Níckolas de Aguiar Alves

Notes on Functional Analysis

Draft as of: December 27, 2023

São Bernardo do Campo – SP
2023

Níckolas de Aguiar Alves

Notes on Functional Analysis

São Bernardo do Campo – SP
2023

AGUIAR ALVES, Níckolas de

Notes on Functional Analysis / Níckolas de Aguiar Alves — São Bernardo do Campo, 2023.

viii, 76 p.: 0 il.

1. functional analysis. I. Title.

Abstract

This document is a personal study notebook about functional analysis. It is intended to guide my studies of the subject and work as a manner of registering my thoughts, doubts, and solutions. Hopefully, it will also be of use to other students. The book by Oliveira (2018) is the main reference in the sense I'll follow its structure and order closely, but hopefully, my comments will be somewhat more authentic. Our main accompanying references are the books by Kreyszig (1978) and Reed and Simon (1980), but some other books on functional analysis that I might often check are those by Brezis (2011), Conway (2007), Rudin (1991), and Yosida (1995), among others.

These notes are written from the point of view of a theoretical physicist interested in learning more mathematics. Hence, they might have a style somewhat different from usual mathematics texts. I'll typically try to write to an audience of similar-minded people who already have some necessary prerequisites, but might like to be convinced that the definitions make sense. Well, at least that's how I usually like to learn mathematics.

Talking about prerequisites,, I shall assume you to be familiar with topology, linear algebra, measure theory, and so on. These topics can be studied, for example, in the books by Axler (2015), Folland (1999), Geroch (1985), Munkres (2000), and Simon (2015).

Keywords: functional analysis.

Abbreviations

PDE partial differential equation

RHS right-hand side

SOT strong operator topology

WOT weak operator topology

Contents

Abstract iii

Abbreviations v

1 Banach Spaces 1
- 1.1 What's the matter with infinite dimensions? 1
- 1.2 Completing Normed Spaces 17
- 1.3 Separable Spaces 21
- 1.4 Linear Operators 24
- 1.5 Space of Bounded Operators 27
- 1.6 Dual Space 32
- 1.7 Banach Fixed Point Theorem 35
- 1.8 Baire Theorem 41
- 1.9 Principle of Uniform Boundedness 43
- 1.10 Open Mapping Theorem 45
- 1.11 Closed Graph Theorem 49
- 1.12 Zorn's Lemma 51
- 1.13 Hahn–Banach Theorem 53
- 1.14 Bidual 62
- 1.15 Adjoint Operators in Normed Spaces 64
- 1.16 Weak Convergence 67

Bibliography 75

One

Banach Spaces

1.1 What's the matter with infinite dimensions?

From a practical perspective, a physicist can think of functional analysis as being the theory of infinite-dimensional vector spaces. These spaces are interesting, among other reasons, because they often arise as the spaces of solutions to linear partial differential equations (PDEs), for example. Since PDEs occur everywhere in physics, it is not surprising that functional analysis ends up being fairly useful.

One might then wonder what is special about infinite dimension that makes it so more complex than the finite-dimensional cases. One can think about it by noticing that an infinite linear combination will look something like

$$\psi = \sum_{n=1}^{+\infty} a_n \phi_n, \tag{1.1.1}$$

for example, where ϕ_n are vectors and a_n are scalars in some convenient field (typically the real or complex numbers). What is tricky about Eq. (1.1.1) is noticing that what it actually means is

$$\psi = \lim_{N \to +\infty} \sum_{n=1}^{N} a_n \phi_n, \tag{1.1.2}$$

and it is not immediately obvious—in fact, it is not obvious at all—what is the sense in which this limit should be taken. In other words, it is not clear what is the topology we should be considering. Furthermore, once we do pick a topology, many complicated features may and will arise due to the difficult properties it brings along. Hence, infinite-dimensional vector spaces will often require us to look at their topological properties at the same time we analyze their linear properties, and this will lead us into interesting and complex results.

With this in mind, we desire to somehow introduce a topology in some vector space. One way of doing so is by employing a norm.

Definition 1.1 [Normed Spaces]**:**

Let V be a $\mathbb{K}$-vector space, where $\mathbb{K}$ is either $\mathbb{R}$ or $\mathbb{C}$. A *norm* on V is a function $\|\cdot\|\colon V \to \mathbb{R}$ such that the following properties hold:

i. $\|\lambda v\| = |\lambda|\|v\|$, for any $\lambda \in \mathbb{K}, v \in V$ (scaling);

ii. $\|u + v\| \leq \|u\| + \|v\|$, for any $u, v \in V$ (triangle inequality, or subadditivity);

iii. if $\|v\| = 0$, then $v = 0$ (separates points).

If $\|\cdot\|$ is a norm on V, the pair $(V, \|\cdot\|)$ is said to be a *normed space*. If a function $\|\cdot\| : V \to \mathbb{R}$ satisfies the scaling law and the triangle inequality, we call it a *seminorm*. ♠

This is an interesting definition because it will lead us to some nice aspects. For example,

i. the norm will induce a metric, and hence a normed space is a metric space;

ii. the topology induced by the norm will be such that the algebraic operations in the normed space are continuous;

iii. the norm itself will be continuous.

Let us prove these statements.

Proposition 1.2:
Let $(V, \|\cdot\|)$ be a normed space over $\mathbb{K}$. Then

$$d(u, v) = \|u - v\| \tag{1.1.3}$$

defines a metric on V. □

Proof:
We want to prove two properties to ensure that d is indeed a metric. They are

i. given $u, v \in V$, $d(u, v) = 0$ if, and only if, $u = v$;

ii. given $u, v \in V$, $d(u, v) \leq d(u, w) + d(v, w)$, for any $w \in V$.

These two conditions are sufficient to characterize d as a metric, and they imply symmetry and positivity.

Begin by noticing that $d(u, v) = \|u - v\|$ implies $d(u, v) = 0$ if, and only if, $u - v = 0$. Hence, $d(u, v) = 0$ if, and only if, $u = v$, as required.

Furthermore, given $u, v, w \in V$,

$$\begin{aligned} d(u, v) &= \|u - v\|, && (1.1.4a) \\ &= \|u - w + w - v\|, && (1.1.4b) \\ &\leq \|u - w\| + \|w - v\|, && (1.1.4c) \\ &= \|u - w\| + \|v - w\|, && (1.1.4d) \\ &= d(u, w) + d(v, w), && (1.1.4e) \end{aligned}$$

proving the triangle inequality. ■

Proposition 1.3:

Let $(V, \|\cdot\|)$ be a normed space. The linear operations (addition and scalar multiplication) are continuous in the norm-induced topology. □

Proof:

Let us begin by considering addition: $+\colon V \times V \to V$. We want to show that

$$\forall\, x, y, u, v \in V, \forall\, \epsilon > 0, \exists\, \delta > 0; \sqrt{\|x-u\|^2 + \|y-v\|^2} < \delta \Rightarrow \|x+y-u-v\| < \epsilon. \tag{1.1.5}$$

Pick $\delta = \epsilon$. Notice that the triangle inequality implies

$$\|x+y-u-v\| \leq \|x-u\| + \|y-v\|, \tag{1.1.6}$$

and hence

$$\begin{aligned} \|x+y-u-v\|^2 &\leq \|x-u\|^2 + 2\|x-u\|\|y-v\| + \|y-v\|^2, && (1.1.7a)\\ &\leq \|x-u\|^2 + \|y-v\|^2, && (1.1.7b)\\ &< \delta^2, && (1.1.7c) \end{aligned}$$

and hence

$$\|x+y-u-v\| < \delta = \epsilon \tag{1.1.8}$$

whenever $\sqrt{\|x-u\|^2 + \|y-v\|^2} < \delta$.

Next we want to prove the scalar product is continuous. To do so we should prove that

$$\forall\, u, v \in V, \forall\, \lambda, \mu \in \mathbb{K}, \forall\, \epsilon > 0, \exists\, \delta > 0; \sqrt{|\lambda-\mu|^2 + \|u-v\|^2} < \delta \Rightarrow \|\lambda u - \mu v\| < \epsilon. \tag{1.1.9}$$

Notice that

$$\begin{aligned} \|\lambda u - \mu v\| &= \|\lambda u - \lambda v + \lambda v - \mu v\|, && (1.1.10a)\\ &\leq \|\lambda u - \lambda v\| + \|\lambda v - \mu v\|, && (1.1.10b)\\ &= |\lambda|\|u-v\| + |\lambda-\mu|\|v\|, && (1.1.10c)\\ &\leq (|\lambda| + \|v\|)\delta. && (1.1.10d) \end{aligned}$$

Hence, by picking

$$\delta = \frac{\epsilon}{|\lambda| + \|v\|} \tag{1.1.11}$$

we get to the desired result. ■

Proposition 1.4:

Let $(V, \|\cdot\|)$ be a normed space. The norm $\|\cdot\|\colon V \to \mathbb{R}$ is continuous. □

Proof:

Let $u, v \in V$. Suppose $\|v\| \leq \|u\|$. Notice that the triangle inequality implies

$$\|u\| \leq \|v\| + \|u - v\|. \tag{1.1.12}$$

Therefore, we see that

$$\big|\|u\| - \|v\|\big| \leq \|u - v\|. \tag{1.1.13}$$

To prove the norm is continuous we must show that

$$\forall\, u, v \in V, \forall\, \epsilon > 0, \exists\, \delta > 0; \|u - v\| < \delta \Rightarrow \big|\|u\| - \|v\|\big| < \epsilon. \tag{1.1.14}$$

This follows directly from Eq. (1.1.13) by picking $\delta = \epsilon$. ■

A norm is not the only possibility for introducing a topology in a vector space. We could have pursued other possibilities and obtained other sorts of spaces. For example, locally convex spaces generalize normed spaces by employing a family of seminorms to generate a topology (see, *e.g.*, Reed and Simon 1980, Chap. V). Still, normed spaces tend to be particularly convenient, since they have a rich theory and many interesting cases. While the proof of the pudding is in the eating, the proof of a definition is in its theorems, and so we shall see how normed spaces can be useful.

Let us provide a concrete example.

Example 1.5 [Uniform Norm]**:**

Let Ω be a compact subspace of a Hausdorff topological space. Let $\mathscr{C}^0(\Omega)$ denote the space of continuous maps $f\colon \Omega \to \mathbb{K}$, where $\mathbb{K}$ is either the real line or the complex plane. Consider the map $\|\cdot\|_\infty\colon \mathscr{C}^0(\Omega) \to \mathbb{R}$ defined by

$$\|\psi\|_\infty = \sup_{p\in\Omega} |\psi(p)| = \max_{p\in\Omega} |\psi(p)|. \tag{1.1.15}$$

$\|\cdot\|_\infty$ is actually a norm on $\mathscr{C}^0(\Omega)$. Indeed, if $\|\psi\|_\infty = 0$, then $|\psi(p)| \leq 0, \forall\, p \in \Omega$, which implies $\psi(p) = 0$. Furthermore, if $\psi, \phi \in \mathscr{C}^0(\Omega)$, then

$$\begin{aligned}
\|\psi + \phi\|_\infty &= \sup_{p\in\Omega} |\psi(p) + \phi(p)|, && \text{(1.1.16a)}\\
&\leq \sup_{p\in\Omega} |\psi(p)| + \sup_{p\in\Omega} |\phi(p)|, && \text{(1.1.16b)}\\
&= \|\psi\|_\infty + \|\phi\|_\infty. && \text{(1.1.16c)}
\end{aligned}$$

Moreover, given $\lambda \in \mathbb{K}$ and $\psi \in \mathscr{C}^0(\Omega)$,

$$\begin{aligned}
\|\lambda\psi\|_\infty &= \sup_{p\in\Omega} |\lambda\psi(p)|, && \text{(1.1.17a)}\\
&= |\lambda| \sup_{p\in\Omega} |\psi(p)|, && \text{(1.1.17b)}
\end{aligned}$$

$$= |\lambda| \|\psi\|_\infty. \tag{1.1.17c}$$

$\|\cdot\|_\infty$ is known as the uniform convergence norm, or simply uniform norm. An interesting property it has is that it turns $\mathscr{C}^0(\Omega)$ into a complete metric space. Let us show this as well by following Conway (2007, p. 65).

Let $(\psi_n)_{n\in\mathbb{N}}$ be a Cauchy sequence on $\mathscr{C}^0(\Omega)$. Then we know that

$$\forall\, \epsilon > 0, \exists\, N \in \mathbb{N}; \|\psi_n - \psi_m\|_\infty < \epsilon, \forall\, n, m > N. \tag{1.1.18}$$

However, due to the norm's definition, this implies that

$$\forall\, \epsilon > 0, \exists\, N \in \mathbb{N}; |\psi_n(p) - \psi_m(p)| < \epsilon, \forall\, p \in \Omega, \forall\, n, m > N. \tag{1.1.19}$$

Therefore, at each point, we have a Cauchy sequence on $\mathbb{K}$, $(\psi_n(p))_{n\in\mathbb{N}}$. Since the real line and the complex plane are complete, it follows that the pointwise limit

$$\psi(p) = \lim_{n\to+\infty} \psi_n(p) \tag{1.1.20}$$

exists for each $p \in \Omega$.

Fix $p \in \Omega$. Notice now that, for $n, m > N$,

$$\begin{aligned}
|\psi(p) - \psi_m(p)| &= |\psi(p) - \psi_n(p) + \psi_n(p) - \psi_m(p)|, && (1.1.21a)\\
&\leq |\psi(p) - \psi_n(p)| + |\psi_n(p) - \psi_m(p)|, && (1.1.21b)\\
&\leq |\psi(p) - \psi_n(p)| + \|\psi_n - \psi_m\|_\infty, && (1.1.21c)\\
&< |\psi(p) - \psi_n(p)| + \epsilon. && (1.1.21d)
\end{aligned}$$

We may then take the limit $n \to +\infty$ and the supremum over $p \in \Omega$ on both sides of the expression. This leads us to

$$\|\psi - \psi_m\|_\infty < \epsilon. \tag{1.1.22}$$

Therefore, we have just shown that

$$\forall\, \epsilon > 0, \exists\, N \in \mathbb{N}; \|\psi - \psi_m\|_\infty < \epsilon, \forall\, m > N. \tag{1.1.23}$$

Therefore, $(\psi_n)_{n\in\mathbb{N}}$ converges in the uniform norm to ψ. Hence, $\psi_n \to \psi$ uniformly, which ensures ψ is continuous (see, *e.g.*, Munkres 2000, Theorem 21.6). Therefore, $\psi \in \mathscr{C}^0(\Omega)$, and we have thus shown that $\mathscr{C}^0(\Omega)$ is indeed complete in the uniform norm. ♥

Complete normed spaces are often more interesting and convenient than other normed spaces. The reason for that is that we are ensured expressions such as Eq. (1.1.2) on page 1 will have a proper meaning, since they will converge whenever they are Cauchy (if they are not Cauchy, they would not converge anyway). Hence, complete normed spaces deserve a special name of their own.

Definition 1.6 [Banach Space]**:**

A *Banach space* is a complete normed space, where completeness is meant in the norm-induced metric. ♠

Note that Example 1.5 on page 4 shows that $\mathscr{C}^0(\Omega)$ is a Banach space with respect to the uniform norm $\|\cdot\|_\infty$. Let us see some more examples.

Example 1.7 [$\mathbb{K}^n$]**:**

Let $\mathbb{K}$ be either $\mathbb{R}$ or $\mathbb{C}$. Then $\mathbb{K}^n$ is a Banach space in many different norms. Given $1 \le p < \infty$ let us define

$$\|u\|_p = \left(\sum_{i=1}^n |u_i|^p\right)^{\frac{1}{p}}, \tag{1.1.24}$$

where $u = (u_1, \dots, u_n)$. Define also

$$\|u\|_\infty = \max_{1\le i\le n} |u_i|. \tag{1.1.25}$$

Let us show that these norms are indeed norms and that they are indeed complete.

We begin by proving they are norms. Let $1 \le p < \infty$. Firstly notice that

$$\|u\|_p^p = \sum_{i=1}^n |u_i|^p. \tag{1.1.26}$$

The scaling law follows at once, and so does the fact that $\|\cdot\|_p$ separates points, because it is a sum of non-negative numbers. Let us show the triangle inequality. The proof follows from the triangle inequality for the absolute value in the case $p = 1$, so we'll focus on $p \ge 1$. Notice that in this case $x \mapsto |x|^p$ is a convex function, *i.e.*, given $\lambda \in [0,1]$, it holds that $|\lambda x + (1-\lambda)y|^p \le \lambda|x|^p + (1-\lambda)|y|^p$, for any $x, y \in \mathbb{R}$. Notice then that

$$\begin{aligned}
\|u+v\|_p^p &= \sum_{i=1}^n |u_i + v_i|^p, && (1.1.27a)\\
&\le \sum_{i=1}^n \big||u_i| + |v_i|\big|^p, && (1.1.27b)\\
&= 2^p \sum_{i=1}^n \left|\frac{1}{2}|u_i| + \frac{1}{2}|v_i|\right|^p, && (1.1.27c)\\
&\le 2^{p-1}\sum_{i=1}^n |u_i|^p + 2^{p-1}\sum_{i=1}^n |v_i|^p, && (1.1.27d)\\
&= 2^{p-1}\|u\|_p^p + 2^{p-1}\|v\|_p^p. && (1.1.27e)
\end{aligned}$$

Therefore,

$$\|u+v\|_p \le 2^{1-\frac{1}{p}}\left(\|u\|_p^p + \|v\|_p^p\right)^{\frac{1}{p}}. \tag{1.1.28}$$

Since $x \mapsto x^{\frac{1}{p}}$ is concave, it follows that

$$\|u+v\|_p \le 2^{1-\frac{1}{p}}\left(\|u\|_p^p + \|v\|_p^p\right)^{\frac{1}{p}}, \tag{1.1.29a}$$

$$= 2^{1-\frac{1}{p}} 2^{\frac{1}{p}} \left(\frac{1}{2} \|u\|_p^p + \frac{1}{2} \|v\|_p^p \right)^{\frac{1}{p}}, \tag{1.1.29b}$$

$$= 2 \left(\frac{1}{2} \|u\|_p^p + \frac{1}{2} \|v\|_p^p \right)^{\frac{1}{p}}, \tag{1.1.29c}$$

$$\leq \|u\|_p + \|v\|_p, \tag{1.1.29d}$$

which proves the triangle inequality.

The proof that $\|\cdot\|_\infty$ is a norm is straightforward and follows from the properties of the absolute value.

Let us now prove completeness. We shall later show that, for finite-dimensional vector spaces, all norms induce the same topology (Theorem 1.19 on page 13), and hence it suffices to prove completeness for a single norm.

Let $(u_m)_{m\in\mathbb{N}}$ be a Cauchy sequence on $(\mathbb{K}^n, \|\cdot\|_\infty)$. Then

$$\forall \epsilon > 0, \exists N \in \mathbb{N}; \|u_l - u_m\|_\infty < \epsilon, \forall l, m > N. \tag{1.1.30}$$

However, due to the definition of $\|\cdot\|_\infty$, this implies that each coordinate of (u_m) defines a Cauchy sequence on $\mathbb{K}$. Hence, each coordinate has a limit. We can define u to be the vector on $\mathbb{K}^n$ which has each of these limits as a coordinate, and it is straightforward to prove that both $u \in \mathbb{K}^n$ (by construction) and that $u_m \to u$ in the $\|\cdot\|_\infty$ norm.

It is also interesting to notice how the previous norms relate to each other. Namely, why is it that we used the notation $\|\cdot\|_\infty$? It can actually be understood as a limit of the other norms. Indeed, fix $1 \leq p < \infty$ and $u \in V$. Let $1 \leq j \leq n$ be such that $|u_j| = \|u\|_\infty$. We assume $u \neq 0$, as the vanishing case is straightforward. Notice that

$$\|u\|_p = \left(\sum_{i=1}^n |u_i|^p \right)^{\frac{1}{p}}, \tag{1.1.31a}$$

$$= \left(\sum_{i\neq j} |u_i|^p + |u_j|^p \right)^{\frac{1}{p}}, \tag{1.1.31b}$$

$$= |u_j| \left(\sum_{i\neq j} \left| \frac{u_i}{u_j} \right|^p + 1 \right)^{\frac{1}{p}}, \tag{1.1.31c}$$

$$\geq |u_j|, \tag{1.1.31d}$$

$$= \|u\|_\infty. \tag{1.1.31e}$$

Furthermore,

$$\|u\|_p = \left(\sum_{i=1}^n |u_i|^p \right)^{\frac{1}{p}}, \tag{1.1.32a}$$

$$\leq \left(n |u_j|^p \right)^{\frac{1}{p}}, \tag{1.1.32b}$$

$$= n^{\frac{1}{p}} \left|u_j\right|, \tag{1.1.32c}$$
$$= n^{\frac{1}{p}} \|u\|_\infty. \tag{1.1.32d}$$

Therefore, for each $1 \le p < \infty$ and for each $u \in V$, it holds that

$$\|u\|_\infty \le \|u\|_p \le n^{\frac{1}{p}} \|u\|_\infty. \tag{1.1.33}$$

We can then see that, for fixed $u \in V$,

$$\lim_{p \to +\infty} \|u\|_p = \|u\|_\infty, \tag{1.1.34}$$

which explains our choice of notation. ♥

Example 1.8 [Integral Norm]**:**
Let $a, b \in \mathbb{R}$, $a < b$. Consider the space $\mathscr{C}^0([a,b])$. It is straightforward to show that

$$\|\psi\|_1 = \int_a^b |\psi(t)| \,\mathrm{d}t \tag{1.1.35}$$

is a norm on $\mathscr{C}^0([a,b])$.

Notice that $\|\cdot\|_1$ is bounded by $\|\cdot\|_\infty$. Indeed,

$$\|\psi\|_1 = \int_a^b |\psi(t)| \,\mathrm{d}t\,, \tag{1.1.36a}$$
$$\le \int_a^b \|\psi\|_\infty \,\mathrm{d}t\,, \tag{1.1.36b}$$
$$= (b-a)\|\psi\|_\infty. \tag{1.1.36c}$$

$\|\cdot\|_\infty$, on the other hand, is not bounded by $\|\cdot\|_1$. Indeed, suppose there was $A > 0$ such that $\|\psi\|_\infty \le A\|\psi\|_1$, for all $\psi \in \mathscr{C}^0([a,b])$. Then consider any function ψ such that $\|\psi\|_\infty = A$, but supported on an interval with size $\frac{1}{2A}$ or smaller—such a function can be shown to exist by using Urysohn's Lemma (see, *e.g.*, Munkres 2000, Theorem 33.1). Then this function has $\|\psi\|_1 \le \frac{1}{2}$, which means that $\|\psi\|_\infty > A\|\psi\|_1$.

It is also interesting to show that $(\mathscr{C}^0([a,b]), \|\cdot\|_1)$ is not complete. Hence, it is an example of a normed space that is not a Banach space. To see this, let us pick $a = 0$ and $b = 1$ for simplicity (the general case is analogous). Consider the sequence of continuous functions with

$$\psi_n(x) = \begin{cases} 1, & \text{if } x \ge \frac{1}{2} + \frac{1}{n}, \\ 0, & \text{if } x \le \frac{1}{2} - \frac{1}{n}, \\ \frac{2nx+2-n}{4}, & \text{otherwise.} \end{cases} \tag{1.1.37}$$

A direct calculation shows that

$$\|\psi_n - \psi_m\|_1 = \frac{|m-n|(m+mn+n)}{4m^2n^2}, \tag{1.1.38}$$

which can be used to show that (ψ_n) is a Cauchy sequence. Hence, for each $x \in [0,1]$, $(\psi_n(x))$ is a Cauchy sequence of real numbers and we can define $\psi(x) = \lim_{n\to+\infty} \psi_n(x)$ pointwise. This leads to

$$\psi(x) = \begin{cases} 1, & \text{if } x > \frac{1}{2}, \\ 0, & \text{if } x < \frac{1}{2}, \\ \frac{1}{2}, & \text{if } x = \frac{1}{2}. \end{cases} \tag{1.1.39}$$

One can then show that

$$\left\|\psi - \psi_n\right\|_1 = \frac{1}{4n}, \tag{1.1.40}$$

and hence $\psi_n \to \psi$ in the topology induced by $\|\cdot\|_1$. However, ψ is not continuous, and hence $\psi \notin \mathscr{C}^0([0,1])$. There are no other candidates for the limit, since they would violate the pointwise limits. Therefore, $\mathscr{C}^0([0,1])$ is not complete. ♥

We can make an infinite-dimensional version of $\mathbb{K}^n$ by considering spaces of sequences on $\mathbb{K}$.

Example 1.9 [ℓ^p]**:**
Consider $\mathbb{K}^{\mathbb{N}}$, the space of all sequences $u \colon \mathbb{N} \to \mathbb{K}$. We shall denote $u = (u_1, u_2, \dots)$. We can define some normed subspaces of $\mathbb{K}^{\mathbb{N}}$.

For example, let $1 \le p < \infty$. Define

$$\ell^p(\mathbb{N}) = \left\{ u \in \mathbb{K}^{\mathbb{N}};\, \|u\|_p = \left(\sum_{i=1}^{+\infty} |u_i|^p \right)^{\frac{1}{p}} < \infty \right\}. \tag{1.1.41}$$

Define also

$$\ell^\infty(\mathbb{N}) = \left\{ u \in \mathbb{K}^{\mathbb{N}};\, \|u\|_\infty = \sup_{1 \le i < +\infty} |u_i| < \infty \right\}. \tag{1.1.42}$$

It is possible to show that $\ell^p(\mathbb{N})$ and $\ell^\infty(\mathbb{N})$ are Banach spaces.

Similar definitions and results hold for $\ell^p(\mathbb{Z})$. In fact, we can be even more general. Consider some index set Λ and the space $\mathbb{K}^\Lambda$ of functions $\psi \colon \Lambda \to \mathbb{K}$. We may then define

$$\ell^p(\Lambda) = \left\{ \psi \in \mathbb{K}^\Lambda;\, \|\psi\|_p = \left(\sum_{\lambda \in \Lambda} |\psi_\lambda|^p \right)^{\frac{1}{p}} < \infty \right\}, \tag{1.1.43}$$

where the sum assumes that there are only countably many non-vanishing terms. We can define also

$$\ell^\infty(\Lambda) = \left\{ \psi \in \mathbb{K}^\Lambda;\, \|\psi\|_\infty = \sup_{\lambda \in \Lambda} |\psi_\lambda| < \infty \right\}. \tag{1.1.44}$$

These are too Banach spaces. ♥

Instead of only sequences, we can go on further and consider function spaces.

Example 1.10 [L^p]**:**

Let $(\Omega, \mathscr{A}, \mu)$ be a measure space. Then, for $1 \leq p \leq \infty$, the $L^p(\Omega, \mu)$ spaces are Banach spaces with the norms

$$\left\|\psi\right\|_p = \int \left(\int_\Omega \left|\psi(x)\right|^p \, \mathrm{d}\mu(x)\right)^{\frac{1}{p}} \tag{1.1.45}$$

and

$$\left\|\psi\right\|_\infty = \operatorname*{ess\,sup}_{x \in \Omega} \left|\psi(x)\right|. \tag{1.1.46}$$

This is a well-known result in measure theory. For its proof, and the relevant definitions, see, *e.g.*, the book by Folland (1999, Theorems 6.6 and 6.8). ♥

Remark [Minkowski and Hölder Inequalities]*:*

It is useful to recall two interesting results about L^p norms. The triangle inequality in this context is often known as the Minkowski inequality,

$$\left\|\psi + \phi\right\|_p \leq \left\|\psi\right\|_p + \left\|\phi\right\|_p, \tag{1.1.47}$$

For a proof, see, *e.g.*, Folland (1999, Theorems 6.5 and 6.8).

Furthermore, given $1 < p < \infty$, we say that $1 < q < \infty$ is *conjugate* to p if $\frac{1}{p} + \frac{1}{q} = 1$. By convention we take 1 and ∞ to be conjugates. The Hölder inequality states that

$$\left\|\psi\phi\right\|_1 \leq \left\|\psi\right\|_p \left\|\phi\right\|_q, \tag{1.1.48}$$

where q is conjugate to p. For a proof, see, *e.g.*, Folland (1999, Theorems 6.2 and 6.8). ♣

Let us recall some more definitions and notation from linear algebra that will be useful.

Definition 1.11 [Span]**:**

Let V be a vector space and let $X \subseteq V$ be a subset of V. The *span* of X, denoted $\operatorname{span}(X)$, is the vector space comprised of all finite linear combinations of elements of X. ♠

Definition 1.12 [Hamel Basis]**:**

Let V be a vector space. We say a set $B \subseteq V$ is a *Hamel basis* for V if, and only if, B is linearly independent and $V = \operatorname{span}(B)$. ♠

Definition 1.13 [Dimension of a Vector Space]**:**

Let V be a vector space. If it has a finite Hamel basis with n elements, we say that the *algebraic dimension* (or simply the dimension) of V, $\dim V$, is $\dim V = n$. If $V = \{0\}$, we say it has dimension zero and write $\dim V = 0$. If there is no finite Hamel basis, we say the dimension is infinite and write $\dim V = \infty$. ♠

Remark:

One might worry about the existence of a Hamel basis in a general vector space in order to make sense of the previous definition. Furthermore, one might worry about the existence of different Hamel bases with different numbers of elements. We shall later use Zorn's lemma to prove the existence of a Hamel basis in any vector space (Theorem 1.110 on page 52). As for ensuring different (finite) bases always have the same number of elements, see, *e.g.*, the book by Axler (2015, Theorem 2.35). ♣

Definition 1.14 [Total Subset]**:**
Let V be a normed space. A subset $X \subset V$ is said to be *total* if, and only if, $\operatorname{span}(X)$ is dense in V. ♠

Definition 1.15 [Equivalent Norms]**:**
Let V be a vector space and let $\|\cdot\|_1$ and $\|\cdot\|_2$ be norms on V. We say these two norms are *equivalent* if, and only if, there are constants $A, B > 0$ such that

$$A\|u\|_1 \leq \|u\|_2 \leq B\|u\|_1, \forall\, u \in V. \tag{1.1.49}$$

♠

The notion of equivalence of norms is interesting because it holds if, and only if, the norms induce the same topology on the normed space. Hence, checking for equivalence between norms is a simple way of checking for topological equivalence.

Theorem 1.16:
Let V be a vector space and let $\|\cdot\|_1$ and $\|\cdot\|_2$ be norms on V. $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent if, and only if, they induce the same topology on V. □

Proof:
We know both norms generate metric topologies, and hence we know both topologies are generated by open balls. To prove they generate the same topology we can then simply check whether an open ball in one of the topologies is an open set in the other one, and vice versa.

We begin then by assuming the norms to be equivalent, *i.e.*, we assume the existence of constants $A, B > 0$ such that

$$A\|u\|_1 \leq \|u\|_2 \leq B\|u\|_1, \forall\, u \in V. \tag{1.1.50}$$

Let $\mathscr{B}_1(u;\epsilon)$ be the open ball (in the $\|\cdot\|_1$-induced topology) centered at $u \in V$ with radius ϵ. Pick $v \in \mathscr{B}_1(u;\epsilon)$. We want to show v is an interior point of $\mathscr{B}_1(u;\epsilon)$ in the $\|\cdot\|_2$-induced topology. To do so, notice that

$$\|u - v\|_1 < \epsilon. \tag{1.1.51}$$

Pick now $\delta = A(\epsilon - \|u - v\|_1) > 0$. Let $w \in \mathscr{B}_2(v, \delta)$. It follows from Eq. (1.1.50) on the preceding page that

$$\begin{aligned}\|u-w\|_1 &\le \|u-v\|_1 + \|v-w\|_1, &\text{(1.1.52a)}\\ &\le \|u-v\|_1 + \frac{1}{A}\|v-w\|_2, &\text{(1.1.52b)}\\ &< \|u-v\|_1 + \frac{\delta}{A}, &\text{(1.1.52c)}\\ &= \epsilon. &\text{(1.1.52d)}\end{aligned}$$

Hence, $\mathcal{B}_2(v,\delta) \subseteq \mathcal{B}_1(u,\epsilon)$, meaning $\mathcal{B}_1(u,\epsilon)$ is open in the $\|\cdot\|_2$-induced topology. The argument could be repeated exchanging the norms, and hence it is proven that the equivalence of norms implies they generate equivalent topologies.

On the other hand, suppose that the topologies are equivalent. In this case, we know there is some $\epsilon > 0$ such that

$$\mathcal{B}_2(0,\epsilon) \subseteq \mathcal{B}_1(0,1). \tag{1.1.53}$$

Pick $u \in V$ with $u \neq 0$. Then define v through

$$v = \frac{\epsilon u}{2\|u\|_2}. \tag{1.1.54}$$

Notice that

$$\|v\|_1 = \frac{\epsilon\|u\|_1}{2\|u\|_2}. \tag{1.1.55}$$

However, $\|v\|_2 = \frac{\epsilon}{2} < \epsilon$, which implies $\|v\|_1 < 1$. Therefore,

$$\|u\|_1 < \frac{2\|u\|_2}{\epsilon}. \tag{1.1.56}$$

The argument also works when we switch the norms, and hence the result is proven. ■

Notation [Open Balls]*:*

When proving Theorem 1.16 on the previous page we introduced the notation $\mathcal{B}(x,\epsilon)$ for the open ball centered at x with radius ϵ. We might also write $\mathcal{B}_d(x,\epsilon)$ to indicate we mean the open ball in the metric d, or even $\mathcal{B}_\epsilon(x)$ to save space and mean $\mathcal{B}(x,\epsilon)$. In any case, the expression should be clear from context. ♦

Example 1.17 [Unequivalent Norms in a Space of Polynomials]**:**

Consider the real polynomials on the interval $[0,1]$ and equip them with the L^p norms for $p = 1$ and $p = \infty$. These norms are not equivalent. An example is given by considering the polynomials x^{n-1}, which have $\|x^{n-1}\|_\infty = 1$, but $\|x^{n-1}\|_1 = \frac{1}{n}$. Hence, for $n \to \infty$, we get $\|x^{n-1}\|_\infty \to 1$, but $\|x^{n-1}\|_1 \to 0$, meaning the expression defining the equivalence of norms can't possibly hold. ♥

Let us prove one more property about the topology of finite-dimensional spaces that will later be useful: a generalization of the Heine–Borel theorem for other finite-dimensional vector spaces, possibly over the complex numbers.

Theorem 1.18:
Let V be a finite-dimensional normed space. A subset $X \subseteq V$ is compact if, and only if, it is closed and bounded. □

Proof:
It holds that, in a metric space, all compact sets are closed and bounded (see, for instance, Kreyszig 1978, Lemma 2.5-2). We must then only prove that all closed and bounded sets are compact.

In a metric space, compactness and sequential compactness coincide (see, *e.g.*, Munkres 2000, Theorem 28.2). Hence, we just need to prove that every sequence in a closed and bounded set in a normed space has a convergent subsequence.

To do so, we can simply introduce a finite basis and employ the Bolzano–Weierstrass theorem. For a more detailed proof, see, for example, the book by Kreyszig (1978, Theorem 2.5-3). ■

To further motivate how infinite-dimensional vector spaces differ from the finite-dimensional case, let us state some results we shall later prove.

i. All norms are equivalent on a finite-dimensional vector space, but this does not hold in infinite-dimensional spaces (Example 1.17 and Theorem 1.19 on the facing page and on the current page).

ii. The closed ball $\overline{\mathscr{B}(0,1)}$ on a normed space is compact if, and only if, the space if finite-dimensional (Theorem 1.23 on page 15).

iii. All linear transformations from a finite-dimensional normed space to itself are continuous, but this does not hold in infinite-dimensional spaces (Propositions 1.55 and 1.56 on page 27).

iv. For an infinite-dimensional normed space, it is possible to have dense proper subspaces, and even have different such subspaces that intersect solely at the origin. It is also possible to have linear maps defined on a dense subspace that cannot be linearly extended to the whole space.

Cross reference (and check?)

Theorem 1.19:
Let V be a finite-dimensional vector space. All norms on V are equivalent. □

Proof:
Let $n = \dim V$. Pick a basis $(e_i)_{1 \le i \le n}$. We define $\|\cdot\|_1$ through

$$\|u\|_1 = \sum_{i=1}^{n} |u_i|, \tag{1.1.57}$$

where the u_i are the coefficients such that

$$u = \sum_{i=1}^{n} u_i e_i. \tag{1.1.58}$$

If we prove that an arbitrary norm $\|\cdot\|$ on V is equivalent to $\|\cdot\|_1$ the result follows. Let then $\|\cdot\|$ be any norm on V.

We begin by noticing that

$$\begin{aligned} \|u\| &= \left\| \sum_{i=1}^{n} u_i e_i \right\|, && \text{(1.1.59a)} \\ &\leq \sum_{i=1}^{n} \|u_i e_i\|, && \text{(1.1.59b)} \\ &= \sum_{i=1}^{n} |u_i| \|e_i\|, && \text{(1.1.59c)} \\ &\leq \left(\max_{1 \leq i \leq n} \|e_i\|\right) \sum_{j=1}^{n} \left|u_j\right|, && \text{(1.1.59d)} \\ &= \left(\max_{1 \leq i \leq n} \|e_i\|\right) \|u\|_1, && \text{(1.1.59e)} \\ &= B\|u\|_1. && \text{(1.1.59f)} \end{aligned}$$

To prove the other side of the inequality we proceed by contradiction. Suppose that $\forall \epsilon > 0, \exists u_\epsilon \in V$ such that $\epsilon\|u_\epsilon\|_1 > \|u_\epsilon\|$. We can define $v_\epsilon = \frac{u_\epsilon}{\|u_\epsilon\|_1}$, and hence $\forall \epsilon > 0, \exists v_\epsilon \in V$ such that $\epsilon > \|v_\epsilon\|$, but $\|v_\epsilon\|_1 = 1$.

Pick then, for each $m \in \mathbb{N}$, a $v_m \in V$ such that $\|v_m\|_1 = 1$, but $\|v_m\| < \frac{1}{m}$. Since the sphere $\mathcal{S}_1(0, 1)$ is compact in the $\|\cdot\|_1$-induced topology, we know that (v_m) has a convergent subsequence in $\mathcal{S}_1(0, 1)$. Let the limit of this subsequence be v. Since $\mathcal{S}_1(0, 1)$ is compact, we know $\|v\|_1 = 1$. However, notice that $\|v\| = \lim_{m\to\infty} \|v_m\| = 0$. This is a contradiction, since one of the properties of a norm is that it vanishes if, and only if, its argument is the zero vector. Two norms can't disagree on whether a given vector is the zero vector, and hence we are forced to conclude our assumption was a contradiction. Therefore, there is some $A > 0$ such that

$$A\|u\|_1 \leq \|u\|. \tag{1.1.60}$$

This concludes the proof. ∎

Notation [Spheres in a Topology]*:*

When proving Theorem 1.19 on the preceding page we introduced the notation $\mathcal{S}(x, \epsilon)$ for the sphere centered at x with radius ϵ. Our notation for spheres is analogous to our notation for open balls. ♦

Corollary 1.20:

All finite-dimensional normed spaces are Banach spaces. Hence, any finite-dimensional subspace of a normed space is closed. □

Proof:

Since Theorem 1.19 on page 13 ensures all norms on a finite-dimensional vector space are equivalent, it suffices to consider any particular one. We can then use coordinates and do the proof by analogy with Example 1.7 on page 6. ■

Example 1.21:

The Stone–Weierstrass theorem (see, *e.g.*, Simon 2015, Theorem 2.5.2) ensures the space of real polynomials is dense in $\mathscr{C}^0([a,b])$. This implies that $\mathscr{C}^0([a,b])$ is infinite-dimensional. Notice also that this is an example of a dense proper subset, since not all continuous functions are polynomials. ♥

Let us now show that the unit closed ball is compact if, and only if, the space is finite-dimensional. To do so, we'll want to find a sequence with no convergent subsequence. This can be done by employing the Riesz Lemma.

Lemma 1.22 [Riesz]**:**

Let $(V, \|\cdot\|)$ *be a normed space and let* $U \subseteq V$ *be a closed proper subspace. For each* $\alpha \in (0,1)$ *there is a vector* $v \in V \setminus U$ *such that* $\|v\| = 1$ *and* $\inf_{u \in U} \|v - u\| \geq \alpha$. □

Proof:

We begin by picking any $w \in V \setminus U$. For this choice of vector, define the constant c to be $c = \inf_{u \in U} \|w - u\|$. Since U is closed, we know $V \setminus U$ is open, and hence $c > 0$.

Pick then $d > c$. It is ensured that there is some $x \in U$ such that $c \leq \|w - x\| \leq d$. We can then use this to define the vector $v = \frac{w-x}{\|w-x\|}$. Notice that $v \in V$ and $\|v\| = 1$. Furthermore, notice that, given any $u \in U$,

$$\|v - u\| = \left\| \frac{w - x}{\|w - x\|} - u \right\|, \tag{1.1.61a}$$

$$= \frac{1}{\|w - x\|} \|w - x - \|w - x\| u\|, \tag{1.1.61b}$$

$$= \frac{1}{\|w - x\|} \|w - (x + \|w - x\| u)\|. \tag{1.1.61c}$$

Notice that $(x + \|w - x\| u) \in U$. As a consequence,

$$\|v - u\| \geq \frac{c}{\|w - x\|}, \tag{1.1.62a}$$

$$\geq \frac{c}{d}. \tag{1.1.62b}$$

For a given α, we can simply pick d large enough so that $\alpha = \frac{c}{d}$. This concludes the proof. ■

Theorem 1.23:

Let V *be a normed space. The closed unit ball* $\overline{\mathscr{B}(0,1)}$ *is compact if, and only if,* V *is finite-dimensional.* □

Proof:

Suppose V is finite-dimensional. Then Theorem 1.18 on page 13 implies the closed unit ball is compact.

To prove the converse, assume V is infinite-dimensional. We shall give a sequence in $\overline{\mathscr{B}(0,1)}$ that has no convergent subsequences. To do so, pick a vector $u_1 \in \mathcal{S}(0,1)$. Notice the subspace spanned by this vector is finite-dimensional, and hence Corollary 1.20 on page 14 ensures it is a closed subspace. The Riesz Lemma then ensures the existence of $u_2 \in \mathcal{S}(0,1)$ such that $\|u_1 - u_2\| \geq \frac{1}{2}$.

We can repeat this construction indefinitely. Suppose we have built a n-dimensional subspace U. We can then use the Riesz Lemma again to obtain a new vector $u_{n+1} \in \mathcal{S}(0,1)$ and hence construct an infinite sequence. It is always possible to add more vectors because V is infinite-dimensional, and hence has an infinite Hamel basis. Meanwhile, the u_i describe a finite Hamel basis for U, ensuring there is always enough space in $V \setminus U$ to accommodate new vectors.

The sequence thus constructed can't have any convergent subsequences since, by construction, given any $i \neq j$ it holds that $\left\|u_i - u_j\right\| \geq \frac{1}{2}$. This proves $\overline{\mathscr{B}(0,1)}$ is not compact, and hence concludes the proof. ■

Example 1.24 [Non-Closed Subspaces]**:**

Let us give some examples of subspaces of Banach spaces that are not closed. Due to Corollary 1.20 on page 14 we know these subspaces will need to be infinite-dimensional.

As a first example, pick $V = \ell^1(\mathbb{Z})$. Let U be the subspace of sequences with finitely many non-vanishing entries. We can construct a Cauchy sequence in this subspace, but with a limit that has infinitely many entries. For example, consider the sequence of sequences $(x_m)_{m\in\mathbb{N}}$ such that

$$x_m(n) = \begin{cases} \frac{1}{2^{|n|}}, & \text{if } |n| \leq m, \\ 0, & \text{otherwise.} \end{cases} \tag{1.1.63}$$

This sequence is a Cauchy sequence. Indeed, notice that

$$\|x_m - x_l\|_1 = \frac{2^{m+1} - 2^{l+1}}{2^{l+m}} \tag{1.1.64}$$

for any $l, m > 0$. This expression can be made arbitrarily small by taking l and m sufficiently large, and hence the sequence is Cauchy. Nevertheless, its limit is

$$x_\infty(n) = \frac{1}{2^n}, \tag{1.1.65}$$

which does not lie on U.

Now pick $V = \ell^\infty(\mathbb{Z})$. We can make a similar example by picking U to be the space of sequences with finitely many non-vanishing entries. This time we can pick different sequences, though. For example, consider the sequence of sequences

$$x_m(n) = \begin{cases} 0, & \text{if } n < 0 \text{ or } n > m, \\ \frac{\lfloor 10^n \pi \rfloor}{10^n}, & \text{otherwise.} \end{cases} \tag{1.1.66}$$

In this case, the limiting sequence is not summable, but it is bounded. It is given by successive approximations to π.

Another example occurs in the space $L^1(\mathbb{R})$ of integrable functions. We can pick a sequence of functions with support in a compact set. We then pick a smaller compact interval inside this compact set and repeat the construction given on Example 1.8 on page 8. ♥

1.2 Completing Normed Spaces

It is often more interesting to work with Banach spaces than other normed spaces for mere technical convenience. It is useful to know we are working in a space without "holes". However, what happens if we are interested in a space that, at least in principle, is not complete? As it turns out, it is possible to find a larger space that is complete and includes a copy of the original space. Our present goal is to state this result properly and then prove it. To do so, it is useful to recall some constructions involving metric spaces, since these completion procedures occur in ordinary metric spaces as well.

Definition 1.25 [Isometry]**:**

Let (M_1, d_1) and (M_2, d_2) be metric spaces. A function $f\colon M_1 \to M_2$ is said to be an *isometry* if, and only if, $d_2(f(x), f(y)) = d_1(x, y)$ for all $x, y \in M_1$. ♠

Remark:

Notice that a necessary condition for a function to be an isometry is that it is injective. Indeed, suppose $x \neq y$ are such that $f(x) = f(y)$. Then $d_2(f(x), f(y)) = 0$, but $d_1(x, y) \neq 0$. ♣

Definition 1.26 [Isometric Spaces]**:**

Two metric spaces (M_1, d_1) and (M_2, d_2) are said to be *isometric* if, and only if, there exists a bijective isometry $f\colon M_1 \to M_2$. ♠

Isometric spaces are indistinguishable from a topological perspective. In particular, if one of them is complete, so is the other.

Proposition 1.27:

Let (M_1, d_1) and (M_2, d_2) be isometric metric spaces. Assume (M_2, d_2) is complete. Then so is (M_1, d_1). □

Proof:

We'll sketch the proof. Let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in (M_1, d_1). We want to show it converges to a point $x \in M_1$. We can simply consider the sequence $(f(x_n))$ in M_2, pick the point $y = \lim_{n\to\infty} f(x_n) \in M_2$ that exists there ($(f(x_n))$ is Cauchy, since the spaces are isometric) and then notice that $x_n \to f^{-1}(y)$ in M. ■

Let us now prove that any metric space is isometric to a dense subset of a complete metric space.

Theorem 1.28:

Let (M,d) be a metric space. Then it holds that (M,d) is isometric to a dense subset of a complete metric space $(\tilde{M},\tilde{d})$, known as the completion of (M,d). Furthermore, any two completions of (M,d) are isometric. □

Proof:

The proof is similar to Cantor's construction of the real numbers, which involves representing each real number as an equivalence class of Cauchy sequences of rational numbers (see, *e.g.*, Tao 2022, Sec. 5.3).

Consider the Cauchy sequences on M. We shall introduce an equivalence relation $\sim$ such that

$$(x_n) \sim (y_n) \Leftrightarrow \lim_{n\to+\infty} d(x_n,y_n) = 0. \tag{1.2.1}$$

It can be checked that this is indeed an equivalence relation (the triangle inequality can be used to prove transitivity). Let $\tilde{M}$ be the set of equivalence classes of $\sim$.

Let us introduce a metric in $\tilde{M}$. Given $\tilde{x},\tilde{y} \in \tilde{M}$ define

$$\tilde{d}(\tilde{x},\tilde{y}) = \lim_{n\to+\infty} d(x_n,y_n), \tag{1.2.2}$$

where the sequences (x_n) and (y_n) are representatives of the equivalence classes $\tilde{x}$ and $\tilde{y}$, respectively. We want to show that $\tilde{d}$ is well-defined and that it is a metric.

We begin by showing that the limit on the right-hand side (RHS) of Eq. (1.2.2) exists. Notice that the triangle inequality implies

$$d(x_n,y_n) \leq d(x_n,x_m) + d(x_m,y_m) + d(y_m,y_n), \tag{1.2.3}$$

and hence

$$\left|d(x_n,y_n) - d(x_m,y_m)\right| \leq d(x_n,x_m) + d(y_m,y_n). \tag{1.2.4}$$

Since (x_n) and (y_n) are Cauchy sequences, it holds that $(d(x_n,y_n))_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$. Since $\mathbb{R}$ is complete, the limit in Eq. (1.2.2) exists. Let us now prove it does not depend on the choice of representative.

Let $(x'_n) \sim (x_n)$ and $(y'_n) \sim (y_n)$. Then notice how

$$\lim_{n\to+\infty} d(x'_n,y'_n) \leq \lim_{n\to+\infty} d(x'_n,x_n) + \lim_{n\to+\infty} d(x_n,y_n) + \lim_{n\to+\infty} d(y_n,y'_n), \tag{1.2.5a}$$

$$\leq \lim_{n\to+\infty} d(x_n,y_n), \tag{1.2.5b}$$

and analogously we prove the reverse inequality. Therefore, both cases coincide and the function in Eq. (1.2.2) is indeed well-defined. It can be seen to be a metric by simply using the properties of the metric d.

Let us now find an isometry. Consider the function $f\colon M \to \tilde{M}$ that takes x to the equivalence class represented by $(x, x, x, \ldots)$. f is then an isometry with a dense range in $\tilde{M}$. Indeed, notice that, given $x, y \in M$, we have

$$\tilde{d}(f(x), f(y)) = \lim_{n\to+\infty} d(f(x)_n, f(y)_n), \tag{1.2.6a}$$

$$= \lim_{n\to+\infty} d(x, y), \tag{1.2.6b}$$

$$= d(x, y), \tag{1.2.6c}$$

which proves it is an isometry. Let us prove that its range is dense in $\tilde{M}$.

Let $\tilde{x} \in \tilde{M}$. Pick $\epsilon > 0$. We want to find $\tilde{y} \in \tilde{M}$ of the form $f(z)$ ($z \in M$) such that $\tilde{d}(\tilde{x}, \tilde{y}) < \epsilon$. Let (x_n) be a representative of $\tilde{x}$. Pick $m \in \mathbb{N}$. Notice that

$$\tilde{d}(\tilde{x}, f(x_m)) = \lim_{n\to+\infty} d(x_n, f(x_m)_n), \tag{1.2.7a}$$

$$= \lim_{n\to+\infty} d(x_n, x_m). \tag{1.2.7b}$$

Since (x_n) is a Cauchy sequence, we get $\tilde{d}(\tilde{x}, f(x_m)) < \epsilon$ for sufficiently large m.

Notice that, since all isometries are injective, it holds already that M is isometric to a subset of $\tilde{M}$. Therefore, our next goal is to prove completeness and uniqueness up to isometry of $(\tilde{M}, \tilde{d})$.

Let $(\tilde{x}_n)$ be a Cauchy sequence in $\tilde{M}$. Since we know $\operatorname{Ran} f$ to be dense in $\tilde{M}$ we know that, for every $n \in \mathbb{N}$, there is $y_n \in M$ such that $\tilde{d}(\tilde{x}_n, f(y_n)) < \frac{1}{n}$. Notice now that

$$d(y_n, y_m) = \tilde{d}(f(y_n), f(y_m)), \tag{1.2.8a}$$

$$\leq \tilde{d}(f(y_n), \tilde{x}_n) + \tilde{d}(\tilde{x}_n, \tilde{x}_m) + \tilde{d}(\tilde{x}_m, f(y_m)), \tag{1.2.8b}$$

$$\leq \frac{1}{n} + \tilde{d}(\tilde{x}_n, \tilde{x}_m) + \frac{1}{m}. \tag{1.2.8c}$$

Therefore, since $(\tilde{x}_n)$ is Cauchy in $\tilde{M}$, (y_n) is a Cauchy sequence in M. Therefore, there is some element $\tilde{y} \in \tilde{M}$ represented by (y_n). Notice that

$$\tilde{d}(\tilde{x}_n, \tilde{y}) = \lim_{m\to+\infty} d((\tilde{x}_n)_m, y_m), \tag{1.2.9a}$$

$$\leq \lim_{m\to+\infty} d((\tilde{x}_n)_m, f(y_n)_m) + \lim_{m\to+\infty} d(f(y_n)_m, y_m), \tag{1.2.9b}$$

$$= \tilde{d}(\tilde{x}_n, f(y_n)) + \lim_{m\to+\infty} d(y_n, y_m), \tag{1.2.9c}$$

$$< \frac{1}{n} + \lim_{m\to+\infty} d(y_n, y_m). \tag{1.2.9d}$$

Hence, $\tilde{x}_n \to \tilde{y}$ as $n \to +\infty$, proving $\tilde{M}$ is indeed complete.

To prove uniqueness, suppose there is some other metric space (M', d') that is a completion of (M, d) with isometry $g\colon M \to M'$. Notice that $f \circ g^{-1}\colon \operatorname{Ran} g \to \operatorname{Ran} f$ is a bijective isometry. Since $\operatorname{Ran} g$ and $\operatorname{Ran} f$ are dense, $f \circ g^{-1}$ admits a unique continuous extension to an isometry between M' and $\tilde{M}$. This concludes the proof. ■

Example 1.29 [Completion of $\mathbb{Q}$]**:**

If we complete the rational numbers, $\mathbb{Q}$, with respect to the standard metric in $\mathbb{Q}$ we end up with the real numbers. This is the Cantor construction of the real numbers (see, *e.g.*, Tao 2022, Sec. 5.3). ♥

Example 1.30 [Completion of the Discrete Metric]**:**

Consider a set M endowed with the discrete metric,

$$d(x, y) = \begin{cases} 0, & \text{if } x = y, \\ 1, & \text{if } x \neq y. \end{cases} \tag{1.2.10}$$

In this case, notice a sequence is Cauchy if, and only if, it is eventually constant. As a consequence, any set endowed with the discrete metric is automatically complete. ♥

One might complain, though, that a Banach space is more than merely a complete metric space. Indeed: it also has linear structure. We want to show that one can complete a normed space to a Banach space without spoiling the original linear structure. Let us prove it.

Theorem 1.31:

Let $(V, \|\cdot\|)$ be a normed space. Then it is isomorphic to a dense subspace of a Banach space $(V', \|\cdot\|')$. $(V', \|\cdot\|')$ is said to be the completion of $(V, \|\cdot\|)$ and any two completions of the same normed space are isomorphic. □

Proof:

We know that $(V, \|\cdot\|)$ is endowed with a natural metric $d(x, y) = \|x - y\|$, which also has the property $d(x, y) = d(x + z, y + z) = \lambda d(\lambda^{-1} x, \lambda^{-1} y), \forall\, z \in V$. This metric space can be completed to a metric space (V', d'). We know (V, d) is dense in (V', d').

Let us now show that V' is also a vector space. We already know the subspace V is a vector space, so we only need to show that we can add and multiply by scalars the vectors of V'.

Let $x, y \in V'$. Then, since V is dense, there are sequences $x_n, y_n \in V$ such that $x_n \to x$ and $y_n \to y$. We define addition and the scalar product in V' through

$$x + \alpha y = \lim_{n\to\infty} x_n + \alpha \lim_{n\to\infty} y_n, \tag{1.2.11a}$$

$$= \lim_{n\to\infty} (x_n + \alpha y_n). \tag{1.2.11b}$$

Since (x_n) and (y_n) converge, so does $(x_n + \alpha y_n)$.

We now get to the norm. To show we completed the normed space to a Banach space it is enough to notice that

$$d'(x, y) = d'\left(\lim_{n\to+\infty} x_n, \lim_{m\to+\infty} y_m\right), \tag{1.2.12a}$$

$$= \lim_{n\to+\infty} \lim_{m\to+\infty} d'(x_n, y_m), \tag{1.2.12b}$$

$$= \lim_{n\to+\infty} \lim_{m\to+\infty} d(x_n, y_m). \tag{1.2.12c}$$

We can then use this result to show that $d'(x,y) = d'(x+z, y+z) = \lambda d'(\lambda^{-1}x, \lambda^{-1}y), \forall\, z \in V'$. And define the norm through $\|x\| = d'(x, 0)$.

Finally, we must prove uniqueness. Suppose there is a second completion V''. Then, of course, the copy of V inside each completion is isomorphic to V itself, and hence to one another. If there is some $x' \in V' \setminus V$, then there is a sequence (x_n) with $x_n \to x'$ and the same sequence can be used in V'' to get $x_n \to x''$. Through this procedure, one can fill in the details to show that V' and V'' are isomorphic. ■

1.3 Separable Spaces

Hamel bases are always available, but they are often uncountable in infinite dimensions. As a consequence, their utility is limited. It would be interesting for us to obtain a notion of basis that is more practical.

Definition 1.32 [Schauder Basis]**:**

Let V be a normed space. A *Schauder basis* is a sequence $(v_n) \in V$ such that, for every vector $v \in V$, there is a unique sequence $(\alpha_n) \in \mathbb{K}$ such that

$$v = \sum_{n=1}^{+\infty} \alpha_n v_n = \lim_{N \to +\infty} \sum_{n=1}^{N} \alpha_n v_n. \tag{1.3.1}$$

Since the sequence of scalars must be unique, the basis must be linearly independent. Furthermore, notice that finite-dimensional spaces are contemplated within this definition. ♠

Definition 1.33 [Separable]**:**

A metric space is said to be *separable* if, and only if, it admits a countable dense subset. ♠

Example 1.34 [Discrete Metric]**:**

A space endowed with the discrete metric is separable if, and only if, it is countable. ♥

Proposition 1.35:

Let V be a normed space.

i. *If V admits a Schauder basis, it is separable.*

ii. *V is separable if, and only if, it admits a linearly independent countable total subset.* □

Proof:

i. Suppose V admits a Schauder basis (v_n). Consider the set

$$X = \left\{ \sum_{n=1}^{+\infty} \alpha_n v_n ; \alpha_n \in \mathbb{Q}, \forall\, n \in \mathbb{N} \right\}. \tag{1.3.2}$$

Since every $v \in V$ can be written as $\sum_{n=1}^{+\infty} \alpha_n v_n$ for real numbers α_n and $\mathbb{Q}$ is dense in $\mathbb{R}$, it holds that X is dense in V. Hence, since X is countable, V is separable.

ii. Suppose V admits a linearly independent countable total subset (v_n). Then X as defined on Eq. (1.3.2) is countable and dense in V. Hence, V is separable.

Suppose now V is separable and (v_n) is a dense sequence in V. We want to build a linearly independent countable total subset. We define (u_n) inductively by having u_1 be the first non-vanishing element of (v_n) and u_{i+1} being the first element of (v_n) such that $(u_n)_{n=1}^{i+1}$ is linearly independent. Through this construction, we ensure that (v_n) and (u_n) span the same (dense) subspace, but now with a linearly independent set.

■

The existence of a Schauder basis is enough for a normed space to be separable. Nevertheless, it should be mentioned that, as shown by Enflo (1973), not every separable Banach space admits a Schauder basis. Many interesting applications and results of functional analysis focus on separable spaces, which considerably simplifies the theory.

Proposition 1.36:

Let X be a subset of a normed space V. X is separable if, and only if, $\operatorname{span}(X)$ *is separable.*

□

Proof:

Suppose X is separable. Then it admits a countable dense subset. Taking finite rational linear combinations of elements of this countable dense subset yields a countable dense subset of $\operatorname{span}(X)$.

Suppose now that $\operatorname{span}(X)$ is separable. Then it admits a countable dense subset. By intersecting this set with X, one gets a countable dense subset of X. Hence, X is separable. ■

Example 1.37 [$\mathbb{K}^n$]**:**

Since the rational numbers are dense in $\mathbb{R}$, it follows that $\mathbb{K}^n$ is separable for $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Alternatively, this can be seen by noticing that the canonical basis for $\mathbb{K}^n$ is a Schauder basis. ♥

Example 1.38 [$\ell^p(\mathbb{N})$]**:**

$\ell^p(\mathbb{N})$ is separable. This can be seen by noticing that the canonical basis for $\ell^p(\mathbb{N})$ is a Schauder basis or by noticing that sequences of rational numbers are dense in $\ell^p(\mathbb{N})$. ♥

Example 1.39 [$\ell^\infty(\mathbb{N})$]**:**

$\ell^\infty(\mathbb{N})$ is not separable. Indeed, let $v = (v_1, v_2, \ldots)$ be a sequence of zeros and ones. It holds that $v \in \ell^\infty(\mathbb{N})$, for $\|v\|_\infty \leq 1$. There is a bijection between such sequences and the numbers in $[0,1]$ given by the binary basis representation

$$v \mapsto \sum_{n=1}^{+\infty} \frac{v_n}{2^n} = \frac{v_1}{2^1} + \frac{v_2}{2^2} + \ldots. \tag{1.3.3}$$

Since there are uncountably many numbers in $[0,1]$, we see there are uncountably many such sequences of zeros and ones. Furthermore, in the norm $\|\cdot\|_\infty$, any two such sequences that are not equal are a distance 1 apart.

Consider now each of these sequences as the center of a ball of radius $\frac{1}{3}$. There are uncountably many such balls and no two of them intersect. If X is to be a dense subset of $\ell^\infty(\mathbb{N})$, then it must have at least one element in each such ball. Therefore, it must be uncountable.

The canonical "basis" now fails to be a Schauder basis (otherwise, the space would be separable). To see this directly, denote the canonical basis by (e_n). One would expect that the element $v = (1, 1, 1, \ldots)$ would be written as

$$v = \sum_{n=1}^{+\infty} e_n, \tag{1.3.4}$$

but notice that

$$\left\| v - \sum_{n=1}^{N} e_n \right\| = 1 \tag{1.3.5}$$

for every N, and hence the series does not actually converge. ♥

Example 1.40 [A Schauder basis that is not a Hamel basis]**:**

Let V be the subspace of $\ell^\infty(\mathbb{N})$ formed by the sequences with finitely many non-vanishing entries. The canonical basis (e_n) is a Schauder basis and a Hamel basis as well, since $V = \operatorname{span}(\{e_n\})$. Consider, however, the sequence

$$v_n = \frac{e_n}{n} - \frac{e_{n+1}}{n+1}. \tag{1.3.6}$$

This is not a Hamel basis, for the only possible representation of e_1 is the infinite series

$$e_1 = \sum_{n=1}^{+\infty} v_n. \tag{1.3.7}$$

Nevertheless, this is yet a Schauder basis, for each element $u = (u_1, \ldots, u_n, 0, 0, \ldots)$ of V can be written as

$$u = \sum_{k=1}^{n} a_k v_k + a_n \sum_{k=n+1}^{+\infty} v_k, \tag{1.3.8}$$

where the coefficients a_k are the solutions of the linear system defined by $u_1 = a_1$ and $u_j = (a_j - a_{j-1})/j$ for $2 \leq j \leq n$. Hence, even if the Hamel bases are enumerable, there might be a Schauder basis that is not a Hamel basis. ♥

Example 1.41 [$\mathscr{C}^0([a,b])$]**:**

Proposition 1.35 on page 21 implies $\mathscr{C}^0([a,b])$ is separable, for (x^n) is total as a consequence of the Stone–Weierstrass theorem. Since $\mathscr{C}^0([a,b])$ is dense in $L^p([a,b])$ for $1 \leq p < \infty$, it follows that $L^p([a,b])$ is too separable. ♥

1.4 Linear Operators

To further understand a structure in Mathematics, it is usual to study "structure-preserving functions", or "morphisms" in the language of Category Theory (Geroch 1985), and hence so will we do here.

Definition 1.42 [Linear Operator]**:**

A *linear operator* between the $\mathbb{K}$-vector spaces U and V is a function $T\colon \operatorname{Dom} T \subseteq U \to V$, where $\operatorname{Dom} T$ is a vector subspace of U and it holds that $T(u + \lambda v) = Tu + \lambda Tv$ for all $u, v \in \operatorname{Dom} T$ and all $\lambda \in \mathbb{K}$. ♠

Notice that $T(0) = 0$ for all linear operators. Furthermore, operators with the same domain define a vector space with the pointwise operations. Two simple examples of operators are the identity operator, $\mathbb{1}u = u$, and the null operator, $Tu = 0$.

Example 1.43 [Differentiation]**:**

Let V be the space of polynomials defined on $\mathbb{R}$. Then define the operator $T\colon V \to V$ through $Tf = f'$. T is a linear operator. ♥

Example 1.44 [Integration]**:**

Let $V = \mathscr{C}^0([a,b])$. An operator $T\colon V \to V$ can be defined through

$$Tf(x) = \int_a^x f(y)\,\mathrm{d}y\,. \tag{1.4.1}$$

♥

Definition 1.45 [Range and Kernel]**:**

We define the *range* of an operator $T\colon U \to V$ to be the set $\operatorname{Ran} T = T(\operatorname{Dom} T)$. We define the *kernel*, or *null space*, of T through $\operatorname{Ker} T = \{u \in \operatorname{Dom} T; Tu = 0\}$. ♠

Proposition 1.46:

The range and the kernel of a linear operator are vector spaces. □

Proof:

Let $T\colon U \to V$ be an operator. Let us prove that $\operatorname{Ran} T$ is a vector space.

Firstly, $0 \in \operatorname{Ran} T$, for $T(0) = 0$. Secondly, if $u, v \in \operatorname{Ran} T$ and $\lambda \in \mathbb{K}$, then $u + \lambda v \in \operatorname{Ran} T$. Indeed, there are $x, y \in U$ such that $u = Tx$ and $v = Ty$, and therefore $u + \lambda v = T(x + \lambda y)$. This concludes the proof.

Next, let us prove the kernel is also a vector space. Firstly, we recognize that $T(0) = 0$, and hence $0 \in \operatorname{Ker} T$. Furthermore, suppose $u, v \in \operatorname{Ker} T$ and $\lambda \in \mathbb{K}$. Then it is straightforward to check that $T(u + \lambda v) = 0$, and hence $u + \lambda v \in \operatorname{Ker} T$. ■

Proposition 1.47:

Let $T: U \to V$ be an operator. If $\dim(\operatorname{Dom} T) = n < \infty$, *then* $\dim(\operatorname{Ran} T) \leq n$. □

Proof:

If $(v_k)_{k=1}^m$ is a basis for $\operatorname{Ran} T$, then the vectors u_k such that $Tu_k = v_k$ are linearly independent. Hence, $\dim(\operatorname{Ran} T) = m \leq n$. ■

Proposition 1.48:

The inverse operator of $T: U \to V$, $T^{-1}: \operatorname{Ran} T \to \operatorname{Dom} T$, exists if, and only if, $\operatorname{Ker} T = \{0\}$. *If it exists, then it is a linear operator.* □

Proof:

If T^{-1} exists, then it is necessary that $\operatorname{Ker} T = \{0\}$. Otherwise, $T^{-1}(0)$ would be multivalued. Furthermore, it must be linear. Indeed, suppose $u, v \in \operatorname{Ran} T$ are such that $u = Tx$ and $v = Ty$ and let $\lambda \in \mathbb{K}$. Then

$$
\begin{aligned}
T^{-1}(u + \lambda v) &= T^{-1}(Tx + \lambda Ty), && (1.4.2a)\\
&= T^{-1}(T(x + \lambda y)), && (1.4.2b)\\
&= x + \lambda y, && (1.4.2c)\\
&= T^{-1}(u) + \lambda T^{-1}(v). && (1.4.2d)
\end{aligned}
$$

Suppose now, on the other hand, that $\operatorname{Ker} T = \{0\}$. For the inverse of $T: \operatorname{Dom} T \to \operatorname{Ran} T$ to exist it suffices for T to be injective, since it is already surjective. Notice that the inverse of this operator is the inverse of $T: U \to V$. Hence, it suffices for us to prove that $\operatorname{Ker} T = \{0\}$ implies T is injective.

Suppose $Tu = Tv$ for $u \neq v$. Then $T(u - v) = 0$ and $u - v \in \operatorname{Ker} T$. Hence, if $\operatorname{Ker} T = \{0\}$, it follows that $Tu = Tv$ implies $u = v$. This proves $\operatorname{Ker} T = \{0\}$ implies the injectivity of T, and hence its inversibility. ■

When we introduce a norm topology on a vector space, we end up with a very interesting theory of linear operators, as exemplified by the following theorem.

Theorem 1.49:

Let $T: U \to V$ be a linear operator between the normed spaces U and V. Then the following propositions are equivalent:

i. $\sup_{\|v\| \leq 1} \|Tv\| < \infty$, i.e., *the range of the unit ball is limited;*

ii. *there is $C > 0$ such that $\|Tv\| \leq C\|v\|$ for all $v \in U$;*

iii. *T is uniformly continuous;*

iv. *T is continuous;*

v. *T is continuous at the origin.* □

Proof:

i. ⇒ ii. Let $C = \sup_{\|v\|\leq 1} \|Tv\|$. Notice that, for any $v \in U$, this implies $\left\|T\left(\frac{v}{\|v\|}\right)\right\| \leq C$, and hence $\|Tv\| \leq C\|v\|$;

ii. ⇒ iii. if $u, v \in U$, then $\|Tu - Tv\| = \|T(u - v)\| \leq C\|u - v\|$;

iii. ⇒ iv. straightforward;

iv. ⇒ v. straightforward;

v. ⇒ i. there is $\delta > 0$ such that $\|Tv\| < 1$ for every $\|v\| < \delta$. Hence, if $\|v\| \leq 1$, then $\|\delta v\| \leq \delta$, $\|T(\delta v)\| \leq 1$, and hence $\|Tv\| \leq \frac{1}{\delta}$. ■

Theorem 1.49 on the preceding page motivates us to introduce new terminology.

Definition 1.50 [Bounded Operators]**:**
We say a linear operator between normed spaces is *bounded* if, and only if, it is continuous. Furthermore, we denote the space of bounded operators from U to V by $\mathscr{B}(U, V)$. For simplicity, we also write $\mathscr{B}(U) = \mathscr{B}(U, U)$. ♠

Remark:
Notice that the meaning of a bounded operator differs from the meaning of a bounded function, *i.e.*, of a function with bounded range. One can show that any non-vanishing operator is an unbounded function. ♣

Example 1.51 [Identity]**:**
The identity operator $\mathbb{1}: U \to U$ is a bounded operator, for $\|\mathbb{1}u\| = \|u\| \leq \|u\|$. ♥

Example 1.52 [Null Operator]**:**
The null operator $T: U \to U$ is a bounded operator, for $\|Tu\| = 0 \leq \|u\|$. ♥

Example 1.53 [Differentiation]**:**
Consider the vector space V of polynomials in $[0, 1]$ with the norm $\|\cdot\|_\infty$. Notice that $\|x^n\| = 1$. Define the differentiation operator $T: V \to V$ through $Tf = f'$. Then notice that $\|Tx^n\| = n\left\|x^{n-1}\right\| = n$. Hence, for any $C > 0$, there is some n such that $\|Tx^n\| > C\|x^n\|$. Therefore, T is unbounded.

Given that differentiation is an operator of practical interest, this shows that our interest will not be limited to bounded operators. ♥

Example 1.54 [Integration]**:**
Let $V = \mathscr{C}^0([a, b])$. Define the operator $T: V \to V$ through

$$Tf(x) = \int_a^x f(y) \, \mathrm{d}y \, . \tag{1.4.3}$$

♥

T is bounded. Indeed,

$$\|Tf\|_\infty = \sup_{a\le x\le b} \int_a^x f(y)\,\mathrm{d}y\,, \tag{1.4.4a}$$
$$\le \sup_{a\le x\le b} \int_a^x \|f\|_\infty\,\mathrm{d}y\,, \tag{1.4.4b}$$
$$\le (b-a)\|f\|_\infty, \tag{1.4.4c}$$

which concludes the proof by means of Theorem 1.49 on page 25.

Proposition 1.55:
Let $T\colon U \to V$ be an operator between normed spaces. If $\dim U < \infty$, *then T is bounded.* □

Proof:
Equip U with the norm

$$\|u\| = \|u\|_U + \|Tu\|_V. \tag{1.4.5}$$

Theorem 1.19 on page 13 ensures $\|\cdot\|$ and $\|\cdot\|_U$ are equivalent. Hence, there is $C > 0$ such that $\|u\| \le C\|u\|_U$ for all $u \in U$. Since $\|u\| \ge \|Tu\|_V$, it follows that $\|Tu\|_V \le C\|u\|_U$ for all $u \in U$. Therefore, T is bounded. ■

Proposition 1.56:
Let U be a normed space such that every operator $T\colon U \to U$ is bounded. Then $\dim U < \infty$. □

Proof:
Let $(u_\lambda)_{\lambda\in\Lambda}$ be a Hamel basis for U. By contradiction, we assume Λ to be infinitely large. Consider a countable subset of the basis given by $(u_n)_{n=1}^{+\infty}$. Define $T\colon U \to U$ by

$$Tu_n = nu_n. \tag{1.4.6}$$

Then, since n can be arbitrarily large, T cannot be bounded. ■

1.5 Space of Bounded Operators

We can make $\mathscr{B}(U,V)$ into a normed space. To do so, we equip it with the pointwise operations to make it into a vector space. It is still necessary to equip it with a norm. To do so, we use the following definition.

Definition 1.57 [Norm of a Bounded Operator]**:**
Let $T\colon U \to V$ be a bounded operator. We define its *norm* through

$$\|T\| = \sup_{\substack{u\in\operatorname{Dom} T\\ \|u\|\le 1}} \|Tu\|. \tag{1.5.1}$$

Unless explicitly mentioned otherwise, we will always assume this is the norm equipped on $\mathscr{B}(U,V)$. The topology induced on $\mathscr{B}(U,V)$ by this norm is known as the *uniform operator topology* or *norm topology*. ♠

It is straightforward to check that this indeed defines a norm. We could, however, have picked a few other equivalent definitions.

Proposition 1.58:
Let $T \in \mathscr{B}(U,V)$. Then

$$\|T\| = \inf_{u\in U}\{C > 0; \|Tu\| \le C\|u\|\} = \sup_{\|u\|=1}\|Tu\| = \sup_{u\neq 0}\frac{\|Tu\|}{\|u\|}. \tag{1.5.2}$$

□

Proof:
Let

$$I = \inf_{u\in U}\{C > 0; \|Tu\| \le C\|u\|\}, \tag{1.5.3}$$

$$S_1 = \sup_{\|u\|\le 1}\|Tu\|, \tag{1.5.4}$$

$$S_2 = \sup_{\|u\|=1}\|Tu\|, \tag{1.5.5}$$

$$S_3 = \sup_{u\neq 0}\frac{\|Tu\|}{\|u\|}. \tag{1.5.6}$$

Notice that $S_2 \le S_1$. Furthermore, $\frac{1}{\|u\|}\|Tu\| = \left\|T\left(\frac{u}{\|u\|}\right)\right\|$ implies $S_3 \le S_2$. However, for $\|u\| \le 1$, $\|Tu\| \le \frac{1}{\|u\|}\|Tu\|$, and therefore $S_1 \le S_3$. Hence, $S_1 = S_2 = S_3$.

Next, notice that

$$\|Tu\| \le S_3\|u\|, \forall\, u \in U, \tag{1.5.7}$$

which establishes that $I \le S_3$. Furthermore, from the definition of supremum, we have that for all $n \in \mathbb{N}$ there is $u_n \in U$ such that

$$\frac{\|Tu_n\|}{\|u_n\|} \ge S_3 - \frac{1}{n}. \tag{1.5.8}$$

However, notice that by definition of I it holds that

$$I \ge \frac{\|Tu_n\|}{\|u_n\|} \ge S_3 - \frac{1}{n}, \forall\, n \in \mathbb{N}. \tag{1.5.9}$$

Hence, $I \ge S_3$ and it follows that $I = S_3$. ■

Notice that Proposition 1.58 implies that

$$\|Tu\| \le \|T\|\|u\|, \forall\, u \in U. \tag{1.5.10}$$

Proposition 1.59:

Let $S\colon U \to V$ and $T\colon V \to W$ be bounded linear operators. Assume the operator product TS is well-defined. Then TS is bounded and $\|TS\| \le \|T\|\|S\|$. □

Proof:

We first notice that

$$\|Su\| \le \|S\|\|u\|, \tag{1.5.11}$$
$$\|Tv\| \le \|T\|\|v\|. \tag{1.5.12}$$

Hence,

$$\|TSu\| \le \|T\|\|Su\|, \tag{1.5.13a}$$
$$\le \|T\|\|S\|\|u\|. \tag{1.5.13b}$$

Therefore, TS is bounded. From Proposition 1.58 on the facing page it follows that $\|TS\| \le \|T\|\|S\|$. ■

Example 1.60 [Identity and Null Operator]**:**

The only operator with vanishing norm is the null operator. If $U \ne \{0\}$, then the norm of the identity operator $\mathbb{1}\colon U \to U$ is $\|\mathbb{1}\| = 1$. ♥

Theorem 1.61:

Let U be a normed space and V be a Banach space. Then $\mathscr{B}(U, V)$ is a Banach space. □

Proof:

We know $\mathscr{B}(U, V)$ is a normed space. Hence, we only need to prove that it is complete.

Let (T_n) be a Cauchy sequence of operators. We know that

$$\forall\, \epsilon > 0, \exists\, N \in \mathbb{N}; \|T_n - T_m\| < \epsilon, \forall\, n, m > N. \tag{1.5.14}$$

Notice that, given $u \in U$, we have a sequence $(T_n u)$ in V. This is a Cauchy sequence. Indeed, from the fact that (T_n) is Cauchy it follows that

$$\forall\, \epsilon > 0, \exists\, N \in \mathbb{N}; \|T_n u - T_m u\| < \epsilon\|u\|, \forall\, n, m > N. \tag{1.5.15}$$

Since u is fixed, so is $\|u\|$, and hence $(T_n u)$ is Cauchy.

We can then define an operator $T\colon U \to V$ through

$$Tu = \lim_{n\to+\infty} T_n u, \tag{1.5.16}$$

for the completeness of V ensures the limit exists. Now we want to prove that T is bounded and that $T_n \to T$ in the uniform operator topology.

To show that T is bounded, we will use the fact that the norm is continuous. Notice that

$$\|Tu\| = \left\|\lim_{n\to+\infty} T_n u\right\|, \tag{1.5.17a}$$
$$= \lim_{n\to+\infty} \|T_n u\|, \tag{1.5.17b}$$
$$\leq \lim_{n\to+\infty} \|T_n\|\|u\|. \tag{1.5.17c}$$

Since

$$|\|T_n\| - \|T_m\|| \leq \|T_n - T_m\|, \tag{1.5.18}$$

it follows from the fact that (T_n) is a Cauchy sequence that $(\|T_n\|)$ is a Cauchy sequence, and hence it has a limit C. Hence,

$$\|Tu\| \leq C\|u\|, \tag{1.5.19}$$

proving T is bounded.

Now we want to prove that

$$\forall\, \epsilon > 0, \exists\, N \in \mathbb{N}; \|T - T_n\| < \epsilon, \forall\, n > N. \tag{1.5.20}$$

To do so, we notice that the continuity of the norm implies that for all $\epsilon > 0$, there is some $N \in \mathbb{N}$ such that, for all $u \in U$,

$$\|Tu - T_n u\| = \lim_{m\to+\infty} \|T_m u - T_n u\|, \tag{1.5.21a}$$
$$< \epsilon\|u\| \tag{1.5.21b}$$

for all $n > N$. Hence, for all $\epsilon > 0$, there is some $N \in \mathbb{N}$ such that

$$\sup_{u\neq 0} \frac{\|Tu - T_n u\|}{\|u\|} < \epsilon \tag{1.5.22}$$

for all $n > N$. Hence, $T_n \to T$ in the uniform operator topology. ■

The converse of Theorem 1.61 on the previous page is proven in Proposition 1.128 on page 62.

In a metric space, a uniformly continuous function can be continuously extended to the closure of its domain. Since bounded operators are uniformly continuous, we expect to be able to do the same with them in a normed space.

Definition 1.62 [Extension and Restriction]**:**

Let $f\colon X \to Z$ and $g\colon Y \to Z$ be mappings between sets. f is said to be an *extension* of g, or g is said to be a *restriction* of f, if, and only if, $Y \subseteq X$ and $f(y) = g(y), \forall\, y \in Y$. One writes $f|_Y = g$. ♠

Theorem 1.63 [Bounded Linear Extension]**:**

Let $T : U \to V$ be an operator from the normed space U into the Banach space V. Assume the domain of T, $\operatorname{Dom} T$, is dense in U. Then there is a single extension $\bar{T}$ of T with $\bar{T} \in \mathscr{B}(U, V)$. It holds that $\|\bar{T}\| = \|T\|$. □

Proof:

Let $u \in U$. Since $\operatorname{Dom} T$ is dense in U, we know there is a sequence (u_n) in $\operatorname{Dom} T$ such that $u_n \to u$. Define

$$\bar{T}u = \lim_{n\to+\infty} Tu_n. \tag{1.5.23}$$

To show that the limit exists, it suffices to prove that (Tu_n) is Cauchy, since V is already a Banach space. Hence, we want to prove that

$$\forall\, \epsilon > 0, \exists\, N \in \mathbb{N}; \|Tu_n - Tu_m\| < \epsilon, \forall\, n, m > N. \tag{1.5.24}$$

This follows from the fact that

$$\|Tu_n - Tu_m\| \leq \|T\|\|u_n - u_m\| \tag{1.5.25}$$

and from the fact that (u_n) is a Cauchy sequence (for it converges).

Therefore, $\bar{T}$ defined in such a way is well-defined. It remains to prove that it is bounded with $\|\bar{T}\| = \|T\|$. Using the continuity of the norm we see that

$$\begin{aligned}
\|\bar{T}u\| &= \left\|\lim_{n\to+\infty} Tu_n\right\|, && (1.5.26a)\\
&= \lim_{n\to+\infty} \|Tu_n\|, && (1.5.26b)\\
&\leq \lim_{n\to+\infty} \|T\|\|u_n\|, && (1.5.26c)\\
&= \|T\| \lim_{n\to+\infty} \|u_n\|, && (1.5.26d)\\
&= \|T\|\|u\|. && (1.5.26e)
\end{aligned}$$

Hence, $\bar{T}$ is bounded with $\|\bar{T}\| \leq \|T\|$. However, notice that

$$\begin{aligned}
\|T\| &= \sup_{\substack{u\in\operatorname{Dom} T\\ \|u\|=1}} \|Tu\|, && (1.5.27a)\\
&\leq \sup_{\substack{u\in U\\ \|u\|=1}} \|\bar{T}u\|, && (1.5.27b)\\
&= \|\bar{T}\|, && (1.5.27c)
\end{aligned}$$

and hence $\|\bar{T}\| \geq \|T\|$. We thus conclude that $\|\bar{T}\| = \|T\|$. ■

Corollary 1.64:

Let $T \in \mathscr{B}(U, V)$, where U is a normed space and V is a Banach space. If $\bar{U}$ is the completion of U, there is a single $\bar{T} \in \mathscr{B}(\bar{U}, V)$ that extends T continuously. It holds that $\|\bar{T}\| = \|T\|$. □

Proof:

Follows from Theorem 1.63. ■

1.6 Dual Space

There is a particular case of interest of $\mathscr{B}(U,V)$: that in which $V = \mathbb{K}$, with U being a $\mathbb{K}$-normed space.

Definition 1.65 [Dual Space]**:**
Let U be a $\mathbb{K}$-normed space, with $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. We define the *dual space* of U as the Banach space $U^* = \mathscr{B}(U,\mathbb{K})$. Each element of U^* is said to be a *continuous linear functional* on U. ♠

The dual is ensured to be a Banach space due to Theorem 1.61 on page 29 and the fact that $\mathbb{K}$ is a Banach space.

Example 1.66 [Integration]**:**
The integral over $\mathscr{C}^0([a,b])$ is an element of the dual of $\mathscr{C}^0([a,b])$. Indeed, notice that

$$\psi \mapsto \int_a^b \psi(t)\,\mathrm{d}t \le (b-a)\|\psi\|_\infty \tag{1.6.1}$$

is linear and bounded. ♥

Example 1.67 [L^p Spaces]**:**
Let $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Each $\phi \in L^q(\Omega,\mu)$ defines an element of the dual of $L^p(\Omega,\mu)$, for Hölder's inequality it holds that $\phi\psi \in L^1(\Omega,\mu)$ for all $\psi \in L^p(\Omega,\mu)$, and the mapping

$$\psi \mapsto \int_\Omega \phi\psi\,\mathrm{d}\mu \le \|\psi\|_p\|\psi\|_q \tag{1.6.2}$$

is linear and continuous. Hence, $L^q(\Omega,\mu) \subseteq L^p(\Omega,\mu)^*$. It turns out that, in fact, $L^q(\Omega,\mu) = L^p(\Omega,\mu)^*$ (see, for example, Folland 1999, Theorem 6.15). The result also holds for $p = 1$ if the measure μ is σ-finite. ♥

Let us give explicit proofs for the case of $\ell^p(\mathbb{N})$.

Proposition 1.68:
$\ell^1(\mathbb{N})^* = \ell^\infty(\mathbb{N})$. □

Proof:
To each element of $\ell^1(\mathbb{N})^*$ we will associate an element of $\ell^\infty(\mathbb{N})$. This mapping will be a linear isometry, establishing thus an isomorphism between $\ell^1(\mathbb{N})^*$ and $\ell^\infty(\mathbb{N})$.

Let $\{e_n\}$ be the canonical basis for $\ell^1(\mathbb{N})$ and let $f \in \ell^1(\mathbb{N})^*$. Given $u = (u_n)_{n=1}^\infty = \sum_{n=1}^{+\infty} u_n e_n \in \ell^1(\mathbb{N})$, we have that

$$f(u) = \sum_{n=1}^{+\infty} u_n f(e_n) = \sum_{n=1}^{+\infty} u_n \alpha_n, \tag{1.6.3}$$

where $\alpha_n = f(e_n)$. Thus,

$$|\alpha_n| = |f(e_n)|, \tag{1.6.4a}$$
$$\leq \|f\| \|e_n\|, \tag{1.6.4b}$$
$$= \|f\|. \tag{1.6.4c}$$

Hence, if we define the sequence $\alpha = (\alpha_n)_{n=1}^{\infty}$, it follows that $\|\alpha\|_\infty \leq \|f\|$, and thus $\alpha \in \ell^\infty(\mathbb{N})$. On the other hand, notice that

$$|f(u)| \leq \sum_{n=1}^{+\infty} |u_n||\alpha_n|, \tag{1.6.5a}$$
$$\leq \|\alpha\|_\infty \|u\|_1, \tag{1.6.5b}$$

which implies $\|f\| \leq \|\alpha\|_\infty$. Thus, $\|f\| = \|\alpha\|_\infty$, meaning the linear mapping $f \mapsto \alpha$ is an isometry between $\ell^1(\mathbb{N})$ and a subspace of $\ell^\infty(\mathbb{N})$. The goal is now to prove this mapping is surjective.

Pick now $\beta = (\beta_n)_{n=1}^{\infty} \in \ell^\infty(\mathbb{N})$. Define $g\colon \ell^1(\mathbb{N}) \to \mathbb{K}$ through

$$g(u) = g\left(\sum_{n=1}^{+\infty} u_n e_n\right) = \sum_{n=1}^{+\infty} u_n \beta_n. \tag{1.6.6}$$

Since

$$|g(u)| \leq \sum_{n=1}^{+\infty} |u_n||\beta_n|, \tag{1.6.7a}$$
$$\leq \|\beta\|_\infty \|u\|_1, \tag{1.6.7b}$$

it follows that $g \in \ell^1(\mathbb{N})^*$. Hence, the mapping is surjective. ■

Proposition 1.69:

Let $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Then $\ell^p(\mathbb{N})^* = \ell^q(\mathbb{N})$. □

Proof:

The proof is similar to that for Proposition 1.68 on the preceding page and we shall employ a similar notation. We want to show that for each $f \in \ell^p(\mathbb{N})^*$ there is a $\alpha \in \ell^q(\mathbb{N})$ associated in a linear, isometric, and bijective way.

Let $\{e_n\}$ be the canonical Schauder basis for $\ell^p(\mathbb{N})$ and $f \in \ell^p(\mathbb{N})^*$. Then

$$f(u) = \sum_{n=1}^{+\infty} u_n f(e_n), \tag{1.6.8a}$$
$$= \sum_{n=1}^{+\infty} u_n \alpha_n, \tag{1.6.8b}$$

where $\alpha_n = f(e_n)$.

Define the sequence $u^{(n)} = \left(u_i^{(n)}\right)_{i=1}^{\infty}$ through $u_i^{(n)} = \frac{|\alpha_i|^q}{\alpha_i}$ if $\alpha_i \neq 0$ and $1 \leq i \leq n$ and $u_i^{(n)} = 0$ otherwise. One then obtains

$$f(u^{(n)}) = \sum_{i=1}^{n} |\alpha_i|^q. \tag{1.6.9}$$

On the other hand,

$$\begin{aligned} f(u^{(n)}) &\leq \|f\| \left\|u^{(n)}\right\|_p, &(1.6.10a)\\ &= \|f\| \left(\sum_{i=1}^{+\infty} \left|u_i^{(n)}\right|^p\right)^{\frac{1}{p}}, &(1.6.10b)\\ &= \|f\| \left(\sum_{i=1}^{n} |\alpha_i|^{(q-1)p}\right)^{\frac{1}{p}}, &(1.6.10c)\\ &= \|f\| \left(\sum_{i=1}^{n} |\alpha_i|^{(q-1)p}\right)^{\frac{1}{p}}, &(1.6.10d)\\ &= \|f\| \left(\sum_{i=1}^{n} |\alpha_i|^{q}\right)^{\frac{1}{p}}. &(1.6.10e) \end{aligned}$$

Eqs. (1.6.9) and (1.6.10) lead to

$$\begin{aligned} \sum_{i=1}^{n} |\alpha_i|^q &= f(u^{(n)}), &(1.6.11a)\\ &\leq \|f\| \left(\sum_{i=1}^{n} |\alpha_i|^{q}\right)^{\frac{1}{p}}, &(1.6.11b) \end{aligned}$$

and thus

$$\left(\sum_{i=1}^{n} |\alpha_i|^{q}\right)^{\frac{1}{q}} \leq \|f\|. \tag{1.6.12}$$

Since this holds for all $n \in \mathbb{N}$, it follows that $\|\alpha\|_q \leq \|f\|$, and hence $\alpha \in \ell^q(\mathbb{N})$.

Using Hölder's inequality we have that

$$\begin{aligned} |f(u)| &= \left|\sum_{n=1}^{+\infty} u_n \alpha_n\right|, &(1.6.13a)\\ &\leq \|u\|_p \|\alpha\|_q, &(1.6.13b) \end{aligned}$$

and thus $\|f\| \leq \|\alpha\|_q$, which establishes that $\|f\| = \|\alpha\|_q$. Hence, we found a linear isometry $f \mapsto \alpha$ between $\ell^p(\mathbb{N})^*$ and a subspace of $\ell^q(\mathbb{N})$. It remains to prove this mapping is surjective.

Let $\beta \in \ell^q(\mathbb{N})$. Define $g\colon \ell^p(\mathbb{N}) \to \mathbb{K}$ through

$$g\left(\sum_{n=1}^{+\infty} u_n e_n\right) = \sum_{n=1}^{+\infty} u_n \beta_n. \tag{1.6.14}$$

Notice that Hölder's inequality now leads to

$$\left|g\left(\sum_{n=1}^{+\infty} u_n e_n\right)\right| = \left|\sum_{n=1}^{+\infty} u_n \beta_n\right|, \tag{1.6.15a}$$

$$\leq \|u\|_p \|\beta\|_q, \tag{1.6.15b}$$

and thus $g \in \ell^p(\mathbb{N})^*$. ■

Corollary 1.70:

Let $1 < p < \infty$*.* $\ell^p(\mathbb{N})$ *is reflexive.* □

Proof:

Follows from Proposition 1.69 on page 33. ■

Proposition 1.71:

Let U be a finite-dimensional normed space. Then U^* is also finite-dimensional and it holds that $\dim U^* = \dim U$. □

Proof:

Let $\{u_1, \dots, u_n\}$ be a basis for U. Define f_i through $f_i(u_j) = \delta_{ij}$, where δ_{ij} is the Kronecker delta. The f_i's are linearly independent and any functional $f \in U^*$ can be written as a single linear combination of f_i's. Hence, they form a basis with the same number of elements as $\{u_1, \dots, u_n\}$. Thus, $\dim U = \dim U^*$. ■

1.7 Banach Fixed Point Theorem

The Banach Fixed Point Theorem is a result concerning metric spaces, but many of its applications occur in normed spaces and don't concern only linear transformations. Since it is simple, but powerful, it is worth revisiting.

Firstly, let us recall the concept of a fixed point.

Definition 1.72 [Fixed Point]**:**

Let X be a set. A *fixed point* of a mapping $f\colon X \to X$ is a point $x \in X$ such that $f(x) = x$. ♠

One of the interests of fixed points comes through the fact that they provide a manner for solving differential equations.

Example 1.73 [Differential Equation]**:**

Suppose one wants to find a differentiable function ψ defined on $[t_0 - a, t_0 + a]$ $(a > 0)$ that satisfies the differential equation $\psi'(t) = F(t, \psi(t))$ with the initial condition $\psi(t_0) = \psi_0$, with F continuously differentiable in a neighborhood of (t_0, ψ_0). Then, from the Fundamental Theorem of Calculus, it holds that we are actually looking for all the fixed points of the operator $\Phi\colon \mathscr{C}^\infty([t_0 - a, t_0 + a]) \to \mathscr{C}^\infty([t_0 - a, t_0 + a])$ defined through

$$(\Phi\psi)(t) = \psi_0 + \int_{t_0}^{t} F(s, \psi(s))\,\mathrm{d}s, \qquad t \in [t_0 - a, t_0 + a]. \tag{1.7.1}$$

♥

Suppose you have a mapping $f\colon M \to M$ from a metric space (M, d) into itself. A way to search for fixed points of f is to start at a random point, say x_0, and repeatedly apply f to obtain a sequence $x_n = f(x_{n-1})$. If f is continuous and the sequence converges to some x, then one obtains a fixed point, for

$$\begin{aligned} x &= \lim_{n\to\infty} x_n, && (1.7.2a)\\ &= \lim_{n\to\infty} f(x_{n-1}), && (1.7.2b)\\ &= f(\lim_{n\to\infty} x_{n-1}), && (1.7.2c)\\ &= f(x). && (1.7.2d) \end{aligned}$$

This is called the successive approximations method. For complete metric spaces, a sufficient condition for it to work is given in the following definition.

Definition 1.74 [Contraction]:
Let (M_1, d_1) and (M_2, d_2) be metric spaces. A mapping $f\colon M_1 \to M_2$ is said to be a *contraction* if, and only if, there is a constant $0 \leq \alpha < 1$ such that, for all $x, y \in M_1$, $d_2(f(x), f(y)) \leq \alpha d_1(x, y)$. ♠

Proposition 1.75:
Every contraction between metric spaces is uniformly continuous. □

Proof:
Let (M_1, d_1) and (M_2, d_2) be metric spaces. Let $f\colon M_1 \to M_2$ be a contraction. Then there is a constant $0 \leq \alpha < 1$ such that, for all $x, y \in M_1$, $d_2(f(x), f(y)) \leq \alpha d_1(x, y)$. We want to prove that

$$\forall\, \epsilon > 0, \exists\, \delta > 0; d_1(x, y) < \delta \Rightarrow d_2(f(x), f(y)) < \epsilon, \forall\, y \in M_1. \tag{1.7.3}$$

Given ϵ, simply let $\delta = \frac{\epsilon}{\alpha}$. Then $d_1(x, y) < \delta$ implies

$$\begin{aligned} d_2(f(x), f(y)) &\leq \alpha d_1(x, y), && (1.7.4a)\\ &< \alpha\delta, && (1.7.4b)\\ &= \epsilon. && (1.7.4c) \end{aligned}$$

This concludes the proof. ■

Theorem 1.76 [Banach Fixed Point Theorem]**:**
Let R be a closed subset of the complete metric space (M, d). If the mapping $f: R \to R$ is a contraction, then f has one, and only one, fixed point in R. □

Proof:
Let us first prove uniqueness. Suppose there are two different fixed points of f, x and y. Notice that $d(x, y) > 0$. Since f is a contraction, we have that $d(f(x), f(y)) \leq \alpha d(x, y) < d(x, y)$, for $0 \leq \alpha < 1$. Hence, $d(f(x), f(y)) < d(x, y)$. It is thus impossible that $f(x) = x$ and $f(y) = y$ at the same time, so at least one of them can't be a fixed point.

This proves there is at most one fixed point, but we still need to prove there is at least one. To prove its existence, we shall use the method of successive approximations. Let $x_0 \in M$. Define $x_n = f(x_{n-1})$ for $n \geq 1$. We must prove this sequence converges. If it does, since f is uniformly continuous, it will follow that its limit x is a fixed point for f. Since the metric space is complete, it suffices for us to prove that the sequence is Cauchy.

We want to prove that

$$\forall\, \epsilon > 0, \exists\, N \in \mathbb{N}; d(x_n, x_m) < \epsilon, \forall\, n, m > N. \tag{1.7.5}$$

Inductively, one has $d(x_1, x_2) \leq \alpha d(x_0, x_1)$, $d(x_2, x_3) \leq \alpha d(x_1, x_2) \leq \alpha^2 d(x_0, x_1)$, and $d(x_n, x_{n+1}) \leq \alpha^n d(x_0, x_1)$. Hence,

$$\begin{aligned} d(x_n, x_{n+m}) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \cdots + d(x_{n+m-1}, x_{n+m}), && (1.7.6a) \\ &\leq \alpha^n (1 + \alpha + \cdots + \alpha^{m-1}) d(x_0, x_1), && (1.7.6b) \\ &< \frac{\alpha^n}{1-\alpha} d(x_0, x_1). && (1.7.6c) \end{aligned}$$

Since $\alpha^n \to 0$ as $n \to +\infty$, (x_n) is a Cauchy sequence. Therefore, the result is proven. ■

Corollary 1.77:
Consider the situation (and notation) of Theorem 1.76 and its proof. The sequence (x_n) converges to the fixed point x with error (and "convergence speed") at the n-th iteration given by

$$d(x_n, x) \leq \frac{\alpha^n}{1-\alpha} d(x_0, x_1). \tag{1.7.7}$$

□

Proof:
Take $m \to \infty$ on Eq. (1.7.6) and use the fact that the metric is continuous. ■

Example 1.78:
It is necessary that $\alpha < 1$. For $\alpha = 1$ one can lose both existence and uniqueness of the fixed point. For example, the translation $Tu = u + 1$ has no fixed points and the identity $\mathbb{1}u = u$ has infinitely many. ♥

Example 1.79:
Uniformity in the contraction, characterized by $0 \leq \alpha < 1$, might be essential for the

existence of a fixed point. As paradigmatic examples, pick the functions $\psi, \phi\colon [1,+\infty) \to [1,+\infty)$ defined by $\psi(t) = t + \frac{1}{t}$ and $\phi(t) = t + e^{-t}$. They obey $|\psi(t) - \psi(s)| < |t-s|$ and $|\phi(t) - \phi(s)| < |t-s|$ for every $t, s \in [1,+\infty)$, but have no fixed points. ♥

Corollary 1.80:

Let R be a closed subset of the complete metric space (M, d) and let $f\colon R \to R$ be a mapping. If, for some $m \in \mathbb{N}$, $f^m\colon R \to R$ is a contraction, then f has one, and only one, fixed point in R. Furthermore, for every x_0, the sequence $(f^n(x_0))_{n=1}^{+\infty}$ converges to this fixed point. □

Proof:

Let us first prove uniqueness. Suppose x is a fixed point for f. Then it is too a fixed point for f^m. Since f^m has only one fixed point due to the Banach Fixed Point Theorem, it follows that f can have at most one fixed point.

Next we prove existence. We know f^m has a fixed point—let us call it x. Notice that

$$f^m\big(f(x)\big) = f\big(f^m(x)\big), \tag{1.7.8a}$$
$$= f(x), \tag{1.7.8b}$$

for f and f^m commute. Hence, $f(x)$ is a fixed point for f^m. However, since f^m has a single fixed point, it must hold that $f(x) = x$. Therefore, x is a fixed point for f.

Our next step is to prove that the sequence $(f^n(x_0))_{n=1}^{+\infty}$ converges to this fixed point x regardless of the initial point x_0. To do so, let

$$M = \max_{0 \le k \le m-1} d\big(f^m(x_0), x\big). \tag{1.7.9}$$

Notice that we can now write any n in the form $n = jm + k$ with $j \in \mathbb{N}$ and $0 \le k \le m-1$. Hence,

$$d\big(f^n(x_0), x\big) = d\Big(f^{jm}\big(f^k(x_0)\big), x\Big), \tag{1.7.10a}$$
$$= d\Big(f^{jm}\big(f^k(x_0)\big), f^{jm}(x)\Big), \tag{1.7.10b}$$
$$\le \alpha^j d\big(f^k(x_0), x\big), \tag{1.7.10c}$$
$$\le \alpha^j M. \tag{1.7.10d}$$

For $n \to +\infty$, $j \to +\infty$ and one has $d\big(f^n(x_0), x\big) \to 0$. ■

Let us now discuss a few common applications of the Banach Fixed Point Theorem: Fredholm integral equations, affine equations in Banach spaces, and ordinary differential equations over $\mathbb{R}$. Typically, the difficulty in applying the theorem comes in finding spaces or norms in which the operator of interest is a contraction.

Proposition 1.81:

Let $K\colon Q \to \mathbb{R}$ be a continuous function in the region $Q = [a,b] \times [a,b] \times \mathbb{R}$ satisfying the Lipschitz condition

$$|K(t,s,u) - K(t,s,v)| \le L|u - v|, \quad (t,s,u), (t,s,v) \in Q, \tag{1.7.11}$$

for some $L > 0$. If $\phi \in \mathscr{C}^0([a,b])$, the Fredholm nonlinear integral equation

$$\psi(t) = \int_a^b K(t, s, \psi(s))\,\mathrm{d}s + \phi(t) \tag{1.7.12}$$

has a single solution $\psi \in \mathscr{C}^0([a,b])$ if $L(b-a) < 1$. □

Proof:

We want to prove that the mapping $S\colon \mathscr{C}^0([a,b]) \to \mathscr{C}^0([a,b])$ defined by

$$(S\psi)(t) = \int_a^b K(t, s, \psi(s))\,\mathrm{d}s + \phi(t) \tag{1.7.13}$$

has a single fixed point. To do so, let us first show that it is a contraction in the $\|\cdot\|_\infty$ norm. Indeed, given $\psi, \varphi \in \mathscr{C}^0([a,b])$, it holds for every $t \in [a,b]$ that

$$\begin{aligned} \left|S\psi(t) - S\varphi(t)\right| &\leq \int_a^b \left|K(t, s, \psi(s)) - K(t, s, \varphi(s))\right| \mathrm{d}s\,, && (1.7.14a)\\ &\leq L \int_a^b \left|\psi(s) - \varphi(s)\right| \mathrm{d}s\,, && (1.7.14b)\\ &\leq L \int_a^b \left\|\psi - \varphi\right\|_\infty \mathrm{d}s\,, && (1.7.14c)\\ &\leq L(b-a)\left\|\psi - \varphi\right\|_\infty, && (1.7.14d) \end{aligned}$$

and hence

$$\left\|S\psi - S\varphi\right\|_\infty \leq L(b-a)\left\|\psi - \varphi\right\|_\infty. \tag{1.7.15}$$

Therefore, S is a contraction and the Banach Fixed Point Theorem ensures it has one, and only one, fixed point when $L(b-a) < 1$. ■

Example 1.82:

Let $T \in \mathscr{B}(V)$, where V is a Banach space. If $\|T\| < 1$ and $v \in V$, there is a single $u \in V$ such that $u - Tu = v$. Indeed, we are looking for fixed points of $S\colon V \to V$ defined by $Su = Tu + v$. Notice that

$$\begin{aligned} \|Su - Sw\| &= \|Tu - Tw\|, && (1.7.16a)\\ &\leq \|T\|\|u - w\|, && (1.7.16b) \end{aligned}$$

which is therefore a contraction. The Banach Fixed Point Theorem ensures the result. ♥

Theorem 1.83 [Picard]**:**

Let $U = [t_0 - b, t_0 + b] \times [\psi_0 - r, \psi_0 + r]$, $r, b > 0$. Let $F\colon U \to \mathbb{R}$ be a continuous function satisfying the Lipschitz condition

$$|F(t,u) - F(t,v)| \leq L|u - v|, \quad L > 0 \tag{1.7.17}$$

for $(t,u),(t,v)\in U$. *Let us write* $M=\max_{(t,u)\in U}|F(t,u)|$. *The Cauchy problem*

$$\begin{cases}\dfrac{\mathrm{d}\psi}{\mathrm{d}t}(t)=F(t,\psi(t)),\\ \psi(t_0)=\psi_0\end{cases} \tag{1.7.18}$$

with $\psi_0\in\mathbb{R}$ *has a single differentiable solution* $\psi\colon I\to\mathbb{R}$, *with the interval* $I=[t_0-a,t_0+a]$ *being such that*

$$0<a<\min\left\{\frac{r}{M},\frac{1}{L},b\right\}. \tag{1.7.19}$$

□

Proof:

Let $V=\{\psi\in\mathscr{C}^1(I);|\psi(t)-\psi_0|\le r\}$. V is a Banach space because it is a closed subspace of $\mathscr{C}^0(I)$. We want to find a fixed point of the operator $S\colon V\to V$ defined through

$$(S\psi)(t)=\psi_0+\int_{t_0}^t F(s,\psi(s))\,\mathrm{d}s. \tag{1.7.20}$$

S is well-defined from V to itself. Notice that

$$\begin{aligned}|(S\psi)(t)-\psi_0| &= \left|\int_{t_0}^t F(s,\psi(s))\,\mathrm{d}s\right|, && (1.7.21a)\\ &\le \int_{t_0}^t |F(s,\psi(s))|\,\mathrm{d}s, && (1.7.21b)\\ &\le \int_{t_0}^t M\,\mathrm{d}s, && (1.7.21c)\\ &\le (t-t_0)M. && (1.7.21d)\end{aligned}$$

Since $a\le\frac{r}{M}$, it holds that $|(S\psi)(t)-\psi_0|\le r$.

Let $\psi,\phi\in V$. Then

$$\begin{aligned}|(S\psi)(t)-(S\phi)(t)| &\le \int_{t_0}^t |F(s,\psi(s))-F(s,\phi(s))|\,\mathrm{d}s, && (1.7.22a)\\ &\le L\int_{t_0}^t |\psi(s)-\phi(s)|\,\mathrm{d}s, && (1.7.22b)\\ &\le L\int_{t_0}^t \|\psi-\phi\|_\infty\,\mathrm{d}s, && (1.7.22c)\\ &\le L(t-t_0)\|\psi-\phi\|_\infty. && (1.7.22d)\end{aligned}$$

Hence,

$$\|S\psi-S\phi\|_\infty\le L(t-t_0)\|\psi-\phi\|_\infty \tag{1.7.23}$$

and S is a contraction for $a<\frac{1}{L}$. The Banach Fixed Point Theorem then ensures existence and uniqueness of a fixed point. ■

1.8 Baire Theorem

The Baire Theorem, or Baire Category Theorem, is a result concerning metric spaces that underpins some of the main results of the theory of Banach spaces. To introduce it, we follow the discussions given by Lima (2017, Sec. 7.7) and Oliveira (2018, Cap. 6).

One might want to express that a certain property holds in "most" of a set. The simplest way to do so would be to say it holds in a subset of a certain cardinality, but this does not explore any extra structure we could equip on the set. A sophisticated way is to say, as one does in measure theory, that the property holds almost everywhere with respect to some measure μ. Our goal now will be to give a definition of a similar idea within topology.

Notice that for a property to hold in "most" of a set, it must fail only in an "insignificant" subset. So we settle to define what "insignificant" could mean topologically. The most basic requirement one could impose for a set to be topologically insignificant is that it has a null interior—or equivalently that it has a dense complement.

We also desire that subsets and countable unions of "insignificant" sets to be "insiginificant". This rules out a definition based solely on a set having a null interior, for $\mathbb{Q}$ and $\mathbb{R} \setminus \mathbb{Q}$ have null interior, but their union, $\mathbb{R}$, does not.

An improvement in the definition can then be obtained if we pursue the idea that a set is "insignificant" if its closure has null interior. This time, $\operatorname{int}\left(\overline{X}\right) = \emptyset$ and $\operatorname{int}\left(\overline{Y}\right) = \emptyset$ lead to $\operatorname{int}\left(\overline{X \cup Y}\right) = \operatorname{int}\left(\overline{X} \cup \overline{Y}\right) = \emptyset$. Nevertheless, it is still not true that $\operatorname{int}\left(\overline{X_n}\right) = \emptyset$ for each $n \in \mathbb{N}$ implies $\operatorname{int}\left(\overline{\bigcup_n X_n}\right) = \emptyset$. Indeed, $\mathbb{Q}$ is a countable union of singletons, each of which have null interior. Yet, $\overline{\mathbb{Q}} = \mathbb{R}$.

An appropriate definition of "insignificant" is, however, given by the concept of a meager set.

Definition 1.84 [Rare, Meager, and Nonmeager]**:**
Let M be a topological space. Let $A \subseteq M$. A is said to be

i. *rare* (or *nowhere dense*) in M if, and only if, the interior of its closure is empty;

ii. *meager* (or of *first category*) in M if, and only if, it is contained in a countable union of rare sets;

iii. *nonmeager* (or of *second category*) in M if, and only if, it is not meager in M. ♠

Proposition 1.85:
Let M be a topological space. The following statements are equivalent:

i. *the countable union of rare closed sets in M is a set with null interior,* i.e., *the countable union of closed sets with null interior also has null interior;*

ii. *the countable intersection of open dense sets in M is a dense set in M;*

iii. *all meager sets in M have null interior;*

iv. *the complement of any meager set in M is dense in M;*

v. *any non-empty open set in M is nonmeager in M.* □

Proof:
i. and ii. can be seem to be equivalent by considering the complements of the sets in consideration. It is straightforward to see that iii. is equivalent to iv. and v. At last, it follows from the definition of meager set that i. and iii. are equivalent. ■

Definition 1.86 [Baire Space]**:**
A topological space in which one of the conditions on Proposition 1.85 on the preceding page (and hence all of them) holds is said to be a *Baire space*. ♠

Definition 1.87 [G_δ, F_σ, Residual Set]**:**
A subset of a topological space is a G_δ if it is the countable intersection of open sets. It is an F_σ it it is the countable union of closed sets. It is *residual*, or *generic*, if it is a dense G_δ. ♠

Hence, in a Baire space M, every G_δ obtained by the countable intersection of dense open sets is too dense. Furthermore, if M is the countable union of a family of sets, at least one of them has a non-empty interior.

Theorem 1.88 [Baire]**:**
All complete metric spaces are Baire spaces. □

Proof:
We'll prove condition ii. of Proposition 1.85 on the previous page. Let M be a complete metric space and let $O = \bigcup_{n=1}^{+\infty} O_n$, with O_n open and dense in M. We shall prove that O is too dense, *i.e.*, if $\mathscr{B}_1$ is an arbitrary open ball in M, then $O \cap \mathscr{B}_1 \neq \emptyset$.

$O_1 \cap \mathscr{B}_1 \neq \emptyset$ and it is an open set. Hence, it contains the closure of a smaller open ball $\mathscr{B}_2$ of radius smaller than $\frac{1}{2}$. Again, $O_2 \cap \mathscr{B}_2 \neq \emptyset$ is open and contains the closure of a smaller open ball $\mathscr{B}_3$ of radius smaller than $\frac{1}{3}$. One thus obtains a sequence of open balls such that

$$(O_1 \cap \mathscr{B}_1) \supseteq \mathscr{B}_2 \supseteq \mathscr{B}_3 \cdots \tag{1.8.1}$$

and such that $\overline{\mathscr{B}_{n+1}} \subseteq O_n \cap \mathscr{B}_n$, with $\mathscr{B}_n$ ($n > 1$) having radius smaller than $\frac{1}{n}$. The centers of these open balls define a Cauchy sequence and, since M is complete, there is a single $x \in M$ such that $\{x\} = \bigcap_{n=2}^{+\infty} \overline{\mathscr{B}_n}$. Hence, $x \in O \cap \mathscr{B}_1$. ■

As an application of Baire's Theorem, let us prove the following result.

Proposition 1.89:
Let V be a Banach space. If $\dim V = \infty$, *then any Hamel basis for V is uncountable.* □

Proof:
Suppose there is a countable Hamel basis (e_n) for V. We can assume $\|e_n\| = 1$ for every

n. For each pair of natural numbers n and m define

$$E_{n,m} = \left\{ \left(\sum_{j=1}^{n} \alpha_j e_j \right) \in V;\ \sum_{j=1}^{n} \left|\alpha_j\right| \le m \right\}. \tag{1.8.2}$$

Notice these are closed sets and $V = \bigcup_{n,m} E_{n,m}$. Since V is a complete metric space, the Baire Category Theorem ensures that V is nonmeager in itself (property v. of Proposition 1.85 on page 41). Hence, at least one of the $E_{n,m}$ has a non-vanishing interior. Therefore, for this $E_{n,m}$ there is an open ball $\mathscr{B}(u;\epsilon)$ with $\mathscr{B}(u;\epsilon) \subseteq E_{n,m}$. However, this is impossible, since $u + \frac{\epsilon}{2} e_{n+1} \in \mathscr{B}(u;\epsilon)$, but $u + \frac{\epsilon}{2} e_{n+1} \notin E_{n,m}$. Hence, by contradiction, a Hamel basis cannot be countable. ■

1.9 Principle of Uniform Boundedness

The next few sections will now be dedicated to some interesting consequences of Baire's Theorem. We shall start in this section with the so-called Principle of Uniform Boundedness ("principle" is due to historical reasons), which yields conditions for the norms of an arbitrary family of bounded operators to have a finite upper bound. This is the first of the four cornerstones of functional analysis in normed spaces—the others being the Hahn–Banach Theorem (Section 1.13), the Open Mapping Theorem (Section 1.10), and the Closed Graph Theorem (Section 1.11).

Theorem 1.90 [Principle of Uniform Boundedness]**:**

Let $\{T_\lambda\}_{\lambda\in\Lambda}$ be an arbitrary family of operators on $\mathscr{B}(U,V)$, where U is a Banach space and V is a normed space. If the family is bounded pointwise, i.e., *if for each $u \in U$ one has*

$$\sup_{\lambda\in\Lambda} \|T_\lambda u\| < \infty, \tag{1.9.1}$$

then the family is uniformly bounded, i.e., *it holds that*

$$\sup_{\lambda\in\Lambda} \|T_\lambda\| < \infty. \tag{1.9.2}$$

□

Proof:

Consider the sets $E_k = \{u \in U; \|T_\lambda u\| \le k, \forall\, \lambda \in \Lambda\}$. Notice this is the intersection of the closed sets $T_\lambda^{-1}(\overline{\mathscr{B}_Y(0;k)})$ (which are closed due to the T_λ's being continuous). As a consequence, the E_k's are closed sets. Furthermore, $U = \bigcup_{k=1}^{+\infty} E_k$. Baire's Theorem thus ensures that there is at least one m such that E_m has non-empty interior. Let then $\mathscr{B}_U(u_0,r) \subseteq E_m$ be an open ball. It holds that $\|T_\lambda u\| \le m$ for all $u \in \mathscr{B}_U(u_0,r)$.

Let $v \in U$, $\|v\| = 1$. Then $u = u_0 + \frac{r}{2}v$ is an element of $\mathscr{B}_U(u_0, r)$. As a consequence, we see that

$$\begin{aligned} \|T_\lambda v\| &= \frac{2}{r}\|T_\lambda(u - u_0)\|, && (1.9.3a)\\ &\le \frac{2}{r}\big(\|T_\lambda u\| + \|T_\lambda u_0\|\big), && (1.9.3b)\\ &\le \frac{4m}{r}, && (1.9.3c) \end{aligned}$$

for every $\lambda \in \Lambda$. Therefore,

$$\begin{aligned} \|T_\lambda\| &= \sup_{\|v\|=1} \|T_\lambda v\|, && (1.9.4a)\\ &\le \frac{2m}{r}, && (1.9.4b) \end{aligned}$$

for every $\lambda \in \Lambda$. Finally, this implies that

$$\sup_{\lambda\in\Lambda} \|T_\lambda\| \le \frac{2m}{r} < \infty, \tag{1.9.5}$$

as desired. ■

Corollary 1.91:

Let U be a Banach space. A subset $H \subseteq U^*$ is bounded if, and only if, for every $u \in U$ it holds that $\sup_{f\in H} |f(u)| < \infty$. □

Proof:

If the subset is bounded, then $\sup_{f\in H} \|f\| < \infty$, which implies that, for each $u \in U$,

$$\begin{aligned} |f(u)| &\le \|f\|\,\|u\|, && (1.9.6a)\\ \sup_{f\in H} |f(u)| &\le \sup_{f\in H} \|f\|\,\|u\|, && (1.9.6b)\\ &< \infty. && (1.9.6c) \end{aligned}$$

On the other hand, assume that, for each $u \in U$, $\sup_{f\in H} |f(u)| < \infty$. Thus, the family H of bounded operators is bounded pointwise. The Principle of Uniform Boundedness then implies it is uniformly bounded, which means H is bounded. ■

Corollary 1.92 [Banach–Steinhaus Theorem]**:**

Let U be a Banach space and V be a normed space. Let (T_n) be a sequence in $\mathscr{B}(U, V)$ such that for every $u \in U$ the limit

$$Tu \equiv \lim_{n\to+\infty} T_n u \tag{1.9.7}$$

exists. Then $\sup_n \|T_n\| < \infty$ *and* $T \in \mathscr{B}(U, V)$. □

Proof:

It is straightforward to see that T is a linear operator. Since the limit $\lim_{n\to+\infty} T_n u$ exists for each u, it must hold for each $u \in U$ that $\sup_n \|T_n u\| < \infty$. The Principle of Uniform Boundedness then ensures that $\sup_n \|T_n\| < \infty$.

Next, we notice that

$$\|T_n u\| \leq \|T_n\|\|u\|, \tag{1.9.8a}$$

$$\leq \sup_n \|T_n\|\|u\|, \tag{1.9.8b}$$

$$\lim_{n\to+\infty} \|T_n u\| \leq \left(\sup_n \|T_n\|\right)\|u\|, \tag{1.9.8c}$$

$$\|Tu\| \leq \left(\sup_n \|T_n\|\right)\|u\|, \tag{1.9.8d}$$

and hence T is bounded. ■

Remark:

We adopt the terminology used by Conway (2007, Sec. III.14) and Oliveira (2018, Chap. 7). Sometimes, the term "Banach–Steinhaus Theorem" is used to refer to the Principle of Uniform Boundedness (Brezis 2011, Theorem 2.2; Reed and Simon 1980, p. 81). Other authors might state the Principle of Uniform Boundedness only for countable families of operators (Kreyszig 1978, Theorem 4.7-3). ♣

1.10 Open Mapping Theorem

The next big theorem in our journey through Banach spaces is the Open Mapping Theorem. As we shall define, an open mapping is a mapping that maps open sets to open sets. The Open Mapping Theorem gives conditions for a bounded operator to be an open mapping, and can later be used to give conditions for the inverse of a bounded operator to be bounded as well. Furthermore, it serves as yet another example of why Banach spaces have a richer theory than other normed spaces.

Definition 1.93 [Open Mapping]**:**

Let M_1 and M_2 be topological spaces and let $f : M_1 \to M_2$ be a mapping. f is said to be *open* if, and only if, $f(O)$ is open whenever O is open. ♠

Not all continuous invertible maps are open. Let us give an example.

Example 1.94:

Consider the identity map from $\mathbb{R}^n$ with the discrete topology to $\mathbb{R}^n$ with the standard topology. This map is continuous and invertible, but it is not open. Hence, its inverse is not continuous. ♥

Before proving the Open Mapping Theorem, however, we shall follow Kreyszig (1978, Sec. 4.12) and first prove an auxiliary lemma.

Lemma 1.95 [Open Unit Ball]**:**

Let U and V be Banach spaces. If $T \in \mathscr{B}(U,V)$ and $\operatorname{Ran} T = V$. *Then $T(\mathscr{B}(0;1))$ contains an open ball about $0 \in V$.* □

Proof:

We shall use the shorthand $\mathscr{B}_0 = \mathscr{B}(0;1)$. The proof will come in three steps. Namely,

i. firstly, we prove that the closure of the image of the open ball $\mathscr{B}_1 = \mathscr{B}(0;\frac{1}{2})$ contains an open ball $\mathscr{B}^*$;

ii. next, we prove that $\overline{T(\mathscr{B}_n)}$ contains an open ball V_n about $0 \in V$, with $\mathscr{B}_n = \mathscr{B}(0;2^{-n})$;

iii. finally, we prove that $T(\mathscr{B}_0)$ contains an open ball about $0 \in V$.

Let us now fill in the details.

i. Given a subset $A \subseteq U$ we shall write (for $\lambda \in \mathbb{K}$ and $v \in U$)

$$\lambda A = \{u \in U; u = \lambda a, a \in A\}, \tag{1.10.1}$$

$$A + v = \{u \in U; u = a + v, a \in A\}. \tag{1.10.2}$$

We shall use a similar notation for V.

Consider now $\mathscr{B}_1 = \mathscr{B}(0;\frac{1}{2}) \subseteq U$. Notice that any $u \in U$ lies in some $k\mathscr{B}_1$ for sufficiently large k (say $k > 2\|u\|$). Since $\operatorname{Ran} T = V$ and T is linear, we find that

$$V = \operatorname{Ran} T, \tag{1.10.3a}$$

$$= T(U), \tag{1.10.3b}$$

$$= T\left(\bigcup_{k=1}^{+\infty} k\mathscr{B}_1\right), \tag{1.10.3c}$$

$$= \bigcup_{k=1}^{+\infty} T(k\mathscr{B}_1), \tag{1.10.3d}$$

$$= \bigcup_{k=1}^{+\infty} kT(\mathscr{B}_1), \tag{1.10.3e}$$

$$= \bigcup_{k=1}^{+\infty} \overline{kT(\mathscr{B}_1)}. \tag{1.10.3f}$$

At this stage, we can use Baire's Theorem to ensure that there is some m such that $\overline{mT(\mathscr{B}_1)}$ has a non-empty interior. It then follows from a scaling argument that $\overline{T(\mathscr{B}_1)}$ has a non-empty interior, and therefore contains an open ball $\mathscr{B}^* = \mathscr{B}(u_0,\epsilon)$. Thus, it follows that

$$\mathscr{B}^* - u_0 = \mathscr{B}(0,\epsilon) \subseteq \overline{T(\mathscr{B}_1)} - u_0. \tag{1.10.4}$$

ii. We will now show that $\mathscr{B}^* - u_0 \subseteq \overline{T(\mathscr{B}_0)}$ by proving that $\overline{T(\mathscr{B}_1)} - u_0 \subseteq \overline{T(\mathscr{B}_0)}$.

Let $u \in \overline{T(\mathscr{B}_1)} - u_0$. Then $u + u_0 \in \overline{T(\mathscr{B}_1)}$. Hence, we know there is a sequence $(v_n) = (Tw_n)$ in $T(\mathscr{B}_1)$ with $v_n \to u + u_0$. Similarly, there is a sequence $(x_n) = (Ty_n)$ in $T(\mathscr{B}_1)$ with $x_n \to u_0$. Notice that $w_n, y_n \in \mathscr{B}_1 = \mathscr{B}(0; \frac{1}{2})$, and therefore

$$\begin{aligned} \|w_n - y_n\| &\leq \|w_n\| + \|y_n\|, && (1.10.5a)\\ &< \frac{1}{2} + \frac{1}{2}, && (1.10.5b)\\ &= 1, && (1.10.5c) \end{aligned}$$

which implies $w_n - y_n \in \mathscr{B}_0$. Hence,

$$\begin{aligned} T(w_n - y_n) &= Tw_n - Ty_n, && (1.10.6a)\\ &= v_n - x_n, && (1.10.6b)\\ &\to u, && (1.10.6c) \end{aligned}$$

which proves $u \in \overline{T\mathscr{B}_0}$. Hence,

$$\mathscr{B}^* - u_0 = \mathscr{B}(0; \epsilon) \subseteq \overline{T(\mathscr{B}_1)} - u_0 \subseteq \overline{T(\mathscr{B}_0)}, \tag{1.10.7}$$

which means $\overline{T(\mathscr{B}_0)}$ contains an open ball about $0 \in V$. By a scaling argument, the same holds for $\overline{T(\mathscr{B}_n)}$, $n \in \mathbb{N}$. More specifically,

$$V_n = \mathscr{B}(0; \epsilon 2^{-n}) \subseteq \overline{T(\mathscr{B}_n)}. \tag{1.10.8}$$

iii. At last, we shall prove that

$$V_1 = \mathscr{B}\left(0; \frac{\epsilon}{2}\right) \subseteq T(\mathscr{B}_0). \tag{1.10.9}$$

Let then $u \in V_1$. We know $V_1 \subseteq \overline{T(\mathscr{B}_1)}$, and thus $u \in \overline{T(\mathscr{B}_1)}$. Hence, there is some $v \in T(\mathscr{B}_1)$ with $\|u - v\| < \frac{\epsilon}{4}$. $v \in T(\mathscr{B}_1)$ implies the existence of $x_1 \in \mathscr{B}_1$ with $v = Tx_1$, and hence we have

$$\|u - Tx_1\| < \frac{\epsilon}{4}. \tag{1.10.10}$$

We thus have

$$u - Tx_1 \in V_2 \subseteq \overline{T(\mathscr{B}_2)}. \tag{1.10.11}$$

Repeating the argument we can find $x_2 \in \mathscr{B}_2$ with

$$u - T(x_1 + x_2) \in V_3 \subseteq \overline{T(\mathscr{B}_3)}. \tag{1.10.12}$$

Through successive approximations we are able to build a sequence (x_n) such that $x_n \in \mathscr{B}_n$ and

$$u - T(\sum_{k=1}^{n} x_k) \in V_{n+1} \subseteq \overline{T(\mathscr{B}_{n+1})}. \tag{1.10.13}$$

Let $z_n = \sum_{k=1}^{n} x_k$. Notice that

$$\|z_n - z_m\| \leq \sum_{k=m+1}^{n} \|x_k\|, \tag{1.10.14a}$$
$$< \sum_{k=m+1}^{n} \frac{1}{2^k}, \tag{1.10.14b}$$

which proves (z_n) is Cauchy. Hence, (z_n) converges, since U is Banach. Let's say $z_n \to z$. It holds that $z \in \mathscr{B}_0$, for

$$\|z\| \leq \sum_{k=1}^{+\infty} \|x_k\|, \tag{1.10.15a}$$
$$< \sum_{k=1}^{+\infty} \frac{1}{2^k}, \tag{1.10.15b}$$
$$= 1. \tag{1.10.15c}$$

Continuity of T ensures that $Tz_n \to Tz$, and it is straightforward to check that $Tz = u$. Hence, $u \in T(\mathscr{B}_0)$. This proves $V_1 \subseteq T(\mathscr{B}_0)$ and concludes the proof. ■

Theorem 1.96 [Open Mapping]**:**
Let U and V be Banach spaces. If $T \in \mathscr{B}(U, V)$ and $\operatorname{Ran} T = V$*, then T is open.* □

Proof:
We want to prove that, for every open set $O \subseteq U$, $T(O)$ is too open. To do so, we shall prove that given $v = Tu \in T(O)$ there is an open ball about v contained in $T(O)$.

Let $v = Tu \in T(O)$. Since O is open, it contains some open ball centered about u. Therefore, $O - u$ contains some open ball centered about 0. Let r be its radius and define $k = \frac{1}{r}$. Then $k(O - u)$ contains the open unit ball. Lemma 1.95 on page 46 thus implies that $T(k(O - u)) = k[T(O) - v]$ contains an open ball about the origin. It follows that $T(O) - v$ contains an open ball about the origin, and hence $T(O)$ contains an open ball about v, as desired to show. Thus, since v was arbitrary, $T(O)$ is open. ■

Corollary 1.97 [Inverse Mapping Theorem]**:**
Let U and V be Banach spaces. Suppose $T \in \mathscr{B}(U, V)$ is a bijection. Then $T^{-1} \in \mathscr{B}(V, U)$.
□

Proof:
Proposition 1.48 on page 25 ensures that the inverse T^{-1} is a linear operator. The Open Mapping Theorem ensures T is open, and hence O being open implies $T(O)$ being open. Notice now that the preimage of O under T^{-1} is $T(O)$. Thus, T being open implies T^{-1} is continuous, and therefore bounded. ■

1.11 Closed Graph Theorem

The third big theorem of Banach space theory is the Closed Graph Theorem. It states an equivalence between an operator being bounded and its graph being closed. Furthermore, it motivates the definition of a closed operator, which intuitively is an operator that not necessarily is continuous, but behaves relatively well under limiting procedures. Namely, suppose one has a sequence $u_n \to u$ and an operator T with $Tu_n \to v$. Closed operators are such that $v = Tu$ will hold whenever there is such a v (*i.e.*, whenever Tu_n converges), hence avoiding the weird case in which v exists, but $v \neq Tu$. While unbounded operators are also of interest in practical applications of functional analysis, they are typically closed.

To start this section, let us firstly notice that a product of normed spaces U and V, $U \times V$, can be naturally given a vector space structure. This is obtained by defining

$$(u_1, v_1) + \lambda(u_2, v_2) = (u_1 + \lambda u_2, v_1 + \lambda v_2). \tag{1.11.1}$$

This product can also be endowed with the norm given by

$$\|(u, v)\| = \|u\|_U + \|v\|_V, \tag{1.11.2}$$

and hence $U \times V$ can be considered a normed space.

Definition 1.98 [Graph]**:**
Let $T: U \to V$ be an operator between vector spaces. The *graph* of T is the vector subspace

$$\Gamma(T) = \left\{(u, Tu); u \in \operatorname{Dom} T\right\} \tag{1.11.3}$$

of $U \times V$. ♠

Definition 1.99 [Closed Operator]**:**
Let $T: U \to V$ be a linear operator between normed spaces. We say T is *closed* if, and only if, for every convergent sequence (u_n) in $\operatorname{Dom} T$ ($u_n \to u \in U$) with (Tu_n) also convergent ($Tu_n \to v$), it holds that $u \in \operatorname{Dom} T$ and $Tu = v$. Equivalently, T is closed if, and only if, $\Gamma(T)$ is a closed vector subspace of $U \times V$. ♠

Proposition 1.100:
Let U and V be Banach spaces. An operator $T \in \mathscr{B}(U, V)$ is closed if, and only if, $\operatorname{Dom} T$ *is closed.* □

Proof:
If $\operatorname{Dom} T$ is closed, let $u_n \to u$ in $\operatorname{Dom} T$ with $Tu_n \to v$. Since $u \in \operatorname{Dom} T$, continuity of T ensures $Tu = v$.

If T is closed, then a sequence (u_n) in $\operatorname{Dom} T$ that converges to $u \in U$ is ensured to have $u \in \operatorname{Dom} T$, and hence $\operatorname{Dom} T$ is closed. ■

Notice then that boundedness does not imply closedness. Similarly, closedness does not imply boundedness, as illustrated by the following example.

Example 1.101:

Let $U = \mathscr{C}^0([0,1])$ and $V = \mathscr{C}^1([0,1])$. Define the operator $T\colon U \to U$ with $\operatorname{Dom} T = V$ through $Tf = f'$. Then T is closed, but it is not bounded.

Indeed, Example 1.53 on page 26 proves that T is unbounded.

Let us now prove that T is closed. Let (f_n) be a sequence in V converging to f. Let $Tf_n \to g$. Convergence in U is uniform convergence, and hence

$$\int_0^t g(s)\,\mathrm{d}s = \int_0^t \lim_{n\to+\infty} f_n'(s)\,\mathrm{d}s, \tag{1.11.4a}$$
$$= \lim_{n\to+\infty} \int_0^t f_n'(s)\,\mathrm{d}s, \tag{1.11.4b}$$
$$= \lim_{n\to+\infty} (f_n(t) - f_n(0)), \tag{1.11.4c}$$
$$= f(t) - f(0), \tag{1.11.4d}$$

and hence

$$f(t) = f(0) + \int_0^t g(s)\,\mathrm{d}s, \tag{1.11.5}$$

which implies $f'(t) = g(t)$. Thus, T is closed. ♥

It is easier to check that an operator is closed than it is to check that it is bounded, and hence it is useful to have a criterion ensuring that closed operators are bounded. This is where the Closed Graph Theorem comes in. To prove it, we shall first prove an auxiliary lemma.

Lemma 1.102:

Consider a linear operator $T\colon U \to V$ between normed spaces. Let $\pi_1\colon \Gamma(T) \to \operatorname{Dom} T$ and $\pi_2\colon \Gamma(T) \to \operatorname{Ran} T$ be the natural projections given by $\pi_1((u, Tu)) = u$ and $\pi_2((u, Tu)) = Tu$. π_1 and π_2 are linear and continuous. □

Proof:

It is straightforward to check that the projections are linear. To show that they are continuous we shall prove they are bounded. Firstly, we notice that

$$\|\pi_1((u, Tu))\|_U = \|u\|_U, \tag{1.11.6a}$$
$$\leq \|u\|_U + \|Tu\|_V, \tag{1.11.6b}$$

and hence π_1 is bounded. An analogous proof applies to π_2. ■

Theorem 1.103 [Closed Graph Theorem]**:**

Let U and V be Banach spaces. Consider an operator $T\colon U \to V$ with closed domain $\operatorname{Dom} T$. *T is bounded if, and only if, it is closed.* □

Proof:

One of the implications has already been proven in Proposition 1.100 on the preceding page. We shall then focus on proving the other implication.

Suppose T is closed. We know from Lemma 1.102 on the preceding page that the canonical projections $\pi_1 : \Gamma(T) \to \operatorname{Dom} T$ and $\pi_2 : \Gamma(T) \to \operatorname{Ran} T$ are linear and continuous. Furthermore, it is straightforward to notice that π_1 establishes a bijection between $\Gamma(T)$ and $\operatorname{Dom} T$. $\Gamma(T)$ is closed in $U \times V$, characterizing it as a Banach space. The Open Mapping Theorem, through Corollary 1.97 on page 48, then ensures π_1^{-1} is continuous. Noticing then that

$$T = \pi_2 \circ \pi_1^{-1} \tag{1.11.7}$$

it follows that T is bounded. ■

Remark:

One might think that a given operator failed to be closed because it was considered in a small domain, but that taking the closure of the graph, $\overline{\Gamma(T)}$, one would get a closed operator. This, however, may not be the case, for $\overline{\Gamma(T)}$ may fail to be the graph of an operator. If $\overline{\Gamma(T)}$ is the graph of an operator, T is said to be closable and the operator with graph $\overline{\Gamma(T)}$ is said to be its closure. ♣

1.12 Zorn's Lemma

The last of the four big theorems of Banach space theory is the Hahn–Banach Theorem. This theorem is an extension theorem for linear functionals—it states conditions under which one can extend a functional defined on a subspace to the whole space. Later we will be able to apply this result when studying adjoint operators and weak convergences , but in addition to these sorts of applications the result is also important because it ensures that normed spaces have a rich theory of duality, *i.e.*, it ensures the dual of a normed space has a rich structure.

Crossref

However, in order to prove the Hahn–Banach Theorem, we will need Zorn's Lemma. Zorn's Lemma is a result from set theory, rather than functional analysis, which we shall recapitulate now. We won't give the details of the proof, but rather give references to it.

We begin by defining a partially ordered set.

Definition 1.104 [Partial Ordering]**:**

Let X be a set. A *partial ordering* in X is a binary relation $\preccurlyeq$ in X such that

i. $x \preccurlyeq x$ for all $x \in X$ (reflexivity);

ii. $x \preccurlyeq y$ and $y \preccurlyeq z$ implies $x \preccurlyeq z$, for all $x, y, z \in X$ (transitivity);

iii. $x \preccurlyeq y$ and $y \preccurlyeq x$ imply $x = y$, for all $x, y \in X$ (antisymmetry).

If $\preccurlyeq$ is a partial ordering in X, the pair $(X, \preccurlyeq)$ is said to be a *partially ordered set*, or simply a *poset*.

The term *partial* ordering comes from the fact that certain pairs of elements might not be comparable, *i.e.*, there can be $x, y \in X$ such that neither $x \preccurlyeq y$ nor $y \preccurlyeq x$ hold. If

$x \preccurlyeq y$ or $y \preccurlyeq x$ hold, x and y are said to be *comparable*. Otherwise, they are said to be *incomparable*. ♠

Definition 1.105 [Total Ordering]:

A *total* ordering is a partial ordering such that every two elements are comparable. A set endowed with a total ordering is said to be a *totally ordered set* or a *chain*. ♠

Definition 1.106 [Upper Bounds and Maximal Elements]:

Let $(X, \preccurlyeq)$ be a poset. We say $x \in X$ is a *maximal element* if for every $y \in X$ with $x \preccurlyeq y$ it follows that $x = y$. $z \in X$ is said to be an *upper bound* for $Y \subseteq X$ if it holds that $y \preccurlyeq z$ for all $y \in Y$. ♠

Depending on Y and X, maximal elements and upper bounds may or not exist. Furthermore, a maximal element is not necessarily an upper bound.

Example 1.107 [$\mathbb{R}$]:

Consider $\mathbb{R}$ with the usual ordering $\leq$. $(\mathbb{R}, \leq)$ is a totally ordered set, but it has no maximal element. ♥

Example 1.108 [$\mathbb{P}(X)$]:

Let X be a set and consider its power set $\mathbb{P}(X)$. $(\mathbb{P}(X), \subseteq)$ is a partially ordered set. It has a single maximal element: X. ♥

Lemma 1.109 [Zorn]:

Let X be a non-empty poset. If all totally ordered subsets of X have an upper bound, X has a maximal element. □

Zorn's Lemma is actually equivalent to the Axiom of Choice with Zermelo–Frankel set theory. Details of its proof and of this equivalence can be found in books about set theory, such as the ones by Ciesielski (1997, Theorem 4.3.4), Schindler (2014, Theorem 2.10), and Suppes (1972, Sec. 8.2).

As an application of Zorn's Lemma, let us prove that every non-trivial vector space—*i.e.*, every vector space that contains at least one non-vanishing element—admits a Hamel basis.

Theorem 1.110:

Every non-trivial vector space admits a Hamel basis. □

Proof:

Let $U \neq \{0\}$ be a vector space. Let $X \subseteq \mathbb{P}(U)$ be the set of all linearly independent subsets of U. It is straightforward to see that $X \neq \emptyset$ and that $(X, \subseteq)$ is a poset.

If X has a maximal element M, it is a Hamel basis for U. Indeed, we know that $\operatorname{span}(M) \subseteq U$. If $\operatorname{span}(M) \neq U$, then there is some $u \in U$ such that $M \cup \{u\}$ is linearly independent, but this contradicts the assumption that M is a maximal element. Hence, we must conclude that if X has a maximal element, it is a Hamel basis for U.

We then need only to prove that such a maximal element exists. This is where Zorn's Lemma comes into play: it suffices to prove that all totally ordered subsets of X have an upper bound. This, however, is straightforward: an upper bound is given by the union of all elements of the totally ordered subset. This then concludes the proof. ■

1.13 Hahn–Banach Theorem

With Zorn's Lemma in hand, we are ready to start discussing the concepts that underlie the Hahn–Banach Theorem. As mentioned before, this is an extension theorem for linear functionals, and it essentially establishes that we have a rich theory of duality in normed spaces. It, however, requires that the functionals to be extended satisfy a certain boundedness property stated in terms of a sublinear functional. We shall then first define what this means.

Definition 1.111 [Sublinear Functional]**:**

Let U be a vector space. We say a function $p\colon U \to \mathbb{R}$ is a *sublinear functional* if, for every $u, v \in U$ and for every $\lambda \geq 0$, it satisfies

i. $p(u + v) \leq p(u) + p(v)$;

ii. $p(\lambda u) = \lambda p(u)$. ♠

Example 1.112 [Norms and Seminorms]**:**

Notice that norms and seminorms in a vector space are examples of sublinear functionals. ♥

Notice that a sublinear functional need not be positive, although norms and seminorms are.

We shall prove two versions of the Hahn–Banach Theorem: one for real vector spaces and another that applies to both real and complex vector spaces. We start with the simpler one.

Theorem 1.113 [Real Hahn–Banach]**:**

Let U be a real vector space and let $p\colon U \to \mathbb{R}$ be a sublinear functional. Let $V \subseteq U$ be a vector subspace and let $f\colon V \to \mathbb{R}$ be a linear functional dominated by p, i.e., *a linear functional such that $f(v) \leq p(v), \forall\, v \in V$. Then f has a linear extension $\bar{f}\colon U \to \mathbb{R}$ that is too dominated by p,* i.e., *a linear extension $\bar{f}\colon U \to \mathbb{R}$ such that $\bar{f}(u) \leq p(u), \forall\, u \in U$. $\bar{f}$ is said to be the Hahn–Banach extension of f.* □

Proof:

Let $G = \{g_\lambda\}_{\lambda\in\Lambda}$ be the family of linear extensions $g_\lambda\colon V_\lambda \to \mathbb{R}$ of f (with $V \subseteq V_\lambda \subseteq U$) that satisfy $g_\lambda(v) \leq p(v), \forall\, v \in V_\lambda$. $f \in G$, so $G \neq \emptyset$. Furthermore, G admits a poset structure: simply define $g_\lambda \preccurlyeq g_\mu$ whenever g_μ is a linear extension of g_λ.

Take a chain $C = \{g_\lambda\}_{\lambda\in M} \subseteq G$. Define g_C through

$$g_C(v) = g_\lambda(v), \text{ if } v \in V_\lambda, \forall\, \lambda \in M. \tag{1.13.1}$$

This defines a linear functional g_C with $\operatorname{Dom} g_C = \bigcup_{\lambda\in M} V_\lambda$ that extends all g_λ in C and satisfies $g_C(v) \leq p(v), \forall\, v \in \operatorname{Dom} g_C$. Hence, g_C is an upper bound for C. Since this construction holds for any chain $C \subseteq G$, it follows from Zorn's Lemma that G has a maximal element. Let $\bar{f}$ be this maximal element. Since $\bar{f} \in G$, it holds that $\bar{f}$ is dominated by p. We want now to prove that $\operatorname{Dom} \bar{f} = U$.

Suppose, by contradiction, that there is some $u \in U \setminus \operatorname{Dom} \bar{f}$. We will prove that, in this case, there is a linear extension F of $\bar{f}$ to $W = \operatorname{span}\big(\operatorname{Dom} \bar{f} \cup \{u\}\big)$ that satisfies $F(w) \leq p(w), \forall\, w \in W$, hence contradicting the fact that $\bar{f}$ is maximal.

Let F be an arbitrary extension of $\bar{f}$ to W—this is guaranteed to exist because it suffices to attribute an arbitrary value to $F(u)$. Notice then that each $w \in W$ can be written in the form $w = v + \lambda u$ for some $\lambda \in \mathbb{R}$ and some $v \in \operatorname{Dom} \bar{f}$. Hence, for any $w \in W$ we have that

$$F(w) = F(v + \lambda u), \tag{1.13.2a}$$
$$= F(v) + \lambda F(u), \tag{1.13.2b}$$
$$= \bar{f}(v) + \lambda F(u). \tag{1.13.2c}$$

Therefore, what we need to prove is that it is possible to choose $F(u)$ in a manner that leads to $F(w) \leq p(w), \forall\, w \in W$.

Given any $v_1, v_2 \in \operatorname{Dom} \bar{f}$ one has

$$\bar{f}(v_1) + \bar{f}(v_2) = \bar{f}(v_1 + v_2), \tag{1.13.3a}$$
$$\leq p(v_1 + v_2), \tag{1.13.3b}$$
$$\leq p(v_1 - u) + p(v_2 + u), \tag{1.13.3c}$$

and hence

$$\bar{f}(v_1) - p(v_1 - u) \leq p(v_2 + u) - \bar{f}(v_2). \tag{1.13.4}$$

Hence, there is $\alpha \in \mathbb{R}$ such that

$$\sup_{v_1\in\operatorname{Dom}\bar{f}} \big[\bar{f}(v_1) - p(v_1 - u)\big] \leq \alpha \leq \inf_{v_2\in\operatorname{Dom}\bar{f}} \big[p(v_2 + u) - \bar{f}(v_2)\big]. \tag{1.13.5}$$

With this construction in mind, pick $F(u) = \alpha$. Then it holds that $F(w) \leq p(w), \forall\, w \in W$. To see this, write $w = v + \lambda u$ once again. If $\lambda = 0$, the inequality follows immediately. If $\lambda > 0$, one has

$$F(w) = \bar{f}(v) + \lambda F(u), \tag{1.13.6a}$$
$$= \bar{f}(v) + \lambda\alpha, \tag{1.13.6b}$$
$$\leq \bar{f}(v) + \lambda\Big[p\Big(\frac{v}{\lambda} + u\Big) - \bar{f}\Big(\frac{v}{\lambda}\Big)\Big], \tag{1.13.6c}$$

$$
\begin{aligned}
&= p(v + \lambda u), && (1.13.6d)\\
&= p(w). && (1.13.6e)
\end{aligned}
$$

If $\lambda < 0$, then

$$
\begin{aligned}
F(w) &= F(v + \lambda u), && (1.13.7a)\\
&= F(v - |\lambda| u), && (1.13.7b)\\
&= \bar{f}(v) - |\lambda|\alpha, && (1.13.7c)\\
&\leq \bar{f}(v) - |\lambda|\left[\bar{f}\left(\frac{v}{|\lambda|}\right) - p\left(\frac{v}{|\lambda|} - u\right)\right], && (1.13.7d)\\
&\leq |\lambda| p\left(\frac{v}{|\lambda|} - u\right), && (1.13.7e)\\
&= p(v - |\lambda| u), && (1.13.7f)\\
&\leq p(v + \lambda u). && (1.13.7g)
\end{aligned}
$$

Hence, indeed we have $F(w) \leq p(w), \forall\, w \in W$. Since F extends $\bar{f}$ linearly, this contradicts the maximality of $\bar{f}$ and we are forced to conclude that $\operatorname{Dom} \bar{f} = U$, concluding the proof. ■

This proof fails for complex vector spaces, since it explicitly employs the ordering available in $\mathbb{R}$, which is absent in $\mathbb{C}$. Hence, in order to obtain a similar result for complex spaces we will need to make a few changes.

Definition 1.114 [Real Linear Functional]:
Let U be a complex vector space. A *real linear functional* in U is a functional $f: U \to \mathbb{R}$ such that $f(u + \lambda v) = f(u) + \lambda f(v)$ for all $u, v \in U$ and all $\lambda \in \mathbb{R}$. ♠

Lemma 1.115:
Let $z \in \mathbb{C}$. Then $z = \operatorname{Re}(z) - i \operatorname{Re}(iz)$. □

Proof:
Straightforward. ■

The previous lemma serves as motivation for the following one, which will be useful in proving the complex generalization of the Hahn–Banach Theorem.

Lemma 1.116:
Let U be a complex vector space.

i. *If $h: U \to \mathbb{R}$ is a real linear functional, then $f: U \to \mathbb{C}$ defined through*

$$f(u) = h(u) - ih(iu), u \in U, \tag{1.13.8}$$

is a complex linear functional.

ii. *If $f\colon U \to \mathbb{C}$ is a complex linear functional, then there is a real linear functional h satisfying Eq. (1.13.8) on the previous page.*

In both cases, one has $h = \operatorname{Re} f$. □

Proof:

Firstly, let $h\colon U \to \mathbb{R}$ be a real linear functional. We want to prove that Eq. (1.13.8) on the preceding page defines a complex linear functional. To do so, we want to prove that for any $\alpha \in \mathbb{C}$ and any $u, v \in U$ it holds that

$$f(u + \alpha v) = f(u) + \alpha f(v). \tag{1.13.9}$$

To do so, let $\alpha = \lambda + i\mu$, with $\lambda, \mu \in \mathbb{R}$. Then

$$\begin{aligned}
f(u + \alpha v) &= h(u + \alpha v) - ih(iu + i\alpha v), && (1.13.10a)\\
&= h(u) + h(\alpha v) - ih(iu) - ih(i\alpha v), && (1.13.10b)\\
&= h(u) + h(\lambda v + i\mu v) - ih(iu) - ih(i\lambda v - \mu v), && (1.13.10c)\\
&= h(u) + \lambda h(v) + \mu h(iv) - ih(iu) - i\lambda h(iv) + i\mu h(v), && (1.13.10d)\\
&= h(u) - ih(iu) + (\lambda i\mu)h(v) - i(\lambda i\mu)h(iv), && (1.13.10e)\\
&= h(u) - ih(iu) + \alpha h(v) - i\alpha h(iv), && (1.13.10f)\\
&= f(u) + \alpha f(v), && (1.13.10g)
\end{aligned}$$

which proves f is a complex linear functional.

On the other hand, now let $f\colon U \to \mathbb{C}$ be a complex linear functional. We want to show there is a real linear functional h satisfying Eq. (1.13.8) on the previous page. Define $h(u) = \operatorname{Re} f(u), \forall\, u \in U$. Notice then that

$$\begin{aligned}
\operatorname{Re} f(iu) + i \operatorname{Im} f(iu) &= f(iu), && (1.13.11a)\\
&= if(u), && (1.13.11b)\\
&= i \operatorname{Re} f(u) - \operatorname{Im} f(u). && (1.13.11c)
\end{aligned}$$

Therefore, $\operatorname{Re} f(iu) = -\operatorname{Im} f(u)$. Using this, we notice that

$$\begin{aligned}
f(u) &= \operatorname{Re} f(u) + i \operatorname{Im} f(u), && (1.13.12a)\\
&= h(u) - i \operatorname{Re} f(iu), && (1.13.12b)\\
&= h(u) - ih(iu), && (1.13.12c)
\end{aligned}$$

which concludes the proof. ■

We are now ready to state, and prove, a generalized version of the Hahn–Banach theorem.

Theorem 1.117 [Complex Hahn–Banach]**:**

Let U be a $\mathbb{K}$vector space ($\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$) and let $p\colon X \to [0, +\infty)$ be a function such that

i. *$p(u+v) \le p(u) + p(v)$, for all $u, v \in U$;*

ii. *$p(\lambda u) = |\lambda| p(u)$, for all $u \in U$ and all $\lambda \in \mathbb{K}$.*

Let V be a $\mathbb{K}$-vector subspace of U and let $f\colon V \to \mathbb{K}$ be a linear functional with $|f(v)| \le p(v)$ for all $v \in V$—i.e., *f is dominated by p. Then f has a linear extension $\bar{f}\colon U \to \mathbb{K}$ that is too dominated by p,* i.e., *a linear extension $\bar{f}\colon U \to \mathbb{K}$ such that $\bar{f}(u) \le p(u), \forall\, u \in U$. $\bar{f}$ is said to be the Hahn–Banach extension of f.* □

Proof:

Let $h = \operatorname{Re} f$, which is a real linear functional in V by Lemma 1.116 on page 55. Notice that

$$h(v) \le |f(v)| \le p(v), \forall\, v \in V. \tag{1.13.13}$$

Hence, the real Hahn–Banach Theorem implies there is a real linear extension $\bar{h}\colon U \to \mathbb{R}$ with $\bar{h}(u) \le p(u)$ for all $u \in U$.

If f is real, the proof is over. Suppose then that f is complex, in which case we know from Lemma 1.116 on page 55 that $f(v) = h(v) - ih(iv)$. Define then $\bar{f}(u) = \bar{h}(u) - i\bar{h}(iu)$, which extends f and, due to Lemma 1.116 on page 55, is a complex linear functional. We want to prove $|\bar{f}(u)| \le p(u)$ for all $u \in U$.

If $\bar{f}(u) = 0$, the result is immediate, for $p(u) \ge 0$. Suppose then $\bar{f}(u) \neq 0$. Then there is $\theta \in [0, 2\pi)$ such that $\bar{f}(u) = e^{i\theta}|\bar{f}(u)|$. Since $\bar{f}$ is linear, it follows that

$$\begin{aligned}
|\bar{f}(u)| &= \bar{f}(e^{-i\theta}u), && (1.13.14a)\\
&= \operatorname{Re} \bar{f}(e^{-i\theta}u), && (1.13.14b)\\
&= \bar{h}(e^{-i\theta}u), && (1.13.14c)\\
&\le p(e^{-i\theta}u), && (1.13.14d)\\
&= p(u), && (1.13.14e)
\end{aligned}$$

which concludes the proof. ■

To exhibit the usefulness of the Hahn–Banach Theorem, let us prove a few corollaries.

Corollary 1.118:

Let U be a normed vector space and let $V \subseteq U$ be a vector subspace. Let $f \in V^*$. Then there is $\bar{f} \in U^*$ with $\|\bar{f}\|_{U^*} = \|f\|_{V^*}$. □

Proof:

Pick $p(u) = \|f\|_{V^*}\|u\|_U$. Then the functional f is dominated by p. The complex Hahn–Banach Theorem then ensures the existence of a linear functional $\bar{f}\colon U \to \mathbb{K}$ with $|\bar{f}(u)| \le \|f\|_{V^*}\|u\|_U$. This not only proves $\bar{f}$ is bounded, and hence $\bar{f} \in U^*$, but also proves $\|\bar{f}\|_{U^*} \le \|f\|_{V^*}$. However, notice that

$$\|f\|_{V^*} = \sup_{\substack{v \in V\\ \|v\|=1}} |f(v)| = \sup_{\substack{v \in V\\ \|v\|=1}} |\bar{f}(v)| \le \sup_{\substack{u \in U\\ \|u\|=1}} |\bar{f}(u)| = \|\bar{f}\|_{U^*}. \tag{1.13.15}$$

Hence, $\|\bar{f}\|_{U^*} = \|f\|_{V^*}$. ■

Corollary 1.118 on the previous page is sometimes referred to as being the Hahn–Banach Theorem.

Corollary 1.119:

Let $u \in U$, with U being a normed space and $u \neq 0$. Then there is $f \in U^*$ with $f(u) = \|f\|_{U^*}\|u\|$. In particular, f can be chosen such that $\|f\|_{U^*} = 1$. □

Proof:

Let $V = \operatorname{span}(\{u\})$. Define $g \in V^*$ by imposing $g(u) = \|u\|$. Notice this implies $\|g\|_{V^*} = 1$. Corollary 1.118 on the preceding page then ensures the existence of $f \in U^*$ with $f(u) = \|u\| = \|f\|_{U^*}\|u\|$. This concludes the proof. ■

Corollary 1.120:

Let U be a normed space and $u, v \in U$, $u \neq v$. Then there is $f \in U^*$ such that $f(u) \neq f(v)$ (separates points). □

Proof:

Let $V = \operatorname{span}(\{u, v\})$. If V is two-dimensional, define $g: V \to \mathbb{K}$ by specifying $g(u) \neq g(v)$ and imposing linearity. If V is one-dimensional, simply specify $g(u)$ and impose linearity and it will hold that $g(u) \neq g(v)$. Since V is finite-dimensional, g is bounded and hence $g \in V^*$. Using Corollary 1.118 on the previous page we can extend g to U and thus find a functional $f \in U^*$ with $f(u) = g(u) \neq g(v) = f(v)$. ■

Definition 1.121 [Family of Functionals that Separates Points]**:**

A family of functionals G is said to *separate points* of a set X if, for each distinct $x, y \in X$, there is $f \in G$ with $f(x) \neq f(y)$. ♠

Notice that Corollary 1.120 states that, if U is a normed space, U^* separates points of U.

Corollary 1.122:

Let U be a normed space. If $u \in U$ is such that $f(u) = 0$ for all $f \in U^*$, then $u = 0$. □

Proof:

Suppose, by contradiction, that $u \neq 0$. Then we can construct a functional $f \in U^*$ such that $f(u) \neq 0$. Indeed, let $V = \operatorname{span}(\{u\})$. Define $g: V \to \mathbb{K}$ by assigning a value $g(u) \neq 0$ and imposing linearity. Since V is finite-dimensional, g is bounded and therefore $g \in V^*$. Hence, Corollary 1.118 on the previous page implies the existence of $f \in U^*$ with $f(u) = g(u) \neq 0$. This contradicts the hypothesis that $h(u) = 0$ for all $h \in U^*$, and hence we are forced to conclude that $u = 0$. ■

Corollary 1.123:

Let U be a normed space. If $u \in U$, then

$$\|u\| = \sup_{0 \neq f \in U^*} \frac{|f(u)|}{\|f\|} = \max_{0 \neq f \in U^*} \frac{|f(u)|}{\|f\|}. \tag{1.13.16}$$

□

Proof:

If $f \in U^*$, then

$$|f(u)| \leq \|f\| \|u\|, \tag{1.13.17}$$

from which it follows that

$$\frac{|f(u)|}{\|f\|} \leq \|u\|, \tag{1.13.18a}$$

$$\sup_{0 \neq f \in U^*} \frac{|f(u)|}{\|f\|} \leq \|u\|. \tag{1.13.18b}$$

However, Corollary 1.119 on the facing page ensures the existence of $g \in U^*$ with $g(u) = \|u\|$ and $\|g\| = 1$. Hence,

$$\|u\| = \frac{|g(u)|}{\|g\|}, \tag{1.13.19a}$$

$$\leq \sup_{0 \neq f \in U^*} \frac{|f(u)|}{\|f\|}. \tag{1.13.19b}$$

Therefore, we conclude that

$$\|u\| = \sup_{0 \neq f \in U^*} \frac{|f(u)|}{\|f\|} \tag{1.13.20}$$

and the existence of g allows us to swap the “sup” for a “max”. ■

Corollary 1.124:

Let U be a $\mathbb{K}$-normed space and let $\{u_1, \dots, u_n\} \subseteq U$ be a linearly independent set and $\{a_1, \dots, a_n\} \subseteq \mathbb{K}$. There is $f \in U^*$ such that $f(u_i) = a_i$ for each $1 \leq i \leq n$. □

Proof:

Let $V = \operatorname{span}(\{u_1, \dots, u_n\})$. Define $g \colon V \to \mathbb{K}$ by imposing $g(u_i) = a_i$ for each $1 \leq i \leq n$. Since V is finite-dimensional, g is bounded, and hence $g \in V^*$. Hence, Corollary 1.118 on page 57 implies the existence of $f \in U^*$ with the desired properties. ■

Proposition 1.125:

Let U be a normed space, $V \subset U$ be a closed proper subspace of U, and $u \in U \setminus V$. If

$$\delta = d(u, V) = \inf_{v \in V} \|u - v\|, \tag{1.13.21}$$

then there is $f \in U^*$ satisfying $\|f\| = 1$, $f(u) = \delta$, and $f|_V = 0$. □

Proof:

Since V is closed, $\delta > 0$. Let $W = \operatorname{span}(\{V, u\})$. Define $g\colon W \to \mathbb{K}$ by imposing $g(u) = \delta$, $g|_V = 0$, and linearity. Given $w \in W$, we can write $w = \lambda u + v$ for $\lambda \in \mathbb{K}$ and $v \in V$. For such a w we have

$$\begin{aligned} g(w) &= g(\lambda u + v), && (1.13.22a)\\ &= \lambda\delta. && (1.13.22b) \end{aligned}$$

Notice that

$$\begin{aligned} \left|g(\lambda u + v)\right| &= |\lambda|\delta, && (1.13.23a)\\ &\leq |\lambda| \left\| u + \frac{v}{\lambda} \right\|, && (1.13.23b)\\ &= \|\lambda u + v\|, && (1.13.23c) \end{aligned}$$

which implies $\|g\| \leq 1$.

For each $v \in V$, one has

$$\begin{aligned} \|g\| &\geq \frac{\left|g(u - v)\right|}{\|u - v\|}, && (1.13.24a)\\ &= \frac{\delta}{\|u - v\|}. && (1.13.24b) \end{aligned}$$

Therefore,

$$\begin{aligned} \|g\| &\geq \sup_{v \in V} \frac{\delta}{\|u - v\|}, && (1.13.25a)\\ &= \frac{\delta}{\inf_{v \in V} \|u - v\|}, && (1.13.25b)\\ &= \frac{\delta}{\delta}, && (1.13.25c)\\ &= 1. && (1.13.25d) \end{aligned}$$

Hence, $\|g\| = 1$.

Finally, we can apply Corollary 1.118 on page 57 to obtain $f \in U^*$ satisfying the desired conditions. ■

Corollary 1.126:

Let U be a normed space. A vector subspace V of U is dense in U if, and only if, the only element of U^* that vanishes in V is the null functional. □

Proof:

Suppose V is not dense in U. Then $\overline{V} \subset U$ and we can use Proposition 1.125 on the previous page to prove the existence of a functional that vanishes on $\overline{V}$, and hence on V, but is not the null functional. This prove the contrapositive of the statement "if the only element of U^* that vanishes in V is the null functional, then V is dense in U".

Suppose now that V is dense in U. Let $f \in U^*$ with $f|_V = 0$. We want to prove that $f(u) = 0, \forall\, u \in U$. Since V is dense in U, given $u \in U$ there is a sequence (v_n) in V with $v_n \to u$. Since f is bounded, it is continuous, and hence we have that

$$f(u) = f(\lim_{n\to+\infty} v_n), \tag{1.13.26a}$$
$$= \lim_{n\to+\infty} f(v_n), \tag{1.13.26b}$$
$$= \lim_{n\to+\infty} 0, \tag{1.13.26c}$$
$$= 0. \tag{1.13.26d}$$

Hence, f is the null functional. ■

Proposition 1.127:

Let U be a normed space. If U^* is separable, then so is U. □

Proof:

Assume U^* is separable. Then there is a dense sequence (f_n) in U^*. Pick $u_n \in U$ with $\|u_n\| = 1$ and $|f_n(u_n)| \geq \frac{\|f_n\|}{2}$ for each n and let $V = \overline{\operatorname{span}\left(\{u_n\}_{n=1}^{+\infty}\right)}$, which is separable (Proposition 1.35 on page 21). Our goal is then to prove that $V = U$ and to do that we shall prove that the only functional that vanishes in V is the null functional. The result then follows from Corollary 1.126 on the preceding page.

Suppose f vanishes in V. Let $f_{n_i} \to f$. For each n_i

$$\left\|f - f_{n_i}\right\| \geq \left|(f - f_{n_i})(u_{n_i})\right|, \tag{1.13.27a}$$
$$= \left|f_{n_i}(u_{n_i})\right|, \tag{1.13.27b}$$
$$\geq \frac{\left\|f_{n_i}\right\|}{2}. \tag{1.13.27c}$$

Hence,

$$\|f\| \leq \left\|f - f_{n_i}\right\| + \left\|f_{n_i}\right\|, \tag{1.13.28a}$$
$$\leq 3\left\|f - f_{n_i}\right\|, \tag{1.13.28b}$$
$$\to 0, i \to +\infty, \tag{1.13.28c}$$

which shows $f = 0$. Hence, U is separable. ■

Not every separable space has a separable dual, though. For example, Proposition 1.68 on page 32 establishes that $\ell^1(\mathbb{N})^* = \ell^\infty(\mathbb{N})$, but $\ell^1(\mathbb{N})$ is separable and $\ell^\infty(\mathbb{N})$ isn't.

As another application of the Hahn–Banach Theorem we shall prove the converse of Theorem 1.61 on page 29.

Proposition 1.128:

Suppose $U \neq \{0\}$ is a normed space and let V be also a normed space. If $\mathscr{B}(U,V)$ is a Banach space, then so is V. □

Proof:

Let (v_n) be a Cauchy sequence in V. Pick $u_0 \in U$ with $\|u_0\| = 1$. Then Corollary 1.119 on page 58 ensures the existence of $f \in U^*$ with $\|f\| = f(u_0) = 1$. Since $v_n = f(u_0)v_n$, we are motivated to define $T_n \in \mathscr{B}(U,V)$ through $T_n u = f(u_0)v_n$. (T_n) is a Cauchy sequence, for

$$\begin{aligned}\|(T_n - T_m)u\| &= |f(u)|\|v_n - v_m\|, &(1.13.29a)\\ &\leq \|u\|\|v_n - v_m\|. &(1.13.29b)\end{aligned}$$

Therefore, there is $T \in \mathscr{B}(U,V)$ with $T_n \to T$. Notice then that

$$\begin{aligned}\|v_n - Tu_0\| &= \|T_n u_0 - Tu_0\|, &(1.13.30a)\\ &\leq \|T_n - T\|\|u_0\|, &(1.13.30b)\end{aligned}$$

and it follows that $v_n \to Tu_0$, proving V is complete. ■

1.14 Bidual

To further apply the Hahn–Banach Theorem, we shall discuss a bit about the so-called bidual of a normed space.

Definition 1.129 [Second Dual or Bidual]**:**

Let U be a normed space. Since U^* is a Banach space, it makes sense to define the *second dual*, or the *bidual*, of U as the space $U^{**} = (U^*)^*$. ♠

There is a natural identification of elements of a normed space U with the elements of its bidual. Simply take $u \in U$ and map it to $\hat{u} \in U^{**}$ given by

$$\hat{u}(f) \equiv f(u), \forall f \in U^*. \tag{1.14.1}$$

This is known as the *canonical mapping* from U to U^{**}.

Proposition 1.130:

The canonical mapping from a normed space to its bidual is a linear isometry. □

Proof:

Let $u, v \in U$ and $\lambda \in \mathbb{K}$. Notice that, given $f \in U^*$,

$$\begin{aligned}\widehat{(u + \lambda v)}(f) &= f(u + \lambda v), &(1.14.2a)\\ &= f(u) + \lambda f(v), &(1.14.2b)\\ &= \hat{u}(f) + \lambda\hat{v}(f), &(1.14.2c)\end{aligned}$$

which proves ˆ is linear. To show it is an isometry notice that

$$\|\hat{u}\| = \sup_{0\neq f\in U^*} \frac{|\hat{u}(f)|}{\|f\|}, \tag{1.14.3a}$$
$$= \sup_{0\neq f\in U^*} \frac{|f(u)|}{\|f\|}, \tag{1.14.3b}$$
$$= \|u\|, \tag{1.14.3c}$$

where we used Corollary 1.123 on page 59. ■

Definition 1.131 [Reflexive Space]**:**

We say a normed space U is *reflexive* if, and only if, the canonical mapping ˆ: $U \to U^{**}$ is surjective. In other words, a normed space is reflexive when it is isometric to its bidual, with the isomorphism being given by the canonical mapping between them. ♠

It is necessary that the isometry is given specifically by the canonical mapping, as there are non-reflexive Banach spaces which are isomorphic to their biduals (James 1951).

Proposition 1.132:

All reflexive normed spaces are Banach spaces. □

Proof:

Let U be a reflexive normed space. U is isometric to U^{**}, but U^{**} is the dual of U^*. Since the dual is always a Banach space, U^{**} is a Banach space, and hence so is U. ■

Proposition 1.133:

Any finite-dimensional normed space is reflexive. □

Proof:

It follows from the fact that (Proposition 1.71 on page 35)

$$\dim U = \dim U^* = \dim U^{**} \tag{1.14.4}$$

and from the fact that the canonical mapping is linear. By taking a basis of U, the canonical mapping yields a basis of U^{**} and it is possible to establish a bijection between both spaces. ■

Proposition 1.134:

All closed vector subspaces of a reflexive normed space are too reflexive. □

Proof:

Let U be a reflexive normed space and let $V \subseteq U$ be a closed proper subspace of U. If $f \in U^*$, then $f|_V \in V^*$. Corollary 1.118 on page 57 implies $V^* = \{f|_V; f \in U^*\}$. Hence, for every $h \in V^{**}$ it suffices to consider $h(f|_V)$ and our goal is to find $u_h \in V$ such that $h = \hat{u}_h$.

Define the linear functional $H\colon U^* \to \mathbb{K}$ such that $H(f) = h(f|_V)$. Notice that

$$\begin{aligned} |H(f)| &= |h(f|_V)|, && (1.14.5a)\\ &\le \|h\|\|f|_V\|, && (1.14.5b)\\ &\le \|h\|\|f\|, && (1.14.5c)\end{aligned}$$

and hence $H \in U^{**}$. Since U is reflexive, there is some $u_h \in U$ such that $\hat{u}_h = H$. By construction, it holds that

$$h(f|_V) = H(f) = \hat{u}_h(f) = f(u_h), \forall\, f \in U^*. \tag{1.14.6}$$

We want to prove that $u_h \in V$.

Suppose $u_h \notin V$. Then Proposition 1.125 on page 59 ensures the existence of $f \in U^*$ such that $f(u_h) \neq 0$ and $f|_V = 0$, which contradicts Eq. (1.14.6). Therefore, it is necessary that $u_h \in V$.

Finally, we see that $h(f|_V) = f(u_h) = f|_V(u_h)$ for each $f \in U^*$. Since $V^* = \{f|_V; f \in U^*\}$, it follows that $h(g) = g(u_h)$ for all $g \in V^*$. Thus, $h = \hat{u}_h$. Therefore, the canonical mapping is surjective and V is reflexive. ■

1.15 Adjoint Operators in Normed Spaces

We will now discuss the Banach space adjoint of an operator defined between two normed spaces*. These operators are related to the solutions of equations involving linear operators, as discussed by Kreyszig (1978, Sec. 8.5).

Firstly let U and V be normed spaces. Let $T\colon U \to V$ be some bounded operator. We want to define the Banach space adjoint $T^\times$ of T. To do so, let g be any bounded linear functional on V. Notice that, using T and g, we can define a bounded linear functional f in U. To do so, we define

$$f(u) = g(Tu), \forall\, u \in U. \tag{1.15.1}$$

It is straightforward to check linearity—it follows from linearity of g and T. As for boundedness, notice that

$$\begin{aligned} |f(u)| &= |g(Tu)|, && (1.15.2a)\\ &\le \|g\|\|Tu\|, && (1.15.2b)\\ &\le \|g\|\|T\|\|u\|. && (1.15.2c)\end{aligned}$$

Hence, f is bounded with $\|f\| \le \|g\|\|T\|$.

Notice that this construction can be performed for many different functionals g. Through this construction, we establish a map $g \mapsto f$. The Banach space adjoint of T is the operator $T^\times$ such that $T^\times g = f$.

*While we adopt the terminology "Banach space adjoint", as used by Reed and Simon (1980, p. 185), it is important to notice that the notion can be defined even in incomplete normed spaces.

Definition 1.135 [Banach Space Adjoint]**:**

Let U and V be normed spaces and $T \in \mathscr{B}(U,V)$. The *Banach space adjoint* of T is the operator $T^\times \colon V^* \to U^*$ defined by

$$(T^\times g)(u) = g(Tu), \forall\, g \in V^*, \forall\, u \in U. \tag{1.15.3}$$

♠

Proposition 1.136:

Let U and V be normed spaces. If $T \in \mathscr{B}(U,V)$, then $T^\times \in \mathscr{B}(V^*,U^*)$ and $\|T^\times\| = \|T\|$. □

Proof:

It is straightforward to check that $T^\times$ is linear and the discussion at the beginning of the section proves that $\|T^\times g\| \le \|T\|\|g\|$, which proves $T^\times \in \mathscr{B}(V^*,U^*)$ and $\|T^\times\| \le \|T\|$. If $T = 0$, then $T^\times = 0$ and the proof is done. Suppose then that $T \ne 0$.

Corollary 1.119 on page 58 implies that, given $0 \ne u_0 \in U$ with $Tu_0 \ne 0$ there is $f \in V^*$ such that $0 \ne f(Tu_0) = \|Tu_0\|$ and $\|f\| = 1$. Hence,

$$\begin{aligned}
\|Tu_0\| &= f(Tu_0), && (1.15.4a)\\
&= \left|(T^\times f)(u_0)\right|, && (1.15.4b)\\
&\le \|T^\times\|\|f\|\|u_0\|, && (1.15.4c)\\
&= \|T^\times\|\|u_0\|, && (1.15.4d)
\end{aligned}$$

so $\|T\| \le \|T^\times\|$. Hence, $\|T\| = \|T^\times\|$. ■

Example 1.137:

Given a basis for $\mathbb{K}^n$, an operator $T \in \mathscr{B}(\mathbb{K}^n)$ can be represented by a matrix. In this case, $T^\times$ is represented by the transpose of that matrix. ♥

Example 1.138:

Notice that, for any normed space U, the Banach space adjoint of the identity operator $\mathbb{1}_U \colon U \to U$ is the identity operator $\mathbb{1}_{U^*} \colon U^* \to U^*$. ♥

Proposition 1.139:

Let U and V be normed spaces. Let $T, S \in \mathscr{B}(U,V)$ and $\lambda \in \mathbb{K}$. Then $(T+\lambda S)^\times = T^\times + \lambda S^\times$. If W is a third normed space, $R \in \mathscr{B}(V,W)$ and the product RS is well-defined, then $(RS)^\times = S^\times R^\times$. If $T^{-1} \in \mathscr{B}(V,U)$, then $\left(T^{-1}\right)^\times = (T^\times)^{-1}$. □

Proof:

Firstly we notice that, for all $g \in V^*$ and for all $u \in U$ it holds that

$$\begin{aligned}
\left((T+\lambda S)^\times g\right)(u) &= g((T+\lambda S)u), && (1.15.5a)\\
&= g(Tu + \lambda Su), && (1.15.5b)\\
&= g(Tu) + \lambda g(Su), && (1.15.5c)\\
&= (T^\times g)(u) + \lambda (S^\times g)(u), && (1.15.5d)\\
&= ((T^\times + \lambda S^\times)g)(u). && (1.15.5e)
\end{aligned}$$

Next, we notice that for all $f \in W^*$ and all $u \in U$ it holds that

$$\big((RS)^\times f\big)(u) = f(RSu), \tag{1.15.6a}$$
$$= (R^\times f)(Su), \tag{1.15.6b}$$
$$= (S^\times R^\times f)(u), \tag{1.15.6c}$$

and hence $(RS)^\times = S^\times R^\times$.

Using this last result, we notice that

$$\mathbb{1}_{U^*} = (T^{-1}T)^\times = T^\times \big(T^{-1}\big)^\times, \tag{1.15.7}$$

and the uniqueness of the inverse implies that $\big(T^{-1}\big)^\times = (T^\times)^{-1}$. ■

Proposition 1.140:

Let U and V be normed spaces. For $T \in \mathscr{B}(U,V)$, $T^\times$ is injective if, and only if, $\operatorname{Ran} T$ *is dense in V.* □

Proof:

An operator is injective if, and only if, its kernel is the trivial vector space (this comes from Proposition 1.48 on page 25).

Suppose then that $\operatorname{Ker} T^\times = \{0\}$. Then

$$0 = (T^\times g)(u) = g(Tu) \tag{1.15.8}$$

for all $u \in U$ implies $g = 0$. Hence, the only functional that vanishes on $\operatorname{Ran} T$ is the null functional. Therefore, Corollary 1.126 on page 60 implies $\operatorname{Ran} T$ is dense in V.

Running the argument in reverse yields the other implication. ■

Example 1.141:

Let $S_r \colon \ell^1(\mathbb{N}) \to \ell^1(\mathbb{N})$ be the right shift operator, defined by

$$S_r(u_1, u_2, u_3, \dots) = (0, u_1, u_2, u_3, \dots). \tag{1.15.9}$$

Its Banach space adjoint $S_r^\times \colon \ell^\infty(\mathbb{N}) \to \ell^\infty(\mathbb{N})$ is the left shift operator

$$S_l(u_1, u_2, u_3, \dots) = (u_2, u_3, u_4, \dots). \tag{1.15.10}$$

Indeed, let $u = (u_n)_{n=1}^\infty \in \ell^1(\mathbb{N})$ and $\alpha = (\alpha_n)_{n=1}^\infty \in \ell^\infty(\mathbb{N})$. Then

$$(S_r^\times \alpha)(u) = \alpha(S_r u), \tag{1.15.11a}$$
$$= \sum_{n=1}^\infty \alpha_n (S_r u)_n, \tag{1.15.11b}$$
$$= \sum_{n=2}^\infty \alpha_n u_{n-1}, \tag{1.15.11c}$$

$$
\begin{aligned}
&= \sum_{n=1}^{\infty} \alpha_{n+1} u_n, &(1.15.11\text{d})\\
&= \sum_{n=1}^{\infty} (S_e\alpha)_n u_n, &(1.15.11\text{e})\\
&= (S_e\alpha)(u), &(1.15.11\text{f})
\end{aligned}
$$

as desired. ♥

Proposition 1.142:

Let U and V be Banach spaces. Suppose the linear operators $T: U \to V$ and $S: V^* \to U^*$ satisfy

$$g(Tu) = (Sg)(u), \forall\, u \in U, \forall\, g \in V^*. \tag{1.15.12}$$

Then it follows that both T and S are bounded and $S = T^\times$. □

Proof:

To show T is bounded we shall use the Closed Graph Theorem. Let $u_n \to u$ in U and $Tu_n \to v$ in V. Then, for each $g \in V^*$,

$$
\begin{aligned}
g(v) &= \lim_{n\to+\infty} g(Tu_n), &(1.15.13\text{a})\\
&= \lim_{n\to+\infty} (Sg)(u_n), &(1.15.13\text{b})\\
&= (Sg)(u), &(1.15.13\text{c})\\
&= g(Tu). &(1.15.13\text{d})
\end{aligned}
$$

Corollary 1.120 on page 58 then implies $v = Tu$, and hence the graph of T is closed. Thus, T is bounded. Since $T^\times g = Sg, \forall\, g \in V^*$, it follows that $T^\times = S$, which is thus bounded (Proposition 1.136 on page 65). ■

1.16 Weak Convergence

While so far we have always considered convergence in the sense of the norm, this is not the only possibility we can choose. Hence, we shall now introduce the notions of strong and weak convergences.

Definition 1.143 [Weak Convergence]**:**

Let U be a normed space. A sequence (u_n) in U is said to *converge weakly* to $u \in U$ if $\lim_{n\to+\infty} f(u_n) = f(u)$ for every $f \in U^*$. If (u_n) converges weakly to u we write $u_n \rightharpoonup u$, $u_n \xrightarrow{w} u$, or $\text{w-}\lim_{n\to+\infty} u_n = u$. ♠

Remark [Strong Convergence]*:*

Convergence of (u_n) to u in the norm will be referred to as *strong convergence* and denoted by $u_n \to u$, $u_n \xrightarrow{s} u$, or $\text{s-}\lim_{n\to+\infty} u_n = u$. ♣

Proposition 1.144:

Let U be a normed space and suppose $u_n \rightharpoonup u$. Then the limit u is unique and the sequence (u_n) is bounded. Furthermore, every subsequence of (u_n) converges weakly to u. □

Proof:

Suppose $u_n \rightharpoonup v$. We want to prove that $u = v$. Notice that, given any $f \in U^*$,

$$f(u - v) = f(u) - f(v), \tag{1.16.1a}$$
$$= \lim_{n\to+\infty} f(u_n) - \lim_{n\to+\infty} f(u_n), \tag{1.16.1b}$$
$$= \lim_{n\to+\infty} 0, \tag{1.16.1c}$$
$$= 0, \tag{1.16.1d}$$

and Corollary 1.122 on page 58 implies $u = v$.

To prove boundedness, we use Proposition 1.130 on page 62. For each $f \in U^*$, $\hat{u}_n(f) = f(u_n)$, which is convergent, and thus bounded. From the Uniform Boundedness Principle it follows that

$$\sup_{n\in\mathbb{N}} \|u_n\| = \sup_{n\in\mathbb{N}} \|\hat{u}_n\| < \infty. \tag{1.16.2}$$

Finally, we need to prove every subsequence of (u_n) converges weakly to u. This follows immediately from the fact that $(f(u_n))$ is a convergent sequence in $\mathbb{K}$, and hence all of its subsequences converge to the same value. ■

Notice that

$$\|f(u_n) - f(u)\| \leq \|f\| \|u_n - u\|, \tag{1.16.3}$$

so that strong convergence implies weak convergence with the same limits—justifying the terminology. As we shall see, though, there are examples in which weak convergence does not imply strong convergence.

Proposition 1.145:

In finite dimension, weak and strong convergence coincide. □

Proof:

We already know that strong convergence implies weak convergence with the same limits, so now we have to prove the converse. Suppose U is a finite dimensional normed space and (u_n) is a sequence in U with $u_n \rightharpoonup u$. We want to prove that $u_n \to u$.

Let $\{e_n\}$ be a basis of U. For concreteness, let $k = \dim U$. We can then write

$$u_n = \sum_{i=1}^{k} \alpha_i^{(n)} e_i \tag{1.16.4}$$

and

$$u = \sum_{i=1}^{k} \alpha_i e_i. \tag{1.16.5}$$

Consider the dual basis of $\{e_i\}$, *i.e.*, the basis of U^* given by $\{f_i\}$ with $f_i(e_j) = \delta_{ij}$, where δ_{ij} is the Kronecker delta. Since $u_n \rightharpoonup u$, we know that, for each $1 \le i \le k$,

$$\lim_{n\to+\infty} \alpha_i^{(n)} = \lim_{n\to+\infty} f_i(u_n) = f_i(u) = \alpha_i. \tag{1.16.6}$$

Thus,

$$\begin{aligned} \|u_n - u\| &= \left\| \sum_{i=1}^{k} (\alpha_i^{(n)} - \alpha_i) e_i \right\|, && (1.16.7a)\\ &\le \sum_{i=1}^{k} \left|\alpha_i^{(n)} - \alpha_i\right| \|e_i\|, && (1.16.7b)\\ &\to 0, n \to \infty. && (1.16.7c) \end{aligned}$$

Hence, $u_n \to u$. ■

Proposition 1.146:

Let U be a normed space. Consider a sequence (u_n) in U. $u_n \rightharpoonup u$ if, and only if, $f(u_n) \to f(u)$ for f in a dense set in U^* and bounded $\{\|u_n\|\}$. □

Proof:

One of the implications follows from Proposition 1.144 on the preceding page. We'll focus on proving the other one. Let $W = \{f \in U^*; f(u_n) \to f(u)\}$ and $\infty > C \ge \|u_n\|$ for every n. By assumption, W is dense in U.

If $g \in U^*$, given $\epsilon > 0$ there is some $f \in W$ with $\|f - g\| < \epsilon$. Thus,

$$\begin{aligned} |g(u_n) - g(u)| &\le |g(u_n) - f(u_n)| + |f(u_n) - f(u)| + |f(u) - g(u)|, && (1.16.8a)\\ &\le \epsilon\|u_n\| + |f(u_n) - f(u)| + \epsilon\|u\|. && (1.16.8b) \end{aligned}$$

For sufficiently large n, one has $|f(u_n) - f(u)| < \epsilon$ and thus $|g(u_n) - g(u)| < \epsilon(1 + C + \|u\|)$. Since ϵ is arbitrary, it follows that $g(u_n) \to g(u)$, and thus $u_n \rightharpoonup u$. ■

Example 1.147 [$\ell^1(\mathbb{N})$]**:**

It was shown by Schur (1921) that in $\ell^1(\mathbb{N})$ weak and strong convergence coincide.

To see this, it suffices to prove that weak convergence implies strong convergence, as the converse is always true. Let then (u^n) be a sequence in $\ell^1(\mathbb{N})$ that converges weakly. We can assume it converges weakly to zero.

If (u^n) does not converge strongly to zero, then there is a subsequence of (u^n), which will also be denoted by (u^n), satisfying $\|u^n\|_1 \ge 3\epsilon_0$ for every $n \in \mathbb{N}$ and for some $\epsilon_0 > 0$. However, we know this sequence converges weakly to zero, and thus

$$\lim_{n\to+\infty} \sum_{i=1}^{+\infty} v_i u_i^n = 0, \forall\, v = (v_1, v_2, v_3, \ldots) \in \ell^1(\mathbb{N})^* = \ell^\infty(\mathbb{N}). \tag{1.16.9}$$

If we pick for v the elements e^i of the canonical basis of $\ell^1(\mathbb{N}) \subseteq \ell^\infty(\mathbb{N})$ we get $\lim_{n\to+\infty} u_i^n = 0$ for fixed i. Thus, for each fixed m, $\sum_{i=1}^{m} |u_i^n| < \epsilon_0$ for sufficiently large n. Furthermore, given k, $\sum_{i=M}^{+\infty} \left|u_i^k\right| < \epsilon_0$ for M sufficiently large. We shall use these properties in what follows.

Let $m_0 = n_0 = 1$. Define the strictly growing sequences (m_k) and n_k in the following manner. n_k is the smallest integer larger than n_{k-1} such that

$$\sum_{i=1}^{m_{k-1}} \left|u_i^{n_k}\right| < \frac{\epsilon_0}{2} \tag{1.16.10}$$

and m_k is the smallest integer larger than m_{k-1} such that

$$\sum_{i=m_k+1}^{+\infty} \left|u_i^{n_k}\right| < \frac{\epsilon_0}{2}. \tag{1.16.11}$$

Notice that in these inequalities we use explicitly that $u^n \in \ell^1(\mathbb{N})$. Furthermore, both n_k and m_k tend to infinity.

Now let $v \in \ell^\infty(\mathbb{N})$ be as follows. Let $v_1 = 1$ and for $m_{k-1} < i \le m_k$ let $v_i = 0$ if $u_i^{n_k} = 0$ and $v_i = \bar{u}_i^{n_k} / \left|u_i^{n_k}\right|$ if $u_i^{n_k} \neq 0$ (the bar denotes complex conjugation). Notice that $\|v\|_\infty = 1$.

Hence, using the fact that $|a| - |b| \le |a - b|$ for $a, b \in \mathbb{K}$,

$$
\begin{aligned}
3\epsilon_0 - \left|\sum_{i=1}^{+\infty} v_i u_i^{n_k}\right| &\le \sum_{i=1}^{+\infty} \left|u_i^{n_k}\right| - \left|\sum_{i=1}^{+\infty} v_i u_i^{n_k}\right|, && (1.16.12a)\\
&\le \left|\sum_{i=1}^{+\infty} \left(\left|u_i^{n_k}\right| - v_i u_i^{n_k}\right)\right|, && (1.16.12b)\\
&\le \left|\left(\sum_{i=1}^{m_{k-1}} + \sum_{i=m_k+1}^{+\infty}\right)\left(\left|u_i^{n_k}\right| - v_i u_i^{n_k}\right)\right|, && (1.16.12c)\\
&\le \left(\sum_{i=1}^{m_{k-1}} + \sum_{i=m_k+1}^{+\infty}\right)\left(\left|u_i^{n_k}\right| + \left|v_i u_i^{n_k}\right|\right), && (1.16.12d)\\
&\le \epsilon_0 + \epsilon_0, && (1.16.12e)\\
&= 2\epsilon_0, && (1.16.12f)
\end{aligned}
$$

and thus

$$\left|\sum_{i=1}^{+\infty} v_i u_i^{n_k}\right| \ge \epsilon_0 \tag{1.16.13}$$

for every $k > 1$, which contradicts the weak convergence of the sequence $u^{n_k} \rightharpoonup 0$. Hence, by contradiction, the proof is complete. ♥

We can also define different notions of convergence for sequences of bounded operators.

Definition 1.148 [Topologies on Bounded Operators]**:**

Let U and V be normed spaces and (T_n) be a sequence in $\mathscr{B}(U, V)$. Let $T\colon U \to V$ be linear. We say

i. T_n *converges uniformly*, or *in the norm*, or in the *uniform operator topology*, to T if

$$\|T_n - T\| \to 0, \tag{1.16.14}$$

and we write $T_n \to T$ or $\lim_{n\to+\infty} T_n = T$.

ii. T_n *converges strongly*, or in the *strong operator topology (SOT)*, to T if

$$\|T_n u - Tu\|_V \to 0, \forall\, u \in U \tag{1.16.15}$$

and we write $T_n \xrightarrow{s} T$ or $\text{s-lim}_{n\to+\infty} T_n = T$.

iii. T_n *converges weakly*, or in the *weak operator topology (WOT)*, to T if

$$\left|f(T_n u) - f(Tu)\right| \to 0, \forall\, u \in U, \forall\, f \in V^* \tag{1.16.16}$$

and we write $T_n \rightharpoonup T$, $T_n \xrightarrow{w} T$, or $\text{w-lim}_{n\to+\infty} T_n = T$. ♠

Notice that strong operator convergence is strong convergence in the codomain of the operator, while weak operator convergence is weak convergence in the codomain of the operator.

Proposition 1.149:

Let U and V be normed spaces and let (T_n) be a sequence in $\mathcal{B}(U,V)$. If (T_n) converges to a linear operator $T\colon U \to V$ uniformly, strongly, or weakly, the limit is unique. □

Proof:

We will prove each of the three cases separately. To establish notation, suppose that (T_n) converges to both T and S in the sense being considered at each time. We want to show that $T = S$ in all three cases.

We begin with uniform convergence. Notice that

$$\|T - S\| \leq \|T - T_n\| + \|T_n - S\|. \tag{1.16.17}$$

However, the RHS can be made arbitrarily small by choosing a sufficiently large n. Therefore, $\|T - S\| = 0$. Hence, $T = S$.

Suppose now that convergence is meant in the SOT. Then, $\forall\, u \in U$,

$$\|Tu - Su\|_V \leq \|Tu - T_n u\|_V + \|T_n u - Su\|_V, \tag{1.16.18}$$

and once again the RHS can be made arbitrarily small. Thus, $\|(T - S)u\|_V = 0, \forall\, u \in U$. Hence, $(T - S)u = 0$ for every $u \in U$. This implies $T = S$.

At last, assume convergence is meant in the WOT. Then, $\forall\, u \in U, \forall\, f \in V^*$,

$$\left|f(Tu) - f(Su)\right| \leq \left|f(Tu) - f(T_n u)\right| + \left|f(T_n u) - f(Su)\right|. \tag{1.16.19}$$

One more time we can make the RHS arbitrarily small, leading to $f(Tu) = f(Su)$ for all $u \in U$ and all $f \in V^*$. Corollary 1.122 on page 58 implies $Tu = Su$ for all $u \in U$. Hence, $T = S$. ■

Proposition 1.150:

Let U and V be normed spaces and let (T_n) be a sequence in $\mathcal{B}(U,V)$. Let $T\colon U \to V$ be a linear operator. If $\lim_{n\to+\infty} T_n = T$, *then* $\text{s-lim}_{n\to+\infty} T_n = T$. *If* $\text{s-lim}_{n\to+\infty} T_n = T$, *then* $\text{w-lim}_{n\to+\infty} T_n = T$. □

Proof:

We begin by proving uniform convergence implies strong convergence. We notice that, for any $u \in U$,

$$\|T_n u - Tu\|_V \leq \|T_n - T\| \|u\|, \tag{1.16.20}$$

so $\|T_n - T\| \to 0$ implies $\|T_n u - Tu\| \to 0$ for all $u \in U$.

Next we show strong convergence implies weak convergence. Notice that, for all $u \in U$, $f \in V^*$,

$$\left|f(T_n u) - f(Tu)\right| \leq \|f\| \|T_n u - Tu\|, \tag{1.16.21}$$

meaning $\|T_n u - Tu\| \to 0$ for all $u \in U$ implies $\left|f(T_n u) - f(Tu)\right| \to 0$ for all $u \in U$ and all $f \in V^*$. ■

Example 1.151:

Let $\{e_i\}$ be the canonical basis in $\ell^1(\mathbb{N})$ and define $P_n \colon \ell^1(\mathbb{N}) \to \ell^1(\mathbb{N})$ through

$$P_n(u_1, u_2, u_3, \ldots) = (u_1, u_2, \ldots, u_n, 0, 0, \ldots). \tag{1.16.22}$$

Notice that

$$\|P_n u - \mathbb{1}u\| = \sum_{i=n+1}^{+\infty} |u_i|, \tag{1.16.23}$$

the RHS of which can be made arbitrarily small for sufficiently large n. Therefore, $P_n \xrightarrow{s} \mathbb{1}$. However, for all n one has $\|P_n u - \mathbb{1}u\| \leq \|u\|$ and

$$\|P_n e_{n+1} - \mathbb{1}e_{n+1}\| = \|e_{n+1}\| = 1, \tag{1.16.24}$$

and thus $\|P_n - \mathbb{1}\| = 1$. Hence, P_n does not converge uniformly to $\mathbb{1}$ (and since the limit must be $\mathbb{1}$, it doesn't converge uniformly at all). ♥

Example 1.152:

Consider the sequence of linear operators $T_n \colon \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ defined by

$$T_n u = (\underbrace{0, \ldots, 0}_{n \text{ entries}}, u_{n+1}, u_{n+2}, u_{n+3}, \ldots) \tag{1.16.25}$$

converges strongly to zero, but not uniformly.

Indeed, notice that

$$\|T_n u\| = \left(\sum_{i=n+1}^{+\infty} |u_i|^2\right)^{\frac{1}{2}}, \tag{1.16.26}$$

which can be made arbitrarily small for sufficiently large n. Hence, $T_n \xrightarrow{s} 0$. However, notice that $\|T_n u\| \leq \|u\|$ and $\|T_n e_{n+1}\| = \|e_{n+1}\| = 1$, where $\{e_i\}$ is the canonical basis in $\ell^2(\mathbb{N})$. Hence, $\|T_n\| = 1$, meaning T_n does not converge to zero uniformly. ♥

Example 1.153:

Consider the sequence of linear operators $T_n\colon \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ defined by

$$T_n u = (\underbrace{0, \dots, 0}_{n \text{ entries}}, u_1, u_2, u_3, \dots) \tag{1.16.27}$$

converges weakly to zero, but not strongly.

Let $u, f \in \ell^2(\mathbb{N})$ (recall that $\ell^2(\mathbb{N})$ is its own dual). Denote by $\{e_i\}$ the canonical basis of $\ell^2(\mathbb{N})$. Then

$$\left|f(T_n u)\right| = \left|f\left(\sum_{i=1}^{+\infty} u_i e_{n+i}\right)\right|, \tag{1.16.28a}$$

$$= \left|\sum_{i=1}^{+\infty} u_i f(e_{n+i})\right|, \tag{1.16.28b}$$

$$\leq \left(\sum_{i=1}^{+\infty} |u_i|^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{+\infty} \left|f(e_{n+i})\right|^2\right)^{\frac{1}{2}}, \tag{1.16.28c}$$

$$\leq \left(\sum_{i=1}^{+\infty} |u_i|^2\right)^{\frac{1}{2}} \left(\sum_{i=n+1}^{+\infty} \left|f(e_i)\right|^2\right)^{\frac{1}{2}}. \tag{1.16.28d}$$

where we used the Hölder inequality on Eq. (1.16.28c). The RHS of this equation is (the square root of) a convergent series times the remainder of a convergent series, and hence it can be made arbitrarily small by picking smaller portions of the remainder. Hence, $\left|f(T_n u)\right| \to 0$ and we conclude $T_n \rightharpoonup 0$.

Consider though $\|T_n u\|$. We know $\|T_n u\| = \|u\|$ for all $u \in U$, and hence T_n does not converge strongly. ♥

Lemma 1.154:

Let U be a Banach space and V be a normed space. Suppose (T_n) is a sequence in $\mathcal{B}(U, V)$. If T_n converges strongly to the operator $T\colon U \to V$, then $T \in \mathcal{B}(U, V)$. □

Proof:

Linearity of T follows immediately from linearity of each T_n. We want to prove boundedness. We know $\|T_n u\|$ converges for each $u \in U$, so $\|T_n u\|$ is bounded for each $u \in U$. The Principle of Uniform Boundedness then ensures that $\sup_{n\in\mathbb{N}} \|T_n\| \leq c < \infty$. Then $\|T_n u\| \leq \|T_n\|\|u\| \leq c\|u\|$, from which it follows that $\|Tu\| \leq c\|u\|$, and hence $\|T\| \leq c$. ■

Proposition 1.155:

Let U and V be Banach spaces. A sequence (T_n) is strongly convergent in $\mathcal{B}(U, V)$ if, and only if,

i. *$\{\|T_n\|\}$ is bounded;*

ii. *$(T_n u)$ is a Cauchy sequence for u in a total set in U.* □

Proof:
If $T_n \xrightarrow{s} T$, the first item follows from the Principle of Uniform Boundedness and the second item is straightforward. Let us then focus on proving the converse.

Let $\|T_n\| \leq c$ for all n and suppose $(T_n u)$ is a Cauchy sequence for $u \in M$, with M being a total set in U. Pick $u \in U$. We want to prove that $(T_n u)$ converges strongly in V. Let $\epsilon > 0$. Since M is total, there is $v \in \operatorname{span}(M)$ such that

$$\|u - v\| < \frac{\epsilon}{3c}. \tag{1.16.29}$$

Since $v \in \operatorname{span}(M)$, $(T_n v)$ is a Cauchy sequence as a consequence of the second requirement. Hence, there is $N \in \mathbb{N}$ such that

$$\|T_n v - T_m v\| < \frac{\epsilon}{3}, \forall\, m, n > N. \tag{1.16.30}$$

Using these two inequalities, we notice that for every $\epsilon > 0$ there is $N \in \mathbb{N}$ such that

$$\begin{aligned}
\|T_n u - T_m u\| &\leq \|T_n u - T_n v\| + \|T_n v - T_m v\| + \|T_m v - T_m u\|, && \text{(1.16.31a)}\\
&\leq \|T_n\|\|u - v\| + \|T_n v - T_m v\| + \|T_m\|\|u - v\|, && \text{(1.16.31b)}\\
&< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3}, && \text{(1.16.31c)}\\
&= \epsilon, && \text{(1.16.31d)}
\end{aligned}$$

for all $m, n > N$. Since V is complete, $(T_n u)$ converges in V, proving strong operator convergence. ∎

Bibliography

Axler, Sheldon (2015). *Linear Algebra Done Right*. 3rd ed. Undergraduate Texts in Mathematics. Cham: Springer. DOI: 10.1007/978-3-319-11080-6 (cit. on pp. iii, 11).

Brezis, Haim (2011). *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Universitext. New York: Springer. DOI: 10.1007/978-0-387-70914-7 (cit. on pp. iii, 45).

Ciesielski, Krzysztof (1997). *Set Theory for the Working Mathematician*. London Mathematical Society Student Texts 39. Cambridge University Press. DOI: 10.1017/CBO9781139173131 (cit. on p. 52).

Conway, John B. (2007). *A Course in Functional Analysis*. 2nd ed. Graduate Texts in Mathematics 96. New York, NY: Springer. DOI: 10.1007/978-1-4757-4383-8 (cit. on pp. iii, 5, 45).

Enflo, Per (1973). "A Counterexample to the Approximation Problem in Banach Spaces". *Acta Mathematica* 130, pp. 309–317. DOI: 10.1007/BF02392270 (cit. on p. 22).

Folland, Gerald B. (1999). *Real Analysis: Modern Techniques and Their Applications*. New York: Wiley (cit. on pp. iii, 10, 32).

Geroch, Robert (1985). *Mathematical Physics*. Chicago Lectures in Physics. Chicago: University of Chicago Press (cit. on pp. iii, 24).

James, Robert C. (1951). "A Non-Reflexive Banach Space Isometric With Its Second Conjugate Space". *Proceedings of the National Academy of Sciences* 37.3, pp. 174–177. DOI: 10.1073/pnas.37.3.174 (cit. on p. 63).

Kreyszig, Erwin (1978). *Introductory Functional Analysis With Applications*. New York, NY: Wiley (cit. on pp. iii, 13, 45, 64).

Lima, Elon Lages (2017). *Espaços Métricos*. 5th ed. Projeto Euclides. Rio de Janeiro: IMPA (cit. on p. 41).

Munkres, James (2000). *Topology*. Upper Saddle River: Prentice Hall, Inc (cit. on pp. iii, 5, 8, 13).

Oliveira, César R. de (2018). *Introdução à Análise Funcional*. 2nd ed. Projeto Euclides. Rio de Janeiro: IMPA (cit. on pp. iii, 41, 45).

Reed, Michael C. and Barry Simon (1980). *Methods of Modern Mathematical Physics*. Vol. 1: *Functional Analysis*. Revised and Enlarged Edition. New York: Academic Press (cit. on pp. iii, 4, 45, 64).

Rudin, Walter (1991). *Functional Analysis*. 2nd ed. New York, NY: McGraw-Hill (cit. on p. iii).

Schindler, Ralf (2014). *Set Theory: Exploring Independence and Truth*. Universitext. Cham: Springer. DOI: 10.1007/978-3-319-06725-4 (cit. on p. 52).

Schur, I. (1921). "Über Lineare Transformationen in Der Theorie Der Unendlichen Reihen." *Journal für die reine und angewandte Mathematik* 151, pp. 79–111. DOI: 10.1515/crll.1921.151.79 (cit. on p. 69).

Simon, Barry (2015). *A Comprehensive Course in Analysis*. Vol. 1: *Real Analysis*. Providence, RI: American Mathematical Society. DOI: 10.1090/simon/001 (cit. on pp. iii, 15).

Suppes, Patrick (1972). *Axiomatic Set Theory*. New York: Dover Publications (cit. on p. 52).

Tao, Terence (2022). *Analysis I*. 4th ed. Texts and Readings in Mathematics 37. Singapore: Springer. DOI: 10.1007/978-981-19-7261-4 (cit. on pp. 18, 20).

Yosida, Kôsaku (1995). *Functional Analysis*. 6th ed. Classics in Mathematics. Berlin: Springer (cit. on p. iii).