_k)) + \rho(f_n(x_k), f(x_k)) \\
\leq \delta_n + \rho(f(x), f_n(x_k)) + \delta_n.
$$

Therefore,

$$
\limsup_{k \to \infty} \rho(f(x), f(x_k)) \leq \limsup_{k \to \infty} \rho(f_n(x), f_n(x_k)) + 2\delta_n = 2\delta_n,
$$

wherein we have used the continuity of $f_n$ for the last equality. Thus we have shown

$$
\limsup_{k \to \infty} \rho(f(x), f(x_k)) \leq 2\delta_n
$$

which upon passing to the limit as $n \to \infty$ shows $\limsup_{k \to \infty} \rho(f(x), f(x_k)) = 0$. This suffices to show $\lim_{k \to \infty} f(x_k) = f(x)$.

Third Proof. Let $x \in X$ and $\varepsilon > 0$ be given. Choose $n \in \mathbb{N}$ so that $\delta_n \leq \varepsilon$ and let $g := f_n$. Since $g$ is continuous there exists $\delta > 0$ such that $\rho(g(x), g(x')) \leq \varepsilon$ when $d(x, x') \leq \delta$. So if $d(x, x') \leq \delta$, then

$$
\rho(f(x), f(x')) \leq \rho(f(x), g(x)) + \rho(g(x), f(x')) \\
\leq \rho(f(x), g(x)) + \rho(g(x), g(x')) + \rho(g(x'), f(x')) \\
\leq \delta_n + \rho(g(x), g(x')) + \delta_n \leq \varepsilon + \varepsilon + \varepsilon = 3\varepsilon.
$$

As $\varepsilon > 0$ and $x \in X$ were arbitrary, we have shown $f$ is continuous on $X$. 

Example 6.50 (Non-uniform convergence). For an example of nonuniform convergence, suppose that $g(x) = \max(1 - 4x^2, 0)$ and $f_n(x) := g(x - 3n)$ for all $n$, see Figure 6.7 Notice that for each $x \in \mathbb{R}_+$, $f_n(x) = 0$ for a.a. $n$ and therefore $\lim_{n \to \infty} f_n(x) = 0$. On the other hand, $\|f_n\|_u = 1$ for all $n$ so $f_n$ converges to $0$ pointwise but not uniformly in $x$. Nevertheless the limiting function is still continuous. This is not always the case as you will see in the next exercise.

Exercise 6.24. Let $f_n : [0, 1] \to \mathbb{R}$ be defined by $f_n(x) = x^n$ for $x \in [0, 1]$. Show $f(x) := \lim_{n \to \infty} f_n(x)$ exists for all $x \in [0, 1]$ and find $f$ explicitly. Show that $f_n$ does not converge to $f$ uniformly.

6.4 Density and Separability

Definition 6.51. Let $(X,d)$ be a metric space. We say $A \subset X$ is dense in $X$ if for all $x \in X$, there exists $\{x_n\}_{n=1}^\infty \subset A$ such that $x = \lim_{n \to \infty} x_n$. [In words, all points in $X$ are limit points of sequences in $A$.] A metric space is said to be separable if it contains a countable dense subset, $A$.

Example 6.52. The spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ with their Euclidean metrics are separable. Indeed we can take $D = \mathbb{Q}^n$ and $D = (\mathbb{Q} + i\mathbb{Q})^n$ respectively for the
countable dense subsets. For example given $k \in \mathbb{N}$ and $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, we may choose $q^k = (q^k_1, \ldots, q^k_n) \in D$ such that $|x_i - q^k_i| \leq \frac{1}{k}$ for $1 \leq i \leq n$. We then have,

$$d(x, q^k) = \sqrt{\sum_{i=1}^{n} |x_i - q^k_i|^2} \leq \sqrt{n} \cdot \frac{1}{k} \to 0 \text{ as } k \to \infty.$$ 

The $\mathbb{C}^n$ case now follows from this case as $\mathbb{C}^n$ is really $\mathbb{R}^{2n}$ in disguise.

**Example 6.53.** Let $Y := \mathbb{R} \setminus \mathbb{Q}$ which we equip with the usual metric, $d(y, y') = |y - y'|$ for all $y, y' \in Y$. I now claim that $Y$ is separable. We can no longer use $\mathbb{Q}$ as the countable dense subset of $Y$ since $\mathbb{Q}$ is not contained in $Y$! On the other hand, for each $q \in \mathbb{Q}$ we may choose $y_n(q) \in Y$ such that $\lim_{n \to \infty} y_n(q) = q$. Then if $y \in Y$ and $\varepsilon = \frac{1}{\pi} > 0$ is given, we may choose $q \in \mathbb{Q}$ such that $|y - q| \leq \frac{1}{\pi}$. We then take $a_k := y_n(q)$ for some large $n$ so that $|a_k - q| \leq \frac{1}{\pi}$. It then follows that $|y - a_k| \leq \frac{1}{\pi} k \to 0$ as $k \to \infty$ and this shows that $A := \bigcup_{q \in \mathbb{Q}} \{y_n(q) : n \in \mathbb{N}\}$ is a dense subset of $Y$. Moreover $A$ is countable, why?

**Remark 6.54.** An equivalent way to say that $A \subset X$ is dense is to say $d_A \equiv 0$, i.e. $d_A(x) = 0$ for all $x \in X$. Indeed if $x \in X$, you should show that $d_A(x) = 0$ iff there exists $a_n \in A$ such that $d(x, a_n) \to 0$ as $n \to \infty$.

**Exercise 6.25.** Suppose that $(X, d)$ is a separable metric space and $Y$ is a non-empty subset of $X$ which is also a metric space by restricting $d$ to $Y$. Show $(Y, d)$ is separable. [Hint: suppose that $A \subset X$ be a countable dense subset of $X$. For each $a \in A$ choose $\{y_n(a)\}_{n=1}^{\infty} \subset Y$ so that $d_Y(a) \leq d(a, y_n(a)) \leq d_Y(a) + \frac{1}{n}$. Now show $A_Y := \bigcup_{a \in A} \{y_n(a) : n \in \mathbb{N}\}$ is a countable dense subset of $Y$.]

**Exercise 6.26.** Let $n \in \mathbb{N}$. Show any non-empty subset $Y \subset \mathbb{C}^n$ equipped with the metric,

$$d(x, y) = \|y - x\| \text{ for all } x, y \in Y$$

is separable, where $\|\cdot\|$ is either $\|\cdot\|_1$, $\|\cdot\|_1$, or $\|\cdot\|_2$.

**Exercise 6.27.** For $x, y \in \mathbb{R}$, let

$$d(x, y) := \begin{cases} 
0 & \text{if } x = y \\
1 & \text{if } x \neq y
\end{cases}$$

Then $(\mathbb{R}, d)$ is a non-separable metric space. (This metric space is also complete.)

**Exercise 6.28.** Suppose $(X, \rho)$ and $(Y, d)$ are metric spaces and $A$ is a dense subset of $X$.

1. Show that if $F : X \to Y$ and $G : X \to Y$ are two continuous functions such that $F = G$ on $A$ then $F = G$ on $X$.

2. Now suppose that $(Y, d)$ is complete and $f : A \to Y$ is a function which is uniformly continuous (Notation 6.43). Recall this means for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$d(f(a), f(b)) \leq \varepsilon \text{ for all } a, b \in A \text{ with } \rho(a, b) \leq \delta. \quad \text{(6.7)}$$

Show there is a unique continuous function $F : X \to Y$ such that $F = f$ on $A$. [Hint: Define $F(x) = \lim_{n \to \infty} f(x_n)$ where $\{x_n\}_{n=1}^{\infty} \subset A$ is chosen to converge $x \in X$. You must show the limit exists and is independent of the choice of sequence $\{x_n\}_{n=1}^{\infty} \subset A$ which converges for $x$.]

3. Let $X = \mathbb{R} = Y$ and $A = \mathbb{Q} \subset X$, find a function $f : \mathbb{Q} \to \mathbb{R}$ which is continuous on $\mathbb{Q}$ but does not extend to a continuous function on $\mathbb{R}$.

### 6.5 Test 2: Review Topics

1. Understand the basic properties of complex numbers.
2. Countability. Key facts are that countable union of countable sets is countable and the finite product of countable sets is countable.
3. Definitions of metric and normed spaces and their basic properties which in the end of the day typically follow from the triangle inequality.
4. Be aware of different norms, $\|\cdot\|_1$, $\|\cdot\|_1$, and $\|\cdot\|_2$.
5. Understand the notion of limits of sequences, Cauchy sequences, completeness, limits and continuity of functions.
6. Know what is meant by pointwise and uniform convergence. You should be able to compute pointwise limits and know how to test if the limit is uniform or not. A key theorem is the uniform limit of continuous functions is still continuous.
Series and Sums in Banach Spaces

Definition 7.1. Suppose \((X, \| \cdot \|)\) is a normed space and \(\{x_n\}_{n=1}^{\infty}\) is a sequence in \(X\). Then we say \(\sum_{n=1}^{\infty} x_n\) converges in \(X\) iff \(\lim_{N \to \infty} \sum_{n=1}^{N} x_n\) exists in \(X\) otherwise we say \(\sum_{n=1}^{\infty} x_n\) diverges. We often let \(S_N := \sum_{n=1}^{N} x_n\) and refer to \(\{S_N\}_{N=1}^{\infty} \subset X\) as the sequence of partial sums.

Theorem 7.3 (Telescoping Series / Fundamental Theorem of Summation). Suppose \(\{a_n\}_{n=1}^{\infty}\) and \(\{b_n\}_{n=1}^{\infty}\) are sequences in \([0, \infty)\). If \(a_n \leq b_n\) for all \(n\), then
\[
\sum_{n=1}^{\infty} a_n \leq \sum_{n=1}^{\infty} b_n
\]
where we allow for these sums to be infinite. Moreover if \(a_n \leq b_n\) for a.a. \(n\) then \(\sum_{n=1}^{\infty} b_n < \infty\) implies \(\sum_{n=1}^{\infty} a_n < \infty\) and if \(\sum_{n=1}^{\infty} a_n = \infty\) then \(\sum_{n=1}^{\infty} b_n = \infty\).

Proof. Let \(A_k := \sum_{n=1}^{k} a_n\) and \(B_k := \sum_{n=1}^{k} b_n\). Then a simple induction argument shows that \(A_k \leq B_k\) for all \(k\) and therefore
\[
\sum_{n=1}^{\infty} a_n = \lim_{k \to \infty} A_k \leq \lim_{k \to \infty} B_k = \sum_{n=1}^{\infty} b_n
\]
by the sandwich lemma.

Theorem 7.7 (Telescoping Series / Fundamental Theorem of Summation). Let \(\{f(n)\}_{n=1}^{\infty} \subset X\) be a sequence, then
\[
\sum_{n=1}^{N} [f(n+1) - f(n)] = f(N+1) - f(1) \text{ for all } N \in \mathbb{N}
\]
and \(\sum_{n=1}^{\infty} [f(n+1) - f(n)]\) is convergent in \(X\) iff \(\lim_{N \to \infty} f(N)\) exists in \(X\) in which case,
\[
\sum_{n=1}^{\infty} (f(n+1) - f(n)) = \lim_{N \to \infty} f(N) - f(1).
\]

We also have,
\[
\sum_{n=M}^{N} [f(n+1) - f(n)] = f(N+1) - f(M) \text{ for } N \geq M.
\]

Proof. When \(N = 3\) we have,
\[
\sum_{n=1}^{3} (f(n+1) - f(n)) = (f(2) - f(1)) + (f(3) - f(2)) + (f(4) - f(3))
= f(4) - f(1)
\]
In general, Eqs. (7.1) and (7.2) are easily verified by a simple induction argument. The rest of the theorem is now evident.

Example 7.4 (Geometric Series). Suppose that \(f(n) = \alpha^n\) where \(\alpha \in \mathbb{C}\). Then
\[
f(n+1) - f(n) = \alpha^{n+1} - \alpha^n = \alpha^n (\alpha - 1)
\]
and we find,
\[
(\alpha - 1) \sum_{n=1}^{N} \alpha^n = \alpha^{N+1} - \alpha.
\]
If \(\alpha \neq 1\) it follows that
\[
\sum_{n=1}^{N} \alpha^n = \frac{\alpha^{N+1} - \alpha}{\alpha - 1}
\]
and if \(|\alpha| < 1\), it follows that \(|\alpha^{N+1}| = |\alpha|^{N+1} \to 0\) as \(N \to \infty\) and therefore
\[
\sum_{n=1}^{\infty} \alpha^n = \frac{\alpha}{1 - \alpha}.
\]

Theorem 7.5 (Integral test). Let \(f : (0, \infty) \to \mathbb{R}\) be a \(C^1\) function such that \(f'(x) \geq 0\) and \(f'(x)\) is decreasing in \(x\) (i.e. \(f''(x) \leq 0\) if it exists) and \(f(\infty) := \lim_{x \to \infty} f(x)\), see Figure 7.1. (As \(f\) is an increasing function, \(f(\infty)\) exists in \((-\infty, \infty).\)) Then
\[ f(N+1) - f(M) \leq \sum_{n=M}^{N} f'(n) \leq [f(N+1) - f(M)] + f'(M) - f'(N+1) \] (7.3)

for all \( M \leq N \). Letting \( N \to \infty \) in these inequalities also gives,

\[ f(\infty) - f(M) \leq \sum_{n=M}^{\infty} f'(n) \leq f(\infty) - f(M) + f'(M) \] (7.4)

and in particular, \( \sum_{n=1}^{\infty} f'(n) < \infty \) iff \( f(\infty) < \infty \).

**Proof.** By the mean value theorem, there exists \( c_n \in (n-1, n) \) such that

\[ f(n) - f(n-1) = f'(c_n) \] (1)

Since \( f' \) is decreasing, it follows that

\[ f'(n) \leq f'(c_n) \leq f'(n-1) \] for all \( n \),

or equivalently\(^1\) that

\[ f'(n) \leq f(n) - f(n-1) \leq f'(n-1) \] for all \( n \).

Summing these inequalities on \( n \), using

\[ \sum_{n=M+1}^{N+1} [f(n) - f(n-1)] = f(N+1) - f(M), \]

shows,

\[ \sum_{n=M+1}^{N+1} f'(n) \leq f(N+1) - f(M) \leq \sum_{n=M+1}^{N+1} f'(n) - \sum_{n=M+1}^{N+1} f'(n-1) = \sum_{n=M}^{N} f'(n). \]

This inequality is equivalent to Eq. (7.3) because it gives,

\[ f(N+1) - f(M) \leq \sum_{n=M}^{N} f'(n) = \sum_{n=M+1}^{N+1} f'(n) + f'(M) - f'(N+1) \]

\[ \leq [f(N+1) - f(M)] + f'(M) - f'(N+1). \]

**Corollary 7.6 \((p - \text{series})\).** Let \( p \in \mathbb{R}, \) then

\[ \sum_{n=1}^{\infty} \frac{1}{n^p} = \begin{cases} \infty & \text{if } p \leq 1 \\ < \infty & \text{if } p > 1 \end{cases} \]

**Proof.** If \( p \leq 0 \), then \( \lim_{n \to \infty} \frac{1}{n^p} \neq 0 \) and hence the series diverges and we may assume that \( p > 0 \). For \( p > 1 \), let \( f(x) = -x^{p-1} \) so that \( f'(x) = (p-1) x^{-p} \) which is decreasing in \( x \) and therefore by Theorem 7.5,

\[ f(\infty) - f(M) \leq \sum_{n=M}^{\infty} \frac{p-1}{n^p} \leq f(\infty) - f(M) + f'(M) = \frac{1}{M^{p-1}} + \frac{p-1}{M^p} < \infty \]

and so

\[ \frac{1}{(p-1) M^{p-1}} \leq \sum_{n=M}^{\infty} \frac{1}{n^p} \leq \frac{1}{(p-1) M^{p-1}} + \frac{1}{M^p}. \]

In particular for \( M = 1 \),

\[ \frac{1}{p-1} \leq \sum_{n=1}^{\infty} \frac{1}{n^p} \leq \frac{1}{p-1} + 1. \]

Since

\[ \frac{1}{p-1} \leq \sum_{n=1}^{\infty} \frac{1}{n^p} \leq \sum_{n=1}^{\infty} \frac{1}{n} \]
we may let $p \downarrow 1$ in order to show, $\infty \leq \sum_{n=1}^{\infty} \frac{1}{n^p}$. This complete the proof as for any $p < 1$ we have,

$$\infty \leq \sum_{n=1}^{\infty} \frac{1}{n} \leq \sum_{n=1}^{\infty} \frac{1}{n^p}.$$ 

\[\boxed{\text{Exercise 7.1.} \text{ Take } f(x) = \ln x \text{ in Theorem 7.6 in order to directly conclude that } \sum_{n=1}^{\infty} \frac{1}{n} = \infty.}\]

\[\text{Exercise 7.2.} \text{ Let } 0 \leq p < \infty. \text{ Prove,} \]

$$\sum_{n=2}^{\infty} \frac{1}{n (\ln n)^p} = \begin{cases} \infty & \text{if } 0 \leq p \leq 1 \\ < \infty & \text{if } p > 1 \end{cases}.$$ 

by applying Theorem 7.5 with the following functions;

1. for $0 \leq p < 1$ take $f(x) = (\ln x)^{1-p}$,
2. for $p = 1$ take $f(x) = \ln \ln x$, and
3. for $p > 1$ take $f(x) = -(\ln x)^{1-p}$.

\[\text{Theorem 7.7.} \text{ Let } (X, \| \cdot \|) \text{ be a Banach space and } \{x_k\}_{k=1}^{\infty} \subset X \text{ be a sequence. Then;}

1. $\sum_{k=1}^{\infty} x_k$ converges iff

   $$\left\| \sum_{k=m}^{n} x_k \right\| \to 0 \text{ as } n, m \to \infty \text{ with } n \geq m.$$ 

2. If $\sum_{k=1}^{\infty} x_k$ converges then $\lim_{k \to \infty} x_k = 0$ or alternatively if $\lim_{k \to \infty} x_k \neq 0$

   then $\sum_{k=1}^{\infty} x_k$ diverges.

3. If $\sum_{k=1}^{\infty} x_k$ converges then $\lim_{N \to \infty} \sum_{k=N}^{\infty} x_k = 0$, i.e. the $N$ - tail, $\sum_{k=N}^{\infty} x_k$, of a

   convergent series, $\sum_{k=1}^{\infty} x_k$, go to zero as $N \to \infty.$

\[\text{Remark:} \text{ the only place we use } X \text{ is complete is in the implication (\Rightarrow) in item 1. All of the remaining statements hold for any normed space } X.\]

\[\text{Proof.} \text{ Let } S_n := \sum_{k=1}^{n} x_k \text{ so that } \sum_{k=1}^{\infty} x_k \text{ converges iff } \lim_{n \to \infty} S_n \exists\text{ iff } \{S_n\}_{n=1}^{\infty} \text{ is Cauchy since } X \text{ is a Banach space which gives item 1. since}\]

$$S_n - S_{m-1} = \sum_{k=m}^{n} x_k.$$ 

For the second item apply the first with $n = m + 1$. For the third item let $S := \sum_{k=1}^{\infty} x_k$, then $\lim_{N \to \infty} S_N = S$ and so by very definition,

$$\sum_{k=N}^{\infty} x_k = S - S_{N+1} \to 0 \text{ as } N \to \infty.$$ 

\[\text{Exercise 7.3.} \text{ Let } (X, d) \text{ be a metric space. Suppose that } \{x_n\}_{n=1}^{\infty} \subset X \text{ is a sequence and set } \varepsilon_n := d(x_n, x_{n+1}). \text{ Show that for } m > n \text{ that}

$$d(x_n, x_m) \leq \sum_{k=n}^{m-1} \varepsilon_k \leq \sum_{k=n}^{\infty} \varepsilon_k.$$ 

Conclude from this that if

$$\sum_{k=1}^{\infty} \varepsilon_k = \sum_{n=1}^{\infty} d(x_n, x_{n+1}) < \infty$$

then $\{x_n\}_{n=1}^{\infty}$ is Cauchy. Moreover, show that if $\{x_n\}_{n=1}^{\infty}$ is a convergent sequence and $x = \lim_{n \to \infty} x_n$ then

$$d(x, x_n) \leq \sum_{k=n}^{\infty} \varepsilon_k.$$ 

\[\text{Theorem 7.8 (Absolute Convergence Implies Convergence).} \text{ Let } (X, \| \cdot \|) \text{ be a Banach space } \{x_k\}_{k=1}^{\infty} \subset X \text{ be a sequence. Then } \sum_{k=1}^{\infty} \| x_k \| < \infty \text{ implies}

$$\sum_{k=1}^{\infty} x_k \text{ is convergent. [We say } \sum_{k=1}^{\infty} x_k \text{ is absolutely convergent if } \sum_{k=1}^{\infty} \| x_k \| < \infty.]}$$

\[\text{Exercise 7.4.} \text{ Prove Theorem 7.8.} \text{ Namely if } (X, \| \cdot \|) \text{ is a Banach space and } \{x_k\}_{k=1}^{\infty} \subset X \text{ is a sequence, then } \sum_{k=1}^{\infty} \| x_k \| < \infty \text{ implies } \sum_{k=1}^{\infty} x_k \text{ is convergent.}\]

\[\text{Proposition 7.9 (Alternating Series Test).} \text{ If } \{a_k\}_{k=1}^{\infty} \subset [0, \infty) \text{ is a non-increasing sequence (i.e. } a_k \geq a_{k+1} \text{ for all } k) \text{ such that } \lim_{k \to \infty} a_k = 0, \text{ then}

$$s := \sum_{k=1}^{\infty} (-1)^{k+1} a_k \text{ is convergent. Moreover, for all } n \in \mathbb{N}$$

$$s - \sum_{k=1}^{n} (-1)^{k+1} a_k = \sum_{k=n+1}^{\infty} (-1)^{k+1} a_k \leq a_{n+1}. \quad (7.5)$$
In parenthesis above are non-negative. From these inequalities we learn that wherein we have used the fact that 

\[ \sum_{k=1}^{\infty} (-1)^{k+1} a_k \] 

and hence 

\[ S_n \] 

In summary, 

\[ S_n := \lim_{n \to \infty} S_{2n} \] 

exists in \((-\infty, \infty]\) and 

\[ S_{2n+1} := \lim_{n \to \infty} S_{2n+1} \] 

exists in \((-\infty, \infty]\). 

Since \( S_{2n+1} - S_{2n} = a_{2n+1} \to 0 \) as \( n \to \infty \) it follows that \( s_{ev} = s_{odd} \in \mathbb{R} \). It is now easy to conclude (You prove!) that 

\[ S = \sum_{k=1}^{\infty} (-1)^{k+1} a_k = \lim_{n \to \infty} S_n = s_{ev} = s_{odd} \in \mathbb{R}. \]

In summary, \( S_{2n} \uparrow S \) while \( S_{2n+1} \downarrow S \) as \( n \to \infty \) and in particular, for all \( n \in \mathbb{N} \), 

\[ S_{2n+1} - a_{2n} = S_{2n} \leq S \leq S_{2n+1} = S_{2n} + a_{2n+1}. \]

These inequalities then implies 

\[ 0 \leq S - S_{2n} = \sum_{k=2n+1}^{\infty} (-1)^{k+1} a_k \leq a_{2n+1} \]

and 

\[ 0 \leq S_{2n+1} - S = \sum_{k=2n}^{\infty} (-1)^{k+1} a_k \leq a_{2n}. \]

which imply Eq. (7.6).

**Alternatively,** here is another way to think about the error estimates; 

\[ 0 \leq a_1 - a_2 = S_2 \leq S \leq S_1 \leq a_1 \]

so that \( 0 \leq S \leq a_1 \). Applying these results to \( \sum_{k=n+1}^{\infty} (-1)^{k+1} a_k \) shows, 

\[ s - \sum_{k=1}^{n} (-1)^{k+1} a_k \]

Example 7.10. By the alternating series test, 

\[ \sum_{n=1}^{\infty} (-1)^n \frac{1}{n^p} \]

is convergent for all \( p > 0 \). On the other hand, by Corollary 7.6 the series is absolutely convergent if \( p > 1 \).

**Theorem 7.11 (Uniform convergence and the Weierstrass M-test).** Suppose that \((X, \| \cdot \|)\) is a Banach space, \((Y, d)\) is a metric space, for each \( n \in \mathbb{N} \), 

\[ f_n : Y \to X \text{ is a function, and there exists } \{M_n\}_{n=1}^{\infty} \subset [0, \infty) \text{ satisfying:} \]

\[ \sup_{y \in Y} \|f_n(y)\|_X \leq M_n \quad \forall \ n \in \mathbb{N} \text{ and } \sum_{n=1}^{\infty} M_n < \infty. \]

Then \( S_N(y) := \sum_{n=1}^{N} f_n(y) \) converges absolutely and uniformly to \( S(y) = \sum_{n=1}^{\infty} f_n(y) \). Moreover, if we further assume that \( \{f_n\}_{n=1}^{\infty} \subset C(Y, X) \) (i.e. \( f_n \) is continuous for all \( n \)), then the function \( S : Y \to X \) is also continuous.

**Proof.** For any \( y \in Y \), 

\[ \sum_{n=1}^{\infty} \|f_n(y)\|_X \leq \sum_{n=1}^{\infty} M_n < \infty \]

and therefore \( S(y) = \sum_{n=1}^{\infty} f_n(y) \) converges absolutely. Moreover we have, 

\[ \|S(y) - S_N(y)\| = \lim_{M \to \infty} \|S_M(y) - S_N(y)\| = \lim_{M \to \infty} \left\| \sum_{n=N+1}^{M} f_n(y) \right\| \leq \liminf_{M \to \infty} \sum_{n=N+1}^{M} \|f_n(y)\| = \sum_{n=N+1}^{\infty} \|f_n(y)\| \leq \sum_{n=N+1}^{\infty} M_n. \]
As the last member of this inequality does not depend on $y$ we have,
\[
\sup_{y \in Y} \|S(y) - S_N(y)\| \leq \sum_{n=N+1}^{\infty} M_n \to 0 \text{ as } N \to \infty
\]
because tails of convergent series vanish, Theorem 7.11. The continuity of $S$ now forms the continuity of $S_N$, the uniform convergence just proved, and Theorem 6.48.

**Theorem 7.12 (Root test).** Suppose that \( \{a_n\}_{n=1}^{\infty} \) is a sequence in \( \mathbb{C} \) and let \( \alpha := \limsup_{n \to \infty} |a_n|^{1/n} \). Then

1. If \( \alpha < 1 \) then \( \sum_{n=1}^{\infty} |a_n| < \infty \) and \( \sum_{n=1}^{\infty} a_n \) is absolutely convergent.
2. If \( \alpha > 1 \), then \( \limsup_{n \to \infty} |a_n| = \infty \) and \( \sum_{n=1}^{\infty} a_n \) diverges.
3. If \( \alpha = 1 \), the test fails, i.e. you must work harder!

**Proof.** We take each item in turn.

1. If \( \alpha < 1 \), let \( \beta \in (0,1) \), then \( |a_n|^{1/n} \leq \beta \) for a.a. \( n \) which implies that \( |a_n| \leq \beta^n \) for a.a. \( n \) and so the result follows by the comparison Theorem 7.2 as
\[
\sum_{n=1}^{\infty} \beta^n = \frac{\beta}{1 - \beta} < \infty.
\]

2. If \( \alpha > 1 \) and \( \beta \in (1, \alpha) \), then \( |a_n|^{1/n} \geq \beta \) i.o. \( n \) and hence \( |a_n| \geq \beta^n \) i.o. \( n \).

As \( \beta^n \to \infty \) as \( n \to \infty \) it follows that \( \limsup_{n \to \infty} |a_n| = \infty \).

3. From Corollary 7.6 we know that \( \sum_{n=1}^{\infty} \frac{1}{n^\beta} = \infty \) while \( \sum_{n=1}^{\infty} \frac{1}{n^\alpha} < \infty \). However from Lemma 3.30
\[
\lim_{n \to \infty} \left( \frac{1}{n} \right)^{1/n} = 1 = \lim_{n \to \infty} \left( \frac{1}{n^\beta} \right)^{1/n}
\]
which shows the test has failed. In fact, if \( 0 < p < \infty \) and \( k \in \mathbb{N} \) such that \( k \geq p \), then \( \frac{1}{n^k} \leq \frac{1}{n^p} \leq 1 \) which implies
\[
\left( \frac{1}{n^k} \right)^{1/n} \leq \left( \frac{1}{n^p} \right)^{1/n} \leq (1)^{1/n} = 1.
\]

By Lemma 3.30 we know \( \lim_{n \to \infty} \left( \frac{1}{n} \right)^{1/n} = 1 \) and therefore by the sandwich lemma, \( \lim_{n \to \infty} \left( \frac{1}{n^p} \right)^{1/n} = 1 \).

**Exercise 7.5.** For every \( p \in \mathbb{N} \), show \( \sum_{n=0}^{\infty} (n)^{n/p} z^n \) is convergent iff \( z = 0 \).

**Theorem 7.13 (Ratio test).** Suppose that \( \{a_n\}_{n=1}^{\infty} \) is a sequence in \( \mathbb{C} \) such that \( a_n \neq 0 \) for a.a. \( n \). Then

1. If \( \alpha := \limsup_{n \to \infty} |\frac{a_{n+1}}{a_n}| < 1 \) then \( \sum_{n=1}^{\infty} |a_n| < \infty \) and \( \sum_{n=1}^{\infty} a_n \) is absolutely convergent.
2. If \( |\frac{a_{n+1}}{a_n}| \geq 1 \) for a.a. \( n \) then \( \lim_{n \to \infty} |a_n| > 0 \) and \( \sum_{n=1}^{\infty} a_n \) diverges.
3. If \( \liminf_{n \to \infty} |\frac{a_{n+1}}{a_n}| > 1 \) then \( \lim_{n \to \infty} |a_n| = \infty \) and \( \sum_{n=1}^{\infty} a_n \) diverges.
4. If \( \limsup_{n \to \infty} |\frac{a_{n+1}}{a_n}| = 1 \), the test fails, i.e. you must work harder!

**Proof.** We take each item in turn.

1. If \( \alpha < 1 \), let \( \beta \in (0,1) \), then \( |\frac{a_{n+1}}{a_n}| \leq \beta \) for a.a. \( n \), i.e. there exists \( N \in \mathbb{N} \) such that \( |a_{n+1}| \leq \beta |a_n| \) for all \( n \geq N \). A simple induction argument shows,
\[
|a_n| \leq |a_N| \beta^{n-N} = \beta^{-N} |a_N| \beta^n \quad \text{for } n \geq N.
\]
The result follows by the comparison Theorem 7.2 and the fact that
\[
\sum_{n=N}^{\infty} \beta^{-N} |a_N| \beta^n = \frac{|a_N|}{1 - \beta} < \infty.
\]

2. Suppose there exists \( N \in \mathbb{N} \) such that \( |\frac{a_{n+1}}{a_n}| \geq 1 \) for all \( n \geq N \). This inequality says that \( |a_n| \) is non-decreasing for large \( n \) and therefore \( \lim_{n \to \infty} |a_n| \geq |a_m| \) for any \( m \geq N \). We may now choose \( m \geq N \) such that \( |a_m| \neq 0 \).

3. If \( \liminf_{n \to \infty} |\frac{a_{n+1}}{a_n}| > \beta > 1 \), then \( |\frac{a_{n+1}}{a_n}| \geq \beta \) for a.a. \( n \). Working as above there exists \( N \in \mathbb{N} \) such that \( |a_N| \neq 0 \) and \( |a_n| \geq |a_N| \beta^{-n-N} \) for all \( n \geq N \). From this it follows that \( \lim_{n \to \infty} |a_n| = \infty \).

4. From Corollary 7.6 we know that \( \sum_{n=1}^{\infty} \frac{1}{n^p} < \infty \) iff \( p > 1 \). However
\[
\lim_{n \to \infty} \left( \frac{1}{n^p} \right)^{1/n} = \lim_{n \to \infty} \left( \frac{n}{n+1} \right)^p = \left[ \lim_{n \to \infty} \left( \frac{n}{n+1} \right) \right]^p = 1^p = 1
\]
which shows the test has failed and does not resolve when \( \sum_{n=1}^{\infty} \frac{1}{n^p} < \infty \).

In what follows let \( (Z, \| \cdot \|) \) be a complex Banach space. For example, \( Z = \mathbb{C} \) and \( \|z\| = |z| \) is an important special case.
Remark 7.14. The root and ratio tests extend to series in Banach spaces \( \left( \{a_n\}_{n=1}^{\infty} \subseteq Z \right) \) with the only changes being that we replace \( |a_n| \) by \( \|a_n\| \) and \( \frac{|a_{n+1}|}{|a_n|} \) by \( \frac{\|a_{n+1}\|}{\|a_n\|} \) wherever these occur. The point is that we can apply the root and ratio test to the sequence of real numbers \( \{\|a_n\|\}_{n=1}^{\infty} \) in order to make conclusion about the absolute convergence or divergence of \( \sum_{n=1}^{\infty} a_n \).

Definition 7.15. Given \( z_0 \in \mathbb{C} \) and \( \{a_n\}_{n=0}^{\infty} \subseteq Z \), the series of the form

\[
\sum_{n=0}^{\infty} a_n (z - z_0)^n
\]

is called a power series. If \( z_0 = 0 \) we call it a Maclaurin series, i.e. a series of the form

\[
\sum_{n=0}^{\infty} a_n z^n.
\]

The radius of convergence of either of these series is defined to be 
\( R := \frac{1}{\alpha} \in [0, \infty] \) where

\[
\alpha := \limsup_{n \to \infty} \|a_n\|^{1/n} \in [0, \infty].
\]

By definition, \( \frac{1}{\alpha} := \infty \) in this formula.

The next theorem shows that \( R \) is the critical radius governing the convergence of Eqs. (7.7) and (7.8).

Proposition 7.16. If \( R \) is the radius of convergence of a power series in Eq. (7.7) then:

1. If \( |z - z_0| < R \), the series converges.
2. If \( |z - z_0| > R \), the series diverges.
3. If \( |z - z_0| = R \), the series may or may not converge.

We may also characterize \( R \) as

\[
R := \sup \left\{ |z - z_0| : z \in \mathbb{C} \text{ and } \sum_{n=0}^{\infty} a_n (z - z_0)^n \text{ converges} \right\}.
\]

Proof. Let

\[
\rho(z) := \limsup_{n \to \infty} \|a_n (z - z_0)^n\|^{1/n} = |z - z_0| \cdot \limsup_{n \to \infty} \|a_n\|^{1/n} = \frac{|z - z_0|}{R}.
\]

By the root test, Theorem 7.12 we know that power series in Eq. (7.7) converges if \( \rho(z) < 1 \) and diverges if \( \rho(z) > 1 \) and the test fails if \( \rho(z) = 1 \).

Corollary 7.17. If \( \mu = \lim_{n \to \infty} \frac{\|a_{n+1}\|}{\|a_n\|} \) exists, the radius \( (R) \) of convergence of a power series in Eq. (7.7) is 
\( R = \frac{1}{\mu} = \lim_{n \to \infty} \frac{\|a_n\|}{\|a_{n+1}\|}. \)

Proof. If \( \mu = \lim_{n \to \infty} \frac{\|a_{n+1}\|}{\|a_n\|} \), then

\[
\rho(z) := \lim_{n \to \infty} \frac{\|a_{n+1} (z - z_0)^{n+1}\|}{\|a_n (z - z_0)^n\|} = \frac{|z - z_0|}{R} \lim_{n \to \infty} \frac{\|a_{n+1}\|}{\|a_n\|} = \frac{\mu}{R} \cdot |z - z_0|.
\]

The ratio test in Theorem 7.13 now implies \( \sum_{n=0}^{\infty} a_n (z - z_0)^n \) converges if \( |z - z_0| < \frac{1}{\mu} \) (\( \rho(z) < 1 \)) and diverges if \( |z - z_0| > \frac{1}{\mu} \) (\( \rho(z) > 1 \)), i.e. \( R = \frac{1}{\mu} \) from Eq. (7.7).

Theorem 7.18. If the radius of convergence \( (R) \) of a power series \( \sum_{n=0}^{\infty} a_n (z - z_0)^n \) is positive \( (R > 0) \), then the functions

\[
S : D(z_0, R) := \{ z \in \mathbb{C} : |z - z_0| < R \} \to Z
\]

defined by \( S(z) := \sum_{n=0}^{\infty} a_n (z - z_0)^n \) is continuous and the series is uniformly convergent on \( D(z_0, R) \) for all \( \rho < R \).

Proof. For any \( \rho < R \) let \( M_n := \|a_n\| \rho^n \). Then \( \sum_{n=0}^{\infty} M_n \) is finite \( \leq \|a_n\| \rho^n = M_n \).

It follows from the Weierstrass M - test, Theorem 7.11 that the series is uniformly convergent and \( S \) is continuous on \( D(z_0, \rho) \) for all \( \rho < R \). Since \( D(z_0, R) = \cup_{0<\rho<R} D(z_0,\rho) \), we may conclude that \( S \) is continuous on \( D(z_0, R) \).

Exercise 7.6. Show that each of the following power series have an infinite radius of convergence and hence define continuous functions from \( \mathbb{C} \) to \( \mathbb{C} \):

1. \( \exp(z) := \sum_{n=0}^{\infty} \frac{z^n}{n!} \),
2. \( \sin(z) := \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!} \), and
3. \( \cos(z) := \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!} \).

Proposition 7.19. Suppose that \( (X, \| \cdot \|) \) is a normed space and \( \{a_n\}_{n=1}^{\infty} \), 
\( \{b_n\}_{n=1}^{\infty} \subseteq X \) such that \( A := \sum_{n=1}^{\infty} a_n \) and \( B := \sum_{n=1}^{\infty} b_n \) exist in \( X \). Then
for all \( \lambda \in F \) (\( F = \mathbb{R} \) or \( \mathbb{C} \) if \( X \) is a real or complex vector space respectively), we have the series, \( \sum_{n=1}^{\infty} (a_n + \lambda b_n) \), is convergent and

\[
\sum_{n=1}^{\infty} (a_n + \lambda b_n) = A + \lambda B.
\]
The easy proof of this proposition is left to the reader. Our next goal is to consider the multiplication of series. In this case we will need a multiplication on $X$ itself. We start with the case where $X = \mathbb{C}$ or $X = \mathbb{R}$. Let $\{a_n\}_{n=0}^\infty$, $\{b_n\}_{n=0}^\infty \subset \mathbb{C}$ and let us consider the power series,

$$A(z) := \sum_{n=0}^\infty a_n z^n \quad \text{and} \quad B(z) := \sum_{n=0}^\infty b_n z^n.$$ 

Working formally,

$$A(z) B(z) = \left( \sum_{n=0}^\infty a_n z^n \right) \left( \sum_{m=0}^\infty b_m z^m \right) = \sum_{m,n=0}^\infty a_n b_m z^{n+m} = \sum_{k=0}^\infty \left( \sum_{m+n=k} a_n b_m \right) z^k.$$

Thus we expect $A(z) B(z)$ to be represented by a power series

$$C(z) = \sum_{k=0}^\infty c_k z^k \quad \text{where} \quad c_k := \sum_{m+n=k} a_n b_m = \sum_{m=0}^k a_{k-m} b_m.$$ 

Theorem 7.20 (Multiplication of Series). Suppose that $\{a_n\}_{n=0}^\infty$, $\{b_n\}_{n=0}^\infty \subset \mathbb{C}$ and $\sum_{n=0}^\infty a_n$ and $\sum_{n=0}^\infty b_n$ are absolutely convergent series. Let $\{c_k\}_{k=0}^\infty$ be defined by

$$c_k := \sum_{m=0}^k a_{k-m} b_m. \quad (7.9)$$

Then $\sum_{k=0}^\infty c_k$ converges absolutely and

$$\sum_{k=0}^\infty c_k = \left( \sum_{n=0}^\infty a_n \right) \left( \sum_{n=0}^\infty b_n \right). \quad (7.10)$$

**Proof.** Let $J_N = \{0, 1, 2, \ldots, N\}$, $J_N := \{(m, n) \in J_N^2 \mid m + n > N\}$, and

$$R_N = \sum_{(m, n) \in J_N} |a_n b_m|.$$ 

I now claim that $\lim_{N \to \infty} R_N = 0$. To see this let $A := \sum_{n=0}^\infty |a_n|$, $B := \sum_{m=0}^\infty |b_m|$ and observe that $m + n > N$ implies either $m > N/2$ or $n > N/2$, see Figure 7.3. Hence we find,

$$R_N \leq \sum_{\frac{N}{2} < m \leq N} \sum_{n=0}^N |a_n b_m| + \sum_{\frac{N}{2} < n \leq N} \sum_{m=0}^N |a_n b_m| \leq \sum_{\frac{N}{2} < m \leq N} |b_m| \sum_{n=0}^N |a_n| + \sum_{\frac{N}{2} < n \leq N} |a_n| \sum_{m=0}^N |b_m| \leq \sum_{m > \frac{N}{2}} |b_m| \cdot A + \sum_{n > \frac{N}{2}} |a_n| \cdot B \to 0 \quad \text{as} \quad N \to \infty$$

because the tails of convergent series tend to zero.

We now have the identity,
\[
\sum_{n=0}^{N} a_n \cdot \sum_{m=0}^{N} b_m = \sum_{(m,n) \in I_N \times J_N} a_n b_m = \sum_{m+n \leq N} a_n b_m + r_N = \sum_{k=0}^{N} c_k + r_N,
\]

where \( r_N := \sum_{(m,n) \in I_N} a_n b_m \). Since \( |r_N| \leq R_N \to 0 \), it follows that
\[
\sum_{k=0}^{\infty} c_k = \lim_{N \to \infty} \sum_{k=0}^{N} c_k = \lim_{N \to \infty} \left[ \sum_{n=0}^{N} a_n \cdot \sum_{m=0}^{N} b_m - r_N \right] = \lim_{N \to \infty} \sum_{n=0}^{N} a_n \cdot \lim_{N \to \infty} \sum_{m=0}^{N} b_m = \left( \sum_{n=0}^{\infty} a_n \right) \cdot \left( \sum_{n=0}^{\infty} b_n \right).
\]

Lastly observe that
\[
|c_k| \leq \sum_{m=0}^{k} |a_{k-m}| \cdot |b_m| =: \tilde{c}_k
\]
and so applying what we have just proved to \( \{a_n\}_{n=0}^{\infty} \) and \( \{b_n\}_{n=0}^{\infty} \) shows
\[
\sum_{k=0}^{\infty} |c_k| \leq \sum_{k=0}^{\infty} \tilde{c}_k = \left( \sum_{n=0}^{\infty} |a_n| \right) \cdot \left( \sum_{n=0}^{\infty} |b_n| \right) < \infty.
\]

Here is the short summary of the argument;

\[
\left| \sum_{m,n=0}^{N} a_n \cdot b_m - \sum_{k=0}^{N} c_k \right| = \left| \sum_{(m,n) \in I_N} a_n \cdot b_m \right| \\
\leq \sum_{(m,n) \in I_N} |a_n \cdot b_m| \\
\leq \sum_{m \leq N, \ 0 \leq n \leq N} |a_n \cdot b_m| + \sum_{n \leq N, \ 0 \leq m \leq N} |a_n \cdot b_m| \\
\leq \sum_{m \geq \frac{N}{2}} |b_m| \cdot A + \sum_{n \geq \frac{N}{2}} |a_n| \cdot B \to 0 \text{ as } N \to \infty.
\]

\begin{corollary}
Let \( \{a_n\}_{n=0}^{\infty}, \{b_n\}_{n=0}^{\infty} \subset \mathbb{C} \) and \( c_k := \sum_{m+n=k} a_m b_n \). If
\[
R := \min \left( \frac{1}{\limsup_{n \to \infty} |a_n|^{1/n}}, \frac{1}{\limsup_{n \to \infty} |b_n|^{1/n}} \right) > 0,
\]
then \( \sum_{k=0}^{\infty} c_k z^k \) is convergent for \( |z| < R \) and
\[
\sum_{k=0}^{\infty} c_k z^k = \left( \sum_{n=0}^{\infty} a_n z^n \right) \left( \sum_{m=0}^{\infty} b_m z^m \right).
\]
\end{corollary}

\begin{proof}
For \( |z| < R \), the root test shows
\[
\sum_{n=0}^{\infty} |a_n z^n| < \infty \quad \text{and} \quad \sum_{m=0}^{\infty} |b_m z^m| < \infty.
\]
The result now follows by applying Theorem \ref{prop:root-test} with \( a_n \to |a_n z^n| \) and \( b_n \to |b_n z^n| \).
\end{proof}

\begin{example}
For \( |z| < 1 \) we have
\[
(1 - z) \sum_{n=0}^{\infty} z^n = 1.
\]
\end{example}

This shows that power series, \( \sum_{k=0}^{\infty} c_k z^k \), in Corollary \ref{corr:power-series} may in fact have radius of convergence larger than \( R \) in Eq. \ref{eq:7.12}.

\begin{theorem}
The function \( \exp : \mathbb{C} \to \mathbb{C} \) satisfies the following properties:
\begin{enumerate}
\item \( \exp(0) = 1 \),
\item \( \exp(w) \exp(z) = \exp(w + z) \) for all \( w, z \in \mathbb{C} \),
\item \( \exp(w) \) is never 0 and \( \exp(w) \)^{-1} = \exp(-w) \) for all \( w \in \mathbb{C} \).
\end{enumerate}
\end{theorem}

\begin{proof}
The first assertion is clear and item 3. easily follows from item 2. with \( z = -w \). We now prove item 2. by applying Theorem \ref{prop:root-test} with \( a_n := \frac{1}{n!} z^n \) and \( b_n := \frac{1}{n!} w^n \). In this case,
\[
c_k = \frac{1}{m!} z^m \cdot \frac{1}{(k-m)!} w^{k-m} = \frac{1}{k!} \sum_{m=0}^{k} \binom{k}{m} z^m w^{k-m} = \frac{1}{k!} (z + w)^k,
\]
wherein we have used the Binomial theorem for the last equality, see Exercise \ref{ex:binomial-theorem}. It now follows from Theorem \ref{prop:root-test} that
\[
\exp(w) \exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!} \cdot \sum_{n=0}^{\infty} \frac{w^n}{n!} = \sum_{k=0}^{\infty} \frac{1}{k!} (z + w)^k = \exp(w + z).
\]
\end{proof}
Definition 7.24 (Euler’s Number). Euler’s number is defined by
\[
e := \exp (1) = \sum_{m=0}^{\infty} \frac{1^m}{m!} = \sum_{m=0}^{\infty} \frac{1}{m!}.
\]
If \( m \in \mathbb{N} \), then
\[
e = \exp (1) = \exp \left( \frac{1}{m} \right) = \left[ \exp \left( \frac{1}{m} \right) \right]^m.
\]
As \( \exp \left( \frac{1}{m} \right) > 0 \) we may conclude that \( \exp \left( \frac{1}{m} \right) = e^{1/m} \). Now if \( n \in \mathbb{N}_0 \) we find,
\[
e^{\frac{1}{m}} = \left[ \exp \left( \frac{1}{m} \right) \right]^n = \exp \left( \frac{n}{m} \right).
\]
As
\[
e^{-\frac{1}{m}} = \frac{1}{e^{\frac{1}{m}}} = \frac{1}{\exp \left( \frac{1}{m} \right)} = \exp \left( -\frac{n}{m} \right)
\]
we have in fact shown that \( e^q = \exp (q) \) for all \( q \in \mathbb{Q} \). Owing to this identity, we will often write \( e^z \) for \( \exp (z) \).

With the definitions in Exercise 7.6, we have
\[
\exp (iz) = \sum_{n=0}^{\infty} \frac{(iz)^n}{n!} = \sum_{n=0}^{\infty} \frac{(iz)^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{(iz)^{2n+1}}{(2n+1)!} = \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n}}{(2n)!} + i \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n+1}}{(2n+1)!} = \cos (z) + i \sin (z),
\]
i.e.
\[
\exp (iz) = \cos (z) + i \sin (z)
\]
for all \( z \in \mathbb{C} \).

This is called Euler’s formula. If \( z = \theta \in \mathbb{R} \), it states that
\[
\Re \exp (i \theta) = \cos \theta \quad \text{and} \quad \Im \exp (i \theta) = \sin \theta.
\]
Thus if \( \alpha, \theta \in \mathbb{R} \), we have the following addition formulas;
\[
\cos (\theta + \alpha) = \Re \exp (i (\theta + \alpha)) = \Re \left[ \exp (i \theta) \exp (i\alpha) \right] = \Re \left[ (\cos \theta + i \sin \theta) \cdot (\cos \alpha + i \sin \alpha) \right] = \cos \theta \cos \alpha - \sin \theta \sin \alpha
\]
and
\[
\sin (\theta + \alpha) = \Im \exp (i (\theta + \alpha)) = \Im \left[ \exp (i \theta) \exp (i\alpha) \right] = \Im \left[ (\cos \theta + i \sin \theta) \cdot (\cos \alpha + i \sin \alpha) \right] = \cos \theta \sin \alpha + \sin \theta \cos \alpha.
\]

We also define
\[
\sinh (z) = \frac{e^z - e^{-z}}{2} = \sum_{n=0}^{\infty} \frac{z^{2n+1}}{(2n+1)!}
\]
and
\[
\cosh (z) = \frac{e^z + e^{-z}}{2} = \sum_{n=0}^{\infty} \frac{z^{2n}}{(2n)!}.
\]
The following identities are now easily verified;
\[
\cos (-z) = \cos (z) \quad \text{and} \quad \cosh (-z) = \cosh (z) \quad \text{and} \quad \sinh (-z) = -\sinh (z) \quad \text{and} \quad \sin (z) = \sin \theta \cos \alpha + \sin \alpha \cos \theta \cos \alpha.
\]

Exercise 7.7 (Addition Formulas). Prove the addition formulas in Eqs. (7.13) and (7.14) extend to \( \alpha, \theta \in \mathbb{C} \), i.e. show for all \( w, z \in \mathbb{C} \) that
\[
\cos (w + z) = \cos (w) \cos (z) - \sin (w) \sin (z)
\]
and
\[
\sin (w + z) = \sin (w) \sin (z) + \cos (w) \cos (z).
\]

Exercise 7.8. Suppose that \( p (t) \) and \( q (t) \) are non-zero polynomials of \( t \in \mathbb{R} \) with (possibly) complex coefficients. Further assume that \( q (n) \neq 0 \) for all \( n \in \mathbb{N}_0 \). Show for any sequence \( \{a_n\}_{n=0}^{\infty} \subset \mathbb{C} \) that the power series
\[
\sum_{n=0}^{\infty} \frac{p (n)}{q (n)} a_n z^n
\]
and
\[
\sum_{n=0}^{\infty} a_n z^n
\]
have the same radius of convergence.

Theorem 7.25 (Hilbert Schmidt norm). Let \( Z = \mathbb{C}^{J \times J} \) denote the space of \( N \times N \) complex matrices, \( A = (A_{ij})_{i,j=1}^{N} \) with \( A_{ij} \in \mathbb{C} \). We let
\[
\| A \|_2 := \left\| \sum_{i,j=1}^{N} |A_{ij}|^2 \right\|
\]
which is called the Hilbert Schmidt norm on \( Z \). This norm satisfies,
1. \( \|I\|_2 = \sqrt{N} \) where \( I \) is the \( N \times N \) identity matrix.

2. \( \|AB\|_2 \leq \|A\|_2 \cdot \|B\|_2 \) for all \( A, B \in \mathbb{Z} \).

3. \( \|A^n\|_2 \leq \|A\|_{2n}^n \) for all \( n \in \mathbb{N} \).

4. If \( A_n \to A \) and \( B_n \to B \) then \( A_nB_n \to AB \) and \( A_n + B_n \to A + B \) as \( n \to \infty \). Thus matrix multiplication and addition are continuous operations on \( (Z, \|\cdot\|_2) \).

5. If \( (X, d) \) is a metric space and \( f : X \to Z \) and \( g : X \to Z \) are continuous functions then so is \( f \cdot g \) (order matters) and \( f + g \).

6. The functions \( f(A) = A^n \) is continuous on \( Z \) for all \( n \in \mathbb{N}_0 \), where by convention \( A^0 = I \).

7. \( Z, \|\cdot\|_2 \) is a Banach space.

**Proof.** Item 1 is clear. From the definition of matrix multiplication and the Cauchy-Schwarz inequality we find

\[
\left| (AB)_{ij} \right|^2 \leq \sum_{k=1}^{N} |A_{ik}|^2 \cdot \sum_{k=1}^{N} |B_{kj}|^2.
\]

Therefore,

\[
\|AB\|_2^2 = \sum_{i,j=1}^{n} \left| (AB)_{ij} \right|^2 \leq \sum_{k=1}^{N} \left( \sum_{k=1}^{N} |A_{ik}|^2 \cdot \sum_{k=1}^{N} |B_{kj}|^2 \right)
\]

\[
= \sum_{i,k=1}^{N} |A_{ik}|^2 \cdot \sum_{j,k=1}^{N} |B_{kj}|^2 = \|A\|_2^2 \cdot \|B\|_2^2
\]

which proves item 2. Item 3. follows by an easy induction argument.

Item 4. Let \( \delta A_n := A_n - A \) and \( \delta B_n := B_n - B \) so that \( A_n = A + \delta A_n \) and \( B_n = B + \delta B_n \) where \( \|\delta A_n\|_2 \to 0 \) and \( \|\delta B_n\|_2 \to 0 \) as \( n \to \infty \). Then

\[
\|A_nB_n - AB\|_2 = \|(A + \delta A_n)(B + \delta B_n) - AB\|_2
\]

\[
= \|\delta A_nB + A\delta B_n + \delta A_n\delta B_n\|_2
\]

\[
\leq \|\delta A_nB\|_2 + \|\delta B_n\|_2 + \|\delta A_n\| \cdot \|\delta B_n\|_2
\]

which converges and is continuous.

Item 5. Suppose that \( \{x_n\}_{n=1}^{\infty} \subset X \) and \( x_n \to x \) as \( n \to \infty \), then using item 4.

\[
\lim_{n \to \infty} (f \cdot g)(x_n) = \lim_{n \to \infty} [f(x_n)g(x_n)] = f(x)g(x) = (f \cdot g)(x).
\]

This shows \( f \cdot g \) and \( f + g \) are continuous.

Item 6. follows from item 5 with \( X = Z \) along with an induction argument.

Item 7. follows from the fact that \( (Z, \|\cdot\|_2) \) is \( C^N \) with the \( 2 \) - norm in a slight disguise.

**Definition 7.26 (Matrix Functions).** If \( f(z) = \sum_{n=0}^{\infty} c_n z^n \) for some \( c_n \in \mathbb{C} \) and \( A \in C^{JN} \times JN \), then we define

\[
f(A) := \sum_{n=0}^{\infty} c_n A^n \text{ provided the sum is convergent.}
\]

For example,

\[
\sin(A) := \sum_{n=0}^{\infty} \frac{(-1)^{2n+1}}{(2n+1)!} A^{2n+1}.
\]

**Theorem 7.27.** Suppose that \( \{c_n\}_{n=1}^{\infty} \) is a sequence in \( \mathbb{C} \) and \( R := \left( \limsup_{n \to \infty} |c_n|^{1/n} \right)^{-1} \). If \( A \in C^{JN} \times JN \) with \( \|A\| < R \), then

\[
f(A) := \sum_{n=0}^{\infty} c_n A^n
\]

is convergent and \( f : B_0(R) := \{ A : \|A\| < R \} \to C^{JN \times JN} \) is continuous. Moreover, for any \( \rho < R \), the sum converges absolutely and uniformly on \( B_0(\rho) \).

**Proof.** Let \( \rho \in (\|A\|, R) \) and let \( M_\rho := \|c_n\| \rho^n \). Then \( \limsup_{n \to \infty} M_\rho^{1/n} = \rho/R < 1 \) and therefore by the Root test, \( \sum_{n=0}^{\infty} M_\rho < \infty \). Since, for \( n \geq 1 \),

\[
|c_n A^n| = |c_n| \|A^n\| \leq |c_n| \|A\|^n \leq M_\rho,
\]

the results are now again a consequence of the Weierstrass M - test in Theorem 7.11.

**Theorem 7.28 (Matrix Exponentials and etc.).** The series for \( e^A \), \( \sin(A) \), \( \cos(A) \), and \( \sinh(A) \), and \( \cosh(A) \) are all absolutely convergent define continuous functions from \( C^{JN \times JN} \) to \( C^{JN \times JN} \) in the Hilbert Schmidt norm.

**Proof.** Since the corresponding numerical series all of infinite radius of convergence the result follows from Theorem 7.27.
Exercise 7.9 (Inverting perturbations of the identity). For $\|A\|_2 < 1$ in $\mathbb{C}^{J_N \times J_N}$, $I - A$ is invertible and

$$(I - A)^{-1} = \sum_{n=0}^{\infty} A^n$$

where the sum is absolutely convergent. Moreover the function $A \to (I - A)^{-1}$ is continuous on the ball, $B := \{A : \|A\|_2 < 1\}$ and

$$\| (I - A)^{-1} \|_2 \leq \sum_{n=0}^{\infty} \|A^n\|_2 \leq \|A\|_2^{\infty} \cdot \sqrt{N}.$$  

Exercise 7.10 (Binomial Theorem). Suppose that $A, B \in \mathbb{C}^{J_N \times J_N}$ are commuting matrices, i.e., $AB = BA$. For any $n \in \mathbb{N}$, show by induction that

$$\sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k} = (A + B)^n.$$  \hspace{1cm} (7.14)

Exercise 7.11 (Matrix Addition Formula). If $A, B \in \mathbb{C}^{J_N \times J_N}$ are commuting matrices, show $\exp(A + B) = \exp(A) \cdot \exp(B)$. \textbf{Hint:} mimic the proof of Theorem 7.23

Example 7.29 ($\exp(A + B) \neq \exp(A) \cdot \exp(B)$). Let $\theta \in \mathbb{R}$ and

$$A = \begin{bmatrix} 0 & \theta \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 0 \\ -\theta & 0 \end{bmatrix}.$$  

Since $A^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = B^2$ it follows that

$$\exp(A) = I + A = \begin{bmatrix} 1 & \theta \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \exp(B) = I + B = \begin{bmatrix} 1 & 0 \\ -\theta & 1 \end{bmatrix}.$$  

and in particular,

$$\exp(A) \exp(B) = \begin{bmatrix} 1 & \theta \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\theta & 1 \end{bmatrix} = \begin{bmatrix} 1 - \theta^2 & \theta \\ -\theta & 1 \end{bmatrix}.$$  

Let us now compute $\exp(A + B)$. First observe that

$$A + B = \theta J \quad \text{where} \quad J := \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$  

Since

$$J^2 = -I \quad \text{where} \quad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

it follows that $J^{2n} = (-I)^n = (-1)^n I$ and $J^{2n+1} = J^2 J = (-1)^n J$ for all $n \in \mathbb{N}$. Therefore,

$$\exp(A + B) = \exp(\theta J) = \sum_{n=0}^{\infty} \frac{\theta^n}{n!} J^n$$

$$= \sum_{n=0}^{\infty} \frac{1}{(2n)!} \theta^{2n} J^{2n} + \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} \theta^{2n+1} J^{2n+1}$$

$$= \left( \sum_{n=0}^{\infty} \frac{1}{(2n)!} \theta^{2n} (-1)^n \right) \cdot I + \left( \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} \theta^{2n+1} (-1)^n \right) J$$

$$= \cos \theta \cdot I + \sin \theta \cdot J = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}.$$  

From these formula it is evident (using facts that you know about $\sin \theta$ and $\cos \theta$) that $\exp(A) \exp(B) \neq \exp(A + B)$ unless $\theta = 0$. There is no contradiction to Exercise 7.11 since

$$[A, B] = \begin{bmatrix} 0 & \theta \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ -\theta & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ -\theta & 0 \end{bmatrix} \begin{bmatrix} 0 & \theta \\ 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} -\theta^2 & 0 \\ 0 & \theta^2 \end{bmatrix} = -\theta^2 \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

which is zero iff $\theta = 0$.

Exercise 7.12. Let $(X, \|\cdot\|)$ be a normed space and $T : X \to X$ be a linear transformation and suppose $T$ is continuous at $0 \in X$. Show;

1. There exists $C < \infty$ such that $\|T(x)\| \leq C \|x\|$ for all $x \in X$. \textbf{Hint:} for sake of contradiction suppose that no such $C < \infty$ exists. Then construct a sequence $\{y_n\}_{n=1}^{\infty} \subset X$ such that $\lim_{n \to \infty} y_n = 0$ while $\|T(y_n)\| = 1$ for all $n$.

2. Show $T$ is continuous on all of $X$.  

More Sums and Sequences

Warning: this chapter is a bit rough and will be edited later when it becomes needed.

8.1 Rearrangements

[The stuff about sums of positive numbers could go much earlier in fact.]

Definition 8.1. If $A$ is a countable set and $a : A \to [0, \infty]$ is a function, let

$$\sum_{x \in A} a(x) := \sup_{A_0 \subset_f A} \sum_{x \in A_0} a(x).$$

Notice that if $0 \leq a(x) \leq b(x)$, then $\sum_{x \in A} a(x) \leq \sum_{x \in A} b(x)$. Moreover if $B \subset A$, then

$$\sum_{x \in B} a(x) = \sum_{x \in A} a(x) 1_B(x).$$

Theorem 8.2 (Monotone Convergence Theorem for Sums). If $A$ is a countable set and $a_n : A \to [0, \infty]$ is an increasing sequence of functions, then

$$\lim_{n \to \infty} \sum_{x \in A} a_n(x) := \sum_{x \in A} \lim_{n \to \infty} a_n(x).$$

Proof. This is an easy consequence of the sup – sup theorem. Indeed, $\sum_{x \in A} a_n(x)$ is an increasing sequence of numbers, therefore,

$$\lim_{n \to \infty} \sum_{x \in A} a_n(x) = \sum_{n} \sup_{x \in A} \sum_{x \in A_0} a_n(x)$$

$$= \sup_{A_0 \subset_f A} \sum_{x \in A_0} a_n(x) = \sup_{A_0 \subset_f A} \lim_{n \to \infty} \sum_{x \in A_0} a_n(x)$$

$$= \sup_{A_0 \subset_f A} \lim_{n \to \infty} a_n(x) = \sum_{x \in A} \lim_{n \to \infty} a_n(x).$$

Corollary 8.3. Again suppose that $A$ is a countable set and $a : A \to [0, \infty]$ is a function. If $A_n \subset A$ and $A_n \uparrow A$, then

$$\sum_{A} a(x) = \lim_{n \to \infty} \sum_{A_n} a(x).$$

Proof. Apply Theorem 8.2 with $a_n(x) := a(x) 1_{A_n}(x).$

Corollary 8.4. If $A = A \cup B$ with $A \cap B = \emptyset$, then

$$\sum_{x \in A} a(x) = \sum_{x \in A} a(x) + \sum_{x \in B} a(x).$$

Proof. Choose $A_n \subset_f A$ such that $A_n \uparrow A$. Then $A_n \cap A \uparrow A$ and $A_n \cap B \uparrow B$ and hence,

$$\sum_{x \in A} a(x) = \lim_{n \to \infty} \sum_{x \in A_n} a(x)$$

$$= \lim_{n \to \infty} \left[ \sum_{x \in A_n \cap A} a(x) + \sum_{x \in A_n \cap B} a(x) \right]$$

$$= \lim_{n \to \infty} \sum_{x \in A_n \cap A} a(x) + \lim_{n \to \infty} \sum_{x \in A_n \cap B} a(x)$$

$$= \sum_{x \in A} a(x) + \sum_{x \in B} a(x).$$

Corollary 8.5. If $A$ is a countable set, $a : A \to [0, \infty]$ is a function, and

$$\Lambda = \sum_{n=1}^{\infty} A_n,$$

then

$$\sum_{A} a(x) = \sum_{n=1}^{\infty} \sum_{A_n} a(x).$$

Proof. We have

$$\sum_{n=1}^{\infty} \sum_{A_n} a(x) = \lim_{N \to \infty} \sum_{n=1}^{N} \sum_{A_n} a(x) = \lim_{N \to \infty} \sum_{x \in \bigcup_{n=1}^{N} A_n} a(x) = \sum_{x \in A} a(x).$$
Corollary 8.6. If \( \{x_n\}_{n=1}^{\infty} \) is any enumeration of \( \Lambda \) then
\[
\sum_{x \in \Lambda} a(x) = \sum_{n=1}^{\infty} a(x_n).
\]

**Proof.** \( \Lambda = \sum_{n=1}^{\infty} \{x_n\} \) and therefore
\[
\sum_{x \in \Lambda} a(x) = \sum_{n=1}^{\infty} \sum_{x \in \{x_n\}} a(x) = \sum_{n=1}^{\infty} a(x_n).
\]

Corollary 8.7 (Tonelli Theorem for Sums). Suppose that \( \Lambda \) is a countable set and \( a_n : \Lambda \to [0,\infty] \) is a sequence of functions, then
\[
\sum_{n=1}^{\infty} \sum_{x \in \Lambda} a_n(x) = \sum_{x \in \Lambda} \sum_{n=1}^{\infty} a_n(x).
\]

**Proof.** This is a simple consequence of MCT,
\[
\sum_{n=1}^{\infty} \sum_{x \in \Lambda} a_n(x) = \lim_{N \to \infty} \sum_{n=1}^{N} \sum_{x \in \Lambda} a_n(x) = \lim_{N \to \infty} \sum_{n=1}^{N} \sum_{x \in \Lambda} a_n(x)
\]
\[
= \sum_{x \in \Lambda} \lim_{N \to \infty} \sum_{n=1}^{N} a_n(x) = \sum_{x \in \Lambda} \sum_{n=1}^{\infty} a_n(x).
\]

Lemma 8.8 (Fatou’s Lemma). If \( \Lambda \) is a countable set and \( a_n : \Lambda \to [0,\infty] \) is a sequence of functions, then
\[
\sum_{x \in \Lambda} \liminf_{n \to \infty} a_n(x) \leq \liminf_{n \to \infty} \sum_{x \in \Lambda} a_n(x).
\]

**Proof.** By definition and MCT,
\[
\sum_{x \in \Lambda} \liminf_{n \to \infty} a_n(x) = \sum_{x \in \Lambda} \liminf_{n \to \infty} \inf_{k \geq n} a_k(x)
\]
\[
= \lim_{n \to \infty} \sum_{x \in \Lambda} \inf_{k \geq n} a_k(x)
\]
\[
= \liminf_{n \to \infty} \sum_{x \in \Lambda} \inf_{k \geq n} a_k(x)
\]
\[
\leq \liminf_{n \to \infty} \sum_{x \in \Lambda} a_n(x).
\]

Theorem 8.9 (Dominated Convergence Theorem for Sums I). If \( \Lambda \) is a countable set and \( a_n : \Lambda \to [0,\infty] \) is a sequence of functions such that \( \sum_{x \in \Lambda} \sup_{n} a_n(x) < \infty \) and \( \lim_{n \to \infty} a_n(x) = 0 \) for all \( x \in \Lambda \), then
\[
\lim_{n \to \infty} \sum_{x \in \Lambda} a_n(x) := 0.
\]

**Proof.** For all \( \varepsilon > 0 \) there \( \Lambda_\varepsilon \subset \Lambda \) such that \( \sum_{x \not\in \Lambda_\varepsilon} \sup_{n} a_n(x) \leq \varepsilon \). Therefore,
\[
\sum_{x \in \Lambda} a_n(x) = \sum_{x \in \Lambda_\varepsilon} a_n(x) + \sum_{x \not\in \Lambda_\varepsilon} a_n(x)
\]
\[
\leq \sum_{x \in \Lambda_\varepsilon} a_n(x) + \varepsilon
\]
and therefore,
\[
\limsup_{n \to \infty} \sum_{x \in \Lambda} a_n(x) \leq \limsup_{n \to \infty} \sum_{x \in \Lambda_\varepsilon} a_n(x) + \varepsilon = \lim_{n \to \infty} \sum_{x \in \Lambda_\varepsilon} a_n(x) + \varepsilon = \varepsilon.
\]
Since \( \varepsilon > 0 \) is arbitrary, it follows that
\[
\limsup_{n \to \infty} \sum_{x \in \Lambda} a_n(x) = 0.
\]

**Alternative Proof.** Let \( \alpha(x) := \sup_{n} a_n(x) \), then \( \alpha(x) - a_n(x) \geq 0 \) for all \( x \in \Lambda \). Therefore by Fatou’s Lemma 8.8
\[
\sum_{x \in \Lambda} \liminf_{n \to \infty} [\alpha(x) - a_n(x)] \leq \liminf_{n \to \infty} \sum_{x \in \Lambda} [\alpha(x) - a_n(x)]
\]
\[
= \sum_{x \in \Lambda} \alpha(x) + \liminf_{n \to \infty} \left[ - \sum_{x \in \Lambda} a_n(x) \right]
\]
\[
= \sum_{x \in \Lambda} \alpha(x) - \limsup_{n \to \infty} \left[ \sum_{x \in \Lambda} a_n(x) \right]
\]
while similarly,
\[
\sum_{x \in \Lambda} \liminf_{n \to \infty} [\alpha(x) - a_n(x)] = \sum_{x \in \Lambda} \left( \alpha(x) + \liminf_{n \to \infty} [-a_n(x)] \right)
\]
\[
= \sum_{x \in \Lambda} \left( \alpha(x) - \limsup_{n \to \infty} a_n(x) \right) = \sum_{x \in \Lambda} \alpha(x).
\]
This then implies that
\[
\sum_{x \in A} a(x) \leq \sum_{x \in A} a(x) - \limsup_{n \to \infty} \left[ \sum_{x \in A} a_n(x) \right]
\]
and hence that
\[
\limsup_{n \to \infty} \left[ \sum_{x \in A} a_n(x) \right] \leq 0
\]
as before. \hfill \Box

Now suppose that \((Z, \| \cdot \|)\) is a Banach space and \(a : A \to Z\) is a function such that \(\sum_{x \in A} \|a(x)\| < \infty\).

**Theorem 8.10.** Let \(\ell^1(A, Z)\) be the space of functions \(a : A \to Z\) such that
\[
\|a\|_1 := \sum_{x \in A} \|a(x)\| < \infty.
\]

Then:
1. \((\ell^1(A, Z), \| \cdot \|_1)\) is a Banach space.
2. The set \(D = \{a : A \to Z : \# \{x \in A : a(x) \neq 0\} < \infty\}\) is dense subspace of \(\ell^1(A, Z)\).
3. For \(a \in D\) we may define
\[
\sum_{x \in A} a(x) = \sum_{x : a(x) \neq 0} a(x) \in Z.
\]

Since
\[
\left\| \sum_{x \in A} a(x) \right\| \leq \sum_{x \in A} \|a(x)\| = \|a\|_1,
\]
this linear operator has a unique continuous extension to a linear operator \(\Sigma : \ell^1(A, Z) \to Z\).

**Proof.** We take each item in turn.
1. Let \(\{a_n\}_{n=1}^\infty\) be a Cauchy sequence in \(\ell^1(A, Z)\). Since
\[
\|a_n(x) - a_m(x)\| \leq \|a_n - a_m\|_1 \to 0 \text{ as } m, n \to \infty,
\]
it follows that \(\lim_{n \to \infty} a_n(x) =: a(x)\) exists for all \(x \in A\). Moreover, using Fatou’s Lemma 8.8,
\[
\|a - a_n\|_1 = \sum_{x \in A} \|a(x) - a_n(x)\|
\]
\[
= \sum_{x \in A} \liminf_{m \to \infty} \|a_m(x) - a_n(x)\|
\]
\[
\leq \liminf_{m \to \infty} \sum_{x \in A} \|a_m(x) - a_n(x)\|
\]
\[
= \liminf_{m \to \infty} \|a_m - a_n\|_1 \to 0 \text{ as } n \to \infty.
\]
2. Choose \(A_n \subset_f A\) such that \(A_n \uparrow A\) and set \(a_n(x) = a(x) 1_{A_n}(x)\) for all \(n \in \mathbb{N}\) and \(x \in A\). Then
\[
\lim_{n \to \infty} \|a - a_n\|_1 = 0 \text{ by DCT.}
\]
3. This is a consequence of the BLT theorem. \hfill \Box

**Theorem 8.11 (DCT II).** Suppose that \(\{a_n\}_{n=1}^\infty \subset \ell^1(A, Z)\) and \(A(x) := \sup_n \|a_n(x)\|\) satisfy:
1. \(a(x) := \lim_{n \to \infty} a_n(x)\) existing in \(Z\) for all \(x \in A\), [Conflict of notation] and
2. \(\sum_{x \in A} A(x) < \infty\),

then
\[
\lim_{n \to \infty} \|a - a_n\|_1 = 0 \text{ and } \lim_{n \to \infty} \sum_{x \in A} a_n(x) = \sum_{x \in A} a(x).
\]

**Proof.** It suffices to prove the first assertion since the sum operation is continuous relative to the \(\| \cdot \|_1\) norm on \(\ell^1(A, Z)\). Since
\[
\|a(x)\| = \lim_{n \to \infty} \|a_n(x)\| \leq A(x),
\]
it follows that
\[
\|a(x) - a_n(x)\| \leq \|a(x)\| + \|a_n(x)\| \leq 2A(x).
\]
Since
\[
\lim_{n \to \infty} \|a(x) - a_n(x)\| = \left\| a(x) - \lim_{n \to \infty} a_n(x) \right\| = \|a(x) - a(x)\| = 0 \text{ for all } x \in A,
\]
we may use the dominated convergence theorem for sums (Theorem 8.9) to conclude,
\[
\lim_{n \to \infty} \|a - a_n\|_1 = \lim_{n \to \infty} \sum_{x \in A} \|a(x) - a_n(x)\| = \sum_{x \in A} \lim_{n \to \infty} \|a(x) - a_n(x)\| = 0.
\]
\hfill \Box

**Corollary 8.12.** Suppose \((Z, \| \cdot \|)\) and \(a \in \ell^1(A, Z)\). If \(A_n \uparrow A\), then
\[
\lim_{n \to \infty} \sum_{x \in A_n} a(x) = \sum_{x \in A} a(x).
\]
Proof. Let \( a_n (x) := 1_{A_n} (x) a (x) \) for all \( n \). Then \( \|a_n (x)\| \leq \|a (x)\| \) and \( \sum_{x \in A} \|a (x)\| < \infty \). Since \( \lim_{n \to \infty} a_n (x) = a (x) \) for all \( x \in A \), it follows by DCT that
\[
\lim_{n \to \infty} \sum_{x \in A} a (x) = \lim_{n \to \infty} \sum_{x \in A} a_n (x) = \sum_{x \in A} \lim_{n \to \infty} a_n (x) = \sum_{x \in A} a (x).
\]

\[\text{Corollary 8.13. Suppose \((Z, \|\|)\) is a Banach space, \(A\) is a countable set, and \(a \in \ell^1 (A, Z)\). If \(A = \bigcup_{n=1}^\infty A_n\), then}
\[
\sum_{x \in A} a (x) = \sum_{n=1}^\infty \sum_{x \in A_n} a (x).
\]

[We permit \(A_n = \emptyset\) for some \(n\) in this result with the convention that \(\sum_{x \in \emptyset} a (x) = 0\).]

Proof. We have
\[
\sum_{n=1}^\infty \sum_{x \in A} a (x) = \lim_{N \to \infty} \sum_{n=1}^N \sum_{x \in A} a (x) = \lim_{N \to \infty} \sum_{x \in A} a (x) = \sum_{x \in A} a (x).
\]

because \(\bigcup_{n=1}^N A_n \uparrow A\) as \(N \to \infty\). [Bruce: we need to prove the finite additivity here, namely that if \(A = A \cup B\), then
\[
\sum_{x \in A} a (x) = \sum_{x \in A} a (x) + \sum_{x \in B} a (x).
\]

But this is simple, since \(a = 1_Aa + 1_Ba\) and
\[
\sum_{x \in A} 1_Aa = \sum_{x \in A} a
\]

where this last equality follow by choosing \(A_n \subset_f A\) such that \(A_n \uparrow A\) so that
\[
\sum_{x \in A} 1_Aa = \lim_{n \to \infty} \sum_{x \in A_n} 1_Aa = \lim_{n \to \infty} \sum_{x \in A_n \cap A} a = \sum_{x \in A} a.
\]

\[\text{Corollary 8.14. If} \sum_{n=1}^\infty \sum_{x \in A} \|a_n (x)\| < \infty, \text{then}
\[
\sum_{n=1}^\infty \sum_{x \in A} a_n (x) = \sum_{x \in A} \sum_{n=1}^\infty a_n (x) \in Z.
\]\n
Proof. First observe that
\[
\left\| \sum_{n=1}^N a_n (x) \right\| \leq \sum_{n=1}^N \|a_n (x)\| \leq \sum_{n=1}^\infty \|a_n (x)\| \in \ell^1 (A, Z)
\]

and \(\lim_{N \to \infty} \sum_{n=1}^N a_n (x) = \sum_{n=1}^\infty a_n (x)\). Therefore by DCT,
\[
\sum_{n=1}^\infty a_n (x) = \lim_{N \to \infty} \sum_{n=1}^N a_n (x) = \lim_{N \to \infty} \sum_{n=1}^N \sum_{x \in A} a_n (x)
\]

\[\sum_{x \in A} \lim_{N \to \infty} \sum_{n=1}^N a_n (x) = \sum_{x \in A} \sum_{n=1}^\infty a_n (x).
\]

Better proof. By assumption,
\[
\sum_{n=1}^\infty \|a_n\| < \infty \implies \sum_{n=1}^\infty a_n \text{ converges in } \ell^1 (A, Z).
\]

Therefore,
\[
\sum_{n=1}^\infty \left( \sum_{n=1}^\infty a_n \right) = \sum_{n=1}^\infty \sum_{n=1}^\infty a_n.
\]

\[\text{Exercise 8.1 (Differentiating past an infinite sum). Suppose that}
\[
\sum_{n=0}^\infty \sup_{x} |f_n' (x)| < \infty \text{ and } \sum_{n=0}^\infty f_n (0) \text{ exists.}
\]

Then \(S (x) := \sum_{n=0}^\infty f_n (x)\) exists and
\[
S' (x) = \sum_{n=0}^\infty f_n' (x).
\]

\[\text{Exercise 8.2 (Differentiating past a limit). Suppose } \lim_{n \to \infty} f_n (x) = f (x) \text{ and } f_n' \to g \text{ uniformly. Show } f' (x) = g (x), \text{ i.e. we have in this case that}
\]
\[
d \frac{d}{dx} \lim_{n \to \infty} f_n (x) = \lim_{n \to \infty} d \frac{d}{dx} f_n (x).
\]

\[\text{Exercise 8.3. Suppose that}
\]
\[
f (z) := \sum_{n=0}^\infty a_n z^n
\]

has radius of convergence \(R\). Show \(f' (z) = \sum_{n=0}^\infty n a_n z^n\) for all \(|z| < R\) and the radius of convergence of \(f'\) is still \(R\).
Theorem 8.15 (Changing the center of a power series). Suppose that \( A(z) = \sum_{n=0}^{\infty} \frac{a_n}{n!} z^n \) has radius of convergence \( R > 0 \). For \( |z| < R \) and \( |h| < R - |z| \) we have

\[
A(z + h) = \sum_{m=0}^{\infty} \frac{1}{m!} a_m(z) h^m
\]

where

\[
a_m(z) := \sum_{n=m}^{\infty} \frac{a_n}{(n-m)!} z^{n-m} = A^{(m)}(z).
\]

Proof. Working formally for the moment,

\[
A(z + h) = \sum_{n=0}^{\infty} \frac{a_n}{n!} (z + h)^n = \sum_{n=0}^{\infty} \frac{a_n}{n!} \left( \sum_{m=0}^{n} \binom{n}{m} z^{n-m} h^m \right)
\]

\[
= \sum_{0 \leq m \leq n < \infty} \frac{a_n}{m! \cdot (n-m)!} z^{n-m} h^m
\]

\[
= \sum_{m=0}^{\infty} \frac{1}{m!} \left[ \sum_{n=m}^{\infty} \frac{a_n}{(n-m)!} z^{n-m} \right] h^m
\]

which gives the result provided the interchange of order of sums was permissible. To check this we consider,

\[
\sum_{0 \leq m \leq n < \infty} \frac{|a_n|}{m! \cdot (n-m)!} |z^{n-m}| |h|^m
\]

\[
= \sum_{n=0}^{\infty} \left[ \frac{|a_n|}{n!} (|z| + |h|)^n \right] < \infty
\]

as we know for all \( 0 \leq \rho < R \) that \( \sum_{n=0}^{\infty} \frac{|a_n|}{n!} \rho^n < \infty \). Hence we may apply Fubini’s theorem for sums in order to conclude the result.

8.2 Double Sequences

In this chapter we will consider doubly indexed sequences, \( \{S_{m,n}\}_{m,n=1}^{\infty} \), of complex numbers. To be more precise \( \{S_{m,n}\}_{m,n=1}^{\infty} \) is simply a function from \( \mathbb{N}^2 \) to \( \mathbb{C} \). In this chapter we are interested in the following limits;

\[
\lim_{m \to \infty} S_{m,n}, \quad \lim_{n \to \infty} S_{m,n}, \quad \lim_{m \to \infty} S_{m,n}, \quad \lim_{n \to \infty} S_{m,n}, \quad \lim_{m \wedge n \to \infty} S_{m,n},
\]

where the last two limits are defined as follows.

```
1 Much of what we will say holds for sequences taking values in “complete metric spaces” to be covered later.
```

Definition 8.16. Suppose that \( \{S_{m,n}\}_{m,n=1}^{\infty} \) is a sequence of complex numbers (or more generally elements of a metric space). We say \( \lim_{m \wedge n \to \infty} S_{m,n} = L \) iff for all \( \varepsilon > 0 \) there exists \( N \in \mathbb{N} \) such that

\[
|L - S_{m,n}| \leq \varepsilon \quad \text{for all} \quad m \wedge n \geq N.
\]

We say \( \lim_{m \vee n \to \infty} S_{m,n} = L \) exists iff for all \( \varepsilon > 0 \) there exists \( N \in \mathbb{N} \) such that

\[
|L - S_{m,n}| \leq \varepsilon \quad \text{for all} \quad m \vee n \geq N.
\]

Clearly \( \lim_{m \vee n \to \infty} S_{m,n} = L \) implies \( \lim_{m \wedge n \to \infty} S_{m,n} = L \) but the converse is not true. We are mostly interested in finding sufficient conditions in order for iterated limits to be equal, i.e. for

\[
\lim_{m \to \infty} \lim_{n \to \infty} S_{m,n} = \lim_{n \to \infty} \lim_{m \to \infty} S_{m,n}.
\]

Example 8.17 (Switching Limits is Dangerous I). If \( S_{m,n} = \frac{1}{1 + \frac{m}{n}} \), then \( \lim_{m \to \infty} S_{m,n} = 0 \) (but not uniformly in \( n \)) so that \( \lim_{n \to \infty} \lim_{m \to \infty} S_{m,n} = 0 \) while \( \lim_{n \to \infty} S_{m,n} = 1 \) and \( \lim_{m \to \infty} \lim_{n \to \infty} S_{m,n} = 1 \). In order to visualize better what is going on here let us make the change of variables, \( x = \frac{1}{m} \) and \( y = \frac{1}{n} \), i.e. let

\[
S(x, y) = S_{m,n} = \frac{1}{1 + \frac{y}{x}} = \frac{x}{x + y}
\]

whose plot appears in Figure 8.17.
A plot of \( S_{m,n} = \frac{1+\pi}{1+m+n} \) in terms of the variables \( \frac{1}{m} \) and \( \frac{1}{n} \).

With this change of variables, \( m \to \infty \iff x \to 0 \) and \( n \to \infty \iff y \to 0 \) and \( m = \infty \) and \( n = \infty \) correspond to \( x = 0 \) and \( y = 0 \) respectively. It is now quite clearly that \( \lim_{x \to 0} A(x,y) = 0 \) for all \( y > 0 \) and \( \lim_{y \to 0} A(x,y) = 1 \) for all \( x > 1 \).

Example 8.18 (Switching Limits is Dangerous II). Let

\[
S_{m,n} := \frac{(-1)^{m+1}}{1 + \frac{m}{n}} = -\frac{\cos(\pi m)}{1 + \frac{m}{n}}.
\]

Then \( \lim_{m \to \infty} S_{m,n} = 0 \) (but not uniformly in \( n \)) so that \( \lim_{n \to \infty} \lim_{m \to \infty} S_{m,n} = 0 \). We also have \( \lim_{n \to \infty} S_{m,n} = (-1)^m \) and \( \lim_{m \to \infty} \lim_{n \to \infty} S_{m,n} = \lim_{m \to \infty} (-1)^m \) does not exist. In order to visualize better what is going on here let us again make the change of variables, \( x = \frac{1}{m} \) and \( y = \frac{1}{n} \), i.e. let

\[
S(x,y) = S_{m,n} = -x \cos(\pi/y) - y \cos(\pi/x)
\]

whose plot appears in Figure 8.18.

Example 8.19 (Switching Limits is Dangerous III). If

\[
S_{m,n} := \frac{(-1)^n}{m} + \frac{(-1)^m}{n} = -\frac{\cos(n \pi)}{m} - \frac{\cos(m \pi)}{n}
\]

and we let \( x = \frac{1}{m} \) and \( y = \frac{1}{n} \). Then

\[
S(x,y) = S_{m,n} = -x \cos(\pi/y) - y \cos(\pi/x)
\]

whose plot appears in Figure 8.19.
A plot of $-\frac{\cos(n\pi)}{m} - \frac{\cos(m\pi)}{n}$ in terms of the variables $\frac{1}{m}$ and $\frac{1}{n}$.

In this case, $\lim_{m,n \to \infty} S_{m,n} = 0$ exists but $\lim_{m \to \infty} S_{m,n}$ and $\lim_{n \to \infty} S_{m,n}$ do not exist!

**Remark 8.20.** However, if $\lim_{m,n \to \infty} S_{m,n}$ exists then we do always have,

$$\lim_{m,n \to \infty} S_{m,n} = \lim_{m \to \infty} \limsup_{n \to \infty} S_{m,n} = \lim_{m \to \infty} \liminf_{n \to \infty} S_{m,n}.$$ 

On the other hand if $\lim_{m,n \to \infty} S_{m,n}$ exists then we will have

$$\lim_{m,n \to \infty} S_{m,n} = \lim_{m \to \infty} \lim_{n \to \infty} S_{m,n} = \lim_{n \to \infty} \lim_{m \to \infty} S_{m,n}.$$ 

One way to avoid these types of behaviors is to assume $S_{m,n} \geq 0$ and is increasing in each index.

**Example 8.21 (Monotonicity is good).** If

$$S_{m,n} := \frac{1}{1 + \frac{1}{m} + \frac{1}{n}}$$

and we let $x = \frac{1}{m}$ and $y = \frac{1}{n}$. Then

$$S(x, y) = S_{m,n} = \frac{1}{1 + y + x^2}$$

whose plot appears in Figure 8.21.

This example is covered by Theorem 8.22 below.

**Theorem 8.22 (Equality of monotone iterated limits).** Suppose that $S_{m,n} \geq 0$ for all $m, n \in \mathbb{N}$ and $S_{m+1,n} \geq S_{m,n}$ and $S_{m,n+1} \geq S_{m,n}$ for all $m, n \in \mathbb{N}$. Then

$$L := \lim_{n \to \infty} \lim_{m \to \infty} S_{m,n} = \lim_{m \to \infty} \lim_{n \to \infty} S_{m,n} = \sup_{(m,n)} S_{m,n}$$

where all limits exist but may take on the value infinity. Moreover

$$\lim_{m,n \to \infty} S_{m,n} = L.$$ 

**Proof.** Because $S_{m,n}$ is increasing in each of its variables, we find,

$$\lim_{m \to \infty} \lim_{n \to \infty} S_{m,n} = \sup_{m \in \mathbb{N}} S_{m,n} \quad \text{and} \quad \lim_{n \to \infty} \lim_{m \to \infty} S_{m,n} = \sup_{n \in \mathbb{N}} S_{m,n}$$

and therefore Eq. (8.2) follows from Corollary 8.2 of the sup-sup Theorem. So it only remains to show $\lim_{m,n \to \infty} S_{m,n} = L$.

Cases 1. If $L = \infty$, then for any $N \in \mathbb{N}$, there exists $(m_N, n_N)$ such that $S_{m_N, n_N} \geq N$ and since $S_{m,n}$ is increasing in each of its variables it follows that

$$S_{m,n} \geq N \quad \text{for all} \quad m \geq m_N \quad \text{and} \quad n \geq n_N.$$
and it follows that \( \lim_{m,n \to \infty} S_{m,n} = \infty. \)

Case 2. \( L < \infty. \) In this case if \( \varepsilon > 0 \) is given, there exists \((m_\varepsilon, n_\varepsilon)\) such that \( S_{m_\varepsilon, n_\varepsilon} \geq L - \varepsilon. \) Since \( S_{m,n} \) is increasing in each of its variables it follows that

\[
L \geq S_{m,n} \geq L - \varepsilon \quad \text{for all } m \geq m_\varepsilon \text{ and } n \geq n_\varepsilon
\]

and \( |L - S_{m,n}| \leq \varepsilon \) for all \( m, n \geq m_\varepsilon \land n_\varepsilon. \) So by definition, \( \lim_{m,n \to \infty} S_{m,n} = L. \)

**Exercise 8.4.** If \( \{f_n(x)\}_{n=1}^\infty \) is a sequence of increasing continuous functions on \( \mathbb{R} \) such that \( f_n(x) \uparrow f(x) < \infty \) as \( n \to \infty, \) then \( f(x) \) is continuous and increasing as well.

**Exercise 8.5.** Show \( \lim_{m,n \to \infty} S_{m,n} = L \) if for all \( \varepsilon > 0 \) there exists \( M, N \in \mathbb{N} \) such

\[
|L - S_{m,n}| \leq \varepsilon \quad \text{for all } m \geq M \text{ and } n \geq N.
\]

**Lemma 8.23.** Suppose that \( \{S_{m,n}\}_{m,n=1}^\infty \) is Cauchy in the sense that for all \( \varepsilon > 0 \) there exists \( M, N \in \mathbb{N} \) such that

\[
|S_{m,n} - S_{m',n'}| \leq \varepsilon \quad \text{for all } m, m' \geq M \text{ and } n, n' \geq N.
\]

Then \( \lim_{m,n \to \infty} S_{m,n} = L \) exists.

**Proof.** Let \( s_n := S_{n,n}. \) Then the assumption shows that \( \{s_n\}_{n=1}^\infty \) is a Cauchy sequence and hence convergent. Let \( L := \lim_{n \to \infty} s_n. \) Now take \( m' = n' \) and then let \( n' \to \infty \) in Eq. (8.3) in order to learn,

\[
|S_{m,n} - L| \leq \varepsilon \quad \text{for all } m \geq M \text{ and } n \geq N.
\]

From this we conclude that \( \lim_{m,n \to \infty} S_{m,n} = L. \)

Another way to ensure that the iterated limits are equal is to assume some uniformity in one of the limits as in the next key theorem. [This theorem will be used in one guise or another repeatedly throughout these notes.]

**Theorem 8.24.** Suppose that \( \{S_{m,n}\}_{m,n=1}^\infty \) is a sequence of complex numbers (or more generally elements of a complete metric space \((X, \rho))\). Assume that

\[
S_{m,\infty} := \lim_{n \to \infty} S_{m,n} \quad \text{exists uniformly in } m \quad \text{and}
\]

\[
S_{\infty,n} := \lim_{m \to \infty} S_{m,n} \quad \text{exists pointwise in } n.
\]

Then \( \lim_{m \to \infty} \lim_{n \to \infty} S_{m,n}, \lim_{n \to \infty} \lim_{m \to \infty} S_{m,n}, \text{ and } \lim_{m,n \to \infty} S_{m,n} \) all exist and are equal, i.e.

\[
L := \lim_{m \to \infty} \lim_{n \to \infty} S_{m,n} = \lim_{n \to \infty} \lim_{m \to \infty} S_{m,n} = \lim_{m,n \to \infty} S_{m,n}.
\]

[In words, the existence of limits in both variables along with uniformity in one of the variables implies the iterated limits exists and agree.]

**Proof.** Let \( \varepsilon > 0 \) be given. Choose \( N \in \mathbb{N} \) such that

\[
\sup_m |S_{m,\infty} - S_{m,n}| \leq \varepsilon \quad \text{for all } n \geq N.
\]

Now choose \( M \in \mathbb{N} \) such that

\[
|S_{\infty,N} - S_{m,n}| \leq \varepsilon \quad \text{for all } m \geq M.
\]

Then for \( n \geq N \) and \( m \geq M \) we have,

\[
|S_{m,n} - S_{\infty,N}| \leq |S_{m,n} - S_{m,\infty}| + |S_{m,\infty} - S_{\infty,N}|
\]

\[
\leq |S_{m,n} - S_{m,\infty}| + |S_{m,\infty} - S_{m,n}| + |S_{m,n} - S_{\infty,N}|
\]

\[
\leq \varepsilon + \varepsilon + \varepsilon = 3\varepsilon.
\]

Therefore it follows that

\[
|S_{m,n} - S_{m',n'}| \leq 6\varepsilon \quad \text{for all } m, m' \geq M \text{ and } n, n' \geq N.
\]

Hence \( \{S_{m,n}\}_{m,n=1}^\infty \) is Cauchy and therefore by Lemma 8.23 we know that \( L := \lim_{m,n \to \infty} S_{m,n} \) exists, i.e. for all \( \varepsilon > 0 \) there exists \( M, N \in \mathbb{N} \) such that

\[
|S_{m,n} - L| \leq \varepsilon \quad \text{for all } m \geq M \text{ and } n \geq N.
\]

Letting \( m \to \infty \) above then shows,

\[
|S_{\infty,n} - L| \leq \varepsilon \quad \text{for all and } n \geq N
\]

and therefore \( \lim_{n \to \infty} S_{\infty,n} = L. \) Similarly one shows \( \lim_{m \to \infty} S_{m,\infty} = L \) as well.

**Metric space proof.** Let \( \varepsilon > 0 \) be given. Choose \( N \in \mathbb{N} \) such that

\[
\sup_m \rho(S_{m,\infty}, S_{m,n}) \leq \varepsilon \quad \text{for all } n \geq N.
\]

Now choose \( M \in \mathbb{N} \) such that

\[
\rho(S_{\infty,N}, S_{m,n}) \leq \varepsilon \quad \text{for all } m \geq M.
\]

Then for \( n \geq N \) and \( m \geq M \) we have,

\[
\rho(S_{m,n}, S_{\infty,N}) \leq \rho(S_{m,n}, S_{m,\infty}) + \rho(S_{m,\infty}, S_{\infty,N})
\]

\[
\leq \rho(S_{m,n}, S_{m,\infty}) + \rho(S_{m,\infty}, S_{m,n}) + \rho(S_{m,n}, S_{\infty,N})
\]

\[
\leq \varepsilon + \varepsilon + \varepsilon = 3\varepsilon.
\]

Therefore it follows that
\[ \rho(S_{m,n}, S_{m',n'}) \leq 6\varepsilon \text{ for all } m, m' \geq M \text{ and } n, n' \geq N. \]

Hence \( \{S_{m,n}\}_{m,n=1}^{\infty} \) is Cauchy and therefore by Lemma 8.23 we know that \( L := \lim_{m,n \to \infty} S_{m,n} \) exists, i.e. for all \( \varepsilon > 0 \) there exists \( M, N \in \mathbb{N} \) such that

\[ \rho(S_{m,n}, L) \leq \varepsilon \text{ for all } m \geq M \text{ and } n \geq N. \]

Letting \( m \to \infty \) above then shows,

\[ \rho(S_{\infty,n}, L) \leq \varepsilon \text{ for all } n \geq N \]

and therefore \( \lim_{n \to \infty} S_{\infty,n} = L \). Similarly one shows \( \lim_{m \to \infty} S_{m,\infty} = L \) as well.

**Remark 8.25.** The previous theorem is rather easy to understand intuitively as the following picture indicates the strategy of the proof.

### 8.3 Iterated Limits

**Theorem 8.26.** Let \( \{a_{m,n}\}_{m,n=1}^{\infty} \) be a sequence of complex numbers and assume that

\[ \lim_{m \to \infty} a_{m,n} = A_n \quad \text{exists uniformly in } n \]
\[ \lim_{n \to \infty} A_n = L \quad \text{exists} \]

then \( \lim_{m,n \to \infty} a_{m,n} = L \) and in particular,

\[ \lim_{m \to \infty} \lim_{n \to \infty} a_{m,n} = \lim_{n \to \infty} \lim_{m \to \infty} a_{m,n}. \]

**Idea.** The first assumption guarantees the rows of \( a_{m,*} \) converge for large \( m \). The second assertion say that \( A_n \) is a sequence of Cauchy sequences and therefore by Lemma 8.23 we know that \( \lim_{n \to \infty} A_n = L \) for large \( n \). Thus we must have

\[ \lim_{m,n \to \infty} a_{m,n} = L. \]

**Theorem 8.27.** Let \( \{a_{m,n}\}_{m,n=1}^{\infty} \) be a sequence of complex numbers and assume that

\[ \lim_{m \to \infty} a_{m,n} = a_n \quad \text{exists uniformly in } n \]
\[ \lim_{n \to \infty} a_{m,n} = b_m \quad \text{exists} \]
\[ \lim_{m \to \infty} b_m = B \quad \text{exists}. \]

Then, \( \lim_{n \to \infty} a_n = B \). In other words under the above conditions,

\[ \lim_{m \to \infty} \lim_{n \to \infty} a_{m,n} = \lim_{n \to \infty} \lim_{m \to \infty} a_{m,n}. \]

Moreover \( \lim_{m,n \to \infty} a_{m,n} = B \) as well.

**Idea.** As above we know that \( a_{m,*} \) converges uniformly for large \( m \). We are also given that \( \lim_{n \to \infty} a_{m,n} = B \) for large \( n \). Thus \( B = \lim_{n \to \infty} a_{m,n} = \lim_{n \to \infty} a_n \) for large \( m \). Hence we may now apply the previous theorem.

**[Details]** We have

\[ |B - a_n| \leq |B - b_m| + |b_m - a_{m,n}| + |a_{m,n} - a_n| \]
\[ \leq |B - b_m| + |b_m - a_{m,n}| + \sup_k |a_{m,k} - a_k| \]

and then taking \( \limsup_{n \to \infty} \) of this inequality shows,

\[ \limsup_{n \to \infty} |B - a_n| \leq |B - b_m| + \sup_k |a_{m,k} - a_k| \to 0 \quad \text{as } m \to \infty. \]

We now prove the second assertion. For this we have

\[ |B - a_{m,n}| \leq |B - a_n| + |a_n - a_{m,n}| \]
\[ \leq |B - a_n| + \sup_k |a_k - a_{m,k}|. \]

Thus given \( \varepsilon > 0 \) there exists \( M \in \mathbb{N} \) and \( N \in \mathbb{N} \) such that

\[ |B - a_{m,n}| \leq |B - a_n| + \sup_k |a_k - a_{m,k}| \leq \varepsilon + \varepsilon \]

for \( n \geq N \) and \( m \geq M \). This proves the stronger statement that \( \lim_{m,n \to \infty} a_{m,n} = B \), i.e. \( a_{m,n} \) is a convergent sequence as long as both \( m, n \) are sufficiently large.

**Exercise 8.6.** Suppose that \( f_n \to f \) uniformly and \( f_n \) is continuous for all \( n \), then \( f \) is continuous. [Use sequential notion of continuity here.]
8.4 Double Sums and Continuity of Sums

Here are a couple of very useful consequences of these theorems.

Theorem 8.28 (Monotone convergence theorem for sums). Let \( \{a_{k,n}\}_{k,n=1}^{\infty} \) be a sequence of non-negative numbers, assume that \( a_{k,n+1} \geq a_{k,n} \) for all \( k,n \). Then

\[
\lim_{n \to \infty} \sum_{k=1}^{\infty} a_{k,n} = \sum_{k=1}^{\infty} \lim_{n \to \infty} a_{k,n} = \lim_{m \to \infty} \sum_{k=1}^{m} a_{k,n}.
\]

Proof. Let \( S_{m,n} := \sum_{k=1}^{m} a_{k,n} \), then \( \{S_{m,n}\}_{m,n=1}^{\infty} \) satisfies the hypothesis of Theorem 8.22 and the conclusions now follows from that Theorem upon noting that

\[
\lim_{n \to \infty} \sum_{k=1}^{\infty} a_{k,n} = \lim_{n \to \infty} S_{m,n} \quad \text{and} \quad \lim_{k \to \infty} a_{k,n} = \lim_{m \to \infty} \lim_{n \to \infty} S_{m,n}.
\]

Theorem 8.29 (Tonelli’s Theorem for sums). Let \( \{a_{k,l}\}_{k,l=1}^{\infty} \) be any sequence of non-negative numbers, then

\[
\sum_{k=1}^{\infty} \sum_{l=1}^{\infty} a_{k,l} = \sum_{l=1}^{\infty} \sum_{k=1}^{\infty} a_{k,l}.
\]

Proof. Apply Theorem 8.22 with \( S_{m,n} := \sum_{k=1}^{m} \sum_{l=1}^{n} a_{k,l} \).

Theorem 8.30 (Dominated convergence theorem for sums). Let \( \{a_{k,n}\}_{k,n=1}^{\infty} \) be a sequence of complex numbers such that \( \lim_{n \to \infty} a_{k,n} = a_k \) exists for all \( n \) and there exists a summable dominating sequence, \( \{M_k\} \), such that \( |a_{k,n}| \leq M_k \) for all \( k,n \in \mathbb{N} \). Then

\[
\lim_{n \to \infty} \sum_{k=1}^{\infty} a_{k,n} = \sum_{k=1}^{\infty} \lim_{n \to \infty} a_{k,n} = \lim_{m \to \infty} \sum_{k=1}^{m} a_{k,n}.
\] (8.4)

Proof. Let \( S_{m,n} := \sum_{k=1}^{m} a_{k,n} \), then \( \{S_{m,n}\}_{m,n=1}^{\infty} \) satisfies the hypothesis of Theorem 8.27. Indeed,

\[
|S_{m,n} - \sum_{k=1}^{m} a_{k,n}| = |\sum_{k=m+1}^{\infty} a_{k,n}| \leq \sum_{k=m+1}^{\infty} |a_{k,n}| \leq \sum_{k=m+1}^{\infty} M_k
\]

and hence,

\[
\sup_{n} \left| S_{m,n} - \sum_{k=1}^{\infty} a_{k,n} \right| \leq \sum_{k=m+1}^{\infty} M_k \to 0 \quad \text{as} \quad m \to \infty.
\]

Thus the hypothesis of Theorem 8.27 are satisfied and so we may conclude,

\[
\lim_{m \to \infty} \lim_{n \to \infty} S_{m,n} = \lim_{n \to \infty} \lim_{m \to \infty} S_{m,n} = \lim_{m \to \infty} \lim_{n \to \infty} S_{m,n} = \sum_{k=1}^{\infty} a_k
\]

which is exactly Eq. (8.4).

Theorem 8.31 (Fubini’s Theorem for sums). Let \( \{a_{k,l}\}_{k,l=1}^{\infty} \) be any sequence of complex numbers. If \( \sum_{k,l=1}^{\infty} |a_{k,l}| < \infty \), then

\[
\sum_{k=1}^{\infty} \sum_{l=1}^{\infty} a_{k,l} = \sum_{k=1}^{\infty} a_{k,l} = \sum_{l=1}^{\infty} \sum_{k=1}^{\infty} a_{k,l}.
\]

8.5 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 \( 2^X \) for those \( \alpha \subset X \).

Definition 8.32. 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 8.33. Suppose that \( X = \{1,2,3,\ldots\} \) and \( a : X \to [0,\infty] \), then

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

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

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

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

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

$$\sum_{\alpha} a \leq \sum_{\alpha} a(n).$$

**Remark 8.34.** 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 $\{x : a(x) \geq \varepsilon\}$ must be finite for otherwise $\sum_{X} a = \infty$. Thus

$$\{x \in X : a(x) > 0\} = \bigcup_{k=1}^{\infty} \{x : a(x) \geq 1/k\}$$

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

**Lemma 8.35.** Suppose that $a, b : X \to [0, \infty]$ are two functions, then

$$\sum_{X} (a + b) = \sum_{X} a + \sum_{X} b$$

for all $\lambda \geq 0$.

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

$$\sum_{\alpha} (a + b) = \sum_{\alpha} a + \sum_{\alpha} b \leq \sum_{X} a + \sum_{X} b$$

which after taking sups over $\alpha$ shows that

$$\sum_{X} (a + b) \leq \sum_{X} a + \sum_{X} b.$$

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

$$\sum_{\alpha} a + \sum_{\beta} b \leq \sum_{\alpha \cup \beta} a + \sum_{\alpha \cup \beta} b = \sum_{\alpha \cup \beta} (a + b) \leq \sum_{X} (a + b).$$

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

$$\sum_{X} a + \sum_{X} b \leq \sum_{X} (a + b).$$

**Lemma 8.36.** Let $X$ and $Y$ be sets, $R \subset X \times Y$ and suppose that $a : R \to \mathbb{R}$ is a function. Let $x R := \{y \in Y : (x, y) \in R\}$ and $R_y := \{x \in X : (x, y) \in R\}$. Then

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

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

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

$$\sup_{x \in X} \sup_{y \in x R} a(x, y) = \sup_{x \in X} N_x \leq M. \tag{8.5}$$

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

$$a(x, y) \leq N_x \leq \sup_{x \in X} \sup_{y \in x R} a(x, y)$$

and therefore

$$M = \sup_{(x, y) \in R} a(x, y) \leq \sup_{x \in X} \sup_{y \in x R} a(x, y). \tag{8.6}$$

Equations (8.5) and (8.6) show that

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

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

**Theorem 8.37 (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 proofs.

**First proof.** Let

$$\mathcal{Y} := \{A \subset X : A \subset \subset X\}.$$

Then
Fig. 8.1. The \( x \) and \( y \) – slices of a set \( R \subset X \times Y \).

\[
\lim_{n \to \infty} \sum_X f_n = \sup_n \sum_X f_n = \sup \sum_{\alpha \in 2^X} \sum_{n \to \infty} f_n = \sup_{\alpha \in 2^X} \sup_n \sum_X f_n = \sum_X \lim f_n
\]

\[
= \sup_{\alpha \in 2^X} \sum f = \sum_X f.
\]

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

\[
S_n \leq S \leq m
\]

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

\[
A := \lim_{n \to \infty} \sum_X f_n \leq \sum_X f.
\]  

(8.7)

Noting that \( \sum_\alpha f_n \leq \sum_X f_n = S_n \leq A \) for all \( \alpha \subset \subset X \) and in particular,

\[
\sum_\alpha f_n \leq A \quad \text{for all } n \text{ and } \alpha \subset \subset X.
\]

Letting \( n \) tend to infinity in this equation shows that

\[
\sum_\alpha f \leq A \quad \text{for all } \alpha \subset \subset X
\]

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

\[
\sum_X f \leq A = \lim_{n \to \infty} \sum_X f_n
\]

which combined with Eq. (8.7) proves the theorem.

Lemma 8.38 (Fatou’s Lemma for Sums). Suppose that \( f_n : X \to [0, \infty] \) is a sequence of functions, then

\[
\sum_X \liminf_{n \to \infty} f_n \leq \liminf_{n \to \infty} \sum_X f_n.
\]

Proof. Define \( g_k := \inf_{n \geq k} f_n \) so that \( g_k \uparrow \liminf_{n \to \infty} f_n \) as \( k \to \infty \). Since \( g_k \leq f_n \) for all \( n \geq k \),

\[
\sum_X g_k \leq \sum_X f_n \quad \text{for all } n \geq k
\]

and therefore

\[
\sum_X g_k \leq \liminf_{n \to \infty} \sum_X f_n \quad \text{for all } k.
\]

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

\[
\sum_X \liminf_{n \to \infty} f_n = \sum_X \lim_{k \to \infty} g_k \leq \sum_X \lim_{n \to \infty} f_n
\]

Remark 8.39. If \( A = \sum_X a < \infty \), then for all \( \varepsilon > 0 \) there exists \( \alpha_\varepsilon \subset \subset X \) such that

\[
A \geq \sum_\alpha a \geq A - \varepsilon
\]

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

\[
\left| A - \sum_\alpha a \right| \leq \varepsilon
\]

(8.9)

for all \( \alpha \subset \subset X \) containing \( \alpha_\varepsilon \). Indeed, choose \( \alpha_\varepsilon \) so that \( \sum_\alpha a \geq A - \varepsilon \).

8.6 Sums of complex functions

Definition 8.40. 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 \( \alpha_\varepsilon \subset \subset X \) such that for all \( \alpha \subset \subset X \) containing \( \alpha_\varepsilon \) we have

\[
\left| A - \sum_\alpha a \right| \leq \varepsilon.
\]
The following lemma is left as an exercise to the reader.

**Lemma 8.41.** Suppose that \( a, b : X \to \mathbb{C} \) are two functions such that \( \sum_X a \) and \( \sum_X b \) exist, then \( \sum_X (a + \lambda b) = \sum_X a + \lambda \sum_X b \).

**Definition 8.42 (Summable).** We call a function \( a : X \to \mathbb{C} \) summable if
\[
\sum_X |a| < \infty.
\]

**Proposition 8.43.** 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
\[
|\sum_X a| \leq \sum_X |a|.
\]

**Proof.** If \( \sum_X |a| < \infty \), then \( \sum_X (\Re a)^\pm < \infty \) and \( \sum_X (\Im a)^\pm < \infty \) and hence by Remark 8.39 these sums exists in the sense of Definition 8.40. Therefore by Lemma 8.41, \( \sum_X a \) exists and
\[
\sum_X a = \sum_X (\Re a)^+ - \sum_X (\Re a)^- + i \left( \sum_X (\Im a)^+ - \sum_X (\Im a)^- \right).
\]

Conversely, if \( \sum_X |a| = \infty \) then, because \( |a| \leq |\Re a| + |\Im a| \), we must have
\[
\sum_X |\Re a| = \infty \text{ or } \sum_X |\Im a| = \infty.
\]

Thus it suffices to consider the case where \( a : X \to \mathbb{R} \) is a real function. Write \( a = a^+ - a^- \) where
\[
a^+(x) = \max(a(x), 0) \text{ and } a^-(x) = \max(-a(x), 0).
\]
Then \( |a| = a^+ + a^- \) and
\[
\sum_X |a| = \sum_X a^+ + \sum_X a^- = \infty
\]
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 \( X' = \{x \in X : a(x) \geq 0\} \), then we know that \( \sum_{X'} a = \infty \) which means there are finite subsets \( \alpha_n \subset X' \subset X \) such that \( \sum_{\alpha_n} a \geq n \) for all \( n \). Thus if \( \alpha \subset X \) 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 \( \rho \geq 0 \) and \( \theta \in \mathbb{R} \), then
\[
\left| \sum_X a \right| = \rho e^{-i\theta} \sum_X a = \sum_X e^{-i\theta} a
\]
\[
= \sum_X \Re [e^{-i\theta} a] \leq \sum_X (\Re [e^{-i\theta} a])^+
\]
\[
\leq \sum_X |\Re [e^{-i\theta} a]| \leq \sum_X |e^{-i\theta} a| \leq \sum_X |a|.
\]

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

**Remark 8.44.** Suppose that \( X = \mathbb{N} \) and \( a : \mathbb{N} \to \mathbb{C} \) is a sequence, then it is not necessarily true that
\[
\sum_{n=1}^{\infty} a(n) = \sum_{n \in \mathbb{N}} a(n).
\]
(8.11)

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. (8.11) is valid.

**Theorem 8.45 (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
\[
|f_n(x)| \leq g(x) \text{ for all } x \in X \text{ and } n \in \mathbb{N}
\]
and that \( g \) is summable. Then
\[
\lim_{n \to \infty} \sum_{x \in X} f_n(x) = \sum_{x \in X} f(x).
\]
(8.13)
Proof. Notice that \(|f| = \lim |f_n| \leq g\) 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,
\[
\sum_X (g \pm f) = \sum_X \lim \inf_{n \to \infty} (g \pm f_n) \leq \lim \inf_{n \to \infty} \sum_X (g \pm f_n) = \sum_X g + \lim \inf_{n \to \infty} \left( \pm \sum_X f_n \right).
\]
Since \(\lim \inf_{n \to \infty} (-a_n) = -\lim sup_{n \to \infty} a_n\), we have shown,
\[
\sum_X g \pm \sum_X f \leq \sum_X g + \left\{ \lim \inf_{n \to \infty} \sum_X f_n - \lim sup_{n \to \infty} \sum_X f_n \right\}
\]
and therefore
\[
\lim \sup_{n \to \infty} \sum_X f_n \leq \sum_X f \leq \lim \inf_{n \to \infty} \sum_X f_n.
\]
This shows that \(\lim_{n \to \infty} \sum_X f_n\) exists and is equal to \(\sum_X f\).

Proof. (Second Proof.) Passing to the limit in Eq. (8.12) shows that \(|f| \leq g\) and in particular that \(f\) is summable. Given \(\varepsilon > 0\), let \(\alpha \subset X\) such that
\[
\sum_{X \setminus \alpha} g \leq \varepsilon.
\]
Then for \(\beta \subset \subset X\) such that \(\alpha \subset \beta\),
\[
\left| \sum_{\beta} f - \sum_{\beta} f_n \right| = \left| \sum_{\beta} (f - f_n) \right|
\leq \sum_{\beta} |f - f_n| = \sum_{\alpha} |f - f_n| + \sum_{\beta \setminus \alpha} |f - f_n|
\leq \sum_{\alpha} |f - f_n| + 2 \sum_{\beta \setminus \alpha} g
\leq \sum_{\alpha} |f - f_n| + 2 \varepsilon.
\]
and hence that
\[
\left| \sum_{\beta} f - \sum_{\beta} f_n \right| \leq \sum_{\alpha} |f - f_n| + 2 \varepsilon.
\]
Since this last equation is true for all such \(\beta \subset \subset X\), we learn that
\[
\sum_X f - \sum_X f_n \leq \sum_{\alpha} |f - f_n| + 2 \varepsilon
\]
which then implies that
\[
\limsup_{n \to \infty} \sum_X f - \sum_X f_n \leq \limsup_{n \to \infty} \sum_{\alpha} |f - f_n| + 2 \varepsilon = 2 \varepsilon.
\]
Because \(\varepsilon > 0\) is arbitrary we conclude that
\[
\limsup_{n \to \infty} \sum_X f - \sum_X f_n = 0.
\]
which is the same as Eq. (8.13). ■

Remark 8.46. Theorem 8.45 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, \(|f_n| \leq g_n\) and \(\sum_X g_n \to \sum_X g\) as \(n \to \infty\). Then \(f\) is summable and Eq. (8.13) still holds. For the proof we use Fatou’s Lemma to again conclude
\[
\sum_X (g \pm f) = \sum_X \lim \inf_{n \to \infty} (g_n \pm f_n) \leq \lim \inf_{n \to \infty} \sum_X (g_n \pm f_n) = \sum_X g + \lim \inf_{n \to \infty} \left( \pm \sum_X f_n \right)
\]
and then proceed exactly as in the first proof of Theorem 8.45.

8.7 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 8.47. 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,
\[
\sum_F a = \sum_X a.
\]

Theorem 8.48 (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 \( A \subset X \times Y \). Then for any \( \alpha \subset X \) and \( \beta \subset Y \) such that \( A \subset \alpha \times \beta \), we have
\[ \sum_{A} a \leq \sum_{\alpha \times \beta} \sum_{\alpha} a \leq \sum_{\alpha \times \beta} \sum_{Y} a \leq \sum_{X} \sum_{Y} a, \]
i.e. \( \sum_{A} a \leq \sum_{X} \sum_{Y} a \). Taking the sup over \( A \) 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 \( \beta_{n}^{x} \subset \subset Y \) such that \( \beta_{n}^{x} \uparrow Y \) as \( n \uparrow \infty \) and
\[ \sum_{y \in Y} a(x, y) = \lim_{n \to \infty} \sum_{y \in \beta_{n}^{x}} a(x, y). \]
If \( \alpha \subset X \) is a given finite subset of \( X \), then
\[ \sum_{y \in Y} a(x, y) = \lim_{n \to \infty} \sum_{y \in \beta_{n}^{x}} a(x, y) \]
for all \( x \in \alpha \)
where \( \beta_{n} := \bigcup_{x \in \alpha} \beta_{n}^{x} \subset \subset Y \). Hence
\[ \sum_{x \in \alpha} \sum_{y \in Y} a(x, y) = \lim_{n \to \infty} \sum_{y \in \beta_{n}} \sum_{x \in \alpha} a(x, y) = \lim_{n \to \infty} \sum_{x \in \alpha} \sum_{y \in \beta_{n}} a(x, y) \]
\[ = \lim_{n \to \infty} \sum_{(x, y) \in \alpha \times \beta_{n}} a(x, y) \leq \sum_{X \times Y} a. \]
Since \( \alpha \) is arbitrary, it follows that
\[ \sum_{x \in X} \sum_{y \in Y} a(x, y) = \sup_{\alpha \subset X} \sum_{x \in \alpha} \sum_{y \in Y} a(x, y) \leq \sum_{X \times Y} a \]
which completes the proof. ■

Theorem 8.49 (Fubini’s Theorem for Sums). Now suppose that \( a : X \times Y \to \mathbb{C} \) is a summable function, i.e. by Theorem 8.48 any one of the following equivalent conditions hold:
1. \( \sum_{X \times Y} |a| < \infty \),
2. \( \sum_{X} \sum_{Y} |a| < \infty \) or
3. \( \sum_{X} \sum_{Y} |a| < \infty \). Then
\[ \sum_{X \times Y} a = \sum_{X} \sum_{Y} a = \sum_{X} \sum_{Y} a. \]

Proof. If \( a : X \to \mathbb{R} \) is real valued the theorem follows by applying Theorem 8.48 to \( a^{+} \) – 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 \( a \).

8.8 \( \ell^{p} \) – 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
\[ \|f\|_{p} := \left( \sum_{x \in X} |f(x)|^{p} \mu(x) \right)^{1/p} \]
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 8.50. 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) \( \|f + g\| \leq \|f\| + \|g\| \) for all \( f, g \in Z \).
3. (Positive definite) \( \|f\| = 0 \) implies \( f = 0 \).

A function \( p : Z \to [0, \infty) \) satisfying properties 1. and 2. but not necessarily 3. above will be called a semi-norm on \( Z \).

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 8.51. 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 8.57 below. ■
Proposition 8.52. 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

$$F(s) = \int_0^s f'(t) \, dt$$

and $G(t) = \int_0^t g'(t) \, dt$. Then for all $s,t \geq 0$,

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

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:

$$\int_0^{g(t)} (t - f(\sigma)) \, d\sigma = \int_0^t (t - f(g(\tau))) g'(\tau) \, d\tau = \int_0^t (t - \tau) g'(\tau) \, d\tau = \int_0^t g(\tau) \, d\tau = G(t).$$

**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 8.2, $[0, s] \times [0, t] \subset A_s \cup B_t$. (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

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

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

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

Combining this with $h(0) = 0$ we see that $h(s)$ takes its maximum at some point $s \in (0, g(t)]$ 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:

$$\int_0^{g(t)} (t - f(\sigma)) \, d\sigma = \int_0^t (t - f(g(\tau))) g'(\tau) \, d\tau = \int_0^t (t - \tau) g'(\tau) \, d\tau = \int_0^t g(\tau) \, d\tau = G(t).$$

**Definition 8.53.** 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:

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

(8.14)

**Lemma 8.54.** Let $p \in (1, \infty)$ and $q := \frac{p}{p - 1} \in (1, \infty)$ be the conjugate exponent. Then

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

(8.15)

with equality if and only if $t^q = s^p$. (See Example ?? below for a generalization of the inequality in Eq. (8.15).)
Proof. Let \( F(s) = \frac{r}{p} \) for \( p > 1 \). Then \( f(s) = s^{p-1} = t \) and \( g(t) = t^{p-1} = t^{q-1} \), wherein we have \( q-1 = p/(p-1)-1 = 1/(p-1) \). Therefore \( G(t) = t^q/q \) and hence by Proposition 8.52

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

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 8.52, 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 \( s = t^{1/p} \). Since

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

it follows from the first derivative test that

\[
\max h = \max \left\{ h(0), h \left( t^{1/p} \right) \right\} = \max \left\{ 0, \frac{t^q}{q} \right\} = \frac{t^q}{q}.
\]

So we have shown

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

Theorem 8.55 (Hölder’s inequality). Let \( p, q \in [1, \infty) \) be conjugate exponents. For all \( f, g : X \to \mathbb{F} \),

\[
\|fg\|_1 \leq \|f\|_p \cdot \|g\|_q.
\]  

(8.16)

If \( p \in (1, \infty) \) and \( f \) and \( g \) are not identically zero, then equality holds in Eq. (8.16) iff

\[
\left( \frac{|f|}{\|f\|_p} \right)^p = \left( \frac{|g|}{\|g\|_q} \right)^q.
\]  

(8.17)

Proof. The proof of Eq. (8.16) for \( p \in \{1, \infty\} \) is easy and will be left to the reader. The cases where \( \|f\|_p = 0 \) or \( \|g\|_q = 0 \) or \( \infty \) are easily dealt with and are also left to the reader. So we will assume that \( p \in (1, \infty) \) and \( 0 < \|f\|_p, \|g\|_q < \infty \). Letting \( s = |f(x)|/\|f\|_p \) and \( t = |g|/\|g\|_q \) in Lemma 8.54 implies

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

with equality iff

\[
\frac{|f(x)|^p}{\|f\|_p^p} = s^p = t^q = \frac{|g(x)|^q}{\|g\|_q^q}.
\]  

(8.18)

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

\[
\|fg\|_1 \leq \frac{1}{p} + \frac{1}{q} = 1
\]  

with equality if Eq. (8.18) holds for all \( x \in X \), i.e. iff Eq. (8.17) holds.

** Definition 8.56. For a complex number \( \lambda \in \mathbb{C} \), let

\[
\text{sgn}(\lambda) = \left\{ \begin{array}{ll}
\frac{\lambda}{|\lambda|} & \text{if } \lambda \neq 0 \\
0 & \text{if } \lambda = 0.
\end{array} \right.
\]

For \( \lambda, \mu \in \mathbb{C} \) we will write \( \text{sgn}(\lambda) \equiv \text{sgn}(\mu) \) if \( \text{sgn}(\lambda) = \text{sgn}(\mu) \) or \( \lambda \mu = 0 \).

Theorem 8.57 (Minkowski’s Inequality). If \( 1 \leq p \leq \infty \) and \( f, g \in \ell^p(\mu) \) then

\[
\|f + g\|_p \leq \|f\|_p + \|g\|_p.
\]  

(8.19)

Moreover, assuming \( f \) and \( g \) are not identically zero, equality holds in Eq. (8.19) iff

\[
\text{sgn}(f) \equiv \text{sgn}(g) \text{ when } p = 1 \text{ and }
f = cg \text{ for some } c > 0 \text{ when } p \in (1, \infty).
\]

Proof. For \( p = 1 \),

\[
\|f + g\|_1 = \sum_X |f + g| \mu \leq \sum_X (|f| \mu + |g| \mu) = \sum_X |f| \mu + \sum_X |g| \mu
\]

with equality iff

\[
|f| + |g| = |f + g| \iff \text{sgn}(f) \equiv \text{sgn}(g).
\]

For \( p = \infty \),

\[
\|f + g\|_\infty = \sup_X |f + g| \leq \sup_X (|f| + |g|) \leq \sup_X |f| + \sup_X |g| = \|f\|_\infty + \|g\|_\infty.
\]

Now assume that \( p \in (1, \infty) \). Since
it follows that
\[
\|f + g\|_p \leq 2\|f\|_p + 2\|g\|_p
\]
Eq. (8.19) is easily verified if \(\|f + g\|_p = 0\), so we may assume \(\|f + g\|_p > 0\). Multiplying the inequality,
\[
|f + g|^p = |f + g| |f + g|^{p-1} \leq |f| |f + g|^{p-1} + |g| |f + g|^{p-1}
\]
by \(\mu\), then summing on \(x\) and applying Holder’s inequality on each term gives
\[
\sum_X |f + g|^p \mu \leq \sum_X |f| |f + g|^{p-1} \mu + \sum_X |g| |f + g|^{p-1} \mu
\]
\[
\leq (\|f\|_p + \|g\|_p) \|f + g|^{p-1}\|_q^q.
\]
(8.21)

Since \(q(p-1) = p\), as in Eq. (8.14),
\[
\|f + g|^{p-1}\|_q^q = \sum_X (|f + g|^{p-1})^q \mu = \sum_X |f + g|^p \mu = \|f + g\|_p^p.
\]
(8.22)

Combining Eqs. (8.21) and (8.22) shows
\[
\|f + g\|_p \leq (\|f\|_p + \|g\|_p) \|f + g\|_p^p \quad (8.23)
\]
and solving this equation for \(\|f + g\|_p\) (making use of Eq. (8.14)) implies Eq. (8.19). Now suppose that \(f\) and \(g\) are not identically zero and \(p \in (1, \infty)\). Equality holds in Eq. (8.19) iff equality holds in Eq. (8.23) iff equality holds in Eq. (8.21) and Eq. (8.20). The latter happens iff
\[
\text{sgn}(f) = \text{sgn}(g)
\]
and
\[
\left( \frac{|f|}{\|f\|_p} \right)^p = \frac{|f + g|^p}{\|f + g\|_p} = \left( \frac{|g|}{\|g\|_p} \right)^p.
\]
(8.24)

wherein we have used
\[
\left( \frac{|f + g|^{p-1}}{\|f + g|^{p-1}\|_q^q} \right)^q = \frac{|f + g|^p}{\|f + g\|_p^p}.
\]
Finally Eq. (8.24) is equivalent to \(|f| = c|g|\) with \(c = (\|f\|_p/\|g\|_p) > 0\) and this equality along with \(\text{sgn}(f) = \text{sgn}(g)\) implies \(f = cg\). ■

8.9 Exercises

Exercise 8.7. 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
\[
D = \cap_{M=1}^{\infty} \cup_{N=1}^{\infty} \cap_{n \geq N} \{x \in X : f_n(x) \geq M\}.
\]
(8.25)

Exercise 8.8. 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 (8.25). (Hint: use the Cauchy criteria for convergence.)

8.9.1 Limit Problems

Exercise 8.9. Show \(\liminf_{n \to \infty} (-a_n) = -\limsup_{n \to \infty} a_n\).

Exercise 8.10. Suppose that \(\limsup_{n \to \infty} a_n = M \in \mathbb{R}\), 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 8.11. Show that
\[
\limsup_{n \to \infty} (a_n + b_n) \leq \limsup_{n \to \infty} a_n + \limsup_{n \to \infty} b_n
\]
(8.26)
provided that the right side of Eq. (8.26) 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 8.12. Suppose that \(a_n \geq 0\) and \(b_n \geq 0\) for all \(n \in \mathbb{N}\). Show
\[
\limsup_{n \to \infty} (a_n b_n) \leq \limsup_{n \to \infty} a_n \cdot \limsup_{n \to \infty} b_n,
\]
(8.27)
provided the right hand side of (8.27) is not of the form \(0 \cdot \infty\) or \(\infty \cdot 0\).

Exercise 8.13. Prove Lemma 8.11
8.9.2 Monotone and Dominated Convergence Theorem Problems

Exercise 8.15. Let \( M < \infty \), show there are polynomials \( p_n(t) \) and \( q_n(t) \) for \( n \in \mathbb{N} \) such that
\[
\lim_{n \to \infty} \sup_{0 \leq t \leq M} \left| \sqrt{t} - q_n(t) \right| = 0 \tag{8.28}
\]
and
\[
\lim_{n \to \infty} \sup_{|t| \leq M} |t| - p_n(t) = 0 \tag{8.29}
\]
using the following outline.

1. Let \( f(x) = \sqrt{1 - x^2} \) for \( |x| \leq 1 \) and use Taylor’s theorem with integral remainder (see Eq. ?? of Appendix ??), or analytic function theory if you know it, to show there are constants \( c_n > 0 \) for \( n \in \mathbb{N} \) such that
\[
\sqrt{1 - x^2} = 1 - \sum_{n=1}^{\infty} c_n x^n \quad \text{for all} \quad |x| < 1. \tag{8.30}
\]
2. Let \( \tilde{q}_n(x) := 1 - \sum_{n=1}^{m} c_n x^n \). Use (8.30) to show \( \sum_{n=1}^{\infty} c_n = 1 \) and conclude from this that
\[
\lim_{m \to \infty} \sup_{|x| \leq 1} |\sqrt{1 - x^2} - \tilde{q}_m(x)| = 0. \tag{8.31}
\]
3. Conclude that \( q_n(t) := \sqrt{M} \tilde{q}_n(1 - t/M) \) and \( p_n(t) := q_n(t^2) \) for \( n \in \mathbb{N} \) are polynomials verifying Eqs. (8.28) and (8.29) respectively.

Notation 8.58 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 8.16. 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 \( \lambda \in \mathbb{C} \) if for all sequences \( \{ u_n \} \to u_0 \) in \( U \setminus \{ u_0 \} \) 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 8.17. 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 \( u \to f(u, y) \) is continuous on \( U \).
2. There is a summable function \( g : Y \to [0, \infty) \) such that
\[
|f(u, y)| \leq g(y) \quad \text{for all} \quad y \in Y \quad \text{and} \quad u \in U.
\]
3. Show there are polynomials \( \tilde{q}_n(x) := \sqrt{M} \tilde{q}_n(1 - t/M) \) and \( p_n(t) := q_n(t^2) \) for \( n \in \mathbb{N} \) are polynomials verifying Eqs. (8.28) and (8.29) respectively.

Exercise 8.18. 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
\[
\left| \frac{\partial f(u, y)}{\partial u} \right| \leq g(y) \quad \text{for all} \quad y \in Y \quad \text{and} \quad u \in J.
\]
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 \( \sum_{y \in Y} |f(u, y)| < \infty \).
4) Let \( F(u) := \sum_{y \in Y} f(u, y) \), show \( F \) is differentiable on \( J \) and that
\[
\hat{F}(u) = \sum_{y \in Y} \frac{\partial}{\partial u} f(u, y).
\]

(Hint: Use the mean value theorem.)

Exercise 8.19 (Differentiation of Power Series). Suppose \( R > 0 \) and \( \{a_n\} \) 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 8.18 \( f(x) := \sum_{n=0}^{\infty} a_n x^n \) is continuously differentiable for \( x \in (-R, R) \) and
\[
f'(x) = \sum_{n=0}^{\infty} n a_n x^{n-1} = \sum_{n=1}^{\infty} n a_n x^{n-1}.
\]

4 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 \( U \).
Exercise 8.20. Show the functions
\[ e^x := \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad \sin x := \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \quad \text{and} \quad \cos x := \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} \]
are infinitely differentiable and they satisfy
\[ \frac{d}{dx} e^x = e^x \quad \text{with} \quad e^0 = 1 \]
\[ \frac{d}{dx} \sin x = \cos x \quad \text{with} \quad \sin(0) = 0 \]
\[ \frac{d}{dx} \cos x = -\sin x \quad \text{with} \quad \cos(0) = 1. \]

1. Use the product and the chain rule to show,
\[ \frac{d}{dx} \left[ e^{-x} e^{(x+\psi)} \right] = 0 \]
and conclude this, that \( e^{-x} e^{(x+\psi)} = e^{\psi} \) for all \( x, y \in \mathbb{R} \). In particular taking \( y = 0 \) this implies that \( e^{-x} = 1/e^x \) and hence that \( e^{(x+\psi)} = e^x e^{\psi} \).
Use this result to show \( e^x \uparrow \infty \) as \( x \uparrow \infty \) and \( e^x \downarrow 0 \) as \( x \downarrow -\infty \).
Remark: since \( e^x \geq \sum_{n=0}^{\infty} \frac{x^n}{n!} \) when \( x \geq 0 \), it follows that \( \lim_{x \to \infty} \frac{x^n}{n!} = 0 \) for any \( n \in \mathbb{N} \), i.e. \( e^x \) grows at a rate faster than any polynomial in \( x \) as \( x \to \infty \).
2. Use the product rule to show
\[ \frac{d}{dx} (\cos^2 x + \sin^2 x) = 0 \]
and use this to conclude that \( \cos^2 x + \sin^2 x = 1 \) for all \( x \in \mathbb{R} \).

Exercise 8.22. 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
\[ F(t,x) = \sum_{n=-\infty}^{\infty} a_n e^{-i n^2 t} e^{i n x}, \]
where as usual \( e^{ix} = \cos(x) + i \sin(x) \), this is motivated by replacing \( x \) in Eq. (8.35) by \( ix \) and comparing the result to Eqs. (8.36) and (8.37).

1. \( F(t,x) \) is continuous for \( (t,x) \in [0,\infty) \times \mathbb{R} \). \textbf{Hint:} Let \( Y = \mathbb{Z} \) and \( u = (t,x) \) and use Exercise 8.17.
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 \( x \in \mathbb{R} \). \textbf{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 \( \partial^2 F(t,x)/\partial x^2 \) via Exercise 8.18. In computing the \( t \) derivative, you should let \( \varepsilon > 0 \) and apply Exercise 8.18 with \( t = u = \varepsilon \) and then afterwards let \( \varepsilon \downarrow 0 \).
3. \( F \) satisfies the heat equation, namely
\[ \partial F(t,x)/\partial t = \partial^2 F(t,x)/\partial x^2 \] for \( t > 0 \) and \( x \in \mathbb{R} \).

8.9.3 \( \ell^p \) Exercises

Exercise 8.23. Generalize Proposition 8.52 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 \( g = f^{-1}, b = f(0) \geq 0, \)
\[ F(s) = \int_0^s f(s') ds' \] and \( G(t) = \int_0^t g(t') dt' \).
Then for all \( s, t \geq 0, \)
\[ st \leq F(s) + G(t) \leq F(s) + G(t) \]
and equality holds iff \( t = f(s) \). In particular, taking \( f(s) = e^s \), prove Young’s inequality stating
\[ st \leq e^s + (t \vee 1) \ln (t \vee 1) - (t \vee 1) \leq e^s + t \ln t - t, \]
where \( s \vee t := \min(s,t) \). \textbf{Hint:} Refer to Figures 8.3 and 8.4.

Exercise 8.24. Using differential calculus, prove the following inequalities
1. For \( y > 0 \), let \( g(x) := xy - e^x \) for \( x \in \mathbb{R} \). Use calculus to compute the maximum of \( g(x) \) and use this prove Young’s inequality;
\[ xy \leq e^x + y \ln y - y \quad \text{for} \ x \in \mathbb{R} \ 	ext{and} \ y > 0. \]
2. For \( p > 1 \) and \( y \geq 0 \), let \( g(x) := xy - x^p / p \) for \( x \geq 0 \). Again use calculus to compute the maximum of \( g(x) \) and show that your result gives the following inequality;
\[ xy \leq x^p + y^q \quad \text{for all} \ x, y \geq 0, \]
where \( q = \frac{p}{p-1} \), i.e. \( \frac{1}{q} = 1 - \frac{1}{p} \).
3. Suppose now that \( u : [0, \infty) \to [0, \infty) \) is a \( C^1 \) function such that: \( u(0) = 0 \), \( \lim_{x \to \infty} \frac{u(x)}{x} = \infty \), and \( u'(x) > 0 \) for all \( x > 0 \). Show

\[
xy \leq u(x) + v(y) \quad \text{for all } x, y \geq 0,
\]

where \( v(y) = y(u')^{-1}(y) - u\left((u')^{-1}(y)\right) \). **Hint:** consider the function, \( g(x) := xy - u(x) \).
Topological Considerations

9.1 Closed and Open Sets

Let \((X, d)\) be a metric space.

**Definition 9.1.** Let \((X, d)\) be a metric space. The **open ball** \(B(x, \delta) \subset X\) centered at \(x \in X\) with radius \(\delta > 0\) is the set

\[
B(x, \delta) := \{y \in X : d(x, y) < \delta\}.
\]

We will often also write \(B(x, \delta)\) as \(B_x(\delta)\). We also define the **closed ball** centered at \(x \in X\) with radius \(\delta > 0\) as the set \(C_x(\delta) := \{y \in X : d(x, y) \leq \delta\}\).

---

**Fig. 9.1.** Balls in \(\mathbb{R}^2\) corresponding to the 1 – norm, 2 – norm, 5 – norm, and \(\frac{1}{2}\) – “norm.”

**Fig. 9.2.** The ball in \(C([0,1], \mathbb{R})\) of radius 1/4 centered at \(f(x) = \sin (x^2)\) are all the continuous functions whose graphs lie between the green envelope.
Definition 9.2. A set $E \subset X$ is **bounded** if $E \subset B(x, R)$ for some $x \in X$ and $R < \infty$. A set $F \subset X$ is **closed** if every convergent sequence $(x_n)_{n=1}^{\infty}$ which is contained in $F$ has its limit back in $F$. A set $V \subset X$ is **open** iff $V^c$ is closed. We will write $F \subset X$ to indicate $F$ is a closed subset of $X$ and $V \subset \alpha X$ to indicate $V$ is an open subset of $X$. We also let $\tau_d$ denote the collection of open subsets of $X$ relative to the metric $d$.

**Lemma 9.3.** If $f : X \to \mathbb{R}$ is a continuous function and $k \in \mathbb{R}$, then the following sets are closed,

$$A := \{ x \in X : f(x) \leq k \}, \quad B := \{ x \in X : f(x) = k \}, \quad \text{and} \quad C := \{ x \in X : f(x) \geq k \}.$$ 

**Proof.** The proof that $A$, $B$, and $C$ are closed all go the same way so let me just check that $A$ is closed. To this end, suppose that $(x_n)_{n=1}^{\infty}$ is a sequence in $A$ such that $x := \lim_{n \to \infty} x_n$ exists in $X$. Since $x_n \in A$, $f(x_n) \leq k$ and therefore,

$$k \geq \lim_{n \to \infty} f(x_n) = f(x)$$

wherein the last equality we have used the definition of $f$ being continuous. By definition of $A$ it then follows that $x \in A$ and so we have checked that $A$ is closed.

**Example 9.4 (Closed Balls).** Closed balls are closed. Indeed, we have seen $f(y) := d(x, y)$ is continuous and therefore

$$C_x(\delta) := \{ y \in X : d(x, y) \leq \delta \} = \{ y \in X : f(y) \leq \delta \}$$

is a closed set. Notice that $\{ x \} = C_x(0)$ is a closed set for all $x \in X$.

**Example 9.5.** The following subsets of $\mathbb{C}$ are closed:

1. $\{ z \in \mathbb{C} : a \leq \text{Im } z \leq b \}$ for all $a \leq b$ in $\mathbb{R}$.
2. $\{ z \in \mathbb{C} : a \leq \text{Re } z \leq b \}$ for all $a \leq b$ in $\mathbb{R}$.
3. $\{ z \in \mathbb{C} : \text{Im } z = 0$ and $a \leq \text{Re } z \leq b \}$ for all $a \leq b$ in $\mathbb{R}$.

**Example 9.6 (Open Balls).** Open balls in metric spaces are open sets. Indeed let $f(y) := d(x, y)$, then

$$B_x(\delta)^c := X \setminus B_x(\delta) = \{ y \in X : d(x, y) \geq \delta \} = \{ y \in X : f(y) \geq \delta \}$$

which is closed since $f$ is continuous. Thus $B_x(\delta)$ is open.

**Theorem 9.7.** The closed subsets of $(X, d)$ have the following properties;

1. $X$ and $\emptyset$ are closed.
2. If $\{ C_\alpha \}_{\alpha \in I}$ is a collection of closed subsets of $X$, then $\cap_{\alpha \in I} C_\alpha$ is closed in $X$.
3. If $A$ and $B$ are closed sets then $A \cup B$ is closed.

**Proof.** 3. Let $\{ z_n \}_{n=1}^{\infty} \subset A \cup B$ such that $\lim_{n \to \infty} z_n = z$ exists. Then $z_n \in A$ i.o. or $z_n \in B$ i.o. For sake of definiteness say $z_n \in A$ i.o. in which case we may choose a subsequence, $w_k := z_{n_k} \in A$ for all $k$. Since $\lim_{k \to \infty} w_k = z$ and $A$ is closed it follows that $z \in A$ and hence $z \in A \cup B$. Thus we have shown $A \cup B$ is closed.

**Exercise 9.1.** Prove item 2. of Theorem 9.7. If $\{ C_\alpha \}_{\alpha \in I}$ is a collection of closed subsets of $X$, then $\cap_{\alpha \in I} C_\alpha$ is closed in $X$.

**Exercise 9.2.** Give an example of a collection of closed subsets, $\{ A_n \}_{n=1}^{\infty}$, of $\mathbb{C}$ such that $\cup_{n=1}^{\infty} A_n$ is not closed.

**Corollary 9.8.** Let $(X, d)$ be a metric space. Then the collection of open subsets, $\tau_d$, of $X$ satisfy;

1. $X$ and $\emptyset$ are in $\tau_d$.
2. $\tau_d$ is closed under taking arbitrary unions, i.e. if $\{ U_\alpha \}_{\alpha \in I}$ is a collection of open sets then $\cup_{\alpha \in I} U_\alpha$ is open.
3. $\tau_d$ is closed under taking finite intersections, i.e. if $U$ and $V$ are open sets then $U \cap V$ is open as well.

**Exercise 9.3.** Prove Corollary 9.8.

**Exercise 9.4.** Let $U$ be a subset of a metric space $(X, d)$. Show the following are equivalent;

1. $U$ is open.
2. for all $z \in U$ there exists $\rho > 0$ such that $B_z(\rho) \subset U$.
3. $U$ can be written as a union of open balls.

**Exercise 9.5 (Redundant now).** Show that $V \subset X$ is open iff for every $x \in V$ there is a $\delta > 0$ such that $B_x(\delta) \subset V$. In particular show $B_x(\delta)$ is open for all $x \in X$ and $\delta > 0$. **Hint:** by definition $V$ is not open iff $V^c$ is not closed.

**Exercise 9.6.** Let $(X, d)$ be a metric space and $\{ x_1, \ldots, x_n \}$ be a finite subset of $X$. Show $X \setminus \{ x_1, \ldots, x_n \}$ is an open subset.

**Exercise 9.7.** Let $(X, d)$ be a complete metric space. Let $A \subset X$ be a subset of $X$ viewed as a metric space using $d|_{A \times A}$. Show that $(A, d|_{A \times A})$ is complete iff $A$ is a closed subset of $X$. 

Therefore, 
\[ |d_A(x) - d_A(y)| \leq d(x,y) \quad \forall x, y \in X \] (9.1)
and in particular if \( x_n \to x \) in \( X \) then \( d_A(x_n) \to d_A(x) \) as \( n \to \infty \). Moreover the set \( F_\varepsilon := \{ x \in X | d_A(x) \geq \varepsilon \} \) is closed in \( X \).

**Proof.** Let \( a \in A \) and \( x, y \in X \), then
\[ d_A(x) \leq d(x, a) \leq d(x, y) + d(y, a). \]

Take the infimum over \( a \) in the above equation shows that
\[ d_A(x) \leq d(x, y) + d_A(y) \quad \forall x, y \in X. \]

Therefore, \( d_A(x) - d_A(y) \leq d(x, y) \) and by interchanging \( x \) and \( y \) we also have that
\[ d_A(y) - d_A(x) \leq d(x, y) \]
which implies Eq. (9.1). If \( x_n \to x \) in \( X \), then by Eq. (9.1),
\[ |d_A(x) - d_A(x_n)| \leq d(x, x_n) \to 0 \text{ as } n \to \infty \]
so that \( \lim_{n \to \infty} d_A(x_n) = d_A(x) \). Now suppose that \( \{ x_n \}_{n=1}^\infty \subset F_\varepsilon \) and \( x_n \to x \) in \( X \), then
\[ d_A(x) = \lim_{n \to \infty} d_A(x_n) \geq \varepsilon \]
for all \( n \). This shows that \( x \in F_\varepsilon \) and hence \( F_\varepsilon \) is closed. ■

**Definition 9.10.** A subset \( A \subset X \) is a **neighborhood** of \( x \) if there exists an open set \( V \subset X \) such that \( x \in V \subset A \). We will say that \( A \subset X \) is an open **neighborhood** of \( x \) if \( A \) is open and \( x \in A \).

**Example 9.11.** Let \( x \in X \) and \( \delta > 0 \), then \( C_x(\delta) \) and \( B_x(\delta)^C \) are closed subsets of \( X \). For example if \( \{ y_n \}_{n=1}^\infty \subset C_x(\delta) \) and \( y_n \to y \in X \), then \( d(y_n, x) \leq \delta \) for all \( n \) and using Corollary 6.46 it follows \( d(y, x) \leq \delta \), i.e. \( y \in C_x(\delta) \). A similar proof shows \( B_x(\delta)^C \) is closed, see Exercise 9.5.

**Lemma 9.12 (Approximating open sets from the inside by closed sets).** Let \( U \subset X \) be an open set and \( F_\varepsilon := \{ x \in X | d_U(x) \geq \varepsilon \} \subset X \) be as in Lemma 9.9. Then \( F_\varepsilon \cup U \) as \( \varepsilon \downarrow 0 \).

**Proof.** It is clear that \( d_U(x) = 0 \) for \( x \in U^C \) so that \( F_\varepsilon \subset U \) for each \( \varepsilon > 0 \) and hence \( \cup_{\varepsilon > 0} F_\varepsilon \subset U \). Now suppose that \( x \in U \). By Exercise 9.5 there exists an \( \varepsilon > 0 \) such that \( B_x(\varepsilon) \subset U \), i.e. \( d(x, y) \geq \varepsilon \) for all \( y \in U^C \). Hence \( x \in F_\varepsilon \) and we have shown that \( U \subset \cup_{\varepsilon > 0} F_\varepsilon \). Finally it is clear that \( F_\varepsilon \subset F_{\varepsilon'} \) whenever \( \varepsilon' \leq \varepsilon \). ■

**Definition 9.13.** Given a set \( A \) contained in a metric space \( X \), let \( \bar{A} \subset X \) be the **closure** of \( A \) defined by
\[ \bar{A} := \{ x \in X : \exists \{ x_n \} \subset A \exists \ x = \lim_{n \to \infty} x_n \}. \]

That is to say \( \bar{A} \) contains all limit points of \( A \). We say \( A \) is dense in \( X \) if \( A = X \), i.e. every element \( x \in X \) is a limit of a sequence of elements from \( A \). A metric space is said to be **separable** if it contains a countable dense subset, \( D \).

**Lemma 9.14.** For any \( A \subset X \), then
1. \( A = \bar{A} \) if \( A \) is closed.
2. \( \bar{A} = \{ x : d_A(x) = 0 \} \) and \( \bar{A} \) is closed.
3. \( \bar{A} = \{ x \in X : A \cap B_x(\rho) \neq \emptyset \ \forall \ \rho > 0 \}. \)
4. \( d_A(x) > 0 \) for all \( x \in A^C \) if \( A \) is closed.

**Proof.** 1. We always have \( A \subset \bar{A} \). If \( A \) is closed we can not leave \( A \) by taking limits and hence \( \bar{A} \subset A \), i.e. \( A = \bar{A} \) if \( A \) is closed.

2. Let \( F := \{ x : d_A(x) = 0 \} \) which is a closed set since \( d_A \) is continuous. If \( x \in F \) (i.e. \( d_A(x) = 0 \)), there exists \( x_n \in A \) such that \( d(x, x_n) \leq 1/n \) for all \( n \in \mathbb{N} \). Thus \( \lim_{n \to \infty} x_n = x \) and so \( x \in \bar{A} \). This shows \( F \subset \bar{A} \). Conversely if \( x \in A \), there exists \( \{ x_n \} \subset A \) such that \( \lim_{n \to \infty} x_n = x \) and so
\[ d_A(x) = d_A \left( \lim_{n \to \infty} x_n \right) = \lim_{n \to \infty} d_A(x_n) = \lim_{n \to \infty} 0 = 0 \]
which shows \( x \in F \).

3. Since \( A \cap B_x(\rho) \neq \emptyset \) happens iff \( d_A(x) < \rho \) we see that \( A \cap B_x(\rho) \neq \emptyset \ \forall \ \rho > 0 \) iff \( d_A(x) < \rho \) for all \( \rho > 0 \), i.e. \( d_A(x) = 0 \).

4. If \( A \) is closed then \( A = \bar{A} = \{ x \in X : d_A(x) = 0 \} \) and therefore \( A^C = \{ x \in X : d_A(x) > 0 \} \). ■

**Exercise 9.8.** If \( A \) is a non-empty subset of \( X \), then \( d_A = d_{\bar{A}} \).

**Exercise 9.9.** Given \( A \subset X \), show \( \bar{A} \) is a closed set and in fact
\[ \bar{A} = \cap \{ F : A \subset F \subset X \text{ with } F \text{ closed} \}. \] (9.2)

That is to say \( \bar{A} \) is the smallest closed set containing \( A \).

**Definition 9.15.** Let \( (X, d) \) be a metric space and \( A \) be a subset of \( X \).

1. The **closure** of \( A \) is the smallest closed set \( \bar{A} \) containing \( A \), i.e.
\[ \bar{A} := \cap \{ F : A \subset F \subset X \}. \]

(Because of Exercise 9.9 this is consistent with Definition 9.13 for the closure of a set in a metric space.)
2. The interior of $A$ is the largest open set $A^o$ contained in $A$, i.e.

$$A^o = \{ V \in \tau : V \subset A \}.$$  

3. $A \subset X$ is a neighborhood of a point $x \in X$ if $x \in A^o$.

4. The accumulation points of $A$ is the set

$$\text{acc}(A) = \{ x \in X : V \cap (A \setminus \{ x \}) \neq \emptyset \text{ for all } V \in \tau_x \}.$$  

5. The boundary of $A$ is the set $\text{bd}(A) := \overline{A} \setminus A^o$.

6. $A$ is dense in $X$ if $\overline{A} = X$ and $X$ is said to be separable if there exists a countable dense subset of $X$.

**Lemma 9.16.** Let $(X, d)$ be a metric space and $A$ be a subset of $X$, then

$$A^o = \{ x \in X : B_x(\rho) \subset A \text{ for some } \rho > 0 \}.$$  

**Proof.** Let $V := \{ x \in X : B_x(\rho) \subset A \text{ for some } \rho > 0 \}$. If $B_x(\rho) \subset A$ and $y \in B_x(\rho)$ and $\delta = \rho - d(x, y)$, then $B_y(\delta) \subset B_x(\rho) \subset A$ which shows that $y \in V$, i.e. $B_x(\rho) \subset V$. Thus we may write

$$V = \bigcup \{ B_x(\rho) : B_x(\rho) \subset A \}.$$  

This shows that $V$ is an open subset of $A$. Moreover if $W$ is another open subset of $A$, then

$$W = \bigcup \{ B_x(\rho) : B_x(\rho) \subset W \} \subset \bigcup \{ B_x(\rho) : B_x(\rho) \subset A \} = V$$

so that $V$ is the largest open subset contained in $A$. This completes the proof.

**Exercise 9.10.** Let $(X, d)$ be the normed space and $d(y, x) := \| y - x \|$. Show:

1. $C_x(\rho^o) = B_x(\rho)$,
2. $B_x(\rho) = C_x(\rho)$,
3. $\text{bd}(C_x(\rho)) = \text{bd}(C_x(\rho)) = \{ y \in X : \| y - x \| = \rho \}$.

**Example 9.17 (Words of Caution).** Let $(X, d)$ be a metric space. It is always true that $C_x(\varepsilon^c) \subset C_x(\varepsilon)$ since $C_x(\varepsilon)$ is a closed set containing $B_x(\varepsilon)$. However, it is not always true that $B_x(\varepsilon^c) = C_x(\varepsilon)$. For example let $X = \{ 1, 2 \}$ and $d(1, 2) = 1$, then $B_1(1) = \{ 1 \}$, $B_1(1) = \{ 1 \}$ while $C_1(1) = X$. For another counterexample, take

$$X = \{ (x, y) \in \mathbb{R}^2 : x = 0 \text{ or } x = 1 \}$$

with the usually Euclidean metric coming from the plane. Then

$$B_{(0, 0)}(1) = \{ (0, y) \in \mathbb{R}^2 : |y| < 1 \},$$
$$\overline{B}_{(0, 0)}(1) = \{ (0, y) \in \mathbb{R}^2 : |y| \leq 1 \},$$
$$C_{(0, 0)}(1) = \overline{B}_{(0, 0)}(1) \cup \{(1, 0)\}.$$  

**Exercise 9.11.** Suppose that $A$ and $B$ are subsets of a metric space, show $A \cup B = \overline{A} \cup \overline{B}$.

**Exercise 9.12.** Given an example showing that $\cup_{n=1}^{\infty} A_n$ need not be equal to $\mathbb{R}$.

**Remark 9.18.** The relationships between the interior and the closure of a set are:

$$(A^o)^c = \bigcap \{ V^c : V \in \tau \text{ and } V \subset A \} = \bigcap \{ C : C \text{ is closed } C \supseteq A^c \} = \overline{A}^c$$

and similarly, $\overline{A}^c = (A^c)^o$.

**Proposition 9.19.** If $A$ is a subset of a metric space, $(X, d)$, then $\text{bd}(A)$ may be computed using either:

1. $\text{bd}(A) = \overline{A} \cap \overline{A}^c$,
2. $\text{bd}(A) = \{ x \in X : d_A(x) = 0 = d_{A^c}(x) \}$,
3. $\text{bd}(A) = \{ x \in X : B_x(\rho) \subset A \neq \emptyset \neq B_x(\rho) \cap A^c \text{ for all } \rho > 0 \}$, or
4. $\text{bd}(A) = \{ x \in X : \exists \{ x_n \} \subset A \text{ and } \{ y_n \} \subset A^c \exists \lim_{n \to \infty} x_n = x = \lim_{n \to \infty} y_n \}$.

**Proof.** From the definition of $\text{bd}(A)$ and the relation, $(\overline{A})^c = (A^o)^c$, we find

$$\text{bd}(A) = \overline{A} \setminus A^o = \overline{A} \cap (A^o)^c = \overline{A} \cap \overline{A}^c.$$  

Item 2 now follow from item 1. and the fact that $\overline{A} = \{ x \in X : d_A(x) = 0 \}$. I leave it to the reader to check that $2. \implies 3. \implies 4. \implies 1.$

**Exercise 9.13.** If $D$ is a dense subset of a metric space $(X, d)$ and $E \subset X$ is a subset such that to every point $x \in D$ there exists $\{ x_n \}_{n=1}^{\infty} \subset E$ with $x = \lim_{n \to \infty} x_n$, then $E$ is also a dense subset of $X$. If points in $E$ well approximate every point in $D$ and the points in $D$ well approximate the points in $X$, then the points in $E$ also well approximate all points in $X$.

**Exercise 9.14.** Suppose $(X, d)$ is a metric space which contains an uncountable subset $A \subset X$ with the property that there exists $\varepsilon > 0$ such that $d(a, b) \geq \varepsilon$ for all $a, b \in A$ with $a \neq b$. Show that $(X, d)$ is not separable.

**Exercise 9.15.** Let $Y = BC(\mathbb{R}, \mathbb{C})$ be the Banach space of compact valued functions on $\mathbb{R}$ equipped with the uniform norm, $||f||_u := \sup_{x \in \mathbb{R}} |f(x)|$. Further let $C_0(\mathbb{R}, \mathbb{C})$ denote those $f \in C(\mathbb{R}, \mathbb{C})$ that vanish at infinity, i.e. $\lim_{x \to \pm \infty} f(x) = 0$. Also let $C_c(\mathbb{R}, \mathbb{C})$ denote those $f \in C(\mathbb{R}, \mathbb{C})$ with compact support, i.e. there exists $N < \infty$ such that $f(x) = 0$ if $|x| \geq N$. Show $C_0(\mathbb{R}, \mathbb{C})$ is a closed subspace of $Y$ and that

$$C_c(\mathbb{R}, \mathbb{C}) = C_0(\mathbb{R}, \mathbb{C}).$$
9.2 Continuity Revisited

Suppose now that \((X, \rho)\) and \((Y, d)\) are two metric spaces and \(f : X \to Y\) is a function.

**Definition 9.20.** A function \(f : X \to Y\) is **continuous at** \(x \in X\) if for all \(\varepsilon > 0\) there is a \(\delta > 0\) such that

\[
d(f(x), f(x')) < \varepsilon \quad \text{provided that} \quad \rho(x, x') < \delta.
\]

(9.4)

The function \(f\) is said to be **continuous** if \(f\) is continuous at all points \(x \in X\).

The following lemma gives two other characterizations of continuity of a function at a point.

**Lemma 9.21 (Local Continuity Lemma).** Suppose that \((X, \rho)\) and \((Y, d)\) are two metric spaces and \(f : X \to Y\) is a function defined in a neighborhood of a point \(x \in X\). Then the following are equivalent:

1. \(f\) is continuous at \(x \in X\).
2. For all neighborhoods \(A \subseteq Y\) of \(f(x)\), \(f^{-1}(A)\) is a neighborhood of \(x \in X\).
3. For all sequences \(\{x_n\}_{n=1}^{\infty} \subseteq X\) such that \(x = \lim_{n \to \infty} x_n\), \(\{f(x_n)\}\) is convergent in \(Y\) and

\[
\lim_{n \to \infty} f(x_n) = f \left( \lim_{n \to \infty} x_n \right).
\]

**Proof.** \(1 \implies 2\). If \(A \subseteq Y\) is a neighborhood of \(f(x)\), there exists \(\varepsilon > 0\) such that \(B_{f(x)}(\varepsilon) \subseteq A\) and because \(f\) is continuous there exists a \(\delta > 0\) such that Eq. (9.4) holds. Therefore

\[
B_x(\delta) \subseteq f^{-1} \left( B_{f(x)}(\varepsilon) \right) \subseteq f^{-1}(A)
\]

showing \(f^{-1}(A)\) is a neighborhood of \(x\).

\(2 \implies 3\). Suppose that \(\{x_n\}_{n=1}^{\infty} \subseteq X\) and \(x = \lim_{n \to \infty} x_n\). Then for any \(\varepsilon > 0\), \(B_{f(x)}(\varepsilon) \subseteq f^{-1}(A)\) is a neighborhood of \(f(x)\) and so \(f^{-1} \left( B_{f(x)}(\varepsilon) \right) \subseteq f^{-1}(A)\). Let \(x_n \to x\), it follows that \(x_n \in B_{f(x)}(\delta) \subseteq f^{-1} \left( B_{f(x)}(\varepsilon) \right)\) for a.a. \(n\) and this implies \(f(x_n) \in B_{f(x)}(\varepsilon)\) for a.a. \(n\), i.e. \(d(f(x), f(x_n)) < \varepsilon\) for a.a. \(n\). Since \(\varepsilon > 0\) is arbitrary it follows that \(\lim_{n \to \infty} f(x_n) = f(x)\).

\(3 \implies 1\). We will show not 1. \(\implies\) not 3. If \(f\) is not continuous at \(x\), there exists an \(\varepsilon > 0\) such that for all \(n \in \mathbb{N}\) there exists a point \(x_n \in X\) with \(\rho(x_n, x) < \frac{\varepsilon}{2}\) yet \(d(f(x_n), f(x)) \geq \varepsilon\). Hence \(x_n \to x\) as \(n \to \infty\) yet \(f(x_n)\) does not converge to \(f(x)\).

Here is a global version of the previous lemma.

**Lemma 9.22 (Global Continuity Lemma).** Suppose that \((X, \rho)\) and \((Y, d)\) are two metric spaces and \(f : X \to Y\) is a function defined on all of \(X\). Then the following are equivalent:

1. \(f\) is continuous.
2. \(f^{-1}(V) \in \tau_d\) for all \(V \in \tau_d\), i.e. \(f^{-1}(V)\) is open in \(X\) if \(V\) is open in \(Y\).
3. \(f^{-1}(C)\) is closed in \(X\) if \(C\) is closed in \(Y\).
4. For all convergent sequences \(\{x_n\} \subseteq X\), \(\{f(x_n)\}\) is convergent in \(Y\) and

\[
\lim_{n \to \infty} f(x_n) = f \left( \lim_{n \to \infty} x_n \right).
\]

**Proof.** Since \(f^{-1}(A^c) = [f^{-1}(A)]^c\), it is easily seen that 2. and 3. are equivalent. So because of Lemma 9.21 it only remains to show 1. and 2. are equivalent. If \(f\) is continuous and \(V \subseteq Y\) is open, then for every \(x \in f^{-1}(V)\), \(V\) is a neighborhood of \(f(x)\) and so \(f^{-1}(V)\) is a neighborhood of \(x\). Hence \(f^{-1}(V)\) is a neighborhood of all of its points and from this and Exercise 9.5 it follows that \(f^{-1}(V)\) is open. Conversely, if \(x \in X\) and \(A \subseteq Y\) is a neighborhood of \(f(x)\) then there exists \(V \subseteq X\) such that \(f(x) \in V \subseteq A\). Hence \(x \in f^{-1}(V) \subseteq f^{-1}(A)\) and by assumption \(f^{-1}(V)\) is open showing \(f^{-1}(A)\) is a neighborhood of \(x\). Therefore \(f\) is continuous at \(x\) and since \(x \in X\) was arbitrary, \(f\) is continuous.

**Exercise 9.16.** Use Example 9.28 and Lemma 9.22 to recover the results of Example 9.11.

The next result shows that there are lots of continuous functions on a metric space \((X, d)\).

**Lemma 9.23 (Urysohn’s Lemma for Metric Spaces).** Let \((X, d)\) be a metric space and suppose that \(A\) and \(B\) are two disjoint closed subsets of \(X\). Then

\[
f(x) = \frac{d_B(x)}{d_A(x) + d_B(x)} \quad \text{for} \quad x \in X
\]

defines a continuous function, \(f : X \to [0, 1]\), such that \(f(x) = 1\) for \(x \in A\) and \(f(x) = 0\) if \(x \in B\).

**Proof.** By Lemma 9.9 \(d_A\) and \(d_B\) are continuous functions on \(X\). Since \(A\) and \(B\) are closed, \(d_A(x) > 0\) if \(x \notin A\) and \(d_B(x) > 0\) if \(x \notin B\). Since \(A \cap B = \emptyset\), \(d_A(x) + d_B(x) > 0\) for all \(x\) and \((d_A + d_B)^{-1}\) is continuous as well. The remaining assertions about \(f\) are all easy to verify.

Sometimes Urysohn’s lemma will be used in the following form. Suppose \(F \subseteq V \subseteq X\) with \(F\) being closed and \(V\) being open, then there exists \(f \in C(X, [0, 1])\) such that \(f = 1\) on \(F\) while \(f = 0\) on \(V^c\). This of course follows from Lemma 9.23 by taking \(A = F\) and \(B = V^c\).
9.3 Metric spaces as topological spaces (Not required Reading!)

Let \((X,d)\) be a metric space and let \(\tau = \tau_d\) denote the collection of open subsets of \(X\). (Recall \(V \subset X\) is open iff \(V^c\) is closed iff for all \(x \in V\) there exists an \(\varepsilon = \varepsilon_x > 0\) such that \(B(x,\varepsilon_x) \subset V\) iff \(V\) can be written as a (possibly uncountable) union of open balls.) Although we will stick with metric spaces in this chapter, it will be useful to introduce the definitions needed here in the more general context of a general “topological space,” i.e., a space equipped with a collection of “open sets.”

**Definition 9.24 (Topological Space).** Let \(X\) be a set. A **topology** on \(X\) is a collection of subsets \(\tau\) of \(X\) with the following properties:

1. \(\tau\) contains both the empty set \(\emptyset\) and \(X\).
2. \(\tau\) is closed under arbitrary unions.
3. \(\tau\) is closed under finite intersections.

The elements \(V \in \tau\) are called **open** subsets of \(X\). A subset \(F \subset X\) is said to be **closed** if \(F^c\) is open. I will write \(V \subseteq X\) to indicate that \(V \subset X\) and \(V \in \tau\) and similarly \(F \subseteq X\) will denote \(F \subset X\) and \(F\) is closed. Given \(x \in X\) we say that \(V \subset X\) is an **open neighborhood** of \(x\) if \(V \in \tau\) and \(x \in V\). Let \(\tau_x = \{V \in \tau : x \in V\}\) denote the collection of open neighborhoods of \(x\).

**Definition 9.25 (Continuity at a point in topological terms).** Let \((X,\tau_X)\) and \((Y,\tau_Y)\) be topological spaces. A function \(f : X \to Y\) is **continuous at a point** \(x \in X\) if for every open neighborhood \(V\) of \(f(x)\) there is an open neighborhood \(U\) of \(x\) such that \(U \subset f^{-1}(V)\). See Figure 9.3.

**Definition 9.26 (Global continuity in topological terms).** Let \((X,\tau_X)\) and \((Y,\tau_Y)\) be topological spaces. A function \(f : X \to Y\) is **continuous** if

\[
 f^{-1}(\tau_Y) := \{f^{-1}(V) : V \in \tau_Y\} \subset \tau_X.
\]

We will also say that \(f\) is \(\tau_X/\tau_Y\) –continuous or \((\tau_X,\tau_Y)\) – continuous. Let \(C(X,Y)\) denote the set of continuous functions from \(X\) to \(Y\).

**Exercise 9.17.** Show \(f : X \to Y\) is continuous (Definition ??) iff \(f\) is continuous at all points \(x \in X\).

**Exercise 9.18.** Show \(f : X \to Y\) is continuous iff \(f^{-1}(C)\) is closed in \(X\) for all closed subsets \(C\) of \(Y\).

**Definition 9.27.** A map \(f : X \to Y\) between topological spaces is called a **homeomorphism** provided that \(f\) is bijective, \(f\) is continuous and \(f^{-1} : Y \to X\) is continuous. If there exists \(f : X \to Y\) which is a homeomorphism, we say that \(X\) and \(Y\) are homeomorphic. (As topological spaces \(X\) and \(Y\) are essentially the same.)

**Example 9.28.** The function \(d_A\) defined in Lemma 9.9 is continuous for each \(A \subset X\). In particular, if \(A = \{x\}\), it follows that \(y \in X \to d(y,x)\) is continuous for each \(x \in X\).

9.4 Exercises

**Exercise 9.19.** Show that \((X,d)\) is a complete metric space iff every sequence \(\{x_n\}_{n=1}^{\infty} \subset X\) such that \(\sum_{n=1}^{\infty} d(x_n,x_{n+1}) < \infty\) is a convergent sequence in \(X\). You may find it useful to prove the following statements in the course of the proof.

1. If \(\{x_n\}\) is Cauchy sequence, then there is a subsequence \(y_j := x_{n_j}\) such that \(\sum_{j=1}^{\infty} d(y_{j+1},y_j) < \infty\).
2. If \(\{x_n\}_{n=1}^{\infty}\) is Cauchy and there exists a subsequence \(y_j := x_{n_j}\) of \(\{x_n\}\) such that \(x = \lim_{j \to \infty} y_j\) exists, then \(\lim_{n \to \infty} x_n\) also exists and is equal to \(x\).

**Exercise 9.20.** Suppose that \(f : [0,\infty) \to [0,\infty)\) is a \(C^2\) – function such that \(f(0) = 0\), \(f' > 0\) and \(f'' \leq 0\) and \((X,\rho)\) is a metric space. Show that \(d(x,y) = f(\rho(x,y))\) is a metric on \(X\). In particular show that

\[
 d(x,y) := \frac{\rho(x,y)}{1 + \rho(x,y)}
\]

is a metric on \(X\). (Hint: use calculus to verify that \(d(a+b) \leq d(a) + d(b)\) for all \(a,b \in [0,\infty)\).)
Exercise 9.21. Let \( \{(X_n, d_n)\}_{n=1}^{\infty} \) be a sequence of metric spaces, \( X := \prod_{n=1}^{\infty} X_n \), and for \( x = (x(n))_{n=1}^{\infty} \) and \( y = (y(n))_{n=1}^{\infty} \) in \( X \) let

\[
d(x, y) = \sum_{n=1}^{\infty} 2^{-n} \frac{d_n(x(n), y(n))}{1 + d_n(x(n), y(n))}.
\]

Show:

1. \((X, d)\) is a metric space,
2. a sequence \( \{x_k\}_{k=1}^{\infty} \subset X \) converges to \( x \in X \) iff \( x_k(n) \to x(n) \) in \( X_n \) as \( k \to \infty \) for each \( n \in \mathbb{N} \) and
3. \( X \) is complete if \( X_n \) is complete for all \( n \).

9.5 Sequential Compactness

Suppose that \((X, d)\) and \((Y, \rho)\) are metric spaces.

Definition 9.29. As subset \( K \subset X \) is (sequentially) compact if every sequence \( \{z_n\}_{n=1}^{\infty} \subset K \) has a convergent subsequence, \( \{w_k := z_{n_k}\}_{k=1}^{\infty} \) such that \( \lim_{k \to \infty} w_k \in K \).

Example 9.30. Suppose that \( F \subset X \) is an unbounded set, i.e. for all \( n \in \mathbb{N} \) there exists \( z_n \in F \) such that \( d(z_n, 0) \geq n \). The sequence \( \{z_n\}_{n=1}^{\infty} \) and all of its subsequences are unbounded and therefore not Cauchy in \( X \) and hence not convergent in \( X \). This shows that sequentially compact sets must be bounded.

Example 9.31. Suppose that \( F \subset X \) is not closed. Then there exists \( \{z_n\}_{n=1}^{\infty} \subset F \) such that \( z := \lim_{n \to \infty} z_n \notin F \). Moreover, although every subsequence of \( \{z_n\}_{n=1}^{\infty} \) is convergent, they all still converge to \( z \notin F \). This shows that a sequentially compact set must be closed.

Lemma 9.32 (Bolzano–Weierstrass property for \( \mathbb{C}^D \)). Let \( D \in \mathbb{N} \). Every bounded sequence, \( \{z(n)\}_{n=1}^{\infty} \subset \mathbb{C}^D \), has a convergent subsequence.

Proof. By assumption there exists \( M < \infty \) such that \( \|z(n)\| = d(z(n), 0) \leq M \) for all \( n \in \mathbb{N} \). Writing \( z(n) = (z_1(n), \ldots, z_D(n)) \in \mathbb{C}^D \). Since \( |z_i(n)| \leq \|z(n)\| \) it follows that \( \{z_i(n)\}_{n=1}^{\infty} \) is a bounded sequence in \( \mathbb{C} \). Hence by the Bolzano–Weierstrass property for \( \mathbb{C} \) we replace \( z(n) \) by a subsequence \( z(n_k) \) such that \( \lim_{k \to \infty} z(n_k) = z_1 \) exists. We may now replace the original \( z \) by this new subsequence and then find a further subsequence \( z(n_k) \) such that \( \lim_{k \to \infty} z_i(n_k) = z_i \) exists for \( i = 1, 2 \). We may continue this way inductively to find a subsequence such that \( \lim_{k \to \infty} z_i(n_k) = z_i \) exists for all \( 1 \leq i \leq D \). It then follows that \( \lim_{k \to \infty} \|z - z(n_k)\| = 0 \) as desires where \( z := (z_1, \ldots, z_D) \).

Exercise 9.22 (Extreme value theorem). Let \( K \) be sequentially compact subset of \( X \) and \( f : K \to \mathbb{R} \) be a continuous function. Show \( -\infty <
Exercise 9.23 (Uniform Continuity). Let $K$ be sequentially compact subset of $X$ and $f : K \to \mathbb{C}$ be a continuous function. Show that $f$ is uniformly continuous, i.e. if $\varepsilon > 0$ there exists $\delta > 0$ such that $|f(z) - f(w)| < \varepsilon$ if $w, z \in K$ with $|w - z| < \delta$. Hint: first argue that there exists $(z_n)_{n=1}^\infty \subset K$ such that $f(z_n) \uparrow \sup_{x \in K} f(x)$ as $n \to \infty$.

Exercise 9.24. If $(X, d)$ is a metric and $K \subset X$ is sequentially compact. Show subset, $C \subset K$, which is closed is sequentially compact as well.

Exercise 9.25. If $K \subset \mathbb{R}$ is sequentially compact then $\sup (K) \in K$, i.e. $\sup(K) = \max (K)$.

Exercise 9.26. Let $(X, d)$ and $(Y, \rho)$ be metric spaces, $K \subset X$ be a sequentially compact set, and $f : K \to Y$ be a continuous function. Show $f(K)$ is sequentially compact in $Y$. In particular, for $C \subset K$ closed, we have $f(C)$ is closed and in fact sequentially compact in $Y$.

Exercise 9.27. Let $f : [a, b] \to [c, d]$ be a strictly increasing continuous function such that $f(a) = c$ and $f(b) = d$ and $g := f^{-1} : [c, d] \to [a, b]$ as in Exercise 6.20. Give one or better yet two alternative proofs that $g$ is continuous based on sequentially compactness arguments.

**Definition 9.35.** Let $Z$ be a vector space. We say that two norms, $|\cdot|$ and $\|\|$, on $Z$ are equivalent if there exists constants $\alpha, \beta \in (0, \infty)$ such that

$$\|f\| \leq \alpha |f| \text{ and } |f| \leq \beta \|f\| \text{ for all } f \in Z.$$  

**Theorem 9.36.** Let $Z$ be a finite dimensional vector space. Then any two norms $|\cdot|$ and $\|\|$ on $Z$ are equivalent. (This is typically not true for norms on infinite dimensional spaces.)

**Proof.** Let $\{f_i\}_{i=1}^n$ be a basis for $Z$ and define a new norm on $Z$ by

$$\left\| \sum_{i=1}^n a_i f_i \right\|_2 := \left( \sum_{i=1}^n |a_i|^2 \right)^{1/2} \text{ for } a_i \in \mathbb{F}.$$  

By the triangle inequality for the norm $|\cdot|$, we find

$$\left( \sum_{i=1}^n |a_i| \right)^2 \leq \sum_{i=1}^n |a_i|^2 \leq M \left\| \sum_{i=1}^n a_i f_i \right\|_2$$  

where $M = \sqrt{\sum_{i=1}^n |f_i|^2}$. Thus we have $|f| \leq M \|f\|_2$ for all $f \in Z$ and this inequality shows that $|\cdot|$ is continuous relative to $\|\|_2$. Since the normed space $(Z, \|\|_2)$ is homeomorphic and isomorphic to $\mathbb{F}^n$ with the standard euclidean norm, the closed bounded set, $S := \{ f \in Z : \|f\|_2 = 1 \} \subset Z$, is a sequentially compact subset of $Z$ relative to $\|\|_2$. Therefore by Exercise 9.22 there exists $f_0 \in S$ such that

$$m = \inf \{ |f| : f \in S \} = |f_0| > 0.$$  

Hence given $0 \neq f \in Z$, then $\frac{f}{\|f\|_2} \in S$ so that

$$m \leq \left\| \frac{f}{\|f\|_2} \right\|_2 = \frac{|f|}{\|f\|_2} \frac{1}{m} \|f\|.$$  

or equivalently

$$\|f\|_2 \leq \frac{1}{m} |f|.$$  

This shows that $|\cdot|$ and $\|\|_2$ are equivalent norms. Similarly one shows that $|\cdot|$ and $\|\|_2$ are equivalent and hence so are $|\cdot|$ and $\|\|_2$.

**Corollary 9.37.** If $(Z, \|\|)$ is a finite dimensional normed space, then $A \subset Z$ is sequentially compact iff $A$ is closed and bounded relative to the given norm, $\|\|$.

**Corollary 9.38.** Every finite dimensional normed vector space $(Z, \|\|)$ is complete. In particular any finite dimensional subspace of a normed vector space is automatically closed.

**Proof.** If $\{f_n\}_{n=1}^\infty \subset Z$ is a Cauchy sequence, then $\{f_n\}_{n=1}^\infty$ is bounded and hence has a convergent subsequence, $g_k = f_{n_k}$, by Corollary 9.37. It is now routine to show $\lim_{n \to \infty} f_n = f := \lim_{k \to \infty} g_k$.

9.6 Open Cover Compactness

**Definition 9.39.** Let $A$ be a subset of a metric space, $(X, d)$. An open cover of $A$ is a collection of $\mathcal{U}$, of open subsets of $X$ such that $A \subset \bigcup_{V \in \mathcal{U}} V$.

**Definition 9.40.** The subset $A \subset X$ is (open cover) **compact** if every open cover (Definition 9.39) of $A$ has a finite sub-cover, i.e. if $\mathcal{U}$ is an open cover of $A$ there exists finite subcollection, $\mathcal{U}_0 \subset \mathcal{U}$ such that $\mathcal{U}_0$ is still a cover of $A$. (We will write $A \subset \subset X$ to denote that $A \subset X$ and $A$ is compact.) A subset $A \subset X$ is **precompact** if $A$ is compact.

1 We will simply say $A$ is compact in this case.
Remark 9.41. If \((X, d)\) is a metric space, we will see below in Theorem 10.7 that \(X\) is sequentially compact iff it is open cover compact. We will not use this fact in this section however. In the future though we will just refer to compact sets with out the any extra adjectives.

Example 9.42. Suppose that \(A\) is an unbounded subset of \(X\). Pick \(a_1 \in A\) and then choose \(\{a_n\}_{n=2}^\infty\) inductively so that \(d(a_1, \ldots, a_n, a_{n+1}) \geq 1\) for all \(n\). This sequence then has the property that \(d(a_k, a_l) \geq 1\) for all \(k \neq l\) and from this it follows that \(F = \{a_1, a_2, \ldots\}\) is a closed set. We then define an open cover of \(A\) by taking,

\[
\mathcal{U} = \{F^c, B_{a_1}(1/3), B_{a_2}(1/3), B_{a_3}(1/3), \ldots\}.
\]

This cover has no finite subcover. Therefore \(A\) can not be compact.

Lemma 9.43. Suppose that \(K \subset X\) is a compact set, then \(K\) is closed.

Proof. We will show that \(K\) is open. To this end suppose \(x \in K^c\). Then let \(\varepsilon := \frac{1}{2}d(x, k) > 0\) for all \(k \in K\). Then it follows that \(B_x(\varepsilon) \cap B_k(\varepsilon) = \emptyset\) for all \(k \in K\). As \(\{B_k(\varepsilon_k)\}_{k \in K}\) is an open cover \(K\), there exists \(A \subset_f K\) such that \(K \subset \bigcup_{k \in A} B_k(\varepsilon_k)\). If we now let \(\delta := \min_{k \in A} \varepsilon_k > 0\), then

\[
B_x(\delta) \cap B_k(\varepsilon_k) \subset B_x(\varepsilon_k) \cap B_k(\varepsilon_k) = \emptyset
\]

for all \(k \in A\) and therefore

\[
B_x(\delta) \cap K \subset B_x(\delta) \cap \bigcup_{k \in K} B_k(\varepsilon_k) = \emptyset.
\]

Theorem 9.44. The compact subsets of \(\mathbb{R}^n\) are the closed and bounded sets.

Proof. Let us first suppose that \(K = [-M, M]^n\) for some positive integer \(M\) and for sake of contradiction let us suppose that \(K\) is an open cover of \(K\) with no finite subcover. For any \(k \in \mathbb{N}\) let \(A_k := [-M, M) \cap \{\ell 2^{-k} : \ell \in \mathbb{Z}\}\) and for \(x \in A_k^n\) let

\[
C_k^x := [x_1, x_1 + 2^{-k}] \times \cdots \times [x_n, x_n + 2^{-k}].
\]

We now let \(C^k = \{C_k^x : x \in A_k^n\}\). The cubes have the following properties;

1. \(K = \bigcup_{C^k \in C} C\) for all \(k\) and if \(C \in C^k\), then \(C = \bigcup_{F \in C^{k+1}} F\).
2. \(\text{diam}(C) = \sqrt{n} \cdot 2^{-k}\) for all \(C \in C^k\).

By item 1. there must be a \(C_1 \subset C^1\) such that \(\mathcal{U}\) has no finite subcover of \(C_1\). Similarly there exists \(C_2 \subset C_1\) such that \(\mathcal{U}\) has no finite subcover of \(C_2\). Continuing this way inductively we may construct \(C_k \subset C^k\) such that \(C_1 \supset C_2 \supset C_3 \supset \ldots\) and \(\mathcal{U}\) has no finite subcover of \(C_k\) for any \(k\). Choose a point \(z_k \in C_k\) for all \(k \in \mathbb{N}\). Since \(\text{diam}(C_k) = \sqrt{n} \cdot 2^{-k} \to 0\) one learns that \(\{z_k\}_{k=1}^\infty\) is a Cauchy sequence and hence convergent. Let \(z = \lim_{k \to \infty} z_k\) which is in \(C_k\) for all \(k\) as each of these cubes are closed sets. Since \(\mathcal{U}\) is an open cover of \(K\), there exists \(V \in \mathcal{U}\) such that \(z \in V\). As \(V\) is open, there exists \(\varepsilon > 0\) such that \(B_{\varepsilon}(z) \subset V\). On the other hand because \(\text{diam}(C_k) \to 0\) as \(k \to \infty\), it follows that \(C_k \subset B_{\varepsilon}(z) \subset V\) for all \(k\) sufficiently large which violates the condition that \(\mathcal{U}\) has no finite subcover of \(C_k\) for any \(k\).

Now suppose that \(K\) is a general closed and bounded subset of \(\mathbb{R}^n\) and \(\mathcal{U}\) is an open cover of \(K\). Since \(K\) is bounded, \(K \subset [-M, M]^n\) for some \(M \in \mathbb{N}\). Since \(\mathcal{U} := \mathcal{U} \cup \{K^c\}\) is an open cover of \([-M, M]^n\) it has a finite subcover, say \(\{U_1, \ldots, U_l\} \cup \{K^c\}\). As \(K \cap K^c = \emptyset\) we must have that \(U_i \cap K = \emptyset\) and \(U_i \subset \mathcal{U}\) is a finite subcover of \(K\).

Proposition 9.45. Suppose that \(K \subset X\) is a compact set and \(F \subset K\) is a closed sub- \(\mathcal{U}\) is an open cover of \(K\). Since \(K\) is bounded, \(K \subset [-M, M]^n\) for some \(M \in \mathbb{N}\). Since \(\mathcal{U} := \mathcal{U} \cup \{K^c\}\) is an open cover of \([-M, M]^n\) it has a finite subcover, say \(\{U_1, \ldots, U_l\} \cup \{K^c\}\). As \(K \cap K^c = \emptyset\) we must have that \(U_i \cap K = \emptyset\) and \(U_i \subset \mathcal{U}\) is a finite subcover of \(K\).

Exercise 9.28. Suppose \(f : X \to Y\) is continuous and \(K \subset X\) is compact, then \(f(K)\) is a compact subset of \(Y\). Give an example of continuous map, \(f : X \to Y\), and a compact subset \(K\) of \(Y\) such that \(f^{-1}(K)\) is not compact.

Exercise 9.29 (Extreme value theorem). Let \((X, d)\) be a compact metric space and \(f : X \to \mathbb{R}\) be a continuous function. Show \(-\infty \leq \inf f \leq \sup f < \infty\) and there exists \(a, b \in X\) such that \(f(a) = \inf f\) and \(f(b) = \sup f\). Hint: use Exercise 10.2 and Theorem 9.44.

Exercise 9.30 (Uniform Continuity). Let \((X, d)\) be a compact metric space, \((Y, \rho)\) be a metric space and \(f : X \to Y\) be a continuous function. Show that \(f\) is uniformly continuous, i.e. if \(\varepsilon > 0\) there exists \(\delta > 0\) such that \(\rho(f(x), f(y)) < \varepsilon\) if \(x, y \in X\) with \(d(x, y) < \delta\).

Exercise 9.31. Suppose \(f : X \to Y\) is continuous and \(K \subset X\) is compact, then \(f(K)\) is a compact subset of \(Y\). Give an example of continuous map, \(f : X \to Y\), and a compact subset \(K\) of \(Y\) such that \(f^{-1}(K)\) is not compact.

Exercise 9.32 (Dini’s Theorem). Let \(X\) be a compact topological space and \(f_n : X \to [0, \infty)\) be a sequence of continuous functions such that \(f_n(x) \downarrow 0\) as \(n \to \infty\) for each \(x \in X\). Show that in fact \(f_n \downarrow 0\) uniformly in \(x\), i.e. \(\sup_{x \in X} f_n(x) \downarrow 0\) as \(n \to \infty\). Hint: Given \(\varepsilon > 0\), consider the open sets \(V_\varepsilon := \{x \in X : f_n(x) < \varepsilon\}\).
Definition 9.46. A collection $\mathcal{F}$ of closed subsets of a topological space $(X, \tau)$ has the finite intersection property if $\cap F_0 \neq \emptyset$ for all $F_0 \subset \subset \mathcal{F}$.

The notion of compactness may be expressed in terms of closed sets as follows.

Proposition 9.47. A topological space $X$ is compact iff every family of closed sets $\mathcal{F} \subset 2^X$ having the finite intersection property satisfies $\cap \mathcal{F} \neq \emptyset$.

Proof. The basic point here is that complementation interchanges open and closed sets and open covers go over to collections of closed sets with empty intersection. Here are the details.

($\Rightarrow$) Suppose that $X$ is compact and $\mathcal{F} \subset 2^X$ is a collection of closed sets such that $\cap \mathcal{F} = \emptyset$. Let

$$\mathcal{U} = \mathcal{F}^c := \{C^c : C \in \mathcal{F}\} \subset \tau,$$

then $\mathcal{U}$ is a cover of $X$ and hence has a finite subcover, $\mathcal{U}_0$. Let $\mathcal{F}_0 = \mathcal{U}_0^c \subset \subset \mathcal{F}$, then $\cap \mathcal{F}_0 = \emptyset$ so that $\mathcal{F}$ does not have the finite intersection property.

($\Leftarrow$) If $X$ is not compact, there exists an open cover $\mathcal{U}$ of $X$ with no finite subcover. Let

$$\mathcal{F} = \mathcal{U}^c := \{U^c : U \in \mathcal{U}\},$$

then $\mathcal{F}$ is a collection of closed sets with the finite intersection property while $\cap \mathcal{F} = \emptyset$. $\blacksquare$
General Metric Space Compactness Reuslts

This chapter may be skipped on first reading.

10.1 Compactness

**Definition 10.1.** The subset \( A \) of a topological space \((X, \tau)\) is said to be **compact** if every open cover (Definition 9.39) of \( A \) has finite a sub-cover, i.e. if \( \mathcal{U} \) is an open cover of \( A \) there exists \( \mathcal{U}_0 \subset \mathcal{U} \) such that \( \mathcal{U}_0 \) is a cover of \( A \). (We will write \( A \subseteq X \) to denote that \( A \subset X \) and \( A \) is compact.) A subset \( A \subset X \) is **precompact** if \( A \) is compact.

**Proposition 10.2.** Suppose that \( K \subset X \) is a compact set and \( F \subset K \) is a closed subset. Then \( F \) is compact. If \( \{K_i\}_{i=1}^n \) is a finite collection of compact subsets of \( X \) then \( K = \cup_{i=1}^n K_i \) is also a compact subset of \( X \).

**Proof.** Let \( U \subset \tau \) be an open cover of \( F \), then \( U \cup \{F^c\} \) is an open cover of \( K \). The cover \( U \cup \{F^c\} \) of \( K \) has a finite subcover which we denote by \( U_0 \cup \{F^c\} \) where \( U_0 \subset \cup U \). Since \( F \cap F^c = \emptyset \), it follows that \( U_0 \) is the desired subcover of \( F \). For the second assertion suppose \( U \subset \tau \) is an open cover of \( K \). Then \( U \) covers each compact set \( K_i \) and therefore there exists a finite subset \( U_i \subset \cup \) for each \( i \) such that \( K_i \subset U_i \). Then \( U_0 := \cup_{i=1}^n U_i \) is a finite cover of \( K \).

**Exercise 10.1 (Suggested by Michael Gurvich).** Show by example that the intersection of two compact sets need not be compact. (This pathology disappears if one assumes the topology is Hausdorff, see Definition ?? below.)

**Exercise 10.2.** Suppose \( f : X \to Y \) is continuous and \( K \subset X \) is compact, then \( f(K) \) is a compact subset of \( Y \). Give an example of continuous map, \( f : X \to Y \), and a compact subset \( K \) of \( Y \) such that \( f^{-1}(K) \) is not compact.

**Exercise 10.3 (Dini’s Theorem).** Let \( X \) be a compact topological space and \( f_n : X \to [0, \infty) \) be a sequence of continuous functions such that \( f_n(x) \downarrow 0 \) as \( n \to \infty \) for each \( x \in X \). Show that in fact \( f_n \downarrow 0 \) uniformly in \( x \), i.e. \( \sup_{x \in X} f_n(x) \downarrow 0 \) as \( n \to \infty \). **Hint:** Given \( \varepsilon > 0 \), consider the open sets \( V_{n, \varepsilon} := \{x \in X: f_n(x) < \varepsilon\} \).

**Definition 10.3.** A collection \( \mathcal{F} \) of closed subsets of a topological space \((X, \tau)\) has the **finite intersection property** if \( \cap \mathcal{F} \neq \emptyset \) for all \( \mathcal{F}_0 \subset \mathcal{F} \).

The notion of compactness may be expressed in terms of closed sets as follows.

**Proposition 10.4.** A topological space \( X \) is compact iff every family of closed sets \( \mathcal{F} \subset 2^X \) having the finite intersection property satisfies \( \cap \mathcal{F} \neq \emptyset \).

**Proof.** The basic point here is that complementation interchanges open and closed sets and open covers go over to collections of closed sets with empty intersection. Here are the details. 

\( (\Rightarrow) \) Suppose that \( X \) is compact and \( \mathcal{F} \subset 2^X \) is a collection of closed sets such that \( \cap \mathcal{F} = \emptyset \). Let 

\[ \mathcal{U} = \mathcal{F}^c := \{C^c : C \in \mathcal{F}\} \subset \tau, \]

then \( \mathcal{U} \) is a cover of \( X \) and hence has a finite subcover, \( \mathcal{U}_0 \). Let \( \mathcal{F}_0 = \mathcal{U}_0^c \subset \mathcal{F} \), then \( \cap \mathcal{F}_0 = \emptyset \) so that \( \mathcal{F} \) does not have the finite intersection property.

\( (\Leftarrow) \) If \( X \) is not compact, there exists an open cover \( \mathcal{U} \) of \( X \) with no finite subcover. Let 

\[ \mathcal{F} = \mathcal{U}^c := \{U^c : U \in \mathcal{U}\}, \]

then \( \mathcal{F} \) is a collection of closed sets with the finite intersection property while \( \cap \mathcal{F} = \emptyset \).

**Exercise 10.4.** Let \((X, \tau)\) be a topological space. Show that \( A \subset X \) is compact iff \((A, \tau_A)\) is a compact topological space.

Let \((X, d)\) be a metric space and for \( x \in X \) and \( \varepsilon > 0 \) let 

\[ B_x^d(\varepsilon) := B_x(\varepsilon) \setminus \{x\} \]

be the ball centered at \( x \) of radius \( \varepsilon > 0 \) with \( x \) deleted. Recall from Definition 9.15 that a point \( x \in X \) is an accumulation point of a subset \( E \subset X \) if \( \emptyset \neq E \cap V \setminus \{x\} \) for all open neighborhoods, \( V \), of \( x \). The proof of the following elementary lemma is left to the reader.

**Lemma 10.5.** Let \( E \subset X \) be a subset of a metric space \((X, d)\). Then the following are equivalent:

1. \( x \in X \) is an accumulation point of \( E \).
2. $B_x(\varepsilon) \cap E \neq \emptyset$ for all $\varepsilon > 0$.
3. $B_x(\varepsilon) \cap E$ is an infinite set for all $\varepsilon > 0$.
4. There exists $\{x_n\}_{n=1}^\infty \subset E \setminus \{x\}$ with $\lim_{n \to \infty} x_n = x$.

**Definition 10.6.** A metric space $(X, d)$ is $\varepsilon$-**bounded** ($\varepsilon > 0$) if there exists a finite cover of $X$ by balls of radius $\varepsilon$ and it is **totally bounded** if it is $\varepsilon$-bounded for all $\varepsilon > 0$.

**Theorem 10.7.** Let $(X, d)$ be a metric space. The following are equivalent.

(a) $X$ is compact.

(b) Every infinite subset of $X$ has an accumulation point.

(c) Every sequence $\{x_n\}_{n=1}^\infty \subset X$ has a convergent subsequence.

(d) $X$ is totally bounded and complete.

**Proof.** The proof will consist of showing that $a \Rightarrow b \Rightarrow c \Rightarrow d \Rightarrow a$.

(a $\Rightarrow$ b) We will show that not $b \Rightarrow$ not $a$. Suppose there exists an infinite subset $E \subset X$ which has no accumulation points. Then for all $x \in X$ there exists $\delta_x > 0$ such that $V_x := B_x(\delta_x)$ satisfies $(V_x \setminus \{x\}) \cap E = \emptyset$. Clearly $\mathcal{V} = \{V_x\}_{x \in X}$ is a cover of $X$, yet $\mathcal{V}$ has no finite sub cover. Indeed, for each $x \in X$, $V_x \cap E \subset \{x\}$ and hence if $A \subset X, \cup_{x \in A} V_x$ can only contain a finite number of points from $E$ (namely $A \cap E$). Thus for any $A \subset X, E \not\subset \cup_{x \in A} V_x$ and in particular $X \not\subset \cup_{x \in A} V_x$. (See Figure 10.1)

(b $\Rightarrow$ c) Let $\{x_n\}_{n=1}^\infty \subset X$ be a sequence and $E := \{x_n : n \in \mathbb{N}\}$. If $\#(E) < \infty$, then $\{x_n\}_{n=1}^\infty$ has a subsequence $\{x_{n_k}\}_{k=1}^\infty$ which is constant and hence convergent. On the other hand if $\#(E) = \infty$ then by assumption $E$ has an accumulation point and hence by Lemma [10.5] $\{x_n\}_{n=1}^\infty$ has a convergent subsequence.

(c $\Rightarrow$ d) Suppose $\{x_n\}_{n=1}^\infty \subset X$ is a Cauchy sequence. By assumption there exists a subsequence $\{x_{n_k}\}_{k=1}^\infty$ which is convergent to some point $x \in X$. Since $\{x_n\}_{n=1}^\infty$ is Cauchy it follows that $x_n \to x$ as $n \to \infty$ showing $X$ is complete. We now show that $X$ is totally bounded. Let $\varepsilon > 0$ be given and choose an arbitrary point $x_1 \in X$. If possible choose $x_2 \in X$ such that $d(x_2, x_1) \geq \varepsilon$, then if possible choose $x_3 \in X$ such that $d(x_1, x_2) \geq \varepsilon$ and continue inductively choosing points $\{x_j\}_{j=1}^\infty \subset X$ such that $d(x_j, x_{j+1}) \geq \varepsilon$. (See Figure 10.2) This process must terminate, for otherwise we would produce a sequence $\{x_n\}_{n=1}^\infty \subset X$ which can have no convergent subsequences. Indeed, the $x_n$ have been chosen so that $d(x_n, x_m) \geq \varepsilon > 0$ for every $m \neq n$ and hence no subsequence of $\{x_n\}_{n=1}^\infty$ can be Cauchy.

(d $\Rightarrow$ a) For sake of contradiction, assume there exists an open cover $\mathcal{V} = \{V_a\}_{a \in A}$ of $X$ with no finite subcover. Since $X$ is totally bounded for each $n \in \mathbb{N}$ there exists $A_n \subset X$ such that

$$X = \bigcup_{x \in A_n} B_x(1/n) \subset \bigcup_{x \in A_n} C_x(1/n).$$

Choose $x_1 \in A_1$ such that no finite subset of $\mathcal{V}$ covers $K_1 := C_{x_1}(1)$. Since $K_1 = \cup_{x \in A_1} K_1 \cap C_x(1/2)$, there exists $x_2 \in A_2$ such that $K_2 := K_1 \cap C_{x_2}(1/2)$ can not be covered by a finite subset of $\mathcal{V}$, see Figure 10.3. Continuing this way inductively, we construct sets $K_n = K_{n-1} \cap C_{x_n}(1/n)$ with $x_n \in A_n$ such that no $K_n$ can be covered by a finite subset of $\mathcal{V}$. Now choose $y_n \in K_n$ for each $n$. Since $\{K_n\}_{n=1}^\infty$ is a decreasing sequence of closed sets such that $\text{diam}(K_n) \leq 2/n$, it follows that $\{y_n\}$ is a Cauchy and hence convergent with

$$y = \lim_{n \to \infty} y_n \in \cap_{m=1}^\infty K_m.$$ 

Since $\mathcal{V}$ is a cover of $X$, there exists $V \in \mathcal{V}$ such that $y \in V$. Since $K_n \downarrow \{y\}$ and $\text{diam}(K_n) \to 0$, it now follows that $K_n \subset V$ for some $n$ large. But this violates the assertion that $K_n$ can not be covered by a finite subset of $\mathcal{V}$.

**Corollary 10.8.** Any compact metric space $(X, d)$ is second countable and hence separable by Exercise ??.. (See Example ?? below for an example of a compact topological space which is not separable.)
Suppose that \( \delta > M \).

\[ \text{for} \quad i = 1, 2, \ldots, n \]

which shows that \( d(x, y) \leq \sqrt{n} \delta \). Hence if choose \( \delta < \varepsilon / \sqrt{n} \) we have shows that \( d(x, y) < \varepsilon \), i.e. Eq. (10.1) holds. \( \square \)

**Example 10.10.** Let \( X = \ell^p (\mathbb{N}) \) with \( p \in [1, \infty) \) and \( \mu \in \ell^p (\mathbb{N}) \) such that \( \mu (k) \geq 0 \) for all \( k \in \mathbb{N} \). The set

\[ K := \{ x \in X : |x(k)| \leq \mu (k) \text{ for all } k \in \mathbb{N} \} \]

is compact. To prove this, let \( \{ x_n \}_{n=1}^\infty \subset K \) be a sequence. By compactness of closed bounded sets in \( \mathbb{C} \), for each \( k \in \mathbb{N} \) there is a subsequence of \( \{ x_n (k) \}_{n=1}^\infty \subset C \) which is convergent. By Cantor’s diagonalization trick, we may choose a subsequence \( \{ y_n \}_{n=1}^\infty \) of \( \{ x_n \}_{n=1}^\infty \) such that \( y(k) := \lim_{n \to \infty} y_n (k) \) exists for all \( k \in \mathbb{N} \). Since \( |y_n (k)| \leq \mu (k) \) for all \( n \) it follows that \( |y(k)| \leq \mu (k) \), i.e. \( y \in K \).

Finally

\[ \lim_{n \to \infty} \| y - y_n \|_p = \lim_{n \to \infty} \sum_{k=1}^{\infty} |y(k) - y_n (k)|^p = \sum_{k=1}^{\infty} \lim_{n \to \infty} |y(k) - y_n (k)|^p = 0 \]

wherein we have used the Dominated convergence theorem. (Note

\[ |y(k) - y_n (k)|^p \leq 2^p \mu^p (k) \]

and \( \mu^p \) is summable.) Therefore \( y_n \to y \) and we are done.

Alternatively, we can prove \( K \) is compact by showing that \( K \) is closed and totally bounded. It is simple to show \( K \) is closed, for if \( \{ x_n \}_{n=1}^\infty \subset K \) is a convergent sequence in \( X \), \( x := \lim_{n \to \infty} x_n \), then

\[ |x(k)| \leq \lim_{n \to \infty} |x_n (k)| \leq \mu (k) \forall k \in \mathbb{N}. \]

This shows that \( x \in K \) and hence \( K \) is closed. To see that \( K \) is totally bounded, let \( \varepsilon > 0 \) and choose \( N \) such that \( (\sum_{n=N+1}^{\infty} |\mu (k)|^p)^{1/p} < \varepsilon \). Since

\[ d^2 (x, y) = \sum_{i=1}^{n} (y_i - x_i)^2 \leq n \delta^2 \]

1 The argument is as follows. Let \( \{ n_j \}_{j=1}^\infty \) be a subsequence of \( \mathbb{N} = \{ n_j \}_{j=1}^\infty \) such that \( \lim_{j \to \infty} x_{n_j} (1) \) exists. Now choose a subsequence \( \{ n_j \}_{j=1}^\infty \) of \( \{ n_j \}_{j=1}^\infty \) such that \( \lim_{j \to \infty} x_{n_j} (2) \) exists and similarly \( \{ n_j \}_{j=1}^\infty \) of \( \{ n_j \}_{j=1}^\infty \) such that \( \lim_{j \to \infty} x_{n_j} (3) \) exists. Continue on this way inductively to get

\[ \{ n \}_{n=1}^\infty \supset \{ n_j \}_{j=1}^\infty \supset \{ n_j \}_{j=1}^\infty \supset \{ n_j \}_{j=1}^\infty \supset \ldots \]

such that \( \lim_{j \to \infty} x_{n_j} (k) \) exists for all \( k \in \mathbb{N} \). Let \( m_j := n_j \) so that eventually \( \{ m_j \}_{j=1}^\infty \) is a subsequence of \( \{ n_j \}_{j=1}^\infty \) for all \( k \). Therefore, we may take \( y_j := x_{m_j} \).
uniformly continuous, i.e. if \( \varepsilon > 0 \),

\[
\prod_{k=1}^{N} C_{\mu(k)}(0) \subset C_{\mathbb{N}}
\]
is closed and bounded, it is compact. Therefore there exists a finite subset \( A \subset \prod_{k=1}^{N} C_{\mu(k)}(0) \) such that

\[
\prod_{k=1}^{N} C_{\mu(k)}(0) \subset \bigcup_{z \in A} B_{z}^{N}(\varepsilon)
\]

where \( B_{z}^{N}(\varepsilon) \) is the open ball centered at \( z \in \mathbb{C}^{N} \) relative to the \( \ell^{p}(\{1, 2, 3, \ldots, N\}) \) - norm. For each \( z \in A \), let \( \hat{z} \in X \) be defined by \( \hat{z}(k) = z(k) \) if \( k \leq N \) and \( \hat{z}(k) = 0 \) for \( k \geq N + 1 \). I now claim that

\[
K \subset \bigcup_{z \in A} B_{\hat{z}}(2\varepsilon)
\]

which, when verified, shows \( K \) is totally bounded. To verify Eq. \((10.2)\), let \( x \in K \) and write \( x = u + v \) where \( u(k) = x(k) \) for \( k \leq N \) and \( u(k) = 0 \) for \( k < N \). Then by construction \( u \in B_{\hat{z}}(\varepsilon) \) for some \( \hat{z} \in A \) and

\[
\|u\|_{p} \leq \left( \sum_{k=N+1}^{\infty} |\mu(k)|^{p} \right)^{1/p} < \varepsilon.
\]

So we have

\[
\|x - \hat{z}\|_{p} = \|u + v - \hat{z}\|_{p} \leq \|u - \hat{z}\|_{p} + \|v\|_{p} < 2\varepsilon.
\]

**Exercise 10.5 (Extreme value theorem).** Let \( (X, \tau) \) be a compact topological space and \( f : X \to \mathbb{R} \) be a continuous function. Show \(-\infty < \inf f \leq \sup f < \infty\) and there exists \( a, b \in X \) such that \( f(a) = \inf f \) and \( f(b) = \sup f \).

**Hint:** use Exercise 10.2 and Corollary 10.9

**Exercise 10.6 (Uniform Continuity).** Let \( (X, d) \) be a compact metric space, \( (Y, \rho) \) be a metric space and \( f : X \to Y \) be a continuous function. Show that \( f \) is uniformly continuous, i.e. if \( \varepsilon > 0 \) there exists \( \delta > 0 \) such that \( \rho(f(y), f(x)) < \varepsilon \) if \( x, y \in X \) with \( d(x, y) < \delta \).

**Hint:** you could follow the argument in the proof of Theorem ???

**Theorem 10.11.** Suppose that \( (X, \|\cdot\|) \) is a normed vector in which the unit ball, \( V := B_{0}(1) \), is precompact. Then \( \dim X < \infty \).

**An alternate proof is given in Proposition 10.13.** Since \( V \) is compact, we may choose \( A \subset X \) such that

\[
V \subset \bigcup_{x \in A} \left( x + \frac{1}{2} V \right)
\]

where, for any \( \delta > 0 \),

\[
\delta V := \{ \delta x : x \in V \} = B_{0}(\delta).
\]

Let \( Y := \text{span}(A) \), then Eq. \((10.3)\) implies,

\[
V \subset \hat{V} \subset Y + \frac{1}{2} V.
\]

Multiplying this equation by \( \frac{1}{2} \) then shows

\[
\frac{1}{2} V \subset \frac{1}{2} Y + \frac{1}{4} V = Y + \frac{1}{4} V
\]

and hence

\[
V \subset Y + \frac{1}{2} V \subset Y + \frac{1}{4} V = Y + \frac{1}{4} V.
\]

Continuing this way inductively then shows that

\[
V \subset Y + \frac{1}{2^{n}} V\text{ for all } n \in \mathbb{N}.
\]

Indeed, if Eq. \((10.4)\) holds, then

\[
V \subset Y + \frac{1}{2^{n}} V \subset Y + \frac{1}{2} \left( Y + \frac{1}{2^{n}} V \right) = Y + \frac{1}{2^{n+1}} V.
\]

Hence if \( x \in V \), there exists \( y_{n} \in Y \) and \( z_{n} \in B_{0}(2^{-n}) \) such that \( y_{n} + z_{n} \to x \). Since \( \lim n \to \infty z_{n} = 0 \), it follows that \( x = \lim n \to \infty y_{n} \in \hat{Y} \). Since \( \dim Y \leq \#(A) < \infty \), Corollary 9.38 implies \( \hat{Y} = \hat{Y} \) and so we have shown that \( V \subset Y \).

For any \( x \in X \), \( \frac{1}{2^{n}} \|x\| \in V \subset Y \), we have \( x \in Y \) for all \( x \in X \), i.e. \( X = Y \).

**Lemma 10.12.** Let \( H \) be a normed linear space and \( H_{0} \) a closed proper subspace.

For any \( \varepsilon > 0 \), there exists \( x_{0} \in H \) such that \( \|x_{0}\| = 1 \) and \( \|x - x_{0}\| \geq 1 - \varepsilon \) whenever \( x \in H_{0} \).

**Proof.** Can assume \( \varepsilon < 1 \). Take any \( z_{0} \notin H_{0} \). Let \( d = \inf_{x \in H_{0}} \|x - z_{0}\| \). For any \( \delta > 0 \), there exists \( z \in H_{0} \), such that \( \|z - z_{0}\| \leq d + \delta \). Take \( \delta = \frac{d}{1 - \varepsilon} \). Let \( x_{0} = (z - z_{0})/\|z - z_{0}\| \), where \( z \) is determined for this \( \delta \). Then \( \|x_{0}\| = 1 \), and if \( x \in H_{0} \),

\[
\|x - x_{0}\| = \frac{\|z - z_{0}\| \|x - z + z_{0}\|}{\|z - z_{0}\|} \geq \frac{d}{\|z - z_{0}\|} \geq \frac{d}{d + \delta} = 1 - \varepsilon.
\]

Here is the proof again at a higher level. Choose \( h \in H_{0} \) such that \( d := \text{dist}(z_{0}, H_{0}) \equiv \|h - z_{0}\| \) and then take \( x_{0} := (h - z_{0})/\|h - z_{0}\| \) as above. Then
\[ \text{dist}(x_0, H_0) = \frac{1}{\|h - z_0\|} \text{dist}(h - z_0, H_0) \]

where we have used the easily verified fact that \( \text{dist}(ax + h, H) = |a| \text{dist}(x, H) \) for all \( a \in \mathbb{R} \) and \( h \in H \).

**Proposition 10.13.** A locally compact Banach space is finite dimensional.

**Proof.** We prove that an infinite dimensional Banach space is not locally compact. We construct a sequence \( x_1, x_2, \ldots, x_n, \ldots \) such that \( \|x_n\| = 1, \|x_i - x_j\| \geq 1/2, i \neq j \). Take \( x_1 \) to be any unit vector. Suppose vectors \( x_1, \ldots, x_n \) are constructed. Let \( H_0 \) be the linear span of \( x_1, \ldots, x_n \). By Corollary 9.38, \( H_0 \) is closed. By Lemma 10.12, there exists \( x_{n+1} \) such that \( \|x_i - x_{n+1}\| \geq 1/2, i = 1, \ldots, n \). Now the sequence just constructed has no Cauchy subsequence. Hence the closed unit ball is not compact. Similarly the closed ball of radius \( r > 0 \) is also not compact.

**Exercise 10.7.** Suppose \((Y, \|\cdot\|_Y)\) is a normed space and \((X, \|\cdot\|_X)\) is a finite dimensional normed space. Show every linear transformation \( T : X \to Y \) is necessarily bounded.

### 10.2 Exercises

#### 10.2.1 General Topological Space Problems

**Exercise 10.8.** Let \( V \) be an open subset of \( \mathbb{R} \). Show \( V \) may be written as a disjoint union of open intervals \( J_n = (a_n, b_n) \), where \( a_n, b_n \in \mathbb{R} \cup \{-\infty, \infty\} \) for \( n = 1, 2, \ldots, N \) with \( N = \infty \) possible.

**Exercise 10.9.** Let \((X, \tau)\) and \((Y, \tau')\) be a topological spaces, \( f : X \to Y \) be a function, \( U \) be an open cover of \( X \) and \( \{F_j\}_{j=1}^n \) be a finite cover of \( X \) by closed sets.

1. If \( A \subset X \) is any set and \( f : X \to Y \) is \((\tau, \tau')\) continuous then \( f|_A : A \to Y \) is \((\tau_A, \tau')\) continuous.
2. Show \( f : X \to Y \) is \((\tau, \tau')\) continuous iff \( f|_U : U \to Y \) is \((\tau_U, \tau')\) continuous for all \( U \in \mathcal{U} \).
3. Show \( f : X \to Y \) is \((\tau, \tau')\) continuous iff \( f|_{F_j} : F_j \to Y \) is \((\tau_{F_j}, \tau')\) continuous for all \( j = 1, 2, \ldots, n \).

**Exercise 10.10.** Suppose that \( X \) is a set, \( \{Y_\alpha, \tau_\alpha\} : \alpha \in A \) is a family of topological spaces and \( f_\alpha : X \to Y_\alpha \) is a given function for all \( \alpha \in A \). Assuming that \( S_\alpha \subset \tau_\alpha \) is a sub-base for the topology \( \tau_\alpha \) for each \( \alpha \in A \), show \( S := \bigcup_{\alpha \in A} f_\alpha^{-1}(S_\alpha) \) is a sub-base for the topology \( \tau := \tau(f_\alpha : \alpha \in A) \).

#### 10.2.2 Metric Spaces as Topological Spaces

**Definition 10.14.** Two metrics \( d \) and \( \rho \) on a set \( X \) are said to be equivalent if there exists a constant \( c \in (0, \infty) \) such that \( c^{-1} \rho \leq d \leq c \rho \).

**Exercise 10.11.** Suppose that \( d \) and \( \rho \) are two metrics on \( X \).

1. Show \( \tau_d = \tau_\rho \) if \( d \) and \( \rho \) are equivalent.
2. Show by example that it is possible for \( \tau_d = \tau_\rho \) even though \( d \) and \( \rho \) are inequivalent.

**Exercise 10.12.** Let \((X_i, d_i)\) for \( i = 1, \ldots, n \) be a finite collection of metric spaces and for \( 1 \leq p \leq \infty \) and \( x = (x_1, x_2, \ldots, x_n) \) and \( y = (y_1, \ldots, y_n) \) in \( X := \prod_{i=1}^n X_i \), let

\[
\rho_p(x, y) = \left\{ \begin{array}{ll}
\left( \sum_{i=1}^n |d_i(x_i, y_i)|^p \right)^{1/p} & \text{if } p \neq \infty \\
\max_i d_i(x_i, y_i) & \text{if } p = \infty
\end{array} \right.
\]

1. Show \((X, \rho_p)\) is a metric space for \( p \in [1, \infty] \). **Hint:** Minkowski’s inequality.
2. Show for any \( p, q \in [1, \infty] \), the metrics \( \rho_p \) and \( \rho_q \) are equivalent. **Hint:** This can be done with explicit estimates or you could use Theorem 9.36 below.

**Notation 10.15** Let \( X \) be a set and \( \mathcal{P} := \{p_n\}_{n=0}^\infty \) be a family of semi-metrics on \( X \), i.e. \( p_n : X \times X \to [0, \infty) \) are functions satisfying the assumptions of metric except for the assertion that \( p_n(x, y) = 0 \) implies \( x = y \). Further assume that \( p_n(x, y) \leq p_{n+1}(x, y) \) for all \( n \) and if \( p_n(x, y) = 0 \) for all \( n \in \mathbb{N} \) then \( x = y \). Given \( n \in \mathbb{N} \) and \( x \in X \)

\[
B_n(x, \varepsilon) := \{y \in X : p_n(x, y) < \varepsilon\}.
\]

We will write \( \tau(\mathcal{P}) \) form the smallest topology on \( X \) such that \( p_n(x, \cdot) : X \to [0, \infty) \) is continuous for all \( n \in \mathbb{N} \) and \( x \in X \), i.e. \( \tau(\mathcal{P}) := \tau(p_n(x, \cdot) : n \in \mathbb{N} \) and \( x \in X \).

**Exercise 10.13.** Using Notation 10.15 show that collection of balls,

\[
\mathcal{B} := \{B_n(x, \varepsilon) : n \in \mathbb{N}, x \in X \text{ and } \varepsilon > 0\},
\]

forms a base for the topology \( \tau(\mathcal{P}) \). **Hint:** Use Exercise 10.10 to show \( \mathcal{B} \) is a sub-base for the topology \( \tau(\mathcal{P}) \) and then use Exercise ?? to show \( \mathcal{B} \) is in fact a base for the topology \( \tau(\mathcal{P}) \).

**Exercise 10.14.** (A minor variant of Exercise 9.21). Let \( p_n \) be as in Notation 10.15 and

\[
d(x, y) := \sum_{n=0}^\infty 2^{-n} \frac{p_n(x, y)}{1 + p_n(x, y)}.
\]
Show $d$ is a metric on $X$ and $\tau_d = \tau(p)$. Conclude that a sequence $\{x_k\}_{k=1}^\infty \subset X$ converges to $x \in X$ iff
\[
\lim_{k \to \infty} p_n(x_k, x) = 0 \text{ for all } n \in \mathbb{N}.
\]

Exercise 10.15. Let $\{(X_n, d_n)\}_{n=1}^\infty$ be a sequence of metric spaces, $X := \prod_{n=1}^\infty X_n$, and for $x = (x(n))_{n=1}^\infty$ and $y = (y(n))_{n=1}^\infty$ in $X$ let
\[
d(x, y) = \sum_{n=1}^\infty 2^{-n} \frac{d_n(x(n), y(n))}{1 + d_n(x(n), y(n))}.
\]
(See Exercise 9.21.) Moreover, let $\pi_n : X \to X_n$ be the projection maps, show
\[
\tau_d = \otimes_{n=1}^\infty \tau_{d_n} = : \tau(\{\pi_n : n \in \mathbb{N}\}).
\]
That is show the $d$ – metric topology is the same as the product topology on $X$. Suggestions: 1) show $\pi_n$ is $\tau_d$ continuous for each $n$ and 2) show for each $x \in X$ that $d(x, \cdot)$ is $\otimes_{n=1}^\infty \tau_{d_n}$ – continuous. For the second assertion notice that $d(x, \cdot) = \sum_{n=1}^\infty f_n$ where $f_n = 2^{-n} \left( \frac{d_n(x(n), \cdot)}{1 + d_n(x(n), \cdot)} \right) \circ \pi_n$.

10.2.3 Compactness Problems

Exercise 10.16 (Tychonoff’s Theorem for Compact Metric Spaces). Let us continue the Notation used in Exercise 9.21. Further assume that the spaces $X_n$ are compact for all $n$. Show (without using Theorem ?? below) that $(X, d)$ is compact. Hint: Either use Cantor’s method to show every sequence $\{x_m\}_{m=1}^\infty \subset X$ has a convergent subsequence or alternatively show $(X, d)$ is complete and totally bounded. (Compare with Example 10.10 and see Theorem ?? below for the general version of this theorem.)
Connectedness in Metric Spaces

**Definition 11.1.** \((X, \tau)\) is **disconnected** if there exist non-empty open sets \(U\) and \(V\) of \(X\) such that \(U \cap V = \emptyset\) and \(X = U \cup V\). We say \(\{U, V\}\) is a **disconnection** of \(X\). The topological space \((X, \tau)\) is called **connected** if it is not disconnected, i.e. if there is no disconnection of \(X\). If \(A \subset X\) we say \(A\) is connected if \((A, \tau_A)\) is connected where \(\tau_A\) is the relative topology on \(A\). Explicitly, \(A\) is disconnected in \((X, \tau)\) iff there exists \(U, V, U \cap V = \emptyset, A \subset U \cup V\).

The reader should check that the following statement is an equivalent definition of connectivity. A topological space \((X, \tau)\) is connected iff the only sets \(X, \tau\) such that \(U \cap V = \emptyset, U \cap A \neq \emptyset, A \cap U \cap V = \emptyset\) and \(A \subset U \cup V\).

**Remark 11.2.** Let \(A \subset Y \subset X\). Then \(A\) is connected in \(X\) iff \(A\) is connected in \(Y\).

**Proof.** Since
\[
\tau_A := \{V \cap A : V \subset X\} = \{V \cap A \cap Y : V \subset X\} = \{U \cap A : U \subset Y\},
\]
the relative topology on \(A\) inherited from \(X\) is the same as the relative topology on \(A\) inherited from \(Y\). Since connectivity is a statement about the relative topologies on \(A\), \(A\) is connected in \(X\) iff \(A\) is connected in \(Y\).

**Theorem 11.3 (The Connected Subsets of \(\mathbb{R}\)).** The connected subsets of \(\mathbb{R}\) are intervals.

**Proof.** Suppose that \(A \subset \mathbb{R}\) is a connected subset and that \(a, b \in A\) with \(a < b\). If there exists \(c \in (a, b)\) such that \(c \notin A\), then \(U := (−\infty, c) \cap A\) and \(V := (c, \infty) \cap A\) would form a disconnection of \(A\). Hence \((a, b) \subset A\). Let \(\alpha := \inf(A)\) and \(\beta := \sup(A)\) and choose \(\alpha_n, \beta_n \in A\) such that \(\alpha_n < \beta_n\) and \(\alpha_n \downarrow \alpha\) and \(\beta_n \uparrow \beta\) as \(n \to \infty\). By what we have just shown, \((\alpha_n, \beta_n) \subset A\) for all \(n\) and hence \((\alpha, \beta) = \bigcup_{n=1}^{\infty} (\alpha_n, \beta_n) \subset A\). From this it follows that \(A = (\alpha, \beta)\).

Conversely suppose that \(A\) is a sub-interval of \(\mathbb{R}\). For the sake of contradiction, suppose that \(\{U, V\}\) is a disconnection of \(A\) with \(a \in U, b \in V\). After relabelling \(U\) and \(V\) if necessary we may assume that \(a < b\). Since \(A\) is an interval \([a, b] \subset A\). Let \(p = \sup([a, b] \cap U)\), then because \(U\) and \(V\) are open, \(a < p < b\). Now \(p\) cannot be in \(U\) for otherwise \(\sup([a, b] \cap U) > p\) and \(p\) cannot be in \(V\) for otherwise \(p < \sup([a, b] \cap U)\). From this it follows that \(p \notin U \cup V\) and hence \(A \neq U \cup V\) contradicting the assumption that \(\{U, V\}\) is a disconnection of \(A\).

**Alternative proof of the converse.** Because of Theorem 11.7 below, it suffices to assume that \(A\) is an open interval. For the sake of contradiction, suppose that \(\{U, V\}\) is a disconnection of \(A\) with \(a \in U, b \in V\). After relabelling \(U\) and \(V\) if necessary we may assume that \(a < b\). Let \(J_a = (α, β)\) be the maximal open interval in \(U\) which contains \(a\). (See Exercise 11.8 or Remark 11.4 for the structure of open subsets of \(\mathbb{R}\).) If \(β \in U\) we could extend \(J_a\) to the right and still be in \(U\) violating the definition of \(β\). Moreover we can not have \(β \in V\) because in this case \(J_a\) would not be in \(U\). Therefore \(β \notin U \cup V\) as \(A\) and on the other hand \(a < β < b\) and so \(β \in A\) as \(A\) is an interval and we have reached the desired contradiction.

**Remark 11.4 (Structure of open sets in \(\mathbb{R}\)).** Let \(V \subset \mathbb{R}\) be an open set. For \(x \in V\), let \(a_x := \inf\{a : (a, x] \subset V\}\) and \(b_x := \sup\{b : [x, b) \subset V\}\). Since \(V\) is open, \(a_x < x < b_x\) and it is easily seen that \(J_x := (a_x, b_x) \subset V\). Moreover if \(y \in V\) and \(J_x \cap J_y \neq \emptyset\), then \(J_x = J_y\). The collection \(\{J_x : x \in V\}\) is at most countable since we may label each \(J \in \{J_x : x \in V\}\) by choosing a rational number \(r \in J\). Letting \(\{J_n : n < N\}\), with \(N = \infty\) allowed, be an enumeration of \(\{J_x : x \in V\}\), we have \(V = \bigsqcup_{n < N} J_n\) as desired.

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

**Lemma 11.5.** Suppose that \(f : X \to Y\) is a continuous map between topological spaces. Then \(f(X) \subset Y\) is connected if \(X\) is connected.

**Proof.** Suppose \(f : X \to Y\) is a continuous map between topological spaces, the space \(X\) is connected and the space \(Y\) is “\(T_1\),” i.e. \(\{y\}\) is a closed set for all \(y \in Y\) as in Definition 11 below. Furthermore assume \(f\) is locally constant, i.e. for all \(x \in X\) there exists an open neighborhood \(V\) of \(x\) in \(X\) such that \(f|_V\) is constant. Then \(f\) is constant, i.e. \(f(X) = \{y_0\}\) for some \(y_0 \in Y\). To prove this, let \(y_0 \in f(X)\) and let \(W := f^{-1}(\{y_0\})\). Since \(\{y_0\} \subset Y\)
Therefore \( f(x + v) = f(x) \) for all \( |v| < \varepsilon \) and this shows \( f \) is locally constant. Hence, by what we have just proved, \( f \) is constant on \( X \).

**Theorem 11.7 (Properties of Connected Sets).** Let \((X, \tau)\) be a topological space.

1. If \( B \subset X \) is a connected set and \( X \) is the disjoint union of two open sets \( U \) and \( V \), then either \( B \subset U \) or \( B \subset V \).
2. If \( A \subset X \) is connected,
   a) then \( \bar{A} \) is connected.
   
   b) More generally, if \( A \subset \text{acc}(A) \) or \( B \subset \text{bd}(A) \), then \( A \cup B \) is connected as well. (Recall that \( \text{acc}(A) \) — the set of accumulation points of \( A \) was defined in Definition 9.13 above. Moreover by Exercise 9.7, we know that \( \text{acc}(A) \cap A = \text{bd}(A) \setminus \bar{A} \). What we are really showing here is that for any \( B \subset \bar{A} \), then \( \bar{A} \) is connected.)
3. If \( \{E_\alpha\}_{\alpha \in A} \) is a collection of connected sets such that \( E_\alpha \cap E_\beta \neq \emptyset \) for all \( \alpha, \beta \in A \) then \( Y := \bigcup_{\alpha \in A} E_\alpha \) is connected as well.
4. Suppose \( A, B \subset X \) are non-empty connected subsets of \( X \) such that \( A \cap B \neq \emptyset \), then \( A \cup B \) is connected in \( X \).
5. Every point \( x \in X \) is contained in a unique maximal connected subset \( C_x \) of \( X \) and this subset is closed. The set \( C_x \) is called the connected component of \( x \).

**Proof.**

1. Since \( B \) is the disjoint union of the relatively open sets \( B \cap U \) and \( B \cap V \), we must have \( B \cap U = B \) or \( B \cap V = B \) for otherwise \( \{B \cap U, B \cap V\} \) would be a disconnection of \( B \).

---

2. a) Let \( Y = \bar{A} \) be equipped with the relative topology from \( X \). Suppose that \( U, V \subset_o Y \) form a disconnection of \( Y = \bar{A} \). Then by 1. either \( A \subset U \) or \( A \subset V \). Say that \( A \subset U \). Since \( U \) is both open and closed in \( Y \), it follows that \( Y = \bar{A} \subset U \). Therefore \( V = \emptyset \) and we have a contradiction to the assumption that \( \{U, V\} \) is a disconnection of \( Y = \bar{A} \). Hence we must conclude that \( Y = \bar{A} \) is connected as well.

b) Now let \( Y = A \cup B \) with \( B \subset \text{acc}(A) \), then 

\[
\bar{A}^Y = \bar{A} \cap Y = (A \cup \text{acc}(A)) \cap Y = A \cup B.
\]

Because \( A \) is connected in \( Y \), by (2a) \( Y = A \cup B = \bar{A}^Y \) is also connected.

3. Let \( Y := \bigcup_{\alpha \in A} E_\alpha \). By Remark 11.2 we know that \( E_\alpha \) is connected in \( Y \) for each \( \alpha \in A \). If \( \{U, V\} \) were a disconnection of \( Y \), by item 1, either \( E_\alpha \subset U \) or \( E_\alpha \subset V \) for all \( \alpha \). Let \( A = \{ \alpha \in A : E_\alpha \subset U \} \) then \( U = \bigcup_{\alpha \in A} E_\alpha \) and \( V = \bigcup_{\alpha \in A \setminus A} E_\alpha \). (Notice that neither \( A \) or \( A \setminus A \) can be empty since \( U \) and \( V \) are not empty.) Since 

\[
\emptyset = U \cap V = \bigcup_{\alpha \in A, \beta \in A \setminus A} (E_\alpha \cap E_\beta) \neq \emptyset.
\]

we have reached a contradiction and hence no such disconnection exists.

4. (A good example to keep in mind here is \( X = \mathbb{R} \), \( A = (0, 1) \) and \( B = [1, 2) \).)

For sake of contradiction suppose that \( \{U, V\} \) were a disconnection of \( Y = A \cup B \). By item 1 either \( A \subset U \) or \( A \subset V \), say \( A \subset U \) in which case \( B \subset V \). Since \( Y = A \cup B \) we must have \( A = U \) and \( B = V \) and so we may conclude: \( A \) and \( B \) are disjoint subsets of \( Y \) which are both open and closed. This implies 

\[
A = \bar{A}^Y = \bar{A} \cap Y = \bar{A} \cap (A \cup B) = A \cup (\bar{A} \cap B)
\]

and therefore 

\[
\emptyset = A \cap B = [A \cup (\bar{A} \cap B)] \cap B = \bar{A} \cap B \neq \emptyset
\]

which gives us the desired contradiction.

**Alternative proof.** Let \( A' := A \cup [B \cap \bar{A}] \) so that \( A \cup B = A' \cup B \). By item 2b. we know \( A' \) is still connected and since \( A' \cap B \neq \emptyset \) we may now apply item 3. to finish the proof.

5. Let \( C \) denote the collection of connected subsets \( C \subset X \) such that \( x \in C \). Then by item 3., the set \( C_x := \cup C \) is also a connected subset of \( X \) which contains \( x \) and clearly this is the unique maximal connected set containing \( x \). Since \( C_x \) is also connected by item 2) and \( C_x \) is maximal, \( C_x = C_x \), i.e. \( C_x \) is closed.

---

\(^1\) One may assume much less here. What we really need is for any \( \alpha, \beta \in A \) there exists \( \{\alpha_i\}_{i=0}^n \) in \( A \) such that \( \alpha_0 = \alpha, \alpha_n = \beta \), and \( E_{\alpha_i} \cap E_{\alpha_{i+1}} \neq \emptyset \) for all \( 0 \leq i < n \).

Moreover if we make use of item 4. it suffices to assume that 

\[
E_{\alpha_i} \cap E_{\alpha_{i+1}} \subset E_{\alpha_i} \cap E_{\alpha_{i+1}} \neq \emptyset \text{ for all } 0 \leq i < n.
\]
Proposition 11.10. Let \( X \) be a topological space.

1. If \( X \) is path connected then \( X \) is connected.
2. If \( X \) is connected and locally path connected, then \( X \) is path connected.
3. If \( X \) is any connected open subset of \( \mathbb{R}^n \), then \( X \) is path connected.

Proof. The reader is asked to prove this proposition in Exercises 11.4 – 11.6 below.

Proposition 11.11 (Stability of Connectedness Under Products). Let \((X_{\alpha}, \tau_{\alpha})\) be connected topological spaces. Then the product space \( X_A = \prod_{\alpha \in A} X_{\alpha} \) is connected with the product topology is connected.

Proof. Let us begin with the case of two factors, namely assume that \( X \) and \( Y \) are connected topological spaces, then we will show that \( X \times Y \) is connected as well. Given \( x \in X \), let \( f_x : Y \to X \times Y \) be the map \( f_x(y) = (x, y) \) and notice that \( f_x \) is continuous since \( \pi_Y \circ f_x(y) = y \) and \( \pi_X \circ f_x(y) = x \) are continuous maps. From this we conclude that \( \{x\} \times Y \) is connected by Lemma 11.10. A similar argument shows \( X \times \{y\} \) is connected for all \( y \in Y \).

Let \( p = (x_0, y_0) \in X \times Y \) and \( C_p \) denote the connected component of \( p \). Since \( \{x_0\} \times Y \) is connected and \( p \in \{x_0\} \times Y \) it follows that \( \{x_0\} \times Y \subseteq C_p \) and hence \( C_p \) is also the connected component \( \{x_0, y_0\} \) for all \( y_0 \in Y \). Similarly, \( X \times \{y\} \subseteq C_{(x_0, y)} \) is connected, and therefore \( X \times \{y\} \subseteq C_p \). So we have shown \( (x, y) \in C_p \) for all \( x \in X \) and \( y \in Y \), see Figure 11.1. By induction the theorem holds whenever \( A \) is a finite set, i.e. for products of a finite number of connected spaces.

For the general case, again choose a point \( p \in X_A = X^A \) and again let \( C = C_p \) be the connected component of \( p \). Recall that \( C_p \) is closed and therefore \( C_p \) is a proper subset of \( X_A \), then \( X_A \setminus C_p \) is a non-empty open set. By the definition of the product topology, this would imply that \( X_A \setminus C_p \) contains a non-empty open set of the form

\[
V := \cap_{\alpha \in A} \pi_{\alpha}^{-1}(V_{\alpha}) = V_A \times X_{\alpha \setminus A} \subset X_A \setminus C_p
\]

where \( A \subset A \) and \( V_\alpha \in \tau_\alpha \) for all \( \alpha \in A \).

On the other hand, let \( \varphi : X_A \to X_A \setminus C_p \) by \( \varphi(y) = x \) where

\[
x_\alpha = \begin{cases} y_\alpha & \text{if } \alpha \in A \\ p_\alpha & \text{if } \alpha \notin A. \end{cases}
\]

If \( \alpha \in A \), \( \pi_\alpha \circ \varphi(y) = y_\alpha = \pi_\alpha(y) \) and if \( \alpha \in A \setminus A \) then \( \pi_\alpha \circ \varphi(y) = p_\alpha \) so that in every case \( \pi_\alpha \circ \varphi : X_A \to X_A \setminus C_p \) is continuous and therefore \( \varphi \) is continuous. Since \( X_A \) is a product of a finite number of connected spaces and so is connected and thus so is the continuous image, \( \varphi(X_A) = X_A \times \{p_\alpha\}_{\alpha \in A \setminus A} \subset X_A \). Since \( p \in \varphi(X_A) \) we must have

\[
X_A \times \{p_\alpha\}_{\alpha \in A \setminus A} \subset C_p.
\]

Hence it follows from Eqs. 11.1 and 11.2 that

\[
V_A \times \{p_\alpha\}_{\alpha \in A \setminus A} = \left( X_A \times \{p_\alpha\}_{\alpha \in A \setminus A} \right) \cap V \subset C_p \cap [X_A \setminus C_p] = \emptyset
\]

which is a contradiction since \( V_A \times \{p_\alpha\}_{\alpha \in A \setminus A} \neq \emptyset \).
11.0.4 Connectedness Problems

Exercise 11.1. Show any non-trivial interval in \( \mathbb{Q} \) is disconnected.

Exercise 11.2. Suppose \( a < b \) and \( f : (a, b) \to \mathbb{R} \) is a non-decreasing function. Show if \( f \) satisfies the intermediate value property (see Theorem 11.8), then \( f \) is continuous.

Exercise 11.3. Suppose \( -\infty < a < b \leq \infty \) and \( f : [a, b) \to \mathbb{R} \) is a strictly increasing continuous function. Using the intermediate value theorem, one sees that \( f([a,b)) \) is an interval and since \( f \) is strictly increasing it must of the form \( [c,d) \) for some \( c \in \mathbb{R} \) and \( d \in \mathbb{R} \) with \( c < d \). Show the inverse function \( f^{-1} : [c,d) \to [a, b) \) is continuous and is strictly increasing. In particular if \( n \in \mathbb{N} \), apply this result to \( f(x) = x^n \) for \( x \in [0, \infty) \) to construct the positive \( n^{th} \) root of a real number. Compare with Exercise ??.

Exercise 11.4. Prove item 1. of Proposition 11.10 i.e. if \( X \) is path connected then \( X \) is connected.. Hint: show \( X \) is not connected implies \( X \) is not path connected.

Exercise 11.5. Prove item 2. of Proposition 11.10 i.e. if \( X \) is connected and locally path connected, then \( X \) is path connected. Hint: fix \( x_0 \in X \) and let \( W \) denote the set of \( x \in X \) such that there exists \( \sigma \in C([0,1],X) \) satisfying \( \sigma(0) = x_0 \) and \( \sigma(1) = x \). Then show \( W \) is both open and closed.

Exercise 11.6. Prove item 3. of Proposition 11.10 i.e. if \( X \) is any connected open subset of \( \mathbb{R}^n \), then \( X \) is path connected.

Exercise 11.7. Let

\[
X := \{ (x, y) \in \mathbb{R}^2 : y = \sin(x^{-1}) \text{ with } x \neq 0 \} \cup \{ (0,0) \}
\]

equipped with the relative topology induced from the standard topology on \( \mathbb{R}^2 \). Show \( X \) is connected but not path connected.
Appendix: Notation and Logic

The following abbreviations along with their negations are used throughout these notes.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
<th>Negation</th>
</tr>
</thead>
<tbody>
<tr>
<td>∀</td>
<td>for all</td>
<td>∃</td>
</tr>
<tr>
<td>∃</td>
<td>there exists</td>
<td>∀</td>
</tr>
<tr>
<td>, or “space” then</td>
<td>∃</td>
<td>∀, or “space”</td>
</tr>
<tr>
<td>∋</td>
<td>such that</td>
<td>∈</td>
</tr>
<tr>
<td>a.a.</td>
<td>almost always</td>
<td>i.o.</td>
</tr>
<tr>
<td>i.o.</td>
<td>infinitely often</td>
<td>a.a.</td>
</tr>
<tr>
<td>=</td>
<td>equals</td>
<td>≠</td>
</tr>
<tr>
<td>≠</td>
<td>not equals</td>
<td>=</td>
</tr>
<tr>
<td>≤</td>
<td>less than or equal</td>
<td>&gt;</td>
</tr>
<tr>
<td>&gt;</td>
<td>greater than</td>
<td>≤</td>
</tr>
</tbody>
</table>

Here are some examples.

1. $a_n = b_n$ i.o. $n \iff \# \{ n : a_n = b_n \} = \infty$. The negation of $\# \{ n : a_n = b_n \} = \infty$ is $\# \{ n : a_n = b_n \} < \infty \iff a_n \neq b_n$ for a.a. $n$.

2. $\lim_{n \to \infty} a_n = L$ is by definition the statement:
   \[ \forall \varepsilon > 0 \exists \, \exists N \in \mathbb{N} \exists \forall n \geq N, \ |L - a_n| \leq \varepsilon. \]
   This may also be written as $\forall \varepsilon > 0, \ |L - a_n| \leq \varepsilon$ for a.a. $n$.

3. The negation of the previous statement is $\lim_{n \to \infty} a_n \neq L$ which translates to $\exists \varepsilon > 0 \exists \forall N \in \mathbb{N}, \exists n \geq N \exists |L - a_n| > \varepsilon$.
   This last statement is also equivalent to;
   $\exists \varepsilon > 0 \exists |L - a_n| > \varepsilon$ i.o. $n$.
   It is sometimes useful to reformulate this last statement as; there exists $\varepsilon > 0$ and an increasing function $\mathbb{N} \ni k \rightarrow n_k \in \mathbb{N}$ such that $|L - a_{n_k}| > \varepsilon$ for all $k \in \mathbb{N}$. 
Appendix: More Set Theoretic Properties (highly optional)

B.1 Appendix: Zorn’s Lemma and the Hausdorff Maximal Principle (optional)

Definition B.1. A partial order \( \leq \) on \( X \) is a relation with following properties:

1. If \( x \leq y \) and \( y \leq z \) then \( x \leq z \).
2. If \( x \leq y \) and \( y \leq x \) then \( x = y \).
3. \( x \leq x \) for all \( x \in X \).

Example B.2. Let \( Y \) be a set and \( X = 2^Y \). There are two natural partial orders on \( X \):

1. Ordered by inclusion, \( A \leq B \) is \( A \subset B \) and
2. Ordered by reverse inclusion, \( A \leq B \) if \( B \subset A \).

Definition B.3. Let \( (X, \leq) \) be a partially ordered set we say \( X \) is linearly or totally ordered if for all \( x, y \in X \) either \( x \leq y \) or \( y \leq x \). The real numbers \( \mathbb{R} \) with the usual order \( \leq \) is a typical example.

Definition B.4. Let \( (X, \leq) \) be a partial ordered set. We say \( x \in X \) is a maximal element if for all \( y \in X \) such that \( y \geq x \) implies \( y = x \), i.e. there is no element larger than \( x \). An upper bound for a subset \( E \) of \( X \) is an element \( x \in X \) such that \( x \geq y \) for all \( y \in E \).

Example B.5. Let
\[
X = \{ \emptyset, \{1\}, \{1, 2\}, \{1, 2, 3\}, \{2, 4\}, \{2\} \}
\]
ondered by set inclusion. Then \( b \) and \( d \) are maximal elements despite that fact that \( b \notin d \) and \( d \notin b \). We also have;

1. If \( E = \{a, c, e\} \), then \( E \) has no upper bound.
2. If \( E = \{a, e\} \), then \( b \) is an upper bound.
3. If \( E = \{e\} \), then \( b \) and \( d \) are upper bounds.

Theorem B.6. The following are equivalent.

1. The axiom of choice: to each collection \( \{X_\alpha\}_{\alpha \in A} \) of non-empty sets there exists a “choice function,” \( x : A \to \prod_{\alpha \in A} X_\alpha \) such that \( x(\alpha) \in X_\alpha \) for all \( \alpha \in A \), i.e. \( \prod_{\alpha \in A} X_\alpha \neq \emptyset \).

2. The Hausdorff Maximal Principle: Every partially ordered set has a maximal (relative to the inclusion order) linearly ordered subset.

3. Zorn’s Lemma: If \( X \) is partially ordered set such that every linearly ordered subset of \( X \) has an upper bound, then \( X \) has a maximal element.

Proof. (2 \( \Rightarrow \) 3) Let \( X \) be a partially ordered subset as in 3 and let \( F = \{E \subset X : E \text{ is linearly ordered}\} \) which we equip with the inclusion partial ordering. By 2, there exist a maximal element \( E \in F \). By assumption, the linearly ordered set \( E \) has an upper bound \( x \in X \). The element \( x \) is maximal, for if \( y \in Y \) and \( y \geq x \), then \( E \cup \{y\} \) is still an linearly ordered set containing \( E \). So by maximality of \( E \), \( E = E \cup \{y\} \), i.e. \( y \in E \) and therefore \( y \leq x \) showing which combined with \( y \geq x \) implies that \( y = x \).

(3 \( \Rightarrow \) 1) Let \( \{X_\alpha\}_{\alpha \in A} \) be a collection of non-empty sets, we must show \( \prod_{\alpha \in A} X_\alpha \) is not empty. Let \( G \) denote the collection of functions \( g : D(g) \to \prod_{\alpha \in A} X_\alpha \) such that \( D(g) \) is a subset of \( A \), and for all \( \alpha \in D(g) \), \( g(\alpha) \in X_\alpha \). Notice that \( G \) is not empty, for we may let \( \alpha_0 \in A \) and \( x_0 \in X_{\alpha_0} \) and then set \( D(g) = \{\alpha_0\} \) and \( g(\alpha_0) = x_0 \) to construct an element of \( G \). We now put a partial order \( \leq \) as follows. We say that \( f \leq g \) for \( f, g \in G \) provided that \( D(f) \subset D(g) \) and \( f|_{D(f)} = g|_{D(f)} \). If \( \Phi \subset G \) is a linearly ordered set, let \( \Phi(h) = \bigcup_{\alpha \in \Phi} D(g) \) for \( \alpha \in D(g) \) let \( h(\alpha) = g(\alpha) \). Then \( h \in \Phi \) is an upper bound for \( \Phi \). So by Zorn’s

---

1. If \( X \) is a countable set we may prove Zorn’s Lemma by induction. Let \( \{x_n\}_{n=1}^\infty \) be an enumeration of \( X \), and define \( E_n \subset X \) inductively as follows. For \( n = 1 \) let \( E_1 = \{x_1\} \), and if \( E_n \) has been chosen, let \( E_{n+1} = E_n \cup \{x_{n+1}\} \) if \( x_{n+1} \) is an upper bound for \( E_n \) otherwise let \( E_{n+1} = E_n \). The set \( E = \bigcup_{n=1}^\infty E_n \) is a linearly ordered (you check) subset of \( X \) and hence by assumption \( E \) has an upper bound, \( x \in X \). I claim that his element is maximal, for if there exists \( y = x_m \in X \) such that \( y \geq x \), then \( x_m \) would be an upper bound for \( E_{m-1} \) and therefore \( y = x_m \in E_m \subset E \). That is to say if \( y \geq x \), then \( y \in E \) and hence \( y \leq x \), so \( y = x \). (Hence we may view Zorn’s lemma as a “jazzy” up version of induction.)

2. Similarly one may show that \( 3 \Rightarrow 2 \). Let \( F = \{E \subset X : E \text{ is linearly ordered}\} \) and order \( F \) by inclusion. If \( M \subset F \) is linearly ordered, let \( E = \bigcup M \). If \( x, y \in E \) then \( x \in A \) and \( y \in B \) for some \( A, B \subset M \). Now \( M \) is linearly ordered by set inclusion so \( A \subset B \) or \( B \subset A \) i.e. \( x, y \in A \) or \( x, y \in B \). Since \( A \) and \( B \) are linearly order we must have either \( x \leq y \) or \( y \leq x \). That is to say \( E \) is linearly ordered. Hence by 3, there exists a maximal element \( E \in F \) which is the assertion in 2.
Lemma there exists a maximal element \( h \in \mathcal{G} \). To finish the proof we need only show that \( D(h) = A \). If this were not the case, then let \( \alpha_0 \in A \setminus D(h) \) and \( x_0 \in X_{\alpha_0} \). We may now define \( D(h) = D(h) \cup \{ \alpha_0 \} \) and 
\[
\hat{h}(\alpha) = \begin{cases} 
  h(\alpha) & \text{if } \alpha \in D(h) \\
  x_0 & \text{if } \alpha = \alpha_0.
\end{cases}
\]

Then \( h \leq \hat{h} \) while \( h \neq \hat{h} \) violating the fact that \( h \) was a maximal element.

(1 \( \Rightarrow \) 2) Let \((X, \leq)\) be a partially ordered set. Let \( \mathcal{F} \) be the collection of linearly ordered subsets of \( X \) which we order by set inclusion. Given \( x_0 \in X \), \( \{x_0\} \in \mathcal{F} \) is linearly ordered set so that \( \mathcal{F} \neq \emptyset \). Fix an element \( P_0 \in \mathcal{F} \). If \( P_0 \) is not maximal there exists \( P_1 \in \mathcal{F} \) such that \( P_0 \not\subseteq P_1 \). In particular we may choose \( x \not\in P_0 \) such that \( P_0 \cup \{x\} \in \mathcal{F} \). The idea now is to keep repeating this process of adding points \( x \in X \) until we construct a maximal element \( P \) of \( \mathcal{F} \). We now have to take care of some details. We may assume with out loss of generality that \( \mathcal{F} \) is a non-empty set. For \( P \in \mathcal{F} \), let \( P^* = \{x \in X : P \cup \{x\} \in \mathcal{F}\} \). As the above argument shows, \( P^* \neq \emptyset \) for all \( P \in \mathcal{F} \). Using the axiom of choice, there exists \( f \in \prod_{P \in \mathcal{F}} P^* \).

We now define \( g : \mathcal{F} \to \mathcal{F} \) by
\[
g(P) = \begin{cases} 
  P & \text{if } P \text{ is maximal} \\
  P \cup \{f(x)\} & \text{if } P \text{ is not maximal.}
\end{cases}
\]

(B.1)

The proof is completed by Lemma B.7 below which shows that \( g \) must have a fixed point \( P \in \mathcal{F} \). This fixed point is maximal by construction of \( g \).

\begin{lemma}
\( g : \mathcal{F} \to \mathcal{F} \) defined in Eq. (B.1) has a fixed point.\footnote{Here is an easy proof if the elements of \( \mathcal{F} \) happened to all be finite sets and there existed a set \( P \in \mathcal{F} \) with a maximal number of elements. In this case the condition that \( P \subseteq g(P) \) would imply that \( P = g(P) \), otherwise \( g(P) \) would have more elements than \( P \).}
\end{lemma}

\begin{proof}
The idea of the proof is as follows. Let \( P_0 \in \mathcal{F} \) be chosen arbitrarily. Notice that \( \Phi = \{g^n(P_0)\}_{n=0}^\infty \subset \mathcal{F} \) is a linearly ordered set and it is therefore easily verified that \( P_1 = \bigcup_{n=0}^\infty g^n(P_0) \in \mathcal{F} \). Similarly we may repeat the process to construct \( P_2 = \bigcup_{n=0}^\infty g^n(P_1) \in \mathcal{F} \) and \( P_3 = \bigcup_{n=0}^\infty g^n(P_2) \in \mathcal{F} \), etc. etc. Then take \( P_\infty = \bigcup_{n=0}^\infty P_n \) and start again with \( P_0 \) replaced by \( P_\infty \). Then keep going this way until eventually the sets stop increasing in size, in which case we have found our fixed point. The problem with this strategy is that we may never win. (This is very reminiscent of constructing measurable sets and the way out is to use measure theoretic like arguments.)

Let us now start the formal proof. Again let \( P_0 \in \mathcal{F} \) and let \( \mathcal{F}_1 = \{P \in \mathcal{F} : P_0 \subseteq P\} \). Notice that \( \mathcal{F}_1 \) has the following properties:

1. \( P_0 \in \mathcal{F}_1 \).
2. If \( \Phi \subset \mathcal{F}_1 \) is a totally ordered (by set inclusion) subset then \( \cup \Phi \in \mathcal{F}_1 \).
3. If \( P \in \mathcal{F}_1 \) then \( g(P) \in \mathcal{F}_1 \).

Let us call a general subset \( \mathcal{F}' \subset \mathcal{F} \) satisfying these three conditions a tower and let
\[
\mathcal{F}_0 = \cap \{\mathcal{F}' : \mathcal{F}' \text{ is a tower}\}.
\]

Standard arguments show that \( \mathcal{F}_0 \) is still a tower and clearly is the smallest tower containing \( P_0 \). (Morally speaking \( \mathcal{F}_0 \) consists of all of the sets we were trying to constructed in the “idea section” of the proof.) We now claim that \( \mathcal{F}_0 \) is a linearly ordered subset of \( \mathcal{F} \). To prove this let \( \Gamma \subset \mathcal{F}_0 \) be the linearly ordered set
\[
\Gamma = \{C \in \mathcal{F}_0 : \text{for all } A \in \mathcal{F}_0 \text{ either } A \subset C \text{ or } C \subset A\}.
\]

Shortly we will show that \( \Gamma \subset \mathcal{F}_0 \) is a tower and hence that \( \mathcal{F}_0 = \Gamma \). That is to say \( \mathcal{F}_0 \) is linearly ordered. Assuming this for the moment let us finish the proof.

Let \( P \equiv \cup \mathcal{F}_0 \) which is in \( \mathcal{F}_0 \) by property 2 and is clearly the largest element in \( \mathcal{F}_0 \). By 3. it now follows that \( P \subset g(P) \subset \mathcal{F}_0 \) and by maximality of \( P \), we have \( g(P) = P \), the desired fixed point. So to finish the proof, we must show that \( \Gamma \) is a tower. First off it is clear that \( P_0 \in \Gamma \) so in particular \( \Gamma \) is not empty. For each \( C \in \Gamma \) let
\[
\Phi_C := \{A \in \mathcal{F}_0 : \text{either } A \subset C \text{ or } C \subset A\}.
\]

We will begin by showing that \( \Phi_C \subset \mathcal{F}_0 \) is a tower and therefore that \( \Phi_C = \mathcal{F}_0 \).
1. \( \Phi_C \subset \mathcal{F}_0 \) since \( P_0 \subset C \) for all \( C \in \Gamma \subset \mathcal{F}_0 \). 2. If \( \Phi_C \subset \mathcal{F}_0 \) is totally ordered by set inclusion, then \( \mathcal{F}_0 := \cup \Phi_C \neq \mathcal{F}_0 \). We must show \( \Phi \subset \mathcal{F}_0 \), that is that \( \mathcal{F}_0 \subset C \) or \( C \subset A \). Now if \( A \subset C \) for all \( A \in \Phi \), then \( \Phi \subset C \) and hence \( \Phi \subset \mathcal{F}_0 \). On the other hand if there is some \( A \in \Phi \) such that \( g(C) \subset A \) then clearly \( g(C) \subset \Phi \) and again \( \Phi \subset \mathcal{F}_0 \). 3. Given \( A \in \Phi_C \) we must show \( g(A) \in \Phi_C \), i.e. that
\[
g(A) \subset C \text{ or } C \subset g(A). \tag{B.2}
\]

There are three cases to consider: either \( A \not\subseteq C \), \( A = C \), or \( g(C) \subset A \). In the case \( A = C \), \( g(C) = g(A) \subset g(A) \) and if \( g(C) \subset A \) then \( g(C) \subset A \subset g(A) \) and Eq. (B.2) holds in either of these cases. So assume that \( A \not\subseteq C \). Since \( C \in \Gamma \), either \( g(A) \subset C \) (in which case we are done) or \( C \subset g(A) \). Hence we may assume that
\[
A \not\subseteq C \subset g(A).
\]

Now if \( C \) were a proper subset of \( g(A) \) it would then follow that \( g(A) \setminus A \) would consist of at least two points which contradicts the definition of \( g \). Hence we
must have \( g(A) = C \subset C \) and again Eq. (B.2) holds, so \( \Phi_C \) is a tower. It is now easy to show \( \Gamma \) is a tower. It is again clear that \( P_0 \in \Gamma \) and Property 2. may be checked for \( \Gamma \) in the same way as it was done for \( \Phi_C \) above. For Property 3., if \( C \in \Gamma \) we may use \( \Phi_C = F_0 \) to conclude for all \( A \in F_0 \), either \( A \subset C \subset g(C) \) or \( g(C) \subset A \), i.e. \( g(C) \in \Gamma \). Thus \( \Gamma \) is a tower and we are done. ■

Here is an example of using Zorn’s lemma.

**Proposition B.8.** Suppose that \( X \) and \( Y \) are non-empty sets, then either there exists an injective function, \( f : X \to Y \), or an injective function \( g : Y \to X \). In other words, either \( \text{card}(X) \leq \text{card}(Y) \) or \( \text{card}(Y) \leq \text{card}(X) \).

**Proof.** Let \( \mathcal{F} \) be the collection of injective functions, \( u : D(u) \to Y \) where \( D(u) \) is a non-empty subset of \( X \). We say that \( u \leq v \) for \( u, v \in \mathcal{F} \) provided \( D(u) \subset D(v) \) and \( u = v|_{D(u)} \). One now checks that \( (\mathcal{F}, \leq) \) is a partially ordered set such that every linearly ordered subset of \( \mathcal{F} \) has an upper bound. Therefore, by an application of Zorn’s lemma, \( \mathcal{F} \) has a maximal element, \( U \).

If \( D(U) = X \), we take \( f = U \) and we have constructed an injective map from \( X \) to \( Y \). If \( D(U) \neq X \), then \( \text{Ran}(U) := U(D(U)) = Y \). [Indeed, if not we could find \( x \in X \setminus D(U) \) and \( y \in Y \setminus \text{Ran}(U) \) and then extend \( U \) to \( D(U) \cup \{x\} \) by setting \( U(x) = y \). The extended \( U \) is still injective and hence would violate the maximality of \( U \).] In this case we take \( g := U^{-1} : Y \to \mathcal{D}(U) \subset X \). ■

### B.2 Cardinality

In mathematics, the essence of counting a set and finding a result \( n \), is that it establishes a one to one correspondence (or bijection) of the set with the set of numbers \( \{1, 2, \ldots, n\} \). A fundamental fact, which can be proved by mathematical induction, is that no bijection can exist between \( \{1, 2, \ldots, n\} \) and \( \{1, 2, \ldots, m\} \) unless \( n = m \); this fact (together with the fact that two bijections can be composed to give another bijection) ensures that counting the same set in different ways can never result in different numbers (unless an error is made). This is the fundamental mathematical theorem that gives counting its purpose; however you count a (finite) set, the answer is the same. In a broader context, the theorem is an example of a theorem in the mathematical field of (finite) combinatorics—hence (finite) combinatorics is sometimes referred to as “the mathematics of counting.”

Many sets that arise in mathematics do not allow a bijection to be established with \( \{1, 2, \ldots, n\} \) for any natural number \( n \); these are called infinite sets, while those sets for which such a bijection does exist (for some \( n \)) are called finite sets. Infinite sets cannot be counted in the usual sense; for one thing, the mathematical theorems which underlie this usual sense for finite sets are false for infinite sets. Furthermore, different definitions of the concepts in terms of which these theorems are stated, while equivalent for finite sets, are inequivalent in the context of infinite sets.

The notion of counting may be extended to them in the sense of establishing (the existence of) a bijection with some well understood set. For instance, if a set can be brought into bijection with the set of all natural numbers, then it is called “countably infinite.” This kind of counting differs in a fundamental way from counting of finite sets, in that adding new elements to a set does not necessarily increase its size, because the possibility of a bijection with the original set is not excluded. For instance, the set of all integers (including negative numbers) can be brought into bijection with the set of natural numbers, and even seemingly much larger sets like that of all finite sequences of rational numbers are still (only) countably infinite. Nevertheless there are sets, such as the set of real numbers, that can be shown to be “too large” to admit a bijection with the natural numbers, and these sets are called “uncountable.” Sets for which there exists a bijection between them are said to have the same cardinality, and in the most general sense counting a set can be taken to mean determining its cardinality. Beyond the cardinalities given by each of the natural numbers, there is an infinite hierarchy of infinite cardinalities, although only very few such cardinalities occur in ordinary mathematics (that is, outside set theory that explicitly studies possible cardinalities).

Counting, mostly of finite sets, has various applications in mathematics. One important principle is that if two sets \( X \) and \( Y \) have the same finite number of elements, and a function \( f : X \to Y \) is known to be injective, then it is also surjective, and vice versa. A related fact is known as the pigeonhole principle, which states that if two sets \( X \) and \( Y \) have finite numbers of elements \( n \) and \( m \) with \( n > m \), then any map \( f : X \to Y \) is not injective (so there exist two distinct elements of \( X \) that \( f \) sends to the same element of \( Y \)); this follows from the former principle, since if \( f \) were injective, then so would its restriction to a strict subset \( S \) of \( X \) with \( m \) elements, which restriction would then be surjective, contradicting the fact that for \( x \) in \( X \) outside \( S \), \( f(x) \) cannot be in the image of the restriction. Similar counting arguments can prove the existence of certain objects without explicitly providing an example. In the case of infinite sets this can even apply in situations where it is impossible to give an example; for instance there must exists real numbers that are not computable numbers, because the latter set is only countably infinite, but by definition a non-computable number cannot be precisely specified.

The domain of enumerative combinatorics deals with computing the number of elements of finite sets, without actually counting them; the latter usually being impossible because infinite families of finite sets are considered at once, such as the set of permutations of \( \{1, 2, \ldots, n\} \) for any natural number \( n \).
We have

**Theorem B.11 (Schröder-Bernstein Theorem).** If card \((X) ≤ card (Y)\) and card \((Y) ≥ card (X)\) if there exists a surjective map \(g : Y → X\), We say card \((X) = card (Y)\) if there exists a bijections, \(f : X → Y\).

**Proposition B.10.** We have card \((X) ≤ card (Y)\) iff card \((Y) ≥ card (X)\).

**Proof.** If \(f : X → Y\) is an injective map, define \(g : Y → X\) by \(g\{f(x)\} = f^{-1}\), and \(g\{\emptyset\} = x_0 \in X\) chosen arbitrarily. Then g : Y → X is surjective.

If \(g : Y → X\) is a surjective map, then \(x_0 := g^{-1}(\{x\}) \neq \emptyset\) for all \(x \in X\) and so by the axiom of choice there exists \(f \in \prod_{x \in X} Y_x\). Thus \(f : X → Y\) such that \(f(x) \in Y_x\) for all \(x\). As the \(\{Y_x\}_{x \in X}\) are pairwise disjoint, it follows that \(f\) is injective.

**Theorem B.11 (Schröder-Bernstein Theorem).** If card \((X) ≤ card (Y)\) and card \((Y) ≤ card (X)\), then card \((X) = card (Y)\). Stated more explicitly; if there exists injective maps \(f : X → Y\) and \(g : Y → X\), then there exists a bijective map, \(h : X → Y\).

**Proof.** Starting with an \(x \in X\) we may form the sequence of “ancestors” of \(X\), namely ancestor \((x) := (x, y_1, x_1, y_2, \ldots)\) where \(y_1 = g^{-1}(x)\), \(x_1 = f^{-1}(y_1)\), \(y_2 = g^{-1}(x_1)\), i.e.

\[
x \xrightarrow{g^{-1}} y_1 \xrightarrow{f^{-1}} x_1 \xrightarrow{g^{-1}} y_2 \xrightarrow{f^{-1}} \ldots
\]

We continue this process of inverse iterates as long as it is possible, i.e. we can construct \(y_{n+1}\) if \(x_n \in g\) (and \(x_{n+1}\) if \(y_{n+1}\) \(\in f\)). There are now three possibilities;

1. ancestor \((x)\) has infinite length so the process never gets stuck in which case we say \(x \in X_∞\), read as start in \(X\) and end never get stuck.
2. ancestor \((x)\) is finite and the last term in the sequence is in \(X\), in which case we say \(x \in X_X\) (read as start in \(X\) and get stuck in \(X\)).
3. ancestor \((x)\) is finite and the last term in the sequence is in \(Y\), in which case we say \(x \in X_Y\) (read as start in \(X\) and end in \(Y\)).

In this way we partition \(X\) into three disjoint sets, \(X_∞, X_X,\) and \(X_Y\). Similarly we may partition \(Y\) into \(Y_∞, Y_Y,\) and \(Y_Y\). Let us now observe that

1. \(f(X_∞) = Y_∞\). Indeed if \(x \in X_∞\) then \(f(x) = \langle x, \text{ancestor}(x) \rangle\) is an infinite sequence, i.e. \(f(x) \in Y_∞\). Moreover if \(y \in Y_∞\), then \(y = \langle y, \text{ancestor}(x) \rangle\) where \(f(x) = y\) so that \(x \in X_∞\) and \(y \in f(X_∞)\). Thus we have shown \(f : X_∞ → Y_∞\) is a bijection, i.e. card \((X_∞) = card (Y_∞)\).
2. \(f(X_X) = Y_X\). Indeed if \(x \in X_X\) then again \(f(x) = \langle x, \text{ancestor}(x) \rangle\) which ends in \(X\) so that \(f(x) \in Y_X\). Moreover if \(y \in Y_X\), then \(y = \langle y, \text{ancestor}(x) \rangle\) where \(f(x) = y\) so that \(x \in X_X\) and \(y \in f(X_X)\). Thus we have shown \(f : X_X → Y_X\) is a bijection, i.e. card \((X_X) = card (Y_X)\).
3. By the same argument as in item 2. it follow that \(g : Y_Y → X_Y\) is a bijection, i.e. card \((X_Y) = card (Y_Y)\).

The last three statements imply card \((X) = card (Y)\). We may in fact define a bijection, \(h : X → Y\), by

\[
h(x) = \begin{cases} f(x) & \text{if } x \in X_∞ \cup X_X \\ g^{-1}(x) & \text{if } x \in X_Y. \end{cases}
\]

**Definition B.12.** We say card \((X) < card (Y)\) if card \((X) ≤ card (Y)\) and card \((X) \neq card (Y)\), i.e. card \((X) < card (Y)\) if there exists an injective map, \(f : X → Y\), but not bijective map exists. Similarly we say card \((Y) > card (X)\) if card \((Y) ≥ card (X)\) and card \((Y) \neq card (X)\), i.e. card \((Y) > card (X)\) if there exists a surjective map \(g : Y → X\) but no bijective map exists.

**Proposition B.13.** For any non-empty set \(X\), card \((X) < card (2^X)\).

**Proof.** Define \(f : X → 2^X\) by \(f(x) = \{x\}\). Then \(f\) is an injective map and hence card \((X) ≤ card (2^X)\). Now suppose that card \((X) = card (2^X)\). Let \(X_0 = \{x ∈ X : x ∉ g(x)\} \subset X\). I claim that \(X_0 ∉ g(X)\).

Indeed suppose there exists \(x_0 \in X\) such that \(g(x_0) = X_0\). If \(x_0 ∈ X_0\), then \(x_0 ∉ g(x_0) = X_0\) which is impossible. Similarly if \(x_0 ∉ X_0\), then \(x_0 ∈ X_0\) and again we have reached a contradiction. Thus we must conclude that \(X_0 ∉ g(X)\). Thus there are no surjective maps, \(g : X → 2^X\) so that card \((X) ≠ card (2^X)\).

**Proposition B.14.** If card \((X) < card (Y)\) and card \((Y) ≤ card (Z)\), then card \((X) < card (Z)\).

**Proof.** If there exists an injective map, \(f : Z → X\) then composing this with and injective map, \(g : X → Y\) gives an injective map, \(g ∘ f : Z → X\) and there for card \((Z) ≤ card (X)\). But this would imply that card \((X) = card (Z)\).

**Definition B.15.** Let \(A_n : = \{1,2, \ldots, n\}\) for all \(n ∈ N\) and write \(n\) for card \((A_n)\).

**Proposition B.16.** We have card \((A_m) < card (A_n)\) for all \(m < n\). Moreover if \(∅ ∉ A_n\) then card \((X) = card (A_k)\) for some \(k < n\).
Proof. If \( f : A_1 \to A_2 \), then either \( f(1) = 1 \) or \( f(1) = 2 \). In either case \( f \) is injective but not bijective so that \( \text{card}(A_2) < \text{card}(A_1) \). Let \( S_n \) be the statement that \( \text{card}(A_k) < \text{card}(A_l) \) for all \( 1 \leq k < l \leq n \) and for any proper subset \( X \subset A_n \) we have \( \text{card}(X) = \text{card}(A_m) \) for some \( m < n \). Then we have just shown that \( S_2 \) is true. So suppose that \( S_n \) is now true. As \( f : A_k \to A_l \) defined by \( f(m) = m \) for all \( m \in A_k \) is a injection when \( k < l \) we always have \( \text{card}(A_k) \leq \text{card}(A_l) \). Now suppose that \( \text{card}(A_k) = \text{card}(A_{k+1}) \) for some \( k \leq n \). Then there exists a bijection, \( f : A_{n+1} \to A_k \). In this case \( f \) is a proper subset of \( A_k \) and therefore \( \text{card}(A_k) < \text{card}(A_k) \) but on the other hand card \( (f(A_n)) = \text{card}(A_n) \geq \text{card}(A_k) \) which is a contradiction. So no such bijection can exists and we have shown \( \text{card}(A_k) < \text{card}(A_{k+1}) \) for all \( k \leq n \). Finally suppose that \( X \subset A_n \) is proper subset. If \( X \subset A_n \) then \( \text{card}(X) = \text{card}(A_k) \) for some \( k \leq n \) by the induction hypothesis. On the other hand if \( n+1 \in X \), let \( X' := X \setminus \{n+1\} \not\subset A_n \). Therefore by the induction hypothesis \( \text{card}(X') = \text{card}(A_k) \) for some \( k < n \). It is then clear that \( \text{card}(X) = \text{card}(A_{k+1}) \) where \( k+1 < n \), indeed we map \( X := X' \cup \{n+1\} \to A_k \cup \{k+1\} = A_{k+1} \).

Example B.17. \( \text{card}(A_n \setminus \{k\}) = n-1 \) for \( k \in A_n \). Indeed, let \( f : A_{n-1} \to A_n \setminus \{k\} \) be defined by

\[
f(x) = \begin{cases} x & \text{if } x < k \\ x+1 & \text{if } x \geq k \end{cases}
\]

Then \( f \) is the desired bijection. More generally if \( X \subset Y \) and \( \text{card}(X) = m < n = \text{card}(Y) \), then \( \text{card}(Y \setminus X) = n-m \) and if \( X \) and \( Y \) are finite disjoint sets then \( \text{card}(X \cup Y) = \text{card}(X) + \text{card}(Y) \). Similarly, \( \text{card}(X \times Y) = \text{card}(X) \cdot \text{card}(Y) \).

Proposition B.18. If \( f : A_n \to A_n \) is a map, then the following are equivalent,
1. \( f \) is injective,
2. \( f \) is surjective,
3. \( f \) is bijective.

Moreover \( \text{card}(\text{Bijec}(A_n)) = n! \).

Proof. If \( n = 1 \), the only map \( f : A_1 \to A_1 \) is \( f(1) = 1 \). So in this case there is nothing to prove. So now suppose the proposition holds for level \( n \) and \( f : A_{n+1} \to A_{n+1} \) is a given map.

If \( f : A_{n+1} \to A_{n+1} \) is an injective map and \( f(A_{n+1}) \) is a proper subset of \( A_{n+1} \), then \( \text{card}(A_{n+1}) < \text{card}(f(A_{n+1})) = \text{card}(A_{n+1}) \) which is absurd. Thus \( f \) is injective implies \( f \) is surjective.

Conversely suppose that \( f : A_{n+1} \to A_{n+1} \) is surjective. Let \( g : A_{n+1} \to A_{n+1} \) be a right inverse, i.e. \( f \circ g = \text{id} \), which is necessarily injective, see the proof of Proposition B.10. By the pervious paragraph we know that \( g \) is necessarily surjective and therefore \( f = g^{-1} \) is a bijection.

It now only remains to prove \( \text{card}(\text{Bijec}(A_n)) = n! \) which we again do by induction. For \( n = 1 \) the result is clear. So suppose it holds at level \( n \). If \( f : A_{n+1} \to A_{n+1} \) is a bijection. Given \( 1 \leq k \leq n+1 \) let

\[
\text{Bij}_{k}(A_{n+1}) := \{f \in \text{Bij}(A_{n+1}) : f(n+1) = k\}.
\]

For \( f \in \text{Bij}_{k}(A_{n+1}) \), we have \( f : A_n \to A_{n+1} \setminus \{k\} \cong A_n \) is a bijection. Thus \( \text{Bij}_{k}(A_{n+1}) \cong \text{Bij}(A_n) \) and

\[
\text{Bij}(A_{n+1}) = \sum_{k=1}^{n+1} \text{Bij}_{k}(A_{n+1})
\]

we have

\[
\text{card}(\text{Bij}(A_{n+1})) = \sum_{k=1}^{n+1} \text{card}(\text{Bij}_{k}(A_{n+1}))
\]

\[
= \sum_{k=1}^{n+1} \text{card}(\text{Bij}(A_n)) = \sum_{k=1}^{n+1} n!
\]

\( = (n+1)! = (n+1)! \).

Theorem B.19. Suppose that \( X \) is a set. Then \( \text{card}(J_n) \leq \text{card}(X) \) for all \( n \in \mathbb{N} \) if \( f \text{ card}(\mathbb{N}) \leq \text{card}(X) \).

Proof. Since \( \text{card}(J_n) \leq \text{card}(\mathbb{N}) \) for all \( n \in \mathbb{N} \) it suffices to prove \( \text{card}(J_n) \leq \text{card}(X) \) for all \( n \in \mathbb{N} \) implies \( \text{card}(\mathbb{N}) \leq \text{card}(X) \). The intuitive idea is as follows.

Suppose we have constructed \( f_n : J_n \to X \) which is injective. If \( f_n \) were bijective we would have \( \text{card}(J_n) = \text{card}(X) \) and in particular \( \text{card}(J_m) > \text{card}(J_n) = \text{card}(X) \) for all \( m > n \). Thus there exists \( x \in X \setminus f_n(J_n) \) and we then define \( f_{n+1} : J_{n+1} \to X \) so that \( f_{n+1}(n+1) = x \) and \( f_{n+1}|_{J_n} = f_n \). This process continues indefinitely and so we may construct injective maps \( f_n : J_n \to X \) such that \( f_m = f_n|_{J_m} \) for all \( m \leq n \). We then define \( f(m) := f_n(m) \) where \( n \in \mathbb{N} \) is any integer such that \( n \geq m \). In this way we construct a function, \( f : \mathbb{N} \to X \) such that \( f|_{J_n} = f_n \) for all \( n \). This function is easily seen to be injective.

Formalities Version 1. Consider the collection of injective maps \( f : D(f) \subset \mathbb{N} \to X \), where \( D(f) \) is either \( J_n \) for some \( n \in \mathbb{N} \) or is \( \mathbb{N} \). We say \( f \leq g \) if \( D(f) \subset D(g) \) and \( f = g|_{D(f)} \). It is easy to see that every linearly ordered collection of such maps has an upper bound and so by Zorn’s lemma (see
Theorem B.6), there exists a maximal element, \( f \). If \( D(f) \neq \mathbb{N} \) then \( D(f) = J_n \) for some \( n \). By the last paragraph we could extend \( f \) to injective map on \( J_{n+1} \) violating the maximality of \( f \). Thus \( D(f) = \mathbb{N} \) and we have found an injective map from \( \mathbb{N} \) to \( X \).

**Formalities Version 2.** (This argument will avoid the use of Zorn’s Lemma.) By assumption, for each \( n \in \mathbb{N} \) there exists an injective map, \( f_n : J_n \to X \). We now let \( Y := \bigcup_{n \in \mathbb{N}} f_n(J_n) \subset X \). We may construct a surjective map (but not necessarily injective map) \( F : \mathbb{N} \to Y \). From this map we then define \( \psi : Y \to \mathbb{N} \) by \( \psi(y) := \min F^{-1}(\{y\}) \) so that \( \psi : Y \to \mathbb{N} \) is now injective. Suppose for the sake of contradiction that \( \psi(Y) \subset J_N \) for some \( N \in \mathbb{N} \), i.e. \( \psi(Y) \) is a bounded set. Then using our above arguments, we know that \( \text{card} (\psi(Y)) = \text{card} (J_k) \) for some \( k \leq N \). On the other hand, \( f_n : J_n \to Y \) being injective implies \( \text{card} (\psi(Y)) \geq \text{card} (J_n) \) for all \( n \in \mathbb{N} \). As both of these statements can not be correct at the same time we conclude that \( \psi(Y) \) is unbounded. We may now apply Lemma 5.26 in order to see that \( \text{card} (Y) = \text{card} (\psi(Y)) = \text{card} (\mathbb{N}) \). From this it follows that \( \text{card} (\mathbb{N}) \leq \text{card} (X) \).

**Alternate Proof.** By assumption, there exists an injective map, \( f_n : J_n \to X \) for each \( n \in \mathbb{N} \). By replacing \( X \) by \( X_0 := \bigcup_{n \in \mathbb{N}} f_n(J_n) \) we may assume that \( X = \bigcup_{n \in \mathbb{N}} f_n(J_n) \). As \( X \) is the countable union of finite sets it follows that there exists a surjective map, \( f : \mathbb{N} \to X \) by item 2 of Theorem 5.27. Let \( g : X \to \mathbb{N} \) be defined by \( g(x) := \min f^{-1}(\{x\}) \) for all \( x \in X \) and let \( S := g(\mathbb{N}) \). To finish the proof we need only show that \( S \) is unbounded. If \( S \) were bounded, then we would find \( k \in \mathbb{N} \) such that \( J_k \sim X \). However this is impossible since \( \text{card} J_n \leq \text{card} X = \text{card} J_k \) would imply \( n \leq k \).
Appendix: Math 140A Topics

C.1 Summary of Key Facts about Real Numbers

1. The real numbers, \( \mathbb{R} \), is the unique (up to order preserving field isomorphism) ordered field with the least upper bound property or equivalently which is Cauchy complete.
2. Informally the real numbers are the rational numbers with the (irrational) hole filled in.
3. Monotone bounded sequence always converge in \( \mathbb{R} \).
4. A sequence converges in \( \mathbb{R} \) iff it is Cauchy.
5. Cauchy sequences are bounded.
6. \( \mathbb{N} \) is unbounded from above in \( \mathbb{R} \).
7. For all \( \varepsilon > 0 \) in \( \mathbb{R} \) there exists \( n \in \mathbb{N} \) such that \( \frac{1}{n} < \varepsilon \).
8. \( \mathbb{Q} \) and \( \mathbb{R} \setminus \mathbb{Q} \) is dense in \( \mathbb{R} \). In particular, between any two real numbers \( a < b \), there are infinitely many rational and irrational numbers.
9. Decimal numbers map (almost 1-1) into the real numbers by taking the limit of the truncated decimal number.
10. If \( a, b, \varepsilon \in \mathbb{R} \), then
   a) \( a \leq b \) by showing that \( a \leq b + \varepsilon \) for all \( \varepsilon > 0 \).
   b) \( a = b \) by proving \( a \leq b \) and \( b \leq a \) or
   c) \( a = b \) by showing \( |b - a| \leq \varepsilon \) for all \( \varepsilon > 0 \).
11. A number of standard limit theorems hold, see Theorem 3.13.
12. Unlike limits, \( \limsup \) and \( \liminf \) always exist. Moreover we have;
\[
\liminf_{n \to \infty} a_n \leq \limsup_{n \to \infty} a_n
\]
with equality iff \( \lim_{n \to \infty} a_n \) exists in which case
\[
\liminf_{n \to \infty} a_n = \lim_{n \to \infty} a_n = \limsup_{n \to \infty} a_n.
\]
We may allow the values of \( \pm \infty \) in these statements.
13. If \( b_k := \{a_{n_k}\}_{k=1}^{\infty} \) is a convergent subsequence of \( \{a_n\} \), then
\[
\liminf_{n \to \infty} a_n \leq \lim_{k \to \infty} b_k \leq \limsup_{n \to \infty} a_n
\]
and we may choose \( \{b_k\} \) so that \( \lim_{k \to \infty} b_k = \limsup_{n \to \infty} a_n \) or \( \lim_{k \to \infty} b_k = \liminf_{n \to \infty} a_n \).
14. Bounded sequences of real numbers always have convergence subsequences.
15. If \( S \subset \mathbb{R} \) and \( A := \sup(S) \), then there exists \( \{a_n\}_{n=1}^{\infty} \subset S \) such that \( a_n \leq a_{n+1} \) for all \( n \) and \( \lim_{n \to \infty} a_n = \sup(S) \).
16. If \( S \subset \mathbb{R} \) and \( A := \inf(S) \), then there exists \( \{a_n\}_{n=1}^{\infty} \subset S \) such that \( a_{n+1} \leq a_n \) for all \( n \) and \( \lim_{n \to \infty} a_n = \inf(S) \).

C.2 Test 2 Review Topics:

1. Understand the basic properties of complex numbers.
2. Coutability. Key facts are that countable union of countable sets is countable and the finite product of countable sets is countable.
3. Definitions of metric and normed spaces and their basic properties which in the end of the day typically follow from the triangle inequality.
4. You should know that metrics and norms are continuous functions that satisfy,
\[
||x|| - ||y|| \leq ||x - y|| \quad \text{and}
\]
\[
|d(x, y) - d(x', y')| \leq d(x, x') + d(y, y').
\]
5. Be aware of different norms, \( ||\cdot||_1 \), \( ||\cdot||_2 \).
6. Understand the notion of limits of sequences, Cauchy sequences, completeness, limits and continuity of functions.
7. Know what is meant by pointwise and uniform convergence. You should be able to compute pointwise limits and know how to test if the limit is uniform or not. A key theorem is the uniform limit of continuous functions is still continuous.

C.3 After Test 2 Review Topics:

Let \( (X, \|\cdot\|) \) be a Banach space.

1. Know \( \sum_{n=1}^{\infty} x_n = \lim_{N \to \infty} \sum_{n=1}^{N} x_n \) if the limit exists.
2. Know \( \sum_{n=1}^{\infty} x_n \) converges absolutely iff \( \sum_{n=1}^{\infty} ||x_n|| < \infty \) and absolute convergence implies convergence in a Banach space.
3. Telescoping series and geometric series.
5. Absolute convergence tests: 1) integral test, 2) root test, 3) ratio test often combined with the 4) comparison test.
7. $n^{th}$ – term test for divergence.
8. Cauchy criteria for convergence and the fact that tails of convergent series tend to zero. i.e. tails of convergent series tend to zero.
10. Uniform convergence of sums and the Weierstrass $M$ – test
11. Power series including radius of convergence notion.
12. The exponential function and its relatives, sin, sinh, cos, cosh.
References


Index

Cauchy, 5, 18
Cauchy sequence in a metric space, 16
Complete

Metric space, 46
Function continuous, 49
continuous at a point, 49