Real Analysis

Hanan Ather

Fall 2023

Contents

⋆ These notes were created during my review process to aid my own understanding and not written for the purpose of instruction. I originally wrote them only for myself, and they may contain typos and errors[a]. *No professor has verified or confirmed the accuracy of these notes.* With that said, I've decided to share these notes on the off chance they are helpful to anyone else.

[a]Any corrections are greatly appreciated.

§1 September 7, 2023

§1.1 Metric spaces

Definition 1.1 (Metric Space) A metric space is a nonempty set X together with a mapping $d : X \times X \to \mathbb{R}^+ = [0, \infty)$ with the properties:

1. $d(x, y) = 0$ if and only if $x = y \quad (x, y \in X)$
2. $d(x, y) = d(y, x)$ for all $x, y \in X$
3. $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in X$

Example 1.2 (Natural Metric)**.** Let X be any nonempty set of real numbers and define d by

$$d(x, y) = |x - y|, \quad x, y \in X.$$

This is the usual definition of distance between two points.

Proof. To prove that this is a metric, we need to show that $d(x, y)$ satisfies the three metric properties (M1, M2, and M3).

1. **Non-negativity and Identity of Indiscernibles (M1):**
   We have $d(x, y) = |x - y| \geq 0$ for all $x, y \in X$. Furthermore, $d(x, y) = 0$ if and only if $x = y$.
2. **Symmetry (M2):**
   $d(x, y) = |x - y| = |y - x| = d(y, x)$ for all $x, y \in X$.
3. **Triangle Inequality (M3):**
   $d(x, z) = |x - z|$
   $\leq |x - y| + |y - z|$
   $= d(x, y) + d(y, z)$ for all $x, y, z \in X$.

□

Example 1.3 (Distance in $\mathbb{R}^2$)**.** Let X be any nonempty set of points in the plane (so X may be considered as a subset of $\mathbb{R}^2$) and define d by

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2},$$

where $x = (x_1, x_2)$ and $y = (y_1, y_2)$ are any two points of X.

Proof. To prove that $d(x, y)$ is a metric, we need to verify that it satisfies the three metric properties (M1, M2, and M3).

1. **Non-negativity and Identity of Indiscernibles (M1):**
   Note that $(x_1 - y_1)^2 \geq 0$ and $(x_2 - y_2)^2 \geq 0$ for all $x, y \in X$. Thus,

   $$d(x, y) = (x_1 - y_1)^2 + (x_2 - y_2)^2 \geq 0.$$

   Furthermore, $d(x, y) = 0$ if and only if $x_1 = y_1$ and $x_2 = y_2$, which implies $x = y$.

2. **Symmetry (M2):** is obvious.

3. **Triangle Inequality (M3):**

□

Definition 1.4 In a metric space (X, d), an **open ball** of radius r centered at a point a is defined as the set of all points x in X such that $d(x, a) < r$. Formally, the open ball $B(a, r)$ is given by

$$B(a, r) = \{x \in X \mid d(x, a) < r\}.$$

Definition 1.5 In a metric space (X, d), an **closed ball** of radius r centered at a point a is defined as the set of all points x in X such that $d(x, a) \leq r$.

$$\overline{B(a)_r} = \{x \in X \mid d(x, a) \leq r\}.$$

Example 1.6. We can define an alternative metric d' on the same set $X = \mathbb{R}^2$ as follows:

$$d'(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|\}.$$

This metric d' has a different interpretation compared to the Euclidean metric. In d', the distance between two points x and y is defined as the maximum of the absolute differences of their respective coordinates. This is often referred to as the *Chebyshev distance* or L_∞ *metric*, and it essentially measures the "greatest single axis" distance between two points.

Proof.

(M1): $d'(x, y) = 0 \iff |x_1 - y_1| = 0$ and $|x_2 - y_2| = 0 \iff x_1 = y_1$ and $x_2 - y_2 = 0$

(M2): Obvious.

□

Example 1.7 (Euclidean metric)**.** Let X be any nonempty subset of $\mathbb{R}^n$, meaning that X consists of ordered n-tuples of real numbers. We define the distance d between any two points x and y in X by

$$d(x, y) = \sqrt{\sum_{k=1}^{n} (x_k - y_k)^2},$$

where $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$.

⋆ This mapping d is commonly known as the *Euclidean metric* for such a set X. Henceforth, when we refer to the metric space $\mathbb{R}^n$ (as opposed to merely the set $\mathbb{R}^n$), we imply that the metric space (X, d) of this example is being considered with $X = \mathbb{R}^n$. In other words, any reference to the metric space $\mathbb{R}^n$ will always imply that the Euclidean metric is being used.

The term *Euclidean space* is often used synonymously with the metric space $\mathbb{R}^n$.

Theorem 1.8 (Cauchy-Schwarz Inequality) — For any points $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n)$ in $\mathbb{R}^n$, the following inequality holds:

$$\left(\sum_{k=1}^{n} a_k b_k\right)^2 \leq \left(\sum_{k=1}^{n} a_k^2\right)\left(\sum_{k=1}^{n} b_k^2\right)$$

Proof. To prove the Cauchy–Schwarz Inequality, we introduce the function $\psi(u)$ defined by

$$\psi(u) = \sum_{k=1}^{n} (a_k u + b_k)^2 .$$

It is evident that $\psi(u)$ is a quadratic form in u, having the general form $Au^2 + 2Bu + C$.

Being a sum of squares, $\psi(u)$ is non-negative for all u, i.e., $\psi(u) \geq 0$. Hence, the discriminant of this quadratic form, $(2B)^2 - 4AC$, cannot be positive.

Dividing the discriminant by 4, we have

$$\frac{(2B)^2}{4} - AC = B^2 - AC \leq 0.$$

This can be rewritten as

$$\left(\sum_{k=1}^{n} a_k b_k\right)^2 - \left(\sum_{k=1}^{n} a_k^2\right)\left(\sum_{k=1}^{n} b_k^2\right) \leq 0,$$

which proves the Cauchy–Schwarz Inequality.

We can use this inequality as a foundation for other inequalities and results in real analysis and vector spaces.

□

Theorem 1.9 — For any points $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n)$ in $\mathbb{R}^n$, we have the following inequality:

$$\left(\sum_{k=1}^{n} (a_k + b_k)^2\right) \leq \left(\sum_{k=1}^{n} a_k^2\right) + \left(\sum_{k=1}^{n} b_k^2\right).$$

Proof. Taking the square roots of both sides of the Cauchy–Schwarz inequality gives:

$$\sqrt{\left(\sum_{k=1}^{n} a_k^2\right)\left(\sum_{k=1}^{n} b_k^2\right)} \geq \left|\sum_{k=1}^{n} a_k b_k\right|.$$

□

Example 1.10 (l^p-metric on $\mathbb{R}^n$)**.** The L^p metric, also known as the L^p distance, is defined for $p \geq 1$ and provides a generalized notion of distance between points in $\mathbb{R}^n$. For two points $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$, the L^p distance $d_p(x, y)$ is defined as:

$$d_p(x, y) = \left(\sum_{k=1}^{n} |x_k - y_k|^p\right)^{1/p}.$$

Proof. The verification (M3) is difficult requires the *Holder inequality.* □

Example 1.11 (l^∞–metric on $\mathbb{R}^n$)**.** The L^∞ metric, also known as the Chebyshev distance or infinity norm, is defined for points $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ in $\mathbb{R}^n$ as:

$$d_\infty(x, y) = \max_{1 \leq k \leq n} |x_k - y_k|.$$

Definition 1.12 (l_2 space) The ℓ_2 space is a vector space consisting of all sequences $x = (x_1, x_2, x_3, \ldots)$ of real or complex numbers for which the ℓ_2 norm is finite. The ℓ_2 norm $\|x\|_2$ is defined as:

$$\|x\|_2 = \left(\sum_{n=1}^{\infty} |x_n|^2 \right)^{1/2} < \infty.$$

A sequence is in ℓ_2 if it is "square-summable," meaning that the sum of the squares of its elements is finite.
A natural metric on the ℓ_2 space can be defined using the ℓ_2 norm. Given two sequences $x = (x_1, x_2, x_3, \ldots)$ and $y = (y_1, y_2, y_3, \ldots)$ in ℓ_2, the distance $d(x, y)$ between x and y is defined as:

$$d(x, y) = \|x - y\|_2 = \left(\sum_{n=1}^{\infty} |x_n - y_n|^2 \right)^{1/2}.$$

This metric is well-defined because the sequences x and y are both in ℓ_2, making their ℓ_2 norms and the distance $d(x, y)$ finite.

Definition 1.13 (l_p space) The ℓ^p space is a generalization of the ℓ^2 space and consists of all sequences $x = (x_1, x_2, x_3, \ldots)$ of real or complex numbers for which the ℓ^p norm is finite. The ℓ^p norm $\|x\|_p$ is defined as:

$$\|x\|_p = \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{1/p} < \infty,$$

where $1 \leq p < \infty$. A metric d_p on the ℓ^p space can be naturally defined using the ℓ^p norm. For two sequences x and y in ℓ^p, the distance $d_p(x, y)$ is:

$$d_p(x, y) = \left(\sum_{n=1}^{\infty} |x_n - y_n|^p \right)^{1/p}.$$

Example 1.14. For the space $C[a, b]$, consisting of all continuous functions defined on the interval $[a, b]$, the uniform metric d is defined as follows. Given two functions f and g in $C[a, b]$, the distance $d(f, g)$ between f and g is:

$$d(f, g) = \sup_{x \in [a,b]} |f(x) - g(x)|.$$

Here, sup denotes the supremum, or the least upper bound, of the set $\{|f(x) - g(x)| : x \in [a, b]\}$.

Example 1.15. The $C^1[a, b]$ space consists of all functions that are continuously differentiable on the closed interval $[a, b]$. A common metric d used in $C^1[a, b]$ is the C^1 metric, defined as follows. Given two functions f and g in $C^1[a, b]$, the distance $d(f, g)$ between f and g is:

$$d(f, g) = \sup_{x \in [a,b]} |f(x) - g(x)| + \sup_{x \in [a,b]} |f'(x) - g'(x)|.$$

This metric sums the uniform distance between the functions f and g and the uniform distance between their first derivatives f' and g'.

§2 September 11, 2023

§2.1 Examples of metric spaces

Another metric for the set $\mathbb{R}^n$ is the mapping d_1, where

$$d_1(x, y) = \sum_{k=1}^{n} |x_k - y_k|.$$

This also reduces to the metric of Example (1) when $n = 1$.

Both the Euclidean metric and the metric d_1 just defined are special cases of the metric d_p, where

$$d_p(x, y) = \left(\sum_{k=1}^{n} |x_k - y_k|^p \right)^{\frac{1}{p}},$$

with $p \geq 1$. The verification of (M3) for this mapping for general values of p requires a discussion of the Hölder inequality and the Minkowski inequality.

(6) A third metric for the set $\mathbb{R}^n$ is given by the mapping d_∞, where

$$d_\infty(x, y) = \max_{1 \leq k \leq n} |x_k - y_k|.$$

When $n = 1$, we again obtain the metric of Example (1), while when $n = 2$ we obtain that of Example (3). The method of Example (3) is used in showing that d_∞ is a metric.

§2.2 Hölder and Minkowski inequalities

Example 2.1. Let $X = \mathbb{R}^n$ and $d_p(x, y) = (\sum_{k=1}^{n} |x_k - y_k|^p)^{\frac{1}{p}}$ for $p \geq 1$.

To prove (M3), we use:

Hölder inequality

$$\left| \sum_{k=1}^{n} a_k b_k \right| \leq \left(\sum_{k=1}^{n} |a_k|^p \right)^{\frac{1}{p}} \cdot \left(\sum_{k=1}^{n} |b_k|^q \right)^{\frac{1}{q}}$$

where $\frac{1}{p} + \frac{1}{q} = 1$ (extension of Cauchy-Schwarz inequation).

Goal: Prove inequality:

$$\left(\sum_{k=1}^{n} |a_k + b_k|^p \right)^{\frac{1}{p}} \leq \left(\sum_{k=1}^{n} |a_k|^p \right)^{\frac{1}{p}} + \left(\sum_{k=1}^{n} |b_k|^p \right)^{\frac{1}{p}} \quad \text{(Minkowski)}$$

Then (M3) follows taking $a_k = x_k - y_k$ and $b_k = y_k - z_k$.

Exercise: Check this. (Hint: Look at case $p = 2$, last time)

Proof. **Proof of (1)**:

$$\left| \sum_{k=1}^{n} a_k b_k \right|^p = |a_k b_k| \, |a_k + b_k|^{p-1} \leq (|a_k| + |b_k|) \, |a_k + b_k|^{p-1}$$

$$\leq \left(\sum_{k=1}^{n} |a_k|^p \right)^{\frac{1}{p}} \cdot \left(\sum_{k=1}^{n} |a_k + b_k|^{(p-1)q} \right)^{\frac{1}{q}} \quad \text{by Hölder's inequality with } \frac{1}{p} + \frac{1}{q} = 1$$

Taking the sum of the previous two inequalities, we get:

$$\sum_{k=1}^{n} |a_k b_k|^p \leq \left(\sum_{k=1}^{n} |a_k|^p \right)^{\frac{1}{p}} + \left(\sum_{k=1}^{n} |b_k|^p \right)^{\frac{1}{p}}$$

Also, we know that:

$$\sum_{k=1}^{n}|a_k+b_k|^p \le \sum_{k=1}^{n}(|a_k|+|b_k|)|a_k+b_k|^{p-1}$$

Hence, combining the last two inequalities:

$$\sum_{k=1}^{n}|a_k+b_k|^p \le \left(\sum_{k=1}^{n}|a_k+b_k|^p\right)^{\frac{1}{p}} \cdot \left(\sum_{k=1}^{n}(|a_k|^p+|b_k|^p)\right)^{\frac{1}{p}}$$

$$\le \left(\sum_{k=1}^{n}|a_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{n}|b_k|^p\right)^{\frac{1}{p}} \quad \text{by Minkowski's inequality}$$

□

§2.3 ℓ_2-spaces

Definition 2.2 Let ℓ_2 be the set of sequences $x = \{x_k\}_{k\ge1}$ with $x_k \in \mathbb{C}$ for all $k \ge 1$ such that $\sum_{k=1}^{\infty}|x_k|^2 < \infty$. We define

$$d(x,y) = \sqrt{\sum_{k=1}^{\infty}|x_k-y_k|^2}$$

where $x = \{x_k\}_{k\ge1}$ and $y = \{y_k\}_{k\ge1}$.

Lemma 2.3 — We justify that it makes sense to define $\sqrt{\sum_{k=1}^{\infty}|x_k-y_k|^2}$ as a metric by proving:

a) $d(x,y) < \infty$ for all $x, y \in \ell_2$.

b) $d(x,z) \le d(x,y) + d(y,z)$ for all $x, y, z \in \ell_2$.

Proof. a) If $|x_k - y_k| \le |x_k| + |y_k|$, then $|x_k-y_k|^2 \le (|x_k|+|y_k|)^2$ for any $k = 1, \ldots, n$, for any $n \ge 1$ fixed. We take the sum for all $k = 1, \ldots, n$ and then square root.

$$\sqrt{\sum_{k=1}^{n}|x_k-y_k|^2} \le \sqrt{\sum_{k=1}^{n}(|x_k|+|y_k|)^2}$$

$$\le \sqrt{\sum_{k=1}^{n}|x_k|^2} + \sqrt{\sum_{k=1}^{n}|y_k|^2}$$

for any $n \ge 1$ by Minkowski's inequality (p=2).

Hence $\sum_{k=1}^{n}|x_k-y_k|^2 \le M^2 < \infty$ for all $n \ge 1$, and $\sum_{k=1}^{\infty}|x_k-y_k|^2 \le M^2$.

So $d(x,y) = \sqrt{\sum_{k=1}^{\infty}|x_k-y_k|^2} \le M < \infty$.

b) $|x_k - z_k| = |x_k - y_k + y_k - z_k|$. Take the square $|x_k-z_k|^2 \le (|x_k-y_k|+|y_k-z_k|)^2$. Take the sum for $k = 1, \ldots, n$, then the square root. We get:

$$\sqrt{\sum_{k=1}^{n}|x_k-z_k|^2} \le \sqrt{\sum_{k=1}^{n}(|x_k-y_k|+|y_k-z_k|)^2}$$

$$\leq \sqrt{\sum_{k=1}^{n} |x_k - y_k|^2} + \sqrt{\sum_{k=1}^{n} |y_k - z_k|^2}$$

for any $n \geq 1$ by Minkowski's inequality (p=2).

Then, taking $n \to \infty$ we get:

$$d(x, z) \leq d(x, y) + d(y, z)$$

□

Remark 2.4. Let ℓ_p be the set of all sequences $x = \{x_k\}_{k\geq 1}$ with $x_k \in \mathbb{C}$ such that $\sum_{k=1}^{\infty} |x_k|^p < \infty$. If $p \geq 1$, we define

$$d(x, y) = \left(\sum_{k=1}^{\infty} |x_k - y_k|^p \right)^{\frac{1}{p}}$$

where $x = \{x_k\}_{k\geq 1}$ and $y = \{y_k\}_{k\geq 1}$. Then d is a distance on ℓ_p (Exercise).

Remark 2.5. If $p \in (0, 1)$, we define

$$d(x, y) = \sum_{k=1}^{\infty} |x_k - y_k|^p$$

This is also a distance. (M3) is verified using another inequality:

$$|a + b|^p \leq |a|^p + |b|^p$$

for any $a, b \in \mathbb{R}$ (sub-additivity) if $p \in (0, 1)$.

Replacement for Minkowski inequality is:

$$\sum_{k=1}^{n} |a_k + b_k|^p \leq \sum_{k=1}^{n} |a_k|^p + \sum_{k=1}^{n} |b_k|^p$$

Remark 2.6. Let ℓ_∞ be the space of sequences $x = \{x_k\}_{k\geq 2}$, with $x_k \in \mathbb{C}$, such that $\{x_k\}$ is bounded, i.e., $\exists M > 0$ such that $|x_k| \leq M$ for all $k \geq 1$. We define

$$d(x, y) = \sup_{k\geq 1} |x_k - y_k|$$

Exercise: check that d is a distance on ℓ_∞.

Example 2.7. Let $X = \mathbb{C}$. Consider the distance function $d(x, y) = \frac{|x-y|}{\sqrt{(1+|x|^2)(1+|y|^2)}}$, which is called the chordal distance.

Proof. The proof that d is a distance is omitted. □

Example 2.8. Let $X = C[a, b]$ where $C[a, b] = \{f : [a, b] \to \mathbb{R} \mid f \text{ is continuous}\}$. Recall (MAT2125) that a function $f : [a, b] \to \mathbb{R}$ is continuous at $x_0 \in [a, b]$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that for all $x \in (x_0 - \delta, x_0 + \delta)$ we have $|f(x) - f(x_0)| < \varepsilon$, i.e., $|x - x_0| < \delta$. f is continuous on $[a, b]$ if f is continuous at every $x \in [a, b]$.

Observation: If f is continuous at x_0 then $\lim_{x\to x_0} f(x) = f(x_0)$.

Let $d(f, g) = \max_{t\in[a,b]} |f(t) - g(t)|$. Note: $|x - y|$ is a continuous function on $[a, b]$, and $[a, b]$ is compact, hence the maximum is attained, and d is well-defined. d is called the uniform metric.

§3 September 14, 2023

§3.1 Examples of Metric Spaces - Continued

Example 3.1. Let $X = C[a,b]$ and define $d(f,g) = \max_{t\in[a,b]} |f(t) - g(t)|$.

Proof. We will prove that this is indeed a metric:

(M1) $x = y \Rightarrow d(x,y) = 0$.
Assume $d(x,y) = 0$. We want to show that $x = y$. Assume there exists $t_0 \in [a,b]$ such that $x(t_0) \neq y(t_0)$. Then $d(x,y) = \max_{t\in[a,b]} |x(t) - y(t)| \geq |x(t_0) - y(t_0)| > 0$, which gives a contradiction.

(M2) $d(x,y) = d(y,x)$ (clear) since $|x(t) - y(t)| = |y(t) - x(t)|$ for all $t \in [a,b]$.

(M3) $|x(t) - z(t)| \leq |x(t) - y(t)| + |y(t) - z(t)|$. Then, taking the maximum, we obtain: $d(x,z) \leq d(x,y) + d(y,z)$.

□

Example 3.2. Let $X = C[a,b]$ and define a metric on X by

$$d(x,y) = \int_a^b |x(t) - y(t)|\, dt.$$

Proof. We check that the metric axioms (M1)–(M3) hold.

1. **(M1)** If $x = y$ then $d(x,y) = 0$. Assume that $d(x,y) = 0$. We show that $x = y$. Assume that there exists $t_0 \in [a,b]$ such that $x(t_0) \neq y(t_0)$. Let $f = |x - y|$, clearly f is continuous. Hence for every $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|f(t) - f(t_0)| < \varepsilon$$

for all t such that $|t - t_0| < \delta$. This is equivalent to: $f(t_0) - \varepsilon < f(t) < f(t_0) + \varepsilon$. Taking $\varepsilon = \frac{f(t_0)}{2}$, then

$$f(t) > \frac{f(t_0)}{2}$$

for $t \in (t_0 - \delta, t_0 + \delta)$, and hence

$$d(x,y) \geq \int_{t_0-\delta}^{t_0+\delta} |x(t) - y(t)|\, dt \geq 2\delta \frac{f(t_0)}{2} > 0,$$

which is a contradiction.

2. **(M2)** is clear.

3. **(M3)**: $|x(t) - z(t)| \leq |x(t) - y(t)| + |y(t) - z(t)|$ for all $t \in [a,b]$. Hence

$$\int_a^b |x(t) - z(t)|\, dt \leq \int_a^b |x(t) - y(t)|\, dt + \int_a^b |y(t) - z(t)|\, dt$$

This proves that $d(x,z) \leq d(x,y) + d(y,z)$.

□

Example 3.3. Let $X = C[a,b]$ and define a metric on X by

$$d(x,y) = \left(\int_a^b (x(t) - y(t))^2\, dt \right)^{\frac{1}{2}}.$$

In verifying (M1), the same note as in Example (12) is relevant. For the triangle inequality, an integral version of the Cauchy–Schwarz inequality must first be obtained. See Exercise 2.4(6). We will denote this metric space by $C_2[a,b]$.

Proof notes. The proof for d_p being a metric is omitted; it uses:

- Hölder's inequality: $\int_a^b |f(t)g(t)|\,dt \leq \left(\int_a^b |f(t)|^p\,dt\right)^{\frac{1}{p}} \left(\int_a^b |g(t)|^q\,dt\right)^{\frac{1}{q}}$ where $\frac{1}{p}+\frac{1}{q}=1$.
- Minkowski's inequality: $\left(\int_a^b |f(t)+g(t)|^p\,dt\right)^{\frac{1}{p}} \leq \left(\int_a^b |f(t)|^p\,dt\right)^{\frac{1}{p}} + \left(\int_a^b |g(t)|^p\,dt\right)^{\frac{1}{p}}$.

□

Example 3.4. Our final example shows that a metric may be defined for any nonempty set X, without any specification as to the nature of its elements. We define d by

$$d(x,y) = \begin{cases} 0, & \text{if } x = y, \\ 1, & \text{if } x \neq y, \end{cases}$$

where $x, y \in X$. It is a simple matter to check that (M1), (M2) and (M3) are satisfied. This metric is called the *discrete metric*, or the *trivial metric*, for X, and serves a useful purpose as a provider of counterexamples. What is not true in this metric space cannot be true in metric spaces generally.

§3.2 Convergence in metric spaces

A sequence has been defined as a mapping from $\mathbb{N}$ into some set X. If a metric d has been defined for X, we may speak of sequences in the metric space (X,d).

Definition 3.5 (Convergence) A sequence $\{x_n\}$ in a metric space (X,d) is said to **converge** to an element $x \in X$ if for any number $\varepsilon > 0$ there exists a positive integer N such that

$$d(x_n, x) < \varepsilon \quad \text{whenever } n > N.$$

Then x is called the *limit* of the sequence, and we write $x_n \to x$ or $\lim_{n\to\infty} x_n = x$ (adding 'n to ∞' when needed for clarification) [a].

[a] An alternative way of putting this is to require that the real-valued sequence $\{d_n\}$, where $d_n = d(x_n, x)$, converge with limit 0. Thus $x_n \to x$ if and only if $d(x_n, x) \to 0$.

Two important points must be noticed about the definition. First, the element x to which the sequence $\{x_n\}$ in X converges must itself be an element of X. Secondly, the metric by which the convergence is defined must be the metric of the metric space (X,d): the fact that $d(x_n,x)\to 0$ does not imply that $d'(x_n,x)\to 0$, where d' is a different metric for the same set X.

Observations

1. If $X = \mathbb{R}$, then $x_n \to x$ if for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$|x_n - x| < \varepsilon \quad \text{whenever } n > N.$$

2. $x_n \to x$ if and only if $d(x_n, x) \to 0$.

Remarks

1. The limit x has to be an element of X. For instance, $x_n = 1/n$ does not have a limit in $X = (0,1)$.
2. If d and d' are two metrics on X and $d(x_n,x)\to 0$, we cannot necessarily conclude that $d'(x_n,x)\to 0$.

Example 3.6. Consider the set $X = [0,1]$ with the standard metric $d(x,y) = |x-y|$. We define another metric d' on X by

$$d'(x,y) = \begin{cases} 1 & \text{if } x \neq y, \\ 0 & \text{if } x = y. \end{cases}$$

Let $\{x_n\} = \frac{1}{n}$ with $x = 0$. Then we have that $d(x_n, x) = |x_n - x| = \frac{1}{n} \to 0$ as $n \to \infty$. However, $d'(x_n, x) = 1 \not\to 0$ as $n \to \infty$.

In particular, $d'(x_n, x) \not\to 0$ even as $d(x_n, x) \to 0$.

Proposition 3.7. If $\{x_n\}$ is a convergent sequence, then its limit is unique.

Proof. Suppose that $d(x_n, x) \to 0$ and $d(x_n, y) \to 0$. We show that $x = y$.

We have:

$$0 \le d(x,y) \le d(x, x_n) + d(x_n, y).$$

Since $d(x, x_n) \to 0$ and $d(x_n, y) \to 0$, it follows by the triangle inequality that $d(x,y) = 0$, hence $x = y$ by the property (M1) of a metric space. □

Theorem 3.8 — We summarize a few results :

(a) Let $x_n = (x_{n1}, \dots, x_{nm}) \in \mathbb{C}^m$ (or $\mathbb{R}^m$) and $x = (x_1, \dots, x_m) \in \mathbb{C}^m$ (or $\mathbb{R}^m$). Then $x_n \to x$ in Euclidean distance $d(x,y) = \sqrt{\sum_{k=1}^m |x_k - y_k|^2}$ if and only if $x_{nk} \to x_k$ for all $k = 1, \dots, m$.

(b) Let $x_n = \{x_{nk}\}_{k \in \mathbb{N}}$ be an element of ℓ_2, i.e., $x_{nk} \in \mathbb{C}$ for all $k \in \mathbb{N}$ and $\sum_{k=1}^\infty |x_{nk}|^2 < \infty$. Let $x = \{x_k\}_{k \in \mathbb{N}}$ be an element of ℓ_2. If $x_n \to x$ in ℓ_2, equipped with distance $d(x,y) = \sqrt{\sum_{k=1}^\infty |x_k - y_k|^2}$, then $x_{nk} \to x_k$ for all $k \in \mathbb{N}$. The converse is not true.

(c) Let $\{x_n\}$ be a sequence in $C[a,b]$ and $x \in C[a,b]$. Then $x_n \to x$ in $C[a,b]$ equipped with the distance $d(x,y) = \max_{t \in [a,b]} |x(t) - y(t)|$ if and only if $\{x_n\}$ converges uniformly to x, i.e., for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $|x_n(t) - x(t)| < \varepsilon$ for all $t \in [a,b]$ whenever $n \ge N$.

Proof. We have to show that $x_n \to x \iff x_{nk} \to x_k$ for all $k = 1, \dots, m$.

"$\Rightarrow$'" $0 \le |x_{nk} - x_k| \le \sqrt{\sum_{j=1}^m |x_{nj} - x_j|^2} = d(x_n, x)$. Hence $x_{nk} \to x_k$.

"$\Leftarrow$" Assume that $x_{nk} \to x_k$ for all $k = 1, \dots, m$. Then $|x_{nk} - x_k|^2 \to 0$ for all $k = 1, \dots, m$ and hence $\sum_{k=1}^m |x_{nk} - x_k|^2 \to 0$. We infer that $d(x_n, x) = \sqrt{\sum_{k=1}^m |x_{nk} - x_k|^2} \to 0$.

Hence $x_n \to x$ in $\mathbb{R}^m$ or $\mathbb{C}^m$.

Converse is not true, i.e., it is possible to construct a sequence $\{x_n\}$ in ℓ_2 such that $x_{nk} \to x_k$ for all $k \in \mathbb{N}$, but $x_n \not\to x$ in ℓ_2 (does not converge to x).

Here is an example:

$$x_1 = (1, 0, 0, \dots), \quad x = (0, 0, 0, \dots).$$

Then $x_{nk} \to x_k = 0$ for all $k \in \mathbb{N}$. But $d(x_n, x) = \sqrt{\sum_{k=1}^\infty |x_{nk} - x_k|^2} = 1 \not\to 0$.

Notice that $|x_n(t) - x(t)| < \varepsilon$ for all $t \in [a,b]$ is equivalent to:

$d(x_n, x) = \max_{t \in [a,b]} |x_n(t) - x(t)| < \varepsilon$.

Hence $\{x_n\}$ converges uniformly to $x \Rightarrow \forall \varepsilon > 0\ \exists N \in \mathbb{N}$ such that $d(x_n, x) < \varepsilon\ \forall n \ge N$, which is equivalent to $d(x_n, x) \to 0$. □

§4 September 18, 2023

§4.1 Convergence in Metric Spaces - Continued

Definition 4.1 (Cauchy sequence) A sequence $\{x_n\}$ in a metric space (X, d) is a **Cauchy sequence** if $\forall \epsilon > 0\ \exists N \in \mathbb{N}$ such that

$$d(x_n, x_m) < \epsilon \quad \forall n, m > N.$$

Remark 4.2. (Obs.: Th. 1.7.13) In $X = \mathbb{R}$, $\{x_n\}$ converges $\Leftrightarrow$ $\{x_n\}$ is a Cauchy sequence.

Theorem 4.3 — In any metric space, a convergent sequence is a Cauchy Sequence.

Proof. Let $\{x_n\}$ be a convergent sequence in a metric space (X, d). Let $\epsilon > 0$ be arbitrary. Then $\exists N \in \mathbb{N}$ such that $d(x_n, x) < \frac{\epsilon}{2}\ \forall n > N$, where $x = \lim_{n\to\infty} x_n$ in X. Then $\forall n, m > N$,

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

□

Remark 4.4. The converse is NOT true.

Definition 4.5 (Complete Metric Space) If every Cauchy sequence converges in a metric space, then the space is said to be complete

Obs 1. $\mathbb{R}$ is complete.
Obs 2. $\mathbb{Q}$ is not complete. Here is an example of a Cauchy sequence in $\mathbb{Q}$ which does not converge in $\mathbb{Q}$:
Let x_n be the decimal expansion of π truncated after the n-th decimal point, for example, $x_1 = 3.1$, $x_2 = 3.14$, $x_3 = 3.141$, etc. We have that $|x_n - x_m| < 10^{-N}$ for all $n, m > N$. The sequence $\{x_n\}$ is Cauchy since $\forall \epsilon > 0\ \exists N \in \mathbb{N}$ such that $10^{-N} < \epsilon$. However, $\{x_n\}$ is not convergent in $\mathbb{Q}$.

Example 4.6. Let $X = \ell_2$ and consider sequences $\{x_n\}$ such that

$$x_1 = (1, 0, 0, \ldots), \quad x_2 = (0, 1, 0, \ldots), \quad x_3 = (0, 0, 1, \ldots), \ldots$$

Then $\{x_n\}$ is *not* a Cauchy sequence. By Theorem 2.5.6, $\{x_n\}$ is not convergent.
For $n \neq m$,

$$d(x_n, x_m) = \sqrt{\sum_{k=1}^{\infty} |x_{nk} - x_{mk}|^2} = \sqrt{2}.$$

§4.2 Examples on completeness

Example 4.7. $\mathbb{R}$ is complete.

Example 4.8. (2) Let (X, d) be the metric space $\mathbb{C}$, consisting of the set of all complex numbers with the natural metric $d(x, y) = |x - y|$ $(x, y \in \mathbb{C})$. We will show that $\mathbb{C}$ is a complete metric space. Let $\{x_n\}$ be a Cauchy sequence in $\mathbb{C}$. For each $n \in \mathbb{N}$, write $x_n = u_n + iv_n$, where u_n and v_n are real numbers and $i = \sqrt{-1}$. Because $\{x_n\}$ is a Cauchy sequence, for any $\varepsilon > 0$ there is a positive integer N such that $|x_n - x_m| < \varepsilon$ when $m, n > N$. But

$$|u_n - u_m| = |\mathrm{Re}(x_n - x_m)| \leq |x_n - x_m|,$$
$$|v_n - v_m| \leq |x_n - x_m|,$$

so $\{u_n\}$ and $\{v_n\}$ are Cauchy sequences in $\mathbb{R}$. Since $\mathbb{R}$ is complete, these sequences are convergent, and we can write $\lim u_n = u$ and $\lim v_n = v$, say, for some real numbers u, v. Put $x = u + iv$. Then $x \in \mathbb{C}$. Furthermore, $x = \lim x_n$, because

$$0 \le d(x_n, x) = |x_n - x| = |(u_n + iv_n) - (u + iv)| = |(u_n - u) + i(v_n - v)| \le |u_n - u| + |v_n - v| < \varepsilon$$

for any $\varepsilon > 0$, provided n is large enough. Hence we have proved that the Cauchy sequence $\{x_n\}$ is convergent, so $\mathbb{C}$ is a complete metric space.

This proof has been written out in full detail. A similar process is followed in Examples (3) and (5) below. The general technique is to take a Cauchy sequence in the space, postulate a natural limit for the sequence, show that it is an element of the space, and then verify that it is indeed the limit.

Example 4.9. The metric space ℓ_2 is complete. Let $\{x_n\}$ be a Cauchy sequence in ℓ_2. We must show that the sequence converges. For each $n \in \mathbb{N}$, write $x_n = (x_{n1}, x_{n2}, \dots)$. By definition of the space ℓ_2, the series $\sum_{k=1}^{\infty} |x_{nk}|^2$ converges for each n. Since $\{x_n\}$ is a Cauchy sequence, for any $\varepsilon > 0$ there is a positive integer N such that

$$\sqrt{\sum_{k=1}^{\infty} |x_{nk} - x_{mk}|^2} < \varepsilon$$

when $m, n > N$, using the definition of the metric for ℓ_2. That is,

$$\sum_{k=1}^{\infty} |x_{nk} - x_{mk}|^2 < \varepsilon^2, \quad m, n > N,$$

so we must have

$$|x_{nk} - x_{mk}| < \varepsilon, \quad m, n > N,$$

for each $k \in \mathbb{N}$. Then, for each k, $\{x_{nk}\}$ is a Cauchy sequence in $\mathbb{C}$ so $\lim_{n\to\infty} x_{nk}$ exists since $\mathbb{C}$ is complete. Write $\lim_{n\to\infty} x_{nk} = x_k$ and set $x = (x_1, x_2, \dots)$. We will show that $x \in \ell_2$ and that $\{x_n\}$ converges to x. This will then mean that ℓ_2 is complete. We note first that for any $r = 1, 2, \dots,$

$$\sum_{k=1}^{r} |x_{nk} - x_{mk}|^2 < \varepsilon^2, \quad m, n > N,$$

so that, keeping n fixed and using the fact that $\lim_{m\to\infty} x_{mk} = x_k$,

$$\sum_{k=1}^{r} |x_{nk} - x_k|^2 < \varepsilon^2, \quad n > N,$$

by Theorem 1.7.7. For points $(a_1, a_2, \dots, a_r), (b_1, b_2, \dots, b_r), (c_1, c_2, \dots, c_r) \in \mathbb{C}^r$, the triangle inequality in $\mathbb{C}^r$ gives us

$$\sqrt{\sum_{k=1}^{r} |a_k - c_k|^2} \le \sqrt{\sum_{k=1}^{r} |a_k - b_k|^2} + \sqrt{\sum_{k=1}^{r} |b_k - c_k|^2}.$$

Replacing a_k by x_k, b_k by x_{nk} and c_k by 0, we have

$$\sqrt{\sum_{k=1}^{r} |x_k|^2} \le \sqrt{\sum_{k=1}^{r} |x_k - x_{nk}|^2} + \sqrt{\sum_{k=1}^{r} |x_{nk}|^2} \le \varepsilon + \sqrt{\sum_{k=1}^{\infty} |x_{nk}|^2} \le \varepsilon + \sqrt{M_n},$$

if $n > N$. The convergence of the final series here thus implies the convergence of $\sum_{k=1}^{\infty} |x_k|^2$, so that indeed $x \in \ell_2$. Moreover, an inequality a few lines back shows further that

$$\sqrt{\sum_{k=1}^{\infty} |x_{nk} - x_k|^2} < \varepsilon, \quad n > N,$$

and this implies that the sequence $\{x_n\}$ converges to x.

Example 4.10. metric spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ are complete. This is easily shown by adapting the method of Example (3).

The

Example 4.11. The metric space $C[a, b]$ is complete. Let $\{x_n\}$ be a Cauchy sequence in $C[a, b]$. Then, for any $\varepsilon > 0$, we can find N so that, using the definition of the metric for this space,

$$\max_{a \leq t \leq b} |x_n(t) - x_m(t)| < \varepsilon$$

when $m, n > N$. Certainly then for each particular t in $[a, b]$ we have

$$|x_n(t) - x_m(t)| < \varepsilon, \quad m, n > N,$$

so $\{x_n(t)\}$ is a Cauchy sequence in $\mathbb{R}$. But $\mathbb{R}$ is complete, so the sequence $\{x_n(t)\}$ converges to a real number, which we will write as $x(t)$, for each t in $[a, b]$. This determines a function x, defined on $[a, b]$. In the preceding inequality, fix n (and let $m \to \infty$) to give

$$|x_n(t) - x(t)| < \varepsilon, \quad n > N.$$

The N here is independent of t in $[a, b]$, so we have shown that the sequence $\{x_n\}$ converges uniformly on $[a, b]$ to x. Using the theorem that the uniform limit of a sequence of continuous functions is itself continuous (Theorem 1.10.3), our limit function x must be continuous on $[a, b]$. That is, $x \in C[a, b]$. Furthermore, uniform convergence on $[a, b]$ is equivalent to convergence in $C[a, b]$ (Theorem 2.4.3(c)). Thus the Cauchy sequence $\{x_n\}$ converges to x, completing the proof that $C[a, b]$ is complete.

§5 September 21, 2023

Definition 5.1 (Subspace) Let (X, d) be a metric space and $S \subseteq X$. The restriction of d on S is defined by:

$$d_S : S \times S \to \mathbb{R} \text{ given by } d_S(x, y) = d(x, y)$$

(S, d_S) is called a **subspace** of X.

Obs: (S, d_S) is a metric space.

Definition 5.2 (Sequentially closed) S is called sequentially closed if for any sequence $\{x_n\}$ in S for which $x = \lim x_n$ exists in X, we have: $x \in S$.

Obs: If $X = \mathbb{R}$, $S \subseteq \mathbb{R}$ is sequentially closed if and only if S is closed.

- $[a, b]$ is sequentially closed in $\mathbb{R}$ (i.e., S^c is open).
- $\{z \in \mathbb{C} \mid |z| \leq c\}$ is sequentially closed in $\mathbb{C}$.

Theorem 5.3 — Let (X, d) be a metric space and $S \subseteq X$. Then S is complete if and only if S is sequentially closed.

Proof. '⇒" Assume that S is complete. We prove that S is sequentially closed. Let $\{x_n\}$ be a sequence in S such that $x_n \to x \in X$.
By Theorem 2.5.6, $\{x_n\}$ is a Cauchy sequence. Since S is complete, there exists $y \in S$ such that $x_n \to y$. By the uniqueness of the limit, $x = y$. Since $y \in S$, we infer that $x \in S$. "⇐" Let $\{x_n\}$ be a Cauchy sequence in S. Since X is complete, there exists $x \in X$ such that $\lim_{n\to\infty} x_n = x$. Since S is sequentially closed, $x \in S$. □

Definition 5.4 Next we define two notions for metric spaces:

(a) The diameter of a set S is:

$$\delta(S) = \sup\{d(x, y) \mid x, y \in S\}$$

(b) S is bounded if $S = \varnothing$ or $\delta(S) < \infty$.

§5.1 Sequential Continuity

- Let (X, d) and (Y, d') be two metric spaces, and $A : X \to Y$ an arbitrary map. We will denote Ax the value assigned by A to point $x \in X$, instead of $A(x)$.
- Let (Z, d'') be another metric space and $\beta : Y \to Z$ be another map. We write βA instead of $\beta \circ A$ for the composition map defined by: $\beta A : X \to Z$, $x \mapsto \beta(Ax)$.

$$X \xrightarrow{A} Y \xrightarrow{\beta} Z$$

$$x \mapsto Ax \mapsto \beta Ax$$

- If $A : X \to X$ is a function from X to X, then $A^2 = A \circ A : X \to X$ is the function given by $A^2x = A(Ax)$. In general, by recurrence, we define $A^n : X \to X$ given by:

$$A^n x = A(A^{n-1}x) \text{ for all } x \in X$$

- The identity map is defined by $I : X \to X$, $Ix = x$.

- Associativity: If $A : X \to Y$, $B : Y \to Z$, $C : Z \to W$, then

$$C(\beta A) = (CB)A$$

Example 5.5. Let $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$ for $n, m \in \mathbb{N}$. Let $A = (a_{ij})$ where $1 \le i \le m$ and $1 \le j \le n$. Then $A : \mathbb{R}^n \to \mathbb{R}^m$ is a map between the two metric spaces X and Y.

Definition 5.6 Let (X, d) and (Y, d') be two metric spaces. A function $A : X \to Y$ is *sequentially continuous at point* $x \in X$ if for any sequence $\{x_n\}$ in X for which $\lim_{n\to\infty} x_n = x$, $\{Ax_n\}$ is also a convergent sequence in Y and $\lim_{n\to\infty} Ax_n = Ax$. We say that A is *sequentially continuous on* X if it is sequentially continuous at every $x \in X$. Usually, we drop the word "sequentially".

Example 5.7. Let $C[a, b]$ be the set of all continuous functions on $[a, b]$, endowed with the uniform distance:

$$d(f, g) = \max_{t\in[a,b]} |f(t) - g(t)|$$

Let $A : C[a, b] \to \mathbb{R}$ defined by

$$Af = \int_a^b f(t)\, dt, \quad \forall f \in C[a, b].$$

We show that A is continuous (in the sense of Definition 3.1.1).

Let $\{f_n\}$ be a sequence in $C[a, b]$ such that $f_n \to f$ in the uniform distance, for some $f \in C[a, b]$. We have to show that:

$$\lim_{n\to\infty} Af_n = Af \text{ (in } \mathbb{R})$$

Let $\varepsilon > 0$ be arbitrary. Since $d(f_n, f) \to 0$, there exists $N \in \mathbb{N}$ such that $d(f_n, f) < \varepsilon$ for all $n > N$. This means that

$$|Af_n - Af| < \varepsilon \text{ for all } n > N.$$

Then $\forall n > N$,

$$|A_{f_n} - A_f| = \left| \int_a^b f_n(t)\, dt - \int_a^b f(t)\, dt \right| = \left| \int_a^b (f_n(t) - f(t))\, dt \right|$$

$$\le \int_a^b |f_n(t) - f(t)|\, dt \quad (1)$$

$$< \int_a^b \varepsilon\, dt = \varepsilon(b - a) \quad (2)$$

Using the following inequality:

$$\left| \int_a^b f(t)\, dt \right| \le \int_a^b |f(t)|\, dt \quad (3)$$

We also used the following inequality:

$$\text{if } f(t) \le g(t)\, \forall t \in [a, b] \text{ then } \int_a^b f(t)\, dt \le \int_a^b g(t)\, dt \quad (4)$$

This proves that $A_{f_n} \to A_f$ in $\mathbb{R}$.

§5.2 Contraction mapping and fixed point

A basic question in mathematics is to solve equation $A(x) = y$. We will learn how to solve this equation in the format $A(x) = x$.

Definition 5.8 (Fixed point, contraction) Let (X, d) be a metric space and a map $A : X \to X$.

(a) A point $x \in X$ is called a *fixed point* of A if $A(x) = x$.

(b) The map A is a *contraction*, if there exists $\alpha \in (0, 1)$ such that

$$d(A(x), A(y)) \leq \alpha d(x, y) \quad \forall x, y \in X$$

α is called the *contraction constant* of A.

Remark: Solving the equation $A(x) = x$ is equivalent to finding the fixed points of A.

Theorem 5.9 — If $A : X \to X$ is a contraction on a metric space (X, d), then A is continuous.

Proof. We have to prove that A is (sequentially) continuous at every point $x \in X$. Let $x \in X$ be arbitrary, and let $\{x_n\}$ be a sequence in X such that $x_n \to x$. By the contraction property (3)

$$0 \leq d(A(x_n), A(x)) \leq \alpha d(x_n, x)$$

By the Squeeze Criterion, $d(A(x_n), A(x)) \to 0$. □

Remark 1: Theorem 3.2.2 continues to hold if (3) holds for some α (not necessarily $\alpha \in (0, 1)$).
Remark 2: The converse of Theorem 3.2.2 is not necessarily true; if $A : X \to X$ is continuous, A does not have to be a contraction (example: $A = I$).

§6 September 25, 2023

§6.1 Fixed Point Theorem

Theorem 6.1 (Fixed Point Theorem) — Every contraction mapping on a complete metric space has a unique fixed point.

Proof. Let A be a contraction mapping with contraction constant α, on a complete metric space (X, d). Take any point $x_0 \in X$ and define the sequence $\{x_n\}$ in X recursively by

$$x_n = A(x_{n-1}), \quad \text{for } n \in \mathbb{N}.$$

Hence, we have

$$x_1 = A(x_0), \quad x_2 = A^2(x_0), \quad x_3 = A^3(x_0), \quad \ldots, \quad x_n = A^n(x_0).$$

We aim to prove that $\{x_n\}$ is a Cauchy sequence. Observe that for any integer $k > 1$,

$$d(x_k, x_{k-1}) = d(A^k(x_0), A^{k-1}(x_0)) \leq \alpha d(A^{k-1}(x_0), A^{k-2}(x_0)) \leq \ldots \leq \alpha^{k-1} d(A(x_0), x_0).$$

For $1 \leq m < n$, we find that

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, x_{n-1}) + d(x_{n-1}, x_{n-2}) + \ldots + d(x_{m+1}, x_m) \\ &\leq \alpha^{n-1} d(x_1, x_0) + \alpha^{n-2} d(x_1, x_0) + \ldots + \alpha^m d(x_1, x_0) \\ &= \alpha^m (1 + \alpha + \alpha^2 + \ldots + \alpha^{n-m-1}) d(x_1, x_0) \\ &< \frac{\alpha^m}{1 - \alpha} d(x_1, x_0), \end{aligned}$$

by the geometric series sum, since $0 < \alpha < 1$. As m approaches infinity, α^m approaches zero, which implies $d(x_n, x_m)$ can be made arbitrarily small for sufficiently large m and n. Therefore, $\{x_n\}$ is a Cauchy sequence. Because X is complete, there exists $x = \lim x_n$.

To show x is a fixed point of A, observe that for any n,

$$0 \leq d(A(x), x) \leq d(A(x), A(x_{n-1})) + d(A(x_{n-1}), x) \leq \alpha d(x, x_{n-1}) + d(x_n, x),$$

which implies $d(A(x), x) = 0$ as n goes to infinity because $d(x_n, x)$ tends to zero. Hence, $A(x) = x$, and x is a fixed point of A.

To prove uniqueness, suppose y is another fixed point such that $A(y) = y$. Then

$$d(x, y) = d(A(x), A(y)) \leq \alpha d(x, y),$$

and since $\alpha < 1$, this inequality holds only if $d(x, y) = 0$, i.e., $x = y$. Therefore, A has a unique fixed point. □

Theorem 6.2 — Let A be a mapping on a complete metric space, and suppose that A^n is a contraction for some integer $n \in \mathbb{N}$. Then A has a unique fixed point.

Proof. Let the metric space be X. According to the Fixed Point Theorem, the mapping A^n has a unique fixed point $x \in X$ such that $A^n(x) = x$. Observing that

$$A^n(A(x)) = A^{n+1}(x) = A(A^n(x)) = A(x),$$

we conclude that $A(x)$ is also a fixed point of A^n. Since A^n can have only one fixed point, it must be that $A(x) = x$, and thus x is also a fixed point of A.

Now, any fixed point y of A is also a fixed point of A^n since

$$A^n(y) = A^{n-1}(A(y)) = A^{n-1}(y) = \cdots = A(y) = y.$$

It follows that x is the only fixed point of A, as any other fixed point y would also be a fixed point of A^n, contradicting the uniqueness of x. □

§6.2 Applications of Fixed Point Theorem

Example 6.3. Let f be a function with domain $[a, b]$ and range a subset of $[a, b]$. Suppose there is some positive constant $K < 1$ such that

$$|f(x_1) - f(x_2)| \leq K|x_1 - x_2|,$$

for any points $x_1, x_2 \in [a, b]$. (Then f is said to satisfy a Lipschitz condition, with Lipschitz constant K.) The Fixed Point Theorem assures us that the equation $f(x) = x$ has a unique solution for x in $[a, b]$.

This is because f may be considered as a mapping from the metric space consisting of the closed interval $[a, b]$ with the natural metric into itself, and this metric space is complete because it is a closed subspace of $\mathbb{R}$ (Theorem 2.7.3). Second, the Lipschitz condition, with $0 < K < 1$, states that this mapping f is a contraction. Hence f has a unique fixed point.

If f is a differentiable function on $[a, b]$, with range a subset of $[a, b]$, and if there is a constant K such that

$$|f'(x)| \leq K < 1,$$

for all x in $[a, b]$, then again the equation $f(x) = x$ has a unique solution for x in $[a, b]$. This follows from the mean value theorem of differential calculus: for any $x_1, x_2 \in [a, b]$, with $x_1 < x_2$, there is at least one point c, $x_1 < c < x_2$, such that

$$|f(x_1) - f(x_2)| = |f'(c)(x_1 - x_2)| = |f'(c)||x_1 - x_2| \leq K|x_1 - x_2|,$$

and so f satisfies the Lipschitz condition with constant $K < 1$.

As an example, consider the function $f(x) = \frac{x^5 - x}{2} + \frac{1}{2}$, for $0 \leq x \leq 1$. The given equation is equivalent to the equation $f(x) = x$, so we seek information about the fixed points of f. The domain and range of f is $[0, 1]$, and we have

$$|f'(x)| = |5x^4 - x| \leq 5x^4 + x \leq \frac{5}{16} + \frac{1}{2} < 1$$

for all x in $[0, 1]$. All the required conditions are met, so f has a single fixed point, which is the required root of the original equation.

To find the root, we can iteratively apply f starting from $x_0 = 0$. The first three iterates are $x_1 = f(0) = 0.25$, $x_2 = f(0.25) \approx 0.2197$, $x_3 = f(0.2197) \approx 0.2264$, and the subsequent iterates converge to the root, which to three decimal places, is 0.225.

§7 September 28, 2023

§7.1 Advanced Applications of Fixed Point Theorem

Example 7.1. Consider the system of n linear equations in n unknowns $x_1, x_2, \ldots, x_n$, where a_{jk} and b_j are real numbers for each j and k. Introducing the $n \times n$ matrix $A = (a_{jk})$ and the column vectors $x = (x_1, x_2, \ldots, x_n)^T$, $b = (b_1, b_2, \ldots, b_n)^T$, the system can be written in matrix form as $Ax = b$, and must be solved for x. Letting $C = (c_{jk})$ be the matrix $I - A$, where I is the $n \times n$ identity matrix, the equation may be written $(I - C)x = b$, or

$$Cx + b = x.$$

Considering the elements of $\mathbb{R}^n$ to be column vectors, we define a mapping $M : \mathbb{R}^n \to \mathbb{R}^n$ by

$$Mx = Cx + b,$$

so that our matrix equation is replaced by the equation

$$Mx = x.$$

Hence the solutions of the original system are related to the fixed points of the mapping M. Since $\mathbb{R}^n$ is a complete metric space, there will be just one solution if M is a contraction mapping.

Let $y = (y_1, y_2, \ldots, y_n)^T$ and $z = (z_1, z_2, \ldots, z_n)^T$ be two points of $\mathbb{R}^n$ and let d denote the Euclidean metric:

$$d(y, z) = \sqrt{\sum_{j=1}^{n} (y_j - z_j)^2}.$$

Since My is the vector $Cy + b$, with j-th component $\sum_{k=1}^{n} c_{jk} y_k + b_j$ (for $j = 1, 2, \ldots, n$), and similarly for Mz, we have

$$\begin{aligned} d(My, Mz) &= \sqrt{\sum_{j=1}^{n} \left(\sum_{k=1}^{n} c_{jk} y_k + b_j - \sum_{k=1}^{n} c_{jk} z_k + b_j \right)^2} \\ &= \sqrt{\sum_{j=1}^{n} \left(\sum_{k=1}^{n} c_{jk} (y_k - z_k) \right)^2} \\ &\leq \sqrt{\sum_{j=1}^{n} \left(\sum_{k=1}^{n} c_{jk}^2 \right) \left(\sum_{k=1}^{n} (y_k - z_k)^2 \right)}, \end{aligned}$$

where the last inequality follows from the Cauchy-Schwarz inequality. Continuing from the Cauchy-Schwarz inequality, Theorem 2.2.1, we have

$$d(My, Mz) \leq \sqrt{\sum_{j=1}^{n} \left(\sum_{k=1}^{n} c_{jk}^2 \right)} \, d(y, z),$$

so certainly M will be a contraction if

$$0 < \sum_{j=1}^{n} \sum_{k=1}^{n} c_{jk}^2 < 1.$$

In terms of the original matrix A, this condition requires that a_{jk} be near 0 when $j \neq k$ and near 1 when $j = k$.

Different sufficient conditions for M to be a contraction can be obtained by choosing different metrics on the set $\mathbb{R}^n$, as long as the resulting metric space is complete. We are totally free to take whichever metric best serves our purpose. For instance, with the metric d_∞, where

$$d_\infty(y,z) = \max_{1\le k\le n} |y_k - z_k|,$$

we know that $(\mathbb{R}^n, d_\infty)$ is complete (Exercise 2.9(7)), and

$$\begin{aligned} d_\infty(My, Mz) &= \max_{1\le j\le n} \left| \sum_{k=1}^{n} c_{jk} y_k + b_j - \sum_{k=1}^{n} c_{jk} z_k + b_j \right| \\ &= \max_{1\le j\le n} \left| \sum_{k=1}^{n} c_{jk}(y_k - z_k) \right| \\ &\le \max_{1\le j\le n} \sum_{k=1}^{n} |c_{jk}||y_k - z_k| \\ &\le \max_{1\le j\le n} \sum_{k=1}^{n} |c_{jk}| \cdot \max_{1\le k\le n} |y_k - z_k| \\ &= \max_{1\le j\le n} \sum_{k=1}^{n} |c_{jk}| \cdot d_\infty(y,z), \end{aligned}$$

so that M will be a contraction under this metric if

$$0 < \max_{1\le j\le n} \sum_{k=1}^{n} |c_{jk}| < 1,$$

that is, if the sums of the absolute values of the elements in the rows of C are all less than 1 (and C has at least one nonzero element).

A third condition is obtained in Exercise 3.5(3). It only takes one of these conditions for M to be a contraction. The conditions to be satisfied to ensure the existence of the unique fixed point.

Once M is known to be a contraction, its fixed point can be found, at least approximately, by iteration. If x_0 is any column vector, then we have successively

$$\begin{aligned} x_1 &= Mx_0 = Cx_0 + b, \\ x_2 &= Mx_1 = Cx_1 + b = C(Cx_0 + b) + b = C^2x_0 + Cb + b, \\ x_3 &= Mx_2 = Cx_2 + b = C^3x_0 + C^2b + Cb + b, \end{aligned}$$

and so on, the sequence $\{x_n\}$ converging to the unique solution x of $Ax = b$, where $A = I - C$. There are of course other tests for whether a system of linear equations has solutions, and other methods of finding them. However, the above is very simple. The tests essentially require only the operation of addition on the elements of C or their squares, and, if either condition is satisfied, the solution may be obtained to any desired degree of accuracy (subject to computational precision) in terms of powers of C. There is no need to determine the rank, determinant or inverse of any matrix. It must be realised, though, that we have only obtained sufficient conditions: if none of the conditions is met, solutions may still exist.

As a simple example, consider the system of equations

$$\begin{aligned} 16x - 3y + 4z &= 7, \\ 6x + 7y - 4z &= 4, \\ y + 4z &= 15. \end{aligned}$$

Dividing the equations respectively by 16, 8 and 4 gives the equivalent system

$$
\begin{aligned}
x - \frac{3}{16}y + \frac{1}{4}z &= \frac{7}{16}, \\
\frac{3}{4}x + \frac{7}{8}y - \frac{1}{2}z &= \frac{1}{2}, \\
\frac{1}{4}y + z &= \frac{15}{4}.
\end{aligned}
$$

In the notation above, we have

$$
A = \begin{pmatrix} 1 & -\frac{3}{16} & \frac{1}{4} \\ \frac{3}{4} & \frac{7}{8} & -\frac{1}{2} \\ 0 & \frac{1}{4} & 1 \end{pmatrix}, \quad C = I - A = \begin{pmatrix} 0 & \frac{3}{16} & -\frac{1}{4} \\ -\frac{3}{4} & \frac{1}{8} & \frac{1}{2} \\ 0 & -\frac{1}{4} & 0 \end{pmatrix},
$$

and we find that the sum of the squares of the elements of C is $\frac{253}{256}$, less than 1, so our system possesses a unique solution which may be found sufficiently.

Example 7.2. Our third application of the fixed point theorem is to prove an important theorem on the existence of a solution to the first-order differential equation

$$
\frac{dy}{dx} = f(x, y) \tag{1}
$$

with initial condition $y = y_0$ when $x = x_0$. The result is a form of Picard's theorem.

Two conditions are imposed on f: first, f is continuous in some rectangle $\{(x, y) : |x - x_0| \leq a,\ |y - y_0| \leq b\}$; second, f satisfies a Lipschitz condition on y, uniformly in x, in the rectangle. The latter means that there is a positive constant K such that

$$
|f(x, y_1) - f(x, y_2)| \leq K|y_1 - y_2| \tag{2}
$$

for any x in $[x_0 - a, x_0 + a]$ and any y_1, y_2 in $[y_0 - b, y_0 + b]$. Since f is continuous in the rectangle, it must be bounded there (see Section 1.9), so there is a positive constant M such that $|f(x, y)| \leq M$.

Under these conditions, we will prove that there is a positive number h such that in $[x_0 - h, x_0 + h]$ there is a unique solution to the differential equation.

Write the differential equation equivalently in integral form as

$$
y(x) = y_0 + \int_{x_0}^{x} f(t, y)\, dt, \tag{3}
$$

incorporating the initial condition. Let h be a number satisfying

$$
h > 0, \quad h < \frac{1}{K'}, \quad h \leq a, \quad h \leq \frac{b}{M}. \tag{4}
$$

Denote by J the closed interval $[x_0 - h, x_0 + h]$ and write $C[J]$ for $C[x_0 - h, x_0 + h]$. Let F be the subset of $C[J]$ consisting of continuous functions defined on J for which

$$
|y(x) - y_0| \leq b, \quad x \in J, \quad y \in C[J]. \tag{5}
$$

Referring to Figure 9, F is the set of all continuous functions with graphs in the shaded rectangle. Impose the uniform metric on F, so that F becomes a subspace of the complete metric space $C[J]$.

We will show that F is a closed subspace, so that, by Theorem 2.7.3, F is a complete metric space. Let $\{y_n\}$ be a sequence of functions in F which, as a sequence in $C[J]$, converges. Write $y = \lim y_n$ (so $y \in C[J]$). By definition of the uniform metric, given $\varepsilon > 0$ we can find a positive integer N such that

$$
\max_{x \in J} |y_n(x) - y(x)| < \varepsilon, \quad n > N. \tag{6}
$$

Also, for each $x \in J$ and each $n \in \mathbb{N}$,

$$|y_n(x) - y_0| \leq b. \tag{7}$$

Hence, for each $x \in J$, and $n > N$,

$$|y(x) - y_0| \leq |y(x) - y_n(x)| + |y_n(x) - y_0| < \varepsilon + b. \tag{8}$$

But ε is arbitrary, so we must have

$$|y(x) - y_0| \leq b \tag{9}$$

for all $x \in J$. This shows that $y \in F$, so F is a closed subspace of $C[J]$.

Now define a mapping A on F by the equation $Ay = z$, where $y \in F$ and

$$z(x) = y_0 + \int_{x_0}^{x} f(t, y(t))\, dt, \quad x \in J. \tag{10}$$

We will show that $z \in F$ and that A is a contraction mapping. Then the fixed point theorem will imply that A has a unique fixed point. That is, we will have shown the existence of a unique function $y \in F$ such that $Ay = y$, which means

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))\, dt, \quad x \in J. \tag{11}$$

This will complete the proof of the existence on J of a unique solution of our differential equation.

To show that $z \in F$, we see that, for $x \in J$,

$$|z(x) - y_0| = \left| \int_{x_0}^{x} f(t, y)\, dt \right| \leq \int_{x_0}^{x} |f(t, y)|\, dt \leq M|x - x_0| \leq Mh \leq b.$$

Thus $z \in F$ (so A maps F into itself). To show that A is a contraction, take $y, \tilde{y} \in F$. Set $z = Ay$, $\tilde{z} = A\tilde{y}$. Let d denote the uniform metric. Then, for $x \in J$,

$$\begin{aligned} |z(x) - \tilde{z}(x)| = \left| \int_{x_0}^{x} (f(t, y) - f(t, \tilde{y}))\, dt \right| &\leq \int_{x_0}^{x} |f(t, y) - f(t, \tilde{y})|\, dt \leq K \int_{x_0}^{x} |y(t) - \tilde{y}(t)|\, dt \\ &\leq K \cdot \max_{x \in J} |y(x) - \tilde{y}(x)| \cdot |x - x_0| \leq Kh \max_{x \in J} |y(x) - \tilde{y}(x)|, \end{aligned}$$

where $\max_{x \in J} |y(x) - \tilde{y}(x)| = d(y, \tilde{y})$. Thus,

$$d(Ay, A\tilde{y}) \leq \alpha d(y, \tilde{y}),$$

where $\alpha = Kh$. But $0 < \alpha < 1$ and so A is a contraction.

It is easy to check that this result may be applied successfully to, for example, the linear first-order differential equation

$$\frac{dy}{dx} + P(x)y = Q(x), \quad y(x_0) = y_0,$$

to ensure a unique solution in some interval about x_0, provided the functions P and Q are continuous.

An example of a differential equation where it cannot be applied is the equation

$$\frac{dy}{dx} = 2|y|^{1/2}, \quad y(0) = 0.$$

It is impossible to satisfy the Lipschitz condition for small values of $|y|$: the inequality $||y_1|^{1/2} - |y_2|^{1/2}| \leq K|y_1 - y_2|$ cannot hold for any constant K if we take $y_2 = 0$ and $|y_1| < 1/K^2$. In fact, this equation has at least two solutions for x in any interval containing 0. These are the functions defined by the equations

$$y = \begin{cases} x^2, & x \geq 0, \\ -x^2, & x < 0. \end{cases}$$

Example 7.3. The differential equation in (3) was considered by first transforming it into an integral equation. We intend now to study two standard types of integral equations, in each case obtaining conditions which ensure a unique solution.

(a) Any equation of the form

$$x(s) = \lambda \int_a^b k(s,t)x(t)\,dt + f(s), \quad a \leq s \leq b,$$

involving two given functions k (of two variables) and f, an unknown function x, and a nonzero constant λ, is called a *Fredholm integral equation* (of the second kind).

Suppose f is continuous on the interval $[a,b]$, and k is continuous on the square $[a,b] \times [a,b]$. Then k is bounded: there exists a positive constant M so that, in the square, $|k(s,t)| \leq M$.

Take any continuous function x on $[a,b]$ and define a mapping A on $C[a,b]$ by $y = Ax$, where

$$v(s) = \lambda \int_a^b k(s,t)x(t)\,dt + f(s)$$

We will obtain a condition for A to be a contraction. Note in the first place that, since k, x and f are continuous, so is y, and so indeed A maps the complete metric space $C[a,b]$ into itself. Now, if d denotes the uniform metric of $C[a,b]$, and if $y_1 = Ax_1$, $y_2 = Ax_2$ ($x_1, x_2 \in C[a,b]$), then

$$d(y_1,y_2) = \max_{a\leq s\leq b} |y_1(s) - y_2(s)|$$

$$= \max_{a\leq s\leq b} \left| \lambda \int_a^b k(s,t)(x_1(t) - x_2(t))\,dt \right|$$

$$\leq |\lambda| \cdot \max_{a\leq s\leq b} \int_a^b |k(s,t)||x_1(t) - x_2(t)|\,dt$$

$$\leq |\lambda| M(b-a) \cdot \max_{a\leq s\leq b} |x_1(s) - x_2(s)|$$

$$= |\lambda| M(b-a) d(x_1,x_2),$$

and hence A is a contraction mapping provided

$$|\lambda| < \frac{1}{M(b-a)}.$$

Thus, provided the constant λ satisfies this inequality, we are assured that the original Fredholm integral equation has a unique solution. This solution may be found by iteration, taking any function in $C[a,b]$ as starting point.

As an example, consider the equation

$$x(s) = \frac{1}{2}\int_0^1 stx(t)\,dt + \frac{5s}{6}.$$

In the above notation, $\lambda = \frac{1}{2}$, $a = 0$, $b = 1$, $k(s,t) = st$, $f(s) = \frac{5s}{6}$. For $s,t \in [0,1]$, we have $|k(s,t)| = st \leq 1$, so take $M = 1$. The inequality for λ is satisfied, so a unique solution is assured. To find it, let us take as starting point the function x_0 where $x_0(s) = 1$, $0 \leq s \leq 1$. Then we obtain

$$x_1(s) = \frac{1}{2}\int_0^1 st\,dt + \frac{5s}{6} = \frac{13s}{12},$$

$$x_2(s) = \frac{1}{2}\int_0^1 \frac{13st}{12}\,dt + \frac{5s}{6} = \frac{73s}{72},$$

$$x_3(s) = \frac{1}{2}\int_0^1 \frac{73st}{72}\,dt + \frac{5s}{6} = \frac{433s}{432},$$

and we are led to suggest

$$x_n(s) = \frac{2 \cdot 6^n + 1}{2 \cdot 6^n} s, \quad n \in \mathbb{N}.$$

This should be verified by mathematical induction. The solution of the integral equation is $\lim x_n$: the function x, where $x(s) = 0 \leq s \leq 1$.

§8 October 2, 2023

Recall: If $\{x_n\}$ is a sequence in a metric space (X, d) and $n_1 < n_2 < n_3 < \cdots$ are integers, then $\{x_{n_k}\}$ is a subsequence of $\{x_n\}$.

If $\lim_{n\to\infty} x_n = x$ then $\lim_{n_k\to\infty} x_{n_k} = x$ for any subsequence $\{x_{n_k}\}$. The converse is not true. If $\{x_{n_k}\}$ d.n.c. limit$x_{n_k} = x$, then $\lim_{n\to\infty} x_n$ may not be x. Example: $\{x_n\}_{n\in\mathbb{N}}$ is given by: $2, \frac{1}{2}, 3, \frac{1}{3}, 4, \frac{1}{4}, \ldots$

However, we can saw the following:

Theorem 8.1 — In any metric space, a Cauchy sequence having a convergent subsequence is itself convergent, with the same limit. [a]

[a] If $\{x_n\}$ is a Cauchy sequence and $\{x_n\}$ has a convergent subsequence $\{x_{n_k}\}$ to a limit x, then $\lim_{n\to\infty} x_n = x$.

Proof. Let $\{x_n\}$ be a Cauchy sequence in metric space (X, d), and let $\{x_{n_k}\}$ be a convergent subsequence of $\{x_n\}$. Set $x = \lim_{k\to\infty} x_{n_k}$ be the limit of the convergent subsequence. Now for any $\varepsilon > 0$, there exists a $K \in \mathbb{N}$ such that $d(x_{n_k}, x) < \frac{1}{2}\varepsilon$ when $k > K$. As $\{x_n\}$ is a Cauchy sequence, there exists $N \in \mathbb{N}$ such that for all $n, m > N$, $d(x_n, x_m) < \frac{1}{2}\varepsilon$ when $n, m > N$. We may assume that $K > N$. Which implies that $n_k \geq k > K > N$ and we have

$$d(x_n, x) \leq d(x_n, x_{n_k}) + d(x_{n_k}, x) = \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon$$

□

Before we introduced completeness because of the need to categorize those metric spaces which have the propriety of Cauchy convergence for real numbers. In a similar way we now discuss an other property: *compactness*. The Bolzano–Weierstrass theorem says that there exists a convergent subsequence of any (real-valued) sequence, as long as it is bounded. This notions leads us to our definition of compactness.

Definition 8.2 A subset of a metric space is a called **sequentially compact** if every sequence in the subset has a convergent subsequence.

Theorem 8.3 — If a metric space is compact, then it is complete.

Proof. This follows directly from the previous theorem. A complete metric space is one in which every Cauchy sequence is convergent. From the previous theorem we know if Cauchy sequence has a convergent subsequence, it is convergent. And in the definition of compactness is every sequence in the subset has a convergent subsequence. Therefore, every Cauchy sequence would be convergent in this space. □

Remark: The converse of Theorem 4.1.5 is not true.
Ex: $X = \mathbb{R}$ is sequentially closed, but X is not sequentially compact.

Definition 8.4 **Recall:** $S \subseteq X$ is bounded if

$$\delta(S) = \sup\{d(x, y); x, y \in S\} < \infty \quad \text{or} \quad S = \emptyset.$$

This is equivalent to saying that $\exists M > 0$ s.t. $S \subseteq B_M(x_0)$ where $B_M(x_0) = \{x \in X; d(x, x_0) < M\}$ is the ball of radius M and center x_0.

Theorem 8.5 — Let (X, d) be a metric space and $S \subseteq X$. If S is sequentially compact then S is sequentially closed.

Proof. Let $\{x_n\}$ be a sequence in S such that $x = \lim_{n\to\infty} x_n$ exists in X. We have to prove that $x \in S$. Since S is sequentially compact, there exists a subsequence $\{x_{n_k}\}$ such that $y = \lim_{k\to\infty} x_{n_k} \in S$. But $\lim_{n\to\infty} x_n = x$. By uniqueness of the limit $x = y$. Hence $x \in S$. □

Remark: The converse is not true.
Ex: $X = \mathbb{R}$ is sequentially closed, but X is not sequentially compact.
Recall (Definition 2.8.1), $S \subseteq X$ is bounded if

$$\delta(S) = \sup\{d(x, y) \mid x, y \in S\} < \infty$$

or $S = \emptyset$.

This is equivalent to saying that there exists $M > 0$ such that $S \subseteq B_M(x_0)$ for some $x_0 \in X$, where $B_M(x_0) = \{x \in X \mid d(x, x_0) < M\}$ is the ball of radius M and center x_0.

Theorem 8.6 — Every compact set in a metric space is bounded.

Proof. Suppose that $S \neq \emptyset$. Assume by contradiction that S is not bounded. The idea is to construct a sequence $\{x_k\}$ in S which does not have a convergent subsequence. Let $x \in S$ be arbitrary. Note that it is impossible to have $d(x_n, x) < 1 \ \forall x_n \in S$, since otherwise $\delta(S) = 2$ (by triangular inequality $d(x, y) \leq d(x, x_n) + d(x_n, y) < 1 + 1$ for all $x, y \in S$). Hence, $\exists x_2 \in S$ s.t. $d(x_2, x) \geq 1$. Denote $d_1 = 1$, and $d_2 = d_1 + d(x_2, x) = 1 + d(x_2, x)$. It is impossible to have $d(x_3, x) < d_2 \ \forall x_3 \in S$ since otherwise $\delta(S) \leq d_2$, hence, $\exists x_3 \in S$ s.t. $d(x_3, x) \geq d_2$. Denote $d_3 = d_1 + d(x_3, x) = 1 + d(x_3, x)$. We continue in this manner. We construct the sequence $\{x_k\}$ in S and the sequence $\{d_n\}$ in $\mathbb{R}$ such that

$$d(x_n, x) > d_{n-1} \quad \text{and} \quad d_n = d_1 + d(x_n, x) \geq d_{n-1}$$

for all $n \geq 2$, we have:

$$\frac{d_n - d_{n-1}}{d_n - 1} \leq d(x_n, x) \leq d(x_{n+1}, x) + d(x_n, x_{n+1}) = d(x_{n+1}, x) + (d_n - 1)$$

(by triangular inequality) Therefore, $\forall n \geq 2, d(x_n, x_{n+1}) \geq 1$ Hence, $\{x_k\}$ is not a Cauchy sequence. Therefore, $\{x_k\}$ cannot have a convergent subsequence (since it cannot have any Cauchy subsequence) This contradicts our assumption that S is sequentially compact.

□

Discussion about the case $X = \mathbb{R}$. **Recall:** Theorem 1.7.11 (Bolzano-Weierstrass Theorem for sequences) In $\mathbb{R}$, any bounded sequence has a convergent subsequence.

Lemma 8.7 — A set $S \subseteq \mathbb{R}$ is sequentially compact if and only if it is sequentially closed and bounded[a].

[a] Discussion about the case $X = \mathbb{R}$

Proof. By Th. 4.1.4 and 4.1.5, we only have to prove the "if" part. Let $S \subseteq \mathbb{R}$ be sequentially closed and bounded. Let $\{x_n\}$ be a sequence in S. Since S is bounded, $\{x_n\}$ is also bounded. By Theorem 1.7.11, $\{x_n\}$ has a convergent subsequence $\{x_{n_k}\}$ to a limit x. Because S is sequentially closed, $x \in S$. This proves that S is sequentially compact. □

Theorem 8.8 — A subset of $\mathbb{R}^n$ is compact if and only if it is closed and bounded.

Proof. *Omitted. Not required for class.* □

§9 October 12, 2023

§9.1 Compactness Theorems

- A set $S \subset X$ is sequentially compact if for all sequences $\{x_n\} \subset S$ there exists sub sequence $\{x_{n_k}\}_k$ such that $x_{n_k} \to x \in S$ as $k \to \infty$
- If x is a sequentially compact then it is complete
- If $S \subset X$ is sequentially compact then S is sequentially closed.
- If $S \subset X$ is sequentially compact the S is bounded

⋆ **Remark:** A set $S \subseteq X$ is bounded if and only if $\exists x_0 \in X \ \exists r > 0$ such that

$$S \subseteq B_r(x_0) := \{x \in X; d(x, x_0) < r\}$$

- Suppose that S is bounded, i.e.

$$\delta(S) = \sup\{d(x, y)|x, y \in S\} < \infty$$

(Assume that $S \neq \emptyset$) Take $x_0 \in S$ arbitrarily. Then: $\forall x \in S$

$$d(x, x_0) < \delta(S) + 1 = n. \text{ This means that } S \subseteq B_n(x_0)$$

- Suppose that $S \subseteq B_n(x_0)$ for some $x_0 \in X$ and $n > 0$ Assume that $S \neq \emptyset$. For any $x, y \in S$, by triangular inequality

$$d(x, y) \leq d(x, x_0) + d(x_0, y) < 2n. \text{ This proves that } \delta(S) \leq 2n < \infty.$$

§9.2 Arzelà–Ascoli Theorem

Next we turn to the compact sets of $C[a, b]$. The motivation for the Arzelà–Ascoli theorem is that it gives us the necessary and sufficient conditions to decide weather every sequence of a family real-valued continuous functions defined on a closed and bounded interval has uniformly convergent subsequences.

Definition 9.1 Let F be a family (or set) of functions, each with domain D.

- We say that F is uniformly bounded on D if there is a positive number M such that $|f(x)| \leq M$ for all $f \in F$ and $x \in D$.
- We say that F is equicontinuous on D if, given any number $\varepsilon > 0$, there exists a number δ, for any $f \in F$

$$|f(x') - f(x'')| < \varepsilon, \quad \text{whenever } x', x'' \in D \text{ and } |x' - x''| < \delta$$

Remark:

1. Note that (1) is equivalent to:

$$\sup_{x \in D} |f(x)| \leq M \quad \forall f \in F$$

This means that

$$d(f, 0) = \max_{x \in [a,b]} |f(x) - 0(x)| \leq M \quad \forall f \in F$$

which is the same thing as saying that

$$F \subseteq B_{M+1}(0)$$

This means that F is uniformly bounded if and only if F is bounded

2. If F is equicontinuous then f is uniformly continuous $\forall f \in F$.

Example 9.2. Let $f_n : [0,1] \to \mathbb{R}$, $f_n(x) = x^n$ for all $n \geq 1$ integer. Let $F = \{f_n; n \geq 1\}$. We claim that F is NOT equicontinuous.

Proof. To see this, let $\varepsilon \in (0,1)$ be arbitrary. We have to find $\delta > 0$ such that for all $x \in [0,1]$ with $|1-x| < \delta$, we have

$$|1^n - x^n| = |1 - x^n| < \varepsilon \quad \forall n \geq 1 \quad (\text{Take } x = \tfrac{1}{n} \text{ in (2)})$$

This would mean that $x^n > 1 - \varepsilon \; \forall n \geq 1$, i.e., $x > (1-\varepsilon)^{\frac{1}{n}} \; \forall n \geq 1$.

Hence if $x > 1 - \delta$ then $x > (1-\varepsilon)^{\frac{1}{n}} \; \forall n \geq 1$.

This would be possible if and only if

$$1 - \delta > (1-\varepsilon)^{\frac{1}{n}} \quad \forall n \geq 1,$$

which means that $1-\delta$ is larger than the limit of $(1-\varepsilon)^{\frac{1}{n}}$ as n approaches infinity, which contradicts the fact that $\delta > 0$. ☐

Theorem 9.3 (Arzelà–Ascoli theorem) — A subset F of metric space $C[a,b]$ is compact if F is closed, uniformly bounded, and equicontinuous.

Proof. Omitted. Reading Material. ☐

A simple sufficinent condition for a family of functions to be equicontinuous is that all functions of the family satisfy Liptschitz condition with the same Liptschitz constant. A family $F \subset C[a,b]$ is equicontinious if $\exists K > 0$ such that

$$|f(x) - f(x')| \leq K|x' - x|, \quad \forall x, x' \in [a,b], \quad \forall f \in F$$

§9.3 Application to approximation theory

Theorem 9.4 — Let $A : X \to Y$ be a continuous mapping between metric spaces X and Y, and let S be a non-empty compact subset of X. Then the image $A(S)$ is compact subset of Y. [a]

[a] This theorem basically says that image under continuous mapping of a compact set is again a compact set

Proof. Let $\{y_n\}$ be a sequence in $A(S)$. For each $n \in \mathbb{N}$, there is at least one point $w \in S$ such that $Aw = y_n$. Choose one and call it x_n. Then x_n is a sequence in S and $Ax_n = y_n$. Since S is compact, $\{x_n\}$ has a convergent subsequence $\{x_{n_k}\}$, with limit x. Then $x \in S$, so $Ax \in A(S)$. Now, $A_{n_k} = y_{n_k}$ and $Ax_{n_k} \to Ax$ since A is continuous, so $\{y_{n_k}\}$ is a convergent subsequence of $\{y_n\}$. Hence, $A(S)$ is compact. ☐

Theorem 9.5 — If f is a real-valued continuous function mapping on metric space X and S is any nonempty compact set in X, then there exist points x_M and x_m in S such that

$$f(x_M) = \max_{s \in S} f(x) \quad \text{and} \quad f(x_m) = \min_{s \in S} f(x)$$

Proof. By previous theorem, $f(S) \subseteq \mathbb{R}$ is a compact set (in $\mathbb{R}$). And $f(S)$ is sequentially closed (in $\mathbb{R}$) and bounded. From MAT-2125 we prove that $\sup f(S) \in f(S)$ and $\inf f(S) \subseteq f(S)$ Hence $\exists x_{\max} \in S$ s.t. $\sup f(S) = f(x_{\max})$ and $\exists x_{\min} \in S$ s.t. $\inf f(S) = f(x_{\min})$. ☐

§10 October 16, 2023

§10.1 Application to approximation theory continued

Theorem 10.1 — Given a nonempty compact subset S of a metric space (X, d) and a point $x \in X$, there exists a point $p \in S$ such that $d(p, x)$ is a minimum.[a]

[a]The point p is called a *best approximation* in S of the point x in X

Proof. Let us define $f : X \to \mathbb{R}$ by $f(y) = d(y, x)$, for any $y \in X$. We show that f is continuous. Let $\{y_n\} \subseteq X$ such that $y_n \to y \in X$. Then

$$0 \leq |f(y_n) - f(y)| = |d(y_n, x) - d(y, x)| \leq d(y_n, y)$$

due to the triangular inequality. Hence $f(y_n) \to f(y)$. By Theorem 3.2, f attains its minimum, so that there exists $p \in X$ for which $f(p) = \min_{y \in S} f(y)$. □

Remark: p is called the best approximation of x in S.

Example 10.2 (Application of Ascoli's theorem)**.** Consider a family F of functions f defined as

$$f(x) = a \sin bx + c \cos dx, \quad 0 \leq x \leq \pi,$$

where the coefficients a, b, c, and d are chosen from a closed interval $[-M, M]$.

First, observe that the functions in F are uniformly bounded. For any f in F and any x in $[0, \pi]$, we have:

$$|f(x)| \leq |a| + |c| \leq 2M.$$

Additionally, the derivatives of functions in F satisfy:

$$|f'(x)| \leq |ab| + |cd| \leq 2M^2.$$

Thus, F is equicontinuous.

Given this, the family F can be treated as a subset of the continuous functions on the interval $[0, \pi]$. Due to its equicontinuity and boundedness, F is compact in this function space.

As a result, for any continuous function g defined on $[0, \pi]$, we can find a function within F that comes closest to g in the sense of minimizing the maximum deviation over the interval. Formally, there exist values of a, b, c, and d in $[-M, M]$ such that:

$$\max_{0 \leq x \leq \pi} |g(x) - (a \sin bx + c \cos dx)|$$

is minimized. A function in F achieving this minimal deviation is called a *minimax approximation* to g. It's important to note that this approximating function might not be unique.

§10.2 Topological Spaces

Definition 10.3 (Topological Spaces) A *topology* on a nonempty set X is a collection $\mathcal{T}$ of subsets of X satisfying the following properties:

1. $X \in \mathcal{T}$ and $\emptyset \in \mathcal{T}$ (Property T1).
2. The union of any subcollection $\mathcal{S}$ of $\mathcal{T}$ belongs to $\mathcal{T}$. Formally, for any subcollection $\mathcal{S}$ of $\mathcal{T}$:
$$\bigcup_{T \in \mathcal{S}} T \in \mathcal{T} \quad \text{(Property T2)}.$$
3. The intersection of any two sets T_1 and T_2 in $\mathcal{T}$ belongs to $\mathcal{T}$. Formally:
$$T_1 \cap T_2 \in \mathcal{T} \quad \text{whenever} \quad T_1, T_2 \in \mathcal{T} \quad \text{(Property T3)}.$$

The pair $(X, \mathcal{T})$ is called a *topological space*. The sets T in $\mathcal{T}$ are referred to as the *open sets* in $(X, \mathcal{T})$. A subset S of X is said to be *closed* in $(X, \mathcal{T})$ if its complement, denoted by $X \setminus S$ or $\sim S$, is an open set in $(X, \mathcal{T})$.

Example 10.4 (Standard Topology on $\mathbb{R}$)**.** The most common topology on the real numbers is the *standard topology*, which is generated by the open intervals. In this topology, a set $U \subseteq \mathbb{R}$ is open if for every point $x \in U$, there exists an open interval (a, b) such that $x \in (a, b) \subseteq U$.

Example 10.5 (Lower Limit Topology on $\mathbb{R}$)**.** Another example is the *lower limit topology* on $\mathbb{R}$, which is generated by the half-open intervals of the form $[a, b)$. A set is open in this topology if it can be expressed as a union of such half-open intervals.

Example 10.6 (Why Infinite Intersections Aren't Always Open)**.** While the arbitrary union of open sets is open (by definition of a topology), the same isn't true for infinite intersections. To see why, consider the standard topology on $\mathbb{R}$. Take the nested sequence of open intervals:

$$I_n = \left(-\frac{1}{n}, \frac{1}{n} \right), \quad n = 1, 2, 3, \ldots$$

The intersection of all these intervals is $\{0\}$, which is not an open set in the standard topology on $\mathbb{R}$. Hence, while each I_n is open, their infinite intersection is not. This demonstrates the necessity of the limitation in the definition of a topology that only finite intersections of open sets are guaranteed to be open.

⋆ Given any set X, there are two fundamental topologies: The *discrete topology* on X, denoted $T_{\max}$, is the collection of all subsets of X. Formally,

$$T_{\max} = \mathcal{P}(X),$$

where $\mathcal{P}(X)$ represents the power set of X. In this topology, every subset of X is considered open. The *indiscrete topology* or *trivial topology* on X, denoted $T_{\min}$, consists of only the empty set and X itself. Formally,

$$T_{\min} = \{\emptyset, X\}.$$

In this topology, no set other than the entire set and the empty set is considered open. Clearly, $T_{\max}$ and $T_{\min}$ satisfy the properties of a topology, and as their names suggest, they represent the largest and smallest collections of subsets of X that can be considered as topologies.

Definition 10.7 (Weaker and Stronger Topologies) Given two topologies $\mathcal{T}_1$ and $\mathcal{T}_2$ on a set X, if

$$\mathcal{T}_1 \subseteq \mathcal{T}_2,$$

then $\mathcal{T}_1$ is said to be *weaker* (or *coarser*) than $\mathcal{T}_2$, and $\mathcal{T}_2$ is said to be *stronger* (or *finer*) than $\mathcal{T}_1$. Given any topology $\mathcal{T}$ on X, it must hold that

$$\mathcal{T}_{\text{min}} \subseteq \mathcal{T} \subseteq \mathcal{T}_{\text{max}},$$

where $\mathcal{T}_{\text{min}}$ is the *indiscrete topology* and $\mathcal{T}_{\text{max}}$ is the *discrete topology*. Thus, among all possible topologies on a set X, the indiscrete topology is always the weakest (or coarsest), and the discrete topology is always the strongest (or finest).

Example 10.8 (Illustrative Example of Topologies on a Finite Set)**.** Consider the set $X = \{1, 2, 3, 4, 5\}$ and the collections:

$$\begin{aligned}
\mathcal{T}_1 &= \{\emptyset, \{1\}, \{2\}, \{1, 2\}, X\}, \\
\mathcal{T}_2 &= \{\emptyset, \{1\}, \{2\}, \{1, 2\}, \{1, 2, 3\}, \{1, 2, 3, 4\}, X\}, \\
\mathcal{T}_3 &= \{\emptyset, \{1\}, \{1, 2\}, \{1, 2, 3\}, X\}, \\
\mathcal{T}_4 &= \{\emptyset, \{1\}, \{2\}, \{1, 2\}, \{2, 3, 4\}, X\}.
\end{aligned}$$

It is evident that $\mathcal{T}_1$ and $\mathcal{T}_2$ are topologies on X. $\mathcal{T}_1$ is weaker than $\mathcal{T}_2$ since $\mathcal{T}_1 \subseteq \mathcal{T}_2$. On the other hand, $\mathcal{T}_3$ is another topology for X that is weaker than $\mathcal{T}_2$ but is neither weaker nor stronger than $\mathcal{T}_1$. For the topological space $(X, \mathcal{T}_2)$, the closed sets are:

$$X, \{2, 3, 4, 5\}, \{1, 3, 4, 5\}, \{3, 4, 5\}, \{4, 5\}, \{5\} \text{ and } \emptyset.$$

The set $\{2, 3\}$ is neither open nor closed in this topology. The interior of $\{2, 3\}$ is $\{2\}$ and its closure is $\{2, 3, 4, 5\}$. However, $\mathcal{T}_4$ is not a topology on X because the union $\{1\} \cup \{2, 3, 4\} = \{1, 2, 3, 4\}$ is not an element of $\mathcal{T}_4$, violating the property that arbitrary unions of open sets should remain open.

Definition 10.9 (Interior and Closure in Topological Spaces) Let X be a topological space.

Interior: The *interior* of a subset S of X is the union of all open sets contained in S. It is denoted by $\text{int}(S)$ or S°.

Closure: The *closure* of a subset S of X is the intersection of all closed sets containing S. It is denoted by $\text{cl}(S)$ or $\overline{S}$.

Every metric space has an associated topology, called the *metric topology*, derived from its metric. This enables us to study metric spaces in the context of topological spaces.

Definition 10.10 (Metric Topology in Metric Spaces) Let (X, d) be a metric space.

- **Open Ball:** For a point x_0 in X and a positive real number r, the set

$$\{x : x \in X, d(x, x_0) < r\}$$

is called an *open ball* in X. It represents the set of all points in X that are less than r distance away from x_0. This set is denoted by $b(x_0, r)$ and is referred to as the open ball with center x_0 and radius r.

- **Open Sets:** A subset T of X is called *open* if $T = \emptyset$ or for every point in T, there exists an open ball centered at that point which is completely contained in T.

- **Metric Topology:** The metric topology $\mathcal{T}_d$ for X is the collection of all open sets, as defined above. Thus, a subset T of X is in $\mathcal{T}_d$ if and only if for each x in T, there exists some radius r such that the open ball $b(x, r)$ is a subset of T.

[a]

[a] It can be shown that the collection $\mathcal{T}_d$ of open sets satisfies the properties of a topology, making $(X, \mathcal{T}_d)$ a topological space. By convention, when we refer to a metric space as a topological space, it is implied that we are considering its associated metric topology.

§11 October 19, 2023

Recall that if (X, d) is a metric space, a set U is open if $\forall x \in U, \exists \epsilon > 0$ such that $B(x, \epsilon) \subseteq U$ or $U = \varnothing$.

- **Real Numbers $\mathbb{R}$:**
  - Metric: $d(x, y) = |x - y|$
  - Open ball: $B(x_0, \epsilon) = (x_0 - \epsilon, x_0 + \epsilon)$
  - **Example:** For $x_0 = 2$ and $\epsilon = 0.5$, the open ball is $B(2, 0.5) = (1.5, 2.5)$.
- **Euclidean Plane $\mathbb{R}^2$:**
  - Open ball: Set of all points inside the circle of radius ϵ centered at (x_0, y_0) without the boundary.
  - **Example:** For $(x_0, y_0) = (1, 2)$ and $\epsilon = 1$, the open ball is the interior of the circle of radius 1 centered at $(1, 2)$.
- **Euclidean 3-space $\mathbb{R}^3$:**
  - Open ball: Set of all points inside the sphere of radius ϵ centered at (x_0, y_0, z_0) without the boundary.
  - **Example:** For $(x_0, y_0, z_0) = (1, 2, 3)$ and $\epsilon = 1$, the open ball is the interior of the sphere of radius 1 centered at $(1, 2, 3)$.
- **Continuous Functions $C[a, b]$:**
  - Metric: $d(f, g) = \sup_{x \in [a,b]} |f(x) - g(x)|$
  - Open ball: Set of all functions g such that $d(f_0, g) < \epsilon$.
  - **Example:** Let $f_0(x) = x$ on $[0, 1]$. The open ball of radius $\epsilon = 0.5$ centered at f_0 contains functions g on $[0, 1]$ with $\sup_{x \in [0,1]} |x - g(x)| < 0.5$.

§11.1 Closed Sets

Recall: Let (X, d) be a metric space. A set $S \subset X$ is said to be sequentially closed if for every sequence $\{x_n\}$ in S that converges in X, we have $\lim_{n \to \infty} x_n = x$ implies $x \in S$. If (X, τ) is a topological space, a set $S \subseteq X$ is called closed if its complement S^c in X is open.

Theorem 11.1 — Let (X, d) be a metric space and let τ_d denote the metric topology induced by d. A set $S \subseteq X$ is closed in the metric topology if and only if it is sequentially closed in X.

Proof. Only if part: Suppose that $S \subseteq X$ is closed. We have to prove that S is sequentially closed. Let $\{x_n\}$ be a sequence in S such that $x = \lim_{n \to \infty} x_n$ exists. We want to show that $x \in S$. Suppose that $x \in S^c$. Since S is closed, S^c is open. Hence, $\exists \epsilon > 0$ such that $B(x, \epsilon) \subseteq S^c$. Since $x_n \to x$, there exists $N \in \mathbb{N}$ such that $x_n \in B(x, \epsilon)$ for all $n > N$. So $x_n \in S^c$ $\forall n > N$. This is a contradiction since $\{x_n\}$ is a sequence in S. Recall that: $x_n \to x$ means that $d(x_n, x) \to 0$.

" $\implies$ " : Now suppose that $S \subseteq X$ is sequentially closed. We have to prove that S is closed, i.e., S^c is open. Suppose that S^c is not open. This means that there exists an $x \in S^c$ such that for all $\delta > 0$, $B(x, \delta) \nsubseteq\in S^c$. [Complete the proof] □

⋆ Given a topological space (X, τ) it is not always possible to find a distance d on X such that $\tau = \tau_d$. If this is possible, we that τ is **metrizable**.

Definition 11.2 Let (X, τ) be a topological space.

(a) If $x \in X$ and $U \subseteq X$ is an open set such that $x \in U$, then we say that U is a neighbourhood of x.

(b) A point $x \in X$ is called a cluster point (or accumulation point) of a set $S \subseteq X$ if every neighbourhood U of x intersects S in at least one point other than x (when $x \in S$).

(c) The set of all cluster points of S is called the derived set (or limit point set) of S and is denoted by S'.

(d) A point $x \in X$ is called an adherent point of S if, for every neighbourhood U of x, $U \cap S \neq \varnothing$. This implies that x can be either in S or not in S[a].

[a] So, every cluster point is indeed a adherent (closure) point. However, not all closure points are cluster points because a adherent (closure) point could be a point in S that doesn't have other points from S arbitrarily close to it.

Lemma 11.3 — Let (X, τ) be a topological space and $S \subseteq X$. Then $x \in \bar{S} \iff x$ is an adherent point.

Proof. Complete the proof. Its pretty long. □

Lemma 11.4 — Let $(X, \mathcal{T})$ be a topological space and $S \subseteq X$. Then $x \in \bar{S}$ (where $\bar{S}$ is the closure of S) if and only if x is an adherent point of S, where the closure $\bar{S}$ of S was defined as follows:

$$\bar{S} = \bigcap \{F \mid F \text{ closed}, F \supseteq S\} \quad (\text{i.e., } \bar{S} \text{ is the smallest closed set which contains } S)$$

Proof. ($\Rightarrow$) Suppose that x is an adherent point of S. We need to prove that $x \in \bar{S}$. By contradiction, suppose that $x \in (\bar{S})^c$. Note that

$$(\bar{S})^c = \bigcup \{F^c \mid F \text{ closed}, F \supseteq S\} = \bigcup \{U \mid U \text{ open}, U \subseteq S^c\}$$

by De Morgan's laws. This means that there exists an open set U with $U \subseteq S^c$ such that $x \in U$. Hence $U \cap S = \emptyset$. This contradicts the definition of x being an adherent point of S; namely, we were able to find a neighborhood U of x such that $U \cap S = \emptyset$.

($\Leftarrow$) Suppose that $x \in \bar{S}$. We want to prove that x is an adherent point. By contradiction, suppose that x is not an adherent point by contradiction, suppose that x is not an adherent point. Then, there exists a neighborhood U of x such that $U \cap S = \emptyset$. Recall: U is open and $x \notin U$. $U \cap S = \emptyset$ is equivalent to saying that $S \subseteq U^c$ and U^c is closed and contains S, hence $\bar{S} \subseteq U^c$. From here, we deduce that $U \subseteq \bar{S}^c$. Hence $x \in \bar{S}^c$ (since $x \in U$). This is a contradiction. □

§12 October 30, 2023

Recall: $(X, \mathcal{T})$ is a topological space, $S \subseteq X$.

- We say that x is a cluster point of S if $(U \setminus \{x\}) \cap S \neq \emptyset$ for any neighborhood U of x.
- We say that x is an adherence point of S if $U \cap S \neq \emptyset$ for any neighborhood U of x.

Lemma 12.1 — Assume that (X, d) is a metric space endowed with the metric topology. Then $x \in X$ is an adherent point of S if and only if there is a sequence $\{x_n\}$ in S such that $\lim_{n\to\infty} x_n = x$.

Proof. a) Suppose that x is an adherent point. For any $n \geq 1$, there exists $x_n \in B(x, \frac{1}{n}) \cap S$. Then $d(x_n, x) < \frac{1}{n}$ for any n. So $\lim_{n\to\infty} x_n = x$.

b) Suppose that there exists a sequence $\{x_n\}$ in S such that $\lim_{n\to\infty} x_n = x$. We have to show that x is an adherent point of S. Let U be a neighborhood of x. Then there exists $\delta > 0$ such that $B(x, \delta) \subseteq U$. Since $x_n \to x$, there exists $N \in \mathbb{N}$ such that $x_n \in B(x, \delta)$ for all $n \geq N$. So $x_n \in U \cap S$ for all $n \geq N$. Hence $U \cap S \neq \emptyset$. □ □

§12.1 Closed Sets continued

Theorem 12.2 — A set S in a topological space is closed if and only if it contains cluster points, that is $S \supseteq S'$[a].

[a] S' is the set of all cluster point i.e., the derived set

Proof. Prove this. □

§12.2 Compact Sets

In any metric space (X, d) containing at least two distinct points x and y, we can always find open balls centered at x and y that are disjoint. For example, take the open balls $B(x, r)$ and $B(y, r)$, with $r < \frac{1}{2}d(x, y)$. Not all topological spaces possess this property. However, this property is fundamental for many analyses in topology. Spaces with this property are given a special name: they are called *Hausdorff spaces* or *T2 spaces*.

Definition 12.3 (Compactness) A subset S of a topological space is said to be *compact* if for every collection $\{U_\alpha\}$ of open sets such that $S \subseteq \bigcup_\alpha U_\alpha$ (i.e., the union of the sets in the collection contains S), there exists a finite sub-collection $\{U_{\alpha_1}, U_{\alpha_2}, \ldots, U_{\alpha_n}\}$ such that $S \subseteq \bigcup_{i=1}^n U_{\alpha_i}$.

Theorem 12.4 (Equivalence of Compactness Theorem) — Let (X, d) be a metric space. Then, a subset $S \subseteq X$ is compact if and only if it is sequentially compact.

Proof. Omitted. □

Definition 12.5 (Hausdorff Space) A topological space $(X, \mathcal{T})$ is called a *Hausdorff space*, and $\mathcal{T}$ is called a *Hausdorff topology*, if for every pair of distinct points $x, y \in X$, there exists a neighbourhood U_x of x and a neighbourhood U_y of y such that $U_x \cap U_y = \emptyset$.

In essence, a space X is termed Hausdorff when any two distinct points within it can be enclosed by non-overlapping neighbourhoods. It's established that all metric spaces possess this Hausdorff property. Likewise, any set endowed with the discrete topology, denoted as $\mathcal{T}_{\max}$, is also Hausdorff. In contrast, the indiscrete topology, $\mathcal{T}_{\min}$, doesn't meet the Hausdorff criterion unless it's defined over a set with a single point or none. Amidst the diverse spectrum of topological properties, the Hausdorff characteristic stands out as a fundamental trait shared by numerous, albeit not all, topological spaces.

Lemma 12.6 (Metric Spaces are Hausdorff) — Let (X, d) be a metric space, and let τ_d be the topology induced by d. Then, the topological space (X, τ_d) is a Hausdorff space.

Proof. Let $x, y \in X$ be distinct points. Since $d(x, y) > 0$, we can choose $\epsilon = \frac{d(x,y)}{2}$. Consider the open balls $B(x, \epsilon)$ and $B(y, \epsilon)$. Clearly, $x \in B(x, \epsilon)$ and $y \in B(y, \epsilon)$. Moreover, $B(x, \epsilon)$ and $B(y, \epsilon)$ are disjoint because for any point z in their intersection, we would have $d(x, z) < \epsilon$ and $d(y, z) < \epsilon$, which contradicts the triangle inequality. Thus, we have found disjoint open neighborhoods for x and y, and so (X, τ_d) is Hausdorff. □

Theorem 12.7 (Compact Sets in Hausdorff Spaces) — Let (X, τ) be a Hausdorff topological space. Then, every compact subset K of X is closed.

Proof. To show that K is closed, it suffices to show that its complement $X \setminus K$ is open. Let $x \in X \setminus K$. For each $y \in K$, since (X, τ) is Hausdorff, there exist open sets U_y and V_y such that $x \in U_y$, $y \in V_y$, and $U_y \cap V_y = \emptyset$. The collection $\{V_y\}$ forms an open cover for K. Since K is compact, there exists a finite subcollection $\{V_{y_1}, V_{y_2}, \ldots, V_{y_n}\}$ that covers K. Let $U = \cap_{i=1}^{n} U_{y_i}$. Then, U is an open set containing x that does not intersect K. Hence, $X \setminus K$ is open, and K is closed. □

Theorem 12.8 (Heine-Borel Theorem) — Let $S \subseteq \mathbb{R}$. The subset S is compact in $\mathbb{R}$ with the standard topology if and only if S is both closed and bounded.

Proof. ($\Rightarrow$) **Direction:** Suppose S is compact.

- *Boundedness:* If S were not bounded, for each $n \in \mathbb{N}$, pick $x_n \in S$ such that $|x_n| > n$. This would produce a sequence without a convergent subsequence in S, which contradicts the compactness of S.
- *Closedness:* Let (x_n) be a sequence in S that converges to x. Since S is compact, the sequence has a convergent subsequence that also converges to x. Thus, x must be in S, implying S is closed.

($\Leftarrow$) **Direction:** Suppose S is both closed and bounded. By Bolzano-Weierstrass, any sequence in S has a convergent subsequence. Since S is closed, the limit of this subsequence also lies in S. Hence, S is compact by the sequential characterization of compactness. □

§12.3 Continuity in Topological spaces

In calculus, we typically define continuity in terms of limits and the behavior of functions on the real numbers. However, in the broader setting of topological spaces, we use the open set definition, which captures the essence of continuity in a more general setting.

Definition 12.9 (Sequential Continuity in Metric Spaces) Let (X, d) and (Y, d') be two metric spaces. A function $f : X \to Y$ is said to be *sequentially continuous* at a point $x \in X$ if for every sequence $\{x_n\}$ in X that converges to x (i.e., $x_n \to x$ as $n \to \infty$), the sequence $\{f(x_n)\}$ in Y converges to $f(x)$ (i.e., $f(x_n) \to f(x)$ as $n \to \infty$.

Definition 12.10 (Continuity on $\mathbb{R}$) Let $f : \mathbb{R} \to \mathbb{R}$ be a function and let c be a point in its domain. The function f is said to be *continuous at* c if, for every $\epsilon > 0$, there exists a $\delta > 0$ such that, for all x in $\mathbb{R}$ satisfying $0 < |x - c| < \delta$, we have $|f(x) - f(c)| < \epsilon$.

The function f is said to be *continuous on* $\mathbb{R}$ if it is continuous at every point c in its domain.

Definition 12.11 Let $(X, \mathcal{T})$ be a topological space.

(a) We say that a sequence $\{x_n\}$ in X converges to a limit $x \in X$ if for any neighbourhood U of x, there exists $N \in \mathbb{N}$ such that $x_n \in U$ for all $n > N$.

(b) Let $(Y, \mathcal{T}')$ be another topological space and $f : X \to Y$. We say that f is sequentially continuous at $x \in X$ if, for every sequence $\{x_n\} \subset X$ such that $x_n \to x$, we have $f(x_n) \to f(x)$.

§13 November 2, 2023

§13.1 Continuous functions in topological spaces

Let (X, Y), (Y, Y') be topological spaces; $A : X \to Y$.

- A is necessary continuous if for every sequence $x_n \to x$ in X then $Ax_n \to Ax$ in Y.
- A is continuous at x if for every neighborhood V of Ax there exists a neighborhood U of x such that if $y \in U$ then $Ay \in V$.

(Recall: $f : \mathbb{R} \to \mathbb{R}$ is continuous at x if for every $\varepsilon > 0$ there exists $\delta > 0$ such that if $y - x| < \delta$ then $|f(y) - f(x)| < \varepsilon$.)[1]

Theorem 13.1 — Let A be a function from the topological space X to the topological space Y. The following statements are considered equivalent:

1. A is continuous over X.
2. The preimage of every open set in Y under A is an open set in X.
3. The preimage of every closed set in Y under A is a closed set in X.

Proof. ⋆ The equivalence of these statements offers an alternative definition of continuity: a function is continuous if the preimage of every open set in the codomain is an open set in the domain. The proof follows the logical sequence: (1) ⇒ (2) ⇒ (3) ⇒ (2) ⇒ (1). To demonstrate that statement 1 implies statement 2, assume that A is continuous and consider any open set T in Y. For any element x in X such that Ax is in T, there exists a neighborhood around x that is completely contained within the preimage of T under A. As a result, this preimage is a union of open sets and therefore open itself. To show that statement 2 implies statement 3, start with a closed set S in Y, so the complement of S is open. By assumption, its preimage under A is open in X, which implies that the preimage of S is closed in X. To conclude that statement 3 implies statement 2, use a similar argument by interchanging the roles of open and closed sets. Finally, to confirm that statement 2 implies statement 1, take any point x in X and a neighborhood V of Ax in Y. The preimage of V is an open set that contains x, establishing the continuity of A at x.

(a) ⇒ (b) Let T be an open set in Y. Since A is continuous at x, for any $x \in X$, there exists a neighborhood U_x of x such that $U_x \subseteq A^{-1}(T)$. **Claim:** $\bigcup_{x \in X} U_x = A^{-1}(T)$
Hence, $A^{-1}(T)$ is an open set, since it is the union of open sets.
(b) ⇒ (a) Let $x \in X$ be arbitrary and V a neighborhood of Ax. Then by (b), $A^{-1}(V)$ is open in X. Moreover, $Ax \in V$ and so $x \in A^{-1}(V)$. So, $A^{-1}(V)$ is a neighborhood of x. Take $U = A^{-1}(V)$. Clearly, $U \subseteq A^{-1}(V)$, and A is continuous.
(b) ⇒ (c) Let S be a closed subset in Y. Then S^c is an open set in Y. By (b), $A^{-1}(S^c)$ is open in X.
Recall: Theorem 5.4.2 (a): $A^{-1}(S^c) = [A^{-1}(S)]^c$
In our case, $A^{-1}(S^c) = [A^{-1}(S)]^c$ is open. Hence, $A^{-1}(S)$ is closed.
(c) ⇒ (b) Same argument as for (b) ⇒ (c), swapping the words "closed" and "open". □

Theorem 13.2 — Let $(X, \mathcal{T})$ and $(Y, \mathcal{T}')$ be topological spaces, and let $A : X \to Y$. If A is continuous on X, then A is sequentially continuous.

Proof. Let $x \in X$ and $\{x_n\}$ be a sequence in X such that $x_n \to x$. We need to prove that $Ax_n \to Ax$. Let V be a neighborhood of Ax. By the continuity of A, $A^{-1}(V)$ is an open set in X containing x, which means that $x \in A^{-1}(V)$.
Hence, $A^{-1}(V)$ is a neighborhood of x. Since $x_n \to x$, there exists $N \in \mathbb{N}$ such that $x_n \in A^{-1}(V)$ for all $n \geq N$. This means that $Ax_n \in V$ for all $n \geq N$. Thus, $Ax_n \to Ax$. □

[1]**Remark:** (1) says $U \subseteq A^{-1}(V) = \{x \in X; Ax \in V\}$.

Theorem 13.3 — Let (X, d) and (Y, d') be metric spaces, and $A : X \to Y$. If A is sequentially continuous on X, then A is continuous on X.

Proof. Let $x \in X$ be arbitrary, $y = Ax$ and V a neighborhood of y. Since Y is endowed with the metric topology, V is open in this topology, which means that there exists $\varepsilon > 0$ such that $B_y(y, \varepsilon) \subseteq V$.

Suppose that A is not continuous at x. Then there is no $\delta > 0$ such that $B_x(x, \delta) \subseteq A^{-1}(B_y(y, \varepsilon))$ (because if there exists such a δ, then $B_x(x, \delta) \subseteq A^{-1}(B_y(y, \varepsilon)) = A^{-1}(V)$ and then A would be continuous at x).

This means that for all $\varepsilon > 0$, $B_x(x, \varepsilon) \not\subseteq A^{-1}(B_y(y, \varepsilon))$ (i.e., there exists $x_\varepsilon \in B_x(x, \varepsilon)$ and $x_\varepsilon \notin A^{-1}(B_y(y, \varepsilon))$.

Take $\delta = \frac{1}{n}$ and denote $x_\delta = x_{\frac{1}{n}} \in B_x(x, \frac{1}{n})$ and $x_\delta \notin A^{-1}(B_y(y, \varepsilon))$. This means that $d(x, x_\delta) < \frac{1}{n}$ for all $n \geq 1$. Hence $x_\delta \to x$. Since A is sequentially continuous, $Ax_\delta \to Ax = y$. This means that there exists $N \in \mathbb{N}$ such that $Ax_\delta \in B_y(y, \varepsilon)$ for all $n \geq N$. Therefore $x_\delta \in A^{-1}(B_y(y, \varepsilon))$ for all $n \geq N$. This is a contradiction.

Hence, A must be continuous at x. □

§13.2 Normed Vector Spaces

Definition 13.4 A vector space over a field F (where $F = \mathbb{R}$ or $F = \mathbb{C}$) is a set $X \neq \emptyset$ together with two binary operations:

- $+ : X \times X \to X$ which maps (x, y) to $x + y$,
- $\cdot : F \times X \to X$ which maps (α, x) to αx,

such that the following properties hold:

1. Associativity of addition: $(x + y) + z = x + (y + z)$ for all $x, y, z \in X$
2. Commutativity of addition: $x + y = y + x$
3. Identity element of addition: There exists $0 \in X$ such that $x + 0 = x$ for all x
4. Inverse elements of addition: For every $x \in X$ there exists $-x \in X$ such that $x + (-x) = 0$
5. Compatibility of scalar multiplication with field multiplication: $a(bx) = (ab)x$ for all $a, b \in F$ and $x \in X$
6. Identity element of scalar multiplication: There exists $1 \in F$ such that $1 \cdot x = x$ for all x
7. Distributivity of scalar multiplication with respect to vector addition: $a(x+y) = ax+ay$ for all $a \in F$, $x, y \in X$
8. Distributivity of scalar multiplication with respect to field addition: $(a+b)x = ax+bx$ for all $a, b \in F$, $x \in X$

Definition 13.5 A normed vector space (or simply normed space) is a vector space together with a map $\|\cdot\| : X \to [0, \infty)$, called a norm, that the following properties hold:

(N1) $\|x\| = 0$ if and only if $x = 0$.

(N2) $\|ax\| = |a| \cdot \|x\|$ for all $a \in \mathbb{F}$, $x \in X$.

(N3) (Triangular inequality) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.

Remark 13.6. If $(X, \|\cdot\|)$ is a normed space, then (X, d) is a metric space, where $d(x, y) = \|x-y\|$, properties $(M1)$ to $(M3)$ hold:

(M1) $d(x, y) = 0$ if and only if $x = y$ (clear due to $(N1)$).

(M2) $d(x, y) = \|x - y\| = \|y - x\| = d(y, x)$.

(M3) $d(x, z) \leq \|x - y\| + \|y - z\| = d(x, y) + d(y, z)$.

A ball in a normed space is:

$$B(x, \varepsilon) = \{y \in X; \|y - x\| < \varepsilon\}$$

Example 13.7. Examples of normed spaces include:

- $X = \mathbb{R}^n$ with the Euclidean norm $\|x\| = \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}}$.
- $X = C^n$ with the maximum norm $\|x\|_\infty = \max_{i=1,\ldots,n} |x_i|$.
- $X = \ell^p$ for $p \geq 1$ with the ℓ^p norm $\|x\|_p = (\sum_{i=1}^{\infty} |x_i|^p)^{\frac{1}{p}}$.

Remark 13.8. Question: If (X, d) is a metric space, does it imply we have a norm $\|x\|$ on X such that $d(x, y) = \|x - y\|$ for all $x, y \in X$? No.

Suggestion: Take $\|x\| = d(x, 0)$ and check properties (M1)–(M3). (M1) clearly holds.

(M2): If $\|ax\| = d(ax, 0)$, we don't know if this is equal to $|a|d(x, 0)$.

Normed spaces are contained in the set of metric spaces, which in turn are contained in the set of topological spaces.

Example 13.9. Example of a metric space that is not a normed space: Let $X = \ell^2 = \left\{x = \{x_i\}_{i=1}^{\infty} \mid x_i \in \mathbb{C}, \sum_{i=1}^{\infty} |x_i|^2 < \infty\right\}$. Then ℓ^2 is a normed space with the norm $\|x\| = \left(\sum_{i=1}^{\infty} |x_i|^2\right)^{\frac{1}{2}}$.

§14 November 6, 2023

§14.1 Normed Spaces

Example 14.1. For $X = \mathbb{R}^n$, the norm is defined as

$$\|x\| = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{\frac{1}{2}}.$$

For $X = \ell^2$, the norm is given by

$$\|x\| = \left(\sum_{i=1}^{\infty} |x_i|^2 \right)^{\frac{1}{2}}.$$

For $X = C[a, b]$, we have several norms:

- The uniform norm: $\|x\| = \max_{t \in [a,b]} |x(t)|$.
- The L^1 norm: $\|x\|_1 = \int_a^b |x(t)| dt$.
- The L^p norm for $p \geq 1$: $\|x\|_p = \left(\int_a^b |x(t)|^p dt \right)^{\frac{1}{p}}$.

§14.2 Convergence in Normed Spaces

A sequence $\{x_n\}$ in a normed space $(X, \|\cdot\|)$ converges to $x \in X$ if

$$\forall \varepsilon > 0 \; \exists N \in \mathbb{N} \text{ such that } \|x_n - x\| < \varepsilon \; \forall n \geq N.$$

A sequence $\{x_n\}$ is Cauchy if

$$\forall \varepsilon > 0 \; \exists N \in \mathbb{N} \text{ such that } \|x_n - x_m\| < \varepsilon \; \forall n, m \geq N.$$

Remark 14.2. Observation: If $\{x_n\}$ is convergent then $\{x_n\}$ is Cauchy.

Definition 14.3 A Banach space is a complete normed space.

Recall: "Complete" means that every Cauchy sequence is convergent.
Examples:

- $\mathbb{R}^n$ with Euclidean norm are Banach spaces.
- ℓ^2 is a Banach space.
- $C[a, b]$ with uniform norm is a Banach space.

Remark 14.4. If X is a normed space and $A : X \to X$ is a map, then A is a contraction if there exists $\alpha \in (0, 1)$ such that:

$$\|Ax - Ay\| \leq \alpha \|x - y\| \quad \forall x, y \in X.$$

⋆ A contraction is a continuous map: if $x_n \to x$ then $Ax_n \to Ax$ sequentially because $\|Ax_n - Ax\| \leq \alpha \|x_n - x\| \to 0$.

Theorem 14.5 (Fixed Point Theorem) — Let $(X, \|\cdot\|)$ be a Banach space. If $A: X \to X$ is a contraction, then there is a unique $x \in X$ such that $Ax = x$.

Let $(X, \|\cdot\|)$ be a normed space. A subset $S \subseteq X$ is sequentially compact if for every sequence $\{x_n\} \subseteq S$ there is a convergent subsequence $\{x_{n_k}\}$ whose limit x is also in S.[2]

A subset $S \subseteq X$ is compact if every covering of S has a finite subcover.

(a) If S is compact then S is closed (Th 5.3.3).

(b) S is compact if and only if S is sequentially compact (Th 5.3.3).

(c) If S is sequentially compact then S is bounded (Th 4.1.5).

Definition 14.6 Let $(X, \|\cdot\|)$ be a normed space and a sequence $\{x_k\} \subseteq X$.

1. We say that the series $\sum_{k=1}^{\infty} x_k$ converges if the partial sum sequence $\{\sum_{k=1}^{n} x_k\}_{n\geq 1}$ converges to a limit $x \in X$. In this case, we say that x is the sum of the series and we write $\sum_{k=1}^{\infty} x_k = x$.
2. We say that the series $\sum_{k=1}^{\infty} x_k$ converges absolutely if the series $\sum_{k=1}^{\infty} \|x_k\|$ (of non-negative real numbers) converges.

Theorem 14.7 (Theorem from MAT 2125) — If $X = \mathbb{R}$ and $\sum_{k=1}^{\infty} x_k$ converges absolutely, then $\sum_{k=1}^{\infty} x_k$ converges.

Remark 14.8. Observation: The converse may not be true.

Theorem 14.9 — Let $(X, \|\cdot\|)$ be a normed space. Then X is a Banach space if and only if every absolutely convergent series in X is convergent.

Proof. "Only If" Part: Suppose that X is a Banach space. We have to prove that every absolutely convergent series is convergent. Let $\sum_{k=1}^{\infty} x_k$ be an absolutely convergent series in X. By definition, $\sum_{k=1}^{\infty} \|x_k\|$ converges in $\mathbb{R}$. Hence the partial sum sequence $\{S_n\}$, where $S_n = \sum_{k=1}^{n} \|x_k\|$, converges to a limit $S \in \mathbb{R}$.

Then $\{S_n\}$ is a Cauchy sequence in $\mathbb{R}$, i.e., $\forall \varepsilon > 0\ \exists N \in \mathbb{N}$ such that $|S_n - S_m| < \varepsilon\ \forall n, m > N$.

Note that for any $n > m > N$,

$$S_n - S_m = \sum_{k=m+1}^{n} \|x_k\| \geq 0,$$

so $S_n \geq S_m$. Let $x_n = \sum_{k=1}^{n} x_k$. Then $x_n - x_m = \sum_{k=m+1}^{n} x_k$, and so, by the triangular inequality,

$$\|x_n - x_m\| \leq \sum_{k=m+1}^{n} \|x_k\| = S_n - S_m < \varepsilon$$

for all $n > m > N$. This proves that $\{x_n\}$ is a Cauchy sequence in X. Since X is a Banach space, $\{x_n\}$ is convergent. Hence, the series $\sum_{k=1}^{\infty} x_k$ is convergent.

"If" Part: Suppose that every absolutely convergent series in X is convergent. We have to prove that X is a Banach space. Let $\{x_n\}$ be a Cauchy sequence in X. We will use the following result: Theorem : Let (X, d) be a metric space, if $\{x_n\}$ is a Cauchy sequence in X with a convergent subsequence $\{x_{n_k}\}$ with limit x, then $x_n \to x$ as $n \to \infty$.

Since $\{x_n\}$ is Cauchy, for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $\|x_n - x_m\| < \varepsilon\ \forall n, m > N$. Take $\varepsilon = \frac{1}{2^k}$ for $k \in \mathbb{N}$ and let $N_k = N_{\frac{1}{2^k}}$. We may assume that $N_1 < N_2 < N_3 < \ldots$. Choose

[2] S may not be a vector subspace of X.

arbitrary integers $n_1 < n_2 < n_3 < \dots$ with $n_k > N_k$ for all $k \in \mathbb{N}$. Then $\|x_{n_k} - x_{n_m}\| < \frac{1}{2^k}$, due to the inequality applied to $\varepsilon = \frac{1}{2^k}$. Hence $\sum_{k=1}^{\infty} \|x_{n_k} - x_{n_{k+1}}\| < \infty$. This proves that the series $\sum_{k=1}^{\infty}(x_{n_k} - x_{n_{k+1}})$ is absolutely convergent.
By hypothesis, this series is convergent. Let $S_m = \sum_{k=1}^{m}(x_{n_k} - x_{n_{k+1}})$ be the partial sum sequence of the series. We prove that $\{S_m\}$ converges to a limit $\Delta \in X$. Therefore, $x_{n_m} = (x_{n_m} - x_{n_1}) + x_{n_1} = S_m + x_{n_1}$ converges as $m \to \infty$ to $S + x_{n_1}$. This means that $\{x_{n_m}\}$ is a convergent subsequence of $\{x_n\}$. By the aforementioned theorem, we infer that $\{x_n\}$ converges. Hence X is a Banach space. □

Corollary 14.10

Let $(X, \|\cdot\|)$ be a normed space. Then X is not complete if and only if there exists an absolutely convergent series which is not convergent.

Example 14.11. Consider $X = C[0,2]$ endowed with the L^1 norm: $\|x\|_1 = \int_0^2 |x(t)|dt$. Define the function $f_k(t)$ as:

$$f_k(t) = \begin{cases} 1 - \frac{t^2}{2}, & \text{if } 0 \le t \le \frac{2}{k} \\ 0, & \text{if } \frac{2}{k} < t \le 2 \end{cases} \quad \text{for } t \in [0,2]$$

The series $\sum_{k=1}^{\infty} f_k$ is absolutely convergent with respect to $\|\cdot\|_1$, but not convergent.

§15 November 9, 2023

§15.1 Homeomorphisms

Definition 15.1 (Homeomorphism) Let X and Y be topological spaces. A function $f : X \to Y$ is called a *homeomorphism* if it satisfies the following conditions:

1. f is a bijection (one-to-one and onto).
2. f is continuous.
3. The inverse function $f^{-1} : Y \to X$ is also continuous.

If such a function f exists, then X and Y are said to be homeomorphic.

Recall: A bijection is a function that is both one-to-one and onto, and it always possesses an inverse function. Therefore, there exists a one-to-one correspondence between the points of two spaces if they are homeomorphic. Furthermore, since continuous mappings have the property that the inverse images of open sets are open, and both a homeomorphism and its inverse are continuous, it follows that there is also a one-to-one correspondence between the open sets of the two homeomorphic spaces.

For these reasons, in topology, two spaces that are homeomorphic are considered to be essentially identical. This is because homeomorphisms preserve the topological structure of spaces, meaning that the properties that are purely topological (like connectedness, compactness, continuity, etc.) are invariant under homeomorphisms.

Definition 15.2 A property P is called a *topological property* if the fact that P holds in a topological space $(X, \mathcal{T})$ implies that P holds in any topological space $(Y, \mathcal{T}')$ which is homeomorphic to $(X, \mathcal{T})$.

⋆ A *topological property* is a feature of a space that is preserved under homeomorphisms. In other words, it is a property shared by all topological spaces that are homeomorphic to each other. These properties are intrinsic to the space's structure, irrespective of how the space is embedded or represented in a larger space.

Topological properties are fundamental to the study of topology, which can be thought of as the examination of properties that remain invariant under continuous deformations (like stretching, bending, but not tearing or gluing). This is why topology is sometimes colloquially referred to as 'rubber sheet geometry'. In this analogy, a topological space can be envisioned as drawn on a rubber sheet, where homeomorphic transformations are akin to stretching or bending the sheet without tearing or cutting it.

For instance, compactness is a topological property because if one space is compact, any space homeomorphic to it is also compact. However, completeness, which is often considered in the context of metric spaces, is not a topological property. This is evident from the fact that there are homeomorphic metric spaces where one is complete and the other is not. [3]

Example 15.3 (Completeness is **not** a topological property.)**.** The function $A : [1, \infty) \to (0, 1]$ defined by $A(x) = \frac{1}{x}$ is a homeomorphism between $[1, \infty)$ and $(0, 1]$. Both spaces $[1, \infty)$ and $(0, 1]$ are endowed with the standard distance $d(x, y) = |x - y|$. This distance induces the metric topologies $\mathcal{T}$ and $\mathcal{T}'$ respectively, where:

- $X = [1, \infty)$ is endowed with topology $\mathcal{T}$.
- $Y = (0, 1]$ is endowed with topology $\mathcal{T}'$.
- The space $[1, \infty)$ is complete: any Cauchy sequence is convergent within the space.

[3] A classic example of this concept is the topological equivalence of a circle and an ellipse, or a circle and a rectangle. Despite their different geometric shapes, from a topological perspective, they are identical because they can be transformed into one another through continuous deformation without tearing or gluing.

- The space $(0, 1]$ is **not** complete: the sequence $\{\frac{1}{n}\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $(0, 1]$, but it does not converge within $(0, 1]$.

⋆ A function is non-injective (not 1-to-1) if different elements in the domain map to the same element in the codomain. Here are two examples:

1. The squaring function $f : \mathbb{R} \to \mathbb{R}$, defined by $f(x) = x^2$, is not injective because $f(1) = f(-1) = 1$.
2. The sine function $g : \mathbb{R} \to [-1, 1]$, defined by $g(x) = \sin(x)$, is not injective since $g(x) = g(x + 2\pi)$ for all $x \in \mathbb{R}$.

Theorem 15.4 — Let $A : X \to Y$ be a continuous mapping between topological spaces X and Y, and let S be a compact subset of X. Then $A(S)$ is a compact subset of Y.

Proof. For the proof, let $\mathcal{V}$ be an open covering of $A(S)$. Since A is continuous, $A^{-1}(V)$ is an open set in X, for each $V \in \mathcal{V}$. We will show that $\mathcal{U} = \{A^{-1}(V) : V \in \mathcal{V}\}$ is an open covering of S. If $x \in S$, then $Ax \in A(S)$ so that $Ax \in V$ for some $V \in \mathcal{V}$. Then $x \in A^{-1}(V)$. So indeed $\mathcal{U}$ is an open covering of S. Since S is compact, there is a finite subcovering $\{A^{-1}(V_1), A^{-1}(V_2), \ldots, A^{-1}(V_n)\}$, say, chosen from $\mathcal{U}$. If $y \in A(S)$, then $y = Ax$ for some $x \in S$, and $x \in A^{-1}(V_k)$ for some $k = 1, 2, \ldots, n$. Then $Ax = y \in V_k$. This shows that $\{V_1, V_2, \ldots, V_n\}$ is a finite subcovering of $A(S)$, chosen from $\mathcal{V}$. Hence $A(S)$ is compact. □

Theorem 15.5 (Compact-Hausdorff Homeomorphism) — Let $(X, \mathcal{T})$ and $(Y, \mathcal{T}')$ be topological spaces. Assume that:

- X is compact,
- Y is Hausdorff, i.e., $\forall x, y \in Y$ there exist neighbourhoods U of x and V of y such that $U \cap V = \emptyset$.

If $A : X \to Y$ is a continuous bijection, then A is a homeomorphism.

Proof. We have to show that A^{-1} is continuous, i.e., $(A^{-1})^{-1}(T) \in \mathcal{T}$ for any $T \in \mathcal{T}$. Recall $A^{-1} : Y \to X$ (by Theorem 5.5.1). Let $T \subseteq X$ be an open set. We have to prove that $A(T)$ is open in Y. Then T^c is a closed set in X. Moreover, since T^c is a subset of X, which is compact by our hypothesis, by Lemma 1, T^c is compact. By Theorem 5.5.2, $A(T^c)$ is compact in Y. Recall Theorem 5.3.3: In a Hausdorff space, any compact set is closed. Since Y is Hausdorff, by Theorem 5.5.2, $A(T^c)$ is closed in Y. Hence $\sim A(T^c)$, which is $A(T)$, is open in Y. This is because A is a bijection, i.e., $A(T)^c = A(T^c)$. Use Theorem 5.4.2(d), which states that for any function f, we have $f^{-1}(D)^c = f^{-1}(D^c)$. We apply this theorem with $f = A^{-1}$ and $D = T^c$. This proves that $A(T)$ is open. □

Lemma 15.6 (Closed Subset Compactness) — In any topological space $(X, \mathcal{T})$, any closed subset of a compact set is compact.

Proof. Let $S \subseteq X$ be a compact set and $T \subseteq S$ be a closed set. We have to prove that T is compact. Let $\mathcal{V}$ be an open covering of T, i.e.,

$$T \subseteq \bigcup_{V \in \mathcal{V}} V \quad \text{and} \quad V \in \mathcal{V}.$$

Note that

$$S = S \times X = T \cup T^c = \left(\bigcup_{V \in \mathcal{V}} V \right) \cup T^c,$$

where T^c is the complement of T in S and is open since T is closed. So $\mathcal{V} \cup \{T^c\}$ is an open covering of S. Since S is compact, there exists a finite subcovering $\{T_1, \ldots, T_n\}$ of S, chosen from $\mathcal{V} \cup \{T^c\}$. We have two cases:

a) If the list $T_1, \ldots, T_n$ does not include T^c, then $T_i \in \mathcal{V}$ for $i = 1, \ldots, n$ and

$$T \subseteq S \subseteq \bigcup_{i=1}^{n} T_i;$$

so $\{T_1, \ldots, T_n\}$ is a finite subcovering of T, with sets chosen from $\mathcal{V}$.

b) If $T^c = T_n$ for some $i = 1, \ldots, n$, say $T^c = T_n$, then

$$T \subseteq S \subseteq \left(\bigcup_{i=1}^{n-1} T_i \right) \cup T^c,$$

which implies that

$$T \subseteq \left(\bigcup_{i=1}^{n-1} T_i \right);$$

so $\{T_1, \ldots, T_{n-1}\}$ is a finite subcovering of T, with sets chosen from $\mathcal{V}$.

□

§16 November 13, 2023

§16.1 Connectedness

Definition 16.1 (Disjoint) Two sets A and B are said to be *disjoint* if $A \cap B = \emptyset$. We sometimes write $A \sqcup B$ for the union of A and B when these sets are disjoint. So the statement '$X = A \sqcup B$' means that $X = A \cup B$ and $A \cap B = \emptyset$.

Definition 16.2 (Separation, connected, disconnected) Let $(X, \mathcal{T})$ be a topological space.

a) A *separation* (or *partition*) of a subset $S \subseteq X$ is a pair (T_1, T_2) of nonempty, disjoint open sets in X such that:
   (i) $T_1 \cap S \neq \emptyset$ and $T_2 \cap S \neq \emptyset$,
   (ii) $S = (T_1 \cap S) \cup (T_2 \cap S)$ [a]

b) A set S is *disconnected* if it has a separation.

c) A set S is *connected* if it has no separation.

[a] $S \subseteq T_1 \cup T_2$ and $T_1 \cap T_2 \cap S = \emptyset$.

(a) Assume $S = X$. Then a separation of X is a pair (T_1, T_2) of disjoint open sets such that $X = T_1 \cup T_2$, $T_1 \neq \emptyset$, $T_2 \neq \emptyset$. X is connected if it cannot be written as $X = T_1 \cup T_2$ with T_1, T_2 open, disjoint, non-empty. Note that $T_1 = T_2^c$ is a closed set (since it is the complement of an open set) and $T_2 = T_1^c$ is a closed set. X is disconnected $\Leftrightarrow$ there exists a subset T_1 of X which is both open and closed (hence $T_2 = T_1^c$; then $X = T_1 \cup T_2$). X is connected $\Leftrightarrow$ $\emptyset$ and X are the only subsets of X which are both open and closed.

(b) In any topological space $(X, \mathcal{T})$, $\{x\}$ is connected, for any $x \in X$.

Theorem 16.3 (Continuous Image Connectedness Theorem) — Let $A : X \to Y$ be a continuous mapping between topological spaces. If S is a connected subset of X then $A(S)$ is a connected subset of Y.

Proof. To prove this, suppose there exists a separation (T_1, T_2) of $A(S)$. Then we will show that (S_1, S_2), where $S_1 = A^{-1}(T_1)$ and $S_2 = A^{-1}(T_2)$, is a separation of S, contradicting the fact that S is connected. Certainly, S_1 and S_2 are open sets in X, since T_1 and T_2 are open in Y and A is continuous. If $x \in S_1 \cap S_2$, then we easily see that $Ax \in T_1 \cap T_2$. But $T_1 \cap T_2 = \emptyset$, so $S_1 \cap S_2 = \emptyset$. We know that $T_1 \cap A(S) \neq \emptyset$. Take any point $y \in T_1 \cap A(S)$ and say $y = Ax$. Then $x \in A^{-1}(T_1) = S_1$ and $x \in S$, so $S_1 \cap S \neq \emptyset$, and similarly $S_2 \cap S \neq \emptyset$. Finally, suppose $x \in S$, so that $Ax \in A(S) = (T_1 \cap A(S)) \cup (T_2 \cap A(S))$. If $Ax \in T_1 \cap A(S)$ then $x \in A^{-1}(T_1 \cap A(S)) = A^{-1}(T_1) \cap A^{-1}(A(S))$, by Theorem 5.4.2(c). In particular, $x \in A^{-1}(T_1) = S_1$, so $x \in S_1 \cap S$. If $Ax \in T_2 \cap A(S)$, then we proceed similarly, and conclude that $x \in (S_1 \cap S) \cup (S_2 \cap S)$, so that $S \subseteq (S_1 \cap S) \cup (S_2 \cap S)$. The reverse inclusion is obvious, so $S = (S_1 \cap S) \cup (S_2 \cap S)$. We have shown that (S_1, S_2) is a separation of S, as required. $\square$

§16.2 Finite-dimensional normed vector spaces

Definition 16.4 (Vector Space Terminology) Let V be a vector space over the field $\mathbb{F}$ (where $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$), and let $\{v_1, v_2, \dots, v_n\}$ be a subset of V.

(a) A *linear combination* of $v_1, v_2, \dots, v_n$ is a vector x of the form:

$$x = \sum_{k=1}^{n} \alpha_k v_k,$$

where $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{F}$. The scalars α_k are called the *coefficients* of v_k, for $k = 1, 2, \dots, n$.

(b) The set $\{v_1, v_2, \dots, v_n\}$ is *linearly independent* if the only scalars $\alpha_1, \alpha_2, \dots, \alpha_n$ for which

$$\sum_{k=1}^{n} \alpha_k v_k = 0$$

are $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. If there exists a non-trivial combination (i.e., not all α_k are zero) that sums to the zero vector, then the set is *linearly dependent*.

(c) The *span* of $\{v_1, v_2, \dots, v_n\}$, denoted by $\mathrm{Sp}\{v_1, v_2, \dots, v_n\}$, is the set of all linear combinations of $v_1, v_2, \dots, v_n$ and is a subspace of V called the *subspace spanned (or generated)* by $v_1, v_2, \dots, v_n$.

(d) The set $\{v_1, v_2, \dots, v_n\}$ is a *basis* for V if it is linearly independent and spans V, i.e., $\mathrm{Sp}\{v_1, v_2, \dots, v_n\} = V$. The vector space V is then said to be *finite-dimensional* with *dimension* n. If there does not exist any finite set that is a basis for V, then V is *infinite-dimensional*.

Example 16.5. Consider the vector space $V = \mathbb{R}^n$. The k-th standard basis vector is denoted by e_k and is defined as the vector in $\mathbb{R}^n$ that has a 1 in the k-th position and 0's elsewhere, that is,

$$e_k = (0, \dots, 0, 1, 0, \dots, 0),$$

where the 1 is in the k-th position. The set of vectors $\{e_1, \dots, e_n\}$, where each e_k is a standard basis vector, forms a basis for V. This set is linearly independent, and every vector in V can be uniquely expressed as a linear combination of these basis vectors. Therefore, $\{e_1, \dots, e_n\}$ is called the standard basis of $\mathbb{R}^n$.

Theorem 16.6 (Infinite-Dimensional Subspace Theorem) — A vector space is infinite-dimensional if it has an infinite-dimensional subspace.

Proof. Let W be an infinite-dimensional subspace of a vector space V, and suppose that V is finite-dimensional, with dimension n, say. By what was just said, there exists a set of n linearly independent vectors in W, which, since they belong also to V, must be a basis for V. Every vector in V, which includes all those in W, is expressible as a linear combination of these basis vectors, so they span W. Hence that set of n vectors is also a basis for W, contradicting the fact that W is infinite-dimensional. □

Proposition 16.7. $C[a, b]$ is infinite-dimensional.

Proof. The space $P[a, b]$ of polynomial functions defined on $[a, b]$ is infinite-dimensional. To see this, suppose by contradiction that $\{v_1, \dots, v_n\}$ is a basis for $P[a, b]$. Let $K = \max\{\deg(v_i) \mid i = 1, \dots, n\}$, where $\deg(v_i)$ is the degree of polynomial v_i. Then any polynomial with degree larger than K cannot be written as a linear combination of $v_1, \dots, v_n$. This contradicts the fact that $\{v_1, \dots, v_n\}$ is a basis for $P[a, b]$. So $P[a, b]$ is infinite-dimensional. By Theorem 1.11.4, $C[a, b]$ is infinite-dimensional, since $P[a, b]$ is a subspace of $C[a, b]$. □

Theorem 16.8 — The map $\|\cdot\|_\infty : V \to \mathbb{R}_+$, given by

$$\|x\|_\infty = \max_{1\le k\le n} |\alpha_k|, \quad \text{where } x = \sum_{k=1}^{n} \alpha_k v_k \in V,$$

is a norm on V.

Proof. We need to check axioms (N1)–(N3) in the definition of a normed space. Since $\{v_1, \dots, v_n\}$ is a basis of V, $\|\cdot\|_\infty$ is well-defined (since the α_k are uniquely determined). For $x = \sum_{k=1}^n \alpha_k v_k$, we clearly have $x = 0$ if and only if $(\alpha_k = 0\,\forall k)$ if and only if $\|x\|_\infty = 0$. So (N1) is satisfied.

Now, for any scalar α, we have

$$\|\alpha x\|_\infty = \left\| \sum_{k=1}^{n} (\alpha\alpha_k) v_k \right\|_\infty = \max_{1\le k\le n} |\alpha\alpha_k| = |\alpha| \max_{1\le k\le n} |\alpha_k| = |\alpha| \cdot \|x\|_\infty.$$

Thus, (N2) is satisfied.

Finally, we must prove the triangle inequality. Let $y = \sum_{k=1}^n \beta_k v_k$ be a second vector in V. For $k = 1, \dots, n$, we have

$$|\alpha_k + \beta_k| \le |\alpha_k| + |\beta_k| \le \max_{1\le k\le n} |\alpha_k| + \max_{1\le k\le n} |\beta_k| = \|x\| + \|y\|.$$

Therefore,

$$\|x+y\|_\infty = \left\| \sum_{k=1}^{n} (\alpha_k + \beta_k) v_k \right\|_\infty = \max_{1\le k\le n} |\alpha_k + \beta_k| \le \|x\|_\infty + \|y\|_\infty.$$

So (N3) is satisfied. □

Theorem 16.9 — Convergence in a finite-dimensional vector space with the norm $\|\cdot\|_\infty$ is equivalent to componentwise convergence. In other words, if $\{x_m\}$ is a sequence in a finite-dimensional vector space with this norm, and $x_m = \sum_{k=1}^n \alpha_{mk} v_k$, then

$$\{x_m\}_{m=1}^\infty \text{ converges} \iff \{\alpha_{mk}\}_{m=1}^\infty \text{ converges for each } k = 1, 2, \dots, n.$$

(Here the convergence of $\{\alpha_{mk}\}_{m=1}^\infty$ is in $\mathbb{R}$ or $\mathbb{C}$.)

Proof. Suppose $x_m \to x$, with $x = \sum_{k=1}^n \alpha_k v_k$. Then for all $\varepsilon > 0$, there exists an $N > 0$ such that

$$m > N \implies \|x_m - x\|_\infty = \max_{1\le k\le n} |\alpha_{mk} - \alpha_k| < \varepsilon.$$

Therefore, for each $k = 1, \dots, n$,

$$m > N \implies |\alpha_{mk} - \alpha_k| < \varepsilon.$$

Thus, $\{\alpha_{mk}\}_{m=1}^\infty$ converges to α_k for each $k = 1, 2, \dots, n$.

Now suppose that for each $k = 1, 2, \dots, n$, the sequence $\{\alpha_{mk}\}_{m=1}^\infty$ converges to, say, α_k. Then for all $\varepsilon > 0$ and $k = 1, 2, \dots, n$, there exists an N_k such that

$$m > N_k \implies |\alpha_{mk} - \alpha_k| < \varepsilon.$$

Let $N = \max\{N_1, \dots, N_n\}$ and $x = \sum_{k=1}^n \alpha_k v_k$. Then

$$m > N \implies \|x_m - x\|_\infty = \max_{1\le k\le n} |\alpha_{mk} - \alpha_k| < \varepsilon.$$

Thus $\{x_m\}$ converges to x. □

§17 November 20, 2023

§17.1 Finite Dimensional Subspaces

Let V be a vector space of dimension n over a field, and let $\{v_1, \dots, v_n\}$ be a basis for V. For any vector $x \in V$, represented uniquely as $x = \sum_{k=1}^{n} \alpha_k v_k$, we define the ∞-norm (or maximum norm) of x as

$$||x||_\infty = \max_{1 \le k \le n} |\alpha_k|,$$

where α_k are the coefficients of x in the basis $\{v_1, \dots, v_n\}$.

Lemma 17.1 — If $||x_m - x|| \to 0 \iff \alpha_{mk} \to \alpha_k$ for all $k = 1, .., n$. Here $x_m = \sum_{k=1}^{n} \alpha_{mk} v_k$ and $x = \sum_{k=1}^{n} \alpha_k v_k$

Lemma 17.2 — Let V be a vector space of dimension n over a field, with a basis $\{v_1, \dots, v_n\}$. Consider the sequences $\{x_m\}$ and $\{x\}$ in V, where $x_m = \sum_{k=1}^{n} \alpha_{mk} v_k$ and $x = \sum_{k=1}^{n} \alpha_k v_k$. Then, under the ∞-norm $|| \cdot ||_\infty$, the sequence $\{x_m\}$ converges to x (i.e., $||x_m - x||_\infty \to 0$) if and only if for each k from 1 to n, the sequence of coefficients $\{\alpha_{mk}\}$ converges to α_k (i.e., $\alpha_{mk} \to \alpha_k$). [a]

[a] This lemma states that in a finite-dimensional vector space with a specific basis, for a sequence of vectors to converge to a given vector under the ∞-norm, it is necessary and sufficient that the sequence of each individual coefficient (associated with the basis vectors) of these vectors converges to the corresponding coefficient of the limit vector.

This convergence behavior is directly related to the nature of the ∞-norm, which focuses on the maximum absolute value among the coefficients of the vector representation.

Theorem 17.3 — Let V be a vector space of dimension n. Consider the set

$$Q = \{x \in V \mid ||x||_\infty \le 1\}.$$

Then, Q is a compact set in V (when V is equipped with the ∞-norm).

Proof. Not required for final
We assume V is a complex vector space (the case of a real vector space is almost identical). We prove the result by induction on the dimension n of V. First suppose that $n = 1$ and that $\{v\}$ is a basis of V. Let

$$Z = \{a \in \mathbb{C} \mid |a| \le 1\} = B_{\mathbb{C}}(0, 1),$$

which we know is compact in $\mathbb{C}$ (it is a closed and bounded subset of $\mathbb{C}$, which is the same as $\mathbb{R}^2$ topologically). Define

$$A : Z \to V, \quad Av = av.$$

Then $A(Z) = B_V(0, 1)$. Since the continuous image of a compact set is compact (Theorem 4.4.7), it suffices to show that A is continuous. Suppose $\{a_m\}$ is a sequence in Z converging to $a \in Z$. Then

$$\|Aa_m - Aa\|_\infty = \|(a_m - a)v\|_\infty = |a_m - a| \to 0.$$

Hence $Aa_m \to Aa$ and so A is continuous.

Now assume the proposition is true for $n < h - 1$ for some $h > 1$. We want to show that it is true for $n = h$. Let

$$B_i = B_V(0, 1)$$

when $n = i$, $i \in \mathbb{N}_+$. So our inductive assumption is that B_n is compact for $n < h - 1$ and we wish to show that B_h is compact.

Let $\{x_m\}$ be a sequence in B_h and let

$$x_m = \sum_{j=1}^{h} \alpha_{mj} v_j$$

for $m \in \mathbb{N}_+$. The sequence $\{\alpha_{m1} v_1\}_{m=1}^{\infty}$ is a sequence in B_1, which is compact. Therefore, it has a convergent subsequence $\{\alpha_{mk_1} v_1\}_{k=1}^{\infty}$. So

$$\left\{ \sum_{j=2}^{h} \alpha_{mk,j} v_j \right\}_{k=1}^{\infty}$$

is a subsequence of $\{x_m\}$ such that the sequence of coefficients of v_1 converges. Then

$$\left\{ \sum_{j=2}^{h} \alpha_{mk,j} v_j \right\}_{k=1}^{\infty} \tag{5.1}$$

is a sequence in Span$\{v_2, \ldots, v_h\}$, which is a vector space of dimension $h-1$. Since, for $k \in \mathbb{N}_+$,

$$\left\| \sum_{j=2}^{h} \alpha_{mk,j} v_j \right\|_{\infty} = \max_{2 \le j \le h} |\alpha_{mk,j}| \le \max_{1 \le j \le h} |\alpha_{mk,j}| = \|x_{mk}\|_{\infty} \le 1,$$

it is a sequence in B_{h-1}. Since B_{h-1} is compact by the inductive hypothesis, the sequence (5.1) has a convergent subsequence. By Theorem 5.3.2, this subsequence converges componentwise. Therefore the corresponding subsequence of $\{x_m\}$ also converges componentwise (since we had already chosen a subsequence in which the first component converged) and hence converges.

□

Definition 17.4 (Equivalent Norms) Two norms $||\cdot||_1$ and $||\cdot||_2$ on a vector space V are said to be *equivalent* if there exist constants $a, b > 0$ such that for all $x \in V$,

$$a||x||_1 \le ||x||_2 \le b||x||_1.$$

This means that the two norms are equivalent if they induce the same topology on V, i.e., a sequence converges in one norm if and only if it converges in the other.

Theorem 17.5 (Equivalence of Norms in Finite-Dimensional Spaces) — In any finite-dimensional vector space V, any two norms $||\cdot||_a$ and $||\cdot||_b$ are equivalent. That is, there exist positive constants c and C such that for all vectors $x \in V$,

$$c||x||_a \le ||x||_b \le C||x||_a.$$

[a]

[a] This implies that all norms on a finite-dimensional vector space induce the same topological structure, meaning that notions of convergence, continuity, and compactness are the same under any norm.

Proof. We will only prove that any norm $\|\cdot\|$ for our vector space V is equivalent to the norm $\|\cdot\|_{\infty}$. That is, we will show that there exist positive numbers a and b such that

$$a\|x\|_{\infty} \le \|x\| \le b\|x\|_{\infty}$$

for any $x \in V$. This readily implies the theorem, but the details are left as an exercise.

In Theorem 6.5.1, we showed that the subset $Q = \{x : \|x\|_{\infty} \le 1\}$ of V is compact. It is another simple exercise to use this fact, in conjunction with Exercise 4.5(3), to conclude that the set $Q' = \{x : \|x\|_\infty = 1\}$ in V is also compact. On any normed space, the norm is a continuous mapping (Exercise 6.4(3)(c)) so we may invoke Theorem 4.3.2 to ensure the existence of points x_M and x_m in Q' such that

$$\|x_M\| = \max_{x \in Q'} \|x\|, \quad \|x_m\| = \min_{x \in Q'} \|x\|.$$

Thus $\|x_m\| \leq \|x\| \leq \|x_M\|$ for all $x \in Q'$. Also, since $\|x_m\|_\infty = 1$, we cannot have $x_m = 0$, so $\|x_m\| > 0$. For any nonzero vector $x \in V$, we have

$$\frac{1}{\|x\|_\infty} x = \frac{1}{\|x\|_\infty} x \in Q'.$$

Hence, for $x \neq 0$,

$$\|x_m\| \leq \left\| \frac{1}{\|x\|_\infty} x \right\| \leq \|x_M\|.$$

We thus have

$$\|x_m\| \|x\|_\infty \leq \|x\| \leq \|x_M\| \|x\|_\infty,$$

or

$$a\|x\|_\infty \leq \|x\| \leq b\|x\|_\infty,$$

where $a = \|x_m\| > 0$ and $b = \|x_M\|$, and this is clearly true also when $x = 0$. Hence the norms $\|\cdot\|$ and $\|\cdot\|_\infty$ are equivalent. □

Theorem 17.6 (Finite-Dimensional Spaces are Banach Spaces) — Every finite-dimensional normed vector space is a Banach space. That is, if V is a vector space of finite dimension n equipped with any norm $||\cdot||$, then V is complete with respect to this norm. In other words, every Cauchy sequence in V converges to an element in V.

Proof. Let V be a finite-dimensional vector space with basis $\{v_1, \ldots, v_n\}$. By Theorem 6.5.3, all norms are equivalent. So, it is enough to prove that $(V, \|\cdot\|_\infty)$ is a Banach space. Let $\{x_m\}$ be a Cauchy sequence in V with respect to $\|\cdot\|_\infty$. Then

$$x_m = \sum_{k=1}^{n} \alpha_{mk} v_k \text{ with } \alpha_{mk} \text{ scalars.}$$

For any $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$\|x_m - x_j\|_\infty < \varepsilon \text{ for all } m, j > N.$$

Note that $x_m - x_j = \sum_{k=1}^{n} \alpha_{mk} v_k - \sum_{k=1}^{n} \alpha_{jk} v_k = \sum_{k=1}^{n} (\alpha_{mk} - \alpha_{jk}) v_k$, and hence

$$\|x_m - x_j\|_\infty = \max_{1 \leq k \leq n} |\alpha_{mk} - \alpha_{jk}| < \varepsilon \text{ for all } m, j > N.$$

This means that for all $k = 1, \ldots, n$ fixed,

$$|\alpha_{mk} - \alpha_{jk}| < \varepsilon \text{ for all } m, j > N.$$

Hence, $\{\alpha_{mk}\}_{m \in \mathbb{N}}$ is a Cauchy sequence in $\mathbb{F}$ (where $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$). Since $\mathbb{F}$ is complete, there exists $\alpha_k \in \mathbb{F}$ such that

$$\alpha_{mk} \to \alpha_k \text{ as } m \to \infty, \text{ for } k = 1, \ldots, n.$$

By Lemma 1, we have that $x_m \to x$ in $(V, \|\cdot\|_\infty)$ where $x = \sum_{k=1}^{n} \alpha_k v_k$. Hence $\{x_m\}$ converges in $(V, \|\cdot\|_\infty)$. This proves that $(V, \|\cdot\|_\infty)$ is a complete space. □

This theorem is significant because it ensures that many of the convenient properties of finite-dimensional spaces (like $\mathbb{R}^n$) hold more generally in any space with a finite number of dimensions, regardless of the specific norm used. It's a fundamental difference between finite-dimensional and infinite-dimensional spaces, where completeness is not guaranteed and depends heavily on the chosen norm.

Next theorem is a key result in the context of finite-dimensional normed vector spaces. It connects the concepts of compactness, sequential closedness, and boundedness.

Theorem 17.7 — Let S be a subset of a finite-dimensional normed vector space V. Then S is compact if and only if it is sequentially closed and bounded.

This theorem is a variation of the Heine-Borel theorem, adapted to the language of sequential closedness. In finite-dimensional spaces, the concepts of closed and bounded sets and sequentially closed sets are often interchangeable in terms of leading to compactness, which is one of the fundamental distinctions between finite and infinite-dimensional spaces.

§17.2 Approximation Theory

Let (X, d) be a metric space, and let $S \subseteq X$ be a compact subset. For any point $x \in X$, there exists a point $p \in S$ such that

$$d(p, x) = \min_{y \in S} d(y, x),$$

which means p is the closest point in S to x, or the best approximation of x in S.

Theorem 17.8 — Let $(X, ||\cdot||)$ be a normed vector space, and let $S \subseteq X$ be a finite-dimensional subspace of X. For a fixed point $x \in X$, and assuming $S \neq \emptyset$, there exists a point $p \in S$ such that

$$||p - x|| = \min_{y \in S} ||y - x||,$$

which means p is the closest point in S to x, and is called the best approximation of x in S.

Proposition 17.9. Let f be a function in $C[0, 1]$, the space of continuous functions on the interval $[0, 1]$, and let r be an integer greater than 1. Then, there exists a polynomial p of degree less than r such that

$$||p - f|| = \min_{g \in S_r} ||g - f||,$$

where $||f|| = \max_{t \in [0,1]} |f(t)|$ and S_r is the set of all polynomials of degree less than r.

The theorem essentially states that for any continuous function on $[0, 1]$, there exists a polynomial of degree less than r that is the best approximation of f in the uniform norm, among all polynomials of degree less than r. This is a fundamental result in approximation theory and is related to the Weierstrass approximation theorem.

§18 November 23, 2023

Theorem 18.1 (6.6.1) — Let X be a normed vector space and $S \subset X$ a finite-dimensional subspace, $S \neq \varnothing$. Fix $x \in X$. Then there exists $p \in S$ such that

$$\|p - x\| = \min_{y \in S} \|y - x\|.$$

Proof. Let $p_0 \in S$ be arbitrary. We consider the following subset of S:

$$Y = \{y \in S; \|y - x\| \leq \|p_0 - x\|\}.$$

We claim that $Y \neq \varnothing$. Assume that $Y = \varnothing$. Then $\|y - x\| > \|p_0 - x\|$ for all $y \in S$ which is a contradiction since $p_0 \in S$. Thus, there exists $p \in Y$ such that

$$\|p - x\| = \min_{y \in Y} \|y - x\|.$$

To prove that Y is compact, we need to show that Y is closed and bounded. Recall from Theorem 6.5.5 that a subset of any finite-dimensional vector space is compact if it is closed and bounded.

Y is closed: Let $\{y_k\} \subset Y$ be such that $y_k \to y$. We have to prove that $y \in Y$. For any $\varepsilon > 0$, there exists N such that for all $k > N$,

$$\|y_k - x\| < \|p_0 - x\| + \varepsilon.$$

Since ε is arbitrary, we must have $\|y - x\| \leq \|p_0 - x\|$. Hence $y \in Y$.

Y is bounded: For all $y \in Y$,

$$\|y\| = \|y - x + x\| \leq \|y - x\| + \|x\| \leq \|p_0 - x\| + \|x\| := M.$$

Hence $Y \subset B(0, M)$ where $B(0, M)$ is the closed ball centered at the origin with radius M.

By Theorem 4.3.3, Y is compact. Therefore, the set Y is closed and bounded, and by the previous argument, contains the point p which minimizes the distance to x. This completes the proof. □

Definition 18.2 (6.6.2) A normed vector space X is strictly convex if the equation

$$\|x + y\| = \|x\| + \|y\|$$

holds only when $x = \beta y$ for some $\beta > 0$ and $x, y \in X \setminus \{0\}$.

Note: If $x = \beta y$ then

$$\|x + y\| = \|\beta y + y\| = \|(\beta + 1)y\| = (\beta + 1)\|y\| = \beta\|y\| + \|y\| = \|x\| + \|y\|.$$

This illustrates that in a strictly convex space, the triangle inequality becomes an equality if and only if the vectors are linearly dependent and pointing in the same direction.

a) The space $C[a, b]$ endowed with the norm

$$\|f\|_2 = \left(\int_a^b |f(t)|^2 dt\right)^{\frac{1}{2}}$$

is strictly convex (see Exercise 6.10.(5), Assignment 4).

b) The space $C[a,b]$ endowed with the uniform norm

$$\|f\| = \max_{t\in[a,b]} |f(t)|$$

is **not** strictly convex.

Here is an example: let $f(t) = bt$ and $g(t) = t^2$ for $t \in [a,b]$. Assume $0 < a < b$.

$$\|f\| = \max_{t\in[a,b]} |bt| = b \cdot \max_{t\in[a,b]} t = b^2$$

$$\|g\| = \max_{t\in[a,b]} |t^2| = b^2$$

$$\|f+g\| = \max_{t\in[a,b]} |bt + t^2| = \max_{t\in[a,b]} |t(b+t)| = b^2 + b^3 = 2b^2 = \|f\| + \|g\|$$

Hence, condition (2) holds. But $f \neq \beta g$ for some $\beta > 0$.

Theorem 18.3 (6.6.3) — Let X be a strictly convex normed vector space, and $S \subseteq X$ a finite-dimensional subspace, $S \neq \varnothing$. Fix $x \in X$. Then, there exists a unique $p \in S$ such that

$$\|p - x\| = \min_{y\in S} \|y - x\|.$$

Proof. **Case 1:** $x \in S$. Then $\min_{y\in S} \|y - x\| = \|x - x\| = 0$, so the minimum is achieved for $p = x$, which is unique.

Case 2: $x \notin S$. The existence of p is given by Theorem 6.6.1. Assume that there exists another point $p' \in S$ such that

$$\|p' - x\| = \|p - x\| = \min_{y\in S} \|y - x\| =: d.$$

Now, since S is a vector space, $\frac{1}{2}(p + p') \in S$ and

$$d \leq \|x - \frac{1}{2}(p + p')\| = \frac{1}{2}\|x - p\| + \frac{1}{2}\|x - p'\| \leq \frac{1}{2}\|x - p\| + \frac{1}{2}\|x - p'\| = d.$$

Hence $\|x - \frac{1}{2}(p + p')\| = d$ so $\frac{1}{2}(p + p')$ is also a best approximation of x. (This averaging process can be continued indefinitely to show the existence of infinitely many best approximations in a normed space once there are two different best approximations.) It follows that

$$\|x - \frac{1}{2}(p + p')\| = \frac{1}{2}\|x - p\| + \frac{1}{2}\|x - p'\|,$$

from which, since X is strictly convex,

$$x - p = \beta(x - p')$$

for some number $\beta > 0$. If $\beta \neq 1$, we get

$$x = \frac{1}{1-\beta} p - \frac{\beta}{1-\beta} p'.$$

This is impossible since it represents x as belonging to the vector space S, whereas $x \notin S$. So we must have $\beta = 1$. Thus $p = p'$ and we have proved that the best approximation is unique. □

§19 November 27, 2023

Definition 19.1 Let X, Y be normed vector spaces and $S \subseteq X$ a subset. A map $A : S \to Y$ is *uniformly continuous* on S if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$\|Ax' - Ax''\| \leq \varepsilon \text{ for all } x', x'' \in S \text{ with } \|x' - x''\| < \delta.$$

Theorem 19.2 — Let X, Y be normed vector spaces and $S \subseteq X$ a compact subset of X, $S \neq \emptyset$. If $A : S \to Y$ is continuous on S, then A is also uniformly continuous.

Proof. Suppose that A is not uniformly continuous on S. This means that there is some number $\epsilon > 0$ such that, regardless of the value of δ, there are points $x', x'' \in S$ with $\|x' - x''\| < \delta$ but for which $\|Ax' - Ax''\| \geq \epsilon$. Take $\delta = 1/n$, for $n = 1, 2, \dots$ in turn, and for each n let x'_n, x''_n be points in S (known to exist by our supposition) such that

$$\|x'_n - x''_n\| < \frac{1}{n} \quad \text{and} \quad \|Ax'_n - Ax''_n\| \geq \epsilon.$$

As S is compact, the sequence $\{x'_n\}$ has a convergent subsequence $\{x'_{n_k}\}$, with limit x, say. Take any number $\eta > 0$. There exists a positive integer K such that $\|x'_{n_k} - x\| < \eta/2$ when $k > K$. We may suppose $K > 2/n$ for such k, $n_k \geq k > 2/n$ and

$$\|x''_{n_k} - x\| \leq \|x''_{n_k} - x'_{n_k}\| + \|x'_{n_k} - x\| < \frac{1}{n_k} + \frac{\eta}{2} < \eta,$$

so that $\{x''_{n_k}\}$ is a convergent subsequence of $\{x''_n\}$, also with limit x. Further, the sequence $Ax'_{n_1}, Ax''_{n_1}, Ax'_{n_2}, Ax''_{n_2}, \dots$ in Y must converge with limit Ax. Hence there is an integer N such that, when $k > N$,

$$\|Ax'_{n_k} - Ax''_{n_k}\| \leq \|Ax'_{n_k} - Ax\| + \|Ax''_{n_k} - Ax\| < \frac{\eta}{2} + \frac{\eta}{2} = \eta,$$

and this gives us a contradiction. Thus A is indeed uniformly continuous on S. □

Corollary 19.3
Let X be a finite-dimensional normed vector space, and $S \subseteq X$ be a closed and bounded subset of X. Let Y be a normed vector space and $A : S \to Y$. If A is continuous on S then A is also uniformly continuous on S.

Proof. By Theorem 6.5.5, S is compact. The conclusion follows from Theorem 6.8.2. □

§19.1 The Weierstrass approximation theorem

Theorem 19.4 (Weierstrass Approximation Theorem) — For any $f \in C[0, 1]$ and for any $\epsilon > 0$, there exists a polynomial p such that

$$\|p - f\| < \epsilon, \quad \text{i.e. } |p(x) - f(x)| < \epsilon \quad \forall x \in [0, 1].$$

Recall: $\|f\| = \max_{x \in [0,1]} |f(x)|$ is the uniform norm on $C[0, 1]$.

Proof. We define Bernstein polynomial for f, of degree n, by:

$$P_n(x) = \sum_{k=0}^{n} \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1 - x)^{n-k}$$

Then

$$f(x) = f(0) \cdot (1-x)^n + \sum_{k=1}^{n-1} \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k} + f(1) \cdot x^n$$

and

$$|f(x) - P_n(x)| \le \sum_{k=0}^{n} |f(x) - f\left(\frac{k}{n}\right)| \cdot \binom{n}{k} x^k (1-x)^{n-k} \quad (5)$$

Since f is continuous on a compact interval, f is uniformly continuous. Let $\epsilon > 0$ be arbitrary. Then $\exists \delta > 0$ s.t.

$$|f(x') - f(x'')| < \frac{\epsilon}{2} \quad \text{for all } x', x'' \in [0,1] \text{ with } |x' - x''| < \delta \quad (6)$$

Fix $x \in [0,1]$. Consider

$$S_1 = \left\{ k \in \{0,1,\dots,n\} \mid \left|\frac{k}{n} - x\right| < \delta \right\}, \quad S_2 = \left\{ k \in \{0,1,\dots,n\} \mid \left|\frac{k}{n} - x\right| \ge \delta \right\}$$

Clearly $S_1 \cup S_2 = \{0,1,\dots,n\}$ and $S_1 \cap S_2 = \varnothing$. Note that:

$$\text{If } k \in S_1 \text{ then } |f(x) - f\left(\frac{k}{n}\right)| < \frac{\epsilon}{2} \text{ due to (6)}$$

$$\text{If } k \in S_2 \text{ then } |f(x) - f\left(\frac{k}{n}\right)| \le |f(x) - f(0)| + |f\left(\frac{k}{n}\right) - f(0)| \le 2||f||_\infty$$

Coming back to (5), we obtain:

$$|f(x) - P_n(x)| \le \sum_{k\in S_1} |f(x) - f\left(\frac{k}{n}\right)| \cdot \binom{n}{k} x^k (1-x)^{n-k} + \sum_{k\in S_2} |f(x) - f\left(\frac{k}{n}\right)| \cdot \binom{n}{k} x^k (1-x)^{n-k}$$

$$\le \frac{\epsilon}{2} \sum_{k\in S_1} \binom{n}{k} x^k (1-x)^{n-k} + 2||f||_\infty \sum_{k\in S_2} \binom{n}{k} x^k (1-x)^{n-k} \quad (7)$$

For the first sum, we use:

$$\sum_{k\in S_1} \binom{n}{k} x^k (1-x)^{n-k} \le \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} = 1$$

For the second sum, we use:

$$\begin{aligned}
\sum_{k\in S_2} \binom{n}{k} x^k (1-x)^{n-k} &= \frac{1}{\delta^2} \sum_{k\in S_2} \left(\frac{k}{n} - x\right)^2 \binom{n}{k} x^k (1-x)^{n-k} \\
&\le \frac{1}{\delta^2} \sum_{k=0}^{n} \left(\frac{k}{n} - x\right)^2 \binom{n}{k} x^k (1-x)^{n-k} \\
&= \frac{1}{\delta^2} \cdot \frac{1}{n^2} \sum_{k=0}^{n} (k - nx)^2 \binom{n}{k} x^k (1-x)^{n-k} \\
&= \frac{x(1-x)}{\delta^2 n}.
\end{aligned}$$

Returning to (7), we get:

$$|f(x) - P_n(x)| \le \frac{\epsilon}{2} \cdot 1 + 2\|f\|_\infty \cdot \frac{x(1-x)}{\delta^2 n}$$

For any $\epsilon > 0$ fixed, there exists $N \in \mathbb{N}$ such that

$$2\|f\|_\infty \cdot \frac{x(1-x)}{\delta^2 n} < \frac{\epsilon}{2} \quad \text{for all } n > N.$$

Hence, for all $n > N$,

$$|f(x) - P_n(x)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

□

§20 November 30, 2023

§20.1 Bounded Linear Mappings

Definition 20.1 (Linear map) Let X and Y be vector spaces. A map $A : X \to Y$ is **linear** if

$$A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 A x_1 + \alpha_2 A x_2, \quad \forall \alpha_1, \alpha_2 \in \mathbb{F}, \forall x_1, x_2 \in X.$$

Definition 20.2 (Bounded map) Let X and Y be normed vector spaces. A map $A : X \to Y$ is **bounded** if there exists $K > 0$ such that

$$\|Ax\| \le K\|x\|, \quad \forall x \in X.$$

Observation: A function $f : X \to \mathbb{R}$ is bounded if there exists $K > 0$ such that

$$|f(x)| \le K, \quad \forall x \in X.$$

Definition 20.3 (Operator) Let X and Y be normed vector spaces. A map $A : X \to Y$ which is linear and bounded is called an **operator**.

Theorem 20.4 — Let X and Y be normed vector spaces and $A : X \to Y$ be a linear map. If A is continuous at $x_0 \in X$, then A is continuous on X.

Proof. Let $x_n \in X$ be arbitrary. Let $\{x_n\}$ be a sequence in X s.t. $x_n \to x$. Then $x_n - x \to 0$ and hence $x_n - x + x_0 \to x_0$. Since A is continuous at x_0, $A(x_n - x + x_0) \to Ax_0$ (sequentially).

By linearity of A, $A(x_n - x + x_0) = Ax_n - Ax + Ax_0$. Hence $Ax_n - Ax \to 0$, i.e., $Ax_n \to Ax$. Hence A is (sequentially) continuous at x. □

Theorem 20.5 — Let X and Y be normed vector spaces. Let $A : X \to Y$ be a linear map. Then A is continuous on X if and only if A is bounded.

Proof. ($\Rightarrow$) Suppose that A is bounded, i.e., there exists $K > 0$ such that

$$\|Ax\| \le K\|x\|, \quad \forall x \in X. \tag{1}$$

Let $x_n \in X$ be arbitrary and $\{x_n\}$ be a sequence in X with $x_n \to x$. Let $\varepsilon > 0$ be arbitrary. Then there exists N such that if $n > N$, $\|x_n - x\| < \varepsilon$. Then

$$\|Ax_n - Ax\| = \|A(x_n - x)\| \le K\|x_n - x\| < K\varepsilon, \quad \text{for all } n > N,$$

which proves that $Ax_n \to Ax$, i.e., A is continuous at x. By Theorem 2.1, A is continuous on X.

($\Leftarrow$) Assume that A is continuous on X. We want to prove that A is bounded. Suppose by contradiction that A is not bounded; this means that $\forall K > 0$ there exists $x_k \in X$ s.t. $\|Ax_k\| > K\|x_k\|$. We apply this for $K = n \in \mathbb{N}$. Then we find $x_n \in X$ s.t.

$$\|Ax_n\| > n\|x_n\|, \quad \text{for all } n \in \mathbb{N}.$$

Note that since A is linear, $A0 = 0$ because $A0 = A(x - x) = Ax - Ax = 0$. Hence $x_n \neq 0$, and therefore $\|x_n\| \neq 0$, using (N). Define

$$y_n = \frac{1}{n\|x_n\|} x_n.$$

Then

$$Ay_n = A\left(\frac{1}{n\|x_n\|}x_n\right) = \frac{1}{n\|x_n\|}Ax_n$$

and

$$\|Ay_n\| = \left\|\frac{1}{n\|x_n\|}Ax_n\right\| = \frac{\|Ax_n\|}{n\|x_n\|} > 1, \quad \text{for all } n.$$

On the other hand,

$$\|y_n\| = \left\|\frac{1}{n\|x_n\|}x_n\right\| = \frac{1}{n} \to 0, \text{ hence } y_n \to 0.$$

Since A is continuous, $Ay_n \to A0 = 0$, hence there exists N s.t. $\|Ay_n\| < 1$ for all $n > N$. This is a contradiction. □

Example 20.6. example of an operator on a normed space X is the mapping A defined by

$$Ax = \beta x, \quad x \in X,$$

for some fixed scalar β. It is indeed linear, since

$$A(\alpha_1 x_1 + \alpha_2 x_2) = \beta(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 Ax_1 + \alpha_2 Ax_2,$$

for $x_1, x_2 \in X$, scalars α_1, α_2. And it is bounded, since

$$\|Ax\| = \|\beta x\| = |\beta|\|x\|$$

so $\|Ax\| \leq K\|x\|$ for some constant K (such as $|\beta|$ or any larger number). If $\beta = 1$, A is the identity operator or unit operator on X and is denoted by I. Thus I maps every element of X into itself. If $\beta = 0$, A is called the zero operator on X and maps every element of X into θ.

Example 20.7. For a second example, we take the mapping $A : C[a, b] \to C[a, b]$ defined by the equation $Ax = y$ where

$$y(s) = \lambda \int_a^b k(s,t)x(t)\,dt, \quad x \in C[a,b], \quad a \leq s \leq b.$$

Here, k is a function of two variables, which is continuous in the square $[a, b] \times [a, b]$, and λ is a given nonzero real number. The mapping A is linear, since, for $x_1, x_2 \in C[a, b]$, scalars α_1, α_2, and any $s \in [a, b]$,

$$(A(\alpha_1 x_1 + \alpha_2 x_2))(s) = \lambda \int_a^b k(s,t)(\alpha_1 x_1(t) + \alpha_2 x_2(t))\,dt$$

$$= \alpha_1 \lambda \int_a^b k(s,t)x_1(t)\,dt + \alpha_2 \lambda \int_a^b k(s,t)x_2(t)\,dt$$
$$= (\alpha_1 Ax_1)(s) + (\alpha_2 Ax_2)(s);$$

that is, $A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 Ax_1 + \alpha_2 Ax_2$. Also, A is bounded. To see this, let M be a positive constant such that $|k(s, t)| \leq M$ for (s, t) in the square. Then

$$\|Ax\| = \|y\| = \max_{a\leq s\leq b} |y(s)| = \max_{a\leq s\leq b} \left|\lambda \int_a^b k(s,t)x(t)\,dt\right|$$

$$\leq |\lambda| \max_{a\leq s\leq b} \int_a^b |k(s,t)||x(t)|\,dt$$
$$\leq |\lambda| M \max_{a\leq s\leq b} |x(t)| \cdot (b-a)$$
$$= |\lambda| M(b-a)\|x\|.$$

Thus, for $K = |\lambda| M(b - a)$, say, we have $\|Ax\| \leq K\|x\|$ for all $x \in C[a, b]$, so A is bounded. This verifies that A is indeed an operator.

§21 December 4, 2023

§21.1 Norm of operators

Theorem 21.1 — Let X and Y be normed vector spaces, and $A : X \to Y$ an operator. Let

$$a = \inf\{K > 0; \|Ax\| \leq K\|x\|, \forall x \in X\},$$
$$b = \sup\{\|Ax\|; x \in X, x \neq 0\},$$
$$c = \sup\{\frac{\|Ax\|}{\|x\|}; x \in X, \|x\| = 1\},$$
$$d = \sup\{\|Ax\|; x \in X, \|x\| \leq 1\}.$$

Then:

(a) $\|Ax\| \leq a\|x\|, \forall x \in X$,

(b) $a = b = c = d$.

Proof. Recall (MAT 2125): If $S \subseteq \mathbb{R}$ and $a = \inf S$, then:

- $a \leq x$ for all $x \in S$
- For every $\varepsilon > 0$ there exists $x_\varepsilon \in S$ such that $a \leq x_\varepsilon < a + \varepsilon$

In our case, $S = \{k > 0 \mid \|Ax\| \leq k\|x\| \text{ for all } x \in X\}$. Let ε_0 be arbitrary. By the definition above, $k_\varepsilon \in S$ such that $a \leq k_\varepsilon < a + \varepsilon$.
$k_\varepsilon \in S$ means that $\|Ax\| \leq k_\varepsilon\|x\|$ for all $x \in X$. Hence, for any $x \in X$ fixed, $\|Ax\| < (a + \varepsilon)\|x\|$ for all $\varepsilon > 0$. Letting $\varepsilon \to 0$, we obtain: $\|Ax\| \leq a\|x\|$.

We will show that $a \leq b \leq c \leq d \leq a$ by definition, $b \geq \|Ax\|$ for all $x \in X, x \neq 0$. So $\|A\| \leq b$ for all $x \neq 0$. Hence, $b \in S$ and $\|Ax\| \leq k\|x\|$ for all $x \in X$ and therefore $b \geq a$, by property a) of the inf given above.

Fix $x \in X, x \neq 0$. Then $\frac{\|Ax\|}{\|x\|} \leq c$, and since this is true for any $x \in X, x \neq 0$, we infer that $b \leq c$. Here we use the fact that if $\alpha \leq c$ for all x, then $\sup S \leq c$ ($S \subseteq \mathbb{R}$).

Here we use the fact that if $S_1 \subseteq S_2 \subseteq \mathbb{R}$ then $\sup S_1 \leq \sup S_2$. In our case, $S_1 = \{\|Ax\| \mid x \in X, \|x\| = 1\}$ and $S_2 = \{\|Ax\| \mid x \in X, \|x\| \leq 1\}$, hence $\sup S_1 = c \leq \sup S_2 = d$.

By a), for any vector $x \in X$ with $\|x\| \leq 1$, we have $\|Ax\| \leq a\|x\| \leq a$. Taking the supremum over all such x, we obtain $d \leq a$. Therefore, $a \leq b \leq c \leq d \leq a$, and thus $a = b = c = d$ and X is a normed space with $\|A\| = a$. □

Definition 21.2 Let $A : X \to Y$ be an operator between normed vector spaces X and Y. The norm of A is:

$$\|A\| = \inf\{k > 0 \mid \|Ax\| \leq k\|x\| \text{ for all } x \in X\} \quad (= a \text{ in Th 7.2.1})$$

Theorem 21.3 — Let $B(X, Y)$ be the set of all operators $A : X \to Y$, where X and Y are normed vector spaces. Then $B(X, Y)$ is a vector space, equipped with the following operations:

$$(A_1 + A_2)x = A_1x + A_2x \quad \text{for all } x \in X$$
$$(\alpha A)x = \alpha Ax \quad \text{for all } \alpha \in \mathbb{F}, x \in X$$

Proof: Exercise: Check that the 8 axioms of Definition 1.1.1 are satisfied.

Remark 21.4. By Theorem 7.2.1 a), $\|Ax\| \leq \|A\|\|x\|$ for all $x \in X$ (2).

Example 21.5. Let $A : X \to X$ be given by $Ax = \beta x$ for some $\beta \in \mathbb{F}$. Then

$$\|A\| = \sup\{\|Ax\| \mid x \in X, \|x\| = 1\} = \sup\{|\beta x\| \mid x \in X, \|x\| = 1\} = |\beta|$$

Use the fact that $a = c$ in Th 7.2.1.
In particular, if $\beta = 1$, we obtain the identity operator $I : A \to A$ given by $Ix = x$, and $\|I\| = 1$.

Theorem 21.6 — $B(X, Y)$ with the operator norm $\|\cdot\|$ is a normed vector space.

Proof. We prove the three properties.

(M1): If $A = 0$ then clearly $\|A\| = 0$ since $\|Ax\| = \|0x\| = 0\ \forall x$.

Assume that $\|A\| = 0$. Note that $0 = \|A\| = \sup\left\{\frac{\|Ax\|}{\|x\|} \mid x \in X, x \neq 0\right\}$. Hence $\frac{\|Ax\|}{\|x\|} = 0$ for all $x \in X, x \neq 0$. This means that $\|Ax\| = 0$, and hence $Ax = 0$ for all $x \in X, x \neq 0$. Clearly $A0 = 0$. Hence $A = 0$.

- **(M2)**: $\|(\alpha A)x\| = \|\alpha Ax\| = |\alpha|\|Ax\| \leq |\alpha|k\|x\|\ \forall x \in X, x \neq 0$ (assuming $\alpha \neq 0$ and $A \neq 0$). Hence $\|\alpha A\| \leq |\alpha|\|A\|$. We want to prove that $\|\alpha A\| \geq |\alpha|\|A\|$. Let $x \in X$ be arbitrary. Then $\|\alpha Ax\| = \|\alpha^{-1}(\alpha A)x\| \leq |\alpha^{-1}|\|(\alpha A)x\| \leq |\alpha^{-1}| \cdot |\alpha|\|A\|\|x\| = \|A\|\|x\|$ for all x. Hence, $\|\alpha A\| \leq |\alpha|\|A\|$ so $\|\alpha A\| = |\alpha|\|A\|$.

- **(M3)**: $\|(A_1 + A_2)x\| = \|A_1x + A_2x\| \leq \|A_1x\| + \|A_2x\| \leq \|A_1\|\|x\| + \|A_2\|\|x\|$ for all $x \in X$. Here $A_1, A_2 \neq 0$ so that $A_1 + A_2 \neq 0$. This proves that $\|A_1 + A_2\| \leq \|A_1\| + \|A_2\|$.

So, $B(X, Y)$ is a normed vector space. □

Theorem 21.7 — If Y is a Banach space, then $B(X, Y)$ is also a Banach space.

Proof. We use Theorem 6.2.2 (from textbook): we show that every absolutely convergent series in $B(X, Y)$ is convergent.

Let $\sum_{i=1}^{\infty} A_i$ be an absolutely convergent series in $B(X, Y)$. This means that $\sum_{i=1}^{\infty} \|A_i\|$ converges in $\mathbb{R}$. For $x \in X$, define

$$y_n = \sum_{i=1}^{n} A_i x.$$

Let $\varepsilon > 0$ be arbitrary. Since $\{\sum_{i=1}^{n} \|A_i\|\}_{n=1}^{\infty}$ is a Cauchy sequence, there exists $N \in \mathbb{N}$ such that for $n > m > N$,

$$\left|\sum_{i=1}^{n} \|A_i\| - \sum_{i=1}^{m} \|A_i\|\right| = \left|\sum_{i=m+1}^{n} \|A_i\|\right| < \varepsilon.$$

Note that for $n > m > N$,

$$\|y_n - y_m\| = \left\|\sum_{i=m+1}^{n} A_i x\right\| \leq \sum_{i=m+1}^{n} \|A_i x\| \leq \sum_{i=m+1}^{n} \|A_i\|\|x\| < \varepsilon\|x\|.$$

This proves that $\{y_n\}$ is a Cauchy sequence in Y. Since Y is a Banach space, $\lim_{n\to\infty} y_n$ exists in Y.

Basically, $y = \sum_{i=1}^{\infty} A_i x$. We define $A : X \to Y$ by setting $Ax = y$. Note that A is linear and bounded, which is left as an exercise.

Finally, we show that $A = \sum_{i=1}^{\infty} A_i$ in $B(X, Y)$, i.e. $\sum_{i=1}^{n} A_i \to A$ in $B(X, Y)$. Let $\varepsilon > 0$ be arbitrary. Recall that there exists $N \in \mathbb{N}$ such that for $n > N$,

$$\|A - \sum_{i=1}^{n} A_i\| < \varepsilon.$$

This shows (4). □