Undergraduate Analysis Tools

November 9, 2012  File:unal.tex
## Part II: Banach and Metric Spaces

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1.1</td>
<td>Review of Vector Spaces and Subspaces</td>
<td>43</td>
</tr>
<tr>
<td>6.1.2</td>
<td>Normed Spaces</td>
<td>43</td>
</tr>
<tr>
<td>6.2</td>
<td>Sequences in Metric Spaces</td>
<td>46</td>
</tr>
<tr>
<td>6.3</td>
<td>General Limits and Continuity in Metric Spaces</td>
<td>48</td>
</tr>
</tbody>
</table>

## Part III: Appendices

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Appendix: Notation and Logic</td>
<td>69</td>
</tr>
<tr>
<td>B</td>
<td>Appendix: More Set Theoretic Properties (highly optional)</td>
<td>71</td>
</tr>
<tr>
<td>B.1</td>
<td>Appendix: Zorn’s Lemma and the Hausdorff Maximal Principle (optional)</td>
<td>71</td>
</tr>
<tr>
<td>B.2</td>
<td>Cardinality</td>
<td>73</td>
</tr>
<tr>
<td>B.3</td>
<td>Formlities of Counting</td>
<td>74</td>
</tr>
</tbody>
</table>

**References**

<table>
<thead>
<tr>
<th>Page</th>
<th>Index</th>
</tr>
</thead>
</table>
Part I

Numbers
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 \( S = \mathbb{N}. \) \( \blacksquare \)

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. \( S = \emptyset. \) \( \blacksquare \)

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 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
\[
|q| = \begin{cases} 
q & \text{if } q \geq 0 \\
-q & \text{if } q < 0.
\end{cases}
\]
Recall that, for all \( a, b \in \mathbb{Q}, \)
\[
|a + b| \leq |a| + |b|, \quad |ab| = |a||b|, \quad \text{and } \left| \frac{1}{a} \right| = \frac{1}{|a|} \quad \text{when } a \neq 0.
\]
It is also often useful to keep in mind that 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
\[
||b| - |a|| \leq |b - a|. \tag{1.1}
\]

Proof. Since both sides of Eq. \( \text{(1.1)} \) are symmetric in \( a \) and \( b, \) we may assume that \( |b| \geq |a| \) so that \( ||b| - |a|| = |b| - |a| \). Since
\[
|b| = |b - a + a| \leq |b - a| + |a|,
\]
it follows that
\[
||b| - |a|| = |b - a| \leq |b - a|.
\]
The proof of the previous lemma illustrates one of the key techniques of adding 0 to an expression. In this case we added 0 in the form of \(-a + a\) to \(b\). The next remark records a couple of other very important "tricks" in this subject. Taking to heart the following remarks will greatly aid the student in real analysis.

**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. prove \(a = b\) 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 if \(a > b\) 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 \(\varepsilon > 0\).  

---

1 Absolute values will be discussed in more generality in Section 2.2 below.
2 We will see that the real numbers have these same properties as well.

## 1.1 Limits in \(\mathbb{Q}\)

In this course we will often use the abbreviations, i.o. and a.a. which stand for **infinitely often** and **almost always** respectively. For example, \(a_n \leq b_n\) a.a. \(n\) means there exists an \(N \in \mathbb{N}\) such that \(a_n \leq b_n\) for all \(n \geq N\) while \(a_n \leq b_n\) i.o. \(n\) means for all \(N \in \mathbb{N}\) there exists a \(n \geq N\) such that \(a_n \leq b_n\). So for example, \(1/n \leq 1/100\) for a.a. \(n\) while and \((-1)^n \geq 0\) i.o. By the way, it should be clear that if something happens for a.a. \(n\) then it also happens i.o. \(n\).

**Definition 1.9.** A sequence \(\{a_n\}_{n=1}^{\infty} \subset \mathbb{Q}\) converges to \(0 \in \mathbb{Q}\) if for all \(\varepsilon > 0\) in \(\mathbb{Q}\) there exists \(N \in \mathbb{N}\) such that \(|a_n| \leq \varepsilon\) for all \(n \geq N\). Alternatively put, for all \(\varepsilon > 0\) we have \(|a_n| \leq \varepsilon\) for a.a. \(n\). This may also be stated as for all \(M \in \mathbb{N}\), \(|a_n| \leq \frac{1}{M}\) for a.a. \(n\).

**Definition 1.10.** A sequence \(\{a_n\}_{n=1}^{\infty} \subset \mathbb{Q}\) converges to \(a \in \mathbb{Q}\) if \(|a - a_n| \to 0\) as \(n \to \infty\), i.e. if for all \(N \in \mathbb{N}\), \(|a - a_n| \leq \frac{1}{N}\) for a.a. \(n\). As usual if \(\{a_n\}_{n=1}^{\infty}\) converges to \(a\) we will write \(a_n \to a\) as \(n \to \infty\) or \(a = \lim_{n \to \infty} a_n\).

**Proposition 1.11.** If \(\{a_n\}_{n=1}^{\infty} \subset \mathbb{Q}\) converges to \(a \in \mathbb{Q}\), then \(\lim_{n \to \infty} |a_n| = |a|\).

**Proof.** From Lemma 1.6 we have,

\[
|a| - |a_n| \leq |a - a_n|
\]

Thus if \(\varepsilon > 0\) is given, by definition of \(a_n \to a\) there exists \(N \in \mathbb{N}\) such that \(|a - a_n| < \varepsilon\) for all \(n \geq N\). From the previously displayed equation, it follows that \(|a - |a_n|| < \varepsilon\) for all \(n \geq N\) and hence we may conclude that \(\lim_{n \to \infty} |a_n|\) exists and is equal to \(|a|\).

**Lemma 1.12 (Convergent sequences are bounded).** If \(\{a_n\}_{n=1}^{\infty} \subset \mathbb{Q}\) converges to \(a \in \mathbb{Q}\), then there exists \(M \in \mathbb{Q}\) such that \(|a_n| \leq M\) for all \(n \in \mathbb{N}\).

**Proof.** Taking \(\varepsilon = 1\) in the definition of \(a = \lim_{n \to \infty} a_n\) implies there exists \(N \in \mathbb{N}\) such that \(|a_n - a| \leq 1\) for all \(n \geq N\). Therefore,

\[
|a_n| = |a_n - a + a| \leq |a_n - a| + |a| \leq 1 + |a| \text{ for } n \geq N.
\]

We may now take \(M := \max\left\{|a_n|_{n=1}^{N}\} \cup \{1 + |a|\}\right\} \).  

**Theorem 1.13.** If \(\{a_n\}_{n=1}^{\infty} \subset \mathbb{Q}\) converges to \(a \in \mathbb{Q} \setminus \{0\}\), then

\[
\lim_{n \to \infty} \frac{1}{a_n} = \frac{1}{a}.
\]

(1.2)

It is possible that \(a_n = 0\) for small \(n\) so that \(\frac{1}{a_n}\) is not defined but for large \(n\) this can not happen and therefore it makes sense to talk about the limit which only depends on the tail of the sequences.
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\footnote{The idea is very simple here. If $a_n$ is near $a$ and $a \neq 0$ then $a_n$ must stay away from zero. You should draw the picture to go along with the proof.} 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,

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

As $\varepsilon > 0$ is arbitrary it follows that $2\varepsilon/|a|^2 > 0$ 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 routing, 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,

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

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

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

for all $n \geq N(\delta)$.\footnote{The idea is very simple here. If $a_n$ is near $a$ and $a \neq 0$ then $a_n$ must stay away from zero. You should draw the picture to go along with the proof.}

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

\vspace{10pt}

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 all $m, n \in \mathbb{N}$, 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

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

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

\[
\sum_{k=n}^{m} \alpha^k = \alpha^n \frac{1 - \alpha^{m-n+1}}{1 - \alpha} \leq \frac{\alpha^n}{1 - \alpha}.
\] (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

\[
S = \sum_{k=n}^{m} \alpha^k = \frac{\alpha^{m+1} - \alpha^n}{\alpha - 1} \text{ if } \alpha \neq 1.
\] (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,

\[
0 \leq S_n - S_m = \sum_{k=m+1}^{n} \frac{1}{k!} = \sum_{j=1}^{n-m} \frac{1}{(m+j)!} \\
\leq \frac{1}{m!} \left[ \frac{1}{m+1} + \left( \frac{1}{m+1} \right)^2 + \cdots + \left( \frac{1}{m+1} \right)^{n-m} \right] \\
\leq \frac{1}{m! \cdot m^{n-m}} \cdot \frac{1}{m+1} = \frac{1}{m \cdot m!}.
\] (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,

\[
\frac{1}{(m+1)!} \leq S_n - S_m \leq \frac{1}{m \cdot m!} \text{ for all } n > m. \quad (1.7)
\]

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 \( \mathbb{Q} \)!

Plotting the partial sums \( \sum_{k=0}^{n} \frac{1}{k!} \) (black curve) and \( (1 + \frac{1}{n})^n \) (red curve) which are both converging to “e.”

Example 1.18 (Square roots need not exist). The square root, \( \sqrt{2} \), of 2 does not exist in \( \mathbb{Q} \). Indeed, if \( \sqrt{2} = \frac{n}{m} \) 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 \( \{q_n\}_{n=1}^{\infty} \subset \mathbb{Q} \) which is not convergent in \( \mathbb{Q} \).

1. Recall that there is no element \( q \in \mathbb{Q} \) such that \( q^2 = 2 \). To each \( n \in \mathbb{N} \) let 
   \[
   m_n \in \mathbb{N} \text{ be chosen so that } \frac{m_n^2}{n^2} < 2 < \frac{(m_n + 1)^2}{n^2} \tag{1.8}
   \]
   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 \( S \subset \mathbb{N}_0 \) 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{1} 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 \( x \in \mathbb{N}_0 \cap S \) 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

\[ p_{s(x)}(s(y)) := s(p_x(s(y))) = s(p_x(y)) = s(p_{s(x)}(y)) \]

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 Natural, integer, and rational Numbers

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,

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

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

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. $s(x) = x + 1 = 1 + x$;
3. $x + y = y + x$;
4. $(x + y) + z = x + (y + z)$.
5. The induction hypothesis may now be written as; if $S \subset \mathbb{N}_0$ 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.

Definition 1.24. Given $x, y \in \mathbb{N}_0$, we say $x < y$ if $y = x + n$ for some $n \in \mathbb{N}$ and $x \leq y$ if $y = x + n$ 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) $y < x$, 3) $x = y$.

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 $y \in \mathbb{N}_0$ 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. $y = 0 < x$. 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 1.25 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 $\mathbb{N}_0$.
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 $S := \{x \in \mathbb{N}_0 : M(x, \cdot) = M(\cdot, x)\}$. 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

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

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.

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. $x \cdot 0 = 0 = 0 \cdot x$,
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.

Remark 1.30 (Base 10 counting). The typical method of counting is to use base 10 enumeration of $\mathbb{N}_0$. The rules are;

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

0 := 0,
1 := 1, 2 := 1 + 1, 3 := 2 + 1, 4 := 3 + 1, 5 := 4 + 1,
6 := 5 + 1, 7 := 6 + 1, 8 := 7 + 1, 9 := 8 + 1, and 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: (subtraction and division are defined implicitly in terms of the inverse operations of addition and multiplication;)

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. $(-x) \cdot y = -(xy)$.

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. $0 \cdot x = 0 = x + 1 = x = (0 + 1) \cdot x = 1 \cdot x = x$.

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. $(-x) y + xy = (-x + x) \cdot y = 0 \cdot y = 0 \implies (-x) y = -(xy)$.

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 \( 0 + 1 = 1 + 0 = 0 \) and \( 0 \cdot 1 = 1 \cdot 0 = 0 \cdot 0 = 0 \) and \( 1 \cdot 1 = 1 \). 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. \( \left( \frac{m}{n} \right)^{-1} = \frac{n}{m} \) if \( m \neq 0 \) and \(-m/n = -m/n \).

3. \( F = \mathbb{Q}(t) \) where

\[
\mathbb{Q}(t) = \left\{ \frac{p(t)}{q(t)} : p(t) \text{ and } q(t) \text{ are polynomials over } \mathbb{Q} \right\}.
\]

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 \( 0 \cdot 1_F = 0_F \),

\[
n1_F := \underbrace{1_F + \cdots + 1_F}_{n \text{ times}} \text{ if } n \in \mathbb{N} \text{ and } \quad (-n)1_F := -(n1_F) \text{ if } n \in \mathbb{N}.
\]

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

\[
\varphi(n) = \underbrace{1_F + \cdots + 1_F}_{n \text{ times}}.
\]

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

\[
\varphi((m + 1) + n) = \varphi(m + n + 1) = \varphi(m) + \varphi(n + 1)
\]

\[
= \varphi(m) + \varphi(n) + 1_F
\]

\[
= \varphi(m) + 1_F + \varphi(n) = \varphi(m + 1) + \varphi(n)
\]

which shows \( m + 1 \in S \). Therefore by induction, \( S = \mathbb{N}_0 \) and \( \varphi(m + n) = \varphi(m) + \varphi(n) \) for all \( m, n \in \mathbb{N}_0 \).

If \( m \in \mathbb{N}_0 \), 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 \( F \) is an **ordered field** if there exists, \( P \subseteq F \), called the positive elements, such that

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

2. \( P + P \subseteq P \) and \( P \cdot P \subseteq P \).

**Lemma 2.8.** Let \((F, P)\) be an ordered field, then:

1. For all \( x \in F \setminus \{0\} \), \( x^2 \in P \). 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 \( x^2 \in P \cdot P \subseteq P \) while if \( x \in -P \) then \( -x \in P \) and \( x^2 = (-x)^2 \in P \). For item 3. we have \( x \cdot x^{-1} = 1 \) \( \square \)

**Example 2.9.** The field \( F = \{0, 1\} \) 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 \( F = \mathbb{Q} \) 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 \( \mathbb{N} \subseteq P \). 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 \( F = \mathbb{Q}(t) \) and

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

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, \( \frac{t^2 - 25t + 7}{t^4 - 10t^3 + 25t^2 - 10t + 2} \in P \) while \( \frac{t^2 + 25t + 7}{t^4 - 10t^3 + 25t^2 - 10t + 2} \notin -P \).

Notice that \( t > n \) for all \( n \in \mathbb{N} \) and \( \frac{t}{n} \leq 1 \) for all \( n \in \mathbb{N} \). This 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

\[
a \land b := \min(a, b) = \begin{cases} a & \text{if } a \leq b \\ b & \text{if } b \leq a \end{cases}
\]

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 \( \{a_i\}_{i=1}^n \subseteq F \) we let

\[
a_1 \land \cdots \land a_n := \min(a_1, \ldots, a_n) \quad \text{and} \quad a_1 \lor \cdots \lor a_n := \max(a_1, \ldots, a_n)
\]

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

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

\[
\varphi(x + y) = \varphi(x) + \varphi(y) \quad \text{and} \quad \varphi(xy) = \varphi(x) \varphi(y).
\]

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

\[
\varphi \left( \frac{m}{n} \right) = \frac{m}{n} \cdot 1_F := m1_F \cdot (n1_F)^{-1} \tag{2.1}
\]

where \( n1_F := 1_F + \cdots + 1_F \) and \( (-n)1_F := -(n1_F) \) for all \( n \in \mathbb{N} \) and \( 0 \cdot 1_F = 0_F \). 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 \( F \).

[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, \( \varphi : \mathbb{Z} \rightarrow F \), given by \( \varphi(m) = m \cdot 1_F \). So for \( m \in \mathbb{Z} \) and \( n \in \mathbb{N} \) we must have

\[
\varphi \left( \frac{m}{n} \right) \cdot n1_F = \varphi \left( \frac{m}{n} \right) \cdot \varphi(n) = \varphi \left( \frac{m}{n} \cdot n \right) = \varphi(m) = m1_F
\]

which forces us to define \( \varphi \) as in Eq. (2.1). Notice that is easy to verify by induction that \( n1_F = \varphi(n) \in P \) for all \( n \in \mathbb{N} \) and in particular \( n1_F \neq 0 \) for \( n \in \mathbb{N} \). In particular if \( x = m/n > 0 \) then \( \varphi \left( \frac{m}{n} \right) = m1_F \cdot (n1_F)^{-1} \in P \) 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

\[
(km)1_F \cdot (kn1_F)^{-1} = m1_F \cdot (n1_F)^{-1}.
\]

By cross multiplying, this will happen iff

\[
(km)1_F \cdot (n1_F) = ((kn)1_F) \cdot m1_F
\]
which is the case as \( \varphi : \mathbb{Z} \to \mathbb{F} \) is a ring homomorphism.

**Homomorphism property.** We have

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

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

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

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

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. \( m/n = 0 \).

---

**End of Lecture 3, 10/3/2012**

**Notation 2.15** If \( (\mathbb{F}, P) \) is an ordered field we write \( x > y \) iff \( x - y \in P \). We also write \( x \geq y \) iff \( x > y \) or \( x = y \).

Notice that if \( x, y \in \mathbb{F} \) then either \( x - y = 0 \) (i.e. \( x = y \)), or \( x - y \in P \) (i.e. \( x > y \)), or \( x - y \in -P \) (i.e. \( y - x \in P \) and \( y > x \)). Also in this notation we have \( P = \{ x \in \mathbb{F} : x > 0 \} \), \( -P = \{ x \in \mathbb{F} : x < 0 \ \text{ i.e. } 0 > x \} \).

**Lemma 2.16.** Suppose that \( x < y \) and \( y < z \) and \( a > 0 \). Then \( x < z \) and \( ax < ay \).

**Proof.** By assumption \( y - x \in P \) and \( z - y \in P \), therefore \( z - x = (y - x) + (z - y) \in P \), i.e. \( z > x \). Moreover, \( a \in P \) and \( (y - x) \in P \) implies

\[
P \ni a(y - x) = ay - ax.
\]

That is \( ay > ax \).

---

**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. \( x, y - x \in P \), then \( x^2 < y^2 \).

**Proof.** By Lemma 2.16 we know \( x \cdot x < x \cdot y \) 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,

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

from which it follows that \( \sqrt{y} - \sqrt{x} \in P \) if \( (y - x) \in P \). 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

1. \(|x| \geq 0\)
2. \(|xy| = |x| \cdot |y|\)
3. \(|x + y| \leq |x| + |y|\).

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

Definition 2.24. Let \((\mathbb{F}, P)\) be an ordered field and \(S\) be a subset of \(\mathbb{F}\).

1. We say that \(S \subset \mathbb{F}\) 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 \(S\).
2. 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 \(S\).

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
   \[\sup (S) = a_1 \lor \cdots \lor a_n := \max (a_1, \ldots, a_n)\]
   \[\inf (S) = a_1 \land \cdots \land a_n := \min (a_1, \ldots, a_n)\].
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 \(\inf (S) = \frac{1}{2}\).
4. Let
   \[S = \{1, 1.4, 1.41, 1.411, 1.4112, 1.41121, 1.411213, 1.4112135, \ldots \}\]
   where I am getting these digits from the decimal expansion of \(\sqrt{2}\);
   \[\sqrt{2} \approx 1.41421356237095504880168872409698078569671875376948073176667973799\].

In this case \(S\) is bounded above by 2, or 1.42, or 1.415, etc. Nevertheless \(\sqrt{2} = \sup (S)\) does not exist 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 \(\frac{1}{m} t\) 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:

1. \(\alpha\) is a proper subset of \(\mathbb{Q}\), i.e. \(\alpha \neq \emptyset\) and \(\alpha \neq \mathbb{Q}\),
2. if \(p \in \alpha\) and \(q \in \mathbb{Q}\) and \(q < p\), then \(q \in \alpha\),
3. if \(p \in \alpha\), then there exists \(r \in \alpha\) with \(r > p\).

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 \(\lim_{n \to \infty} S_n := \sum_{k=0}^n \frac{1}{k} \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

\[m - S_n = |m - S_n| \geq \varepsilon\ i.o. n.\]

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.

We then define \( \alpha \) and \( \beta \) to be the equivalence class of the constant sequence \( \{ a_n \} \) and \( \{ b_n \} \) respectively. The set of real numbers will be denoted by \( \{ \alpha, \beta \} \).

Definition 3.1. An order preserving field isomorphism between two ordered fields, \( (F_1, P_1) \) and \( (F_2, P_2) \), is a bijection, \( f : F_1 \to F_2 \) such that \( f(0) = 0 \), \( f(1) = 1 \), \( f(P_1) = P_2 \), and

\[
f(x + y) = f(x) + f(y) \quad \text{and} \quad f(xy) = f(x)f(y) \quad \text{for all} \quad x, y \in F_1.
\]

Definition 3.2. An ordered field \( (F, P) \) is has the least upper bound property (or is complete) if every non-empty subset, \( S \subseteq F \), which is bounded from above possesses a least upper bound in \( F \). [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 \( \{ a_n \}_{n=1}^\infty \sim \{ b_n \}_{n=1}^\infty \) if

\[
\lim_{n \to \infty} |a_n - b_n| = 0.
\]

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 \( \mathbb{P} := \{ \alpha \in \mathbb{R} : \alpha > 0 \} \), where we say \( \alpha > 0 \) iff there exists an \( N \in \mathbb{N} \) such that \( a_n > \frac{1}{N} \) for a.a. \( n \).
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 \( \mathbb{R} \), see [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 3.6 at the end of this chapter. 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 \( t \) 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 \( \mathbb{N} \).

\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. \( n \).}
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 $a = b + \frac{1}{n}$ 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 $a \leq b$.

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

**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 $s$ is a lower bound for $S$. Moreover if $\varepsilon > 0$ is given there exists $s_\varepsilon \in S$ such that $-s_\varepsilon \geq -m - \varepsilon$, i.e. $s_\varepsilon \leq m + \varepsilon$. This shows that any lower bound, $k$ 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 $\mathbb{R}$ based on Cauchy sequences of rational numbers.

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

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

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

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

\[ |M - x_n| \leq \varepsilon \text{ for all } n \geq N_\varepsilon. \]

- End of Lecture 5, 10/8/2012

**Theorem 3.15.** If $\{x_n\} \subset \mathbb{R}$ is bounded from above and $x_n$ is non-decreasing, then $\lim_{n \to \infty} x_n = \sup_n x_n$. Similarly if $\{x_n\} \subset \mathbb{R}$ is bounded from below and $x_n$ is non-increasing, then $\lim_{n \to \infty} x_n = \inf_n x_n$. 

**Proof.** Let $M := \sup_n x_n$, 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.
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 \( \alpha \subset \mathbb{Q} \) 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 \( M \leq m \).

2. 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 \( \alpha_m \).

**Proof 2.** [Here is a slight rewriting of the above argument.] Choose \( y_n \in \alpha_m \) such that \( y_n \uparrow M \) as \( m \to \infty \). Choose \( n \in \mathbb{N} \) so that \( m - M > \frac{1}{n} \). Then \( y_m + \frac{1}{n} \uparrow M + \frac{1}{n} < m \) 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 \( \alpha \subset \mathbb{Q} \) 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 \( \alpha_m := \{ y \in \mathbb{Q} : y < m \} \).

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 (\( n^{th} \) 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 \( y > x \).

---

\(^3\) 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 $A := \{ t \in \mathbb{R}^+ : t^n \leq x \}$. If $t = \frac{1}{1+n} \in (0, 1)$, then $t^n \leq t \leq x$ so that $t \in A$ and $A \neq \emptyset$. If $t = 1 + x$, then $t^n = (1 + x)^n \geq 1 + nx > x$ and therefore $A$ is bounded from above. Hence we may define $y := \sup A$. We will now show that $y^n = x$.

By Theorem 3.14 there exists $t_k \in A$ such that $t_k \uparrow y$ as $k \to \infty$. By definition of $A$, $t_k^n \leq x$ 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,

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

Hence for sufficiently large $m$ we will have $(y + \frac{1}{m})^n < x$. But this shows that $y + \frac{1}{m} \in A$ which violates $y$ being an upper bound for $A$. 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 $\mathbb{R} := \mathbb{R} \cup \{ \pm \infty \}$, i.e. it is $\mathbb{R}$ with two new points called $\infty$ and $-\infty$. We use the following conventions, $\pm \infty \cdot a = \pm \infty$ 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:

$\infty - \infty$, $-\infty + \infty$, $\infty/\infty$, $0 \cdot \infty$, and $\infty \cdot 0$.

A sequence $a_n \in \mathbb{R}$ 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 $A \subset \mathbb{R}$, let $\sup A$ and $\inf A$ denote the least upper bound and greatest lower bound of $A$ respectively. The convention being that $\sup A = \infty$ if $\infty \in A$ or $A$ is not bounded from above and $\inf A = -\infty$ if $-\infty \in A$ or $A$ 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 $A$ is a subset of $\mathbb{R}$ such that $A = \bigcup_{\alpha \in I} A_{\alpha}$ where $A_{\alpha} \subset A$ and $I$ is some index set. Then

$$\sup A = \sup_{\alpha \in I} A_{\alpha}.$$

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 $M := \sup A$ and $M_\alpha := \sup A_{\alpha}$ for all $\alpha \in I$. As $A_{\alpha} \subset A$ we have $M_\alpha \leq M$ for all $\alpha \in I$ and therefore $\sup_{\alpha \in I} M_\alpha \leq M$. Conversely, if $\lambda \in A$, then $\lambda \in M_\alpha$ for some $\alpha \in I$ and therefore $\lambda \leq M_\alpha$. From this it follows that $\lambda \leq \sup_{\alpha \in I} M_\alpha$ and as $\lambda \in A$ is arbitrary we may conclude that $M = \sup A \leq \sup_{\alpha \in I} M_\alpha$.

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

$$\sup_{x \in X \ y \in Y} S(x, y) = \sup_{y \in Y} \sup_{x \in X} S(x, y).$$

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

$$\sup_{m \in \mathbb{N}} \sup_{n \in \mathbb{N}} S_{m,n} = \sup_{(m,n)} S_{m,n} = \sup_{m,n} S_{m,n}.$$

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

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

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 $\mathbb{R}$, 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 $\mathbb{R}$ are those which oscillate too much.
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. \( \{a_n b_n\}_{n=1}^{\infty} \) is convergent and
\[
\lim_{n \to \infty} (a_n b_n) = \lim_{n \to \infty} a_n \cdot \lim_{n \to \infty} b_n
\]
(provided the right hand side is not of the form \( \infty \cdot 0 \) or \( 0 \cdot (\pm \infty) \)).

Before going to the proof consider the simple example where \( a_n = n \) and \( b_n = -\alpha n + c \) with \( \alpha > 0 \) and \( c \in \mathbb{R} \). Then
\[
\lim (a_n + b_n) = \begin{cases} 
\infty & \text{if } \alpha < 1 \\
c & \text{if } \alpha = 1 \\
-\infty & \text{if } \alpha > 1 
\end{cases}
\]
while
\[
\lim_{n \to \infty} a_n + \lim_{n \to \infty} b_n = "\infty - \infty".
\]

This shows that the requirement that the right side of Eq. \((3.1)\) is not of form \( \infty - \infty \) is necessary in Lemma \((3.2)\). Similarly by considering the examples \( a_n = n \) and \( b_n = n^{-\alpha} \) with \( \alpha > 0 \) shows the necessity for assuming right hand side of Eq. \((3.2)\) 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,
\[
|a + b - (a_n + b_n)| = |a - a_n + b - b_n| \leq |a - a_n| + |b - b_n| \leq 2\varepsilon
\]
6 This example shows that if you formally arrive at an expression like \( \infty - \infty \), then you should work harder to decide what it really means!

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

**Exercise 3.2.** Show \( \lim_{n \to \infty} a^n = \infty \) and \( \lim_{n \to \infty} \frac{1}{a^n} = 0 \) 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, \( a^n \) 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
\[
\liminf_{n \to \infty} x_n = \liminf_{n \to \infty} \{x_k : k \geq n\} \text{ and } \limsup_{n \to \infty} x_n = \limsup_{n \to \infty} \{x_k : k \geq n\}.
\]

We will also write \( \liminf \) for \( \lim \inf \) and \( \limsup \) for \( \lim \sup \).

**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
\[
\liminf_{n \to \infty} x_n = \sup \inf \{x_k : k \geq n\} \text{ and } \limsup_{n \to \infty} x_n = \inf \sup \{x_k : k \geq n\}.
\]

Owing to the following exercise, one may reduce properties of the \( \liminf \) to those of the \( \limsup \).

**Exercise 3.5.** Show \( \liminf_{n \to \infty} (-a_n) = - \limsup_{n \to \infty} a_n \).
Exercise 3.6. Let \( \{a_n\}_{n=1}^{\infty} \) be the sequence given by,
\[
(-1, 2, 3, -1, 2, 3, -1, 2, 3, \ldots).
\]

Find \( \limsup_{n \to \infty} a_n \) and \( \liminf_{n \to \infty} a_n \).

Exercise 3.7. If \( \{a_n\}_{n=1}^{\infty} \) and \( \{b_n\}_{n=1}^{\infty} \) are two sequences such that \( a_n \leq b_n \) for a.a. \( n \), then
\[
\limsup_{n \to \infty} a_n \leq \limsup_{n \to \infty} b_n \quad \text{and} \quad \liminf_{n \to \infty} a_n \leq \liminf_{n \to \infty} b_n.
\] (3.5)

The following proposition contains some basic properties of \( \liminf \) and \( \limsup \).

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 \( \infty - \infty \). (3.6)
4. If \( a_n \geq 0 \) and \( b_n \geq 0 \) for all \( n \in \mathbb{N} \), then
\[
\limsup_{n \to \infty} (a_n b_n) \leq \limsup_{n \to \infty} a_n \cdot \limsup_{n \to \infty} b_n,
\] (3.7)
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
\[
\inf\{a_k : k \geq n\} \leq \sup\{a_k : k \geq n\} \forall n,
\]
and therefore by the Sandwich theorem, \( \liminf_{n \to \infty} a_n \leq \limsup_{n \to \infty} a_n \).

2. \( \Leftarrow \) Let \( A \) be defined as above. 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 \( \lim_{k \to \infty} a_k = -\infty \).

( \( \Rightarrow \)) Conversely, suppose that \( \limsup_{n \to \infty} a_n = A \in \mathbb{R} \). 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.
\[
A - \varepsilon \leq a_n \leq A + \varepsilon \quad \text{for all} \quad n \geq N(\varepsilon).
\]
From this we learn that
\[
A - \varepsilon \leq \inf_{k \geq n} a_k \leq \sup_{k \geq n} a_k \leq A + \varepsilon
\]
and so passing to the limit as \( n \to \infty \) implies
\[
A - \varepsilon \leq \lim_{n \to \infty} \inf_{n \to \infty} a_n \leq \lim_{n \to \infty} \sup_{n \to \infty} a_n \leq A + \varepsilon.
\]
Since \( \varepsilon > 0 \) is arbitrary, it follows that
\[
A \leq \liminf_{n \to \infty} a_n \leq \limsup_{n \to \infty} a_n \leq A,
\]
i.e. that \( A = \liminf_{n \to \infty} a_n = \limsup_{n \to \infty} a_n \).

If \( A = \infty \), then for all \( M > 0 \) there exists \( N = N(M) \) such that \( a_n \geq M \) for all \( n \geq N \). This shows 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
\[
\limsup_{n \to \infty} (a_n + b_n) \leq \limsup_{n \to \infty} a_n + \limsup_{n \to \infty} b_n
\] (3.8)
provided the right side of Eq. (3.8) is well defined, i.e. no \( -\infty + \infty = \infty \) or \( -\infty - \infty = -\infty \), etc.

Exercise 3.9. Suppose that \( a_n \geq 0 \) and \( b_n \geq 0 \) for all \( n \in \mathbb{N} \). Show
\[
\limsup_{n \to \infty} (a_n b_n) \leq \limsup_{n \to \infty} a_n \cdot \limsup_{n \to \infty} b_n.
\] (3.9)
provided the right hand side of (3.9) is not of the form \( 0 \cdot \infty \) or \( \infty \cdot 0 \).

Definition 3.29. A sequence, \( \{a_n\}_{n=1}^{\infty} \), of positive numbers is said to have sub-geometric growth iff for all \( \alpha > 1 \) there exists \( c(\alpha) < \infty \) such that
\[
a_n \leq c(\alpha) a^n \quad \text{for} \quad n \in \mathbb{N}.
\]
Lemma 3.30. Suppose \( \{a_n\}_{n=1}^\infty \) is a sequence of positive numbers having sub-
geometric growth such that \( a_n \geq 1 \) for a.a. \( n \). [For example, by Exercise 3.3 the hypothesis is satisfied if \( 1 \leq a_n \leq n^p \) for a.a. \( n \) for some \( p \in \mathbb{N} \).] Then \( \lim_{n \to \infty} (a_n)^{1/n} = 1 \) and \( \lim_{n \to \infty} \left( \frac{1}{a_n} \right)^{1/n} = 1 \).

Proof. Let \( \alpha > 1 \) be given and choose \( \beta \in (1, \alpha) \). By assumption there exists \( c(\beta) < \infty \) such that \( a_n \leq c(\beta) \beta^n \) for all \( n \in \mathbb{N} \). From this we conclude,

\[
a_n \leq c(\beta) \beta^n = c(\beta) \left( \frac{\beta}{\alpha} \right)^n \alpha^n.
\]

Since \( \lim_{n \to \infty} c(\beta) \left( \frac{\beta}{\alpha} \right)^n = 0 \) we may now conclude that \( 1 \leq a_n \leq \alpha^n \) for a.a. \( n \) and this implies,

\[
1 = 1^{1/n} \leq (a_n)^{1/n} \leq (\alpha^n)^{1/n} = \alpha \text{ for a.a. } n.
\]

and hence

\[
1 \leq \liminf_{n \to \infty} (a_n)^{1/n} \leq \limsup_{n \to \infty} (a_n)^{1/n} \leq \alpha.
\]

As \( \alpha > 1 \) is arbitrary, it follows that

\[
1 = \liminf_{n \to \infty} (a_n)^{1/n} = \limsup_{n \to \infty} (a_n)^{1/n},
\]

i.e. \( \lim_{n \to \infty} (a_n)^{1/n} = 1 \). Lastly,

\[
\lim_{n \to \infty} \left( \frac{1}{a_n} \right)^{1/n} = \lim_{n \to \infty} \frac{1}{(a_n)^{1/n}} = \frac{1}{\lim_{n \to \infty} (a_n)^{1/n}} = 1.
\]

\[ \blacksquare \]

End of Lecture 7, 10/12/2012.

Exercise 3.10. If \( a_n \geq 0 \), then \( \lim_{n \to \infty} a_n = 0 \) if and only if \( \limsup_{n \to \infty} a_n = 0 \).

Proposition 3.31. 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 \( B \) is not bounded from above and \( \sup B = \limsup_{n \to \infty} a_n \). If \( B = \emptyset \) so that \( \sup B = -\infty \), then for all \( y \in \mathbb{R} \) we must have \( a_n < y \) for a.a. \( n \). This then implies \( \limsup_{n \to \infty} a_n \leq y \) for all \( y \in \mathbb{R} \) from which we conclude that \( \limsup_{n \to \infty} a_n = -\infty \). So let us now assume that \( B \neq \emptyset \) and \( \{a_n\}_{n=1}^\infty \) is bounded in which case \( B \) is bounded from above. Let us set \( \beta := \sup B \in \mathbb{R} \) and \( a^* := \limsup_{n \to \infty} a_n \).

If \( y > \beta \), then \( a_n < y \) for a.a. \( n \) from which it follows that \( a^* := \limsup_{n \to \infty} a_n \leq y \). We may now let \( y \downarrow \beta \) in order to see that \( a^* \leq \beta \)[^7] Now suppose that \( y < \beta \), then \( a_n \geq y \) for a.a. \( n \) and hence \( a^* = \limsup_{n \to \infty} a_n \geq y \). Letting \( y \uparrow \beta \) then shows \( a^* \geq \beta \). Thus we have shown \( a^* = \beta \).

\[ \blacksquare \]

Theorem 3.32. 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 subsequence \( \{a_{n_k}\}_{k=1}^\infty \) of \( \{a_n\}_{n=1}^\infty \) such that \( \liminf_{n \to \infty} a_{n_k} = \liminf_{n \to \infty} a_n \). Moreover, every convergent subsequence, \( \{b_k := a_{n_k}\}_{k=1}^\infty \) of \( \{a_n\}_{n=1}^\infty \), satisfies

\[
\lim_{k \to \infty} a_{n_k} \leq \limsup_{n \to \infty} b_k \leq \limsup_{n \to \infty} a_n.
\]

Proof. Let me prove the last assertion first. Suppose that \( b_k := a_{n_k} \) is some convergent subsequence of \( \{a_n\}_{n=1}^\infty \). We then 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,

\[
\liminf_{n \to \infty} a_n = \liminf_{k \to \infty} \inf_{n \geq n_k} a_n \leq \limsup_{k \to \infty} b_k \leq \limsup_{n \to \infty} \sup_{n \geq n_k} a_n = \limsup_{n \to \infty} a_n.
\]

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 of \( \{\inf_{n \geq k} a_n\}_{k=1}^\infty \) and \( \{\sup_{n \geq k} a_n\}_{k=1}^\infty \) respectively and therefore converge to the same limits respectively, see Lemma 3.25.

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

i) \( A \in \mathbb{R} \), then by Proposition 3.31 for all \( k \in \mathbb{N} \) we have \( A - \frac{1}{k} \leq a_n \) for infinitely many \( n \). In particular we can choose \( n_1 < n_2 < n_3 < \ldots \) inductively so that \( A - \frac{1}{k} \leq a_{n_k} \) for all \( k \). Since

\[
A - \frac{1}{k} \leq a_{n_k} \leq \sup_{m \geq n_k} a_m
\]

and the limit as \( k \to \infty \) of both extremes of this inequality are \( A \), it follow from the sandwich inequality that \( \lim_{k \to \infty} a_{n_k} = A \).

[^7]: This can be done more formally by choosing a sequence \( \{y_k\}_{k=1}^\infty \) such that \( y_k \downarrow \alpha \) so that \( a^* \leq y_k \). Now let \( k \to \infty \) to conclude \( a^* \leq \limsup_{k \to \infty} y_k = \alpha \).
ii) If $A = \limsup_{n \to \infty} a_n = \infty$, then $\sup_{k \geq n} a_k = \infty$ for all $n \in \mathbb{N}$ which implies for all $M < \infty$ that $a_k \geq M$ i.o. $k$. Working similarly to case i) we can choose $n_1 < n_2 < n_3 < \ldots$ so that $a_{n_k} \geq k$ for all $k$ and therefore $\lim_{k \to \infty} a_{n_k} = \infty$.

iii) Finally suppose that $A = \limsup_{n \to \infty} a_n = -\infty$ 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.

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

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

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

In particular, for $m, n \geq k \geq N$ we have $a_m \leq a_n + \varepsilon$ and hence $b \leq \sup_{m \geq k} a_m \leq a_n + \varepsilon$ 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 $b \leq a$.

\section*{Corollary 3.33 (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 $\mathbb{R}$.

\textbf{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,

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

It now follows from Theorem 3.32 that there exists a subsequence, $\{a_{n_k}\}_{k=1}^{\infty}$, of $\{a_n\}_{n=1}^{\infty}$ such that

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

\section*{Theorem 3.34 ($\mathbb{R}$ 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,

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

\textbf{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.

\textbf{First proof.} By Corollary 3.33 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 $L$.

\textbf{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

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

Notice that this exercise gives another proof of item 2. of Proposition 3.28 in the case all limits are real valued.

\begin{exercise} \label{exercise11} Suppose that $\{a_n\}_{n=1}^{\infty}$ is a sequence of real numbers and let $A := \{y \in \mathbb{R} : a_n \geq y \text{ for a.a. } n\}$.

Then $\sup A = \limsup_{n \to \infty} a_n$ with the convention that $\sup A = -\infty$ if $A = \emptyset$.

\begin{exercise} \label{exercise12} Suppose that $\{a_n\}_{n=1}^{\infty}$ is a sequence of real numbers. Show $\limsup_{n \to \infty} a_n = a^* \in \mathbb{R}$ iff for all $\varepsilon > 0$,

\[ a_n \leq a^* + \varepsilon \text{ for a.a. } n \text{. and } \]

\[ a^* - \varepsilon \leq a_n \text{ i.o. } n \text{.} \]

Similarly, show $\liminf_{n \to \infty} a_n = a_* \in \mathbb{R}$ iff for all $\varepsilon > 0$,

\[ a_n \leq a_* + \varepsilon \text{ i.o. } n \text{. and } \]

\[ a_* - \varepsilon \leq a_n \text{ for a.a. } n. \]

\end{exercise}

\section*{3.3 Partitioning the Real Numbers}

\textbf{Notation 3.35 (Intervals)} For $a, b \in \mathbb{R}$ with $a < b$ we define,

\[ (a, b) := \{x \in \mathbb{R} : a < x < b\}, \]

\[ [a, b] := \{x \in \mathbb{R} : a \leq x \leq b\}, \]

\[ (a, b) := \{x \in \mathbb{R} : a < x \leq b\}, \text{ and } \]

\[ [a, b) := \{x \in \mathbb{R} : a \leq x < b\}. \]

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 3.36 (Pairwise disjoint unions) If $X$ is a set and $A_{\alpha} \subset X$ for $\alpha \in I$, we write $X = \sum_{\alpha \in I} A_{\alpha}$ to mean: $X = \cup_{\alpha \in I} A_{\alpha}$ and $A_{\alpha} \cap A_{\beta}$ for all $\alpha \neq \beta$.

Exercise 3.14. 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 3.37 (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 $\tilde{S} := \{ s - k + 1 : s \in S \} \subset \mathbb{N}$ and hence by the well-ordering principle, $\min(\tilde{S}) := m \in \mathbb{N}$ 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,

$$s_0 = m + k - 1 \leq s \text{ for all } s \in S,$$

which is to say $s_0 = \min(S)$.

Proposition 3.38. Suppose that $\{ S_n \}_{n=-\infty}^{\infty} \subset \mathbb{R}$ such that $S_n < S_{n+1}$ 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 3.14. 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 $x \not\leq S_{n_0-1}$, 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. (3.10). The proof of the second equality is similar and will be omitted.

Proposition 3.39. Suppose that $-\infty < a < b < \infty$ and $\{ S_n \}_{n=0}^{N} \subset [a, b]$ such that $a = S_0 < S_1 < \cdots < S_{N-1} < S_N = b$, then

$$[a, b] = \sum_{n=1}^{N} (S_{n-1}, S_n).$$

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 3.38 and so will be omitted.

3.4 The Decimal Representation of a Real Number

Lemma 3.40 (Geometric Series). Let $\alpha \in \mathbb{R}$ or $\alpha \in \mathbb{Q}$, $m, n \in \mathbb{Z}$ and $S := \sum_{k=m}^{n} \alpha^k$. 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 $\alpha = 1$,

$$S = \sum_{k=m}^{n} 1^k = m - n + 1.$$

If $\alpha \neq 1$, then

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

Solving for $S$ gives

$$S = \sum_{k=m}^{n} \alpha^k = \frac{\alpha^{m+1} - \alpha^n}{\alpha - 1} \text{ if } \alpha \neq 1. \ (3.11)$$

Taking $\alpha = 10^{-1}$ in Eq. (3.11) implies

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

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

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

Definition 3.41 (Decimal Numbers). Let $\mathbb{D}$ denote those sequences $\alpha \in \{0, 1, 2, \ldots, 9\}_{\mathbb{N}}$ with the following properties:

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

A decimal number is then an expression of the form

$$\alpha_{-N} \alpha_{-N+1} \cdots \alpha_0 \alpha_1 \alpha_2 \alpha_3 \cdots$$

For example

$$52 + \sqrt{2} \equiv 53.41421356237309504880168872420969807856967187537694807\cdots$$
To every decimal number $\alpha \in \mathbb{D}$ is the sequence $a_n = a_n(\alpha)$ defined for $n \in \mathbb{N}$ by

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

Since for $m > n$,

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

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 3.42.** 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

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

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

$$[m + \frac{\alpha_1}{10}, m + \frac{\alpha_1 + 1}{10}) = \sum_{l=0}^{9} \left\{ m + \frac{\alpha_1}{10} + \frac{l}{100} \right\} [0, 10^{-2})$$

and set $\alpha_2 = l$ for $x \in [m + \frac{\alpha_1}{10} + \frac{l}{100}, m + \frac{\alpha_1 + 1}{10} + \frac{l+1}{100})$. Continuing this way inductively we construct $\{\alpha_k\}_{k=1}^{\infty}$ such that

$$x \in [m + \sum_{j=1}^{k} \frac{\alpha_j}{10^j}, m + \sum_{j=1}^{k-1} \frac{\alpha_j}{10^j} + \frac{\alpha_k + 1}{10^k}).$$

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

**Remark 3.43.** The representation of $x \geq 0$ as a decimal number may not be unique. For example,

$$0.999 = \sum_{k=1}^{\infty} 9 \cdot 10^{-k} = 9 \cdot \frac{1}{10} = 0.9 = \frac{9}{10}.$$  

[Or note that

$$1 = 0.999 \ldots 9 = 0.000 \ldots 1 = 10^{-n} \to 0 \text{ as } n \to \infty.$$]

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 $\mathbb{R}$ if 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

$$\lim_{n \to \infty} a_n = \lim_{n \to \infty} a_n = \limsup_{n \to \infty} a_n.$$  

We may allow the values of $\pm \infty$ in these statements.
13. If \( bk := \{a_{nk}\}_{k=1}^\infty \) is a convergent subsequence of \( \{a_n\} \), then

\[
\lim \inf a_n \leq \lim_{k \to \infty} b_k \leq \lim \sup a_n
\]

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

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

\[
a_n \leq a_{n+1} \text{ for all } n \text{ and } \lim_{n \to \infty} a_n = \sup (S).
\]

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

\[
a_{n+1} \leq a_n \text{ for all } n \text{ and } \lim_{n \to \infty} a_n = \inf (S).
\]

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 3.15. Prove the following properties of \( \mathbb{R} \).

1. Show addition and multiplication in Theorem 3.6 are well defined.

2. Show \((\mathbb{R},+,-)\) satisfies the axioms of a field. \textbf{Hint:} for constructing multiplicative inverses, make use of Proposition 3.44 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 a.a. \( n \). 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

\[
\alpha^{-1} = \{a^{-1}_n\}_{n=1}^\infty.
\]

3. Show \( i: \mathbb{Q} \to \mathbb{R} \) is 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 3.44. 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 \( a_n \geq b_n + \varepsilon \) for a.a. \( n \) in which case \( \alpha > \beta \), or
3. there exists \( \varepsilon = \frac{1}{N} > 0 \) such that \( b_n \geq a_n + \varepsilon \) for a.a. \( n \) in which case \( \beta > \alpha \).

\textbf{Proof.} If case 1. does not hold then there exists \( \delta > 0 \) 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) \( a_n - b_n \geq \delta \text{ i.o. } n \),

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

Since \( \{a_n\} \) and \( \{b_n\} \) are Cauchy sequences, there exists \( N \in \mathbb{N} \) such that

\[
|a_n - a_m| \geq \frac{\delta}{3} \text{ and } |b_n - b_m| \geq \frac{\delta}{3} \text{ for all } m,n \geq N.
\]

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,

\[
\delta \leq a_m - b_m = a_m - a_n + a_n - b_n + b_n - b_m
\]

\[
\leq |a_m - a_n| + a_n - b_n + |b_n - b_m|
\]

\[
= \frac{\delta}{3} + a_n - b_n + \frac{\delta}{3}
\]

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

Corollary 3.45. Suppose that \( \alpha := \{(a)_n\}_{n=1}^\infty \) and \( \beta := \{(b)_n\}_{n=1}^\infty \) are real numbers, then \( \alpha \geq \beta \text{ iff for all } N \in \mathbb{N} \),

\[
a_n - b_n \geq -\frac{1}{N} \text{ for a.a. } n.
\]

Alternatively put, \( \alpha \geq \beta \text{ iff for all } N \in \mathbb{N} \),

\[
b_n \leq a_n + \frac{1}{N} \text{ for a.a. } n.
\]

\textbf{Proof.} If \( \alpha = \beta \), then \( \lim_{n \to \infty} (a_n - b_n) = 0 \) and therefore Eq. (3.12) holds. If \( \alpha > \beta \), then in fact \( a_n - b_n \geq \varepsilon > 0 \) 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. (3.12) were to also hold we could conclude for each \( N \in \mathbb{N} \) that

\[
a_n \geq b_n - \frac{1}{N} \geq a_n + \varepsilon - \frac{1}{N} \text{ for a.a. } n.
\]

This leads to a contradiction as soon as we choose \( N \) so large as to make \( 1/N < \varepsilon \). Thus if Eq. (3.12) holds we must have \( \alpha \geq \beta \). ■

Proposition 3.46. Suppose that \( \lambda \in \mathbb{R} \), \( \{a_n\}_{n=1}^\infty \) be a Cauchy sequence in \( \mathbb{Q} \), and \( \alpha := \{(a)_n\}_{n=1}^\infty \). 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 $\lambda_n \leq a_k + \frac{1}{2N}$ for all $n$. Because $\{a_k\}$ is Cauchy, we may conclude there exists $M \in \mathbb{N}$ such that $\lambda_n \leq a_k + \frac{1}{2N}$ for all $n, k \geq M$.

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

$$a_n + \frac{1}{N} \leq \lambda_n \leq a_n + \frac{1}{2N} \Rightarrow \frac{1}{2N} \geq \frac{1}{N}$$

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 3.46 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 $i(a_m) := i\left(\frac{k_m}{2^n}\right)$ 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

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

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

If $\lambda \in A$, then $\lambda \leq i(a_n)$ for all $n$ and so by Proposition 3.46 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 3.46 we learn that

$$\alpha = \{\{a_n\}_{n=1}^{\infty}\} = \{\{a_n - 2^{-n}\}_{n=1}^{\infty}\} \leq \beta.$$  

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

Theorem 3.47 (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(q_{1F}) = q_{1G}. \text{ We know that } \{q \cdot 1_F : q \in \mathbb{Q}\} \text{ and } \{q \cdot 1_G : q \in \mathbb{Q}\} \text{ are dense copies of } \mathbb{Q} \text{ inside of } \mathbb{F} \text{ and } \mathbb{G} \text{ respectively. Now for general } a \in \mathbb{F} \text{ choose } q_n, p_n \in \mathbb{Q} \text{ that } q_n 1_F \uparrow a \text{ and } p_n 1_F \downarrow a. \text{ Since } \varphi \text{ is order preserving we must have } q_n 1_G = \varphi(q_n 1_F) \text{ is increasing and } p_n 1_G = \varphi(p_n 1_F) \text{ is decreasing. Moreover, since } p_n - q_n \to 0 \text{ we must have } \lim_{n \to \infty} \varphi(q_n 1_F) = \lim_{n \to \infty} \varphi(p_n 1_F). \text{ Since } \varphi(q_n 1_F) \leq \varphi(a) \leq \varphi(p_n 1_F) \text{ for all } n \text{ it then follows that } \varphi(a) = \lim_{n \to \infty} q_n 1_G = \lim_{n \to \infty} p_n 1_G \text{ and we have shown } \varphi \text{ is uniquely determined.}

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

$$|q_n 1_F - q_m 1_F| = |q_n - q_m| 1_F \text{ and } |q_n 1_G - q_m 1_G| = |q_n - q_m| 1_G.$$  

Thus if $\{q_n 1_F\}_{n=1}^{\infty}$ is convergent in $\mathbb{F}$ iff $\{q_n 1_G\}_{n=1}^{\infty}$ is convergent in $\mathbb{G}$. Thus for any $a \in \mathbb{F}$ we choose $q_n \in \mathbb{Q}$ such that $q_n 1_F \to a$ and then define $\varphi(a) := \lim_{n \to \infty} q_n 1_G$. One now checks that this formula is well defined (independent of the choice of $\{q_n\} \subset \mathbb{Q}$ such that $q_n 1_F \to a$) 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 $q_n 1_F \uparrow a$ and $p_n 1_F \downarrow b$. Then $q_n 1_G \leq p_n 1_G$ for all $n$ and letting $n \to \infty$ shows,

$$\varphi(a) = \lim_{n \to \infty} q_n 1_G \leq \lim_{n \to \infty} p_n 1_G = \varphi(b).$$

The other properties of $\varphi$ are proved similarly.
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 $2$ and in particular $1$岛屿 the usual rule for vector addition. As is standard we will write
\[
Re z = x and Im z = y.
\]

Solving these equations as follows
\[
a(4.4) + b(4.5) \implies (a + b^2)c = a \implies Re w = c = \frac{a}{a^2 + b^2}
\]
\[
-b(4.4) + a(4.5) \implies (a + b^2)d = -b \implies Im w = d = \frac{b}{a^2 + b^2}.
\]
gives implies the result in Eq. (4.3).

Probably the most painful thing to check directly is the associative law, namely that $[z_1z_2]z_3 = z_1[z_2z_3]$ for all $z_1, z_2, z_3 \in \mathbb{C}$. This is equivalent to showing for all $a, b, u, v, x, y \in \mathbb{R}$ that
\[
[(a + ib)(u + iv)](x + iy) = (a + ib)[(u + iv)(x + iy)].
\]

We do this by working out both sides as follows;
\[
LHS = [(au - bv) + i(au + bv)](x + iy)
= (au - bv)x - (av + bu)y + i[(av + bu)x + (au - bv)y];
RHS = (a + ib)[(ux - vy) + i(uy + vx)]
= a(ux - vy) - b(uy + vx) + i[b(ux - vy) + a(uy + vx)].
\]
The reader should now easily see that both of these expressions are in fact equal. The remaining axioms of a field are checked similarly.

- End of Lecture 9, 10/17/2012.
- Test 1 took place of lecture 10, 10/22/2012.

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
\[
|z|^2 \equiv \bar{z}z = a^2 + b^2.
\]
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. $\overline{zw} = \bar{z}\bar{w}$ and $\bar{z} + \bar{w} = \bar{z + w}$.
3. $|\bar{z}| = |z|$.
4. \( |zw| = |z||w| \) and in particular \( |z^n| = |z|^n \) for all \( n \in \mathbb{N} \).
5. \( \text{Re} z \leq |z| \) and \( |\text{Im} z| \leq |z| \)
6. \( |z+w| \leq |z| + |w| \).
7. \( z = 0 \) iff \( |z| = 0 \).
8. If \( z \neq 0 \) then

\[
z^{-1} := \frac{\bar{z}}{|z|^2}
\]

(also written as \( \frac{1}{z} \)) is the inverse of \( z \).
9. \( |z^{-1}| = |z|^{-1} \) and more generally \( |z^n| = |z|^n \) for all \( n \in \mathbb{Z} \).

**Proof.** 1. and 3. are geometrically obvious as well as easily verified.

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. \( |zw|^2 = zw\bar{w} = z\bar{w}\bar{w} = |z|^2|w|^2 \) 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

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

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

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

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

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

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

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

**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

\[
\frac{1}{u} \frac{1}{v} = \frac{1}{uwv}, \text{ i.e. } u^{-1}v^{-1} = (uv)^{-1}
\]

\[
\frac{u}{v} = \frac{uw}{vz} \text{ and }
\]

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

[These statements hold in any field.]
Theorem 4.10. The rest follow using 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 \( |z - z_n| \to 0 \) 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.

Proof. By assumption there exists \( M < \infty \) 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 3.33 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 3.33 again to find an an increasing function \( N \ni l \to k_l \in \mathbb{N} \) such that \( \lim_{l \to \infty} b_{n_k} = B \) exists. We now let \( w_l := z_{n_k} \) 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,

\[
|w_l - (A + iB)| = |a_{n_k} - A + i(b_{n_k} - B)| \\
\leq |a_{n_k} - A| + |b_{n_k} - B| \to 0 \quad \text{as} \quad l \to \infty.
\]

Notation 4.12 (Euclidean Spaces) Let \( \mathbb{C}^n := \{a := (a_1, \ldots, a_n) : a_i \in \mathbb{C}\} \) and \( \mathbb{R}^n := \{a := (a_1, \ldots, a_n) : a_i \in \mathbb{R}\} \subset \mathbb{C}^n \). For \( a, b \in \mathbb{C}^n \) we let \( \bar{a} = (\bar{a_1}, \ldots, \bar{a_n}) \),

\[
a \cdot b := a_1 b_1 + \cdots + a_n b_n = \sum_{i=1}^{n} a_i b_i, \quad \text{and} \quad \|a\| = \|a\|_2 = \sqrt{a \cdot \bar{a}} = \sqrt{\sum_{i=1}^{n} |a_i|^2}.
\]

Exercise 4.1. If \( a, b \in \mathbb{C}^n \) and \( \lambda \in \mathbb{C} \), then \( a \cdot b = b \cdot a \),

\[
(\lambda a) \cdot b = a \cdot (\lambda b) = \lambda (a \cdot b) \quad \text{and} \quad \|\lambda a\| = |\lambda| \|a\| = |\lambda| \|\bar{a}\|.
\]

If we further assume that \( c \in \mathbb{C}^n \), then \( (a + b) \cdot c = a \cdot c + b \cdot c \).

4.1 A Matrix Perspective (Optional)

Here is a way to understand some of the basic properties of \( \mathbb{C} \) using your knowledge of linear algebra. Let \( M_z : \mathbb{C} \to \mathbb{C} \) denote multiplication by \( z = a + ib \).

We now identify \( \mathbb{C} \) with \( \mathbb{R}^2 \) by

\[
\mathbb{C} \ni c + id \cong \begin{pmatrix} c \\ d \end{pmatrix} \in \mathbb{R}^2.
\]

Using this identification, the product formula

\[
zw = (ac - bd) + i(bc + ad),
\]

becomes
\[ M_z w = \begin{pmatrix} ac - bd \\ bc + ad \end{pmatrix} = \begin{pmatrix} a - b \\ b - a \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} \]

so that

\[ M_z = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = aI + bJ \]

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

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

4. the associativity of complex multiplication follows from the associativity properties of matrix multiplication,
5. \( M_z^* = aI - bJ = M_{\bar{z}} \) and in particular
6. \( M_{\bar{z}w} = (M_z M_w)^* = M_w^* M_z^* = M_{\bar{w}} M_{\bar{z}} = M_{\bar{w} \bar{z}}, \)
7. \( M_{\bar{z}w} M_z = M_{\bar{z}z} = M_{|z|^2} = \det(M_z), \)
8. \( |wz| = \det(M_{wz}) = \det(M_{w} M_{z}) = \det(M_{w}) \det(M_{z}) = |w| |z|, \)
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

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

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

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

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 \( \mathbb{Z} = \mathbb{N}_0 \cup \{-\mathbb{N}_0\} \) – 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 \( f : X \to Y \). If \( X = \mathbb{N} \), we will say that \( f \in Y^N \) is a sequence with values in \( Y \) and often write \( f_n \) for \( f(n) \) and express \( f \) as \( \{f_n\}_{n=1}^\infty \). If \( X = \{1, 2, \ldots, N\} \), we will write \( Y^N \) in place of \( Y^{\{1, 2, \ldots, N\}} \) and denote \( f \in Y^N \) by \( f = (f_1, f_2, \ldots, f_N) \) where \( f_n = f(n) \).

**Notation 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 \( \pi_\alpha : X_A \to X_\alpha \) 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

\[
A \triangle B := (B \setminus A) \cup (A \setminus B).
\]

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

\[
\bigcup_{\alpha \in I} A_\alpha := \{x \in X : \exists \alpha \in I \exists x \in A_\alpha\}
\]

and

\[
\bigcap_{\alpha \in I} A_\alpha := \{x \in X : x \in A_\alpha \forall \alpha \in I\}.
\]

**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 \( \bigcup_{\alpha \in I} A_\alpha \) 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

\[
\{A_n \text{ a.a.}\} := \{x \in X : x \in A_n \forall n \text{ sufficiently large}\}.
\]

(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

\[
\{A_n \text{ i.o.}\} = \bigcap_{N=1}^\infty \bigcup_{n \geq N} A_n.
\]

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

\[
\{A_n \text{ a.a.}\} = \bigcap_{N=1}^\infty \bigcup_{n \geq N} A_n.
\]
For example, suppose that 

\[ f \in \mathcal{F} \]

inclusion map \( f \) holds. To prove this notice that 

Example 5.7. 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 

\[ x \in f^{-1}(B^c) \iff f(x) \notin B \iff x \notin f^{-1}(B) \iff x \in [f^{-1}(B)]^c. \]

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 \( [f(A)]^c = \{1\}^c = \{2\} \).

Definition 5.6. 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 

\[ x \in f^{-1}(B^c) \iff f(x) \notin B \iff x \notin f^{-1}(B) \iff x \in [f^{-1}(B)]^c. \]

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 \( [f(A)]^c = \{1\}^c = \{2\} \).

Definition 5.6. If \( f : X \to Y \) is a function and \( E \subset 2^Y \), let 

\[ f^{-1}(E) := \{ x \in X : f(x) \in E \}. \]

If \( G \subset 2^X \), let 

\[ f^* G := \{ A \in 2^X : f^{-1}(A) \in G \}. \]

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 \( i_A(x) = x \) for all \( x \in A \) and 

\[ E_A = i_A^{-1}(E) = \{ A \cap E : E \in E \}. \]

5.1.1 Exercises

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

Exercise 5.1. \(( \cap_{i \in I} A_i)^c = \cup_{i \in I} A_i^c \).

Exercise 5.2. Suppose that \( B \subset Y \), show that \( B \setminus (\cup_{i \in I} A_i) = \cap_{i \in I} (B \setminus A_i) \).

Exercise 5.3. \( f^{-1}(\cup_{i \in I} A_i) = \cup_{i \in I} f^{-1}(A_i) \).

Exercise 5.4. \( f^{-1}(\cap_{i \in I} A_i) = \cap_{i \in I} f^{-1}(A_i) \).

Exercise 5.5. Find a counterexample which shows that \( f(C \cap D) = f(C) \cap f(D) \) need not hold.

5.2 Cardinality

In this section, \( X \) and \( Y \) be sets.

Definition 5.10 (Cardinality). We say \( \text{card}(X) \leq \text{card}(Y) \) if there exists an injective map, \( f : X \to Y \) and \( \text{card}(Y) \leq \text{card}(X) \) if there exists a surjective map \( g : Y \to X \). We say \( \text{card}(X) = \text{card}(Y) \) (also denoted as \( X \sim Y \)) if there exists a bijections, \( f : X \to Y \).

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 [B,3] and for the second see Theorem B.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 \( f : J_n \rightarrow J_{n+1} \setminus \{k\} \) 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. \( \square \)

Alternatively. If \( n = 1 \), then \( J_2 = \{1, 2\} \) and either \( J_2 \setminus \{k\} = J_1 \) 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 \( J_{n+2} \setminus \{k\} = (J_{n+1} \setminus \{k\}) \cup \{n+2\} \sim J_n \cup \{n+2\} \sim J_n \cup \{n+1\} \sim J_{n+1} \). \( \square \)

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

Proof. If \( f : J_n \rightarrow J_m \) were injective, then \( f|_{J_{n+1}} : J_{n+1} \rightarrow J_m \) would be injective as well. Therefore it suffices to show there is no injective map, \( f : J_{m+1} \rightarrow J_m \) 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, \( f : J_2 \rightarrow J_1 \) is the function, \( f(1) = 1 = f(2) \) which is not injective.

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

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

Proof. As \( J_m \subset J_n \) if \( m \leq n \) and \( J_m = J_n \) 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, \( g : J_m \rightarrow J_n \). According to Lemma 5.16 this can only happen if \( m \leq n \). If \( \text{card}(J_n) = \text{card}(J_m) \) then \( \text{card}(J_m) \leq \text{card}(J_n) \) and \( \text{card}(J_m) \leq \text{card}(J_n) \) and hence \( m \leq n \) and \( n \leq m \), i.e. \( m = n \). \( \blacksquare \)

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 \( \#(S) = k \leq n \). Moreover if \( \#(S) = n \), then \( S = X \).

Proof. It suffices to assume that \( X = J_n \) and \( S \subset J_n \). 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 \( f(k) \geq k \) for all \( k \) that is defined, this process has to stop after at most \( n \) steps. Say it stops at \( k = s \) so \( S_{s+1} = \emptyset \). Then \( f : J_s \rightarrow S \) 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 \( f : J_n \rightarrow S \) is the identity map in this case, i.e. \( S = J_n \).

Proof 2. We prove this by induction on \( n \). When \( n = 1 \) the only non-empty subset of \( S \) of \( J_1 \) itself. Thus \( \#(S) = 1 \) and \( S = J_1 \). Now suppose that the result hold for some \( n \in \mathbb{N} \) and \( S \subset J_{n+1} \). If \( n+1 \notin S \), then \( S \subset J_n \) and by the induction hypothesis we know \( \#(S) = k \leq n + 1 \). 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 \( \#(S') = k < n \), i.e. there exists a bijection, \( f' : J_k \rightarrow S' \) and \( S' = J_n \) is \( k = n \). Therefore \( f : J_{k+1} \rightarrow S \) given by \( f = f' \) on \( J_k \) and \( f(k+1) = n+1 \) is a bijection from \( J_{k+1} \) to \( S \). Therefore \( \#(S) = k+1 \leq n + 1 \) with equality if \( S' = J_n \) which happens iff \( S = J_{n+1} \).

Proposition 5.19. If \( f : J_n \rightarrow J_n \) 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 \( f : J_1 \rightarrow J_1 \) is \( f(1) = 1 \). So in this case there is nothing to prove. So now suppose the proposition holds for level \( n \) and \( f : J_{n+1} \rightarrow J_{n+1} \) 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 \( \text{card}(J_{n+1}) = \text{card}(f(J_{n+1})) \) 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. \( f \circ g = \text{id} \), which is necessarily injective, see the proof of Proposition 5.18. By the previous paragraph we know that \( g \) is necessarily surjective and therefore \( f = g^{-1} \) is a bijection.

**Theorem 5.20.** A subset \( S \subseteq \mathbb{N} \) 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 \( S \subseteq J_n \) for some \( n \in \mathbb{N} \) and hence \( S \) is a finite set by Proposition 5.18. Also observe that if \( \#(S) = n = \text{sup}(S) \), then \( S \subseteq J_n \) and \( \#(S) = n = \#(J_n) \). Thus it follows from Proposition 5.18 that \( S = J_n \).

Conversely suppose that \( S \subseteq \mathbb{N} \) 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 \( \text{sup}(S) = k \geq 1 \) with equality iff \( S = J_1 \).

Suppose the truth of the statement for some \( n \in \mathbb{N} \) and let \( S \subseteq \mathbb{N} \) 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, \( \text{sup}(S \setminus \{k\}) \geq n \) with equality iff \( S \setminus \{k\} = J_n \). If \( \text{sup}(S \setminus \{k\}) > n \) then \( \text{sup}(S) \geq \text{sup}(S \setminus \{k\}) \geq n + 1 \) as desired. If \( \text{sup}(S \setminus \{k\}) = n \) then \( S \setminus \{k\} = J_n \) therefore \( S \supseteq k > n \). Hence it follows that \( \text{sup}(S) = k \geq n + 1 \).

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

**Proof.** If \( S \) is bounded we know \( \text{card}(S) = \text{card}(J_k) \) for some \( k \in \mathbb{N} \) which would violate the hypothesis that \( \text{card}(J_n) \leq \text{card}(S) \) for all \( n \in \mathbb{N} \). Conversely if \( \text{card}(S) \leq \text{card}(J_n) \) for some \( n \in \mathbb{N} \), then there exists and injective map, \( f : S \to J_n \). Therefore \( \text{card}(S) = \text{card}(f(S)) = \text{card}(J_k) \) 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 \( \#(X \times Y) = \#(X) \cdot \#(Y) \).

### 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 \( \text{card}(X) \leq \text{card}(\mathbb{N}) \). 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 \( f : X \to \mathbb{N} \) 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, \( f : \mathbb{N} \to X \). Moreover, if \( S \) is a subset of \( \mathbb{N} \) we may list its elements in increasing order so that either

\[
S = \{n_1 < n_2 < \cdots < n_k\} \text{ for some } k \in \mathbb{N} \text{ or }
S = \{n_1 < n_2 < \cdots < n_k < \cdots\} .
\]

In the first case \( \text{card}(X) = \text{card}(J_k) \) while in the second \( \text{card}(S) = \text{card}(\mathbb{N}) \). Define \( f(j) := n_j \) to set up the bijections between \( J_k \) and \( \mathbb{N} \) and \( S \).

The above arguments demonstrate that the following statements are equivalent:

1. \( X \) is countable, i.e. there exists a surjective map \( f : \mathbb{N} \to X \).
2. \( \text{card}(X) \leq \text{card}(\mathbb{N}) \), i.e. there exists an injective map, \( g : X \to \mathbb{N} \).
3. \( \text{card}(S) = \text{card}(\mathbb{N}) \) 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 finite 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 \( f : \mathbb{N} \to S \) defined by \( f(k) = n_k \) defines a bijection.
5.4 Countable and Uncountable Sets

Theorem 5.27. The following properties hold:

1. If \( X \) and \( Y \) are countable, then \( X \times Y \) is countable.
2. If \( X \) is a set and \( \text{card} J_n \leq \text{card} X \) for all \( n \in \mathbb{N} \), then \( \text{card} J \leq \text{card} X \).
3. If \( X \) is a subset of a countable set \( Y \) and \( \text{card} \ Y \leq \text{card} X \), then \( \text{card} \ X \) is countable.
4. If \( \{ X_n \}_{n \in \mathbb{N}} \) is a countable family of sets, then \( \bigcup_{n \in \mathbb{N}} X_n \) is countable.

Proof. We take each item in turn.

1. Put the elements of \( \mathbb{N} \times \mathbb{N} \) 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 \( \mathbb{N} \times \mathbb{N} \) into the following list form,

\[\mathbb{N} \times \mathbb{N} = \{ (1,1), (2,1), (1,2), (3,1), (2,2), (1,3), (4,1), \ldots, (1,4), \ldots \} .\]

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

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

4. To see this let \( f : \mathbb{N} \to X \) be a surjective map and let \( g(x) := \min f^{-1}(\{x\}) \) for all \( x \in X \). Then \( g : X \to \mathbb{N} \) is an injective map. Let \( S := g(X) \), then \( g : X \to S \subseteq \mathbb{N} \) is a bijection. So it remains to show \( S \sim \mathbb{N} \) or \( S \sim J_n \) for some \( n \in \mathbb{N} \). If \( S \) is unbounded, then \( S \sim \mathbb{N} \) 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 so 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 \( k \leq n \), i.e. there exists a bijection, \( f : J_k \to S \setminus \{ n + 1 \} \). 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 \( S \subseteq \mathbb{N} \) with \( S \sim J_n \) is bounded. Now suppose that \( S \subseteq \mathbb{N} \) with \( S \sim J_{n+1} \). Then \( f(J_n) \sim J_n \) and hence \( f(J_n) \) is bounded in \( \mathbb{N} \). Then \( \text{max} f(J_n) \cap (f(n) + 1) \) is an upper bound for \( S \). This completes the inductive argument.

6. For each \( n \in \mathbb{N} \) there exists and injection, \( f_n : J_n \to X \). By replacing \( X \) by \( X_0 := \bigcup_{n \in \mathbb{N}} f_n(J_n) \) we may assume that \( X = \bigcup_{n \in \mathbb{N}} f_n(J_n) \). Thus there exists a surjective map, \( f : \mathbb{N} \to X \) by item 3. Let \( g : X \to \mathbb{N} \) be defined by \( g(x) := \min f^{-1}(\{x\}) \) 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 \( J_k \sim S \sim X \). However this is impossible since \( \text{card} J_n \leq \text{card} X = \text{card} J_k \) would imply \( n \leq k \) even though \( n \) can be chosen arbitrarily in \( \mathbb{N} \).

7. If \( g : X \to \mathbb{N} \) is an injective map then \( g|A : A \to \mathbb{N} \) is an injective map and therefore \( A \) is countable.

Lemma 5.28. If \( X \) is a countable set which contains \( Y \subseteq X \) 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, \( (g \circ f)(J_n) \subseteq g(X) \) for all \( n \) implies \( \text{card} J_n \leq \text{card} (g(X)) \) for all \( n \) which implies \( g(X) \) is unbounded by Corollary 3.21.

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 \( f(n) = (f_1(n), f_2(n), f_3(n), \ldots) \). 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 \( Y^X \subset Y^X \) for any subset \( Y_0 \subset Y \), if \( Y_0^X \) is uncountable then so is \( Y^X \). In this way we may assume \( Y_0 \) is a two point set which may as well be \( Y_0 = \{0,1\} \). Moreover, since \( X \) is an infinite set we may find an injective map \( x : \mathbb{N} \to X \) and use this to set up an injection, \( i : 2^\mathbb{N} \to 2^X \) by setting \( i(A) := \{x_n : n \in \mathbb{N}\} \subset X \) for all \( A \subset \mathbb{N} \). If \( 2^X \) were countable we could find a surjective map \( f : 2^\mathbb{N} \to \mathbb{N} \) 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 \( \{0,1\} := \{a \in \mathbb{R} : 0 < a < 1\} \) is uncountable while \( \mathbb{Q} \cap (0,1) \) is countable. More generally, for any \( a < b \) in \( \mathbb{R} \), \( \text{card} (\mathbb{Q} \cap (a, b)) = \text{card} (\mathbb{N}) \) while \( \text{card} (\mathbb{Q} \cap (a, b)) > \text{card} (\mathbb{N}) \).

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 \( A_m := \{ \frac{n}{m} : n \in \mathbb{N} \} \) with \( n < m \}. \). Since \( \mathbb{Q} \cap (0,1) = \bigcup_{m=1}^\infty X_m \) 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 5.29.

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. \( \Rightarrow \) 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 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 \( f' : J_{n+1} \to X \) by \( f'(j_n) = f \) and \( f'(n+1) = x \). Then \( f' : J_{n+1} \to X \) is injective and hence \( \text{card} (J_{n+1}) \leq \text{card} (X) \).

2. \( \iff \) 3. This is the content of Theorem 5.19.

3. \( \Rightarrow \) 4. Let \( x_1 \in X \) and \( f : \mathbb{N} \to X \) be an injective map such that \( f(1) = x_1 \). We now define a bijections, \( \psi : X \to X \setminus \{x_1\} \) by

\[
\psi(x) = \begin{cases} 
 x & \text{if } x \notin f(\mathbb{N}) \\
 f(i+1) & \text{if } x = f(i) \in f(\mathbb{N}).
\end{cases}
\]

4. \( \Rightarrow \) 1. We will prove the contrapositive. If \( X \) is a finite and \( x \in X \), we have seen that \( \text{card} (X \setminus \{x\}) < \text{card} (X) \), namely \( \#(X \setminus \{x\}) = \#(X) - 1 \).

The next two theorems summarizes the properties of cardinalities that have been proven above.

Theorem 5.35 (Cardinality/Counting Summary I). Given a non-empty set \( X \), then one and only one of the following statements holds;

1. There exists a unique \( n \in \mathbb{N} \) such that \( \text{card} (X) = \text{card} (J_n) \),
2. \( \text{card} (X) = \text{card} \mathbb{N} \),
3. \( \text{card} (X) > \text{card} \mathbb{N} \).

Cases 2. or 3. hold iff \( \text{card} (J_n) \leq \text{card} X \) for all \( n \in \mathbb{N} \) which happens iff \( \text{card} \mathbb{N} \leq \text{card} X \).

If \( X \) satisfies case 1. we say \( X \) is a finite set. If \( X \) satisfies case 2 we say \( X \) is a countably infinite set and if \( X \) satisfies case 3. we say \( X \) is an uncountably infinite set.

Theorem 5.36 (Cardinality/Counting Summary II). Let \( X \) and \( Y \) be sets and \( S \) be a subset of \( \mathbb{N} \).
1. If $S \subset \mathbb{N}$ is an unbounded set, then $\text{card}(S) = \text{card}(\mathbb{N})$.
2. If $S \subset \mathbb{N}$ is a bounded set, then $\text{card}(S) = \text{card}(\mathbb{J}_n)$ for some $n \in \mathbb{N}$.
3. If $\{X_k\}_{k=1}^{\infty}$ are subsets of $X$ such that $\text{card}(X_k) \leq \text{card}(\mathbb{N})$, then $\text{card}(\bigcup_{k=1}^{\infty} X_k) \leq \text{card}(\mathbb{N})$.
4. If $X$ and $Y$ are sets such that $\text{card}(X) \leq \text{card}(\mathbb{N})$ and $\text{card}(Y) \leq \text{card}(\mathbb{N})$, then $\text{card}(X \times Y) \leq \text{card}(\mathbb{N})$.
5. $\text{card}(\mathbb{Q}) = \text{card}(\mathbb{N}) = \text{card}(\mathbb{Z})$.
6. For any $a < b$ in $\mathbb{R}$, $\text{card}(\mathbb{Q} \cap (a, b)) = \text{card}(\mathbb{N})$ while $\text{card}(\mathbb{Q}^c \cap (a, b)) > \text{card}(\mathbb{N})$.

### 5.4.1 Exercises

**Exercise 5.9.** Show that $\mathbb{Q}^n$ is countable for all $n \in \mathbb{N}$.

**Exercise 5.10.** Let $\mathbb{Q}[t]$ be the set of polynomial functions, $p$, such that $p$ has rational 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

$$p(t) = \sum_{k=0}^{n} a_k t^k \text{ for all } t \in \mathbb{R}. \quad (5.1)$$

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

**Definition 5.37 (Algebraic Numbers).** A real number, $x \in \mathbb{R}$, is called an 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.
Part II

Banach 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: There is a element $0 \in Z$ such that $0 + v = v$ 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. $\mathbb{C}$ 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
   $Z := R^T$. 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
   $J_n := \{1, \ldots, n\}$. 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 $\mathbb{R}^R$;

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 $\lambda \in \mathbb{R}$ 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

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

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

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

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

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

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 $\rho = y/x$. 

\[\blacksquare\]
Theorem 6.15 (Cauchy-Schwarz Inequality). For \( a, b \in \mathbb{C}^n \), \( |a \cdot b| \leq \|a\|_2 \cdot \|b\|_2 \).

**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

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

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

Theorem 6.16 (Triangle Inequality). The function, \( \|\cdot\|_2 \), 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;

\[
\|a + b\|_2^2 = (a + b) \cdot (\overline{a + b}) = (a + b) \cdot (\overline{a} + \overline{b}) = (a + b) \cdot \overline{a} + (a + b) \cdot \overline{b} = \|a\|_2^2 + \|b\|_2^2 + 2 \text{Re}(a \cdot \overline{b}) \\
\leq \|a\|_2^2 + \|b\|_2^2 + 2 \|a\|_2 \cdot \|b\|_2 \quad \text{(Theorem 6.15)}
\]

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

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

**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

\[
a \cdot b := \sum_{i=1}^n a_i b_i \rho_i \text{ and } \|a\| := \sqrt{\sum_{i=1}^n |a_i|^2 \rho_i}.
\]

Show, for all \( a, b \in \mathbb{C}^n \) that \( |a \cdot b| \leq \|a\| \cdot \|b\| \) and that \( \|\cdot\| \) is a norm on \( \mathbb{C}^n \).

**[Hint: reduce to the case where \( \rho_i = 1 \) for all \( i \).]**

- End of Lecture 12, 10/26/2012.

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 function 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 \( f \in Z \),

\[
\int_0^1 |f(t)| \, dt \leq \int_0^1 |f(t)| \, dt.
\]

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

\[
\int_0^1 f(t) \, dt = \int_0^1 \pm f(t) \, dt \leq \int_0^1 |f(t)| \, dt \iff \left| \int_0^1 f(t) \, dt \right| \leq \int_0^1 |f(t)| \, dt.
\]

**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 \( f, g \in Z \).

\footnote{The notion of continuity will be formally developed shortly.}
Remark 6.17.\footnote{\textit{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
   $$\int_0^1 f(t)g(t) \, dt \leq \|f\|_2 \cdot \|g\|_2,$$

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 6.18.\footnote{\textit{An interpretation of }$\|f\|_2$.} Let us suppose that $f(t)$ is the voltage across a 1 Ohm resistor. By Ohms law the current through this resistor is $f(t)/1 = f(t)$ and the power dissipated by the resistor at time $t$ is (Voltage-CURRENT) $f(t)^2$. 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 $\|a_i\| \leq \|b\|$ whenever $0 \leq a_i \leq b_i$ for $1 \leq i \leq n$.\footnote{For example we could take $\|\cdot\|$ to be $\|\cdot\|_u$, $\|\cdot\|_1$, or $\|\cdot\|_2$ on $\mathbb{R}^n$. Not all norms satisfy this assumption though. For example, take $\|(x,y)\| = |x-y| + |y|$ on $\mathbb{R}^2$, then $\|(x,y)\|$ is decreasing in $x$ when $0 \leq x < y$.} 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. \[\text{[We started Section 5.1 above as well this day.]}\]

6.2 Sequences in Metric Spaces

Definition 6.19. A sequence $\{x_n\}_{n=1}^\infty$ in a metric space $(X,d)$ is said to be \textbf{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 6.19 is necessarily unique.

Definition 6.20 (Cauchy sequences). A sequence $\{x_n\}_{n=1}^\infty$ in a metric space $(X,d)$ is \textbf{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

$$d(x_n,x_m) \leq \varepsilon \text{ when } n,m \geq N(\varepsilon).$$

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 $(\mathbb{Q},d)$ – Cauchy but not $(\mathbb{Q},d)$ – 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 6.21. A metric space $(X,d)$ (or normed space $(X,\|\cdot\|)$) is \textbf{complete} if all Cauchy sequences are convergent sequences. A complete normed space is called a \textbf{Banach space}.

Lemma 6.22. Let $X$ be a non-empty set and

$$\|f\|_u := \sup_{x \in \mathbb{R}} |f(x)| \text{ for all } f \in \mathbb{C}^X.$$ 

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

$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\|_u \quad (6.2)$$

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. (6.2) implies

$$|f - f_m(x)| \leq \lim_{n \to \infty} \|f_n - f_m\|_u \quad (6.2)$$
and taking the supremum over \( x \in X \) of this inequality implies
\[
\| f - f_m \|_u \leq \lim_{n \to \infty} \inf \| f_n - f_m \|_u \to 0 \text{ as } m \to \infty
\]
showing \( f_m \to f \) in \( Z \).

**Definition 6.23.** We say that two norms, \( \| \cdot \|_a \) and \( \| \cdot \|_b \), on a vector space \( X \) are equivalent if there are constants \( C_1, C_2 \in (0, \infty) \) such that
\[
\| x \|_a \leq C_1 \| x \|_b \text{ and } \| x \|_b \leq C_2 \| x \|_a \quad \text{for all } x \in X.
\]
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
\[
d_a(x,y) \leq C_1 d_b(x,y) \quad \text{and} \quad d_b(x,y) \leq C_2 d_a(x,y) \quad \text{for all } x, y \in X.
\]

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

**Corollary 6.24.** 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
\[
d_b(x_n,x_m) \leq C_2 d_a(x_n,x_m) \to 0 \text{ as } m,n \to \infty
\]
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 \( d_b(x,x_n) \to 0 \) as \( n \to \infty \). Since
\[
d_a(x,x_n) \leq C_1 d_b(x,x_n) \to 0 \text{ as } n \to \infty
\]
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
\[
\| z \|_u \leq \| z \|_1 \leq n \| z \|_u,
\]
\[
\| z \|_1 \leq \sqrt{n} \| z \|_2, \quad \text{(Hint: use Cauchy Schwarz.)}
\]
\[
\| z \|_2 \leq \sqrt{\| z \|_u \cdot \| z \|_1} \leq \| z \|_1 \quad \text{and}
\]
\[
\| z \|_2 \leq \sqrt{n} \| z \|_u.
\]

It follows from these inequalities that \( \| \cdot \|_1 \), \( \| \cdot \|_u \), and \( \| \cdot \|_2 \) are equivalent norms on \( \mathbb{C}^n \).

**Theorem 6.25 (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 6.22 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
\[
d_{m,n} := \sup_{x \in X} d(f_n(x), f_m(x)) \to 0 \text{ as } m,n \to \infty.
\]
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 6.22.

**Exercise 6.12.** Let \( Z \) denote the continuous function 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
\[
\| f \|_u \leq C \| f \|_2 \quad \text{for all } f \in Z.
\]
**[Hint: consider the sequence, } f_n(t) = t^n.\]

**Example 6.26.** 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
\[
g(t) := \begin{cases} 2t & \text{if } t \leq 1/2 \\ 1 & \text{if } 1/2 < t \leq 1 \end{cases}
\]
and then set \( f_n(t) = g(t)^n \) for 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
\[
||f_n - h||_1 = \int_0^1 (2t)^n \, dt = \frac{1}{2} \left( \frac{1}{n+1} \right)
\]
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
\[
\lim f(t) = \lim h(t) = 1 \quad \text{and} \quad \lim f(t) = \lim h(t) = 0
\]
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.

**Definition 6.27.** Let \((X,d)\) be a metric space. We say \( A \subseteq X \) is **dense** in \( X \) if for all \( x \in X \), there exists \( \{x_n\}_{n=1}^\infty \subseteq A \) 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, \( D \).

**Exercise 6.13.** 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.

**Exercise 6.14.** Let \( n \in \mathbb{N} \). Show any non-empty subset \( Y \subseteq \mathbb{C}^n \) equipped with the metric,
\[
d(x,y) = ||y - x||
\]
is separable, where \( ||\cdot|| \) is either \( ||\cdot||_u \), \( ||\cdot||_1 \), or \( ||\cdot||_2 \).

**Exercise 6.15.** For \( x, y \in \mathbb{R} \), let
\[
d(x,y) := \begin{cases} 
1 & \text{if } x = y \\
0 & \text{if } x \neq y
\end{cases}
\]

### 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 6.28 (Limits of functions).** If \( x_0 \in X \) and \( f : X \setminus \{x_0\} \to Y \) is a function, then we say \( \lim_{x \to x_0} f(x) = y_0 \in Y \) iff for all \( \varepsilon > 0 \) there is a \( \delta = \delta(\varepsilon, x_0) > 0 \) such that
\[
d(f(x),y_0) \leq \varepsilon \text{ provided that } 0 < \rho(x,x_0) \leq \delta(\varepsilon, x_0).
\]
[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 6.29 (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 \( \{x_n\}_{n=1}^\infty \subseteq X \setminus \{x_0\} \) such that \( \lim_{n \to \infty} x_n = x_0 \).

**Proof.** Suppose that \( \lim_{x \to x_0} f(x) = y_0 \in Y \) and \( \{x_n\}_{n=1}^\infty \subseteq X \setminus \{x_0\} \) with \( \lim_{n \to \infty} x_n = x_0 \). Then all \( \varepsilon > 0 \) there is a \( \delta > 0 \) such that Eq. (6.4) holds. Since \( x_n \to x_0 \) as \( n \to \infty \) it follows that \( 0 < \rho(x_n,x_0) \leq \delta \) for a.a. \( n \) and therefore \( d(f(x_n),y_0) \leq \varepsilon \) for a.a. \( n \). As \( \varepsilon > 0 \) was arbitrary it follows that \( \lim_{n \to \infty} d(f(x_n),y_0) = 0 \), i.e. that \( \lim_{n \to \infty} f(x_n) = y_0 \).

Conversely if \( \lim_{x \to x_0} f(x) \neq y_0 \), then there exists \( \varepsilon > 0 \) such that for any \( \delta = \frac{1}{n} > 0 \) there exists \( x_n \in X \setminus \{x_0\} \) such that \( d(f(x_n),y_0) \geq \varepsilon \) while \( \rho(x_n,x_0) < \frac{1}{n} \). We then have \( \lim_{n \to \infty} x_n = x \) while \( \lim_{n \to \infty} f(x_n) \neq y_0 \).
Definition 6.30 (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
\[
\| f(x) - f(x') \| < \varepsilon \quad \text{provided that} \quad \rho(x, x') < \delta.
\] (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 6.31 (Sequential Continuity). A function \( f : X \to Y \) is continuous at \( x \in X \) if
\[
\lim_{n \to \infty} f(x_n) = f(x) \quad \text{for all} \quad \{x_n\}_{n=1}^{\infty} \subset X \text{ with} \quad \lim_{n \to \infty} x_n = x.
\]
We say \( f \) is sequentially continuous on \( X \) if it is continuous at all points in \( X \).

Corollary 6.32. Continuity and sequential continuity are the same notions.

Proof. This follows rather directly from Theorem 6.29.

Example 6.33. 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 6.34. Let \( f : \mathbb{R} \to \mathbb{R} \) be the function defined by
\[
f(x) = \begin{cases} 
1 & \text{if } x \in \mathbb{Q} \\
0 & \text{if } x \not\in \mathbb{Q}.
\end{cases}
\]
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 6.16. 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 6.17. 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 6.18. 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 6.19. 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 \( X \).

Exercise 6.20. Show that the following functions from \( \mathbb{C} \) to \( \mathbb{C} \) are continuous functions then \( |f| \) is continuous and
\[
F := \sum_{m,n=0}^{N} a_{m,n} f^m \cdot \bar{f}^n
\]
is continuous.

• End of Lecture 16, 11/5/2012.

Definition 6.36 (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
\[
\lim_{x \to x_0^+} f(x) = y_0 \iff \forall \varepsilon > 0 \exists \delta = \delta(\varepsilon) > 0 \ni d(f(x), y_0) \leq \varepsilon \text{ if } 0 < x - x_0 \leq \delta
\]
and
\[
\lim_{x \to x_0^-} f(x) = y_0 \iff \forall \varepsilon > 0 \exists \delta = \delta(\varepsilon) > 0 \ni d(f(x), y_0) \leq \varepsilon \text{ if } 0 < x_0 - x \leq \delta.
\]

Theorem 6.37 (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 \( \{x_n\}_{n=1}^{\infty} \subset (x_0, b) \) such that \( \lim_{n \to \infty} x_n = x_0 \).
3. \( \lim_{n \to \infty} f(x_n) = y_0 \) for all \( \{x_n\}_{n=1}^{\infty} \subset (x_0, b) \) such that \( x_n \downarrow x_0 \), i.e. \( x_{n+1} \leq x_n \) and \( x_n < x_n \) 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 \( \{x_n\}_{n=1}^\infty \subset (a, x_0) \) such that \( \lim_{n \to \infty} x_n = x_0 \).

c. \( \lim_{n \to \infty} f(x_n) = y_0 \) for all \( \{x_n\}_{n=1}^\infty \subset (a, x_0) \) such that \( x_n \uparrow x_0 \), i.e. \( x_{n+1} \geq x_n \) and \( x_n < x_0 \) 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. \( \implies \) 2.) If \( \lim_{x \to x_0} f(x) = y_0 \) then for all \( \epsilon > 0 \) there exists \( \delta = \delta (\epsilon) > 0 \) such that \( d(f(x), y_0) < \epsilon \) if \( 0 < x - x_0 < \delta \). Hence if \( \{x_n\}_{n=1}^\infty \subset (a, x_0) \) such that \( x_n \to x_0 \) then \( 0 < x_n - x_0 < \delta \) for a.a. \( n \) and hence \( d(f(x_n), y_0) < \epsilon \) for a.a. \( n \). Since \( \epsilon > 0 \) 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. \( \implies \) 3.)

(3. \( \implies \) 1.) If \( \lim_{x \to x_0} f(x) \neq y_0 \) there exists \( \epsilon > 0 \) such that for all \( \delta = \frac{1}{n} > 0 \) there exists \( x_n \in (x_0, b) \) such that \( 0 \leq x_n - x_0 \leq \frac{1}{n} \) while \( d(f(x_n), y_0) \geq \epsilon \). Then take \( x_n = \min (x_1', \ldots, x_n') \). Then \( x_n \downarrow x \) while \( d(f(x_n), y_0) \geq \epsilon \), i.e., \( \lim_{n \to \infty} f(x_n) \neq y_0 \). The equivalent of statements a.-c. are proved similarly and so the proofs will be omitted.

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 \( \epsilon > 0 \) is given, then choose \( \delta_+ (\epsilon) > 0 \) such that

\[
d(f(x), y_0) \leq \epsilon \quad \text{when} \quad 0 < x - x_0 < \delta_+ (\epsilon) \quad \text{or} \quad 0 < x_0 - x < \delta_+ (\epsilon) .
\]

If we take \( \delta (\epsilon) := \min (\delta_+(\epsilon), \delta_- (\epsilon)) \), then we will have

\[
d(f(x), y_0) \leq \epsilon \quad \text{when} \quad 0 < |x - x_0| \leq \delta (\epsilon)
\]

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

Corollary 6.38 (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 \( x_0 \in X \) 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 \( \ref{thm:continuity} \).

Exercise 6.20 (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 6.21 (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 6.22 (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 6.23. 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 6.39 Let \((X, \rho) \) and \((Y, d) \) be metric spaces and \( f : X \to Y \) be a function.

1. We say \( f \) is uniformly continuous, iff for all \( \epsilon > 0 \) there exists \( \delta (\epsilon) > 0 \) such that

\[
\forall x, x' \in X \quad \rho(x, x') \leq \delta \implies d(f(x), f(x')) \leq \epsilon.
\]

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

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

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

\[
\forall x \in X \quad \rho(x, x_0) \leq \delta \implies d(f(x), f(x_0)) \leq \epsilon.
\]

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

\[
\rho(x, x') \leq \delta \implies d(f(x), f(x')) \leq C \rho(x, x') \leq C \delta = \epsilon
\]

which shows \( f \) is uniformly continuous.

Lemma 6.40 (Distance to a Set). For any non empty subset \( A \subset X \), let

\[
d_A(x) := \inf \{ d(x, a) \mid a \in A \},
\]

then

\[
|d_A(x) - d_A(y)| \leq d(x, y) \quad \forall x, y \in X.
\]

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

\footnote{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.}
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. (6.6).

\[\square\]

**Corollary 6.41.** 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 6.40 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. \[\square\]

**Exercise 6.24.** 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 6.39). Recall this means for every \(\varepsilon > 0\) there exists a \(\delta > 0\) such that
\[
d(f(a), f(b)) \leq \varepsilon \quad \text{for all } a, b \in A \text{ with } \rho(a, b) \leq \delta.
\]

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 \(\mathbb{R}\).

**Exercise 6.25 (Continuity of integration).** Let \(Z = C([0, 1], \mathbb{R})\) be the continuous functions from \([0, 1]\) to \(\mathbb{R}\) and \(\|\cdot\|_u\) be the uniform norm, \(\|f\|_u := \sup_{0 \leq t \leq 1} |f(t)|\). Define \(K : Z \to Z\) by
\[
K(f)(x) := \int_0^x f(t) \, dt \quad \text{for all } x \in [0, 1].
\]

Show that \(K\) is a Lipschitz function. In more detail, show
\[
\|K(f) - K(g)\|_u \leq \|f - g\|_u \quad \text{for all } f, g \in Z.
\]

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 \(\int_0^t f(t) \, dt \leq \int_0^t g(t) \, dt\) for all \(x \in X\).
4. From 3. it follows that \(\left|\int_0^x f(t) \, dt\right| \leq \int_0^x |f(t)| \, dt\).

**Exercise 6.26 (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 \(\|p\|_u := \sup_{0 \leq t \leq 1} |p(t)|\). 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 \(p \in Z\).
we will let $X$ as well. In particular if $x \to \delta > 0$ then $f$ is continuous at $x$ as well.

Definition 6.42 (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 $\lim_{n \to \infty} f_n(x) = f(x)$ for all $x \in X$, i.e. provided

$$\lim_{n \to \infty} d(f(x), f_n(x)) = 0 \text{ for each } x \in X.$$

Definition 6.43 (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

$$\delta_n := \sup_{x \in X} d(f(x), f_n(x)) \to 0 \text{ as } n \to \infty.$$

Hopefully it is clear to the reader the uniform convergence implies pointwise convergence. The next theorem is a basic fact about uniform convergence which does not hold in general for pointwise convergence. The proof of this theorem is a bit subtle but well worth mastering as the method will arise over and over again.

Theorem 6.44 (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 6.2.

**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 $x := \lim_{k \to \infty} x_k \in X$. So assume we are given such a sequence $\{x_k\}_{k=1}^\infty$. Then for any $n \in \mathbb{N}$ we have,

$$\rho(f(x), f(x_k)) \leq \rho(f(x), f_n(x)) + \rho(f_n(x), f(x_k))$$

$$\leq \rho(f(x), f_n(x)) + \rho(f_n(x), f_n(x_k)) + \rho(f_n(x_k), f(x_k))$$

$$\leq \delta_n + \rho(f_n(x), f_n(x_k)) + \delta_n.$$

Therefore,

$$\limsup_{k \to \infty} \rho(f(x), f(x_k)) \leq \limsup_{k \to \infty} \rho(f(x), f_n(x)) + 2\delta_n = 2\delta_n,$$

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

$$\rho(f(x), f(x')) \leq \rho(f(x), g(x)) + \rho(g(x), f(x'))$$

$$\leq \rho(f(x), g(x)) + \rho(g(x), g(x')) + \rho(g(x'), f(x'))$$

$$\leq \delta_n + \rho(g(x), g(x')) + \delta_n \leq \varepsilon + \varepsilon = 3\varepsilon.$$
As $\varepsilon > 0$ and $x \in X$ were arbitrary, we have shown $f$ is continuous on $X$. ■

**Exercise 6.28.** 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.
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 $X$ if $s := \lim_{N \to \infty} \sum_{n=1}^{N} x_n$ 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.

If $X = \mathbb{R}$ and $x_n \geq 0$, then $\sum_{n=1}^{\infty} x_n$ diverges if $\lim_{N \to \infty} \sum_{n=1}^{N} x_n = \infty$ and so we will write $\sum_{n=1}^{\infty} x_n = \infty$ in this case to indicate that $\sum_{n=1}^{\infty} x_n$ diverges to infinity.

Theorem 7.2 (Comparison Theorem). Suppose that $\{a_n\}_{n=1}^{\infty}$ 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 7.3 (Telescoping Series / Fundamental Theorem of Summation). Let $\{f(n)\}_{n=1}^{\infty} \subset X$ be a sequence, then
\[
\sum_{n=1}^{N} [f(n+1) - f(n)] = f(N+1) - f(1) \quad \text{for all } N \in \mathbb{N} \quad (7.1)
\]
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,
\[
\sum_{n=1}^{\infty} (f(n+1) - f(n)) = \lim_{N \to \infty} f(N) - f(1).
\]

Proof. When $N = 3$ we have,
\[
\sum_{n=1}^{3} (f(n+1) - f(n)) = (f(x_2) - f(1)) + (f(x_3) - f(x_2)) + (f(4) - f(x_3)) = f(4) - f(1).
\]
In general, Eq. (7.1) is 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
\[
\sum_{n=1}^{\infty} \alpha^n = \frac{\alpha}{1 - \alpha}.
\]

Proposition 7.5. Suppose that $f : (0, \infty) \to \mathbb{R}$ is a $C^1$ functions such that $f'(x) \leq 0$ and $f'(x)$ is increasing in $x$ (i.e. $f''(x) \geq 0$ if it exists). Setting $a_n := -f'(n) \geq 0$ for all $n \in \mathbb{N}$ we find,
\[
L \leq \sum_{n=1}^{\infty} a_n \leq L + a_1 \quad (7.2)
\]
where
\[
L := f(1) - f(\infty) \geq 0 \quad \text{(with } L = \infty \text{ allowed here)}
\]
and $f(\infty) := \lim_{x \to \infty} f(x)$ which exists because $f$ is decreasing.
Example 7.6 (p-series). Let $p > 1$, $\alpha := p - 1 > 0$, and $f (x) = x^{-\alpha}$ so that $f' (x) = -\alpha x^{-\alpha - 1}$ and $f'' (x) = \alpha (\alpha - 1) x^{-\alpha - 2} \geq 0$. In this case $L = 1 - 0 = 1$ and $a_n = \frac{\alpha}{n^p}$ and we have,

or equivalently,

$$1 \leq \sum_{n=1}^{\infty} \frac{\alpha}{n^p} \leq p$$

Notice that

$$\frac{1}{p - 1} \leq \sum_{n=1}^{\infty} \frac{1}{n^p} \leq \sum_{n=1}^{\infty} \frac{1}{n}$$

and so letting $p \downarrow 1$ shows $\sum_{n=1}^{\infty} \frac{1}{n} = \infty$.

Exercise 7.1. Take $f (x) = -\ln x$ in Proposition 7.5 in order to directly conclude that $\sum_{n=1}^{\infty} \frac{1}{n} = \infty$.

Theorem 7.7. Let $(X, \| \cdot \|)$ be a Banach space and $\{ x_k \}_{k=1}^{\infty} \subset X$ 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 $N \to \infty$.

Proof. Let $S_n := \sum_{k=1}^{n} x_k$ so that $\sum_{k=1}^{\infty} x_k$ converges iff $\lim_{n \to \infty} S_n$ exists iff $\{ S_n \}_{n=1}^{\infty}$ is Cauchy since $X$ 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,

$$\sum_{k=N}^{\infty} x_k = S - S_{N+1} \to 0 \text{ as } N \to \infty.$$ 

Proof.
Exercise 7.2. Let \((X, d)\) be a metric space. Suppose that \(\{x_n\}_{n=1}^{\infty} \subset X\) is a sequence and set \(\varepsilon_n := d(x_n, x_{n+1})\). Show that for \(m > n\) 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
\[
d(x, x_n) \leq \sum_{k=n}^{\infty} \varepsilon_k.
\]

Proposition 7.8 (Alternating Series Test). If \(\{a_k\}_{k=1}^{\infty} \subset [0, \infty)\) is a non-increasing sequence (i.e. \(a_k \geq a_{k+1}\) for all \(k\)) such that \(\lim_{k \to \infty} a_k = 0\), then \(s := \sum_{k=1}^{\infty} (-1)^k a_k\) is convergent. Moreover, for all \(n \in \mathbb{N}\)
\[
\left| s - \sum_{k=1}^{n} (-1)^k a_k \right| = \sum_{k=n+1}^{\infty} (-1)^k a_k \leq a_{n+1}.
\]

Exercise 7.3. Prove the alternating series test. That is if \(\{a_k\}_{k=1}^{\infty} \subset [0, \infty)\) is a non-increasing sequence (i.e. \(a_k \geq a_{k+1}\) for all \(k\)) such that \(\lim_{k \to \infty} a_k = 0\), then \(s := \sum_{k=1}^{\infty} (-1)^k a_k\) is convergent. Moreover, for all \(n \in \mathbb{N}\)
\[
\left| s - \sum_{k=1}^{n} (-1)^k a_k \right| = \sum_{k=n+1}^{\infty} (-1)^k a_k \leq a_{n+1}.
\]

[Hint: first show \(S_{2n+1} \geq S_{2n-1}\) and \(S_{2(n+1)} \leq S_{2n}\) for \(n \in \mathbb{N}\) where \(S_n := \sum_{k=1}^{n} (-1)^k a_k\).

Theorem 7.9 (Absolute Convergence Implies Convergence). Let \((X, \|\cdot\|)\) be a Banach space \(\{x_k\}_{k=1}^{\infty} \subset X\) be a sequence. Then \(\sum_{k=1}^{\infty} \|x_k\| < \infty\) then \(\sum_{k=1}^{\infty} x_k\) is convergent. We say \(\sum_{k=1}^{\infty} x_k\) is absolutely convergent if \(\sum_{k=1}^{\infty} \|x_k\| < \infty\).

Exercise 7.4. Prove Theorem 7.9. Namely if \((X, \|\cdot\|)\) is a Banach space and \(\{x_k\}_{k=1}^{\infty} \subset X\) is a sequence, then \(\sum_{k=1}^{\infty} \|x_k\| < \infty\) implies \(\sum_{k=1}^{\infty} x_k\) is convergent.

Theorem 7.10 (Weierstrass M-test). Suppose that \((X, \|\cdot\|)\) is a Banach space, \((Y, d)\) is a metric space, \(\{f_n\}_{n=1}^{\infty} \subset C(Y, X)\), and \(\{M_n\}_{n=1}^{\infty} \subset [0, \infty)\) satisfy
\[
\sup_{y \in Y} \|f_n(y)\|_X \leq M_n \forall n \in \mathbb{N} \text{ and } \sum_{n=1}^{\infty} M_n < \infty.
\]
Then \(S_N(y) := \sum_{n=1}^{N} f_n(y)\) converges absolutely and uniformly to \(S(y) = \sum_{n=1}^{\infty} f_n(y)\) and the function \(S : Y \to X\) is 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,
\[
\|S(y) - S_N(y)\| = \lim_{M \to \infty} \|S_M(y) - S_N(y)\| = \lim_{M \to \infty} \left\| \sum_{n=N+1}^{M} f_n(y) \right\| \leq \liminf_{M \to \infty} \sum_{n=N+1}^{M} \|f_n(y)\| \leq \sum_{n=N+1}^{\infty} M_n.
\]
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 7.7. The continuity of \(S\) now follows form the continuity of \(S_N\), the uniform convergence just proved, and Theorem 6.43.

Theorem 7.11 (Root test). Suppose that \(\{a_n\}_{n=1}^{\infty} \subset \mathbb{C}\) 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 \(\beta \in (\alpha, 1)\), 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
\[
\sum_{n=1}^{\infty} \beta^n = \frac{\beta}{1 - \beta} < \infty.
\]
Exercise 7.5. For every $p \in \mathbb{N}$, show $\sum_{n=0}^{\infty} (n/n)^p \cdot z^n$ is convergent iff $z = 0$.

Theorem 7.12 (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 $\alpha := \lim \sup_{n \to \infty} |a_{n+1}/a_n| < 1$ then $\sum_{n=1}^{\infty} |a_n| < \infty$ and $\sum_{n=1}^{\infty} a_n$ is absolutely convergent.
2. If $|a_{n+1}/a_n| \geq 1$ for a.a. $n$ then $\lim_{n \to \infty} |a_n| > 0$ and $\sum_{n=1}^{\infty} a_n$ diverges.
3. If $\lim \inf_{n \to \infty} \left| a_{n+1}/a_n \right| > 1$ then $\lim_{n \to \infty} |a_n| = \infty$ and $\sum_{n=1}^{\infty} a_n$ diverges.
4. If $\lim \sup_{n \to \infty} \left| a_{n+1}/a_n \right| = 1$, the test fails, i.e. you must work harder!

Proof. We take each item in turn.

1. If $\alpha < 1$, let $\beta \in (\alpha, 1)$, then $|a_{n+1}/a_n| \leq \beta$ 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,

$$|a_n| \leq |a_N| \beta^{n-N} = \beta^{-N} |a_N| \beta^n$$

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 $|a_{n+1}/a_n| \geq 1$ 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 $\lim \inf_{n \to \infty} \left| a_{n+1}/a_n \right| > 1$, then $|a_{n+1}/a_n| \geq \beta$ 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 Example 7.6 we know that $\sum_{n=1}^{\infty} 1/n = \infty$ while $\sum_{n=1}^{\infty} 1/n^2 < \infty$. However,

$$\lim_{n \to \infty} \left( \frac{1}{n+1} \right) = 1 = \lim_{n \to \infty} \left( \frac{1}{n+1} \right)^2$$

which shows the test has failed.

Exercise 7.6. Let $z \in \mathbb{C}$. Show the following series are absolutely convergent;

1. $e^z := \sum_{n=0}^{\infty} \frac{z^n}{n!}$, 
2. $\sin (z) := \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!}$, and
3. $\cos (z) := \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!}$.

The reader should verify Euler's formula

$$e^{iz} = \cos (z) + i \sin (z).$$

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;

$$\cos (z) = \cos (z) \quad \text{and} \quad \cosh (z) = \cosh (z),$$

$$\sin (z) = -\sin (z) \quad \text{and} \quad \sinh (z) = -\sinh (z),$$

$$e^z = \cosh (z) + \sinh (z),$$

$$\sin (z) = \frac{e^{iz} - e^{-iz}}{2i} = \frac{1}{i} \sinh (iz),$$

and

$$\cos (z) = \frac{e^{iz} + e^{-iz}}{2} = \cosh (iz).$$

Theorem 7.13 (Hilbert Schmidt norm). Let $Z = \mathbb{C}^{\mathbb{N} \times \mathbb{N}}$ denote the space of $N \times N$ complex matrices, $A = (A_{ij})^{N}_{i,j=1}$ with $A_{ij} \in \mathbb{C}$. We let

$$\|A\|_2 := \left( \sum_{i,j=1}^{N} |A_{ij}|^2 \right)^{1/2}$$

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 \( A, B \in \mathbb{C} \).

3. \( \|A^n\|_2 \leq \|A\|_2^n \) 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 \((\mathbb{R}, \|\cdot\|_1)\).

5. If \((X, d)\) is a metric space and \( f : X \to \mathbb{Z} \) and \( g : X \to \mathbb{Z} \) are continuous functions then so is \( f \cdot g \) (order matters) and \( f + g \).

6. The functions \( f(A) = A^n \) is continuous on \( \mathbb{Z} \) for all \( n \in \mathbb{N}_0 \), where by convention \( A^0 = I \).

**Proof.** Item 1. is clear. From the definition of matrix multiplication and the Cauchy Schwarz inequality we find

\[
\left\| (AB)_{ij} \right\|^2 = \left\| \sum_{k=1}^{N} A_{ik} B_{kj} \right\|^2 \leq \sum_{k=1}^{N} \left| A_{ik} \right|^2 \cdot \sum_{k=1}^{N} \left| B_{kj} \right|^2.
\]

Therefore,

\[
\|AB\|_2^2 = \sum_{i,j=1}^{n} \left( (AB)_{ij} \right)^2 \leq \sum_{i,j=1}^{n} \left( \sum_{k=1}^{N} \left| A_{ik} \right|^2 \cdot \sum_{k=1}^{N} \left| B_{kj} \right|^2 \right) = \sum_{i,k=1}^{N} \left| A_{ik} \right|^2 \cdot \sum_{j,k=1}^{N} \left| B_{kj} \right|^2 = \|A\|_2^2 \cdot \|B\|_2^2
\]

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

\[
\|A_nB_n - AB\|_2 \leq \|(A + \delta A_n) (B + \delta B_n) - AB\|_2 = \|\delta A_nB + A\delta B_n + \delta A_n \delta B_n\|_2
\]

\[
\leq \|\delta A_n B\|_2 + \|A\delta B_n\|_2 + \|\delta A_n \delta B_n\|_2
\]

\[
\leq \|\delta A_n\|_2 \cdot \|B\|_2 + \|A\|_2 \cdot \|\delta B_n\|_2 + \|\delta A_n\|_2 \cdot \|\delta B_n\|_2 \to 0 \text{ as } n \to \infty.
\]

Similarly,

\[
\|A_n + B_n - (A + B)\|_2 \to 0 \text{ as } n \to \infty.
\]

Item 5. Suppose that \( \{x_n\}_{n=1}^{\infty} \subset X \) and \( x_n \to x \) as \( n \to \infty \), then using item 4.,

\[
\lim_{n \to \infty} (f \cdot g) (x_n) = \lim_{n \to \infty} [f (x_n) g (x_n)] = f (x) g (x) = (f \cdot g) (x)
\]

and

\[
\lim_{n \to \infty} (f + g) (x_n) = \lim_{n \to \infty} [f (x_n) + g (x_n)] = f (x) + g (x) = (f + g) (x).
\]

This shows \( f \cdot g \) and \( f + g \) are continuous.

Item 6. follows from item 5 with \( X = \mathbb{Z} \) along with an induction argument.

**Definition 7.14 (Matrix Functions).** If \( f(z) = \sum_{n=0}^{\infty} c_n z^n \) for some \( c_n \in \mathbb{C} \) and \( A \in \mathbb{C}^{J_N \times J_N} \), then we define

\[
f(A) := \sum_{n=0}^{\infty} c_n A^n
\]

provided the sum is convergent.

For example,

\[
sin (A) := \sum_{n=0}^{\infty} \frac{(-1)^{2n+1}}{n!} A^{2n+1}.
\]

**Theorem 7.15 (Matrix Exponentials).** The series for \( e^A \) is absolutely convergent defined a continuous function from \( \mathbb{C}^{J_N \times J_N} \) to \( \mathbb{C}^{J_N \times J_N} \) in the Hilbert Schmidt norm. In the same way one can prove analogous statements for \( \cos (A) \), \( \sinh (A) \), and \( \cosh (A) \). These result apply to the case where \( N = 1 \) in which case \( \mathbb{C}^{J_N \times J_N} = \mathbb{C} \).

**Proof.** We wish to apply the Weierstrass \( M \)-test of Theorem 7.10. Fix for the moment \( K < \infty \) and suppose that \( \|A\|_2 \leq K \). Let \( M_n := \frac{1}{n!} K^n \) for \( n \geq 1 \) and \( M_0 = \sqrt{N} \) then

\[
\sum_{n=0}^{\infty} M_n = e^K - 1 + \sqrt{N} < \infty
\]

and when \( \|A\|_2 \leq K \) we have

\[
\left\| \frac{1}{n!} A^n \right\|_2 = \frac{1}{n!} \|A^n\|_2 \leq \frac{1}{n!} \|A\|_2^n \leq M_n.
\]

Therefore we may apply Theorem 7.10 in order to conclude that \( e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n \) is convergent absolutely and uniformly on \( C_K := \{ A \in \mathbb{C} : \|A\|_2 \leq K \} \). The function, \( A \to e^A \), is therefore continuous on \( C_K \) for all \( K < \infty \). As \( K < \infty \) is arbitrary, it follows that \( Z \ni A \to e^A \in Z \) is continuous on all of \( Z \).

**Exercise 7.7 (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

$$\left\| (I - A)^{-1} \right\|_2 \leq \sum_{n=0}^{\infty} \|A^n\|_2 \leq \frac{\|A\|_2}{1 - \|A\|_2} + \sqrt{N}.$$
Topological Considerations

8.1 Closed and Open Sets

Let \((X, d)\) be a metric space.

**Definition 8.1.** 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 \(B(x, \delta) = B_x(\delta)\). We also define the closed ball centered at \(x \in X\) with radius \(\delta > 0\) as the set \(C_x(\delta) := \{y \in X : d(x, y) \leq \delta\}\).

**Definition 8.2.** 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 if \(V^c\) is closed. We will write \(F \subset X\) to indicate \(F\) is a closed subset of \(X\) and \(V \subset_o X\) to indicate the \(V\) 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 8.3.** If \(f : X \rightarrow \mathbb{R}\) 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 8.4 (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 8.5.** 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. \(\{z \in \mathbb{C} : \text{Im} z = 0 \text{ and } a \leq \text{Re} z \leq b\}\) for all \(a \leq b\) in \(\mathbb{R}\).

**Exercise 8.1.** Prove item 2. of Theorem 8.7. If \(\{C_\alpha\}_{\alpha \in I}\) is a collection of closed subsets of \(X\), then \(\bigcap_{\alpha \in I} C_\alpha\) is closed in \(X\).

**Exercise 8.2.** Given an example of a collection of closed subsets, \(\{A_n\}_{n=1}^{\infty}\), of \(\mathbb{C}\) such that \(\bigcup_{n=1}^{\infty} A_n\) is not closed.

**Corollary 8.8.** 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 \(\bigcup_{\alpha \in I} U_\alpha\) is open.
Therefore, take the infimum over a
\[ d_X, \in \mathbb{F} \]
the set \( A \)
Lemma 8.9.
For any non empty subset \( A \neq \emptyset \) of \( \mathbb{C} \) and \( S \neq \emptyset \) is open for all whenever \( S \) is a finite subset of \( \mathbb{C} \).
Exercise 8.7. Let \((X, d)\) be a complete metric space. Let \( A \subset X \) be a subset of \( X \) viewed as a metric using \( d|_{A \times A} \). Show that \((A, d|_{A \times A})\) is complete iff \( A \) is a closed subset of \( X \).
Lemma 8.9. For any non empty subset \( A \subset X \), let \( d_A(x) := \inf\{d(x,a)|a \in A\} \), then
\[ |d_A(x) - d_A(y)| \leq d(x, y) \quad \forall x, y \in X \] (8.1)
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. (8.1). If \( x_n \to x \) in \( X \), then by Eq. (8.1),
\[ |d_A(x) - d_A(x_n)| \leq d(x, x_n) \to 0 \quad \text{as} \quad n \to \infty \]
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
\[ d_A(x) = \lim_{n \to \infty} d_A(x_n) \geq \varepsilon \]
since \( d_A(x_n) \geq \varepsilon \) for all \( n \). This shows that \( x \in F_\varepsilon \) and hence \( F_\varepsilon \) is closed. \( \blacksquare \)

Definition 8.10. A subset \( A \subset X \) is a neighborhood of \( x \) if there exists an open set \( V \subset \mathbb{C} X \) 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 8.11. 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 6.41 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 8.5.

Lemma 8.12 (Approximating open sets from the inside by closed sets). Let \( A \) be a closed subset of \( X \) and \( F_\varepsilon := \{x \in X|d_A(x) \geq \varepsilon\} \subset X \) be as in Lemma 8.9. Then \( F_\varepsilon \uparrow A^c \) as \( \varepsilon \downarrow 0 \).
Proof. It is clear that \( d_A(x) = 0 \) for \( x \in A \) so that \( F_\varepsilon \subset A^c \) for each \( \varepsilon > 0 \) and hence \( \cup_{\varepsilon>0} F_\varepsilon \subset A^c \). Now suppose that \( x \in A^c \subset A \). By Exercise 8.5 there exists an \( \varepsilon > 0 \) such that \( B_{x}(\varepsilon) \subset A^c \), i.e. \( d(x, y) \geq \varepsilon \) for all \( y \in A \). Hence \( x \in F_\varepsilon \) and we have shown that \( A^c \subset \cup_{\varepsilon>0} F_\varepsilon \). Finally, it is clear that \( F_\varepsilon \subset F_{\varepsilon'} \) whenever \( \varepsilon' \leq \varepsilon \).

Definition 8.13. Given a set \( A \) contained in a metric space \( X \), let \( \bar{A} \subset X \) be the closure of \( A \) defined by
\[ \bar{A} := \{x \in X : \exists \{x_n\} \subset A \ni x = \lim_{n \to \infty} x_n\}. \]
That is to say \( \bar{A} \) contains all limit points of \( A \). We say \( A \) is dense in \( X \) if \( \bar{A} = X \), 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, \( D \).

Exercise 8.8. Given \( A \subset X \), show \( \bar{A} \) is a closed set and in fact
\[ \bar{A} = \cap\{F : A \subset F \subset X \text{ with } F \text{ closed}\}. \] (8.2)
That is to say \( \bar{A} \) is the smallest closed set containing \( A \).

Exercise 8.9. 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 8.10. 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 8.11. Let $Y = BC(\mathbb{R}, \mathbb{C})$ be the Banach space of continuous bounded complex 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 $f \in C(\mathbb{R}, \mathbb{C})$ such 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 $C_c(\mathbb{R}, \mathbb{C}) = C_0(\mathbb{R}, \mathbb{C})$.

8.2 Exercises

Exercise 8.12. Show that $(X, d)$ is a complete metric space if 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 a 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 a Cauchy sequence 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 8.13. 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 $f(a + b) \leq f(a) + f(b)$ for all $a, b \in [0, \infty)$.)

Exercise 8.14. 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

$$d(x, y) = \sum_{n=1}^\infty 2^{-n} \frac{d_n(x(n), y(n))}{1 + d_n(x(n), y(n))}.$$  \hfill (8.3)

Show:

1. $(X, d)$ is a metric space,
2. a sequence $\{x_k\}_{k=1}^\infty \subset X$ converges to $x \in X$ if $x_k(n) \to x(n) \in X_n$ as $k \to \infty$ for each $n \in \mathbb{N}$ and
3. $X$ is complete if $X_n$ is complete for all $n$.

8.3 Sequential Compactness

Suppose that $(X, d)$ and $(Y, \rho)$ are metric spaces.

Definition 8.14. 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 8.15. 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 $d(x, z_n) \geq n$. 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 compact sets must be bounded.

Example 8.16. 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 compact set must be closed.

Lemma 8.17 (Bolzano–Weierstrass property for $C^D$). Let $D \in \mathbb{N}$. Every bounded sequence, $\{z(n)\}_{n=1}^\infty \subset C^D$, 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 $z(n) = (z_1(n), \ldots, z_D(n)) \in C^D$. 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 $\lim_{k \to \infty} z(n_k) = z_i$ exists for $i = 1, 2$. We may continue this way inductively to find a subsequence such that $\lim_{k \to \infty} z(n_k) = z_i$ exists for all $1 \leq i \leq D$. It then follows that $\lim_{k \to \infty} \|z - z(n_k)\| = 0$ as desires where $z := (z_1, \ldots, z_D)$.

Theorem 8.18 (Bolzano–Weierstrass / Heine–Borel theorem). As subset $K \subset C^D$ is compact iff it is closed and bounded.

Proof. In light of Examples 8.15 and 8.16 we are left to show that closed and bounded subsets are compact. So let $K \subset C^D$ be a closed and bounded set and $\{z_n\}_{n=1}^\infty$ be any sequence in $K$. According to Lemma 8.17 $\{z_n\}_{n=1}^\infty$ has a convergent subsequence, $\{w_k := z_{n_k}\}_{k=1}^\infty$. Since $w_k \in K$ for all $k$ and $K$ is closed it necessarily follows that $\lim_{k \to \infty} w_k \in K$ which shows $K$ is compact.

Example 8.19 (Warning!). It is not true that a closed and bounded subset of an arbitrary metric space $(X, d)$ is necessarily compact. For example let $Z$ denote the vector space of continuous functions on $[0, 1]$ with values in $\mathbb{R}$ and for $f \in Z$ let $\|f\| = \sup_{t \in [0, 1]} |f(t)|$. Then the set $C := \{f_n\}_{n=0}^\infty$ where
Exercise 8.15 (Extreme value theorem). Let $K$ be compact subset of $X$ and $f : K \to \mathbb{R}$ be a continuous function. Show $-\infty < \inf_{x \in K} f(x) \leq \sup_{x \in K} f(x) < \infty$ and there exists $a, b \in K$ such that $f(a) = \inf_{x \in K} f(x)$ and $f(b) = \sup_{x \in K} f(x)$. 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 8.16 (Uniform Continuity). Let $K$ be 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: prove the contrapositive.

Exercise 8.17. If $(X, d)$ is a metric space and $K \subset X$ is compact. Show subset, $C \subset K$, which is closed is compact as well.

Exercise 8.18. If $K \subset \mathbb{R}$ is compact then $\sup(K) \in K$, i.e. $\sup(K) = \max(K)$.

Exercise 8.19. Let $(X, d)$ and $(Y, \rho)$ be metric spaces, $K \subset X$ be a compact set, and $f : K \to Y$ be a continuous function. Show $f(K)$ is compact in $Y$. In particular, for $C \subset K$ closed, we have $f(C)$ is closed and in fact compact in $Y$.

Exercise 8.20. 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 8.17. Give one or better yet two alternative proofs that $g$ is continuous based on compactness arguments.

Definition 8.20. Let $Z$ be a vector space. We say that two norms, $\| \cdot \|$ and $\| \cdot \| \,$ 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 8.21. Let $Z$ be a finite dimensional vector space. Then any two norms $| \cdot |$ and $\| \cdot \| \,$ 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

$$\left\| \sum_{i=1}^{n} a_i f_i \right\|_2 := \sqrt{\sum_{i=1}^{n} |a_i|^2 \text{ for } a_i \in \mathbb{F}}.$$ 

By the triangle inequality for the norm $| \cdot |$, we find

$$\left| \sum_{i=1}^{n} a_i f_i \right| \leq \sum_{i=1}^{n} |a_i| \left| f_i \right| \leq \sqrt{\sum_{i=1}^{n} |f_i|^2} \sum_{i=1}^{n} |a_i|^2 \leq M \left\| \sum_{i=1}^{n} a_i f_i \right\|_2,$$

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 $\| \cdot \|_2$. Since the normed space $(Z, \| \cdot \|_2)$ 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 compact subset of $Z$ relative to $\| \cdot \|_2$. Therefore by Exercise 8.15 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 $m < \frac{f}{\|f\|_2} \in S$ so that

$$m \leq \frac{|f|}{\|f\|_2} \leq \frac{1}{\|f\|_2} \|f\|_2$$

or equivalently

$$\|f\|_2 \leq \frac{1}{m} |f|.$$ 

This shows that $| \cdot |$ and $\| \cdot \|_2$ are equivalent norms. Similarly one shows that $| \cdot |$ and $\| \cdot \|_2$ are equivalent and hence so are $| \cdot |$ and $\| \cdot \|_2$.
Corollary 8.22. If $(Z, \|\cdot\|)$ is a finite dimensional normed space, then $A \subset Z$ is compact iff $A$ is closed and bounded relative to the given norm, $\|\cdot\|$.

Corollary 8.23. Every finite dimensional normed vector space $(Z, \|\cdot\|)$ 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 8.22. It is now routine to show $\lim_{n \to \infty} f_n = f := \lim_{k \to \infty} g_k$. □
Part III

Appendices
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 exits</td>
<td>∀</td>
</tr>
<tr>
<td>, or “space” then</td>
<td>∃</td>
<td>∀</td>
</tr>
<tr>
<td>∋</td>
<td>such that , or “space”</td>
<td>a.a.</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:
\[
\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall n \geq N, \ |L - a_n| \leq \varepsilon.
\]
This may also be written as:
\[
\forall \varepsilon > 0, \ |L - a_n| \leq \varepsilon \text{ for a.a. } n.
\]

3. The negation of the previous statement is \( \lim_{n \to \infty} a_n \neq L \) which translates to:
\[
\exists \varepsilon > 0 \forall N \in \mathbb{N}, \exists n \geq N \exists |L - a_n| > \varepsilon.
\]
This last statement is also equivalent to:
\[
\exists \varepsilon > 0 \exists |L - a_n| > \varepsilon \text{ i.o. } n.
\]
It is sometimes useful to reformulate this last statement as; there exists \( \varepsilon > 0 \) and an increasing function \( \mathbb{N} \ni k \to n_k \in \mathbb{N} \) such that
\[
|L - a_{n_k}| > \varepsilon \text{ for all } k \in \mathbb{N}.
\]
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

\[
X = \{ a = \{1\} \ b = \{1, 2\} \ c = \{3\} \ d = \{2, 4\} \ e = \{2\}\}
\]

ordered by set inclusion. Then \( b \) and \( d \) are maximal elements despite that fact that \( b \not\leq d \) and \( d \not\leq b \). 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,” \( x : A \rightarrow \prod_{\alpha \in A} X_\alpha \) 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[3].

Proof. (2 \( \Rightarrow \) 3) Let \( X \) be a partially ordered subset as in 3 and let \( F = \{ E \subset X : E \) 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 a 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 \( g : D(g) \rightarrow \prod_{\alpha \in A} X_\alpha \) 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 \( x_0 \in X_\alpha \) 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 on \( G \) as follows. We say that \( f \leq g \) for \( f, g \in G \) provided that \( D(f) \subset D(g) \) and \( f = g|_{D(f)} \). If \( \Phi \subset G \) is a linearly ordered set, let \( D(h) = \cup_{g \in \Phi} D(g) \) and for \( \alpha \in D(g) \) let \( h(\alpha) = g(\alpha) \). Then \( h \in G \) 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 \( E_0 \subset X \) inductively as follows. For \( n = 1 \) let \( E_1 = \{x_1\} \), and if \( E_n \) have 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 \( E = \cup_{n=1}^\infty E_n \) 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 \( y = x_m \in E \subset E \). 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 “jazzed” up version of induction.)

2 Similarly one may show that 3 \( \Rightarrow \) 2. Let \( F = \{E \subset X : E \) is linearly ordered\} and order \( F \) by inclusion. If \( M \subset F \) is linearly ordered, let \( E = \cup M = \bigcup_{A \in M} A \). If \( x, y \in E \) then \( x \in A \text{ and } y \in B \) for some \( A, B \subset M \). Now \( M \) is linearly ordered by set inclusion so \( A \subset B \text{ or } B \subset A \) i.e. \( x, y \in A \text{ or } x, y \in B \). 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

\[
\tilde{h}(\alpha) = \begin{cases} h(\alpha) & \text{if } \alpha \in D(h) \\ x_0 & \text{if } \alpha = \alpha_0. \end{cases}
\]

Then \( h \leq \tilde{h} \) while \( h \neq \tilde{h} \) 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 \( P_0 \not\subset P_1 \). In particular we may choose \( x \notin P_0 \) 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 without loss of generality that \( \mathcal{F} \neq \emptyset \). Notice that \( \mathcal{F} \neq \emptyset \) if \( \mathcal{F} \) 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 \( g \).

Lemma B.7. The function \( g : \mathcal{F} \to \mathcal{F} \) defined in Eq. (B.1) has a fixed point.\(^3\)

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. Then take \( P_\infty = \bigcup_{n=0}^\infty P_n \) and start again with \( P_0 \) replaced by \( P_\infty \). Then keeping 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 \( \mathcal{F}_1 = \{P \in \mathcal{F} : P_0 \subset P\} \). 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 \( P \cap g(P) \in \mathcal{F}_0 \) 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 \( \Phi_C := \{A \in \mathcal{F}_0 : \text{ either } A \subset C \text{ or } g(C) \subset A\} \).

We will begin by showing that \( \Phi_C \subset \mathcal{F}_0 \) is a tower and therefore that \( \mathcal{F}_0 = \mathcal{F}_0 \).

1. \( P_0 \in \Phi_C \) since \( P_0 \subset C \) for all \( C \in \Gamma \subset \mathcal{F}_0 \). 2. If \( \Phi \subset \Phi_C \subset \mathcal{F}_0 \) is totally ordered by set inclusion, then \( A_\Phi := \cup \Phi \in \mathcal{F}_0 \). We must show \( A_\Phi \in \Phi_C \), that is that \( A_\Phi \subset C \) or \( C \subset A_\Phi \). Now if \( A \subset C \) for all \( A \in \Phi \), then \( A_\Phi \subset C \) and hence \( A_\Phi \in \Phi_C \). On the other hand if there is some \( A \in \Phi \) such that \( g(C) \subset A \) then clearly \( g(C) \subset A_\Phi \) and again \( A_\Phi \in \Phi_C \). 3. Given \( A \in \Phi_C \) we must show \( g(A) \in \Phi_C \), i.e. that

\[
g(A) \subset C \text{ or } g(C) \subset g(A).
\]

(B.2)

There are three cases to consider: either \( A \not\subset C \), \( A = C \), 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 \( A \not\subset C \). 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

\[
A \not\subset C \subset g(A).
\]

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 \( \Phi_C = \mathcal{F}_0 \) to conclude for all \( A \in \mathcal{F}_0 \), 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 \( u \preceq v \) for \( u,v \in \mathcal{F} \) provided \( D(u) \subset D(v) \) and \( u = v|_{D(u)} \). One now checks that \( (\mathcal{F},\preceq) \) 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 \( g := U^{-1} : Y \to D(U) \subset X \).

### 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 exist 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 \( n \).
B.3 Formalities of Counting

**Definition B.9.** We say card \((X) \leq \text{card}(Y)\) if there exists an injective map, \(f : X \to Y\) and card \((Y) \geq \text{card}(X)\) if there exists a surjective map \(g : Y \to X\). We say \(\text{card}(X) = \text{card}(Y)\) if there exists bijections, \(f : X \to Y\).

**Proposition B.10.** We have \(\text{card}(X) \leq \text{card}(Y)\) 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 \(g/\text{range}(f) = x_0 \in X\) chosen arbitrarily. Then \(g : Y \to X\) is surjective.

If \(g : Y \to X\) is a surjective map, then \(x_j := g^{-1}(\{x\}) \neq \emptyset\) 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 B.11 (Schröder-Bernstein Theorem).** 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.** Starting with an \(x \in X\) we may form the sequence of \(\text{"ancestors"}\) of \(x\), namely ancestor \((x) := (x, y_1, x_1, y_2, \ldots)\) where \(y_1 = g^{-1}(x)\), \(x_1 = f^{-1}(y_1)\), \(y_2 = g^{-1}(x_1)\), \ldots.

\[
x \xrightarrow{g^{-1}} y_1 \xrightarrow{f^{-1}} x_1 \xrightarrow{g^{-1}} y_2 \xrightarrow{f^{-1}} \ldots
\]

We continue this process of inverse iterates as long as it is possible, i.e. we can construct \(y_{n+1}\) if \(x_n \in Y(\text{and } x_{n+1} \in f(X))\). There are now three possibilities;

1. ancestor \((x)\) has infinite length so the process never gets stuck in which case we say \(x \in X_\infty\), 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 end 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, \(X_\infty, X_X, \text{and } X_Y\). Similarly we may partition \(Y\) into \(Y_\infty, Y_Y, \text{and } Y_Y\). Let us now observe that,

1. \(f(X_\infty) = Y_\infty\). Indeed if \(x \in X_\infty\) then ancestor \((f(x)) = (x, \text{ancestor}(x))\) is an infinite sequence, i.e. \(f(x) \in Y_\infty\). Moreover if \(y \in Y_\infty\), then ancestor \((y) = (y, \text{ancestor}(x))\) where \(f(x) = y\) so that \(x \in X_\infty\) and \(y \in f(X_\infty)\). Thus we have shown \(f : X_\infty \to Y_\infty\) is a bijection, i.e. \(\text{card}(X_\infty) = \text{card}(Y_\infty)\).
2. \(f(X_X) = Y_Y\). Indeed if \(x \in X_X\) then again ancestor \((f(x)) = (x, \text{ancestor}(x))\) which ends in \(X\) so that \(f(x) \in Y_Y\). Moreover if \(y \in Y_Y\), then ancestor \((y) = (y, \text{ancestor}(x))\) where \(f(x) = y\) so that \(x \in X_X\) ends in \(X\), i.e. \(y \in f(X_X)\). Thus we have shown \(f : X_X \to Y_Y\) is a bijection, i.e. \(\text{card}(X_X) = \text{card}(Y_Y)\).
3. By the same argument as in item 2. it follow that \(g : Y_Y \to X_X\) is a bijection, i.e. \(\text{card}(X_X) = \text{card}(Y_Y)\).

The last three statement implies \(\text{card}(X) = \text{card}(Y)\). We may in fact define a bijection, \(h : X \to Y\), by

\[
h(x) = \begin{cases} f(x) & \text{if } x \in X_\infty \cup X_X \\ g^{-1}(x) & \text{if } x \in X_Y \end{cases}
\]

**Definition B.12.** We say \(\text{card}(X) < \text{card}(Y)\) if \(\text{card}(X) \leq \text{card}(Y)\) and \(\text{card}(X) \neq \text{card}(Y)\), i.e. \(\text{card}(X) < \text{card}(Y)\) if there exists an injective map, \(f : X \to Y\), but not bijective map exists. Similarly we say \(\text{card}(Y) > \text{card}(X)\) if \(\text{card}(Y) \geq \text{card}(X)\) and \(\text{card}(Y) \neq \text{card}(X)\), i.e. \(\text{card}(Y) > \text{card}(X)\) if there exists a surjective map \(g : Y \to X\) but no bijective map exists.

**Proposition B.13.** For any non-empty set \(X\), \(\text{card}(X) < \text{card}(2^X)\).

**Proof.** Define \(f : X \to 2^X\) by \(f(x) = \{x\}\). Then \(f\) is an injective map and hence \(\text{card}(X) \leq \text{card}(2^X)\). Now suppose that \(g : X \to 2^X\) is any map. Let \(X_0 = \{x \in X : x \notin g(x)\}\). I claim that \(X_0 \notin g(X)\).

Indeed suppose there exists \(x_0 \in X\) such that \(g(x_0) = X_0\). If \(x_0 \in X_0\), then \(x_0 \notin g(x_0) = X_0\) which is impossible. Similarly if \(x_0 \notin X_0 = g(x_0)\) then \(x_0 \in X_0\) and again we have reached a contradiction. Thus we must conclude that \(X_0 \notin g(X)\). Thus there are no surjective maps, \(g : X \to 2^X\) so that \(\text{card}(X) \neq \text{card}(2^X)\).

**Proposition B.14.** If \(\text{card}(X) < \text{card}(Y)\) and \(\text{card}(Y) \leq \text{card}(Z)\), then \(\text{card}(X) < \text{card}(Z)\).

**Proof.** If there exists an injective map, \(f : Z \to X\) then composing this with and injective map, \(g : X \to Y\) gives an injective map, \(g \circ f : Z \to X\) and there for \(\text{card}(Z) \leq \text{card}(X)\). But this would imply that \(\text{card}(X) = \text{card}(Z)\).

**Definition B.15.** Let \(A_n := \{1, 2, \ldots, n\}\) for all \(n \in \mathbb{N}\) and write \(n\) for \(\text{card}(A_n)\).

**Proposition B.16.** We have \(\text{card}(A_m) < \text{card}(A_n)\) for all \(m < n\). Moreover if \(\emptyset \neq X \subseteq A_n\) then \(\text{card}(X) = \text{card}(A_k)\) 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 an 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 \( f(A_n) \) is a proper subset of \( A_k \) and therefore \( \text{card} (f(A_n)) < \text{card} (A_k) \) but on the other hand \( \text{card} (f(A_n)) = \text{card} (A_n) \geq \text{card} (A_k) \) which is a contradiction. So no such bijection can exists and we have shown \( \text{card} (A_k) < \text{card} (A_{n+1}) \) 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 \( X' := X \setminus \{ n + 1 \} \neq A_n \). 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 previous 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 \( f_n : J_n \to X \) is injective. If \( f_n \) were bijective we would have \( \text{card} (J_n) = \text{card} (X) \) and in particular \( \text{card} (J_n) > \text{card} (J_n) = \text{card} (X) \) 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 \( f_n = f_n | J_m \) for all \( m \leq n \). We then define \( f(m) := f_n (m) \) where \( n \in \mathbb{N} \) is an integer such that \( n \geq m \). In this way we construct a function, \( f : N \to X \) 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 $\operatorname{card}(\psi(Y)) = \operatorname{card}(J_k)$ for some $k \leq N$. On the other hand, $f_n : J_n \to Y$ being injective implies $\operatorname{card}(\psi(Y)) \geq \operatorname{card}(J_n)$ 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 $\operatorname{card}(Y) = \operatorname{card}(\psi(Y)) = \operatorname{card}(\mathbb{N})$. From this it follows that $\operatorname{card}(\mathbb{N}) \leq \operatorname{card}(X)$.

**Alternate Proof.** By assumption, there exists and 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 $\operatorname{card} J_n \leq \operatorname{card} X = \operatorname{card} J_k$ would imply $n \leq k$. \[\square\]
References

Index

Cauchy, \[ \text{[18]} \]
Cauchy sequence
in a metric space, \[ \text{[16]} \]
Complete

Metric space, \[ \text{[46]} \]
Function
continuous, \[ \text{[49]} \]
continuous at a point, \[ \text{[49]} \]