Functional Analysis and Applications IV

MATH 4371/42920

Michaelmas Term 2023

Amit Einav

Durham University

Department of Mathematical Sciences

The notes presented here were written for the Graduate Level Course 'Functional Analysis' given in the academic year of 23/24, and cover the first part of it, given during the Michaelmas Term.

Contents

Introduction - What is Functional Analysis?

Simply put, functional analysis is the study of infinite dimensional vector spaces.

Many times in mathematics, we are looking for structures and/or symmetries that will help us approach and understand a certain given problem. This is extremely prevalent when considering problems that have connection to physics, biology, chemistry, optimisation and many other topics. One extremely useful such structure is a *linear structure*, i.e. the structure of vector spaces and linear transformation, which you were introduced to in Linear Algebra I. Here are but a few examples where you have seen this structure:

- *Solving a (finite) system of linear equations.*
- *Solving a (finite) system of linear ODEs.* The set of all solutions to ODEs of the form
$$y^{(n)}(x) + p_{n-1}(x)y^{(n-1)}(x) + \dots p_1(x)y(x) = 0$$
forms a vector space. Moreover, one can show that it is an $n-$dimensional vector space - implying that one can find n functions, $y_1(x),\dots,y_n(x)$ such that *any* solution to the above ODE will be of the form
$$y(x) = \alpha_1 y_1(x) + \dots + \alpha_n y_n(x)$$
for some scalars $\alpha_1,\dots,\alpha_n$.
- *Function optimisation.* As you have seen in Calculus I, classification of extreme points rely heavily on our understanding of the Hessian matrix.

There is one common thread in all the above examples, however, and that is that the linear structure you've seen so far is *finite dimensional*. This is rarely the case when considering a complicated (and not approximated) real life problem. Let us consider an example:

Example. Let $\mathscr{P}[x]$ be the set of all real polynomials. As you've seen in Linear Algebra I, $\mathscr{P}[x]$ is a vector space. It is, however, not finite dimensional. Indeed, if it were finite dimensional we would have been able to find some $n \in \mathbb{N}$ and n polynomials, $p_1(x),\dots,p_n(x)$ such that *any* polynomial in $\mathscr{P}[x]$ could have been written as
$$p(x) = \alpha_1 p_1(x) + \dots + \alpha_n p_n(x)$$
for some scalars $\alpha_1,\dots,\alpha_n$. This, however, implies that the degree of all polynomials can't be more than $\max_{i=1,\dots,n} \deg(p_i)$ which is a contradiction. Thus, $\mathscr{P}[x]$ *can't* be a finite dimensional vector space.

While $\mathscr{P}[x]$ is not finite dimensional, it doesn't seem so far fetched for us to conjecture that it is "infinite dimensional". After all, the set of polynomials

$$1, x, x^2, \dots, x^n, \dots$$

is linearly independent and every polynomial in $\mathscr{P}[x]$ can be written as a *finite* linear combination of these polynomials.
We do need to be a bit careful. What does it mean for a set of *infinitely* many vectors to be independent? What does it mean to span the vector space in that case? Do we always mean that we can only use finitely many of our "basis vector"? If $\{e_1, e_2, \dots, e_n, \dots\}$ spans the vector space V can we write

$$x = \sum_{n\in\mathbb{N}} \alpha_n e_n?$$

In what sense is the above sum defined?.
It is exactly here where we need to bring notions from Analysis into the mix as the above infinite sum can be thought of as *the limit* of the partial sums sequence, $S_N = \sum_{n=1}^{N} \alpha_n e_n$.

Combining the linear structure of a vector space with a notion of limits and closeness is where the study of infinite dimensional spaces, the study of *Functional Analysis*, begins.

List of Notations

In our notes we shall use the following notations:

$\mathbb{R}$ - The field of real numbers.

$\mathbb{R}_+$ - The set of non-negative real numbers, $\mathbb{R}_+ = \{x \in \mathbb{R} | x \geq 0\}$.

$\mathbb{C}$ - The field of complex numbers.

$\mathbb{Q}$ - The set of rational numbers.

$\mathbb{Z}$ - The set of integers.

$\mathbb{N}$ - The set of natural numbers (i.e. positive integers).

$\mathbb{N}^*$ - The set of natural numbers and $\{0\}$, $\mathbb{N}^* = \mathbb{N} \cup \{0\}$.

$\mathbb{R}^n$ - The set of $n-$tuples, where $n \in \mathbb{N}$ is given, $(x_1, \ldots, x_n)$ with $\{x_i\}_{i=1,\ldots,n} \subset \mathbb{R}$.

$\mathbb{C}^n$ - The set of $n-$tuples, where $n \in \mathbb{N}$ is given, $(z_1, \ldots, z_n)$ with $\{z_i\}_{i=1,\ldots,n} \subset \mathbb{C}$.

$d(x, y)$ - The value of the metric d on x and y.

$\|x\|$ - The norm of the x.

$|z|$ - The absolute value of the (real or complex) number z.

$\overline{z}$ - The complex conjugate of the number z.

$\langle x, y \rangle$ - The inner product (real or complex) of x and y.

On occasion we will call a function T between two sets a map, or an operator (the latter are mostly used in the context of normed and inner product spaces, while the former is used more frequently in the context of metric and general topological spaces).

$\mathscr{D}(T)$ - The domain of the function/map/operator T. i.e. the set on which T is defined.

$\mathscr{R}(T)$ - The range (sometimes known as image) of the function/map/operator T, i.e. the set of possible values of T:

$$\mathscr{R}(T) = \{y \mid T(x) = y \text{ for some } x \in \mathscr{D}(T)\}$$

We will use the notation $U \subset X$ to indicate that U is a subset (that may equal to) of X.

CHAPTER 1

Banach and Hilbert Spaces

In this chapter we will define the basic structures of Functional Analysis, Banach and Hilbert spaces, and explore their properties.

1.1. Basic notions of distances

We start by reminding ourselves three notions of distances

Definition 1.1.1. Let X be a non-empty set. A function $d : X \times X \to \mathbb{R}_+$ is called a *metric* if it satisfies the following conditions:

m 1 $d(x, y) = 0$ if and only if $x = y$ (Positivity).
m 2 $d(x, y) = d(y, x)$ for all $x, y \in X$ (Symmetry).
m 3 $d(x, y) \le d(x, z) + d(z, y)$ for all $x, y, z \in X$ (Triangle inequality).

The couple (X, d) is called a *metric space*.

In certain cases, the metric, i.e. the distance, is induced from a length - a norm.

Definition 1.1.2. Let $\mathcal{X}$ be a vector space over $\mathbb{F}$, be it $\mathbb{R}$ or $\mathbb{C}$. A function $\|\cdot\| : \mathcal{X} \to \mathbb{R}_+$ is called a *norm* if it satisfies the following conditions:

n 1 $\|x\| \ge 0$ for all $x \in \mathcal{X}$ and $\|x\| = 0$ if and only if $x = 0$ (Positivity).
n 2 $\|\alpha x\| = |\alpha| \, \|x\|$ for all $\alpha \in \mathbb{F}$ and all $x \in \mathcal{X}$ (Homogeneouity).
n 3 $\left\|x + y\right\| \le \|x\| + \left\|y\right\|$ for all $x, y \in \mathcal{X}$ (Triangle inequality).

The couple $(\mathcal{X}, \|\cdot\|)$ is called a *normed space*.

As expected, a norm induces a metric:

Theorem 1.1.3. *Let $(\mathcal{X}, \|\cdot\|)$ be a normed space. Define the function $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_+$ by*

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

Then d is a metric on $\mathcal{X}$. We call it the metric induced by the norm $\|\cdot\|$.
Unless stated otherwise, the metric structure in a normed space will always be the one induced from the norm.

An immediate question we might ask ourselves is: When is a given metric induced by a norm? The answer to this question is provided in the following theorem:

Theorem 1.1.4. *Let $(\mathcal{X}, d)$ be a metric space where $\mathcal{X}$ is a vector space over $\mathbb{R}$ or $\mathbb{C}$. Then the metric d is induced by a norm if and only if*

(i) $d(x,y) = d\left(x+z, y+z\right)$ *for any* $x,y,z \in \mathcal{X}$.
(ii) $d\left(\alpha x, \alpha y\right) = |\alpha|\, d(x,y)$ *for any* $x,y \in \mathcal{X}$ *and scalar* α.

In that case the norm which induces the metric is given by

$$\|x\| = d(x,0).$$

Thinking of $\mathbb{R}^n$ as a canonical example to many of the notions that pertain to distances, we recall that the length of a vector (its norm) is defined by using the geometric notion of the dot product. This could also be generalised:

Definition 1.1.5. Let $\mathcal{X}$ be a vector space over $\mathbb{R}$ or $\mathbb{C}$. A function $\langle\cdot,\cdot\rangle : \mathcal{X}\times\mathcal{X} \to \mathbb{R}$ or $\mathbb{C}$ (respectively) is called an *inner product* if it satisfies the following conditions:

- ***p*** 1 $\langle x,x\rangle \geq 0$ for all $x \in \mathcal{X}$ and $\langle x,x\rangle = 0$ if and only if $x = 0$ (Positivity).
- ***p*** 2 $\left\langle x+y,z\right\rangle = \langle x,z\rangle + \left\langle y,z\right\rangle$ for any $x,y,z \in \mathcal{X}$ (Addition of the first component).
- ***p*** 3 $\left\langle \alpha x, y\right\rangle = \alpha\left\langle x,y\right\rangle$ for any $x,y \in \mathcal{X}$ and any scalar α (Scalar multiplication of the first component).
- ***p*** 4 $\left\langle x,y\right\rangle = \overline{\left\langle y,x\right\rangle}$ (Symmetry/Hermitian property).

The couple $(\mathcal{X}, \langle\cdot,\cdot\rangle)$ is called an *inner product* space and sometimes a *pre-Hilbert space.*

Again, as expected, we find that an inner product induces a norm:

Theorem 1.1.6. *Let* $(\mathcal{X}, \langle\cdot,\cdot\rangle)$ *be an inner product space. Define the function* $\|\cdot\| : \mathcal{X} \to \mathbb{R}_+$ *by*

$$\|x\| = \sqrt{\langle x,x\rangle}.$$

Then $\|\cdot\|$ *is a norm on* $\mathcal{X}$. *We call it the norm induced by the inner product* $\langle\cdot,\cdot\rangle$.

Similarly to before, we can ask ourselves the following: When is a given norm induced by an inner product? The answer to this question is provided in the following theorem:

Theorem 1.1.7. *Let* $(\mathcal{X}, \|\cdot\|)$ *be a normed space over* $\mathbb{R}$ *or* $\mathbb{C}$. *Then the norm* $\|\cdot\|$ *is induced by an inner product if and only if*

$$\left\|x+y\right\|^2 + \left\|x-y\right\|^2 = 2\,\|x\|^2 + 2\left\|y\right\|^2. \tag{1.1}$$

In that case the inner product which induces the norm is given by

$$\left\langle x,y\right\rangle = \frac{\left\|x+y\right\|^2 - \left\|x-y\right\|^2}{4} \tag{1.2}$$

when $\mathcal{X}$ *is a vector field over* $\mathbb{R}$ *and*

$$\left\langle x,y\right\rangle = \frac{\left\|x+y\right\|^2 - \left\|x-y\right\|^2}{4} + i\left(\frac{\left\|x+iy\right\|^2 - \left\|x-iy\right\|^2}{4}\right) \tag{1.3}$$

when $\mathcal{X}$ *is a vector field over* $\mathbb{C}$. *Equation* (1.1) *is known as the parallelogram identity while equations* (1.2) *and* (1.3) *are known as the polarisation identities.*

There will be many cases in our modules that we would like to look at a subset/subspace of a certain space of linear space. In that, unless stated otherwise, *we will always considered the induced metric/norm/inner product obtained by restriction the metric/norm/inner product of the entire space to the subset/subspace.*

1.2. Metric spaces prerequisites

In this short section we remind ourselves a few important notions and theorems from the theory of metric spaces.

Definition 1.2.1. Let (X,d) be a metric space and let $\{x_n\}_{n\in\mathbb{N}}$ be a given sequence in X. We say that $\{x_n\}_{n\in\mathbb{N}}$ is a *Cauchy sequence* (or *Cauchy* in short) if for every $\varepsilon>0$ there exists $n_0\in\mathbb{N}$ such that for any $n,m\geq n_0$ we have that

$$d(x_n,x_m)<\varepsilon.$$

Being a Cauchy sequence means that the elements of the entire sequence are as close as we want to each other, as long as we let the index be large enough. This is reminiscent of the notion of a convergence of a sequence, just without knowing what the limit is. It won't surprise us to find out that:

Lemma 1.2.2. *Let (X,d) be a metric space. Then any converging sequence is Cauchy.*

The converse, however, depends on the metric space we're dealing with and motivates the following definition:

Definition 1.2.3. We say that a metric space (X,d) is *complete* if every Cauchy sequence in X converges (to an element of X).

REMARK 1.2.4 (The importance of completeness and the Cauchy criteria). Completeness and Cauchy sequences are not arbitrary concepts but quite a practical one. A prime example to their use, which we will be able to generalise in a sense, is the study of series.
We remind ourselves that we say that $\sum_{n\in\mathbb{N}}a_n$ converges (i.e. makes sense) if the partial sum sequence, $\{S_N\}_{N\in\mathbb{N}}$, defined by

$$S_N=\sum_{n=1}^{N}a_n$$

converges. A common problem when studying partial sums is that more often than not we can't compute them explicitly. Take, for example, the series $\sum_{n\in\mathbb{N}}\frac{1}{n^2}$. Since $\mathbb{R}$ and $\mathbb{C}$ are complete, however, we don't need to find a limit to show that the partial sum sequence converges - we only need to show that it is Cauchy to conclude that it must convergence. This indeed holds for our example as we can show that

$$|S_N-S_M|\leq\frac{1}{\min(M,N)}$$

which is less than ε as long as $N,M>\frac{1}{\varepsilon}$.

Example 1.2.5 *(Completeness of Euclidean spaces)*. The inner product spaces $\mathbb{R}^n$ and $\mathbb{C}^n$, are complete with respect to the standard distance (induced by the dot product).

Example 1.2.6 *($\mathbb{Q}$ is not complete)*. The set $\mathbb{Q}$ with the induced metric from $\mathbb{R}$ is not a complete space.

In general, when trying to show that a given metric space is complete we have three hurdles to overcome:

- We need to find/guess the limit, x, for an arbitrary Cauchy sequence, $\{x_n\}_{n\in\mathbb{N}}$.
- We need to show that $x \in X$.
- We need to show that $d(x_n, x) \underset{n\to\infty}{\longrightarrow} 0$.

It is important to note that while Cauchy sequences don't necessarily converge, they do share some properties with converging sequences. In particular we have that

Theorem 1.2.7. *Let (X, d) be a metric space and let $\{x_n\}_{n\in\mathbb{N}} \subset X$ be Cauchy. Then $\{x_n\}_{n\in\mathbb{N}}$ is bounded*[1].

The proof of this theorem is almost identical to the proof of the same result for converging sequences.

REMARK 1.2.8. In our module we will almost exclusively consider sequences in a normed space $(\mathscr{X}, \|\cdot\|)$. In that cases, boundedness is equivalent to having some $M > 0$ such that

$$\sup_{n\in\mathbb{N}} \|x_n\| \le M.$$

In many cases in our module we would consider subsets and subspaces of a given metric/normed/inner product spaces. The question of whether these subsets/subspaces are complete is answered by the following theorem:

Theorem 1.2.9. *Let (X, d) be a complete metric space and let M be a subset of X. Then (M, d) is complete if and only if M is closed.*

To prove the above we will rely on the following known theorem:

Theorem 1.2.10. *Let (X, d) be a metric space and let $U \subset X$ be a given set. Then*

(i) $x \in \overline{U}$ if and only if there exists a sequence of points $\{x_n\}_{n\in\mathbb{N}} \subset U$ that converges to x.

(ii) U is closed if and only if every converging sequence of points from U converges to a point in U.

PROOF OF THEOREM 1.2.9. We start by assuming that M is closed and consider a Cauchy sequence $\{x_n\}_{n\in\mathbb{N}}$ in M. Since $M \subset X$ and the metric on M and X are identical we find that $\{x_n\}_{n\in\mathbb{N}}$ is Cauchy in X. Using the fact that X is complete we conclude that $\{x_n\}_{n\in\mathbb{N}}$ converges to an element in X which we will denote by x. According to Theorem 1.2.10 M contains all of the limits of sequences of elements from M since M is closed. Thus, $x \in M$ and (since, again, the metrics on X and M are identical) $\{x_n\}_{n\in\mathbb{N}}$ converges to $x \in M$. This shows that M is

[1] boundedness in a metric space means that for some $x_0 \in X$ we have that $\sup_{n\in\mathbb{N}} d(x_n, x_0) < \infty$.

complete.
Conversely, assume that M is complete. We will use Theorem 1.2.10 to show that M is closed by showing that every converging sequence of elements from M has its limit in M. Indeed, if $\{x_n\}_{n\in\mathbb{N}} \subset M$ converges (in X) to an element x, then according to Lemma 1.2.2 we know that it is Cauchy. Since M is complete and $\{x_n\}_{n\in\mathbb{N}} \subset M$ we find that it must converge to an element $y \in M$. This implies that that $\{x_n\}_{n\in\mathbb{N}}$ converges to y in X and due to the uniqueness of the limit in metric spaces we have that $x = y$. Thus $x \in M$ which shows the desired result. The proof is now complete. □

We end this section by reminding ourselves the notions of *density* and *separability*. These notions were introduced in Analysis III for normed spaces, which is how we will use them in our module, but they can easily be defined on metric spaces.

Definition 1.2.11. A set A in a metric space (X, d) is called *dense* if $\overline{A} = X$.

Definition 1.2.12. A metric space (X, d) is called *separable* if there exists a dense countable set in X.

How do we check if a set is dense or that a metric space is separable? We have the next lemmas to help us

Lemma 1.2.13. *Let (X, d) be a metric space and let A be a set in X. Then A is dense if and only if for every $x \in X$ there exists a sequence of elements from A, $\{x_n\}_{n\in\mathbb{N}}$, that converges to x.*

The proof of the above follows from the sequential criteria for the closedness of sets in metric spaces, Theorem 1.2.10. We leave the proof as an exercise.

Similarly we have that

Lemma 1.2.14. *Let (X, d) be a metric space. Then X is separable if and only if there exists a sequence $\{x_n\}_{n\in\mathbb{N}}$ such that every $x \in X$ is a limit of a subsequence of $\{x_n\}_{n\in\mathbb{N}}$ that converges to x.*

1.3. Banach and Hilbert spaces

We start our study of Banach and Hilbert spaces by recalling the basic definitions and properties of Banach and Hilbert spaces. Many concepts and results are defined and hold equivalently for the field $\mathbb{R}$ or $\mathbb{C}$. In these cases we will use the letter $\mathbb{F}$ to express either fields.

Definition 1.3.1. Let $(\mathscr{X}, \|\cdot\|)$ be a normed space. We say that $\mathscr{X}$ is a *Banach space* if it is complete under the metric induced by $\|\cdot\|$.

Definition 1.3.2. Let $(\mathscr{H}, \langle\cdot,\cdot\rangle)$ be an inner product space. We say that $\mathscr{H}$ is a *Hilbert space* if it is complete under the metric induced by $\langle\cdot,\cdot\rangle$.

Throughout this module we will usually (but not always) use $\mathscr{X}$ to denote a Banach space and $\mathscr{H}$ to denote a Hilbert space.

REMARK 1.3.3. Any Hilbert space is also a Banach space when one considers the norm induced by the given inner product $\langle\cdot,\cdot\rangle$. The converse isn't always true, and Theorem 1.1.6 gives us a necessary and sufficient condition for it to be true.

Example 1.3.4 *(Euclidean spaces with $p-th$ norm).* Let $1 \le p \le \infty$ be given and define a function $\|\cdot\|_p : \mathbb{F}^n \to \mathbb{R}_+$ by

$$\|\boldsymbol{a}\|_p = \left(\sum_{i=1}^{n} |a_i|^p\right)^{\frac{1}{p}}$$

when $1 \le p < \infty$, and

$$\|\boldsymbol{a}\|_\infty = \max_{i=1,\dots,n} |a_i|$$

when $p = \infty$, where $\boldsymbol{a} = (a_1, \dots, a_n)$. Then $\|\cdot\|_p$ is a norm on $\mathbb{F}^n$ for any $1 \le p \le \infty$. Moreover, $\left(\mathbb{F}^n, \|\cdot\|_p\right)$ is a Banach space.
When $p = 2$ the norm is induced by the (standard) inner product

$$\langle \boldsymbol{a}, \boldsymbol{b}\rangle_2 = \sum_{i=1}^{n} a_i \overline{b_i}.$$

REMARK 1.3.5. To show that $\|\cdot\|_p$ is indeed a norm one uses the *finite Hölder inequality*

$$\sum_{i=1}^{n} |a_i b_i| \le \left(\sum_{i=1}^{n} |a_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |b_i|^q\right)^{\frac{1}{q}}.$$

where p and q are Hölder conjugate to prove the *finite Minkoski's inequality*

$$\left(\sum_{i=1}^{n} |a_i + b_i|^p\right)^{\frac{1}{p}} \le \left(\sum_{i=1}^{n} |a_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |b_i|^p\right)^{\frac{1}{p}}$$

for $1 \le p < \infty$. This shows the triangle inequality. The case $p = \infty$ is immediate.

Example 1.3.6 *(L^p spaces).* For any Lebesgue measurable set E and any $p \in [1, \infty]$ the space $L^p(E)$ is a Banach space with respect to the norm

$$\|f\|_p = \left(\int_E |f(x)|^p \, dx\right)^{\frac{1}{p}}$$

when $1 \le p < \infty$ and

$$\|f\|_\infty = \operatorname*{ess\,sup}_{x\in E} |f(x)|$$

when $p = \infty$. When $p = 2$ the norm is induced by the inner product

$$\langle f, g\rangle_2 = \int_E f(x)\overline{g(x)}\,dx,$$

making $L^2(E)$ into a Hilbert space.

Example 1.3.7 *(Continuous functions on a bounded Interval).* Consider the set

$$C([a,b],\mathbb{F}) = \left\{f : [a,b] \to \mathbb{F} \;\middle|\; f \text{ is continuos on } [a,b]\right\},$$

where $-\infty < a < b < \infty$. Then $C([a,b],\mathbb{F})$ is a vector space over $\mathbb{F}$ with respect to pointwise addition and scalar multiplication, where the additive zero is the functions $f \equiv 0$ and the additive inverse of $f \in C([a,b],\mathbb{F})$ is given by $-f$. Moreover, the function $\|\cdot\|_\infty : C([a,b],\mathbb{F}) \to \mathbb{R}_+$ given by

$$\|f\|_\infty = \sup_{x\in[a,b]} |f(x)| = \max_{x\in[a,b]} |f(x)|$$

is a norm on $C([a,b],\mathbb{F})$. Moreover, it can be shown that the normed space $(C([a,b],\mathbb{F}), \|\cdot\|_\infty)$ is a Banach space.
For the sake of brevity we will drop the underlying field $\mathbb{F}$ from this point onwards and write $C[a,b]$. The underlying field, be it $\mathbb{R}$ or $\mathbb{C}$, will be clear from context.

An important example which we will see frequently in our module is the ℓ_p spaces, an infinite dimensional generalisation of $\left(\mathbb{F}^n, \|\cdot\|_p\right)$ from Example 1.3.4.

Example 1.3.8 *(ℓ_p spaces).* For a given $p \in [1,\infty]$ we consider the set

$$\ell_p(\mathbb{N},\mathbb{F}) = \left\{\{a_n\}_{n\in\mathbb{N}} \subset \mathbb{F} \;\middle|\; \sum_{n\in\mathbb{N}} |a_n|^p < \infty\right\}$$

when $1 \le p < \infty$ and

$$\ell_\infty(\mathbb{N},\mathbb{F}) = \left\{\{a_n\}_{n\in\mathbb{N}} \subset \mathbb{F} \;\middle|\; \sup_{n\in\mathbb{N}} |a_n| < \infty\right\}$$

when $p = \infty$. From this point onwards we will drop the underlying field $\mathbb{F}$ and write $\ell_p(\mathbb{N})$. It will be assumed to be known by context.
We can define a linear structure on $\ell_p(\mathbb{N})$ by defining addition and scalar multiplication in the following way:

- For a given $\boldsymbol{a}, \boldsymbol{b} \in \ell_p(\mathbb{N})$ we define $\boldsymbol{a}+\boldsymbol{b}$ as the sequence
$$(\boldsymbol{a}+\boldsymbol{b})_n = a_n + b_n.$$
- For a given $\boldsymbol{a} \in \ell_p(\mathbb{N})$ and a scalar α we define $\alpha\boldsymbol{a}$ as the sequence
$$(\alpha\boldsymbol{a})_n = \alpha a_n,$$

These operators are well defined and it is straight forward to show that the element $\mathbf{0}$ defined by

$$\mathbf{0}_n = 0$$

belongs to $\ell_p(\mathbb{N})$ for any $p \in [1,\infty]$ and is the additive zero of the addition operation, and that for any $\boldsymbol{a} \in \ell_p(\mathbb{N})$ the element $-\boldsymbol{a} \in \ell_p(\mathbb{N})$ is its additive inverse. Thus, $\ell_p(\mathbb{N})$ is in fact a vector space for any $p \in [1,\infty]$.
We can go one step further and define the function $\|\cdot\|_p : \ell_p(\mathbb{N}) \to \mathbb{R}_+$ by

$$\|\boldsymbol{a}\|_p = \left(\sum_{n\in\mathbb{N}} |a_n|^p\right)^{\frac{1}{p}}$$

when $1 \leq p < \infty$, and

$$\|\boldsymbol{a}\|_\infty = \sup_{n\in\mathbb{N}} |a_n|$$

when $p = \infty$. It is possible to show that this function is indeed a norm which, just like in Example 1.3.4, relies on special and very useful inequalities: the *discrete Hölder inequality*

$$\sum_{n\in\mathbb{N}} |a_n b_n| \leq \left(\sum_{n\in\mathbb{N}} |a_n|^p\right)^{\frac{1}{p}} \left(\sum_{n\in\mathbb{N}} |b_n|^q\right)^{\frac{1}{q}}.$$

where p and q are Hölder conjugates and the *discrete Minkoski's inequality*

$$\left(\sum_{n\in\mathbb{N}} |a_n + b_n|^p\right)^{\frac{1}{p}} \leq \left(\sum_{n\in\mathbb{N}} |a_n|^p\right)^{\frac{1}{p}} + \left(\sum_{n\in\mathbb{N}} |b_n|^p\right)^{\frac{1}{p}}.$$

when $1 \leq p < \infty$.
The space $\left(\ell_p(\mathbb{N}), \|\cdot\|_p\right)$ is more than a normed space. It is in fact a Banach space and when $p = 2$ its norm is induced by the inner product

$$\langle \boldsymbol{a}, \boldsymbol{b}\rangle_2 = \sum_{n\in\mathbb{N}} a_n \overline{b_n}.$$

making $\ell_2(\mathbb{N})$ into a Hilbert space.

REMARK 1.3.9. The underlying set of indexes we have used to define $\ell_p(\mathbb{N})$ was the natural numbers $\mathbb{N}$. There is nothing, however, to stop us from defining the space ℓ_p on a general index set $\mathcal{I}$ in a similar way. For instance

$$\ell_p^{\mathbb{F}}(\mathbb{Z}) = \left\{ \{a_n\}_{n\in\mathbb{Z}} \subset \mathbb{F} \;\middle|\; \sum_{n\in\mathbb{Z}} |a_n|^p < \infty \right\}$$

when $1 \leq p < \infty$ and

$$\ell_\infty^{\mathbb{F}}(\mathbb{Z}) = \left\{ \{a_n\}_{n\in\mathbb{N}} \subset \mathbb{F} \;\middle|\; \sup_{n\in\mathbb{Z}} |a_n| < \infty \right\}$$

when $p = \infty$.
When $\mathcal{I}$ is *uncountable* the sum $\sum_{i\in\mathcal{I}} |a_i|^p$ is well defined as a supremum of all possible *countable* sub-sums of the expression, as you've seen in Analysis III. In fact, one can show that if $\sum_{i\in\mathcal{I}} |a_i|^p$ is finite then $a_i = 0$ for all but countable many $i \in \mathcal{I}$.
It is customary to indicate the index of the space by writing $\ell_p(\mathcal{I})$.

We have seen many examples of Banach and Hilbert spaces. The next example will give us a non-trivial normed space which is *not* a Banach space.

Example 1.3.10 *(The space of polynomial is not a Banach space)*. Consider the linear space of polynomials with coefficients in the field $\mathbb{F}$ and which are defined over the bounded closed interval $[a, b]$ for some $-\infty < a < b < \infty$. As it is a subspace of $C[a, b]$ we know that $(\mathscr{P}[x], \|\cdot\|_\infty)$ is a normed space over $\mathbb{F}$. It is, however, not a Banach space as the uniform limit of polynomials is not always a polynomial. Indeed, the seqeunce of polynomials

$$p_n(x) = \sum_{j=0}^{n} \frac{x^j}{j!}$$

converges uniformly on *any* bounded interval to e^x, which doesn't belong to $\mathscr{P}[x]$.

Exercise 1.3.11. Consider the vector space $C[0,1]$ over $\mathbb{R}$ and define the function $\|\cdot\|_1 : C[0,1] \to \mathbb{R}_+$ by

$$\|f\|_1 = \int_0^1 |f(x)|\, dx.$$

Show that $(C[0,1], \|\cdot\|_1)$ is a normed space but not a Banach space.

We conclude this section with an immediate consequence of the completeness theorem for metric spaces, Theorem 1.2.9:

Theorem 1.3.12. *Let $\mathscr{M}$ be a subspace of a Banach space $\mathscr{X}$ or a Hilbert space $\mathscr{H}$. Then $\mathscr{M}$ is a Banach space or a Hilbert space respectively if and only if $\mathscr{M}$ is closed.*

1.4. Basic properties of Banach and Hilbert spaces

We continue our study by exploring a few basic properties of normed and inner product spaces, and reminding ourselves important notions in the theory of Hilbert spaces (which were shown in Analysis III).

Theorem 1.4.1. *Let $(\mathscr{X}, \|\cdot\|)$ be a normed space. Then*

(i) for any $x, y \in \mathscr{X}$

$$\big|\|x\| - \|y\|\big| \le \|x - y\| \qquad \text{(Reverse triangle inequality).}$$

(ii) The norm is a continuous function from $\mathscr{X}$ to $\mathbb{R}_+$. In other words, if $x_n \underset{n\to\infty}{\longrightarrow} x$ then

$$\|x_n\| \underset{n\to\infty}{\longrightarrow} \|x\|.$$

(iii) If $\{x_n\}_{n\in\mathbb{N}} \subset \mathscr{X}$ and $\{y_n\}_{n\in\mathbb{N}} \subset \mathscr{X}$ converge to $x \in \mathscr{X}$ and $y \in \mathscr{X}$ respectively, and if $\{\alpha_n\}_{n\in\mathbb{N}}$ and $\{\beta_n\}_{n\in\mathbb{N}}$ are sequences of scalars that converge to α and β respectively then

$$\alpha_n x_n + \beta_n y_n \underset{n\to\infty}{\longrightarrow} \alpha x + \beta y.$$

In particular, the addition and scalar multiplication operations are continuous.

(iv) If there is an inner product, $\langle\cdot,\cdot\rangle$, that induces the norm $\|\cdot\|$ (i.e. $\mathscr{X}$ is in fact an inner product space) then the inner product is a continuous function from $\mathscr{X} \times \mathscr{X}$ to its underlying field. In other words if $x_n \underset{n\to\infty}{\longrightarrow} x$ and $y_n \underset{n\to\infty}{\longrightarrow} y$ then

$$\langle x_n, y_n\rangle \underset{n\to\infty}{\longrightarrow} \langle x, y\rangle.$$

PROOF OF THEOREM 1.4.1. Using the standard triangle inequality we find that for any $x, y \in \mathscr{X}$

$$\|x\| = \|x - y + y\| \leq \|x - y\| + \|y\|$$

and equivalently (since $\|z\| = \|-z\|$)

$$\|y\| \leq \|x - y\| + \|x\|.$$

As such

$$\big|\|x\| - \|y\|\big| = \max\{\|x\| - \|y\|, \|y\| - \|x\|\} \leq \|x - y\|$$

showing (i).
Using the reverse triangle inequality we have that

$$0 \leq |\|x_n\| - \|x\|| \leq \|x_n - x\|,$$

which according to the pinching lemma implies that

$$\|x_n\| \underset{n\to\infty}{\longrightarrow} \|x\|$$

showing (ii).
Moving to (iii), we notice that

$$\begin{aligned}\|\alpha_n x_n + \beta_n y_n - (\alpha x + \beta y)\| &= \|(\alpha_n x_n - \alpha_n x) + (\alpha_n x - \alpha x) + (\beta_n y_n - \beta_n y) + (\beta_n y - \beta y)\| \\ &\leq |\alpha_n| \|x_n - x\| + |\alpha_n - \alpha| \|x\| + |\beta_n| \|y_n - y\| + |\beta_n - \beta| \|y\|.\end{aligned}$$

If $\{\alpha_n\}_{n\in\mathbb{N}}$ and $\{\beta_n\}_{n\in\mathbb{N}}$ converge to α and β respectively then these sequences must be bounded, i.e. there exists $A, B > 0$ such that $\sup_{n\in\mathbb{N}} |\alpha_n| \leq A$ and $\sup_{n\in\mathbb{N}} |\beta_n| \leq B$. In this case we have that

$$\|\alpha_n x_n + \beta_n y_n - (\alpha x + \beta y)\| \leq A\|x_n - x\| + \underbrace{|\alpha_n - \alpha| \|x\|}_{\underset{n\to\infty}{\longrightarrow} 0} + B\|y_n - y\| + \underbrace{|\beta_n - \beta| \|y\|}_{\underset{n\to\infty}{\longrightarrow} 0}.$$

If, in addition, $\{x_n\}_{n\in\mathbb{N}}$ and $\{y_n\}_{n\in\mathbb{N}}$ converge to x and y respectively then (iii) follows immediately from the above.
Lastly, (iv) was shown in Analysis III and relies on Cauchy-Schwartz inequality and the fact that converging sequences must be bounded. □

We end this short section with the definition of orthogonality and a few related concepts and theorems.

Definition 1.4.2. Let $(\mathscr{H}, \langle\cdot,\cdot\rangle)$ be an inner product space.

(i) We say that x is *orthogonal* to y, and write $x \perp y$ if

$$\langle x, y\rangle = 0.$$

(ii) We say that a set M is *orthogonal* if every two elements of it are orthogonal.

(iii) We say that two sets, A and B, are *orthogonal* if for any $x \in A$ and $y \in B$ we have that $x \perp y$.

(iv) Given a subset M of $\mathscr{H}$ we define the *orthogonal complement* of M, $M^\perp$ to be the set

$$M^\perp = \{x \in \mathscr{H} \mid x \perp y,\ \forall y \in M\}$$

Theorem 1.4.3 *(Pythagoras' theorem). Let* $(\mathscr{H}, \langle\cdot,\cdot\rangle)$ *be an inner product space. If* $\{x_1,\ldots,x_n\}\subset\mathscr{H}$ *and* $x_i \perp x_j$ *for* $i \neq j$ *then*

$$\left\|\sum_{i=1}^{n} x_i\right\|^2 = \sum_{i=1}^{n} \|x_i\|^2. \tag{1.4}$$

Lemma 1.4.4. *Let* $\mathscr{H}$ *be an inner product space and let* M *be a subset of* $\mathscr{H}$*. Then*

(i) $M^\perp$ *is a subspace of* $\mathscr{H}$.
(ii) $M^\perp$ *is a closed set.*
(iii) $M^\perp = \overline{M}^\perp$.
(iv) $M^\perp = \left(\operatorname{span} M\right)^\perp$.
(v) $M^\perp = \overline{\operatorname{span} M}^\perp$.
(vi) Let $M^{\perp\perp}$ *be defined as* $\left(M^\perp\right)^\perp$*. Then, if* $\mathscr{M}$ *is a subspace of* $\mathscr{H}$ *we have that* $\mathscr{M}^{\perp\perp} = \overline{\mathscr{M}}$.

Theorem 1.4.5. *Let* $\mathscr{H}$ *be a Hilbert space and let* $\mathscr{M}$ *be a closed subspace of* $\mathscr{H}$*. Then:*

(i) For any $x \in \mathscr{H}$ *there exists a unique vector* $x_\parallel$ *in* $\mathscr{M}$ *such that*

$$\left\|x - x_\parallel\right\| = \inf_{v\in\mathscr{M}} \|x - v\| = \min_{v\in\mathscr{M}} \|x - v\|.$$

We denote this vector by $P_{\mathscr{M}}(x)$ *and call it the orthogonal projection of* x *on* $\mathscr{M}$.

(ii) For any $x \in \mathscr{H}$ *we have that* $x - P_{\mathscr{M}}(x) \in \mathscr{M}^\perp$.
(iii) $\mathscr{H} = \mathscr{M} \oplus \mathscr{M}^\perp$.

The above theorem has been shown in Analysis III.

REMARK 1.4.6. In some cases we'll encounter, the subspace we'll consider, $\mathscr{M}$, will not necessarily be closed. Property (iii) of the previous theorem can be modified to read as

$$\mathscr{H} = \overline{\mathscr{M}} \oplus \mathscr{M}^\perp. \tag{1.5}$$

This proof is a simple application of (iii) from Lemma 1.4.4, (iii) from Theorem (1.4.5), and the fact that if $\mathscr{M}$ is a subspace then so is $\overline{\mathscr{M}}$.

1.5. Bases in functional analysis

In this section we will investigate what it means to be a basis for infinite dimensional normed spaces. We start by recalling the definition of linear combinations, dependence and independence.

Definition 1.5.1. Given a vector space $\mathscr{X}$ over a field $\mathbb{F}$ and vectors $x_1,\ldots,x_n \in \mathscr{X}$ a *linear combination of* $x_1,\ldots,x_n$ is a vector of the form

$$\alpha_1 x_1 + \ldots \alpha_n x_n,$$

where $\alpha_1, \dots, \alpha_n \in \mathbb{F}$. Given a set $\mathscr{M} \subset \mathscr{X}$ (finite or infinite) we define the *span of* $\mathscr{M}$, span$\mathscr{M}$, as the set of all (finite) linear combinations of vectors form $\mathscr{M}$, i.e.

$$\text{span}\mathscr{M} = \left\{ x = \sum_{i=1}^{n} \alpha_i x_i \;\middle|\; \text{for some } n \in \mathbb{N} \text{ and some } \alpha_1, \dots, \alpha_n \in \mathbb{F},\; x_1, \dots, x_n \in \mathscr{M} \right\}.$$

It is important to note that *linear combinations always include only* finitely *many vectors.*

Definition 1.5.2. Given a vector space $\mathscr{X}$ over a field $\mathbb{F}$ and vectors $x_1, \dots, x_n \in \mathscr{X}$ we say that $x_1, \dots, x_n$ are *linearly independent*, or in short independent, if the equality

$$\alpha_1 x_1 + \dots \alpha_n x_n = 0$$

implies that $\alpha_1 = \dots = \alpha_n = 0$. Otherwise we say that the vectors are *linearly dependent*, or dependent in short.
We say that a given set $\mathscr{M} \subset \mathscr{X}$ (finite or infinite) is *linearly independent*, or independent in short, if every finite set of vectors from $\mathscr{M}$ are linearly independent. Otherwise we say that $\mathscr{M}$ is *linearly dependent*, or dependent in short.

In our module we will consider three possible type of bases in an infinite dimensional normed space:

- Algebraic bases (Hamel basis) - a purely algebraic construction that doesn't require a norm. These bases always exist but are not extremely useful in the context of Functional Analysis.
- Schauder bases - an independent *sequence* of vectors, $\{e_n\}_{n\in\mathbb{N}}$, such that every element in our space, x, can be written *uniquely* as an *infinite* "linear combination" of these vectors, i.e.
$$x = \sum_{n\in\mathbb{N}} \alpha_n e_n.$$
Such bases don't always exist, but are fundamental when they do.
- Orthonormal bases - special bases in Hilbert spaces that rely on the notion of orthogonality. Such bases always exist, and when they are countable they are automatically Schauder bases (though never Hamel bases when the dimension is not finite). You have encountered these bases in Analysis III.

1.5.1. Hamel bases. We start with the algebraic bases of vector spaces - the so-called Hamel basis.

Definition 1.5.3. Let $\mathscr{X}$ be a vector space over $\mathbb{F}$. We say that a set $\mathscr{B}$ is a *Hamel basis* for $\mathscr{X}$ if $\mathscr{B}$ is independent and

$$\text{span}\mathscr{B} = \mathscr{X}.$$

If the set $\mathscr{B}$ is finite then we say that $\mathscr{X}$ is *finite dimensional* and we denote its *dimension* as $\dim\mathscr{X} = \#\mathscr{B}$. Otherwise, we say that $\mathscr{X}$ is *infinite dimensional* and write $\dim\mathscr{X} = \infty$. The trivial space (which has no basis) is said to be of dimension zero.

REMARK 1.5.4. One can show that the notion of dimension is well defined by proving that if $\mathscr{B}_1$ and $\mathscr{B}_2$ are Hamel bases for $\mathscr{X}$ then they have the same cardinality.

Example 1.5.5 *(Hamel basis for Euclidean spaces)*. The set of vectors

$$\boldsymbol{e}_1 = (1, 0, \ldots, 0, 0)$$

$$\boldsymbol{e}_2 = (0, 1, \ldots, 0, 0)$$

$$\vdots$$

$$\boldsymbol{e}_n = (0, 0, \ldots, 0, 1)$$

is a Hamel basis for $\mathbb{R}^n$ over $\mathbb{R}$ and $\mathbb{C}^n$ over $\mathbb{C}$.

Example 1.5.6 *(Polynomials of degree less or equal to n)*. Denoting by $\mathscr{P}_n[x]$ the set of all polynomials of degree less or equal to n over $\mathbb{F}$ (to be understood from the context of the settings) we have that

$$\mathscr{B} = \left\{1, x, \ldots, x^n\right\}$$

is a Hamel basis for $\mathscr{P}_n[x]$.

Example 1.5.7 *(Polynomials of any degree)*. Similarly to the above example, if we denote by $\mathscr{P}[x]$ the set of all polynomials over $\mathbb{F}$ then

$$\mathscr{B} = \left\{1, x, \ldots, x^n, \ldots\right\}$$

is a Hamel basis for $\mathscr{P}[x]$. Since $\mathscr{B}$ is not finite, $\mathscr{P}[x]$ is an infinite dimensional vector space.

Example 1.5.8 *(A Hamel basis for ℓ_p?)*. As we saw, $\ell_p(\mathbb{N})$ seems like a natural "extension" of $\left(\mathbb{F}^n, \|\cdot\|_p\right)$. As these spaces have a relatively straight forward Hamel basis, we can try to "extend" it to this infinite dimensional space. Consider the set (of sequences) $\mathscr{B} = \{\boldsymbol{e}_n\}_{n\in\mathbb{N}} \subset \ell_p(\mathbb{N})$

$$(\boldsymbol{e}_n)_k = \begin{cases} 1, & k = n, \\ 0, & k \neq n, \end{cases} \tag{1.6}$$

or more explicitly:

$$\boldsymbol{e}_n = \left(0, \ldots, 0, \underbrace{1}_{n-\text{th poistion}}, 0, \ldots\right).$$

While quite natural, $\mathscr{B}$ is *not* a Hamel basis for $\ell_p(\mathbb{N})$. Indeed, the element

$$\boldsymbol{a} = \left(1, \frac{1}{2}, \ldots, \frac{1}{2^{n-1}}, \ldots\right)$$

which is $\ell_p(\mathbb{N})$ for any $p \in [1, \infty]$ can't be written as a linear combination of elements in $\mathscr{B}$.

The above example illustrates some of the failings of the Hamel basis. intuitively speaking, $\{\boldsymbol{e}_n\}_{n\in\mathbb{N}}$ does seem like a "basis" for $\ell_p(\mathbb{N})$ as we can write, *at least formally,*

$$\boldsymbol{a} = (a_1, a_2, \dots, a_n, \dots) = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + \cdots + a_n\boldsymbol{e}_n + \cdots = \sum_{n\in\mathbb{N}} a_n\boldsymbol{e}_n.$$

Can we make sense of this?

Lemma 1.5.9. *Let* $\boldsymbol{a} = \{a_n\}_{n\in\mathbb{N}} \in \ell_p(\mathbb{N})$ *for some* $1 \le p < \infty$. *Then the partial sum sequence* $\{S_N(\boldsymbol{a})\}_{N\in\mathbb{N}}$ *defined by*

$$S_N(\boldsymbol{a}) = \sum_{n=1}^{N} a_n\boldsymbol{e}_n$$

is also in $\ell_p(\mathbb{N})$ *and converges to* $\boldsymbol{a}$ *in norm as N goes to infinity.*

PROOF. The fact that $S_N(\boldsymbol{a}) \in \ell_p(\mathbb{N})$ for all $N \in \mathbb{N}$ is immediate from the fact that $\ell_p(\mathbb{N})$ is a vector space and that for any $N \in \mathbb{N}$, $S_N(\boldsymbol{a})$ is a (finite) linear combination of elements form $\ell_p(\mathbb{N})$.
To conclude the convergence we notice that as

$$S_N(\boldsymbol{a}) = (a_1, a_2, \dots, a_N, 0, 0, \dots)$$

we have that

$$\|\boldsymbol{a} - S_N(\boldsymbol{a})\|_p^p = \sum_{N+1}^{\infty} |a_n|^p \underset{N\to\infty}{\longrightarrow} 0$$

since the series $\sum_{n\in\mathbb{N}} |a_n|^p$ converges. This concludes the proof. □

Lemma 1.5.9 will give us the means to consider, and consequently define, a notion of a basis that bypasses the finiteness required by Hamel bases.

REMARK 1.5.10. It is worth to note that while writing $\boldsymbol{a} = \sum_{n\in\mathbb{N}} a_n\boldsymbol{e}_n$ holds formally for all $\boldsymbol{a} \in \ell_p(\mathbb{N})$ with $p \in [1,\infty]$, Lemma 1.5.9 holds *only* when $p < \infty$. Indeed, consider the element $\boldsymbol{a} \in \ell_\infty(\mathbb{N})$ given by

$$\boldsymbol{a} = (1, 1, \dots, 1, \dots).$$

We have that

$$S_N(\boldsymbol{a}) = \left(1, 1, \dots, \underbrace{1}_{n-\text{th poistion}}, 0, 0, \dots\right)$$

and

$$\boldsymbol{a} - S_N(\boldsymbol{a}) = \left(0, 0, \dots, \underbrace{0}_{n-\text{th poistion}}, 1, 1, \dots\right).$$

Consequently,

$$\|\boldsymbol{a} - S_N(\boldsymbol{a})\|_\infty = 1,$$

which doesn't go to zero.
This is a good indication that we need to be careful with formal writing and our intuition when infinite dimension in involved.

While slightly problematic, Hamel basis have the advantage of *always existing*. This is expressed in the following theorem:

Theorem 1.5.11. *Every non-zero vector space $\mathcal{X}$ has a Hamel basis.*

Intuitively speaking we would like to "build up" our Hamel basis by creating increasing sets of independent vectors where each set includes one additional vector which is independent of the span of the previous set. This process, however, is *inductive* and would imply that our "final" set *must be our initial set plus countably many vectors* (in most cases our initial set will just be one vector). Not only is it problematic in the sense that we don't know the cardinality of our original space apriori, but we might also reach a situation like that of the standard basis in $\ell_p(\mathbb{N})$, or of $\left\{1, x, x^2, \dots\right\}$ in $C[a,b]$ - our suspected Hamel basis *can't span everything*. This is not a coincidence and we will see (much) later that:

Theorem 1.5.12. *Every Hamel basis of a Banach space is uncountable.*

To prove Theorem 1.5.11 we need to rely on an argument that is not inductive. In most cases where this happen we end up using Zorn's lemma. We leave this as a (difficult) exercise.

In our next subsection we will consider a notion of basis that is motivated from the intuition that led us to Lemma 1.5.9 - Schauder basis.

1.5.2. Schauder bases.

Definition 1.5.13. Let $\mathcal{X}$ be a Banach space. A set $\mathcal{B}$ is called a *Schauder basis for $\mathcal{X}$* if $\mathcal{B}$ is a *countable* set of independent vectors, i.e. $\mathcal{B} = \{e_n\}_{n\in\mathbb{N}} \subset \mathcal{X}$ where $\{e_n\}_{n\in\mathbb{N}}$ are independent, and for every $x \in \mathcal{X}$ there exists a *unique* sequence of scalars, $\{\alpha_n(x)\}_{n\in\mathbb{N}}$ such that the partial sum sequence

$$S_N(x) = \sum_{n=1}^{N} \alpha_n(x) e_n$$

converges to x as N goes to infinity, i.e.

$$\lim_{N\to\infty} \left\| \sum_{n=1}^{N} \alpha_n(x) e_n - x \right\| = 0. \tag{1.7}$$

REMARK. The fact that the partial sum series, $\{S_N(x)\}_{N\in\mathbb{N}}$ converges to x gives meaning to the notation

$$x = \sum_{n\in\mathbb{N}} \alpha_n(x) e_n.$$

From this point onwards, when we write $x = \sum_{n\in\mathbb{N}} \alpha_n e_n$ we will mean that the partial sum sequence $S_N = \sum_{n=1}^{N} \alpha_n e_n$ converges (in norm) to x as N goes to infinity.

Example 1.5.14 *(Schauder basis for ℓ_p with $1 \le p < \infty$).* The set $\mathcal{B} = \{\boldsymbol{e}_n\}_{n\in\mathbb{N}}$, where $\boldsymbol{e}_n$ were defined in (1.6) is a Shauder basis for $\ell_p(\mathbb{N})$ for any $1 \le p < \infty$. $\mathcal{B}$ is called *the standard basis for ℓ_p*.

Example 1.5.15 *(Schauder basis for ℓ_∞?).* From Remark 1.5.10 we have the inkling that the standard basis $\mathscr{B}$ defined above is *not* a Schauder basis for $\ell_\infty(\mathbb{N})$. In fact, $\ell_\infty(\mathbb{N})$, *can't* have a Schauder basis, as we will see later.

Example 1.5.16 *(Schauder basis for $C[0,1]$).* The space $C[0,1]$ has a Schauder basis. One choice for such a basis is the Faber-Schauder basis given by

$$f_0(x) = 1,$$
$$f_1(x) = x,$$

and for $2^{k-1} < n \le 2^k$ with $k \ge 1$ we define

$$f_n(x) = \begin{cases} 2^k\left(x - \left(2^{-k}(2n-2)-1\right)\right) & x \in \left[2^{-k}(2n-2)-1, 2^{-k}(2n-1)-1\right] \\ 1 - 2^k\left(x - \left(2^{-k}(2n-1)-1\right)\right) & x \in \left[2^{-k}(2n-1)-1, 2^{-k+1}n-1\right] \\ 0 & \text{otherwise} \end{cases}$$

If we draw these functions for $n \ge 2$ we will get tents of height 1 which "sweep" across the interval $[0,1]$.
The "natural" candidate $\mathscr{B} = \left\{1, x, x^2, \dots\right\}$ is not a Scahuder basis for $C[a,b]$ but it can shown that span$\mathscr{B}$, which is the space of polynomials of any order $\mathscr{P}[x]$, is dense in $C[a,b]$. This is known as *Weierstrass' Approximation Theorem.*

The above exercises show that it is not immediate that Schauder bases exist, even in "simple" cases. While such bases follow our intuition more than the Hamel bases, there are two properties that they demand:

- a *countable* set of vectors whose linear combination can *approximate* any vector in the space.
- a unique representation of the aforementioned linear combination - a power series like expansion where the coefficients of the partial sums *do not* change as we refine our approximation. This is, in many cases, a delicate point.

The first property of Schauder basis reminds us of the notions of density and separability. Indeed, if $\mathscr{B} = \{e_n\}_{n\in\mathbb{N}}$ is a Schauder basis then, since $\{S_N(x)\}_{N\in\mathbb{N}}$ converges to x for any $x \in \mathscr{X}$ and since $S_N(x) \in \text{span}\mathscr{B}$ for any $N \in \mathbb{N}$ we can conclude that

$$\overline{\text{span}\mathscr{B}} = \mathscr{X}$$

which means that if $\mathscr{X}$ has a Schauder basis we can always find a dense subspace of it that is spanned by a countable set. Separability, however, is not as immediate as span$\mathscr{B}$ is not countable (it contains a "copy" of the underlying field $\mathbb{F}$ which is always uncountable). But, as $\mathbb{F}^n$ is separable for any $n \in \mathbb{N}$[2], and as any element in $\mathscr{X}$ can be approximated by an element in span$\mathscr{B}$, which is a finite linear combination of elements from $\mathscr{B}$,we are not too surprised to find out the following:

[2] The countable sets

$$\mathbb{Q}^n = \left\{(q_1, \dots, q_n) \mid q_1, \dots, q_n \in \mathbb{Q}\right\}$$
$$(\mathbb{Q} + i\mathbb{Q})^n = \left\{(p_1 + iq_1, \dots, p_n + iq_n) \mid p_1, q_1, \dots, p_n, q_n \in \mathbb{Q}\right\}$$

are dense in $\mathbb{R}^n$ and $\mathbb{C}^n$ respectively

Lemma 1.5.17. *Let $\mathscr{X}$ be a normed space over $\mathbb{F}$. If M is a countable set and* span$M$ *is dense in $\mathscr{X}$ then $\mathscr{X}$ is separable. Consequently, if $\mathscr{X}$ has a Schauder basis then it is separable.*

PROOF. We will outline the proof, but leave the details as an exercise.
Step 1: Since M is countable we can find a sequence $\{e_n\}_{n\in\mathbb{N}}$ such that $M = \{e_n\}_{n\in\mathbb{N}}$. Define the countable set

$$\mathscr{M}_n = \begin{cases} \left\{\sum_{i=1}^n q_i e_i \mid q_i \in \mathbb{Q}\right\}, & \mathbb{F} = \mathbb{R}, \\ \left\{\sum_{i=1}^n q_i e_i \mid q_i \in \mathbb{Q} + i\mathbb{Q}\right\}, & \mathbb{F} = \mathbb{C}, \end{cases}$$

and the set $\mathscr{M} = \cup_{n\in\mathbb{N}}\mathscr{M}_n$. $\mathscr{M}_n$ is countable for any $n \in \mathbb{N}$ and consequently $\mathscr{M}$ is also countable.
Step 2: We can show that $\mathscr{M}$ is dense in span$M$.
Step 3: We can conclude that $\overline{\mathscr{M}} = \overline{\text{span}M} = \mathscr{X}$ which shows the separability of $\mathscr{X}$. □

We can sum up our recent observations in the next corollary:

Corollary 1.5.18. Let $\mathscr{X}$ be an infinite dimensional Banach space. If $\mathscr{X}$ has a Schauder basis, $\mathscr{B} = \{e_n\}_{n\in\mathbb{N}}$, then span$\mathscr{B}$ is dense in $\mathscr{X}$. Moreover, $\mathscr{X}$ is separable.

A very natural question at this point is: Does every separable Banach space have a Shcauder basis? One could imagine that we can construct a basis inductively, due to the separability, but surprisingly, the answer to this question is *No*. This was shown by Per Enflo in 1973, where they have constructed a separable Banach space that had no Schauder basis. The issue one encounters when trying to use an inductive argument is not in finding a countable set of independent vectors whose closure of a span is the entire space - it is in the uniqueness of the representation.

Corollary 1.5.18 does gives us a simple criterion to identify when a Banach space *doesn't* have a Schauder basis:

Corollary 1.5.19. Any non-separable Banach space can't have a Schauder basis.

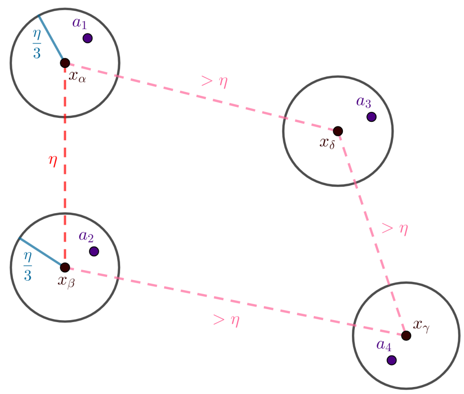
In order to be able to utilise the above corollary we present the following criterion for lack of separability:

Theorem 1.5.20. *Let X be a metric and assume that there exists $\eta > 0$ and an uncountable set $\{x_\alpha\}_{\alpha\in\mathscr{I}}$ such that for any $\alpha \neq \beta \in \mathscr{I}$*

$$d\left(x_\alpha, x_\beta\right) > \eta.$$

Then $\mathscr{X}$ is not separable.

The main idea of the proof is the following: Since the elements $\{x_\alpha\}_{\alpha\in\mathscr{I}}$ are at least $\eta-$far apart, balls with radii $\frac{\eta}{3}$ around these elements are mutually disjoint and must include distinct elements from any given dense set. Thus, no dense set can be countable



Theorem 1.5.20 can be used to show that ℓ_∞ can't have a Schauder basis.

1.5.3. Orthonormal bases. The last type of bases we'll consider are bases that are unique to the setting of inner product spaces - orthonormal bases.

Definition 1.5.21. Let $\mathscr{H}$ be an inner product space. We say that a set M is *orthonormal* if it is orthogonal and every element in M has length 1, i.e. for any $x, y \in M$ we have that

$$\langle x, y \rangle = \begin{cases} 0, & x \neq y, \\ 1, & x = y. \end{cases}$$

Orthonormal sets have a few nice properties, expressed in the following lemma:

Lemma 1.5.22. *Let $\mathscr{H}$ be an inner product space and let M be an orthonormal set. Then*

(i) M is independent.

(ii) If

$$x = \sum_{i=1}^{n} \alpha_i e_i$$

for some $\{e_1, \ldots, e_n\} \subset M$ and $x \in \mathscr{H}$ then $\alpha_i = \langle x, e_i \rangle$. Moreover if

$$x = \sum_{n \in \mathbb{N}} \alpha_n e_n,$$

for a sequence of elements $\{e_n\}_{n\in\mathbb{N}} \subset M$ then the same formula for α_i holds. In this case we call the coefficients

$$\alpha_n = \langle x, e_n \rangle \tag{1.8}$$

the Fourier coefficients of x with respect to the orthonormal sequence $\{e_n\}_{n\in\mathbb{N}}$.

(iii) For any $x \in \mathscr{H}$ and $\{e_1, \ldots, e_n\} \subset M$

$$\left\| x - \sum_{i=1}^{n} \langle x, e_i \rangle e_i \right\|^2 = \|x\|^2 - \sum_{i=1}^{n} |\langle x, e_i \rangle|^2. \tag{1.9}$$

Consequently, if $\{e_n\}_{n\in\mathbb{N}}$ is a sequence of elements in M then

$$\sum_{n\in\mathbb{N}} |\langle x, e_n \rangle|^2 \leq \|x\|^2. \tag{1.10}$$

This inequality is knows as Bessel's inequality.

PROOF. We start by showing (ii).
Assume that

$$x = \sum_{i=1}^{n} \alpha_i e_i.$$

Then, due to the linearity of the inner product and the orthonormality of $\{e_1, \dots, e_n\}$, we have that for any $j \in \{1, \dots, n\}$

$$\langle x, e_j \rangle = \left\langle \sum_{i=1}^{n} \alpha_i e_i, e_j \right\rangle = \sum_{i=1}^{n} \alpha_i \langle e_i, e_j \rangle = \sum_{i=1}^{n} \alpha_i \delta_{i,j} = \alpha_j$$

where $\delta_{i,j}$ is the Kronecker delta. Thus, the first half of the claim is shown. To prove the second half we notice that by the definition of the infinite sum and the continuity of the inner product we have that

$$\langle x, e_j \rangle = \left\langle \lim_{N\to\infty} \sum_{n=1}^{N} \alpha_n e_n, e_j \right\rangle = \lim_{N\to\infty} \left\langle \sum_{n=1}^{N} \alpha_n e_n, e_j \right\rangle$$

$$= \lim_{N\to\infty} \sum_{n=1}^{N} \alpha_n \langle e_n, e_j \rangle = \lim_{N\to\infty} \begin{cases} \alpha_j & j \le N \\ 0 & j > N \end{cases} = \alpha_j$$

for any $j \in \mathbb{N}$. We conclude that (ii) indeed holds.
Next we turn our attention to (i). Assume that $\{e_1, \dots, e_n\} \subset M$ and that there exist scalars $\alpha_1, \dots, \alpha_n$ such that

$$0 = \alpha_1 e_1 + \cdots + \alpha_n e_n.$$

According to (ii) we have that for any $i \in \{1, \dots, n\}$

$$\alpha_i = \langle 0, e_i \rangle = 0,$$

and as such $\{e_1, \dots, e_n\}$ are independent. Since the choice of $\{e_1, \dots, e_n\}$ was arbitrary M must also be independent.
Lastly, we turn our attention to (iii). We have that

$$\left\| x - \sum_{i=1}^{n} \langle x, e_i \rangle e_i \right\|^2 = \|x\|^2 - 2\mathrm{Re}\left(\left\langle x, \sum_{i=1}^{n} \langle x, e_i \rangle e_i \right\rangle \right) + \left\| \sum_{i=1}^{n} \langle x, e_i \rangle e_i \right\|^2. \tag{1.11}$$

Using the conjugate linearity in the second component of the inner product we find that

$$\left\langle x, \sum_{i=1}^{n} \langle x, e_i \rangle e_i \right\rangle = \sum_{i=1}^{n} \overline{\langle x, e_i \rangle} \langle x, e_i \rangle = \sum_{i=1}^{n} |\langle x, e_i \rangle|^2,$$

and utilising to Pythagoras' theorem and the fact that if $e_i \perp e_j$ then $\alpha e_i \perp \beta e_j$ for any scalars α and β, we find that

$$\left\| \sum_{i=1}^{n} \langle x, e_i \rangle e_i \right\|^2 = \sum_{i=1}^{n} \|\langle x, e_i \rangle e_i\|^2 = \sum_{i=1}^{n} |\langle x, e_i \rangle|^2 \|e_i\|^2 = \sum_{i=1}^{n} |\langle x, e_i \rangle|^2.$$

Combining these two observations with (1.11) shows that

$$\left\| x - \sum_{i=1}^{n} \langle x, e_i \rangle e_i \right\|^2 = \|x\|^2 - 2\mathrm{Re}\left(\underbrace{\sum_{i=1}^{n} |\langle x, e_i \rangle|^2}_{\text{real number}} \right) + \sum_{i=1}^{n} |\langle x, e_i \rangle|^2$$

$$= \|x\|^2 - \sum_{i=1}^{n} |\langle x, e_i \rangle|^2 ,$$

which is exactly (1.9).
To conclude the second half of the claim, and the proof of the theorem, we notice that due to (1.9) we have that for any sequence of elements in M, $\{e_n\}_{n\in\mathbb{N}}$,

$$\sum_{i=1}^{n} |\langle x, e_i \rangle|^2 = \|x\|^2 - \left\| x - \sum_{i=1}^{n} \langle x, e_i \rangle e_i \right\|^2 \leq \|x\|^2 ,$$

for any $n \in \mathbb{N}$. As such, according to Weierstrass' $M-$test, we conclude that $\sum_{i=1}^{\infty} |\langle x, e_i \rangle|^2 < \infty$ and that

$$\sum_{n\in\mathbb{N}} |\langle x, e_n \rangle|^2 \leq \|x\|^2 .$$

The proof is now complete □

These nice additional properties make orthonormal sets quite appealing to be considered as basis of some form. Looking back at the definition of Schauder bases we see that if $\mathscr{B} = \{e_n\}_{n\in\mathbb{N}}$ is a countable orthonormal set then

- According to (i) $\mathscr{B}$ is independent.
- According to (ii) if $S_N(x) = \sum_{n=1}^{N} \alpha_n(x) e_n$ converges to x then $\alpha_n(x)$ must be the Fourier coefficients (1.8), i.e. the uniqueness of the coefficients in the expansion is established.
- According to (iii) we have that any $x \in \mathscr{H}$

$$\left\| x - \sum_{n=1}^{N} \langle x, e_n \rangle e_n \right\|^2 = \|x\|^2 - \sum_{n=1}^{N} |\langle x, e_n \rangle|^2 ,$$

  which implies that

$$x = \sum_{n\in\mathbb{N}} \langle x, .e_n \rangle e_n \Leftrightarrow \lim_{N\to\infty} \left\| x - \sum_{n=1}^{N} \langle x, e_n \rangle e_n \right\| = 0 \Leftrightarrow \sum_{n\in\mathbb{N}} |\langle x, e_n \rangle|^2 = \|x\|^2 .$$

We conclude the following theorem

Theorem 1.5.23. *Let $\mathscr{H}$ be a Hilbert space and let $\mathscr{B} = \{e_n\}_{n\in\mathbb{N}}$ be countable orthonormal set. Then $\mathscr{B}$ is a Schauder basis for $\mathscr{H}$ if and only if*

$$x = \sum_{n\in\mathbb{N}} \langle x, e_n \rangle e_n \tag{1.12}$$

for any $x \in \mathscr{H}$, or equivalently if

$$\sum_{n\in\mathbb{N}} |\langle x, e_n \rangle|^2 = \|x\|^2 \tag{1.13}$$

for any $x \in \mathscr{H}$. Equality (1.13) *is known as Parseval's identity.*

REMARK 1.5.24. While we considered the situation where $x = \sum_{n\in\mathbb{N}} \langle x, .e_n\rangle e_n$ in the above, it is important to note that If $\mathscr{H}$ is a Hilbert space and $\mathscr{B} = \{e_n\}_{n\in\mathbb{N}}$ is an orthonormal set, then for any $x \in \mathscr{H}$ the vector

$$\widetilde{x} = \sum_{n\in\mathbb{N}} \langle x, .e_n\rangle e_n$$

is well defined $\mathscr{H}$. Let us prove this claim: Consider the following partial sums in $\mathscr{H}$ and $\mathbb{R}$ respectively,

$$S_N(x) = \sum_{n=1}^{N} \langle x, e_n\rangle e_n, \qquad s_N = \sum_{n=1}^{n} |\langle x, e_n\rangle|^2 .$$

Since $\mathscr{B}$ is an orthonormal set we can use Pythagoras' theorem to conclude that

$$\begin{aligned}\|S_N(x) - S_M(x)\|^2 &= \left\| \sum_{\min\{N,M\}+1}^{\max\{N,M\}} \langle x, e_n\rangle e_n \right\|^2 \\ &= \sum_{\min\{N,M\}+1}^{\max\{N,M\}} |\langle x, e_n\rangle|^2 = |s_N - s_M| .\end{aligned}$$

Due to Bessel's inequality we know that $\{s_N\}_{N\in\mathbb{N}}$ converges and as such is Cauchy. The above identity implies that $\{S_N(x)\}_{N\in\mathbb{N}}$ is also Cauchy and since $\mathscr{H}$ is complete it converges to an element in $\mathscr{H}$. This is the element which we denote by $\widetilde{x}$.

Example 1.5.25. the space $\ell_2(\mathbb{N})$, which we have shown to be a Hilbert space, has an orthonormal Schauder basis given by the standard basis $\{\boldsymbol{e}_n\}$.

REMARK 1.5.26. It is quite straight forward to see that not all Schauder basis in a Hilbert space are orthonormal. However, as we'll see soon, using the Gram-Schmidt procedure we will be able to transform such a basis to an orthonormal one. Here we really use the countability of the basis, as otherwise we wouldn't be able to use the *inductive* Gram-Schmidt procedure.

So far we have only considered the case where our orthonormal set, $\mathscr{B}$, is countable *á la Schauder*, yet the power of the notion of orthonormality is in the fact that we can extend our setting to the *uncountable* case as well. To do so we start by noticing the following:

Lemma 1.5.27. *Let $\mathscr{H}$ be an inner product space and let $\mathscr{B} = \{e_\alpha\}_{\alpha\in\mathscr{I}}$ be orthonormal. If $\mathscr{I}$ is uncountable, then for any $x \in \mathscr{H}$ we have that $\langle x, e_\alpha\rangle \neq 0$ for at most a countable subset of $\mathscr{B}$, $\{e_{\alpha_n}\}_{n\in\mathbb{N}}$.*

This was mentioned in Analysis III.

An immediate corollary of the above is the following theorem:

Theorem 1.5.28. *Let $\mathscr{H}$ be a Hilbert space and let $\mathscr{B} = \{e_\alpha\}_{\alpha\in\mathscr{I}}$ be an orthonormal set. Then for any $x \in \mathscr{H}$ the vector*

$$\widetilde{x} = \sum_{\alpha\in\mathscr{I}} \langle x, e_\alpha\rangle e_\alpha,$$

is well defined, where the above sum is to be understood as the standard infinite sum over the $\alpha-s$ *such that* $\langle x, e_\alpha \rangle \neq 0$ *when* $\mathcal{I}$ *is uncountable. Moreover,*

$$\|\widetilde{x}\|^2 = \sum_{\alpha \in \mathcal{I}} |\langle x, e_\alpha \rangle|^2 \leq \|x\|^2 \tag{1.14}$$

with the same convention on the summation. This inequality is known as (the generalised) Bessel inequality.
Lastly, we have that $x = \sum_{\alpha \in \mathcal{I}} \langle x, e_\alpha \rangle e_\alpha$ *if and only if*

$$\|x\|^2 = \sum_{\alpha \in \mathcal{I}} |\langle x, e_\alpha \rangle|^2 . \tag{1.15}$$

This identity is known as (the generalised) Parseval's identity

At long last we are able to define the notion of an orthonormal basis:

Definition 1.5.29. Let $\mathcal{H}$ be a Hilbert space. We say that a set $\mathcal{B} = \{e_\alpha\}_{\alpha \in \mathcal{I}}$ is an orthonormal basis for $\mathcal{H}$ if $\mathcal{B}$ is orthonormal and every $x \in \mathcal{H}$ satisfies

$$x = \sum_{\alpha \in \mathcal{I}} \langle x, e_\alpha \rangle e_\alpha .$$

Example 1.5.30. For any set $\mathcal{I}$, countable or uncountable, the Hilbert space $\ell_2(\mathcal{I})$ has an orthonormal basis $\mathcal{B} = \{\boldsymbol{e}_\alpha\}_\alpha$ where

$$(\boldsymbol{e}_\alpha)_\gamma = \begin{cases} 1, & \gamma = \alpha, \\ 0, & \gamma \neq \alpha. \end{cases}$$

This is, in a sense, the prototype of *all* Hilbert spaces.

Theorem 1.5.31. *Let* $\mathcal{H}$ *be a Hilbert space and let* $\mathcal{B} = \{e_\alpha\}_{\alpha \in \mathcal{I}}$ *be an orthonormal set in* $\mathcal{H}$*. Then the following are equivalent:*

(i) $\mathcal{B}$ *is an orthonormal basis, i.e. for any* $x \in \mathcal{H}$ *we have that*

$$x = \sum_{\alpha \in \mathcal{I}} \langle x, e_\alpha \rangle e_\alpha .$$

(ii) Parseval's identity is satisfied for every $x \in \mathcal{H}$*, i.e.*

$$\|x\|^2 = \sum_{\alpha \in \mathcal{I}} |\langle x, e_\alpha \rangle|^2 .$$

(iii) $\mathcal{B}^\perp = \{0\}$.
(iv) $\overline{\operatorname{span} \mathcal{B}} = \mathcal{H}$.

REMARK 1.5.32. We say that a set M in a Banach space $\mathcal{X}$ is *total* if $\overline{\operatorname{span} M} = \mathcal{X}$, i.e. if $\operatorname{span} M$ is dense in $\mathcal{X}$. Condition (iv) of the above theorem states that $\mathcal{B} = \{e_\alpha\}_{\alpha \in \mathcal{I}}$ is an orthonormal basis for $\mathcal{H}$ if and only if it is a total orthonormal set.

Parts of Theorem 1.5.31 were proved in Analysis III but we'll include the proof here for completion.

PROOF OF THEOREM 1.5.31. (i) and (ii) are equivalent according to Theorem 1.5.28, and (iii) and (iv) are equivalent due to the fact that

$$\mathcal{H} = \overline{\text{span}\mathcal{B}} \oplus \left(\text{span}\mathcal{B}\right)^{\perp} = \overline{\text{span}\mathcal{B}} \oplus \mathcal{B}^{\perp}$$

according to Lemma 1.4.4 and Remark 1.4.6.
This implies that by showing that (ii) implies (iii) and that (iii) implies (i) we will conclude the proof.
(ii) $\Rightarrow$ (iii): Let $x \in \mathcal{B}^{\perp}$. Using Parseval's identity we find that

$$\|x\|^2 = \sum_{\alpha\in\mathcal{I}} |\langle x, e_\alpha\rangle|^2 = \sum_{\alpha\in\mathcal{I}} 0 = 0.$$

Thus $\mathcal{B}^{\perp} = \{0\}$.
(iii) $\Rightarrow$ (i): Let $x \in \mathcal{H}$ and define

$$\widetilde{x} = \sum_{\alpha\in\mathcal{I}} \langle x, e_\alpha\rangle\, e_\alpha.$$

By the continuity of the inner product we have that for any $\beta \in \mathcal{I}$

$$\left\langle x - \widetilde{x}, e_\beta \right\rangle = \left\langle x, e_\beta \right\rangle - \left\langle \sum_{\alpha\in\mathcal{I}} \langle x, e_\alpha\rangle\, e_\alpha, e_\beta \right\rangle$$

$$\left\langle x, e_\beta \right\rangle - \sum_{\alpha\in\mathcal{I}} \langle x, e_\alpha\rangle \underbrace{\left\langle e_\alpha, e_\beta \right\rangle}_{\delta_{\alpha,\beta}} = \left\langle x, e_\beta \right\rangle - \left\langle x, e_\beta \right\rangle = 0$$

which shows that $x - \widetilde{x} \in \mathcal{B}^{\perp}$. As $\mathcal{B}^{\perp} = \{0\}$ we conclude that

$$x = \widetilde{x} = \sum_{\alpha\in\mathcal{I}} \langle x, e_\alpha\rangle\, e_\alpha,$$

which shows (i). The proof is now complete. □

An immediate corollary of the theorem is the following:

Corollary 1.5.33. Let $\mathcal{H}$ be an infinite dimensional Hilbert space. The following are equivalent:

(i) $\mathcal{H}$ is separable.
(ii) $\mathcal{H}$ has a Schauder basis.
(iii) $\mathcal{H}$ has a countable orthonormal basis $\mathcal{B} = \{e_n\}_{n\in\mathbb{N}}$.

PROOF. Recalling Corollary 1.5.18, Theorem 1.5.23 and the definition of an orthonormal basis we conclude that (ii) implies (i) and that (iii) implies (ii) respectively. We will conclude the proof of the theorem when we'll show that (i) implies (iii).
Let $M = \{x_n\}_{n\in\mathbb{N}}$ be a dense set in $\mathcal{H}$. Using Gram-Schmidt procedure, *which is allowed due to the countability of* M, we find a countable orthonormal set $\mathcal{B} = \{e_n\}_{n\in\mathbb{N}}$ such that

$$\text{span}M = \text{span}\mathcal{B}.$$

Since M is dense and $M \subset \text{span}M$ we conclude that span$M$ is dense and consequently

$$\overline{\text{span}\mathcal{B}} = \overline{\text{span}M} = \mathcal{H}.$$

This shows, according to theorem 1.5.31, that $\mathscr{B}$ is an orthonormal basis for $\mathscr{H}$, completing the proof. $\square$

REMARK 1.5.34. A careful look at the proof of Corollary 1.5.33 gives us a method to find the orthonormal basis for $\mathscr{H}$ in many cases.
If $\mathscr{H} = \overline{\text{span} M}$ where $M = \{x_n\}_{n\in\mathbb{N}}$ is a countable set then by performing the Gram-Schmidt procedure on $\{x_n\}_{n\in\mathbb{N}}$ we'll find that

$$\mathscr{H} = \overline{\text{span} M} = \overline{\text{span}\{x_1, x_2, \ldots,\}} = \overline{\text{span}\mathscr{B}}$$

i.e. we can find an orthonormal basis for $\mathscr{H}$ in this case by performing Gram-Schmidt on the *countable generator of a dense subspace of* $\mathscr{H}$.

There are many other examples that illustrate this process. These examples include the Legendre polynomials in $L^2[-1,1]$, Hermite functions in $L^2(-\infty,\infty)$, and Laguerre functions in $L^2[0,\infty)$.

Example 1.5.35 *(Fourier series in $L^2[-\pi,\pi]$).* Consider the space $L^2[-\pi,\pi]$ with the inner product

$$\langle f, g\rangle_{L^2[-\pi,\pi]} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}dx,$$

and consider the following orthonormal sets

$$\mathscr{B}_{\mathbb{R}} = \left\{1, \sqrt{2}\sin(x), \sqrt{2}\cos(x), \ldots, \sqrt{2}\sin(nx), \sqrt{2}\cos(nx), \ldots\right\}$$

$$\mathscr{B}_{\mathbb{C}} = \left\{e^{inx}\right\}_{n\in\mathbb{Z}}$$

where $\mathscr{B}_{\mathbb{R}}$ is considered when our space is over $\mathbb{R}$ and $\mathscr{B}_{\mathbb{C}}$ is considered when our space is over $\mathbb{C}$.
One can show that the span of $\mathscr{B}_{\mathbb{R}}$ and $\mathscr{B}_{\mathbb{C}}$ are dense in $L^2[-\pi,\pi]$ over $\mathbb{R}$ and $\mathbb{C}$ respectively, and as such they are orthonormal basis for the appropriate Hilbert space. The L^2−series associated to these bases

$$f = \frac{a_0}{2} + \sum_{n\in\mathbb{N}} a_n\cos(nx) + b_n\sin(nx)$$

where

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)dx, \qquad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)dx,$$

and

$$f = \sum_{n\in\mathbb{Z}} c_n e^{inx}$$

where

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx,$$

are known as the *Fourier Series* and *Complex Fourier Series* of a function $f \in L^2[-\pi,\pi]$. These series are so important that we have named the general coefficients in the expansion of a vector $x \in \mathscr{H}$ with respect to an orthonormal basis $\mathscr{B} = \{e_\alpha\}_{\alpha\in\mathscr{I}}$ after them.
These series have also been extensively explored (pointwise convergence, uniform convergence, the convergence of the derivatives), and you have seen some

results about them in Analysis III. You can find more information about them, and the so-called *Fourier Transform*, in the mathematical field known as *Harmonic Analysis*.
It is worth to mention that Fourier series and Parseval's identity give us a powerful tool to compute many series, such as $\sum_{n\in\mathbb{N}} \frac{1}{n^2}$.

We conclude this section with the following important theorem, that emphasise further the strength of having an inner product on a linear space:

Theorem 1.5.36. *Every non-trivial Hilbert space has an orthonormal basis.*

The proof is quite similar to the proof that every linear space has a Hamel basis, and you have seen it in Analysis III.

1.6. Topological difference between finite and infinite dimensional normed spaces

The notion of a topology. When we have studied metric spaces we have seen how important the notion of open sets is and how we can use it to define and test notions of convergence and compactness, amongst other things. This led people to consider a notion of open sets on sets *that do not have a metric setting*. A family of sets on a certain set X, usually denoted by τ, which satisfy the fundamental properties (which we see in open sets in metric spaces)

- Any union of sets in τ is a set in τ,
- Finite intersection of sets in τ is a set in τ,
- X and the empty set $\varnothing$ are in τ,

is known as *a topology of* X. Sets in τ are called *open sets* and the couple (X,τ) is known as a *topological space*. With a topology (i.e. the family of open sets) in hand we can go forward and define/explore notions of convergence, accumulation, continuity, and compactness.

In this last section of our chapter we will consider some of the topological differences , i.e. differences that pertain to convergence, notions of openness and closedness, and compactness, between finite and infinite dimensional normed vector spaces.
We will restrict our topological consideration in this chapter to normed spaces.

Equivalent topologies in normed spaces.

Definition 1.6.1. Let X be a set and let τ_1 and τ_2 be two topologies on X. We say that τ_1 is equivalent to τ_2 if $\tau_1 = \tau_2$. In other words, the topologies are equivalent if any open set in one topology is an open set in the other.

Given two norms on a linear space, $\|\cdot\|_1$ and $\|\cdot\|_2$, what does it mean that the topologies generated by $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent?

Definition 1.6.2. Let $\mathscr{X}$ be a vector space and let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on $\mathscr{X}$. We say that $\|\cdot\|_1$ *and* $\|\cdot\|_2$ *are equivalent* if there exist $c_1, c_2 > 0$ such that for all $x \in \mathscr{X}$

$$\frac{1}{c_2}\|x\|_1 \leq \|x\|_2 \leq c_1\|x\|_1 . \tag{1.16}$$

Theorem 1.6.3. *Let $\mathcal{X}$ be a vector space and let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on $\mathcal{X}$. The topologies generated by $\|\cdot\|_1$ and $\|\cdot\|_2$ on $\mathcal{X}$ are equivalent, i.e. a set is open with respect to the metric induced by $\|\cdot\|_1$ if and only if it is open with respect to the metric induced by $\|\cdot\|_2$, if and only if $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent.*

The proof of the above relies on the following important observation:

Lemma 1.6.4. *Let $\mathcal{X}$ be a vector space and let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on $\mathcal{X}$. Then, there exists $c > 0$ such that for all $x \in \mathcal{X}$*

$$\|x\|_2 \leq c\,\|x\|_1 \tag{1.17}$$

if and only if one (and as such all) of the following equivalent conditions holds:

(i) For any $x \in \mathcal{X}$ with $\|x\|_1 < 1$ we have that $\|x\|_2 < c$.
(ii) For any $x \in \mathcal{X}$ with $\|x\|_1 \leq 1$ we have that $\|x\|_2 \leq c$.
(iii) For any $x \in \mathcal{X}$ with $\|x\|_1 = 1$ we have that $\|x\|_2 \leq c$.

PROOF. The fact that $\|x\|_2 \leq c\,\|x\|_1$ implies (i)-(iii) is immediate. We shall show the converse by showing that

$$\text{(i)} \Rightarrow \text{(ii)} \Rightarrow \text{(iii)} \Rightarrow (1.17)$$

Assume that (i) holds and let $x \in \mathcal{X}$ be such that $\|x\|_1 \leq 1$. For any $\varepsilon > 0$ define $x_\varepsilon = \frac{x}{1+\varepsilon}$. We have that $\|x_\varepsilon\|_1 = \frac{\|x\|_1}{1+\varepsilon} < 1$. Consequently,

$$\frac{\|x\|_2}{1+\varepsilon} = \|x_\varepsilon\|_2 < c$$

or equivalently

$$\|x\|_2 < (1+\varepsilon)\,c.$$

As ε is arbitrary we see that by taking ε to zero we find that $\|x\|_2 \leq c$, which is (ii).
Next, we notice that if $\|x\|_1 = 1$ then $\|x\|_1 \leq 1$ which shows that (ii) implies (iii).
Lastly, we show that (iii) implies (1.17). We start by noticing that if $x = 0$ then (1.17) is satisfied automatically. Thus, we can assume that $x \neq 0$.
For a given $x \neq 0$ we define $y = \frac{x}{\|x\|_1}$. We have that $\|y\|_1 = 1$ and consequently

$$\frac{\|x\|_2}{\|x\|_1} = \|y\|_2 \leq c$$

or equivalently

$$\|x\|_2 \leq c\,\|x\|_1.$$

The proof is thus complete. □

REMARK 1.6.5. A similar proof to the above shows that the condition

$$\|x\|_2 \leq c\,\|x\|_1$$

for some $c > 0$ and all $x \in \mathcal{X}$ is equivalent to any (and all) of the following:

(i) For any $x \in \mathcal{X}$ with $\|x\|_2 > 1$ we have that $\|x\|_1 > \frac{1}{c}$.
(ii) For any $x \in \mathcal{X}$ with $\|x\|_2 \geq 1$ we have that $\|x\|_1 \geq \frac{1}{c}$.
(iii) For any $x \in \mathcal{X}$ with $\|x\|_2 = 1$ we have that $\|x\|_1 \geq \frac{1}{c}$.

REMARK 1.6.6. The proof of Lemma 1.6.4 *relies heavily* on the scaling property of the norm. This observation is extremely important and will be used again in the next section.
We would also like to note that one can rephrase conditions (i)-(iii) in more geometrically. Defining

$$B_r^{\|\cdot\|}(x_0) = \{x \in \mathcal{X} \mid \|x - x_0\| < r\},$$

$$\overline{B}_r^{\|\cdot\|}(x_0) = \overline{B_r^{\|\cdot\|}(x_0)} = \{x \in \mathcal{X} \mid \|x - x_0\| \le r\},$$

and

$$\mathbb{S}_r^{\|\cdot\|}(x_0) = \partial B_r^{\|\cdot\|}(x_0) = \{x \in \mathcal{X} \mid \|x - x_0\| = r\}$$

we have that

$$\text{(i)} \;\Leftrightarrow\; B_1^{\|\cdot\|_1}(0) \subset B_c^{\|\cdot\|_2}(0),$$

$$\text{(ii)} \;\Leftrightarrow\; \overline{B}_1^{\|\cdot\|_1}(0) \subset \overline{B}_c^{\|\cdot\|_2}(0)$$

and

$$\text{(iii)} \;\Leftrightarrow\; \mathbb{S}_1^{\|\cdot\|_1}(0) \subset \overline{B}_c^{\|\cdot\|_2}(0).$$

Using the *linear structure* of the space, we can take the above geometric interpretation one step further:

Lemma 1.6.7. *Let $\mathcal{X}$ be a vector space and let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on $\mathcal{X}$. Then, there exists $c > 0$ such that for all $x \in \mathcal{X}$*

$$\|x\|_2 \le c\,\|x\|_1$$

if and only if for any $x_0 \in \mathcal{X}$ and any $\varepsilon > 0$ we have that

$$B_{\frac{\varepsilon}{c}}^{\|\cdot\|_1}(x_0) \subset B_\varepsilon^{\|\cdot\|_2}(x_0). \tag{1.18}$$

PROOF. We start by reminding ourselves that Lemma 1.6.4 implies that the condition $\|x\|_2 \le c\,\|x\|_1$ for all $x \in \mathcal{X}$ is equivalent to

$$\|x\|_1 < 1 \;\Rightarrow \|x\|_2 < c \qquad \forall x \in \mathcal{X}.$$

Given an $\varepsilon > 0$, we clearly see that the above is equivalent to

$$\|x\|_1 < \varepsilon \;\Rightarrow \|x\|_2 < c\varepsilon \qquad \forall x \in \mathcal{X},$$

due to the scaling of the norm. By replacing x with $x - x_0$ for any given $x_0 \in \mathcal{X}$ we see that the above is equivalent to[3]

$$\|x - x_0\|_1 < \varepsilon \;\Rightarrow \|x - x_0\|_2 < c\varepsilon \qquad \forall x \in \mathcal{X}.$$

Much like in Remark 1.6.6, we now conclude that

$$\|x\|_2 \le c\,\|x\|_1 \;\; \forall x \in \mathcal{X} \;\Leftrightarrow\; B_\varepsilon^{\|\cdot\|_1}(x_0) \subset B_{c\varepsilon}^{\|\cdot\|_2}(x_0) \;\; \forall x_0 \in \mathcal{X},\, \forall \varepsilon > 0.$$

Replacing ε with $\frac{\varepsilon}{c}$ yields the desired result. □

REMARK 1.6.8. Note that choosing $\varepsilon = c$ and $x_0 = 0$ gets us back to Lemma 1.6.4 and Remark 1.6.6. The linear structure of the space and the adherence of the norms to it allowed us to push this condition much more.

[3] choosing $x_0 = 0$ gives us back the original inequality.

We are now ready to consider the proof of Theorem 1.6.3.

SKETCH OF THE PROOF OF THEOREM 1.6.3. We give a sketch of the proof and leave some details to the reader.
Step 1: According to Lemma 1.6.7 we have that $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent if and only if there exists $c_1, c_2 > 0$ such that for any $\varepsilon > 0$ and any $x_0 \in \mathcal{X}$

$$B^{\|\cdot\|_1}_{\frac{\varepsilon}{c_1}}(x_0) \subset B^{\|\cdot\|_2}_{\varepsilon}(x_0) \qquad \text{and} \qquad B^{\|\cdot\|_2}_{\frac{\varepsilon}{c_2}}(x_0) \subset B^{\|\cdot\|_1}_{\varepsilon}(x_0).$$

Step 2: The topologies of the normed spaces $(\mathcal{X}, \|\cdot\|_1)$ and $(\mathcal{X}, \|\cdot\|_2)$ are equivalent if and only if for any $x_0 \in \mathcal{X}$ and any $r > 0$ there exists $r_1(x_0) > 0$ and $r_2(x_0) > 0$ such that

$$B^{\|\cdot\|_1}_{r_1(x_0)}(x_0) \subset B^{\|\cdot\|_2}_{r}(x_0) \qquad \text{and} \qquad B^{\|\cdot\|_2}_{r_2(x_0)}(x_0) \subset B^{\|\cdot\|_1}_{r}(x_0).$$

When the space involved in normed, one can show that $r_1(x_0)$ and $r_2(x_0)$ can be chosen independently of x_0.
Step 3: Combining Steps 1 and 2 gives us the desired result. □

REMARK 1.6.9. Looking at the definition of equivalence of norms, given by condition (1.16), we notice that by defining $\boldsymbol{c} = \max\{c_1, c_2\}$ we find that for any $x \in \mathcal{X}$

$$\frac{1}{\boldsymbol{c}}\|x\|_1 \le \|x\|_2 \le \boldsymbol{c}\|x\|_1. \tag{1.19}$$

Alternatively, if the above holds then $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent with $c_1 = c_2 = \boldsymbol{c}$. Sometimes condition (1.19) is used instead of (1.16) as it is more symmetric (though yields less optimal constants).

The reason why we are quite keen on the notion of equivalence of topologies/norms is expressed in the following theorem

Theorem 1.6.10. *Let $\mathcal{X}$ be a vector space and let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on $\mathcal{X}$. Then if the topologies of the normed spaces $(\mathcal{X}, \|\cdot\|_1)$ and $(\mathcal{X}, \|\cdot\|_2)$ are equivalent we have that:*

(i) The sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathcal{X}$ converges to $x \in \mathcal{X}$ in $(\mathcal{X}, \|\cdot\|_1)$ if and only if it converges to x in $(\mathcal{X}, \|\cdot\|_2)$.
(ii) The sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathcal{X}$ is Cauchy in $(\mathcal{X}, \|\cdot\|_1)$ if and only if it Cauchy in $(\mathcal{X}, \|\cdot\|_2)$.
(iii) $(\mathcal{X}, \|\cdot\|_1)$ is a Banach space if and only if $(\mathcal{X}, \|\cdot\|_2)$ is.
(iv) Given an additional normed space $\mathcal{Y}$ we have that $f : \mathcal{Y} \to (\mathcal{X}, \|\cdot\|_1)$ is continuous if and only if $f : \mathcal{Y} \to (\mathcal{X}, \|\cdot\|_2)$ is.

Equivalence of norms in finite dimensional spaces. One fundamental difference between being a finite dimensional space and an infinite dimensional space is expressed in the following theorem:

Theorem 1.6.11. *Let $\mathcal{X}$ be a finite dimensional vector space. Then any two norms on $\mathcal{X}$, $\|\cdot\|_1$ and $\|\cdot\|_2$, are equivalent.*

PROOF. We begin by noticing that the notion of equivalence of norms is transitive, i.e. if $\|\cdot\|_1$ is equivalent to $\|\cdot\|_2$ and $\|\cdot\|_2$ is equivalent to $\|\cdot\|_3$ then $\|\cdot\|_1$ is equivalent to $\|\cdot\|_3$. Thus, in order to prove that all norms are equivalent on finite dimensional spaces it is enough to find *a norm* on $\mathcal{X}$ to which all other norms are equivalent to. The norm we will choose will be motivated by the standard Euclidean norm on $\mathbb{R}^n$ and $\mathbb{C}^n$.
Since $\mathcal{X}$ is finite dimensional we can find a finite set of vectors $\{e_1,\dots,e_n\}$ that form a basis for it. Every $x\in\mathcal{X}$ can be *uniquely* written as

$$x=\sum_{i=1}^{n}\alpha_i(x)e_i$$

for some scalars $\alpha_1(x),\dots,\alpha_n(x)$. We define

$$\|x\|_{\text{Euclid}}=\sqrt{\sum_{i=1}^{n}|\alpha_i(x)|^2}$$

and claim that it is a norm on $\mathcal{X}$. Indeed, for all $x\in\mathcal{X}$ we have that $\|x\|_{\text{Euclid}}\geq 0$ and

$$\|x\|_{\text{Euclid}}=0 \quad\Leftrightarrow\quad \alpha_1(x)=\cdots=\alpha_n(x)=0 \quad\Leftrightarrow\quad x=0,$$

where we have used the fact that the set $\{e_1,\dots,e_n\}$ is independent. This shows property ***n*** 1 of the norm.
To show ***n*** 2 we use the uniqueness of the expansion coefficients, $\{\alpha_i(x)\}_{i=1,\dots,n}$, to conclude that for any scalar β we have that if $x=\sum_{i=1}^{n}\alpha_i(x)e_i$ then $\alpha_i\left(\beta x\right)=\beta\alpha_i(x)$ and as such

$$\left\|\beta x\right\|_{\text{Euclid}}=\sqrt{\sum_{i=1}^{n}\left|\alpha_i\left(\beta x\right)\right|^2}=\sqrt{\sum_{i=1}^{n}\left|\beta\right|^2|\alpha_i(x)|^2}=\left|\beta\right|\|x\|_{\text{Euclid}}.$$

Similarly, since $\alpha_i(x+y)=\alpha_i(x)+\alpha_i(y)$ the triangle inequality in the Euclidean space shows that

$$\left\|x+y\right\|_{\text{Euclid}}=\sqrt{\sum_{i=1}^{n}\left|\alpha_i(x)+\alpha_i(y)\right|^2}\leq\sqrt{\sum_{i=1}^{n}|\alpha_i(x)|^2}+\sqrt{\sum_{i=1}^{n}\left|\alpha_i(y)\right|^2}=\|x\|_{\text{Euclid}}+\left\|y\right\|_{\text{Euclid}}$$

for any $x,y\in\mathcal{X}$, which is property ***n*** 3. We have thus shown that $\|\cdot\|_{\text{Euclid}}$ is indeed a norm on $\mathcal{X}$.
Next, we will now show that if $\|\cdot\|$ is a norm on $\mathcal{X}$, then $\|\cdot\|$ is equivalent to $\|\cdot\|_{\text{Euclid}}$.
For any $x\in\mathcal{X}$ we have that

$$\|x\|=\left\|\sum_{i=1}^{n}\alpha_i(x)e_i\right\|\leq\sum_{i=1}^{n}|\alpha_i(x)|\,\|e_i\|\leq\sqrt{\sum_{i=1}^{n}|\alpha_i|^2}\sqrt{\sum_{i=1}^{n}\|e_i\|^2},$$

where we have used the triangle inequality of $\|\cdot\|$ and the Cauchy-Schwartz inequality on $\mathbb{R}^n$. Thus, defining $c_1=\sqrt{\sum_{i=1}^{n}\|e_i\|^2}$, we see that

$$\|x\|\leq c_1\|x\|_{\text{Euclid}}, \tag{1.20}$$

which shows one half of the required equivalence. To show the other inequality we recall that according to Remark 1.6.5 it is enough to show that there exists

$c_2 > 0$ such that if $\|x\|_{\text{Euclid}} = 1$ then $\|x\| \geq \frac{1}{c_2}$.
Defining the function $f : (\mathcal{X}, \|\cdot\|_{\text{Euclid}}) \to \mathbb{R}$ by $f(x) = \|x\|$ we find that due to the reverse triangle inequality and (1.20)

$$\left|f(x) - f(y)\right| = \left|\|x\| - \|y\|\right| \leq \left\|x - y\right\| \leq c_1 \left\|x - y\right\|_{\text{Euclid}},$$

which tell us that f is a Lipschitz function, and in particular is continuous with respect to $\|\cdot\|_2$. We notice that what we want to prove is equivalent to saying that

$$f(x) = \|x\| \geq \frac{1}{c_2}, \qquad \text{if} \qquad x \in \mathbb{S}_1^{\|\cdot\|_{\text{Euclid}}}(0) = \{x \in \mathcal{X} \mid \|x\|_{\text{Euclid}} = 1\}$$

Since f is continuous and $f(x) = 0$ if and only if $x = 0 \notin \mathbb{S}_1^{\|\cdot\|_{\text{Euclid}}}(0)$, we conclude that if we'll show that $\mathbb{S}_1^{\|\cdot\|_{\text{Euclid}}}(0)$ is compact then the extreme value theorem will assure us that f has a non-zero minimum on $\mathbb{S}_1^{\|\cdot\|_{\text{Euclid}}}(0)$. This minimum will be our desired $\frac{1}{c_2}$, and the proof will be completed. We turn our attention to proving the compactness of $\mathbb{S}_1^{\|\cdot\|_{\text{Euclid}}}(0)$.
Defining the bijection

$$I : \left(\mathbb{F}^n, \|\cdot\|_2\right) \to (\mathcal{X}, \|\cdot\|_{\text{Euclid}}), \qquad I(\alpha_1, \dots, \alpha_n) = \sum_{i=1}^{n} \alpha_i e_i$$

we see that by the definition of the norms on $\mathbb{F}^n$ and $\mathcal{X}$

$$\left\|I(\alpha_1, \dots, \alpha_n) - I\left(\beta_1, \dots, \beta_n\right)\right\|_{\text{Euclid}} = \left\|(\alpha_1, \dots, \alpha_n) - \left(\beta_1, \dots, \beta_n\right)\right\|_2$$

and as such the function I (and its inverse) is continuous. Since the unit sphere in $\mathbb{F}^n$ is compact and the image of compact sets under continuous functions is compact we conclude the desired compactness of $\mathbb{S}_1^{\|\cdot\|_{\text{Euclid}}}(0)$. □

An immediate consequence of this theorem is the following.

Theorem 1.6.12. *Any finite dimensional normed space $(\mathcal{X}, \|\cdot\|)$ is complete. Consequently, any finite dimensional subspace $\mathcal{M}$ of a Banach space $\mathcal{X}$ is closed.*

The connection between compactness and the dimension. One of the most fundamental (and useful) concept in metric spaces (and the general topic of topology) is that of compactness. A simple criterion for compactness in $\mathbb{F}^n$ is given by the Heine-Borel theorem

Theorem 1.6.13 *(Heine-Borel). A set K in $\mathbb{F}^n$ is compact (with respect to the standard norm) if and only if it is closed and bounded.*

Will the same hold in a general Banach space? The answer to this question is provided in the following theorem:

Theorem 1.6.14. *Let $(\mathcal{X}, \|\cdot\|)$ be a Banach space. Then the following*

(i) A set K is compact in $(\mathcal{X}, \|\cdot\|)$.
(ii) K is a bounded and closed set in $(\mathcal{X}, \|\cdot\|)$.

are equivalent if and only if $\mathcal{X}$ is finite dimensional.

The proof of the theorem, which we will not pursue, relies on two ingredients:

- The Heine-Borel theorem.
- The fact that in an infinite dimensional vector space $\overline{B}_1^{\|\cdot\|}(0)$ has a sequence that has no converging subsequence.

The latter point relies on the following theorem:

Theorem 1.6.15 *(F. Riesz's Lemma). Let $\mathcal{X}$ be a normed space and let $\mathcal{M}$ be a closed subspace of $\mathcal{X}$. If $\mathcal{M} \neq \mathcal{X}$ then for any $\varepsilon \in (0,1)$ there exists $x \in \mathcal{X}$ of norm 1 such that*

$$\inf_{y\in\mathcal{M}} \|x-y\| \geq 1-\varepsilon.$$

If $\mathcal{M}$ is finite dimensional the above can be improved to

$$\inf_{y\in\mathcal{M}} \|x-y\| \geq 1.$$

CHAPTER 2

Linear Operators and Functionals

As we've seen in Linear Algebra I, the notion of *Linear Operators*, which was equivalent to matrices in the finite dimensional case, is an extremely important one. This remains the same in the infinite dimensional case and such maps play a pivotal role in subjects such as Quantum Mechanics and PDEs. In this chapter we will define and explore these maps, which we will call *operators* from this point onwards.

2.1. Basic properties of linear operators

We begin with the definition of what it means to be a linear operator.

Definition 2.1.1. Let $\mathscr{X}$ and $\mathscr{Y}$ be two vector spaces. A linear operator $T : \mathscr{D}(T) \subset \mathscr{X} \to \mathscr{Y}$ is a map defined on a subspace of $\mathscr{X}$, $\mathscr{D}(T)$, such that for any $x, y \in \mathscr{D}(T)$ and any scalar α

$$T\left(x+y\right) = T(x) + T(y), \qquad T(\alpha x) = \alpha T(x). \tag{2.1}$$

When speaking about linear operators it is sometimes customary to write Tx instead of $T(x)$. We shall adapt this convention in many, though not all, cases.

Example 2.1.2 *(The identity operator)*. For any vector space $\mathscr{X}$, the function $I_{\mathscr{X}} : \mathscr{X} \to \mathscr{X}$ defined by $I_{\mathscr{X}} x = x$ is a linear operator.

Example 2.1.3 *(The zero operator)*. For any two vector spaces, $\mathscr{X}$ and $\mathscr{Y}$, the map $0 : \mathscr{X} \to \mathscr{Y}$ defined by $0x = 0_{\mathscr{Y}}$, where $0_{\mathscr{Y}}$ is the additive zero in $\mathscr{Y}$, is a linear operator.

Example 2.1.4 *(Differentiation of polynomials)*. Consider the space of polynomials over $\mathbb{F}$, $\mathscr{P}[x]$. The operator $D : \mathscr{P}[x] \to \mathscr{P}[x]$ defined by

$$Dp = p',$$

is a linear operator.

Example 2.1.5 *(Integration of continuous functions)*. Consider the space $C[a,b]$ and define an operator $T : C[a,b] \to C[a,b]$ by

$$Tf(x) = \int_a^x f(t)dt.$$

Then T is a linear operator. Note that the fact that Tf is indeed continuous when f is continuous follows from the fundamental theorem of calculus (in fact, Tf is differentiable on (a,b)).

Example 2.1.6 *(Multiplication by continuous functions)*. Consider again the space $C[a,b]$ and let m be a given function in $C[a,b]$. The operator $M : C[a,b] \to C[a,b]$ defined by

$$Mf(x) = m(x)f(x)$$

is a linear operator.

Example 2.1.7 *(Matrices)*. Consider the space $\mathbb{F}^n$ and let A be an $m \times n$ matrix. The operator $T_A : \mathbb{F}^n \to \mathbb{R}^m$ defined by

$$T_A x = Ax$$

is a linear operator. As you've seen in Linear Algebra I, this is in fact how all linear maps "look like" in finite dimension.

Definition 2.1.8. Let $\mathscr{X}$ and $\mathscr{Y}$ be two vector spaces and let $T : \mathscr{D}(T) \subset \mathscr{X} \to \mathscr{Y}$ be a linear operator. The *range* (or *image*) of T, $\mathscr{R}(T)$, is defined as

$$\mathscr{R}(T) = \left\{ y \in \mathscr{Y} \;\middle|\; T(x) = y, \text{ for some } x \in \mathscr{D}(T) \right\}.$$

The *null space* (or *kernel*) of T, $\mathscr{N}(T)$, is defined as

$$\mathscr{N}(T) = \{x \in \mathscr{D}(T) \mid Tx = 0\}.$$

The following lemma is proven in exactly the same way as in Linear Algebra I:

Lemma 2.1.9. *Let $\mathscr{X}$ and $\mathscr{Y}$ be two vector spaces and let $T : \mathscr{D}(T) \subset \mathscr{X} \to \mathscr{Y}$ be a linear operator. Then*

(i) $T0 = 0$.
(ii) $\mathscr{R}(T)$ *is a subspace of* $\mathscr{Y}$.
(iii) $\mathscr{N}(T)$ *is a subspace of* $\mathscr{D}(T)$.
(iv) T *is injective (or one to one) if and only if* $\mathscr{N}(T) = \{0\}$.

We'll end this section with an easy criterion for the existence of an inverse to linear operators that are defined in *finite dimensional spaces*, and another property of inverses

Theorem 2.1.10. *Let $\mathscr{X}$ and $\mathscr{Y}$ be two vectors spaces and let $T : \mathscr{D}(T) \subset \mathscr{X} \to \mathscr{Y}$ be a linear operator. Then*

(i) *If* $\dim \mathscr{D}(T) < \infty$ *then* $\dim \mathscr{R}(T) \leq \dim \mathscr{D}(T) < \infty$.
(ii) *If* $\dim \mathscr{D}(T) < \infty$ *then the following are equivalent:*
  - T *is invertible.*
  - $\dim \mathscr{D}(T) = \dim \mathscr{R}(T)$.

  In particular, if $\dim \mathscr{D}(T) = \dim \mathscr{Y}$ *then* $\mathscr{R}(T) = \mathscr{Y}$ *and* T *is invertible if and only if* $\mathscr{N}(T) = \{0\}$.
(iii) *If* T *is invertible and* $S : \mathscr{R}(T) \to \mathscr{Z}$, *where* $\mathscr{Z}$ *is another vector space, is an invertible linear operator, then* $S \circ T : \mathscr{D}(T) \to \mathscr{R}(S)$ *is an invertible operator and*

$$(S \circ T)^{-1} = T^{-1} \circ S^{-1}.$$

We won't prove any of the above statements as they have been shown in Linear Algebra I. We will just say that part (i) and (ii) are immediate consequences (and in a sense are a part of the proof) of the *Rank-Nullity Theorem.*

2.2. Continuity of linear operators and the notion of boundedness

As we've seen in our previous chapter, Functional Analysis mixes linear structures with analytic concepts. Linear operators are not exempt from this treatment. It won't strike us as a surprise, then, that the first thing we'd like to explore with regards to these operators is the question of their continuity.

The linear structure of the space and the adherence of linear operators to it results in the following:

Theorem 2.2.1. *Let $\mathscr{X}$ and $\mathscr{Y}$ be two normed spaces and let $T : \mathscr{D}(T) \subset \mathscr{X} \to \mathscr{Y}$ be a linear operator. Then the following are equivalent:*

(i) T is continuous at $0 \in \mathscr{D}(T)$.
(ii) T is continuous at some $x_0 \in \mathscr{D}(T)$.
(iii) T is continuous.

PROOF. As (iii) implies (i), it would be sufficient for us to show that (i) $\Leftrightarrow$ (ii) and (i) $\Rightarrow$ (iii) to conclude the proof.
We start by assuming that (i) holds. Let $x_0 \in \mathscr{D}(T)$ be arbitrary and let $\{x_n\}_{n\in\mathbb{N}} \subset \mathscr{D}(T)$ be a sequence that converges to x_0. Then $\{x_n - x_0\}_{n\in\mathbb{N}}$ is a sequence in $\mathscr{D}(T)$ (which is a subspace) that converges to 0. As T is continuous at 0 we have that

$$Tx_n - Tx_0 = T(x_n - x_0) \underset{n\to\infty}{\longrightarrow} T0 = 0$$

or equivalently

$$Tx_n \underset{n\to\infty}{\longrightarrow} Tx_0.$$

Note that as $x_0 \in \mathscr{D}(T)$ was arbitrary, we have proved that (i) implies (iii) which implies (ii).
Assume now that (ii) holds and let $\{x_n\}_{n\in\mathbb{N}} \subset \mathscr{D}(T)$ be a sequence that converges to 0. Then, following the same ideas as above, the sequence $\{x_n + x_0\} \subset \mathscr{D}(T)$ converges to x_0 and due to the continuity of T at x_0 we have that

$$Tx_n + Tx_0 = T(x_n + x_0) \underset{n\to\infty}{\longrightarrow} Tx_0.$$

Consequently

$$Tx_n \underset{n\to\infty}{\longrightarrow} 0 = T0,$$

which shows (i). The proof is now complete. □

Looking at the proof above we notice that we have only used two properties of the linearity of the operator:

- The fact that $\mathscr{D}(T)$ is a subspace.
- The fact that for any $x, y \in \mathscr{D}(T)$ we have that $T(x + y) = Tx + Ty$.

The last property of linear operators, the adherence to scaling, would give us another, extremely useful, property. The next theorem is motivated from our discussion about equivalence of norms, in particular Lemma 1.6.4, and the fact that for any linear operator T, the function $x \to \|Tx\|$ satisfies the scaling property of a norm.

Theorem 2.2.2. *Let $\mathscr{X}$ and $\mathscr{Y}$ be two normed spaces and let $T : \mathscr{D}(T) \subset \mathscr{X} \to \mathscr{Y}$ be a linear operator. Then T is continuous at 0 if and only if there exists $C > 0$ such that for all $x \in \mathscr{D}(T)$ we have that*

$$\|Tx\| \leq C\|x\|.$$

PROOF. If such a C exists, then whenever $x \in \mathscr{D}(T)$ and $\|x\| < \delta$ we have that

$$\|Tx - T0\| = \|Tx\| \leq C\|x\| < C\delta.$$

Thus, for a given $\varepsilon > 0$ we have that for $\delta = \frac{\varepsilon}{C}$, $\|Tx\| < \varepsilon$ whenever $x \in \mathscr{D}(T)$ and $\|x\| < \delta$, which shows the desired continuity at 0.
To show the converse, we notice that since

$$\|T(\alpha x)\| = \|\alpha Tx\| = |\alpha|\,\|Tx\|$$

for any $x \in \mathscr{D}(T)$ and a scalar α, it is enough to show that there exist $\delta > 0$ and $\varepsilon > 0$ such that $\|Tx\| \leq \varepsilon$ when $\|x\| \leq \delta$. Indeed, if that is the case then for any $0 \neq x \in \mathscr{D}(T)$

$$\left\|T\left(\frac{\delta x}{\|x\|}\right)\right\| \leq \varepsilon.$$

Since

$$\left\|T\left(\frac{\delta x}{\|x\|}\right)\right\| = \frac{\delta\|Tx\|}{\|x\|}.$$

the above implies that

$$\|Tx\| \leq \frac{\varepsilon}{\delta}\|x\|,$$

which also clearly holds when $x = 0$, giving us the desired inequality for T with $C = \frac{\varepsilon}{\delta}$. Since $T0 = 0$ finding such $\delta > 0$ and $\varepsilon > 0$ follows immediately from continuity at 0 (in fact for *any* $\varepsilon > 0$ we can find an appropriate $\delta > 0$), and we conclude the proof of our claim. □

Theorem 2.2.2 motivates the next definition:

Definition 2.2.3. Let $\mathscr{X}$ and $\mathscr{Y}$ be two normed spaces and let $T : \mathscr{D}(T) \subset \mathscr{X} \to \mathscr{Y}$ be a linear operator. We say that T is a *bounded operator* if there exists $C > 0$ such that for any $x \in \mathscr{D}(T)$ we have that

$$\|Tx\| \leq C\|x\|. \tag{2.2}$$

Our study of the continuity of linear operators in normed spaces culminates in the next theorem:

Theorem 2.2.4. *Let $\mathscr{X}$ and $\mathscr{Y}$ be two normed spaces and let $T : \mathscr{D}(T) \subset \mathscr{X} \to \mathscr{Y}$ be a linear operator. Then the following are equivalent:*

(i) T is continuous at $x_0 \in \mathscr{D}(T)$.

(ii) *T is continuous at* $0 \in \mathscr{D}(T)$.
(iii) *T is continuous.*
(iv) *T is bounded.*
(v) $\sup_{x\in\mathscr{D}(T),\, x\neq 0} \frac{\|Tx\|}{\|x\|} < \infty$.
(vi) *T is a Lipschitz on $\mathscr{D}(T)$, i.e. there exists $K > 0$ such that*

$$\left\|Tx - Ty\right\| \leq K \left\|x - y\right\|$$

for all $x, y \in \mathscr{D}(T)$.

It is conventional to refer to a linear operator that satisfies one, and as such all, of the above criteria as a bounded (linear) operator.

PROOF. Parts (i) to (iv) are equivalent due to theorems 2.2.1 and 2.2.2. The equivalence of (v) and (vi) to (iv) is immediate by the definition of the boundedness of T and its linearity. Indeed, if (iv) holds then there exists $C > 0$ such that for any $x \in \mathscr{D}(T)$ we have that $\|Tx\| \leq C\|x\|$. As such

$$\sup_{x\in\mathscr{D}(T),\, x\neq 0} \frac{\|Tx\|}{\|x\|} \leq C < \infty.$$

Moreover, for any $x, y \in \mathscr{D}(T)$ we have that $x - y \in \mathscr{D}(T)$ and

$$\left\|Tx - Ty\right\| = \left\|T\left(x - y\right)\right\| \leq C\left\|x - y\right\|$$

which shows that (v) and (vi) hold.
Conversely, if $C_\infty = \sup_{x\in\mathscr{D}(T),\, x\neq 0} \frac{\|Tx\|}{\|x\|} < \infty$ we find, by definition, that for any $x \neq 0$ in $\mathscr{D}(T)$

$$\frac{\|Tx\|}{\|x\|} \leq C_\infty$$

which implies that $\|Tx\| \leq C_\infty \|x\|$ for any $0 \neq x \in \mathscr{D}(T)$. As this inequality is trivial when $x = 0$ we see that (v) implies (iv). Lastly, if (vi)holds then by choosing $y = 0$ we see that

$$\|Tx\| = \|Tx - T0\| \leq K\|x - 0\| = K\|x\|,$$

which shows the validity of (iv). The proof is now complete. □

Let us consider a few examples.

Example 2.2.5 *(The identity operator is bounded).* For any $x \in \mathscr{X}$ we have that

$$\|I_{\mathscr{X}} x\| = \|x\|$$

and as such $I_{\mathscr{X}}$ is bounded.

Example 2.2.6 *(The zero operator is bounded).* For any $x \in \mathscr{X}$ we have that

$$\|0x\| = \left\|0_{\mathscr{Y}}\right\| = 0 \leq \|x\|$$

and as such the zero operator is bounded.

Example 2.2.7 *(Differentiation of polynomials is not bounded).* Consider the derivative operator $D : (\mathscr{P}[x], \|\cdot\|_\infty) \to (C[0,1], \|\cdot\|_\infty)$. For every $n \in \mathbb{N}$ we have that

$$D\left(x^n\right) = nx^{n-1}.$$

Since $\|x^n\|_\infty = \left\|x^{n-1}\right\|_\infty = 1$ for any $n \in \mathbb{N}$ we see that

$$\sup_{p\in\mathscr{P}[x],\ p\neq 0} \frac{\left\|Dp\right\|_\infty}{\left\|p\right\|_\infty} \geq \sup_{n\in\mathbb{N}} \frac{\|D(x^n)\|_\infty}{\left\|x^{n-1}\right\|_\infty} = \sup_{n\in\mathbb{N}} n = \infty.$$

We conclude, then, that D is *not* a bounded operator.

Example 2.2.8 *(Matrices)*. The matrix induced operator $T_A : (\mathbb{R}^n, \|\cdot\|) \to \left(\mathbb{R}^m, \widetilde{\|\cdot\|}\right)$, defined by

$$T_A x = Ax$$

where A is an $m \times n$ matrix, is a bounded operator.

The fact that the linear operator T_A iis bounded is not too shocking. This will *always* happen when we have a linear operator whose domain is finite dimensional:

Theorem 2.2.9. *Let $\mathscr{X}$ and $\mathscr{Y}$ be two normed spaces and let $T : \mathscr{D}(T) \subset \mathscr{X} \to \mathscr{Y}$ be a linear operator. If $\dim \mathscr{D}(T) < \infty$ then T is bounded.*

An additional useful property of bounded operators is the following:

Theorem 2.2.10. *Let $\mathscr{X}$ and $\mathscr{Y}$ be two normed spaces and let $T : \mathscr{D}(T) \subset \mathscr{X} \to \mathscr{Y}$ be a bounded linear operator. If $\mathscr{D}(T)$ is closed then $\mathscr{N}(T)$ is a closed subspace of $\mathscr{D}(T)$.*

PROOF. From Lemma 2.1.9 we known that $\mathscr{N}(T)$ is a subspace of $\mathscr{D}(T)$. To show that it is closed we notice that if $\{x_n\}_{n\in\mathbb{N}} \subset \mathscr{N}(T)$ converges to a point x then since $\mathscr{D}(T)$ is closed we have that $x \in \mathscr{D}(T)$ and since T is continuous

$$0 = Tx_n \underset{n\to\infty}{\longrightarrow} Tx.$$

Thus $Tx = 0$, or equivalently $x \in \mathscr{N}(T)$. □

REMARK 2.2.11. While the null space of a bounded operator is always closed when its domain is, the topological properties of its range are less certain.
Example for an operator with closed range: Since any finite dimension vector space is complete, and as such closed according to Theorem 1.6.12, we conclude that any bounded linear operator with finite dimensional range has a closed range. For example: Let $\mathscr{H}$ be an inner product space and let $\{e_1, \dots, e_n\}$ be an orthonormal set in $\mathscr{H}$. The operator $T : \mathscr{H} \to \mathscr{H}$ defined by $Tx = \sum_{i=1}^n \langle x, e_i\rangle e_i$ is linear and its range is

$$\mathscr{R}(T) = \operatorname{span}\{e_1, \dots, e_n\} = \overline{\operatorname{span}\{e_1, \dots, e_n\}}.$$

Example for an operator whose range is not closed: Consider the space $\ell_1(\mathbb{N})$ with its standard norm and the operator $T : \ell_1(\mathbb{N}) \to \ell_1(\mathbb{N})$ given by

$$T\boldsymbol{a} = \left(a_1, \frac{a_2}{2}, \dots, \frac{a_n}{n}, \dots\right).$$

We have that

$$\|T\boldsymbol{a}\|_1 = \sum_{n\in\mathbb{N}} \left|\frac{a_n}{n}\right| \leq \sum_{n\in\mathbb{N}} |a_n| = \|\boldsymbol{a}\|_1,$$

which both shows that T is well defined and that it is bounded (its linearity is also straight forward to show, but we won't do it here). Denoting by

$$\boldsymbol{a}_n = \left(1, \frac{1}{2}, \dots, \frac{1}{n}, 0, 0, \dots\right), \qquad \boldsymbol{b}_n = \left(1, \frac{1}{4}, \dots, \frac{1}{n^2}, 0, 0, \dots\right),$$

$$\boldsymbol{a} = \left(1, \frac{1}{2}, \dots, \frac{1}{n}, \frac{1}{n+1}, \dots\right), \qquad \boldsymbol{b} = \left(1, \frac{1}{4}, \dots, \frac{1}{n^2}, \frac{1}{(n+1)^2}, \dots\right),$$

we notice that $\{\boldsymbol{a}_n\}_{n\in\mathbb{N}}, \{\boldsymbol{b}_n\}_{n\in\mathbb{N}}$ and $\boldsymbol{b}$ are all in $\ell_1(\mathbb{N})$ but $\boldsymbol{a} \notin \ell_1(\mathbb{N})$. Moreover

$$T(\boldsymbol{a}_n) = \boldsymbol{b}_n \underset{n\to\infty}{\longrightarrow} \boldsymbol{b},$$

which shows that $\boldsymbol{b} \in \overline{\mathscr{R}(T)}$. $\boldsymbol{b}$, however, can't be in $\mathscr{R}(T)$. Indeed had there been $\boldsymbol{x} \in \ell_1(\mathbb{N})$ such that $T\boldsymbol{x} = \boldsymbol{b}$ then we must have had that $\frac{x_n}{n} = b_n$ for all $n \in \mathbb{N}$, which would have meant that $\boldsymbol{x} = \boldsymbol{a} \notin \ell_1(\mathbb{N})$. Thus, $\mathscr{R}(T)$ is not closed.

We conclude this section with an important theorem about the ability to extend bounded operators from $\mathscr{D}(T)$ to its closure (where we can always apply Theorem 2.2.10):

Theorem 2.2.12. *Let $\mathscr{X}$ and $\mathscr{Y}$ be two normed spaces and let $T : \mathscr{D}(T) \subset \mathscr{X} \to \mathscr{Y}$ be a bounded linear operator. If $\mathscr{Y}$ is a Banach space, then T can be uniquely extended to $\overline{\mathscr{D}(T)}$, i.e. there exists a unique linear bounded operator $\widetilde{T} : \overline{\mathscr{D}(T)} \subset \mathscr{X} \to \mathscr{Y}$ such that $\widetilde{T}|_{\mathscr{D}(T)} = T$. Moreover*

$$\sup_{x\in\overline{\mathscr{D}(T)},\, x\neq 0} \frac{\|\widetilde{T}x\|}{\|x\|} = \sup_{x\in\mathscr{D}(T),\, x\neq 0} \frac{\|Tx\|}{\|x\|}. \tag{2.3}$$

PROOF. We will sketch the proof here and leave some details to the reader.
Step 1: For any $x \in \overline{\mathscr{D}(T)} \setminus \mathscr{D}(T)$ we find a sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathscr{D}(T)$ that converges to x. Since

$$\|Tx_n - Tx_m\| \leq C\|x_n - x_m\|$$

for some $C > 0$, we see that $\{Tx_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathscr{Y}$. As $\mathscr{Y}$ is complete there exists $y \in \mathscr{Y}$ such that $Tx_n \underset{n\to\infty}{\longrightarrow} y$. It can be shown that y is independent of the choice of sequence $\{x_n\}_{n\in\mathbb{N}} \in \mathscr{D}(T)$ that converged to x and as such we can define

$$\widetilde{T}x = \begin{cases} Tx, & x \in \mathscr{D}(T), \\ \lim_{n\to\infty} Tx_n, & x \in \overline{\mathscr{D}(T)} \setminus \mathscr{D}(T),\ \{x_n\}_{n\in\mathbb{N}} \subset \mathscr{D}(T) \text{ converges to } x. \end{cases}$$

Step 2: We show that $\widetilde{T}$ is linear.
Step 3: Identity (2.3), which follows from the fact that a supermum over a set equals the supremum over its closure, holds and Theorem 2.2.4 imply the boundedness of $\widetilde{T}$.
Step 4: The uniqueness of the extension follows from the fact that it is continuous on $\overline{\mathscr{D}(T)}$ and equals to T on $\mathscr{D}(T)$. □

2.3. The space of bounded linear operators $B(\mathscr{X},\mathscr{Y})$

The linear properties of linear operators and the linear structure on which they are defined allows us to define a linear structure on the set of linear operators:

Definition 2.3.1. Let $\mathscr{X}$ and $\mathscr{Y}$ be two vector spaces over the same field. We denote by $L(\mathscr{X},\mathscr{Y})$ the set of all linear operators from $\mathscr{X}$ into $\mathscr{Y}$ and define two operation in $L(\mathscr{X},\mathscr{Y})$:
The *operator additions* of two elements $T,S \in L(\mathscr{X},\mathscr{Y})$, denoted by $T+S$, is the operator $T+S:\mathscr{X}\to\mathscr{Y}$ defined by

$$(T+S)\,x = Tx + Sx.$$

The *scalar multiplication of an operator* $T \in L(\mathscr{X},\mathscr{Y})$ by a scalar α, denoted by αT, is the operator $\alpha T:\mathscr{X}\to\mathscr{Y}$ defined by

$$(\alpha T)\,x = \alpha \cdot Tx.$$

The following has been shown in Linear Algebra I:

Theorem 2.3.2. *Let $\mathscr{X}$ and $\mathscr{Y}$ be two vector spaces over a field $\mathbb{F}$. Then $L(\mathscr{X},\mathscr{Y})$ is a vector space with the operator addition and scalar multiplication of an operator. The additive zero is the zero operator and the additive inverse of a linear operator $T \in L(\mathscr{X},\mathscr{Y})$ is the operator $-T = (-1)\cdot T$.*

$L(\mathscr{X},\mathscr{Y})$ is not the space we are truly interested in as it doesn't necessarily adhere to the topology of our normed spaces. We are, however, interested in its following subset:

Definition 2.3.3. Let $\mathscr{X}$ and $\mathscr{Y}$ be two normed spaces over the same field. We denote by $B(\mathscr{X},\mathscr{Y})$ the set of all *bounded* linear operators from $\mathscr{X}$ into $\mathscr{Y}$. In particular, the domain of any operator in $B(\mathscr{X},\mathscr{Y})$ is $\mathscr{X}$.

REMARK 2.3.4. It is important to mention that not only are *unbounded operators* extremely interesting in their own right, but they appear in many of the applications of Functional Analysis. A lot of their study, however, is influenced by the study of bounded operators (which also provide some interesting contrast at many point) so our efforts in studying such operators will be far from wasted.

As $B(\mathscr{X},\mathscr{Y})$ is a subset of $L(\mathscr{X},\mathscr{Y})$ it is automatically imbued with operations of addition and scalar multiplication. The natural question about $B(\mathscr{X},\mathscr{Y})$'s linear structure is answered in the next theorem.

Theorem 2.3.5. *Let $\mathscr{X}$ and $\mathscr{Y}$ be two normed spaces over the same field. Then $B(\mathscr{X},\mathscr{Y})$ is a vector space under the addition of operators and scalar multiplication of an operator operations.*

PROOF. As $B(\mathscr{X},\mathscr{Y})$ is subset of a vector space, it is enough for us to check that

- $B(\mathscr{X},\mathscr{Y})$ is not empty.
- $B(\mathscr{X},\mathscr{Y})$ is closed under addition.

- $B(\mathscr{X},\mathscr{Y})$ is closed under scalar multiplication.

As the zero operator is a bounded operator we conclude that $B(\mathscr{X},\mathscr{Y})$ is never empty. Next we'll consider the closure under addition and scalar multiplication. Let T, S be in $B(\mathscr{X},\mathscr{Y})$. Then there exist $C_1, C_2 > 0$ such that

$$\|Tx\| \leq C_1 \|x\|, \qquad \|Sx\| \leq C_2 \|x\|.$$

Using the triangle inequality, we find that

$$\|(T+S)x\| = \|Tx + Sx\| \leq \|Tx\| + \|Sx\| \leq (C_1 + C_2)\|x\|,$$

which shows that $T+S$ is also a bounded operator, i.e. $T+S \in B(\mathscr{X},\mathscr{Y})$.
Moreover, if α is a scalar and $T \in B(\mathscr{X},\mathscr{Y})$ then

$$\|(\alpha T)x\| = \|\alpha \cdot Tx\| = |\alpha|\,\|Tx\| \leq (|\alpha| C)\|x\|,$$

where $C > 0$ is a constant such that $\|Tx\| \leq C\|x\|$ for all $x \in \mathscr{X}$, showing that $\alpha T \in B(\mathscr{X},\mathscr{Y})$.
We conclude that as $B(\mathscr{X},\mathscr{Y})$ is closed under addition and scalar multiplication, and contains the zero vector, it must be a subspace. □

$B(\mathscr{X},\mathscr{Y})$ is more than a mere subspace - it is in fact a *normed space*.

Theorem 2.3.6. *The function* $\|\cdot\| : B(\mathscr{X},\mathscr{Y}) \to \mathbb{R}_+$ *defined by*

$$\|T\| = \inf\{C > 0 \mid \|Tx\| \leq C\|x\| \ \ \forall x \in \mathscr{X}\}$$

is a norm on $B(\mathscr{X},\mathscr{Y})$. *Moreover,*

$$\|T\| = \sup_{x\neq 0} \frac{\|Tx\|}{\|x\|} = \sup_{\|x\|=1} \|Tx\|. \tag{2.4}$$

Identity (2.4) lies in the heart of the proof of the above theorem, and is important in its own right. We will start by proving a generalisation of it which holds even if the domain of our operator is not the entire space.

Lemma 2.3.7. *Let* $\mathscr{X}$ *and* $\mathscr{Y}$ *be two normed spaces over the same field, and let* $T : \mathscr{D}(T) \subset \mathscr{X} \to \mathscr{Y}$ *be a bounded operator. Then*

$$\|T\| = \sup_{x\in\mathscr{D}(T),\, x\neq 0} \frac{\|Tx\|}{\|x\|} = \sup_{x\in\mathscr{D}(T),\, \|x\|=1} \|Tx\| \tag{2.5}$$

where

$$\|T\| = \inf\left\{C > 0 \,\middle|\, \|Tx\| \leq C\|x\| \ \ \forall x \in \mathscr{D}(T)\right\}. \tag{2.6}$$

PROOF. We have seen in the proof of Theorem 2.2.4 that if T is bounded then

$$\sup_{x\in\mathscr{D}(T),\, x\neq 0} \frac{\|Tx\|}{\|x\|} \leq C < \infty,$$

meaning the middle term in (2.5) is well defined and finite. Moreover

$$\sup_{x\in\mathscr{D}(T),\, \|x\|=1} \|Tx\| = \sup_{x\in\mathscr{D}(T),\, \|x\|=1} \frac{\|Tx\|}{\|x\|} \leq \sup_{x\in\mathscr{D}(T),\, x\neq 0} \frac{\|Tx\|}{\|x\|} < \infty. \tag{2.7}$$

Additionally, since for any $x \neq 0$

$$\frac{\|Tx\|}{\|x\|} = \left\| \frac{1}{\|x\|} \cdot Tx \right\| = \left\| T \left(\underbrace{\frac{x}{\|x\|}}_{\text{unit norm}} \right) \right\|$$

we find that for any $x \neq 0$

$$\frac{\|Tx\|}{\|x\|} \leq \sup_{x \in \mathscr{D}(T),\, \|x\|=1} \|Tx\|,$$

and taking supremum on the above we conclude that

$$\sup_{x \in \mathscr{D}(T),\, x \neq 0} \frac{\|Tx\|}{\|x\|} \leq \sup_{x \in \mathscr{D}(T),\, \|x\|=1} \|Tx\| \tag{2.8}$$

Combining (2.7) and (2.8) gives us

$$\sup_{x \in \mathscr{D}(T),\, x \neq 0} \frac{\|Tx\|}{\|x\|} = \sup_{x \in \mathscr{D}(T),\, \|x\|=1} \|Tx\|.$$

Thus, to conclude our proof, we are only left with showing that

$$\|T\| = \sup_{x \in \mathscr{D}(T),\, x \neq 0} \frac{\|Tx\|}{\|x\|}.$$

As we've mentioned above, if $C > 0$ satisfies $\|Tx\| \leq C\|x\|$ for any $x \in \mathscr{D}(T)$ then

$$\sup_{x \in \mathscr{D}(T),\, x \neq 0} \frac{\|Tx\|}{\|x\|} \leq C,$$

and by taking infimum over such $C-$s we find that

$$\sup_{x \in \mathscr{D}(T),\, x \neq 0} \frac{\|Tx\|}{\|x\|} \leq \|T\|.$$

Conversely, by definition, for any $x \in \mathscr{D}(T)$ that is not the zero vector we have that

$$\|Tx\| = \frac{\|Tx\|}{\|x\|} \cdot \|x\| \leq \left(\sup_{x \in \mathscr{D}(T),\, x \neq 0} \frac{\|Tx\|}{\|x\|} \right) \|x\|,$$

and consequently if $\sup_{x \in \mathscr{D}(T),\, x \neq 0} \frac{\|Tx\|}{\|x\|} > 0$ then

$$\sup_{x \in \mathscr{D}(T),\, x \neq 0} \frac{\|Tx\|}{\|x\|} \in \{C > 0 \mid \|Tx\| \leq C\|x\| \ \ \forall x \in \mathscr{X}\}$$

which implies that

$$\|T\| \leq \sup_{x \in \mathscr{D}(T),\, x \neq 0} \frac{\|Tx\|}{\|x\|},$$

If, on the other hand, $\sup_{x \in \mathscr{D}(T),\, x \neq 0} \frac{\|Tx\|}{\|x\|} = 0$, then for any $x \in \mathscr{D}(T)$ such that $x \neq 0$ we have that

$$0 \leq \|Tx\| \leq \left(\sup_{x \in \mathscr{D}(T),\, x \neq 0} \frac{\|Tx\|}{\|x\|} \right) \|x\| = 0,$$

i.e. $Tx=0$. As $T0=0$ we conclude that $Tx=0$ for any $x\in\mathscr{D}(T)$, or that T is the zero operator. Thus, for any $\varepsilon>0$ we find that

$$0=\|Tx\|\leq\varepsilon\|x\|,\qquad\forall x\in\mathscr{D}(T).$$

This means that

$$\varepsilon\in\{C>0\,|\,\|Tx\|\leq C\|x\|\ \ \forall x\in\mathscr{X}\}$$

for any $\varepsilon>0$ which implies, that

$$0\leq\|T\|=\inf\{C>0\,|\,\|Tx\|\leq C\|x\|\ \ \forall x\in\mathscr{X}\}=0.$$

Thus, again,

$$\|T\|\leq\sup_{x\in\mathscr{D}(T),\,x\neq0}\frac{\|Tx\|}{\|x\|},$$

and the proof is complete. □

PROOF OF THEOREM 2.3.6. We have shown the validity of (2.4) in Lemma 2.3.7 and as such we're only left to show that $\|\cdot\|$ is indeed a norm.
By its definition, $\|T\|\geq0$ for any $T\in B(\mathscr{X},\mathscr{Y})$. We have seen in the proof of Lemma 2.3.7 that if

$$\|T\|=\sup_{x\neq0}\frac{\|Tx\|}{\|x\|}=0$$

then T is the zero operator. Thus property $\boldsymbol{n}1$ is shown.
Next we consider the scaling property $\boldsymbol{n}2$: Using (2.4) again we see that for any scalar α

$$\|\alpha T\|=\sup_{x\neq0}\frac{\|\alpha\cdot Tx\|}{\|x\|}=\sup_{x\neq0}\frac{|\alpha|\,\|Tx\|}{\|x\|}=|\alpha|\sup_{x\neq0}\frac{\|Tx\|}{\|x\|}=|\alpha|\,\|T\|.$$

Lastly, we will show that the triangle inequality, property $\boldsymbol{n}3$, holds - again with the help of (2.4). For any $T,S\in B(\mathscr{X},\mathscr{Y})$ we have that for any $x\neq0$

$$\frac{\|(T+S)x\|}{\|x\|}=\frac{\|Tx+Sx\|}{\|x\|}\leq\frac{\|Tx\|+\|Sx\|}{\|x\|}=\frac{\|Tx\|}{\|x\|}+\frac{\|Sx\|}{\|x\|}$$
$$\leq\sup_{x\neq0}\frac{\|Tx\|}{\|x\|}+\sup_{x\neq0}\frac{\|Sx\|}{\|x\|}=\|T\|+\|S\|.$$

Taking supremum on the left hand side of the above yields

$$\|T+S\|=\sup_{x\neq0}\frac{\|(T+S)x\|}{\|x\|}\leq\|T\|+\|S\|,$$

which concludes the proof. □

REMARK 2.3.8. It is also worth to mention the following consequences of Lemma 2.3.7 and Theorem 2.3.6:

-
- $\|T\|=\inf\left\{C>0\,\middle|\,\|Tx\|\leq C\|x\|\ \ \forall x\in\mathscr{D}(T)\right\}=\min\left\{C\geq0\,\middle|\,\|Tx\|\leq C\|x\|\ \ \forall x\in\mathscr{D}(T)\right\}$. Note that to move from the infimum to the minimum we need to include the case $C=0$ which corresponds to $T=0$.
- For any $x\in\mathscr{D}(T)$ we have that $\|Tx\|\leq\|T\|\,\|x\|$.
- If $\|Tx\|\leq C\|x\|$ for all $x\in\mathscr{D}(T)$ then $\|T\|\leq C$.

- For any $x_0 \in \mathscr{D}(T)$ such that $x_0 \neq 0$ we have that $\|T\| \geq \frac{\|Tx_0\|}{\|x_0\|}$. Alternatively, for any $x_0 \in \mathscr{D}(T)$ such that $\|x_0\| = 1$ we have that $\|T\| \geq \|Tx_0\|$.

Besides being useful in showing that $\|T\|$ is indeed a norm, identities (2.4) and (2.5) are a powerful *computational tool* that allows us to actually compute $\|T\|$. It will be the main (if not sole) tool we will use in this part of the module to do so. Indeed, from the above observation we note that to compute $\|T\|$ it is enough to find $C > 0$ such that

$$\|Tx\| \leq C\|x\|, \quad \forall x \in \mathscr{D}(T)$$

and

$$\|Tx_0\| = C, \quad \text{for some } x_0 \in \mathscr{D}(T) \text{ such that } \|x_0\| = 1.$$

The second condition can be replaced by finding a sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathscr{D}(T)$ such that $\|x_n\| = 1$ for all $n \in \mathbb{N}$ and $\lim_{n\to\infty}\|Tx_n\| = C$.
We finish this remark by mentioning that we've seen that any bounded linear operator T can be extended from its domain $\mathscr{D}(T)$ to its closure. This extension, $\widetilde{T}$, satisfied

$$\sup_{x\in\overline{\mathscr{D}(T)},\, x\neq 0} \frac{\left\|\widetilde{T}x\right\|}{\|x\|} = \sup_{x\in\mathscr{D}(T),\, x\neq 0} \frac{\|Tx\|}{\|x\|}$$

which implies that the the extension has the same norm as the original operator, i.e. that $\left\|\widetilde{T}\right\| = \|T\|$.

Let us consider a few examples:

Example 2.3.9 *(Norm of the identity operator)*. The identity operator satisfies

$$\|I_{\mathscr{X}}\| = 1.$$

Example 2.3.10 *(Norm of the zero operator)*. The zero operator satisfies

$$\|0\| = 0.$$

Example 2.3.11 *(Norm of integration operator)*. The integration operator $T : (C[a,b], \|\cdot\|_\infty) \to (C[a,b], \|\cdot\|_\infty)$ defined by

$$Tf(x) = \int_a^x f(t)dt.$$

satisfies

$$\|T\| = b - a.$$

Since $B(\mathscr{X},\mathscr{Y})$ has been shown to be a normed space, the next natural question in the setting on Functional Analysis is: Is the space a Banach space? This is answered in the following theorem:

Theorem 2.3.12. *Let $\mathscr{X}$ and $\mathscr{Y}$ be a normed spaces over the same field. If $\mathscr{Y}$ is a Banach space then $B(\mathscr{X},\mathscr{Y})$ is a Banach space under the operator norm.*

PROOF. Theorems 2.3.5 and 2.3.6 show that $B(\mathscr{X},\mathscr{Y})$ is a normed space with the operator norm. We are only left to show that the space is complete. Assume that $\{T_n\}_{n\in\mathbb{N}} \subset B(\mathscr{X},\mathscr{Y})$ is Cauchy. Since for any $x \in \mathscr{X}$ we have that

$$\|T_n x - T_m x\| = \|(T_n - T_m)x\| \leq \|T_n - T_m\|\|x\|$$

we conclude that $\{T_n x\}_{n\in\mathbb{N}}$ is Cauchy for any $x\in\mathscr{X}$. Indeed, for a given $0\neq x\in\mathscr{X}$ and $\varepsilon>0$ we can find $n_0(x)\in\mathbb{N}$ such that for any $n,m\geq n_0(x)$

$$\|T_n-T_m\|\leq\frac{\varepsilon}{\|x\|}.$$

Consequently, for any $n,m\geq n_0(x)$

$$\|T_n x-T_m x\|\leq\|T_n-T_m\|\,\|x\|<\varepsilon.$$

When $x=0$ we have that $T_n 0=0$ and as such not only Cauchy - but in fact converges to 0.
Since $\{T_n x\}_{n\in\mathbb{N}}\subset\mathscr{Y}$ is Cauchy for any x and since $\mathscr{Y}$ is complete, $\{T_n x\}_{n\in\mathbb{N}}$ must converge for any $x\in\mathscr{X}$ to some element $y_x\in\mathscr{Y}$. We can therefore define a new map $T:\mathscr{X}\to\mathscr{Y}$ by

$$Tx=\lim_{n\to\infty}T_n x,$$

which we will now show is a linear bounded operator.
Given $x_1,x_2\in\mathscr{X}$ we have that due to the linearity of every element of the sequence $\{T_n\}_{n\in\mathbb{N}}$

$$T(x_1+x_2)=\lim_{n\to\infty}T_n(x_1+x_2)=\lim_{n\to\infty}(T_n x_1+T_n x_2)$$
$$=\lim_{n\to\infty}T_n x_1+\lim_{n\to\infty}T_n x_2=Tx_1+Tx_2.$$

Similarly, for any $x\in\mathscr{X}$ and a scalar α

$$T(\alpha x)=\lim_{n\to\infty}T_n(\alpha x)=\lim_{n\to\infty}\alpha T_n x=\alpha\lim_{n\to\infty}T_n x=\alpha Tx,$$

which shows the desired linearity of T.
To show the boundedness of T we will use the fact that every Cauchy sequence in a metric space is bounded and as such

$$\sup_{n\in\mathbb{N}}\|T_n\|=M<\infty.$$

Consequently, for any $x\in\mathscr{X}$

$$\|Tx\|=\lim_{n\to\infty}\|T_n x\|=\liminf_{n\to\infty}\|T_n x\|\leq\liminf_{n\to\infty}(\|T_n\|\,\|x\|)$$
$$=\left(\liminf_{n\to\infty}\|T_n\|\right)\|x\|\leq M\|x\|,$$

which, according to the definition, shows that T is bounded.
To conclude the proof, we only need to show that the sequence $\{T_n\}_{n\in\mathbb{N}}$ converges to T in the operator norm. Since $\{T_n\}_{n\in\mathbb{N}}$ is Cauchy we can find $n_0\in\mathbb{N}$ for any $\varepsilon>0$ such that $\|T_n-T_m\|<\frac{\varepsilon}{2}$ whenever $n,m\geq n_0$. As such, for any $x\in\mathscr{X}$ and any $n\geq n_0$

$$\|Tx-T_n x\|=\lim_{m\to\infty}\|T_m x-T_n x\|\leq\left(\liminf_{m\to\infty}\|T_m-T_n\|\right)\|x\|\leq\frac{\varepsilon}{2}\|x\|.$$

Consequently, by definition, for any $n\geq n_0$

$$\|T-T_n\|\leq\frac{\varepsilon}{2}<\varepsilon,$$

which shows the desired convergence. The proof is now complete. □

REMARK 2.3.13. A key ingredient of the above proof is the fact that $\{T_n\}_{n\in\mathbb{N}}$ was Cauchy in the *operator norm*, i.e. *uniformly with respect to x*. We have strongly used the fact that

$$\|T_n x - T_m x\| \leq \underbrace{\|T_n - T_m\|}_{\text{small independently of } x} \|x\|$$

in order to show that $\{T_n\}_{n\in\mathbb{N}}$ converges to T in the operator norm. Had we only known that for every $x \in \mathscr{X}$ $\{T_n x\}_{n\in\mathbb{N}}$ converges to Tx (pointwise convergence) we *wouldn't have been able to conclude that* $\{T_n\}_{n\in\mathbb{N}}$ *converges to* T *in the operator norm.* Indeed, consider the case where our space is a Hilbert space, $\mathscr{H}$, with a countable orthonormal basis $\mathscr{B} = \{e_n\}_{n\in\mathbb{N}}$. We define $T_n : \mathscr{H} \to \mathscr{H}$ by

$$T_n x = \sum_{i=1}^{n} \langle x, e_i \rangle e_i.$$

T_n is clearly linear for any $n \in \mathbb{N}$ and according to Pythagoras theorem and Parseval's identity we have that

$$\|T_n x\|^2 = \sum_{i=1}^{n} |\langle x, e_i \rangle|^2 \leq \sum_{i\in\mathbb{N}} |\langle x, e_i \rangle|^2 = \|x\|^2$$

which shows that $\{T_n\} \subset B(\mathscr{H},\mathscr{H})$.
Since $\mathscr{B}$ is an orthonormal basis we have that for any $x \in \mathscr{X}$

$$T_n x = \sum_{i=1}^{n} \langle x, e_i \rangle e_i \underset{n\to\infty}{\longrightarrow} \sum_{i\in\mathbb{N}} \langle x, e_i \rangle e_i = x.$$

We claim, however, that $\{T_n\}_{n\in\mathbb{N}}$ doesn't converge to $I_{\mathscr{H}}$, the identity operator, in norm. The main problem is that the convergence above is *not uniform in x*. Indeed, noticing that $\mathscr{N}\, T_n = \{e_1, \dots, e_n\}^{\perp}$ we see that

$$\|T_n - I_{\mathscr{H}}\| = \sup_{\|x\|=1} \|T_n x - x\| \geq \|T_n e_{n+1} - e_{n+1}\| = \|e_{n+1}\| = 1$$

which shows that no convergence is possible.

Before continuing, it is worth to mention that much like linear operators in finite dimensions, one can compose bounded operators from $\mathscr{X}$ to $\mathscr{Y}$ with ones from $\mathscr{Y}$ to $\mathscr{Z}$. This is expressed in the next theorem:

Theorem 2.3.14. *Let $\mathscr{X}$, $\mathscr{Y}$ and $\mathscr{Z}$ be normed spaces and let $T \in B(\mathscr{X},\mathscr{Y})$ and $S \in B(\mathscr{Y},\mathscr{Z})$. Then $S \circ T \in B(\mathscr{X},\mathscr{Z})$ and*

$$\|S \circ T\| \leq \|S\| \, \|T\|. \tag{2.9}$$

Consequently for any $T \in B(\mathscr{X},\mathscr{X})$ the operator

$$T^n = \underbrace{T \circ T \circ \cdots \circ T}_{n \text{ times}}$$

is well defined, belongs to $B(\mathscr{X},\mathscr{X})$, and satisfies

$$\left\|T^n\right\| \leq \|T\|^n.$$

The last question we ask ourselves, before concluding this section, is whether or not a basis for $B\left(\mathscr{X},\mathscr{Y}\right)$ exists when ones for $\mathscr{X}$ and $\mathscr{Y}$ do. This question is motivated by the following finite dimensional theorem:

Theorem 2.3.15. *Let $\mathscr{X}$ and $\mathscr{Y}$ be two vector spaces and let $T:\mathscr{X}\to\mathscr{Y}$ be a linear operator. If $\mathscr{B}_{\mathscr{X}}=\{e_1,\dots,e_n\}$ is a basis for $\mathscr{X}$ then T is uniquely determined by the set $\{Te_i\}_{i=1,\dots,n}$. If, in addition, $\mathscr{B}_{\mathscr{Y}}=\{u_1,\dots,u_m\}$ is a basis for $\mathscr{Y}$ then the operators $T_{ij}:\mathscr{X}\to\mathscr{Y}$ defined as the linear extension of*

$$T^{(i,j)}e_k=\begin{cases}u_j, & k=i,\\ 0, & k\neq i,\end{cases}$$

form a basis to $B\left(\mathscr{X},\mathscr{Y}\right)$.

The most natural extension to a finite dimensional basis will be a Shcauder basis. If $\mathscr{X}$ is a Banach space with Schauder basis, $\mathscr{B}=\{e_n\}_{n\in\mathbb{N}}$, the first statement of the above theorem remains valid when $T\in B\left(\mathscr{X},\mathscr{Y}\right)$. Indeed, due the continuity of T and the unique expansion

$$x=\sum_{n\in\mathbb{N}}\alpha_n(x)e_n$$

we find that

$$\begin{aligned}Tx&=T\left(\lim_{N\to\infty}\sum_{n=1}^{N}\alpha_n(x)e_n\right)=\lim_{n\to\infty}T\left(\sum_{n=1}^{N}\alpha_n(x)e_n\right)\\ &=\lim_{N\to\infty}\sum_{n=1}^{N}\alpha_n(x)Te_n=\sum_{n\in\mathbb{N}}\alpha_n(x)Te_n,\end{aligned}\tag{2.10}$$

where the existence of the limit $\lim_{N\to\infty}\sum_{n=1}^{N}\alpha_n(x)Te_n$ is assured from the convergence of $\sum_{n=1}^{N}\alpha_n(x)e_n$ to x and the continuity of T.
A basis connection, however, is *not possible* in the general case, even when $\mathscr{Y}$ is extremely simple. We will see this shortly.

2.4. Linear functional and the Dual space

One extremely important type of bounded linear operators is bounded linear *functionals*, i.e. linear operators from our space to its underlying field. The Banach space associated to these operators, known as the *dual space*, is *intimately* connected to the space on which these operators are defined.

Definition 2.4.1. Let $\mathscr{X}$ be a vector space over a field $\mathbb{F}$. A *linear functional* is a linear operator $f:\mathscr{D}\left(f\right)\subset\mathscr{X}\to\mathbb{F}$. When the $\mathscr{X}$ is also normed we say that a functional f is *bounded* if there exists $C>0$ such that for all $x\in\mathscr{D}\left(f\right)$

$$\left|f(x)\right|\leq C\left\|x\right\|.$$

In that case we define the (operator) norm of f as

$$\left\|f\right\|=\inf\left\{C>0\ \middle|\ \left|f(x)\right|\leq C\left\|x\right\|\ \ \forall x\in\mathscr{D}\left(f\right)\right\}.$$

The space $B(\mathscr{X},\mathbb{F})$, which is a Banach space according to Theorem 2.3.12, is known as *the dual space of $\mathscr{X}$* and is denoted by $\mathscr{X}^*$.

Note that for functionals we tend to write $f(x)$ and not fx to avoid confusion.

REMARK 2.4.2. According to theorem 2.3.6 we have that

$$\|f\| = \sup_{x\neq 0} \frac{|f(x)|}{\|x\|} = \sup_{\|x\|=1} |f(x)|.$$

REMARK 2.4.3. One can also define the algebraic dual of a vector space $\mathscr{X}$, $\mathscr{X}'$, as the space $L(\mathscr{X},\mathbb{F})$. While there is no difference between this space and $\mathscr{X}^*$ when $\mathscr{X}$ is finite dimensional, $\mathscr{X}^* \underset{\neq}{\subset} \mathscr{X}'$ in general.

Example 2.4.4 *(Inner product is a bounded linear functional).* Let $\mathscr{H}$ be an inner product space over a field $\mathbb{F}$ and let $y \in \mathscr{H}$. The function $f_y : \mathscr{H} \to \mathbb{F}$ defined by

$$f_y(x) = \langle x, y \rangle$$

is a bounded linear functional. In fact, as you saw in Analysis III (and we will mention it again shortly) *any* bounded linear functional on a Hilbert space is of this form.

Example 2.4.5 *(The "delta" functional).* Consider the Banach space $(C[a,b], \|\cdot\|_\infty)$. For any $x_0 \in [a,b]$ the functional $\delta_{x_0} : C[a,b] \to \mathbb{F}$ defined by

$$\delta_{x_0}(f) = f(x_0).$$

is a bounded linear functional. δ_{x_0} is known as *the delta functional at* x_0. It appears in other contexts as well.

Example 2.4.6 *(The definite integral).* Consider the Banach space $(C[a,b], \|\cdot\|_\infty)$ over $\mathbb{R}$. The functional $I : C[a,b] \to \mathbb{R}$ defined by

$$I(f) = \int_a^b f(x)dx.$$

is a bounded linear funcitonal.

A question one might ask at this point is: Why do we call $\mathscr{X}^*$ the *dual* space? The answer to that question lies in the next theorem, which is an immediate consequence of Theorem 2.3.15.

Theorem 2.4.7. *Let $\mathscr{X}$ be a vector space over a field $\mathbb{F}$ and let $f : \mathscr{X} \to \mathbb{F}$ be a linear functional. If $\mathscr{B}_{\mathscr{X}} = \{e_1, \ldots, e_n\}$ is a basis for $\mathscr{X}$ then f is uniquely determined by the numbers $\{f(e_i)\}_{i=1,\ldots,n}$. Moreover, the functionals $\{f^{(1)}, \ldots, f^{(n)}\}$ defined by*

$$f^{(j)}\left(\sum_{i=1}^n \alpha_i e_i\right) = \alpha_j,$$

form a basis to $\mathscr{X}^*$.

REMARK 2.4.8. The *duality* we refer to in the name "dual space" is the fact that the basis $\{e_1, \ldots, e_n\}$ is *mirrored* in $\mathscr{X}^*$ by the basis $\{f^{(1)}, \ldots, f^{(n)}\}$. A simple way to think about $\{f^{(1)}, \ldots, f^{(n)}\}$ is to notice that

$$f^{(i)}(e_j) = \delta_{i,j}$$

where $\delta_{i,j}$ is the Kronecker delta, and as such the connection between $\{e_1,\dots,e_n\}$ and $\{f^{(1)},\dots,f^{(n)}\}$ is similar to an *orthonormality* condition.
Moreover, the linear operator $\mathcal{D}:\mathcal{X}\to\mathcal{X}^*$ defined as

$$\mathcal{D}\left(\sum_{i=1}^{n}\alpha_i e_i\right)=\sum_{i=1}^{n}\alpha_i f^{(i)}$$

is an injective linear operator between two vector spaces of the same dimension - and as such is a *bijection*.

Can we extend the above to infinite dimensional spaces that have a Schauder basis? Are we able to build a dual basis?

A partial positive result to our question is given in the following theorem, whose proof relies heavily on one of the fundamental theorems of Functional Analysis which we will discuss in the next chapter:

Theorem 2.4.9. *Let $\mathcal{X}$ be a Banach space with Schauder basis $\mathcal{B}=\{e_n\}_{n\in\mathbb{N}}$. Then there exists a unique sequence $\{f^{(n)}\}_{n\in\mathbb{N}}\subset\mathcal{X}^*$ such that $f^{(n)}(e_j)=\delta_{n,j}$.*

The problem is, however, that the above sequence $\{f^{(n)}\}_{n\in\mathbb{N}}$ is *not always* a Schauder basis for $\mathcal{X}^*$. In fact, there are cases where $\mathcal{X}$ has a Shchauder basis while $\mathcal{X}^*$ is not separable, and as such can have no Schauder basis. We will see an example shortly.

While the notion of $\mathcal{X}^*$ seems complicated, there are some simple and prototypical cases where we are able to find $\mathcal{X}^*$ explicitly. We start with the "simplest" case of Hilbert spaces - i.e. we start with Riesz's representation theorem.

Theorem 2.4.10 *(Riesz's representation theorem for Hilbert spaces). Let $\mathcal{H}$ be a Hilbert space and let $f\in\mathcal{H}^*$. Then there exists a unique $y\in\mathcal{H}$ such that*

$$f(x)=f_y(x)=\langle x,y\rangle. \tag{2.11}$$

Moreover, $\|f\|=\|f_y\|=\|y\|$.

This has been shown in Analysis III.

REMARK 2.4.11. Riesz's representation theorem is sometimes written as $\mathcal{H}^*=\mathcal{H}$ which is to be understood as the identification of elements of $\mathcal{H}$ as the "generators" of $\mathcal{H}^*$.

Looking at the above one may wonder if the "equality" $\mathcal{H}^*=\mathcal{H}$ is more than just notational writing. Are $\mathcal{H}$ and $\mathcal{H}^*$ indeed "equal" in some sense, in which case we would conclude that $\mathcal{H}^*$ is in fact Hilbert space and not only a Banach space? The short answer for this question is *Almost?*. A more detailed answer is given in the next theorem.

Theorem 2.4.12. *Let $\mathcal{H}$ be a Hilbert space. Define the map $\mathcal{J}:\mathcal{H}\to\mathcal{H}^*$ by*

$$\mathcal{J}y=f_y$$

where the functional f_y was defined in (2.11). Then $\mathcal{J}$ is a conjugate linear bijection between $\mathcal{H}$ and $\mathcal{H}^*$, i.e. $\mathcal{J}$ is a bijection such that for any $y_1,y_2\in\mathcal{H}$ and a *scalar α we have that*

$$\mathcal{I}\left(y_1+y_2\right)=\mathcal{I}y_1+\mathcal{I}y_2, \qquad \mathcal{I}\left(\alpha y\right)=\overline{\alpha}\mathcal{I}y.$$

Moreover

$$\left\|\mathcal{I}y\right\|_{\mathcal{H}}=\left\|y\right\|_{\mathcal{H}}.$$

Consequently, we can define an inner product on $\mathcal{H}^*$ which induces the norm on the space by

$$\langle f,g\rangle=\overline{\langle\mathcal{I}^{-1}f,\mathcal{I}^{-1}g\rangle_{\mathcal{H}}}, \tag{2.12}$$

making $\mathcal{H}^*$ into a Hilbert space.

REMARK 2.4.13. Two normed spaces, $(\mathcal{X},\|\cdot\|_{\mathcal{X}})$ and $\left(\mathcal{Y},\|\cdot\|_{\mathcal{Y}}\right)$, are called *isometric* if there exists a linear bijection between them $T:\mathcal{X}\to\mathcal{Y}$ such that

$$\|Tx\|_{\mathcal{Y}}=\|x\|_{\mathcal{X}}$$

for any $x\in\mathcal{X}$. In that case $(\mathcal{X},\|\cdot\|_{\mathcal{X}})$ and $\left(\mathcal{Y},\|\cdot\|_{\mathcal{Y}}\right)$ are *topologically isometric* - meaning that all topological properties such as convergence, completeness, separability and etc. are exactly the same. T is known as an *isometry*. One can also extend the above to inner product spaces.
The above theorem tells us that $\mathcal{H}$ and $\mathcal{H}^*$ are not isometric, but are "conjugate isometric". That is still enough to get all the analytic and geometric structure to be (almost) identical.

It is worth to mention that Riesz' representation theorem and the above observation motivate another frequently used notation in Functional Analysis. Given $x\in\mathcal{X}$ and $f\in\mathcal{X}^*$ we denote by $\langle f,x\rangle$ the expression $f(x)$. We, however, will not use this notation in our part of the module.

We continue with two more examples for explicit dual spaces.

Theorem 2.4.14 *(Riesz representation theorem for L^p). Let $E\subset\mathbb{R}^n$ be a measurable set and let $p\in[1,\infty)$ be given. For any bounded linear functional $\mathfrak{I}:L^p(E)\to\mathbb{C}$ there exists $g\in L^q(E)$, where q is the Hölder conjugate of p, such that*

$$\mathfrak{I}\left(f\right)=\int_E f(x)\overline{g(x)}dx.$$

Moreover, $\|\mathfrak{I}\|=\left\|g\right\|_{L^q(E)}$.

The proof of this theorem was shown in Analysis III for $n=1$ and remains the same for $n>1$. Much like in the Hilbert case, it is common to express the above by writing

$$L^p(E)^*=L^q(E)$$

for $p\in[1,\infty)$. While we won't consider the proof here, we'd like to add that it is fairly straight forward to show that $L^\infty(E)^*\neq L^1(E)$.

We will conclude this section with the investigation of the dual spaces of the $\ell_p(\mathbb{N})$ spaces.

Theorem 2.4.15. *Let $1 \leq p \leq \infty$ be given, and let $1 \leq q \leq \infty$ be its Hölder conjugate. For any $\boldsymbol{b} \in \ell_q(\mathbb{N})$ we have that*

$$f_{\boldsymbol{b}}(\boldsymbol{a}) = \sum_{n\in\mathbb{N}} a_n \overline{b_n}$$

is an element of $\ell_p(\mathbb{N})^*$ that satisfies $\left\| f_{\boldsymbol{b}} \right\| \leq \|\boldsymbol{b}\|_q$. Moreover, for any $f \in \ell_p(\mathbb{N})^*$ with $p \neq \infty$ there exists a unique $\boldsymbol{b} \in \ell_q(\mathbb{N})$ such that $f = f_{\boldsymbol{b}}$. In that case we also find that

$$\left\| f \right\| = \left\| f_{\boldsymbol{b}} \right\| = \|\boldsymbol{b}\|_q . \tag{2.13}$$

Much like with Lebesgue spaces, Theorem 2.4.15 is usually expressed by the (slightly misleading) notation

$$\ell_p(\mathbb{N})^* = \ell_q(\mathbb{N}).$$

PROOF. We start by noticing that the discrete Hölder inequality

$$\sum_{n\in\mathbb{N}} |a_n||b_n| \leq \|\boldsymbol{a}\|_p \|\boldsymbol{b}\|_q$$

shows that $f_{\boldsymbol{b}}$ is a well defined map from $\ell_p(\mathbb{N})$ to $\mathbb{F}$, and gives us the estimate

$$\left\| f_{\boldsymbol{b}} \right\| \leq \|\boldsymbol{b}\|_q , \tag{2.14}$$

once we have show that $f_{\boldsymbol{b}}$ is linear. The linearity of $f_{\boldsymbol{b}}$ is straight forward to show due to the absolute convergence of all involved sum. Indeed, for any $\boldsymbol{a}, \boldsymbol{c} \in \ell_p(\mathbb{N})$ we have that

$$f_{\boldsymbol{b}}(\boldsymbol{a}+\boldsymbol{c}) = \sum_{n\in\mathbb{N}} (a_n + c_n)\overline{b_n} = \sum_{n\in\mathbb{N}} a_n\overline{b_n} + \sum_{n\in\mathbb{N}} c_n\overline{b_n} = f_{\boldsymbol{b}}(\boldsymbol{a}) + f_{\boldsymbol{b}}(\boldsymbol{c}).$$

Similarly, for any $\boldsymbol{a} \in \ell_p(\mathbb{N})$ and a scalar α we have that

$$f_{\boldsymbol{b}}(\alpha\boldsymbol{a}) = \sum_{n\in\mathbb{N}} (\alpha a_n)\overline{b_n} = \alpha \sum_{n\in\mathbb{N}} a_n\overline{b_n} = \alpha f_{\boldsymbol{b}}(\boldsymbol{a}).$$

We thus turn our attention to the second part of the statement.
Let f be a functional in $\ell_p(\mathbb{N})^*$ for some $1 \leq p < \infty$. Due to the continuity of f and the fact that $\mathscr{B} = \{\boldsymbol{e}_n\}_{n\in\mathbb{N}}$ defined by

$$(\boldsymbol{e}_n)_k = \begin{cases} 1, & k = n, \\ 0, & k \neq n, \end{cases}$$

is a Schauder basis for $\ell_p(\mathbb{N})$ with the expansion

$$\boldsymbol{a} = \sum_{n\in\mathbb{N}} a_n \boldsymbol{e}_n \qquad \text{where} \quad \boldsymbol{a} = (a_1, a_2, \ldots),$$

we find that for any $\boldsymbol{a} \in \ell_p(\mathbb{N})$

$$f(\boldsymbol{a}) = \sum_{n\in\mathbb{N}} a_n f(\boldsymbol{e}_n) = f_{\boldsymbol{b}}(\boldsymbol{a}),$$

where $\boldsymbol{b} = \left\{\overline{f(\boldsymbol{e}_n)}\right\}_{n\in\mathbb{N}}$. If we'll show that $\boldsymbol{b} \in \ell_q(\mathbb{N})$ we will conclude the *existence* of the representation claimed in the theorem.

At this point we will have to consider two cases: $p = 1$ and $1 < p < \infty$.
$\underline{p = 1}$: In this case we have that $q = \infty$. For a given $N \in \mathbb{N}$ we notice that by choosing $\boldsymbol{b}_\infty^{(N)} = e^{i\operatorname{Arg}(b_N)}\boldsymbol{e}_N$[1], which is in $\ell_1(\mathbb{N})$, we find that

$$|b_N| = \sum_{n\in\mathbb{N}} \left(\boldsymbol{b}_\infty^{(N)}\right)_n \overline{b_n} = f\left(\boldsymbol{b}_\infty^{(N)}\right) \leq \|f\| \underbrace{\left\|\boldsymbol{b}_\infty^{(N)}\right\|_1}_{=1} = \|f\|.$$

As $N \in \mathbb{N}$ was arbitrary we conclude that

$$\|\boldsymbol{b}\|_\infty = \sup_{n\in\mathbb{N}} |b_n| \leq \|f\|.$$

$\underline{1 < p < \infty}$: In this case we have that $1 < q < \infty$. For a given $N \in \mathbb{N}$ we consider the sequence $\boldsymbol{b}_q^{(N)}$ defined by

$$\boldsymbol{b}_q^{(N)} = \begin{cases} e^{i\operatorname{Arg}(b_n)} b_n^{q-1}, & n \leq N, \\ 0, & n > N, \end{cases}$$

and find that

$$\left\|\boldsymbol{b}_q^{(N)}\right\|_p = \left(\sum_{n=1}^N |b_n|^{p(q-1)}\right)^{\frac{1}{p}} = \left(\sum_{n=1}^N |b_n|^q\right)^{\frac{1}{p}}.$$

Since $\boldsymbol{b}_q^{(N)} \in \ell_p(\mathbb{N})$, as a (finite) linear combination of the standard basis of $\ell_p(\mathbb{N})$, we see that

$$\underbrace{\sum_{n\in\mathbb{N}} \left(\boldsymbol{b}_q^{(N)}\right)_n \overline{b_n}}_{=\sum_{n=1}^N |b_n|^{q-1}|b_n| = \sum_{n=1}^N |b_n|^q} \leq \|f\| \left\|\boldsymbol{b}_q^{(N)}\right\|_p = \|f\| \left(\sum_{n=1}^N |b_n|^q\right)^{\frac{1}{p}}.$$

As the above holds for any $N \in \mathbb{N}$ we conclude that[2]

$$\left(\sum_{n=1}^N |b_n|^q\right)^{\frac{1}{q}} \leq \|f\| \qquad \forall N.$$

Taking N to infinity shows that $\boldsymbol{b} \in \ell_q(\mathbb{N})$ and that

$$\|\boldsymbol{b}\|_q \leq \|f\|. \tag{2.15}$$

We conclude that any $f \in \ell_p(\mathbb{N})^*$ can be written as $f_{\boldsymbol{b}}$ for some $\boldsymbol{b} \in \ell_q(\mathbb{N})$ and combining (2.14) and (2.15) shows that

$$\|f\| = \|f_{\boldsymbol{b}}\| = \|\boldsymbol{b}\|_q.$$

We are only left with the uniqueness of the representation. This follows immediately from the above, the fact that if $\boldsymbol{b}_1, \boldsymbol{b}_2 \in \ell_q(\mathbb{N})$ then $\boldsymbol{b}_1 - \boldsymbol{b}_2 \in \ell_q(\mathbb{N})$ and the fact that

$$f_{\boldsymbol{b}_1} - f_{\boldsymbol{b}_2} = f_{\boldsymbol{b}_1 - \boldsymbol{b}_2}.$$

Indeed, if $f_{\boldsymbol{b}_1} = f_{\boldsymbol{b}_2}$ then

$$0 = \|f_{\boldsymbol{b}_1} - f_{\boldsymbol{b}_2}\| = \|f_{\boldsymbol{b}_1 - \boldsymbol{b}_2}\| = \|\boldsymbol{b}_1 - \boldsymbol{b}_2\|_q,$$

[1] where we define $\operatorname{Arg}(0) = 0$

[2] recall that $1 - \frac{1}{p} = \frac{1}{q}$.

which implies that $\boldsymbol{b}_1 = \boldsymbol{b}_2$. □

Theorem 2.4.15 tell us that the dual of $\ell_1(\mathbb{N})$ is $\ell_\infty(\mathbb{N})$ and that for any $f \in \ell_1(\mathbb{N})^*$ there exists $\boldsymbol{b} \in \ell_\infty(\mathbb{N})$ such that $f = f_{\boldsymbol{b}}$ and $\|f_{\boldsymbol{b}}\| = \|\boldsymbol{b}\|_\infty$. This automatically implies that $\ell_1(\mathbb{N})^*$ can not separable.
An immediate consequence of this observation is an answer to the question about the existence of a dual basis to a space with Schauder basis: $\ell_1(\mathbb{N})$ has a Shcauder basis yet its dual is not separable and as such can't have a Schauder basis (let alone the one we found in Theorem 2.4.9).

We are naturally curious about the dual space of $\ell_\infty(\mathbb{N})$. The next theorem indicates that it more complicated than we would have wished:

Theorem 2.4.16. *There exists a functional in $\ell_\infty(\mathbb{N})^*$ that is not of the form $f_{\boldsymbol{b}}$ with $\boldsymbol{b} \in \ell_1(\mathbb{N})$.*

The proof of the above relies on the following theorem, which is given without a proof:

Theorem 2.4.17. *Let $\mathscr{X}$ be a Banach space. If $\mathscr{X}^*$ is separable then so is $\mathscr{X}$.*

2.5. Weak and Weak-∗ Convergence

While $\mathscr{X}$ and $\mathscr{X}^*$ have a normed structure, it is usually too constricting to get some appealing topological properties such as a compactness criterion. Both spaces, however, do feature a "weaker" notion of convergence, openness, and closedness of sets - but not without a price: these topologies are not metrisable, i.e. not arise from an underlying metric. These "weak" topologies, nonetheless, have a lot of interesting properties.

Definition 2.5.1. Let $\mathscr{X}$ be a Banach space and let $\mathscr{X}^*$ be its dual space. We say that $\{x_n\}_{n\in\mathbb{N}} \subset \mathscr{X}$ *converges weakly* to $x \in \mathscr{X}$ and write

$$x_n \underset{n\to\infty}{\overset{w}{\longrightarrow}} x \qquad \text{or} \qquad x_n \underset{n\to\infty}{\rightharpoonup} x$$

if for every $f \in \mathscr{X}^*$ we have that $f(x_n) \underset{n\to\infty}{\longrightarrow} f(x)$.
We say that $\{f_n\}_{n\in\mathbb{N}} \subset \mathscr{X}^*$ *converges weakly-∗* to $f \in \mathscr{X}^*$ and write

$$f_n \underset{n\to\infty}{\overset{w-*}{\longrightarrow}} f \qquad \text{or} \qquad f_n \underset{n\to\infty}{\overset{*}{\rightharpoonup}} f$$

if for every $x \in \mathscr{X}$ we have that $f_n(x) \underset{n\to\infty}{\longrightarrow} f(x)$.

Example 2.5.2 *(Weak convergence in Hilbert spaces)*. Due to Riesz representation theorem we can show that in any Hilbert space $x_n \underset{n\to\infty}{\overset{w}{\longrightarrow}} x$ is equivalent to

$$\langle x_n, y\rangle \underset{n\to\infty}{\longrightarrow} \langle x, y\rangle \qquad \forall y \in \mathscr{H}.$$

Example 2.5.3 *(The Schauder basis in ℓ_p spaces).* We claim that the standard Schauder basis for $\ell_p(\mathbb{N})$ with $1 \le p < \infty$, $\{\boldsymbol{e}_n\}_{n\in\mathbb{N}}$, converges weakly to $\mathbf{0}$ when

$1 < p < \infty$. Indeed, given $f \in \ell_p(\mathbb{N})^*$ we know from Theorem 2.4.15 that there exists $\boldsymbol{b} \in \ell_q(\mathbb{N})$, where $1 < q < \infty$ is the Hölder conjugate of p, such that

$$f(\boldsymbol{a}) = \sum_{n\in\mathbb{N}} a_n \overline{b_n}.$$

As such

$$f(\boldsymbol{e}_n) = \overline{b_n} \underset{n\to\infty}{\longrightarrow} 0 = f(\boldsymbol{0}),$$

since $\sum_{n\in\mathbb{N}} |b_n|^q < \infty$. This shows the desired weak convergence. Note that since for any $n \neq m$

$$\|\boldsymbol{e}_n - \boldsymbol{e}_m\|_p = 2^{\frac{1}{p}}$$

the sequence $\{\boldsymbol{e}_n\}_{n\in\mathbb{N}}$ is *not* Cauchy and consequently can't converge. This shows that in general *weak convergence doesn't imply (strong) convergence*.
Moreover, the above proof of weak convergence also shows us why the claim is *not true* in $\ell_1(\mathbb{N})$: consider the sequence $\boldsymbol{b} = (1,1,1,\dots) \in \ell_\infty(\mathbb{N})$. Then $f_{\boldsymbol{b}} \in \ell_1(\mathbb{N})^*$ and

$$f_{\boldsymbol{b}}(\boldsymbol{e}_n) = 1 \underset{n\to\infty}{\longrightarrow} 1 \neq f_{\boldsymbol{b}}(\boldsymbol{0}).$$

Weak and weak-∗ convergence do enjoy some "familiar" properties:

Theorem 2.5.4. *Let $\mathcal{X}$ be a Banach space and let $\{x_n\}_{n\in\mathbb{N}} \subset \mathcal{X}$ be a sequence that converges weakly to an element $x \in \mathcal{X}$. Then*

- *(i) The limit of $\{x_n\}_{n\in\mathbb{N}}$ is unique.*
- *(ii) Every subsequence of $\{x_n\}_{n\in\mathbb{N}}$ converges weakly to x.*
- *(iii) $\{x_n\}_{n\in\mathbb{N}}$ is bounded in norm.*
- *(iv) Basic limit arithmetic holds for weak convergence.*

The same statements hold for weak-∗ convergence.

We will end this section, and with it this chapter, with an important theorem that is beyond the scope of this module. This theorem, known as the *Banach-Alaoglu Theorem* gives a prototype of compact sets in $\mathcal{X}^*$, where the notion of compactness corresponds to that which concerns itself with the ability to choose a finite open subcover from any open cover of a set. This criterion is not always equivalent to that of sequential compactness.

Theorem 2.5.5 *(Banach-Alaoglu Theorem). The closed unit ball in $\mathcal{X}^*$, $\overline{B}_{\mathcal{X}^*} = \left\{f \in \mathcal{X}^* \;\middle|\; \|f\|_{\mathcal{X}^*} \leq 1\right\}$ is compact in the weak-∗ topology.*

CHAPTER 3

Fundamental Theorems in Functional Analysis

In our final chapter for the first part of the module we will focus our attention on four fundamental theorems in the field of Functional Analysis - the so-called *cornerstones of Functional Analysis*: the Hahn-Banach Theorem, the Banach-Steinhaus Theorem (also known as The Uniform Boundedness Theorem), the Open Mapping Theorem, and the Closed Graph Theorem.

Both the Banach-Steinhaus theorem and the Open Mapping theorem rely on an important observation from the theory of metric spaces, known as *the Baire Category Theorem*, which we will mention but not prove.

3.1. The Hahn-Banach Theorem

In section §2.1 we have seen that we can always extend a bounded linear operator to the closure of its domain in a unique way. It is natural to wonder if we can generalise this further and extend the operator to a larger subspace, the entire space if possible, by relaxing the requirement of uniqueness.

The Hahn-Banach theorem tackles this question for bounded linear functionals and shows that not only can we extend them - but that we can do it in a way that preserves the norm of the original functional.

Our main goal in this section will be to prove the following:

Theorem 3.1.1 *(Hahn-Banach Theorem). Let $\mathscr{X}$ be a normed space and let $\mathscr{Y}$ be a subspace of $\mathscr{X}$. Assume that f is a bounded linear functional on $\mathscr{Y}$. Then there exists a bounded linear extension of f to all of $\mathscr{X}$, $\widetilde{f}:\mathscr{X}\to\mathbb{F}$, such that*

$$\|f\|_{\mathscr{Y}^*}=\|\widetilde{f}\|_{\mathscr{X}^*},$$

where

$$\|f\|_{\mathscr{Y}^*}=\begin{cases}\sup_{y\in\mathscr{Y},\,y\neq0}\frac{|f(y)|}{\|y\|}, & \mathscr{Y}\neq\{0\},\\ 0, & \mathscr{Y}=\{0\}.\end{cases}$$

Before we start the proof of the above theorem, it is worth to note that the Hahn-Banach theorem is much simple to show when one considers Hilbert spaces - another testament to the geometric prowess of such spaces.

Our initial proof of the Hahn-Banach theorem focuses on the case where the underlying field of $\mathscr{X}$, $\mathbb{F}$, is $\mathbb{R}$.

Proof of the Hahn-Banach Theorem over $\mathbb{R}$. We start by showing that we can always extend f to a space that is spanned by an additional vector that is

not in $\mathscr{Y}$ in a way that preserves the norm.
Let $x \in \mathscr{X}$ be a vector that is not in $\mathscr{Y}$ and define

$$\mathscr{Y}_x = \operatorname{span}\{x, \mathscr{Y}\} = \operatorname{span}\{x\} + \mathscr{Y}.$$

Every vector $z \in \mathscr{Y}_x$ can be written uniquely as $z = \alpha_z x + y_z$ for some scalar α_z and some $y_z \in \mathscr{Y}$. Consequently, any linear extension of f from $\mathscr{Y}$ to $\mathscr{Y}_x$, $\widetilde{f}$, must satisfy

$$\widetilde{f}(z) = \widetilde{f}(\alpha_z x + y_z) = \alpha_z \widetilde{f}(x) + f(y_z).$$

The unique representation of vectors in $\mathscr{Y}_x$ implies that the converse also holds: for any scalar c (representing $\widetilde{f}(x)$) the functional

$$\widetilde{f}_c(z) = \alpha_z c + f(y_z), \tag{3.1}$$

is a linear extension of f to $\mathscr{Y}_x$. To show the desired result, then, we need to find a scalar c such that for any $z \in \mathscr{Y}_x$ we have that

$$-\|f\|_{\mathscr{Y}^*} \|z\| \leq \widetilde{f}_c(z) \leq \|f\|_{\mathscr{Y}^*} \|z\|. \tag{3.2}$$

Indeed, if we find that the above holds then $\|\widetilde{f}_c\|_{\mathscr{Y}_x^*} \leq \|f\|_{\mathscr{Y}^*}$ by definition and since

$$\|\widetilde{f}_c\|_{\mathscr{Y}_x^*} = \sup_{z\in\mathscr{Y}_x,\, z\neq 0} \frac{|\widetilde{f}_c(z)|}{\|z\|} \geq \sup_{z\in\mathscr{Y},\, z\neq 0} \frac{|\widetilde{f}_c(z)|}{\|z\|} = \sup_{z\in\mathscr{Y},\, z\neq 0} \frac{|f(z)|}{\|z\|} = \|f\|_{\mathscr{Y}^*}.$$

we can conclude that if (3.1) holds then the extension of f, $\widetilde{f}_c$, has the same norm as f. Note that the last inequality can be extended to show that whenever g extends h we must have that $\|g\|_{\mathscr{D}(g)^*} \geq \|h\|_{\mathscr{D}(h)^*}$.

To simplify matters, we notice that the linearity of $\widetilde{f}_c$ will imply that in order to show (3.2) we only need to consider the right hand side inequality. Indeed, if $\widetilde{f}_c(z) \leq \|f\|_{\mathscr{Y}^*} \|z\|$ for any $z \in \mathscr{Y}_z$ then since $\mathscr{Y}_z$ is a subspace we find that

$$-\widetilde{f}_c(z) = \widetilde{f}_c(-z) \leq \|f\|_{\mathscr{Y}^*} \|-z\| = \|f\|_{\mathscr{Y}^*} \|z\|$$

which implies that $-\|f\|_{\mathscr{Y}^*} \|z\| \leq \widetilde{f}_c(z)$. We conclude that we are looking for $c \in \mathbb{R}$ such that

$$\alpha c + f(y) \leq \|f\|_{\mathscr{Y}^*} \|\alpha x + y\|, \tag{3.3}$$

for all $\alpha \in \mathbb{R}$ and $y \in \mathscr{Y}$. We consider three cases:
$\underline{\alpha = 0}$: In this case the above reads as

$$f(y) \leq \|f\|_{\mathscr{Y}^*} \|y\|$$

for all $y \in \mathscr{Y}$, which is our initial assumption.
$\underline{\alpha > 0}$: In this case we can rearrange (3.3) to read as

$$c + f\left(\frac{y}{\alpha}\right) \leq \frac{1}{\alpha} \|f\|_{\mathscr{Y}^*} \|\alpha x + y\| = \|f\|_{\mathscr{Y}^*} \left\|x + \frac{y}{\alpha}\right\|$$

for all $y \in \mathscr{Y}$, where we have used the linearity of f. Since $\mathscr{Y}$ is a subspace we know that $y \in \mathscr{Y}$ if and only if $\beta y \in \mathscr{Y}$ for any $\beta \neq 0$ and consequently the above inequality is equivalent to requiring that

$$c \leq \|f\|_{\mathscr{Y}^*} \|x+y\| - f(y),$$

for all $y \in \mathscr{Y}$. Notice that α *plays no role here* (we have only used its positivity to reach this point). Defining

$$u_f = \inf_{y\in\mathscr{Y}} \left(\|f\|_{\mathscr{Y}^*} \|x+y\| - f(y) \right) \leq \|f\|_{\mathscr{Y}^*} \|x+0\| - f(0) = \|f\|_{\mathscr{Y}^*} \|x\|$$

we see that if $u_f > -\infty$ and $c \leq u_f$ then (3.3) will hold for any $\alpha > 0$. Thus, it will be enough to show that u_f is finite. Since for any $y \in \mathscr{Y}$

$$f(y) \leq \|f\|_{\mathscr{Y}^*} \|y\| = \|f\|_{\mathscr{Y}^*} \|(y+x) - x\|$$
$$\leq \|f\|_{\mathscr{Y}^*} \|x+y\| + \|f\|_{\mathscr{Y}^*} \|-x\| = \|f\|_{\mathscr{Y}^*} \|x+y\| + \|f\|_{\mathscr{Y}^*} \|x\|$$

we conclude that

$$\|f\|_{\mathscr{Y}^*} \|x+y\| - f(y) \geq - \|f\|_{\mathscr{Y}^*} \|x\|,$$

which implies that $u_f \geq - \|f\| \|x\| > -\infty$ as desired.
$\underline{\alpha < 0}$: Much like the previous case, we can rearrange (3.3) to read as

$$c + f\left(\frac{y}{\alpha}\right) \geq -\left(-\frac{1}{\alpha}\right) \|f\|_{\mathscr{Y}^*} \|\alpha x + y\| = -\|f\|_{\mathscr{Y}^*} \left\|x + \frac{y}{\alpha}\right\|$$

and see that the above holds if and only if for every $y \in \mathscr{Y}$

$$c \geq -\|f\|_{\mathscr{Y}^*} \|x+y\| - f(y).$$

Defining

$$l_f = \sup_{y\in\mathscr{Y}} \left(-\|f\|_{\mathscr{Y}^*} \|x+y\| - f(y) \right) \geq -\|f\|_{\mathscr{Y}^*} \|x+0\| + f(0) = -\|f\|_{\mathscr{Y}^*} \|x\|$$

we see that if $l_f < \infty$ and $c \geq l_f$ then (3.3) will hold for *any* $\alpha < 0$. Much like the previous case, we will now show that l_f is finite. Since for any x and y in $\mathscr{X}$

$$\|y - x\| \geq \|y\| - \|x\|$$

we see that for any $y \in \mathscr{Y}$

$$-\|f\|_{\mathscr{Y}^*} \|y-x\| + f(y) \leq -\|f\|_{\mathscr{Y}^*} \|y\| + \|f\|_{\mathscr{Y}^*} \|x\| + f(y)$$
$$\leq -\|f\|_{\mathscr{Y}^*} \|y\| + \|f\|_{\mathscr{Y}^*} \|x\| + \|f\|_{\mathscr{Y}^*} \|y\| = \|f\|_{\mathscr{Y}^*} \|x\|,$$

which implies that $l_f \leq \|f\|_{\mathscr{Y}^*} \|x\| < \infty$ as desired.
Looking at the three cases we've investigated we conclude that if we could find a number c such that

$$l_f \leq c \leq u_f$$

then $\widetilde{f}_c$ would provide an extension to f which has the same norm as f. This is guaranteed as long as $l_f \leq u_f$. The last inequality is equivalent to showing that for any $y_1, y_2 \in \mathscr{Y}$

$$-\|f\|_{\mathscr{Y}^*} \|x+y_1\| - f(y_1) \leq \|f\|_{\mathscr{Y}^*} \|x+y_2\| - f(y_2),$$

or alternatively, that for any $y_1, y_2 \in \mathscr{Y}$

$$f(y_2 - y_1) = f(y_2) - f(y_1) \leq \|f\|_{\mathscr{Y}_*} \left(\|x + y_1\| + \|x + y_2\|\right).$$

The above indeed holds as for any $y_1, y_2 \in \mathscr{Y}$

$$f(y_2 - y_1) \leq \|f\|_{\mathscr{Y}^*} \|y_2 - y_1\| = \|f\|_{\mathscr{Y}^*} \|(y_2 + x) - (y_1 + x)\|$$
$$\leq \|f\|_{\mathscr{Y}_*} \left(\|x + y_1\| + \|x + y_2\|\right).$$

We conclude that a choice of $c \in \mathbb{R}$ such that $\widetilde{f}_c$ is an extension of f to $\mathscr{Y}_x$ that preserves the norm is possible.
We now proceed to the general case, which will be shown by using Zorn's lemma. Let $\mathcal{M}$ be the set of all linear extensions of f that have the same norm as f, i.e. $g \in \mathcal{M}$ if and only if it is a linear functional such that

$$\mathscr{Y} = \mathscr{D}(f) \subset \mathscr{D}(g), \qquad g(x) = f(x) \quad \forall x \in \mathscr{Y}$$

and

$$\|g\|_{\mathscr{D}(g)^*} = \|f\|_{\mathscr{Y}}.$$

We define a partial order on $\mathcal{M}$ in the following way: $g \leq h$ if h is an extension of g. To be able to use Zorn's lemma we will now show that every chain $\mathcal{C} \subset \mathcal{M}$ has an upper bound. For a given chain $\mathcal{C}$ in $\mathcal{M}$ we define the a functional $\widetilde{g}_{\mathcal{C}}$ with a domain

$$\mathscr{D}(\widetilde{g}_{\mathcal{C}}) = \cup_{g \in \mathcal{C}} \mathscr{D}(g)$$

by

$$\widetilde{g}_{\mathcal{C}}(x) = g(x), \qquad \text{when } x \in \mathscr{D}(g) \text{ for some } g \in \mathcal{C}.$$

We start by mentioning that $\widetilde{g}_{\mathcal{C}}$ is well defined, i.e. doesn't depend on the choice of g. Indeed, assume that $x \in \mathscr{D}(g_1) \cap \mathscr{D}(g_2)$ for some $g_1, g_2 \in \mathcal{C}$. Since $\mathcal{C}$ is a chain we have that either $g_1 \leq g_2$ or $g_2 \leq g_1$. Without loss of generality we can assume that $g_1 \leq g_2$. As this implies that g_2 is an extension of g_1, we see that $x \in \mathscr{D}(g_1) \cap \mathscr{D}(g_2) = \mathscr{D}(g_1)$ and $g_2(x) = g_1(x)$, which shows that the choice of the function with which we define $\widetilde{g}_{\mathcal{C}}$ doesn't matter.
If $\widetilde{g}_{\mathcal{C}}$ is indeed a linear functional that extends f and satisfies $\|\widetilde{g}_{\mathcal{C}}\|_{\mathscr{D}(\widetilde{g}_{\mathcal{C}})^*} = \|f\|_{\mathscr{Y}^*}$, i.e. if $\widetilde{g}_{\mathcal{C}} \in \mathcal{M}$, we have that for any $g \in \mathcal{C}$, $\mathscr{D}(g) \subset \mathscr{D}(\widetilde{g}_{\mathcal{C}})$ and for any $x \in \mathscr{D}(g)$, $\widetilde{g}_{\mathcal{C}}(x) = g(x)$ by definition. Thus, $\widetilde{g}_{\mathcal{C}}$ will be an upper bound to the chain and we would be able to invoke Zorn's lemma.
To show the above we'll start by showing that $\mathscr{D}(\widetilde{g}_{\mathcal{C}})$ is a subspace of $\mathscr{X}$:

- Since $\mathcal{C}$ is non-empty there exists $g \in \mathcal{C}$ and as such, by definition, $\mathscr{D}(g) \subset \mathscr{D}(\widetilde{g}_{\mathcal{C}})$ which shows that $\mathscr{D}(\widetilde{g}_{\mathcal{C}})$ is not empty.
- For given $x_1, x_2 \in \mathscr{D}(\widetilde{g}_{\mathcal{C}})$ we can find $g_1, g_2 \in \mathcal{C}$ such that $x_1 \in \mathscr{D}(g_1)$ and $x_2 \in \mathscr{D}(g_2)$. As $\mathcal{C}$ is a chain we can assume without loss of generality that $g_1 \leq g_2$ and as such $x_1, x_2 \in \mathscr{D}(g_1) \cup \mathscr{D}(g_2) = \mathscr{D}(g_2)$. Since $\mathscr{D}(g_2)$ is a subspace of $\mathscr{X}$ we conclude that $x_1 + x_2 \in \mathscr{D}(g_2) \subset \mathscr{D}(\widetilde{g}_{\mathcal{C}})$.
- For a given $x \in \mathscr{D}(\widetilde{g}_{\mathcal{C}})$ there exists $g \in \mathcal{C}$ such that $x \in \mathscr{D}(g)$. As such, for any $\alpha \in \mathbb{R}$ we have that $\alpha x \in \mathscr{D}(g) \subset \mathscr{D}(\widetilde{g}_{\mathcal{C}})$.

We now turn our attention to showing that $\widetilde{g}_{\mathcal{C}}$ is linear and extends f. As we saw before, if $x_1, x_2 \in \mathscr{D}\left(\widetilde{g}_{\mathcal{C}}\right)$ then there exists $g \in \mathcal{C}$ such that $x_1, x_2 \in \mathscr{D}\left(g\right)$. By the definition of $\widetilde{g}_{\mathcal{C}}$ we have that

$$\widetilde{g}_{\mathcal{C}}\left(x_1 + x_2\right) = g\left(x_2 + x_2\right) \underset{g \in L(\mathscr{D}(g),\mathbb{R})}{=} g\left(x_1\right) + g\left(x_2\right) \underset{x_1,x_2 \in \mathscr{D}(g)}{=} \widetilde{g}_{\mathcal{C}}\left(x_1\right) + \widetilde{g}_{\mathcal{C}}\left(x_2\right),$$

showing the additive property. Similarly for any $x \in \mathscr{D}\left(\widetilde{g}_{\mathcal{C}}\right)$ there exists $g \in \mathcal{C}$ such that $x \in \mathscr{D}\left(g\right)$ and as such for any $\alpha \in \mathbb{R}$

$$\widetilde{g}_{\mathcal{C}}\left(\alpha x\right) = g\left(\alpha x\right) \underset{g \in L(\mathscr{D}(g),\mathbb{R})}{=} \alpha g\left(x\right) \underset{x \in \mathscr{D}(g)}{=} \alpha \widetilde{g}_{\mathcal{C}}\left(x\right),$$

which shows the scaling property and consequently the desired linearity. Moreover, since $\mathscr{D}\left(f\right) \subset \mathscr{D}\left(g\right)$ for any $g \in \mathcal{M}$ we have that $\mathscr{D}\left(f\right) \subset \mathscr{D}\left(\widetilde{g}_{\mathcal{C}}\right)$ and for any $x \in \mathscr{D}\left(f\right)$ we have that

$$\widetilde{g}_{\mathcal{C}}(x) = g(x) = f(x),$$

where $g \in \mathcal{C}$ was chosen arbitrarily.
Lastly, we will show that $\widetilde{g}_{\mathcal{C}}$ has the same norm as f. We have seen before that since $\widetilde{g}_{\mathcal{C}}$ extends f we have that

$$\left\|\widetilde{g}_{\mathcal{C}}\right\|_{\mathscr{D}\left(\widetilde{g}_{\mathcal{C}}\right)^*} \geq \left\|f\right\|_{\mathscr{Y}^*}.$$

To show the reverse inequality we notice that for any $x \in \mathscr{D}\left(\widetilde{g}_{\mathcal{C}}\right)$ we can find $g \in \mathcal{C}$ such that $\widetilde{g}_{\mathcal{C}}(x) = g(x)$. Since $g \in \mathcal{M}$ we find that

$$-\left\|f\right\|_{\mathscr{Y}^*}\|x\| \leq \underbrace{\widetilde{g}_{\mathcal{C}}(x)}_{=g(x)} \leq \left\|f\right\|_{\mathscr{Y}^*}\|x\|$$

and as $x \in \mathscr{D}\left(\widetilde{g}_{\mathcal{C}}\right)$ was arbitrary we conclude that by the definition

$$\left\|\widetilde{g}_{\mathcal{C}}\right\|_{\mathscr{D}\left(\widetilde{g}_{\mathcal{C}}\right)^*} \leq \left\|f\right\|_{\mathscr{Y}^*},$$

from which the equality of norms follows. Thus, $\widetilde{g}_{\mathcal{C}}$ is an element of $\mathcal{M}$ and is an upper bound to $\mathcal{C}$.
As the conditions for Zorn's lemma are satisfied, we know that there exists a maximal element in $\mathcal{M}$ which we will denote by $\widetilde{f}$. By definition $\widetilde{f}$ is an extension of f which has the same norm as f. Therefore, we'll conclude the proof of our theorem if we'll show that $\mathscr{D}\left(\widetilde{f}\right) = \mathscr{X}$. Indeed, if $\mathscr{D}\left(\widetilde{f}\right)$ is not $\mathscr{X}$ then there exists $x \in \mathscr{X}$ that is not in $\mathscr{D}\left(\widetilde{f}\right)$ and, as shown in our first step of the proof, we can extend $\widetilde{f}$ to $\operatorname{span}\left\{x, \mathscr{D}\left(\widetilde{f}\right)\right\}$ in a way the preserves the norm. This extension has a strictly larger domain than $\mathscr{D}\left(\widetilde{f}\right)$ - contradicting its maximality. The proof is thus complete. □

Next we will consider the case where the underlying field is $\mathbb{C}$. The idea of the proof is to utilise the Hahn-Banach theorem over $\mathbb{R}$ to show the general case of $\mathbb{C}$:

PROOF OF THE HAHN-BANACH THEOREM OVER $\mathbb{C}$. We start by noticing that if $f \in L\left(\mathscr{Y}, \mathbb{C}\right)$ then It is straight forward to show that

$$f_1 = \operatorname{Re}\left(f\right), \qquad f_2 = \operatorname{Im}\left(f\right)$$

belong to $L(\mathscr{Y},\mathbb{R})$ (note that we *can't* use complex scalars here!). Due to the linearity of f and the fact that by definition $f = f_1 + if_2$ we have that

$$i\left(f_1(x)+if_2(x)\right) = if(x) = f(ix) = f_1(ix) + if_2(ix)$$

which implies that

$$f_1(x) = f_2(ix), \qquad f_2(x) = -f_1(ix).$$

Thus, we conclude that

$$f(x) = f_1(x) - if_1(ix).$$

Additionally, we notice that for all $x \in \mathscr{Y}$

$$|f_1(x)| \le |f(x)| \le \|f\|_{\mathscr{Y}^*}\|x\|$$

which shows that f_1 if bounded and

$$\|f_1\|_{B(\mathscr{Y},\mathbb{R})} \le \|f\|_{\mathscr{Y}^*}.$$

Since $\mathscr{Y}$ is a vector space over $\mathbb{C}$, it is also a vector space over $\mathbb{R}$. As such we can extend the real valued functional f_1 *over* $\mathbb{R}$ to a functional $\widetilde{f}_1 \in B(\mathscr{X},\mathbb{R})$ with the same norm as f_1. We define

$$\widetilde{f}(x) = \widetilde{f}_1(x) - i\widetilde{f}_1(ix)$$

and claim that it is the desired (not necessarily unique) extension.
We start by showing that it is linear over $\mathbb{C}$. For any $x_1, x_2 \in \mathscr{X}$ we have that

$$\begin{aligned}\widetilde{f}(x_1+x_2) &= \widetilde{f}_1(x_1+x_2) - i\widetilde{f}_1(ix_1+ix_2) = \widetilde{f}_1(x_1)+\widetilde{f}_1(x_2) - i\left(\widetilde{f}_1(ix_1)+\widetilde{f}_1(ix_2)\right)\\ &= \left(\widetilde{f}_1(x_1) - i\widetilde{f}_1(ix_1)\right) + \left(\widetilde{f}_1(x_2) - i\widetilde{f}_1(ix_2)\right) = \widetilde{f}(x_1)+\widetilde{f}(x_2),\end{aligned}$$

showing the additive property. In addition, for any $x \in \mathscr{X}$ and $\alpha = a+ib$ with $a, b \in \mathbb{R}$ we have that

$$\widetilde{f}((a+ib)x) = \widetilde{f}(ax+ibx) \underset{\widetilde{f}\text{ is additive}}{=} \widetilde{f}(ax)+\widetilde{f}(ibx)$$

$$= \widetilde{f}_1(ax) - i\widetilde{f}_1(aix) + \widetilde{f}_1(ibx) - i\widetilde{f}_1(-bx) \underset{\widetilde{f}_1\in L(\mathscr{X},\mathbb{R})}{=} a\left(\widetilde{f}_1(x) - i\widetilde{f}_1(ix)\right) + b\left(\widetilde{f}_1(ix) - i\widetilde{f}_1(-x)\right)$$

$$= a\widetilde{f}(x) + ib\left(\widetilde{f}_1(x) - i\widetilde{f}_1(ix)\right) = (a+ib)\widetilde{f}(x),$$

which shows that scaling property over $\mathbb{C}$, from which we conclude that $\widetilde{f}$ is indeed linear over $\mathbb{C}$.
Next we notice that since $\widetilde{f}_1$ is an extension of f_1, for any $x \in \mathscr{Y}$ we have that

$$\widetilde{f}(x) = f_1(x) - if_1(ix) = f(x),$$

i.e. $\widetilde{f}$ extends f, as desired.
Lastly, we will show that $|\widetilde{f}(x)| \le \|f\|_{\mathscr{Y}^*}\|x\|$. As we saw before, since $\widetilde{f}$ extends f this inequality will imply that $\|\widetilde{f}\| = \|f\|_{\mathscr{Y}^*}$, which will conclude the proof of the theorem.
For a given $x \in \mathscr{X}$ we can find $\theta \in \mathbb{R}$ such that

$$|\widetilde{f}(x)| = e^{i\theta}\widetilde{f}(x) = \widetilde{f}\left(e^{i\theta}x\right).$$

As $\widetilde{f}\left(e^{i\theta}x\right) \in \mathbb{R}$ we have that

$$\widetilde{f}\left(e^{i\theta}x\right) = \mathrm{Re}\left(\widetilde{f}\left(e^{i\theta}x\right)\right) = \widetilde{f}_1\left(e^{i\theta}x\right)$$

and consequently

$$\left|\widetilde{f}(x)\right| = \widetilde{f}_1\left(e^{i\theta}x\right) \leq \left\|\widetilde{f}_1\right\|_{B(\mathscr{X},\mathbb{R})} \left\|e^{i\theta}x\right\| = \left\|f_1\right\|_{B(\mathscr{Y},\mathbb{R})} \|x\| \leq \left\|f\right\|_{\mathscr{Y}^*} \|x\|.$$

The proof is now complete. □

It is worth to note that one can find more general variation of the Hahn-Banach theorem whose proofs are very similar to the ones we gave above. We won't prove them here but will state them, together with the relevant definitions, for completion.

Definition 3.1.2. Let $\mathscr{X}$ be a normed space over a field $\mathbb{F}$ and let p be a function from $\mathscr{X}$ to $\mathbb{R}$. We say that p is *sub-additive* if for any $x, y \in \mathscr{X}$

$$p\left(x+y\right) \leq p(x) + p(y).$$

We say that p is *positive-homogeneous* if for any $x \in \mathscr{X}$ and $\alpha \geq 0$

$$p(\alpha x) = \alpha p(x),$$

and that it is *absolute-homogeneous* if for any $x \in \mathscr{X}$ and $\alpha \in \mathbb{F}$

$$p(\alpha x) = |\alpha|\, p(x).$$

A function p is called a *sublinear functional* if it is sub-additive and positive-homogeneous. It is called a *non-negative sublinear functional* if it is in addition non-negative.
A function p is called a *seminorm* if it is sub-additive and absolute-homogeneous.

Theorem 3.1.3 *(Hahn-Banach Theorem - Sublinear functionals). Let $\mathscr{X}$ be a normed space over $\mathbb{R}$ and let p be a sublinear functional on $\mathscr{X}$. Let $\mathscr{Y}$ be a subspace of $\mathscr{X}$ and let $f \in L\left(\mathscr{Y},\mathbb{R}\right)$ be such that for all $x \in \mathscr{Y}$*

$$f(x) \leq p(x).$$

Then there exists an extension $\widetilde{f} \in L(\mathscr{X},\mathbb{R})$ of f, i.e a linear functional such that $\widetilde{f}(x) = f(x)$ for all $x \in \mathscr{Y}$, that satisfies

$$\widetilde{f}(x) \leq p(x), \qquad \forall x \in \mathscr{X}.$$

Theorem 3.1.4 *(Hahn-Banach Theorem - Seminiorms). Let $\mathscr{X}$ be a normed space over a field $\mathbb{F}$, be it $\mathbb{R}$ or $\mathbb{C}$, and let p be a seminorm on on $\mathscr{X}$. Let $\mathscr{Y}$ be a subspace of $\mathscr{X}$ and let $f \in L\left(\mathscr{Y},\mathbb{F}\right)$ be such that for all $x \in \mathscr{Y}$*

$$\left|f(x)\right| \leq p(x).$$

Then there exists an extension $\widetilde{f} \in L(\mathscr{X},\mathbb{F})$ of f that satisfies

$$\left|\widetilde{f}(x)\right| \leq p(x), \qquad \forall x \in \mathscr{X}.$$

3.2. Applications of the Hahn-Banach Theorem

The applications of the Hahn-Banach theorem are numerous and vast. We will consider a few of them in this short section.

Theorem 3.2.1. *Let $\mathscr{X}$ be a Banach space and let $\mathscr{Y}$ be a finite dimensional subspace of $\mathscr{X}$. Then any linear funcitonal of $\mathscr{Y}$ can be extended to a bounded linear functional of $\mathscr{X}$.*

The above theorem tells us that $\mathscr{X}^*$ has *a lot* of functionals - more than those associated to *all possible* finite subspaces of $\mathscr{X}$.

PROOF. This is a direct consequence of the Hahn-Banach theorem and the fact that every linear funcitonal on a finite dimensional space is bounded. □

Theorem 3.2.2. *Let $\mathscr{X}$ be a Banach space and let $\mathscr{Y}$ be a closed subspace of $\mathscr{X}$. For any $x \notin \mathscr{Y}$ there exists a functional $f_{x,\mathscr{Y}} \in \mathscr{X}^*$ such that*

$$f_{x,\mathscr{Y}}(x) = 1, \qquad f_{x,\mathscr{Y}}|_{\mathscr{Y}} = 0.$$

Moreover, $\left\|f_{x,\mathscr{Y}}\right\| = \frac{1}{d}$ where $d = \inf_{y\in\mathscr{Y}} \left\|x-y\right\| > 0$.

Before we begin the proof we will mention that the fact that $d > 0$ follows from the assumptions that $\mathscr{Y}$ is closed and $x \notin \mathscr{Y}$. Indeed, $d \geq 0$ by definition and had $d = \inf_{y\in\mathscr{Y}} \left\|x-y\right\| = 0$ we could have found a sequence $\left\{y_n\right\}_{n\in\mathbb{N}} \subset \mathscr{Y}$ such that

$$y_n \underset{n\to\infty}{\longrightarrow} x.$$

Since $\mathscr{Y}$ is closed this would have implied that $x \in \mathscr{Y}$ which is a contradiction.

PROOF OF THEOREM 3.2.2. Consider the space $\mathscr{Y}_x = \text{span}\left\{x,\mathscr{Y}\right\} = \text{span}\{x\}+\mathscr{Y}$. We will prove our claim by defining a functional with the required conditions on $\mathscr{Y}_x$ and then using the Hahn-Banach theorem to extend it to the entire space. We start by noticing that since every $z \in \mathscr{Y}_x$ can be written *uniquely* as

$$z = \alpha_z x + y_z$$

with $y_z \in \mathscr{Y}$, any $f \in L\left(\mathscr{Y}_x,\mathbb{F}\right)$ that satisfies $f(x) = 1$ and $f|_{\mathscr{Y}} = 0$ must satisfy

$$f(z) = f\left(\alpha_z x + y_z\right) = \alpha_z f(x) + f\left(y_z\right) = \alpha_z.$$

If we'll show that f is bounded with $\left\|f\right\|_{\mathscr{Y}_x^*} = \frac{1}{d}$ we will conclude the proof.
For any $z \in \mathscr{Y}_x$ such that $\alpha_z \neq 0$ we see that

$$\left|f(z)\right| = \left|f\left(\alpha_z x + y_z\right)\right| = |\alpha_z| = \frac{\left\|\alpha_z x + y_z\right\|}{\left\|x - \frac{y}{\alpha_z}\right\|} = \frac{\|z\|}{\left\|x - \frac{y}{\alpha_z}\right\|} \underset{\frac{y}{\alpha_z}\in\mathscr{Y}}{\leq} \frac{\|z\|}{d}.$$

In the case where $\alpha_z = 0$, i.e. $z \in \mathscr{Y}$, we find that

$$\left|f(z)\right| = 0 \leq \frac{\|z\|}{d}.$$

We conclude that for any $z \in \mathcal{Y}_x$ we have that $\left|f(z)\right| \leq \frac{\|z\|}{d}$ which implies that $f \in B\left(\mathcal{Y}_x, \mathbb{F}\right)$ and

$$\|f\|_{\mathcal{Y}_x^*} \leq \frac{1}{d}.$$

To show the converse inequality we use the definition of d and for any $\varepsilon > 0$ we find $y_\varepsilon \in \mathcal{Y}$ such that

$$d \leq \left\|x - y_\varepsilon\right\| \leq (1+\varepsilon)\, d.$$

Since $x - y_\varepsilon \in \mathcal{Y}_x$

$$\|f\|_{\mathcal{Y}_x} \geq \frac{f\left(x-y_\varepsilon\right)}{\left\|x-y_\varepsilon\right\|} = \frac{1}{\left\|x-y_\varepsilon\right\|} \geq \frac{1}{(1+\varepsilon)\, d}.$$

As $\varepsilon > 0$ was arbitrary we conclude that

$$\|f\|_{\mathcal{Y}_x^*} \geq \frac{1}{d}$$

and with it, the proof. □

Two immediate consequences of the above are the following:

Theorem 3.2.3. *Let $\mathcal{X}$ be a Banach space and let $\mathcal{Y}$ be a subspace of $\mathcal{X}$. Then $\mathcal{Y}$ is dense in $\mathcal{X}$ if and only if any $f \in \mathcal{X}^*$ such that $f|_{\mathcal{Y}} = 0$ must be the zero functional.*

PROOF. Assume that $\mathcal{Y}$ is dense in $\mathcal{X}$ and let $f \in \mathcal{X}^*$ be such that $f|_{\mathcal{Y}} = 0$. As f is bounded, we have that $f|_{\overline{\mathcal{Y}}} = 0$ and since $\overline{\mathcal{Y}} = \mathcal{X}$ we conclude that $f \equiv 0$. Assume now that $\mathcal{Y}$ is not dense in $\mathcal{X}$. As $\overline{\mathcal{Y}}$ is a closed subspace that, by assumption, is not $\mathcal{X}$ we can find a vector x that is not in $\overline{\mathcal{Y}}$. Using Theorem 3.2.2 we find $f_{x,\overline{\mathcal{Y}}} \in \mathcal{X}^*$ which satisfies $f_{x,\overline{\mathcal{Y}}}(x) \neq 0$ and $f_{x,\overline{\mathcal{Y}}}|_{\overline{\mathcal{Y}}} = 0$. Since $\mathcal{Y} \subset \overline{\mathcal{Y}}$ we have found $f \not\equiv 0 \in \mathcal{X}^*$ such that $f|_{\mathcal{Y}} = 0$. This concludes the proof. □

Theorem 3.2.4. *Let $\mathcal{X}$ be a Banach space. For any $x \in \mathcal{X}$ there exists $f_x \in \mathcal{X}^*$ such that $\left\|f_x\right\| = 1$ and $f_x(x) = \|x\|$. Consequently, we have that for any $x \in \mathcal{X}$*

$$\|x\| = \sup_{f \in \mathcal{X}^*,\, f \neq 0} \frac{\left|f(x)\right|}{\|f\|}. \tag{3.4}$$

REMARK 3.2.5. Equation (3.4) is extremely importance as it shows how the norm on $\mathcal{X}$ can be understood as *an operator norm-like expression* which is induced from $\mathcal{X}^*$.

PROOF. We start by noticing that for any $f \in \mathcal{X}^*$

$$f(0) = 0 = \|0\|,$$

so when $x = 0$ we can choose any bounded funcitonal of norm one. Next, we consider $x \neq 0$. Since the trivial subspace $\mathcal{Y} = \{0\}$ is closed we conclude from Theorem 3.2.2 that there exists $\widetilde{f} \in \mathcal{X}^*$ such that $\widetilde{f}(x) = 1$, $\widetilde{f}|_{\{0\}} = 0$ (which always holds) and

$$\|\widetilde{f}\| = \frac{1}{\inf_{y \in \{0\}} \left\|x - y\right\|} = \frac{1}{\|x\|}.$$

Defining $f_x = \|x\| \widetilde{f}$ we see that $f_x \in \mathcal{X}^*$, $f_x(x) = \|x\|$ and $\|f_x\| = 1$.
To show 3.4 we notice that if $f \in \mathcal{X}^*$ is not zero then

$$\frac{|f(x)|}{\|f\|} \leq \frac{\|f\| \|x\|}{\|f\|} = \|x\|$$

which implies that

$$\|x\| \geq \sup_{f \in \mathcal{X}^*,\, f \neq 0} \frac{|f(x)|}{\|f\|}.$$

On the other hand, for any x we have that

$$\|x\| = \frac{|f_x(x)|}{\|f_x\|} \leq \sup_{f \in \mathcal{X}^*,\, f \neq 0} \frac{|f(x)|}{\|f\|}.$$

Combining the above inequalities gives us the desired proof. □

An immediate corollary of Theorem 3.2.4 is that if $\mathcal{X}$ is non-trivial then $\mathcal{X}^*$ is also non-trivial. Since $\mathcal{X} = \{0\}$ automatically implies that every linear functional on $\mathcal{X}$ is the zero functional we conclude that $\mathcal{X}$ *is non-trivial if and only if* $\mathcal{X}^*$ *is non-trivial.*

REMARK 3.2.6. As is evident from all the above, the Hahn-Banach theorem have given us a way to relate properties of $\mathcal{X}$ and $\mathcal{X}^*$ to each other. It is also, in fact, an essential ingredient in showing Theorem 2.4.17 from section §2.4, which stated that if $\mathcal{X}^*$ is separable then so is $\mathcal{X}$.

3.3. The Banach-Steinhaus Theorem (uniform boundedness principle)

The Banach-Steinhaus theorem, sometimes known as the *Uniform Boundedness principle/theorem*, asserts that if a sequence of bounded linear operators is bounded pointwise for any $x \in \mathcal{X}$, i.e. when applied to any $x \in \mathcal{X}$, then it is bounded in the operator norm - which is uniform in x.

Theorem 3.3.1 *(The Banach-Steinhaus Theorem). Let $\mathcal{X}$ be a Banach space and let $\mathcal{Y}$ be a normed space. If $\{T_n\}_{n\in\mathbb{N}} \subset B(\mathcal{X}, \mathcal{Y})$ satisfies that*

$$\sup_{n\in\mathbb{N}} \|T_n x\| < \infty, \qquad \forall x \in \mathcal{X}$$

then

$$\sup_{n\in\mathbb{N}} \|T_n\| < \infty.$$

The proof of this theorem relies on a deep result in the study of metric spaces known as the *Baire's category theorem.* The proof of it, as well as the most general version of it, lies outside the realm of our module. We will only state without proof a version of it that we will use here:

Theorem 3.3.2 *(The Baire Category Theorem - Functional Analysis version). Let $\mathcal{X}$ be a Banach space. Then if $\mathcal{X} = \cup_{n\in\mathbb{N}} M_n$ where $\{M_n\}_{n\in\mathbb{N}}$ are closed sets, then one of the M_n-s must contain an open ball.*

We will usually utilise the above theorem in Functional Analysis in the following way:

- Write the space $\mathcal{X}$ as a countable union of closed sets, usually connected to one or more (potentially a countable family of) linear operators.
- Conclude that one of the sets must contain an open ball.
- Use properties of the linear operators together with the knowledge that one of the closed sets contains a ball to conclude some brilliant result.

PROOF OF THEOREM 3.3.1. For any $k, n \in \mathbb{N}$ we define

$$M_{k,n} = \{x \in \mathcal{X} \mid \|T_n x\| \leq k\}$$

and

$$M_k = \{x \in \mathcal{X} \mid \|T_n x\| \leq k, \text{ for all } n \in \mathbb{N}\} = \cap_{n\in\mathbb{N}} M_{k,n}.$$

We have that $M_{k,n}$ is closed for any k and n in $\mathbb{N}$, and since M_k is the intersection of $M_{k,n}$−s, it must be closed as well. Indeed, if $\{x_j\}_{j\in\mathbb{N}} \subset M_{k,n}$ converges to x then due to the boundedness of T_n and the continuity of the norm

$$\|T_n x\| = \lim_{j\to\infty} \left\|T_n x_j\right\| \underset{x_j\in M_{k,n}}{\leq} k$$

which shows that $x \in M_{k,n}$.
Next we notice that since for any $x \in \mathcal{X}$

$$\sup_{n\in\mathbb{N}} \|T_n x\| < \infty$$

we are able to find $k(x) \in \mathbb{N}$ for any $x \in \mathcal{X}$ such that

$$\sup_{n\in\mathbb{N}} \|T_n x\| \leq k(x).$$

Thus, any $x \in \mathcal{X}$ belongs to $\cup_{k\in\mathbb{N}} M_k$, which implies that $\mathcal{X} = \cup_{k\in\mathbb{N}} M_k$.
According to our variant of Baire's Category theorem there must exist $k_0 \in \mathbb{N}$, $x_0 \in \mathcal{X}$ and $r > 0$ such that

$$B_r(x_0) \subset M_{k_0}.$$

Since for any $x \in \mathcal{X}$ such that $x \neq 0$ we have that $x_0 + \frac{r}{2} \cdot \frac{x}{\|x\|} \in B_r(x_0)$ we find that

$$\left\|T_n\left(x_0 + \frac{r}{2} \cdot \frac{x}{\|x\|}\right)\right\| \leq k_0 \qquad \text{for all } n \in \mathbb{N}.$$

Consequently, for any $x \neq 0$ and any $n \in \mathbb{N}$

$$\frac{\|T_n x\|}{\|x\|} = \left\|T_n\left(\frac{x}{\|x\|}\right)\right\| = \frac{2}{r}\left\|T_n\left(\frac{r}{2} \cdot \frac{x}{\|x\|}\right)\right\| = \frac{2}{r}\left\|T_n\left(x_0 + \frac{r}{2} \cdot \frac{x}{\|x\|}\right) - T_n x_0\right\|$$

$$\leq \frac{2\left\|T_n\left(x_0 + \frac{r}{2} \cdot \frac{x}{\|x\|}\right)\right\| + 2\|T_n x_0\|}{r} \leq \frac{2k_0 + 2k(x_0)}{r},$$

from which we conclude that

$$\|T_n\| = \sup_{x\in\mathcal{X},\, x\neq 0} \frac{\|T_n x\|}{\|x\|} \leq \frac{2k_0 + 2k(x_0)}{r}, \qquad \forall n \in \mathbb{N}$$

As the right hand side of the above inequality is independent of n we find that

$$\sup_{n\in\mathbb{N}} \| T_n \| \leq \frac{2k_0 + 2k(x_0)}{r},$$

and conclude the proof of the theorem. □

An important and extremely useful application of the Banach-Steinhaus theorem is the following:

Theorem 3.3.3. *Let $\mathcal{X}$ be a Banach space and let $\mathcal{Y}$ be a normed space. If $\{T_n\}_{n\in\mathbb{N}} \subset B(\mathcal{X}, \mathcal{Y})$ satisfies that*

$$\lim_{n\to\infty} T_n x \qquad \text{exists for all } x \in \mathcal{X}$$

then the map $T : \mathcal{X} \to \mathcal{Y}$ defined by

$$Tx = \lim_{n\in\mathbb{N}} T_n x$$

is a bounded linear map.

The Banach-Steinhaus theorem is the key to show the boundedness of weak and weak-$*$ sequences which was mentioned in Theorem 2.5.4.

Lemma 3.3.4 *(Boundedness of weak-$*$ converging sequences). Let $\mathcal{X}$ be a Banach space. If $\{f_n\}_{n\in\mathbb{N}} \subset \mathcal{X}^*$ converges weakly-$*$ to $f \in \mathcal{X}^*$ then $\{\|f_n\|\}_{n\in\mathbb{N}}$ is bounded.*

PROOF. By the definition of weak-$*$ convergence we have that for any $x \in \mathcal{X}$

$$f_n(x) \underset{n\to\infty}{\longrightarrow} f(x)$$

which implies (as converging sequences are always bounded) that for any $x \in \mathcal{X}$

$$\sup_{n\in\mathbb{N}} \left| f_n(x) \right| < \infty.$$

Using the Banach-Steinhaus theorem we can conclude that

$$\sup_{n\in\mathbb{N}} \| f_n \| < \infty$$

which is the desired result □

REMARK. The same theorem, and idea as presented above, show that if $\{x_n\}_{n\in\mathbb{N}}$ converges weakly to x then $\{x_n\}_{n\in\mathbb{N}}$ must be bounded. In order to utilise the Banach-Steinhaus theorem, however, we need to somehow think of $\{x_n\}_{n\in\mathbb{N}}$ as a family of bounded linear operators. This is possible by considering the *second dual space*, $\mathcal{X}^{**} = (\mathcal{X}^*)^*$. Indeed, for any $x \in \mathcal{X}$ we can define $\widehat{x} \in \mathcal{X}^{**}$ by

$$\widehat{x}(f) = f(x), \qquad f \in \mathcal{X}^*.$$

It is straight forward to check that $\widehat{x} \in L(\mathcal{X}^*, \mathbb{F})$. The fact that it is bounded follows directly from Theorem 3.2.4. Indeed,

$$\|\widehat{x}\| = \sup_{f\neq 0} \frac{\left|\widehat{x}(f)\right|}{\|f\|} = \sup_{f\neq 0} \frac{\left|f(x)\right|}{\|f\|} = \|x\|.$$

Since we can "lift" $\{x_n\}_{n\in\mathbb{N}}$ to the sequence $\{\widehat{x_n}\}_{n\in\mathbb{N}}$ in $\mathcal{X}^{**}$ and since the weak convergence of $\{x_n\}_{n\in\mathbb{N}}$ implies the weak-$*$ convergence of $\{\widehat{x_n}\}_{n\in\mathbb{N}}$ we find that

$$\sup_{n\in\mathbb{N}} \|x_n\| = \sup_{n\in\mathbb{N}} \|\widehat{x_n}\| < \infty.$$

It is worth to note that the identification of x with $\widehat{x}$ and the fact that their norm are identical gives us a way to embed $\mathcal{X}$ in $\mathcal{X}^{**}$ in a canonical way. This embedding is very important and the question of whether or not it is surjective is the topic of *Reflexive Banach spaces*.

3.4. The Open Mapping Theorem

The second application of the Baire Category theorem we'll discuss in this module is the *Open Mapping Theorem*.

A known criterion for continuity of a map between metric spaces is that the *pre-image* of every open set is an open set. A natural question that one may consider at this point is: What do continuous maps *do* to open sets? Do they take them to open sets? In general the answer to that is *No*.

Example 3.4.1. Consider the function $\sin(x)$ on $\mathbb{R}$. It is simple to check that it takes the open interval $\left(-\frac{2\pi}{3}, \frac{2\pi}{3}\right)$ to the closed interval $[-1,1]$

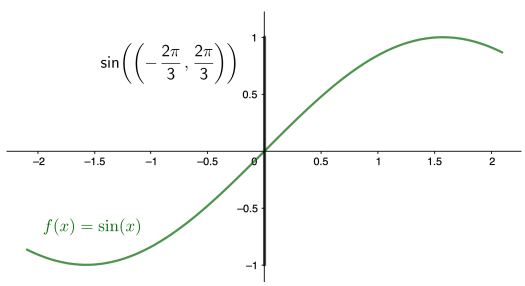


Figure 3.1. The function $\sin(x)$ doesn't take all open sets to open sets.

Maps that do take open sets to open sets deserve special attention:

Definition 3.4.2. Let X and Y be two metric spaces. A map $T : \mathcal{D}(T) \subset X \to Y$ is called *an open map* if $T(U \cap \mathcal{D}(T))$ is open for any open set U.

While continuos maps are not open in general, a subclass of such functions is very natural in this framework: Noticing that when T is an invertible map we have that the pre-image of a set corresponds to the image of that set by the map T^{-1} we conclude that *when $T : \mathcal{D}(T) \subset X \to Y$ is injective, and as such a bijection onto $\mathcal{R}(T)$, we have that T is an open map if and only if $T^{-1} : \mathcal{R}(T) \to \mathcal{D}(T)$ is continuous.*

As we've seen already, the properties of bounded linear operators interact extremely well with the norm induced topology of Banach spaces. The next theorem is yet another example:

Theorem 3.4.3 *(Open Mapping Theorem). Let $\mathscr{X}$ and $\mathscr{Y}$ be two Banach spaces. If $T:\mathscr{X}\to\mathscr{Y}$ is a bounded linear operator onto $\mathscr{Y}$ then T is an open map.*

SKETCH OF THE PROOF OF THEOREM 3.4.3. The proof of the theorem is quite involved and we shall only sketch its steps here.
Step 1: Since T is *onto* $\mathscr{Y}$ we find that

$$\mathscr{Y}=\cup_{n\in\mathbb{N}}n\overline{T\left(B_1^{\|\cdot\|_{\mathscr{X}}}(0)\right)}$$

and using Baire'e category theory we conclude that there exist $\eta>0$ and $y_*\in\mathscr{Y}$ such that

$$B_\eta^{\|\cdot\|_{\mathscr{Y}}}\left(y_*\right)\subset\overline{T\left(B_1^{\|\cdot\|_{\mathscr{X}}}(0)\right)}.$$

Step 2: Using the fact that $B_\delta^{\|\cdot\|_{\mathscr{Y}}}\left(y\right)=y+B_\delta^{\|\cdot\|_{\mathscr{Y}}}(0)$ we see that

$$B_\eta^{\|\cdot\|_{\mathscr{Y}}}(0)\subset\overline{T\left(B_1^{\|\cdot\|_{\mathscr{X}}}(0)\right)}-y_*,$$

from which we can infer that

$$B_\eta^{\|\cdot\|_{\mathscr{Y}}}(0)\subset\overline{T\left(B_2^{\|\cdot\|_{\mathscr{X}}}(0)\right)}.$$

Step 3: The most difficult step is to show that if we reduce by a factor of half the radius of the left hand side ball in the last inclusion of Step 2, we would end up in $T\left(B_2^{\|\cdot\|_{\mathscr{X}}}(0)\right)$ *and not its closure*. In other words:

$$B_{\frac{\eta}{2}}^{\|\cdot\|_{\mathscr{Y}}}(0)\subset T\left(B_2^{\|\cdot\|_{\mathscr{X}}}(0)\right).$$

Step 4: With the above at hand we have that for any $x_0\in\mathscr{X}$ and any $\varepsilon>0$

$$B_{\frac{\varepsilon\eta}{4}}^{\|\cdot\|_{\mathscr{Y}}}(Tx_0)\subset T\left(B_\varepsilon^{\|\cdot\|_{\mathscr{X}}}(x_0)\right),$$

which is enough to show that T is an open map. □

The Open Mapping theorem has many applications, one of which is the Closed Graph Theorem which we will consider in the next section. We end this section with a relatively straight forwards yet extremely important consequence of the theorem. It is used many times in the *Spectral study* of operators (bounded or unbounded).

Theorem 3.4.4 *(Open Mapping Theorem for injective linear operators). Let $\mathscr{X}$ and $\mathscr{Y}$ be two Banach spaces and let $T:\mathscr{X}\to\mathscr{Y}$ be a bounded and injective linear operator. If $\mathscr{R}(T)$ is closed in $\mathscr{Y}$ then the inverse map of T, $T^{-1}:\mathscr{R}(T)\to\mathscr{X}$ is a bounded linear operator.*

PROOF. As we've mentioned before, if $T:\mathscr{D}(T)\to\mathscr{Y}$ is an injective linear operator then so is its inverse $T^{-1}:\mathscr{R}(T)\to\mathscr{D}(T)$. This means that to prove our theorem we only need to show that T^{-1} is bounded, or equivalently that T is an open map. Since $\mathscr{R}(T)$ is closed in $\mathscr{Y}$, and $\mathscr{Y}$ is a Banach space, it must be a Banach space as well. We find that $T:\mathscr{X}\to\mathscr{R}(T)$ satisfies the conditions of the Open Mapping theorem and is conseqeuntly an open map. The proof is now complete. □

REMARK. If $T : \mathscr{X} \to \mathscr{Y}$ is injective and its inverse is bounded then for any $y \in \mathscr{R}(T)$ we have that

$$\left\| T^{-1} y \right\| \le \left\| T^{-1} \right\| \|y\| .$$

Consequently, for any $x \in \mathscr{X}$

$$\|x\| = \left\| T^{-1} (Tx) \right\| \le \left\| T^{-1} \right\| \|Tx\| ,$$

or equivalently, for any $x \in \mathscr{X}$

$$\|Tx\| \ge \frac{\|x\|}{\left\| T^{-1} \right\|}.$$

We thus see that while the boundedness of T implies an upper bound with respect to $\|x\|$ for $\|Tx\|$, the boundedness of T^{-1} implies a *lower bound* with respect to $\|x\|$ for $\|Tx\|$

REMARK 3.4.5. The condition that $\mathscr{R}(T)$ is closed is *necessary*. Recall that we saw in Remark 2.2.11 that the operator $T : \ell_1(\mathbb{N}) \to \ell_1(\mathbb{N})$ defined as

$$T(\boldsymbol{a}) = \left(a_1, \frac{a_2}{2}, \dots, \frac{a_n}{n}, \dots \right)$$

is a bounded operator whose range is not close. It is straight forwards to verify that T is injective and that $T^{-1} : \ell_1(\mathbb{N}) \to \ell_1(\mathbb{N})$ is defined by

$$T^{-1}(\boldsymbol{a}) = (a_1, 2a_2, \dots, na_n, \dots) .$$

T^{-1} is *not* continuous since for every $n \in \mathbb{N}$ we have that $\boldsymbol{e}_n = T(n\boldsymbol{e}_n) \in \mathscr{R}(T) = \mathscr{D}\left(T^{-1}\right)$ and

$$\left\| T^{-1}(\boldsymbol{e_n}) \right\|_1 = \|n\boldsymbol{e}_n\|_1 = n$$

which implies that

$$\sup_{x \in \mathscr{D}(T^{-1}),\, \|x\|=1} \left\| T^{-1} x \right\| \ge \sup_{n \in \mathbb{N}} \left\| T^{-1}(\boldsymbol{e}_n) \right\| = \sup_{n \in \mathbb{N}} n = \infty.$$

3.5. Closed Graph Theorem

In the last section of our notes we will introduce another criterion for the boundedness of a linear operator - a criterion that relies of the geometric notion of a *graph of an operator*.

Definition 3.5.1. Let X and Y be sets and let $T : \mathscr{D}(T) \subset X \to Y$ be a given map. *The graph of* T, denoted by $\mathscr{G}(T)$, is the subset of the set $X \times Y$ defined as

$$\mathscr{G}(T) = \left\{ (x, y) \mid x \in \mathscr{D}(T),\ y = Tx \right\} = \{ (x, Tx) \mid x \in \mathscr{D}(T) \} .$$

When X and Y are not just sets but have some linear and/or topological structure we can induce a linear and/or topological structure on $X \times Y$ and investigate linear and/or topological properties of $\mathscr{G}(T)$.

Theorem 3.5.2. *Let $\mathscr{X}$ and $\mathscr{Y}$ be vector spaces over the same field $\mathbb{F}$. Consider the following operators: $+ : \mathscr{X} \times \mathscr{Y} \to \mathscr{X} \times \mathscr{Y}$ defined by*

$$(x_1, y_1) + (x_2, y_2) = \left(x_1 +_{\mathscr{X}} x_2, y_1 +_{\mathscr{Y}} y_2 \right),$$

where $+_{\mathcal{X}}$ *and* $+_{\mathcal{Y}}$ *are the addition operations on* $\mathcal{X}$ *and* $\mathcal{Y}$ *respectively, and* $\cdot : \mathbb{F} \times (\mathcal{X} \times \mathcal{Y}) \to \mathcal{X} \times \mathcal{Y}$ *defined by*

$$\alpha \cdot (x, y) = (\alpha \cdot_{\mathcal{X}} x, \alpha \cdot_{\mathcal{Y}} y),$$

where $\cdot_{\mathcal{X}}$ *and* $\cdot_{\mathcal{Y}}$ *are the scalar multiplication operations on* $\mathcal{X}$ *and* $\mathcal{Y}$ *respectively. Then* $\mathcal{X} \times \mathcal{Y}$ *is a vector space under these operations. Its additive zero is given by* $0 = (0_{\mathcal{X}}, 0_{\mathcal{Y}})$ *and the additive inverse to* (x, y) *is* $(-x, -y)$*.*
If, in addition, $\mathcal{X}$ *and* $\mathcal{Y}$ *are normed spaces then the function* $\|\cdot\| : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+$ *given by*

$$\|(x, y)\| = \|x\|_{\mathcal{X}} + \|y\|_{\mathcal{Y}}$$

is a norm on $\mathcal{X} \times \mathcal{Y}$*. Moreover, if* $\mathcal{X}$ *and* $\mathcal{Y}$ *are Banach spaces under their respective norm, then so is* $\mathcal{X} \times \mathcal{Y}$ *under* $\|\cdot\|$*.*

From this point onwards, unless stated differently, we will always consider $\mathcal{X} \times \mathcal{Y}$ as the normed space described above.

The first property we notice for $\mathcal{G}(T)$ is its *inherent linearity*, when T is a linear operator which we state without proof

Lemma 3.5.3. *Let* $\mathcal{X}$ *and* $\mathcal{Y}$ *be two vector spaces and let* $T : \mathcal{D}(T) \subset \mathcal{X} \to \mathcal{Y}$ *be a linear operator. Then* $\mathcal{G}(T)$ *is a subspace of* $\mathcal{X} \times \mathcal{Y}$*.*

In order to explore the connection between the topological properties of a map and its associated graph we now define an extremely important notion:

Definition 3.5.4. Let $\mathcal{X}$ and $\mathcal{Y}$ be two normed spaces and let $T : \mathcal{D}(T) \subset \mathcal{X} \to \mathcal{Y}$ be a linear operator. We say that T is a *closed linear operator* if its graph, $\mathcal{G}(T)$, is a closed set in $\mathcal{X} \times \mathcal{Y}$.

REMARK 3.5.5. Since the norm on $\mathcal{X} \times \mathcal{Y}$ is given by

$$\|(x, y)\| = \|x\|_{\mathcal{X}} + \|y\|_{\mathcal{Y}}$$

we see that $\{(x_n, y_n)\}_{n \in \mathbb{N}}$ converges to (x, y), i.e.

$$\|(x_n, y_n) - (x, y)\| = \|x_n - x\|_{\mathcal{X}} + \|y_n - y\|_{\mathcal{Y}} \underset{n \to \infty}{\longrightarrow} 0,$$

if and only if $\{x_n\}_{n \in \mathbb{N}}$ converges to x in $\mathcal{X}$ and $\{y_n\}_{n \in \mathbb{N}}$ converges to y in $\mathcal{Y}$.
Consequently $\mathcal{G}(T)$ is closed in $\mathcal{X} \times \mathcal{Y}$ if and only if

$$x_n \underset{n \to \infty}{\longrightarrow} x \quad \text{and} \quad Tx_n \underset{n \to \infty}{\longrightarrow} y$$

imply that $x \in \mathcal{D}(T)$ and $y = Tx$.

While bounded linear operators are extremely interesting, many of the operators we encounter in applications such as quantum physics and PDEs are *unbounded*. However, almost all the operators we deal with *are closed*. Closed operators enjoy a plethora of beautiful and useful properties.

Example 3.5.6 *(The derivative is a closed linear operator)*. Let

$$D : C^1[0,1] \subset C[0,1] \to C[0,1]$$

be the derivative operator, where $C[0,1]$ is imbued with the norm $\|\cdot\|_\infty$ and $C^1[0,1]$ is the space of all continuously differentiable functions on $[0,1]$. We have seen before that D can't be a bounded linear operator but it is a closed linear operator.

What is the connection between closed operators and bounded operators? If $T:\mathscr{D}(T)\subset\mathscr{X}\to\mathscr{Y}$ is bounded and if $\{x_n\}_{n\in\mathbb{N}}\subset\mathscr{D}(T)$ converges to an element x *that is also in* $\mathscr{D}(T)$ then

$$Tx_n \underset{n\to\infty}{\longrightarrow} Tx.$$

Consequently we conclude that If $\mathscr{D}(T)$ is closed and $T:\mathscr{D}(T)\subset\mathscr{X}\to\mathscr{Y}$ is a bounded linear operator then T must also be a closed linear operator. As we have shown that any bounded linear operator can be extended *uniquely* to a bounded linear operator on the closure of its domain we see that, modulus an extension, every bounded linear operator is closed.

The next theorem shows the opposite direction:

Theorem 3.5.7 *(Closed Graph Theorem). Let $\mathscr{X}$ and $\mathscr{Y}$ be two Banach spaces and let $T:\mathscr{D}(T)\subset\mathscr{X}\to\mathscr{Y}$ be a closed linear operator. If $\mathscr{D}(T)$ is closed then T is bounded.*

PROOF. Consider the map $\Pi:\mathscr{G}(T)\to\mathscr{X}$ defined by

$$\Pi(x,Tx)=x.$$

It is straight forward to show that Π is injective and onto $\mathscr{R}(\Pi)=\mathscr{D}(T)$.
Since $\mathscr{G}(T)$ is a closed subspace of a Banach space it is a Banach space, and as $\mathscr{D}(T)$ is closed Theorem 3.4.4 guarantees that Π has an inverse, $\Pi^{-1}:\mathscr{D}(T)\to\mathscr{G}(T)$ that is bounded. Since

$$\Pi^{-1}x=(x,Tx)$$

and

$$\|x\|+\|Tx\|=\left\|\Pi^{-1}x\right\|\leq\left\|\Pi^{-1}\right\|\|x\|$$

we conclude that

$$\|Tx\|\leq\|x\|+\|Tx\|\leq\left\|\Pi^{-1}\right\|\|x\| \qquad \forall x\in\mathscr{D}(T).$$

This shows the desired boundedness T and conclude the proof. □

REMARK 3.5.8. When we try to show that an operator T is bounded, i.e. continuous, we need to show that if $\{x_n\}_{n\in\mathbb{N}}\subset\mathscr{D}(T)$ converges to $x\in\mathscr{D}(T)$ then $\{Tx_n\}_{n\in\mathbb{N}}$ converges to Tx. When we try to show that an operator T is closed, on the other hand, we don't just assume that $\{x_n\}_{n\in\mathbb{N}}$ converges to x - but also that $\{Tx_n\}_{n\in\mathbb{N}}$ converges to an element y. Unlike boundedness *we have an extra condition to use which might make the investigation simpler.* This makes the Closed Graph Theorem very useful in many cases.

We conclude this part of the module with an immediate corollary of the Closed Graph theorem is:

Corollary 3.5.9. $T\in B(\mathscr{X},\mathscr{Y})$ if and only if $T:\mathscr{X}\to\mathscr{Y}$ is closed.

PROOF. Since $T \in B(\mathcal{X},\mathcal{Y})$ implies that the domain of T is closed we conclude that if $T \in B(\mathcal{X},\mathcal{Y})$ then T must be closed. Conversely, if $T : \mathcal{X} \to \mathcal{Y}$ then its domain is closed (as it is $\mathcal{X}$) and as such the Closed Graph theorem grantees that it is bounded when it is closed. □